A COMPARISON OF MULTIVARIATE METHODS FOR THE DETECTION OF HYBRIDIZATION

Robert P. Adams[1]

Summary

Principal component analysis (PCA), Wells' hybrid distance diagram, principal coordinate analysis (PCO) using weighted Gower metrics, and canonical variate analysis (CVA) were compared using data from artificial crosses of sunfish (*Lepomis*) and putative hybridization in *Juniperus*. The results were similar between the two data sets in that using the weighted characters (F ratios from ANOVA of the parents) produced well resolved parental groups and generally ordinated the hybrids between the parents. The F weighted PCO and Wells' distance diagrams were found to be better suited for the analysis of hybridization than PCA or CVA. Several limitations of PCA and CVA are discussed.

Introduction

The detection and analyses of gene flow is one of the most common problems in systematics. Almost without fail, monographic theses address the question of natural hybridization of the taxa studied. Perhaps it is due to the weight given to hybridization (or crossability) in some species concepts. In chemosystematics, much of this research on hybridization stems from an interest in understanding the genetic control of compounds.

The identification of hybrids continues to present methodological problems. Edgar Anderson (1949) argued that introgressive hybridization might defy any statistical detection and yet assume great biological significance. Questions concerning circular resasoning have recently been raised in the a priori treatment of the unknowns as hybrids (Wells, 1980) and in the treatment of the unknowns as a third a priori group for canonical variate analysis (Neff and Smith, 1979). Maze (1980) criticized Wells' hybrid distance coefficient presentation (Wells, 1980) on the basis that "A competent systematist would never use a one-dimensional representation (a hybrid index) to depict the relationship between three entities." Thus, Maze assumed that an a priori decision would be made concerning the number of taxa involved before a method of analysis were chosen. The recently presented hybrid distance diagram (Wells, 1980) does allow for the unknowns to be either hybrids or a third taxon, however, a priori recognition of parental taxa is required. Neff and Smith (1979) have suggested that a principal component analysis (PCA) of all individuals might resolve the problem by revealing whether the unknowns are a third taxon or hybrids. Although PCA (using a pooled correlation or variance/covariance matrix) is useful for the first iteration of analyses, it is not optimal for analysis of hybridization once groups (ex. parents, unknowns) have been discovered (Kistler, 1976). The reason for this is that each character has the same a priori weight and inter-group distances are not taken into account. Namkoong (1966) noted this with PCA, Anderson's hybrid index, and canonical correlation. The inclusion of non-diagnostic characters has been a serious

Table 1. Characters and F ratios from ANOVA between *L. cyanellus* (20) and *L. macrochirus* (20) and ANOVA between *J. horizontalis* (10) and *J. scopulorum* (15).

Char.	F ratio	Char.	F ratio
Lepomis cyanellus/macrochirus			
1. std. length	3.51	10. l. D I	14.61
2. l. head	33.31	11. l. A III	219.63
3. d. caud. ped	9.65	12. l. supramaxilla	247.34
4. P_2 to D	0.48	13. no. lat. line sc.	38.79
5. l. P_1	34.9	14. no. D spines	0.09
6. l. lower jaw	56.50	15. no. D rays	31.38
7. l. upper jaw	96.04	16. no. A rays	133.39
8. l. D V	478.34	17. no. palatine teeth	24.58
9. l. D III	191.47		
Juniperus horizontalis/scopulorum			
1. α-pinene/α-thujene	38.48	16. δ-cadinene	243.63
2. sabinene	3.67	17. (carane hydrate)	37.48
3. α-phellandrene	111.90	18. sesquiterpene	133.65
4. myrcene	805.03	19. sesquiterpene	59.09
5. α-terpinene	15.21	20. sesquiterpene	65.17
6. β-phellandrene	2.84	21. sesquiterpene	96.50
7. γ-terpinene	10.16	22. elemene alcohol	125.45
8. p-cymene	10.14	23. elemol	60.45
9. linalool	11.46	24. γ-eudesmol	83.35
10. iso-safrole	251.31	25. cadinol isomer	157.02
11. 4-terpineol	27.97	26. α-cadinol	190.35
12. C_{10}-alcohol	120.96	27. α-eudesmol/cadinol isomer	53.89
13. borneol	8.85	28. β-eudesmol	117.60
14. piperitone	85.47	29. acetate II	116.45
15. citronellol	21.64	30. % oil yield	33.35

problem with Anderson's hybrid index and is a limitation for the use of the Wells' hybrid distance (Maze, 1980).

In this paper I would like to expand the Wells' hybrid distance diagram to include character weights (Adams, 1975) and compare it with PCA, principal coordinate analysis (PCO) using weighted characters (Whiffin, 1977), and canonical variate analysis (CVA) for the detection of hybrids. The first of two examples involves additional analysis of artificial hybrids between two species of sunfish (*Lepomis*) initially analyzed by Neff and Smith (1979) and the second example is a case of putative natural hybridization between *Juniperus scopulorum* and *J. horizontalis* (Fassett, 1945; von Rudloff, 1975).

Materials and Methods

Twenty *Lepomis macrochirus* and 20 *L. cyanellus* were collected from a pond at the University of Michigan Botanical Gardens (Neff and Smith, 1979). Of the 40 hybrids previously studied (Neff and Smith, 1979), 10 of each reciprocal cross were randomly chosen giving a set of 20 *L. cyanellus*, 20 *L. macrochirus* and 20 hybrids. The 17 characters used are listed in Table 1 (see also Table 2 of Neff and Smith, 1979). The first 12 characters involved length measurements and the last 5 characters were meristic.

The juniper samples were taken in conjunction with a larger study of *Juniperus*

scopulorum (Adams, 1983) and were as follows: 15 samples of *J. scopulorum* from Durango, Colorado; 10 samples of *J. horizontalis* from the Sasketchewan River bank, Saskatoon, Sask.; 15 samples from Banff, Alb.; 5 samples from 14 mi. S.E. of Bridger, Montana; 3 samples from the Burning Coal Seam, Amidon, North Dakota, and 5 samples of presumably *J. horizontalis* from Grass Range, Montana. Some explanations should be made concerning the juniper collections. The 3 plants from Amidon were collected in addition to the normal population sample. Plant 1910 was about 90 cm high and 300 cm across and appeared morphologically intermediate between *J. horizontalis* and *J. scopulorum*. Plant 1911 was a "sport" or mutant looking tree with foliage in tight balls like a bonsai tree. Plant 1912 was prostrate, about 8 cm tall and 300 cm across and appeared to be typical of *J. horizontalis*. All of the 5 plants from Grass Range, Montana were prostrate (8 to 15 cm tall) and appeared to be morphologically typical of *J. horizontalis*. The 5 plants from south of Bridger, Montana were taken as the most intermediate appearing plants of that population. The plants at Banff ranged from 8 times as wide as tall, to taller than wide.

The foliage from these samples was frozen until voucher specimens and materials for steam distillation were processed. Specimens are on deposit at SRCG. The volatile terpenoids were removed by steam distillation and analyzed by capillary gas/liquid chromatography (see Adams, 1975, for details). Peak identification (Table 1) was based on the recent mass spectral computer search report for these taxa (Adams, von Rudloff, Zanoni, and Hogge, 1981). In a couple of instances, non-resolvable compounds were treated as one (a-pinene/a-thujene; B-eudesmol/cadinol isomer). Analysis of variance (ANOVA) was run on the *J. scopulorum* from Colorado and the *J. horizontalis* from Sasketchewan to eliminate some of the terpenoids with small F ratios that failed to separate the taxa. Thirty compounds with the largest F ratios were then chosen for subsequent analyses.

Principal component analysis follows the programs of Blackrith and Reyment (1971) and Veldman (1967). Principal coordinate analysis (PCO) follows Gower (1966) using mean character differences (MCD or Manhattan metric), scaled by the range (=Gower metric, Gower, 1971) and weighted by F-1 (F ratio from ANOVA of the parental groups as previously formulated, Adams, 1972). Table 1 shows F ratios from ANOVA between *L. cyanellus* and *L. macrochirus* and F ratios for *J. scopulorum* and *J. horizontalis*. Canonical variate analysis (CVA) follows Blackrith and Reyment (1971), Cooley and Lohnes (1971) and Pimentel (1979). Hybrid distance diagrams were based on the original formulation of Wells (1980) and modified as follows for character weighting:

$$^d Ah = \left(\frac{\sum_i^n W_i(A_i - Y_{ih})^2}{\sum_i^n W_i} \right)^{\frac{1}{2}}$$

$$^d Bh = \left(\frac{\sum_i^n W_i(B_i - Y_{ih})^2}{\sum_i^n W_i} \right)^{\frac{1}{2}}$$

$$^d AB = \left(\frac{\sum_i^n W_i(A_i - B_i)^2}{\sum_i^n} \right)^{\frac{1}{2}}$$

where $^d Ah$ = distance from reference point A to individual h; $^d Bh$ = distance from reference point B to individual h; $^d AB$ = distance between extreme reference points A and B. W_i =
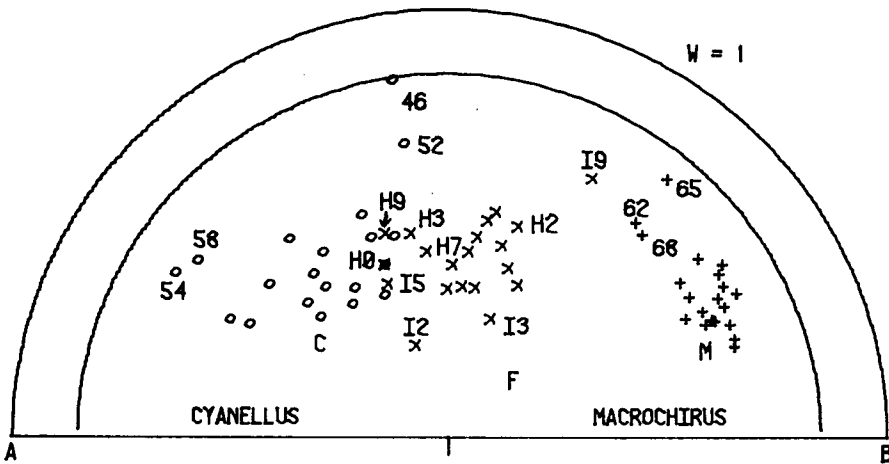
Fig. 1. Wells' distance diagram for *L. cyanellus* (circles), *L. macrochirus* (pluses) and hybrids (X's) using equally weighted character matches (W = 1). The symbols C, M, and F show the position of the average distances for *L. cyanellus*, *L. macrochirus,* and a hypothetical intermediate individual. Some of the outlying hybrids are designated by an H or I number. H designated hybrids have *L. macrochirus* as the maternal parent; I designated hybrids have *L. cyanellus* as the maternal parent. No maternal inheritance is evident, but the hybrids do tend to intergrade with *L. cyanellus*.

weight for character: 1.0 for equal weights or (F ratio − 1.0) for character weighting; A_i, B_b and Y_{ih} are as defined by Wells (1979).

In addition to the mean coordinate points for the reference taxa, a third reference point has been computed as the average for the two parental taxa, which is an OTU with character values exactly intermediate to each of the parental groups. This hypothetical intermediate is designated as "F" on the plots and is useful to see how hybrids or putative hybrids depart from true intermediacy between the parents.

Results and Discussion

Lepomis

Wells' hybrid distance diagram (Fig. 1) shows *L. macrochirus* forming a well defined group with *L. cyanellus* forming a rather diffuse group. The hybrids are somewhat intermediate but merging into *L. cyanellus*. A couple of *L. cyanellus* individuals (#46, #52) appear somewhat separated. Hybrid I9 shows close affinities with *L. macrochirus*. The overall pattern is similar to the one obtained from PCA (Fig. 2) where the *L. macrochirus* is fairly well resolved and the relation of the hybrids to the parents is very similar. It is interesting to note that these two methods are very different in their theoretical basis and yet had similar results in this case. PCA does not assume any a priori knowledge of groups, while the hybrid distance diagram does assume knowledge of the parental groups. If one allows that the a priori identification of parental (or putative parental) taxa is possible, then more powerful methods may be used which optimize discrimination between the putative parental taxa. This in fact, is not an invalid assumption in systematics because taxonomists generally can identify the taxa they are studying. Obtaining samples of reference individuals for each taxon can present a problem for several reasons.
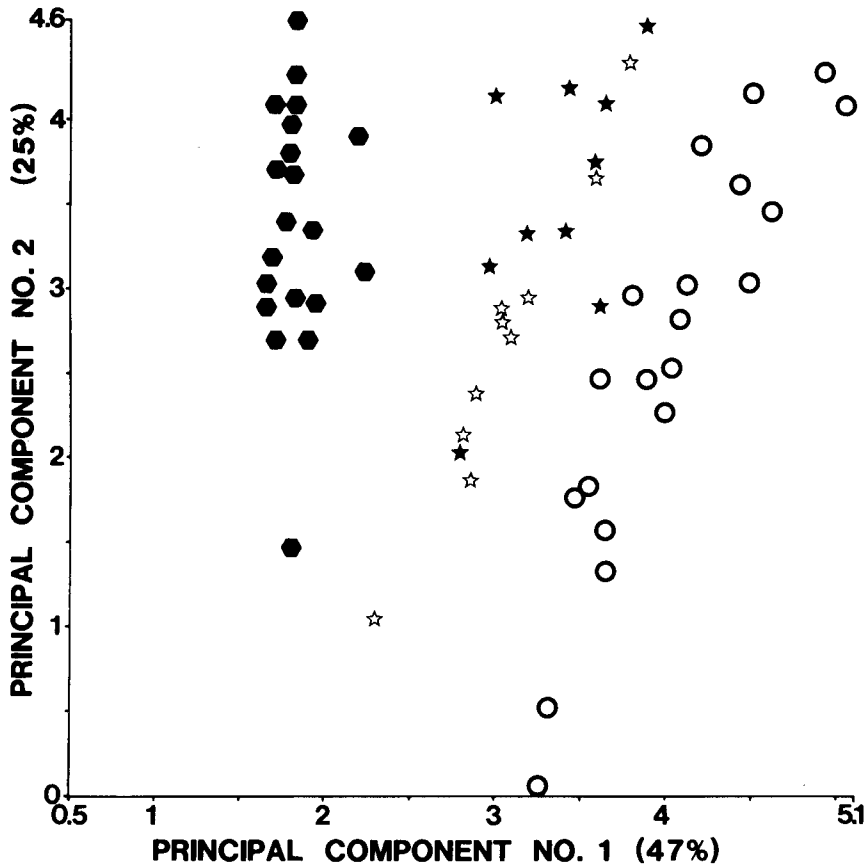
Fig. 2. Principal component analysis of the same individuals as used in Fig. 1 (open circles are *L. cyanellus,* solid hexagons are *L. macrochirus,* open stars are hybrids with *L. cyanellus* as the maternal parent, solid stars are hybrids with *L. macrochirus* as the maternal parent). Note the intergradation of the hybrids with *L. cyanellus* as seen with the unweighted hybrid distances in Fig. 1.

Individuals from sympatric populations may be introgressed leading to a biased or "contaminated" sample. On the other hand individuals from allopatric and perhaps distant populations may be quite differentiated from the actual parents involved in hybridization. Ideally one would like to examine intraspecific variation before selecting a reference population but this is not always feasible.

The modification of Wells' original formulation to include character weights permits the effective use of a priori knowledge about the parental taxa and the characters differentiating them. Figure 3 shows the results of character weighting by F-1 (F from ANOVA between *L. cyanellus* and *L. macrochirus*). Three groupings are now apparent. *Lepomis cyanellus* now clusters much more tightly than when equal weights were used (cf. Fig. 1) and is resolved from the hybrids. Individual hybrid I9 is still some distance from most of the hybrids. Individual hybrid H2 is now moved closer to *L. macrochirus* and individual 65 of *L. macrochirus* has assumed a more intermediate position (toward the hybrids). The latter could suggest gene flow in the sympatric population from which it was sampled. A reexamination of the
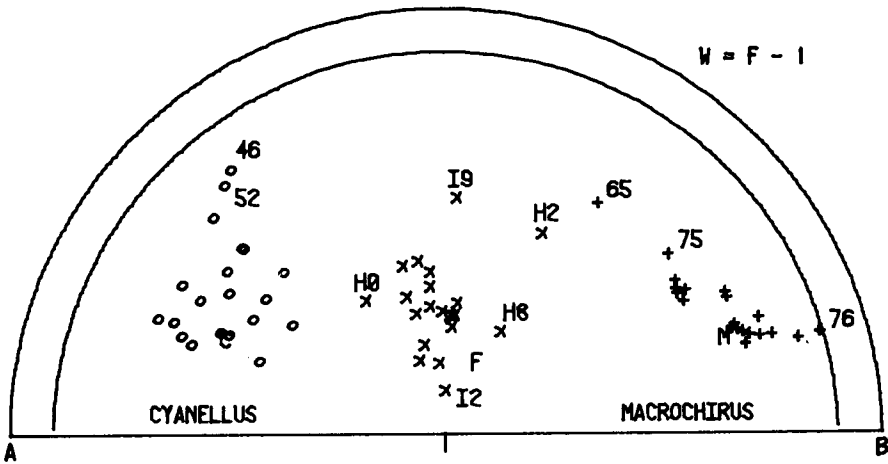
Fig. 3. Wells' hybrid distance diagram using character match weights of F-1 (F from AN-OVA of the two parental groups). Circles are *L. cyanellus,* pluses are *L. macrochirus,* and X's are hybrids as in Fig. 1. The hybrids are now well resolved and the parents form tighter clusters than with unweighted character matches (Fig. 1). Hybrid individual H2 and *L. macrochirus* individual 65 are shown to be more similar than in previous analyses. See text for additional discussion.

average values for parents and hybrids (Neff and Smith, 1979, Table 1) showed that characters 4, 5, 10, and 14 were transgressive to one of the parent means. The average F ratio for these four characters is 12.52, whereas the average F for the remaining 13 characters is 120.30. Thus, it is apparent that the unweighted hybrid distance and PCA were affected by the considerable "noise" in some of the characters as well as the transgressive nature of some of the characters. It should be pointed out however, that the fact that the transgressive characters were not as discriminatory as the other characters may not hold in general. Whiffin (1977) has shown dramatic effects of character weighting (by F ratios) on clustering and principal coordinate analysis (PCO) in an excellent study of hybridization in *Correa.* Similar results were obtained by Kistler (1976) on both real and simulated data. The effects of character weighting by F (and various combinations of F) have also been examined for classification at the specific level (Adams, 1975) and found to be very useful.

Principal coordinate analysis (PCO) has generally been lumped under PCA in most texts or added as kind of an afterthought. Since the measure of association in PCO can be anything, one can avoid the objections of the use of the correlation coefficient as a measure of taxonomic similarity (Eades, 1965). In addition, one can easily allow for none, some, or even complete knowledge of a priori groups in the analysis. PCA was originally designed for analyses of a single group from one population. Since the axes are in the direction of greatest overall variance (among individuals), groups will only be apparent if the distances among groups are making a large contribution to the overall variance. Of course, a PCO using unweighted characters and, for example, Euclidean distances suffers the same limitations as PCA. The problem associated with PCA and unweighted PCO seem to have been generally ignored by most numerical systematists.

Comparison of PCO (similarities = F-1 weighted, Gower metric) of the *Lepomis* data (Fig. 4) with the weighted hybrid distance method (Fig. 3) show considerable correspondence. The ability of PCO to extract a second axis for the hybrids (or a
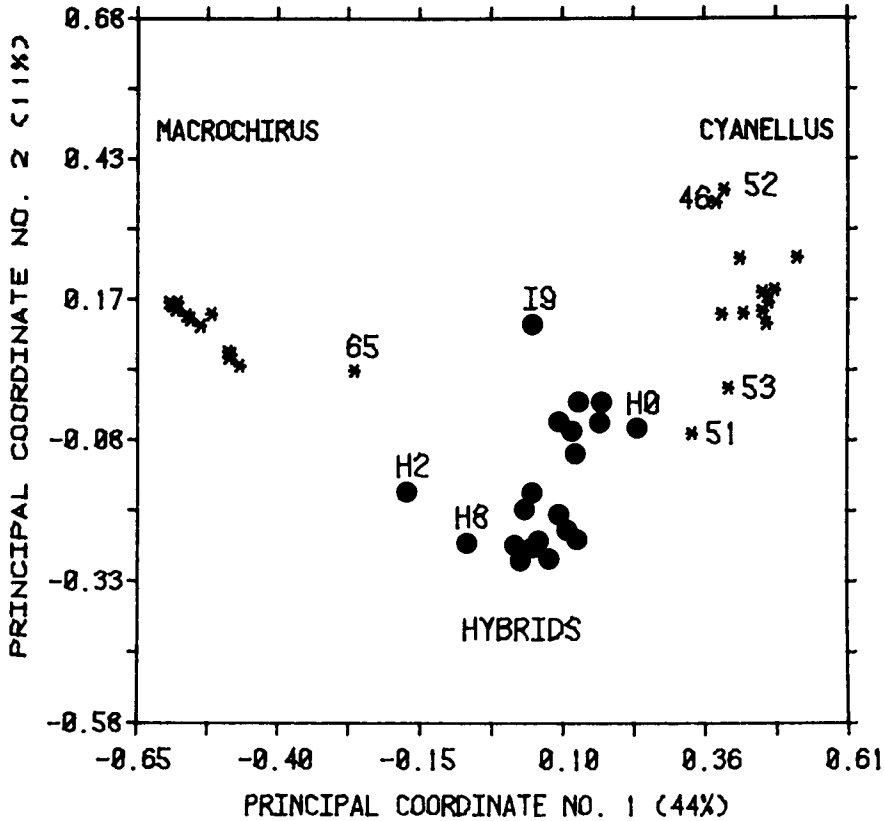
Fig. 4. Principal coordinate analysis (PCO) using F-1 weighted Gower metric similarities as in the weighted Wells' hybrid distance diagram of Fig. 3. The parental groups are even more tightly clustered than with the weighted Wells' hybrid distance (Fig. 3), however the hybrids are clustered about the same and information concerning outliers is preserved (i.e., such as the H2-65 affinity, the distinctness of I9 and individuals 46 and 52). The second axis appears to discriminate between some characteristic(s) of the hybrids and the parents.

third taxon) has been previously shown (Kistler, 1976; Whiffin, 1977) and is reiterated in Fig. 4. The first axis accounted for 44% of the variance among groups. The second axis appears to be due to the difference between the hybrids and parents. In general, the parental groups are more tightly clustered than with the weighted hybrid distances. However, outlier 65 is still present as well as a couple of outliers for *L. cyanellus* (51, 53). Hybrid individuals I9 and H2 are still somewhat removed from the other hybrids. These results are consistent with studies of hybridization in other taxa by Whiffin (1977) and Kistler (1976).

Canonical variate analysis (CVA) of the *L. cyanellus–L. macrochirus* data with 2 a priori groups gave a good separation of the parents (Fig. 5) with the hybrids skewed toward *L. cyanellus*. The overlapping hybrid is H3 and the overlapping *L. cyanellus* is #56. The next closest hybrid is H5. None of these individuals appeared unusual in any of the previous analyses. The principal problem with CVA (at this point) is that this method depicts only the relationship between groups, so the hybrids can only be placed along the axis between the two parental groups when using
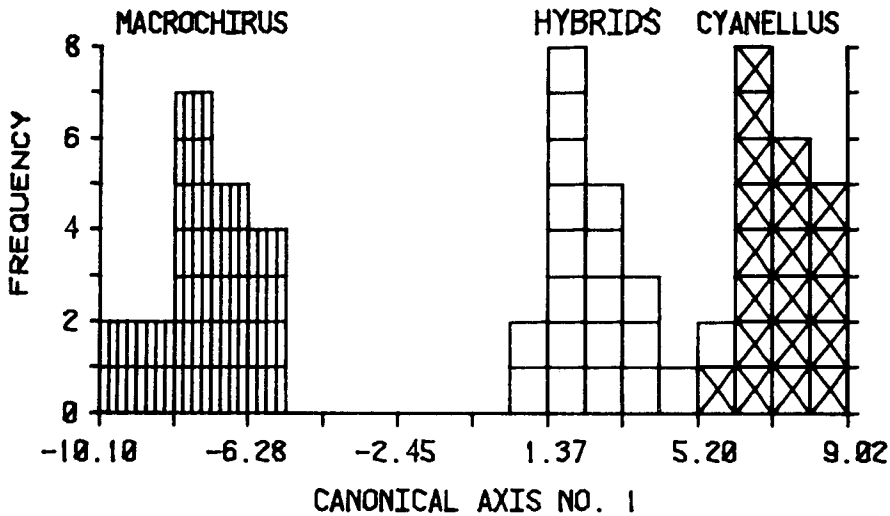
Fig. 5. Histogram plot of the first canonical axis from CVA with 2 a priori groups (*L. cyanellus* and *L. macrochirus*). The hybrids are plotted toward the *L. cyanellus* group as in several of the previous analyses, but fairly well resolved, except for one hybrid and one *L. cyanellus* individual. The distinctness of *L. macrochirus* individual 65 is not shown, nor is the affinity of H0 and 51 of *L. cyanellus* (the overlapping individuals are H3 and 56 of *L. cyanellus*).
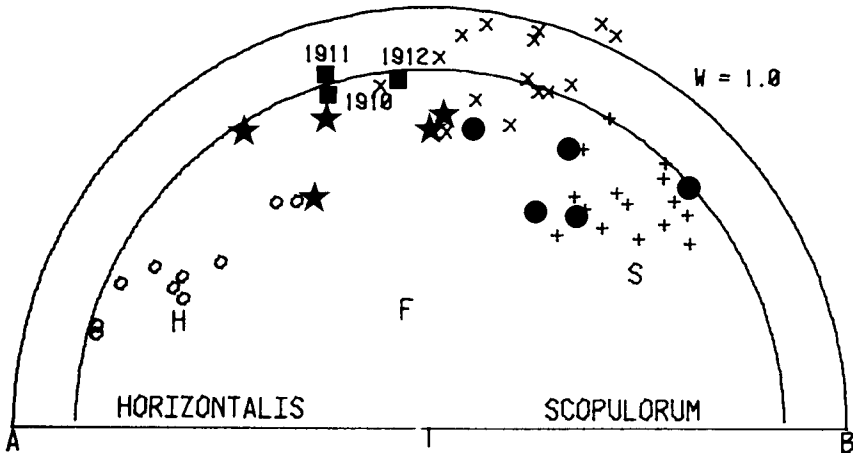

Fig. 6. Unweighted Wells' hybrid distance diagram for *Juniperus horizontalis* (open circles); *J. scopulorum* (pluses); *J. scopulorum* from Banff, Alberta (X's); putative hybrids from Bridger, Montana (solid circles); putative *J. horizontalis* from Grass Range, Montana (stars); and individuals from Amidon, North Dakota (solid squares). Variation appears to be continuous between these species in this analysis.
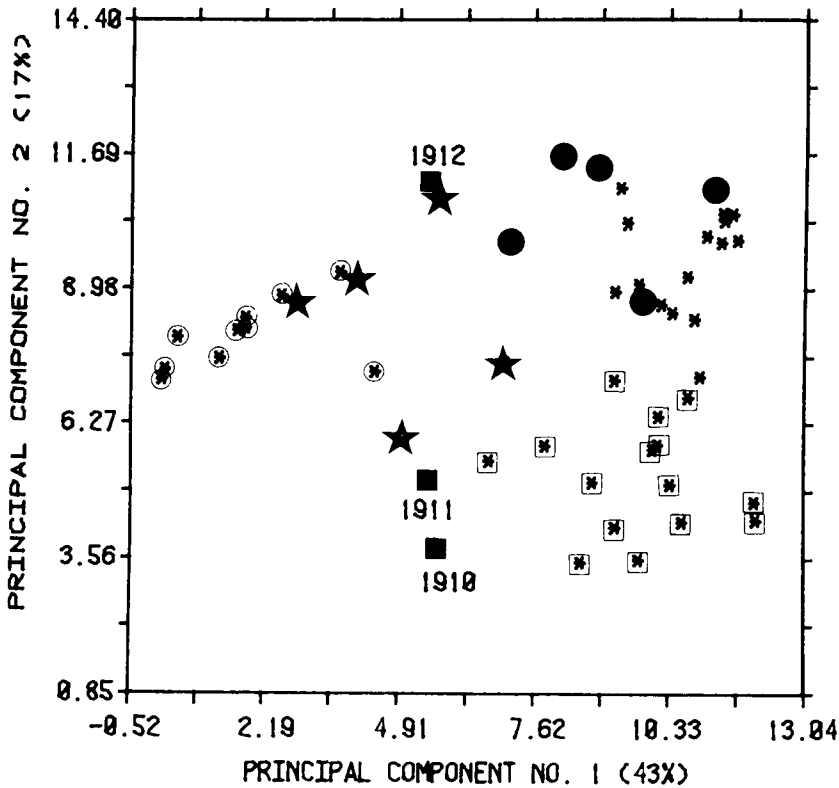
Fig. 7. Principal component analysis of *J. horizontalis* (circled asterisks); *J. scopulorum* (asterisks); *J. scopulorum* from Banff, Alberta (boxed asterisks); putative hybrids from Bridger, Montana (solid circles); putative *J. horizontalis* from Grass Range, Montana (stars); and individuals from Amidon, North Dakota (solid squares). The individuals from Banff are separated from the *J. scopulorum* reference individuals as previously shown. Variation is continuous from *J. horizontalis* to *J. scopulorum*.

2 a priori groups, and yet using 3 a priori groups has been shown to be unacceptable in an analysis with only putative hybrids and no independently known hybrids (Neff and Smith, 1979).

Juniperus

The unweighted hybrid distance diagram for *J. horizontalis*, *J. scopulorum*, 15 plants from Banff, 5 plants from south of Bridger, 5 prostrate plants from Grass Range, and 3 plants from Amidon shows a more or less continuum for *J. horizontalis* to *J. scopulorum* (Fig. 6). The Grass Range individuals appear somewhat intermediate and the Banff individuals seem to vary from intermediate to suggestive of a third taxon. Most of the Bridger individuals cluster with the *J. scopulorum* reference population. PCA of the same data set (Fig. 7) gave similar results except that the prostrate plant from Amidon is now separated on the second axis from the other two plants from Amidon. The plants from Banff are almost completely separated from the reference *J. scopulorum* population.

The F-1 weighted Wells' hybrid distance diagram (Fig. 8) resulted in tighter clusters of the *J. scopulorum* and *J. horizontalis* reference populations and a closer
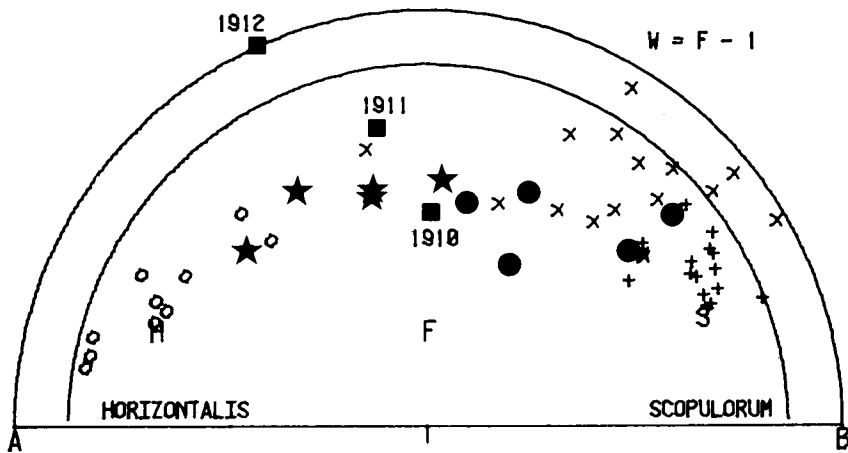
Fig. 8. Wells' hybrid distance diagram with character matches weighted by F-1 (F from ANOVA between the reference populations of *J. horizontalis* and *J. scopulorum*). The symbols are as previously defined (Fig. 5, 6 and 7). The *J. scopulorum* individuals from Banff, Alberta (X's) are clustered with the *J. scopulorum* reference but still show some divergence and a few individuals are intermediate to *J. horizontalis*. See text for further discussion.

association of the individuals from Banff with the *J. scopulorum* reference population. Notice that in general, the intermediate individuals are plotted lower on the diagram and more along a line between the reference populations than in the unweighted analysis (Fig. 6).

Principal coordinate analyses using characters weighted by F-1 gave very tight clusters for the species (Fig. 9). The prostrate individual 1912 is in the direction of *J. horizontalis* but obviously quite different. Individuals 1911 and 1910 both appear to be intermediate. Only a few of the plants from Banff are intermediate; most are suggestive of backcrosses or regional differentiation. Three of the 5 Bridger plants are intermediate. One aspect that is quite apparent is the close relationship of the Banff population with the reference *J. scopulorum* population, although 2 individuals are intermediate and the population shows some affinities toward *J. horizontalis*. The prostrate plant, 1912, from Amidon now shows more of the *J. horizontalis* characteristics. It should be emphasized that the volatile oils are not known to be linked with any morphological characters and thus selection for particular oil types in the field is minimized. Nevertheless, it is surprising to see the strong chemical intermediacy in the Grass Range population which appeared to be morphologically typical *J. horizontalis*.

The results obtained from *Juniperus* using principal coordinate analysis (PCO) with F-1 weighted similarities appear to parallel the *Lepomis* analyses in that the putative parents (or *Lepomis* parents) are more tightly clustered and intermediates are ordinated on a second axis to effect separation from both reference sets.

Canonical variate analysis using *J. horizontalis* and *J. scopulorum* as 2 a priori groups resulted in transgressive mapping of Banff individuals onto the axis (Fig. 10). The intermediate plants from Bridger are plotted much closer to *J. scopulorum*, and three of the intermediate plants from Grass Range were plotted intermediate but one (1901), which was the most intermediate in PCO, etc., was depicted beyond *J. scopulorum*. Several of the Banff individuals were placed to the left of *J. scopulorum* (Fig. 10). The prostrate plant from Amidon (1912) was mapped transgressive to *J. horizontalis*. Even though additive complementation is the general rule for the inheritance of secondary compounds (Alston and Turner, 1963; Hanover, 1966; Irving
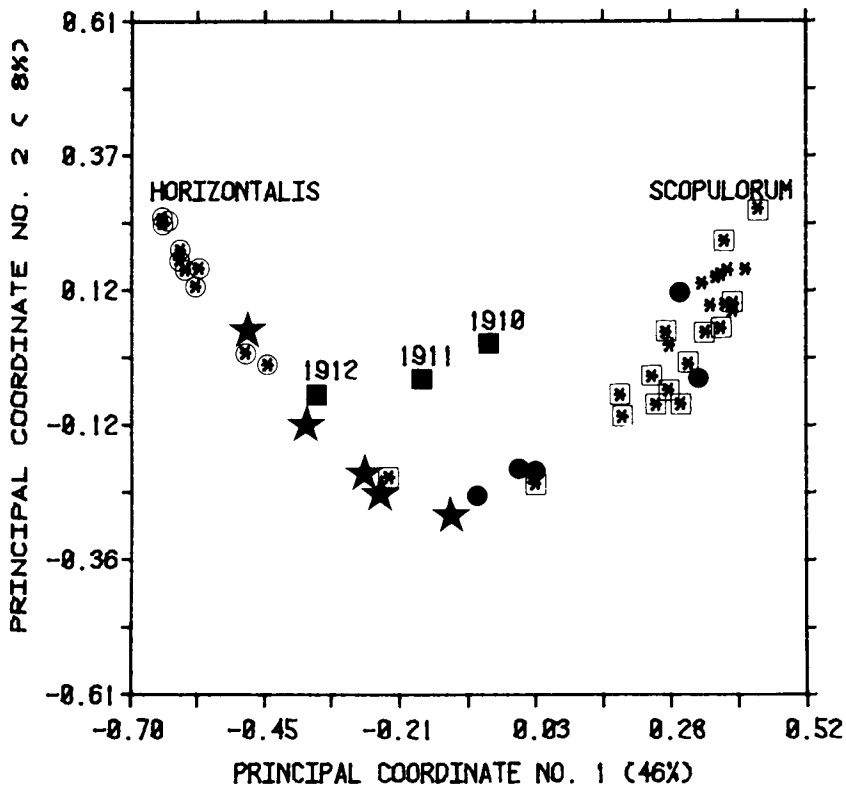
Fig. 9. Principal coordinate analysis (PCO) of *J. horizontalis* (circled asterisks); *J. scopulorum* (asterisks); *J. scopulorum* from Banff, Alberta (boxed asterisks); putative hybrids from Bridger, Montana (solid circles); putative *J. horizontalis* from Grass Range, Montana (stars); and individuals from Amidon, North Dakota (solid squares). Several of the individuals collected as hybrids still appear intermediate between the two species, however there is evidence of a break between most of the *J. scopulorum* individuals and the putative hybrids, with a continuum of variation into *J. horizontalis*. The species reference populations are very well resolved and tightly clustered. See text for additional discussion.

and Adams, 1973; Murray and Lincoln, 1970; Whiffin, 1977), transgressive quantitative inheritance in a few terpenoids is to be expected (Irving and Adams, 1973; Hanover, 1966; Whiffin, 1977). In addition, regional variation can certainly account for terpenoid values that might exceed either of the parental groups.

Examination of the average compound values for *J. horizontalis, J. scopulorum,* and the most extreme (by CVA) individual from Banff (OTU 1807), reveals that 15 of the 30 compounds of OTU 1807 exceed the means of either species. Compounds 2, 3, 6, 18, 20, 23, 27, 28, and 30 are beyond the means of *J. scopulorum*. Compounds 5, 7, 8, and 11 are beyond the means of *J. horizontalis*. Although the F-1 weights gave considerable weight to myrcene (Table 1) in the juniper example, analysis using the square root of F-1 gave similar results. This has also been shown in a taxonomic example (Adams, 1975). The use of the square root of F-1 appears to be acceptable as a weighting function and avoids some of the problems of giving a single character too much weight. Another (similar) weighting function is the added variance function
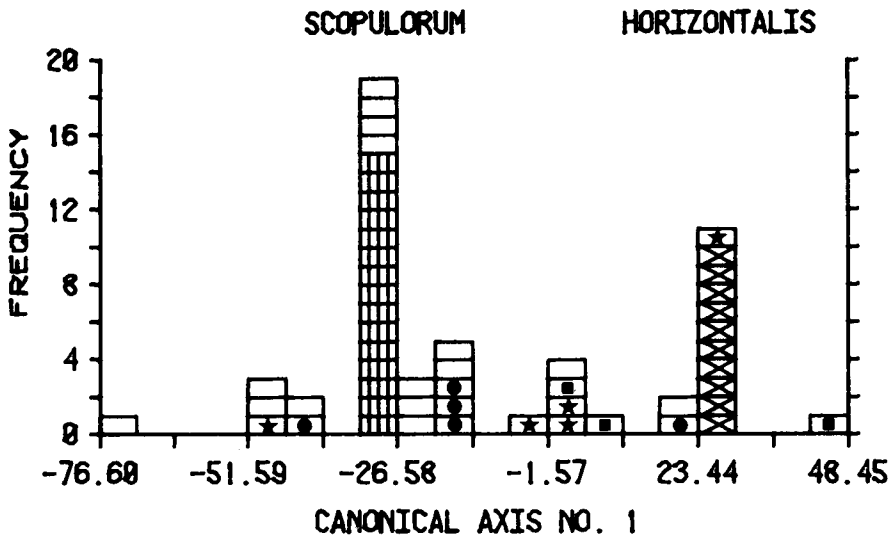
SCOPULORUM   HORIZONTALIS

CANONICAL AXIS NO. 1

Fig. 10. Canonical variate analysis (CVA) based on 2 a priori groups (*J. horizontalis* and *J. scopulorum* from Saskatoon, Saskatchewan and Durango, Colorado) with a posteriori plotting of exemplars on the first canonical axis. Open boxes are individuals from Banff, Alberta. Boxes with closed circles are individuals from Bridger, Montana. Boxes with stars are putative *J. horizontalis* from Grass Range, Montana and boxes with solid squares are the individuals from Amidon, North Dakota. Although the two species reference populations are well resolved, many individuals are shown to be transgressive. Note particularly one of the individuals of the putative *J. horizontalis* from Grass Range, Montana (star) that is transgressive to *J. scopulorum* and the putative hybrid from Bridger, Montana (solid circle) that is plotted next to the *J. horizontalis* population. See text for discussion.

used by Barkworth, McNeill and Maze (1979). It should be useful and deserves further investigation as it limits the weight of individual characters. The placement of OTU 1807, therefore, depends on the relative weights used in CVA. The canonical vector (axis 1) gives some idea as to the use of a character in the formation of the canonical score but this can be more readily understood by examining the sum of products between the canonical vector and the character values (Table 2). The canonical products for the *J. scopulorum* group average show that characters 7, 5, 1, 4, 23, 28, 2, 11, 12, and 9 contribute 14.33, 11.66, 10.63, 7.72, 7.40, 7.42, 6.69, 5.51, 4.55, 3.24 percent, respectively (79.05% of the total toward the group score). It appears that the characters driving the score of OTU 1807 so far negative (positive product differences, Table 2) are compounds 7 and 28 even though these are somewhat counterbalanced by characters 5 and 11 being extreme toward *J. horizontalis*. Comparison of the F ratios with the canonical products for the mean *J. scopulorum* (Tables 1, 2) reveals little correspondence. This is presumedly due to character correlations and the differences for within variances of the characters.

Since differences in the within group correlations might have a large effect on CVA, correlation patterns were examined in the *Lepomis* and *Juniperus* sets. Figure 11 shows the intercorrelation pattern in *L. cyanellus, L. macrochirus,* and the hybrids. Characters 1–6 are not shown as they are all highly correlated with character 7 and thus show the same pattern. Some of the correlations are common to both parents and the hybrids, such as 7-8, 8-11, 9-10, perhaps indicating coherence (Grant,

Table 2. Comparison of canonical variate scores and terpenoid values for *J. horizontalis*, *J. scopulorum* and the most extreme individual from Banff, 1807. Compound numbers are as in Table 1. Compound values for OTU 1807 marked with an asterisk were outside the range of the mean of *J. scopulorum* or *J. horizontalis*.

Cpd.#	Can. vector	Avg. horiz. values	Avg. scop. values	OTU 1807 values	Scop. prod.	1807 prod.	Prod. diff.
1	−55.85	3.60	5.85	4.43	−326.72	−247.42	−79.30
2	−4.13	55.09	50.23	39.68*	−207.45	−163.88	−43.57
3	−28.30	0.58	0.02	0.00*	−0.57	0.00	−0.57
4	133.02	4.43	1.80	2.04	239.44	271.36	−31.92
5	−327.23	0.81	1.11	0.52*	−361.59	−170.17	−191.42
6	218.18	0.26	0.18	0.00*	39.93	0.00	39.93
7	250.49	1.42	1.77	0.65*	444.37	162.83	281.54
8	240.30	0.11	0.24	0.06*	57.19	14.42	42.77
9	−49.23	1.32	2.04	1.55	−100.43	−76.31	−24.12
10	12.45	0.30	5.99	3.13	74.58	38.97	35.61
11	−33.30	3.31	5.13	2.12*	−170.85	−70.60	−100.25
12	302.65	0.01	0.47	0.29	141.03	87.77	53.26
13	−342.66	0.28	0.12	0.15	−39.75	−51.40	11.65
14	−176.00	0.37	0.04	0.14	−7.22	−24.64	17.42
15	16.36	0.02	0.55	2.64*	8.93	43.19	−34.26
16	29.93	3.28	0.28	0.00*	8.47	0.00	8.47
17	51.15	0.29	0.05	0.05	2.51	2.56	−0.05
18	−99.61	0.33	0.04	0.00*	−3.79	0.00	−3.79
19	−108.77	0.27	0.13	0.13	−14.25	−14.14	−0.11
20	167.91	0.14	0.26	0.33*	43.83	55.41	−11.58
21	−94.79	0.01	0.21	0.10	−20.00	−9.48	−10.52
22	−31.30	8.83	0.80	3.43	−24.89	−107.37	82.47
23	−65.92	1.56	3.48	4.55*	−229.40	−299.94	70.54
24	−177.36	0.87	0.35	0.47	−62.25	−83.36	21.11
25	−301.78	1.02	0.15	0.35	−45.87	−105.62	59.75
26	2.71	0.30	0.02	0.15	0.07	0.41	−0.34
27	−34.76	0.12	0.32	1.92*	−11.05	−66.72	55.67
28	297.56	2.98	0.77	0.00*	230.01	0.00	230.01
29	−29.91	0.03	3.06	0.37	−91.49	−11.07	−80.42
30	−221.46	0.21	0.42	0.79*	−93.90	−174.95	81.05
				Total	−521.11	−1000.14	479.03

1979). But several characters have a pattern of correlation in the hybrids (8-15, 8-16, 10-12, 7-17, 7-10, 8-10, 10-12) quite different than seen in either parent.

The correlations within *J. horizontalis* and *J. scopulorum* are shown in Fig. 12. The differences are much larger than in *Lepomis*, with the most obvious difference being the lack of negative correlations in *J. scopulorum*. This is in agreement with the previous report (Flake, Urbatsch and Turner, 1978). The pooled (within) correlation matrix between *J. horizontalis* and *J. scopulorum* (Fig. 12) shows both averaging and cancellation effects. The correlations are generally less than in either individual group except for the sesquiterpene alcohols (22, 24, 25, 28) whose intercorrelations remained high. This is comparable to the results of Flake, Urbatsch and Turner (1978) when the pooled correlation matrix from *J. virginiana* and *J. scopu-*
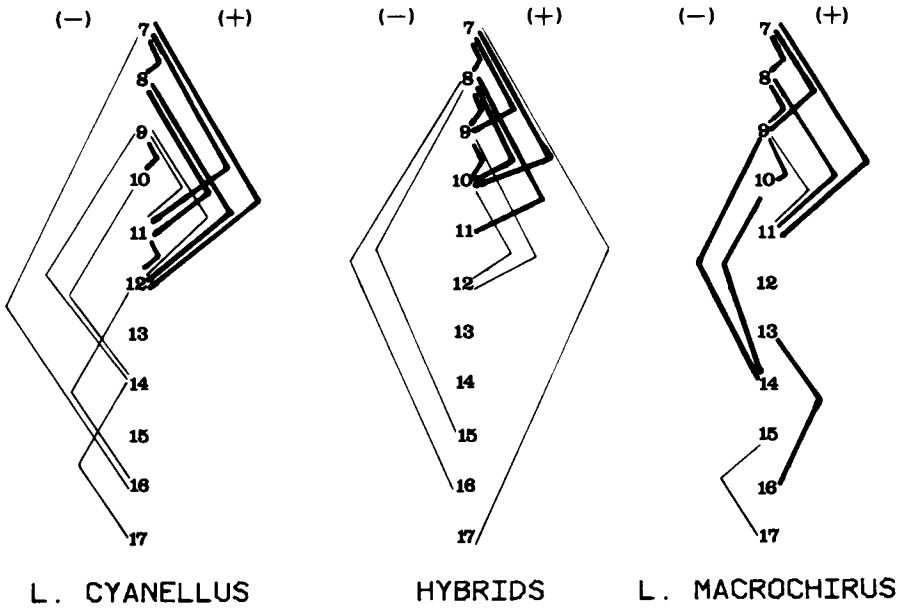
| (−) | (+) | (−) | (+) | (−) | (+) |

L. CYANELLUS   HYBRIDS   L. MACROCHIRUS

Fig. 11. Within group correlation patterns for 11 characters of *L. cyanellus, L. macrochirus* and artificial hybrids. The first six characters are not shown as they are very highly correlated with character 7 and thus show the same patterns. Notice that some of the high correlations are present in both parents and hybrids whereas other correlation patterns are broken down in the hybrids. Wide lines are highly significant and narrow lines are significant from zero.

lorum resulted in much lower correlations except where compounds were highly correlated in both of the individual groups. Of course, a set of artificial hybrids was not available for *J. horizontalis* and *J. scopulorum* to compare with the *Lepomis* hybrids and thus one could only speculate as to coherence in juniper hybrids. In any case, the variance/covariance matrices of a priori groups are quite different and appear to have exerted a considerable effect on the CVA analysis.

Conclusion

PCA was found to be useful in the initial analysis when no assumptions need be made about a priori groups for the construction of initial groups and, of course, for its original purpose, the examination of character correlation patterns. Unweighted Wells' hybrid distances performed very similarly to PCA. If a priori groups can be recognized, other methods involving some discrimination between groups are more powerful. Weighting the characters in the Wells' hybrid distance by F-1 resulted in an apparently robust method which appears to be satisfactory in handling cases of hybridization. PCO using F-1 weighted similarities compared closely to the Wells' hybrid distance diagram, with PCO having perhaps a little tighter clustering of the parents. One advantage of PCO over the Wells' distance diagram is that several axes will be obtained if there is multi-dimensionality to the set of OTUs. Since the amount of variance explained is presented for each axis, one might develop an objective method for the recognition of additional groups if they are segregated onto other
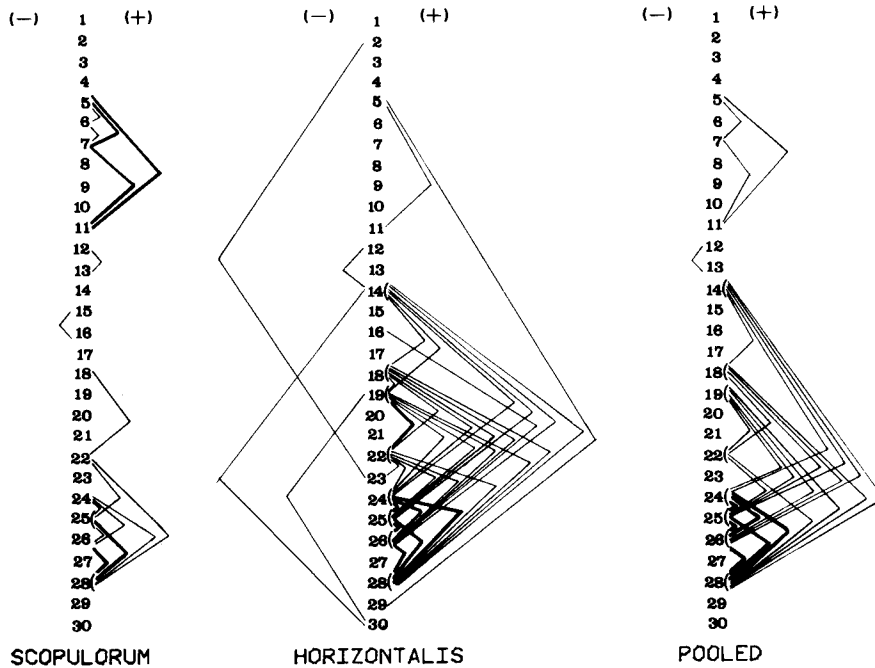
Fig. 12. Within group correlation patterns for 30 terpenoids of *J. horizontalis* and *J. scopulorum* and correlation patterns resulting from pooling individuals of the two taxa. The wide lines are highly significant and the narrow lines are significant from zero. Notice the lack of negative correlations and strong positive correlations between compounds 5, 7 and 11 in *J. scopulorum*. See discussion in text.

axes. CVA with 2 a priori groups cannot allow for a possible third group and if CVA is done with 3 a priori groups, it will tend to form 3 groups obscuring the nature of a hybrid group (Neff and Smith, 1979). Another limitation of using CVA (or discriminant functions) for the detection of hybridization is the unpredictability of analyses of individuals with transgressive character values.

Acknowledgments

Literature Cited

Adams, R. P. 1972. Numerical analysis of some common errors in chemosystematics. *Brittonia* 24: 9–21.
——. 1975. Statistical character weighting and similarity stability. *Brittonia* 27: 305–316.
——. 1983. Intraspecific variation in *Juniperus scopulorum*: Evidence for Pleistocene refugia and recolonization in western North America. *Taxon* (in press).

————, E. von Rudloff, T. A. Zanoni and L. Hogge. 1981. The volatile terpenoids of *Juniperus blancoi* and its affinities with other entire leaf margin junipers of North America. *J. Natural Products* 44.

Anderson, E. 1949. *Introgressive hybridization*. John Wiley and Sons, N.Y.

Barkworth, Mary E., J. McNeill and J. Maze. 1979. A taxonomic study of *Stipa nelsonii* (Poaceae) with a key distinguishing it from related taxa in western North America. *Can. J. Bot.* 57: 2539–2553.

Blackrith, R. E. and R. A. Reyment. 1971. *Multivariate morphometrics*. Acad. Press, London.

Cooley, W. W. and P. R. Lohnes. 1971. *Multivariate data analysis*. John Wiley & Sons, N.Y.

Eades, D. C. 1965. The inappropriateness of the correlation coefficient as a measure of taxonomic resemblance. *Syst. Zool.* 14: 98–100.

Fassett, N. C. 1945. *Juniperus virginiana, J. horizontalis,* and *J. scopulorum* III. Possible hybridization of *J. horizontalis* and *J. scopulorum*. *Bull. Torrey Bot. Club* 72: 42–46.

Flake, R. H., L. Urbatsch and B. L. Turner. 1978. Chemical documentation of allopatric introgression in *Juniperus*. *Syst. Bot.* 3: 129–144.

Gower, J. C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 53: 315–328.

————. 1971. A general coefficient of similarity and some of its properties. *Biometrics* 27: 857–874.

Grant, V. 1979. Character coherence in natural hybrid populations of plants. *Bot. Gaz.* 140: 443–448.

Hanover, J. W. 1966. Genetics of terpenes 1. Gene control of monoterpene levels in *Pinus monticola* Dougl. *Heredity* 21: 73–84.

Irving, R. S. and R. P. Adams. 1973. Genetic and biosynthetic relationships of monoterpenes. *In: Recent advances in phytochemistry* 6: 187–215. V. C. Runeckles and T. J. Mabry, Eds. Acad. Press, N.Y.

Kistler, J. R. 1976. Detection of hybridization in natural populations. M.S. Thesis, Colorado State Univ., Ft. Collins, 215 pp.

Maze, J. 1980. A comment on Wells' distance coefficient as a hybridization index. *Taxon* 29: 667.

Murray, M. J. and D. E. Lincoln. 1970. The genetic basis of acyclic oil constituents in *Mentha citrata* Ehrh. *Genetics* 65: 457–471.

Neff, N. A. and G. R. Smith. 1979. Multivariate analysis of hybrid fishes. *Syst. Zool.* 28: 176–196.

Namkoong, G. 1966. Statistical analysis of introgression. *Biometrics* 22: 448–502.

Pimentel, R. A. 1979. *Morphometrics*. Kendall/Hunt Publishing Co., Dubuque, Iowa.

Veldman, D. J. 1967. *Fortran programming for the behavioral sciences*. Holt, Rinehart and Winston, N.Y.

von Rudloff, E. 1975. Volatile oil analysis in chemosystematic studies of North American conifers. *Biochem. Syst. and Ecol.* 2: 131–167.

Wells, H. 1980. A distance coefficient as a hybridization index: an example using *Mimulus longiflorus* and *M. flemingii* (Scrophulariaceae) from Santa Cruz Island, California. *Taxon* 29: 53–65.

Whiffin, T. 1977. Volatile oils and the study of natural hybridization between *Correa aemula* and *C. reflexa* (Rutaceae). *Aust. J. Bot.* 25: 291–298.