MONOTERPENE PHENOTYPES IN LOBLOLLY PINE POPULATIONS: NATURAL SELECTION TRENDS AND IMPLICATIONS

A.E. Squillace, Harry R. Powers, Jr., and S. V. Kossuth

Abstract.-- The degree of discrepancy between observed proportions of various monoterpene phenotypes involving two or more loci and those expected under random association between loci was studied in 111 populations scattered throughout the range of loblolly pine. Results suggest that some phenotypes are being favored by natural selection while others are being disfavored. Natural selection varies among regions and appears related to variation in resistance to fusiform rust.

Additional keywords: Linkage disequilibrium, fusiform rust.

Previous work has shown that contents of four of the major monoterpenes in cortical oleoresin of loblolly pine (*Pinus taeda* L.) are largely controlled by single genes, with high content being dominant over low in all cases (Squillace et al. 1980). The four loci involved have also been shown to be rather closely linked (Squillace and Swindel, in press). The objectives of the present study are to: (1) examine deviations between observed proportions of phenotypes involving two or more monoterpene loci and those expected under random association between alleles at different loci (linkage disequilibrium) in loblolly pine populations, (2) interpret the results from the standpoint of natural selection, and (3) seek relationships with resistance to fusiform rust (*Cronartium quercuum* f. sp. *fusiforme*).

MATERIALS AND METHODS

In this study we utilized data previously reported by Squillace and Wells (1981) and McRae and Thor (1982). In those reports, cortical monoterpenes were sampled in trees from a total of 111 loblolly pine populations scattered throughout the species range (Table 1). Most of the trees were in seed source study plantations containing trees of known geographic origin. Details on sampling and analytical techniques are available in the two publications cited. In some of the plantations sampled by Squillace and Wells (1981), data on occurrence of fusiform rust infection were also available and were utilized in interpreting results of the disequilibrium analyses.

---

[1] Squillace is Adjunct Professor, University of Florida, and Volunteer, USDA Forest Service, Gainesville, FL. Powers and Kossuth are Project Leaders, Southeastern Forest Experiment Station, USDA Forest Service, Athens, GA and Gainesville, FL, respectively. The study was also sponsored by the Georgia Forestry Commission. Journal Series Paper 6469 of the Florida Agricultural Experiment Station.

Table 1.--Numbers of loblolly pine populations and trees sampled for cortical monoterpene composition by regions of the southern U.S.

Region[a]	Populations	Trees
Western	21	565
Central	85	2141
Northeastern	5	143
Totals	111	2849

[a] Western = populations west of the Mississippi River.
Central = populations between western and northeastern regions.
Northeastern = populations in Virginia, Maryland, and Delaware.

In order to study linkage disequilibrium between pairs of loci, we first determined the frequencies of trees in each phenotypic class. Since dominance occurs at each locus, there are four possible phenotypes for each pair of monoterpenes. For example, designating two monoterpenes as A and B, with capital letters representing high amounts and lower case letters as low amounts, the four possible phenotypes are:

1. AB, which includes genotypes AABB, AABb, AaBB and AaBb
2. Ab, which includes genotypes AAbb and Aabb
3. aB, which includes genotypes aaBB and aaBb
4. ab, which includes only aabb genotypes

An estimate of linkage disequilibrium, $\hat{D}$, is given by the following (Cavalli-Sforza and Bodmer 1971, Hill 1974)

$$\hat{D} = \sqrt{\frac{x_4}{n}} - \frac{\sqrt{(x_2 + x_4)\ (x_3 + x_4)}}{n}$$

in which x_2, x_3, and x_4 are numbers of phenotypes Ab, aB, and ab in the population, respectively, and n = total number.

If the product of the AB and ab phenotypes (coupling types) exceeds the product of Ab and aB phenotypes (repulsion types), D will be positive. If the reverse is true, D will be negative. A positive D indicates that the proportion of coupling types observed is greater than that expected from random association. It can mean, for example, that natural selection is favoring coupling over repulsion types. A negative D would suggest the reverse.

To test the significance of $\hat{D}$, we used the likelihood criterion (K) given in Hill (1974):

$$K = \frac{4\ n\ \hat{D}^2}{\hat{p}\ (2-\hat{p})\ \hat{q}\ (2-\hat{q})}$$, which is a χ^2 distribution, with 1 d.f.,

where $\hat{p}$ = estimated frequency of the A allele = $1- \sqrt{\frac{x_3 + x_4}{n}}$,

and $\hat{q}$ = estimated frequency of the B allele = $1- \sqrt{\frac{x_2 + x_4}{n}}$.

As an example, $\hat{D}$ and K will be computed for β-pinene vs. myrcene in population #2 (Marion Co., FL). The numbers of phenotypes BM, Bm, bM, and bm were 14, 2, 6, and 1, respectively (summed from the second column of table 2). Substituting these values into the above equations, we obtain: $\hat{D} = .009$, $\hat{p} = .448$; $\hat{q} = .639$, and $k = .01$.

Since our data involved four loci, we were also interested in determining evidence of disequilibrium that may occur among four-locus phenotypes. With four loci showing dominance, there are 16 possible phenotypes as shown in Table 2. We could find no procedure in the literature for estimating and testing linkage disequilibrium in such cases and hence used the following procedure to get indications of natural selection favoring or disfavoring each phenotype. Expected frequencies (proportions) of each phenotype were computed on the basis of frequencies of single-locus phenotypes and these were subtracted from observed phenotypic frequencies. We shall designate the differences by D'.

Table 2.--Computation of observed-expected proportions of four-gene phenotypes (D'), for Population #2, Marion Co., FL

Phenotype [a]	Observed #	Observed prop.	Expected prop. [b]	D'
BMLP	2	.087	.101	-.014
BMLp	0	.000	.005	-.005
BMℓP	11	.478	.478	.000
BMℓp	1	.043	.022	.022
BmLP	1	.043	.015	.028
BmLp	0	.000	.001	-.001
BmℓP	1	.043	.072	-.028
Bmℓp	0	.000	.003	-.003
bMLP	0	.000	.044	-.044
bMLp	0	.000	.002	-.002
bMℓP	6	.261	.209	.052
bMℓp	0	.000	.010	-.010
bmLP	1	.043	.007	.037
bmLp	0	.000	.000	.000
bmℓP	0	.000	.031	-.031
bmℓp	0	.000	.001	-.001
Totals	23	.998	1.001	.000

[a] B, M, L, and P represent high amounts of β-pinene, myrcene, limonene, and β-phellandrene, respectively, while lower case letters represent low amounts.

[b] Computed from observed proportions of one-gene phenotypes, summed from column three, above: B = .696, b = .304, M = .870, m = .130, L = .174, ℓ = .826, P = .956, and p = .044. Thus, for example, the expected proportion of BMLP trees is (.696)(.870)(.174)(.956) = .101. See text.

Thus, an estimate of D' for phenotype BMLP was computed as:

$$D'_{BMLP} = f_{BMLP} - f_B \, f_M \, f_L \, f_P,$$

in which f_{BMLP} is the observed frequency of the BMLP phenotype in the population, and f_B, f_M, f_L and f_P are the observed frequencies of B, M, L, and P phenotypes in the same population, respectively. D' values were likewise obtained for the other 15 phenotypes:

$$\begin{aligned} D'_{BMLp} &= f_{BMLp} - f_B \, f_M \, f_L \, f_p, \\ D'_{BM\ell P} &= f_{BM\ell P} - f_B \, f_M \, f_\ell \, f_P, \\ &\vdots \\ D'_{bm\ell p} &= f_{bmlp} - f_b \, f_m \, f_\ell \, f_p. \end{aligned}$$

An example is given in Table 2. The above procedure was also used for getting estimates of disequilibrium for groups of populations (regions). It should be noted that expected four-locus phenotypic frequencies can also be computed from observed frequencies of individual alleles--rather than using observed frequencies of individual phenotypes--with the same result, but the latter procedure is simpler.

The numbers of trees in phenotypes within individual populations were too few to make reliable tests of significance because in most populations many phenotypes were represented by four or less trees, preventing reliable use of chi-square tests. However, we tested significance of average differences for each phenotype across all populations, using a t-test with the null hypothesis that the observed-expected values = zero. When pooling populations within regions, we tested for significance by computing

$$\frac{(\text{observed number-expected number})^2}{\text{expected number}}$$

for each phenotypic class in which expected numbers were five or more and summing these to obtain X^2. Degrees of freedom here were presumed to be total number of four-locus phenotypes minus number of single-locus phenotypes used in computing expected values.

RESULTS

Two-locus Phenotypes

Significant pair-wise linkage disequilibrium was found rather frequently (Table 3). Myrcene vs. limonene was especially notable--84 of the 98 populations permitting this test showed negative D values, with 31 of them being significant. None of the positive values were significant. Curiously, the populations showing positive values were clustered in three areas: southwestern Alabama, southeast Georgia-northeast Florida, and the Carolinas. The results suggest that Mt and mL phenotypes are being favored by natural selection, while ML and mi types are being disfavored in most portions of the species range. Note also that '1) values fore -pinene vs. myrcene and for myrcene vs. f3-phellandrene also tended to be negative, judging from both numbers of negative vs. positive values and significance (Table 3). These findings suggest that BmLP trees are being favored by natural selection, which will be examined further.

Table 3.--Results of linkage disequilibrium ($\hat{D}$) analyses in 111 loblolly pine populations

Loci compared	Range of $\hat{D}$ values	Positive $\hat{D}$ values		Negative $\hat{D}$ values	
		Total no. [a]	No. significant [b]	Total no. [a]	No. significant [b]
β-pinene & myrcene (B,M)	0.08 to -0.27	17	0	50	16
β-pinene & limonene (B,L)	.13 to - .27	43	0	17	8
β-pinene & β-phellandrene (B,P)	.22 to - .23	8	1	43	2
Myrcene & limonene (M,L)	.07 to - .37	14	0	84	31
Myrcene & β-phellandrene (M,P)	.15 to - .27	31	1	46	8
Limonene & β-phellandrene (L,P)	.17 to - .27	48	1	22	8

[a] The total number of positive and negative values in each comparison are less than 111 because in many populations both the observed and expected proportions of phenotypes were zero, negating a test.

[b] Significant at 0.05.

Four-locus Phenotypes

Analyses of four-locus phenotypes in individual populations showed that observed frequencies of some phenotypes exceeded expected values while for others the reverse was true (Table 4). Phenotype BmLP had both the greatest positive average deviation and the greatest proportion of populations showing positive deviations. Values obtained for this phenotype are plotted in Figure 1. Note that positive values prevail in all areas except southwest Alabama, southeast Georgia-northeast Florida, and the Carolinas. Thus, as suggested earlier, BmLP phenotypes seem to be favored by natural selection over most portions of the species range. Phenotype BMtP also showed a significant trend toward positive deviations. Finally, the relatively rare phenotype bMtP also showed a significant but small positive average deviation. These results may be partly a reflection of the very strong negative disequilibrium between myrcene and limonene (Mt and mL types being favored), noted earlier.

Table 4.--Results of analyses of observed vs. expected proportions of each monoterpene phenotype, within populations

Phenotype [a]	No. of populations in test [b]	Percent of populations showing + deviations	Average deviation over all populations	t value [c]
BMLP	102	18	-.039	9.38**
BMLp	71	19	-.005	2.44*
BMℓP	109	68	.030	5.73**
BMℓp	79	55	.004	1.12
BmLP	100	82	.043	9.79**
BmLp	70	22	.002	.88
BmℓP	107	25	-.034	7.33**
Bmℓp	78	31	-.001	.24
bMLP	61	21	-.001	.52
bMLp	45	6	.001	.72
bMℓP	67	66	.012	3.64**
bMℓp	52	21	-.001	.67
bmLP	60	14	-.001	.57
bmLp	45	4	.001	.56
bmℓP	66	12	-.010	4.92**
bmℓp	52	2	-.001	1.68

[a] B, M, L, and P represent high amounts of β-pinene, myrcene, limonene, and β - phellandrene, respectively, while lower case letters represent low amounts.

[b] These values are less than the total (111 populations) because in many cases both the observed and expected proportions were zero, in which case no test was possible.

[c] Test of the hypothesis that the population average is zero, with N-1 d.f. An * = significant at the 0.05 level; ** = significant at the 0.01 level.

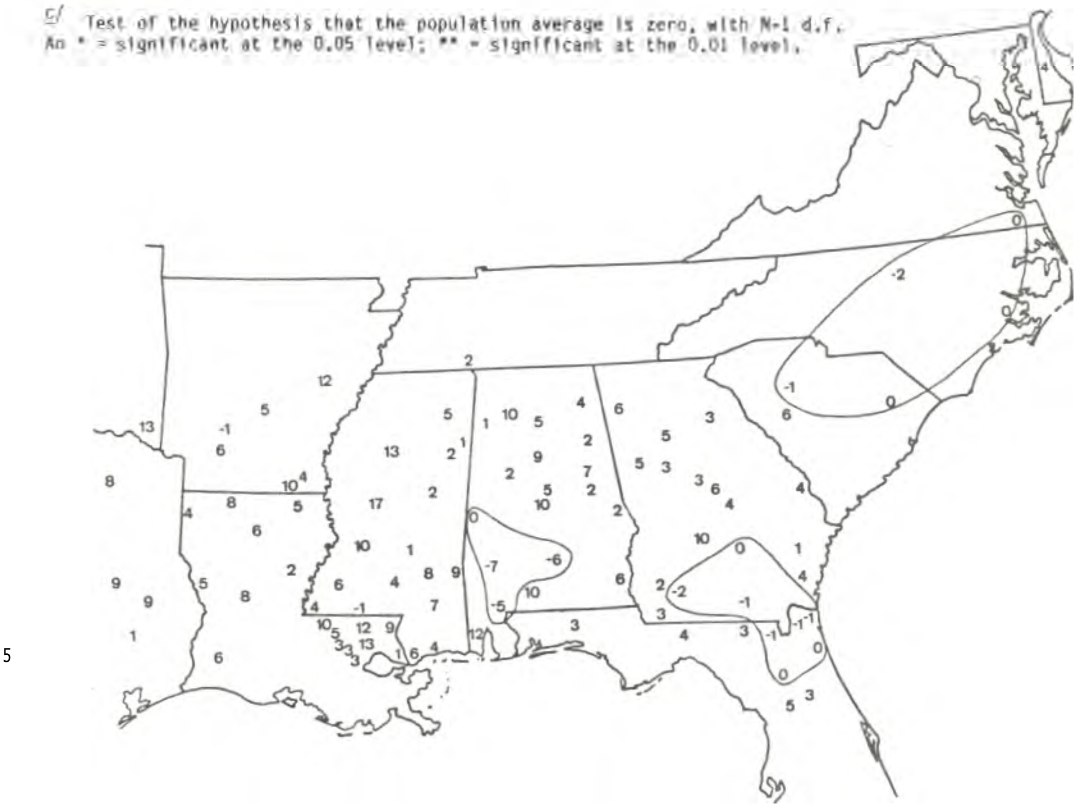
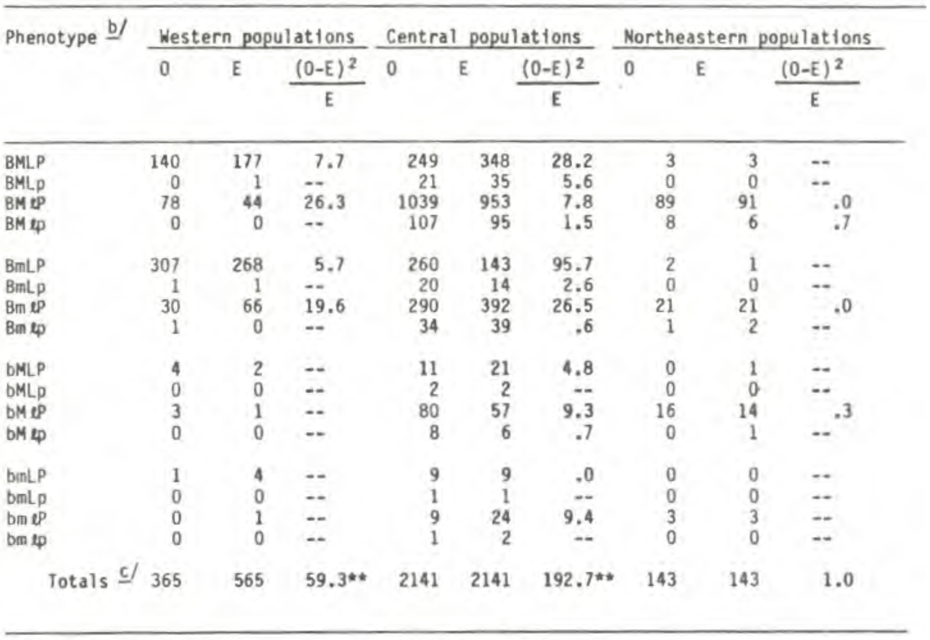


Figure 1.--Observed-expected percent of BmLP phenotypes. Clusters of zero and negative values are outlined.

Analyses in which populations were pooled by regions gave similar results (Table 5). However, the deviations of observed vs. expected values were generally largest in the central region, smaller in the West, and very small in the northeast.

Table 5.-- Observed (0) and expected (E) numbers of trees in each of 16 monoterpene phenotypes within regions[a/] and tests of significance



Phenotype [b/]	Western populations			Central populations			Northeastern populations		
	0	E	$\frac{(0-E)^2}{E}$	0	E	$\frac{(0-E)^2}{E}$	0	E	$\frac{(0-E)^2}{E}$
BMLP	140	177	7.7	249	348	28.2	3	3	--
BMLp	0	1	--	21	35	5.6	0	0	--
BMℓP	78	44	26.3	1039	953	7.8	89	91	.0
BMℓp	0	0	--	107	95	1.5	8	6	.7
BmLP	307	268	5.7	260	143	95.7	2	1	--
BmLp	1	1	--	20	14	2.6	0	0	--
BmℓP	30	66	19.6	290	392	26.5	21	21	.0
Bmℓp	1	0	--	34	39	.6	1	2	--
bMLP	4	2	--	11	21	4.8	0	1	--
bMLp	0	0	--	2	2	--	0	0	--
bMℓP	3	1	--	80	57	9.3	16	14	.3
bMℓp	0	0	--	8	6	.7	0	1	--
bmLP	1	4	--	9	9	.0	0	0	--
bmLp	0	0	--	1	1	--	0	0	--
bmℓP	0	1	--	9	24	9.4	3	3	--
bmℓp	0	0	--	1	2	--	0	0	--
Totals [c/]	365	565	59.3**	2141	2141	192.7**	143	143	1.0

[a/] Western = populations west of the Mississippi River.
Central = populations between western and northeastern regions.
Northeastern = populations in Virginia, Maryland, and Delaware.

[b/] B, M, L, and P represent high amounts of β-pinene, myrcene, limonene, and β-phellandrene, respectively, while lowercase letters represent low amounts.

[c/] The totals of $(0-E)^2/E$ are chi-squares, with 12 d.f. Classes having fewer than five expected values were omitted in computing χ^2.

** = significant at the 0.01 level.

IMPLICATIONS

The results strongly suggest that natural selection in most portions of the range of loblolly pine is favoring certain monoterpene phenotypes and disfavoring others. Here we examine possible reasons.

Note first that the phenotype which seems to be most stroogly favored by selection (BmLP) is very prevalent in the western region, $^{307}/_{565}$ = 54 percent being of this type (Table 5). It is much less prevalent in the central region (12 percent) and almost absent in the northeast. The geographic pattern is shown more clearly in Figure 2. Note further that in many respects it conforms to regional patterns of resistance to fusiform rust (Grigsby 1973, Squillace and Wells 1981). Western trees have large proportions of BmLP trees and are relatively resistant, and this relationship tends to extend into western Mississippi, although to a lesser extent. Populations in the panhandle of Florida and in southeast Georgia-northeast Florida have few BmLP trees and are very susceptible to rust. The pattern fails in the northeast where BmLP trees are scarce and resistance is high. We shall return to this point later.

Figure 2.--Percent of BmLP phenotypes.

With these observations in mind, we hypothesize that natural selection is favoring BmLP trees because they tend to be more resistant to fusiform rust than other types. Data from progeny tests reported by Squillace et al. (1984) also suggest that BmLP trees tend to be more resistant than other types. Perhaps natural selection has been favoring such trees in the West over a longer period than in the central region, explaining its greater prevalence in the West.

A somewhat similar situation seems to exist for phenotype bMLP, which also tended to be favored by selection. The proportion of such trees was 11.2 percent in the northeast, 3.7 percent in the central region, and 0.5 percent in the West (Table 5). As is well known, northeastern populations tend to be relatively resistant [see, for example Grigsby (1973)]. Thus, natural selection for this phenotype may also be a reflection of resistance to rust.

CONCLUSIONS

Natural selection is definitely favoring BmLP trees in both the west and central regions of loblolly pine. This phenotype presently comprises a large proportion of Western populations, which are relatively resistant to fusiform rust. Although it is now rather infrequent in central populations, it is presumably increasing with each generation. We hypothesize that BmLP trees tend to be more resistant to fusiform rust and that this is why they are favored by natural selection. A similar situation seems to occur for bMLP trees, although this is less certain. The latter phenotype is relatively most prevalent in the northeast, where resistance to rust also occurs and it is presumably being favored by natural selection, especially in central populations.

The nature and degree of the relationship between monoterpenes and fusiform rust resistance is still unclear. Although significant relationships were found on a regional basis (Squillace and Wells 1981), we have not yet completed studies comparing trees of different phenotypes within families and populations. We do not believe monoterpene composition actually affects resistance--more likely it may be an indicator of the presence of some types of resistance. It is possible, for example, that the monoterpene phenotypes being favored by natural selection are results of hybridization or introgression with other species, such as shortleaf pine (P. echinata Mill.)(Hare and Switzer 1969) and pond pine (P. serotina Michx.)(Saylor and Kang 1973). Studies of cortical monoterpenes in these species would be desirable. Also, it would be desirable to artificially inoculate trees having different monoterpene phenotypes and observe their reactions to the rust.

LITERATURE CITED

Cavalli-Sforza, L.L. and W.F. Bodmer. 1971. The genetics of human populations. Freeman & Co., San Francisco, CA. 965 p.

Grigsby, N.C. 1973. South Carolina best of 36 loblolly pine seed sources for Southern Arkansas. USDA For. Serv. Res. Pap. SO-89. 10 p.

Hare, R.C. and G.L. Switzer. 1969. Introgression with shortleaf pine may explain rust resistance in western loblolly pine. U.S. Forest Serv. Res. Note SO-88. 2 p.

Hill, W.G. 1974. Estimation of linkage disequilibrium in randomly mating populations. Heredity 33:229-239.

McRae, J. and E. Thor. 1982. Cortical monoterpene variation in 12 loblolly pine provenances planted in Tennessee. For. Sci. 28:732-736.

Saylor, L.C., and K.W. Kang. 1973. A study of sympatric populations of *Pinus taeda* L. and *Pinus serotina* Michx. in North Carolina.

Squillace, A.E., H.R. Powers, Jr., and S.V. Kossuth. 1984. Relationships between cortical monoterpenes and fusiform rust resistance in loblolly pine. Abstract in Southwide Forest Disease Workshop, Long Beach, MS.

Squillace, A.E. and B.F. Swindel. In press. Linkage among genes controlling monoterpene constituent levels in loblolly pine. (Accepted by Forest Science)

Squillace, A.E., and O.O. Wells. 1981. Geographic variation of monoterpenes in cortical oleoresin of loblolly pine. Silvae Genetica 30: 127-135.

Squillace, A.E., O.O. Wells, and D.L. Rockwood. 1980. Inheritance of monoterpene composition in cortical oleoresin of loblolly pine. Silvae Genetica 29:141-151.