TAXONOMIC WORKING SYSTEM BY COMPUTER (SAC) WITH APPLICATION TO JAPANESE ANDRENID BEES

Tadauchi, Osamu

https://doi.org/10.5109/2419

出版情報：ESAKIA. 17, pp.161-182, 1981-11-30. Entomological Laboratory, Faculty of Agriculture, Kyushu University
バージョン：
権利関係：

ESAKIA, (17) : 16-I-182. 1981

TAXONOMIC WORKING SYSTEM BY COMPUTER (SAC) WITH APPLICATION TO JAPANESE ANDRENID BEES*

OSAMU TADAUCHI

Entomological Laboratory, Faculty of Agriculture,
Kyushu University, Fukuoka 812, Japan

Abstract

A taxonomic working program package named SAC (system for Andrena classification) by computer is presented. The system is widely applicable for taxonomic works on organisms by changing user's data sets. The system includes 17 main taxonomic programs, such as character analysis, taxa grouping, specimen identification, key construction, description printing, and distribution mapping. Examples of the use of the SAC system are shown, using lists from a teletype terminal based on two main data sets of Japanese andrenid bees.

Introduction

There are two general reasons for using a computer to process taxonomic data (Sokal and Sneath, 1966; Mello, 1970; Morse, 197413). One is to perform traditional tasks more efficiently, e.g., use of computer for writing or editing keys, printing taxonomic descriptions or mapping out geographical localities. The other is to introduce novel tasks, such as numerical taxonomy, massive data analysis and various procedures for searching large data files in response to particular requests, which were previously impractical or impossible.

Libraries of taxonomic data matrices and computer programs may play a major future role in information processing retrieval in systematic biology and natural history. Many of the potential use of taxonomic data banks are also possible with information networks of data matrices. If an information network of data matrices on important pests is established, it will be of great use in applied entomology. A number of computer-based taxonomic information systems have been developed in recent years (e.g., MSU system by Morse, 1974a; DELTA system by Dallwitz, 1980). The author has also edited a taxonomic working system, which is named a system for Andrena classification, SAC in short. In the present study the SAC system is explained first and

---

* Contribution from the Entomological Laboratory, Faculty of Agriculture, Kyushu University, Fukuoka (Ser. 3, No. 98).

then some of the examples of the system are cited using Japanese andrenid bees.

Outline of the SAC system

The SAC system is a program package of a working system for *Andrena* classification. It is widely applicable for taxonomic works on organisms by means of changing data sets. The system is designed to aid researchers in various information-processing tasks of systematics by providing methods ranging from simple data description to advanced multivariate statistical techniques. The computer language used in the SAC system is only FORTRAN. It is a programming language widely employed in scientific and engineering works.

This system can be practiced readily from a local terminal in a user's laboratory by using a time-sharing system (TSS). It is aimed that even if the user is a beginner, he can have easy access to the SAC system by the guide of a dialogic time sharing system. The minimum equipment, as a terminal, is a character display (Fig. 1) or a teletype terminal (Fig. 2) to access the system. Since FACOM M-200 system on which the SAC system is implemented offers a wide system-support software such as programs, editing procedures, command sequences, the general user of the SAC system has only to study an outline of the TSS and to prepare his own data sets.

The SAC system has three functions and consists of 17 main programs which are interlaced one another. Each module performs a major kind of taxonomic information processing. The various main data processing modules perform the following operations.

1) Making data set
    Data accumulation (SAC 11)
    Data checking (SAC 12)
    Data editing (SAC 13)

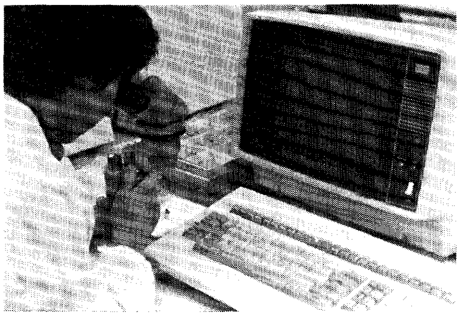


Fig. 1. Character display.

Fig. 2. Teletype terminal.

2) Data analysis
  - Character analysis by PCA (SAC 21)
  - Character analysis by factor analysis (SAC 22)
  - Character analysis by factor analysis with rotation (SAC 23)
  - Character grouping by belonging-coefficient (SAC 24)
  - Discrimination between two taxa by discriminant analysis (SAC 25)
  - Discrimination among three taxa and over by discriminant analysis (SAC 26)
  - Taxa grouping by cluster analysis (SAC 27)

3) Data retrieval
  - Specimen identification (SAC 31)
  - Key construction (SAC 32)
  - Taxon-taxon comparisons (SAC 33)
  - Diagnostic character retrieval (SAC 34)
  - Description printing (SAC 35)
  - Inversion of taxon-by-character lists (SAC 36)
  - Distribution mapping (SAC 37)

Fig. 3 shows a flow of the SAC system.

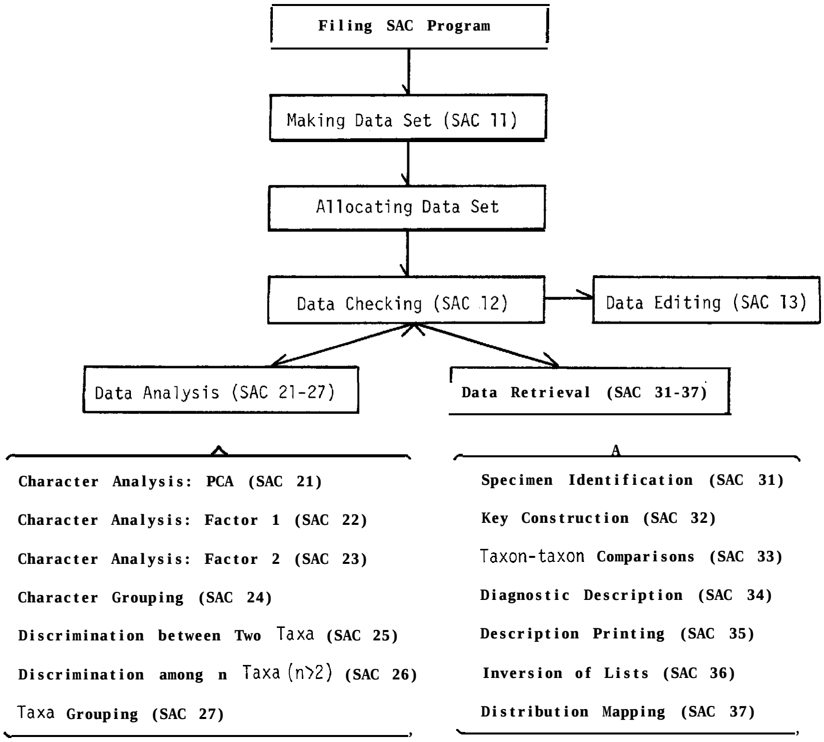


Fig. 3. A flow of the SAC system.

The SAC program can read three kinds of data sets as follows:

1) OTU (taxon) name list
2) OTU (taxon)/character data list
3) Character name list

In this system both simple (dichotomous characters only) and general data (including multistate or quantitative characters) are available. The present programs are limited to data matrices of not more than 120 OTU's and 130 characters. However these limits can be readily changed by redimensioning most of the vectors and arrays in the programs and subroutines.

At the beginning the user must prepare his data sets. He has only to input his own data according to the instructions from the SAC system. After construction of the data sets, he must allocate each data set to each file in the computer by using "command". MAIN program reads a data file into central storage, then activates a taxonomic main program at the user's request. Data are usually analyzed by an interactive "data checking" module. Firstly, the data are examined for unreasonable values. If unreasonable values are found, they are checked, if possible, corrected. An analysis is then performed. If he wants to edit the data, he can call "data editing" module. Since the SAC system has a feature of a program-chaining, the MAIN program can call a user's requested program by inputting the code number of the program. Upon completion of the user's requested program, the program can be stopped or passed back to the MAIN program for activation of a different program.

Examples of the SAC system

Some of the examples of the use of the SAC system are cited, using printout lists from a teletype terminal.

1. Control of a session in TSS

This is an example on the FACOM M-200 system. When a user wants to open a session from a character display or a teletype terminal, he must keypunch a command "LOGON". Commands are languages used for controlling a session in the TSS. There are many kinds of commands available such as editing, executing, and saving programs. After keypunching a command "LOGON", the user must input "TSS", the user's subject number and a maximum storage size required. As the result the system asks the user for his private password, If the password is agreed with that previously registered, a session is opened and a system message "READY" is output at the terminal, This state is called "command mode" in which various commands are able to input. On the other hand, when the user wants to end the session, he must keypunch a command "LOGOFF" in "READY" mode.

Then, the user must allocate his data set to a file in the computer. For instance, when he allocates a data set "SACDATA 2" to a file "FT08F001", he must keypunch "ALLOCATE FILE (FT08F001) DATASET(SACDATA 2. DATA) SHR".

Next, the user use a command "EDIT" for editing a program, followed by a program name (e.g., SAC 1) and the language form (e.g., FORT(FIXED)). If the program named 'SAC 1" is already stored in the computer, a system message "EDIT" is output and the "command mode" is turned to "edit mode" in which various subcommands for editing the program are available. When commands and subcommands are used, either full names or shortened names are usable. For instance, "ALLOC" is the shortened command of "ALLOCATE". Inputting subcommand "RUN" (or shortened "R") for executing the program in the "edit mode", the "FORTRAN GE" compiler is entered. If there is no errors in the program, the system message "END OF COMPILATION" is output and then the "SAC 1" program begins and prints head lines.

After completion of the program with no errors, the system outputs "END OF GO, SEVERITY CODE=O" and turns to the "edit mode". When the user wants to end the mode, he must use a subcommand "END". Then the system changes the "edit mode" to the "READY" mode.

2. Data accumulation (Prints 1 and 2)

After compilation of the program, the SAC system outputs general information as shown in Print 1. Since the system is constructed as a dialogic type, the user must answer to messages from the terminal. For instance, when a first message "DO YOU MAKE DATA SET ? 1=YES 2=NO, ALREADY FILING" followed by five figures and question mark are output, he has only to input either "1" in case of "YES" or "2" in case of "NO" after the question mark. As shown in the example (Print 2) when the system requires instructions from the user, it always outputs a message, followed by five figures and question mark. In the example the user keypunches a number "1" because data sets are not prepared. As a result MAIN program calls "data accumulation" program (SAC 11). After printing documentation, the program asks the user what kind of data set he makes. There are five modes for producing data set. First is making an OTU name list in case of dichotomous characters only, second is producing an OTU name list in case of multistate or quantitative characters, third is an OTU/character data list, fourth is a character name list and the fifth is a stop mode.

Example of making an OTU name list is shown in Print 2. The list consists of two parts, a documentation section and an OTU list. At the beginning the user inputs data set name within 16 letters, the number of OTU's within 3 numerals according to the instruction of the SAC system. In the example he keypunches "ANDRENA TEST 1 21 20". Namely the data sets consist of 20 dichotomous characters chosen from 21 OTU's of andrenid bees. This documentation section provides parameters for storage allocation purposes of the computer system such as the dimensions of the matrix, i.e., the number of characters and the number of OTU's. Next the program requires the user for an OTU code (within 3 numerals) and an OTU name (within 30 letters). The OTU list gives the name for each OTU along with an arbitrary reference number between 1 and 999 assigned to each OTU. In the example OTU code number 1 is "A. (ANDRENA) MIKADO A" and the last OTU coded 21 is "A. (LARANDRENA) ECHIZENIA". This data set is ended by inputting an arbitrary reference number -99 and then the program outputs "END OF TAXON NAME LIST" and "END OF DATA ACCUMULATION". In the example the user forgets to input a species name of No. 16. Afterwards this is corrected.

3. Data editing

There are two methods for editing data sets. One is using a command "EDIT" of the TSS in the system. The other is using "data editing" program in the SAC system. The OTU name list produced in the preceding section has a few mistakes, such as lacking the species name coded No. 16. In such a case, the user can use a subcommand "CHANGE" for editing programs in the "edit mode". The user must keypunch the subcommand "CHANGE", followed by a line number required to correct, incorrect letters, and correct letters, interposing slants between them, respectively. For confirmation the user had better output the whole line corrected, He can know that the mistakes are corrected. After the complete data set is prepared it must be saved in the computer using a subcommand "SAVE".

4. Data sets

In the following sections two different main data sets are employed for the SAC system. One is the simple main data set, which consists of 20 female dichotomous characters chosen from 21 OTU's, one OTU from one subgenus of the genus *Andrena* (Hymenoptera, Andrenidae) of Japan. The other is the general main data set, which is made up of 130 female characters including multistate ones chosen from 85 OTU's of the same genus in Japan.

The data set includes three kinds of taxonomic matrices, one of which is made in the Print 2. They are as follows:

1) OTU name list (data set name: SACDATA 2)
2) OTU/character data list (data set name: SACDATA 4)
3) Character name list (data set name: SACDATA 6)

The last mentioned list is used for several main programs required for printing output in word form such as key construction and description printing. In the dichotomous character list, one character consists of three lines. One is mostly noun phrase such as *clypeus* along with character code number "1" and its state code number "0" for convenience. Another is a following adjective phrase such as *without median impunctate line* along with the character code number "1" and its state code number "1". The other is the opposite adjective phrase of the dichotomous state such as *with median impunctate line* along with the character state number "1" and its opposite state code number "3". Therefore, one dichotomous character is coded as a single phrase followed by two constructing adjective or modified phrases. These three entries are given separate line numbers, then all given the same character number, and then each is given the different state number. Thus, the first couplet of the SACDATA 6 list, *clypeus without median impunctate line vs. clypeus wiih impunctate line,* is coded 11 and 13, respectively.

The simple main data set is used in the following taxonomic programs as follows : data checking (SAC 12) ; specimen identification (SAC 31) ; key construction (SAC 32) ; taxon-taxon comparisons (SAC 33) ; diagnostic description (SAC 34). The general main data set is used in the following main programs : character analysis by PCA (SAC 21) ; character grouping (SAC 24) ; taxa grouping (SAC 27) ; description printing (SAC 35) ; distribution mapping (SAC 37).

5. **Data checking** (SAC 12: Print 3)

After making data set, the SAC program asks the user what kind of data checking he wants. The SAC program contains a syntax-checking option for proof-reading taxonomic data matrices. In the matrix-checking mode, various comments are printed as the matrix is processed. If reading stops because of a syntax error, the user can approximate the location of the error in the data file. Two modes of matrix checking, summary and detailed, are offered. In the detailed mode, a message is printed after every successful read operation, as shown in Print 3. An error can always be readily located when this mode of operation is employed, since error is in the next line of the matrix after the line about which the last message is printed. However, for long matrices the detailed mode of operation produces a huge amount of output, for which a summary mode is offered. In the summary mode, the system prints messages only at intervals throughout the program. Moreover if the matrices need no-checking in the second times or over, the user can select skip mode. When data matrices are read successfully in the detailed mode, they are printed successfully in the order of an OTU list and an OTU/character data list. If there is no syntactic error, concluding two messages "END OF DATA SET READING AND CHECKING" and "DATA SET SYNTACTICALLY ACCEPTABLE" are printed. Next, the user is asked whether the original data sets require any editing or not. When he wants to change something in the data sets, he may keypunch a word "YES". As a result subroutine EDITOR is called and a data editing program begins. Then the user is asked "WHICH PROGRAM DO YOU WANT NEXT ? (O=STOP)". He can select anyone from 14 main taxonomic programs (data analysis : 21-27, and data retrieval : 31-37). He has only to input a code number of the main program which he wants. If he wants to stop the program, he must input a numeral zero.

6. Character analysis by principal component analysis (SAC 21: Print 4)

This is one of the various character analysis programs in the SAC system. Principal component analysis (PCA) computes the principal components explaining the trends in variability between characters, taking the greatest trend first and sequentially removing the next greatest trend until all the variability is accounted for. Component scores of each OTU for the principal components are also computed and dotted onto the first three principal components. Both matrices of correlation and covariance are available for the analysis.

In case of keypunching a code number of a program "21", the system prints headline-information and asks the user for some instructions. Many modes for output-printing are offered, for example, "WILL YOU OUTPUT MEAN AND STANDARD DEVIATION ?, "WILL YOU OUTPUT MATRIX USED ?", "WHICH RESULT DO YOU WANT ? O=EIGENVALUES AND PERCENTAGES ONLY, 1 =ABOVE THE RESULT FOLLOWED BY INFORMATION OF 20 COMPONENTS, 2=ALL **THE RESULTS", "WILL YOU OUTPUT INDIVIDUAL SCORES FOR COMPONENTS** ?", **"WILL YOU DOT INDIVIDUAL SCORES ONTO THE** 1ST **THREE COMPONENTS?"** and "IF DOT, WHICH SYMBOL DO YOU USE? O=NUMBER, 1=*". Print 4 shows the example of printing eigenvalues and percentages up to **30** components on the basis of a correlation matrix of the 130 general characters derived from 85 OTU's. After the analysis the system asks "NEXT ? 1=RECYCLE 2=NEW PROGRAM 3=STOP". If the user wants to continue the same program with different options, he has only to input a numeral code "1" from the terminal. In case of requesting a new program he can return the program to MAIN program by keypunching a numeral code "2". When he wants to stop the SAC program, he can select a stop mode "3".

7. Character grouping by coefficient of belonging (SAC 24)

This program is a character grouping routine by using coefficient of belonging (B-coefficient) based on a correlation matrix. This is based on the program developed by Asano (1971). This coefficient is defined as the ratio of the average of correlation coefficients among characters clustered as a group to the average of correlation coefficients among all the characters. The formula is as follows:

$$B(k) = 100\left(\frac{S}{n_s}\right)\left(\frac{T}{n_t}\right), \quad k=1,2,\ldots.q$$

where $B(k)$ stands for the B-coefficient of characters clustered as a k-group, S is the sum of correlation coefficients among characters in a K-group, T is the sum of correlation coefficients among characters in a k-group and a not-K-group, n_s and n_t are the numbers of the correlation coefficients summed for S and T, respectively. At the beginning of this grouping, two characters which have the highest correlation coefficients are selected. Secondly, the third character is chosen in order to have the maximum total of correlation coefficients when they are combined together. Next, the fourth character is selected in the similar fashion and the coefficient of belonging is calculated. If the coefficient shows a significant drop, the character is excluded. Thus, when the coefficient has a large drop, the characters are grouped except for the last candidate. Eighteen groups are obtained from the result of the character grouping of the 130 general characters in case of the significant drop of 1.0.

8. **Taxa grouping by cluster analysis** (SAC **27:** Print 5)

Cluster analysis is a procedure by which similar individuals or units are gathered as clusters based on similarities or dissimilarities among a set of characters. Print 5 shows an example of SAC 27 program using the general main data set. In this program both eight kinds each of coefficients and clustering methods are available. As to the coefficient, euclidian distance, weighted euclidian distance, standardized euclidian distance, mahalanobis distance, sum of cross products of standardized deviation, correlation, matching and similarity ratio are usable. The last two coefficients can be only used in the dichotomous characters. On the other hand, as to the clustering methods, nearest neighbor (single linkage), furthest neighbor (complete linkage), median, flexible, centroid, group average (UPGMA, unweighted pair-group using arithmetic averages), the Ward, and flexible group-average methods are sevriceable. This program also contains various output options as follows: printing only the cycle number in case of skipping intermediate cycle; printing every intermediate results ; and printing only the results of expected cluster or printing all the results. The user can select his choice. When the user wants only a summary, he can take a result in his expected cluster. Print 5 shows first half lines in a summary of 21 expected clusters by the group average method based on standardized euclidian distance coefficient. If the user wants to know every clustering process, he can select an option of printing every intermediate result. He can get information about which two OTU's cluster together and about which value of similarity or dissimilarity they fuse.

9. Specimen identification (SAC 31: Print 6)

This is one of the data retrieval programs using a threshold search algorithm. The routine is used to provide suggested identifications for individual specimens. It is designed for interactive use at on-line terminal. In this program a set of possibilities is delimited, some characters of the unknown specimen are requested from the user, and the differences between each OTU's data and the specimen's characters are examined. On the basis of the data submitted, possibilities having the lowest tallies are output as the most similar to the specimen. If several possibilities remain likely, the entire process may be continued with additional characters, A user can ask the program to suggest some useful characters for continuing. Several other option are also available after a set of characters has been processed. All these can be submitted by keypunching the appropriate code number in response to the question "NEXT ?". The code numbers are explained in the instructions which can be printed after the message "NEXT ?". One of these options is to request the system to suggest some characters useful for distinguishing the remaining OTU's. Another is to list the remaining possibilities. The other choice is to delete the effects of the last submitted character set.

This program provides two different procedures for identification. They are monothetic and polythetic modes (Mayr, 1969). In the identification program, an initial set of possibilities is gradually reduced by the application of user-selected characters until one or a few OTU's remain as possible identifications. If the procedure requires more than one difference between specimen and OTU's in the matrix in order to eliminate an OTU from the list of possibilities, it is called polythetic. While the opposite procedure requires only one difference between them is called monothetic. This program allows the user to choose either of these two methods. In the monothetic mode the user has only to input a code number "1", while a code number "2 or over" indicates that polythetic elimination is desired for the current unknown specimen. In the polythetic mode, an OTU is not eliminated unless it has accumulated more differences from the unknown specimen than the code number specified.

Print 6 shows an example of identification for the specimen based on the simple data matrix. The program asks the user for several information such as the locality, the collecting date, the collector and the specimen number. For the specimen, it was collected at Fukuoka, on 9. IV. 1978, by 0. Tadauchi. For this specimen, the monothetic mode is used for identification. Next, the program asks for some characteristics of the specimen in question. In this case, the user can put a microscope on the terminal table and examine the characteristics of his specimen by a guide of character manual (Table 1). One character datum is to be formed by connecting a character number and the appropriate state code number. For instance, if the character code number is 3 and its state code number for the specimen is 1, the user can input a numeral "31". In the example, the user gives three features, 13, 43, and 81. These codes stand for *the clypeus with median impunctate line, the malar space well developed to elongate* and *the propodeal enclosure without a carina posteriorly,* respectively. Seven of the 21 OTU's remain after these characters are processed. The program asks for the next operation. The user selects option "2", to list the remaining seven OTU's. After having the list of the seven OTU's, the user requests to suggest useful characters, which are in order of decreasing importance value. Eight characters, coded 18, 20, 9, 19, 3, 5, 6, and 14, are listed. Continuing, the user eliminates three OTU's with the two characters 181

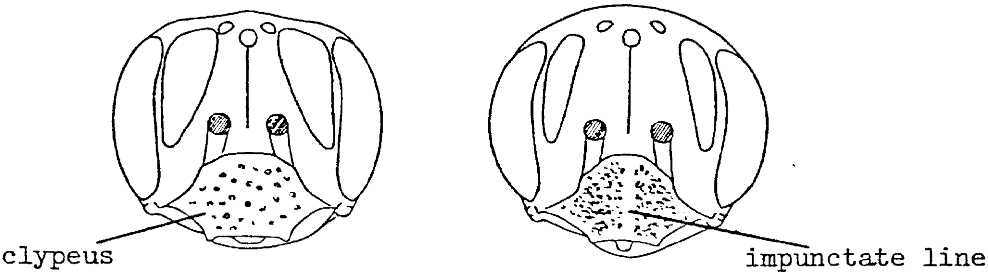
Table 1. An example of a guide to characters.

Genus *Andrena* (Hymenoptera, Andrenidae)

For simple (dichotomous) characters

Please read the following explanations of characters and their states and then examine your specimen. Next, please decide the states of characters of your specimen one by one.

1. Character No. 1.
   State No. 1: ***Clypeus without median impunctate line.***
   No. 2: Clypeus with median impunctate line.



State No. 1 State No. 2

2. Character No. 2.
   State No. 1 : ***Facial fovea narrowly separated from eye margin.***
   No. 2 : Facial fovea widely separated from eye margin.

and ***201.*** Four OTU's are **now** remaining. Next, the user inputs the third character set, 91, and 191. Thus, A. (*Euandrena*) *hebes* remains as a suggested one.

In this program if it can not identify the specimen, it outputs "NO TAXON APPROACHES THIS CHARACTER SET". Then the program asks the user for next instruction.

10. Key construction (SAC 32: Print 7)

SAC 32 program is a key construction routine for producing a dichotomous key to the OTU's in a data matrix, printing the key in phrases or code numbers of the characters at the terminal. After printing a head line, the program asks the user for a title for a key. He can keypunch the title in words as shown in the example (Print 7). Next, he must select printing mode of characters. The example shows first half lines of a production of a key for the 21 OTU's titled "ANDRENA TEST 1 IN PHRASES".

11. Taxon-taxon comparisons (SAC 33: Print 8)

In this program three kinds of OTU's-comparison modes are offered. One of them produces a summary of all possible two-OTU's comparisons for a data matrix. Another mode compares a selected OTU with all the others. The other compares two selected OTU's in detail.

Print 8 shows the comparison of two **OTU's,** with their respective character states for all differences being listed. As shown in the example, two OTU's coded 1 and 7, i.e., *A. (Andrena) mikado* A and *A. (Gymnandrena) watasei* are compared with each other. Six differences between them are indicated, namely for characters coded 5, 6, 9, 18, 19, and 20. With respect to the character coded 5, the former OTU has the state 3, i.e., the *pronotum with strong humeral angle,* while the latter possesses the state 1, i.e., the *pronotum with weak humeral angle or not.*

12. Diagnostic character retrieval (SAC 34: Print 9)

This program has three options as for selection of characters. One is printing characters in the order of the original character list, another is printing by decreasing importance values to the OTU, and the other is listing in an arbitrary sequence specified by the user. In all cases a user can instruct the program a printing limit of characters, i.e., the number of characters. The second option is particularly useful. A character shown by only one OTU in the group is at the top of that OTU's description as a highly diagnostic character, while the opposing state of that character, descriptive of all but one OTU, is listed at the bottom of every other description.

Print 9 shows the example of special character ranked description. It is the example of *A. (Parandrena) yasumatsui.* The five characters of *yasumatsui* are printed in decreasing order of their importance as peculiar features of that OTU when compared with others in the data matrix. Examination of data matrix shows that *yasumatsui* is only one **OTU** among 21 OTU's having *wings with the two submarginal cells.* From this option the above character state is selected in the first place and printed. With respect to the second character, there are five OTU's showing *the pronotal suture not indicated.* Furthermore, as to the fifth character, there are eleven OTU's showing *the propodeum with gradual angle posteriorly.*

13. Description printing (SAC 35: Print 10)

In the preceding routine the mode of printing all the characters provides enough descriptions. However, this program produces more arranged style of descriptions. The user has only to input a code number of an OTU which he wants. The program prints the OTU name and its sex, followed by description as shown in the example (Print 10). Each description in words is followed by its state code of the character. The example shows the first lines of the description of *A. (Andrena) brevihirtiscopa.*

14. **Distribution mapping** (SAC 37 : Print 11)

This program produces a distribution map for an OTU in Japan. The user has only to input an OTU code number which he wants. Print 11 shows an example of the map of *A. (Andrena) brevihirtiscopa.* After inputting a code number of a requested OTU, the program prints the OTU name and produces its distribution map of Japan. In the map two different marks are dotted. One (*) is for the type locality for that taxon, and the other (x) is for a record locality for it.

Summary

1. A taxonomic working program package named SAC (computer system for Andrena classification) is presented. The system is widely applicable for taxonomic works on organisms by changing user's data sets.
2. The system has three functions as follows : 1) making data set; 2) data analysis ; 3) data retrieval. This package includes the following 17 main programs : 1) data accumulation ; 2) data checking ; 3) data editing; 4) character analysis by PCA; 5) character analysis by factor analysis ; 6) character analysis by factor analysis with rotation; 7) character grouping by B-coefficient; 8) discrimination between two taxa by discriminant analysis; 9) discrimination among n taxa ($n>2$) by discriminant analysis ; 10) taxa grouping by cluster analysis; 11) specimen identification ; 12) key construction ; 13) taxon-taxon comparisons ; 14) diagnostic character retrieval; 15) description printing ; 16) inversion of taxon-by-character lists; 17) distribution mapping.
3. A user can easily access to the system using a character display or a teletype terminal on a timesharing computer system (TSS). Since the SAC system is edited as a dialogic type, the user is able to work easily and continuously on his data sets. The user has only to prepare his data sets and learn some information about TSS.
4. The SAC system can read three kinds of data sets as follows: 1) OTU (taxon) name list; 2) OTU/character data list; 3) character name list. In this system both simple (dichotomous characters only) and general (including multistate of quantitative characters) characters can be read.
5. Examples of the use of the SAC system are presented using a teletype terminal based on the two main data sets, namely, simple (20 characters derived from 21 OTU's) and general (130 characters derived from 85 OTU's) characters of the genus *Andrena* (Hymenoptera, Andrenidae) of Japan.

Acknowledgements

I am greately indebted to Prof. Y. Hirashima of Kyushu University for making this study possible and his valuable advices on andrenid bees. My sincere thanks are due to Prof. Ch. Asano of Research Institute of Fundamental Information Science of Kyushu University for his valuable advices on multivariate statistics and computer systems. I also wish to express my hearty thanks to Prof. Emerit. T. Okada of Tokyo Metropolitan University for his kind guidance to numerical taxonomy and encouragement. I am also grateful to Assoc. Prof. K. Morimoto of Kyushu University, Prof. S. Sakai of Daito Bunka University, Assoc. Prof. K. Suzuki of Toyama University for their helpful advices and suggestions.

References

Asano, Ch. 1971. *Factor Analysis.* Kyoritsu Shuppan, Tokyo. 481 pp. (In Japanese)

Dallwitz, M. J. 1974. A flexible program for generating identification keys. *Syst. Zool.*, *23:* *50-57.*

—— *1980.* User's guide to the DELTA system. A general system for coding taxonomic descriptions. *CSIRO Aust. Div. Entomol. Rep., (13) : 1-71.*

Hall, A. V. 1974. Museum specimen record data storage and retrieval. *Taxon,* 23: 23-28.

Mayr, E. 1969. *Principles of Systematic Zoology.* McGraw-Hill, New York. 428 pp.

Mello, J. F. 1970. Paleontological data storage and retrieval. *Proc. North Amer. Paleontol. Conv.,* **1969: 57-71.**

Morse, L. E. 1974a. Computer programs for specimen identification, key construction and description printing using taxonomic data matrices. *Pub. Mus., Michigan State Univ., Biol. Ser., 5 : 1-128.*

—— 1974b. Computer-assisted storage and retrieval of data of taxonomy and systematics. *Taxon, 23:* 29-43.

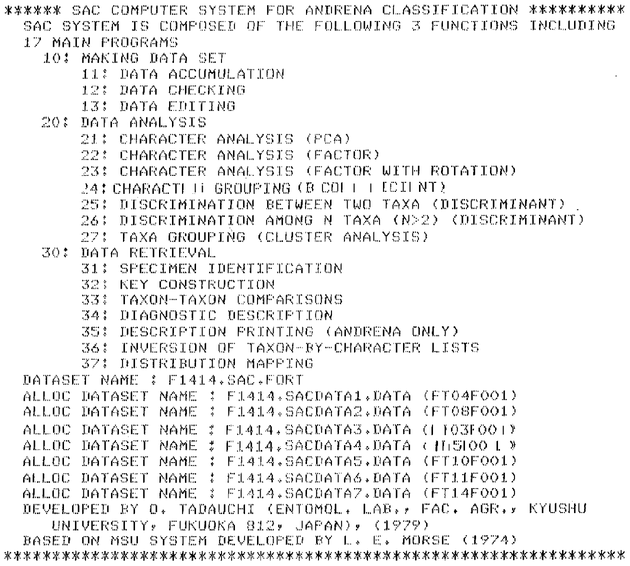
Pankhurst, R. J. 1974. Automated identification in systematics. *Taxon,* 23: 45-51.

—— 1978. The printing of taxonomic descriptions by computer. *Taxon,* 27: 35-38.

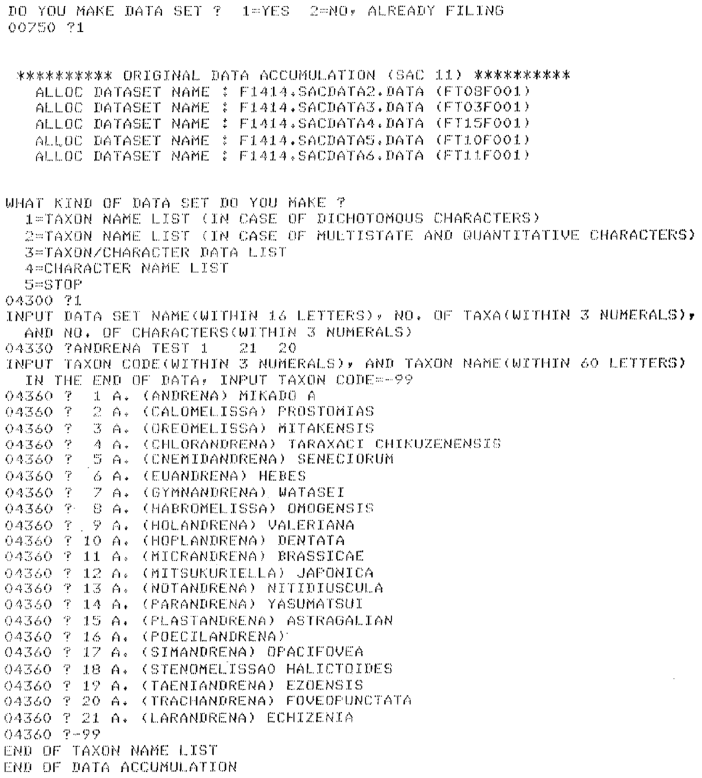
Rogers, D. J., H. S. Fleming, and G. Estabrook 1967. Use of computers in studies of taxonomy and evolution. *Evol. Biol.,* 1 : 169-196.

Sokal, R. R. and P. H. A. Sneath 1966. Efficiency in taxonomy. *Taxon, 15: l-21.*

Print 1 and **Print 2** (below)

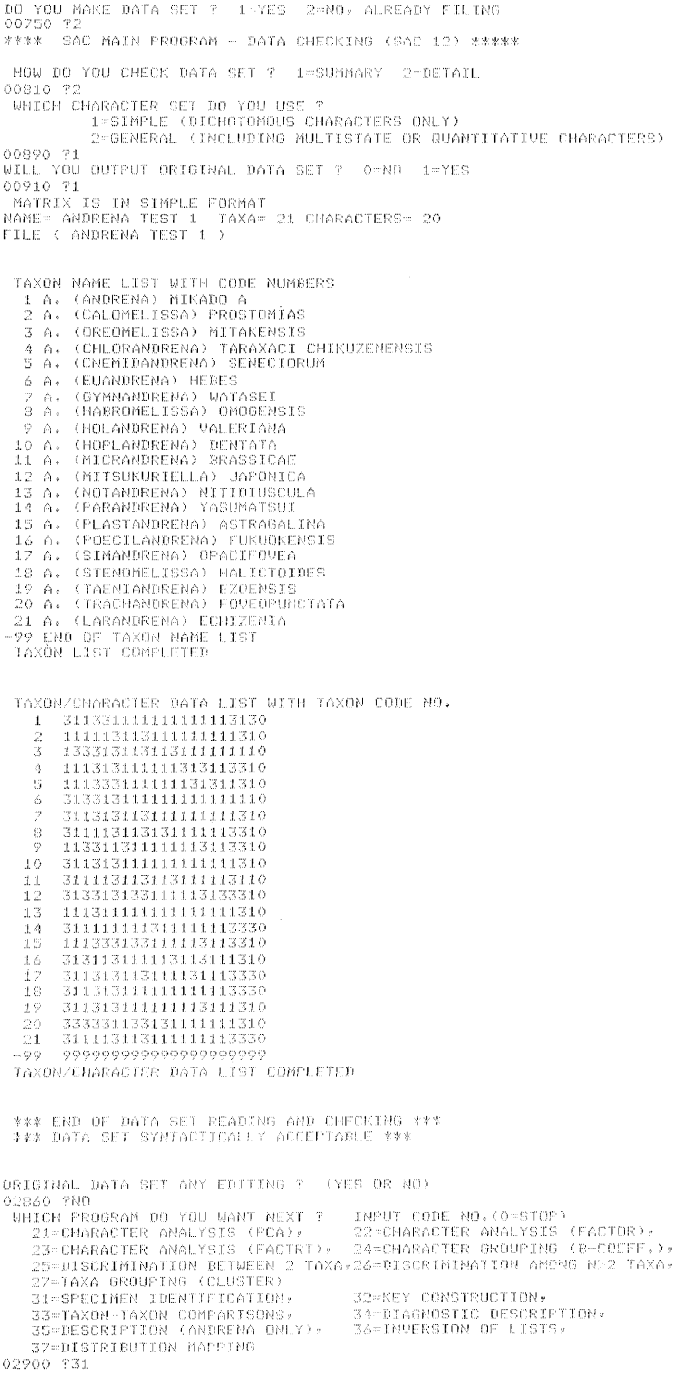


```
****** SAC COMPUTER SYSTEM FOR ANDRENA CLASSIFICATION **********
  SAC SYSTEM IS COMPOSED OF THE FOLLOWING 3 FUNCTIONS INCLUDING
  17 MAIN PROGRAMS
    10: MAKING DATA SET
        11: DATA ACCUMULATION
        12: DATA CHECKING
        13: DATA EDITING
    20: DATA ANALYSIS
        21: CHARACTER ANALYSIS (PCA)
        22: CHARACTER ANALYSIS (FACTOR)
        23: CHARACTER ANALYSIS (FACTOR WITH ROTATION)
        24: CHARACTER GROUPING (B COEFFICIENT)
        25: DISCRIMINATION BETWEEN TWO TAXA (DISCRIMINANT)
        26: DISCRIMINATION AMONG N TAXA (N>2) (DISCRIMINANT)
        27: TAXA GROUPING (CLUSTER ANALYSIS)
    30: DATA RETRIEVAL
        31: SPECIMEN IDENTIFICATION
        32: KEY CONSTRUCTION
        33: TAXON-TAXON COMPARISONS
        34: DIAGNOSTIC DESCRIPTION
        35: DESCRIPTION PRINTING (ANDRENA ONLY)
        36: INVERSION OF TAXON-BY-CHARACTER LISTS
        37: DISTRIBUTION MAPPING
  DATASET NAME : F1414.SAC.FORT
  ALLOC DATASET NAME : F1414.SACDATA1.DATA (FT04F001)
  ALLOC DATASET NAME : F1414.SACDATA2.DATA (FT08F001)
  ALLOC DATASET NAME : F1414.SACDATA3.DATA (FT03F001)
  ALLOC DATASET NAME : F1414.SACDATA4.DATA (FT15F001)
  ALLOC DATASET NAME : F1414.SACDATA5.DATA (FT10F001)
  ALLOC DATASET NAME : F1414.SACDATA6.DATA (FT11F001)
  ALLOC DATASET NAME : F1414.SACDATA7.DATA (FT14F001)
  DEVELOPED BY O. TADAUCHI (ENTOMOL. LAB., FAC. AGR., KYUSHU
     UNIVERSITY, FUKUOKA 812, JAPAN), (1979)
  BASED ON MSU SYSTEM DEVELOPED BY L. E. MORSE (1974)
*****************************************************************
```



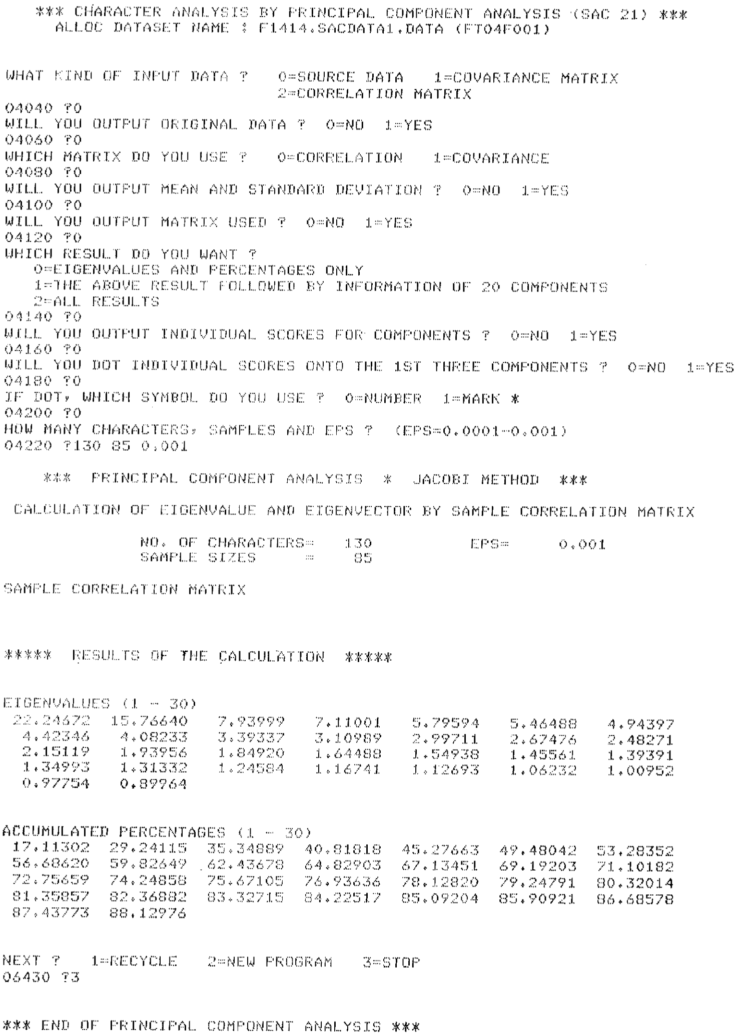
```
DO YOU MAKE DATA SET ?  1=YES  2=NO, ALREADY FILING
00750 ?1


 ********** ORIGINAL DATA ACCUMULATION (SAC 11) **********
   ALLOC DATASET NAME : F1414.SACDATA2.DATA (FT08F001)
   ALLOC DATASET NAME : F1414.SACDATA3.DATA (FT03F001)
   ALLOC DATASET NAME : F1414.SACDATA4.DATA (FT15F001)
   ALLOC DATASET NAME : F1414.SACDATA5.DATA (FT10F001)
   ALLOC DATASET NAME : F1414.SACDATA6.DATA (FT11F001)


WHAT KIND OF DATA SET DO YOU MAKE ?
  1=TAXON NAME LIST (IN CASE OF DICHOTOMOUS CHARACTERS)
  2=TAXON NAME LIST (IN CASE OF MULTISTATE AND QUANTITATIVE CHARACTERS)
  3=TAXON/CHARACTER DATA LIST
  4=CHARACTER NAME LIST
  5=STOP
04300 ?1
INPUT DATA SET NAME(WITHIN 16 LETTERS), NO. OF TAXA(WITHIN 3 NUMERALS),
  AND NO. OF CHARACTERS(WITHIN 3 NUMERALS)
04330 ?ANDRENA TEST 1    21   20
INPUT TAXON CODE(WITHIN 3 NUMERALS), AND TAXON NAME(WITHIN 60 LETTERS)
  IN THE END OF DATA, INPUT TAXON CODE=-99
04360 ?  1 A. (ANDRENA) MIKADO A
04360 ?  2 A. (CALOMELISSA) PROSTOMIAS
04360 ?  3 A. (OREOMELISSA) MITAKENSIS
04360 ?  4 A. (CHLORANDRENA) TARAXACI CHIKUZENENSIS
04360 ?  5 A. (CNEMIDANDRENA) SENECIORUM
04360 ?  6 A. (EUANDRENA) HEBES
04360 ?  7 A. (GYMNANDRENA) WATASEI
04360 ?  8 A. (HABROMELISSA) OMOGENSIS
04360 ?  9 A. (HOLANDRENA) VALERIANA
04360 ? 10 A. (HOPLANDRENA) DENTATA
04360 ? 11 A. (MICRANDRENA) BRASSICAE
04360 ? 12 A. (MITSUKURIELLA) JAPONICA
04360 ? 13 A. (NOTANDRENA) NITIDIUSCULA
04360 ? 14 A. (PARANDRENA) YASUMATSUI
04360 ? 15 A. (PLASTANDRENA) ASTRAGALIAN
04360 ? 16 A. (POECILANDRENA)
04360 ? 17 A. (SIMANDRENA) OPACIFOVEA
04360 ? 18 A. (STENOMELISSAO HALICTOIDES
04360 ? 19 A. (TAENIANDRENA) EZOENSIS
04360 ? 20 A. (TRACHANDRENA) FOVEOPUNCTATA
04360 ? 21 A. (LARANDRENA) ECHIZENIA
04360 ?-99
END OF TAXON NAME LIST
END OF DATA ACCUMULATION
```

Print 3



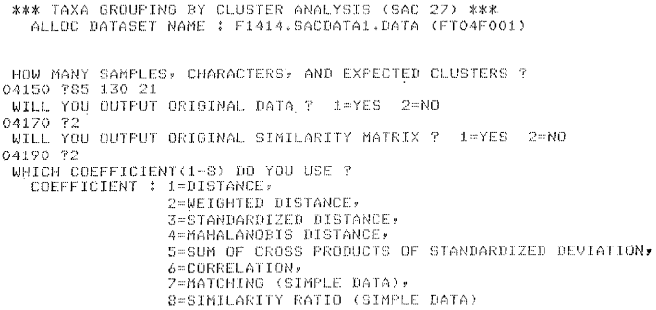
```
DO YOU MAKE DATA SET ?  1=YES  2=NO, ALREADY FILING
00750 ?2
****  SAC MAIN PROGRAM - DATA CHECKING (SAC 12) *****

 HOW DO YOU CHECK DATA SET ?  1=SUMMARY  2=DETAIL
00810 ?2
 WHICH CHARACTER SET DO YOU USE ?
          1=SIMPLE (DICHOTOMOUS CHARACTERS ONLY)
          2=GENERAL (INCLUDING MULTISTATE OR QUANTITATIVE CHARACTERS)
00890 ?1
WILL YOU OUTPUT ORIGINAL DATA SET ?  0=NO  1=YES
00910 ?1
 MATRIX IS IN SIMPLE FORMAT
NAME= ANDRENA TEST 1  TAXA= 21 CHARACTERS= 20
FILE ( ANDRENA TEST 1 )


 TAXON NAME LIST WITH CODE NUMBERS
  1 A. (ANDRENA) MIKADO A
  2 A. (CALOMELISSA) PROSTOMIAS
  3 A. (OREOMELISSA) MITAKENSIS
  4 A. (CHLORANDRENA) TARAXACI CHIKUZENENSIS
  5 A. (CNEMIDANDRENA) SENECIORUM
  6 A. (EUANDRENA) HEBES
  7 A. (GYMNANDRENA) WATASEI
  8 A. (HABROMELISSA) OMOGENSIS
  9 A. (HOLANDRENA) VALERIANA
 10 A. (HOPLANDRENA) DENTATA
 11 A. (MICRANDRENA) BRASSICAE
 12 A. (MITSUKURIELLA) JAPONICA
 13 A. (NOTANDRENA) NITIDIUSCULA
 14 A. (PARANDRENA) YASUMATSUI
 15 A. (PLASTANDRENA) ASTRAGALINA
 16 A. (POECILANDRENA) FUKUOKENSIS
 17 A. (SIMANDRENA) OPACIFOVEA
 18 A. (STENOMELISSA) HALICTOIDES
 19 A. (TAENIANDRENA) EZOENSIS
 20 A. (TRACHANDRENA) FOVEOPUNCTATA
 21 A. (LARANDRENA) ECHIZENIA
-99 END OF TAXON NAME LIST
 TAXON LIST COMPLETED


 TAXON/CHARACTER DATA LIST WITH TAXON CODE NO.
   1   311331111111111113130
   2   111113113111111111310
   3   133313113113111111110
   4   111313111111313113310
   5   111333111111131311310
   6   313313111111111111110
   7   311313113111111111310
   8   311113113131111113310
   9   113311311111113113310
  10   311313111111111111310
  11   311113113113111113110
  12   313313133111113133310
  13   111311111111111111310
  14   311111111311111113330
  15   111333133111113113310
  16   313113111113113111310
  17   311313113111131113330
  18   311313111111111113330
  19   311313111111113111310
  20   333331133131111111310
  21   311113113111111113330
 -99   999999999999999999999
 TAXON/CHARACTER DATA LIST COMPLETED


 *** END OF DATA SET READING AND CHECKING ***
 *** DATA SET SYNTACTICALLY ACCEPTABLE ***


ORIGINAL DATA SET ANY EDITING ?  (YES OR NO)
02860 ?NO
 WHICH PROGRAM DO YOU WANT NEXT ?   INPUT CODE NO.(0=STOP)
   21=CHARACTER ANALYSIS (PCA),      22=CHARACTER ANALYSIS (FACTOR),
   23=CHARACTER ANALYSIS (FACTRT),   24=CHARACTER GROUPING (B-COEFF.),
   25=DISCRIMINATION BETWEEN 2 TAXA,26=DISCRIMINATION AMONG N>2 TAXA,
   27=TAXA GROUPING (CLUSTER)
   31=SPECIMEN IDENTIFICATION,       32=KEY CONSTRUCTION,
   33=TAXON-TAXON COMPARISONS,       34=DIAGNOSTIC DESCRIPTION,
   35=DESCRIPTION (ANDRENA ONLY),    36=INVERSION OF LISTS,
   37=DISTRIBUTION MAPPING
02900 ?31
```

Print 4



```
   *** CHARACTER ANALYSIS BY PRINCIPAL COMPONENT ANALYSIS (SAC 21) ***
     ALLOC DATASET NAME : F1414.SACDATA1.DATA (FT04F001)


WHAT KIND OF INPUT DATA ?   0=SOURCE DATA   1=COVARIANCE MATRIX
                            2=CORRELATION MATRIX
04040 ?0
WILL YOU OUTPUT ORIGINAL DATA ?  0=NO  1=YES
04060 ?0
WHICH MATRIX DO YOU USE ?   0=CORRELATION   1=COVARIANCE
04080 ?0
WILL YOU OUTPUT MEAN AND STANDARD DEVIATION ?  0=NO  1=YES
04100 ?0
WILL YOU OUTPUT MATRIX USED ?  0=NO  1=YES
04120 ?0
WHICH RESULT DO YOU WANT ?
   0=EIGENVALUES AND PERCENTAGES ONLY
   1=THE ABOVE RESULT FOLLOWED BY INFORMATION OF 20 COMPONENTS
   2=ALL RESULTS
04140 ?0
WILL YOU OUTPUT INDIVIDUAL SCORES FOR COMPONENTS ?  0=NO  1=YES
04160 ?0
WILL YOU DOT INDIVIDUAL SCORES ONTO THE 1ST THREE COMPONENTS ?  0=NO  1=YES
04180 ?0
IF DOT, WHICH SYMBOL DO YOU USE ?  0=NUMBER  1=MARK *
04200 ?0
HOW MANY CHARACTERS, SAMPLES AND EPS ?  (EPS=0.0001-0.001)
04220 ?130 85 0.001

    ***  PRINCIPAL COMPONENT ANALYSIS  *  JACOBI METHOD  ***

 CALCULATION OF EIGENVALUE AND EIGENVECTOR BY SAMPLE CORRELATION MATRIX

                NO. OF CHARACTERS=   130           EPS=     0.001
                SAMPLE SIZES     =    85

SAMPLE CORRELATION MATRIX



*****  RESULTS OF THE CALCULATION  *****


EIGENVALUES (1 - 30)
 22.24672  15.76640   7.93999   7.11001   5.79594   5.46488   4.94397
  4.42346   4.08233   3.39337   3.10989   2.99711   2.67476   2.48271
  2.15119   1.93956   1.84920   1.64488   1.54938   1.45561   1.39391
  1.34993   1.31332   1.24584   1.16741   1.12693   1.06232   1.00952
  0.97754   0.89764


ACCUMULATED PERCENTAGES (1 - 30)
 17.11302  29.24115  35.34889  40.81818  45.27663  49.48042  53.28352
 56.68620  59.82649  62.43678  64.82903  67.13451  69.19203  71.10182
 72.75659  74.24858  75.67105  76.93636  78.12820  79.24791  80.32014
 81.35857  82.36882  83.32715  84.22517  85.09204  85.90921  86.68578
 87.43773  88.12976


NEXT ?   1=RECYCLE   2=NEW PROGRAM   3=STOP
06430 ?3


*** END OF PRINCIPAL COMPONENT ANALYSIS ***
```

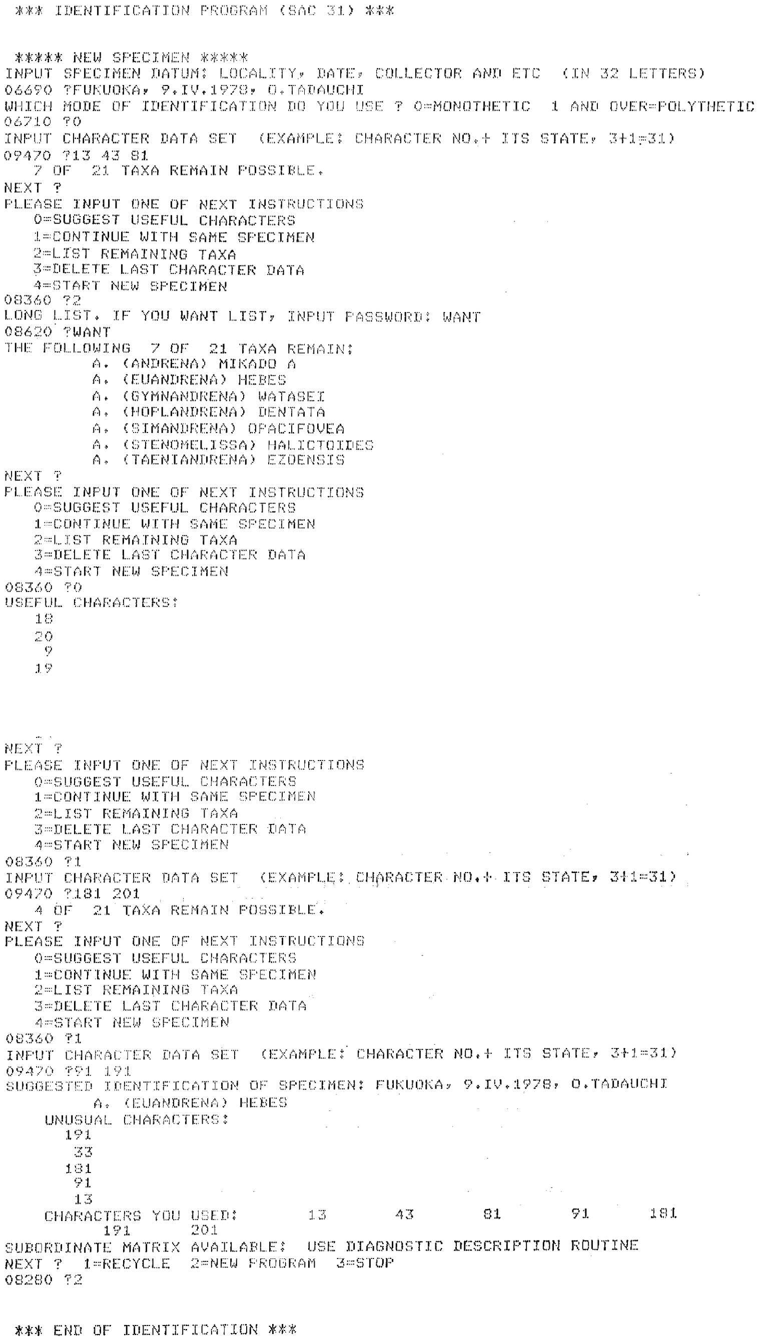
Print 5



```
    *** TAXA GROUPING BY CLUSTER ANALYSIS (SAC 27) ***
      ALLOC DATASET NAME : F1414.SACDATA1.DATA (FT04F001)


    HOW MANY SAMPLES, CHARACTERS, AND EXPECTED CLUSTERS ?
   04150 ?85 130 21
    WILL YOU OUTPUT ORIGINAL DATA ?  1=YES  2=NO
   04170 ?2
    WILL YOU OUTPUT ORIGINAL SIMILARITY MATRIX ?  1=YES  2=NO
   04190 ?2
    WHICH COEFFICIENT(1-8) DO YOU USE ?
      COEFFICIENT : 1=DISTANCE,
                    2=WEIGHTED DISTANCE,
                    3=STANDARDIZED DISTANCE,
                    4=MAHALANOBIS DISTANCE,
                    5=SUM OF CROSS PRODUCTS OF STANDARDIZED DEVIATION,
                    6=CORRELATION,
                    7=MATCHING (SIMPLE DATA),
                    8=SIMILARITY RATIO (SIMPLE DATA)
```

```
04320 ?3
WHICH METHOD  -8) DO YOU USE ?
   METHOD       : 1=NEAREST NEIBOR,
                  2=FURTHEST NEIGHBOR,
                  3=MEDIAN,
                  4=FLEXIBLE,
                  5=CENTROID,
                  6=GROUP AVERAGE,
                  7=WARD,
                  8=FLEXIBLE GROUP AVERAGE


04350 ?6
 WHICH OUTPUT CONTROL M1(1-2), M2(1-2), AND M3(1-2) DO YOU USE ?
   M1 : 1=OUTPUT ONLY INTERMEDIATE CYCLE NO., 2=NO
   M2 : 1=OUTPUT EVERY INTERMEDIATE PROCESS, 2=NO
   M3 : 1=OUTPUT ONLY RESULT IN EXPECTED CLUSTER, 2=ALL RESULTS


04380 ?2


CLUSTER ANALYSIS
         BY THE GROUP-AVERAGE METHOD

                BASED ON STANDARDIZED EUCLIDIAN DISTANCE



         SAMPLE SIZE                    =  85
         MAX. NO. OF CLUSTERS EXPECTED  =  21
         NO. OF CHARACTERS              = 130
         SPECIFIED NO. OF CLUSTERS      =  21

 CLUSTER FUSION POINT IN THIS CYCLE

                        1.256


    CLUSTER NO.  1
         1    2    3    4    5    6    7    8    9   10

        11   12   13   65   66   68

    CLUSTER NO.  2
        14   15   16   17   62

    CLUSTER NO.  3
        18   19   20

    CLUSTER NO.  4
        21   22   23

    CLUSTER NO.  5
        24   25   26   27   70   71   72   73

    CLUSTER NO.  6
        28   29   30   31   32   33   74

    CLUSTER NO.  7
        34

    CLUSTER NO.  8
        35

    CLUSTER NO.  9
        36   37

    CLUSTER NO. 10
        38   39   40   41   75   76

    CLUSTER NO. 11
        42   43   44   45   46   47   77   78   79   80

        81   82

    CLUSTER NO. 12
        48   49

    CLUSTER NO. 13
        50   51   55
```

Print 6



```
 *** IDENTIFICATION PROGRAM (SAC 31) ***


 ***** NEW SPECIMEN *****
INPUT SPECIMEN DATUM: LOCALITY, DATE, COLLECTOR AND ETC  (IN 32 LETTERS)
06690 ?FUKUOKA, 9.IV.1978, O.TADAUCHI
WHICH MODE OF IDENTIFICATION DO YOU USE ? 0=MONOTHETIC  1 AND OVER=POLYTHETIC
06710 ?0
INPUT CHARACTER DATA SET  (EXAMPLE: CHARACTER NO.+ ITS STATE, 3+1=31)
09470 ?13 43 81
   7 OF  21 TAXA REMAIN POSSIBLE.
NEXT ?
PLEASE INPUT ONE OF NEXT INSTRUCTIONS
   0=SUGGEST USEFUL CHARACTERS
   1=CONTINUE WITH SAME SPECIMEN
   2=LIST REMAINING TAXA
   3=DELETE LAST CHARACTER DATA
   4=START NEW SPECIMEN
08360 ?2
LONG LIST. IF YOU WANT LIST, INPUT PASSWORD: WANT
08620 ?WANT
THE FOLLOWING  7 OF  21 TAXA REMAIN:
          A. (ANDRENA) MIKADO A
          A. (EUANDRENA) HEBES
          A. (GYMNANDRENA) WATASEI
          A. (HOPLANDRENA) DENTATA
          A. (SIMANDRENA) OPACIFOVEA
          A. (STENOMELISSA) HALICTOIDES
          A. (TAENIANDRENA) EZOENSIS
NEXT ?
PLEASE INPUT ONE OF NEXT INSTRUCTIONS
   0=SUGGEST USEFUL CHARACTERS
   1=CONTINUE WITH SAME SPECIMEN
   2=LIST REMAINING TAXA
   3=DELETE LAST CHARACTER DATA
   4=START NEW SPECIMEN
08360 ?0
USEFUL CHARACTERS:
   18
   20
    9
   19




NEXT ?
PLEASE INPUT ONE OF NEXT INSTRUCTIONS
   0=SUGGEST USEFUL CHARACTERS
   1=CONTINUE WITH SAME SPECIMEN
   2=LIST REMAINING TAXA
   3=DELETE LAST CHARACTER DATA
   4=START NEW SPECIMEN
08360 ?1
INPUT CHARACTER DATA SET  (EXAMPLE: CHARACTER NO.+ ITS STATE, 3+1=31)
09470 ?181 201
   4 OF  21 TAXA REMAIN POSSIBLE.
NEXT ?
PLEASE INPUT ONE OF NEXT INSTRUCTIONS
   0=SUGGEST USEFUL CHARACTERS
   1=CONTINUE WITH SAME SPECIMEN
   2=LIST REMAINING TAXA
   3=DELETE LAST CHARACTER DATA
   4=START NEW SPECIMEN
08360 ?1
INPUT CHARACTER DATA SET  (EXAMPLE: CHARACTER NO.+ ITS STATE, 3+1=31)
09470 ?91 191
SUGGESTED IDENTIFICATION OF SPECIMEN: FUKUOKA, 9.IV.1978, O.TADAUCHI
          A. (EUANDRENA) HEBES
    UNUSUAL CHARACTERS:
       191
        33
       181
        91
        13
    CHARACTERS YOU USED:        13        43        81        91       181
          191       201
SUBORDINATE MATRIX AVAILABLE:  USE DIAGNOSTIC DESCRIPTION ROUTINE
NEXT ?  1=RECYCLE  2=NEW PROGRAM  3=STOP
08280 ?2


 *** END OF IDENTIFICATION ***
```

Print 7

```
 *** KEY CONSTRUCTION PROGRAM (SAC 32) ***

 WHAT IS A TITLE FOR KEY ? (IN WORDS)
19240 ?KEY TO ANDRENA TEST 1 IN PHRASES
  WHICH OUTPUT MODE DO YOU WANT ?  1=CODE NUMBERS  2=PHRASES
19340 ?2


          KEY TO ANDRENA TEST 1 IN PHRASES


 1. PROPODEUM WITH DISTINCT ANGLE POSTERIORLY

 2. MALAR SPACE LINEAR

 3. CLYPEUS WITHOUT MEDIAN IMPUNCTATE LINE
                              .....  2. A. (CALOMELISSA) PROSTOMIAS

 3. CLYPEUS WITH MEDIAN IMPUNCTATE LINE

 4. 3RD SUBM. CELL RECEIVING 2ND REC.V. AT END OF CELL
                              .....  8. A. (HABROMELISSA) OMOGENSIS

 4. 3RD SUBM. CELL RECEIVING 2ND REC.V. FAR FROM END OF CELL

 5. 1ST INTERCUB. V. ENDING APART FROM STIGMA
                              ..... 21. A. (LARANDRENA) ECHIZENIA

 5. 1ST INTERCUB. V. ENDING NEAR OR CLOSE TO STIGMA
                              ..... 11. A. (MICRANDRENA) BRASSICAE

 2. MALAR SPACE WELL DEVELOPED TO ELONGATE

 6. FACIAL FOVEAE SHALLOW

 7. CLYPEUS WITHOUT MEDIAN IMPUNCTATE LINE
                              ..... 15. A. (PLASTANDRENA) ASTRAGALINA

 7. CLYPEUS WITH MEDIAN IMPUNCTATE LINE

 8. HIND TIBIAE NOT STRONGLY WIDENED APICALLY
                              .....  7. A. (GYMNANDRENA) WATASEI

 8. HIND TIBIAE STRONGLY WIDENED NEAR APICES
                              ..... 17. A. (SIMANDRENA) OPACIFOVEA

 6. FACIAL FOVEAE DEEP

 9. CLYPEUS WITHOUT MEDIAN IMPUNCTATE LINE
                              .....  3. A. (OREOMELISSA) MITAKENSIS

 9. CLYPEUS WITH MEDIAN IMPUNCTATE LINE

10. FACIAL FOVEA NARROWLY SEPARATED FROM EYE MAG.
                              ..... 12. A. (MITSUKURIELLA) JAPONICA
```

Print 8

```
 *** TAXON-TAXON COMPARISON PROGRAMS (SAC 33) ***


 WHICH COMPARISON PROGRAM DO YOU WANT ?  (1-3)


 1=DIAGNOSTIC CHECK OF DATA MATRIX
 2=COMPARISONS OF ONE TAXON WITH OTHERS
 3=COMPARISON OF A PAIR OF TAXA
12880 ?3
 WHICH TWO TAXA
13120 ?1 7
 NONOVERLAPPING DIFFERENCES OF THESE TAXA:
DICHOTOMOUS:
  5:      3     1
  6:      1     3
  9:      1     3
 18:      3     1
 19:      1     3
 20:      3     1

NEXT ?  1=RECYCLE  2=NEW PROGRAM  3=STOP
13850 ?3


** END OF TAXON-TAXON COMPARISONS ***
```

Print 9

```
 *** DIAGNOSTIC DESCRIPTIONS PROGRAM (SAC 34) ***


 WHICH CHARACTER-SEQUENCING MODE DO YOU USE ?
 1=IN DATA MATRIX ORDER
 2=BY DECREASING IMPORTANCE VALUES
 3=SELECTED CHARACTERS ONLY
11800 ?2
 WHICH OUTPUT MODE DO YOU WANT ?  1=CODE NUMBERS  2=PHRASES
11830 ?2
DO YOU WANT ALL TAXA ?  (YES OR NO)
11850 ?NO
 HOW MANY CHARACTERS ?
11890 ?5
IF YOU WANT TO STOP, INPUT -99 FOR TAXON NUMBER


INPUT TAXON NUMBER
12150 ?14
*****
A. (PARANDRENA) YASUMATSUI
   -- RIGHT ?   (YES OR NO)
12270 ?YES

10. WINGS WITH TWO SUBMARGINAL CELLS

 6. PRONOTAL SUTURE NOT INDICATED

20. PROPODEAL CORBICULA WITH HAIRS ANTERIORLY

 4. MALAR SPACE LINEAR

 9. PROPODEUM WITH GRADUAL ANGLE POSTERIORLY


INPUT TAXON NUMBER
12150 ?-99

 [illegible]      1=RECYCLE  2=NEW PROGRAM 3=STOP
1[illegible]     ?3


 [illegible] DIAGNOSTIC DESCRIPTION ***
```

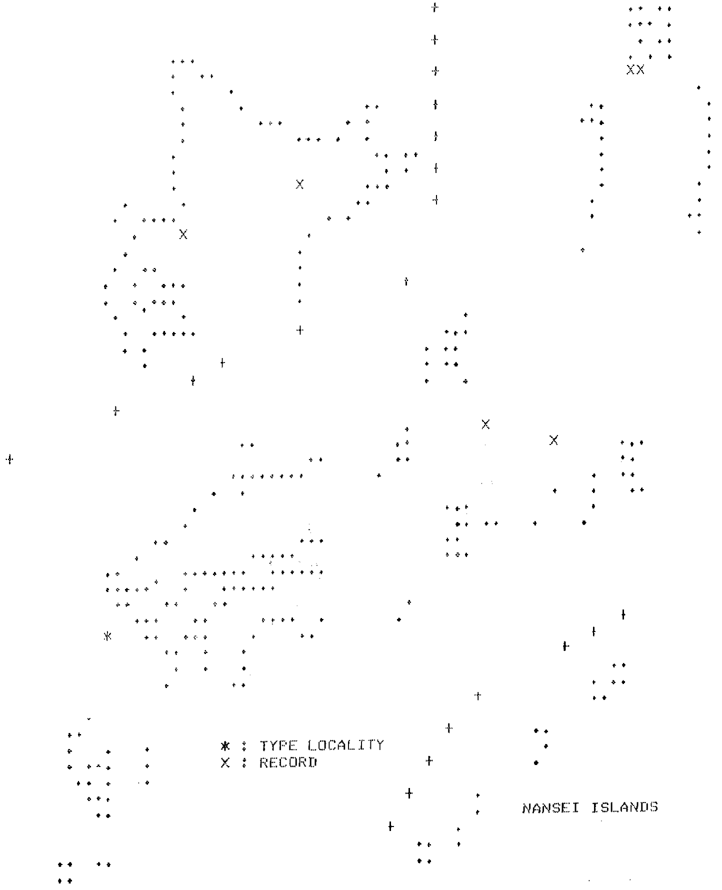
Print 10

```
 *** DESCRIPTION PRINTING (ANDRENA ONLY) (SAC 35) ***
   ALLOC DATASET NAME : F1414.SACDATA1.DATA (FT04F001)


 WHICH TAXON ? (CODE NO. PLEASE.   0 STOP)
10370 ?1
     ANDRENA (ANDRENA) BREBIHIRTISCOPA HIRASHIMA 1962
     FEMALE

     PUBESCENCE:
HAIRS ON CLYPEUS SPARSE BASALLY, 2;
HAIRS ON CLYPEUS WHITISH, 1;
FACIAL FOVEAE BROWNISH, 3;
VERTEX WITH FUSCOUS HAIRS, 2;
HAIRS ON GENAL AREA VERY DENSE, 3;
HAIRS ON GENAL AREA LONG, 2;
HAIRS ON ANTENNAL REGION NOT DENSE, 2;
HAIRS ON MESOSCUTUM NOT DENSE, 2;
HAIRS ON MESOSCUTUM LONG, 2;
HAIRS ON MESOSCUTUM CHIEFLY WHITISH, 1;
HAIRS ON MESOSCUTUM WITHOUT MIXED HAIRS, 1;
HAIRS ON METANOTUM NOT DENSE, 2;
HAIRS ON MESEPISTERNUM WHITISH, 1;
TROCHANTERAL FLOCCUS PERFECTLY ARRANGED, 2;
TROCHANTERAL FLOCCUS DENSE, 2;
TROCHANTERAL FLOCCUS LONG, 3;
FEMORAL FLOCCUS NOT BRANCHED, 1;
FEMORAL FLOCCUS WHITISH, 1;
TIBIAL SCOPA NOT BRANCHED, 1;
TIBIAL SCOPA COMPACT, 2;
TIBIAL SCOPA SHORT, 1;
TIBIAL SCOPA CHIEFLY WHITISH AND PARTLY BROWNISH, 2;
HAIRS ON PROPODEUM NOT DENSE, 2;
DORSAL FRINGE OF PROPODEAL CORBICULA VERY DENSE, 3;
DORSAL FRINGE OF PROPODEAL CORBICULA VERY LONG, 3;
DORSAL FRINGE OF PROPODEAL CORBICULA VERY WELL ARRANGED, 3;
DORSAL FRINGE OF PROPODEAL CORBICULA WELL BRANCHED, 2;
CORBICULA AREA WITH BRANCHED HAIRS ANTERIORLY, 2;
INTERIOR OF CORBICULA WITHOUT SIMPLE HAIRS, 3;
INTERNAL HAIRS OF CORBICULA NOT FOUNDED, 4;
```

```
HAIRS ON FIRST METASOMAL TERGUM SCANTY, 2;
HAIRS ON SECOND METASOMAL TERGUM SPARSE MEDIALLY, 2;
HAIRS ON SECOND METASOMAL TERGUM WHITISH POSTERIORLY, 1;
HAIRS ON THIRD METASOMAL TERGUM LONG POSTERIORLY, 3;
THIRD METASOMAL TERGUM WITH SPARSE, OBSCURE FRINGE POSTERIORLY, 2;
THIRD METASOMAL TERGUM WITH POSTERIOR FRINGE LOOSE, 2;
HAIRS ON THIRD METASOMAL TERGUM DOWNY APICALLY, 1;
CAUDAL FIMBRIA BROWNISH, 2;
HAIRS ON SECOND METASOMAL STERNUM MEDIUM-LENGTH, 2;
APICAL HAIRS ON THIRD METASOMAL STERNUM DENSE, 2;

     COLOUR:
THIRD TO TENTH FLAGELLAR SEGMENTS BENEATH BROWNISH, 2;
TEGULAE DEEP REDDISH, 3;
WINGS BROWNISH SUBHYALINE, 1;
VEINS AND STIGMA PALE BROWNISH, 1;
HIND BASITARSI BROWNISH, 2;
TIBIAL SPURS PALE TO YELLOWISH BROWN, 1;
METASOMAL TERGA BLACK, 1;
METASOMAL TERGA BROWNISH POSTERIORLY, 3;
```

Print 11

```
*** DISTRIBUTION MAPPING (SAC 37) ***

WHICH TAXON ? (CODE NO. PLEASE)
27530 ?1
    ANDRENA (ANDRENA) BREBIHIRTISCOPA HIRASHIMA 1962
```



```
*** END OF MAPPING ***
```