MICHIGAN STATE UNIVERSITY LIBRARIES
3 1293 01019 0951

This is to certify that the

thesis entitled

CLADISTIC AND PHENETIC RELATIONSHIPS WITHIN CYPRIPEDIUM (ORCHIDACEAE) INFERRED FROM FLORAL FRAGRANCE-COMPOUND DATA

presented by

Todd James Barkman

has been accepted towards fulfillment
of the requirements for

M.S. degree in Botany

S. H. Bowman
Major professor

Date 9 August 1993

O-7639 MSU is an Affirmative Action/Equal Opportunity Institution

LIBRARY
Michigan State
University

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.

DATE DUE	DATE DUE	DATE DUE

MSU Is An Affirmative Action/Equal Opportunity Institution
c:\circ\datedue.pm3-p.1

CLADISTIC AND PHENETIC RELATIONSHIPS WITHIN *CYPRIPEDIUM* (ORCHIDACEAE) INFERRED FROM FLORAL FRAGRANCE-COMPOUND DATA

By

Todd James Barkman

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Botany and Plant Pathology

1993

ABSTRACT

CLADISTIC AND PHENETIC RELATIONSHIPS WITHIN *CYPRIPEDIUM* (ORCHIDACEAE) INFERRED FROM FLORAL FRAGRANCE-COMPOUND DATA

By

Todd James Barkman

The qualitative fragrance composition of eight species of *Cypripedium* was compared using phenetic and cladistic techniques to propose relationships within the genus. The clustering technique used in the phenetic analysis is an effective way to objectively assess the similarity in floral fragrance composition between several species, however, phylogenetic inference is limited by problems in accurately coding the chemical data. In contrast, the cladistic analysis may provide accurate estimates of phylogeny as the results are corroborated by several recent independent studies. Characters used in the cladistic analyses were delimited into states which represented enzymatic steps in proposed biogenetic pathways. Phylogenetic conclusions include: retention of *C. arietinum* within the genus, monophyly of the *C. calceolus* complex with a possible inclusion of *C. macranthum*, and a lack of resolution for the relative positions of *C. acaule*, *C. guttatum*, and *C. reginae*. Evolutionary polarity within the genus could not be determined using this data set.

ACKNOWLEDGMENTS

I am appreciative of financial assistance granted by the Mid-America Orchid Congress, Education and Research Committee and the William T. Gillis Memorial Fund.

I thank Dr. John H. Beaman who encouraged me as an undergraduate to pursue a graduate degree at Michigan State University. His advice and patience throughout my academic program have helped in my development as a scientist.

I sincerely thank Dr. Douglas Gage, who has guided me in the analytical chemistry techniques used during my undergraduate and graduate career. The time and help he has given me is invaluable and this project could not have been completed without his guidance.

I offer thanks to Dr. Ralph Taggart who served as a member of my committee and has offered valuable insights to me throughout my career at Michigan State University. I thank Mr. Fred Case and Mrs. Roberta Case for access to their cultivated temperate orchids, including species which could not have been sampled otherwise. Many thanks go to the members of the MSU-NIH Mass Spectrometry facility for technical and moral support, including Jennifer Johnson, Kate Noon, Melinda Berning, and Bev Chamberlain. I offer thanks to Mr. Tom Kalina for encouraging my pursuit of this project.

Finally, I thank Mr. Jack Brechtelsbauer for introducing me to the Orchidaceae and sharing my excitement in this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES; PART 1

LIST OF FIGURES; PART 2

PART 1: PHENETIC RELATIONSHIPS WITHIN *CYPRIPEDIUM* INFERRED FROM FLORAL FRAGRANCE-COMPOUND DATA

INTRODUCTION

Taxonomic history

Cypripedium is a circumboreal genus of slipper orchids consisting of approximately 40 taxa. It is one of five genera in the subfamily Cypripedioideae, all of which are easily recognized by a pouch or slipper-like labellum, which is unique in the Orchidaceae. *Cypripedium* includes several North American taxa that are not difficult to distinguish in most cases. These include *C. acaule*, the mocassin flower, *C. reginae*, the showy lady's slipper, *C. guttatum*, the striped lady's slipper, *C. calceolus*, the yellow lady's slipper, *C. candidum*, the small white lady's slipper, and *C. arietinum*, the rams head lady's slipper. *C. kentuckiense* is a large yellow lady's slipper and may be difficult to distinguish from *C. calceolus* in some cases. *Cypripedium macranthum* is native to Asia and is very distinct from the N. American taxa in terms of floral morphology and color.

Cypripedium has been the subject of numerous studies recently, including those concerned with pollination biology (Newhouse 1976), population genetic structure (Case 1993), and species introgression (Klier 1992). In addition, several species have been the focus of floral fragrance studies (Nilsson 1979; Bergstrom et al. 1992; Holman 1984). Interest in the volatiles of *Cypripedium* is not surprising due to the unusual pollination system of these species, in which floral fragrance appears to play an important role. The pouch-like labellum of the flower acts as a passive trap to insects (primarily bees) which may slip or climb into it (Nilsson 1979). Insects not large enough to escape through the primary orifice of the

pouch may pass through the rear of the labellum. The bee's exit is enabled by the presence of light windows (in some cases), false nectar guides, and trichomes which are used for traction (Newhouse 1976; Stoutamire 1967). The insect visitors do not collect nectar or any other known reward from these visits (although this has been a subject of debate; Newhouse 1976). In such non-rewarding pollination systems, it is surprising that insects would visit flowers. Indeed, Gill (1988) reports 3% pollination success over ten years in a population of *C. acaule*. Contrasting this low figure, Nilsson reports 20% success in *C. calceolus* subsp. *calceolus* (1979). In the same paper, Nilsson (1979) discussed the potential importance of fragrance compounds as attractants in *C. calceolus*. Among the components found, many were similar to those produced by pollinating bees of the genus *Andrena* as nest or territory markers. The similarity between the fragrance composition of *Cypripedium* to flowers of nearby populations of species with rewarding pollination systems such as *Convallaria* (Nilsson 1979) may also serve to attract pollinators. The role of volatile compounds to successful pollination in other orchid genera, particularly tropical euglossine-pollinated species and the importance of these compounds in maintaining a specific flower/pollinator interaction has been reported (Williams & Whitten 1992). Although high specificity of pollinators to flowers due to unique fragrance compositions has not been widely reported in temperate species, it is common in tropical species (Meeuse & Morris 1984). The putative importance of fragrance compounds for modifying insect behavior in *Cypripedium* provides a temperate model system for the systematic investigation of species-specific fragrance composition.

In addition to the widespread interest in the pollination system of *Cypripedium*, the genus has received much taxonomic attention in recent years; however, circumscription of the genus is presently unclear. Atwood (1984) recommended removal of *Cypripedium arietinum*, and placement into a monotypic genus

Criosanthes which had been proposed by Rafinesque in 1818. This suggestion was based upon the unusual spurred labellum, a staminode which resembles the fertile stamens, and separate lateral sepals. Recently, however, M. Case (1993) has shown using allozyme variation data that *Cypripedium arietinum* has a relatively high genetic identity (Nei's genetic identity = .285) with *C. calceolus* and *C. candidum*. Other taxa, including *C. reginae* and *C. acaule* were shown to have a lower identity of 0.072 to these taxa. Additionally, Albert (pers. comm.) has suggested that the taxon remain within *Cypripedium* on the basis of a cladistic analysis utilizing molecular, anatomical, and morphological characters. While *C. arietinum* appears to belong within the genus, its relationship among the other taxa is largely unresolved.

The most complex taxonomic problem within the genus involves the yellow lady's slipper, *C. calceolus*. At present, four varieties have been recognized for this taxon. *Cypripedium calceolus* subsp. *calceolus* is restricted in distribution to Europe. *Cypripedium calceolus* var. *planipetalum* is a dwarfed plant that has the northern-most distribution of the three North American taxa (Atwood 1984). These two infraspecific taxa will not be examined in the present study. The distinctions between *C. calceolus* var. *parviflorum* and *C. calceolus* var. *pubescens* are less obvious, particularly due to high degrees of morphological variation (F. Case 1987). Significant overlap in geographic and ecological distribution occurs for these two varieties (F. Case 1987). In addition, natural hybridization occurs, resulting in swarms of individuals not easily categorized as either taxon (personal observation). Recently, on the basis of allozyme variation, M. Case (1993) has recommended the retention of *C. calceolus* var. *parviflorum* and *C. calceolus* var. *pubescens* within *C. calceolus*, because genetic divergence was low among the varieties. Few workers have suggested a specific status for any of these varieties on the basis of morphological or allozyme allelic differences. Recently, however, Bergstrom et al.

(1992) have published fragrance chemical data which support the separation of *C. calceolus* subsp. *calceolus*, *C. calceolus* var. *parviflorum*, and *C.calceolus* var. *pubescens* into three species. Atwood (1984) also recognized two species including *C. pubescens* (formerly recognized as *C. calceolus* var. *pubescens*) and *C. parviflorum* (formerly *C. calceolus* var. *parviflorum*). A third taxon similar to *C. pubescens* in morphology, *C. kentuckiense*, is recognized as a distinct species on the basis of its geographic distribution and relatively larger size (Atwood 1984). Gradation in floral morphology and color between these varieties has complicated the taxonomic distinctions proposed.

Few relationships have been proposed among the morphologically distinct taxa *C. acaule*, *C. reginae*, *C. guttatum* and *C. macranthum*. In contrast, the *C. calceolus* complex (Atwood 1984) consists of several morphologically similar taxa, including *C. candidum*, *C. calceolus*, and *C. kentuckiense*. Case (1993) has reported relatively high genetic identities for these taxa. Although genetic identities were not calculated between *C. kentuckiense* and the others within the complex due to a small sample size, preliminary study indicates a relatively high identity also to *C. calceolus*.

Use of fragrance in taxonomy

Floral fragrance compounds have been used a few times previously to help resolve taxonomic problems in the Orchidaceae. Gregg (1986) and Borg-Karlson (1990) suggested relationships within complex species groups (*Cycnoches* and *Ophrys*, repectively) based upon evaluation of the patterns of fragrance composition. Qualitative and quantitative distinctions in fragrance composition between putative species corresponded to pollinator differences and floral morphological distinctions in most cases. The fragrance composition of these species appears to be very important in modifying insect behavior, as *Ophrys*

exhibits a pseudocopulatory pollination system in which mimic pheromones are produced (Borg-Karlson 1990), and *Cycnoches* is pollinated by Euglossine bees which collect the fragrance components produced (Gregg 1983). Due to the influence of quantitative and qualitative variation in fragrance chemistry on pollinator specificity, the differences were considered useful for taxonomic distinctions within *Cycnoches* and *Ophrys*.

Previous fragrance work in Cypripedium

Cypripedium has been the object of several floral fragrance investigations. Bergstrom et al. (1992) sampled floral fragrance compound variation among *C. calceolus* subsp. *calceolus*, *C. calceolus* var. *parviflorum*, and *C. calceolus* var. *pubescens*. *Cypripdium calceolus* subsp. *calceolus* has a substantially distinct qualitative fragrance profile from the other two varieties. *Cypripedium calceolus* var. *parviflorum* and *C. calceolus* var. *pubescens* differ primarily in relative quantities of many shared compounds. Holman (1973) identified volatile compounds produced by various organs in *C. reginae*, *C. acaule*, *C. calceolus*, and *C. candidum*. The results of his study illustrate that very different volatiles can be produced from various floral and vegetative organs, with the staminode producing the greatest quantity of volatiles in the flowers of these species. Significant overlap in fragrance components was described between these taxa, but the taxonomic importance of the patterns was not discussed by the author. The previous fragrance work in *Cypripedium* indicates that differences in profiles exist between species. These compositional differences may be useful in taxonomic studies, especially due to the importance of the compounds to the reproductive system of the flowers.

Knowledge of the extent of variation for a potential taxonomic character within any taxon is necessary for comparative study. Floral fragrance composition has been shown to vary extensively within a single taxon as Gregg (1986) and

Kaiser (1993) reported qualitative and quantitative fragrance-compound variation within and between populations of *Cycnoches* and *Gymnadenia*, respectively.

Variation in qualitative fragrance composition between individuals may be attributable to several factors. Kaiser (1993) demonstrated that floral development can affect quantitative fragrance composition. In *Angraecum sesquipedale*, the degree of maturity of a single flower can affect the fragrance composition on consecutively sampled nights. Indeed, studies of volatile indole production in several aroids indicate that the compound is only produced and volatilized for a single day, after which production is either stopped or it is converted to non-volatile products (Chen & Meeuse 1971). Although it has not been studied, post-pollination processes may also cause variation in fragrance production. This would not be surprising as several physiological and chemical processes such as color changes and senescence of the flower occur after pollination. I have noted (unpublished results) quantitative differences in the relative abundances of specific compounds over the life of a flower in *Paphiopedilum delenatii*. Finally, the timing of fragrance release can be very specific for many plant species and appears to be under biorhythmic control (Kaiser 1993). Due to the potential for infraspecific and individual variation in fragrance composition, particularly quantitative, sampling for taxonomic studies must attempt to control for this the data are utilized for taxonomic study.

The use of phenetics and fragrance compounds in taxonomy

Explicit phenetic techniques have rarely been utilized in conjunction with floral fragrance compounds in efforts to determine overall similarity/dissimilarity of taxa. This is surprising due to the utility of these methods in taxonomic studies utilizing morphological as well as chemical information in other groups (Sokal & Sneath 1973; Rodman 1991). Gregg (1986) used ordination methods to

distinguish fragrance chemotypes within a population of *Cycnoches densiflorum*. In her study principal components analysis (PCA) was used to distinguish between high and low α-pinene clones of a population. Holman and Hiemermann (1976) utilized a numerical measure of similarity to compare several species of *Epidendrum*.

Despite these early phenetic studies, clustering methods have not been attempted nor systematically applied in conjunction with fragrance compound data. The use of fragrance compounds in cluster analyses aimed at proposing taxonomic relationships may be problematic for two reasons. Many volatile fragrance compounds are common natural products found in very unrelated organisms such as gymnosperms, angiosperms and fungi, and may not provide phylogenetically useful information (Whitten & Williams 1992). In addition, inferring relationships from these analyses may be misleading because phenograms may reflect pollinator relationships rather than historical relationships (Whitten & Williams 1992). Despite these problems, cluster analysis may represent an objective method for expressing and evaluating the overall qualitative chemical similarity between a group of taxa (Sneath & Sokal 1973). The use of these explicit methods is desirable over more subjective methods involving simple descriptions of compound distributions among taxa.

In the present study nine species of *Cypripedium* were investigated for floral fragrance compound composition. The qualitative compound distributions are analyzed using cluster analysis to examine the taxonomic significance of the chemical data. Additionally, fifteen individuals from two populations of *C. calceolus* var. *parviflorum* were sampled for qualitative fragrance-compound variation. The potential taxonomic importance of relative differences is evaluated. As fragrance compound data have not been previously utilized in conjunction with cluster analyses, an examination of the utility and/or shortcomings of this

technique follows by comparison to results of recent studies utilizing independent data sets.

Materials and Methods

Plant material

All plants were sampled outdoors. *Cypripedium macranthum*, *C. guttatum*, *C. candidum*, *C. calceolus* var. *pubescens*, and *C. kentuckiense* were sampled in the garden of Frederick and Roberta Case (Saginaw Co., MI). The plants were grown under conditions similar to their natural habitats. *Cypripedium calceolus* var. *parviflorum*, *C. arietinum*, *C. acaule*, and *C. reginae* were sampled in natural populations. *Cypripedium. calceolus* var. *parviflorum* clones were sampled from two populations, Rose Lake and Williamston, both in Ingham County, MI. Only one individual was sampled for each taxon except for *C. calceolus* var. *parviflorum*, in which 15 individuals were sampled. Voucher specimens are deposited in the Beal-Darlington Herbarium, Michigan State University (MSC).

Fragrance compound isolation

For field sampling studies, one or two flowers on an intact inflorescence were enclosed in a Plexiglas® box for approximately one hour prior to sampling. The fragrance-laden air was drawn through a sampling cartridge using a portable battery-powered sampling pump. Sorbent tubes containing XAD-2 polymers (SKC) were used to trap volatile components. External atmospheric samples were drawn simultaneously during sampling periods as controls. Sampling was initiated at 9:00 and ended at 18:00 each period with a flow rate of 60 ml/min. Trapped fragrance compounds were removed from the sorbent tubes using 1 ml hexane. The hexane mixture was subsequently concentrated using a slow stream of compressed N_2. The evaporation step was necessary to enable detection of the

small quantities of compounds produced by the taxa studied. Evaporation does not appear to result in the loss of compounds detectable by our instrument; sensitivity was only improved by successive concentration steps of dilute samples (Personal observation).

Fragrance compound identification

Samples were analyzed by gas chromatography-mass spectrometry (GC-MS). Chromatographic separation of fragrance compounds in the sample eluates were performed on an HP 5098 using a DB-Wax capillary column (0.32mm i.d. × 30 m, film thickness 0.25 mm) using a temperature program of 2 minutes at 40° C and a ramp to 240° C at 7° C/minute. The final temperature was held for 2 minutes.

The capillary column was directly interfaced into the ion source of an Hewlett-Packard mass selective detector (MSD). Spectra were acquired from 40 to 300 mass units. The spectra were compared to published spectra available in computer databases and original literature reports. Compound identification was based upon the combination of comparisons of relative retention times, Kovats indices and spectral profiles.

Character coding

Coding of the fragrance compound data is necessary for the calculation of similarity coefficients to be used in the cluster analysis. Erroneous representation of the data in character coding could significantly affect numerical analyses. The use of secondary chemical data in conjunction with clustering techniques is uncommon in botanical classification, so an explanation follows as to how the data have been represented. No attempt was made to represent quantitative variation between taxa in the coding procedure. The quantitative variation which can result from biological and temporal factors outlined above are difficult to control. Although quantitative variation may be taxonomically significant in these species,

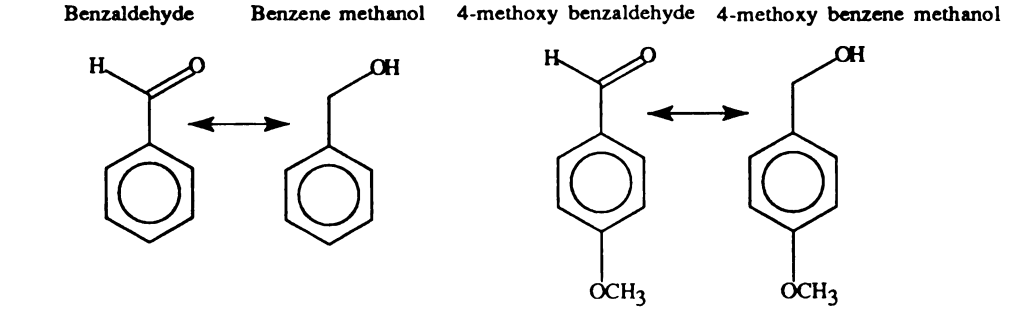
the possibility of variation from sources other than genetic differences between the individuals sampled justified the exclusion of quantitative data from the analyses.

Two methods for coding the chemical data were evaluated to examine the effect this step may have upon the cluster analysis. Figure 1 contrasts the two coding procedures.

Compounds are coded qualitatively as present or absent using the first coding method. One shortcoming of this method is that each compound is treated as an independent character. This may not be a very accurate coding technique because structurally related compounds such as configurational isomers are not biogenetically independent. In addition, compounds with the same basic carbon skeleton but differing only in the nature of the functional group substitution (e.g. alcohol/aldehyde of the same skeleton) are scored the same as would be two compounds from biogenetically distinct classes (e.g. indole/benzaldehyde). Despite the philosophical inconsistencies of this coding procedure, the data were scored this way to examine the effects upon the clustering analysis. No data matrix is presented for this coding procedure because each compound identified in each species (Table 1) is coded as present or absent.

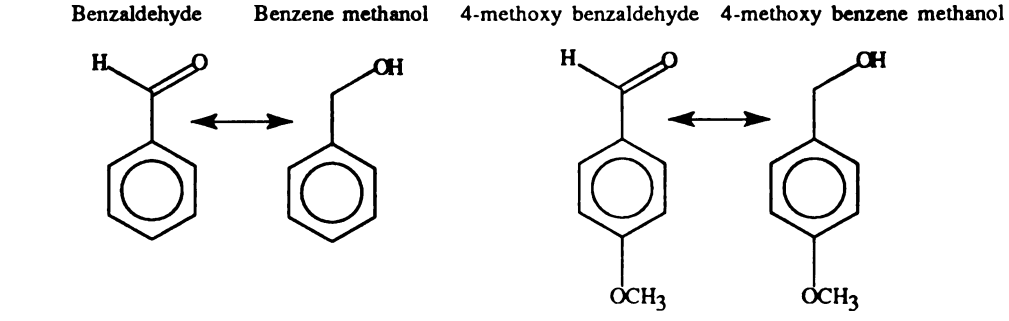
The second coding method was based on the assumption that the biogenesis of some volatiles identified are linked. This coding method is similar to that employed by Seaman and Funk (1983) for sesquiterpene lactone data. In the second method, isomers, both conformational and optical, were treated as identical compounds and were coded as a single character. Criteria used to delimit characters were as follows: fundamental carbon skeletons were coded only once, present or absent, regardless of the state of associated functional groups. Thus, the interconversions between aldehydes and associated alcohols of a single carbon skeleton were not treated as independent characters. In addition, each

Method 1: All compounds coded as separate characters.



Taxon A	1	1	1	1
Taxon B	1	1	0	0

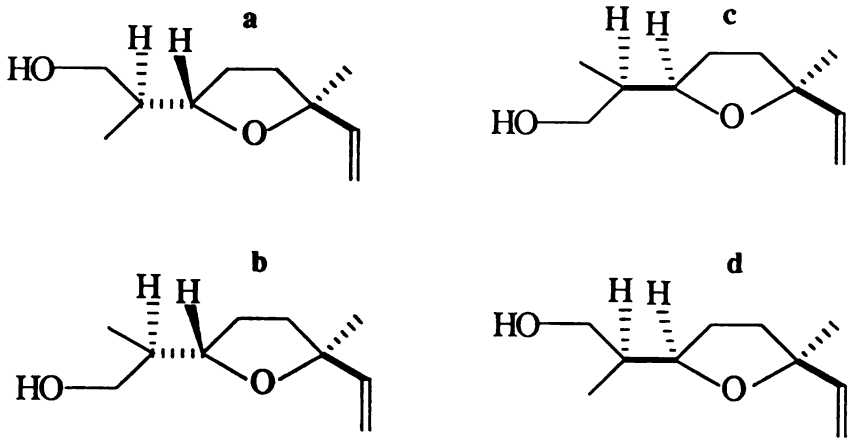
Method 2: Only basic skeletal types and substitutions coded.



Taxon A	1	1
Taxon B	1	0

Figure 1. Character coding for phenetic analyses.

Four conformational isomers of lilac alcohol



	a	b	c	d	
Taxon A	1	1	1	1	**Method 1**: All compounds coded as separate characters.

vs.

	a,b,c,d	
Taxon A	1	**Method 2**: Only basic skeleton coded.

Figure 1, cont'd.

novel substitution of a fundamental carbon skeleton was coded only once. If a single taxon produced an array of compounds, all of which were derivatives of a single carbon skeleton that was not found in any other taxon, then these were all treated as a single character. This exception is justified because the data are uninformative for comparative purposes. Figure 2 shows the character coding used in the data matrix based upon the biogenetic assumptions.

The assessment of overall qualitative chemical similarity, which served as the basis for taxonomic judgement between the taxa in this study, was determined by cluster analysis. The choice of association measure and clustering algorithm, including justification for inclusion of each, follows. The use of computer software facilitated utilization of these techniques. NTSYS version 1.7 (Rohlf 1987) was used for the SAHN (sequential agglomerative hierarchical non-overlapping) clustering analyses.

Association measure

The determination of overall chemical similarity between taxa was calculated by using Jaccard's coefficient of similarity. This coefficient was chosen over others because it is appropriate for use in chemical studies (Sneath & Sokal 1973).

Jaccard $S_J = \Sigma a / a + b + c$

Where a,b,c, are represented by:

	Taxon B	
Taxon A	1	0
1	(1,1) = a	(1,0) = b
0	(0,1) = c	(0,0) = d

1 = presence of compound(s)
0 = absence of compound(s)

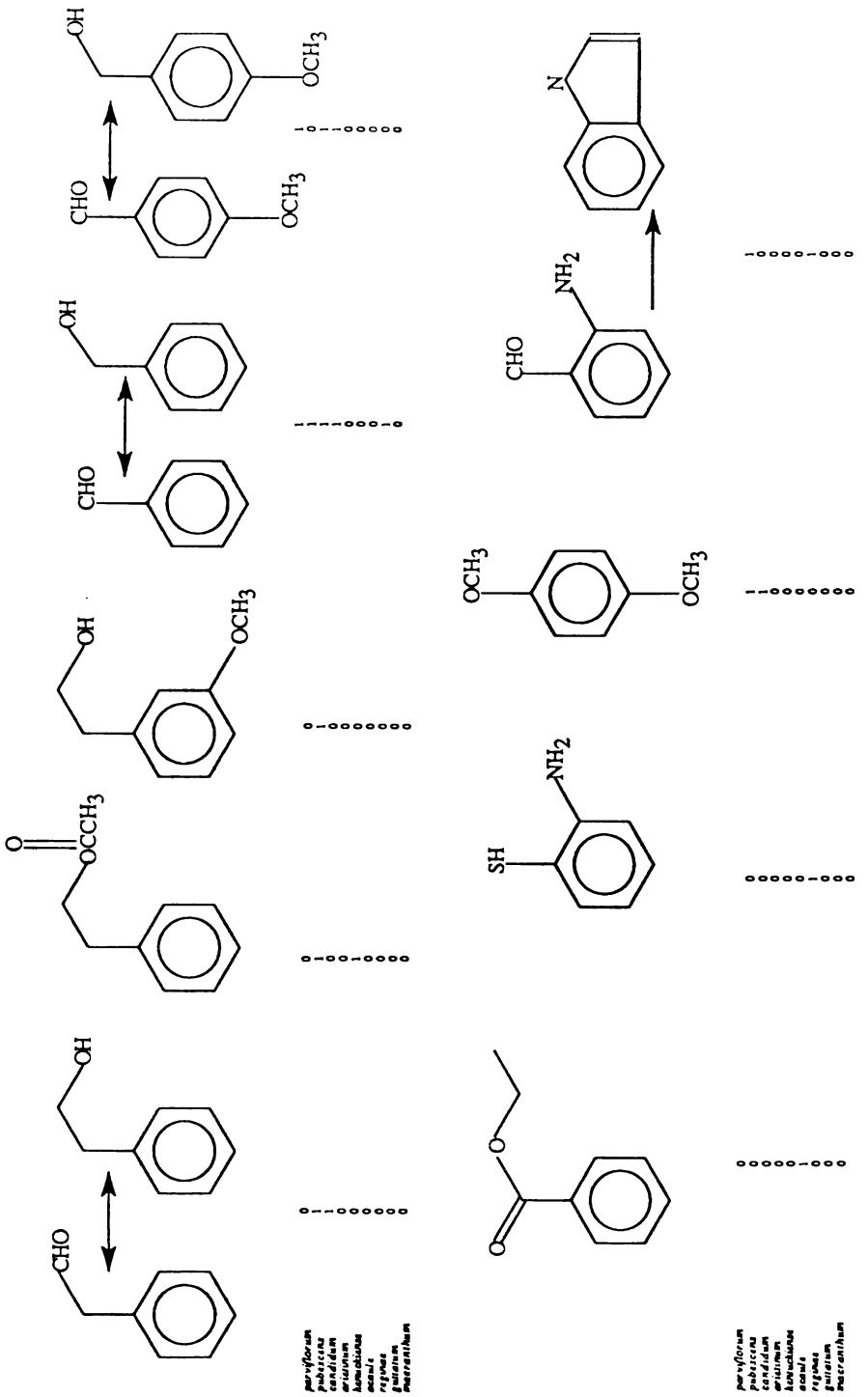


Figure 2. Character matrix with compounds coded according to biogenetic relatedness. 1 = Present; 0 = Absent

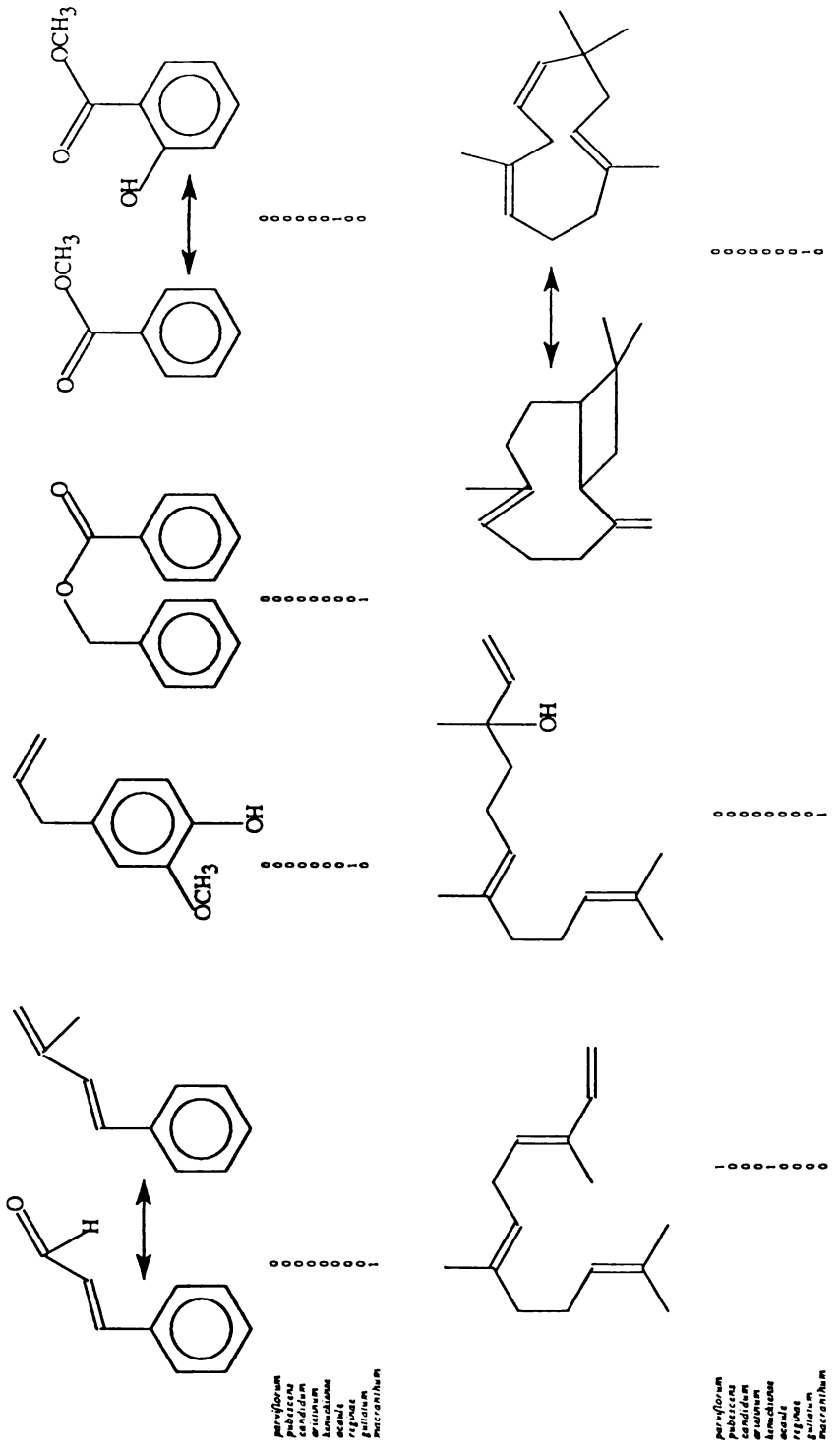


Figure 2, cont'd. Character matrix with compounds coded according to biogenetic relatedness. 1 = Present; 0 = Absent

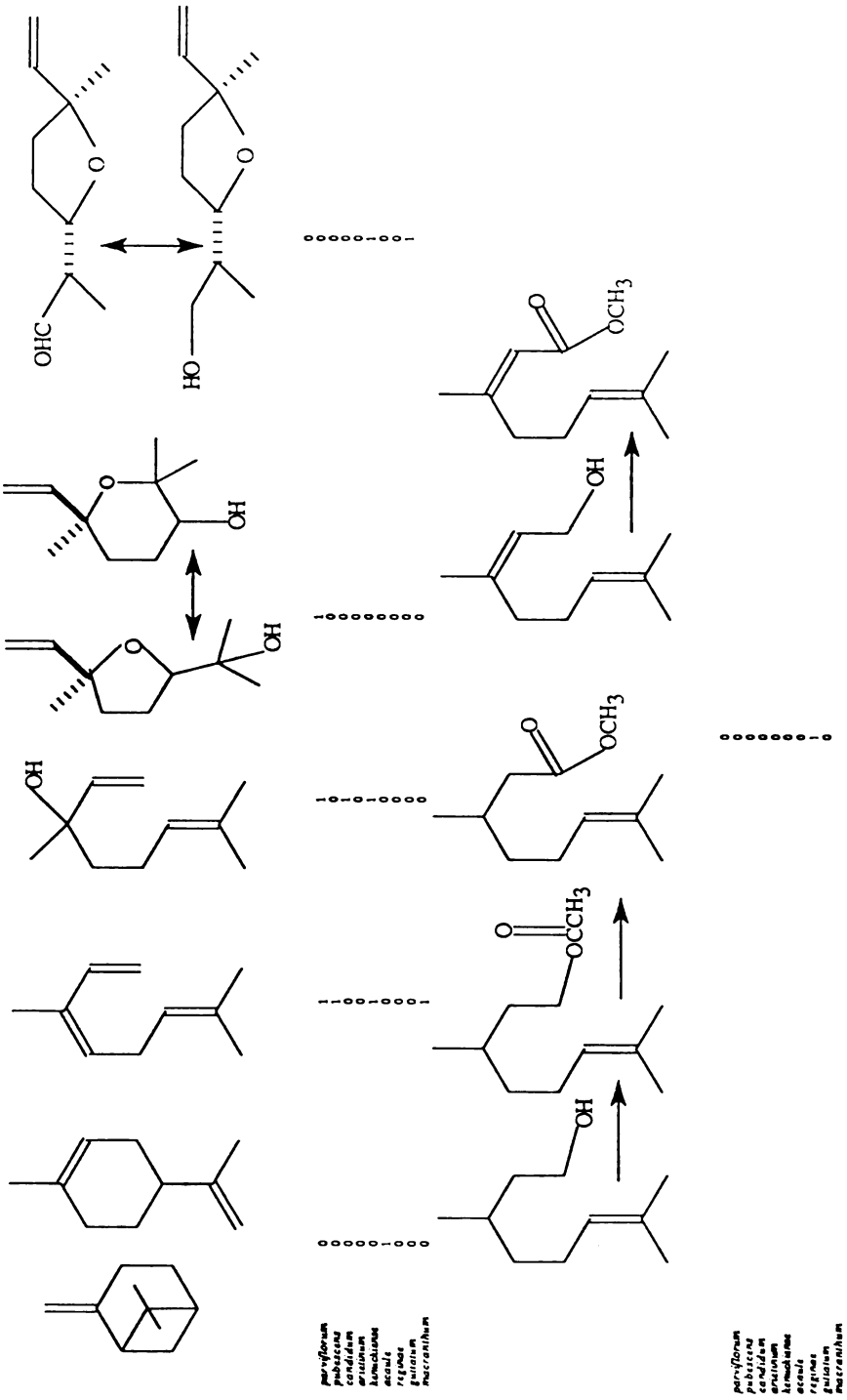


Figure 2, cont'd. Character matrix with compounds coded according to biogenetic relatedness. 1 = Present; 0 = Absent

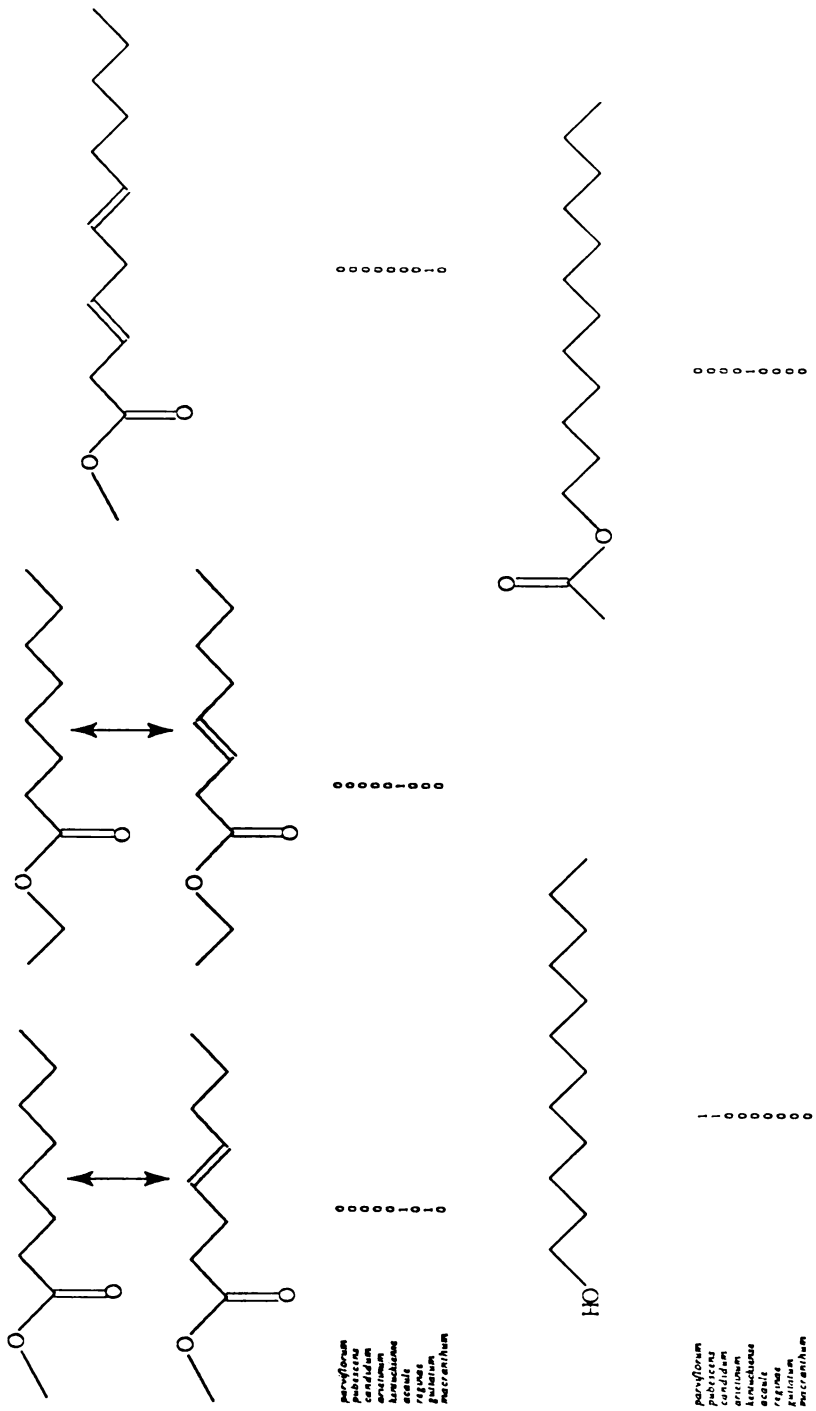


Figure 2, cont'd. Character matrix with compounds coded according to biogenetic relatedness. 1 = Present; 0 = Absent

This coefficient is particularly useful for chemical data because the calculation of similarity is based only upon the shared presence of compounds between two taxa. It does not include shared absences of compounds in the calculation as in the simple matching coefficient of Sokal & Michener (1958). In the case of volatile chemical data or other secondary compounds, the absence of a compound in two taxa should not represent a similarity between them (Stuessey 1990). This coefficient is highly desirable because of the variable nature of volatile chemical data. Further, the absence of a compound in a taxon may not represent the lack of biogenetic machinery to produce it, simply a lack of detectable amounts. The calculated coefficients between pairs of taxa are measures of the chemical similarity between them and are utilized by clustering algorithms.

Clustering algorithms

UPGMA (Unweighted pair group method arithmetic average) was chosen as the clustering algorithm due to its widespread use in phenetic analyses. Admission of a taxon into a cluster is based upon a comparison of its similarity to the average similarity measure of the cluster. Single and complete linkage algorithms are extreme opposites with respect to requirements for admission into clusters and were investigated and compared with the results obtained by UPGMA. Cophenetic values were computed to determine the effectiveness of the clustering as a representation of the original data. This measure allows an objective choice among the various clustering method results (Sneath & Sokal 1973; Duncan & Baum 1981). Taxonomic judgements were based upon the phenogram derived from the algorithm which displayed the highest cophenetic value.

RESULTS

The list of compounds identified in all nine taxa is provided in Table 1. Although quantitative information is not utilized in the cluster analysis, relative ratios of compounds found in taxa are reported. No two taxa are very similar in overall fragrance composition. *Cypripedium reginae* exhibits the least diverse fragrance composition with only two benzoic acid-derived compounds produced, representing the expression of a single biogenetic pathway. *C. arietinum* produced only four compounds, again from a single pathway. All other taxa produce various classes of compounds including mixtures of benzenoids, monoterpenoids, and fatty acid derivatives. All compounds identified are grouped according to biogenetic relatedness. Proposed biogenetic groupings follow: Benzenoid types include phenyl propanoids, anthranilic acid derivatives, benzoic acid derivatives, 1,4 -dimethoxy benzene, and phenyl acetaldehyde derivatives. Fatty acid derivatives include acetates, ethyl and methyl esters, and aliphatic alcohols with straight chain hydrocarbon skeletons. Terpenoids found include monoterpene hydrocarbons, monoterpene alcohols, pyranoid monoterpene alcohols, and sesquiterpene hydrocarbons and alcohols. *Cypripedium acaule* produced the largest number of compounds. The fragrance was dominated by fatty acid esters with decanoic acid methyl ester the most abundant (34.5%). Decanoic acid ethyl ester was the second most abundant (18%). *Cypripedium kentuckiense* produced a fragrance composition in which fatty acid acetates comprise 70% of the mixture. The fragrance of *C. guttatum* was composed of several fatty acid esters, of which octanoic acid methyl ester was dominant (45.3%). Benzene methanol was the second most abundant compound (17.2%). Several mono- and sesquiterpenoids were produced but were minor components. *Cypripedium macranthum* had a fragrance composition comprised almost entirely of the pyranoid monoterpenes, lilac alcohols and lilac aldehydes (83%). The two varieties of *C. calceolus* produced

Table 1. Taxa sampled within *Cypripedium* with compounds listed as percentages of total composition

Compounds listed according to relative retention times

	parviflorum	*pubescens*	*candidum*	*arietinum*	*kentuckiense*	*acaule*	*reginae*	*guttatum*	*macranthum*
beta-pinene 7.6	<.5					<.5			
limonene 8.2						0.5			
cis-ocimene 8.9									1.1
trans-ocimene 9.2		0.6			3.8				
octanoic acid methyl ester 11.9						3.4		45.3	
trans-linalool oxyd 12.5	0.6								
4-octenoic acid methyl ester 12.6						0.6			
octanoic acid ethyl ester 12.7						1.7			
3-octenoic acid ethyl ester 13.4						0.5			
nonanoic acid methyl ester 13.7						0.5		0.5	
benzaldehyde 14.2	<.5	1.4	28.5	13.1				0.5	
cis-linalool oxyd 13.1	2.4								
linalool 14.5	37.7		17.5		0.5				
lilac aldehyde a 14.6						4.2			45.9
lilac aldehyde b 14.8						1			20.1
lilac aldehyde c 14.9						0.4			7.6
citronellic acid methyl ester 14.9								3.1	
lilac aldehyde d 15.3						1.7			6.6
decanoic acid methyl ester 15.4						34.5		3.7	
caryophyllene 15.5								3.9	
4-decenoic acid methyl ester 15.9						3.6		3.8	
benzoic acid methyl ester 15.9							62		
phenylacetaldehyde 16.1		26.8	25.4						
decanoic acid ethyl ester 16.1						18			
citronellyl acetate 16.4								1.4	
3-decenoic acid ethyl ester 16.5						0.9			
benzoic acid ethyl ester 16.6						0.5			
decyl acetate 16.7					3.5				
alpha-humulene 16.7								1.1	
unknown oxygen containing terpene 16.8						0.7		1.4	
unknown oxygen containg terpene 17.0						0.2			
geranic acid methyl ester 17.2								2.8	
lilac alcohol b 17.5						<.5			0.35
isomer of 16.8 Ace- 17.5						0.2			
1,4-dimethoxy benzene 17.6	15.2	<.5							
lilac alcohol d 17.8						<.5			1
citronellol 18.0								5.7	
methyl salicylate 18.3							38		
lilac alcohol a 18.4						0.5			2.4
linalool oxide cis or trans? 18.5	2.6								
nerol 18.5								2.4	
dodecanoic acid methyl ester 18.6						0.7			
lilac alcohol c 19.0									0.5
E-cinnamaldehyde 19.1									0.9
dodecanoic acid ethyl ester 19.2						0.4			
alpha-farnesene 19.3	15.9				13				
phenyl ethyl acetate 19.55		<.5			3.1				
Z-cinnamaldehyde 19.6									0.5
dodecyl acetate 19.7					25.8				
benzene methanol 19.7		2	6.6	74				17.2	
3,6-dodecadienoic acid methyl ester 19.9								3.4	
phenyl ethyl alcohol 20.1		63.6	2.5						
4-phenyl-3-buten-2-one 20.8									0.4
dodecanol 21.67	4.7	2.4							
4-methoxy, benzaldehyde 21.7	6.1		13	7					
nerolidol 21.8	1.8								1.2
tetradecyl acetate 22.5					3.1				
benzyl benzoic acid 23.2									0.5
eugenol 23.5								4.8	
1-hexadecyl acetate 23.5					40				
amino benzaldehyde 23.6	<.5					17.4			
4-methoxy, benzene methanol 24.8	9.7		6.5	5.9					
2-amino benzoic acid ethyl ester 24.9						0.5			
octadecyl acetate 25.0					0.5				
4-methoxy phenyl ethyl alcohol 25.4		3.2							
2-amino benzene thiol 25.7						6.8			
indole 26.9	3.3					1.5			

four compounds in common: benzaldehyde, benzene methanol, 1,4-dimethoxy benzene, and dodecanol. *Cypripedium calceolus* var. *parviflorum* had a fragrance composition dominated by linalool (37.7%), alpha-farnesene (15.9%), and 1,4-dimethoxy benzene (15.2%). In contrast, *C. calceolus* var. *pubescens* produced a composition in which phenyl acetaldehyde (26.8%) and phenyl ethanol (63.6%) dominated. The putatively closely related taxon, *C. candidum*, shared 4 compounds with *C. calceolus var. parviflorum*, and 4 compounds with *C. calceolus var. pubescens*. It is interesting to note that the composition of *C. candidum* is qualitatively intermediate between these two taxa with benzaldehyde comprising 28.5% and phenylacetaldehyde comprising 25.4% of the mixture. In total, 66 compounds were isolated and identified in the nine taxa sampled.

Table 2 lists compounds found in 15 individuals of *C. calceolus* var. *parviflorum* from two populations, Rose Lake and Williamston. Compounds are listed as present (1) or absent (blank) for each clone sampled. Presence of only trace amounts were recorded as (1). In 8 of 15 individuals, dodecanol was produced. Nine of 15 individuals produce cis-linalool oxyd and indole; 11 of 15 individuals produced 4-methoxy benzene methanol; and 14 of 15 individuals produced 1,4-dimethoxy benzene, alpha-farnesene, 4-methoxy benzaldehyde. All individuals sampled produced linalool and benzaldehyde. Figure 3 shows three chromatograms among the clones R1, R0, and R4 sampled respectively. The peaks in the chromatogram representing the presences of linalool, 1,4-dimethoxy benzene, and 4-methoxy benzaldehyde are designated by arrows. Quantitative variation of these primary compounds is extensive between all individuals and no attempt was made to represent this information. Qualitative variation between individuals is evident upon inspection of table however these variant compounds are only minor components of the fragrance. Individuals in the table are listed in chronological order with R0 as the first individual sampled on 5/20/92. R1–3

Table 2. Individual clones sampled from two populations (R=Rose Lake W=Williamston) of *Cypripedium calceolus var. parviflorum*

Compounds listed in order of relative retention times

	R0	R1	R2	R3	R4	R5	R6	R7	R8	R9	W1	W2	W3	W4	W5
beta-pinene	1									1					1
limonene					1										
cis-ocimene										1					
trans-ocimene										1					
1-methoxy-4-methylbenzene												1			
trans-linalool oxyd	1	1					1								
cis-linalool oxyd	1	1	1	1	1	1					1		1		1
benzaldehyde	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
linalool	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1,4-dimethoxy benzene	1	1	1	1	1	1	1	1	1	1		1	1	1	1
alpha-farnesene	1	1	1	1	1	1	1		1	1	1	1	1	1	1
linalool oxide (cis or trans?)	1	1	1				1				1		1		1
Z-cinnamaldehyde													1	1	
benzene methanol		1	1		1		1								
E-cinnamaldehyde													1		
dodecanol	1	1	1	1	1	1	1						1		
4-phenyl-3-buten-2-one													1		
4-methoxy benzaldehyde	1	1	1	1	1	1	1	1	1	1	1	1	1		1
nerolidol	1						1								
benzyl benzoate			1	1											
amino benzaldehyde					1										
3,4,5-tri-methoxy benzene											1	1	1		
4-methoxy-benzene methanol	1	1	1	1	1	1	1			1		1	1		1
indole	1	1	1	1	1	1	1			1					1

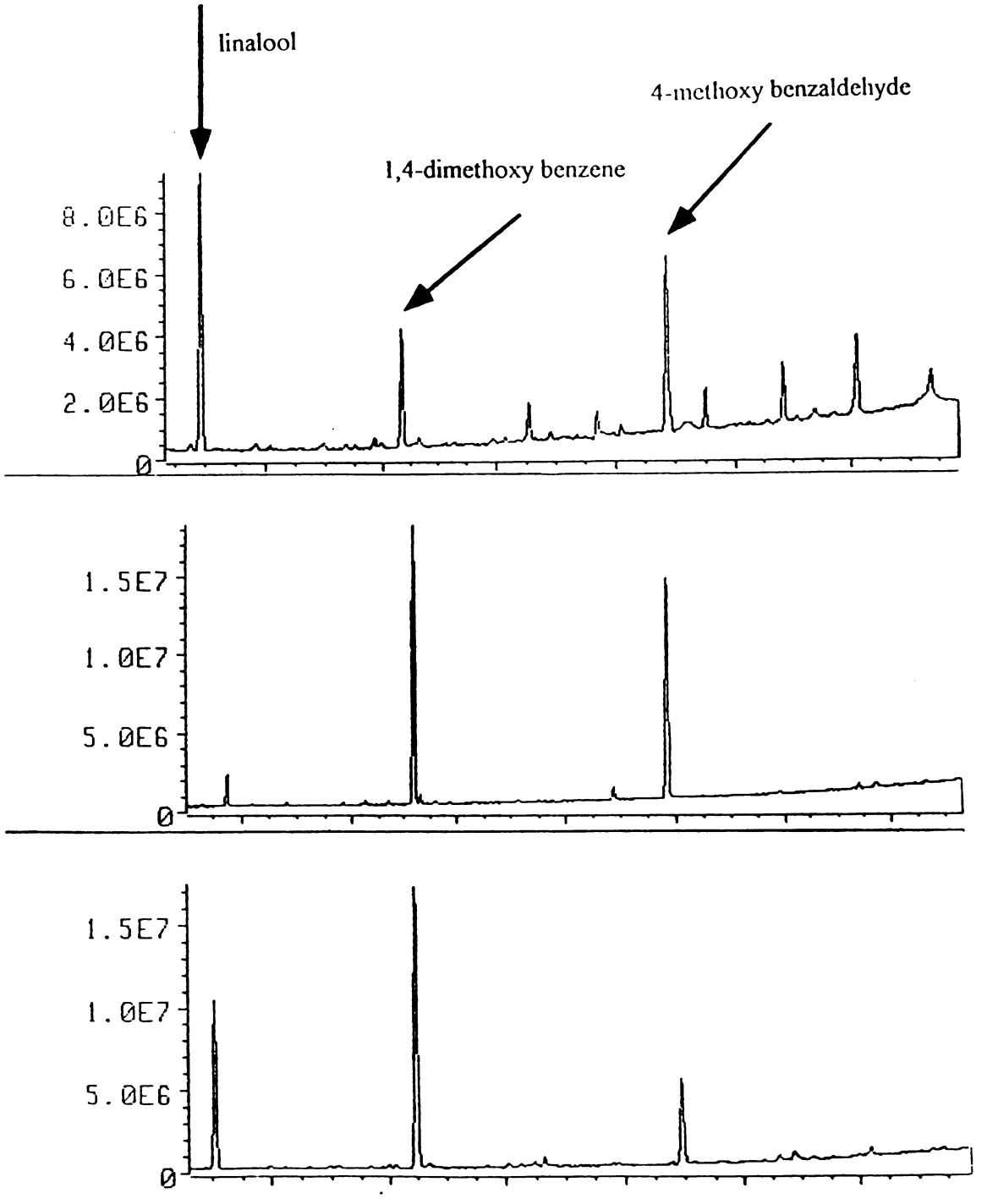


Figure 3. Relative abundances of linalool, 1,4-dimethoxy benzene, and 4-methoxy benzaldehyde in three clones (R1, R0, and R4 respectively) of *Cypripedium calceolus* var. *parviflorum*.

were sampled on 5/22/92, R4–R6 were sampled on 5/23/92, and R7–9 were sampled 5/24/92. W1–2 were sampled on 5/25/92, W3–4 were sampled 5/26/92 and W5 was sampled 5/27/92.

Table 3. Jaccard similarity coefficients calculated using coding method 1.

	par	pub	can	ari	ken	aca	reg	gut	mac
par	1.000								
pub	0.158	1.000							
can	0.250	0.333	1.000						
ari	0.231	0.091	0.429	1.000					
ken	0.100	0.125	0.067	0.000	1.000				
aca	0.051	0.000	0.000	0.000	0.000	1.000			
reg	0.000	0.000	0.000	0.000	0.000	0.000	1.000		
gut	0.037	0.043	0.048	0.059	0.000	0.132	0.000	1.000	
mac	0.038	0.000	0.000	0.000	0.000	0.200	0.000	0.000	1.000

Figure 4 shows the UPGMA phenogram obtained from similarity coefficients of Table 3. The phenogram is comprised of two primary clusters. The first includes taxa traditionally considered part of the *C. calceolus* complex. In this cluster, *C. arietinum* and *C. candidum* are the most similar based upon fragrance chemistry. *Cypripedium calceolus* var. *parviflorum* is the next most similar to these taxa. *Cypripedium calceolus* var. *pubescens* and *C. kentuckiense* are contained within this cluster but are not very similar to each other or to any other taxon. All of these taxa are similar in basic floral structure, possessing linear-lanceolate lateral petals and an inflated labellum with an opening at the top of the pouch.

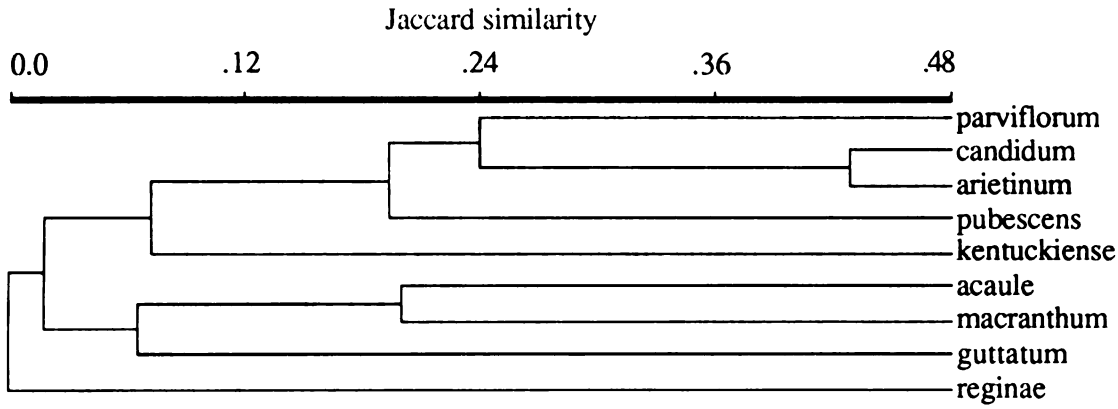


Figure 4. UPGMA-derived phenogram based upon Jaccard similarity coefficient using coding method 1.

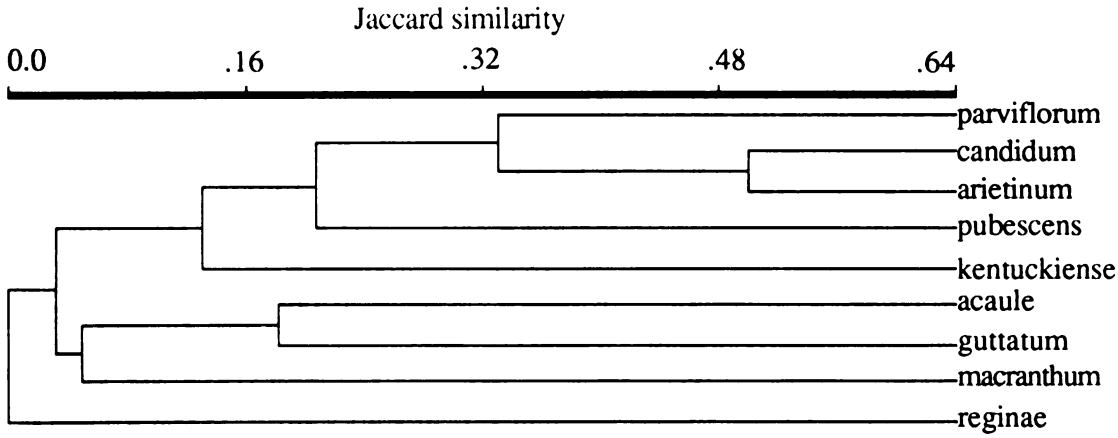


Figure 5. UPGMA-derived phenogram based upon Jaccard similarity coefficient using coding method 2.

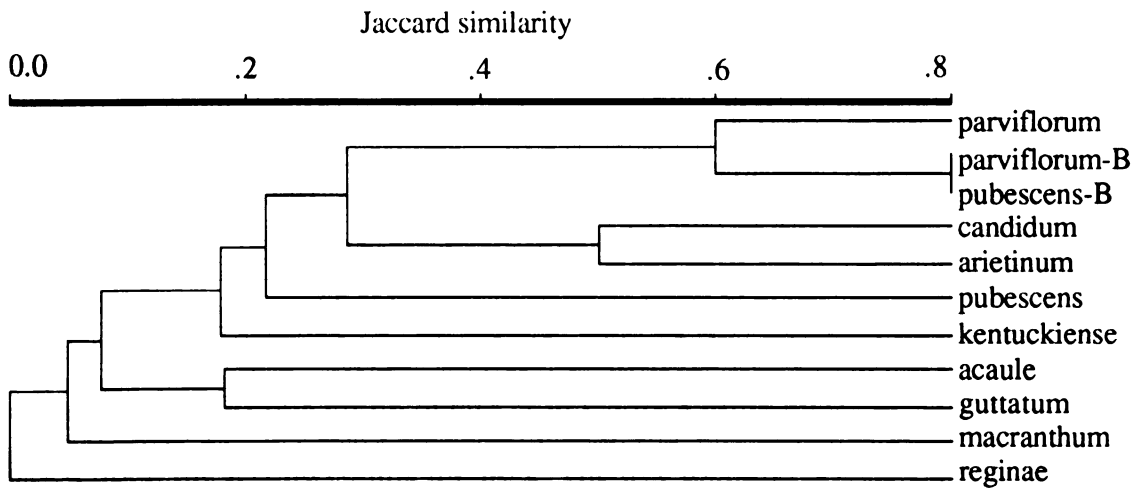


Figure 6. UPGMA-derived phenogram based upon Jaccard similarity coefficient using coding method 2. Analysis includes *C. calceolus* var. *parviflorum* and *C. calceolus* var. *pubescens* sampled by Bergstrom et al. (1992).

Cypripedium arietinum is an exception, however, with a floral morphology distinct within the genus. The other cluster is comprised of morphologically dissimilar taxa with *C. acaule* and *C. macranthum* most similar in terms of fragrance chemistry. *Cypripedium guttatum* is loosely associated with the cluster and *C. reginae* is separate from all other taxa. The cophenetic correlation coefficient (r = 0.92311) for the UPGMA-derived phenogram was the highest obtained for all of the clustering algorithms used.

Table 4. Jaccard similarity coefficients calculated using coding method 2.

	par	pub	can	ari	ken	aca	reg	gut	mac
par	1.000								
pub	0.273	1.000							
can	0.375	0.222	1.000						
ari	0.286	0.125	0.500	1.000					
ken	0.200	0.200	0.125	0.000	1.000				
aca	0.077	0.000	0.000	0.000	0.000	1.000			
reg	0.000	0.000	0.000	0.000	0.000	0.000	1.000		
gut	0.083	0.083	0.111	0.143	0.000	0.182	0.000	1.000	
mac	0.000	0.091	0.000	0.000	0.111	0.091	0.000	0.000	1.000

Figure 5 represents the results of a cluster analysis using the UPGMA algorithm for the same group of taxa as in Figure 4 using similarity coefficients based upon the second coding method. The tree topology is identical with the exception of *C. acaule* now represented as more similar to *C. guttatum* than to *C. macranthum*. The similarity coefficients are higher between other taxa due to the

differences in coding of the data matrix. The cophenetic correlation coefficient is r = 0.91372 and was higher than that obtained by either single or complete linkage.

Figure 6 shows the results of an analysis which included the taxa sampled by Bergstrom et al. (1992). *Cypripedium calceolus* var. *parviflorum* is represented by parviflorum-B and *C. calceolus* var. *pubescens* is represented by pubescens-B. The most interesting result obtained in this cluster analysis is the separation of the two individuals of *C. calceolus* var. *pubescens* in the phenogram. The individuals sampled by Bergstrom et al. (1992) cluster with *C. calceolus* var. *parviflorum* sampled here. The similarity coefficients were calculated from compounds coded according to biogenetic relatedness. The cophenetic correlation coefficient is r = 0.94024

Figure 7 shows the results of the cluster analysis including all previously analyzed taxa with the addition of all individuals of *C. calceolus* var. *parviflorum* sampled from the Rose Lake (R) and Williamston (W) populations. There does not appear to be a separation of the two populations from each other. It is clear that none of the individual clones of *C. calceolus* var. *parviflorum* sampled are more distinct from each other so as to cluster with any other taxon in the study. While the variation between individual clones appears great from inspection of Table 2, when compared to the variation among other taxa within the genus, the differences are not phenetically significant. The cophenetic correlation coefficient is r = 0.97246.

DISCUSSION

Comparison of coding methods

A comparison of the phenograms based upon the two coding methods indicates the effect that the coding procedure can have upon the clustering results

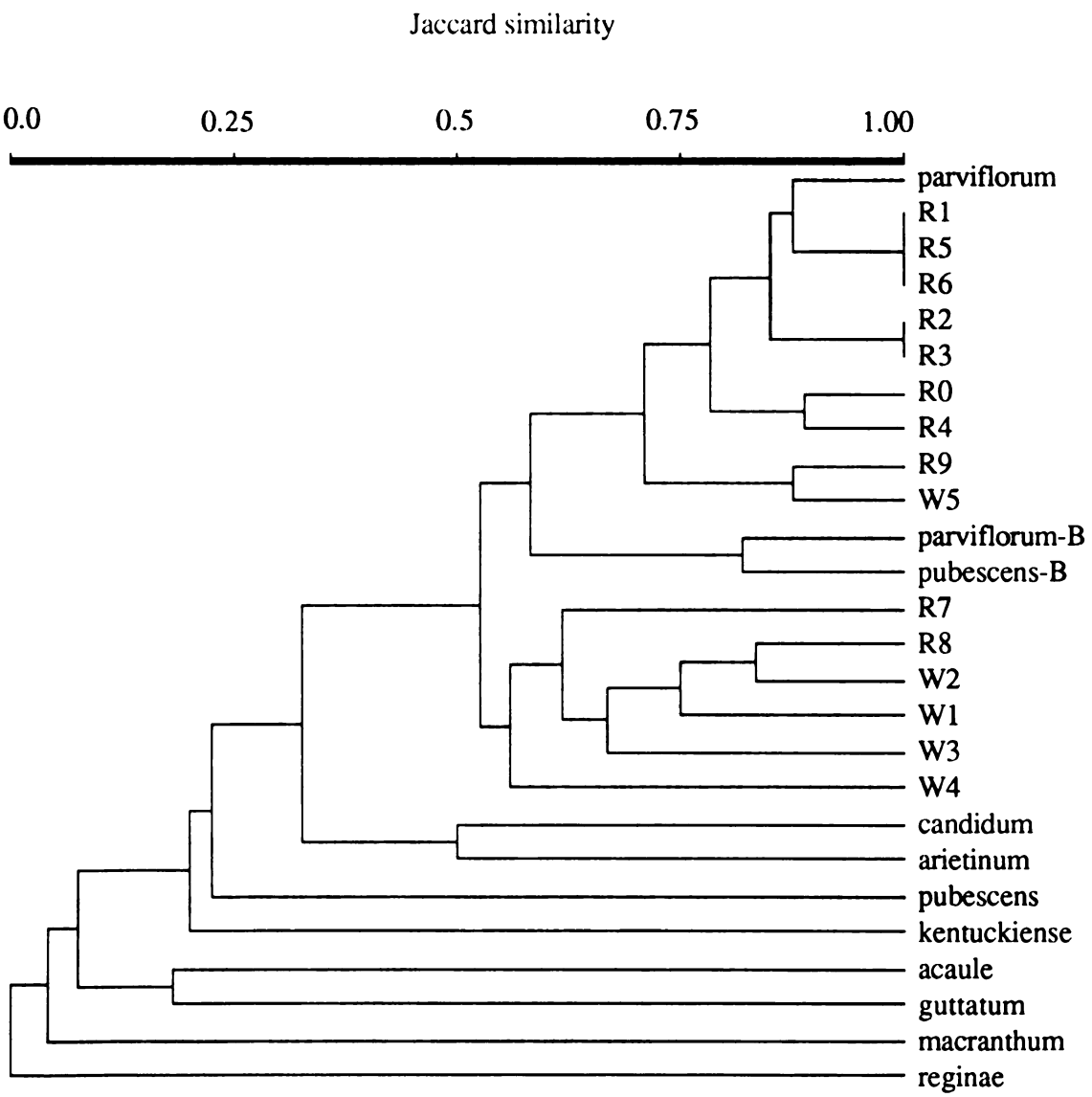


Figure 7. UPGMA-derived phenogram including individuals of *C. calceolus* var. *parviflorum* sampled from two populations (R=Rose Lake, and W=Williamston) using coding method 2.

of an analysis. The two phenograms (Figures 4 and 5) exhibit similar topologies with a single difference being the relative placement of *C. guttatum* and *C. macranthum*. The first coding method, in which isomers were treated as independent characters, resulted in an artificially high measure of similarity for *C. acaule* and *C. macranthum*. These taxa share only a single pyranoid monoterpene pathway, which produces lilac alcohol and lilac aldehyde. However, the expression of this pathway results in the production of four isomers for the alcohol and four isomers for the aldehyde. Coding these eight compounds as independent characters between *C. acaule* and *C. macranthum* resulted in erroneously high similarity values for these taxa and obscured the similarity between *C. acaule* and *C. guttatum*. The results of the second coding method, in which isomers were treated as single dependent characters, indicate the similarity that is present between *C. acaule* and *C. guttatum*. These two taxa share several independent pathways including the fatty acid-ester pathway, and an unidentified oxygenated monoterpene pathway. The co-occurrence of two independent pathways between *C. guttatum* and *C. acaule* is interpreted as a better indicator of relationship rather than the co-occurrence of eight isomers, produced by the same biogenetic pathway.

Taxonomic Implications for *Cypripedium arietinum*

Taxonomic relationships, as indicated by the clustering analysis utilizing the coding method 2, will be discussed and compared with other recent evidence from independent studies. The segregation of *C. arietinum* into *Criosanthes* (Atwood 1984) is not supported by the cluster analysis performed here. *Cypripedium arietinum* clusters closely to members of the *C. calceolus* complex. The clustering of this taxon with members of the complex is congruent with results obtained by Case (1993). Figure 8 shows a comparison of the phenograms of Case (1993) and

of that obtained in this study. Although the relative position of *C. arietinum* differs between the studies, and although the taxon is distinct with respect to its floral morphological features, many independent data sets including those constructed from plastid DNA, allozyme variation, and secondary metabolites, support its inclusion in *Cypripedium.*

Erroneous similarity between *Cypripedium candidum* and *Cypripedium arietinum*

The relatively high similarity displayed between *C. arietinum* and *C. candidum*, shown in the Figure 3 phenogram, requires an explanation, because, on the basis of floral morphology, these taxa would not be predicted to be more similar than *C. candidum* is to *C. calceolus* var. *parviflorum* or *C. calceolus* var *pubescens.* The affinity between *C. arietinum* and *C. candidum* is due to the co-occurrence of benzaldehyde, benzene methanol, 4-methoxy benzaldehyde, and 4-methoxy benzene methanol. These four compounds are likely produced via the same single biogenetic pathway. This is compared to the co-occurrence of these same four compounds plus linalool, representing two independent biogenetic pathways, between *C. calceolus* var. *parviflorum* and *C. candidum.* Calculation of the Jaccard similarity coefficient includes mismatches (0,1) and (1,0) which lower the similarity between any two taxa producing different compounds. Although *C. arietinum* and *C. candidum* only share four biogenetically related compounds representing the expression of a single pathway, they do not possess many differences in overall composition. In contrast, *C. calceolus* var. *parviflorum* and *C. candidum* share two biogenetic pathways yet they also express several unshared pathways. For this reason, two taxa with very different gross morphologies are represented as chemically similar although they share only one major biogenetic pathway. In spite of the different pathways not shared between *C. candidum* and *C. calceolus* var. *parviflorum*, the two taxa are biogenetically more related to each other than *C.*

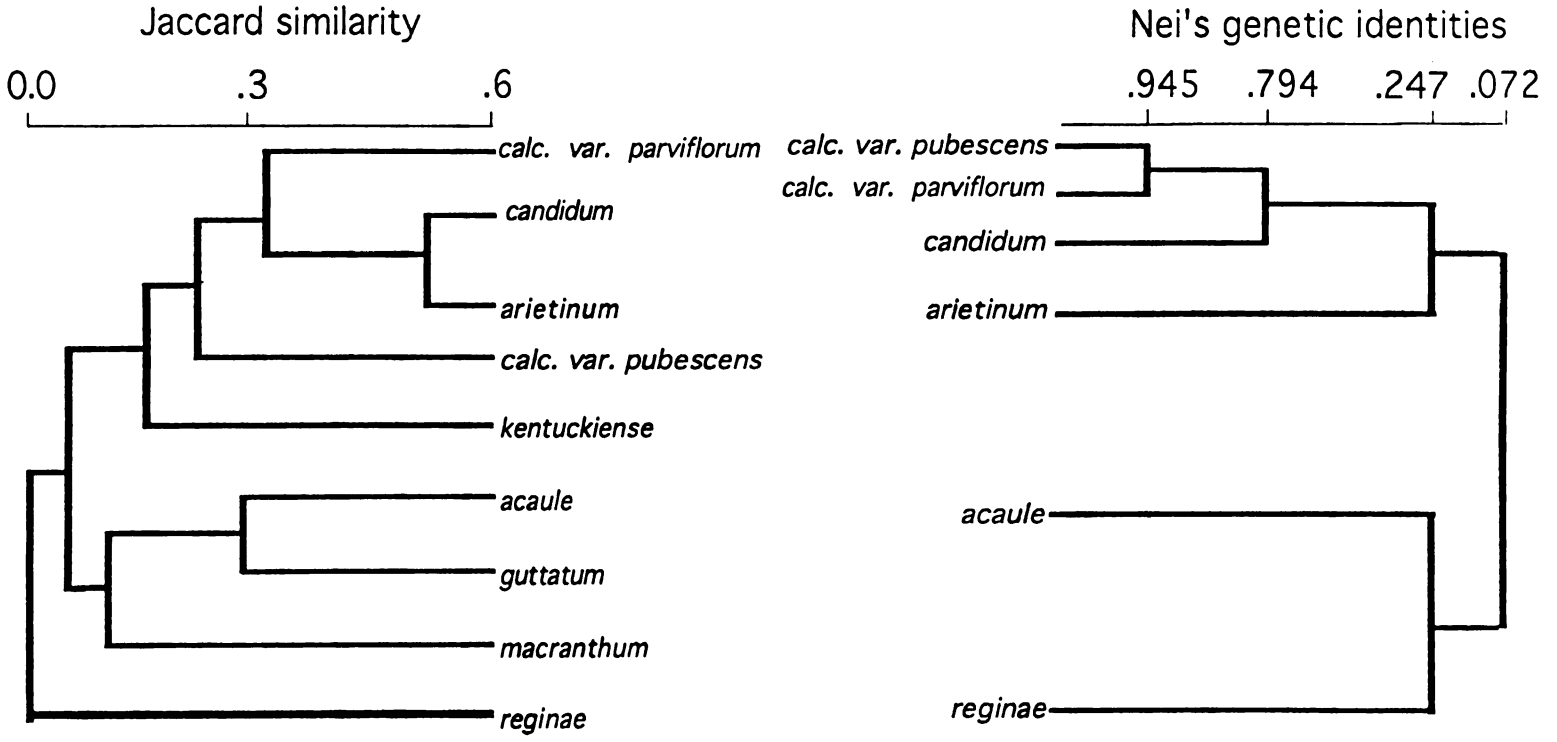


Figure 8. Comparison of UPGMA phenograms based on fragrance similarity (this study) and genetic identities (Case 1993).

arietinum is to either of them, yet in terms of Jaccard similarity, *C. arietinum* and *C. candidum* are the most alike.

Taxonomic implications for *C. calceolus* complex

The taxa composing the cluster of *C. candidum*, *C. calceolus* var. *parviflorum*, *C. calceolus* var. *pubescens*, and *C. kentuckiense* (with the exception of *C. arietinum*) are recognized as the *C. calceolus* complex (Atwood 1984). This grouping of taxa within the complex is generally congruent with relationships presented by Case (1993). While the delimitation of the entire complex is not in question, the specific relationship between *C. calceolus* var. *parviflorum* and *C. calceolus* var. *pubescens* remains ambiguous. The phenogram in Figure 5 supports Atwood (1984) in recognizing *C. calceolus* var. *pubescens* and *C. calceolus* var. *parviflorum* as distinct species, but the phenogram of Figure 6 does not. The separation of *C. calceolus* var. *pubescens* from *C. calceolus* var. *parviflorum* is ambiguous because the clone sampled by Bergstrom et al. clusters with *C. calceolus* var. *parviflorum*, while the clone sampled in this study was distinct from it. The fact that the *C. calceolus* var. *pubescens* individual sampled here is very different in composition from the individual sampled by Bergstrom *et al.* indicates that extensive qualitative fragrance variation exists within this taxon. This pattern should be examined further before taxonomic judgment is reached concerning the separation of *C. calceolus* var. *parviflorum* and *C. calceolus* var. *pubescens* into distinct species.

Taxonomic implications for other taxa

The species in the genus not included in the *C. calceolus* complex form a loose cluster distinct from this complex and are not very similar to each other with respect to fragrance composition. This is concordant with results obtained by Case (1992). Although she did not study *C. macranthum* and *C. guttatum*, *C. reginae* and *C. acaule* only possess a genetic identity of 0.247 with each other. The

relationship displayed between *C. acaule* and *C. guttatum* found in this study is surprising as the taxa are very dissimilar in floral morphology. The placement of *C. macranthum* is distant from all other taxa investigated. Stoutamire (1967) noted a fragrance for this taxon which is similar to that found in *C. acaule* and *C. arietinum*, yet it has a shared similarity coefficient less than 0.1 with *C. acaule* and *C. arietinum*. Little can be proposed for the relationships among these taxa on the basis of floral fragrance alone, yet it appears that none of them are very closely related to each other.

Population Variation

Fragrance profiles for individuals within the same population differ qualitatively and quantitatively. Entire classes of compounds were found to vary between individuals both between and within populations. Monoterpene hydrocarbons, cinnamic acid derivatives, straight chain alcohols, sesquiterpene alcohols, and tri-methoxy benzene are only expressed in a few individuals. Although these variant compounds were often minor components of the fragrance profiles, they do represent variant expressed pathways and hence potentially different genotypes.

The results of the cluster analysis in which all individuals of *C. calceolus* var. *parviflorum* are included indicate that the qualitative variation observed is not taxonomically important, as all individuals cluster more closely with each other than with other taxa in the analysis. This is an important result which demonstrates that although fragrance variation may appear high between individual clones of a taxon, these differences may not have taxonomic importance when compared against variation displayed among closely related taxa. Other taxonomic studies utilizing fragrance data (Gregg 1984; Bergstrom et al. 1992) have not compared fragrance composition across several related species

within the genus sampled. Knowledge of the extent of variation across congenerics allows a more effective evaluation of "significant" fragrance compound variation.

Lack of concordance

The individual of *C. calceolus* var. *pubescens* sampled here differ greatly in qualitative composition in comparison to that sampled by Bergstrom et al. (1992). The differences in trapping techniques prohibit strong inferences from the comparisons. It should be noted, however, that most of the compounds isolated and identified by Bergstrom et al. were also isolated and identified by the technique employed here in different taxa. Thus, differences in affinity of sorbent tubes are assumed to be minimal. The differences in fragrance composition between the individuals of *C. calceolus* in the two studies indicate that extensive genetic variation may exist for the production of fragrance compounds, especially when quantitative differences are considered. In addition, comparison of results obtained by Holman (1983) indicate substantial variation in fragrance composition between the same taxa sampled here. The wide variation in reports of fragrance composition for this taxon emphasizes the need for larger sample sizes if the data are to be used in taxonomic studies.

Figure 8 shows a comparison of the UPGMA-derived phenograms of Figure 5 and Case (1993). While the distinction of two primary groups within the genus is concordant between the studies, specific relationships are not. The relationships proposed by Case (1993) between *C. calceolus*, *C. candidum*, and *C. arietinum* are very different from ours. Due to limitations in the coding procedure and in the calculation of similarity values for floral fragrance data, as discussed above, the specific relationships indicated in the Figure 5 phenogram are not strongly suggestive of evolutionary relationships. The clustering of taxa may more

accurately represent pollinator relationships as suggested by Whitten & Williams (1992). This hypothesis remains to be tested.

Conclusion

The use of floral fragrance-compound data in combination with objective measures of comparison result in clusters of taxa which reflect general taxonomic relationships recently proposed for *Cypripedium*. The "broad" resolution of this technique in the identification of general evolutionary relationships is supported by other data sets. However the "fine" resolution produced by this technique is not. The cluster analysis has been useful in delimiting the *C. calceolus* complex as separate from all other taxa. Relationships among other taxa within *Cypripedium* are difficult to infer given the data accumulated thus far, although the retention of *C. arietinum* within *Cypripedium* appears well supported. The phenetic methods used have several shortcomings when used in conjunction with volatile compound data for the purpose of estimating phylogenetic relationships. The treatment of isomers as independent compounds for use in calculation of similarity coefficients will result in erroneously high values. In addition, the similarity values utilized do not accurately reflect the importance of shared biogenetic pathways if multiple unshared pathways are present between two taxa. The systematic investigation of several putatively related taxa is necessary before taxonomic judgments can be effectively made because the variation between taxa can only be evaluated once variation across other related taxa is considered as well.

PART 2: CLADISTIC RELATIONSHIPS WITHIN *CYPRIPEDIUM* INFERRED FROM FLORAL FRAGRANCE-COMPOUND DATA

INTRODUCTION

The genus *Cypripedium* is composed of approximately 40 species having diverse floral morphologies. The species are fairly similar vegetatively, so little has been proposed about phylogenetic relationships due to a lack of polarizable reproductive character trends. Several species are native to North America and have received considerable taxonomic attention in the last ten years. One group of taxa that is morphologically coherent is the *C. calceolus* complex (Atwood 1984), which includes *C. calceolus* var. *parviflorum*, *C. calceolus* var. *pubescens*, *C. calceolus* var. *planipetalum*, *C. kentuckiense*, *C. montanum*, and *C. candidum*. This complex shares presence of linear-lanceolate lateral petals, an inflated pouch with an orifice on the top, and a white or yellow labellum (Atwood 1983). *Cypripedium calceolus* is noted for extreme morphological variation, and the designated varieties are difficult to distinguish in many cases. In addition, hybridization between the named varieties of *C. calceolus* as well as *C. candidum* results in populations which are taxonomically ambiguous. Introgression confounds taxonomic inference and has been demonstrated between *C. calceolus* var. *pubescens* and *C. candidum* using isozyme data (Klier 1992). While membership of these taxa within the complex seems reasonably certain, it is not known which ones are more evolutionarily derived than others.

Other taxa within *Cypripedium* are a diverse assemblage with much variation in labellum morphology and color. *Cypripedium reginae* has an inflated labellum similar to that found in the *C. calceolus* complex yet differs in that the lateral petals

are oblong-lanceolate and the pouch is pink. *Cypripedium acaule* is distinct in that it has an anterior suture in the pink labellum and has only two leaves. This labellum architecture is very different from any other taxon in the genus. *Cypripedium arietinum* has a diminuitive flower characterized by a spurred pouch and free lateral sepals. The lateral sepals in all other taxa are fused into a synsepal, although various degrees of splitting of the synsepal are found in *C. calceolus* (personal observation). *Cypripedium gutattum*, native to Alaska and Japan, has an unusual striped flower with a tubular labellum similar to that found in the tropical genus *Paphiopedilum*. *Cypripedium macranthum* is an Asian species characterized by an inflated pouch which has a raised ridge around the orifice. The flower color is burgundy and the lateral petals are oblong-lanceolate. Its relationships to other taxa within the genus are unknown, although artificial hybrids have been reported between it and *C. calceolus* (Stoutamire 1967). Phylogenetic relationships remain to be suggested for the morphologically divergent taxa included here.

The first phylogenetic treatment for several North American species of *Cypripedium* was that by Atwood (1984), which also included the tropical slipper orchids, *Paphiopedilum*, *Phragmipedium*, and *Selenipedium*. Atwood proposed phylogenetic relationships within the subfamily Cypripedioideae based upon morphological, anatomical, and karyological data. In his treatment of *Cypripedium*, *C. arietinum*, the rams head lady slipper, is placed in a segregate genus, *Criosanthes*. This decision was based upon unique floral morphological features including a spurred pouch, a staminode which nearly resembles a fertile anther, and free ventral sepals. These characters were interpreted by Atwood as primitive and divergent for the genus. Relationships of the remainder of the genus include a group treated as "most" *Cypripedium* (*C. calceolus* complex, *C. macranthum*, and *C. reginae*) which were interpreted as intermediately advanced and somewhat

separate from the primary evolutionary line that culminates in the most highly derived taxa, *C. acaule* and *C. guttatum*. Atwood (1984) urges: "If this work stimulates botanists into seeking ways for presenting falsifiable hypotheses concerning their [the lady slippers] origin and evolution it shall have served its purpose." This pioneering study of the phylogenetics of the slipper orchids has indeed spurred several other studies within the subfamily and especially of *Cypripedium*.

The most recent work addressing evolutionary relationships of several North American taxa was that by Case (1993). She utilized allozyme-frequency information to show that *C. calceolus* and *C. candidum* are closely related, sharing a genetic identity of 0.794. An interesting result is the relatively high genetic identity of *C. arietinum* shared with *C. calceolus* and *C. candidum* (Nei's genetic identity = 0.247). This value is higher than that obtained for *C. acaule* and *C. reginae* with the other three taxa (0.072). While her study proposes potentially evolutionarily related taxa, it does not distinguish phylogenetically advanced taxa from primitive taxa. Case also investigated evolutionary relationships for infraspecific taxa of *C. calceolus*. A genetic identity of 0.945 was reported between *C. calceolus* var. *parviflorum* and *C. calceolus* var. *pubescens*. This is comparable to identities reported for infraspecific taxa in other genera. Although other workers have proposed recognition of two distinct species, Case recommends their inclusion in a single species.

Albert (1993) has investigated systematic relationships for *Cypripedium* based upon rbcL plastid variation, floral morphology, and vegetative morphology and anatomy. Within the genus, *C. plectrochilon* (putatively related to *C. arietinum*) is most primitive and *C. reginae* is relatively advanced, although the data sets used do not provide good resolution for these relationships within the genus (Albert pers. comm.).

The use of chemical data in conjunction with cladistic methods to infer phylogeny at lower taxonomic levels has proven successful in *Tetragonotheca* and *Iva* (Asteraceae) (Seaman & Funk 1983). Biogenetic transformation series for sesquiterpene lactones were constructed based upon skeletal and substitutional features. Outgroup analysis was used to determine the plesiomorphic and apomorphic states. Synapomorphies were reconstructed using parsimony and resulting cladograms were compared with results from analyses derived from independent data sets such as those based upon morphology.

Although floral fragrance compounds have not been utilized previously for reconstructing phylogeny, there are several advantages to using this type of data over others. Volatile compounds can be objectively collected and their structure unambiguously identified using gas chromatography-mass spectrometry. This analytical technique is relatively easy to use with the only shortcoming being the ambiguity in positively identifying isomers. However, isomers of a particular compound are often chromatographically distinct, making them useful for comparative studies. Compounds can be described, when pathway information is sufficient, as biogenetically simple or complex with ordered enzymatic steps elucidated in many cases. Floral fragrance compounds are often structurally diverse, produced via several independent primary biogenetic pathways, offering many characters for systematic evaluation. Although use of floral fragrance data has many advantages, a major disadvantage is the widespread distribution of identical secondary metabolites in phylogenetically unrelated taxa. Whitten and Williams (1992) report a limited use of fragrance-compound data in reconstructing phylogenies for euglossine-pollinated orchids, due to the widespread distribution of terpenes, which are found to be produced in many plants and fungi.

Volatiles in orchids are ultimately derived via three pathways, mevalonic acid, shikimic acid, and acetic acid (Bergstrom 1991). The mevalonic acid pathway can yield one of two volatile classes of compounds, mono- and sesquiterpenoids. The shikimic acid pathway, via phenyl alanine or not, can yield diverse groups of compounds including benzaldehydes, benzoic acids, phenyl acetic acids, anthranilic acids and cinnamic acids. The acetate pathway yields derivatives such as straight-chain alcohols, acetates, or esters. Although the mevalonic, shikimic, and acetic acid pathways are ubiquitous in distribution, the distribution of their derivative pathways are not and may provide useful phylogenetic information.

Fragrance chemistry within *Cypripedium* has been reported for several North American taxa. In no cases, however, have the data been utilized in a cladistic analysis. This may be due to the complexity in coding fragrance data. The previous studies have compared quantitative and qualitative variation for taxonomic purposes (Bergstrom et al. 1992; Gregg 1984). Taxonomic judgement has been based upon perceived importance of this variation to the pollination system of these taxa. Bergstrom et al. (1992) have discussed the variation in fragrance chemistry in three varieties of *C. calceolus*. They recommended the recognition of three distinct species on the basis of extensive qualitative and quantitative variation. The results of cluster analyses provided in Chapter 1 show that qualitative distributions of biogenetically related floral fragrance compounds are useful in delimiting broadly related groups of taxa only. A different approach is taken here to analyze the distribution of biogenetic pathways and enzymatic steps as potentially informative synapomophies for phylogenetic reconstruction.

The present study is an attempt to suggest relationships within the genus *Cypripedium*, including *C. arietinum*, and to suggest phylogenetic placement for these taxa using floral fragrance data alone. Floral fragrance compounds have been isolated and identified and are used as an independent data set. The coding

step is the most important part of the process in any cladistic analysis. The effective coding of secondary chemical-compound data is dependent upon a knowledge of biogenetic interrelatedness of the compounds to ensure coding of independent characters. Fragrance compounds should be arranged according to biogenetic relatedness as is stressed by Salatino and Gottlieb (1981). Enzymatic steps required for interconversions between compounds are coded as present or absent and were utilized as the data matrix for the analysis. In this coding method, basic biogenetic pathways are treated as characters and specific transformations between compounds are character states. This coding method represents the dependent nature of the data. Parsimony analyses are used to discern cladistic relationships among the taxa. Relationships among the morphologically divergent taxa *C. acaule, C. reginae, C. guttatum,* and *C. macranthum* will be explored and potential associations proposed based upon considerations of other recent findings. Other proposed relationships to be investigated include the controversial position of *C. arietinum* and the questionable hierarchical placement of designated varieties of *C. calceolus.*

MATERIALS AND METHODS

Lists of taxa sampled and compounds identified for each are provided in Chapter 1.

Biogenetic schemes

Schrier (1981) and Croteau (1991) summarize the results of various biosynthetic studies to propose general pathways for the production of most volatile compounds. Most of the pathways have been found to occur in higher plants, although occasionally the only direct evidence derives from experiments with fungi. Despite this work, which addresses the biogenesis of major skeletal

types, little is known about pathways leading to substitutions and skeletal modifications such as reduction or acetylation of functional groups. Although pathways have not been elucidated for these modifications of basic skeletons, pathways have been assumed in most instances. Support for the assumed pathways is based upon the presence of all related compounds in one species. For example, in *C. arietinum*, benzaldehyde, benzene methanol (the reduced aldehyde), 4-methoxy benzaldehyde, and 4-methoxy benzene methanol (methoxy substituted alcohol) are all produced (See figure 2 for assumed biogenetic pathway).

Fatty acids

The biogenesis of fatty acid derivatives is widely reported (Croteau, 1991; Shreier, 1983). Catabolism of fatty acids such as linoleic acid by beta-oxidation can lead to methyl and ethyl esters of mono and dienoic acids (including C-10 and C-12 chains) in ripening Bartlett pear tissue. This is a likely pathway responsible for the derivatives found in *C. acaule* and *C. guttatum*. Many reaction sequences could result in the formation of dodecanol and various aliphatic acetates found in *C. calceolus* and *C. kentuckiense* respectively. Labeling experiments in banana slices have shown fatty acids, such as hexanoic acid, to be reduced to 1-hexanol (Tressl & Drawert, 1973 in Schreier 1984). Analogous experiments have shown the same reaction to occur in labeled decanoic acid. Aliphatic alcohols may also be formed during hydroperoxidation reactions in which fatty acids, liberated from membrane glycerolipids, are cleaved to yield aldehydes of various lengths which may undergo further reduction. Acetylation of the alcohol products could explain the occurrence of aliphatic acetates found in *C. kentuckiense*. Beta-oxidation may produce the aliphatic alcohols and acetates, found in these species of *Cypripedium*. Evidence for this exists as the distribution of volatiles included chain lengths of 8,

10, and 12 which may represent products formed at various stages of catabolism of a single long-chain precursor. Due to the ambiguity in the origin of the fatty acid compounds isolated in several species, an independence of biogenetic pathways was assumed between the two primary types, methyl and ethyl esters, and alcohols and acetates. Figure 1 shows a proposed biogenetic scheme which served as the basis for our coding procedure.

Benzenoids

The biogenesis of benzenoids is well elucidated for higher plants. The anthranilic acid pathway results in the formation amino benzaldehyde and of indole ultimately, both of which are present in *C. calceolus* and *C. acaule*. The dehydroshikimic acid pathway results in the formation of benzoic acid. From benzoic acid, the benzaldehyde and benzene methanol derivatives are easily formed by reduction. Many of the species produce derivatives from benzoic acid, a non-volatile precursor. The cinnamic acid pathway from phenylalanine results in the formation of cinnamaldehyde derivatives and eugenol. *Cypripedium macranthum* and *C. guttatum* express this pathway. It is not clear, however, if the phenyl ethyl derivatives found in *C. candidum*, *C. calceolus*, and *C. kentuckiense* are formed by decarboxylation and reduction of cinnamic acid or if they can derive directly from phenylalanine transformation. For this analysis these are assumed to be derived from cinnamic acid as suggested by Croteau (pers. comm.). Benzyl benzoate biogenesis is uncertain because it may derive from a benzaldehyde precursor or directly from benzoic acid. Due to this ambiguity, this compound has been coded as directly arising from benzoic acid. Likewise, 1,4-dimethoxy benzene biogenesis is unknown and has necessitated a coding procedure which treats it as independently derived from any other compound. Four multistate characters were constructed for the benzenoid compounds; benzoic acid,

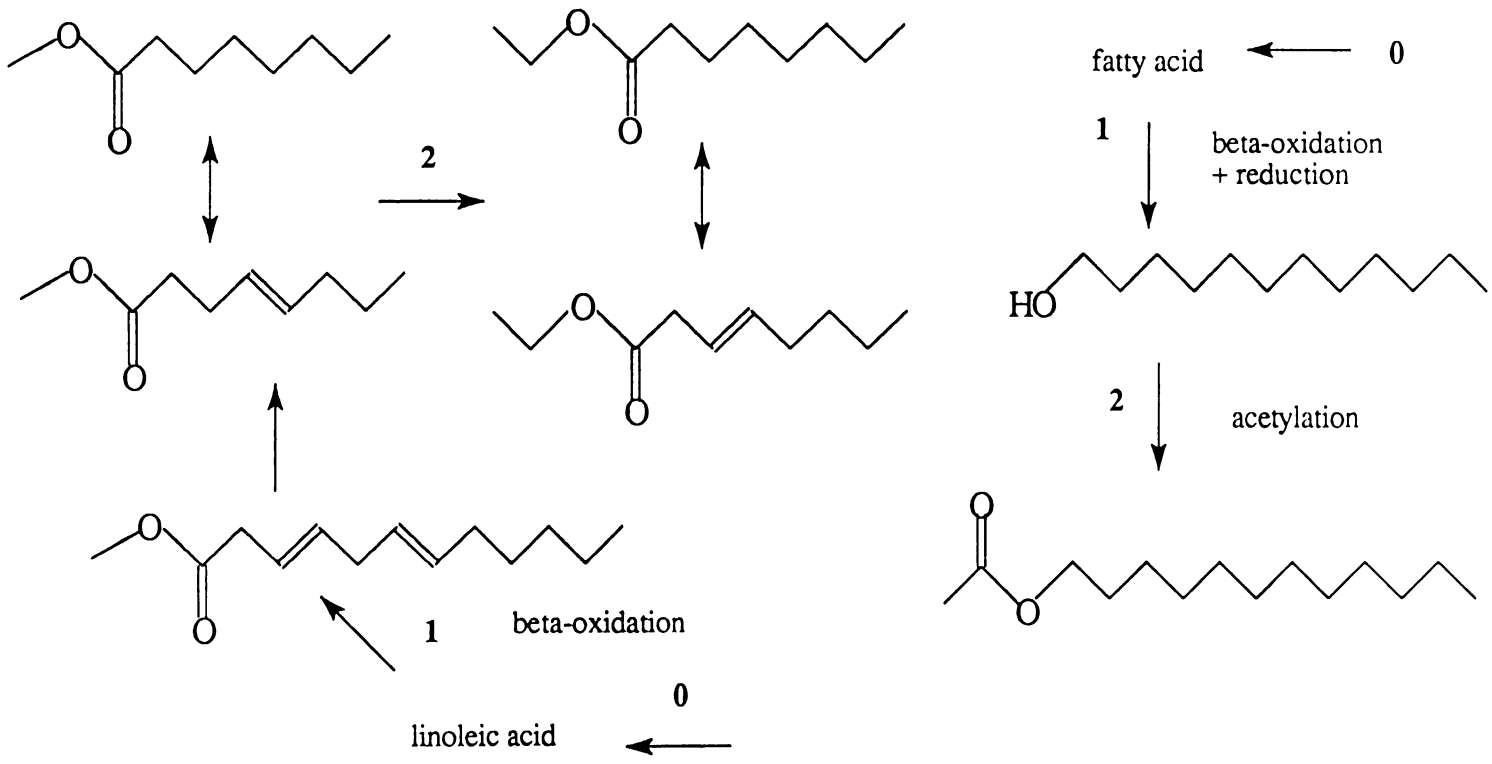


Figure 1. Proposed biogenetic routes for volatile fatty acid derivatives.

Character states representing enzymatic conversions are shown in bold numbers.

cinnamic acid, anthranilic acid, and 1,4-dimethoxy benzene. Figure 2 shows the biogenetic scheme which served as the basis for this coding.

Monoterpenoids

Monoterpenoid biogenesis has received much attention, and many of the reactions involved in the formation of cyclohexanoid monoterpenes have been elucidated (Croteau & Karp 1991). However, exclusive pathways for particular terpenoids do not appear to exist. Geranyl pyrophosphate (GPP), formed by the condensation of dimethyl allyl pyrophosphate (DMAPP) and isopentyl pyrophosphate (IPP), is the ultimate precursor of monoterpene hydrocarbons and alcohols, including acyclic, monocyclic or bicyclic skeletons of each. Linalool is reportedly formed from GPP although it can arise from its isomer linaloyl pyrophosphate (LPP) as well. LPP can produce other mono- and bicyclic terpenes also. Due to the ambiguity of a common precursor which can produce this wide diversity of derivatives, all monoterpenes identified in this study were artificially divided into monterpene hydrocarbons and monoterpene alcohols. A common precursor of GPP is assumed for all derivatives in each of these two classes. Figure 3 shows the proposed biogenetic scheme for both hydrocarbons and alcohols.

Sesquiterpenoids

Sesquiterpene biogenesis is poorly known. Farnesyl pyrophosphate is reported to give rise to all sesquiterpenes, although any of four configurational isomers can give rise to derivative sesquiterpenes. Farnesene results from the reduction of farnesol, a primary product of FPP. Nerolidol results from the dephosphorylation of neryl pyrophosphate, which is an isomer of FPP. Caryophyllene arises by transformation from the original precursor FPP without isomerization to NPP (Croteau & Karp 1991). Figure 4 shows this basic biogenetic scheme.

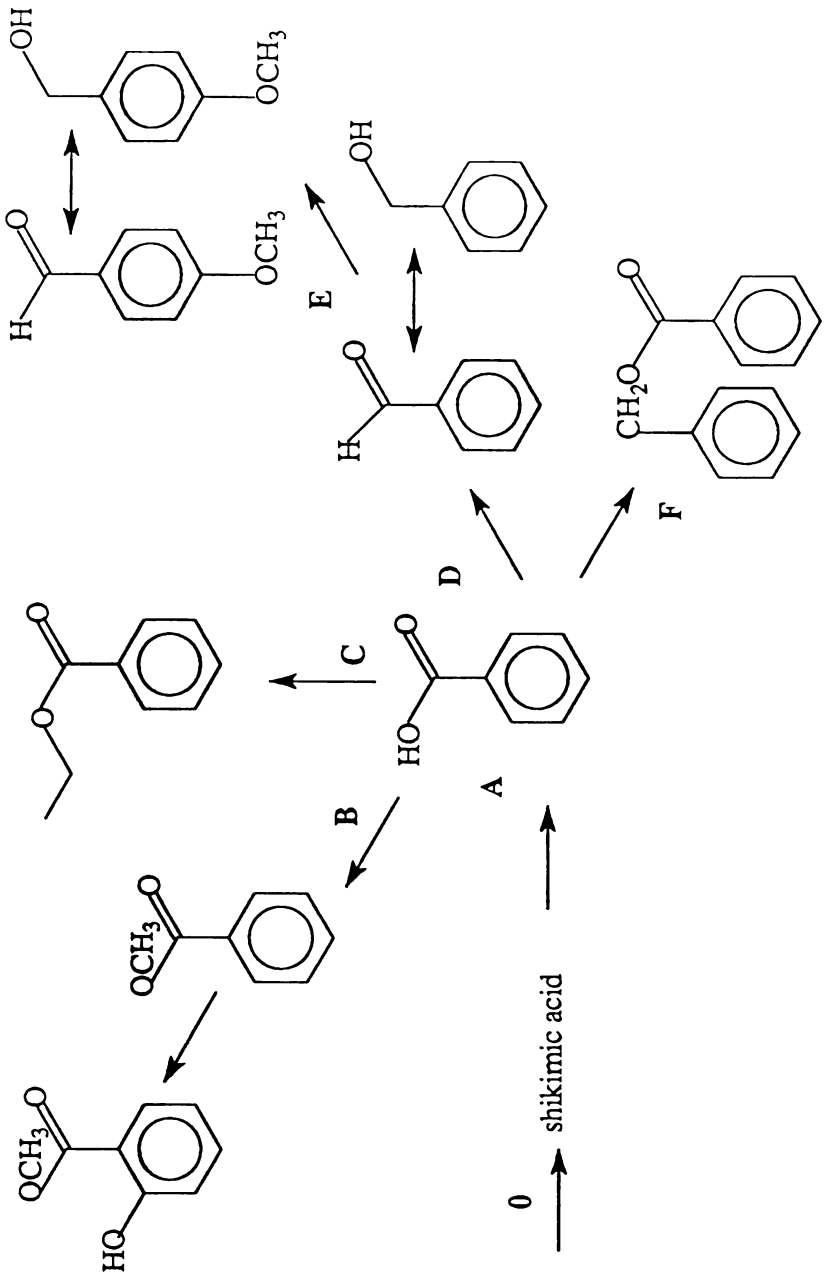


Figure 2. Proposed biogenetic routes for benzoic acid derivatives.

Character states representing enzymatic conversions are shown in bold letters.

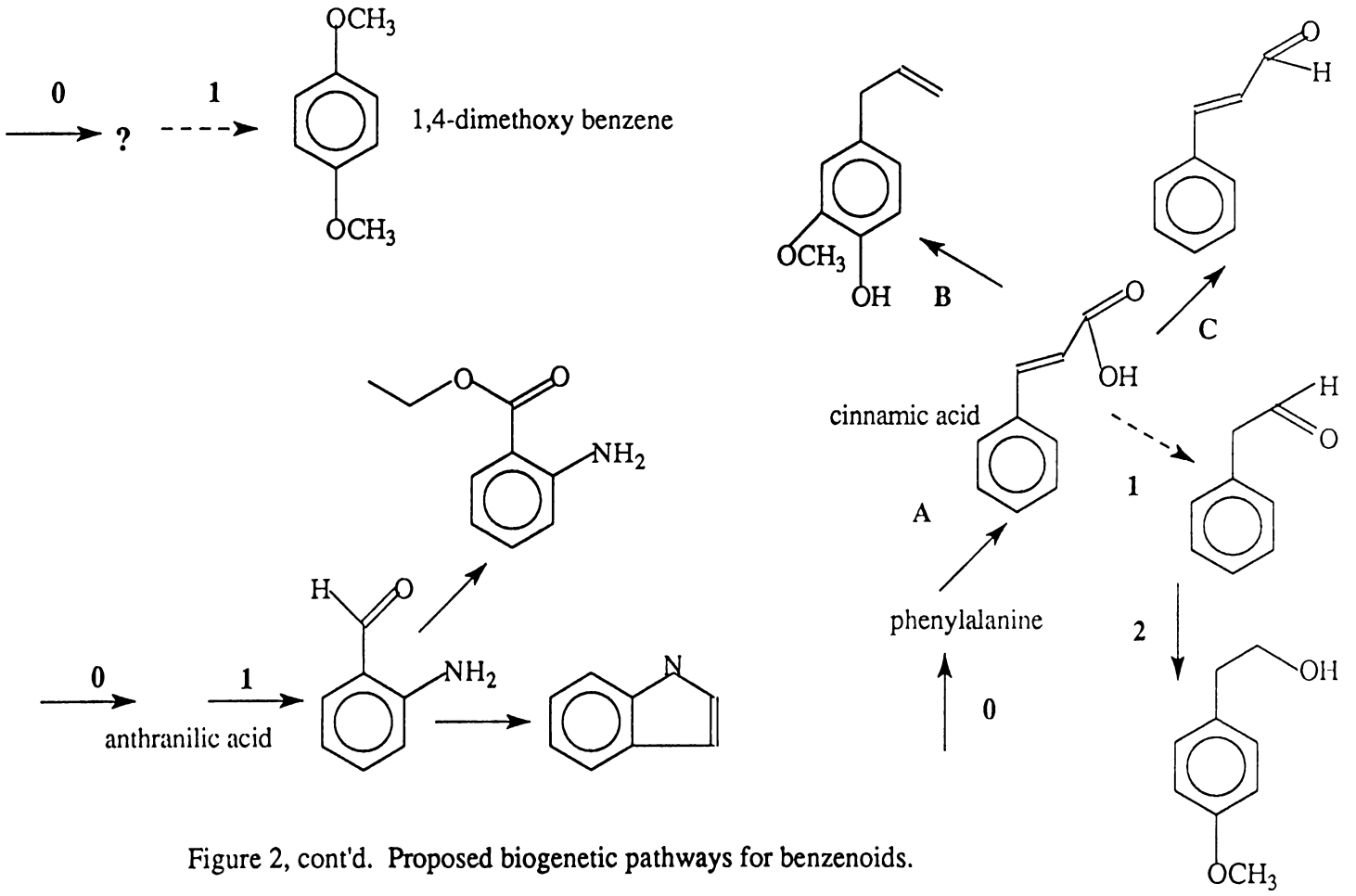


Figure 2, cont'd. Proposed biogenetic pathways for benzenoids.

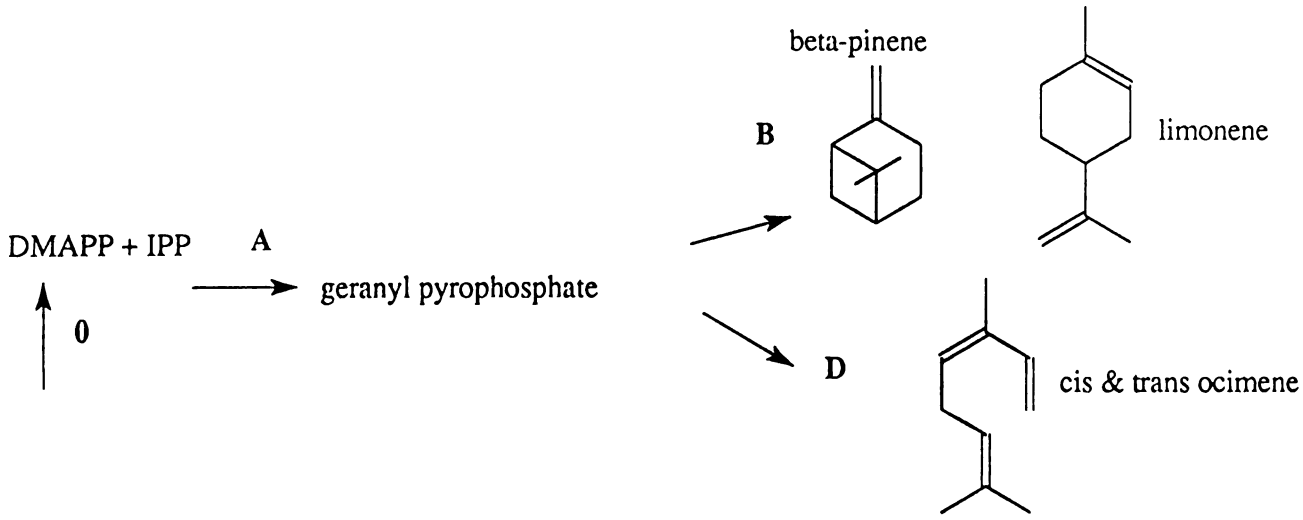


Figure 3. Proposed biogenetic routes for monoterpene hydrocarbons.

Character states representing enzymatic conversions are shown in bold letters.

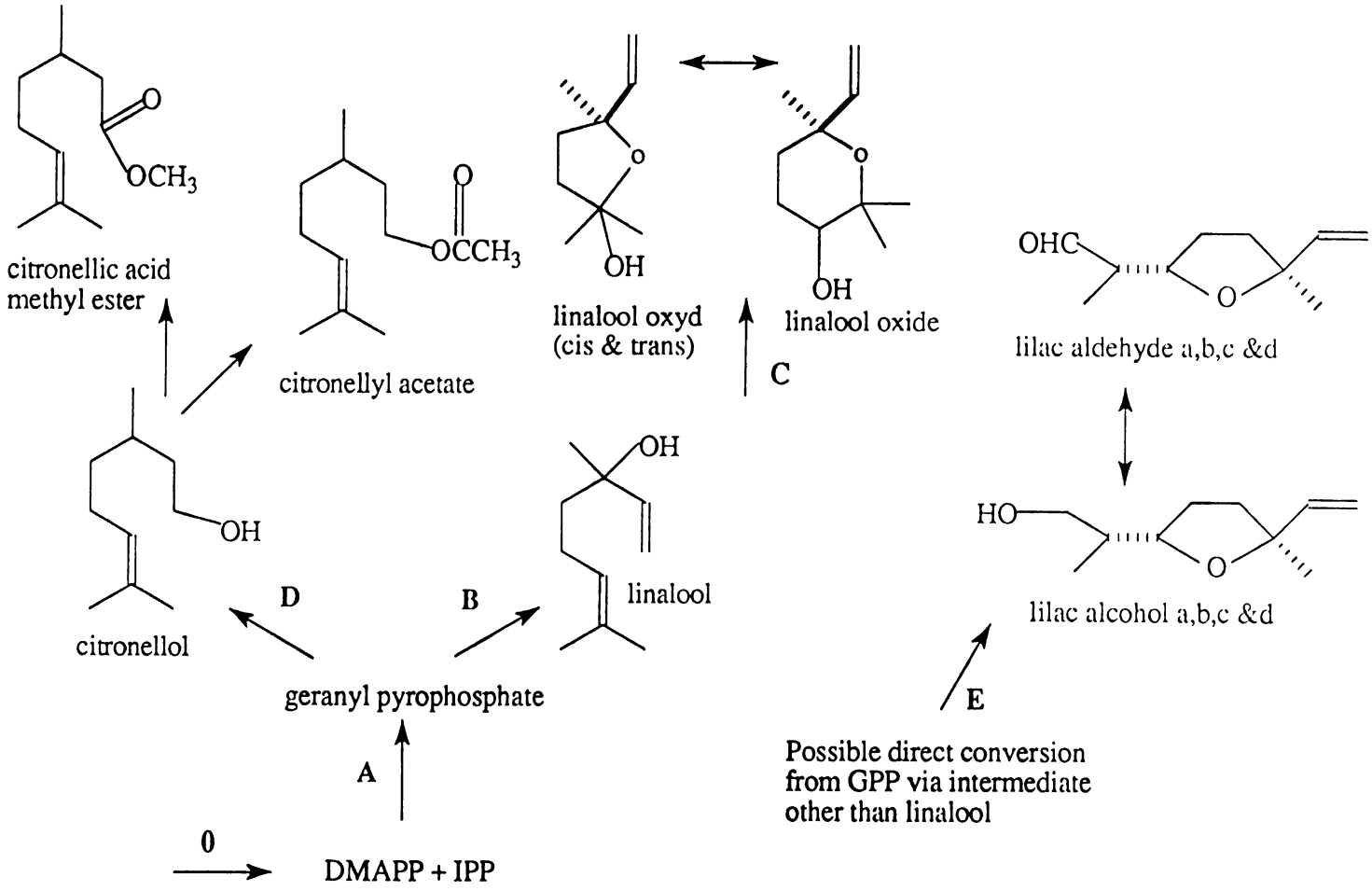


Figure 3, cont'd. Proposed biogenetic routes for monoterpene alcohols.

Unknown biogenesis

In some cases compounds could be derived via alternate pathways. One class of compounds, the lilac alcohols and aldehydes, are of unknown biogenetic origin. The position of the epoxide bridge, between 2-C and 7-C, however allows an inference into its biogenetic origin. This compound may be derived from monoterpene polyols. The position of alcohol substitutions determine where the epoxide bridge may form. Linalool has an alcohol substitution at the 2-C position. No other monoterpene alcohol has alcohol substitution at this position. Despite a potential derivatization from linalool, the pyranoid monoterpenes had to be coded as biogenetically independent from it. The taxa which produce the lilac alcohols do not produce linalool so no biogenetic association was assumed. Figure 3 shows the placement of the enigmatic alcohols within the proposed pathways.

User-defined characters

Figure 5 shows the user-defined character types used in the cladistic analysis. PAUP and MacClade allow these character types, which could not be effectively coded otherwise. These allow the representation of proposed biogenetic pathways as character-state transformations. The transformations between states are ordered, in that the formation of a particular compound is dependent upon the formation and modification of its precursor first. Each character-state represents an enzymatic step in a pathway. The user-defined character states are then reconstructed for the taxa within *Cypripedium* in the most parsimonious manner. The most parsimonious topology represents the distribution of enzymatic steps for the taxa such that independent origins of biogenetic pathways are minimized. Figure 6 shows the data matrix used in the analysis in which the character types are represented by the user-defined characters of Figure 5.

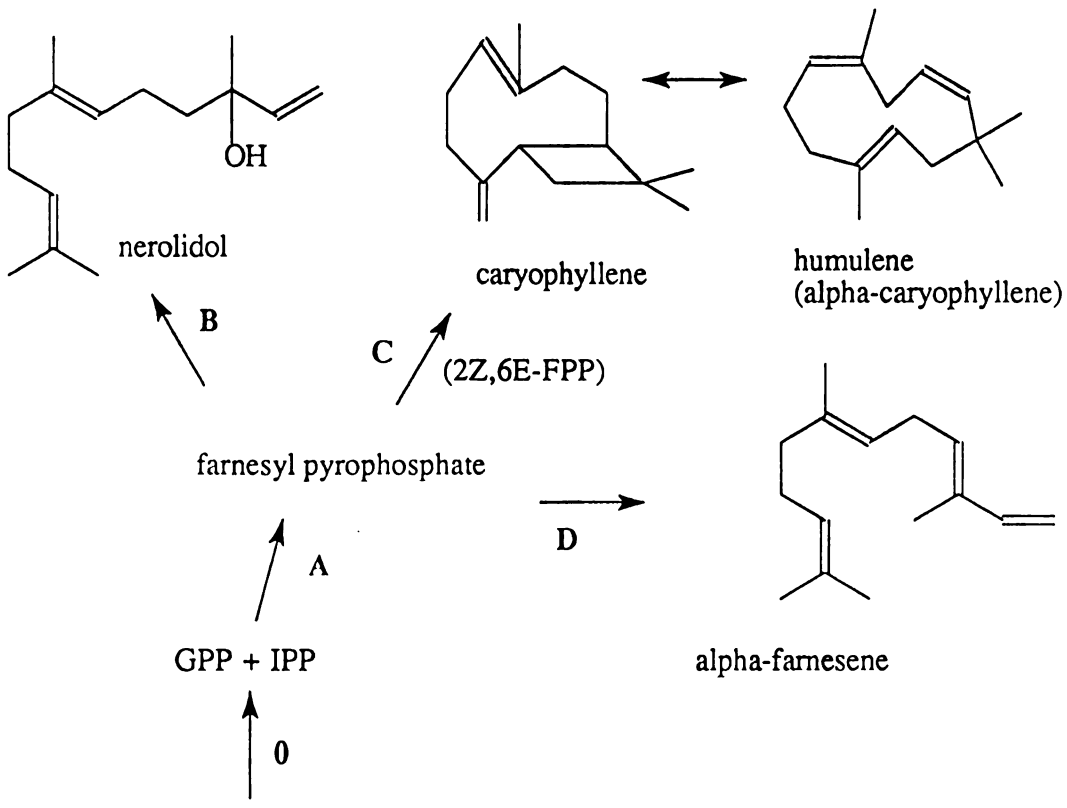


Figure 4. Proposed biogenetic routes for sesquiterpenes.

Character states representing enzymatic conversions are shown in bold letters.

Cladistic Analysis

PAUP (Swofford 1991) was used to generate trees representing cladistic relationships of the taxa. Several assumptions were used to generate the most parsimonious trees.

Character types

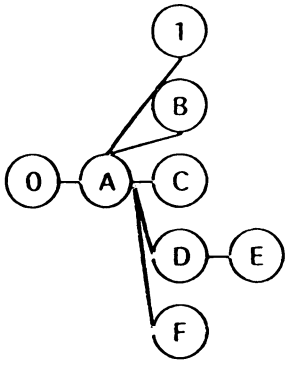
Unordered character states were assumed for all binary coded characters. Multistate characters with states that could be organized into linear transformations were treated as ordered characters. The ordered characters were coded according to the biogenetic pathways by which they are formed. Hence, in a multistate character, with states 0-1-2, the formation of 2 is dependent upon the initial formation of 1. Figure 1, representing the proposed biogenesis of fatty acid derivatives, illustrates the coding of linearly ordered character states. The characters representing biogenetically related compounds which could only be represented by branching transformation series were treated as user-defined character types. The character states are essentially ordered in the sense that the specified path must be followed for transitions between various states.

Rooting

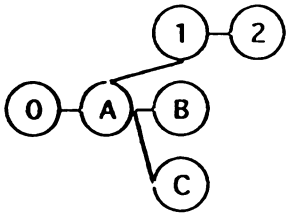
An hypothetical ancestor served as the root of the analysis. All states were assigned as unknown, which is the default setting of PAUP. The assignment of all unknown states allows the most parsimonious assignment of character states to the ingroup under all circumstances. This analysis was undertaken to achieve locally most parsimonious resolutions for the ingroup. The unrooted trees obtained can be rooted anywhere without a change in tree length.

Outgroup analysis was carried out using members of *Paphiopedilum* sampled by Barkman (unpublished). The rooted trees obtained allow hypotheses about

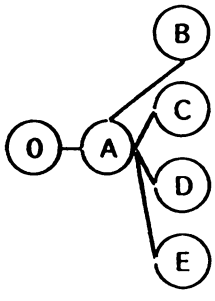
benzaldehyde



cinnamaldehyde



sesquiterpene



monoterpene

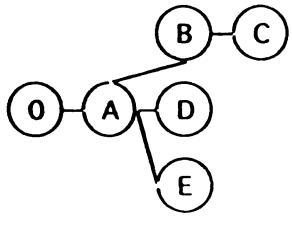


Figure 5. User-defined character-state trees.

bark.pau copy 2 1		1	2	3	4	5	6	7	8	9
		Cinn	dime	anth	benz	mon	mon	sesq	fatty	fatty
1	parviflorum	0	1	1	E	B	C	B	1	0
2	pubescens	2	1	0	D	B	0	0	1	0
3	candidum	1	0	0	E	0	B	0	0	0
4	arietinum	0	0	0	E	0	0	0	0	0
5	kentuckiense	1	0	0	0	B	B	B	2	0
6	acaule	0	0	1	C	D	E	0	0	2
7	reginae	0	0	0	B	0	0	0	0	0
8	guttatum	B	0	0	D	0	D	D	0	1
9	macranthum	C	0	0	F	B	E	C	0	0

Figure 6. Data matrix used in exhaustive search using Wagner parsimony.

ancestral and advanced taxa within the genus. Our outgroup includes representatives from several different sections of *Paphiopedilum* (sensu Cribb 1986). Sections of *Paphiopedilum* included *Brachypetalum*, *Parvisepalum*, *Cochlopetalum*, *Coryopetalum*, and *Barbata*. Species chosen include *P. emersonii*, *P. niveum*, *P. primulinum*, *P. kolopakingii*, *P. argus* and *P. armeniacum*. The choice of several outgroup taxa may allow a better resolution of plesiomorphic states to be used in the analysis and avoid the bias of choosing a single taxon. The ordering of these taxa within the outgroup was based upon the results of a phylogenetic analysis by Albert (1992). The ordering of outgroup taxa is an important consideration in a cladistic analysis because of the effect it can have on character states of the outgroup node.

The initial analyses in which character states were reconstructed with no specification of outgroup taxa resulted in topologies which placed members of *Paphiopedilum* within *Cypripedium*. This analysis was undertaken to see if members of *Paphiopedilum* would be resolved separately from *Cypripedium*. Due to these preliminary findings, *Paphiopedilum* could not be specified as an outgroup to *Cypripedium* with the monophyly of *Cypripedium* remaining intact. For this reason, topological constraints were specified under the search options in PAUP. The topology chosen for the outgroup was used as a constraint for the exhaustive searches of the ingroup. No topological constraints were placed on the taxa within *Cypripedium*, however.

RESULTS

Outgroup analysis

A single tree resulted from the outgroup analysis in which the topology of *Paphiopedilum* was used as a search constraint. Figure 7 shows the results of the analysis. A clade composed of *C. arietinum* and *C. reginae* as sister taxa was basal to

the entire genus. *Cypripedium acaule* and *C. macranthum* were sister taxa in a clade which was in a derived position relative to the *C. arietinum/ C. reginae* clade. The topology supported the monophyly of the *C. calceolus* complex, which was placed in a derived position in the genus. Basal to the complex and monophyletic with it was *C. guttatum*. Although the outgroup analysis resulted in a single most parsimonious topology, the results were abandoned due to the nature of the fragrance compound distributions, which resulted in no distinction between the out- and ingroup. Figure 8 shows the results of a single tree of 119 most parsimonious trees obtained when no outgroup taxa were specified.

Ingroup analysis

Analyses of the ingroup alone were performed to achieve local parsimony. No hypotheses of plesiomorphic and apomorphic charater states could be assumed, however. An exhaustive search was carried out in which all possible topologies were evaluated for most parsimonious reconstructions of the character states. Ten equally parsimonious trees were obtained, all of 45-step length. Figure 9 shows all 10 unrooted trees obtained. Several trends are clear from comparisons of all trees. *Cypripedium calceolus* var. *parviflorum*, *C. calceolus* var. *pubescens*, and *C. kentuckiense* are consistently associated with each other and are supported by seven synapomorphic gains. *Cypripedium reginae* and *C. arietinum* are found together in seven trees although only one synapomorphic gain supports the clade. However, the clade is supported by eight synapomorphic losses. *Cypripedium acaule* and *C. macranthum* are found as sister taxa in a clade in only three of the trees. The position of *C. guttatum* is unstable, because it is occupies a different position in six of the trees. All other trees differ in the relative positions of *C. candidum* and *C. macranthum*. Four of the topologies support the resolution of *C. candidum* with the rest of the members of the *C. calceolus* complex. Alternatively, three of the

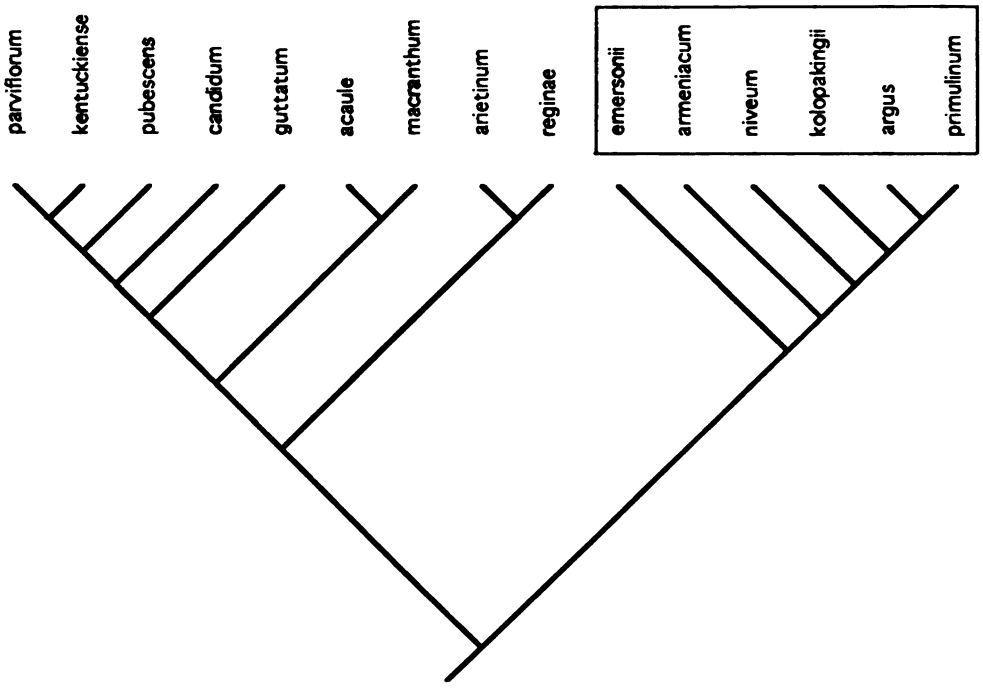


Figure 7. Single most parsimonious tree in which *Paphiopedilum* is held as an outgroup search constraint. (*Paphiopedilum* enclosed within box)

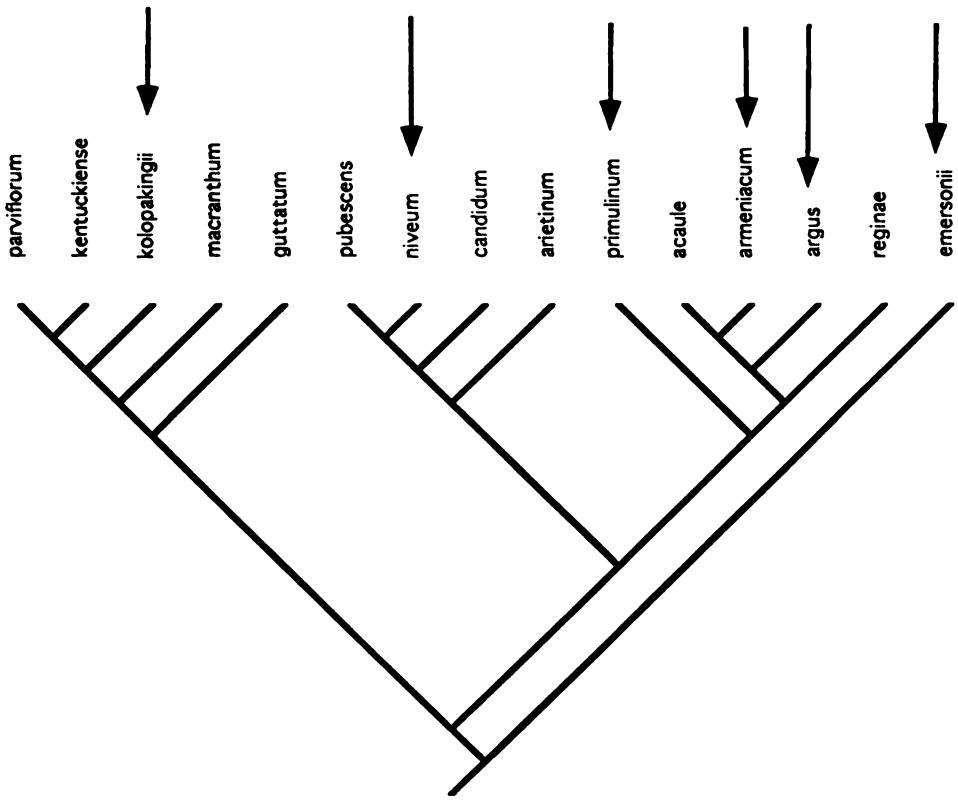


Figure 8. One of 119 most parsimonious reconstructions for *Cypripedium* and *Paphiopedilum* with no outgroup specified. (*Paphiopedilum* is recognized by arrows)

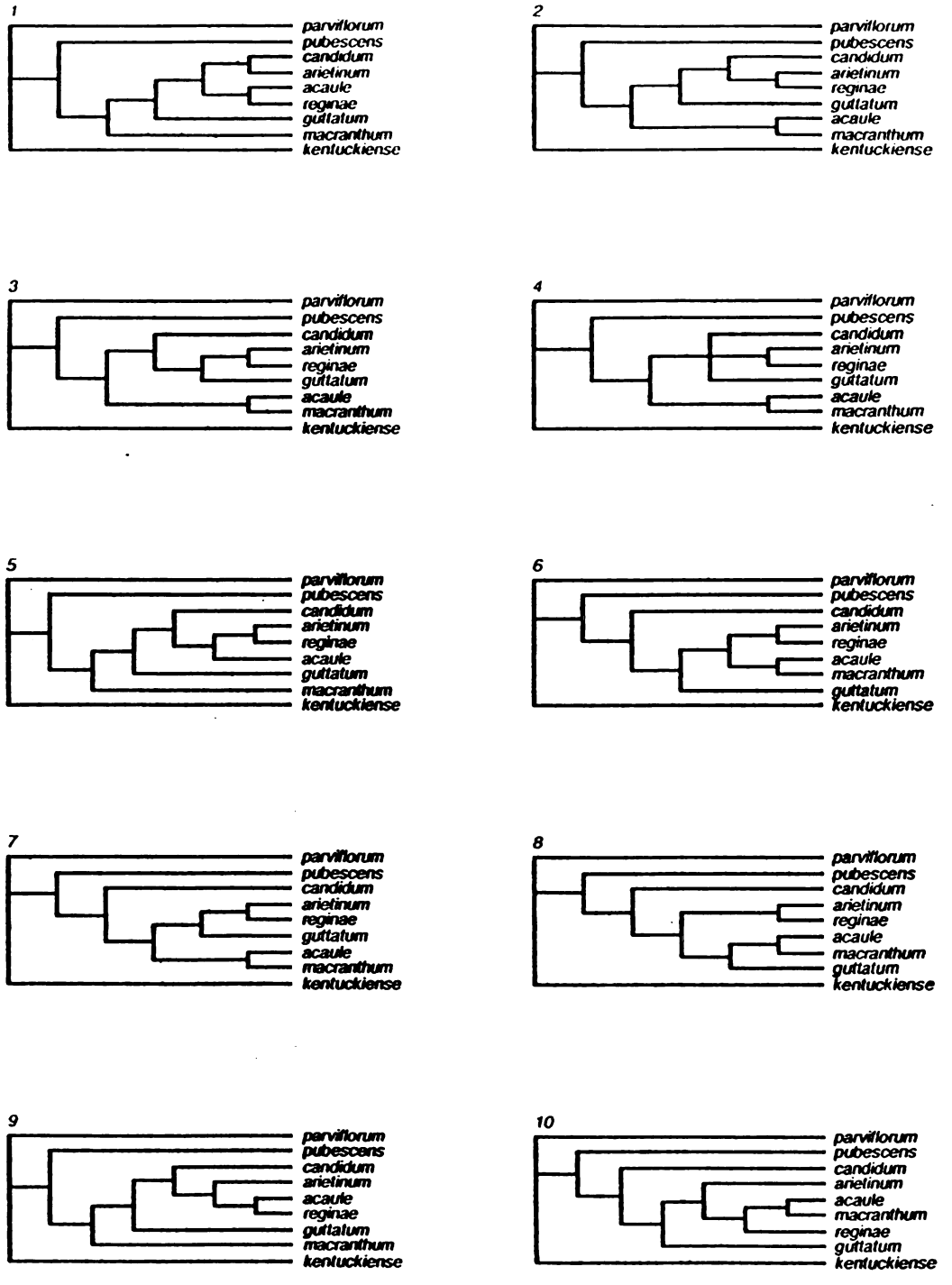


Figure 9. Ten equally parsimonious reconstructions for *Cypripedium*.

topologies support *C. macranthum* with the *C. calceolus* complex. The other three trees do not support any single taxon as sister to the *C. calceolus* complex.

Orientation of unrooted tree

Orientation of the unrooted consensus tree was done with the *C. calceolus* complex kept separate from other taxa. Figure 10 shows the majority rule consensus tree. This orientation follows the results presented in Chapter 1 for a clear separation of these two groups of taxa. Trees are oriented such that the *C. calceolus* complex is monophyletic. MacClade (Maddison & Maddison 1992) was used to trace character evolution on the consensus tree. The nature of the compound distributions is such that few synapomorphic gains have defined any specific lineages. In addition, the entire genus is not supported by any common synapomorphy.

Character state reconstructions and distribution of synapomorphies

Figures 11–19 show character state reconstructions for each character on the consensus topology which includes *C. candidum* within the *C. calceolus* complex. The legend accompanying each reconstruction lists states defined for each character. The enzymatic conversions to which each state corresponds can be found in figures 1–4, representing the proposed biogenetic pathways. The clade composed of *C. reginae* and *C. arietinum* are united by the lack of expression of all biogenetic pathways found in the genus excepting the benzoic acid pathway. The synapomorphic biogenetic pathway shared between them is not very strong as it is only the formation of a common precursor, benzoic acid.

The specific topology of the *C. calceolus* complex is clear in that *C. kentuckiense* and *C. parviflorum* are sister taxa, supported by a synapomorphy of alpha-farnesene and linalool. *Cypripedium calceolus* var. *pubescens* is sister to these two taxa. Together, the three taxa share synapomorphic pathways leading to the production

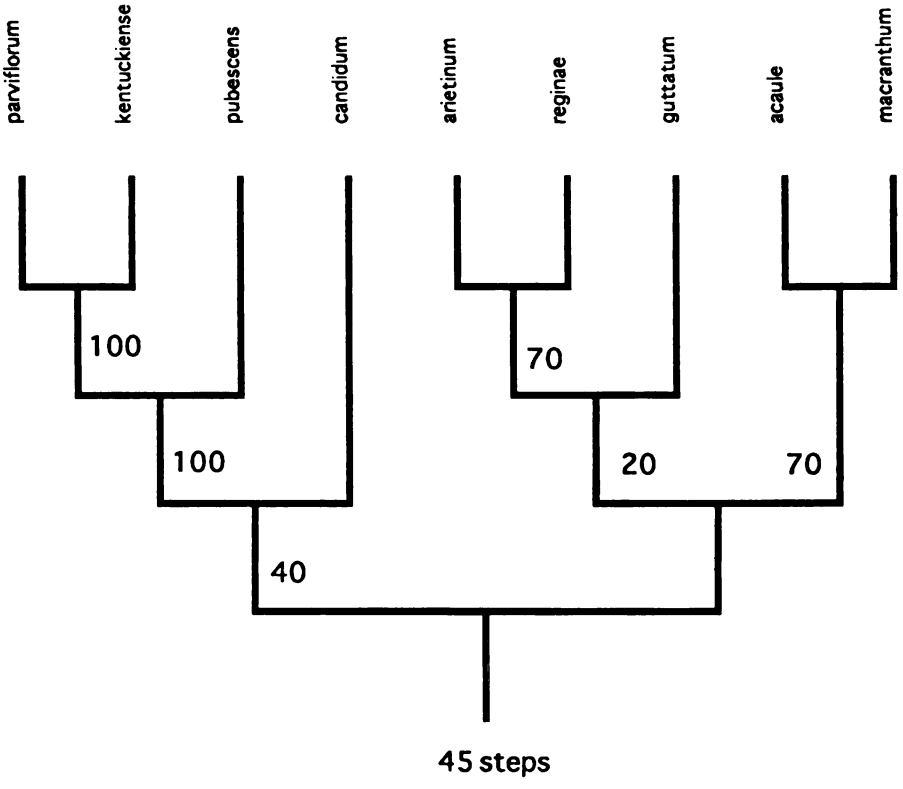


Figure 10. Consensus tree of ten equally parsimonious reconstructions for *Cypripedium*.

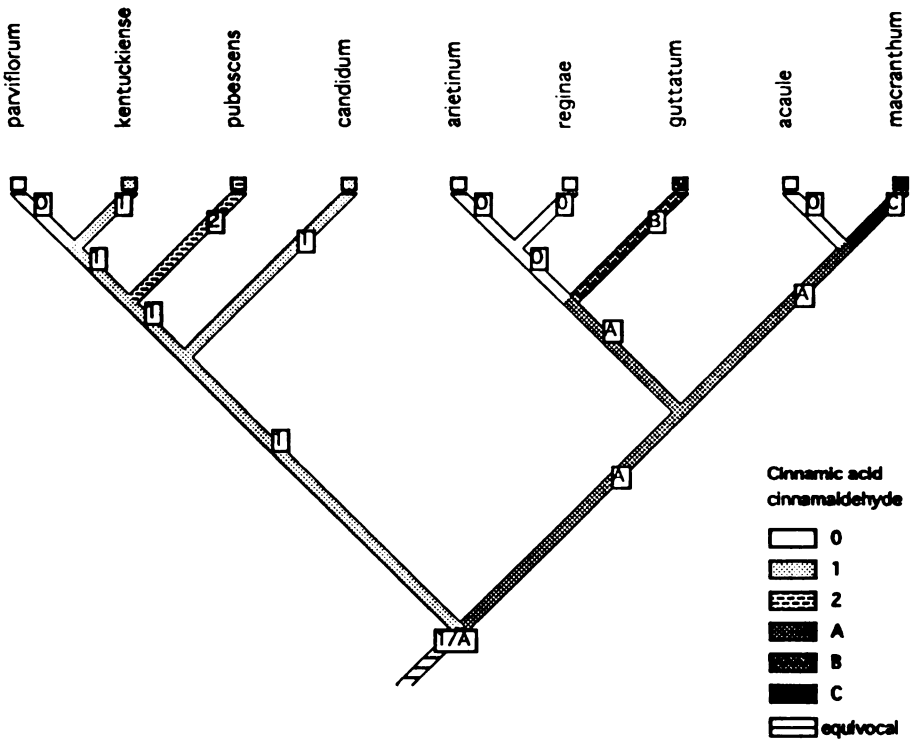


Figure 11. Character state changes for the cinnamic acid pathway.

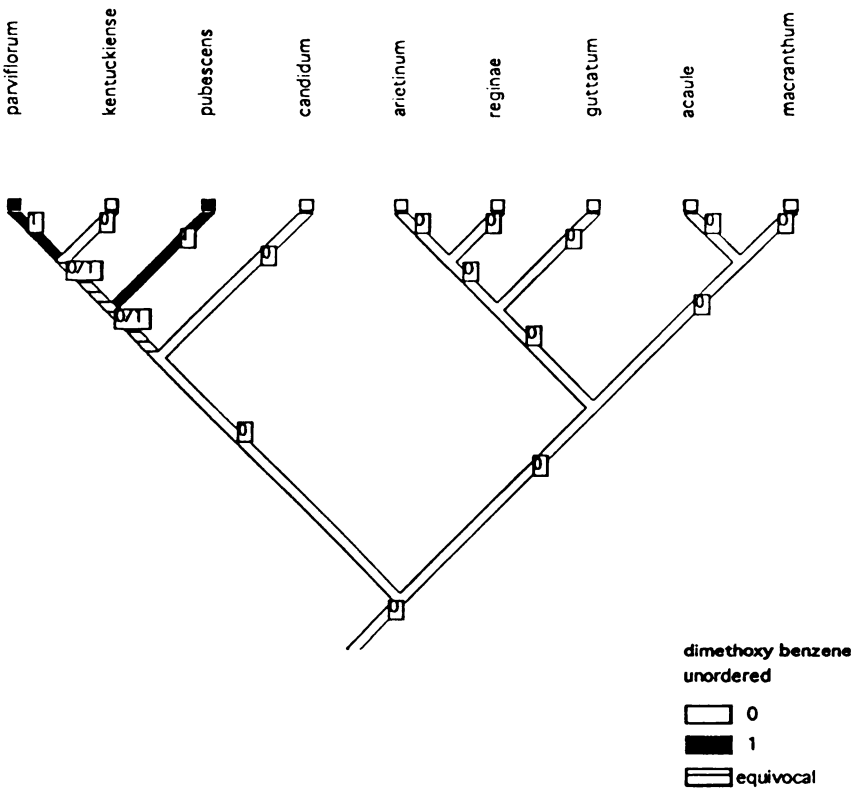


Figure 12. Character state changes for biogenesis of 1,4-dimethoxy benzene.

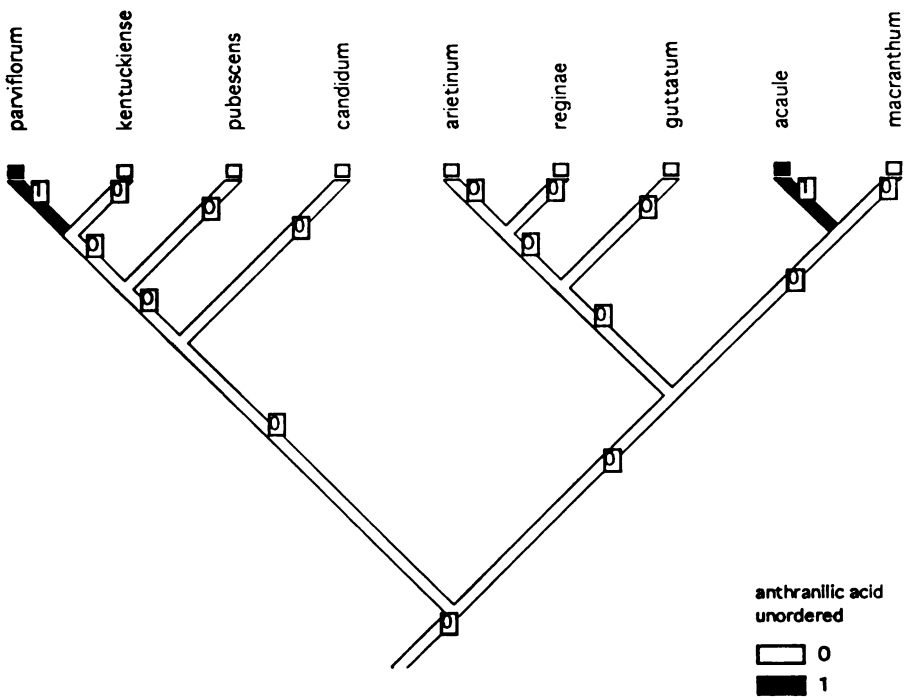


Figure 13. Character state changes for anthranilic acid pathway.

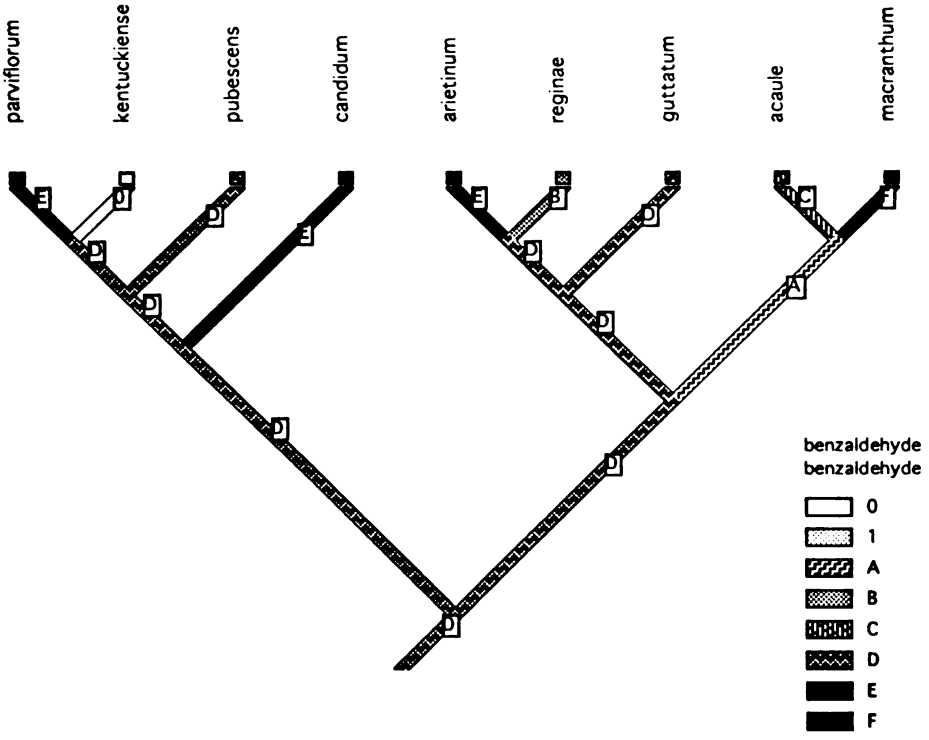


Figure 14. Character state changes for benzoic acid pathway.

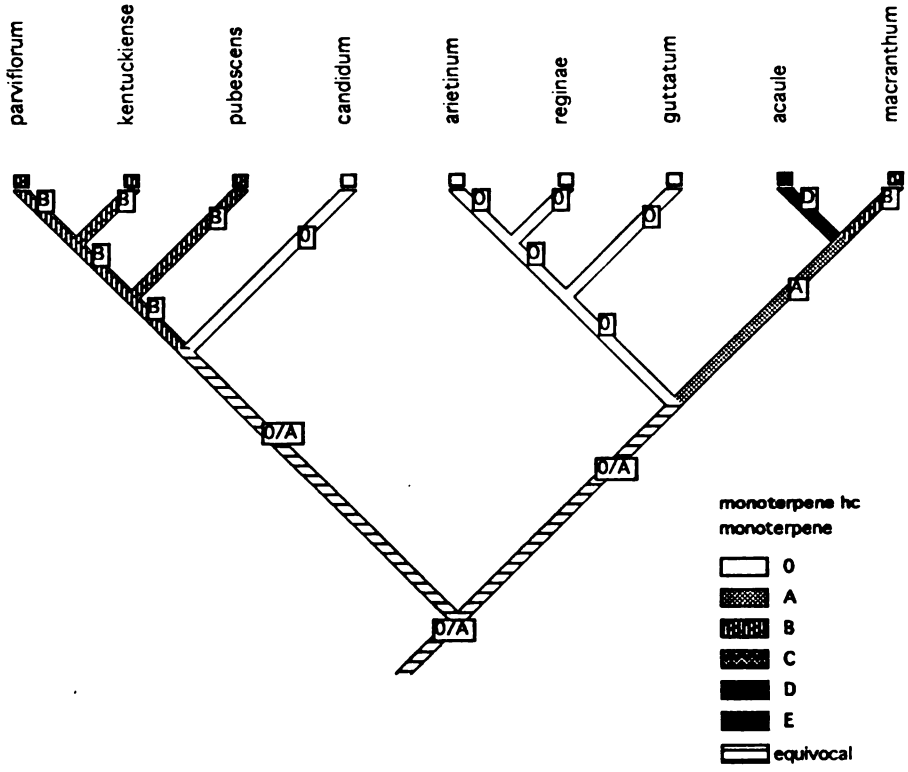


Figure 15. Character state changes for biogenesis of monterpene hydrocarbons.

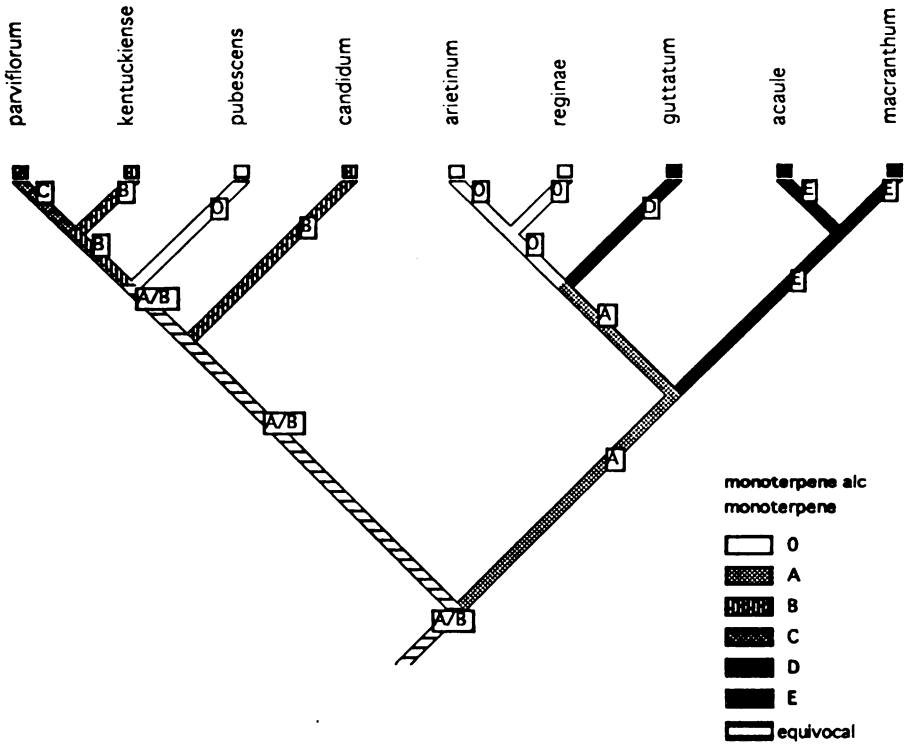


Figure 16. Character state changes for production of monoterpene alcohols.

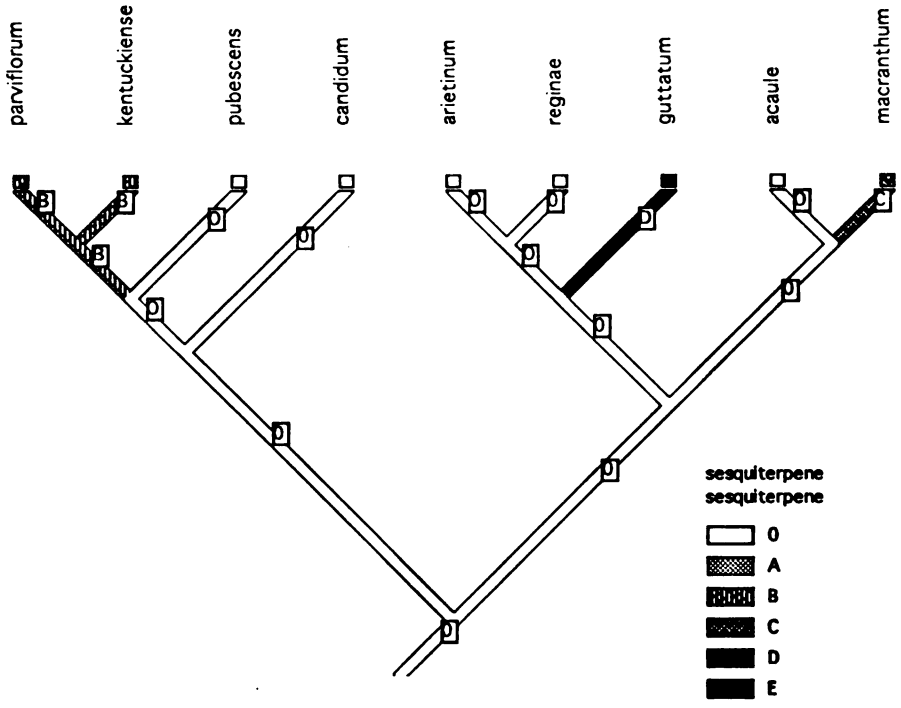


Figure 17. Character state changes for biogenesis of sesquiterpenes.

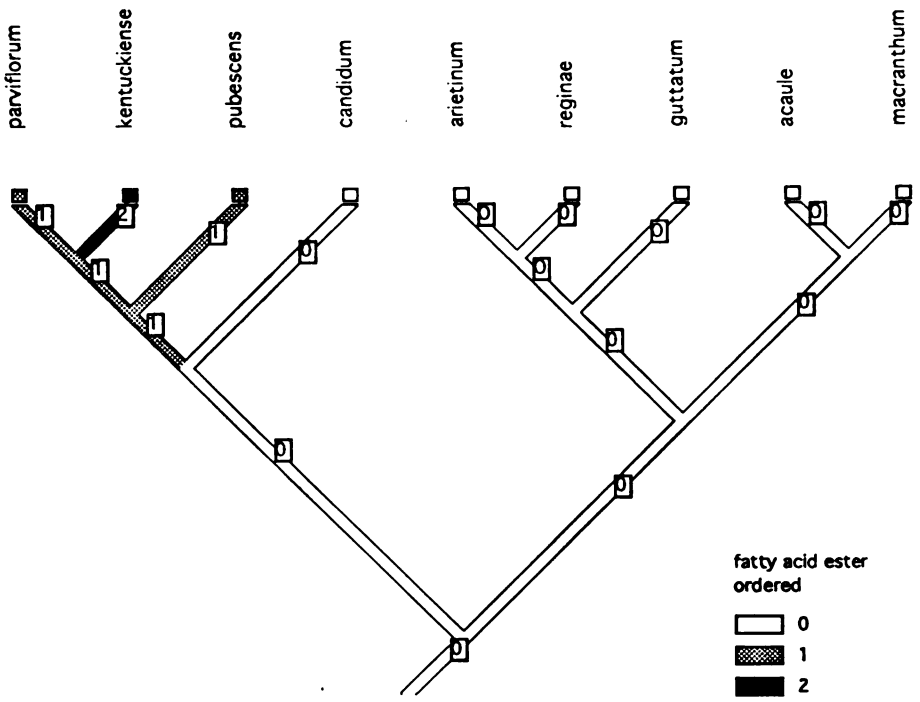


Figure 18. Character state changes for the biogenetic production of aliphatic alcohols and acetates.

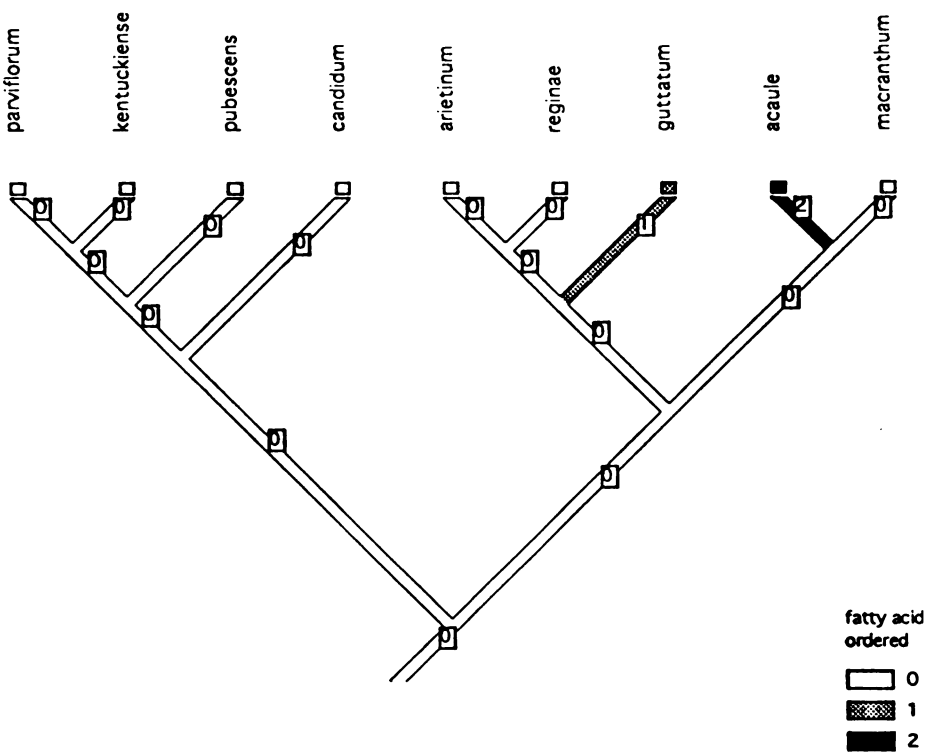


Figure 19. Character state changes for the biogenetic pathway leading to fatty acid ester production.

of cis-ocimene and the fatty acid alcohols and acetates. Strengthening the cladistic association of these taxa is the shared presence of the cinnamic acid pathway, leading to phenyl ethyl derivatives, between *C. calceolus* var. *pubescens* and *C. kentuckiense*. In addition, the shared presences of the biogenetic capability to form 1,4-dimethoxy benzene and benzaldehyde derivatives between *C. calceolus* varieties adds support to the lineage. The basal-most member of the complex, *C. candidum*, derives support for inclusion in the lineage by the shared ability to produce phenyl ethyl derivatives, benzaldehyde derivatives, and linalool.

Synapomorphic gains of biogenetic capability supporting the clade composed of *C. acaule* and *C. macranthum* include the formation geranyl pyrophosphate resulting in lilac alcohol and monoterpene hydrocarbons. In addition, the capability to produce the general precursor, benzoic acid, is shared as well. Only the single synapomorphic gain of the biogenetic capability to produce benzoic acid supports the clade in which *C. guttatum* is basal to *C. arietinum* and *C. reginae*.

Alternative topologies

The primary topological difference between the three general types of trees is that *C. candidum* is excluded from the *C. calceolus* clade in six of the ten trees. If *C. candidum* is included in the *C. calceolus* clade, rather than *C. macranthum*, more synapomorphies support the clade. The relatively higher number of synapomorphic biogenetic steps and the similarity of these taxa with respect to floral and vegetative morphology are the bases for choosing the topologies which include *C. candidum* within the *C. calceolus* clade. Three trees supported the inclusion of *C. macranthum* within the *C. calceolus* complex. Trees were imported into MacClade to evaluate the inclusion of both *C. macranthum* and *C. candidum* within the complex. Tree length was increased by a single step with this new topology (see Figure 20).

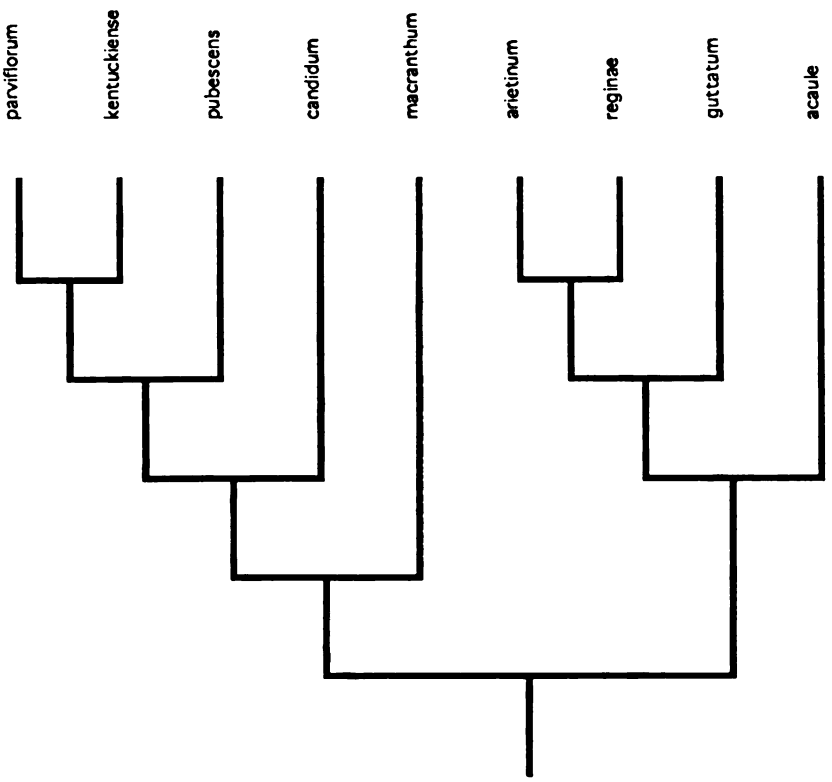


Figure 20. Tree in which *Cypripedium candidum* and *C. macranthum* are included within the *C. calceolus* complex (47 steps).

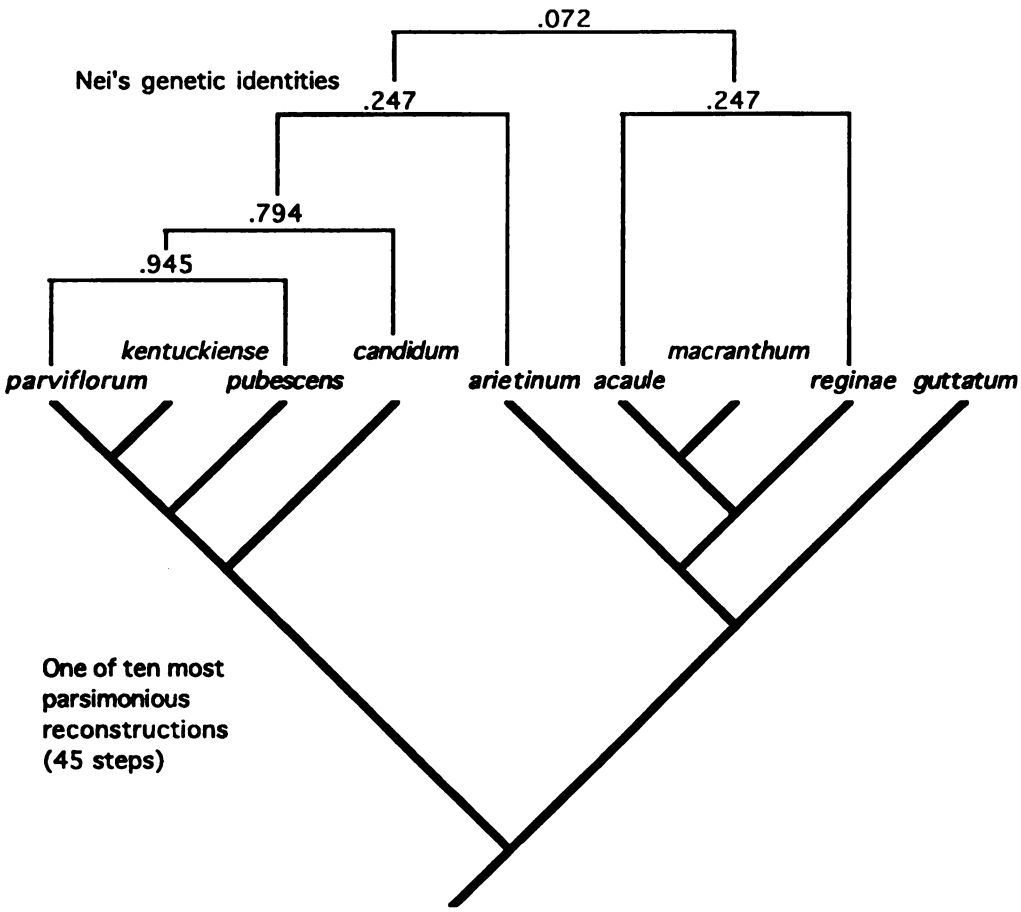


Figure 21. Comparison of genetic identities (Case 1993) and cladogram produced from floral fragrance data for *Cypripedium*.

Phylogenetic interpretations are derived from tree 10 due to a high degree of concordance with results obtained by UPGMA of Case (1993). Figure 21 shows a comparison of the two topologies.

DISCUSSION

Outgroup analysis

The results of the outgroup analysis must be carefully interpreted. Although a single most parsimonious topology was obtained, it may not accurately reflect phylogenetic relationships among taxa in *Cypripedium*. The topology of the outgroup had to be held as a search constraint due to the similarity of biogenetic pathways for many fragrance compounds between *Paphiopedilum* and *Cypripedium*. The initial search, in which no constraining topology was imparted on *Paphiopedilum*, resulted in species of the outgroup occupying sister taxon positions to five different species in the ingroup. The implications of this result are such that if each taxon in the outgroup were used independently as an outgroup, a different topology would result for the ingroup in each analysis. Due to the lack of consistency among the outgroup taxa for character states with respect to the ingroup, the implications of the analysis, a basal most position of *C. arietinum* and *C. reginae*, are not very strongly interpreted.

Phylogenetic Implications

Cypripedium calceolus/C. kentuckiense

The results of the parsimony analysis indicate several trends for evolutionary relationships within *Cypripedium*. First, the *C. calceolus* complex is monophyletic, including *C. kentuckiense, C. calceolus, C. candidum*, and possibly *C. macranthum*. *Cypripedium candidum* is basal within the lineage. This early evolutionary

divergence from the rest of the *C. calceolus* lineage is supported by the genetic identities obtained by Case (1993). Within *C. calceolus, C. calceolus* var. *parviflorum* and *C. calceolus* var. *pubescens* appear to be separate entities. In all trees *C. calceolus* var. *parviflorum* is sister to *C. kentuckiense.* Case (1993) proposed a potential progenitor-derived relationship for *C. kentuckiense* from *C. calceolus* var. *pubescens* on the basis of preliminary allozyme data. However, she does not recognize *C. calceolus* var. *pubescens* and *C. calceolus* var. *parviflorum* as distinct species. The derived sister positions for *C. calceolus* var. *parviflorum* and *C. kentuckiense* as compared to *C. calceolus* var. *pubescens* in all topologies obtained in this study necessitates the same treatment for both taxa. Either both are varieties of *C. calceolus* or both are distinct species. Although taxonomic judgement is reserved in this study, due to the use of a single data set, the topology for these two taxa is such that the same status needs to be applied to both.

Atwood (1984) recommended the recognition of three distinct species on the basis of relative size of the plants and floral features, although much overlap in quantitative characters occurs. Likewise, Bergstrom et al. recommended the recognition of *C. calceolus* var. *parviflorum* and *C. calceolus* var. *pubescens* as distinct species on the basis of differences in floral fragrance composition. One problem with the conclusion reached by Bergstrom et al. (1992) is that only one individual of each taxon was sampled (see Chapter 1). A consideration of the apparently variable nature of floral fragrance-compound data in *Cypripedium* (consider differences between Bergstrom et al. (1992), Holman (1983) and results reported in Chapter 1) and the high genetic identity shared by *C. calceolus* var. *pubescens* and *C. calceolus* var. *parviflorum* obtained by Case (1993), does not support the recognition of separate species. It follows, on the basis of this study, that *C. kentuckiense* should be considered for inclusion within *C. calceolus.*

Cypripedium candidum

The monophyly of the *C. calceolus* complex including *C. candidum* is not clearly resolved. The synapomorphies which link it to the complex or to certain taxa within it are the benzaldehyde pathway, the linalool pathway, and the expression of phenyl ethanol. Additionally, in considering the pathways which separate it from the complex, one is due to the lack of monoterpene hydrocarbon biogenesis in *C. candidum*. The expression of these compounds is reportedly variable within populations (Chapter 1), and Holman (1983) has indeed reported the presence of a monoterpene (probably terpinene) in *C. candidum*. Finally, the overwhelming morphological similarity between *C. candidum* and *C. calceolus* justifies its inclusion within the monophyletic lineage.

Cypripedium macranthum

There is some evidence to suggest the inclusion of *C. macranthum* within the monophyletic *C. calceolus* complex, excluding *C. candidum*. This taxon is not native to North America although it is cultivated by orchid growers. It does not share any striking morphological features with the members of the complex, although the labellum is of the same general construction in that the orifice is dorsal on the inflated pouch. Stoutamire (1967) reported that *C. calceolus* and *C. macranthum* may artificially hybridize. Case (pers. comm.) found *C. macranthum* to share several alleles in common with *C. calceolus*, although this evidence is based upon investigation of a single individual. In spite of these similarities, the shape of the lateral petals and the flower color is very different from any taxon within the complex. Although the inclusion of *C. macranthum* within the clade is suggested by three of the topologies, the primary synapomorphy supporting its inclusion is the monoterpene hydrocarbon ocimene. Other synapomorphies include biogenesis of general precursors shared with certain taxa within the complex, but no identical compounds. This is in contrast to *C. candidum* which shares as many

synapomorphic pathways, and in addition the fragrance compounds produced are identical to those produced by other taxa within the complex. The presence of lilac alcohol in *C. macranthum* may represent shared biogenetic expression with the complex, but the origin of the epoxide alcohol is unknown. The character states between these two taxa appear to be incompatible because the two are never found together with the *C. calceolus* complex. In fact, a topology which includes both of these taxa requires at least two extra steps to accomodate the character-state changes. While support is available for the monophyly of *C. macranthum* with the *C. calceolus* complex, it probably is not more closely related to *C. calceolus* than is *C. candidum.*

The alternative position of *C. macranthum* as sister to *C. acaule* is unlikely. Only one synapomorphy supports the clade, presence of lilac alcohol biogenesis. The two taxa are very different in floral morphology and the presumed relationship based upon fragrance biogenesis may represent a parallelism.

Cypripedium arietinum

The placement of *C. arietinum* with *C. reginae* is supported in seven of the trees. The segregation of *C. arietinum* into *Criosanthes* (Atwood 1984) is not supported by this analysis. The derived sister-taxon placement of *C. arietinum* and *C. reginae* was not expected, however. Little confidence is placed in this sister-taxon relationship due to the support for it deriving primarily from the synapomorphic absence of nearly all fragrance biogenetic expression. Although this topology was supported by seven of the trees, it is not supported by tree 10 which is most congruent with the results of Case (1993). The possibility that the most common reconstructed placement of these taxa is not correct, illustrates the importance of discussing alternative most-parsimonious topologies, not simply relying upon the consensus tree.

Other species

Phylogenetic interpretations for the rest of the morphologically diverse taxa in the genus, based upon the topology most congruent with the results of Case (1993), include a basal position for *C. guttatum* to the rest of the species. *Cypripedium arietinum* is the next taxon to diverge from these other species. The three most derived species within the genus are *C. reginae*, *C. acaule* and *C. macranthum*. The advanced position of *C. macranthum* among these taxa may be artificial due to the possibility of its inclusion within the *C. calceolus* complex. Albert (1993) reports phylogenetic relationships within the genus based upon rbcL variation. He did not sample the *C. calceolus* complex, *C. acaule* or *C. guttatum*. Of the taxa sampled, *C. plectrochilon* (putatively closely related to *C. arietinum*) was basal in the genus, with *C. reginae* more derived (Albert pers. comm.). The topology of tree 10 supports this proposed evolutionary history.

Orienting the tree without an outgroup

The lack of resolution for the ingroup using the specified outgroup leaves the rooting of the tree, based solely on the data at hand, a difficult task. The ad hoc assumption that substituted skeletons represent enzymatic complexity and perhaps an evolutionarily derived state could justify the rooting of the trees with taxa which do not contain substituted compounds. No rooting could be performed in this manner because all taxa appear to possess fragrance compounds which are enzymatically advanced as well as some which are enzymatically simple. Another ad hoc assumption might be that loss of biogenetic complexity is an advanced state (Seaman & Funk 1983). The results of this analysis would indicate that *C. arietinum* and *C. reginae* are the most advanced within the genus using this criterion. This alternative is unlikely given the results of Albert (1993) in which *C. arietinum* and *C. reginae* occupy basal and derived positions respectively.

The distribution of character states upon the branches of the trees of Figures 11–19 illustrates two points. First, the most parsimonious reconstruction of the character states represents the simplest distribution of enzymatic pathways exhibited by these taxa for total fragrance biogenesis. Steps on the cladogram represent the expression of an entire, in the case of binary coded characters, or part of a biogenetic pathway, in the case of multistate characters. Second, the topology chosen to represent the distribution of biogenetic pathways within *Cypripedium* suggests that parts of, or entire pathways can be gained or lost fairly easily within and between lineages.

The initial outgroup analyses using *Paphiopedilum* with *Cypripedium* yielded important information for the use of fragrance data in reconstructing evolutionary history. A limitation to the use of floral fragrance-compound data for evolutionary reconstruction is in the widespread nature of the compounds. The biogenetic pathways for monoterpene hydrocarbon and alcohols are found in many groups of unrelated plants. Within the Orchidaceae only four of 154 taxa sampled in 60 genera by Kaiser (1993) did not produce either of these two classes of volatiles. Additionally, approximately 30 out of 150 taxa did not produce the cinnamic acid derivatives and 15 out of 150 did not produce benzoic acid derivatives. Because of the ubiquitous distribution of these compounds between genera, it is not surprising that the data are of little importance as phylogenetic markers above the level of species. In contrast, the distribution of fatty acid derivatives such as straight chain alcohols and acetates, greater than 10 carbons in length, are very restricted. Only 35 of 150 species sampled produced compounds of these types. The relative distributions of these compounds between taxa are predicted to be more reliable indicators of phylogenetic relationships.

CONCLUSIONS

The use of floral fragrance-compound data appears to be useful in phylogenetic reconstruction. Effective representation of the biogenetic pathways which result in the diversity of fragrance compounds found in *Cypripedium* can be achieved using the user-defined characters of PAUP. This is a better method for representing the data than was the coding procedure utilized for the phenetic analysis of Chapter 1. The relative advantage lies in the ability to express biogenetic dependency between character states in the cladistic analysis. Additionally, the distribution of character states on a most parsimonious reconstruction are easily interpreted as parts of biogenetic pathways in which each step represents a biochemical conversion. The phylogenetic interpretations of the cladistic analysis are highly congruent with those obtained by other recent workers. Results include the monophyly of the *C. calceolus* complex, with the possible inclusion of *C. macranthum*, a relatively basal position for *C. arietinum* and its inclusion within the genus, and derived positions for *C. acaule* and *C. reginae*. The position of *C. guttatum* is unresolved using this data set. The primary shortcoming of the cladistic analysis using floral fragrance data is that outgroup analysis is not possible because of the widespread nature of the biogenetic pathways to produce these compounds in many flowering plants. While conservative phylogenetic interpretations have been made using this single data set, broader conclusions may be possible from an analysis involving many independently derived data sets, including information from morphology, anatomy, cytology, and DNA sequences.

BIBLIOGRAPHY

BIBLIOGRAPHY

Albert, V. A. & Chase, M. W. 1992. *Mexipedium*: A new genus of slipper orchid. Lindleyana 7(3): 172-176.

Atwood, J. T. 1984. The relationships of the slipper orchids (subfamily Cypripedioideae, Orchidaceae). Selbyana 7: 129-247.

Atwood, J. T. 1984. The *Cypripedium calceolus* L. complex in North America. In Proceedings: The 11th. world orchid conference. Ed. Kiat Tan, International Press Co. Singapore.

Bergstrom, G., Birgersson, G., Groth, I. & Nilsson, L.A. 1992. Floral fragrance disparity between three taxa of lady's slipper *Cypripedium calceolus*. Phytochemistry 31(7): 2315-2319.

Borg-Karlson, A.-K. 1990. Chemical and ethological studies of pollination in the genus *Ophrys* (Orchidaceae). Phytochemistry 29(5): 1359-1387.

Case, F.W. 1987. Orchids of the Western Great Lakes Region. Cranbrook Institute of Science, MI.

Case, M. A. 1993. Genetic diversity and evolutionary relationships within and among five *Cypripedium* species: Inferences from allozyme electrophoresis. Ph.D. thesis, Michigan State University, East Lansing, MI.

Chen, J. & Meeuse, J. D. 1971. Production of free indole by some aroids. Acta Bot. Neerl. 20(6): 627-635.

Croteau R. & Karp, F. 1991. Origin of natural odorants. *In Perfumes; Art, Science, and Technology*. Ed. P. M. Muller and D. Lamparsky. Elsevier Applied Science. NY.

Duncan, T. & Baum, B. R. 1981. Numerical phenetics: Its uses in botanical systematics. Ann. Rev. Ecol. Syst. 12: 387-404.

Erman, W. F. 1985. Chemistry of the monoterpenes; an encyclopedic handbook. Part A. In Studies in Organic Chemistry Ed. Paul G. Gassman. Marcel Dekker, Inc. NY, NY.

Gregg, K. B. 1983. Variation in floral fragrances and morphology: Incipient speciation in *Cycnoches*? Bot. Gaz. 144(4): 566-576.

Holman, R. T. 1983. Studies on the fragrances of North American terrestrial orchids. In E. H. Plaxton [ed.], North American Terrestrial Orchids: Symposium II Proceedings and Lectures, Michigan Orchid Society, Livonia.

Holman, R. T. & Heimermann, W. H. 1976. First symposium on the scientific aspects of orchids. eds. H. H. Szmant and J. Wemple, University of Detroit Publ. pp. 75-89.

Kaiser, R. 1993. The Scent of Orchids. Amsterdam, Elsevier.

Klier, K., Leoschke, M. J., & Wendel, J. F. 1991. Hybridization and introgression in white and yellow ladyslipper orchids (*Cypripedium candidum* and *C. pubescens*). J. Heredity 82: 305-318.

Maddison, W. P. and Maddison, D. R. 1992. MacClade: Analysis of phylogeny and character evolution. Version 3.0. Sinauer Associates, Sunderland, Massachusetts.

Meeuse, B. & Morris, S. 1984. The Sex Life of Flowers. Oxford Scientific Films Ltd. New York, NY.

Muller, P. M. & Lamparsky, D. 1991. Perfumes: Art Science and Technology. Elsevier, New York, NY.

Newhouse, C. J. 1976. Pollination biology in seven taxa of Michigan Orchidaceae and a study of *Cypripedium calceolus* in Michigan based on living plants and herbarium specimens. M. S. Thesis, Michigan State University.

Nilsson, L.A. 1979. Anthecological studies on the Lady's slipper, *Cypripedium calceolus*. Bot. Not. 132: 329-347.

Rodman, J. E. 1991. A taxonomic analysis of glucosinolate-producing plants, part 1: Phenetics. Syst. Bot. 16(4): 598-619.

Rohlf, F. J. 1992. Numerical Taxonomy and Multivariate Analysis System. Version 1.70. Exeter Software. New York.

Salatino, A. & Gottlieb, O.R. 1981. Quinolizidine alkaloids as systematic markers of the Genisteae. Biochem. System. Ecol. 9: 267-273.

Schreier, P. 1984. Chromatographic Studies of Biogenesis of Plant Volatiles. Dr. Alfred Huthig Verlag, GmbH Heidelberg.

Schreier P. 1986. Biogeneration of plant aromas. *in* Developments in Food Flavours Ed. Birch G. G. and Lindley M.G. Elsevier Applied Science. NY.

Seaman, F. C. & Funk, V. A. 1983. Cladistic analysis of complex natural products: Developing transformation series from sesquiterpene lactone data. Taxon 32(1): 1-27.

Sneath, P. H. & Sokal, R. R. 1973. Principles of Numerical Taxonomy. San Francisco, W. H. Freeman and Co.

Stoutamire, W. P. 1967. Flower biology of the lady's-slippers. Michigan Bot. 6(1): 159-175.

Stuessy, T. F. 1990. Plant Taxonomy: The Systematic Evaluation of Comparative Data. Columbia University Press. New York.

Swofford, D. L. 1991. PAUP: Phylogenetic analysis using parsimony, Version 3.0s. Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.

Whitten, W. M. & Williams, N. H. 1992. Floral fragrances of *Stanhopea.* Lindleyana 7(3): 130-153.


MICHIGAN STATE UNIV. LIBRARIES
31293010190951