Rev. Brasil. Genet. 14, 2, 425-435 (1991)
(Brazil. J. Genetics)

PHYLOGENETIC STUDIES OF SOME SPECIES OF THE GENUS *Coffea* - I - NUMERICAL ANALYSIS OF FLAVONOID COMPOUNDS

Catalina Romero Lopes[1] and G.P. Shepherd[2]

ABSTRACT

Flavonoid compounds were analyzed in ripe fruit pulp of ten species of *Coffea*, including two cultivars of *C. arabica* and two of *C. canephora*. Three coefficients of similarity: Simple-Matching, Jaccard and Ochiai and three different clustering methods, Single Linkage, Complete Linkage and Unweighted Pair Group, Using Arithmetic Averages (UPGMA), were used to analyze the data.

Jaccard and Ochiai's coefficients of association showed a more coherent result, when compared with taxonomic and hybridization studies. Inclusion of *Psilanthopsis kapakata* in the genus *Coffea*, as *C. kapakata*, is justified by the similarity of this species with other studied species, and clusters clearly approximate the species *C. arabica* and *C. eugenioides*. The latter is one of the possible parents of the allotetraploid species *C. arabica, C. congensis* is the only species whose position remains ambiguous, probably due to the fact that the plants of this species that were introduced into the Campinas collections, were hybrids and not typical of *C. congensis*.

INTRODUCTION

The utilization of chemical characteristics has helped to establish additional criteria for the analysis of groups badly defined taxonomically, has improved knowledge of the phylogeny of genetically important groups, has facilitated the study of population structure, has been used in tests of contamination of hybrid seeds, in the characterization of varieties and cultivars, and in the evaluation of genetic variability of different taxa.

[1] Departamento de Genética, Instituto de Biociências, Universidade Estadual Paulista, Caixa Postal 502, 18610 Botucatu, SP, Brasil. Send correspondence to C.R.L.

[2] Departamento de Botânica, Instituto de Biologia, Universidade de Campinas, 13100 Campinas, SP, Brasil.

The study of chemical compounds as genetic markers has been of great utility in the genus *Coffea*, due to the innumerable difficulties existing in its classification. The most important taxonomic review of this group of plants was done by Chevalier (1947), utilizing characteristics such as geographical distribution, morphology, anatomy and colour of flowers and fruits. More recently, with the exploration of new areas of Africa and Asia, many new species have been described and important taxonomic modifications were introduced by Charrier (1975), Leroy (1980), Carvalho and Monaco (1967) and Carvalho *et al.* (1985). The latter suggested that *C. eugenioides* should be included in the subsection Erythrocoffea, while *C. stenophylla* should be included in the section Pachycoffea. They also suggested the inclusion of the species *Psilanthopsis kapakata* in the genus *Coffea*, due to its high affinity with the species of this genus, although Chevalier (1947) had considered it to be a monotypical species of the genus *Psilanthopsis*. Charrier (1975) observed a high intercrossing rate among the species of the subsection Erythrocoffea, section Eucoffea, and the species of the section Mascarocoffea. Leroy (1980), studied rare species of coffee from Angola and Tanzania, and proposed a complete reclassification of the coffee species that Chevalier (1947) placed in the genus *Coffea* and *Psilanthus*. He distributed them among the genera *Coffea, Psilanthus* and *Nostalachma*, stressing that several species were in a doubtful position. The relationships, among some of the most important species of the genus *Coffea* and *Psilanthopsis*, were classified according to flavonoid compounds affinities by Lopes and Monaco (1979) and Lopes *et al.* (1984). Their results strongly indicated that *Coffea eugenioides* participated in the origin of *C. arabica*, and demonstrated that the other parental species involved might have been *C. liberica* or *C. canephora*, or that *C. liberica* might have participated in the formation of both *C. canephora* and *C. arabica*.

MATERIALS AND METHODS

In the present research, the data obtained from flavonoid analysis were submitted to different methods of numerical taxonomy, by the utilization of three coefficients of similarity, and by three different clustering methods. The differences and similarities obtained in the relationships among the species by the different analyses, are presented and discussed. The methods utilized were applied to permit further comparison between the results obtained with these markers and those obtained with isoenzymatic markers (Lopes, submitted).

The species studied belong to the genus *Coffea*, section Eucoffea, and according to the classification of Chevalier (1947), with modifications introduced by Carvalho and Monaco (1967) and Carvalho *et al.* (1985), are distributed in the following subsections: Erythrocoffea (*C. congensis* Frochner, *C. canephora* Pierre ex

Frochner, *C. arabica* L. and *C. eugenioides* Moore); Mozambicoffea (*C. racemosa* Lour, *C. salvatrix* Swyn et Phill and *C. kapakata* Hirsch); Pachycoffea (*C. liberica* Hiern and *C. dewevrei* De Wild et Durand); Melanocoffea (*C. stenophylla* G. Don.). The species *C. arabica* was represented by the cultivars Arabica, the type of the species according to Linnaeus, and Abissinica, the type of the species according to Chevalier (1947). The species *C. canephora* was represented by the cultivars Kouillou, the type of the species according to Chevalier (1947), and Robusta, because of its greater economic importance. This material belongs to the Coffee germplasm bank of the Genetics Department of the Instituto Agronômico de Campinas, state of São Paulo, Brazil, and the possible origins are related by Lopes and Monaco (1977, 1979). The methodology used for flavonoid extraction and analysis was as presented by Lopes and Monaco (1979).

The species were referred to as operational taxonomic units (OTUs), according to a list of characters (presence or absence of each flavonoid compound). To determine the similarity among the OTUs, three coefficients of association were used, "Simple-Matching" (SM), the Jaccard (J) and the Ochiai (O) (Sneath and Sokal, 1973). The coefficient "Simple-Matching" is defined as a rate resulting from the addition of positive and negative concordance, and the total addition of concordances and discordances; the Jaccard is defined as a rate of the number of positive concordances calculated by the addition of the positive concordances and discordances; and the Ochiai is a coefficient that varies from zero to one and can be interpreted as a distance between two OTUs. It is computed as the ratio between the positive similarities and the square root of the result of the addition of the positive similarities with dissimilarities. The strategy used to accomplish the analysis was sequential, agglomerative, hierarchical and not superimposed. Methods for clustering, Unweighted pair group using arithmetic averages (UPGMA), Complete Linkage (CL), and Single Linkage (SL) were utilized (Sneath and Sokal, 1973).

RESULTS

The results of the "Simple-Matching" coefficient of similarity and the Single Linkage, Complete Linkage and the UPGMA clustering methods analyses are shown, respectively in Figures 1, 2 and 3.

The results obtained for the Jaccard's coefficient of association, are shown in Figures 4, 5 and 6, grouped respectively, by Single Linkage, Complete Linkage and UPGMA methods. Figures 7, 8 and 9 refer to the results obtained from the use of the Ochiai's coefficient of association, grouped by Single Linkage, Complete Linkage and UPGMA methods, respectively.

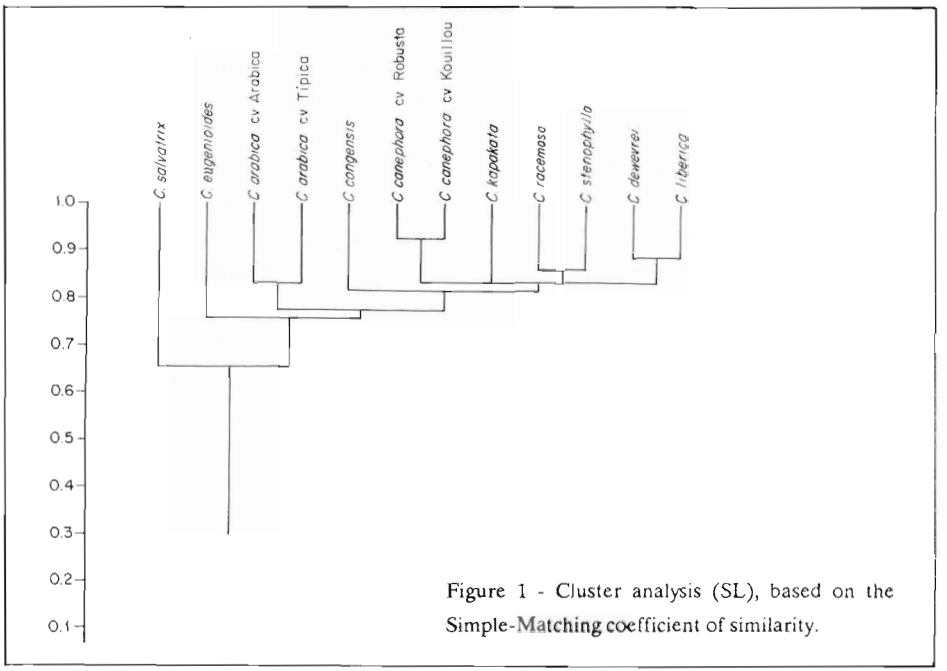

Figure 1 - Cluster analysis (SL), based on the Simple-Matching coefficient of similarity.

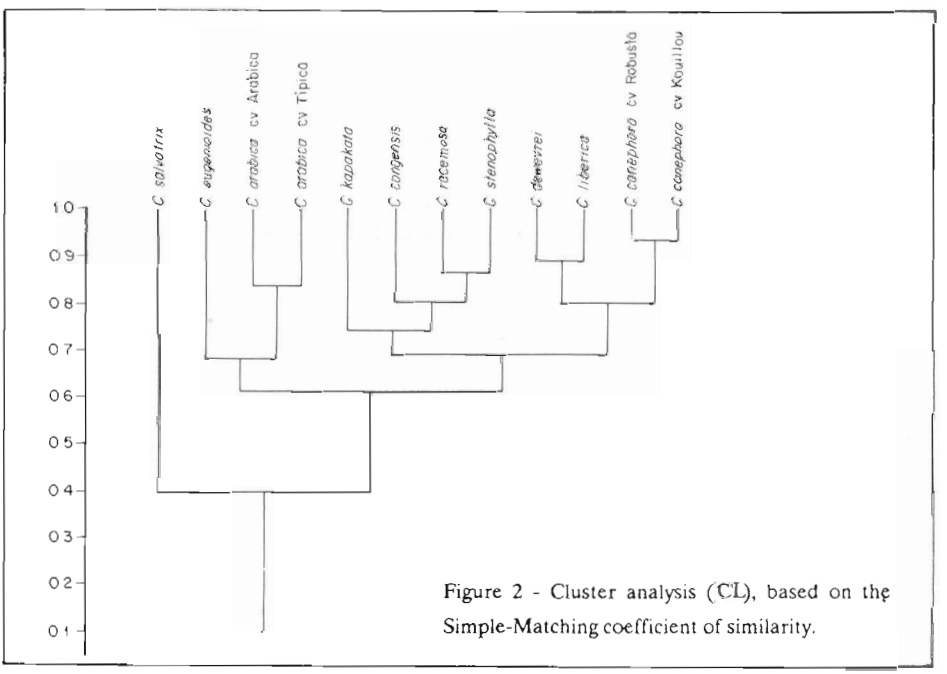

Figure 2 - Cluster analysis (CL), based on the Simple-Matching coefficient of similarity.

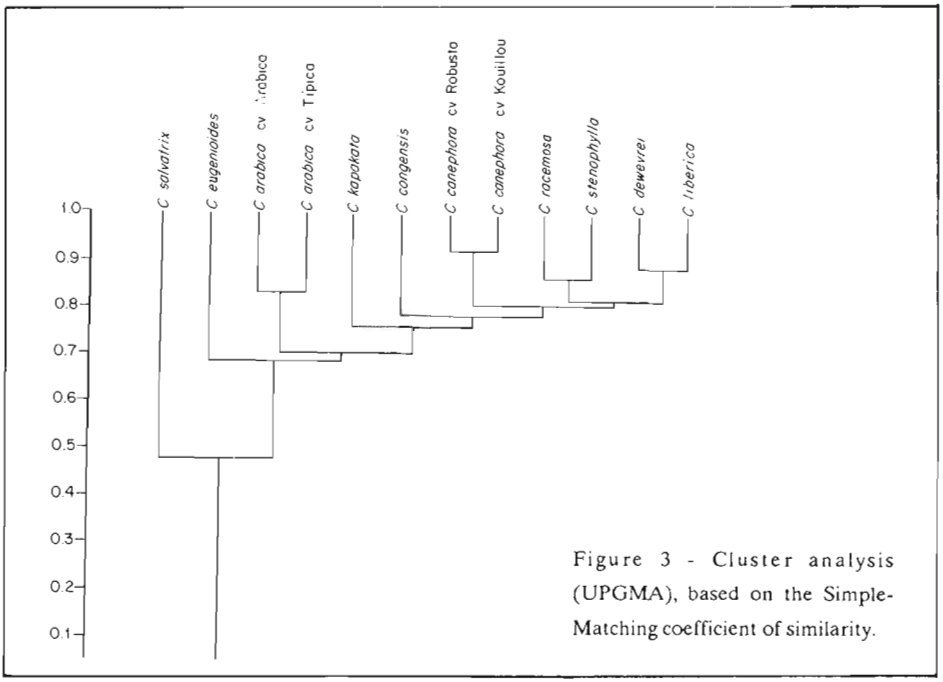

Figure 3 - Cluster analysis (UPGMA), based on the Simple-Matching coefficient of similarity.

DISCUSSION

In all clusters obtained from the different coefficients of association utilized, the two cultivars of *C. arabica* and the two cultivars of *C. canephora* demonstrated a greater similarity than with any of the other species, as it was expected. The great similarity between *C. liberica* and *C. dewevrei* was also confirmed. They both belong to subsection Pachycoffea. Chevalier in his taxonomic review of 1942, considered them as a single species, but as distinct species in the taxonomic review of 1947.

If we observe Figures 1, 4 and 7, in which the coefficients of association used were, respectively, Simple-Matching, Jaccard and Ochiai, and the clustering method was always Single Linkage, we will be able to see that the relationships of affinity among the species are exactly the same, for the coefficients of Jaccard and Ochiai (Figures 4 and 7). In the case of the Simple-Matching (Figure 1), there are only two discrepancies, one in relation to the similarity between the species *C. racemosa* and *C. stenophylla*, higher in this case than in the two previous ones, and the other in relation to the position of *C. congensis*, slightly more distant from *C. arabica*.

Figures 2, 5 and 8, obtained respectively, through Simple-Matching, Jaccard and Ochiai's coefficients of association, but grouped by Complete Linkage, show the same affinity to all species, with discrepancies only in the position of *C. congensis*,

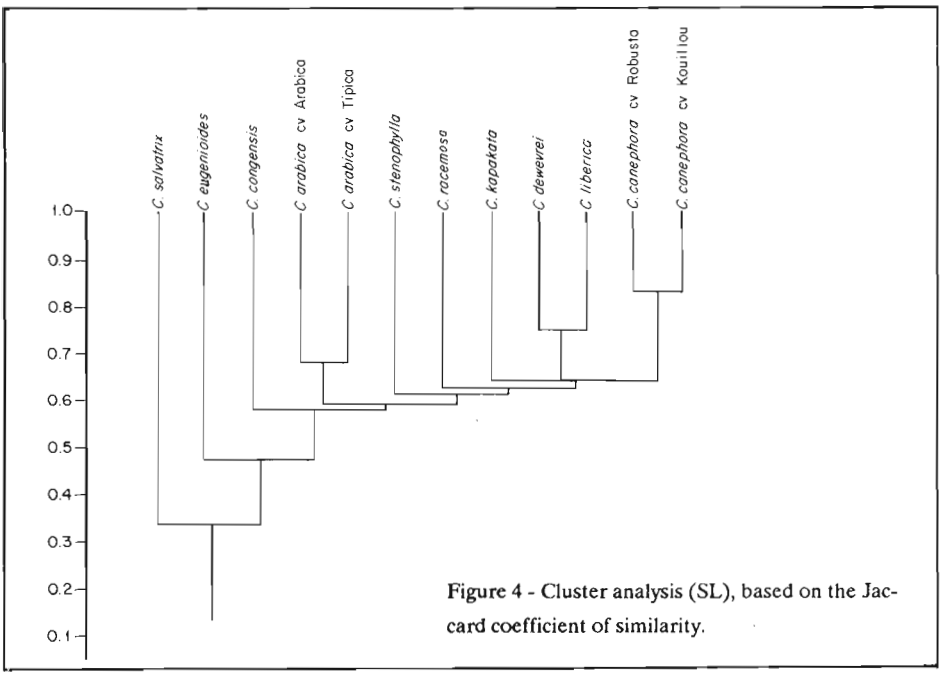

Figure 4 - Cluster analysis (SL), based on the Jaccard coefficient of similarity.

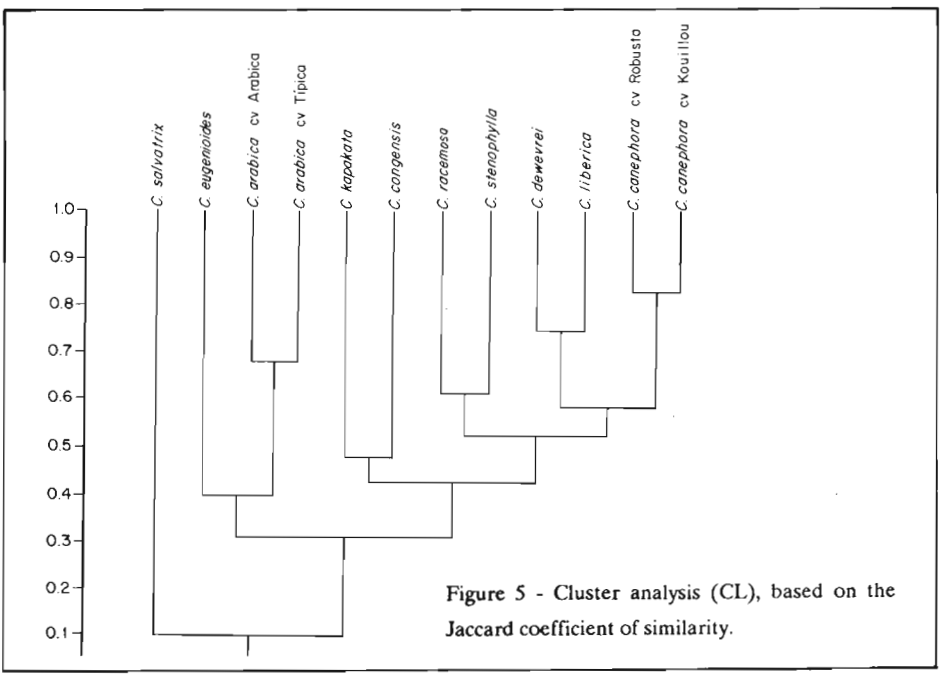

Figure 5 - Cluster analysis (CL), based on the Jaccard coefficient of similarity.

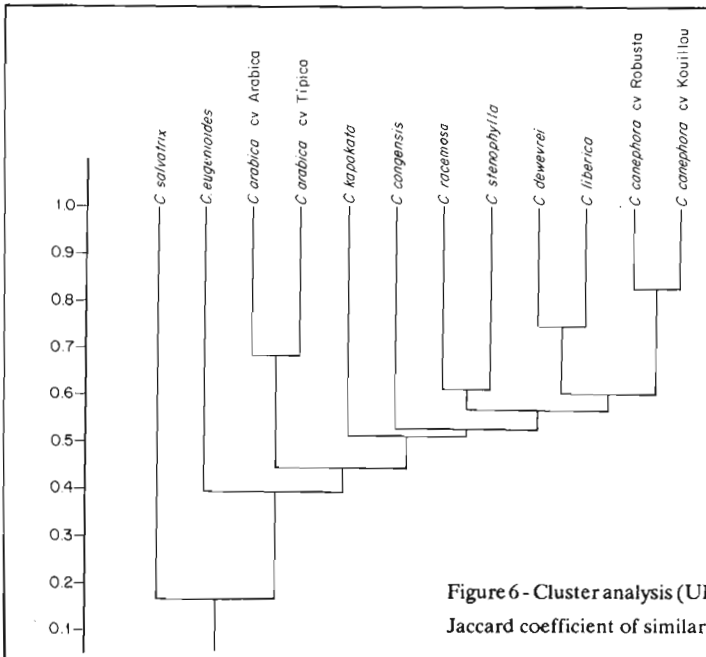

Figure 6 - Cluster analysis (UPGMA), based on the Jaccard coefficient of similarity.

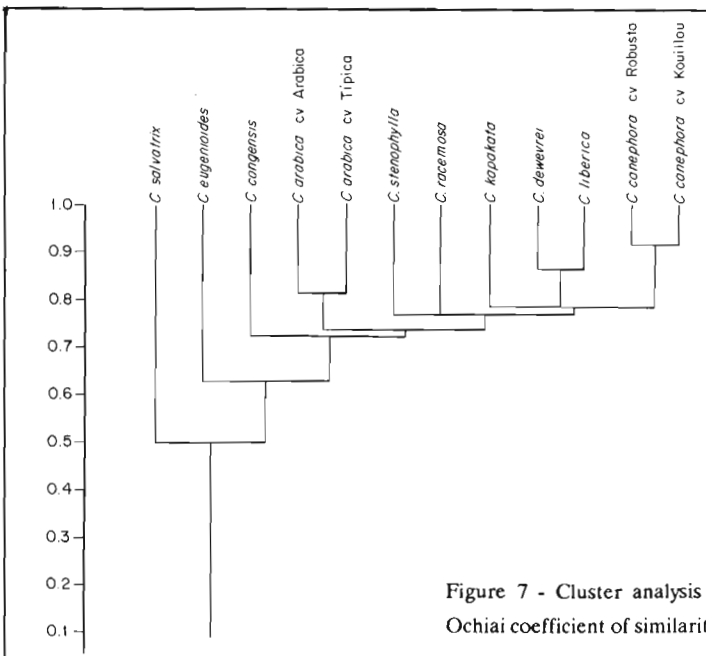

Figure 7 - Cluster analysis (SL), based on the Ochiai coefficient of similarity.

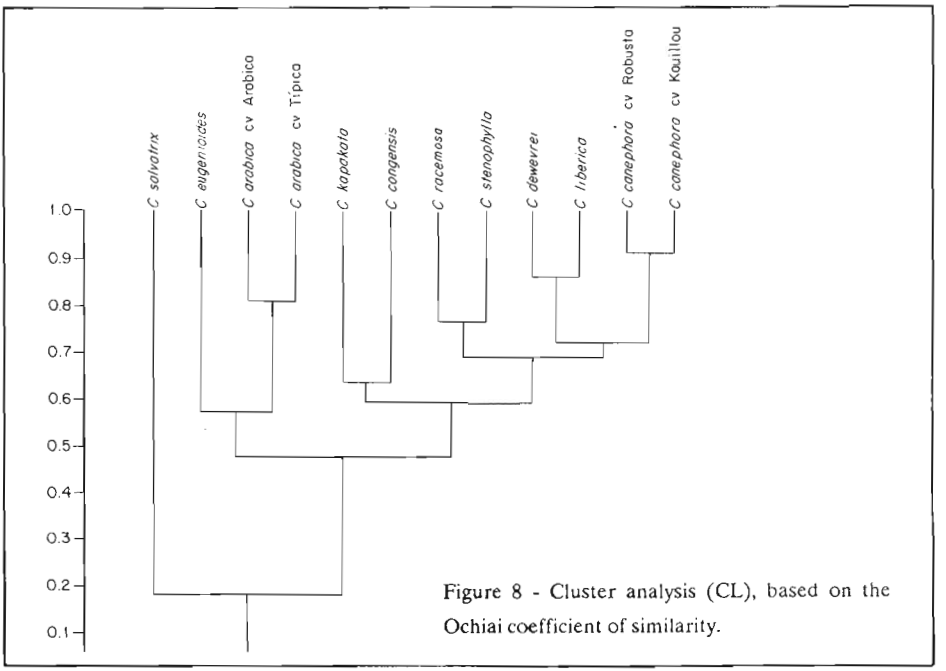

Figure 8 - Cluster analysis (CL), based on the Ochiai coefficient of similarity.

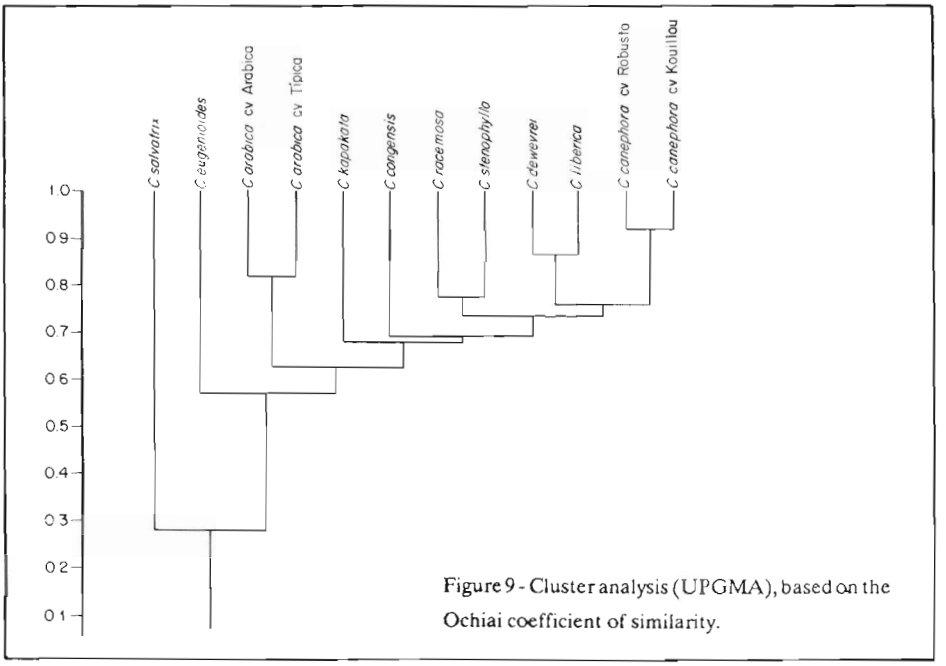

Figure 9 - Cluster analysis (UPGMA), based on the Ochiai coefficient of similarity.

when Simple-Matching was used. In Figures 3, 6 and 9, obtained respectively by the utilization of the same three coefficients of association, but with the UPGMA clustering method, the relations of affinity among the studied species were the same, with only one discrepancy that appears in Figure 3, where Simple-Matching was used, this involves the position occupied by the species *C. racemosa* and *C. stenophylla* that in this case, showed a higher similarity with the grouping formed by the species *C. dewevrei* and *C. liberica*, species that, in the two previous cases, showed a higher similarity with *C. canephora*.

CONCLUSION

The results demonstrated that, although the three coefficients of association used give different weights for the presence of the compounds studied, by pairs of species, and consider or not the absence in common of detected compounds, the final results were extremely similar, even when different clustering methods were used. In the present case, the differences were very small and were concerned with species whose taxonomic position is difficult to evaluate, not only by the morphological characteristics, but also by the interespecific hybridization and chemical data.

Jaccard and Ochiai's coefficients of association, when utilized in the study of flavonoid compounds, showed more coherent results when compared with taxonomic and hybridization studies obtained by several other authors (Chevalier, 1947; Carvalho e Monaco, 1967; Charrier, 1975; Carvalho *et al.*, 1985). The species of the subsection Pachycoffea (*C. liberica* and *C. dewevrei*), are considered closer to the species *C. canephora*, of the subsection Erythrocoffea, according to taxonomical data (Chevalier, 1947), and the results obtained in the present research. The similarity of the species *C. kapakata* with the other species studied explains the success of the obtained results in the interspecific crosses involving that species and justifies the inclusion of that species in the genus *Coffea*, as suggested by Carvalho e Monaco (1967) and Carvalho *et al.* (1985).

All the coefficient of association techniques used, as well as the clustering methods, demonstrated a high similarity between *C. stenophylla* and *C. racemosa*, both of which have small and black fruits. However Chevalier (1947), considered the species as belonging, respectively, to subsection Melanocoffea due to the colour of the fruits and subsection Mozambicoffea due to the occurrence in the Mozambique area.

The only discrepancy observed in the clustering concerned the position of *C. eugenioides*. According to the taxonomic data of Chevalier (1947), that species which has red fruits, belongs to the subsection Mozambicoffea due to its geographical distribution. However all the other reported studies, such as those made with inter-

specific hybridization (Carvalho e Monaco, 1967; Carvalho *et al.*, 1985), seem to indicate that this species is one of the possible parents of the allotetraploid species *Coffea arabica*, of the subsection Erythrocoffea (red fruit). The cluster obtained with the Complete Linkage Method, for the three coefficients of similarity used, clearly approximates the species *C. arabica* and *C. eugenioides*. The occurrence of rare phenolic compounds, that are only common to these two species (Lopes and Monaco, 1979), seems to make clear the participation of *C. eugenioides* in the formation of *C. arabica*. The other procedures of clustering are not so conclusive. The only species, whose position remains ambiguous in the present research, is *C. congensis*, of the subsection Erythrocoffea, according to the taxonomic evaluation of Chevalier (1947). It seems that the introduced plants in the Campinas collection are not a typical *C. congensis*, and there is a strong possibility that they are a hybrid between *C. canephora* cv. Robusta and *C. congensis*, known by the name Congusta (Carvalho, personal communication). *C. salvatrix*, of the subsection Mozambicoffea because of its geographical distribution, in all cases occupies a position distant from the others, but some phenolic compounds common to this species and to *C. eugenoides* (Lopes and Monaco, 1979), confirm the relationship between them, possibly by genetic introgression from *C. salvatrix* into *C. eugenoides*.

ACKNOWLEDGMENTS

We are grateful for the financial support received from the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). We are especially grateful to Dr. Alcides Carvalho for sample collection and for his help.

Publication supported by FAPESP.

RESUMO

Polpas de frutos maduros de 10 espécies do gênero *Coffea*, incluindo dois cultivares de *C. arabica* e dois outros de *C. canephora*, foram estudadas quanto ao seu conteúdo de flavonóides. Os dados obtidos foram analisados através de três coeficientes de similaridade. Simple-Matching, Jaccard e Ochiai e por três diferentes métodos de agrupamento. Single Linkage, Complete Linkage e Unweighted pair group using arithmetic averages.

Os coeficientes de associação Jaccard e Ochiai, quando utilizados no estudo de compostos flavonóides mostraram resultados mais coerentes com os resultados obtidos nos estudos de hibridização interespecífica e de taxonomia. A alta similaridade da espécie *Psilanthopsis kapakata* com as outras espécies estudadas do gênero *Coffea*, justifica sua inclusão neste gênero, como *C. kapakata*. Os agrupamentos obtidos claramente aproximam as espécies *C. arabica* e *C. eugenioides*, confirmando que

esta última espécie é um dos possíveis progenitores da espécie alotetraploide *C. arabica*. A única espécie cuja posição permanece ambígua é *C. congensis*, mas talvez isso se deva ao fato de que os exemplares introduzidos na Coleção de Campinas, não sejam típicos dessa espécie, mas provavelmente híbridos.

REFERENCES

Carvalho, A. and Monaco, L.C. (1967). Genetic relationship of selected *Coffea* species. *Ciênc. Cult. 19*: 151-165.

Carvalho, A., Medina Filho, H.P. and Fazuoli, L.C. (1985). Evolução e melhoramento do cafeeiro. *Colóquio Internacional de Citogenética de Plantas*, 1, Piracicaba.

Charrier, A. (1975). Contribution a l'etude génétique des Mascarocoffea. *7º Coloque Scientifique International sur le cafe.* Hambourg (9/14) juin 1975). ASI, 483-495.

Chevalier, A. (1942). Les caféiers du globe. II. Iconographie des caféiers sauvages et cultivés et des Rubiacées prises pour des caféiers. In: *Encycl. Biol.* (Lechevalier, P., ed.), Paris, 36p. 158 pl.

Chevalier, A. (1947). Les caféiers du globe. III. Systematique des caféiers et faux-caféiers, Maladies et insects nuisibles. In: *Encycl. Biol.* XXVIII, Fas. III (Lechevalier, P., ed.), Paris, 356p.

Leroy, J.F. (1980). Evolution et taxogenese chez les Caféiers (*Coffea* L., *Psilanthus* Hook f. et *Nostolachma* Durand). Hypothése sur leur origine. *C.R. Acad. Sci.* Paris *291*: 593-596.

Lopes, C.R. and Monaco, L.C. (1977). Estudos de quimiotaxonomia em cultivares de *Coffea arabica* L. *Turrialba 27*: 55-61.

Lopes, C.R. and Monaco, L.C. (1979). Chemotaxonomic studies of some species of the genus *Coffea*. *J. Plantat. Crops 7*: 6-14.

Lopes, C.R., Cunha, R.A. and Blotta, L.F. (1984). Estudo taxonômico do gênero *Coffea* com emprego de métodos numéricos. *Turrialba 34*: 421-430.

Sneath, P.H.A. and Sokal, R.R. (1973). *Numerical taxonomy. The principles and practice of numerical classification*. San Francisco, W.H. Freenam, 573p.

(Received March 23, 1990)