



## Genetical approach to systematics and phylogeny of Trypetinae (Diptera, Tephritidae)

ANNA R. MALACRIDA

Dipartimento di Biologia Animale, Università di Pavia, piazza Botta 9, I-27100 Pavia (Italy)

CARMELA R. GUGLIELMINO

Dipartimento di Genetica, Università di Pavia, and Istituto di Genetica e Biologia Evoluzionistica, CNR, Pavia (Italy)

GIULIANO GASPERI

LORELLA BARUFFI

PIERCARLA VILLANI

RICCARDO MILANI

Dipartimento di Biologia Animale, Università di Pavia, piazza Botta 9, I-27100 Pavia (Italy)

### ABSTRACT

Genetic variation at 25 enzyme loci (64 alleles) has been considered in the attempt of an intra-species analysis of *Ceratitis capitata*. Twenty-seven hortologous loci (122 alleles) were selected to elucidate the relationships among *Ceratitis capitata*, *Ceratitis rosa*, *Triribithrum coffeae* and *Capparimya savastanoi* of the Trypetinae subfamily. Two ordination approaches have been used for electrophoretic data: Principal Component Analysis (PCA) and cluster analysis through a tree representation. At the species level, for *C. capitata* ordination by means of PCA enabled the geographic and seasonal intraspecific differentiation to be recognized. At higher levels of taxonomy, when applied to species and genera, PCA has been used as an alternative to cluster analysis. Nei distance and UPGMA procedure have been used in both levels of systematic ordination. For species-genera level, genetic distances have been calculated using also Rogers, Cavalli and Edwards methods (UPGMA and Wagner procedure). The cophenetic correlation, and other measures, have been examined as measures of goodness of fit. All the trees give the same topology. *C. savastanoi* samples present the greatest range of distance. *C. capitata* appears closer to *T. coffeae* than to *C. rosa*. The disagreement between electrophoretic trees and the existing conventional systematic is discussed. Electrophoretic keys for distinguishing the immature stages of Trypetinae are proposed.

**KEY WORDS:** Trypetinae - Electrophoretic data - Principal Component Analysis - Cluster Analysis - Biochemical keys - Phylogenetic trees.

### ACKNOWLEDGEMENTS

We wish to thank Dr. I. M. White of C.A.B. International Institute of Entomology, London, U.K., for his assistance in the specimens

Experimental data and arguments of this article constituted the framework of an invited lecture presented at the Symposium on «Systematics theories, classification methods, and phylogeny» of the LII National Congress of the Unione Zoologica Italiana, held in Camerino (Italy), September 1988.

identification. We also thank him for providing a list of references, a large amount of taxonomic information and fruitful suggestions that have been of essential importance for this work. We are indebted to Dr. Mukiyama (University of Nairobi, Kenya); to Dr. Lindquist and Dr. E. Bush-Petersen (I.A.E.A. Vienna, Austria); to Dr. Quilici (IRAT-CIRAD, Réunion Island); to Dr. Cirio (E.N.E.A., Rome); to Dr. Del Rio (University of Sassari, Italy) for having provided the specimens studied. The manuscript was kindly read and revised by Dr. A.S. Robinson (I.M.B.B., Heraklion, Crete). This work was partially supported by I.A.E.A. (Vienna, Austria), by the Italian Ministero della Pubblica Istruzione Research Contributions (M.P.I. 40%), and by Istituto di Genetica Biochimica ed Evoluzionistica of C.N.R. of Pavia.

### INTRODUCTION

The Tephritidae flies have not been submitted to careful taxonomic revision for many years, most of the information being scattered descriptions based on traditional taxonomy (Cogan & Munro, 1980; Hancock, 1984, 1987). The existence of morphological overlap among the higher taxa categories creates the main difficulty.

The need of alternative approaches has been stressed by many authors (White, 1987), an example being the use of immunological techniques (Kitto, 1983). At lower taxa categories, several independent approaches have already provided significant information both on variability and on affinities between different taxa; members of the subfamily Trypetinae provide an example of this situation. Economically important genera such as *Ceratitis*, *Capparimya*, *Rbagoletis* and *Anastrepha*, are usually treated as members of the Trypetini tribe in the Trypetinae subfamily (Grandi, 1951). *Ceratitis* and allied genera, have often been given a subfamily status, Ceratitinae (Hancock, 1984). However, the current arrangement and the phylogenetic relations of the above mentioned genera are still disputed. New subgenera are proposed to accommodate several species (Hancock, 1984, 1987). An example is given by *Ceratitis rosa* Karsch, regarded now as a member of subgenus *Pterandrus* Bezzi of *Ceratitis* genus. It is noteworthy that, as currently defined, *Ceratitis (Ceratitis) capitata* and *Ceratitis (Pterandrus) rosa* are separated solely on the basis of male secondary sexual characters, with females being inseparable at the generic level (Hancock, 1984). Another example is offered by *Triribithrum coffeae* Bezzi previously included in the *Ceratitis* genus by the original designation of *Ceratitis nigra* Graham (Cogan & Munro, 1980). These examples serve to illustrate the shortcomings of the current arrangement of genera and the need for different methods.

Much taxonomic work has been done on the two genera *Rbagoletis* and *Anastrepha* using a comparison of molecular and morphological data in phenetic and cladistic methods (Berlocher & Bush, 1982; Berlocher, 1983; Feder *et al.*, 1988; Norrbom & Kim, 1988); in addition there are morphological and biochemical keys for the identification of most *Rbagoletis* and *Anastrepha* species (Steyskal, 1977; Berlocher, 1980).

Among the deficiencies in the taxonomic literature related to the Trypetinae subfamily, it is noteworthy that there is no recent work on the *Ceratitis* genus and on its phylogenetic relationships with the other Trypetinae. A major gap in the taxonomic literature on *Ceratitis* genus is the absence of a larval key (White, 1987).

The major fruit pest *Ceratitis capitata* Wied. did not give rise to taxonomic problems, as there is an apparent morphological uniformity within this species. However, *C. capitata* may in reality be a complex of several genetically differentiated populations (Gasperi *et al.*, 1987; Gasperi *et al.*, 1990).

A genetic analysis may offer a suitable approach to clarify the systematics and phylogeny of these species. Using a genetic approach we have attempted an intra-species analysis of *C. capitata*. We have also attempted to elucidate the relationships among the genera *Ceratitis*, *Thrirbitrum*, *Capparimyia*. Finally, we propose an electrophoretic key for the immature stages of these species.

## MATERIALS AND METHODS

### *Intraspecific variation*

Twenty-nine samples of *Ceratitis capitata*, were obtained from different geographic areas and in different seasons (Kenya: 2 samples, Réunion: 2 samples, Sardinia: 3 samples, Procida: 22 samples). The 22 Procida samples were collected in different months from 1983-1986: February (1 sample in 1985), April (1 sample in 1983), June (1 sample in 1985), July (1 sample in 1983; 3 samples in 1985; 1 sample in 1986), August (3 samples in 1985), September (3 samples in 1984; 2 samples in 1985; 5 samples in 1986), December (1 sample in 1984). The samples sizes range from 19 to 33 flies.

The numbers of samples are related to the size of the Procida population which reaches maximum density in the harvest season of host fruits (apricots, peaches, figs etc.): for this reason most of the samples of Procida have been collected in the fruit season.

### *Species and genus level systematics*

Four samples of pupae developed on coffee berries collected in Kenya were a mixture of *Ceratitis capitata* and *Ceratitis rosa*; two samples included also *Thrirbitrum coffeae*. The development on the same berries excludes variability components due to geographic, seasonal, host variation; for these reasons, these samples seemed especially suitable for giving prominence to the phylogenetic relations. Specimens representing *Capparimyia savastanoi* (2 samples) were collected in Pantelleria Island (Italy) on fruits of *Capparis*.

A voucher collection containing two representative specimens of each species was deposited in the British Museum (Natural History) of London.

### *Electrophoretic studies*

Electrophoretic analysis was performed using cellulose acetate gel (Cellogel).

The following 25 enzyme loci (64 alleles) were analyzed in the intraspecific studies on *C. capitata*: *Mpi*, *Est*<sub>6</sub>, *Mdb*<sub>2</sub>, *Hk*<sub>2</sub>, *Est*<sub>1</sub>, *Est*<sub>2</sub>, *Pgi*, *Zw*, *Pgd*, *Fb*, *Had*, *Hk*<sub>1</sub>, *Idb*, *Pgm*, *Got*<sub>1</sub>, *Got*<sub>2</sub>, *Ak*<sub>2</sub>, *Mdb*<sub>1</sub>, *Adb*<sub>2</sub>, *Gpi*, *Pgk*, *Me*, *α Gpdh*, *Aox*, *Ak*<sub>1</sub>.

Twenty-seven hortologous (Ferguson, 1980) loci (122 alleles, that produced consistently interpretable banding patterns in all species) were selected for studies on species and genus level. *Acon*<sub>1</sub> and *Acon*<sub>2</sub> have been included in the above mentioned set of loci. *C.*

*capitata* was used as a standard for each essay because electrophoretic variation in this species is the best documented in literature (Milani *et al.*, 1989).

### *Ordination and phylogenetic analysis*

Two ordination approaches have been used for Trypetinae electrophoretic data: Principal Component Analysis (PCA) and a tree representation based on more than one method. PCA is a multivariate statistical method which allows the representation of populations in a plane or in a three-dimensional space, with a (generally) good approximation of their reciprocal distance. The first two or three principal axes account for a large part of the total variability contained in the original data. The coordinates on these axes can be used as synthetic representatives of the original allelic frequencies.

Cluster analysis through a tree representation was performed using the BIOSYS-1 program of Swofford & Selander (1981). The following genetic distances were calculated: Nei (1972, 1978) distance and unbiased distance, Rogers (1972) and modified Rogers distance (Wright, 1978), Cavalli & Edwards (1967) chord and arc distance, Edwards (1971, 1974) distance. From each distance measure, trees were constructed using Unweighted Pair-Group Arithmetic Average (UPGMA) cluster analysis (Sneath & Sokal, 1973). With Rogers distances we also used Wagner procedure (Farris, 1972). The program provides, along with each dendrogram, the following four measures of goodness of fit: the cophenetic correlation, the Farris (1972) and Prager & Wilson (1976) F statistics, the percent standard deviation (Fitch & Margoliash, 1967).

## RESULTS

### *Species level: Ceratitis capitata*

Figure 1 is the scattergram of all the 29 population samples of *C. capitata*; for 22 of them, belonging to Procida, the month of collection is specified. The first axis, which accounts for 28.5% of the total variation, separates the Procida samples from the rest; the closest is Sardinia, then Kenya and Réunion; but two Procida samples, those collected in the early season, are on the same side as the geographically distinct populations. The second principal axis (ordinate) accounts for 17% of the total variation, which is mainly attributable to the scattering of the Procida samples. The position of the months of collection (i.e., summer at the bottom, winter at the top, with September in between) suggested a seasonal trend as represented by the second axis. A correlation coefficient of 0.62, with  $P < 0.01$ , is obtained from the correlation of the coordinates of the second axis with the sequence of the months in the different years from summer (June with code 1) to spring (April with code 11). The analysis of *Fst* values at loci level across months, demonstrates a very high degree of differentiation for the *Mpi* locus ( $Fst = 0.509$ ) between samples collected from December to April and the ones collected from June to September (Gasperi *et al.*, 1990).

When the genetic distances were represented by a dendrogram (Nei distances and UPGMA method), the geographic and seasonal differentiation intermingle (Fig. 2a): the first split separates Réunion samples, while the second one separates the two Procida early season samples; Kenya and the peripheral populations from Sardinia



TABLE I - Goodness of fit statistics for tree estimation procedures in *Ceratitis*, *Trirhithrum* and *Capparimyia* genera.

Methods	Statistics			
	Farris F	Prager & Wilson F	% Standard deviation	Cophenetic correlation
Nei (1978)	6.57	9.88	11.73	0.973
Nei unbiased (1978)	6.59	9.94	14.21	0.973
Rogers (1972)	1.64	4.47	5.85	0.993
Rogers (Wright, 1978)	1.27	2.98	4.00	0.996
Cavalli & Edwards-cord (1967)	1.18	2.96	4.77	0.996
Cavalli & Edwards-arc (1967)	1.38	3.15	4.97	0.995
Edwards (1971, 1974)	1.73	3.34	5.11	0.994
Wagner tree (Farris, 1972)	1.40	3.82	9.72	0.992

how the results of an electrophoretic approach may depend on the choice of biochemical loci.

*Ceratitis capitata* has a recent history of a very quick, extensive, and initially quite discontinuous colonization of tropical and subtropical countries, even very remote from the original putative source centre (subsaharian tropical regions) of the species (Saul, 1986). Initial stages of intraspecific differentiation caused by geo-climatic pressures or by genetic phenomena such as genetic drift, founder effects and bottlenecks, can be anticipated, even in the presence of the generalized morphological uniformity of this species.

At higher levels of taxonomy, when applied to species and genera, as in Figure 3, PCA can be considered a good alternative method to cluster analysis (Dunn & Everitt, 1982): the relative distance between populations or groups of populations corresponds with the main splits in the trees (Figs 3 and 4). Through PCA we know that almost all (98%) the variability contained in the original data supports the divisions in Figure 3, and that half of it supports the distinction of *Ceratitis rosa*.

#### Ordination by means of a tree

Nei distances and UPGMA clustering procedure have been used in both levels of the systematic ordination; in Figures 2 and 4a, the different range of genetic distances (D) at the species level ( $0 < D < 0.10$ ) with respect to the genera-species level ( $0 < D < 1.80$ ) agrees with the global proportional range of evolutionary time (Time =  $5 \times 10$

D) expected from a Nei neutral model of speciation (Nei, 1972). For species-genera level we analyzed several distance measures (Table I), all with UPGMA clustering methods and one with Wagner procedure. Cophenetic correlation provided generally high values for the goodness of fit, with slightly lower values for Nei distance methods. Nei methods presented the worst value of all the indices of convergence of input matrix with the constrained one; this is as expected, given the non-metric property of Nei distances (i.e., they do not obey the triangle inequality euclidean postulate). However all the trees obtained give the same topology as those chosen for Figure 4. The Wagner tree obtained from Rogers distances in Figure 4c, does not require the assumption of a constant evolutionary rate: the distance measured by the branch lengths are almost the same (given the goodness of fit estimates in Table I) as in the original distance matrix. The root, which is not necessary in a Wagner tree for the above assumption, is here in the middle of the greatest distance which proves to be that separating the *Capparimyia* genus. The two *Capparimyia* populations present the greatest range of Roger's distances (0.841 - 0.716). Distances of *T. coffeae* from *C. capitata* (0.565 - 0.521) appear smaller than those from *C. rosa* (0.739 - 0.698).

The agreement between all electrophoretic trees and the existing conventional systematics appears low. The electrophoretic phylogeny results obtained are a little surprising, as PCA axis I only separates *C. rosa* (Fig. 3).

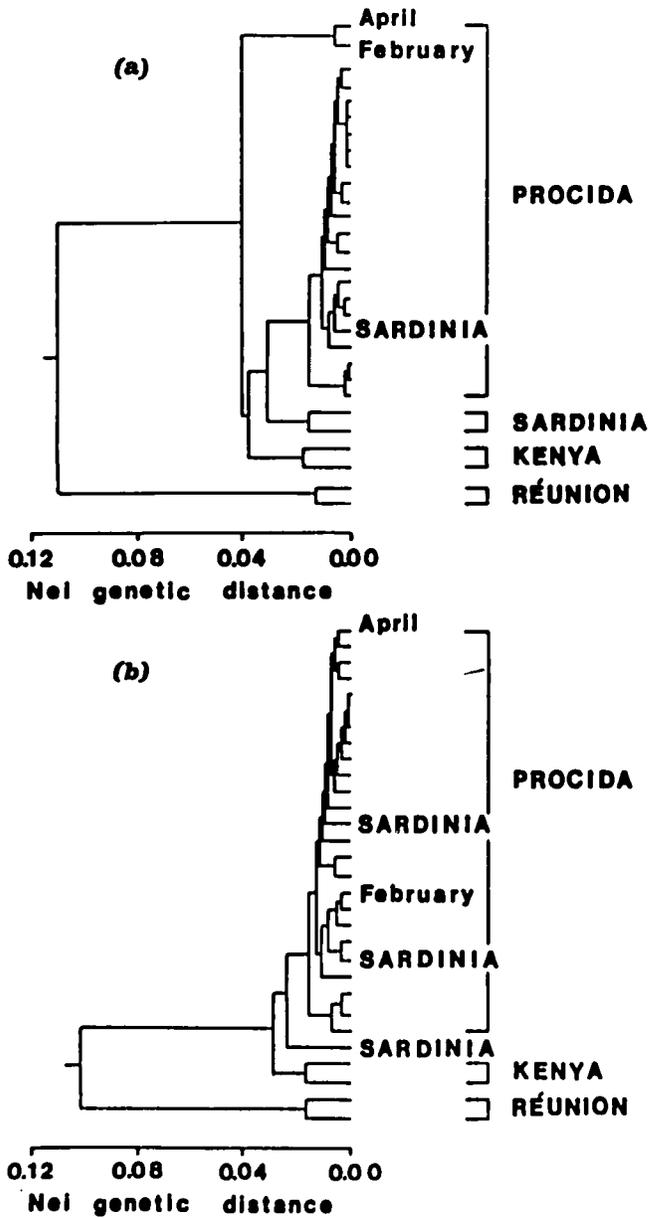


Fig. 2 - UPGMA dendrograms, based on Nei genetic distance data, of the 29 samples of *Ceratitis capitata* considered in Figure 1. The 22 samples from Procida Island were collected in different months in the years 1983-1986. In a) *Mpi* locus is included in the set of loci considered, in b) *Mpi* locus is not considered.

Furthermore, the dendrogram and the PCA plot both place *C. capitata* and *T. coffeae* closer than the two *Ceratitis* species. This result may indicate how poorly external morphology reflects relationships. As previously mentioned, from the morphological point of view, *Ceratitis* (*Ceratitis*) *capitata* and *Ceratitis* (*Pterandrus*) *rosa* are separated on the basis of the male secondary sexual characters, with females being inseparable at the generic level (Hancock, 1984). These disagreements concern taxa which had been assigned to various taxonomic positions.

On the other hand, further studies are necessary to ex-

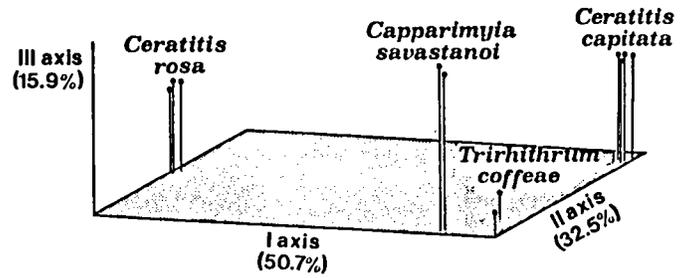


Fig. 3 - Three-dimensional plot of *Ceratitis capitata*, *Ceratitis rosa*, *Trirhithrum coffeae* and *Capparimyia savastanoi* based on PCA analysis at 27 biochemical loci. Points refer to the populations. Percentages give the portion of the overall variance explained by each axis.

clude the possibility that by chance we looked at enzyme loci remaining unaltered in some widely separated lines, but being strongly selected in *C. rosa*.

It has to be stressed that in this electrophoretic study we compared sympatric populations of *C. capitata*, *C.*

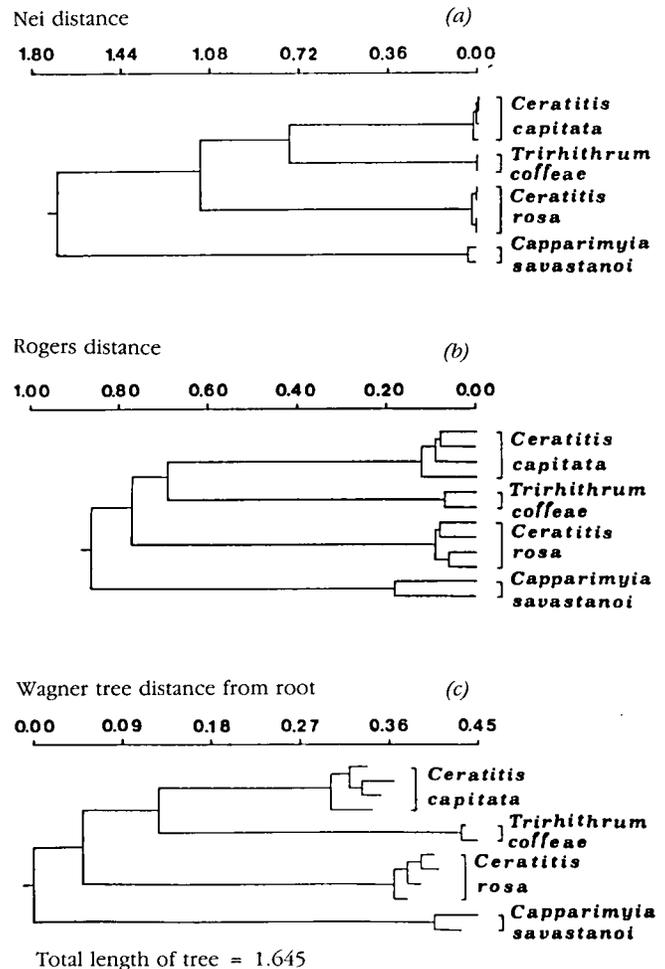


Fig. 4 - Phylogeny of *Ceratitis capitata*, *Ceratitis rosa*, *Trirhithrum coffeae* and *Capparimyia savastanoi*. The dendrograms in a) and b) are derived, respectively, from calculations of Nei and Rogers genetic distances clustered using UPGMA method. In c) Rogers distances have been used for Wagner procedure.

TABLE II - Allele frequencies at enzyme loci which permit discrimination (diagnostic loci) within and between *Ceratitis*, *Trirhithrum* and *Capparimyia* genera.

locus	alleles <sup>a</sup>	<i>Ceratitis capitata</i>	<i>Ceratitis rosa</i>	<i>Trirhithrum coffeae</i>	<i>Capparimyia savastanoi</i>
<i>Adb</i> <sub>2</sub>	130	-	1.000	-	-
	107	-	-	-	1.000
	100	1.000	-	1.000	-
<i>Aox</i>	112	-	1.000	-	1.000
	100	1.000	-	-	-
	96	-	-	1.000	-
<i>Mdb</i> <sub>2</sub>	100	1.000	-	-	-
	74	-	1.000	-	-
	69	-	-	-	1.000
	37	-	-	1.000	-
<i>Got</i> <sub>1</sub>	109	.029	.177	-	-
	126	-	.823	-	-
	117	-	-	1.000	-
	100	.971	-	-	.016
	90	-	-	-	.984
<i>Got</i> <sub>2</sub>	139	-	.111	1.000	-
	107	-	-	-	1.000
	100	1.000	.889	-	-
<i>α-Gpdb</i>	112	-	-	1.000	1.000
	100	1.000	-	-	-
	96	-	.917	-	-
	84	-	.083	-	-

<sup>a</sup> Alleles are labelled according to their relative mobility referred to that of the most common allele of *C. capitata* taken as 100.

*rosa* and *T. coffeae*, collected together in one of their putative original areas (Kenya). This fact would exclude the possibility that phylogenetic relationships were affected by the geographic and/or climatic intraspecific differentiation described, for example, for the *C. capitata* species. However, the systematic controversies on these Tephritidae, in which, very often, species form closely-knit groups (Hancock, 1987) justify the use of new and different approaches to integrate the traditional morphological method.

Tephritid taxonomy is of crucial importance also for applied entomology. Improper identification or the inability to recognize distinct populations can have drastic and costly consequences for pest control management. In this connection, the need for larval keys has often been stressed (White, 1987) as one of the major gaps in the taxonomic literature of Trypetinae. The Trypetinae include the major pest genera in which the larvae are the damaging stage. It is desirable, therefore, to be able to identify the stage which is actually injurious without having to wait for adults to appear. An example is offered by the three species *C. capitata*, *C. rosa*, *T. coffeae* which, in the Afrotropical regions (Kenya, for example), live and compete for the coffee berries.

The identification of suitable diagnostic loci can represent the starting point for constructing an electrophoretic key for the discrimination of the immature stages as previously indicated (Gasperi *et al.*, 1987).

#### REFERENCES

- Berlocher S. H., 1980 - An electrophoretic key for distinguishing species of the genus *Rbagoletis* (Diptera: Tephritidae) as larvae, pupae, or adults. *Entomol. Soc. Am.*, 73: 131-137.
- Berlocher S. H., 1983 - Genetic changes coinciding with the colonization of California by the walnut huskfly, *Rbagoletis completa*. *Evolution*, 38: 906-918.
- Berlocher S. H., Bush G. L., 1982 - An electrophoretic analysis of *Rbagoletis* (Diptera: Tephritidae) phylogeny. *Syst. Zool.*, 31: 136-155.
- Cavalli Sforza L. L., Edwards A. W. F., 1967 - Phylogenetic analysis: models and estimation procedures. *Evolution*, 21: 550-570.
- Cogan B. H., Munro H. K., 1980 - Superfamily Tephritoidea - 40 Family Tephritidae. In: R. W. Crosskey (ed.), *Catalogue of the Diptera of the Afrotropical region*. British Museum (Natural History), London, pp. 518-554.
- Dunn D., Everitt B. s., 1982 - An introduction to mathematical taxonomy. Cambridge Univ. Press., Cambridge.
- Edwards A. W. F., 1971 - Distances between populations on the basis of gene frequencies. *Biometrics*, 27: 873-881.

- Edwards A. W. F., 1974 - Distance measures for phylogenetic trees. *In*: J. F. Crow & C. Denniston (eds.), Genetic distance. Plenum Press, New York, pp. 41-43.
- Farris J. S., 1972 - Estimating phylogenetic trees from distance matrices. *Am. Nat.*, 106: 645-668.
- Feder J. L., Chilcote C. A., Bush G. L., 1988 - Genetic differentiation between sympatric host races of the apple maggot fly *Rhagoletis pomonella*. *Nature (Lond.)*, 336: 61-64.
- Ferguson A., 1980 - Biochemical systematics and evolution. Blackie ed., Glasgow.
- Fitch W. M., Margoliash F., 1967 - Construction of phylogenetic trees. *Science*, 155: 279-284.
- Gasperi G., Malacrida A. R., Guglielmino C. R., Milani R., 1990 - Electrophoretic multilocus analysis for the study of natural populations of the medfly *Ceratitis capitata*. *In*: Genetic sexing of the Mediterranean Fruit Fly. International Atomic Energy Agency, Vienna, pp. 91-96.
- Gasperi G., Malacrida A., Milani R., 1987 - Protein variability and population genetics of *Ceratitis capitata*. *In*: A. P. Economopoulos (ed.), Fruit Flies. Elsevier Science Publ., Amsterdam, pp. 149-157.
- Grandi G., 1951 - *In*: Introduzione allo studio dell'entomologia. Edizioni Agricole, Bologna, Vol. II. pp. 421-437.
- Hancock D. L., 1984 - Ceratitinae (Diptera: Tephritidae) from Malagasy subregion. *J. entomol. Soc. sth. afr.*, 47: 277-301.
- Hancock D. L., 1987 - Notes on some African Ceratitinae (Diptera: Tephritidae), with special reference to the Zimbabwean fauna. *Trans. Zim. Sci. Ass.*, 63: 47-57.
- Kitto G. B., 1983 - An immunological approach to the phylogeny of the Tephritidae. *In*: R. Cavalloro (ed.), Fruit flies of economic importance. Balkema A. A. Publ., Rotterdam, Netherlands, pp. 203-211.
- Milani R., Gasperi G., Malacrida A., 1989 - Biochemical genetics (of *Ceratitis capitata*). *In*: A. S. Robinson & G. H. S. Hooper (eds.), Fruit flies, their biology, natural enemies and control W. P. C., Elsevier Publ., Amsterdam, The Netherlands Vol. 3B; chapter 6.1.4., pp. 35-56.
- Nei M., 1972 - Genetic distance between populations. *Am. Nat.*, 106: 283-292.
- Nei M., 1978 - Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89: 583-590.
- Norrbom A. L., Kim K. C., 1988 - A list of the reported host plants of the species of *Anastrepha* (Diptera: Tephritidae). United States Department of Agriculture, APHIS 81-52.
- Prager E. M., Wilson A. C., 1976 - Congruency of phylogenesis derived from different proteins. A molecular analysis of the phylogenetic position of cracid birds. *J. mol. Evol.*, 9: 45-57.
- Rogers J. S., 1972 - Measures of genetic similarity and genetic distance. *Studies in Genetics, Univ. Texas Publ.*, 7213: 145-153.
- Saul S. H., 1986 - Genetics of the Mediterranean fruit fly (*Ceratitis capitata*) (Wiedemann). *In*: E. R. Gordon (ed.), Agricultural zoology reviews. Intercept, Ponteland, Newcastle upon Tyne, Vol. 1, pp. 73-108.
- Sneath P. H. A., Sokal R. R., 1973 - Numerical taxonomy. W. H. Freeman, San Francisco, 573 pp.
- Steyskal G. C., 1977 - Pictorial key to species of the genus *Anastrepha* (Diptera: Tephritidae). Washington, pp. 1-61.
- Swofford D. L., Selander R. B., 1981 - BIOSYS-1. A computer program for the analysis of allelic variation in genetics. Univ. Illinois, Urbana.
- White I. M., 1989 - The state of fruit fly taxonomy and future research priorities. *In*: R. Cavalloro (ed.), Fruit flies of economic importance, 87. A. A. Balkema, Rotterdam, Neth., pp. 543-552.
- Wright S., 1978 - Evolution and the genetics of populations, vol. 4. Variability within and among natural populations. University of Chicago Press, Chicago.