Amer. J. Bot. 74(10): 1577–1584. 1987.

A NUMERICAL ANALYSIS OF FLAVONOID VARIATION IN ARNICA SUBGENUS AUSTROMONTANA (ASTERACEAE)[1]

STEVEN J. WOLF AND RICHARD WHITKUS

Department of Biology, University of Missouri,
Kansas City, Missouri 64110; and Department of Botany, Ohio State University,
Columbus, Ohio 43210

ABSTRACT

Species of *Arnica* subgenus *Austromontana* produce a total of 23 leaf flavonoids, including simple and methylated flavone and flavonol glycosides as well as highly methylated flavone aglycones and a 6-hydroxylated flavone. Most of the taxa exhibit considerable interpopulational variability, with the number of compounds per population ranging from 2 to 14. Analysis of flavonoid variation in 113 populations representing all 9 species of the subgenus was carried out using cluster analysis, principal components analysis, and binary discriminant analysis. Results indicate the flavonoid profile of the very rare *A. viscosa* is the most distinctive in the subgenus. Although exhibiting considerable interpopulational variability, all populations of *A. gracilis*, a hybrid taxon, form a very distinct and cohesive group, supporting its recognition at the specific level. Additionally, chemical diversification from *A. cordifolia* has taken place largely in the Klamath region of Oregon and California. The range of variability exhibited by *A. cordifolia* is reflected in these Klamath region derivatives.

ARNICA L. subgenus *Austromontana* Maguire consists of nine montane to alpine species restricted primarily to western North America (Wolf and Denford, 1984a). Members of the subgenus are extremely polymorphic, largely due to microspecies formation via apomixis (Wolf, 1980; Wolf and Denford, 1984a). Recent studies based on morphological, cytological, geographical, and chemical analyses concluded that the two widespread species *A. cordifolia* Hook. and its derivative *A. latifolia* Bong., are ancestral species of the subgenus (Wolf and Denford, 1984b, c). From these two species, major evolutionary diversification within the subgenus has taken place largely in the Klamath region of southwestern Oregon and northwestern California (Wolf and Denford, 1984c). Four of the nine species are relatively rare and are restricted to the Klamath region. Additionally, *A. gracilis* Rydbg., an alpine species of the Rocky Mountains and Pacific Northwest, is a hybrid between *A. cordifolia* and *A. latifolia* and has been recognized at the specific level (Wolf and Denford, 1984b).

Flavonoid chemistry has been especially useful in suggesting evolutionary relationships within *Arnica* subgenus *Austromontana*. Species of the subgenus elaborate a total of 23 leaf flavonoids, including simple and methylated flavone and flavonol glycosides as well as highly methylated flavone aglycones and a 6-hydroxylated flavone (Wolf and Denford, 1983, 1984b, c). In general, the rare Klamath region endemics are characterized by reduced flavonoid profiles and/or more methylated aglycones, but the wider ranging species have fewer methylated aglycones and more glycosides (Wolf and Denford, 1984c). Additionally, most of the taxa exhibit considerable interpopulational variability, with the number of compounds per population ranging from 2 to 14. This distribution pattern is probably the result of a combination of factors including founder effect, genetic drift, and a general reduction of flavonoid profile as a result of reproductive isolation in both the rare taxa and apomicts (Wolf and Denford, 1983, 1984c).

Multivariate statistical analysis increasingly has become routine in systematic investigations. This is not surprising since the large number of computer programs available and their ease of use have made the handling of previously unwieldly data sets relatively simple. Numerical techniques are particularly valuable for analyzing variability, both at populational and taxonomic levels, as well as delimiting distinguishing characters and displaying relationships. However, the application of numerical techniques in chemosystematic investigations of flavonoids is a relatively recent phenomenon, and as Bohm, Banek, and Maze (1984) have noted, most such investigations largely have ignored populational variation. Recent studies, such as those by Wolf and Denford (1983), Parker and Maze (1984), McDougal and Parks (1984), and Bohm et al. (1984), have begun to address the significance

[1] Received for publication 15 October 1986; revision accepted 3 February 1987.

TABLE 1. *Distribution of flavonoids in* Arnica *subgenus* Austromontana

Compound	CE (7)[a]	NE (10)	GR (11)	CO (33)	DI (11)	SP (9)	VE (4)	LA (26)	VI (2)
1. A 6-Me		10[b]	7	24	10	6	4	13	2
2. A 7-Me				1	4	1	4		2
3. A 6,7-Me					6				2
4. L 4′-Me		10	9						
5. L 6-Me		6							
6. L 6-OH, 4′-Me									2
7. L 6,4-Me					3	4			2
8. L 3′,6,7-Me									2
9. Q 3-Me			7	2					
10. Q 6-Me			7					3	
11. Q 3′,6-Me									2
12. A 7-O-glu			9						
13. L 7-O-glu			11	15	3				
14. L 6-Me, 7-O-glu			7	7	2				
15. K 3-O-glu	2	6	10	14	11	6	4	9	2
16. K 3-O-gal			11					17	
17. K 6-Me, 3-O-glu	2	8	6	18	8	9	4		
18. Q 3-O-glu	5	10	6	18	9	3		12	2
19. Q 3-O-di glu	5	10	10	32	11	9	4	26	2
20. Q 3-O-gentiobioside	7	10	10	33	11	9	4	26	2
21. Q 6-Me, 3-O-glu			11					18	2
22. Flavone glycoside									2
23. Flavone glycoside									2
Total compounds	5	8	14	10	11	8	6	8	14

Abbreviations: *A. cernua* (CE), *A. cordifolia* (CO), *A. discoidea* (DI), *A. gracilis* (GR), *A. latifolia* (LA), *A. nevadensis* (NE), *A. spathulata* (SP), *A. venosa* (VE), *A. viscosa* (VI), A-apigenin, L-luteolin, Q-quercetin, K-kaempferol, glu-glucose, gal-galactose, Me-OCH_3.

[a] Number of populations surveyed.

[b] Number of populations containing given compound.

of populational variation in flavonoid profiles. Although previous investigations of the flavonoid chemistry of *Arnica* subgenus *Austromontana* have contributed greatly towards an understanding of evolutionary relationships within the group, some previous hypotheses were necessarily tentative due to small sample sizes for some taxa (Wolf and Denford, 1984c). Further analyses of a much larger number of populations has indicated a greater range of variation than previously noted and has allowed for more definitive hypotheses concerning evolutionary relationships within the subgenus. The present study is based on a sampling of 113 populations representing the entire geographical range and all nine species of *Arnica* subgenus *Austromontana* (Table 1). Due to the large number of populations sampled and the considerable variability noted, the data were examined using various clustering and ordination techniques in an effort to both confirm previous hypotheses and, perhaps, provide new insights into evolutionary relationships within the subgenus.

MATERIALS AND METHODS—To form the basic data matrix, 113 populations (OTU's) of *Arnica* subgenus *Austromontana* were scored for the presence or absence of 23 flavonoid compounds (Table 1). Descriptions of population sampling techniques, flavonoid extraction and identification are given in Wolf and Denford (1983, 1984c). A copy of the data matrix is available upon request.

Variation in flavonoid profiles both within and among species was first examined with cluster analysis of the OTU's to see whether populations of the taxa formed cohesive groups. Similarities among OTU's were calculated by the simple matching coefficient of Sokal and Michener (1958):

$$A + D/A + B + C + D,$$

where A is the number of characters whose presence is shared by two OTU's, B and C the number of characters present in one OTU but absent in the other, and D the number of characters whose absence is shared by two OTU's. This coefficient gives equal weight to the shared presence and absence of characters. Because evolutionary diversification in the subgenus has been accompanied by a reduction in flavonoid variation (Wolf and Denford, 1984c), shared absence of compounds is as important as shared presence. Phenograms were produced with average linkage clustering (UPGMA), which gen-

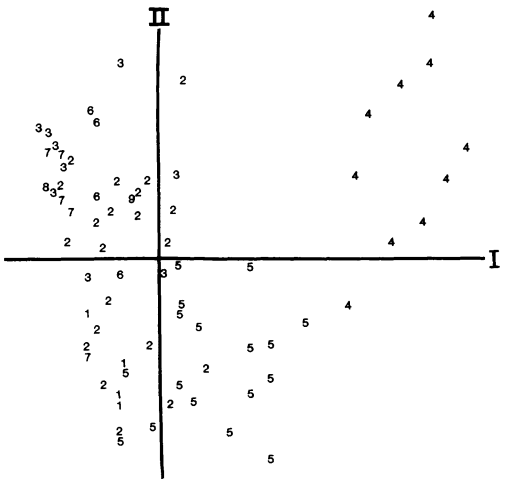

Fig. 2. Projection of 113 populations of *Arnica* subgenus *Austromontana* onto first two principal components. Axes I and II represent 38.6% of the total variation. *A. cernua* (1), *A. cordifolia* (2), *A. discoidea* (3), *A. gracilis* (4), *A. latifolia* (5), *A. nevadensis* (6), *A. spatulata* (7), *A. venosa* (8), and *A. viscosa* (9).

erally gives the least amount of distortion of a similarity matrix (Rohlf, 1970; Sneath and Sokal, 1973).

Clustering methods tend to faithfully represent the similarity matrix at lower levels of clustering, but distort the matrix at higher clustering levels (Sneath and Sokal, 1973). To better represent phenetic relationships at higher levels, ordination techniques are considered more appropriate (Sneath and Sokal, 1973). Phenetic relationships among the taxa, therefore, were explored with principal component analysis (PCA) of the covariance matrix of the data. Results of such an analysis are identical to a principal coordinate analysis of presence-absence data (Gower, 1966; Neff and Marcus, 1980). However, an additional benefit accrues from a PCA in being able to interpret the principal axes in terms of character combinations, a feature absent in principle coordinate analysis. Although a PCA indicates which characters explain the majority of the overall variation among OTU's, it does not separate between taxa vs. within taxa variation since no a priori grouping of OTU's is done. To find those characters which maximally separate taxa, discriminant analysis is appropriate on pre-chosen groups of OTU's (i.e., taxa). Because of various assumptions about the data,

←

Fig. 1. Phenogram for 113 populations of *Arnica* subgenus *Austromontana*. Taxa abbreviations follow Table 1.

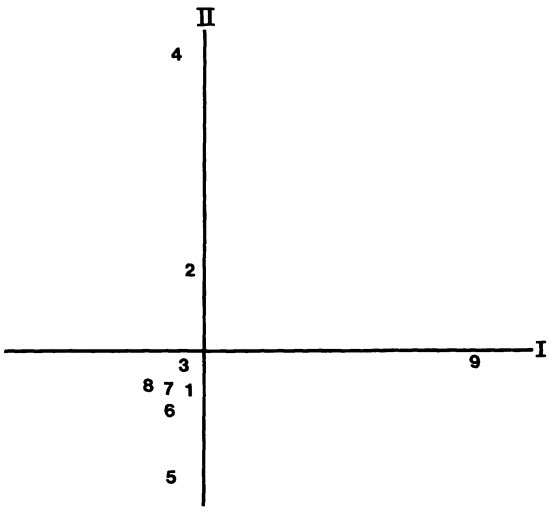

Fig. 3. Projection of the nine species of *Arnica* subgenus *Austromontana* onto the first two axes of binary discriminant analysis. Numbers follow Fig. 2.

discriminant analysis is performed with continuous variables. However, binary discriminant analysis (BDA, Strahler, 1978) is a method for finding combinations of binary variables which are most important for discriminating groups. Thus, to identify those combinations of compounds which maximally separate the taxa, a BDA of the taxa was performed using the method of Strahler (1978).

Cluster analysis was performed with the NTSYS program package (Rohlf, Kishpaugh, and Kirk, 1972), PCA and BDA with the BMDP program package (Dixon, 1981). All calculations were performed on the IBM 3081 at Ohio State University.

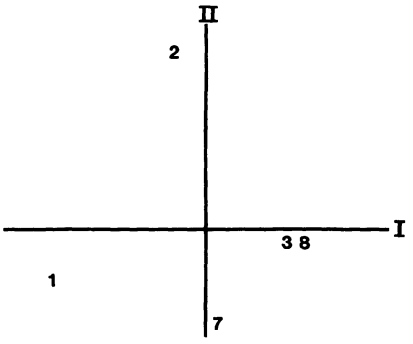

Fig. 5. Projection of *A. cordifolia* and its Klamath derivatives onto the first two axes of binary discriminant analysis. Numbers follow Fig. 2.

To explore different questions of relationships within the subgenus, three sets of analyses were performed. First, all taxa were analyzed simultaneously (113 OTU's) to identify overall patterns of relationships within the subgenus. A second analysis explored the relationships among *A. cordifolia* and its presumed Klamath region derivatives (64 OTU's). Finally, an analysis of *A. cordifolia, A. latifolia* and their hybrid derivative *A. gracilis,* was undertaken (94 OTU's) to see if the chemical data support the recognition of the latter at the specific level.

RESULTS—The phenogram for the analysis of the subgenus (Fig. 1) indicates several clusters comprised of OTU's of a single taxon. Most notable among these are *A. gracilis* and *A. vis-*

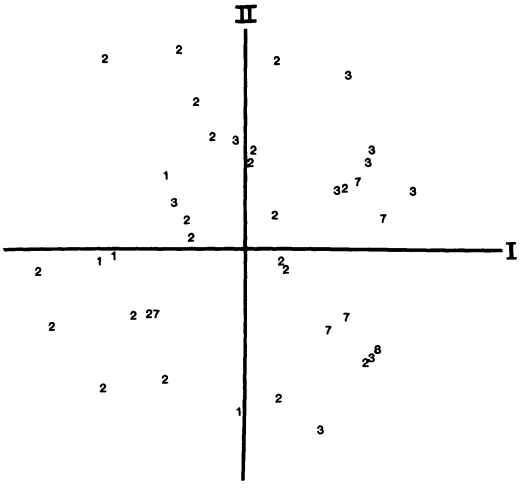

Fig. 4. Projection of 64 populations of *A. cordifolia* and its Klamath derivatives onto the first two principal components. Axes I and II represent 46.1% of the total variation. Numbers follow Fig. 2.

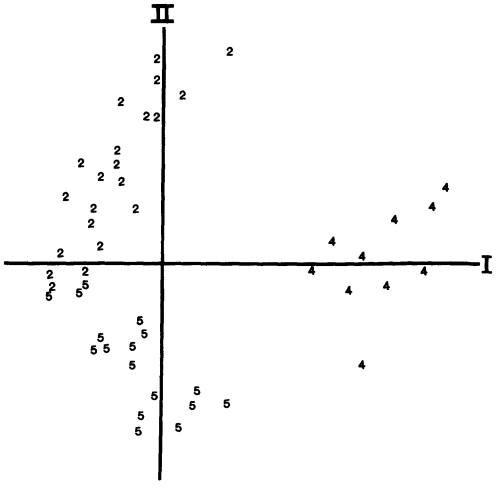

Fig. 6. Projection of 94 populations of *A. cordifolia, A. latifolia,* and *A. gracilis* onto the first two principal components. Axes I and II represent 47.7% of the total variation. Numbers follow Fig. 2.

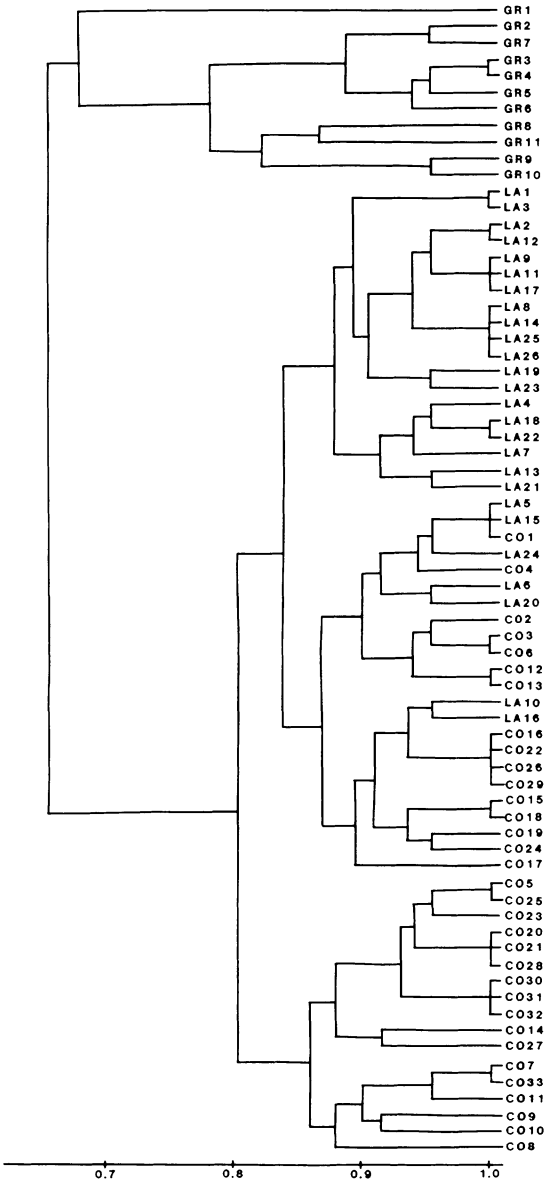

Fig. 7. Phenogram of 94 populations of *A. cordifolia*, *A. latifolia* and *A. gracilis*. Abbreviations follow Table 1.

cosa which are two of the most distinct clusters. Other evident clusters are those comprised of *A. latifolia* and *A. nevadensis*. Shared flavonoid profiles in the remaining OTU's results in their intermixing throughout the remainder of the phenogram.

The overall pattern seen in the phenogram is also exhibited by the PCA (Fig. 2). Again, *A. gracilis* is the only taxon which forms a distinct group on the first two axes. Although *A. viscosa* does not separate out on the first two axes, it does so on the third (not shown),

TABLE 2. *Sorted factor loadings greater than 70% for the first three factors of the BDA of* Arnica *subgenus* Austromontana. *The first three factors account for 76.12% of the total variation*

Character	Factor 1	Factor 2	Factor 3
6	0.987	0.0	0.0
8	0.987	0.0	0.0
11	0.987	0.0	0.0
22	0.987	0.0	0.0
23	0.987	0.0	0.0
14	0.0	0.987	0.0
13	0.0	0.982	0.0
9	0.0	0.905	0.395
12	0.0	0.835	0.499
16	0.0	0.263	0.936
21	0.0	0.0	0.932
10	0.0	0.701	0.703

which accounts for an additional 11.9% of the total variation.

In the BDA (Fig. 3), *A. viscosa* separates from the remaining taxa along the first axis on the basis of compounds 6, 8, 11, 22 and 23 (Table 2), all of which are unique to this taxon. The remaining taxa spread out along axis 2, and to a lesser extent on axis 3 (not shown).

Both the PCA (Fig. 4) and the phenogram (not shown) for the analysis of *A. cordifolia* and its presumed Klamath region derivatives indicate a lack of distinct clusters, indicating a shared overall variational pattern and a lack of unique flavonoid profiles for any one taxon. Further evidence for this pattern is shown in the results of the BDA (Fig. 5) where all the taxa are essentially distinct from one another. Factor loadings (Table 3) indicate that only compound 13 (present in *A. cordifolia* and *A. discoidea*) can clearly discriminate along an axis, while the remaining characters have loadings on two or more axes.

Analysis of *A. latifolia, A. cordifolia* and their hybrid derivative indicates a unique pattern in *A. gracilis*. In the PCA (Fig. 6), *A. gracilis* is separated from its parents on the first axis. The

TABLE 3. *Sorted factor loadings greater than 70% for the first three factors of the BDA of* A. cordifolia *and its Klamath derivatives. The first three factors account for 90.99% of the total variation*

Character	Factor 1	Factor 2	Factor 3
1	0.928	0.334	0.0
15	0.811	−0.464	0.333
17	0.763	−0.525	0.0
2	0.741	−0.379	0.0
13	0.0	0.978	0.0
7	0.368	−0.758	0.0
18	−0.299	0.0	0.940
3	0.470	−0.253	0.842

TABLE 4. *Sorted factor loadings greater than 70% for the two factors of the BDA of* A. cordifolia, A. latifolia, *and* A. gracilis. *These two factors account for all the variation among the three taxa*

Character	Factor 1	Factor 2
14	0.994	0.0
15	0.977	0.0
9	0.959	0.285
13	0.952	−0.306
12	0.907	0.421
4	0.907	0.421
10	0.754	0.656
21	0.0	0.982
16	0.0	0.975
17	0.633	−0.774

compounds which contribute high loadings to the first axis are 16 and 21 (shared by *A. gracilis* and *A. latifolia*), 4 and 12 (unique to *A. gracilis*) and 13 (shared by *A. gracilis* and *A. cordifolia*). *Arnica latifolia* and *A. cordifolia* separate along axis II which has compounds 13 and 17 (found in *A. cordifolia*), and 16 and 21 (found in *A. latifolia*) as those which contribute high loadings. Interestingly, *A. gracilis* falls between these two taxa on axis II, a result of its additive flavonoid profile in these compounds. Like the phenogram of Fig. 1, Fig. 7 shows *A. gracilis* distinct from the parental species.

In the BDA of *A. cordifolia, A. latifolia,* and *A. gracilis* (ordination not shown), compounds of the PCA are seen as a subset of compounds which separate the three taxa (Table 4). As in the PCA, the three taxa are spread along the first axis with *A. gracilis* more displaced than the others. Seven compounds have loadings greater than 75% on the first axis (Table 4), with compounds 4, 12, and 13 also used in the first axis of the PCA. The second axis separates *A. cordifolia* and *A. latifolia* with compounds 16 and 21 (found in *A. latifolia*) and 17 (found in *A. cordifolia*). Again, these compounds contribute high loadings for the PCA.

DISCUSSION—Results of the present study support previous hypotheses concerning evolutionary relationships within *Arnica* subgenus *Austromontana.* Even though most of the species of the subgenus exhibit considerable variation with respect to flavonoid profiles, most of this variation is systematically significant. Previous investigations have suggested that *A. cordifolia* is probably the ancestral species of the subgenus (Maguire, 1943; Wolf and Denford, 1984a, c). It is therefore not surprising that this species forms several clusters with more than one species (Fig. 1), its range of variation encompasses most of the species in the subgenus (Fig. 2), and it takes a central position in the discriminant analysis (Fig. 3). It previously has been suggested that *A. cordifolia* gave rise to *A. discoidea* Benth. via diploid populations in the Klamath region (Wolf and Denford, 1984c). Not only is this hypothesis supported by morphological and cytological evidence, but also by chemical evidence. Three relictual Klamath region diploid populations of *A. discoidea* (DI2, DI3 and DI4), largely an apomictic polyploid complex, cluster with six northwestern populations of *A. cordifolia* (Fig. 1), mostly on the basis of compounds 13 and 14. Additionally, several other populations of these two species form distinct clusters in the phenogram.

Further evolutionary diversification from *A. discoidea* has taken place in the Klamath region giving rise to two rare Klamath endemics, *A. spathulata* Greene and *A. venosa* H. M. Hall (Wolf and Denford, 1984a, c). All four populations of *A. venosa* cluster with *A. discoidea* and most of the populations of *A. spathulata* cluster with the latter (Fig. 1). The three taxa also are in close proximity in the BDA of the subgenus (Fig. 3) and of the Klamath region taxa (Fig. 5). Since *A. cordifolia* is ancestral to all three species, it is not surprising that it clusters with additional populations of each in Fig. 1 and the range of variation in *A. cordifolia* encompasses all these Klamath derivatives in the PCA (Fig. 4).

Based on morphological, cytological, geographical, and flavonoid analyses, it previously has been suggested that *A. nevadensis* A. Gray has been derived from *A. cordifolia* (Maguire, 1943; Wolf and Denford, 1984a, c). However, the previous flavonoid data, based on only two population samples, was inconclusive (Wolf and Denford, 1984c). Further populational sampling has confirmed the close phenetic relationship between these two species. All ten populations of *A. nevadensis* form a very cohesive group and cluster with populations of *A. cordifolia* (Fig. 1).

Previous studies have demonstrated that *A. gracilis* is a hybrid between *A. cordifolia* and *A. latifolia* (Wolf and Denford, 1984b). However, it is morphologically, ecologically, and reproductively distinct and has been recognized at the specific level (Wolf and Denford, 1984b). Results of the present investigation also demonstrate that this species is chemically distinct as well. In both phenograms (Fig. 1, Fig. 7) all populations of *A. gracilis* form a very cohesive group. Additionally, in both PCA's (Fig. 2, Fig. 6) as well as the BDA (not shown), *A. gracilis* is quite distinct from both its presumed parents. These results support the pre-

vious recognition of *A. gracilis* at the specific level (Wolf and Denford, 1984b).

Arnica viscosa A. Gray is the rarest and most morphologically distinctive species in *Arnica,* and probably the most recently evolved species in subgenus *Austromontana* (Wolf and Denford, 1984a, c). It is known from only a few populations on very recent volcanic substrates, at high elevations, largely in the Klamath region. In addition to its unique morphology, this species also has a very distinctive flavonoid profile as evidenced by the cluster analysis (Fig. 1), PCA (Fig. 2, third axis not shown), and BDA of the subgenus (Fig 3). Straley (1980) erected a new subgenus in which he placed *A. viscosa* and *A. venosa.* However, based on morphological, ecological, and chemical analyses, as well as Straley's (1980) own hybridization results, this treatment was rejected by Wolf and Denford (1984a, c). Results of the present study also indicate the artificiality of a new subgenus for these two species. All analyses of the subgenus demonstrate the close phenetic relationship between *A. venosa* and *A. discoidea,* which contrasts considerably with the placement of *A. viscosa.*

As previously noted, some taxa, as well as some populations of *Arnica* subgenus *Austromontana* exhibit very reduced or depleted flavonoid profiles (Wolf and Denford, 1983, 1984c). For example, the flavonoid profile of the rare Klamath endemic *A. cernua* Howell, which consists of only five compounds, is a small subset of its presumed parental species *A. cordifolia* (Wolf and Denford, 1984a). However, since there are as few as two compounds per population in this species, little affinity with the latter is reflected in the numerical analyses. In general, most of the clusters in Fig. 1 above the large *A. latifolia* cluster consist of populations which have very few compounds and, to an extent, cluster on the basis of shared absences of compounds. However, the majority of these clusters do in fact represent the same taxa or closely related taxa. For example, most of the populations of *A. cordifolia, A. cernua,* and *A. latifolia* cluster together. Likewise, the three populations of *A. spatulata* which cluster at the top of Fig. 1 have very reduced flavonoid profiles (i.e., three compounds each), and therefore they do not cluster below with the remaining populations from the Klamath region.

Even though the species of *Arnica* subgenus *Austromontana* exhibit considerable populational variation with respect to flavonoid profiles, most of this variation is systematically significant and confirms previous hypotheses concerning relationships within the subgenus *Arnica cordifolia,* which clusters with several of the species at various levels, appears to be the ancestral species of the subgenus as previously hypothesized (Wolf and Denford, 1984a). The results also support the hypothesis that this species probably has given rise to *A. nevadensis, A. discoidea* (in the Klamath region), and *A. gracilis* (via hybridization with *A. latifolia*). *Arnica discoidea* in turn probably has given rise to the two Klamath region endemics *A. spathulata* and *A. venosa.* The present study also refutes Straley's hypothesis that *A. venosa* and *A. viscosa* are closely related. In addition to morphological differences, the latter's flavonoid profile is the most distinctive within the subgenus. The remaining, nonsystematically significant variation in flavonoid profiles and resulting clusters are likely the result of a combination of factors, including founder effect, genetic drift, and a general reduction of flavonoid profile as a result of reproductive isolation in both the rare taxa and apomicts.

Literature Cited

Bohm, B., H. M. Banek, and J. R. Maze. 1984. Flavonoid variation in North American *Menziesia* (Ericaceae). Syst. Bot. 9: 324–345.

Dixon, W. J. [ed.] 1981. BMDP statistical software. University of California Press, Berkeley.

Gower, J. C. 1966. Some distance properties of latent root and vector methods for systematics. Biometrika 53: 325–338.

Maguire, B. 1943. A monograph of the genus *Arnica.* Brittonia 4: 386–510.

McDougal, K. M., and C. R. Parks. 1984. Elevational variation in foliar flavonoids of *Quercus rubra* L. (Fagaceae). Amer. J. Bot. 71: 301–308.

Neff, N. A., and L. F. Marcus. 1980. A survey of multivariate methods for systematics. New York: privately published.

Parker, W. H., and J. M. Maze. 1984. Intraspecific variation in *Abies lasiocarpa* from British Columbia and Washington. Amer. J. Bot. 71: 1051–1059.

Rohlf, F. J. 1970. Adaptive hierarchial clustering schemes. Syst. Zool. 19: 58–82.

———, J. Kishpaugh, and D. Kirk. 1972. NT-SYS. Numerical taxonomy system of multivariate statistical procedures. State University of New York, Stony Brook.

Sneath, P. H. A., and R. R. Sokal. 1973. Numerical taxonomy. W. H. Freeman, San Francisco.

Sokal, R. R., and C. D. Michener. 1958. A statistical method for evaluating systematic relationships. Univ. Kansas Sci. Bull. 38: 1409–1438.

Strahler, A. H. 1978. Binary discriminant analysis: a new method for investigating species-environment relationships. Ecology 59: 108–116.

Straley, G. B. 1980. Systematics of *Arnica,* subgenus *Austromontana* and a new subgenus *Calarnica* (Asteraceae: Senecioneae). Ph.D. dissertation, University of British Columbia, Vancouver.

Wolf, S. J. 1980. Cytogeographical studies in the genus

Arnica (Compositae: Senecioneae). I. Amer. J. Bot. 67: 300–308.

———, AND K. E. DENFORD. 1983. Flavonoid variation in *Arnica cordifolia*: an apomictic polyploid complex. Biochem. Syst. Ecol. 11: 111–114.

———, AND ———. 1984a. Taxonomy of *Arnica* (Compositae) subgenus *Austromontana*. Rhodora 86: 239–309.

———, AND ———. 1984b. *Arnica gracilis* (Compositae), a natural hybrid between *A. latifolia* and *A. cordifolia*. Syst. Bot. 9: 12–16.

———, AND ———. 1984c. Flavonoid diversity and endemism in *Arnica* subgenus *Austromontana*. Biochem. Syst. Ecol. 12: 183–188.