TAXON 32(1): 1–27. FEBRUARY 1983

CLADISTIC ANALYSIS OF COMPLEX NATURAL PRODUCTS: DEVELOPING TRANSFORMATION SERIES FROM SESQUITERPENE LACTONE DATA

Fred C. Seaman[1] *and V. A. Funk*[2]

Summary

In the Asteraceae, the two most taxonomically significant aspects of the sesquiterpene lactone chemistry are the patterns of skeletal and substitutional diversity. Sesquiterpene lactones can be subdivided into different classes based on their carbon skeletons. Superimposed on each skeleton is a set of substituents which collectively define a specific compound. Both skeletal classes and substitutional features display taxonomically useful patterns of distribution.

A cladistic technique is demonstrated for converting these chemical features into taxonomic characters. Each character is derived via a biosynthetically based evaluation of structural diversity and determination of structural homology. Each set of homologous skeletal and substitutional characters from a taxon is arranged in a biogenetic transformation series. Outgroup comparison is used to determine the polarity of each character transformation series. The distribution of the resulting novel (apomorphic) skeletal and substitutional characters within this taxon are then used to generate a cladogram depicting the chemical synapomorphies. The chemistries of the taxa, *Tetragonotheca* L. and *Iva* L. (Asteraceae: Heliantheae), provide the two examples for the four-step cladistic analysis. The methods employed in this analysis of sesquiterpene lactone distribution may also be applicable to other classes of complex natural products.

Introduction

Complex plant natural products represent a biosynthetically diverse array of compounds which are characterized by skeletal, substitutional and stereochemical variability. Several major classes of natural products included within this group are sesquiterpenes, diterpenes, triterpenes (saponins, cardenolides, etc.), alkaloids (monoterpene-indole alkaloids, aporphine alkaloids, etc.) and isoflavonoid derivatives (pterocarpans). These secondary metabolites express numerous structural novelties, a feature which constitutes a significant resource for generating new taxonomic characters. However, the conversion of this structural variability into discrete taxonomic characters is a difficult task, one requiring a uniform methodology for transforming chemical features into characters. Further, if such characters are to be used in a phylogenetic systematic study, they must be organized into series of homologues, ranked according to degree of novelty, and surveyed for within groups of taxa.

A method for converting raw chemical data into taxonomic characters that incorporates the basic concepts of cladistics is outlined using the sesquiterpene lactones of the Asteraceae. Each character is derived via a biosynthetically based evaluation of structural diversity and determination of structural homology. Homologous characters are then grouped into transformation series and used to generate cladograms based on the distribution of novel or unique homologues.

Sesquiterpene Lactones: Taxonomic Application

Several angiosperm families produce simple germacrane-derived sesquiterpene lactones, but only the Asteraceae is characterized by an extraordinary array of structurally modified and highly substituted compounds. Over 1300 complex structures have been isolated from

[1] Charles B. Harding Laboratories, New York Botanical Garden, Bronx, NY 10458, U.S.A.
[2] Department of Botany, Smithsonian Institution, Washington, DC 20560, U.S.A.

a fraction of the family's total number of species. In order to convert this rapidly growing mass of raw chemical data into taxonomic characters an appropriate methodology must be developed.

Two observations which should influence the choice of methodology relate to (1) sesquiterpene lactone qualitative infraspecific variation and (2) sesquiterpene lactone structural complexity.

Populational sampling of sesquiterpene lactones commonly reveals qualitative infraspecific variation, the extent of which depends on the complexity of the taxon. Thus, simple species-specific chemical complements rarely define taxa at the species level. Although variable, a species' chemistry contains compounds which display a high degree of structural relatedness, reflecting activity in only a limited part of the total system of sesquiterpene lactone biosynthetic pathways. Defining the taxon in terms of this characteristic biogenetic potential is not hindered by the infraspecific variation. In fact, the more details made available about the extent of variability, the more precisely the limits of the taxon's biogenetic capacity can be defined. The breadth of this biogenetic capacity often reflects the complexity of the taxon.

Sesquiterpene lactones display considerable complexity in their carbon skeletons (Fig. 1) and in the types, positions and stereochemistries of the attached substituents. The highly modified structures of many sesquiterpene lactones contain much taxonomically useful information. This complexity suggests that assigning a presence/absence-type character to each compound in a set of related structures is an underutilization of the available structural information. For example, one taxon may produce a complex structure, which with only one minor substitutional exception, is identical to a compound found in a second taxon. If characters are defined on the basis of identity or non-identity, these two compounds must be scored as non-identical even though their biosynthetic relatedness may indicate an alliance between the two taxa. More suitable characters are needed which will communicate degree of structural difference.

If we are to discard these more narrative approaches to chemical character analysis based on species-specific complements and the presence/absence-type scoring of sets of compounds, an alternative method must incorporate some biogenetic basis for character definition. Techniques developed for the analysis of flavonoids (Levy, 1977; Aparecida et al., 1977), neolignans, pyrones (Gottlieb and Kubitzki, 1981) and alkaloids (Gomes and Gottlieb, 1980; Salatino and Gottlieb, 1981) stressed the importance of a biogenetic basis for structural interpretation. In this approach, the distribution of a set of homologous compounds in a group of taxa is analyzed according to the hypothetical biosynthetic relationships of the compounds. Compounds which share common routes of biogenesis frequently co-occur in the same taxon or related taxa and in many cases are restricted to these taxa.

The homology of such biogenetically based characters must be satisfactorily demonstrated. For these characters, homology is defined by the sharing of a biosynthetic pathway. Hypothetical biosynthetic pathways (biogenetic routes) have been proposed for most classes of natural products. Each pathway consists of a sequence of reactions linking homologous compounds to a common precursor. Because such schemes are hardly ever documented for a given set of taxa, how is the hypothesis posed by the reaction sequence tested? The most convincing but rarest evidence supporting such pathways is the in vivo labelling and in vitro enzymological experimental proof of the proposed pathway. Unfortunately few data of this nature are available for sesquiterpene biogenesis.

The second most convincing line of evidence is the laboratory synthesis of a structure from another compound resembling the hypothetical precursor. For example, Fig. 2 illustrates two successful laboratory syntheses which convert different germacranolide precursors into products representing the two major subclasses (Herz, 1977a; Fischer et al., 1979) of the eudesmanolide sesquiterpene lactone skeletal class (Fig. 1, **2**). Presumably, these two syntheses are biomimetic of the biogenetic routes (routes A_1 and A_2, Fig. 2) proposed for

the two sets of eudesmanolide structures. The bulk of experimental evidence supporting sesquiterpene lactone biogenesis is of this type.

A third line of evidence recognized by early sesquiterpene lactone chemists was best described by Ted Geissman (1973):

> The organic chemist, faced with a given structure, can easily derive a reaction sequence by which it might have been derived from a postulated precursor. The probability that the hypothesis represents the reality can be greatly increased if a series of compounds, all occurring together in a single organism, can be shown to be constituent parts of the hypothetical pathway.

Regarding the eudesmanolide biomimetic synthesis example given above, distribution of eudesmanolides in the subtribe Ambrosiinae (discussed in a subsequent section) suggests that the two modes of biogenesis exist and are mutually exclusive. All eudesmanolide-producing taxa but one, *Ambrosia confertiflora* DC., were reported to produce one subclass or the other. Although *Ambrosia confertiflora* was originally reported to produce both types, Rodrigues et al. (1978) recently demonstrated that the route A_1 constituents reported from *A. confertiflora* were probably chromatographic artifacts of a germacrolide 1,10 epoxide. Thus, these reports of route A_1 constituents from *A. confertiflora,* and possibly other taxa, can be translated into reports of novel germacrolide 1,10 epoxides. However, in cases such as *Iva* where the β-1-hydroxy-eudesmanolide is further modified to yield the seco-derivative (Fig. 10) the presence of these derivatives contradicts an artifactual origin. In this case, distribution of eudesmanolides supports the postulated existence of two modes of eudesmanolide biogenesis.

All proposed pathways (biogenetic schemes) must be mechanistically sound, conforming to standard reaction sequences found in other better-established pathways. Also, the hypothetical pathways should include reactions which are energetically compatible with living systems (i.e., no extraordinarily high activation energies).

Any biogenetically based methodology for handling chemical characters requires a flow-chart-like biogenetic scheme showing the sequence of hypothetical precursors, intermediates and products, separated by discrete reaction steps. As the number of reaction steps leading from the precursor increases, there is a corresponding increase in the structural divergence of the product from the biosynthetic precursor. This divergence can take the form of either increased or decreased structural complexity. As shown in Fig. 1, sesquiterpene lactone skeletons correspond to related precursor-, intermediate- and product-like elements situated along one of several biogenetic routes.

The skeletal types of major interest in the Asteraceae (Fig. 1A) are derived directly from a germacrane precursor through a variety of cyclizations, ring fissions, rearrangements and methyl migrations. The lactones derived via an eremophilene intermediate (Fig. 1B) are mostly restricted to the Senecioneae, those derived via an isocedrene intermediate (Fig. 1C) are found in the Mutisieae (skeleton 20, Fig. 1C) and those produced from a cadinane intermediate are found in the Anthemideae (skeletons 23 and 26, Fig. 1C) (Bohlmann and Zdero, 1979a; Seaman, 1982).

Werner Herz (1977a, 1977b) first recommended a useful conceptual framework for organizing skeletal types within this biogenetic scheme: skeletons which are derivable from the farnesyl pyrophosphate hypothetical precursor by the same number of modifications of the carbon skeleton are grouped in one of four columns and assigned the same level of biogenetic complexity. As stressed by Herz, such grouping of skeletons into columns of increasing biogenetic complexity are, in part, superficial and thus should be interpreted with caution. Generally, the lower complexity level skeletons have broader distributions within the family than the higher complexity level skeletons. Superimposed on these skeletons are a myriad of functionalities.

A survey of past taxonomic use yielded a ranking of chemical features according to dominant but not exclusive use (analytic or synthetic, Davis and Heywood, 1963) and

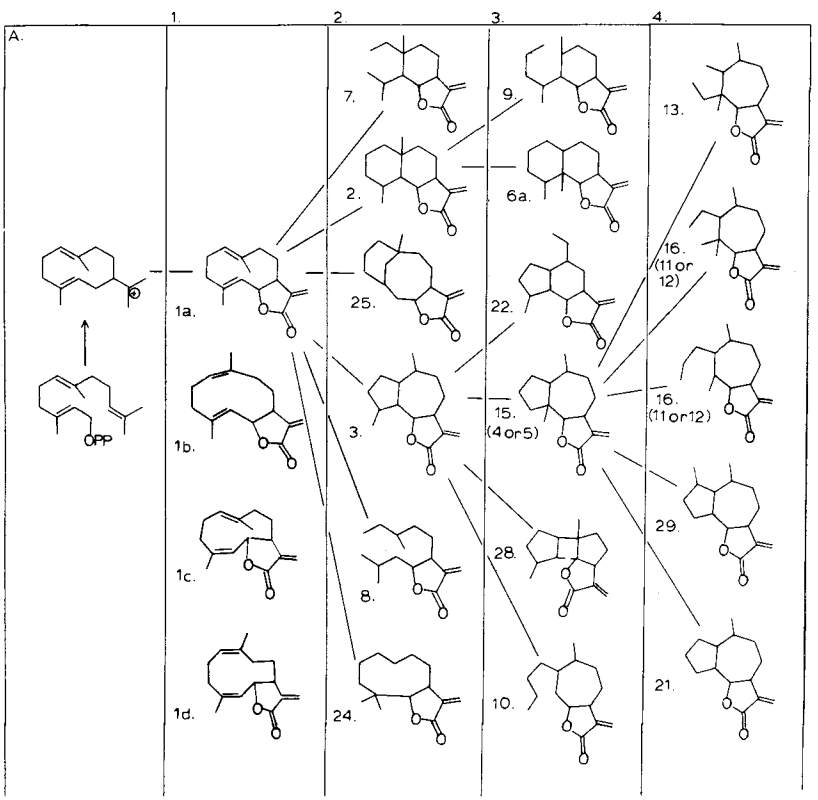


Fig. 1. Hypothetical biogenesis of the major sesquiterpene lactone skeletal types of the Asteraceae. A, Germacrane-derived skeletal types (all structures shown as C-6 lactonized). B, Eremophilane-derived skeletal types (all lactones shown as C-8 lactonized). C, Isocedrene and cadinane-derived skeletal types. Sesquiterpene lactone and furane skeletal code: 1a, Germacrolide. 1b, Melampolide. 1c, Heliangolide. 1d, *cis, cis*-Germacradienolide. 2, Eudesmanolide. 3, Guaianolide. 4, Ambrosanolide. 5, Helenanolide. 6, Eremophilanolide. 7, Elemanolide. 8, Secogermacranolide. 9, Secoeudesmanolide. 10, Xanthanolide. 11, Secoambrosanolide. 12, Secohelenanolide. 13, Norpsilotropin. 14, Bakkenolide. 15, Pseudoguaianolide. 16, Secopseudoguaianolide. 18, Furanoeremophilane. 19, Aromatic Furanoeremophilane. 20, Trixikingolide (Isocedrene-derivative). 21, Nor-helenanolide. 22, Chrymoranolide. 23, Cadinanolide. 24, Trichosalviolide. 25, Disyhamifolide. 26, "Quing Hau Sau." 27, Secofuranoeremophilane. 28, Bourbonolide. 29, Neohelenanolide. 30, Farnesyl derivative lactone. 31, Secoeremophilanolide. 32, Rearranged furanoeremophilane. 33, Furanoeudesmane. 34, Aromatic eremophilanolide.

hierarchical level of application (Fig. 3). Skeletal characters 1 and 2 (Fig. 3) are especially useful as synthetic characters at the higher taxonomic levels (tribal, subtribal). Distribution of shared substitution patterns (especially those including novel functional groups) are also useful in clustering taxa. Selected substitutional characters, when combined with skeletal characters, often permit the grouping of taxa at lower levels than possible when only skeletal features are considered. The presence of a biogenetically advanced compound (character level 4) can be used as a synthetic character to group taxa. The taxon-specific sesquiterpene lactone complement (chemotype; level 5) is an analytic character which can be useful in establishing subspecific or varietal boundaries.

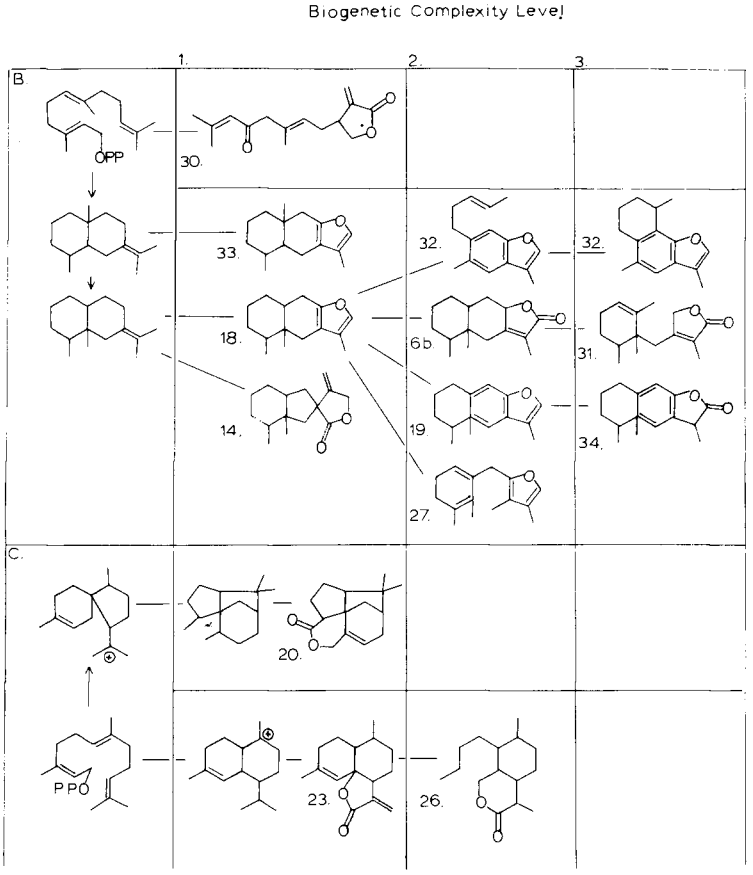


Fig. 1. Continued.

These earlier, mostly narrative, interpretations of sesquiterpene lactone variation demonstrate that variation in skeletal features can be treated initially as a taxonomically significant resource distinct from the substitutional patterns superimposed on the skeleton. Next, the substitutional variation (including stereochemistry of functional-group ring attachment) can be evaluated. The two sets of data can then be jointly considered in determining phylogenetic relationships.

Cladistic Analysis

Because cladistic methodology has been described many times in the recent literature (e.g., Eldredge and Cracraft, 1980; Nelson and Platnick, 1981; Wiley, 1981) we have not included a detailed discussion of the theory. We do, however, provide a brief description to help orient those who are unfamiliar with cladistics.

Cladistics seeks to find natural order. This order is reflected in nature by the presence of groups within groups. The groups are defined by inter-nested sets of novel characters (apomorphies). For a very basic example, given any three taxa, A, B, and C, A and B would be more closely related to one another than either is to C if A and B share a novel character (synapomorphy; Fig. 4). A and B are then considered to be sister-taxa (one is the sister taxon of the other) and the group A-B is the sister taxon or sister group of C. In addition, any group of taxa delimited by a novel character can be referred to as monophyletic, or as a monophyletic group. The importance of the monophyletic group cannot be overemphasized as this indicates a common history for the group. Members of non-monophyletic

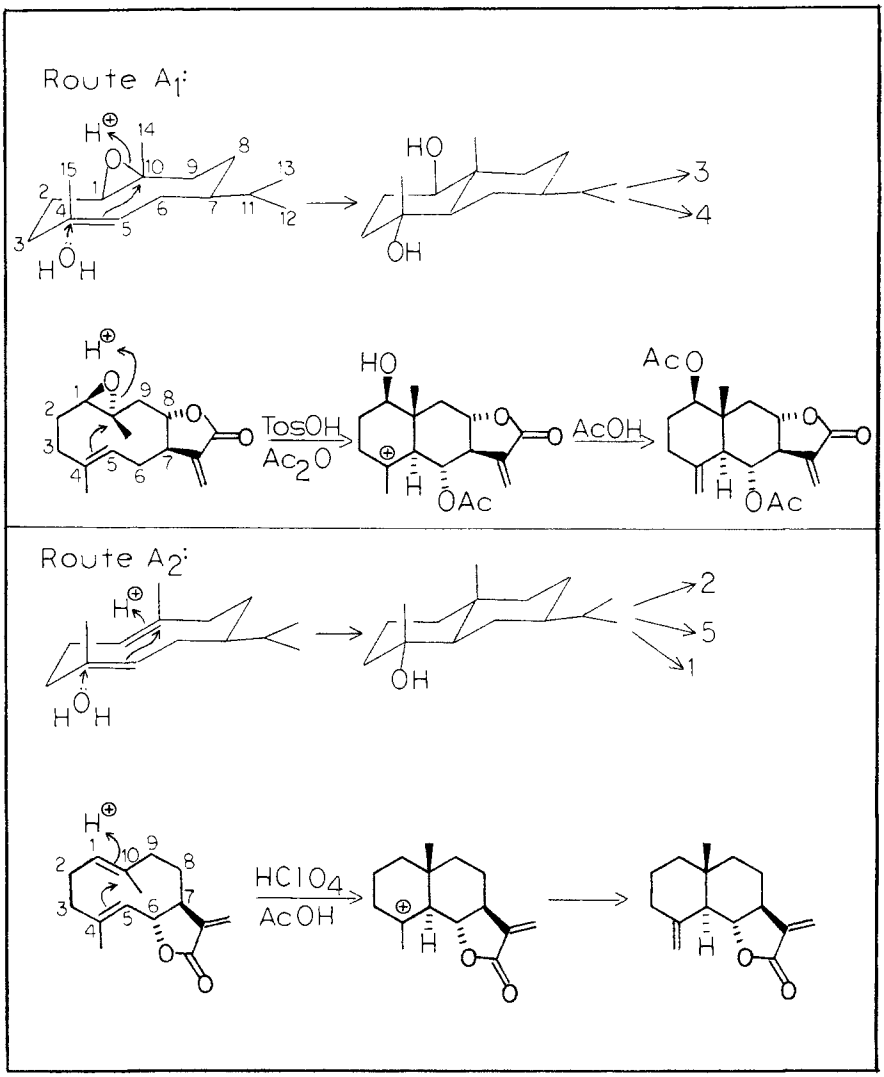


Fig. 2. Two proposed routes of eudesmanolide biogenesis (A_1 and A_2) and corresponding biomimetic laboratory syntheses. The numbers refer to *Iva* compound series shown in Fig. 10: 1. ivanuol series, 2. microcephalin, 3. ivangustin-ivangulin series, 4. ivasperin, 5. ivalin.

groups do not share a common history unique to them and therefore there is none to be discovered. The cladogram (Fig. 4) can be viewed as a hypothesis, and this hypothesis can be tested by searching for additional characters that may falsify the statement that A and B are more closely related to one another than either is to C.

A cladistic analysis involves these steps: 1. Identify the characters; 2. Arrange the characters into transformation series; 3. Determine the polarity of the transformation series; and 4. Construct the cladogram. While all these steps have been examined in the literature (e.g., Nelson and Platnick, 1981; Platnick, 1979; Wiley, 1981) they must be reexamined in order to use them with chemical data.

In step 1, skeletons, skeletal subclasses, and individual compounds are the units recognized as discrete characters. Skeletal characters are defined by the carbon ring systems

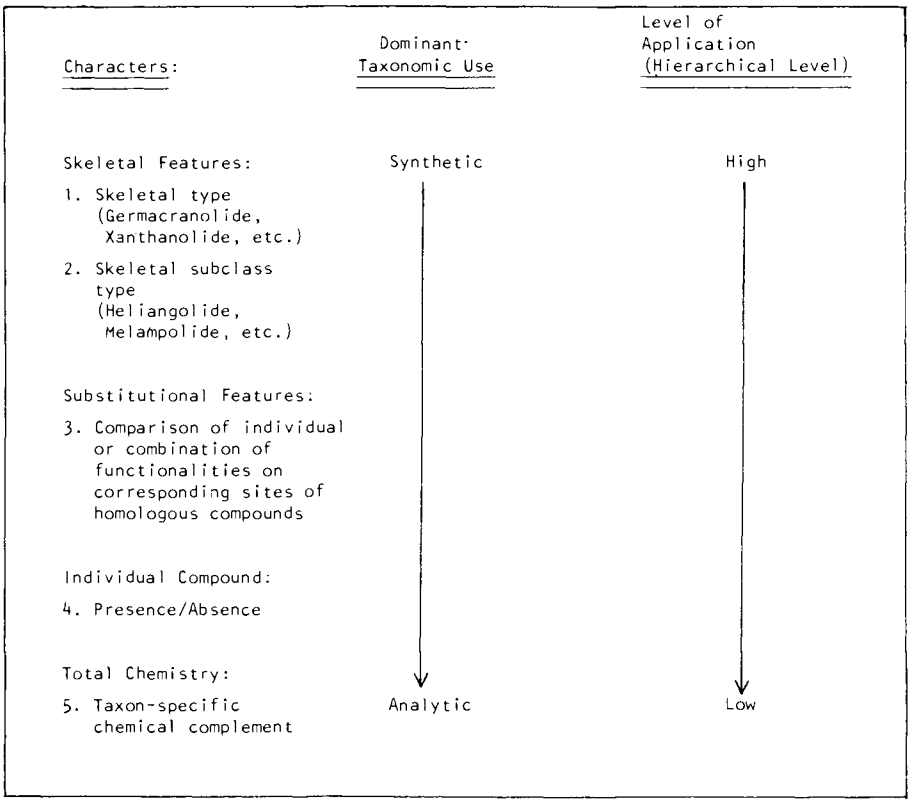


Fig. 3. Sesquiterpene lactone chemical features and the corresponding level of application as taxonomic characters.

upon which various substituents are introduced. Compounds, which are the products of the introduction of specific functionalities onto the carbon skeleton may also serve as characters. The diagnostic feature of the compound is the substitutional pattern which distinguishes that compound from homologous structures. Character analysis is focused on these diagnostic substituents and not on the compound as a whole. Thus, in an operational sense, the compound is a substitutional character.

The arrangement of characters into a transformation series (step 2) is essentially a homology problem. In a morphological analysis one examines the development, anatomy and ontogeny of various structures in order to determine their affinities. For instance, the structures on branches that are commonly called thorns are known from positional, anatomical and ontogenetic information to be non-homologous to one another in some cases. One type of thorn would be placed in the transformation series with leaves, one with branches, etc. When using sesquiterpene lactone data, the determination of homology is based on the hypothetical biosynthetic pathway. For example, given the two compounds, ivangustin and ivangulin (Fig. 10), an examination of the hypothetical biogenetic scheme for these eudesmanolides shows that ivangustin is a precursor of ivangulin and the two compounds therefore belong to the same transformation series (are homologous). All the chemical characters are thus placed into transformation series based on the relevant biogenetic scheme.

Determining the polarity of the transformation series (step 3) is a very different process from placing characters in transformation series. There are two basic types of evidence or

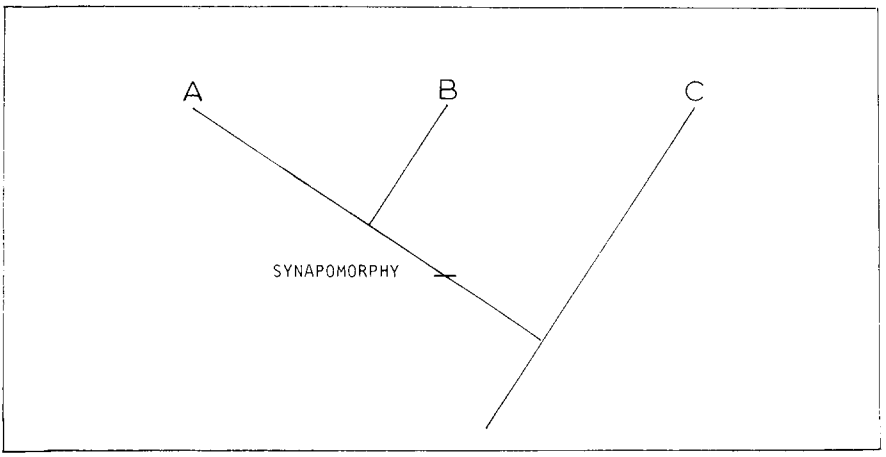
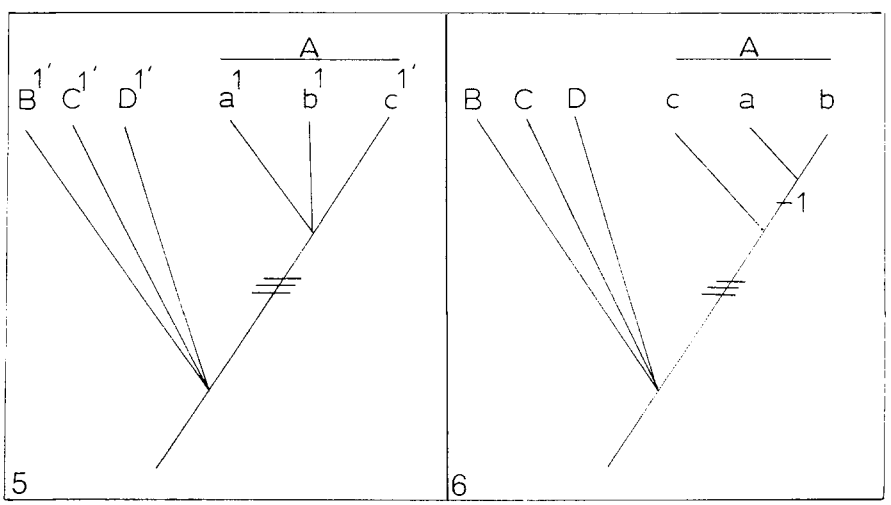


Fig. 4. Simple three-taxon statement cladogram.

argumentation schemes for polarity determination (Nelson, 1973). The first involves direct arguments (those not involving consideration of organisms outside the group) and the second, indirect arguments (those involving consideration of organisms outside the group). Nelson refers specifically to ontogenetic character transformation as the "only valid direct technique of which I am aware." Ontogenetic considerations are not really applicable to sesquiterpene lactone data because plant chemistries do not display any reliable or predictable developmental sequence which corresponds to the plant's life cycle. Rather than observing a progressive, step-wise synthesis of a series of homologous compounds during the course of plant development, overall synthesis seems to be initiated when the sesquiterpene lactone-synthesizing glandular tissues have differentiated.

The indirect arguments that Nelson (1973) refers to are not specified, however, the most useful is outgroup comparison (Watrous and Wheeler, 1981). The technique is explained



Figs. 5, 6. 5: Outgroup comparison method for determining character polarity: Distribution of homologous characters 1 and 1'. *6:* Outgroup comparison method for determining character polarity: Resolution of species relationships.

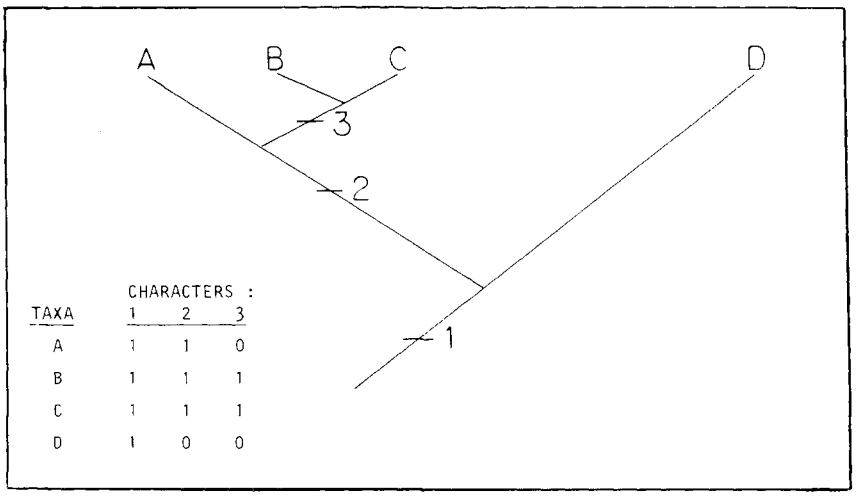


Fig. 7. Construction of a four-taxon statement cladogram based on three characters.

simply using Figs. 5 and 6. If you have four genera A, B, C and D and genus A has three species a, b, and c, and you wish to determine the relationships among the species in genus A, you can examine the distribution of the characters. In particular, one transformation series has two characters, 1 and 1′. Looking at the outgroup, which in this case is made up of the remaining three genera, you find only 1′ is present. This makes 1′ the general (plesiomorphic) character of the transformation series and 1 the novel (apomorphic) character. The apomorphic character 1 can be used to group species a and b together. The final diagram (Fig. 6) then has resolved the relationship problem of the species within genus A based on the distribution of the characters in the outgroup.

Step 4 is the construction of the cladogram. This is done by grouping the taxa based on the sharing of apomorphic characters as was illustrated in Figs. 5 and 6. A simple example is perhaps the easiest method for demonstrating this technique. Figure 7 gives three characters for the four taxa of a genus. If the 1's indicate apomorphies, then the cladogram constructed based on shared apomorphies (synapomorphies) is found in Fig. 7. This is the simplest and most informative method for displaying the information contained in the data matrix.

The application of cladistic analysis to sesquiterpene lactone data will be demonstrated with two examples, one (*Iva* L.) more complex than the other (*Tetragonotheca* L.).

Analysis of Tetragonotheca *Sesquiterpene Lactone Data*

The phylogenetic systematic study of sesquiterpene lactone variation in the four species of *Tetragonotheca* (Melampodiinae, Heliantheae) provides a straightforward example which can be used to demonstrate the four major steps of cladistic analysis. *Tetragonotheca* yielded a total of 26 compounds belonging to the melampolide (germacradienolide) skeletal subclass (Fig. 1A, **1b**) or to the derivative "repandolide" subclass (Fig. 8). The predominantly melampolide chemistry supports Urbatsch and Fischer (in prep.) in their transfer of *Tetragonotheca* to the Heliantheae subtribe Melampodiinae, which is characterized by the melampolide skeleton (Seaman et al., 1980).

In the Melampodiinae (Stuessy, 1977), modification of the melampolide skeleton is expressed as either alteration to yield derivative skeletons (e.g., repandolides, leucantholides) or substitutional alteration. Substitutional modification of the structure is commonly

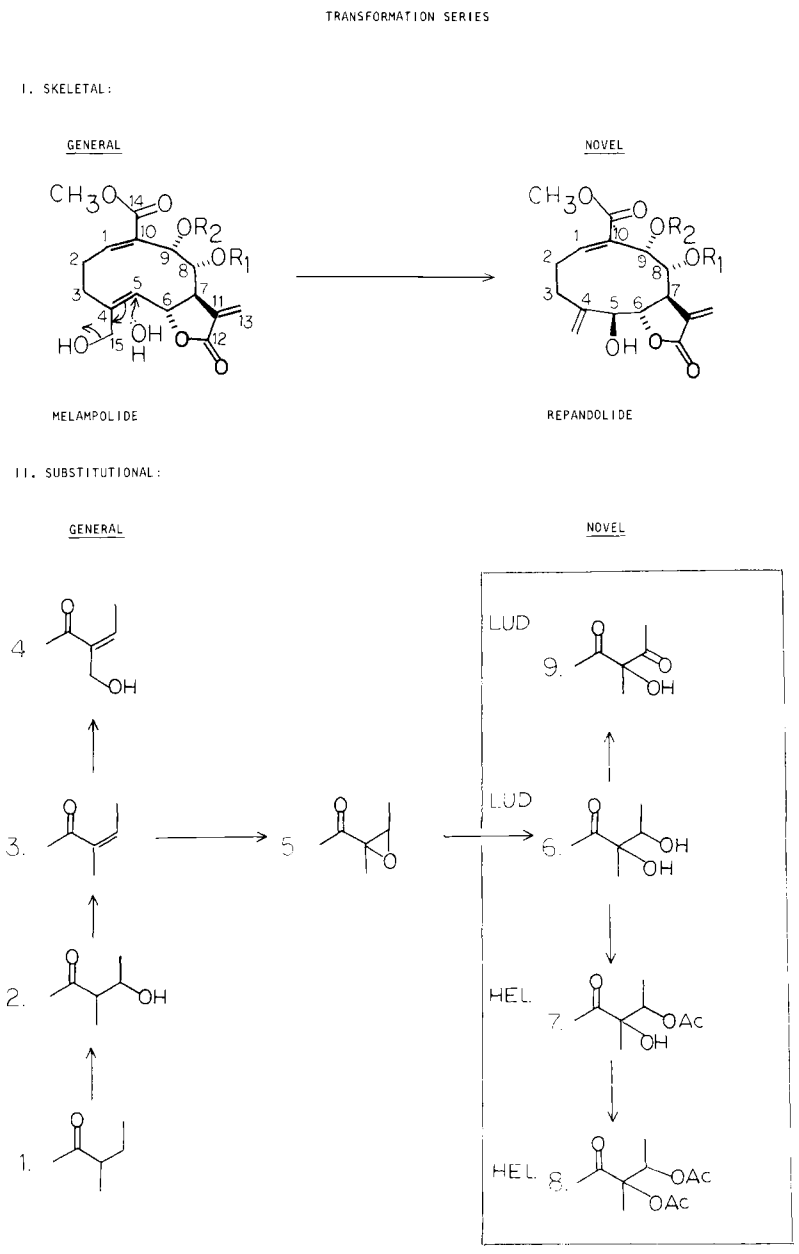


Fig. 8. *Tetragonotheca* sesquiterpene lactone skeletal and substitutional character transformation series. 1. 2-methylbutyrate. 2. 3-hydroxy-2-methylbutyrate. 3. angelate. 4. 5-hydroxyangelate (sarracinate). 5. epoxyangelate. 6. 2,3-dihydroxy-2-methylbutyrate. 7. 2-hydroxy-3-acetoxy-2-methylbutyrate. 8. 2,3-diacetoxy-2-methylbutyrate. 9. 2-hydroxy-3-oxo-2-methylbutyrate.

expressed by oxidation of C-8, C-9 and C-14 and less commonly by oxidation at C-2, C-3 and C-15. Also, epoxidation of the C-4–C-5 double bond is reported in seven of the melampolides isolated from *Melampodium* L., *Smallanthus* Mackenzie in Small (sensu Robinson) and *Tetragonotheca*. The 55 other melampolides isolated from *Melampodium*,

Acanthospermum Schrank, *Smallanthus* and *Tetragonotheca* exist as the unmodified *cis*-1(10)-*trans*-4(5)-germacradienolide. Oxidation of C-15 is reported in 14 of the 61 reported melampolide structures.

Oxidation at C-8 and C-9 is most often expressed by a hydroxyl group which has formed an ester bond with one of a variety of two-, three-, four-, and five-carbon acid sidechains. All 26 *Tetragonotheca* compounds possess oxygen functions at C-8 and C-9; 20 have various carboxylic acids ester-linked at both positions, while the remaining 6 have esters at only C-8 and hydroxyl groups at C-9. Common acids, such as acetic, isobutyric, angelic, tiglic, 2-methylbutyric and epoxyangelic acid occur at these two positions in all genera of the subtribe. At least one member of the 2-methylbutyrate, angelate, epoxyangelate and 5-hydroxyangelate series of common sidechains has been reported from each of the *Tetragonotheca* species. With the exception of 5-hydroxyangelate, these sidechains occur repeatedly as substituents on the compounds of other genera in the subtribe (Seaman et al., 1980). Epoxide cleavage of epoxyangelate and the formation of 2,3-dihydroxy-2-methylbutyrate plus derivative structures (Fig. 8) yield a series of novel sidechains. These novel sidechains are restricted to *T. helianthoides* L. and *T. ludoviciana* (T. & G.) Gray, while only common sidechains are produced by *T. repanda* (Buckley) Small and *T. texana* Gray & Englm. (Fig. 8).

Tetragonotheca sesquiterpene lactones can be categorized as either skeletal or substitutional characters: there are two skeletal characters, the melampolide and repandolide (Fig. 8). Biogenetically, the latter is derived from the former (Seaman et al., 1979). The melampolide character is the most common (22 compounds from three species). The novel sidechain series constitutes on biogenetic and distributional grounds a substitutional character which is distinct from the homologous characters defined by the angelate-related series of common sidechains. Other substitutional characters can be defined on the basis of functionalities (e.g., hydroxylation of C-15) but these will not be discussed in this example.

Cladistic analysis of *Tetragonotheca* can be summarized as follows:

Step 1.—Only two skeletal types, the melampolides and repandolides, are produced in *Tetragonotheca* and are considered separate characters. The two sets of five-carbon acid sidechains, the angelate-related series and the novel 2,3-dihydroxy-2-methylbutyrate-based series comprise the two major substitutional characters.

Step 2.—Biogenetically, the repandolides are derived via a melampolide precursor and, therefore the two can be considered a transformation series of homologues. The two sets of sidechains also possess this precursor-product relationship and are interpreted as a substitutional character transformation series.

Step 3.—The pairs of homologous characters can be ordered as to the general and novel (unique) using the technique of outgroup comparison. As described earlier, the repandolides are far less-broadly distributed, being restricted to two *Tetragonotheca* species. Thus, the presence of repandolides is a novel character absent from the outgroup consisting of the remaining genera of the Melampodiinae. Applying the same biogenetic and distributional criteria to the unique sidechain series, these are also judged to be the novel character of the pair based on their absence from the outgroup.

Step 4.—The members of *Tetragonotheca* have no chemical synapomorphy and are treated here as a monophyletic group based on morphological characters. The four taxa can then be grouped based on shared novel characters (synapomorphies). *Tetragonotheca texana* and *T. repanda* share the ability to synthesize the repandolide skeleton and can therefore be grouped together (Fig. 9).

A second sister group, *Tetragonotheca helianthoides* and *T. ludoviciana*, share the unique capacity to produce the novel sidechain series. The resulting cladogram based on the total sesquiterpene lactone data resolves the 4 species into two groups based on these two major synapomorphies, one skeletal and one substitutional.

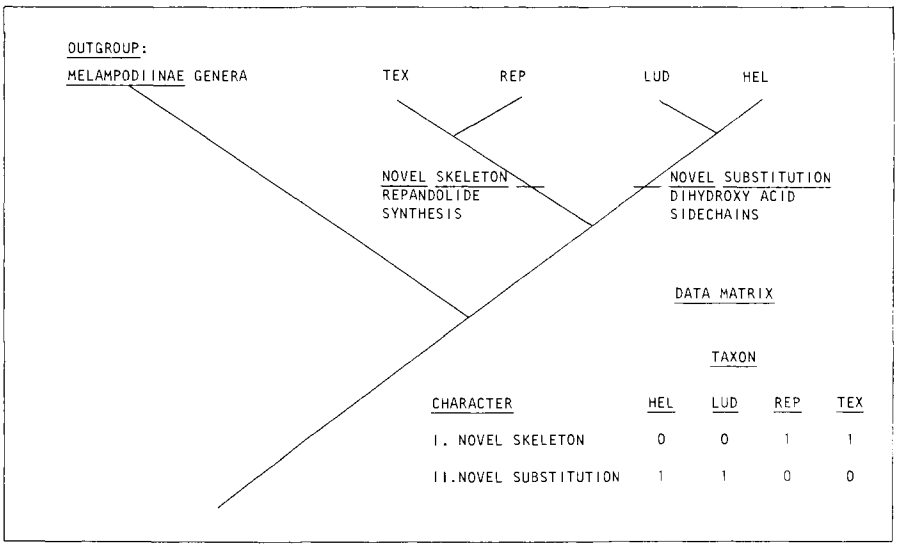


Fig. 9. Tetragonotheca sesquiterpene lactone cladogram. Species code: HEL—*T. helianthoides* Linn., LUD—*T. ludoviciana* Gray, REP—*T. repanda* Small, TEX—*T. texana* Gray et Engelm. 1 = presence, 0 = absence.

Analysis of Iva *Sesquiterpene Lactone Data*

A common feature of sesquiterpene lactone biogenesis not evidenced by *Tetragonotheca* is the frequent loss of the capacity to synthesize biogenetically complex structures. Noting this phenomenon, Mabry and Bohlmann (1977) speculated that among extant Asteraceae taxa, for many classes of compounds, loss of biosynthetic ability may be more common than gain, suggesting that chemically complex taxa are primitive. This observation emphasizes the distinction that must be made between placing characters in a transformation series and determining polarity. Divergence from a precursor-like skeleton is not necessarily representative of an evolutionarily derived or novel character, nor is resemblance to a precursor indicative of a plesiomorphic condition. Loss of biosynthetic capacity, possibly caused by the blocking of a reaction step converting a biosynthetic intermediate to the product, and the resulting synthesis of skeletons resembling biosynthetic precursors, intermediates or products of an altogether different pathway occur commonly in this family. As a consequence of this biosynthetic loss, the expected character polarity is reversed—the presence of precursor or intermediate-like compounds (or the presence of a product of another simultaneously activated pathway) becomes an apomorphic or novel character. Therefore, the presence of the biogenetically complex structure(s) becomes the plesiomorphic or general character.

In simple terms, given (1) the biosynthetic sequence A → B → C → D, and (2) an assortment of related taxa containing either compound C or D, the presence of C can be interpreted as either a biogenetically primitive or derived character. There are two ways to interpret these data. First, C would be the plesiomorphic condition if the ancestor had the ability to produce C but not D. Alternatively, the presence of C (in the absence of D) would be apomorphic if the ancestor had the ability to produce D. The only way this can be unequivocably resolved is through chemical knowledge of the ancestor. Since this is unlikely, the best method is to consult the most closely related group (outgroup). The character (C or D) found in the outgroup becomes plesiomorphic. A major prerequisite for outgroup analysis is an exhaustive, broad survey of chemical distribution. Such a wealth

of sesquiterpene lactone data is available in only a few cases, principally within the Heliantheae subtribe Ambrosiinae.

In the Ambrosiinae, the capacity to produce the biogenetic level three skeletal types, ambrosanolides (Fig. 1, **4**; isolated from *Ambrosia* L., *Parthenium* L., *Iva* and *Hymenoclea* Torrey & Gray) and xanthanolides (Fig. 1, **10**; isolated from *Ambrosia, Parthenium, Parthenice* Gray, *Iva* and *Xanthium* L.) is the dominant feature of the subtribal chemistry. The third most common skeletal type, the secoambrosanolides (Fig. 1, **11**; isolated from *Ambrosia, Hymenoclea,* and *Parthenium*) is a biogenetic level 4 derivative of ambrosanolides. If ambrosanolide (including secoambrosanolides) and/or xanthanolide synthesis is inactivated, either no sesquiterpene lactones are produced or eudesmanolides and/or germacranolides are synthesized. Eudesmanolides (isolated from *Ambrosia* and *Iva*) are major chemical constituents only of the genus *Iva*. The guaianolide-derived complexity level 3 and 4 skeletons and the eudesmanolides appear to be mutually exclusive. The distributional and biogenetic distinctions between the biogenetic level 3 and 4 guaiane-derived skeletal types and eudesmanolides suggest the following characterization of the subtribal chemistry.

Although the Ambrosiinae produces mostly biogenetic level 3 and 4 skeletons, most other Heliantheae subtribes and other tribes of the Asteraceae are characterized by a basic sesquiterpene lactone chemistry which includes germacranolides (Fig. 1, **1a-d**), eudesmanolides (Fig. 1, **2**) and guaianolides (Fig. 1, **3**). A plant's simultaneous synthesis of eudesmanolides and guaianolides is a rare chemical feature which to date has only been reliably reported from two *Iva* species. In a summary of *Artemisia* L. (Anthemideae) sesquiterpene lactone chemistry, Kelsey (1979) stated that of the 93 *Artemisia* taxa sampled for sesquiterpene lactones, the two most ubiquitous skeletons were guaianolides and eudesmanolides, and that only the single species, *A. nova* Nels. simultaneously produced both major skeletal types. However, a review of his chemical listings, references and other portions of the text indicated that only guaianolides and germacranolides were reported from *A. nova*. Whatever the status of *A. nova,* it is apparent that synthesis of eudesmanolides and guaianolides are usually mutually exclusive not only in *Artemisia* but also in the entire family.

In the three eudesmanolide-producing *Ambrosia* taxa and four of the six eudesmanolide-synthesizing *Iva* species, eudesmanolide biogenesis excludes guaianolide or derivative synthesis. In the two exceptional *Iva* species, none of the common guaianolide-derivative skeletal types is reported, only the relatively uncommon (in the subtribe) guaianolides and eudesmanolides. Two of the three eudesmanolide-producing *Ambrosia* species are considered to be derived. The first, *A. confertiflora* DC., produces only one eudesmanolide which is clearly a non-artifact, and this compound is reported from a restricted part of the taxon's range. All other sampled populations (approximately 250) produced germacranolides, ambrosanolides and secoambrosanolides (Renold, 1970; Yoshioka et al., 1970). The second derived taxon, *A. polystachya* DC. is according to Payne (1966) a derivative of *A. arborescens* Mill. which synthesizes ambrosanolides and secoambrosanolides (Seaman, 1982). *Ambrosia camphorata* (Greene) Payne was designated by Payne a member of the least-specialized (core) group of taxa (1964). However, relationships within this group are poorly understood and it is difficult at this time to determine whether or not it represents a relatively derived taxon within this group. Based on these observations it seems most appropriate to treat the presence of eudesmanolides in *Iva* as apomorphic. Hence the presence of these compounds in *Ambrosia* and *Iva* taxa represent parallel evolutionary events in which eudesmanolide synthesis has displaced the more common synthesis of guaiane derivatives (ambrosanolides, secoambrosanolides and xanthanolides). This interpretation suggests that the presence of ambrosanolides, secoambrosanolides, xanthanolides, and guaianolides is primitive within the subtribe and that eudesmanolide synthesis has been activated secondarily in two genera.

All *Iva* (Ambrosiinae) species have been sampled for sesquiterpene lactones (see Seaman, 1982, for references). Each taxon yielded identifiable compounds except *I. hayesiana* Gray,

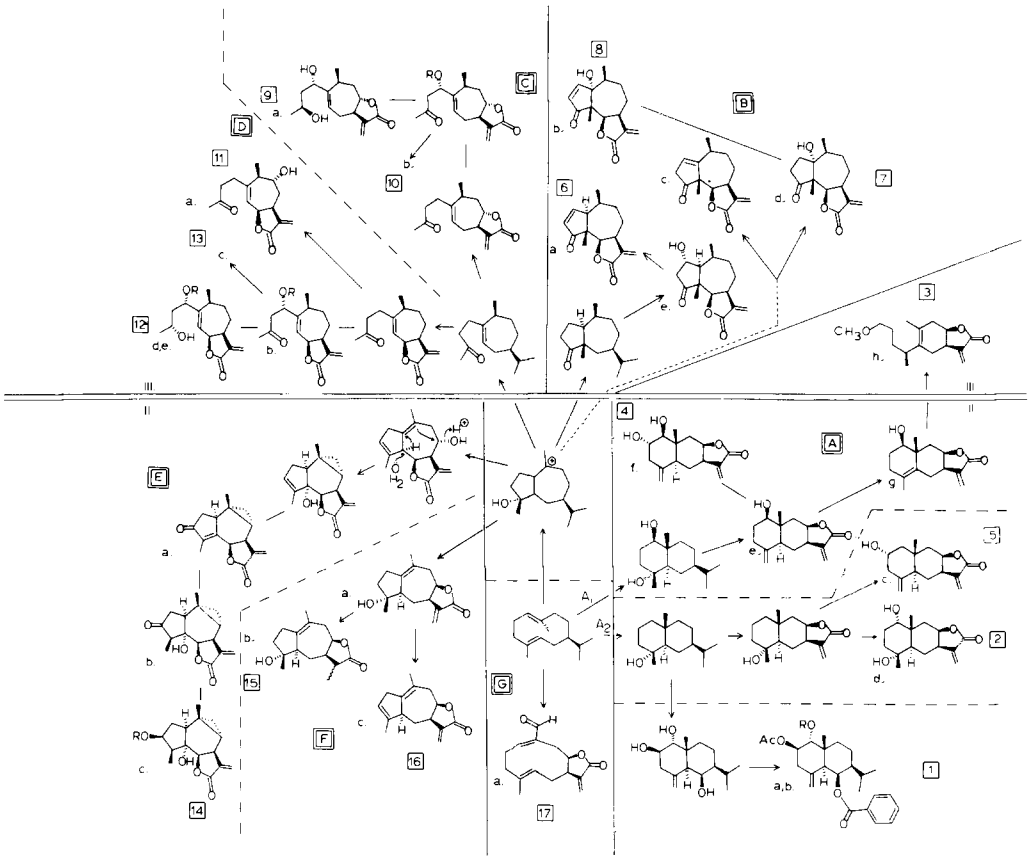


Fig. 10. Hypothetical biogenetic relationships of *Iva* sesquiterpene lactones. A, *Eudesmanolides:* a, ivanuol (R = H). b, ivanuol-1-O-acetate (R = Ac). c, ivalin. d, microcephalin. e, asperilin. f, ivasperin. g, ivangustin. h, ivangulin. B, *Ambrosanolides:* a, ambrosin. b, parthenin. c, neoambrosin. d, coronopilin. e, ivoxanthin. C, *C-8 Lactonized Xanthanolides:* a, ivalbin. b, xanthinin (R = Ac). D, *C-6 Lactonized Xanthanolides:* a, ivalbatin. b, parthemollin (R = H). c, parthemollin acetate (R = Ac). d, ivambrin (R = H). e, apachin (R = Ac). E, *C-6 Lactonized Guaianolides:* a, anhydroivaxillarin. b, ivaxillarin. c, axivalin (R = Ac). F, *C-8 Lactonized Guaianolides:* a, pseudoivalin. b, dihydropseudoivalin, c, ziniolide. G, *Germacranolides:* a, frutescin. The boxed numbers 1–17 refer to individual sequences of related structures. Each sequence can be followed through the series of intermediates connecting an initial structure (which may serve as a point of origin for more than one sequence) and a terminal structure. These numbers and the skeletal type letter (A–G) are used to describe the characteristic chemical features of the taxa included in the cladogram. The published structure for ivanuol and ivanuol acetate (Bohlmann and Zdero, 1979b) is revised to the structure shown here. Bohlmann and Zdero argue that their data indicate a *trans*-decalin type structure; however, their structure which was drawn with a C-10-α-methyl group and C-5-β-hydrogen is incompatible with all other eudesmanolides of the Asteraceae. The alternative *trans*-decalin structure with C-10-β-methyl and C-5-α-hydrogen is biogenetically compatible with other eudesmanolides of the Asteraceae and with the data published for the structures.

which produced compounds that polymerized before identification, and *I. annua* L., which produced a series of non-lactonic sesquiterpene alcohols with isopropyl sidechains.

W. Herz and associates were the first to recognize the taxonomic implications of their analysis of *Iva* sesquiterpene lactones (Farkas et al., 1966; Herz et al., 1964a and 1964b;

Herz et al., 1967). The biogenetic level 2 skeletons (Fig. 1A), guaianolides and eudesmanolides, and the level 3 skeletons, ambrosanolides and xanthanolides, typify the genus. All reported *Iva* structures are grouped in Fig. 10 according to their biogenetic complexity level (2 or 3) and skeletal type (A–G). Within each skeletal type structures are tentatively linked by solid lines to indicate series of biogenetically related compounds. Each skeletal type is indicated by a double-boxed letter and each homologous compound series is labeled with a single boxed number. Lower-case letters indicate the known compounds (see Fig. 10 legend for names).

Iva has more taxa and a considerably more complex chemistry than *Tetragonotheca*. These facts are naturally reflected in the cladistic analysis. The four previously described procedural steps will be used for *Iva*.

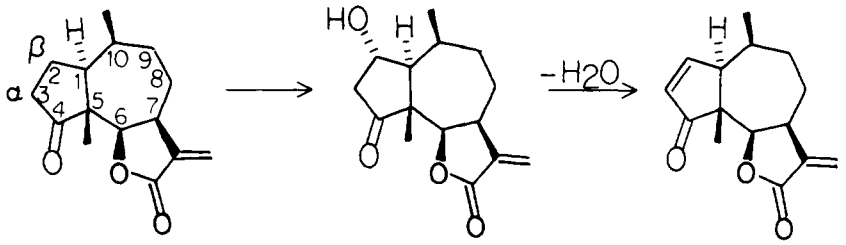
Step 1: Character Identification

Skeletal characters.—Skeletal characters (A–G) were identified using the *Iva* biogenetic scheme (Fig. 10) as a guide. Although these characters are based on the major skeletal classes displayed in Fig. 1, several represent subdivisions of major skeletal types. For example, as mentioned earlier there appear to be two modes of eudesmanolide biogenesis in the Ambrosiinae (Fig. 2). The most common mode (A_1, Fig. 2) yields a series of C-1-β-hydroxy eudesmanolides typified in *Iva* by asperilin, ivasperin and ivangustin. The second mode (A_2, Fig. 2) is represented in *Iva* by microcephalin and ivalin. These two modes of eudesmanolide biogenesis can be used as separate characters. Outgroup analysis indicates that presence of mode A_1 is a novel condition and that its absence is the general condition. As mentioned previously, route A_1 constituents are reported from one member of the outgroup, *Ambrosia confertiflora*, but these reported constituents are probably artifacts. In *Iva*, the co-occurrence of ivangustin and ivangulin disproves an artifactual origin for this β-hydroxy eudesmanolide precursor and its seco-derivative. Similarly, presence of mode A_2 biogenesis is the novel condition and its absence is the general condition. Here, too, mode A_2 biogenesis has parallel expressions in a limited number of *Ambrosia* and *Iva* taxa. In two instances, the xanthanolides and the guaianolides, the reported compounds included C-6 and C-8 lactonized constituents. Consequently, the xanthanolides were divided into two characters, the C-8 *trans*-lactones (Fig. 10, **C**) and the C-6 *cis*-lactones (Fig. 10, **D**); the guaianolides were divided into the C-6 *cis*-lactones with the unique cyclopropane-containing skeleton (Fig. 10, **E**) and the C-8 *cis*-lactones (Fig. 10, **F**).

In two instances, eudesmanolides and ambrosanolides, the skeletal types are subdivided, although each retains its rank as a skeletal character. In the eudesmanolides, two modes of biogenesis (A_1 and A_2) are proposed which share a homology based on the acid-catalyzed Markovnikov-oriented cyclization of a germacrolide precursor. Because they differ principally in terms of the substitution pattern of the germacrolide precursor, the two modes are lumped into a single skeletal character. In the ambrosanolides, the situation is less well defined because the substitution patterns of reaction sequence 7 and 8 products (Fig. 10) only suggest a minor modification of the complex series of methyl and hydride shifts which collectively generate ambrosanolides of reaction sequence 6. Although this proposal is highly speculative, it is felt that the possibility of such a modification should be indicated by a dotted line. At present, these doubts do not warrant subdivision of the ambrosanolide skeletal character.

Substitutional characters.—The substitutional picture for *Iva* is clearly more complicated than it is for *Tetragonotheca*. The complexity of the *Iva* chemistry dictates that substitutional characters be defined by a procedure which reduces the informational load to a manageable number of useful features. In substitutional character definition for *Iva* or any chemically complex genus, this reduction is achieved through a biogenetically based evaluation of compound distribution. This process should generate the maximum number of useful characters but without improperly subdividing legitimate characters into positively-correlated constituent characters.

REACTION SEQUENCE - 6



DAMSIN (D) IVOXANTHIN (I) AMBROSIN (A)

POSSIBLE CHARACTERS OF REACTION SEQUENCE - 6	EXPRESSIONS	
	Presence	Absence
I	1. D + I 2. I*	1. Reaction sequence not detected 2. D present
A	1. D + I + A 2. D + A 3. I + A 4. A*	1. Reaction sequence not detected 2. D present 3. D + I present 4. I present

* In biogenetic terms, D + I = I, because the chemistry of I indicates that D is also present as a precursor which at the time of analysis did not accumulate in detectable quantities. The same is true of the second series, D + I + A = D + A = I + A = A, because I and D are assumed to be present in the A-producing taxon but did not accumulate in detectable quantities at the time of analysis.

Fig. 11. Biogenetic basis for character identification. An example using *Iva* ambrosanolide reaction sequence 6.

First, it is necessary to delimit those structural features which can be individually or collectively defined as a discrete substituent. This is complicated by the causal relationship between the presence of certain functional groups and the secondary introduction of others. For example, the introduction of a double bond into the cyclopentanone ring of damsin (Fig. 11) presumably via an ivoxanthin-like intermediate is mediated by the presence of the keto function which "directs" the oxidation and elimination reactions to yield ambrosin. This type of interrelationship partly accounts for the high frequency of cyclopentenone functions in sesquiterpene lactones. Thus, in substitutional character definition, a decision must be made regarding the interrelationship of neighboring functional groups. The resulting individual functional group or assemblage of related functional groups (e.g., cyclopentenone) can then be treated as discrete substituents.

A second consideration in substitutional character definition is the extent of homology of a particular substitutional feature. Is a substituent such as β-hydroxylation at C-2 a general

substitutional character which is homologous for more than one skeletal type? Would such a substituent on a eudesmanolide be homologous to the same substituent at the corresponding position on a guaianolide? Presumably, a substituent attached to a common precursor is homologous for all derived sesquiterpene lactone skeletal types. Given the absence of information regarding the point of biogenetic origin for substituents, the prudent course is to define a substitutional character only in terms of a particular skeletal type. Thus, as represented in Fig. 10, each substitutional character is restricted to a particular skeleton.

The third and most difficult step in character delimitation is made necessary by the observation that a plant's chemistry often consists of homologous compounds (of a specific skeletal type) with substituents that can be arranged in an order corresponding to a likely reaction sequence. This pattern was noted in early sesquiterpene lactone studies as seen in the previously quoted statement by Geissman. Various *Iva* chemistries can be arranged in such sequences (Fig. 10) based on distributional and mechanistic arguments. For example, the co-occurrence of ivoxanthin and ambrosin in *I. xanthifolia* Nutt. corresponds to the hypothetical biosynthetic sequence, damsin → ivoxanthin → ambrosin, described earlier.

This apparent precursor-product relationship of members of these biogenetic series suggests that skeletal substituents may have to be assembled into discrete biogenetic transformation series of homologous characters based on the most mechanistically sound and parsimonious hypothetical reaction sequence. Thus, a substitutional character series encompasses the entire series of related substituents linked by the reaction sequence.

The value of a biogenetic approach to substitutional character definition is illustrated in Fig. 11, where the two characters, A and 1, are defined by the reaction sequence: damsin (D) → ivoxanthin (1) → ambrosin (A). Clearly, a variety of chemistries can be interpreted as expressions of characters A (D+1+A, D+A, 1+A, A) and 1 (D+1, 1). The rationale for this interpretation is that the presence of ambrosin implies the co-occurrence of ivoxanthin and other precursors even if they cannot be detected at the time of analysis in a particular specimen. Thus in a biogenetic sense, it is difficult to distinguish a character based on the presence of ambrosin from one based on the presence of ambrosin and ivoxanthin.

Other precursor-product-related series of compounds are summarized in Fig. 10 as 17 independent transformation series. As discussed, the presence of each transformation series should be recognized when the terminal member of the sequence is identified, although some latitude in deciding the identity of the terminal product is needed. For example, a compound can undergo acetylation to yield a variety of acetates. Acetylation occurs so readily in the biogenesis of sesquiterpene lactones that it is usually more satisfactory to designate the unacetylated compound and its acetate derivatives as collectively or singly representing the terminal product.

As indicated by the absence of arrows in parts of Fig. 10, the substitutional reaction sequence in several of the 17 characters is unresolved (e.g., substitutional character 14). Fortunately, in these cases the entire sequence of compounds is reported from certain *Iva* species, thus eliminating the need to establish the terminal product and sequence of intermediates in the reaction sequence. The presence of the unresolved series of compounds is both indicative of a complete sequence and the expression of the novel substitutional character.

Step 2: Arrangement into Transformation Series

Skeletal characters.—Sesquiterpene lactone biogenesis follows several alternative, branching routes arising from a common farnesyl precursor (Fig. 1). Apparently, in any given taxon, zero, one or more routes are activated and alternative routes inactivated. The genetic nature of this activation-inactivation process is unknown. Some taxa are also characterized by novel extensions of the common pathways. Examples include the ambrosanolide skeleton of the Ambrosiinae (extension of guaianolide pathway) and the seco-

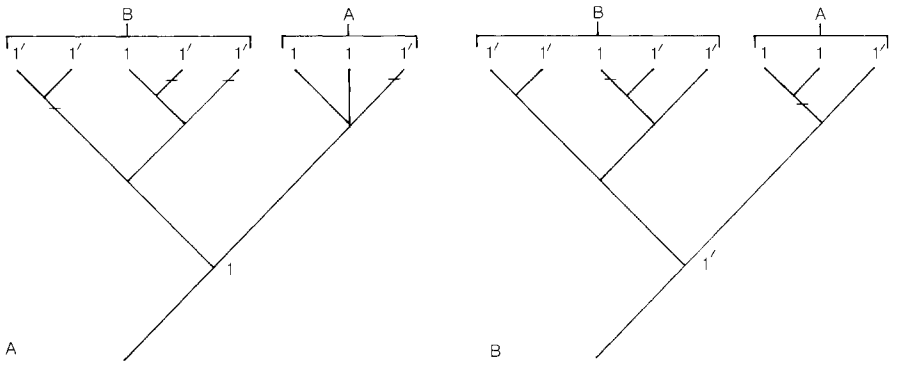


Fig. 12. The most parsimonious interpretation for determining the polarity of a transformation series when both characters are found in the taxon under investigation and in the outgroup. A. When character 1′ is assumed to be apomorphic, four character changes are necessary to explain the distribution of character 1′. B. When character 1 is assumed to be apomorphic, two character changes are necessary to explain the distribution of character 1.

eudesmanolide skeleton of *Iva angustifolia* Nutt. (extension of eudesmanolide pathway). A skeletal biogenetic transformation series is a branching system incorporating the two major eudesmanolide- and guaianolide-based pathways and related minor pathways. Each skeleton (Fig. 1) is a discrete homologue of this biogenetic transformation series. Consequently, the concept of a strictly linear transformation series of homologous characters is not applicable to sesquiterpene lactone skeletal homologues (skeletal homology is defined by origin from a common precursor).

Substitutional characters.—The association of homologous substitutional characters is based on a novel reaction sequence linking homologous compounds. Most of the *Iva* reaction sequences (series 1–17, Fig. 10) are novel, resulting in terminal products that have limited distributions. Structures identical to or analogous to the precursors in these sequences are widely distributed within the investigated taxa and the members of the outgroup. Thus, the apomorphic character is often expressed by the presence of the terminal product(s) of these reaction sequences and the other plesiomorphic characters are usually the precursor- or intermediate-like members of the transformation series.

Step 3: Directionality of the Transformation Series

Skeletal characters.—The comparatively facile activation/inactivation of alternative biogenetic routes complicates the process of character polarity determination by outgroup comparison. Especially if the outgroup is a higher taxon, the likelihood that a novel skeletal character is also expressed in some members of the outgroup is even greater. The parallel expression of a novel character in one part of a large outgroup does not invalidate the character as a possible apomorphy. When the characters of a transformation series are found in both the taxon in question and the outgroup one must rely on the concept of parsimony. Simply stated, parsimony requires that we accept the simplest explanation for the evolution of the character distribution. Figure 12 illustrates how this principle operates. Genus A has three species and is most closely related to genus B which has five species. We are seeking to understand the relationship among the three species of genus A and the only information available is found in the transformation series containing characters 1 and 1′. The outgroup also contains the same two characters. By examining the distribution of the characters in the outgroup we find that when the assumption is made that 1′ is the apomorphic character, four character changes are required to explain the present distribution pattern (Fig. 12A). However, when character 1 is assumed to be apomorphic, only

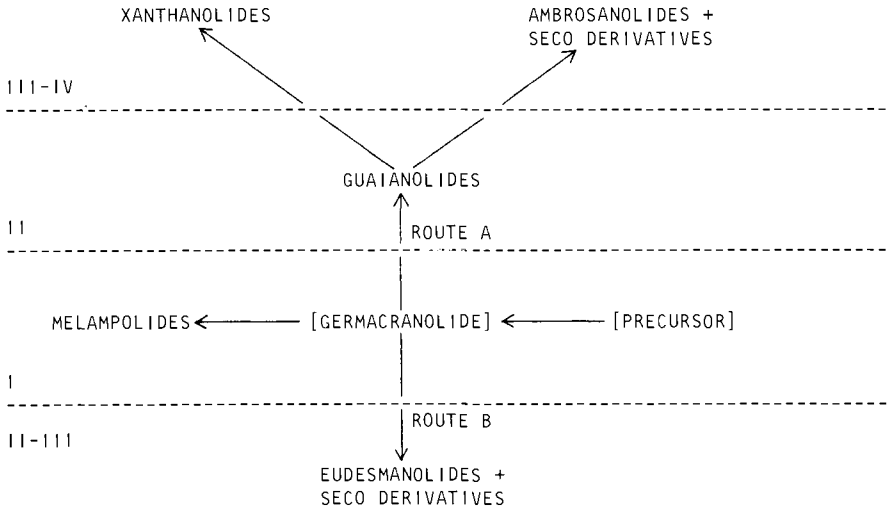


Fig. 13. The major routes of *Iva* sesquiterpene lactone biogenesis. Route A corresponds to the biosynthesis of guaianolide and derivative skeletons. Route B corresponds to eudesmanolide and secoeudesmanolide biogenesis. Roman numerals indicate biogenetic complexity level.

two changes are necessary to explain the distribution (Fig. 12B). The second assumption is favored in the absence of any other information because it requires the fewest assumptions, hence, it is the most parsimonious. It is, of course, preferable to have complete agreement on character distribution when consulting the outgroup but when it does not occur the above procedure can be used. As the number of parallel occurrences in the outgroup increases, there is a rapid decline in the possible utility of this procedure.

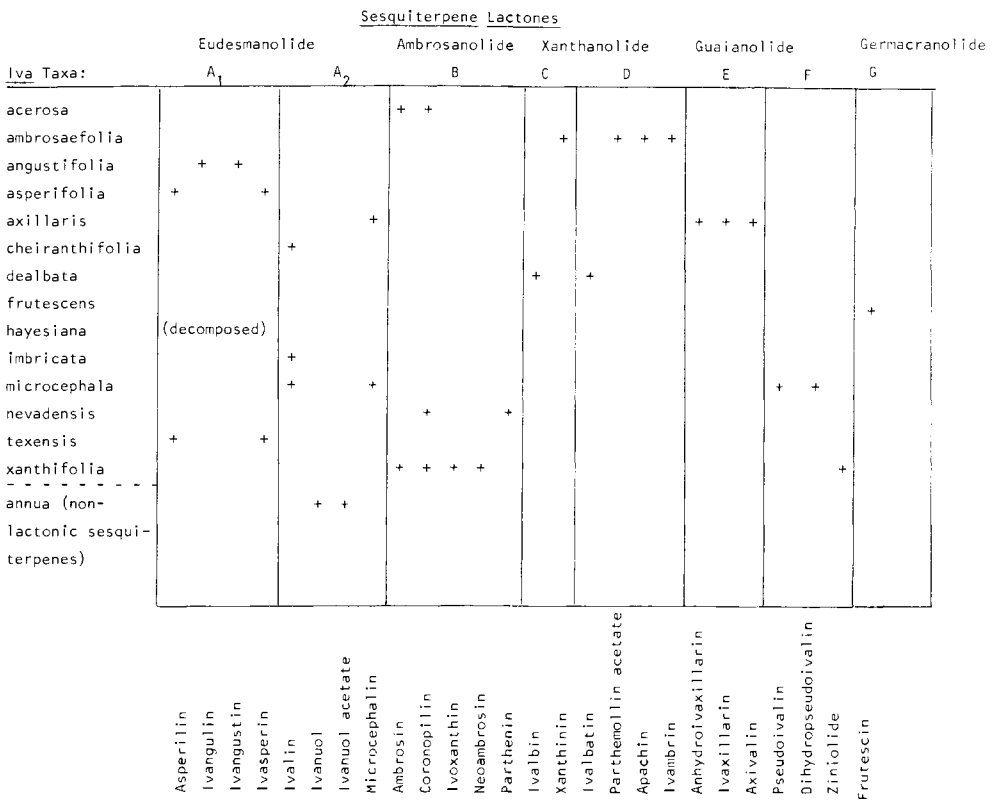
In this investigation, all chemically investigated Ambrosiinae taxa (*Ambrosia, Hymenoclea, Parthenice, Parthenium* and *Xanthium*) not including *Iva* are used collectively as the outgroup.

Outgroup analysis for *Iva* skeletal types indicates that route A skeletons (Fig. 13) (ambrosanolides, xanthanolides and to a lesser extent guaianolides) are plesiomorphic homologues. Their presence in many *Iva* taxa does little to reinforce the notion that *Iva* is monophyletic. Route B skeletons (Fig. 13) appear in the outgroup in only three of the approximately 41 *Ambrosia* species. Eudesmanolides are reported from no other genus of the outgroup (*Parthenium, Parthenice, Xanthium* and *Hymenoclea*). Secoeudesmanolides are unreported from the outgroup. Assuming that their presence in *Ambrosia* is of a parallel origin, we can treat eudesmanolide presence as a good synapomorphic homologue.

The cyclopropane guaianolide skeleton (character E, Fig. 10) possesses a novel skeletal feature (a 5-6-3 tricyclic system) which is unreported from the outgroup, thus, its presence is considered novel. The conformationally unusual melampolide skeleton (character G, Fig. 10) is also absent from the outgroup, indicating that its presence is a novel character. Another possible apomorphy is the presence of C-6 lactonized xanthanolides (character D, Fig. 10) which are reported only twice in the outgroup, in *Parthenice mollis* Gray and *Parthenium fruticosum* Less. The more common C-8 lactonized xanthanolides were isolated from *Ambrosia, Parthenium* and *Xanthium*. All other skeletons are frequently found in the outgroup.

The novel nature of the 17 reaction sequences suggests that the apomorphic substitutional character corresponds to the presence of the terminal product(s). The general or plesiomorphic condition corresponds to the absence of these products. The character state polarity of each substitutional character is discussed below:

Table 1. Sesquiterpene lactones reported for *Iva* L.



1. Within the Ambrosiinae, the presence of the five-membered lactone ring is the general condition. The ivanuol-type eudesmane series of benzoate derivatives which are unique to *Iva* and unreported from the outgroup possess an isopropyl sidechain instead of the five-membered lactone ring. Thus the presence of this eudesmane structure with the unoxidized isopropyl group is considered a novel character.

The presence of the characters defined by the detection of the following compounds is considered novel (apomorphic) and their absence is judged the general condition (plesiomorphic) based on outgroup comparison:

2. Microcephalin; 3. Ivangulin (only reported secoeudesmanolide of the Ambrosiinae); 5. Ivalin; 9. Ivalbin; 11. Ivalbatin; 12. Ivambrin and its acetate derivative, apachin; 14. Entire series of cyclopropyl derivatives; 15. Pseudoivalin and its dihydro derivative (dihydropseudoivalin); 16. Ziniolide; 17. Frutescin (conformationally distinct from melampolides of Melampodiinae).

The polarity of the following characters is less well defined:

4. Ivasperin, a derivative of asperilin, is reported in *Ambrosia polystachya, Iva texensis* Jackson and *I. asperifolia* Less. Thus the character defined by the presence of ivasperin has at least two parallel origins, once in *Ambrosia* and presumably once in *Iva.*

5 and 7. Ambrosin and coronopilin are common constituents of *Ambrosia, Hymenoclea, Iva* and *Parthenium* and information on the placement of the taxa in these groups is not available. As a result, outgroup comparison involving characters associated with these structures yields less significant results.

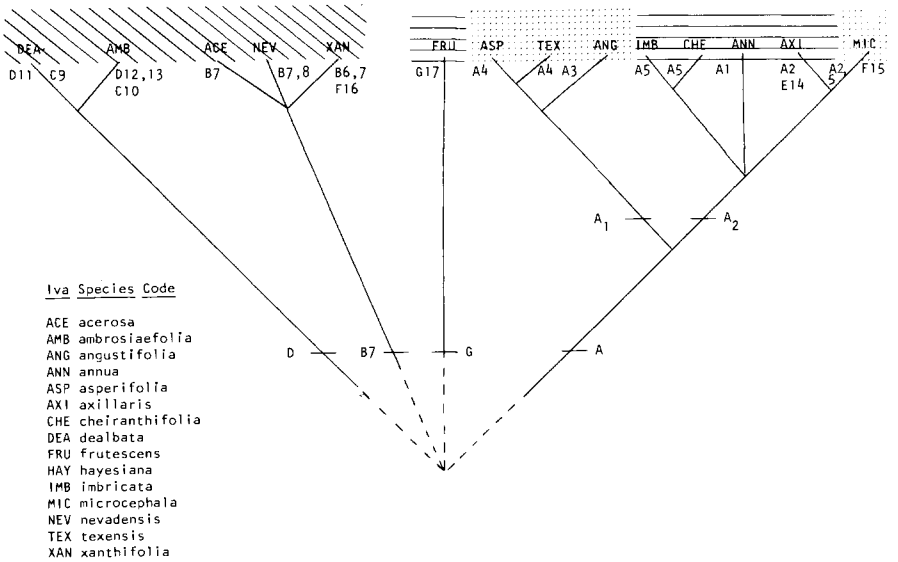


Fig. 14. *Iva* sesquiterpene lactone cladogram. Species code: ACE—*I. acerosa* (Nutt.) Jackson, AMB—*I. ambrosiaefolia* Gray, ANG—*I. angustifolia* Nutt., ASP—*I. asperifolia* Less., AXI—*I. axillaris* Pursh, CHE—*I. cheiranthifolia* H.B.K., DEA—*I. dealbata* Gray, FRU—*I. frutescens* L., HAY—*I. hayesiana* Gray, IMB—*I. imbricata* Walt., MIC—*I. microcephala* Nutt., NEV—*I. nevadensis* M. E. Jones, TEX—*I. texensis* Jackson, XAN—*I. xanthifolia* Nutt. The letter and number code for structural identification is described in Fig. 10. Diagonal lines indicate Jackson's section Cyclachaena, dotted lines indicate section Linearbractea and horizontal lines indicate section Iva.

8. Parthenin biogenesis is reported in only two parallel instances, once in *Parthenium* and once in *Iva*. Despite its single parallel occurrence in the outgroup, it is most parsimonious to assign the novel condition to its presence and the general condition to its absence.

10. Xanthinin also has been reported from *Xanthium* wherein it is a common constituent. Outside *Xanthium* it is only reported from *Iva ambrosiaefolia* Gray. It is most parsimonious to assume that the presence of xanthinin is the novel condition and the absence is the general condition.

13. Character 13 is defined by the acetylation of an intermediate of reaction sequence number 12. Outside *Iva,* this product is only reported in the unacetylated state from *Parthenice mollis.* Its presence is most appropriately characterized as the novel condition.

The distribution of sesquiterpene lactones in *Iva* can be found in Table 1.

Step 4: Construction of the Cladogram

The cladogram for the *Iva* sesquiterpene lactone (STL) data was constructed in the most parsimonious manner based on the synapomorphies indicated in Fig. 14 (results in the cladogram requiring the fewest assumptions).

Results and Discussion

The information content of the cladogram depends on the quality and quantity of the data used to construct it. The diagram reflects the data 100 percent and shows no more (or less) than the data indicate. Polytomies (more than two branches at one node) are either the result of a lack of sufficient information to fully resolve the cladogram or are an indication of possible multiple speciation or hybridization events. A polytomy is a signal

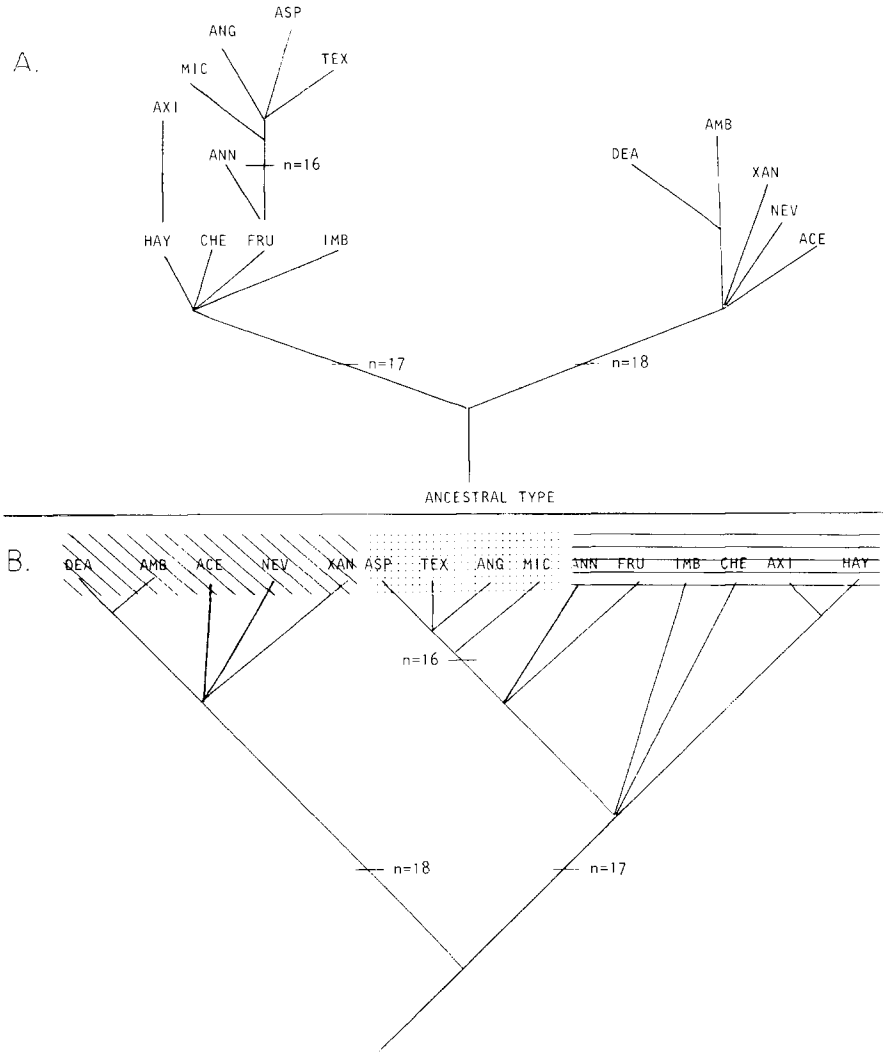


Fig. 15. A. Jackson's (1960) phylogenetic interpretation of *Iva.* B. A reconstruction of Jackson's phylogenetic interpretation of *Iva.* All sister group relationships have been maintained from Jackson's original diagram. Species code is provided in Fig. 14. Diagonal lines indicate Jackson's section Cyclachaena, dotted lines indicate section Linearbractea and horizontal lines indicate section Iva.

to return to the organisms and reexamine them. Sometimes characters have been misinterpreted and must be reinterpreted; sometimes additional characters can be identified (Nelson and Platnick, 1981). The additional data may resolve the polytomies; if not, then they must be left in the diagram because they accurately reflect the data. There are several such areas on the sesquiterpene lactone (STL) cladogram of *Iva.* It should be pointed out at this time that there is no synapomorphy in this data for the genus *Iva.* Each of the four lines has a synapomorphy that is not shared with any other line. *Iva* is treated in one diagram at this time because we have not completed our analysis of the sesquiterpene lactone data for the entire subtribe and so cannot apply this evidence to the resolution of the broader question regarding the disbanding of this genus. In any event, even if the genus

is not monophyletic, polyphylesis does not affect the portrayed sister group relationships because the entire subtribe (minus *Iva*) was used as the outgroup. The four independent groups defined by this analysis are connected at the base by a dashed line indicating the tenative nature of this relationship. Further resolution will have to await the completion of the morphological analysis (M. R. Bolick, in prep.) and the chemical analysis of the rest of the subtribe. The resolution on the rest of the cladogram is good with all sister group relationships being dichotomous except two.

A cladistic study of this type is useful in a number of ways. A close examination of the data in order to delimit and determine the polarity of the transformation series yields new insights into sesquiterpene lactone biogenesis. In addition, because of the high information content of the cladogram, maximum resolution of the patterns of *Iva* sesquiterpene lactone biogenetic capability can be obtained. Finally and probably most significantly, these results can be applied to an analysis of congruence. Congruence is defined by Mickevich (1978) as follows: "Taxonomic congruence is the degree to which classifications of the same organisms postulate the same groupings." The data for the separate classifications can belong to any of the following categories: (1) data of different types (e.g., morphological, micromolecular, macromolecular) from the same organisms, (2) one data set divided into subsets, (3) data from different organisms (e.g., parasite/host data); biogeographical data can be used to produce area cladograms that can be compared with organismal cladograms. A more specific term for the comparison of different types of data from organisms or of the organisms and their distribution is consilience. This term was first introduced by Whewell (1840) who defined it as "when one class of characters predicts others." In any event, congruence deals with examining the predictability of a cladogram by comparing it with another cladogram produced by a separate type of data. This predictability is what we are interested in. If the groups we have indicated in the cladogram are truly natural then they should be able to predict the outcome of a cladistic analysis on other types of data for the same group.

Once the STL cladogram has been constructed the ideal situation would be to compare it with a morphological cladogram and determine whether or not there is congruence. Unfortunately, no such cladogram is available at this time although one is under construction (M. R. Bolick, in prep.). There is, however, a "phylogenetic interpretation" of *Iva* by Jackson (1960) in his revision of the genus (Fig. 15A). This interpretation has been redrawn (Fig. 15B) in the style of the cladogram in order to facilitate comparison. However, all of Jackson's sister group relationships have been maintained.

Comparing the two diagrams (Figs. 14 and 15B), one finds some congruence in the sister group relationships of the taxa but little on higher levels of grouping. The sister group relationships among the terminal taxa change in several cases with the most noticeable being the placement of *I. frutescens* L. and *I. microcephala* Nutt. *Iva frutescens* is placed by itself on the STL cladogram because of its unique chemistry, while Jackson has it as the sister taxon of *I. annua* and the n = 16 group; *I. microcephala* has moved from the n = 16 group to being the sister taxon of *I. axillaris* Pursh. Because the sesquiterpene lactones of *I. hayesiana* polymerized before they could be identified, it has not been included on the STL cladogram.

There is greater discrepancy when one examines the higher level taxa. This lack of congruence is most easily examined by observing the placement of Jackson's three sections of *Iva*. The STL cladogram would appear to indicate a biphyletic origin for the n = 16 group (sect. Linearbractea). This would necessitate parallel evolution in aneuploid reduction. Such a phenomenon is not unusual, in fact it is a relatively common occurrence for taxa in the tribe Heliantheae (Asteraceae). In this tribe, chromosome numbers are notoriously poor indicators of relationship when used without other data (Robinson, 1981). Jackson (1960) does comment that "These species, particularly the latter three (*I. texensis, I. angustifolia* and *I. asperifolia*) are very closely related morphologically. *Iva microcephala*

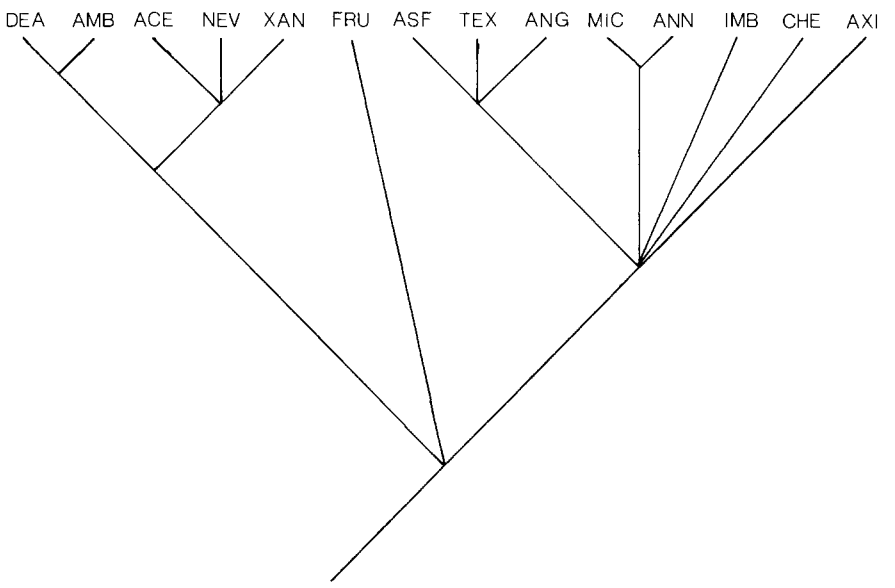


Fig. 16. An Adams Consensus Tree (Adams, 1972) constructed by finding the areas of agreement in the sesquiterpene lactone cladogram of *Iva* (Fig. 14) and Jackson's phylogenetic interpretation of *Iva* (Fig. 15B).

. . . may be considered more primitive than the other three . . . ." This would seem to indicate that Jackson considered *I. microcephala* less closely related to the other three than the three are to each other.

Inconsistencies also exist for the other two sections. For instance, in Jackson's diagram section *Cyclachaena* forms a monophyletic group (Fig. 15B), however the STL cladogram (Fig. 14) is uninformative on this question: there is no synapomorphy for the two components of the section, but neither do the components of the section share a synapomorphy with any other component of the cladogram. One reason the five species were placed together in a section by Jackson is their chromosome number, n = 18. This number is, however, the general condition in the subtribe and is therefore plesiomorphic and not an indicator of a natural group (Hennig, 1966). Another interesting feature of the STL cladogram is the fate of section *Iva*. Some of the species are grouped together, however, *I. frutescens* is very different in its location on the diagram by virtue of its unique chemistry. It is because of

Table 2. Components of the STL cladogram (Fig. 14).

No. of components	No. of terms in each component	Members
4	2	DEA/AMB ASP/TEX IMB/CHE AXI/MIC
2	3	ACE/NEV/XAN ASP/TEX/ANG
1	5	IMB/CHE/ANN/AXI/MIC
1	8	ASP/TEX/ANG/IMB/CHE/ANN/AXI/MIC

Table 3. Components of Jackson's (1960) phylogenetic interpretation (Fig. 15B).

No. of components	No. of terms in each component	Members
1	2	DEA/AMB
1	3	ASP/TEX/ANG
1	4	ASP/TEX/ANG/MIC
1	5	DEA/AMB/ACE/NEV/XAN
1	6	ASP/TEX/ANG/MIC/ANN/FRU
1	9	ASP/TEX/ANG/MIC/ANN/FRU/IMB/CHE/AXI

this unique chemistry that there is no information to indicate to what *I. frutescens* is most closely related. Therefore, based on the chemical evidence it must be viewed as potentially most closely related to any of the terminal taxa (Interpretation 2 of Nelson and Platnick, 1980).

Because of the difficulty of comparing a cladogram with a phylogenetic diagram published without detailed explanation, there is no convenient way to resolve the differences between the two schemes. When an area of discrepancy occurs between the two diagrams there is no way to determine the cause. While all the data used to construct the cladogram are available, little information other than the few comments that can be gleaned from the text are available for Jackson's phylogenetic interpretation. This is not intended to be a criticism of Jackson because the problem is prevalent in the literature. There are, however, two ways the diagrams can be compared without regard to their relative information content: Adams Consensus Trees, and Component Analysis. Computing an Adams Consensus Tree results in one diagram that represents the information they share. To quote Adams (1972):

> The consensus of two or more trees is a tree representing only that information that is shared by all trees. The consensus is a conservative estimate of a compromise classification because any information not represented in all of the rivals is not represented in the consensus.

The Adams Consensus Tree for the STL cladogram (Fig. 14) and Jackson's phylogenetic interpretation (Fig. 15B) is illustrated in Fig. 16. The diagram (Fig. 16) does not have the same branching pattern of either of the original diagrams (Figs. 14 and 15B) because of the very nature of the consensus tree; it only contains either branching points that all of the rivals agree on, or compromise branching points for areas where the rivals disagree.

The second way of examining the two diagrams (Figs. 14 and 15B) is by component analysis (Nelson, 1979). A component is a monophyletic group on the diagram. Every cladogram can be divided up into a certain number of components that have more than one terminal taxon and less than all of the terminal taxa. The STL cladogram has eight components (Fig. 14; Table 2). Jackson's diagram has six components (Fig. 15B; Table 3). There are two ways of looking at these components. If one is looking for replication in order to construct a Nelson Consensus Tree (Nelson, 1979) then there are only two replications in the components of these two figures (14 and 15B): DEA/AMB and ASP/TEX/ANG. However, when comparing cladograms to already published diagrams it is perhaps more fair to look only for areas of contradiction. Viewing it this way, there are really only two components in the Jackson diagram (Table 3) that conflict with any of the components of the STL cladogram (Table 2). Both of these components are the result of the placement of *I. microcephala*. So, while there is little replication, there is also little contradiction.

Additional information will either corroborate the cladogram or change it. When such information is available a summary cladogram for *Iva* can be constructed that will present the best estimate of the relationships among taxa of the genus.

Acknowledgments

We gratefully acknowledge the aid of the following colleagues in developing the ideas contained in this paper: Drs. J. L. Cracraft, N. H. Fischer, T. J. Mabry, G. Nelson, N. Platnick, P. M. Richardson, R. T. Schuh, and D. Young.

Literature Cited

Adams, E. N., III. 1972. Consensus techniques and the comparison of taxonomic trees. *Syst. Zool.* 21: 390–397.

Aparecida, M., H. Cagnin, C. M. R. Gomes, O. R. Gottlieb, M. C. Marx, A. I. da Rocha, M. F. das G. F. da Silva and J. A. Temperini. 1977. Biochemical systematics: Methods and principles. *Plant Syst. Evol. Suppl.* 1: 53–76.

Bohlmann, F. and C. Zdero. 1979a. Neue Sesquiterpene mit anomalen Kohlenstoffgerüst aus der Tribus Mutisieae. *Chem. Ber.* 112: 427–434.

——— and ———. 1979b. Zwei neue Eudesman-derivate aus *Iva annua*. *Phytochemistry* 18: 2034–2035.

Davis, P. H. and V. H. Heywood. 1963. *Principles of angiosperm taxonomy.* D. Van Nostrand Company, Inc., Princeton and New York.

Eldredge, N. and J. Cracraft. 1980. *Phylogenetic patterns and the evolutionary process.* Columbia University Press, New York.

Farkas, L., M. Nogradi, V. Sudarsanam and W. Herz. 1966. Constituents of *Iva* species. V. Isolation, structure and synthesis of nevadensin, a new flavone from *Iva nevadensis* M. E. Jones and *Iva acerosa* (Nutt.) Jackson. *J. Organ. Chem.* (U.S.A.) 31: 3228–3232.

Fischer, N. H., E. J. Olivier and H. D. Fischer. 1979. The biogenesis and chemistry of sesquiterpene lactones. *In:* W. Herz, H. Grisebach and G. W. Kirby (eds.), *Progress in the chemistry of organic natural products,* pp. 48–390. Springer-Verlag, New York.

Funk, V. 1982. The systematics of *Montanoa* Cerv. (Asteraceae, Heliantheae). *Mem. New York Bot. Gard.* 36: 1–131.

Geissman, T. A. 1973. The biogenesis of sesquiterpene lactones of the Compositae. *In:* V. C. Runeckles and T. J. Mabry (eds.), Terpenoids: Structure, biogenesis and distribution. *Recent advances in phytochemistry.* Vol. 6, pp. 65–69. Academic Press, New York and London.

Gomes, C. M. R. and O. R. Gottlieb. 1980. Alkaloid evolution and angiosperm systematics. *Biochem. System. Ecol.* 8: 81–87.

Gottlieb, O. R. and K. Kubitzki. 1981. Chemosystematics of *Aniba*. *Biochem. System. Ecol.* 9: 5–12.

Hennig, W. 1966. *Phylogenetic systematics.* D. Dwight Davis and Rainer Zangerl, tr. University of Illinois Press, Urbana.

Herz, W. 1977a. Biogenetic aspects of sesquiterpene lactone chemistry. *Israel J. Chem.* 16: 32–44.

———. 1977b. Sesquiterpene lactones in the Compositae. *In:* V. H. Heywood, J. B. Harborne and B. L. Turner (eds.), *The biology and chemistry of the Compositae.* Vol. 1, pp. 337–357. Academic Press, London.

———, H. Chikamatsu, N. Viswanathan and V. Sudarsanam. 1967. Constituents of *Iva* species. VIII. Structures of ivalbin, a modified guaianolide from *Iva dealbata* Gray. *J. Organ. Chem.* (U.S.A.) 32: 682–686.

———, G. Högenauer and A. Romo de Vivar. 1964a. Constituents of *Iva* species: III. Structure of microcephalin, a new sesquiterpene lactone. *J. Organ. Chem.* (U.S.A.) 29: 1700–1703.

——— and N. Viswanathan. 1964b. Constituents of *Iva* species. II. The structures of asperilin and ivasperin, two new sesquiterpene lactones. *J. Organ. Chem.* (U.S.A.) 29: 1022–1026.

Jackson, R. C. 1960. A revision of the genus *Iva* L. *Univ. Kansas Sci. Bull.* 41: 793–876.

Kelsey, R. G. and F. Shafizadeh. 1979. Sesquiterpene lactones and systematics of the genus *Artemisia*. *Phytochemistry* 18: 1591–1611.

Levy, M. 1977. Minimum biosynthetic-step indices as measures of comparative flavonoid affinity. *Syst. Bot.* 2: 89–98.

Mabry, T. J. and F. Bohlmann. 1977. Summary of the chemistry of the Compositae. *In:* V. H. Heywood, J. B. Harborne and B. L. Turner (eds.), *The biology and chemistry of the Compositae.* Vol. 2, pp. 1097–1118. Academic Press, London.

Mickevich, M. F. 1978. Taxonomic congruence. *Syst. Zool.* 27: 143–158.

Nelson, G. 1973. The higher-level phylogeny of vertebrates. *Syst. Zool.* 22: 87–91.

———. 1979. Cladistic analysis and synthesis: Principles and definitions, with a historical note on Adanson's *Familles des Plantes* (1763–1764). *Syst. Zool.* 28: 1–21.

——— and N. Platnick. 1980. Multiple branching in cladograms: Two interpretations. *Syst. Zool.* 29: 86–91.

——— and ———. 1981. *Systematics and biogeography: Cladistics and vicariance.* Columbia University Press, New York.

Payne, W. 1964. A re-evaluation of the genus *Ambrosia* (Compositae). *J. Arnold Arb.* 45: 401–438.

———. 1966. Notes on the ragweeds of South America with the description of two new species: *Ambrosia pannosa* and *A. parviflora* (Compositae). Brittonia 18: 28–37.

Platnick, N. 1979. Philosophy and the transformation of cladistics. *Syst. Zool.* 28: 537–546.

Robinson, H. 1981. A revision of the tribal and subtribal limits of the Heliantheae (Asteraceae). *Cont. Smithsonian Bot.* No. 51.

Rodrigues, A. A. S., M. Garcia and J. A. Rabi. 1978. Facile biomimetic synthesis of costunolide-1, 10-epoxide santamarin and reynosin. *Phytochemistry* 17: 953–954.

Renold, W. 1970. The chemistry and infraspecific variation of sesquiterpene lactones in *Ambrosia confertiflora* DC. (Compositae); a chemosystematic study at the populational level. Ph.D. Dissertation, University of Texas, Austin.

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

Seaman, F. 1982. Sesquiterpene lactones as taxonomic characters in the Asteraceae. *Bot. Rev.* 48: 121–194.

———, N. H. Fischer and T. F. Stuessy. 1980. Systematic implication of sesquiterpene lactones in the subtribe Melampodiinae. *Biochem. System. Ecol.* 8: 263–271.

———, G. P. Juneau, D. R. DiFeo, S. Jungk and N. H. Fischer. 1979. Repandin A, B, C, and D, four new germacranolides from *Tetragonotheca repanda* (Compositae). *J. Organ. Chem.* (U.S.A.) 44: 3400–3404.

Stuessy, T. F. 1977. Heliantheae—systematic review. *In:* V. H. Heywood, J. B. Harborne and B. L. Turner (eds.), *The biology and chemistry of the Compositae.* Vol. 2, pp. 621–697. Academic Press, London.

Urbatsch, L. E. and N. H. Fischer. Morphological and chemical features of the genus *Tetragonotheca* and their systematic implications. Manuscript in preparation.

Whewell, W. 1840. *The philosophy of the inductive sciences.* Parker, London.

Wiley, E. O. 1981. *Phylogenetics.* John Wiley and Sons, Inc., Somerset, New Jersey.

Yoshioka, H. W. Renold, N. H. Fischer, A. Higo and T. Mabry. 1970. Sesquiterpene lactones from *Ambrosia confertiflora* (Compositae). *Phytochemistry* 9: 823–832.