GENETIC AND BIOSYNTHETIC RELATIONSHIPS OF MONOTERPENES

ROBERT S. IRVING* and ROBERT P. ADAMS

Botany Department, University of Montana, Missoula, Montana, and Department of Botany and Plant Pathology, Colorado State University, Fort Collins, Colorado

Introduction

In the last one and one-half decades, there has been an increasing interest in secondary plant compounds. Much of this interest has been generated by

* Present address: Biology Department, Louisiana State University at New Orleans, New Orleans, Louisiana.

From: Terpenoids: Structure, Biogensis, and Distribution. Vol. 6. Recent Advances in Phytochemistry series. V.C.Runeckles and T. J. Mabry, Eds. Academic Press, N.Y. 1973. pp 187.-214.

their applicability to the field of systematic biology, where these compounds have served not only to indicate intrapopulational gene flow and population dynamics, but also to establish phenetic and phylogenetic relationships in various plant groups (Critchfield, 1966, 1967; DeWet and Scott, 1965; DeWet, 1967; Emboden and Lewis, 1967; Fujita, 1965a, b, c; Habeck and Weaver, 1969; Scora, 1967a, b, 1970; Stone, 1964; Stone *et al.*, 1965, 1969; Turner, 1967, 1969). With this increasing interest in secondary compounds, especially monoterpenes and sesquiterpenes, details of their modes of inheritance and formation assume major importance. It is of little use to accord phylogenetic or even phenetic significance to chromatographic profiles or to appraise complex introgressing populations until something is known of the genetic and biosynthetic background of these compounds.

With monoterpenes it is becoming increasingly clear that inter- and intrapopulational variation will reflect at least four major variables (Burbott and Loomis, 1967; Loomis, 1969; Adams, 1970; Firmage and Irving, 1971): (1) individual genetic divergence; (2) the broad environmental conditions of the population and the conditions prevailing at the time of sampling; (3) the ontogenetic and phenological stages; and (4) the techniques of extraction and analysis (e.g., *in vitro* rearrangements). Thus, it is important that genetic variation be isolated from other sources of variation before sound systematic conclusions are drawn. There have been few investigations that dealt directly with the inheritance of monoterpenes. Murray (1960a, b) demonstrated that the difference between two *Mentha* species is due to a single gene with two alleles. Forde (1964) using field variation of monoterpenes presented evidence that the major difference between two species of *Pinus* (high and low α-pinene levels) is simply inherited, but that within one species the range of variation suggested a more complex genetic control. Hanover (1966a) working with oleoresin of F_1's in *Pinus monticola* suggests that most of the monoterpenes of this species are under multigenetic control. One compound, 3-carene, proved to be an exception, being seemingly under the control of a single dominant gene (Hanover, 1966b). Zavarin *et al.* (1969) in the F_1, F_2, and backcross progeny of two *Pinus* species (*P. contorta* and *P. banksiana*) suggests that the monoterpene differences are controlled by a limited number of genes with major effects.

Closely coupled to information on monoterpene inheritance is knowledge of the biosynthetic pathways involved. Indeed, it is often difficult to separate genetics from biosynthesis, and both are prerequisite to sound systematic use of terpene data. Currently, a number of biosynthetic models have been proposed based on labeling experiments (Battu and Youngken, 1966; Scora and Mann, 1967; Sandermann, 1962); structural properties and chemical reactivities (Kremers, 1922; Ruzicka, 1953; Reitsema, 1958), and nat-

ural variation (Zavarin, 1970; Zavarin and Cobb, 1970; Zavarin *et al.*, 1971). These numerous schemes although unified in their fundamental premises are, in their details, as diverse as the compounds themselves.

The present work deals with the biosynthesis and the genetics of monoterpenes in experimental populations of three taxa within the mint genus *Hedeoma*. Varying levels of qualitative and quantitative intercorrelations within an experimental F_2 population are used to suggest biosynthetic and genetic relatedness. Further, variation of individual compounds and groups of compounds across the F_2 generation is utilized to provide possible genetic models of inheritance.

The Biology of the *Hedeoma drummondii* Complex

The *Hedeoma drummondii* complex (Labiatae) represents a unique experimental system for addressing questions of monoterpenoid inheritance and biosynthesis. The complex consists of three morphologically defined taxa, *H. drummondii*, *H. reverchonii* var. *reverchonii*, and *H. reverchonii* var. *serpyllifolium*, each possessing a unique terpenoid complement and a high degree of intercrossability. Geographically the complex is centered on the arid limestone outcroppings of the Edwards Plateau region of central Texas (Fig. 1). Here, all three taxa form large sympatric populations on the same chromosome level. The distinguishing morphological characters, although many in number, are, in these regions of sympatry, subtle and intergrading. In fact, there appears to be in each taxa an array of recognizable evolutionary lines or microspecies. It is this recognizability and repetition of variation that had led to much of the classical taxonomic difficulty in dealing with this group. Partial explanation of this complexity lies in their predominant inbreeding behavior which funnels genetic variation into a series of uniform lines. Much of the inter- and intrapopulational variation also appears to be the result of past and present gene exchange between several or all of these taxa. Thus, gene flow, in as yet unknown quantities, is believed to usher in the variation which is then played upon by selection and proliferated by inbreeding. The initial problem was to account for the populational patterns of variation and the origins of the complex itself. Out of these basic systematic and field oriented questions arose the genetic and biosynthetic experiments that are to be discussed. Because morphology did not provide adequate differentiation, the monoterpenes of these taxa were examined. Although the terpenoid data provided the requisite taxonomic separation, concomittantly there had to be greater understanding of the biosynthetic and genetic base of these new characters.

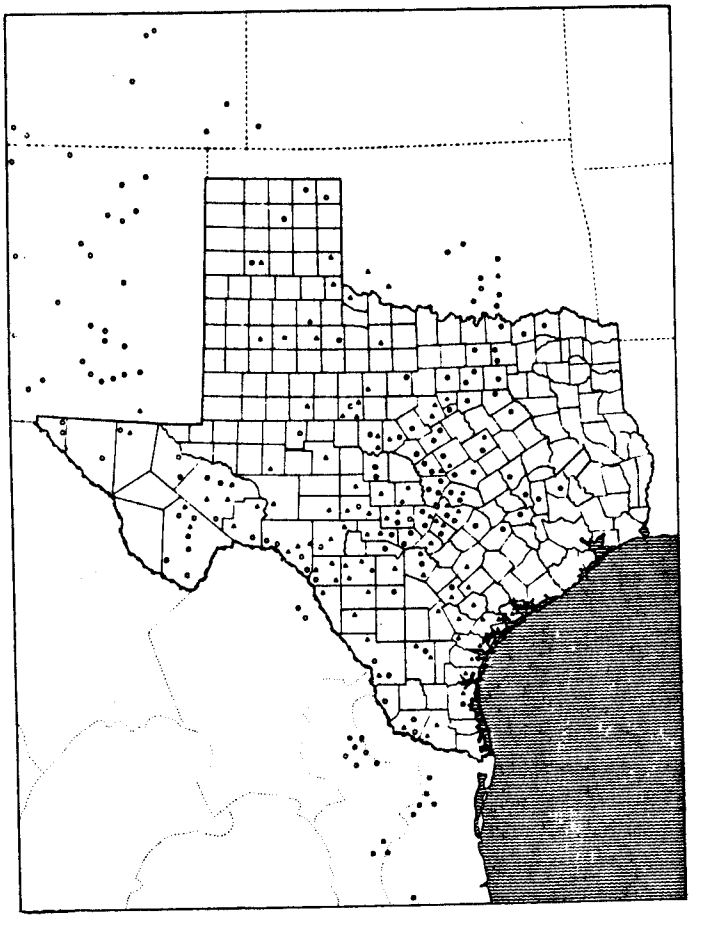


FIG. 1. Distribution of *Hedeome drummondii* complex: *H. drummondii*, open circles; *H. reverchonii* var. *serpyllifolium*, filled triangles; *H. reverchonii* var. *reverchonii*, filled circles.

THE MONOTERPENES OF THE *Hedeoma drummondii* COMPLEX

The salient phytochemical products of the family Labiatae are monoterpenes. Employing standarized, analytical, and preparative gas chromatographic techniques (Irving, 1968), numerous individuals representing several populations were examined as to their monoterpene products (Table 1).

The monoterpenes of *H. drummondii* are ca. 90 percent monocyclic compounds: *d*-limonene, isomenthone, menthone, and pulegone. Although the monocyclic ring theme predominates, *H. drummondii* does not lack the potential to synthesize other ring types in trace amounts (e.g., α-pinene, camphene, sabinene, and myrcene). This pattern, in most of its quantitative

TABLE 1

MONOTERPENES OF THREE PARENTAL TAXA

Monoterpenes	*H. drummondii*	*H. reverchonii* var. *reverchonii*	*H. r.* var *serpyllifolium*
1. α-Pinene	Tr[a]	1.5	8.0
2. Camphene	Tr	1.0	17.0
3. β-Pinene	Tr	Tr	30
4. Sabinene	Tr	Tr	5.0
5. Myrcene	Tr	2.0	2.5
6. *d*-Limonene	1.5	Tr	4.5
7. 1,8-Cineole	Tr	Tr	7.5
8. Terpinene	—	—	2.5
9. Terpinolene	—	Tr	2.0
10. Unknown	—	Tr?	2.0
11. Menthone	1.5	—	—
12. Isomenthone	40.0	—	—
13. Cyclic ketone	1.0	Tr?	Tr
14. Unknown	1.0	Tr?	Tr
15. Tricyclene	—	—	0.5
16. Unknown	1.0	Tr?	4.3
17. Pulegone	50.0	—	—
18. Borneol	—	3.5	26.0
19. Unknown	Tr	1.0	3.0
20. Unknown	—	—	Tr
21. Unknown	?	?	1.0
22. Terpinenol-4	Tr	3.0	1.5
23. Menthol	Tr	—	—
24. *t*-Ocimene	—	4.5	Tr
25. Citronellal	—	22.0	Tr
26. Neral	—	21.0	Tr
27. Geranial	—	31.7	Tr
28. β-Phellandrene	—	Tr	Tr
29. Camphor	—	Tr?	Tr
30. Unknown	—	3.0	—
31. Unknown	Tr	—	—
32. Unknown	—	3.0	—

[a] Tr = < 1 percent.

and all of its qualitative aspects is constant throughout the range of *H. drummondii*. The terpenoid profiles of populations from Montana are remarkably similar to those found in Northern Mexico.

Hedeoma reverchonii var. *serpyllifolium* displays a quite different monoterpene profile with a bicyclic ring theme, the major constituents being borneol, camphene, 1,8-cineole, α- and β-pinenes, tricyclene, and sabinene. As with *H. drummondii*, other ring themes are present in trace amounts.

Although there is quantitative variation found both within and between populations, the overall quantitative and qualitative profile remains invariable.

H. reverchonii var. *reverchonii* displays yet a third distinctive terpenoid pattern with an acyclic monoterpene theme (80–85 percent): neral, geranial, myrcene, *trans*-ocimene, and citronellal. Again there are trace products representing other ring themes, but with the acyclic profile holding from population to population.

Thus, where at the morphological level we have continuous, intergrading characters, at the monoterpene level we have characters that are highly discontinuous and seemingly represent points of major divergence. In light of current biosynthetic views, these three taxa seem to have taken, with the guidance of natural selection, three different biosynthetic routes: acyclic, monocyclic, and bicyclic. Moreover, as we find these patterns intact from taxon to taxon, we might further infer that monoterpenes possess a high degree of heritability and taxonomic reliance. Despite these inferences, however, sound systematic judgments can be made only when nongenetic variation of monoterpenes is reduced and their behavior with gene flow is known. In our work, a series of environmentally controlled breeding experiments was designed to generate the three possible F_1 and F_2 generations.

The Monoterpenes of F_1 Generations

One of the major emphases in micromolecular systematics has been the documentation of hybridization (Alston and Turner, 1963a). The major assumption, stemming primarily from flavanoid data (Alston and Turner, 1963a, b; Smith and Levin, 1963), has been that F_1 hybrids will be recognizable through qualitative and/or quantitative complementation. In monoterpene inheritance this assumption has only been tested quantitatively (Mirov, 1956; Hanover, 1966a; Zavarin *et al.*, 1969).

The experimental *Hedeoma* F_1 hybrids from each of the three combinations of parents yielded the same broad results and confirmed the assumption of complementation. Each F_1 hybrid, although varying in fertility, displayed a complete complementary or additive inheritance pattern of the parental monoterpene products (Table 2). Quantitatively nearly all the monoterpene compounds in the F_1's were intermediate and approached theoretical values based on simple one-to-one addition of the two parental oils. However, in each F_1 type there were transgressive compounds which exceeded the levels found in either of the two parents.

The *H. drummondii* × *H. reverchonni* var *reverchonii* Cross

This cross represents the monocyclic systems of *H. drummondii* in the maternal background of the acyclic *reverchonii* (Table 2). The hybrids were

TABLE 2

MEAN PERCENTAGE OF MONOTERPENES IN F_1 GENERATIONS

Monoterpenes	*Hedeome drummondii* × *H. r.* var. *serpyllifolium*	*H. drummondii* × *H. r.* var. *reverchonii*	*H. r.* var. *serpyllifolium* × *H. r.* var. *reverchonii*
1. α-Pinene	3.4	0.3	3.3
2. Camphene	6.1	0.1	5.2
3. β-Pinene	1.3	0.2	1.0
4. Sabinene	2.6	0.1	3.4
5. Myrcene	1.5	1.1	2.8
6. *d*-Limonene	8.7	5.8	3.0
7. 1,8-Cineole	3.9	Tr	4.1
8. Terpinene	1.2	—	Tr
9. Terpinolene	1.2	—	1.4
10. Unknown	0.4	—	1.7
11. Menthone	11.9	7.8	—
12. Isomenthone	13.5	10.4	—
13. Cyclic ketone	0.3	0.8	Tr
14. Unknown	1.1	0.6	Tr
15. Tricyclene	0.1	—	0.1
16. Unknown	3.8	Tr	5.1
17. Pulegone	16.5	22.5	—
18. Borneol	15.9	1.2	13.6
19. Unknown	3.3	1.5	3.6
20. Unknown	0.3	—	—
21. Unknown	0.8	—	—
22. Terpinenol-4	Tr	3.3	Tr
23. Menthol	Tr	—	—
24. *t*-Ocimene	—	3.4	8.1
25. Citronellal	—	5.3	2.5
26. Neral	—	5.6	10.8
27. Geranial	—	12.5	19.0
28. β-Phellandrene	Tr	0.1	0.1
29. Camphor	Tr	0.5	1.1
30. Unknown	—	11.6	6.4
31. Unknown	—	2.8	—
32. Unknown	—	2.3	2.0

vigorous with an average fertility of 33.2 percent. Their terpenoid profiles were remarkably similar and showed complete complementation; all constituents of both parents were represented in the F_1. Thus, in one plant, and presumably in one gland, there was the coexistence of two disparate pathways: acyclic and monocyclic. Quantitatively, there were three transgressive compounds. Menthone, a trace constituent (0.3 percent) of *H. drum-*

mondii and nonexistent in *H. reverchonii* var. *reverchonii* rose to an average level of 7.8 percent in the F_1 hybrids. Two unidentified trace components (0.5 and 3.5 percent) of *H. reverchonii* var. *reverchonii* rose to 2.9 and 10.6 percent, respectively.

The *H. reverchonii* var. *serpyllifolium* × *H. reverchonii* var *reverchonii* Cross

In this cross a bicyclic system has been incorporated with an acyclic one (Table 2). The hybrids displayed a high degree of fertility (89.0 percent), and as with the previous cross there was qualitative–quantitative complementation and the occurrence of transgressive compounds. However, in this F_1 group there was considerably greater variation which centered principally within the aldehydes, citronellal, neral, and geranial.

The *H. drummondii* × *H. reverchonii* var. *serpyllifolium* Cross

Here a monocyclic system coexists with a bicyclic one with qualitative and quantitative complementation (Table 2). The hybrids themselves were quite healthy despite a lowered fertility of only 16.0 percent. Menthone, entering the system through *H. drummondii*, was noticeably transgressive rising to an average level of 12.8 percent in the F_1 population. Variation between individuals was again centered principally around a carbonyl function in the ketones menthone, isomenthone, and pulegone.

Discussion

With this series of crosses, ostensibly divergent pathways have been shown to coexist within a given plant. At the genetic level, this coexistence is indeed real with every cell of a given F_1 plant possessing the potential for the synthesis of two divergent monoterpene systems. Whether or not this potential is actualized is not as yet known. It may well be that there is a high degree of cellular or intracellular compartmentalization (Loomis, 1972). The transgressive states of several compounds encountered in all the crosses, however, suggest an absence of such compartmentalization, at least at the cellular level. A reasonable model for transgressive character states is that of enzyme complementation (Alston *et al.*, 1965). For example, enzymes for increasing the levels of menthone, or like isomers, may be latent in *H. r. serpyllifolium* and *H. r. reverchonii*. These will then be activated only in the presence of a genome which synthesizes menthone: *H. drummondii*. Such a model would necessitate a decompartmentalized system with enzymatic interchange.

The variation within each F_1 population was primarily with carbonyl functions. Recent biosynthetic models would order such compounds along a reductive sequence (Reitsema, 1958; Loomis, 1969). Such an alignment would remove the products sequentially from their precursor and accord them a greater degree of biosynthetic autonomy. Further, with such a sequence, there would be greater chance of reversibility and sensitivity to environmental change (Loomis, 1969; Firmage, 1971).

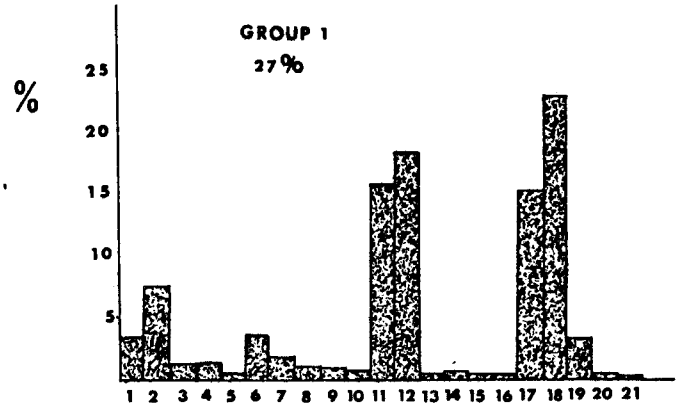
The Monoterpenes of F_2 Generations

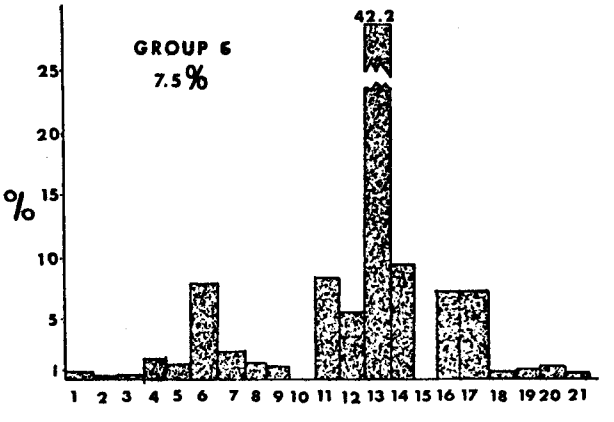
Monoterpene analysis of experimentally controlled F_2 populations can bring insight into three major areas: (1) the quality and quantity of recombination in an F_1 hybrid of divergent parentage; (2) the biogenetic relationships suggested from the variation; and (3) the genetic background of individual or groups of monoterpene compounds. In our work it was further hoped that these patterns of variation would provide direction for subsequent experimental and field research.

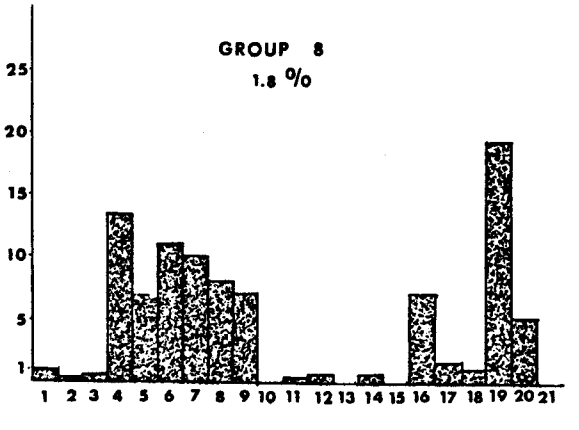
A single F_1 hybrid was chosen from the *H. drummondii* × *H. r. serpyllifolium* cross (monocyclic × bicyclic), and through inbreeding an F_2 population of approximately 300 individuals was established. Of these, 106 were randomly selected for analysis. Of the 45–50 compounds, 21 were singled out for study, the selection based primarily on quantitative "importance" (Table 1). In addition, 10 morphological attributes were examined along with individual plant fertility. The entire generation was subjected to the same environmental regimes as the parents and F_1's and sampling was conducted at comparable developmental stages.

The most difficult appraisal of variation is that of entire patterns, individual to individual. Indeed, it is this type of appraisal which is the heart of the "taxonomic method." However, as the F_2 chromatograms were examined, it was quickly realized that the variation throughout the F_2 generation could be subjectively taxonomized—the patterns were not as variable as one might expect on *a priori* grounds. This conservative variation could be categorized into eight subjective groupings.

The largest group (27.0 percent of the F_2 generation) (Fig. 2) displayed patterns which approximated the qualitative and quantitative patterns of the F_1 hybrid. Morevoer there were two subgroups (8.0 and 19.0 percent, respectively) which also followed the F_1 theme but differed significantly in the levels of pulegone, isomenthone, and menthone. A fourth group was subjectively clustered through depressed levels of the monocyclic ketones and the presence of an unknown ketone at transgressive levels (compound 13). Of particular importance are groups 5 and 6 (22.6 percent) which repre-


GROUP 1
27%
%
25
20
15
10
5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21


42.2
GROUP 6
7.5%
%
25
20
15
10
5
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21


GROUP 8
1.8 %
25
20
15
10
5
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

sent a complete (group 5) or nearly complete (group 6) return to the parental patterns of the *H. drummondii* parent. Group 6 deviated from the *H. drummondii* profile with the presence of the same transgressive ketone present in group 4 (Fig. 3). The monoterpene return to the *H. r.* var. *serpyllifolium* parent (group 7) represented 9.4 percent of the F_2 population. This segregation, back to the parental profiles, was complete both qualitatively and quantitatively. The remaining group, although a small portion of the sample, was instructive in that it represented "radical" recombinants (Fig. 4).

Although these groups are in part subjective, they suggest a high degree of coherency in the monoterpene patterns. Groups of monoterpenes will not breakdown readily through genetic and chromosomal recombination.

To treat the F_2 generation as a variable population and examine individual compounds or groups of compounds out of the individual patterns of variation, two basic techniques were used: (1) intercorrelation analysis which provided correlation coefficients for the 21 monoterpene products, 10 morphological characters, and fertility (Table 3); (2) factor analysis to cluster those characters together which were covarying (Fig. 5). Intercorrelation analysis, as a means of interpreting terpenoid variation in genetic or biosynthetic terms, has been used by several workers (Hanover, 1966a, b; Zavarin, 1970). In an experimental F_2 generation, the correlation coefficients of monoterpenes can assume a high degree of significance as both gene flow and environment have been highly controlled and selective pressures reduced to a minimum. With the above controls, correlations between two monoterpenes, representing concomitantly genetic and biosynthetic relationships, can be one of three intergrading types*: (1) a nonsignificant correlation representing a high degree of genetic and biosynthetic independence; (2) negative correlations representing the synthesis of one compound at the expense of the other or its precursor and, thus, some degree of biogenetic interdepencence; and (3) a large positive correlation signifying a codependence and perhaps a higher degree of common genetic and biosynthetic control.

* For the problem of spurious high correlations through quantification by "percent total oil," see discussion by Zavarin (1970), which we accept.

FIG. 2 (top). Group 1 F_2 pattern, which approximates F_1.

FIG. 3 (middle). Group 6 F_2 pattern, which is essentially a return to *Hedeoma drummondii* except for compound 13 (ketone).

FIG. 4 (bottom). Group 8 F_2 pattern with a high degree of recombination.

TABLE 3

CORRELATION

	1	2	3	4	5	6	7	8	9	10	11	12
1. α-Pinene	1.0	0.94	0.995	0.35	0.32	−0.08	0.36	0.14	0.34	0.66	−0.31	−0.30
2. Camphene		1.0	0.94	0.28	0.24	−0.13	0.27	0.04	0.28	0.73	−0.28	−0.31
3. β-Pinene			1.0	0.30	0.27	−0.11	0.32	0.11	0.29	0.67	−0.29	−0.27
4. Sabinene				1.0	0.99	0.55	0.93	0.75	0.99	0.12	−0.36	−0.47
5. Myrcene					1.0	0.58	0.91	0.76	0.98	0.05	−0.37	−0.48
6. *d*-Limonene						1.0	0.56	0.51	0.54	−0.20	−0.24	−0.39
7. 1,8-Cineole							1.0	0.68	0.93	0.17	−0.34	−0.45
8. Terpinene								1.0	0.75	0.03	−0.24	−0.35
9. Terpinolene									1.0	0.12	−0.36	−0.47
10. Unknown										1.0	−0.06	−0.14
11. Menthone											1.0	0.42
12. Isomenthone												1.0
13. Ketone												
14. Unknown												
15. Tricyclene												
16. Unkown												
17. Pulegone												
18. Borneol												
19. Unknown												
20. Unknown												
21. Unknown												
23. Fertility												
27. Calyx length												

Because of the paucity of information on monoterpene interconversions and the difficulty in separating biosynthesis from genetics, definitive models of biosynthesis cannot be assigned to any of these intercorrelation categories. Moreover, with all schemes of monoterpene interconversion, it is as yet difficult to distinguish the four possible modes of biosynthesis and gene action: (1) nonenzymatic conversions in which the quantitative relationships are the product of reaction rates; (2) nonspecific enzymatic conversions where ubiquitous enzymes will act on a variety of substrates to produce a variety of monoterpene products; (3) specific enzymatic conversions with the enzyme geared to a specific monoterpene substrate and product; and (4) biologically, and perhaps more realistically, *a priori*, a combination of all three. Yet despite these inherent difficulties and the complexity of the system, it may be fruitful to attempt such modeling, fully realizing the risks of oversimplification.

With relatively high negative correlation coefficients two related biosynthetic models are plausible: (1) a sequence or series of interdependent biosynthetic steps; (2) differential enzymatic competitiveness for a common precursor. With a high positive correlation coefficient, a biosynthetic and genetic model based on a common intermediate becomes attractive.

INDIVIDUAL VARIATION

A synopsis of the individual monoterpene variation (Fig. 6) in the F_2 generation revealed a wide range for each monoterpene, and in many in-

COEFFICIENT MATRIX

13	14	15	16	17	18	19	20	21	23	27	
−0.26	−0.12	0.91	−0.04	−0.45	0.84	0.11	−0.19	0.06	−0.03	−0.08	1.
−0.27	−0.15	0.98	−0.07	−0.43	0.93	0.09	−0.22	0.04	−0.06	−0.04	2.
−0.24	−0.09	0.91	−0.09	−0.45	0.83	0.06	−0.23	0.06	−0.04	−0.07	3.
−0.17	−0.18	0.27	0.66	−0.46	0.20	0.85	0.62	0.02	−0.03	−0.06	4.
−0.11	−0.14	0.22	0.68	−0.45	0.13	0.85	0.63	−0.01	−0.01	−0.06	5.
0.18	−0.06	−0.10	0.53	−0.27	−0.16	0.55	0.47	0.13	−0.34	−0.16	6.
−0.18	−0.17	0.25	0.54	−0.45	0.22	0.73	0.48	0.04	−0.12	−0.07	7.
−0.06	−0.15	0.04	0.43	−0.33	0.04	0.62	0.67	−0.00	−0.06	−0.14	8.
−0.17	−0.20	0.26	0.65	−0.45	0.21	0.86	0.61	−0.02	−0.08	−0.07	9.
−0.28	−0.21	0.72	−0.23	−0.34	0.83	−0.04	−0.20	0.08	−0.17	−0.07	10.
−0.09	−0.06	−0.29	−0.30	−0.12	−0.24	−0.32	−0.19	−0.02	0.02	0.11	11.
0.28	−0.16	−0.31	−0.40	−0.17	−0.28	−0.45	−0.21	−0.02	0.02	−0.08	12.
1.0	0.57	−0.27	0.14	−0.31	−0.31	−0.08	0.02	−0.05	0.04	0.07	13.
	1.0	−0.15	0.05	−0.24	−0.20	−0.13	−0.04	−0.05	−0.01	0.08	14.
		1.0	−0.50	−0.41	0.91	0.10	−0.21	0.05	−0.10	−0.03	15.
			1.0	−0.31	−0.17	0.87	0.59	−0.11	−0.03	−0.12	16.
				1.0	−0.36	−0.36	−0.23	−0.09	013	0.06	17.
					1.0	0.04	−0.21	0.10	−0.16	−0.07	18.
						1.0	0.65	−0.06	−0.14	−0.05	19.
							1.0	0.17	0.04	−0.08	20.
								1.0	0.11	−0.01	21.
									1.0	0.21	23.
										1.0	27.

stances the levels were transgressive. Despite this variation, compound to compound, the means of the F_2 population did not differ significantly from those of the F_1's with two exceptions: *d*-limonene and unknown compound 13.

INTERCORRELATION AND FACTOR ANALYSIS

The intercorrelation coefficients are presented in Table 3, and the singles linkage diagram of factor analysis in Fig. 5. The latter allowed us to cluster groups of highly positively intercorrelated compounds and to examine them as units (factors).

Factor one

Factor 1 consists of six compounds: α- and β-pinenes, camphene, tricyclene, borneol, and an unknown bicyclic, compound 10 (Fig. 5). As a unit, it displays a remarkable degree of positive intercorrelation, with the lowest correlation coefficient between pairs 10-1 being 0.7 (Table 3). The highest correlation coefficient existed between α-pinene and β-pinene (0.9953). Structurally, Factor 1 represents the known bicyclic compounds under study. Thus, it seems apparent in view of their high degree of intercorrelation and structural similarity that Factor 1 represents a highly cohesive biosynthetic and perhaps genetic unit. A number of biosynthetic models dealing with these compounds have been formulated from a variety of experimental data (Ruzicka, 1953; Sandermann, 1962; Scora and Mann,

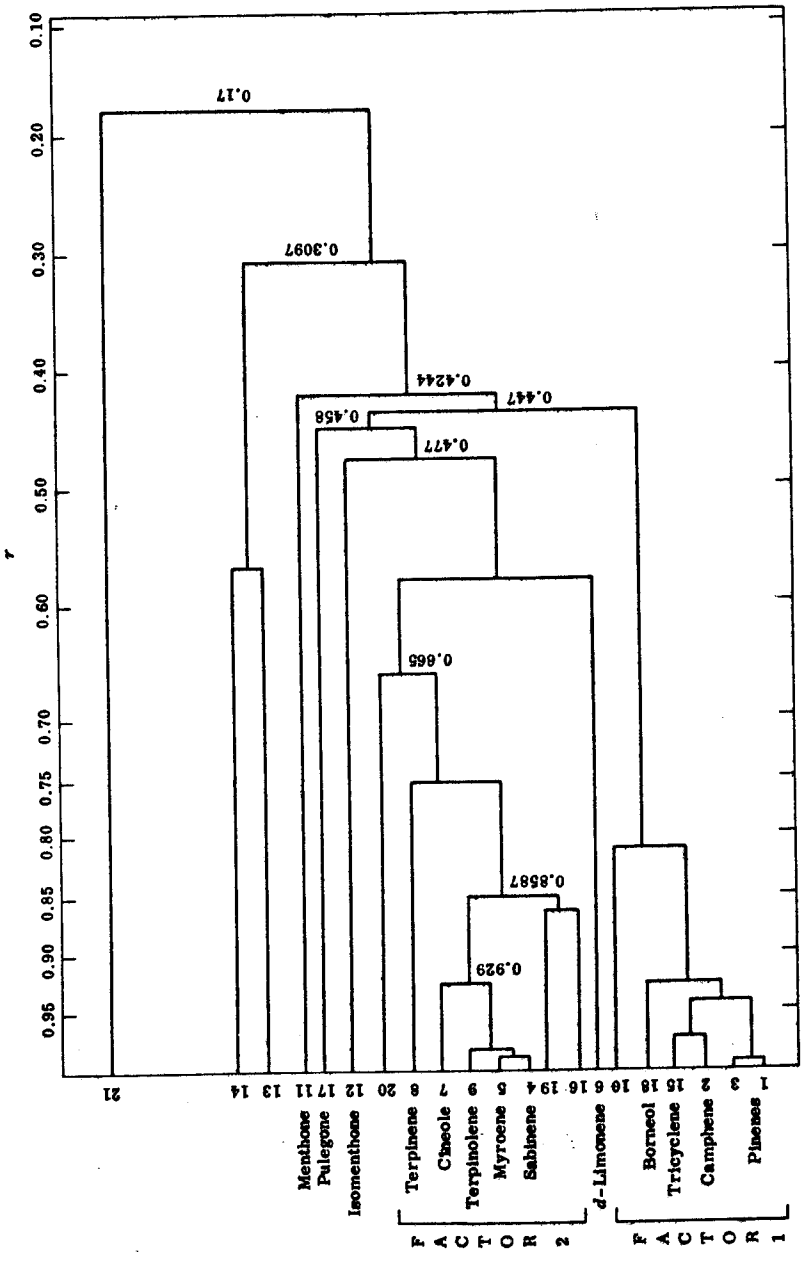


FIG. 5. A single-linkage clustering of monoterpenes in the F_2 generation by correlation coefficients.

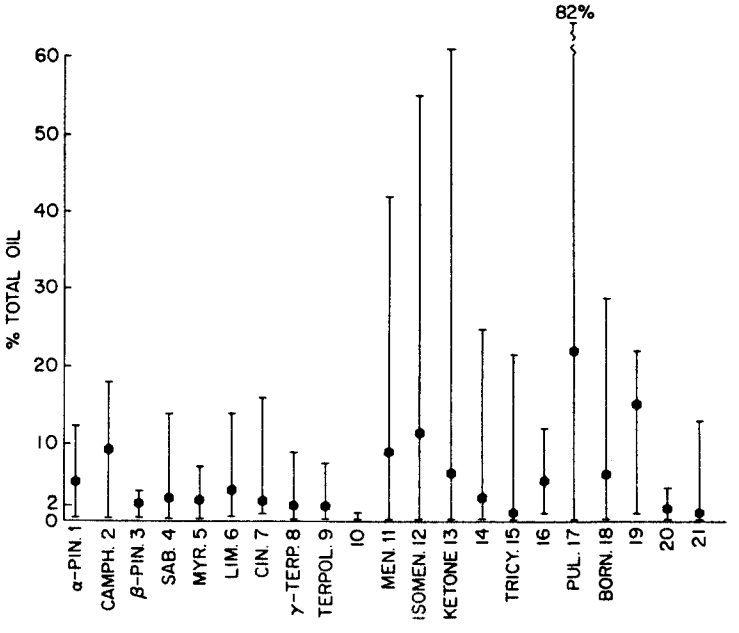


FIG. 6. A synopsis of the total monoterpene variation in the F_2 generation. Height of bar represents the range; hexagon, the mean (x). For names of compounds, see corresponding numbers in Table 3.

1967; Zavarin, 1970; Zavarin and Cobb, 1970). Most of these schemes have biosynthetically arranged the bicyclics around a common intermediate. The interrelationships uncovered in our data would further support such a model. We would modify it, however, to bring the pinenes into closer biosynthetic unity with borneol, camphene, and tricyclene (Fig. 7). This common intermediate theory for Factor 1 would nicely account for the positive intercorrelations and for the genetic patterns discussed in the next section, but its existence is by no means firmly established. What is actually functioning as the intermediate, if any, is in no way determinable from our data, and the carbonium ion is used in ignorance. Indeed, the presence or absence of a common intermediate may not be as of much importance as unique enzymes which can operate on a ubiquitous intermediate.

Factor Two

A second unit of highly positively intercorrelated compounds was revealed through factor analysis: myrcene, sabinene, terpinolene, terpinene, unknowns 16 and 19, and 1,8-cineole (Fig. 5). The structurally unlike pair,

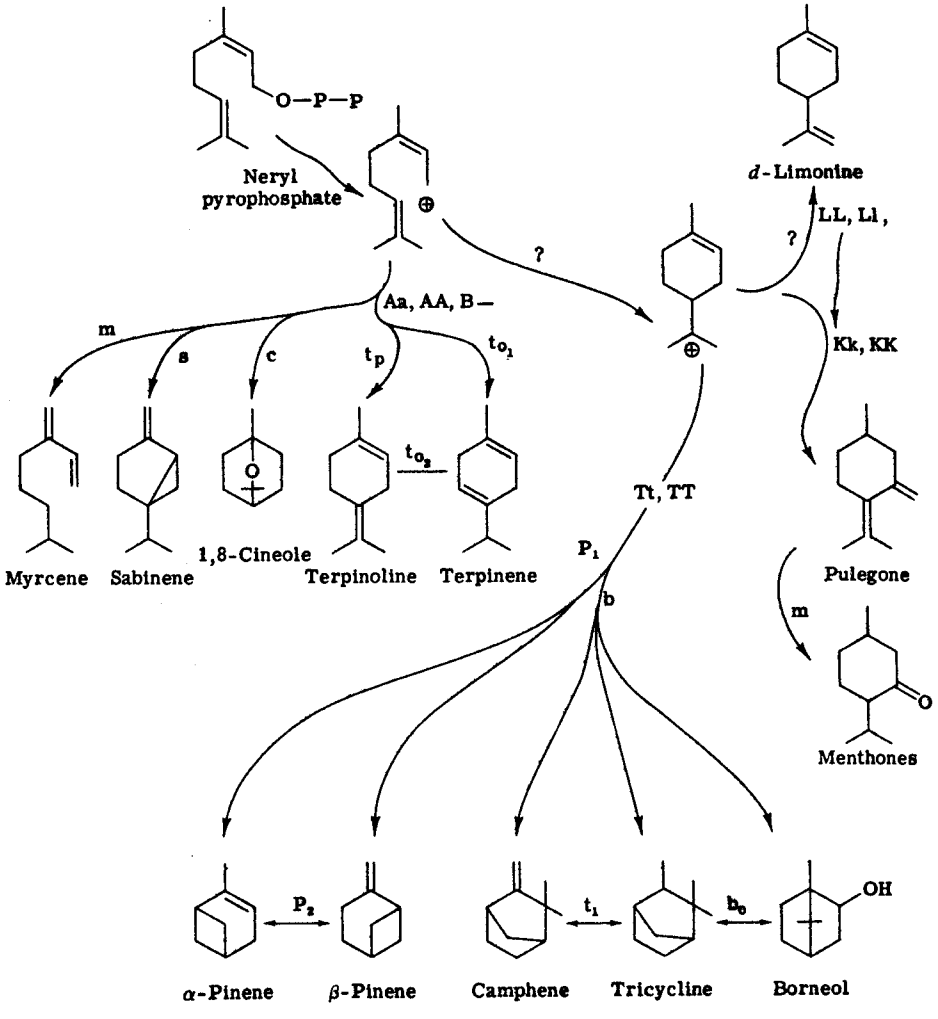


FIG. 7. Biosynthetic and genetic synopsis of the monoterpenes in the *Hedeoma drummondii* complex. Lower case letters represent alternative sites of gene and/or enzymatic action.

sabinene and myrcene, displayed the highest correlation coefficient for this unit, 0.9906 (Table 3) and is at variance with previous data (Zavarin *et al.*, 1971). Although Factor 2 is structurally more diverse than Factor 1, its strong cohesiveness again suggests biosynthetic and/or genetic unity. Indeed, it is interesting to note that two F_2's were quantitatively almost ex-

clusively Factor 2. Again with the high level of positive correlation, it is tempting to model these compounds through a common intermediate. Yet, because of their structural diversity, such a model must be presented with even greater caution. It is conceivable that their unity is the result of diverse but tightly linked enzymes which operate independently (Fig. 7).

Since an orthogonal rotation (verimax) was used to generate Factors 1 and 2, we know that these factors represent independent modes of variation. A comparison of the intercorrelations (Table 3) shows that no compound that is highly loaded on Factor 1 is well correlated with those compounds highly loaded on Factor 2. Interestingly, both Factors, 1 and 2, are the principal constituents of the parental species *H. reverchonii* var. *serpyllifolium*.

d-Limonene

d-Limonene is a major constituent of both *H. r.* var. *serpyllifolium* and *H. drummondii* and displays a number of weak positive correlations with members of Factor 1 (0.6) and negative correlations with Factor 2 (Table 3). Within the *H. drummondii* system, *d*-limonene is highly negatively correlated with isomenthone and other ketones through developmental time (Firmage, 1971). Thus, although not readily apparent in our correlation data, *d*-limonene may be an integral part of the biosynthetic schemes for ketones.

Pulegone, Menthone, and Isomenthone

The three ketones pulegone, menthone, and isomenthone are the major terpene constituents of the *H. drummondii* parent. They display no significant positive correlations either internally or to other monoterpene products. Indeed, if any trend in correlation exists it is a negative one. This is most pronounced with the hydrocarbon fractions (Table 3) and weakly reiterates what has been said by a number of researchers (see Loomis, 1969; Firmage and Irving, 1971), that the oxygenated fraction is made in competition with, or at the expense of, hydrocarbons. Because of their ubiquity a number of biosynthetic schemes have been proposed for these monocyclic ketones (Loomis, 1969; Scora and Mann, 1967). In all, a sequential reduction series is envisioned usually beginning with piperitenone and terminating with menthone or isomenthone (Fig. 7). With such a sequence, negative correlations would be anticipated (Table 3). It also might be anticipated that these would be relatively low, as the system would be an open and reversible one.

Correlations with Morphology and Fertility

The correlation between the terpenes and morphological attributes have interesting systematic implications. None of the 21 monoterpenes had any significant positive or negative correlation to any single or group of morphological characters (Table 3). Thus, monoterpenes and morphology, in this complex of species, will segregate independently in F_2 generations. It is quite conceivable that in natural variation one could obtain, through hybridization, plants morphologically of one species and chemically of another, and indeed this is realized in the field variations encountered in the central Texas populations.

Likewise, fertility (0–100 percent in the F_2) showed no significant correlations; thus, reproductive success of F_2 recombinants (as measured by fertility) would be independent of their terpenoid products (Table 3).

Genetic Analysis of Monoterpenes

In an attempt to examine the genetic component of the monoterpene patterns, frequency distributions were plotted for each of the compounds examined in the F_2 generation. The basic premise is that if a large number

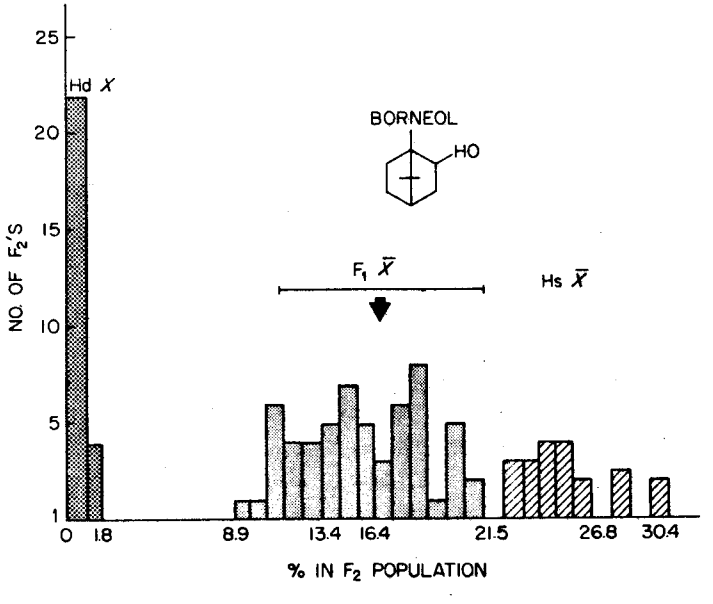


FIG. 8. Frequency distribution of borneol in F_2 generation, as representative of Factor 1 compounds.

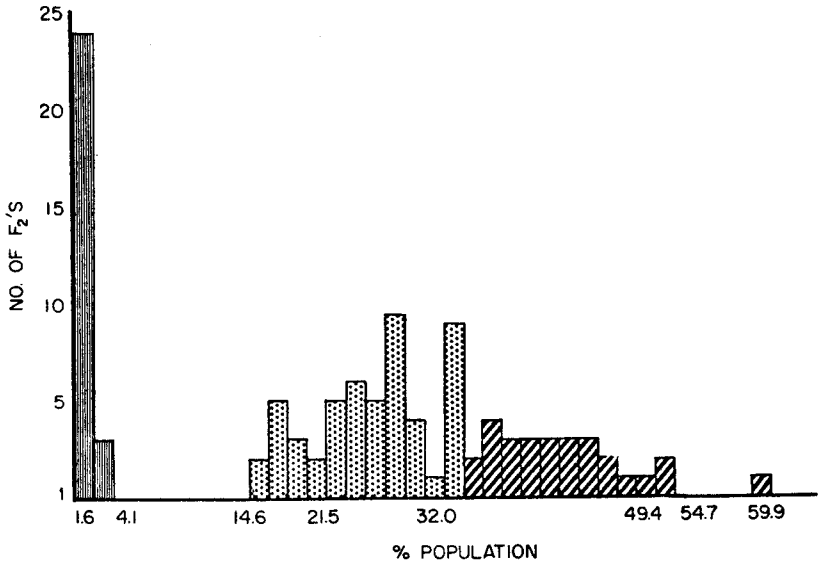


FIG. 9. Frequency distribution of an aggregation of all Factor 1 (bicyclic) compounds: α- and β-pinene, tricyclene, borneol, compound 10, camphene.

of genes govern the quantitative state of a single monoterpene we would expect the frequency distribution to approximate a Gaussian curve, whereas with one or few closely linked genes we would expect a frequency distribution with a high degree of modality. As the parental strains used in this work were highly inbred, and a high degree of initial homozygosity insured, we could also examine for Mendelian and neo-Mendelian ratios.

FACTOR ONE

The first group studied was the bicyclic compounds of Factor 1: α- and β-pinenes, camphene, tricyclene, borneol, unknown compound 10. As would be suggested by their high degree of positive intercorrelation, these compounds are unified genetically; that is, they display nearly identical frequency diagrams and can be treated as a unit. Using borneol and the bicyclic composite as examples, the levels of the two parents fell, as expected, on opposite ends of the frequency distribution and the range and mean of the F_1's at the mid-values (Figs. 8 and 9). Throughout Factor 1 there was a strong bimodality and a suggestion of trimodality (Fig. 9). The numbers of individuals filling the two major modes (low levels and medium plus high levels) have a nearly perfect 1:3 ratio to each other. The second mode (medium and high levels) can be further subdivided into two submodes in

an approximate 2:1 ratio. This is further shown by the data for α-pinene (Fig. 10, Table 4) where there is a strong suggestion of a single gene model of inheritance segregating in classic Mendelian fashion, 1:2:1. Genotypes can be tentative assigned: TT, *H. r.* var. *serpyllifolium*; Tt, F_1 hybrid; tt, *H. drummondii* (Figs. 7 and 10).

As Factor One segregates as a highly intercorrelated unit rather than as individual compounds, the Mendelian pattern of inheritance more likely mirrors the inheritance of a common precursor rather than individual compounds. We could thus envision a single gene or group of tightly linked genes governing this intermediate and the production of individual compounds mediated by nonspecific enzyme steps (Fig. 7). This is, however, only one model out of perhaps several. For example, one could also suggest with the data at hand, a series of linked genes for specific enzymes which segregate and act in concert (Fig. 7). Moreover, as the correlation of Factor 1 compounds was not perfect and each compound was occasionally transgressive over the parent and F_1, a genetic model will necessarily need to include epistasis and/or modifying genes to account fully for the quantitative variation of these compounds in the F_2 generation.

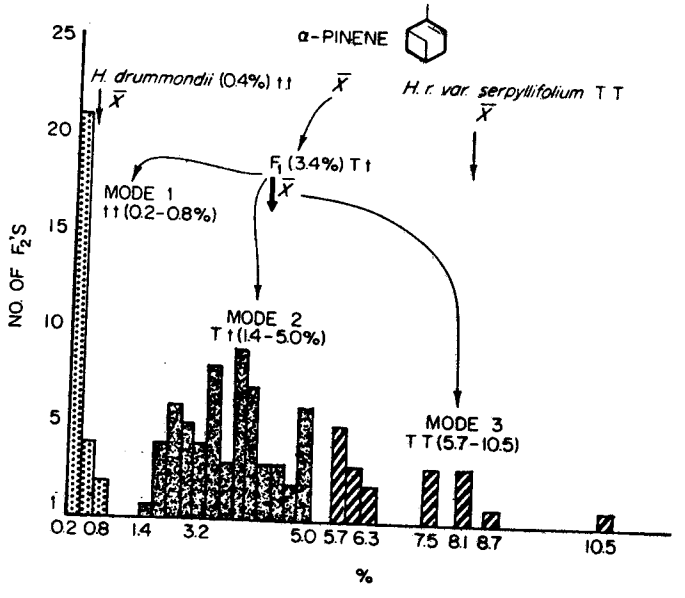


FIG. 10. Frequency of α-pinene in F_2 generation and genetic model of bicyclic inheritance. T is dominant to t.

TABLE 4

CHI-SQUARE ANALYSIS OF THE F_2 SEGREGATION OF α-PINENE, REPRESENTATIVE OF FACTOR 1

Mode	No.	Expected ratio	Actual ratio	X^2
1	27	1	1.0	
2	61	2	2.3	$P = 0.14$
3	18	1	0.7	
2 + 3	79	3	3.0	$P = 1.0$

FACTOR TWO

Myrcene, sabinene, 1,8-cineole, terpinene, terpinolene, and unknowns 16 and 19, are diverse monoterpenes but display a unified inheritance pattern. Thus we are again possibly monitoring the genetics of a common precursor. The frequency distribution of sabinene (Fig. 11) is representative of this group. The distribution is possibly pentamodal. Its inheritance, then, is not that of single gene action, but neither is it that of a large number of genes. A number of genetic models could be developed for Factor 2 using two to four linked or nonlinked genes. Such models would necessarily have to satisfy a number of conditions, however: (1) all compounds in Factor 2 are in at least one or two F_2 plants transgressive to parental variation; (2) the F_2's have a narrow range of variance suggesting a uniform genotype; (3) the parents because of their inbreeding and production of a uniform F_1 should be homozygous for all loci.

A tentative and simple model using two diallelic genes A,a, and B,b can be formulated (Fig. 11). These would be located on separate chromosomes, with A slightly epistatic to B. *H. drummondii* would be represented in the model as aaBB and would produce low quantities of the Factor 2 compounds through the absence of the A allele. *H. r. serpyllifolium* would also be homozygous, AAbb, but would produce significant quantities through AA. A cross of these two homozygous genotypes would produce a genetically uniform F_1, AaBb. Selfing of the latter would yield nine different genotypes (a dihybrid cross) which could be grouped into the five modes of the F_2 frequency distribution and in the approximate ratios (Fig. 11). This array of genotypes in the F_2 would show an increasing production of sabinene (or other Factor 2 compound) with increasing proportions of A, primarily, and B, secondarily, in the genotype. Thus, AABb and AABB would represent the transgressive F_2 genotypes. The expected ratio of such a

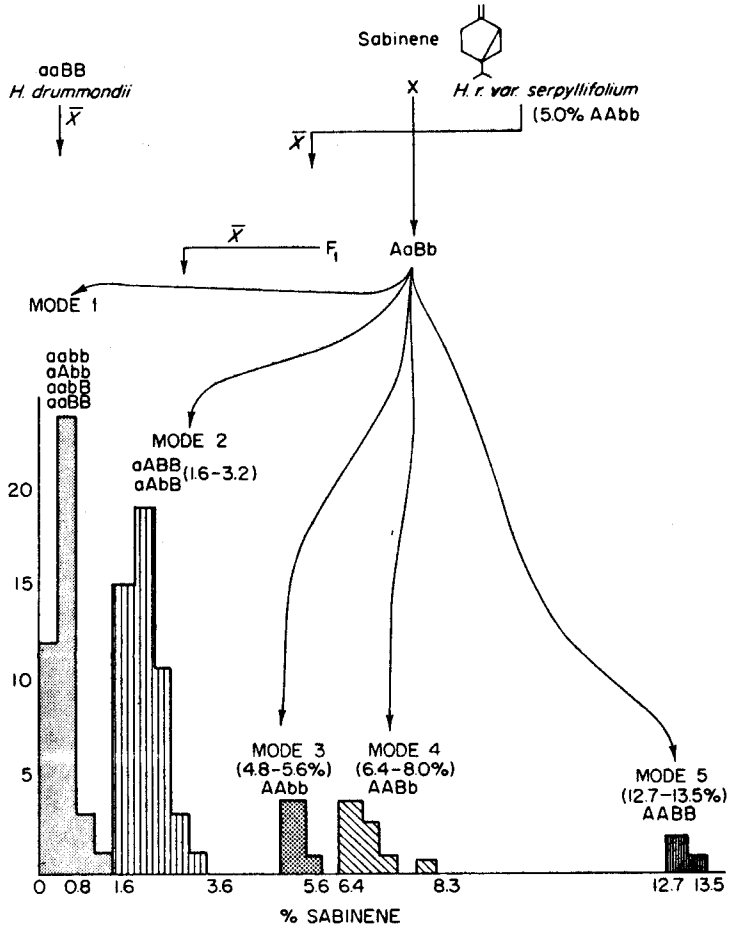


FIG. 11. Frequency of sabinene in F_2 generation and genetic model of inheritance for Factor 2. Model is based on two genes (A,B) with two alleles (A, a and B, b); these are located on separate chromosomes with A slightly epistatic to B. Quantitative level of Factor 2 compound increases with increasing proportion of A in genome (see text).

model, 6:6:1:2:1, approximates the actual ratio obtained (Table 5). No doubt, a better fit could be obtained using more complicated models, but in view of the evidence, further statistical gamesmanship seems unprofitable.

d-LIMONENE

Despite *d*-limonene's seeming importance in terpene biosynthesis, little can be said of its genetic base from our data. The parents showed no sig-

TABLE 5

CHI-SQUARE ANALYSIS OF THE F_2 SEGREGATION OF SABINENE, REPRESENTATIVE OF FACTOR 2

Mode	No.	Expected ratio	Actual ratio	X^2
1 (0–1.2)	40	6	6.1	
2 (1.6–3.2)	49	6	7.4	
3 (4.8–5.6)	5	1	0.8	$P = 0.25$
4 (6.4–8.0)	9	2	1.4	
5 (12.7–13.5)	3	1	0.5	

nificant difference in quantitative levels (3.6 and 4.5 percent). The F_1 and F_2 populations displayed widely varying levels of *d*-limonene which were for the most part transgressive. The F_2 frequency distribution (Fig. 12) offered little suggestion of the *d*-limonene's inheritance save that is perhaps ogligogenic. However, much of the F_2 variation might be due to varying levels of biosynthetic competition as it is perhaps a key compound in monoterpene interconversions.

PULEGONE, ISOMENTHONE, AND MENTHONE

Individually these ketones do not display interpretable frequency patterns, at least, in Mendelian terms. If they are treated as a unit, however

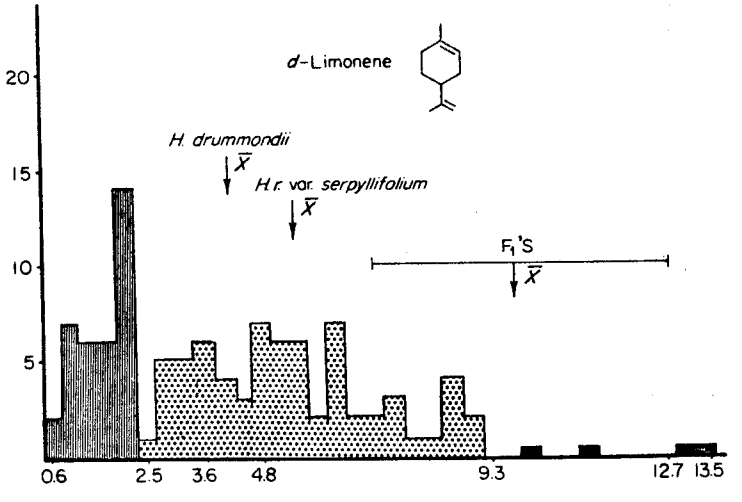


FIG. 12. Frequency of *d*-limonene in F_2 generation.

(Fig. 13), a classic 1:2:1 ratio emerges (Table 6), which suggests one or few tightly linked genes segregating as a Mendelian unit. Using such a model, genotypes could be as follows: kk, *H. r. serpyllifolium*; Kk, F_1, and KK, *H. drummondii*. These alleles would presumably govern the level of a cyclic precursor which would move through a reduction series (Fig. 7).

A Statistical Estimate of Gene Number

Although frequency plots of large controlled F_2 generations are a standard analytical tool for determining genetic background, the minimal num-

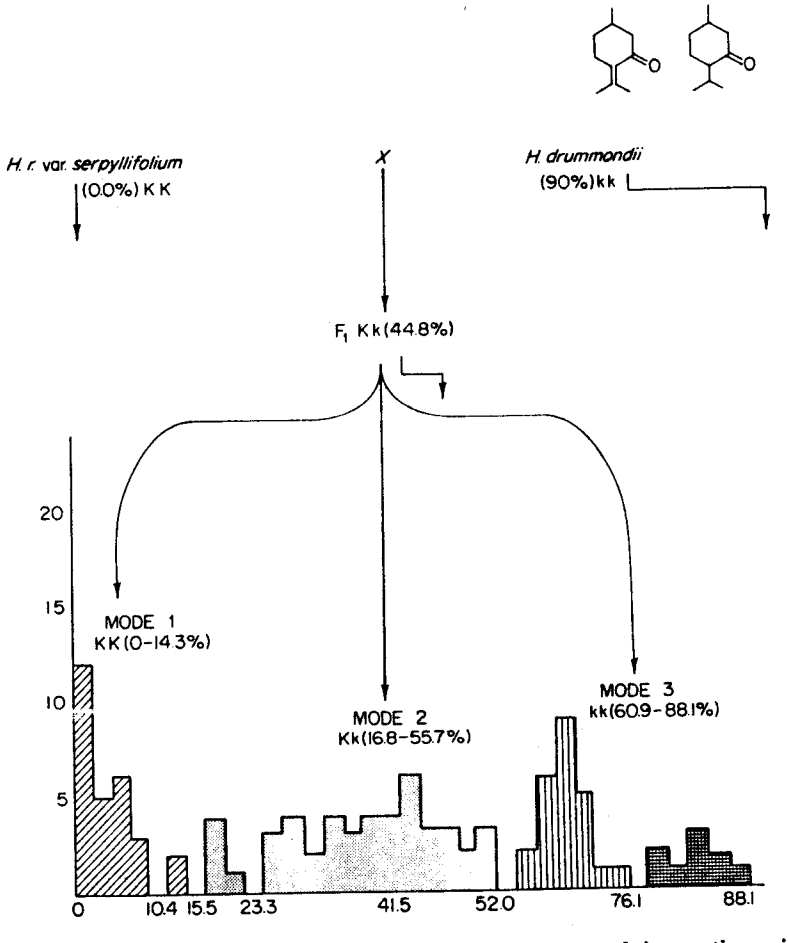


Fig. 13. Frequency distribution of pulegone, menthone, and isomenthone in F_2 generation and accorded genetic model based on one gene with two alleles.

TABLE 6

CHI-SQUARE ANALYSIS OF THE F_2 SEGREGATION OF PULEGONE, ISOMENTHONE, AND MENTHONE

Mode	No.	Expected ratio	Actual ratio	X^2
1	28	1	1.0	
2	45	2	1.8	$P = 0.24$
3	33	1	1.2	
2 + 3	78	3	2.8	$P = 0.9$

ber of genes can also be determined using the formula developed by Sewall Wright (1934) (see also Charles and Goodwin, 1953):

$$N = \frac{(V_{p2} - V_{p1})^2}{8(\sigma_2^2 - \sigma_1^2)}$$

TABLE 7

ESTIMATE OF MINIMAL NUMBER OF GENES

Compound	N
α-Pinene	2
Camphene	3
β-Pinene	2
Sabinene	1
Myrcene	1
d-Limonene	1
1,8-Cineole	4
Terpinene	1
Terpinoline	1
Compound 10	4
Menthone	1
Isomenthone	7
Compound 13 ketone	1
Compound 14	1
Tricyclene	3
Compound 16	1
Pulegone	1
Borneol	2
Compound 19	1
Compound 20	1
Compound 21	0

In the formula, V_{p1} and V_{p2} represent the average of the two parents with respect to a monoterpene character and σ_2, σ_1 the standard deviations of the F_2 and F_1. The validity of N, the minimal number of genes, depends on the number of conditions which must be met by the experimental breeding program (see Charles and Goodwin, 1953).

Making these calculations for the 21 terpenes (Table 7), we have an approximate estimate of minimal genes within the limitations of the formula. In general the results indicate that most terpenes are controlled by a small number of genes. In the specifics there is, however, variance between the two approaches. The strength of Wright's formulation depends, for the most part, on large parental and F_1 sample sizes. In our program we are not able to provide this prerequisite adequately, and these minimal estimates must be treated as only approximations.

Discussion

The problems in the genetics and biosynthesis of monoterpenes are far from solved. In *Hedeoma*, and possibly in other systems, a substantial portion of their biogenetic base is seemingly simplistic, involving a few major genes (Fig. 7). This is in keeping with inferences from several researches (Forde, 1964; Hanover, 1966b; Zavarin *et al.*, 1969). Our data suggest further simplification, but we must proceed with caution in this reduction of complexity; although having their uncomplicated facets, biological systems, as dynamic wholes, may be exceedingly complex. 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. Because of this, in our biogenetic synopsis (Fig. 7) a number of additional and/or alternative sites of gene action have been designated.

Moreover, to evaluate properly genetics through breeding, the organism's limitations must be realized and the nature of the character studied must be understood. Patterns of variation in an F_2 generation take on genetic significance only when there is a cellular environment conducive to free recombination. That is, there must be a substantial degree of normal chromosomal pairing and chromatid exchange in the F_1 parent. In our work, although the F_1 had a lowered fertility, there was a high degree of normal pairing as well as a number of meiotic irregularities (univalent formation). There is then the indication of a partial suppression of recombination, and our conclusions must be tempered accordingly.

To understand the monoterpene as a character, the gap between the gene and its expression must be closed. There must be greater certainty in the

proposed series of interconversions and enzymatics. Lacking this, our genetic and biogenetic models can only stand as guides.

The systematic implications of our data are of considerable importance. Ideally, systematists value those chemical characters which, through their biosynthetic and genetic complexity, can be used to delineate clearly reproductive barriers as well as to provide ample variation with gene flow. The monoterpenes of the *H. drummondii* group, through their perhaps deceptive simplicity, can only partially fulfill these goals. Considering the F_2 segregation, complex patterns of natural interbreeding could only be partly detected by monoterpene analysis. To delimit species by employing these constituents, and their proposed genetic base, would also be dubious. Nevertheless, one of the most important findings in our experimental work was the independent segregation of morphology and terpenes. Thus, an excellent mirror of gene flow would be gained utilizing monoterpene diversity *and* morphology. Moreover, if we are indeed, with monoterpenes, investigating a relatively uncomplicated biogenetic system, it seems that the hope of using biochemical pathways as evolutionary guideposts may well be within our grasp.

Acknowledgments

This work was supported by a National Science Foundation Grant (GB-12910) awarded to the senior author; computer time was furnished by Colorado State University. We wish also to express our sincere thanks to Mrs. Caryn Stone for her technical assistance and patience. Thanks also go to Drs. David V. Clark, B. L. Turner, E. von Rudloff, and Victor C. Runeckles for their helpful discussions, suggestions, and support. We are grateful to others who assisted in many ways: Mr. David Firmage for his laboratory and field work, Linda Nimmo and Evalene Wyatt for their assistance in the preparation of the manuscript, and, to the senior author, to Jody Lubrecht, and to Joellen Irving for their support and encouragement.

References

Adams, R. P. 1970. *Phytochemistry* **9**:397–402.

Alston, R. E., and B. L. Turner. 1963a. "Biochemical Systematics." Prentice-Hall, Englewood Cliffs, New Jersey.

Alston, R. E., and B. L. Turner. 1963b. *Amer. J. Bot.* **50**:159–173.

Alston, R. E., H. Rosler, K. Naifeh, and T. J. Mabry. 1965. *Proc. Nat. Acad. Sci. U.S.* **54**:1453–1465.

Battu, R. G., and H. W. Youngken, Jr. 1966. *Lloydia* **29**:360–367.

Burbott, A. J., and W. D. Loomis. 1967. *Plant Physiol.* **42**:20.

Charles, D. R., and R. H. Goodwin. 1953. *Amer. Natur.* **77**:53–69.

Critchfield, W. B. 1966. *U.S. Forest Serv., Res. Pap.* **NC-6**:34–44.

Critchfield, W. B. 1967. *Silvae Genet.* **16**:89–97.

DeWet, J. M. J. 1967. *Amer. J. Bot.* **54**:384–387.

DeWet, J. M. J., and B. D. Scott. 1965. *Bot. Gaz. (Chicago)* **126**:209–214.

Emboden, W. A., and H. Lewis. 1967. *Brittonia* **19**:152–160.
Firmage, D. 1971. Unpublished observations, Univ. of Montana, Missoula, Montana.
Firmage, D., and R. S. Irving. 1971. *Annu. Phytochem. Soc. Meet., 11th, Monterrey, Mexico.*
Forde, M. B. 1964. *N. Z. J. Bot.* **2**(1): 53–59.
Fujita, Y. 1965a. *Shokubutsugaku Zasshi* 78(924):212–219.
Fujita, Y. 1965b. *Shokubutsugaku Zasshi* 78(925):245–252.
Fujita, Y. 1965c. *Osaka Kogyo Gijutsu Shikenjo Hokoku* No. 306–2.
Habeck, J. R., and T. W. Weaver. 1969. *Can. J. Bot.* **47**:1565–1570.
Hanover, J. W. 1966a. *Heredity* **21**:73–84.
Hanover, J. W. 1966b. *Forest Sci.* **12**:447–450.
Irving, R. S. 1968. Ph.D. Thesis, Univ. of Texas, Austin, Texas.
Kremers, R. E. 1922. *J. Biol. Chem.* **50**:31.
Loomis, W. D. 1969. *In* "Terpenoids in Plants" (J. B. Pridham, ed.). Academic Press, New York.
Loomis, W. D. 1972. *In* "Terpenoids: Structure, Function and Biogenesis" (V. C. Runeckles and T. J. Mabry, eds.). Academic Press, New York.
Mirov, J. T. 1956. *Can. J. Bot.* **34**:443–457.
Murray, M. J. 1960a. *Genetics* **45**:925–929.
Murray, M. J. 1960b. *Genetics* **45**:931–937.
Reitsema, R. H. 1958. *J. Amer. Pharm. Ass.* **47**:267–269.
Ruzicka, L. 1953. *Experientia* **9**:357–396.
Sandermann, W. 1962. *Holzforschung* **16**:65–74.
Scora, R. W. 1967a. *Amer. J. Bot.* **54**:446–452.
Scora, R. W. 1967b. *Taxon* **16**:499–505.
Scora, R. W. 1970. *Taxon* **19**:215–228.
Scora, R. W., and J. D. Mann. 1967. *Lloydia* **30**:236–241.
Smith, D. M., and D. A. Levin. 1963. *Amer. J. Bot.* **50**:952–958.
Stone, D. E. 1964. *Amer. J. Bot.* **51**:687–688.
Stone, D. E., G. A. Adrouny, and S. Adrouny. 1965. *Brittonia* **17**:97–106.
Stone, D. E., G. A. Adrouny, and R. H. Flake. 1969. *Amer. J. Bot.* **56**:928–235.
Turner, B. L. 1967. *Pure Appl. Chem.* **14**:189–213.
Turner, B. L. 1969. *Taxon* **18**:134–151.
Wright, S. 1934. *Genetics* **19**:537–551.
Zavarin, E. 1970. *Phytochemistry* **9**:1049–1063.
Zavarin, E., and F. W. Cobb, Jr. 1970. *Phytochemistry* **9**:2509–2515.
Zavarin, E., W. B. Critchfield, and K. Snajberk. 1969. *Can. J. Bot.* **47**:1443–1453.
Zavarin, E., L. Lawrence, and M. C. Thomas. 1971. *Phytochemistry* **10**:379–393.