Interpretation of the Composition of Coniferous Resin

By J. S. Birks and P. J. Kanowski[1])

Oxford Forestry Institute, Oxford University,
South Parks Road, Oxford, OX1 3RB, England

(Received 8th July 1986)

Abstract

The composition of coniferous oleoresin has been extensively studied since chromatographic analysis made its accurate determination possible. Research has been directed to investigation of the biosynthetic processes responsible for the production of resin constituents and to the genetic interpretation of differences in resin composition. Studies of both aspects have been hampered by the nature of resin compositional data, in which the proportions of all constituents are constrained to sum to unity.

Although progress has been made in elucidating biosynthetic pathways, the specific processes by which genetic control is exercised remain unclear. Nevertheless, there have been many proposals that the production of particular terpenes is controlled by single genes, usually with a dominant/necessive pair of alleles. Frequently it has not been specified whether monogenic control is exercised over the actual quantity of a constituent or the proportion of the total that it comprises. However, the mode of control of the former will not be apparent from compositional data. Monogenic control of the latter is unsubstantiated by the available evidence.

The uncritical citation of previous work and use of inappropriate statistical methods has led to confused and inconsistent results. The nature of resin compositional data has been fundamentally misunderstood. The proportions of particular constituents have often been treated as though they were absolute quantities. However, variation apparent in compositional data may not exist in the absolute quantities, nor is there any direct method of deducing the latter from the former. Composition should be treated as a single vector trait. The cautious application of appropriate methods of classification and analysis is required if we are to understand the reasons for variation in resin composition.

Key words: biosynthesis, compositional data, forest genetics, gas chromatography, multivariate statistical analysis, monoterpenes, oleoresin.

Zusammenfassung

Die Zusammensetzung von Harzen bei Koniferen wird ausführlich untersucht, seitdem durch die chromatographische Analyse eine genaue Bestimmung möglich geworden ist. Die Forschung wurde auf die Untersuchung biosynthetischer Prozesse, die für die Produktion der Harzbestandteile verantwortlich sind, und auf die genetische Interpretation von Unterschieden in der Harzzusammensetzung hin ausgerichtet. Die Untersuchungen beider Aspekte wurden durch die Natur der Daten der Harzzusammensetzung gehemmt, in denen die Proportionen aller Bestandteile als Summe zu einer Einheit zusammengezogen werden.

Obwohl bei der Aufklärung biosynthetischer Wege Fortschritte gemacht worden sind, bleiben die spezifischen Prozesse, bei denen genetische Kontrolle ausgeübt wird, unklar. Nichtsdestoweniger hat es zahlreiche Vorschläge gegeben, daß die Produktion besonderer Terpene durch Einzelgene kontrolliert wird, normalerweise durch ein dominant-rezessives Allelenpaar. Oft wurde nicht spezifiziert, ob die monogenische Kontrolle über die aktuelle Menge eines Bestandteiles oder über die Proportion des Ganzen, das es einschließt, ausgeübt wird. Jedenfalls wird die Art der Kontrolle, ob über die aktuelle Menge oder über die proportionale Zusammensetzung, nicht sichtbar. Die monogenische Kontrolle der letzteren ist durch vorhandene Beweise unbestätigt.

Das unkritische Zitieren früherer Arbeiten und die Verwendung ungeeigneter statistischer Methoden führte zu konfusen und unvereinbaren Resultaten. Die Eigenart der Daten der Harzzusammensetzung wurde grundsätzlich mißverstanden. Die Verhältnisse besonderer Bestandteile wurden oft so behandelt, als handele es sich dabei um absolute Mengen. Weder mag die Variation, die in den Daten aus anteiliger Zusammensetzung besteht, in den absoluten Mengen erscheinen, noch gibt es irgendeine direkte Methode, um von den letzteren auf die ersteren schließen zu können. Die Zusammensetzung sollte wie ein Einzelvektor-Merkmal behandelt werden. Die vorsichtige Anwendung von angemessenen Methoden der Klassifikation und Analyse ist gefordert, wenn wir die Gründe für die Variation in der Harzzusammensetzung verstehen wollen.

1. Introduction

Resins are the "non-volatile products of plants", which exude naturally or can be obtained by incision or infection. Typically, resins are mixtures of compounds such as terpenoids, flavonoids and fatty acids. The terpenoids, comprising C_5 isoprene units, have been the most widely studied compounds, because they are easily identified and quantified by gas chromatography (GC). Terpenes are the subset of terpenoids which have exact multiples of the isoprene unit (*e.g.*, hemiterpenes, C_5; monoterpenes, C_{10}; sesquiterpenes, C_{15}; diterpenes, C_{20}). Essential oils are volatile oils recovered from plant tissue by steam distillation, pressure or solvent extraction which also contain terpenoids. The mixture of resin and essential oil that occurs in many plants is known as oleoresin (Dell and McComb, 1978), which we shall generally abbreviate to resin.

Initially, research on the composition of plant resins was motivated by the commercial importance of those of particular species, such as *Pinus elliottii* Engelm. in the southern United States of America (Roberts *et al.*, 1982). The analysis of the terpene composition of resin as a technique in forest genetics research was made possible by the development of GC, which accurately determines the composition of small samples (Simpson, 1970; Squillace, 1976). The topic has been the subject of numerous articles, a number of reviews (Lever and Burley, 1974; von Rudloff, 1975; Squillace, 1976; Burley and Lockhart, 1985) and conference proceedings (*e.g.*, Seal *et al.* 1977; Rudin, 1979). The purpose of this paper is to draw attention to some of the characteristics of data describing resin composition, and to the consequences for its interpretation. Much of the work based on resin compositional data must be considered suspect because it has not acknowledged the problem discussed below.

As background, we briefly outline the sources of resin in coniferous trees, what is known about resin synthesis,

[1]) Department of Forestry, Queensland, Australia.

what is suggested to control it and how GC results are reported. An appreciation of these topics is fundamental to subsequent discussion.

2. Sources of resin in coniferous trees

Conifers have been the most extensively studied group of forest trees presumably because of their widespread distribution and commercial importance. Coniferous resins may be obtained from distillation of the leaf oils (*e.g.*, ZAVARIN, 1970; VON RUDLOFF, 1975), from cortical tissue (*e.g.*, SQUILLACE, 1977; MEIER and GOGGANS, 1978; MARPEAU *et al.*, 1983), or from xylem tissue (*e.g.*, BANNISTER *et al.*, 1962; 1983; FRANKLIN, 1976; GREEN *et al.*, 1974; 1975; STRAUSS and CRITCHFIELD, 1982). Each of these sources samples independent systems (VON RUDLOFF, 1975).

Nevertheless, the compositions of foliar oils and cortical resins are often similar; the latter are more easily collected (SQUILLACE, 1977) and not subject to the complications associated with distillation (VON RUDLOFF, 1975), but both are commonly used. The composition of xylem resin is often less complex than that of cortical resin (SQUILLACE, 1976), but is easily collected, and is of additional interest because of its commercial value (*e.g.* ROBERTS *et al.*, 1982).

Within an individual tree, the composition of foliar oils may vary with season (VON RUDLOFF, 1975), and that of foliar oils and cortical resins with age of the tissue (SQUILLACE, 1976). The composition of xylem resin has been found to vary with height of sampling in *P. taeda* L. (ROCKWOOD, 1973) and *P. elliottii* (ROBERTS, 1970; FRANKLIN, 1976).

3. Gas chromatographic analysis and data reporting

GC is an efficient, accurate and precise method of determining resin composition (VON RUDLOFF, 1975; LOCKHART, 1985). The analytical results may be reported on a number of bases, defined by SQUILLACE (1976) for the case of monoterpenes as the content of each (1) relative to all others ("total monoterpene basis"), (2) relative to all other constituents of the oleoresin ("oleoresin basis"), or (3) per unit weight or volume of tissue ("tissue basis"). Data defined on the latter basis were referred to by WHITE (1983a) as the "absolute amounts "; there are other absolute measurements which could be defined. Thus, on the first two bases, the proportion of each terpene reported is the proportion of the total, by weight, which each represents; the sum of all constituent proportions totals unity. The results of analyses of the same sample on different bases will not be comparable because, as SQUILLACE (1976) demonstrates, different parameters are being assessed. Nor is there any simple transformation from one to the other, because each is dependent on the composition of the resin, which varies from sample to sample.

Although most authors detail the GC system used in their research, it is not clear that there is general appreciation that the area measured by the integrator is proportional to the weight of the resin constituents. The percentage weight of all monoterpenes is calculated by standardization using the total area of all monoterpenes; percentage weight of total oleoresin is calculated by standardization using the total area of all constituents. The percentage by weight on a monoterpene basis will differ slightly from the percentage by volume on the same basis, as the monoterpenes range in density from about 0.80 to 0.87 g cm^{-3}. The proportion by weight on an oleoresin basis will differ more widely from the proportion by volume on the same basis, as it also involves other constituents of varying densities — for example, common sesquiterpenes range in density from about 0.90 to 0.93 g cm^{-3} (WEAST, 1983).

The need for clarity about what the data represent is illustrated by the work of MEIER and GOGGANS (1978). They apparently intended to measure the actual volume of each monoterpene in a standard volume of resin, and correct application of their technique would have determined this.

What was measured is defined as follows. 4 ml of paramycene, as a standard, and 396 ml of hexane, as solvent, were mixed with 50 ml of resin. The integrator value for paramycene (A_p) is proportional to the weight (W_p) of 4 ml of paramycene. Therefore, $A_p \propto 4 \text{ ml} \times d_p$, where d_p is the density of paramycene. The integrator value for monoterpene X (A_x) is proportional to the weight (W_x) of this monoterpene in 50 ml of resin: $A_x \propto W_x$.

Therefore

$$(1) \quad \frac{A_x}{A_p} = \frac{W_x}{4\ d_p}$$

The value used by MEIER and GOGGANS (1978) in all their analyses was

$$(2) \quad \frac{4\ A_x}{A_p}$$

which is equal to

$$(3) \quad \frac{W_x}{d_p}$$

Because the composition of resin varies from tree to tree, so too do the density and weight of the standard volume of resin. Thus, the values used by MEIER and GOGGANS (1978) were the weights of monoterpenes in samples of oleoresin not standardised by weight. If the values had been correctly standardised they would be equivalent to percentage weights on an oleoresin basis. Alternatively, if the weights of the individual monoterpenes were converted to volumes by dividing each by its own density, the measurements would be equivalent to percentage volumes on an oleoresin basis. Neither of these procedures were undertaken and the results that follow are therefore meaningless.

MEIER and GOGGANS (1978) maintained that they were not measuring percentages. However their measurements would still be relative values if they had been correctly standardised. Their claim that they measured actual volumes and not percentages suggests that they believed, incorrectly, that they had measured absolute values.

Proportional data are independent of the total yield of terpenes, but data reported on a tissue basis are not (SQUILLACE, 1976). WHITE and NILSSON (1984) found that the yield of leaf oil monoterpenes in *P. contorta* DOUGL. was related to the frequency of resin canals in the leaves, which was itself variable and subject to some genetic control. They concluded that, as a result, proportional data should be used to study similarities and differences in populations. SHAW *et al.* (1982) cited the confounding of genetic and environmental effects on yield in reaching the same conclusion. Most studies of coniferous resin composition have been based on the interpretation of data comprising the proportions of individual terpenes detected in a sample of resin. We shall refer to such data as compositional data.

4. Variation in the composition of resin

Variation in the composition of plant resins occurs at all levels of the taxonomic hierarchy. IRVING and ADAMS

(1973) attributed inter- and intra-populational variation in resin composition to four major factors: "(1) individual genetic divergence; (2) the broad environmental conditions of the population and conditions prevailing at the time of sampling; (3) the ontogenetic and phenological stages; and (4) the techniques of extraction and analysis".

The premises underlying analyses of coniferous resin composition were established by the early studies of HANOVER (1966a; b; c), SQUILLACE and FISHER (1966) and SQUILLACE (1971). They are summarized by the statements that "the composition of monoterpenes in the oleoresin of conifers is strongly inherited and often controlled by single genes with major effects" (SQUILLACE, 1976), and is "subject to little environmental variation" (VON RUDLOFF and REHFELDT, 1980). That is, assuming that factors (3) and (4) above are constant, variation in resin composition represents genetic rather than environmental influences.

We suggest that the evidence for this assumption, which has motivated investigation of resin composition in forest genetic research, is not as conclusive as popularly believed. There are two principal reasons for our concern. The first is the poor understanding of the processes by which resin compounds are synthesized and the consequent difficulty involved in separating genetic and biosynthetic effects on resin composition (IRVING and ADAMS, 1973). The second, more important in the context of our discussion, is the nature of resin compositional data.

5. Biosynthetic and genetic influences on resin composition

The biosynthesis of monoterpenes was comprehensively reviewed by CORI (1983). Work to date suggests that the lower (mono- and sesqui-) terpenes are metabolically active, that is, they are not terminal products. Terpenoids are synthesized from mevalonic acid, which is itself formed from photosynthates transported to the site of resin synthesis (LOOMIS and CROTEAU, 1973). In conifers, these sites are the secretory cells adjacent to resin ducts (DELL and MCCOMB, 1978). WHITE's (1983a) work with foliar terpenes of *P. contorta* suggested that synthesis of particular compounds proceeds through a series of transition states, and is " governed by the levels of precursor and activity of enzymes". She suggested that the relative amounts of each terpene were determined by "factors which control enzyme concentration, compartmentation and kinetic properties".

In the absence of similar information specifically for xylem resin, it seems reasonable to assume that such processes also apply to the synthesis of its constituent terpenoids. However, the situation is complicated by the dependence of xylem resin composition on the height of sampling, which FRANKLIN (1976) suggested was a function of distance from the base of the green crown. This is consistent with a pattern of synthesis controlled by metabolic regulators which may diffuse or undergo transport from the crown towards the base of the tree (LARSON, 1960, in FRANKLIN, 1976).

The biosynthetic pathways by which the terpenoid constituents of resin are formed are poorly understood (WHITE, 1983a) because they are difficult to determine directly (LOOMIS and CROTEAU, 1973; MARTIN *et al.*, 1976; WHITE, 1983b). Consequently, investigations have sought to infer biosynthetic relationships from the analysis of compositional data (*e.g.*, ZAVARIN, 1970; IRVING and ADAMS, 1973; ZAVARIN and SNAJBERK, 1973; HILTUNEN, 1975; VON RUDLOFF and REHFELDT, 1980; WHITE, 1983a).

Interpretation of the correlations between individual constituents, which has been the basis of the these indirect approaches, is complicated by the use of proportional data, constrained to sum to unity. ZAVARIN (1970) suggested that the effect of this constraint was unimportant in a multicomponent system. This is incorrect, as subsequent evidence demonstrates. FORREST (1980) followed MOSIMANN's (1962) procedure of calculating the correlation due to constraint (see also SQUILLACE, 1976; SHAW *et al.*, 1982). FORREST found that when constituents were present in proportionally large amounts, most of the correlation was due to constraint, and where the correlation appeared to be real, it was "difficult to assess the size of the adjusted correlation." SQUILLACE (1976) took MOSIMANN's approach a step further, calculating a "net measure of distance". However, even with the example he provided, the biological interpretation of these distances is not apparent.

WHITE (1983a) calculated correlations based on both absolute and relative data, and found that correlations which were evident with the former were obscured by the latter. Correlations between proportions may contain useful information but it is necessary to give some thought to the possible processes which could occur in resin production in order to interpret them.

If the correlation between the absolute quantities of two constituents is one, then the correlation is also one for the proportions derived from any mixture of which the two constituents form a part. Investigation of the ratios of proportions for pairs of constituents would also reveal this phenomenon. The ratio of such a pair of constituents would be constant in all samples of resin.

An example of a possible biosynthetic process which resulted in correlations of one between constituents would be the production from a single precursor of a number of constituents always in the same ratio to each other by weight or volume. The correlation between the values of any pair of constituents is always one, whether between absolute quantities or between proportions derived from a mixture of which the two constituents form a part.

In other situations, when the correlation between two constituents is not one, correlations are not the same for percentages as they would be for absolute quantities. We do not recommend application of MOSIMANN's (1962) correction for constraint. MOSIMANN derived his theory for a multinomial distribution compounded with a multivariate beta-distribution. This does not apply to resin compositional data, which are better described as continuous proportions (STEPHENS, 1982). A correction for constraint applied in this way can only confuse, and would indeed conceal the correlations in the above example.

The biosynthetic processes outlined above are complex and the final composition of resin depends on many factors (CORI, 1983). Although an understanding of them is necessary to determine the meaning of differences in terpene composition (IRVING and ADAMS, 1973; WHITE, 1983b), they have not been extensively studied in forest trees, and genetic interpretations of such differences have proceeded in their absence.

6. Genetic interpretation of differences in resin composition

Due to the premise outlined in section 4 above, "the genetic control of monoterpenes in conifers has received more attention than their biosynthesis" (WHITE, 1983a). One of the topics which has received considerable attention is that of monogenic control of particular resin constituents.

Table 1. — Mono- and sesqui-terpenes for which monogenic control has been proposed.

Species (type of resin)	Terpene	Dominant class (if any)	Author	Definition of class (%)			Constraint (if any)
				Low	Intermediate	High	
Picea abies (cortical)	α- + β-pinene	low	Esteban et al. (1976)	13-39		48-93	
	& δ-3-carene	high		0-8		43-71	
Pinus attenuata x Pinus radiata (xylem)	α- or β-pinene	---	Forde (1964)				
	β-pinene	high	Strauss & Critchfield (1982)[1,2]	0-12		17-87	
		---		0-12	17-50	51-87	
Pinus elliottii (cortical)	β-pinene	high	Squillace (1971)[3]	2-8		21-74	
			Squillace (1977)	0-4		≥8	high limonene
				0-8		≥12	low limonene
	myrcene	high	Squillace (1971)[3]	0-5		6-45	
			Squillace (1977)	0-6		≥9	high limonene
				0-4		≥7	low limonene
	limonene	high	Squillace (1977)	0-8		≥16	
	β-phellandrene	high	Squillace (1977)	0-2		≥4	
Pinus monticola (cortical)	δ-3-carene	high	Hanover (1966c)[1]	0-<2.41		>2.41-15.5	
Pinus pinaster (cortical)	δ-3-carene	high	Baradat et al. (1972)[1,3]	0		20-50	
	myrcene	high	Baradat et al. (1975)[1,3]	0-<6		>6-40	
	caryophyllene	low	Marpeau et al. (1975)[1]	0-<6		>6-23	
	longifolene	high	Marpeau et al. (1975)[1]	0-<6		>6-19	
	limonene	---	Marpeau et al. (1983)[1]	0-<2.88	2.88-<11.9	11.9-25.8	
Pinus sylvestris (foliar)	δ-3-carene	low	Hiltunen et al. (1975)	<10		≥10	
(cortical)	δ-3-carene	---	Yadzani et al. (1982)[1]	0-<10	10-<60	≥60	
	β-pinene	low	Yadzani et al. (1982)[1]	0-<18		≥18	
	β-phellandrene	high	Yadzani et al. (1982)[1]	0-<4		4-45	
	limonene	low	Yadzani et al. (1982)[1]	0-<3	3-<28	≥28	
	myrcene	low	Yadzani et al. (1982)[1]	0-<10		≥10	
Pinus taeda (cortical)	limonene	low	Rockwood (1973)[3]	0-<6		6-61	
		high	Squillace et al.[3] (1980)[1]	0-6		8-62	
	myrcene	high	Rockwood (1973)[3]	0-<6		6-63	
		high	Squillace et al. (1980)[1]	$0-(8-x_1)$[4]		$(8-x_1)-62$[4]	limonene proportion
	β-pinene	high	Squillace et al. (1980)[1]	$0-(5-x_2)$[5]		$(5-x_2)-63$[5]	limonene + β-phellandrene proportion
	β-phellandrene	high	Squillace et al. (1980)[1]	0-1		3-58	
Pinus virginiana (cortical)	δ-3-carene	high	Meier & Goggans (1982)[1]	0-<2.41		>2.41	
Pseudotsuga menziesii (foliar)	camphene group	high	von Rudloff & Rehfeldt (1980)	not specified			

[1]) Monogenic control of the proportion of the constituent proposed.
[2]) Chi-squared tests should be interpreted with caution.
[3]) Heritability value calculated.
[4]) x_1 is a variable dependent on the proportion of limonene.
[5]) x_2 is a variable dependent on the proportion of limonene + β-phellandrene.

Such control was first hypothesized for an interspecific hybrid by Forde (1964) and within a species by Hanover (1966c). The constituents and species for which monogenic control has been proposed are listed in *Table 1*.

Monogenic control of particular resin constituents has been inferred from what White (1983b) referred to as "segregation studies". In these, the frequency distribution of phenotypes — the proportions of a particular resin con-

stituent — is divided into classes, each representing the action of a particular allelic combination. A qualitative trait, class, is thereby being inferred from a quantitative one, proportion. This was FORDE's (1964) approach, but there are many qualifications to her data which, as she notes, do not favour close analysis. Before we discuss subsequent work based on frequency distributions, we shall consider the proposals of HANOVER (1966c) and MEIER and GOGGANS (1982), which were not based on frequency distributions.

6.1 Monogenic control of δ-3-carene in P. monticola and P. virginiana

HANOVER (1966c) hypothesized that the concentration of the monoterpene (δ-)3-carene in the cortical resin of *P. monticola* DOUGL. was "controlled primarily by single dominant and recessive alleles at a locus". His data, the proportion of δ-3-carene in the total oleoresin of 231 seedlings, were derived from 28 full-sib families, the product of matings between seven female and four male parents. The basis of his assertion that "gene segregation for different levels of this terpene is apparent from these data" is not evident. Nor is there any but an arbitrary basis for his classification of δ-3-carene concentration level as low or high. The value of 2.41% which he used to differentiate between classes appears to have been chosen solely because it gave the best fit to his theory. The frequency distribution of δ-3-carene concentration derived from his data (*Fig. 1*) is not bimodal, nor does the value 2.41% have a relationship to any salient feature of the distribution.

There are also inconsistencies with the data presented as his *Table 1*. A parent was classified as homozygous dominant (CC) for the gene controlling δ-3-carene concentration if its progeny were all in the high class, and homozygous recessive (cc) if its progeny were all in the low class; if there was segregation in at least one family, the parent was classified as heterozygous (Cc). HANOVER claimed that "phenotypic values for each of the parents substantiate this assignment". Thus, female parent 181 is classified as cc on the basis of its progeny with male parent 17, also classified cc. However, the phenotypic value of parent 181 is 2.90%, within the high class and therefore apparently Cc or CC. Two offspring of female 214 (cc) crossed with male 17 (cc) have a δ-3-carene concentration within the high class, which is not possible given the parental classification. Expectation and observation cannot be reconciled in these cases.

In the latter case, it is inappropriate to calculate the Chi-squared statistic (his *Table 2*) to test whether the observed ratios differ from those expected. The question here is not whether the ratio approaches theoretical expectation, but whether it is consistent with the theory. A similar issue is raised by the offspring of male parent 17. Hanover suggested that an inhibiting factor, causing "a complete block in 3-carene synthesis", is responsible for the lack of δ-3-carene in the parent. However, only one of its 56 offspring appears to have inherited this inhibiting factor. No explanation of these complications was offered.

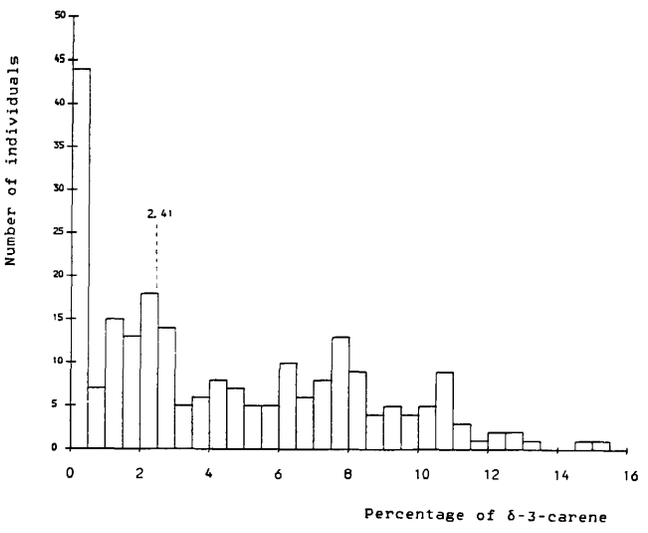

Fig. 1. — Frequency distribution of the percentage of δ-3-carene in cortical oleoresin of *Pinus monticola* (derived from Hanover, 1966c).

The differences between theory, based on an arbitrary class ification, and observation are serious and unresolved. One must conclude that, contrary to the assertion of BARADAT *et al.* (1972) that "les conclusions de HANOVER sur le contrôle monogenétique du 3-carène chez *Pinus monticola* sont irréfutables", the data do not suggest that the concentration of 3-carene in the cortical resin of this species is subject to monogenic control.

MEIER and GOGGANS (1978) proposed that the concentration of δ-3-carene in the cortical resin of *P. virginiana* MILL. was also subject to monogenic control, "under the assumption that high concentration is dominant and that the frequency of the dominant gene is low". They misinterpreted HANOVER's (1966c) criterion for classification, believing it to be based on the proportion of total monoterpenes rather than on the proportion of total oleoresin. The results they presented do not allow derivation of a frequency distribution, but their arbitrary application of an inappropriate criterion to data expressed on a different basis for another species does not engender confidence in their reults. Indeed, the converse proposal, that low concentration of δ-3-carene is dominant over high and that the frequency of the dominant gene is low, is equally plausible from their *Table 5*.

Table 2. — Heritability estimates with standard errors for monoterpene proportions in *Pinus monticola* (from HANOVER, 1966a, *Figure 3*).

Terpene	Narrow sense heritability and standard error estimated from regression on			Broad sense heritability
	male parent	female parent	mid-parent	
α-pinene	0.64 (1.9)	0.68 (1.6)	0.73 (1.3)	0.63
camphene	-0.01 (.05)	0.03 (.06)	-0.001 (.05)	0.38
β-pinene	0.48 (1.1)	1.02 (0.7)	1.21 (0.6)	0.86
δ-3-carene	1.11 (0.6)	0.64 (0.8)	1.01 (0.3)	0.95
limonene	0.29 (1.4)	0.96 (0.5)	1.15 (0.5)	0.87
unknown	0.90 (0.6)	0.57 (0.9)	0.89 (0.9)	0.80

6.2 Monogenic control of particular resin constituents inferred from frequency distributions

The first question raised by many of these studies is the definition of the characteristic which is subject to monogenic control. There are two possibilities — either the absolute quantity or the relative quantity of a particular resin constituent. Although all of the work reported in *Table 1* has been based on proportional data — as indeed it must be (WHITE and NILSSON, 1984) — authors have been vague in their descriptions of the trait for which they were proposing monogenic control. For example, HANOVER (1966c) and YAZDANI *et al.* (1982) use "level" of a terpene; SQUILLACE (1971) refers to "amount" or "quantity"; FORDE (1964), HILTUNEN *et al.* (1975) and SQUILLACE *et al.* (1980) refer to "content". This imprecision of terminology obscures whether they are proposing that absolute or relative quantities of the terpene are subject to monogenic control; those authors who have specifically proposed that proportions of a particular resin constituent were subject to such control are identified in *Table 1*.

Because these terms have been confused, and because some authors proceed as though dealing with absolute quantities, we believe that some have not differentiated between relative and absolute quantities. Therefore we shall consider both possibilities.

6.2.1 Monogenic control of the absolute quantity of a particular resin constituent

The frequency distributions for the proportions of each monoterpene constituent from a sample of trees are compound distributions, dependent on the frequency distributions of the absolute quantities of each of the other constituents of the resin. If the absolute quantity of a terpene were under monogenic control, it is possible to predict the effect that this would have on the frequency distribution of the proportions of that terpene. The only way of using resin compositional data to investigate control of the absolute quantity of a terpene is to then investigate whether the frequency distribution of the former displays the feature predicted. We discuss below some simple models of resin production as a prelude to presentation of a simple model of monogenic control of the absolute quantity of a terpene.

As a simple example, consider a hypothetical system of two constituents. In the population of trees, the production of terpene A is measured by the variable x_1, and that of terpene B is measured by the variable x_2. Assume that, for this populations, x_1 is constant at k_1 units and x_2 is constant at k_2 units. The frequency distributions of x_1 and x_2 would be delta-functions. The frequency distribution of p_1, the proportion of the total that is A, is a delta function at $p_1 = k_1/(k_1 + k_2)$, and that of p_2, the proportion of the total that is B, is a delta-function at $p_2 = k_2/(k_1 + k_2)$. If, in another population of trees, k_2 were different but k_1 the same, both p_1 and p_2 would nevertheless be altered. Therefore, statements such as that of STRAUSS and CRITCHFIELD (1982), "that changes we measured in relative monoterpene levels reflect changes in production of both α- and β-pinene, and not simply a change in one constituent without a change on the part of the other", are incorrect.

Consider also a more complex situation, where x_1 remains constant at k_1 units, but x_2 has a normal distribution in the population of trees, possibly as a result of polygenic control, with mean μ and variance σ^2. The analysis of resin from a tree in the population yields $p_1 = x_1/(x_1 + x_2)$ and $p_2 = x_2/(x_1 + x_2)$. The probability distributions of p_1 and p_2, $f(p_1)$ and $g(p_2)$ can be derived thus:

$$(4) \quad f(p_1) = \frac{k}{\sigma p_1^2 (2\pi)^{\frac{1}{2}}} \exp\left[\frac{-(k(1/p_1 - 1) - \mu)^2}{2\sigma^2}\right]$$

$$(5) \quad g(p_2) = \frac{k}{\sigma(1-p_2)^2 (2\pi)^{\frac{1}{2}}} \exp\left[\frac{-(kp_2/(1-p_2) - \mu)^2}{2\sigma^2}\right]$$

These distributions are both unimodal and skewed. The mode value of p_1 is $k_1/(\mu + k_1)$, and that of p_2 is $\mu/(\mu + k_1)$. Thus, p_1 varies for this population of trees although x_1 is constant. It is impossible to deduce, from the information in $f(p_1)$, that all tree were producing equal quantities of A. *Fig. 2* shows the distributions of $f(p_1)$ and $g(p_2)$ for an example with $k = 2$, $\mu = 5$ and $\sigma^2 = 1$; other examples exhibit similarly skewed distributions.

If both x_1 and x_2 have normal distributions, with means of μ_1 and μ_2 respectively, the mode of $f(p_1)$ would be $\mu_1/(\mu_1 + \mu_2)$, and that of $g(p_2)$ would be $\mu_2/(\mu_1 + \mu_2)$. The distributions of p_1 and p_2 are then unimodal and skewed, with variances dependent on the variances of x_1 and x_2. As in the first example above, it is impossible to determine the quantities of A and B produced from the distributions of p_1 and p_2.

It is now instructive to consider some simplified models, incorporating monogenic control, of the production of constituents of a resin. In the first model, the production of two constituents is simplified to a constant over the population, and production of a third is assumed to be under the control of a dominant/recessive pair of alleles at a single locus. Consider A, B and C as the three constituents of the resin; the quantities of each are measured as x_1, x_2 and x_3 respectively. Assume that A and B are produced in constant quantities, so that x_1 and x_2 equal k_1 and k_2 respectively. The production of C is subject to monogenic control; x_3 has probability α of being k_3 and probability $(1-\alpha)$ of being k_4. The frequency distributions of x_1 and x_2 would each have one peak, and that of x_3 would have two peaks. The frequency distributions of p_1, p_2 and p_3, the proportions of A, B and C would all have two peaks, the areas under which would be in the ratio $\alpha/(1-\alpha)$. The

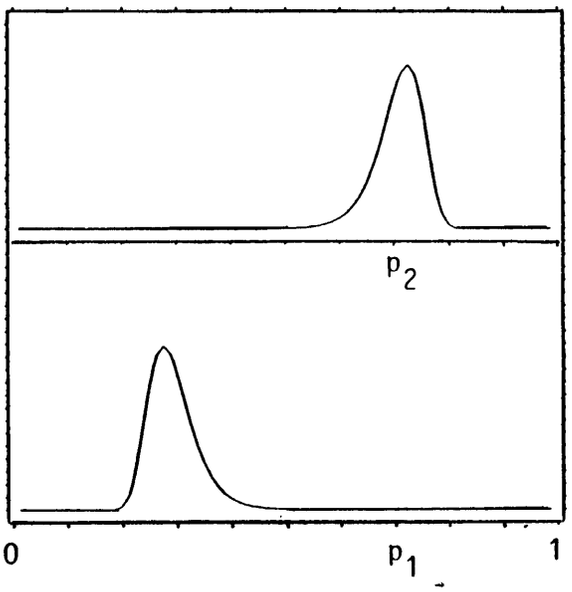

Fig. 2. — The probability distributions of p_1 ($f(p_1)$), and p_2 ($g(p_2)$).

modes of p1 would occur at $k_1/(k_1 + k_2 + k_3)$ and $k_1/(k_1 + k_2 + k_4)$, those of p_2 at $k_2/(k_1 + k_2 + k_3)$ and $k_2/(k_1 + k_2 + k_4)$, and those of p_3 at $k_3/(k_1 + k_2 + k_3)$ and $k_4/(k_1 + k_2 + k_4)$.

Thus, although the frequency distribution of the quantity of only one constituent is bimodal, the frequency distributions of the proportions of all three constituents are bimodal. As with the preceding example, it is not possible to deduce the frequency distributions of x_1, x_2 or x_3 with information only about the frequency distributions of p_1, p_2 and p_3.

As a second example, consider the case where two constituents are subject to monogenic control in the same manner as C above. Then, the frequency distributions for each of p_1, p_2 and p_3 will each show four peaks. An example will demonstrate how complex the situation is becoming. Assume the following:

production of A is constant, $x_1 = 5$ units;

production of B is variable,

$x_2 = 0$ units with probability 0.25
$= 20$ units with probability 0.75;

production of C is variable,

$x_3 = 10$ units with probability 0.50
$= 60$ units with probability 0.50.

Then

$p_1 = 5/15, p_2 = 0, p_3 = 10/65$ with probability 0.125
$p_1 = 5/65, p_2 = 0, p_3 = 60/65$ with probability 0.125
$p_1 = 5/35, p_2 = 20/35, p_3 = 10/25$ with probability 0.375
$p_1 = 5/35, p_2 = 20/85, p_3 = 60/85$ with probability 0.375

If we propose multi-genic control of the production of individual constituents, the corresponding frequency distributions of absolute quantities for the population of trees would be better approximated by normal distributions. The peaks of each constituent in the above example would be broadened and skewed; some peaks would be amalgamated.

These examples give some insight into how the frequency distribution of the proportion of a particular resin constituent of a system depends on the occurrence of the other constituents. Because these frequency distributions are compound, it is not possible to deduce from them the absolute quantity of a particular constituent that a tree is producing. The examples demonstrate how misleading such deductions could be..

6.2.2 Monogenic control of the proportion of a particular resin constituent

The proposal of monogenic control of the proportion of a particular resin constituent implies that, whatever the quantity of resin or total terpenes produced by a tree, the proportion that a particular resin constituent comprises may be allocated to a particular class. If one of the alleles is dominant, there are only two classes, low and high; if not, there is also an intermediate class. Dominance has been proposed for all but two of the cases in *Table 1*.

6.2.2.1 Interpretation of frequency distributions

"Bimodal or skewed frequency distributions ... are often taken as evidence of control by few genes with large effects" (Squillace, 1976). However, because frequency distributions of proportional data are compound, the distributions for particular resin constituents are dependent on those of other constituents. Thus, bimodality in particular frequency distributions may be solely due to constraint (*e.g.*, Gansel and Squillace, 1976; Squillace, 1976; White, 1984). In such circumstances there is no *a priori* reason for nominating any particular constituent as that which is subject to monogenic control.

It is difficult to account for these complications when considering a particular resin constituent in isolation. Consider the work of Yazdani *et al.* (1982), whose proposal of monogenic control for β-pinene was based on the segregation pattern of offspring of the only clone to contain a high "quantity" (by which they mean "proportion") of β-pinene and high proportions of δ-3-carene, or *vice versa* their studies of the mode of inheritance of other constituents because, in those cases, "the pattern of segregation looks very peculiar", that is, it does not agree with their hypothesis. They suggest that "the high content of β-pinene somewhat complicates the picture of interdependence for other monoterpene components", and infer monogenic control of myrcene, limonene and δ-3-carene from crosses whose resin did not contain a high proportion of β-pinene.

There is a number of possible explanations of this result; they are not mutually exclusive. One is that the mode of inheritance differs in trees with different proportions of particular constituents; this is not in accord with the theory of monogenic control. Another possibility, in view of the above discussion, is that frequency distributions of proportions are not necessarily amenable to such interpretations. This example demonstrates that the proportions of constituents purportedly under monogenic control are unable to fulfil Strauss and Critchfield's (1982) expectation to "behave relatively independently of other components". The interpretation of frequency distributions of proportions of particular resin constituents is therefore inherently prone to misinterpretation.

One cannot but wonder whether such misinterpretation led to Esteban *et al.*'s (1976) proposal that "one and the same gene locus" is responsible for low proportions of α- plus β-pinene and high proportions of δ-3-carene, or *vice versa* (the column titles for the former in their *Table 3* should be reversed). They initially propose monogenic control of δ-3-carene only, but subsequently extend their proposal to include, firstly, α-pinene and, finally, the combined proportion of α- and β-pinene. As either the combined pinene proportion or that of δ-3-carene exceeds 50% in each of the clones tested, and as no biosythetic evidence is presented to suggest a relationship between the pinenes and δ-3-carene, it is possible that the level of one is merely responding to the level of the other because of constraint.

6.2.2.2 Definition of qualitative classes

"True qualitative traits are characterized by distinct types with little or no connection by intermediates" (Falconer, 1981). Thus, one expects that discontinuity between the classes should be evident in frequency distributions of proportions from which monogenic control is inferred. Such has not always been the case; consider, for example, the arbitrary classification of β-pinene by Squillace *et al.* (1980) or Yazdani *et al.* (1982), in their *Figures 2* and *1* respectively. Definitions between classes of particular constituents have sometimes been made conditional on the class of others (Squillace, 1977; Squillace *et al.*, 1980) without any biosynthetic explanation, or have been enhanced by omission of individuals perceived to be intermediate. The classification of β-pinene, myrcene, limonene and β-phellandrene by Squillace (1977) and of limonene and β-phellandrene by Squillace *et al.* (1980) omits individuals with proportions of the respective constituent between the classes defined in *Table 1*. Even so, Squillace *et al.* (1980) attribute some of the inconsistencies in the segregation ratios for β-phellandrene "to the difficulty of classification".

Such arbitrary classification and exclusion, without explanation, do not engender faith in the theory proposed for the remaining individuals — as WHITE (1983b) noted, "the interpretation of the mechanism of genetic control depends on how the segregation classes are defined". As we demonstrated in Section 6.1, arbitrary classification is not a firm basis from which to propose monogenic control of the proportion of a particular resin constituent.

6.2.2.3 The range within qualitative classes

A feature of all the classifications reported in *Table 1* is the wide range of phenotypes outside — and sometimes within — the low class. If monogenic control with dominance at a particular locus is proposed, the only variation from two distinct values would be due to non-genetic effects. Environmental effects on resin composition are reportedly small (HANOVER, 1966a; SQUILLACE and FISHER, 1966; GANSEL and SQUILLACE, 1976); therefore, wide ranges within the high class suggest strong effects by other genes (HILTUNEN *et al.*, 1975; WHITE, 1984). In some cases listed in *Table 1*, intermediate classes representing the effect of heterozygotes have been proposed, although their definition appears entirely arbitrary and therefore subject to the concern noted above.

The likelihood that monogenic control is a simplistic interpretation of the mode of genetic control is enhanced by such wide ranges. STRAUSS and CRITCHFIELD (1982) and WHITE (1984) fitted both single and two locus models of gene control to their data, and were unable to choose between them. It seems likely that their findings would be repeated if two locus models were fitted to the other data listed in *Table 1;* as WHITE (1983b) commented, "segregation of progeny into classes appropriate for a one-locus model of genetic control does not exclude the possibility of two (or more) locus control".

6.2.2.4 Tests of segregation

In studies where the segregation of progeny into the qualitative classes of high and low is assessed, the "goodness of fit" between observed and expected ratios is tested against the Chi-squared distribution. The approximation of the test statistic to this distribution is good only for large sample sizes; the test should not be applied where class expectations are less than one, and used only with caution when they are less than five (SNEDECOR and COCHRAN, 1980; MEAD and CURNOW, 1983).

The segregation tests of progeny which have been undertaken to investigate proposals of monogenic control of the proportion of particular resin constituents (*Table 1*) have often been based on too few crosses to be statistically reliable. For example, SQUILLACE *et al.* (1980) note that the evidence they present to contradict ROCKWOOD's (1973) proposal of the dominant class of limonene "cannot be considered conclusive". Similarly, the expected segregation ratios of the five progeny of each of crosses 5 and 11 in STRAUSS and CRITCHFIELD's (1982) *Table 1* are both 3.75:1.25; the corresponding observed ratios are 2:3 and 4:1. Both are judged in accord with expectation by the Chi-squared test. So, too, are the observed ratios of the progeny of crosses 2 and 7, which are 3:9 and 10:0 respectively, corresponding to expected ratios of 6:6 and 7.5:2.5. Common sense would place little faith in such results. It is apparent that other observed ratios could be judged similarly acceptable by the Chi-squared test in such cases, demonstrating that a larger number of progeny is required for any meaningful interpretation. More judicious use of the test is exemplified by the approach of YAZDANI *et al.* (1982).

The failure of self-fertilized or outcrossed progeny to fulfil theoretical expectations is not infrequent, and is usually attributed, without specific justification, to pollen contamination (STRAUSS and CRITCHFIELD, 1982; YAZDANI *et al.*, 1982) or experimental errors (SQUILLACE *et al.*, 1980; STRAUSS and CRITCHFIELD, 1982). More imaginative explanations are required when there are major discrepancies between theory and observation; as well as the difficulties noted previously, YAZDANI *et al.* (1982) found that the proportions of β-phellandrene, myrcene, limonene and δ-3-carene in many self-ed progeny were "unexpected". They attributed this to "the presence of recessive lethals" linked to each of the genes controlling the proportion of each of the constituents. In proposing this complicated mechanism, they fail to consider that all segregation ratios would be affected by the presence of such genes for any particular constituent.

Results such as these echo the suggestions of, for example, HANOVER (1966c) and HILTUNEN *et al.* (1975) that additional modifying genes are implicated in cases where monogenic control has been proposed. One would hope that such inconsistencies would prompt further investigation and, if necessary, reappraisal of the theory. The contrary view, that theory prevails, has been preferred.

6.3 Monogenic control of particular resin constituents: more in hope than in confidence?

MOSIMANN (1962) applied his method of adjusting correlations for the presence of a constraint on the values data could take to a particular example "more in hope than in confidence". This seems also an appropriate summary of the proposals of monogenic control of particular resin constituents. Proposals that proportions of particular resin constituents are subject to monogenic control are qualified by the difficulties of interpreting compound distributions, arbitrary classification of classes, and inconsistencies in experimental results. When models of monogenic control of the proportions of particular resin constituents incorporate dependence on the effects of other genes, the term monogenic control is no longer appropriate. In such cases, monogenic control of the absolute quantity of a particular resin constituent becomes the more plausible theory. However, as we demonstrated above, interpretations of the frequency distributions of proportions for this purpose are fraught with difficulty.

More complex models of genetic control have been demonstrated to be as plausible as models based on single genes. Given the complications noted above, it is instructive to consider the concluding remarks of IRVING and ADAMS (1973), who investigated the "genetic and biosynthetic relationships of monoterpenes" in three taxa of the mint genus *Hedeoma*. They were able to use F_1 and F_2 generation crosses in a more comprehensive study than has been possible with forest trees. Although their data suggested involvement of a few major genes, they cautioned against oversimplification: "pervading our own data are indications of more complex levels of gene and enzymatic action. Epistasis, modifying genes, and enzyme complementation, to which we have alluded, are all characteristic of multigenic systems".

6.4 Quantitative interpretations of resin composition

Most interpretations of resin composition have not acknowledged the conceptual difference between compositional data, comprising a set of proportions, p_i, $i = 1, \ldots n$, $\Sigma p_i = 1$, and a multivariate data set, $x_i, \ldots x_n$, where the same constraint does not apply. For example, it is incorrect to

state that "small amounts of camphene, limonene and α-phellandrene occurred in most trees" (SQUILLACE, 1971) when only relative amounts have been measured. Although they may only comprise a small proportion of the resin, some trees may have been producing large amounts of these constituents relative to other trees. No form of analysis will reveal whether there are differences between observations in the quantity of a particular resin constituent because that has not been measured; what has been assessed is the quantity of a particular resin constituent relative to whatever else is being produced by the tree.

Therefore, it is not possible to make comparisons between observations of the proportions of particular constituents if a comparison of quantities is required. These proportions are not metric traits in the usual sense, neither are they measured independently. Methods of analyzing metric traits, such as analysis of variance, are therefore invalid if applied to such data. Furthermore, a succession of analyses of variance on the individual constituents do not yield meaningful results. Nor can the high, intermediate or low level of a proportion be considered a metric trait. The consequences for the application of quantitative genetic methods are discussed below.

6.4.1 Quantitative genetics and resin composition

Before proposing the theory of monogenic control, HANOVER (1966a) applied the methods of quantitative genetics to investigate genetic control of monoterpene composition, using the proportion of a particular resin constituent as though it were a metric trait. Subsequent workers have followed suit; usually, quantitative modes of inheritance have been postulated for the other constituents of the resin in the cases listed in *Table 1*.

Classical methods of quantitative genetics depend on the assumption that the frequency distribution of a metric trait approximates the normal (FALCONER, 1981). As we are considering neither metric traits nor traits which have a normal distribution, the application of quantitative genetic methods to resin compositional data is inappropriate and the results of such analyses (*e.g.*, HANOVER, 1966a; 1971; FRANKLIN and SNYDER, 1971; SQUILLACE, 1971; BARADAT *et al.*, 1972; HILTUNEN, 1975; THOR *et al.*, 1976; ADAMS, 1986) are invalid.

Inconsistencies in the results of such analyses illustrate this. For example, HANOVER (1966a) was unable to offer a genetic interpretation of many of his results: "it is difficult to understand the genetic basis for higher or lower terpene levels in self-fertilized progeny than in both their parents and the F_1. This unique situation occurred in α-pinene (higher) and 3-carene (lower)". The broad sense heritabilities which he calculated from analysis of variance, and the narrow sense heritabilities from the regressions of F_1 progeny means on both parental and mid-parent values, are not consistent, despite his assertion to the contrary. As is evident in *Table 2*, the various estimates of narrow sense heritability show little agreement, and are in most cases greater than the corresponding estimates of broad sense heritability, which is theoretically impossible.

The calculation of heritabilities for constituents purportedly under monogenic control (*Table 1*) demonstrates a fundamental misunderstanding of the nature of the proposal, presumably resulting from confusion between the quantitative measurement of proportion and the qualitative trait deduced from it. Calculation of heritabilities for these constituents has as much meaning as it would for the colour and shape of Mendel's peas. The results of calculations for quantitatively inherited constituents do not accord with the theory of quantitative genetics because the underlying assumptions are violated. Conclusions about the relative influences of genotype and environment on resin composition which derive from the application of such methods are therefore invalid.

Resin composition should be viewed as a single vector trait, and comparisons should be based on the complete composition. There are appropriate quantitative methods to aid interpretation of such data, but these have been infrequently applied to resin compositional data. A distinction must be drawn between two different aims and the different methods involved in each. The first purpose is that of classification, without calculation of a test statistic. The second is to test hypotheses about groups of observations. In this case, inferences are based on an assumed distribution of both the test statistic and the data set.

6.4.2 Classificatory approaches to interpretation of resin composition

Generally, classificatory approaches are concerned with grouping or reducing the dimensionality of data. Common techniques such as principal components analysis (PCA), cluster analysis and multidimensional scaling are well described elsewhere (*e.g.*, CHATFIELD and COLLINS, 1980). The former two techniques were used by BURLEY and GREEN (1977; 1979), and the former by SHAW *et al.* (1982), to investigate provenance variation in *Pinus* species. Contrary to the latter's assertion, PCA does not make distributional assumptions about the input variables, "though more meaning can generally be given to the components in the case where the observations are assumed to be multivariate normal" (CHATFIELD and COLLINS, 1980). Factor analysis is another classificatory technique which has been applied to resin compositional data (HILTUNEN, 1975). However, the assumptions inherent in the technique (CHATFIELD and COLLINS, 1980) are unlikely to be satisfied by such data.

Because PCA is dependent on scale (CHATFIELD and COLLINS, 1980), the principal components from resin compositional data will derive from the major constituents. For example, the constituents which BURLEY and GREEN (1977) identify as contributing most to the first component are those found in the greatest proportions — α-pinene and δ-3-carene for *P. oocarpa* SCHIEDE, and α-pinene and β-phellandrene for *P. caribaea* MORELET (GREEN *et al.*, 1974; 1975). Consequently, there is little point in undertaking PCA of the raw data. SHAW *et al.* (1982) use the angular transformation, which is appropriate for proportions arising from a binomial distribution, but not for "continuous proportions" (STEPHENS, 1982) such as those we are considering here. AITCHISON (1982; 1983a; b; c) examined the application of PCA to compositional data sets. His suggestion that a transformation based on logarithms might be the most appropriate has yet to be theoretically validated, and cannot be simply applied to many resin compositions because of the frequency of zero values. Two further difficulties are noted by CHATFIELD and COLLINS (1980). Firstly, principal components "are not invariant under linear transformation of the variables" and "are therefore not a unique characteristic of the data". Secondly, because PCA assumes no underlying statistical model, "it is difficult to compare the different components which result from carrying out a PCA on two or more different samples of the same type."

Another appropriate technique is that developed by ESTEBAN *et al.* (1976), specifically to differentiate groups based on their resin composition. Their "interval-method",

which considers the terpene composition as a stochastic vector, proved useful with their clonal data; when phenotypic ranges are greater, little differentiation is possible (KANOWSKI, 1986). The major concern with all these methods is that they should be robust. Different methods, or application of the same methods to slightly different data sets, should not produce contradictory results.

6.4.3 Testing hypotheses about the variation of resin composition

Appropriate statistical analysis of proportional data is a matter of considerable debate (AITCHISON, 1982). The distributions of proportions cannot be assumed to be multivariate normal, or even approximate to it. The search for appropriate transformations continues (AITCHISON, 1982); in the interim, attempts to transform data to approximate such a distribution (*e.g.*, SHAW *et al.* (1982)) have been unsuccessful, and the application of techniques which assume multivariate normality (*e.g.*, multivariate analysis of variance, discriminant analysis, canonical variate analysis) are therefore inappropriate.

An approach suggested by CHATFIELD and COLLINS (1980) and applied by SHAW *et al.* (1982) is to use the component scores in subsequent analyses. The results from such analyses must be interpreted with caution, firstly because the scores depend on the transformation applied, and secondly because their actual meaning may be obscure.

Where the data originate from a designed experiment methods developed from WATSON and WILLIAMS' (1956) analysis of variance on a von Mises distribution may be used. STEPHENS (1982) developed the approach for a two-way classification.

7. Conclusions

Absolute determination of resin composition defies definition and measurement; the only assessment possible is of the proportions of different constituents in a sample of resin, which can be determined accurately and precisely by GC. The properties of this compositional data have obscured the nature of "the gap between the gene and its expression" (IRVING and ADAMS, 1973); indeed, AITCHISON (1983b) suggested that "there has probably been no other form of data analysis where more confusion reigns and where improper or inadequate statistical methods are applied".

The premises which have been fundamental to the interpretation of variation in resin composition have not been justified quantitatively. Proposals for the mode of genetic control of resin composition and its magnitude relative to other influences are similarly unsubstantiated. The utility of resin composition to forest genetics research has yet to be confirmed. The first step in doing so is "to abandon the use of 'standard' methods quite inappropriate to 'nonstandard' data sets such as compositions" (AITCHISON, 1983b). Consideration of resin composition as a single vector trait, and the application of appropriate quantitative methods, such as those to which we have referred, are necessary to determine the extent to which resin phenotype reflects genotype.

Acknowledgements

We thank Drs RICHARD BARNES, JEFFERY BURLEY and COLIN MATHESON for constructive criticism of the manuscript. The work was funded by the Overseas Development Administration of the British Government, under Research Scheme R3881 and by the Conservator of Forests, Queensland, Australia, at the Oxford Forestry Institute, Oxford University.

References

ADAMS, R. P.: Geographic variation in *Juniperus silicicola* and *J. virginiana* of the southeastern United States: multivariate analyses of morphology and terpenoids. Taxon **35**: 61—75, (1986). — AITCHISON, J.: The statistical analysis of compositional data (with discussion). J. Royal Statistical Society, Series B. **44**, 139—177, (1982). AITCHISON, J.: Principal component analysis of compositional data. Biometrika **70**, 57—65, (1983 a). — AITCHISON, J.: The statistical analysis of geochemical compositions. Mathematical Geology **16**, 531—564, (1983 b). — AITCHISON, J.: Reducing the dimensionality of compositional data sets. Mathematical Geology **16**, 617—635, (1983 c). — BANNISTER, M. H. and MCDONALD, I. R. C.: Turpentine composition of the pines of Guadalupe and Cedros Islands, Baja, California. New Zealand J. Botany **21**, 373—377, (1983). — BANNISTER, M. H., WILLIAMS, A. L., MCDONALD, I. R. C. and FORDE, M. B.: Variation of turpentine composition in five populations of *Pinus radiata*. New Zealand J. Science. **5**, 486—495, (1962). — BARADAT, PH., BERNARD-DAGAN, C., FILLON, C., MARPEAU, A. and PAULY, G.: Les terpènes du pin maritime: aspects biologiques et génétiques. II. Hérédité de la teneur en monoterpènes. Annales des Sciences forestières **9**, 307—334, (1972). — BARADAT, PH., BERNARD-DAGAN, C., PAULY, G. and ZIMMERMAN-FILLON, C.: Les terpènes du pin maritime: aspects biologiques et génétiques. III. Hérédité de la teneur en myrcène. Annales des Sciences forestières **32**, 29—54, (1975). — BURLEY, J. and GREEN, C. L.: Variation of gum turpentine between provenances of *Pinus caribaea* MORELET and *Pinus oocarpa* SCHIEDE in Central America. In: D. T. SEAL *et al.* (Eds.): "EEC Symposium on Forest Tree Biochemistry". Commission of European Communities, Luxemborg. 73—108, (1977). — BURLEY, J. and GREEN, C. L.: Relationship of terpenes between exotic and natural populations of *Pinus caribaea* MORELET and *P. oocarpa* SCHIEDE. In: D. RUDIN (Ed.): "Proceedings of the conference on Biochemical Genetics of Forest Trees". Umeå, Sweden, 1978. Swedish Univ. Agricultural Sciences, Dept. of Forest Genetics and Plant Physiology. 118—132, (1979). — BURLEY, J. and LOCKHART, L. A.: Chemical extractives and exudates from trees. Annual Proc. Phytochemical Society Europe **26**, 91—102, (1985). — CHATFIELD, C. and COLLINS, A. J.: Introduction to multivariate analysis. Chapman and Hall, London. 246 pp, (1980). — CORI, O. M.: Enzymic aspects of the biosynthesis of monoterpenes in plants. Phytochemistry **22**, 331—341, (1983). — DELL, B. and MCCOMB, A. J.: Plant resins: their formulation, secretion and possible functions. Advances in Botanical Research **6**, 278—316, (1979). — ESTEBAN, I., BERGMANN, F., GREGORIUS, H.-R. and HUHTINEN, O.: Composition and genetics of monoterpenes from cortical oleoresin of Norway Spruce and their significance for clone identification. Silvae Genetica **25**, 59—66, (1976). — FALCONER, D. S.: Introduction to quantitative genetics. 2nd ed. Longman, London. 340 pp, (1981). — FORDE, M. B.: Inheritance of turpentine composition in *Pinus attenuata* × *radiata* hybrids. New Zealand J. Botany **2**, 53—59, (1964). — FORREST, G. I.: Geographical variation in the monoterpenes of *Pinus contorta* oleoresin. Biochemical Systematics and Ecology **8**, 343—359, (1980). — FRANKLIN, E. C.: Within-tree variation of monoterpene composition and yield in Slash Pine clones and families. Forest Science **22**, 185—191, (1976). — FRANKLIN, E. C. and SNYDER, E. B.: Variation and inheritance of monoterpene composition in Longleaf Pine. Forest Science **17**, 178—179, (1971). — GANSEL, CH. R. and SQUILLACE, A. E.: Geographic variation of monoterpene in cortical oleoresin of Slash Pine. Silvae Genetica **25**, 150—154, (1976). — GREEN, C. L., KEEBLE, B. and BURLEY, J.: Gum turpentine analysis of some *Pinus kesiya, P. caribaea* and *P. oocarpa* provenances. Tropical Science **16**, 195—206, (1974). — GREEN, C. L., KEEBLE, B. and BURLEY, J.: Further gum turpentine analyses of some *Pinus oocarpa, P. caribaea* and *P. kesiya* provenances. Tropical Science **17**, 165—174, (1975). — HANOVER, J. W.: Genetics of terpenes. I. Gene control of monoterpene levels in *Pinus monticola* DOUGL. Heredity **21**, 73—84, (1966a). — HANOVER, J. W.: Environmental variation in the monoterpenes of *Pinus monticola* DOUGL. Phytochemistry **5**, 713—717, (1966b). — HANOVER, J. W.: Inheritance of 3-carene concentration in *Pinus monticola*. Forest Science **12**, 447—450, (1966c). — HANOVER, J. W.: Genetics of terpenes. II. Genetic variances and interrelationships of monoterpene concentrations in *Pinus monticola*. Heredity **27**, 237—245, (1971). — HILTUNEN, R.: Variation and inheritance of some monoterpenes of *Pinus sylvestris*. Planta Medica **28**, 315—323, (1975). — HILTUNEN, R., TIGERSTEDT, P. M. A., JUVONEN, S. and POHJOLA, J.: Inheritance of 3-carene quantity in *Pinus sylvestris* L. Farmaceutiskt Notisblad **84**, 69—72, (1975). — IRVING, R. S. and ADAMS, R. P.: Genetic and biosynthetic relationships of monoterpenes. Recent Advances in Phytochemistry **6**, 187—214, (1973). — KANOWSKI, P. J.: Characterization and interpretation of variation in forest trees: a reappraisal based on a progeny test of

Pinus caribaea Morelet. Unpub. D. Phil. thesis, Dept. of Plant Sciences, Oxford Univ., (1986). — Lever, K. G. and Burley, J.: The application of biochemical methods in forestry. Paper to 10th Commonwealth Forestry Conference, Oxford, 1974. 20 pp., (1974). — Lockhart, L. A.: Investigation of tropical pine resin terpenes. Unpub M. Sc. thesis, Dept. of Plant Sciences, Oxford Univ. 206 pp, (1985). — Loomis, W. D. and Croteau, R.: Biochemistry and physiology of lower terpenoids. Recent Advances in Phytochemistry **6**, 147—185, (1973). — Marpeau, A., Baradat, Ph. and Bernard-Dagan, C.: Les terpènes du pin maritime: aspects biologiques et génétiques. IV. Hérédité de la teneur en deux sesquiterpènes: le longifolène et le caryophyllène. Annales des Sciences forestières **32**, 185—203, (1975). — Marpeau-Bezard, A., Baradat, Ph. and Bernard-Dagan, C.: Les terpènes du pin maritime: aspects biologiques et génétiques. V. Hérédité de la teneur en limonène. Annales des Sciences forestières **40**, 197—216, (1983). — Martin, S. S., Lagenheim, J. H. and Zavarin, E.: Biosynthesis of sesquiterpenes in *Hymenaea* inferred from their quantitative co-occurrence. Phytochemistry **15**, 113—119, (1976). — Mead, R. and Curnow, R. N.: Statistical methods in agriculture and experimental biology. Chapman and Hall, London. 335 pp., (1983). — Meier, R. J. and Goggans, J. F.: Heritabilities and correlations of the cortical monoterpenes of Virginia Pine (*Pinus virginiana* Mill.). Silvae Genetica **27**, 79—84, (1978). — Mosimann, J.: On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions. Biometrika **49**, 65—82, (1962). — Roberts, D. R.: Within-tree variation of monoterpene hydrocarbon composition of Slash Pine oleoresin. Phytochemistry **9**, 809—815, (1970). — Roberts, R. R., Williams, I. L. and Kossuth, S. V.: Double application of paraquat for production of lightwood with high oleoresin content. Forest Products J. **32**, 74—78, (1982). — Rockwood, D. L.: Variation in monoterpene composition of two oleoresin systems of Loblolly Pine. Forest Science **19**, 147—153, (1973). — Rudin, D. (Ed.): Proceedings of the conference on Biochemical genetics of Forest Trees. Umeå, Sweden, 1978. Swedish Univ. of Agricultural Sciences. Dept. of Forest Genetics and Plant Physiology. 201 pp., (1979). — Seal, D. T., Forrest, G. I., Philipson, J. J. and Parkes, H. (Eds.): EEC Symposium on Forest Tree Biochemistry. Brussels, 1977. Commission of European Communities, Luxemborg. 286 pp., (1977). — Shaw, D. V., Yazdani, R. and Muona, O.: Methods for analyzing data on the relative proportions of monoterpenes in conifers. Silvae Fennica **16**, 235—240, (1982). — Snedecor, G. W. and Cochran, W. G.: Statistical methods. 7th ed. Iowa State Univ. Press, Ames. 507 pp., (1980). — Simpson, C.: Gas Chromatography. Kogan Page, London. 117 pp and appendices, (1970). — Squillace, A. E.: Inheritance of monoterpene composition in cortical oleoresin of Slash Pine. Forest Science **17**, 381—389, (1971). — Squillace, A. E.: Analysis of monoterpenes of conifers by gas-liquid chromatography. In: J. P. Miksche (Ed.): "Modern methods in forest genetics". Springer-Verlag, Berlin. 120—137, (1976). — Squillace, A. E.: Use of monoterpene composition in forest genetics research with Slash Pine. In "Proc. 14th Southern Forest Tree Improvement Conference". Gainesville, Fl., 1977. Eastern Tree Seed Laboratory, Macon, Ga. 227—238, (1977). — Squillace, A. E. and Fisher, G. S.: Evidence of the inheritance of turpentine composition in Slash Pine. In "U.S.D.A. Forest Service Research Paper NC-6". 53—59, (1966). — Squillace, A. E., Wells, O. O. and Rockwood, D. L.: Inheritance of monoterpene composition in cortical oleoresin of Loblolly Pine. Silvae Genetica **29**, 141—151, (1980). — Stephens, M. A.: Use of the von Mises distribution to analyse continuous proportions. Biometrika **69**, 197—203, (1982). — Strauss, S. H. and Critchfield, W. B.: Inheritance of β-pinene in xylem oleoresin of Knobcone × Monterey Pine hybrids. Forest Science **28**, 687—696, (1982). — Thor, E., Samman, S. and Gall, W. R.: Variation in monoterpene composition of Eastern White Pine. Applied Polymer Symposium No. **28**, 1331—1339, (1976). — von Rudloff, E.: Volatile leaf oil analysis in chemosystematic studies of North American conifers. Biochemical Systematics and Ecology **2**, 131—167, (1975). — von Rudloff, E. and Rehfeldt, G. E.: Chemosystematic studies in the genus *Pseudotsuga*. VI. Inheritance and geographical variation in the leaf oil terpenes of Douglas-fir from the Pacific Northwest. Canadian J. of Botany **58**, 546—556, (1980). — Weast, R. C. (Ed.).: CRC Handbook of Chemistry and Physics. 64th ed. CRC Press, Boca Raton, Fl. (1983). — Watson, G. S. and Williams, E. J.: On the construction of significance tests on the circle and the sphere. Biometrika **43**, 344—352, (1956). — White, E. E.: Biosynthetic implications of terpene correlations in *Pinus contorta*. Phytochemistry **22**, 1399—1405, (1983a). — White, E. E.: Biochemical genetics of *Pinus contorta*. Swedish Univ. of Agricultural Sciences. Dept. of Forest Genetics and Plant Physiology, Umeå. 40 pp and 6 appendices, (1983b). — White, E. E.: Mode of genetic control of monoterpenes in foliage of controlled crosses of *Pinus contorta*. Silvae Genetica **33**, 115—119, (1984). — White, E. E. and Nilsson, J.-E.: Genetic variation in resin canal frequency and relationship to terpene production in foliage of *Pinus contorta*. Silvae Genetica **33**, 79—84, (1984). — Yazdani, R., Rudin, D., Alden, T., Lindgren, D., Harbom, B. and Ljung, K.: Inheritance pattern of five monoterpene in Scots Pine (*Pinus sylvestris* L.). Hereditas **97**, 261—272, (1982). — Zavarin, E.: Qualitative and quantitative co-occurrence of terpenoids as a tool for elucidation of their biosynthesis. Phytochemistry **9**, 1049—1063, (1970). — Zavarin, E. and Snajberk, K.: Geographic variability of monoterpenes from cortex of *Pseudotsuga menziessii*. Pure and Applied Chemistry **34**, 411—434, (1973).

Characters distinguishing Osier-willow Clones

By A. J. H. Rulkens*)

Agricultural University, Wageningen (Netherlands),
Department of Plant Taxonomy, P.O. Box 8010,
6700 ED Wageningen, Netherlands

(Received 30th October 1986)

Summary

The following characters are very useful to distinguish osier willow clones: sex, colour of the rod in winter, shape of the leaf bud, presence of black bud scales, time of leaf unfolding, colour of the petiole, presence of leaf-like organs on the petiole, glaucousness of under-surface of leaf, length/width ratio of the leaf blade. Of limited use are: hairiness of rod and leaf, position of the rods on the stool, persistance of stipules, length of the leaf blade.

Typical *S. fragilis* L. is often wrongly regarded as *S. alba* × *fragilis* by many authors. Special attention is paid to *S. fragilis* var. *decipiens* (Hoffm.) Koch and the here newly appointed cultivar groups Vitellina and Basfordiana.

Key words: Salix, Salix fragilis, Salix alba, Salix triandra, Salix viminalis, clone distinction and classification.

Zusammenfassung

Die nachfolgenden Merkmale haben sich als sehr nützlich erwiesen, um Flechtweiden-Klone zu unterscheiden: Geschlecht, Winterfarbe der Langtriebe, Form der vegetativen Knospen, Vorhandensein von schwarzen Knospenhüllen, Austriebszeitpunkt der Blätter im Frühjahr, Farbe des Blattstiels, Anwesenheit von blattähnlichen Petiolardrüsen, Glaukfärbung der Blattunterseite, Verhältnis zwischen Länge und Breite der Blattspreite. Weniger nützlich

*) Present address of author: Herenstraat 1, NL-6701 DG Wageningen