Classification of *Aedes* Mosquitoes Using Statistical Methods

F. James Rohlf*
General Sciences Department
IBM Watson Research Center
Yorktown Heights, N. Y. 10598

ABSTRACT The thoracic setae data of Lunt and Nielsen (1971a) is analyzed using cluster analysis and nonmetric multidimensional scaling analyses. The results were compared to previous classifications of *Aedes* mosquitoes.

INTRODUCTION

The purpose of the present paper is to present an analysis of the thoracic setae data of Lunt and Nielsen (1971a) and to compare it with their intuitive analysis (Lunt and Nielsen, 1971b). Even though Lunt and Nielsen (*L&N*) only published information on 14 characters, it was found that their data contained sufficient information to make such a comparison interesting (even though they can not be considered conclusive). Perhaps the study should best be interpreted as an experiment. Comparisons are also made to the results of Rohlf (1963) and Steward (1968). Rohlf's data were also reanalyzed using ordination and graph theoretic methods. Unfortunately Steward's data seem to have been lost so that these newer methods could not be used on them.

Since *L&N* present an extensive intuitive analysis of their data it is possible to make a detailed comparison of their conclusions with a numerical analysis of the data upon which those conclusions were based. This serves as a test of the extent to which intuitive and numerical taxonomic analyses of the same data are concordant. While it is undoubtedly true that *L&N* were influenced by their prior knowledge of many other characteristics of mosquitoes, they state that they have attempted to present a phylogeny and classification as indicated by the adult thoracic setae characters only.

MATERIAL AND METHODS

The data (14 thoracic setal characters, 63 species) used in this study were taken from table 1 of Lunt and Nielsen (1971a). Since means were given for most characters, those values were used directly. Where only ranges of variation were given the midrange values were used. The list of the 63 species and their code numbers used in the present study are given in TABLE I. The arrangement is as given in TABLE II of Lunt and Nielsen (1971b). Some notes are included regarding changes proposed by Zavortink (1972). Five species in Rohlf's (1963) study were not included in the *L&N* study. They are numbered 64 through 68 in the present study. The code numbers used by Rohlf (1963) are also given in TABLE I.

The *L&N* data matrix was standardized by characters and correlation and distance coefficients computed for all pairs of species. The matrices were subjected to UPGMA cluster analysis, principal components analysis, linear adaptive hierarchical cluster analysis (LAHCS, Rohlf, 1970), and nonmetric multidimensional scaling analysis (MDSCALE, Kruskal, 1964).

---

* On leave from the Department of Ecology and Evolution, State University of New York, Stony Brook, N. Y. 11794

The results of a principal components analysis were used as the initial configuration for the MDSCALE analysis. Minimum spanning trees (MST, Rohlf, 1970) and B_k graphs (Rohlf, 1974) were also fitted to the data. These two methods were used to indicate the near neighbor distance relations among the points in the 14 dimensional space. The MST was also superimposed upon the 3 dimensional plots in the figures given below (Figures 3 and 4). General discussions of these methods are given in Sneath and Sokal (1973).

The same computations were also made for the Rohlf (1963b) data (148 adult and larval characters, 48 species). These will be denoted below as the "*A+L*" data. There were not enough species of *Aedes* included in the studies of Rohlf (1967) and Hendrickson and Sokal (1968) to make comparisons worth while.

In addition, a multiple factor analysis was performed on the *L&N* character correlation matrix. Factors were extracted iteratively until the communalities converged to within 0.0001 (this corresponds to the MINRES solution, Harman, 1967). The factors were then rotated to an oblique simple structure configuration using the primary product function plane method (PPFP, Katz and Rohlf, 1975). The hyperplane width parameter, w, was set equal to 0.10. The results were then expressed in terms of the primary axes reference system. General references to factor analysis are Harman (1967) and Mulaik (1972).

The phenograms were drawn using the program given in Rohlf (1975). The perspective plots of the 3-dimensional MDSCALE configurations were prepared using the techniques described in Rohlf (1968).

RESULTS

Figure 1 presents the phenogram which resulted from a UPGMA cluster analysis applied to the *L&N* matrix of distances. The cophenetic correlation (Sokal and Rohlf, 1962) was only 0.741. This indicates that the patterns of similarity among the 63 species cannot be well represented by a hierarchical classification. The cophenetic correlation for the phenogram based upon the correlation matrix (not shown) was even lower, 0.594. Both of these correlations are statistical significant (Rohlf and Fisher, 1968). The results of the UPGMA cluster analysis applied to the *A+L* data are not shown since they are available in Rohlf(1963). The results of the LAHCS cluster analysis are shown in Figure 2 for w = 0.9 (a rather conservative value) for the *L&N* data. Other values of w varying from 0.1 to 10.0 were also tried. Since there were no clear trends in the data (see Figure 3 discussed below) the use of smaller values of w was not warranted. Larger values of w gave results very similar to UPGMA cluster analysis.

Two perspective views of a 3-dimensional model of the best configuration of the 63 points into a 3-dimensional space are given in Figure 3. The coordinates were determined using MDSCALE on the matrix of interpoint distances. The final stress value, S, was 0.108 (which is considered to represent a "good" fit). The solution was very similar to the 3-dimensional principal components solution (which accounted for 81.98 percent of the variance among the species. The near neighbor relations are shown by the MST superimposed on Figure 3. It is shown by the 62 line segments interconnecting the points. Since there were so many points it was difficult to obtain a view of the model in which points and/or labels did not obscure one another. The program tried to place the label within the spheres. If that were not possible then the label was placed at the left if the sphere was partially hidden by other spheres. Otherwise the label was placed at the right.

Figure 4 shows the results of the MDSCALE analysis applied to the *A+L* data. The lower (close up) view is from the rear of the model. For reference some of the edges of the

MST which run off the figure are labelled with the code number of the species that they connect to. The final stress value was S = 0.172 (which is considered to correspond to a fairly good fit). The 48 species are identified with the same system of code numbers used for the *L&N* data. There are 5 species in the *A+L* data set that do not correspond to species in the *L&N* data set. They have been numbered 64 through 68. For this data the first 3 principal coordinate axes (used rather than principal components axes since there were missing data values and there were so many more characters than species, Rohlf, 1972) were only able to explain 32.14 percent of the variation among the species. The configuration obtained (not shown) was, however, again very similar to the MDSCALE solution.

TABLE II gives the correlations between the 14 characters and the 3 axes in Figure 3. This matrix is very similar to the matrix of the first 3 eigenvectors from the character correlation matrix. It can be used to aid in the interpretation of the axes in Figure 3. For example, axis I is highly correlated with all setal characters except the "v" group. It thus represents a "general setal abundance" axis. Species to the left end of axis I in Figure 3 tend to have fewer setae than those toward the right.

TABLES III and IV give the results of a factor analysis applied to the *L&N* matrix of character correlations. Four factors were extracted and rotated to an oblique simple structure using the PPFP method. TABLE III gives the standardized partial regression coefficients (factor "loadings") for the 14 variables on the 4 primary axes. TABLE IV gives the correlations among the 4 primary axes.

DISCUSSION

The classification given in TABLE I will be used as the standard for comparison in the discussion that follows.

Overall, the *L&N* data set does not reveal very distinct structure. Even though the cophenetic correlations are statistically significant there are few clusters visually apparent in Figure 3. One sees only a central elongated cluster with a number of isolated points at the periphery.

It is reassuring to see an analogous structure in Figure 4. If one ignores species 66, 67, and 68 (not present is Figure 3) then one again has a central cluster (not elongated however) and a number of isolated points. Except for *fulvus pallens* (16) the most isolated points in Figure 4 (22, 43, 56, 58, 60, 62, and 63) are among the most isolated points in Figure 3. One also sees agreements between the phenograms in Figures 1 and 2 and the *A+L* phenogram (Figure 1 in Rohlf, 1963, not shown here) in terms of the outlying species.

Within the main cluster in both Figures 3 and 4 one tends to find species in the same species group near one another. There are, however, few apparent gaps to separate the different species groups. This lack of distinct clustering reflects (I believe) the difficulty of classifying *Aedes* by either numerical or traditional methods. These figures should be compared to the ones obtained at the generic level in Culicidae (Rohlf, 1970) or in other taxonomic groups such as the *Hoplitis* complex (see Rohlf, 1968) where the clusters are obvious and the conventional classification is much more stable.

In the present data the phenograms (while giving groups of highly similar species) are much less interesting than the 3-d models shown in Figures 3 and 4. The principal advantage of cluster analysis is that it is not limited to a 3 dimensional space. Since the *L&N* data can be approximated so well in 3 dimensions most of the discussion given below will be based on the relationships shown in Figure 3.

SUBGENERA. Since *Stegomyia* (63) and *Kompia* (57) both have distinctly fewer thoracic setae they are found (as expected) at the extreme left on axis I in Figure 3. They are also clustered together in both Figures 1 and 2. While the *scapularis* group is the closest species group to these subgenera (as also pointed out by *L&N*), *decticus* (37, in the *communis* group) and *atropalpus* (60, listed in the subgenus *Finalya* by *L&N* but placed in the *atropalpus* group of the subgenus *Ochlerotatus* by Zavortink, 1972) are the closest species.

L&N state that the two representatives of the subgenera *Aedes* and *Aedimorphus* are very similar in the number of setae they possess. Yet in Figures 1 - 3 they are not very close. *A. vexans* (59) is centrally located and *cinereus* (58) is found towards the left (reflecting the fact that the latter has a lower mean numbers for all of the 14 setal characters recorded). The *L&N* data do not show the close similarity between these two species found by Rohlf (1963b) and by Steward (1968).

The species *L&N* list for the subgenus *Finlaya* (60-62) are also located towards the left side of axis I in Figure 3 reflecting the generally lower numbers of thoracic setae. The species do not form single cluster in Figures 1 and 2 . In Figure 3 species 61 is plotted between species 60 and 62 which is consistent with *L&N*'s statement that in nearly every case the number of setae in *hendersoni* (61) is intermediate between *atropalpus* (60) and *triseriatus* (62). Species 60 and 62 are also close in Figure 4 based on the $A+L$ data. *L&N* state that on the basis of their study there appears to be no justification for placing *varipalpus* (23) and *sierrensis* (22) in *Finlaya*, yet in Figure 1 species 22 and 61 are clustered together as well as 23 and 62. In Figure 2 species 22, 23, 61, and 62 form a single cluster. These species are close and linked together by the MST in Figure 3 in a manner consistent with the relations shown in Figure 4. Steward (1968) also found that 23, 60, and 62 were closer to each other than they were to any other species. A direct examination of the raw data shows quite clearly that species 22, 23, 61, and 62 have very similar setal counts. In contrast *L&N* state that *Finlaya* appears to be more closely related to the *scapularis* group of the subgenus *Ochlerotatus*. In Figure 3 the members of the *scapularis* group are indeed close (particularly species 26 and 30). In Figure 4 the members of the *scapularis* group are in the same general region as *Finlaya* but not as close as *sierrensis* (22). Zavortink (1972) removed all of the New World species from this subgenus. Species *hendersoni* (61), *triseriatus* (62), and *zoosophus* (65) were placed in the subgenus *Protomacleaya*. Figure 4 is consistent with the placement of species 62 and 65 together. Figure 3 is also consistent in that species 60 is separated somewhat from species 61 and 62.

SPECIES GROUPS. Most of the species groups within the subgenus *Ochlerotatus* are very indistinct. In many cases species in the same species groups are similar to one another, but there is little evidence, however, that there are clear gaps between most of the groups.

The *taeniorhynchus* group (1-4) did not come out as a distinct group in Figures 1, 2, or 3. This was disappointing since it was one of the more distinct clusters found by Rohlf (1963a,b) and Steward (1968). In Figure 3 the species are in the same general region but they are not near neighbors of one another. *L&N* state that *sollicitans* (4) tends to have the highest number of setae. In Figure 3, however, *nigromaculus* (3) is further to the right along axis I (reflecting the fact that it has a higher mean number of setae for 10 out of the 14 characters recorded). In Figure 4 the four species can be seen to be near neighbors of one another and to lie as a group between *Finlaya* and the rest of the *Ochlerotatus* (*scapularis* is the closest species group within *Ochlerotatus*). These results confirm *L&N*'s conclusion that their data do not support the elevation of this group to subgeneric status.

While the *stimulans* group (5 - 14) does not come out as a distinct cluster in Figures 1 and 2, one sees in Figure 3 that all of the species are in the same general region and that many of them are near neighbors of one another. Even *grossbecki* (15) which *L&N* moved in to a separate group has its closest links to members of this group (*i.e.*, *excrucians*, 8).

In Figures 1 - 3 *fulvus pallens* (16) is not as isolated as it is in Figure 4. In Figure 4 its closest neighbors are in the *scapularis* group. This is similar to Figure 3 in that its closest neighbor is *canadensis* (53) which has links to the *scapularis* group. The *L&N* data do not support the elevation of the *fulvus* group to subgeneric rank as was suggested by Rohlf (1963b).

The *dorsalis* group (17-19) formed a cluster in Figure 2 but not in Figure 1. The members of this group are near neighbors as can be seen in Figure 3 (and also in Figure 4). These results clearly support the removal of *canadensis* (53) from this group. *L&N* recommend that *canadensis* be placed in a group by itself rather than in the *communis* group as recommended by Rohlf (1963b). In Figure 3 *canadensis* is plotted between the left end of the *communis* group and the *scapularis* group. It is probably sufficiently separate to warrant its placement in a separate group. Figure 4 shows an analogous placement of *canadensis* although its separation is much less distinct.

The *L&N* data strongly suggest that the *pulchritarsis-varipalpus* group (20-23) be split. *A. sierrensis* (22) and *varipalpus* (23) should be placed with those species that were in *Finlaya* and are now in the subgenus *Protomacleaya* (Zavortink, 1972, see above). The other two species (*monticola* (20) and *muelleri* (21)) should be retained as a species group in *Ochlerotatus*. A direct examination of the raw data themselves shows that this group would be much more homogeneous with species 22 and 23 removed. Of course these conclusions are based on just 14 thoracic setal characters which may not be representative. The recent study by Lunt (1976) showed that despite the similarity in larval morphology (Arnell and Nielsen,1972), *varipalpus* and *monticola* are strikingly different in the characteristics of their protein bands as revealed by electrophoresis. Unfortunately this study was not able to indicate where *monticola* should be placed. Zavortink (1972) removed *muelleri* (21) from this group and placed it in its own monotypic species group.

The *scapularis* group (24-32) is the most distinct species group in the *L&N* data. Most of the species cluster together in Figure 2. In Figure 3 they form a connected but very straggly cluster of points towards the left end of axis I. This figure supports *L&N*'s decision to move *aurifer* (25) and *thibaulti* (30) from the *communis* group. In Figure 4 these two species are closer to *canadensis* and several species in the *communis* group. Rohlf (1963a) found *thibaulti* to be closer to the *scapularis* group using adult characters but not when larval characters were used. As mentioned above the *scapularis* group is also quite close to *Protomacleaya*.

As might be expected the *communis* group (33-52) remains a problem. Members of this group are widely scattered across Figure 1, 2, 3, and 4. The species which *L&N* list as lacking the lower mesepimeral setae (and also having a reduced number of other setae as well) are to be found towards the left of axis I in Figure 3. The five species (33, 40, 43, 47, and 52) that *L&N* list as possessing unusually large numbers of setae are all at the extreme right end of axis I. *L&N* also mention that *decticus* (37) sticks out because it tends to have a lower number of setae. It is plotted towards the extreme left end of axis I (consequentially it is similar to some *Protomacleaya*). If one wished to be consistent then on the basis of the *L&N* data one would have to split the *communis* group into a number of species groups. In particular, species *decticus* (37), *impiger* (40), *nigripes* (43), *punctodes* (47), *rempeli* (49), and *ventrovittis* (52) would have to be placed into groups of their own.

Figure 3 supports Rohlf's (1963b) removal of *canadensis* (53) from the *dorsalis* group. Its nearest neighbors in Figure 3 are members of the *communis* group (*implicatus*, 41) but *L&N* felt that it was sufficiently distinct to place it in a separate group close to the *communis* group.

As *L&N* state, *spencerii* (54) is most similar to members of the *dorsalis* group. In Figure 4 it is reassuring to see *idahoensis* (64) linked to *spencerii* since it is now considered to be only a subspecies of *spencerii*. Their closest links here are with *sticticus* (51) in the *communis* group. The results of Steward (1968) also placed *spencerii* near the *communis* group.

The *rusticus* group (55-56) forms a connected set in both Figures 3 and 4. The reason Rohlf (1963b) placed *trichurus* (56) in a separate group is easily seen in Figure 4. In Figure 3 *bicristatus* (55) is indicated as being one of the more isolated species (but not as extremely as in Figure 4). In both figures the species closest to this group belong to the *communis* group.

Since the *L&N* data set is based on only 14 characters one should not place too much weight on the results based on their data unless they are very consistent with the other studies. Perhaps the only definite change one should make in the classification given in TABLE I is to move *varipalpus* (23) and *sierrensis* (22) to the subgenus *Protomacleaya*.

While it is possible to relate the present classification to the configuration of points in Figures 3 and 4, it is clear that one would not arrive at the present classification using only this information. It is tempting to consider completely redefining the species groups within the *Ochlerotatus* since the present ones seem so arbitrary. However, there seem to be so few distinct gaps in the phenetic spaces depicted in Figures 3 and 4 so that any new groups would not be expected to be much more distinct than the old ones. It may be that a clearer picture would emerge if a larger study were performed which included species from other parts of the world.

FACTOR ANALYSIS. The results of the factor analysis were very interesting despite the fact that so few characters were included. What was most interesting was the morphological "regionalization" of each factor and the suggestion that one of the characters should be redefined. Factor I (TABLE III) has a high loading (standardized partial regression coefficient) only on the number of "v"-group setae. This character has the lowest intercorrelations with the other characters (this can also be seen by the lowness of the correlation of factor I with the other factors , TABLE IV). *A. pionips* (45) and *rempeli* (49) have the highest and lowest projections, respectively, onto this factor. Factor II has its highest loadings with the numbers of anterior lateral, acrostichal, dorsocentral, posterior pronotal, propleural and lower mesepimeral setae. *L&N* observe that all species having the dorsocentral gap (absence of setae in the anterior part of the dorsocentral region) also have the acrostichal gap (absence of most of the acrostichal setae) and lack the lower mesepimeral setae. Since the presence of these "gaps" implies lower numbers of setae, this association corresponds to factor II. While *L&N* point out that the reciprocal relationship is not always true (species having the acrostichal gap may or may not have the dorsocentral gap and lower mesepimeral setae), these traits are highly correlated. *A. nigripes* (43) has the highest projection onto factor II and *purpureipes* (57) and *aegypti* (63) have the lowest projections.

Factor III has its highest loadings on the posterior lateral, dorsocentral, scutellar, anterior pronotal, propleural, posterior pronotal, and sternopleural setae. This association of characters is not mentioned by *L&N*. The correlation between the numbers of posterior lateral and scutellar setae is 0.88, one of the highest in the study. Due to the high correlation between factors II and III the same species have the highest projections on factor III that were found for factor II. It is interesting that on the mesonotum, factor II is correlated with the anterior setal groups and factor III is correlated with the posterior setal groups. The dorsocentral setae group extends over the length of the mesonotum so that it is reasonable to find that the numbers of these setae are correlated with both factors. *L&N* found the dorsocentral gap to be a very useful character. These results suggest that the dorsocentral setal group should be split so that one would have separate setal counts for the anterior and posterior regions. The division should be made somewhere posterior to the scutal angles. If this were done one would expect the anterior counts to be correlated with factor II and the posterior counts to be correlated with factor III.

ACKNOWLEDGEMENTS

This paper is contribution no. 214 of the Program in Ecology and Evolution, State University of New York, Stony Brook. The work was supported in part by a grant (BMS 72-02212 A03) from the National Science Foundation. The statistical computations were performed using NTSYS on Stony Brook's IBM 370/155 computer. The final preparation of the manuscript and figures were done at the IBM Watson Research Center with the technical assistance of their staff. Art Appel helped with the development of the program that made the plots for Figures 3 and 4.

REFERENCES CITED

Arnell, J. H. and L. T. Nielsen. 1972. Mosquito studies (Diptera, Culicidae). XXVII. The varipalpus group of *Aedes (Ochlerotatus)*. *Contr. Amer. Entomol. Inst.*, **8**:1-48.

Harman, H. H. 1967. *Modern factor analysis*, 2nd ed. Univ. Chicago Press:Chicago. 474pp.

Hendrickson, J. A., Jr., and R. R. Sokal. 1968. A numerical taxonomic study of the genus *Psorophora* (Diptera: Culicidae). *Ann. Entomol. Soc. Amer.*, **61**:385-392.

Katz, J. O. and F. J. Rohlf. 1974. Functionplane - A new approach to simple structure rotation. *Psychometrika*, **39**:37-51.

Katz, J. O. and F. J. Rohlf. 1975. Primary product functionplane, an oblique rotation to simple structure. *Multivariate Behavioral Research*, **10**:219-232.

Lunt, S. R. 1976. Protein band studies of fourth instar larvae in the *Aedes varipalpus* group. *Mosquito Systematics*, **8**:200-204.

Lunt, S. R. and L. T. Nielsen. 1971a. The use of thoracic setae as a taxonomic tool and as an aid in establishing phylogenetic relationships in adult female *Aedes* mosquitoes of North America. Part I. *Mosquito Systematics Newsletter* , **3**(2):69-98.

Lunt, S. R. and L. T. Nielsen. 1971b. The use of thoracic setae as a taxonomic tool and as an aid in establishing phylogenetic relationships in adult female *Aedes* mosquitoes of North America. Part II. *Mosquito Systematics Newsletter* , **3**(3):102-121.

Mulaik, S. A. 1972. *The foundations of factor analysis*. McGraw-Hill:New York. 453pp.

Rohlf, F. J. 1963a. Congruence of larval and adult classifications in *Aedes* (Diptera: Culicidae). *Systematic Zool.*, **12**:97-117.

——— 1963b. Classification of *Aedes* by numerical taxonomic methods (Diptera:Culicidae). *Ann. Entomol. Soc. Am.*, **56**:798-804.

——— 1967. Correlated characters in numerical taxonomy. *Systematic Zool.*, **16**:109-126.

——— 1968. Stereograms in numerical taxonomy. *Systematic Zool.*, **17**:246-255.

——— 1970. Adaptive hierarchical clustering schemes. *Systematic Zool.*, **19**:58-82.

——— 1971. Perspectives on the application of multivariate statistics to taxonomy. *Taxon*, **20**:85-90.

——— 1972. Empirical comparison of three ordination techniques in numerical taxonomy. *Systematic Zool.*, **21**:271-280.

——— 1973. Algorithm 76. Hierarchical clustering using the minimum spanning tree. *Computer J.*, **16**:93-95.

——— 1974. Graphs implied by the Jardine-Sibson B_k overlapping clustering methods. *J. Amer. Statistical Assn.*, **69**: 705-710.

——— 1975. Note on Algorithm 81. Dendrogram plot. *Computer J.*, **18**:90-92.

Rohlf, F. J. and D. L. Fisher 1968. Test for hierarchical structure in random data sets. *Systematic Zool.*, **17**:407-412.

Sneath, P. H. A. and R. R. Sokal. 1973. *Numerical taxonomy.* Freeman:San Francisco. 573 pp.

Sokal, R. R. and F. J. Rohlf. 1962. Comparison of dendrograms using objective methods. *Taxon*, **11**:33-40.

Steward, C. C. 1968. Numerical classification of the Canadian species of the genus *Aedes*. *Systematic Zool.*, **17**:426-431.

Zavortink, T. J. 1972. Mosquito studies (Dipter, Culicidae) XXVIII. The New World species formerly placed in *Aedes (Finlaya). Contrib. Amer. Ent. Inst.,* **8**:1-206.

TABLE I. List of species and their classification as given by Lunt and Nielsen (1971). Code numbers used in the present study and in Rohlf (1963) are also given. Species marked with an asterisk were not included in the L&N study.

Species	Present Code No.	Rohlf (1963) Code No.
Subgenus *Ochlerotatus* Lynch-Arribalzaga		
Group 1 (A) *taeniorhynchus* group		
taeniorhynchus (Wiedemann)	1	33
mitchellae (Dyar)	2	20
nigromaculis (Ludlow)	3	22
sollicitans (Walker)	4	28
Group 2 (B) *stimulans* group		
aloponotum Dyar	5	-
barri Rueger	6	-
cantator (Coquillett)	7	7
excrucians (Walker)	8	12
fitchii (Felt & Young)	9	13
flavescens (Müller)	10	14
increpitus Dyar	11	
riparius Dyar & Knab	12	26
squamiger (Coquillett)	13	30
stimulans (Walker)	14	32
Group 3		
grossbecki Dyar & Knab	15	-
Group 4 (C) *fulvus* group		
fulvus pallens Ross	16	15
Group 5 (E) *dorsalis* group		
campestris Dyar & Knab	17	5
dorsalis (Meigen)	18	11
melanimon Dyar	19	-
Group 6 (X) *pulchritarsis-varipalpus* group		
monticola Belkin & McDonald	20	-
muelleri Dyar (now in *muelleri* group)	21	-
sierrensis (Ludlow)	22	40
varipalpus (Coquillett)	23	-
Group 7 (F) *scapularis* group		
atlanticus Dyar & Knab	24	2
aurifer (Coquillett)	25	3
dupreei (Coquillett)	26	-
infirmatus Dyar & Knab	27	-
scapularis (Rondani)	28	27
thelcter Dyar	29	34
thibaulti Dyar & Knab	30	35
tormentor Dyar & Knab	31	36
trivittatus (Coquillett)	32	38

Group 8 (G) *communis* group		
aboriginis Dyar	33	1
abserratus (Felt & Young)	34	-
cataphylla Dyar	35	8
communis (De Geer)	36	9
decticus Howard, Dyar & Knab	37	-
diantaeus Howard, Dyar & Knab	38	10
hexodontus Dyar	39	16
impiger (Walker)	40	18
implicatus Vockeroth	41	-
intrudens Dyar	42	19
nigripes (Zetterstedt)	43	21
niphadopsis Dyar & Knab	44	23
pionips Dyar	45	-
pullatus (Coquillett)	46	24
punctodes Dyar	47	-
punctor (Kirby)	48	25
rempeli Vockeroth	49	-
schizopinax Dyar	50	-
sticticus (Meigen)	51	31
ventrovittis Dyar	52	-
Group 9		
canadensis (Theobald)	53	6
Group 10		
spencerii (Theobald)	54	29
*idahoensis** (Theobald)	64	17
(now a subspecies of *spencerii*)		
Group 11 (H) *rusticus* group		
bicristatus Thurman & Winkler	55	4
trichurus (Dyar)	56	37
Subgenus *Kompia* Aitken		
purpureipes Aitken	57	-
Subgenus *Aedes* Meigen		
cinereus Meigen	58	45
Subgenus *Aedimorphus* Theobald		
vexans (Meigen)	59	44
Subgenus *Finlaya* Theobald		
Group B (*terrens* group)		
atropalpus (Coquillett)	60	39
(now in *Atropalpus* group in *Ochlerotatus*)		
Group H (*geniculatus* group)		
hendersoni Cockerell	61	-
(now in *triseriatus* group in *Protomacleaya*)		
triseriatus (Say)	62	41
(now in *triseriatus* group in *Protomacleaya*)		
Subgenus *Stegomyia* Theobald		
aegypti (Linnaeus)	63	43
*zoosophus** Dyar & Knab	65	42
(now in *zoosophus* group in *Protomacleaya*)		
*albopictus** (Skuse)	66	46
*chemulpoensis** Yamada	67	47
*albolineatus** (Theobald)	68	48

TABLE II. Matrix of Correlations between the MDSCALE axes (shown in Figure 3) and the 14 setal characters from Lunt and Nielsen (1971a).

Variables*			I	II	III
			Axes		
1	Mesonotum	pl	0.861	0.366	0.191
2		al	0.886	-0.251	0.008
3		ac	0.742	-0.218	-0.286
4		v	0.332	0.671	-0.608
5		dc	0.897	0.051	0.131
6	Scutellum		0.894	0.207	0.244
7	Pleuron	apn	0.841	0.126	0.089
8		ppn	0.863	-0.281	0.018
9		ppl	0.922	-0.008	0.071
10		psp	0.894	0.216	0.096
11		pa	0.932	0.117	0.041
12		ume	0.853	0.060	-0.107
13		lme	0.672	-0.402	-0.133
14		stp	0.745	-0.067	0.323

* Codes for the setal groups are as defined by Lunt and Nielsen (1971a).

TABLE III. Primary Pattern Matrix resulting from a factor analysis of the Lunt and Nielsen (1971a) setal data.

	Factors			
Variables*	I	II	III	IV
pl	0.022	-0.280	0.962	0.239
al	-0.093	0.773	-0.003	0.249
ac	0.210	0.923	-0.100	-0.116
v	0.895	-0.005	-0.001	-0.005
dc	0.044	0.469	0.648	-0.189
Scutellum	-0.077	0.027	0.947	0.008
apn	0.064	0.268	0.712	-0.129
ppn	-0.067	0.757	-0.005	0.225
ppl	0.005	0.489	0.436	0.072
psp	0.016	-0.040	0.683	0.320
pa	0.061	0.159	0.368	0.488
ume	0.081	0.115	-0.012	0.852
lme	-0.042	0.700	-0.390	0.458
stp	-0.083	0.157	0.534	0.118

* See TABLE II.

TABLE IV. Correlations among the primary axes for the Lunt and Nielsen (1971a) setal data.

Factors	I	II	III	IV
I	1.000			
II	0.215	1.000		
III	0.414	0.765	1.000	
IV	0.358	0.687	0.747	1.000

APPENDIX

Lunt and Nielsen (1971a) data

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	59.9	11.5	20.2	5.8	40.6	39.1	27.5	6.7	26.4	7.5	19.4	18.8	1.0	17.3
2	52.8	8.7	23.6	4.6	37.7	44.3	29.8	7.4	22.3	7.4	18.0	16.9	1.0	13.6
3	77.6	19.4	24.7	6.1	49.7	63.1	29.7	9.6	24.2	10.4	29.3	25.8	1.1	17.5
4	66.3	11.7	26.2	6.5	50.8	56.7	32.5	8.2	20.8	7.6	22.6	21.2	0.5	13.5
5	60.0	12.5	22.0	6.5	35.5	37.5	34.0	10.5	28.5	11.5	27.0	32.5	0.0	14.5
6	68.8	12.1	25.9	7.8	40.6	51.0	35.3	13.6	28.3	11.5	26.0	27.7	0.0	16.3
7	77.3	14.1	27.7	8.1	48.2	58.7	37.4	11.9	35.2	11.9	27.5	28.5	3.1	19.1
8	62.0	9.1	21.0	7.1	35.4	45.6	33.2	10.2	35.3	9.0	21.7	23.4	0.3	17.7
9	62.0	11.1	22.7	8.0	38.6	47.2	31.1	9.2	30.9	9.6	22.4	24.4	1.3	17.3
10	86.8	14.7	28.5	6.9	49.4	65.9	39.1	11.9	43.7	13.2	30.0	32.0	0.3	23.0
11	64.2	13.6	26.2	6.9	45.4	44.0	32.0	9.5	23.5	9.7	22.8	23.2	3.7	16.1
12	75.3	11.9	22.4	6.3	37.3	52.4	35.7	9.7	35.3	10.1	30.5	28.4	0.0	17.9
13	63.1	19.2	24.3	7.6	47.5	40.8	34.7	8.5	23.7	6.2	22.3	19.8	3.2	13.7
14	73.3	13.7	28.5	7.3	48.0	50.3	36.9	10.0	34.2	12.8	26.1	29.5	3.2	19.9
15	68.1	9.1	11.5	7.4	34.6	50.2	31.5	9.4	39.9	7.8	23.8	24.5	2.3	19.1
16	42.1	12.0	27.8	7.0	33.0	35.9	30.9	6.7	19.8	6.5	15.4	11.0	0.0	17.5
17	82.0	16.5	10.7	6.3	46.7	61.3	35.6	8.2	28.7	13.5	28.3	30.8	4.6	20.7
18	91.1	17.1	13.9	6.3	53.1	63.9	34.6	7.5	28.5	13.9	28.7	32.6	5.0	22.4
19	76.4	12.1	10.6	5.7	41.8	54.5	31.9	7.1	24.2	10.2	25.3	25.6	4.0	19.6
20	67.3	10.5	26.1	6.1	39.8	39.1	34.2	10.3	28.7	8.6	23.6	24.9	0.0	26.1
21	53.0	7.4	23.7	5.7	37.2	39.8	31.1	9.7	20.7	9.5	19.7	23.3	0.0	13.7
22	40.4	8.0	14.0	6.8	33.6	26.1	19.8	5.6	9.6	5.5	15.4	16.1	0.1	14.4
23	45.0	8.2	16.7	6.1	29.2	26.2	22.4	6.4	17.1	5.8	18.0	14.2	0.0	21.0
24	58.4	5.9	8.5	6.1	35.0	40.6	30.1	7.9	23.6	8.6	13.2	15.7	0.0	18.7
25	68.0	6.7	8.8	5.2	32.3	40.7	32.7	9.0	21.1	10.1	20.9	22.5	0.1	13.9
26	39.5	3.7	7.0	4.8	33.4	27.7	23.7	7.5	12.2	6.4	12.5	13.9	0.0	14.6
27	58.4	5.2	5.4	6.0	29.7	37.7	29.5	6.0	19.1	8.6	16.0	14.2	0.0	15.1
28	66.8	5.3	8.1	7.4	31.2	40.4	28.4	7.4	16.2	8.7	21.3	15.9	0.0	19.4
29	66.5	9.7	9.1	7.4	34.0	45.5	33.1	7.6	24.7	7.7	22.1	23.7	0.0	13.3
30	43.8	2.8	7.6	5.3	25.1	29.8	21.9	6.0	18.8	7.2	18.0	18.7	0.1	14.3
31	55.9	7.1	6.3	5.7	32.9	40.9	29.5	6.9	18.5	6.3	13.4	14.4	0.0	18.4
32	74.3	5.9	7.2	6.4	36.6	43.7	30.9	7.1	16.5	8.3	17.8	18.1	0.0	18.9
33	94.5	16.1	29.3	7.4	50.9	52.5	33.4	13.2	35.4	13.0	26.0	32.0	1.7	20.3
34	61.4	7.2	19.9	5.8	34.4	47.7	33.9	10.3	32.9	9.4	21.7	25.6	2.0	14.2
35	56.8	16.8	21.3	6.0	35.3	36.4	26.2	10.9	25.1	8.0	23.2	26.2	5.1	17.7
36	60.3	11.6	26.3	6.9	41.1	43.5	30.7	10.3	27.4	8.5	24.6	30.4	4.2	15.3
37	40.5	6.2	16.3	4.2	22.2	26.8	20.3	5.7	8.3	5.0	12.2	12.5	0.0	6.7
38	53.1	8.0	24.3	5.7	37.1	46.6	29.2	9.2	19.8	9.4	22.2	22.1	0.3	13.2
39	70.5	17.5	25.4	6.4	44.4	47.3	28.5	10.7	29.7	11.7	28.3	32.4	3.6	17.0
40	66.6	25.0	33.8	5.6	51.8	49.6	36.3	21.6	42.1	8.6	30.9	30.7	11.2	24.2
41	55.9	10.1	22.0	6.6	31.9	35.9	26.8	8.0	17.3	7.7	21.2	20.0	2.4	14.7
42	60.2	11.3	25.1	6.2	37.2	46.0	28.1	9.1	19.1	9.2	23.2	21.8	0.7	17.7
43	93.8	38.0	40.9	5.3	61.5	74.9	42.3	28.2	61.7	17.5	34.3	34.1	8.3	29.7
44	67.3	22.7	29.7	6.1	47.1	53.6	30.8	12.0	33.8	9.2	23.6	23.6	4.1	17.5
45	71.1	10.8	31.1	8.2	47.2	51.7	28.8	10.9	33.2	12.6	30.2	33.7	1.7	17.7
46	68.4	14.5	28.8	6.9	40.6	46.4	34.9	15.7	33.5	13.3	29.6	39.3	3.5	16.8
47	111.4	26.7	34.3	7.4	54.4	65.4	38.6	16.4	44.6	16.8	37.9	53.7	4.0	24.4
48	64.7	9.3	23.5	5.8	39.5	44.9	30.1	9.8	27.5	11.8	24.9	26.5	2.7	14.7
49	39.0	12.5	25.5	4.0	43.5	41.5	37.0	8.5	31.5	7.0	14.0	16.5	1.5	13.0
50	55.0	14.4	24.3	6.5	38.3	47.4	28.1	7.7	26.4	9.3	21.6	22.4	2.9	14.9
51	62.2	7.9	19.6	5.1	37.4	37.3	27.5	6.3	17.2	7.2	21.2	16.4	0.2	12.8
52	72.5	25.6	28.2	4.4	50.4	60.2	31.5	13.5	40.1	10.6	34.1	33.8	0.0	31.1
53	55.3	6.3	20.5	7.1	33.9	35.9	24.1	7.7	17.9	8.4	19.0	17.4	0.0	14.6
54	76.6	14.2	6.7	5.9	40.6	54.5	26.2	9.5	26.4	10.4	28.9	30.3	0.1	15.4
55	65.3	24.4	27.1	6.8	39.0	50.3	32.7	14.7	34.2	12.5	32.5	55.0	9.2	22.3
56	70.5	21.8	28.7	6.4	41.7	47.8	37.0	11.6	35.8	10.9	30.9	39.8	6.3	19.1
57	33.0	3.8	2.7	6.0	19.0	21.2	10.1	3.2	4.6	0.0	7.7	13.9	0.0	4.3
58	44.6	6.6	19.3	5.8	29.9	30.2	14.6	5.1	11.2	4.8	12.9	13.0	0.2	10.7
59	64.6	11.0	21.9	7.2	41.9	44.0	27.1	7.0	25.3	8.8	21.8	19.6	0.7	18.2
60	42.6	2.4	5.7	4.4	19.4	31.0	16.3	4.7	7.4	4.9	12.7	16.7	0.0	11.9
61	47.0	4.9	11.1	6.5	30.7	26.8	19.7	4.7	12.7	6.7	14.7	15.1	0.0	16.2
62	49.2	3.2	9.8	5.9	29.4	30.9	21.7	6.1	16.4	8.5	16.8	18.9	0.0	24.1
63	30.7	1.5	2.2	4.9	17.1	17.1	10.2	4.2	5.0	2.6	9.1	4.5	0.0	4.5

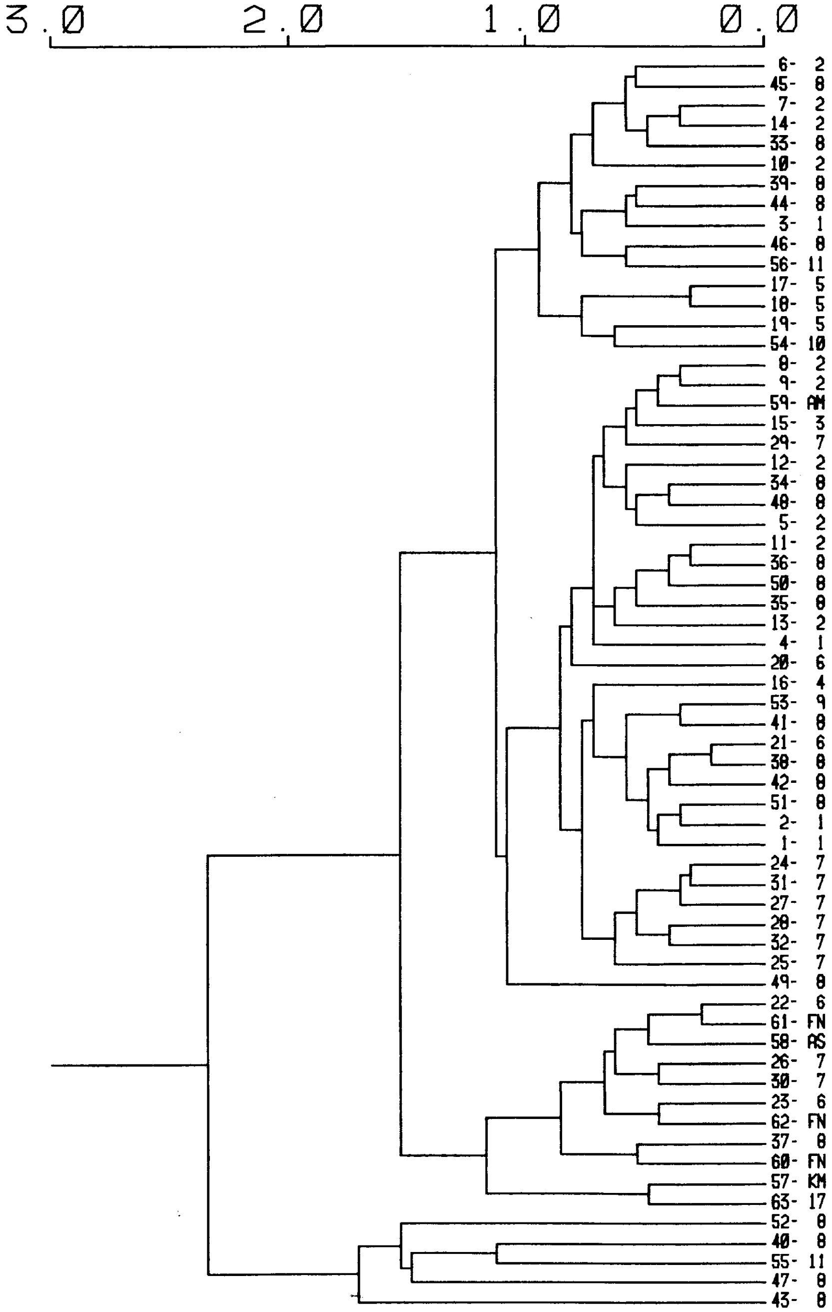

Figure 1. Phenogram based on UPGMA cluster analysis of the matrix of distances between all pairs of species for the Lunt and Nielsen (1971a) data. Abscissa is average Euclidean distance. Species code numbers (and group assignments indicated by the codes) are given in TABLE I. Cophenetic correlation is 0.741.

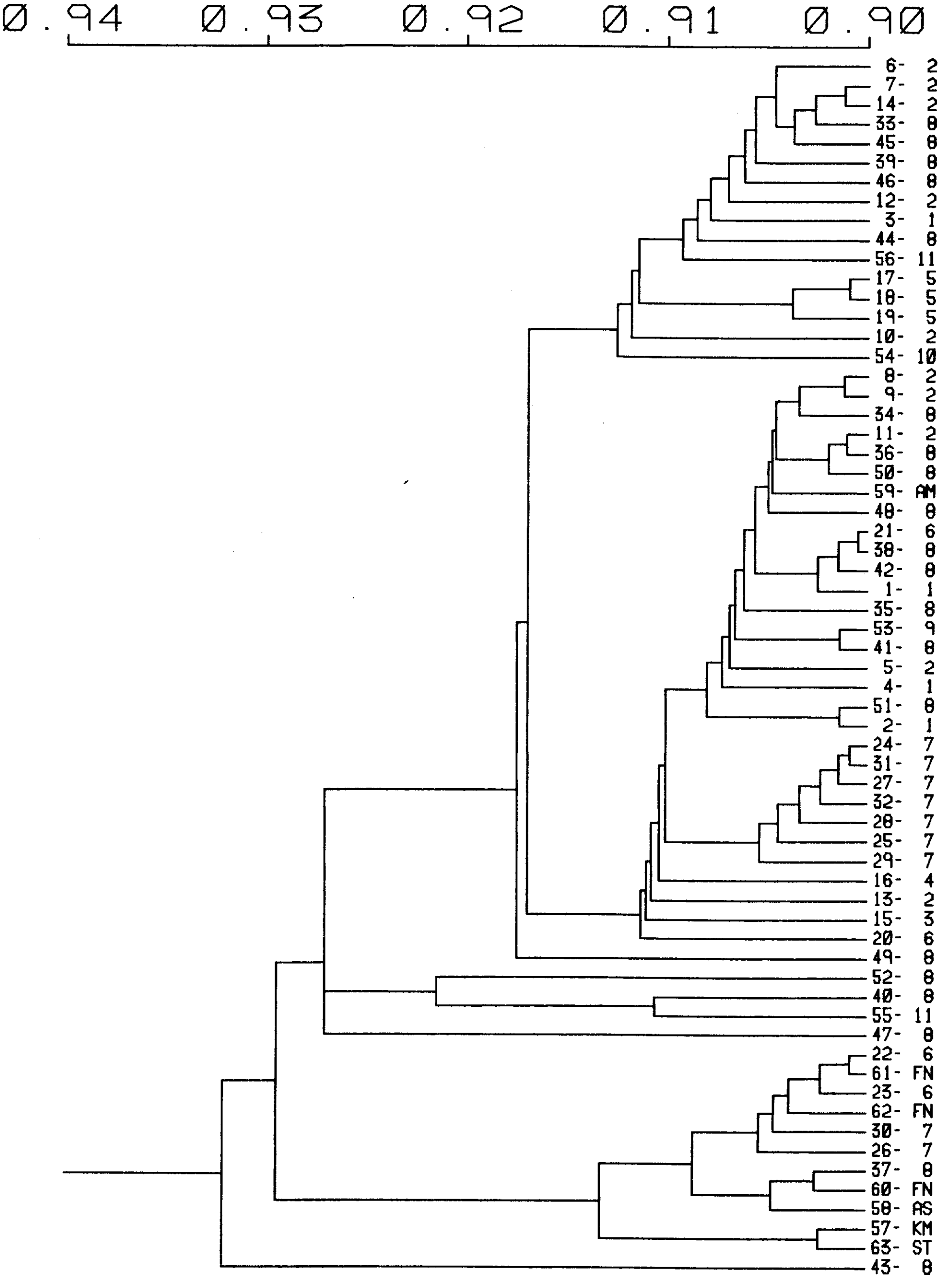

Figure 2. Phenogram based on AHCS cluster analysis based upon the standardized Lunt and Nielsen (1971a) data. The abscissa is (p-volume)$^{1/2p}$. The *w* parameter was set equal to 0.9. Species code numbers are given in TABLE I.

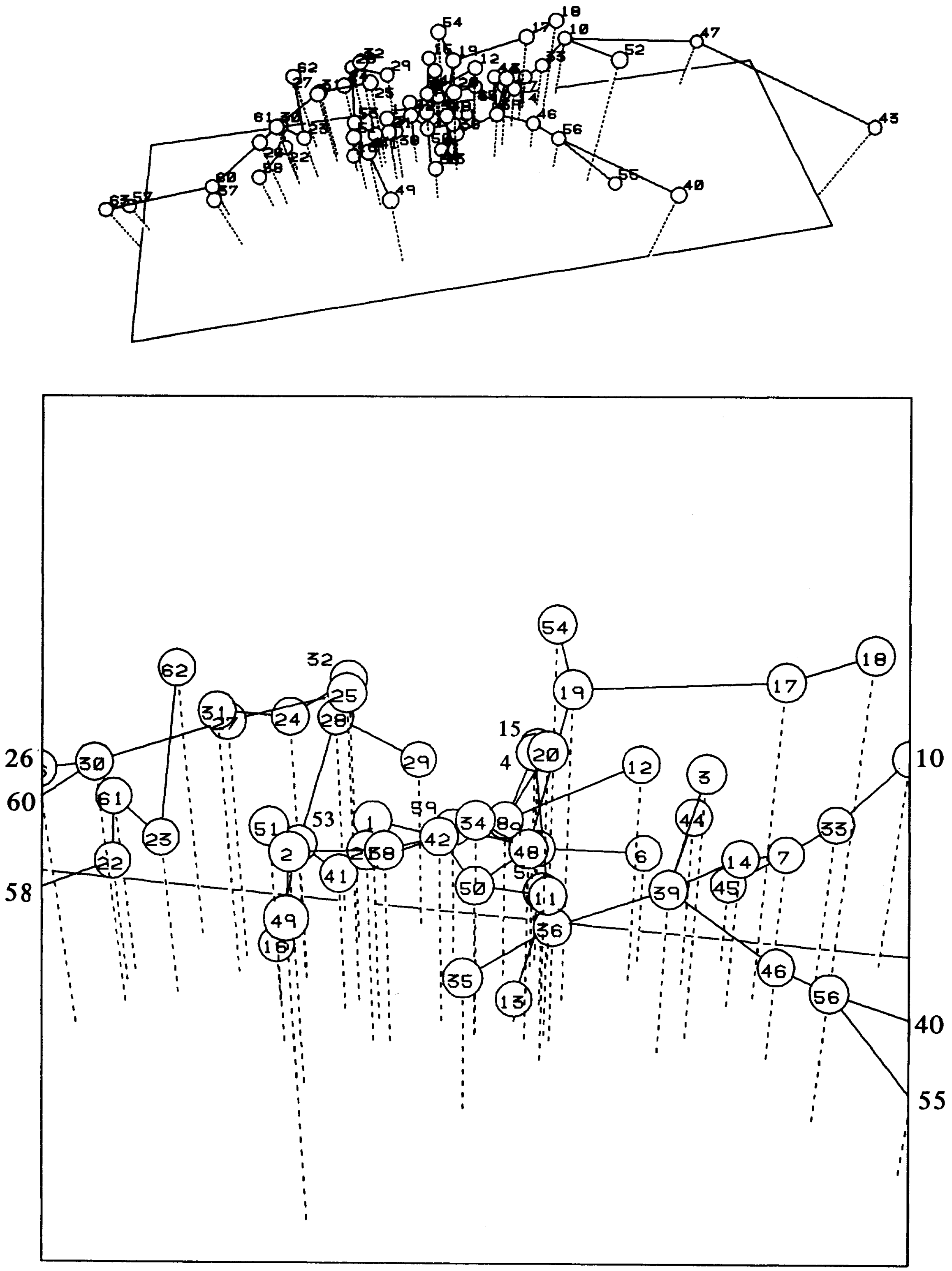

Figure 3. Perspective views of a 3-dimensional model showing the results of an MDSCAL analysis of the distance matrix based on the standardized Lunt and Nielsen (1971a) data. The minimum spanning tree is superimposed upon the figure. Stress = 0.108. Species code numbers are given in TABLE I.

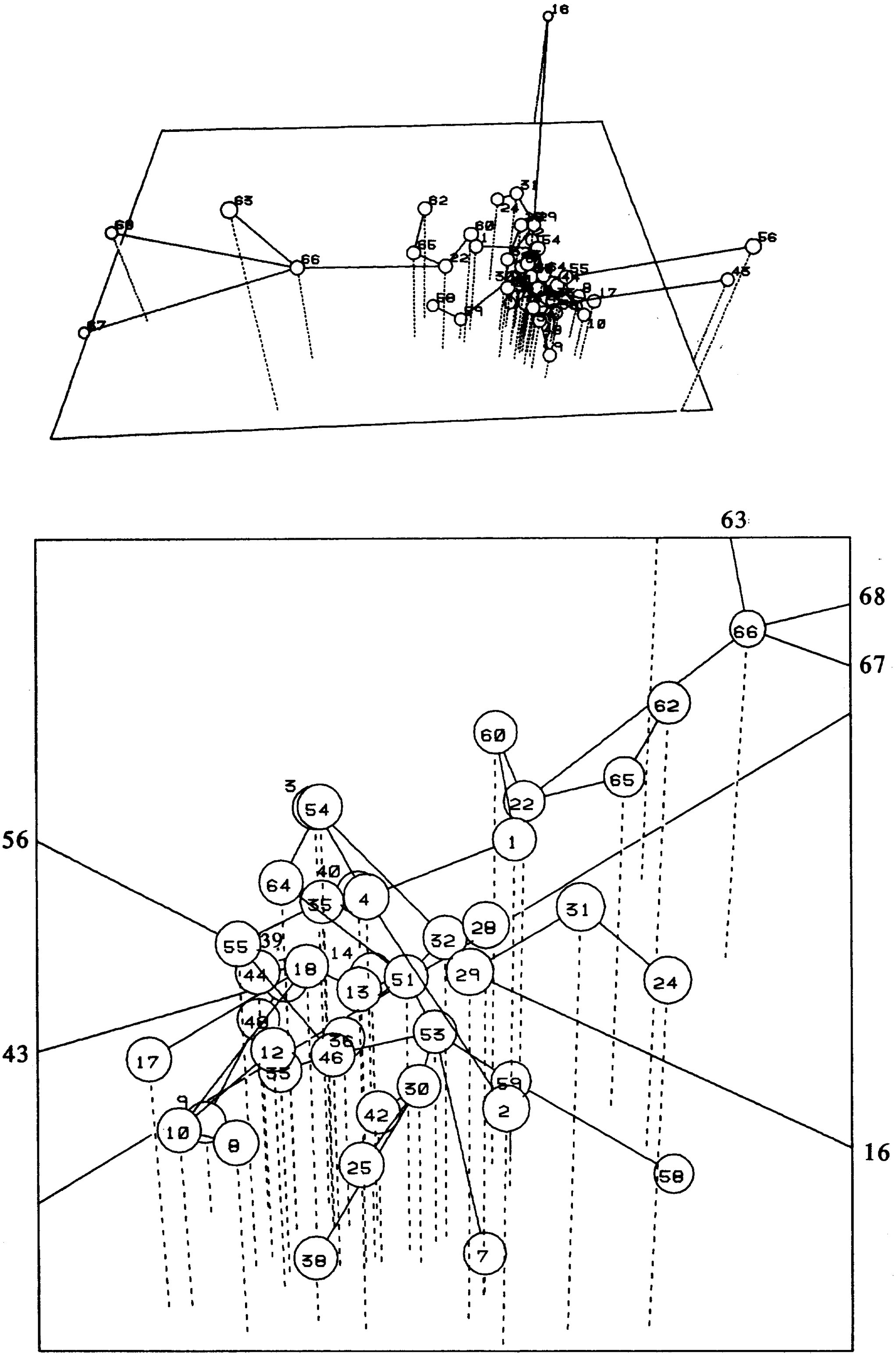

Figure 4. Perspective views of a 3-dimensional model showing the results of an MDSCAL analysis of the distance matrix for the adult + larval data of Rohlf (1963b). The close-up view is from the rear. The minimum spanning tree is superimposed upon the figure. Stress = 0.172. Species code numbers are as given in TABLE I.