Bulletin of Entomological Research (2006) **96**, 279–288 DOI: 10.1079/BER2006425

New contributions towards the understanding of the phylogenetic relationships among economically important fruit flies (Diptera: Tephritidae)

M.D. Segura[1], C. Callejas[2], M.P. Fernández[3] and M.D. Ochando[2]*

[1]Centro Nacional de Biotecnologia, Campus de la Universidad Autonoma de Madrid, Cantoblanco, 28049 Madrid, Spain: [2]Departamento de Genética, Facultad de Biología, Universidad Complutense de Madrid, Avenida Complutense s/n, 28040, Madrid, Spain: [3]Instituto Madrileño de Investigación y Desarrollo Rural, Agrario y Alimentario, Finca El Encín, 28800 Alcalá de Henares, Spain

Abstract

Fruit flies (Diptera: Tephritidae) are a species-rich and economically important group. The phylogenetic relationships among the many taxa are still to be fully resolved and the monophyly of several groups is still to be confirmed. This paper reports a study of the phylogenetic relationships among 23 economically important tephritid species (representing several major lineages of the family) which examines the sequence of a region of mitochondrial DNA encompassing the *cytb*, $tRNA^{Ser}$ and *ND1* genes. Substitutions characteristic of particular taxa were found that could help classify members of the family at any developmental stage. The trees obtained by the maximum parsimony, neighbour joining and maximum likelihood methods were generally compatible with present morphological classification patterns. However, the data reveal some characteristics of the phylogenetic relationships of this family that do not agree with present classifications. The results support the probable non-monophyletic nature of the subfamily Trypetinae and suggest that *Bactrocera cucurbitae* (Coquillet) is more closely related to the genus *Dacus* than to other species of *Bactrocera*.

Keywords: Tephritidae, mitochondrial DNA, genetic distance, molecular phylogeny

Introduction

The family Tephritidae, which includes flies usually known as fruit flies, is one of the dipteran families with the largest number of species. As of July 2000, 4352 species (grouped into 481 genera) were recognized worldwide, many of which are important agricultural pests. The true number of species is much higher since many remain undescribed (USDA, 2004, URL: http://www.sel.barc.usda.gov/diptera/tephriti/TephClas.htm).

Despite the many taxonomic studies performed on this family, most have been based on morphological and ecological traits, and no fully satisfactory classification has yet been produced. Different subfamilies and tribes have been described, but the limits of many taxa (and the relationships among them) are unclear, and the monophyletic nature of some taxa still needs to be confirmed. One of the most

*Author for correspondence
Fax: 34 91 394 4844
E-mail: dochando@bio.ucm.es

https://doi.org/10.1079/BER2006425 Published online by Cambridge University Press

Table 1. Species analysed, their origin, and the year and accession numbers of the corresponding sequences. Code: first letter of genus and species name, 'm' for males and 'h' for females, or 1 or 2 when there are two males, two females or two pupae.

Species	Origin	Year	Accession numbers
Anastrepha ludens (Loew)	Mexico	1999	AY096075
Anastrepha obliqua (Macquart)	Mexico	1999	AY096076
Anastrepha serpentina (Wiedemann)	Mexico	1999	AY096077
Anastrepha striata (Schiner)	Mexico	1999	AY096078
Anastrepha suspensa (Loew)	USA (Florida)	1999	AY096079
Bactrocera cacuminata (Hering)	Australia	1999	AY096080
Bactrocera cucurbitae (Coquillet)	USA (Hawai)	1999	AY096081
Bactrocera dorsalis (Hendel)	USA (Hawai)	1999	AY096082, AY096083
Bactrocera latifrons (Hendel)	USA (Hawai)	1999	AY096084
Bactrocera neohumeralis (Hardy)	Australia	1999	AY096085, AY096086
Bactrocera oleae (Rossi)	Spain	1998	AY096084
Bactrocera tryoni (Frogatt)	Australia	1999	AY096088, AY096089
Bactrocera zonata (Saunders)	Mauritius	1998	AY096090
Ceratitis capitata (Wiedemann)	Spain	1997	AY096091
Ceratitis cosyra (Walker)	South Africa	1999	AY096092, AY096093
Ceratitis rosa (Karsch)	Mauritius	1998	AY096094, AY096095
Ceratitis fasciventris (Bezzi)	Kenya	2000	AY096096
Dacus ciliatus (Loew)	Réunion	1998	AY096097
Dacus demmerezi (Bezzi)	Réunion	1998	AY096098
Neoceratitis cyanescens (Bezzi)	Réunion	1998	AY096099
Rhagoletis cerasi (Linnaeus)	Switzerland	1998	AY096100, AY096101
Rhagoletis pomonella (Walsh)	USA (New York)	1999	AY096102
Toxotrypana curvicauda (Gerstaecker)	USA (Florida)	1998	AY096103

common problems in classifying this family is the existence of a number of closely related sibling species that, morphologically, are not clearly distinct. In addition, nearly all the systematic traits that can be used in the classification of the Tephritidae, especially at levels higher than the genus, are autopomorphies of particular taxa. New criteria for defining the phylogenetic position of each taxon are therefore needed. Molecular data may allow the problem to be tackled in a new way, and provide a tool to test existing classifications.

Mitochondrial DNA (mtDNA) has become a common molecular marker in phylogenetic and population genetics studies in animals. This marker has been used a number of times in insect studies, particularly for tephritid systematics, ranging from examinations of families (Han & McPheron, 1997, 2000) to the study of tribes (Han, 2000; McPheron *et al*., 2000), genera (McPheron & Han, 1997; Muraji & Nakahara, 2001; Jamnongluk *et al*., 2003a; Smith *et al*., 2003) and even species complexes (Jamnongluk *et al*., 2003b). The results obtained have helped decipher aspects of tephritid phylogeny that could not have been solved from morphological and ecological data, and have allowed the validity of molecular information in improving the classification of this family to be tested.

The aim of the present work was to determine the value of a region of mitochondrial DNA for inferring phylogenetic relationships among the taxa of the Tephritidae. This region includes the 3′ end of the cytochrome b gene (*cytb*), a serine transfer RNA gene ($tRNA^{Ser}$), and the 3′ end of subunit 1 of the NADH dehydrogenase gene (*ND1*) (in the negative strand). This region constitutes a novel piece of mtDNA for establishing tephritid relationships. Twenty three economically important species representing several major pest lineages of this family were examined.

Materials and methods

Species samples

Table 1 shows the 23 tephritid species tested, which represent two subfamilies and seven genera of the Tephritidae. All the species sampled are from pest genera and all are pest species. Two individuals from each species were sequenced to take into account possible intraspecific variation and to corroborate data accuracy.

Voucher specimens of all species were deposited at the Genetics Department, Facultad de Ciencias Biológicas, Universidad Complutense de Madrid, Spain.

The outgroup species used were *Drosophila yakuba* Burla (family Drosophilidae, superfamily Ephyroidea), and *Chrysomya chloropyga* (Wiedemann) and *Cochliomyia hominivorax* (Coquerel) (family Calliphoridae, superfamily Oestroidea). Like the Tephritidae (superfamily Tephritoidea), all three belong to the order Diptera, suborder Brachycera, infraorder Muscomorpha. The GenBank sequences used for comparative analyses were those with accession numbers NC 001322, NC 002697 and NC 002660 (http://www.ncbi.nlm.nih.gov/Genbank).

DNA extraction, amplification and sequencing

Total DNA was isolated from individual flies following the method of Reyes *et al*. (1997).

The primers used for PCR amplification were CBF1 (ACATGAATTGGAGCTCGACCAGT) and NR1 (GGTACATTACCTCGGTTTCGTTATGAT). These were designed at the laboratory of Dr C. Fleming (Applied Plant Science, The Queen's University of Belfast, UK), and amplify a 295 bp-long mitochondrial DNA region in *D. yakuba*.

Amplifications were performed in 100 µl reaction volumes containing 4 µl DNA, 10 µl Ecotaq PCR buffer

https://doi.org/10.1079/BER2006425 Published online by Cambridge University Press

(Ecogen), 2 mM $MgCl_2$, 10 pmol CBF1, 10 pmol NR1, 0.2 mM dNTPs and 2.5 U EcoTaq DNA polymerase (Ecogen). The temperature profile included an initial denaturation step of 94°C for 5 min followed by 30 cycles of 94°C for 30 s, 58°C for 1 min and 72°C for 1 min, and a final extension step of 72°C for 6 min.

Double-stranded amplified products were purified using the High Pure PCR Product Purification Kit (Boehringer-Manheim) and used as templates for sequencing reactions. These reactions were performed using an ABI PRISM 377 DNA sequencer (Applied Biosystems, Inc.). Both strands were sequenced to corroborate the sequences. All sequences were sent to GenBank; table 1 shows their accession numbers.

Data analysis

Sequences were aligned using CLUSTAL W software (Thompson *et al*., 1994).

The analysis of nucleotide composition, the overall transition : transversion (ts : tv) ratio and the variable and parsimony-informative positions were calculated using PAUP* 4.0b8 software (Swofford, 2001).

The MEGA package (Kumar *et al*., 2001) was used to calculate the average and pairwise genetic distances (percentage substitutions).

When the two sequences of a species were identical, one was excluded from the analysis. The sequences of genes *cytb* and *ND1* were translated into amino acid sequences using the Editseq program (DNASTAR package, Inc. 1995). The alignment of these sequences was performed using CLUSTAL W software.

With respect to phylogenetic analyses, no single method has proven to be the best in all situations (Swofford *et al*., 1996). Three different methods were therefore used in this study: the maximum parsimony (MP, Fitch, 1971), neighbour joining (NJ, Saitou & Nei, 1987) and maximum likelihood (ML, Felsenstein, 1981) methods.

Parsimony analysis was performed using the PAUP* 4.0b8 program (Swofford, 2001); all traits were treated as unordered and equally weighted, and all gaps were taken as fifth bases. Starting trees were obtained via stepwise analysis involving the random addition of sequences (10 replicates). Bootstrapping was conducted using the heuristic tree bisection reconnection (TBR) swapping search procedure, with a maxtree setting of 100 trees. A strict consensus tree of the 10 most parsimonious trees was calculated.

NJ and ML analyses were conducted using the PHYLIP 3.6a2 software package (Felsenstein, 2001). The nucleotide substitution model used was F84 (with gamma correction). The alpha parameter was calculated using PAUP* 4.0b8 software, employing an iterative approximation (Swofford *et al*., 1996). Support for the NJ topology was tested by bootstrapping (1000 replicates). For the ML analysis, sequences (10 replicates) were added stepwise. Bootstrapping of the analysis was performed with 1000 replicates, using the subtree pruning and regrafting (SPR) procedure.

Results

Sequence analysis

Between 286 and 301 bp were sequenced in the 23 species, amounting to a total of 308 alignable positions, including the gaps inserted to improve alignments. The final alignment is available as a NEXUS file from the authors. The sequence corresponds to positions 11.546–11.840 in the *D. yakuba* mtDNA sequence (Clary & Wolstenholme, 1985). Of 308 sites, 156 were variable and 140 parsimony-informative. The average nucleotide composition across all taxa was A : 45, T : 35, G : 9, C : 11. The ts : tv ratio was 2.48.

The uncorrected sequence divergence among the taxa ranged from 0% (between *Bactrocera tryoni* (Frogatt) and *Bactrocera neohumeralis* (Hardy) and between *Ceratitis rosa* (Karsch) and *Ceratitis fasciventris* (Bezzi)) to 18.3% (between *Rhagoletis cerasi* (Linnaeus) and *Bactrocera cacuminata* (Hering)). The overall mean sequence divergence was 12.1%.

In *D. yakuba*, the analysed sequence corresponded to the last 106 bp of the gene *cytb*, an intergenic region of 6 bp, the $tRNA^{Ser}$ gene and, in the minus strand and overlapping with the *tRNA* gene, the last 145 bp of gene *ND1*.

Variable sites were not randomly distributed across the sequenced region. Substantial variation was seen in the third codon position of genes *cytb* and *ND1* (55.8% and 61.7% respectively) and all indels were on the 3′ end of these genes (on the last 15 sites of gene *cytb* and the last 21 of gene *ND1*). The length of the intergenic region was also quite variable. However, the length of the $tRNA^{Ser}$ gene was quite conserved in these species. The approximate base positions corresponding to the loops and stems in the *tRNA* sequence were estimated by adapting the previously reported secondary structure model of *D. yakuba* (Clary *et al*., 1983). Of the 20 mutations in $tRNA^{Ser}$, nine were in the loops (none in the anticodon loop) and 11 in the stems. Among the latter, there was one deletion at the 5′ end, one insertion and nine substitutions, seven of which are compensated in all the species affected. Thus, only two mutations affected annealing in the stems. Some of the substitutions and indels found in $tRNA^{Ser}$ were common to all the members of the same taxonomic group (table 2). Five mutations affected all the species of *Anastrepha* examined, one substitution affected the two species representing the genus *Rhagoletis*, and substitutions at positions 10 and 15 affected all members of the tribe Toxotrypanini.

The regions of the sequence coding for cytochrome b (Cytb) and subunit 1 of NADH dehydrogenase (ND1) were translated to obtain the amino acid sequences. A 34 amino acid-long sequence corresponding to Cytb was obtained for all the species except *Rhagoletis cerasi* and *Rhagoletis pomonella* (Walsh), which had an additional Leu at the carboxyl end of

Table 2. Nucleotides characteristic of certain taxa and their position in the $tRNA^{Ser}$ nucleotide sequence.

	$tRNA^{Ser}$	
	Position	Nucleotide
Tribe Toxotrypanini	10	T
	15	A
Genus *Anastrepha*	1	(deletion)
	4	T
	12	G
	14	C
	18	C
	19	C
Tribe Carpomyini		
Genus *Rhagoletis*	4	T

https://doi.org/10.1079/BER2006425 Published online by Cambridge University Press

Table 3. Amino acids characteristic of certain taxa and their position in the Cytb and ND1 amino acid sequence.

	Cytb		*ND1*	
	Position	Amino acid	Position	Amino acid
Tribe Toxotrypanini	–	–	289	Cys
	–	–	313	Leu
Genus *Anastrepha*	346	Asp	–	–
	–	–	316	Asp
fraterculus group			304	Val
A. serpentina + *A. striata*	–	–	307	Phe
			304	Ile
Tribe Carpomyini	346	Asn	–	–
	–	–	311	Leu
Tribe Dacini	–	–	297	Phe
Genus *Bactrocera*	–	–	304	Leu
B. cacuminata + *B. dorsalis*	–	–	307	Met
Tribe Ceratitini	–	–	310	Ser
			323	Asn
Genus *Ceratitis*	351	Thr	–	–

the protein. Comparing the amino acid sequence with the sequence of *D. yakuba*, 14 substitutions were found, one of them common to all the tephritids studied. Thus, 13 substitutions were found within the family Tephritidae. As for the nucleotide sequence of $tRNA^{Ser}$, some of these substitutions were common to all the members of some taxonomic groups (table 3). One substitution affected members of the genus *Anastrepha*, while another affected the two *Rhagoletis* species studied. Finally, one substitution affected all the species of *Ceratitis* studied.

Translation of the mtDNA region coding for ND1 rendered a sequence 44 amino acids long in all species except for those representing the genus *Dacus* (one insertion was seen in *Dacus demmerezi* (Bezzi), and one insertion and two deletions in *Dacus ciliatus* (Loew)). After the alignment of these sequences with the sequence of *Drosophila yakuba*, 24 substitutions were observed, 21 of them within the family Tephritidae. Also, some substitutions were shared by all the members of the same taxonomic group (table 3). Two substitutions were common to all the members of tribe Toxotrypanini, two were common to all the species of *Anastrepha*, and two to all members of the *fraterculus* group of species belonging to the genus *Anastrepha*. One shared substitution was seen in the tribe Carpomyini. One substitution was also common to all members of the tribe Dacini and, within this tribe, one substitution was shared by all the *Bactrocera* species. One substitution was common to the species representing the *dorsalis* complex, *B. cacuminata* and *B. dorsalis* (Hendel). Finally, two substitutions were shared by all members of the tribe Ceratitini, and one substitution was common to all members of the genus *Ceratitis* (part of the latter tribe).

Phylogenetic relationships

Maximum parsimony (MP) analysis produced 10 equally parsimonious trees (length TL = 442, consistency index CI = 0.5023 and retention index RI = 0.8151). Figure 1 shows the strict consensus tree of these 10 trees (TL = 454, CI = 0.4890 and RI = 0.8050).

Figures 2 and 3 show the majority rule consensus trees resulting from the neighbour joining (NJ) and maximum likelihood (ML) analyses. The F84 substitution model with gamma correction ($\alpha = 0.2838$) was used in both trees.

All the trees showed similar topologies and grouped together the species belonging to the genera *Bactrocera*, *Dacus*, *Ceratitis* and *Neoceratitis*. The bootstrap support for this group was low in the ML tree (21), but high in the NJ (90) and MP (87) trees. This group was always split into two subgroups, one with the genera *Bactrocera* and *Dacus*, the other with the genera *Ceratitis* and *Neoceratitis*. The associated bootstrap values were high in both the NJ (52) and MP (78) trees.

Three other genera, *Rhagoletis*, *Anastrepha* and *Toxotrypana*, were included in this study. All the trees showed the genera *Anastrepha* and *Toxotrypana* to be more closely related to each other than to *Rhagoletis*. The *Rhagoletis* species were grouped together with *Anastrepha* and *Toxotrypana* in the neighbour joining tree. However, this assemblage was weakly supported (bootstrap value 21). In the MP and ML trees, the two species of *Rhagoletis* formed a separate group with high statistical support (bootstrap values 84 and 74 respectively).

The genera *Toxotrypana* (represented in this study by *T. curvicauda* (Gerstaecker)) and *Anastrepha*, are closely related according to both the NJ and ML trees. The bootstrap support for this group was highly significant in the NJ tree (62) and even higher in the ML tree (76). *Toxotrypana* was placed as a separate group only in the MP tree, although it is more closely related to *Anastrepha* than to *Rhagoletis*.

Finally, species belonging to the same genera were grouped together in all the trees, except for *Bactrocera cucurbitae* (Coquillet). In all three trees, this species was positioned as a group apart from the two main clusters formed by *Dacus* and the other species of *Bactrocera*. The species groups formed were nearly identical in all the trees.

Discussion

The Biosystematic Database of World Diptera (http://www.sel.barc.usda.gov/diptera/names/FamClass.htm) classify the species included in this analysis into two main clusters. The first, the subfamily Dacinae, includes the tribes Dacini (genera *Dacus* and *Bactrocera*) and Ceratitini (genera

https://doi.org/10.1079/BER2006425 Published online by Cambridge University Press

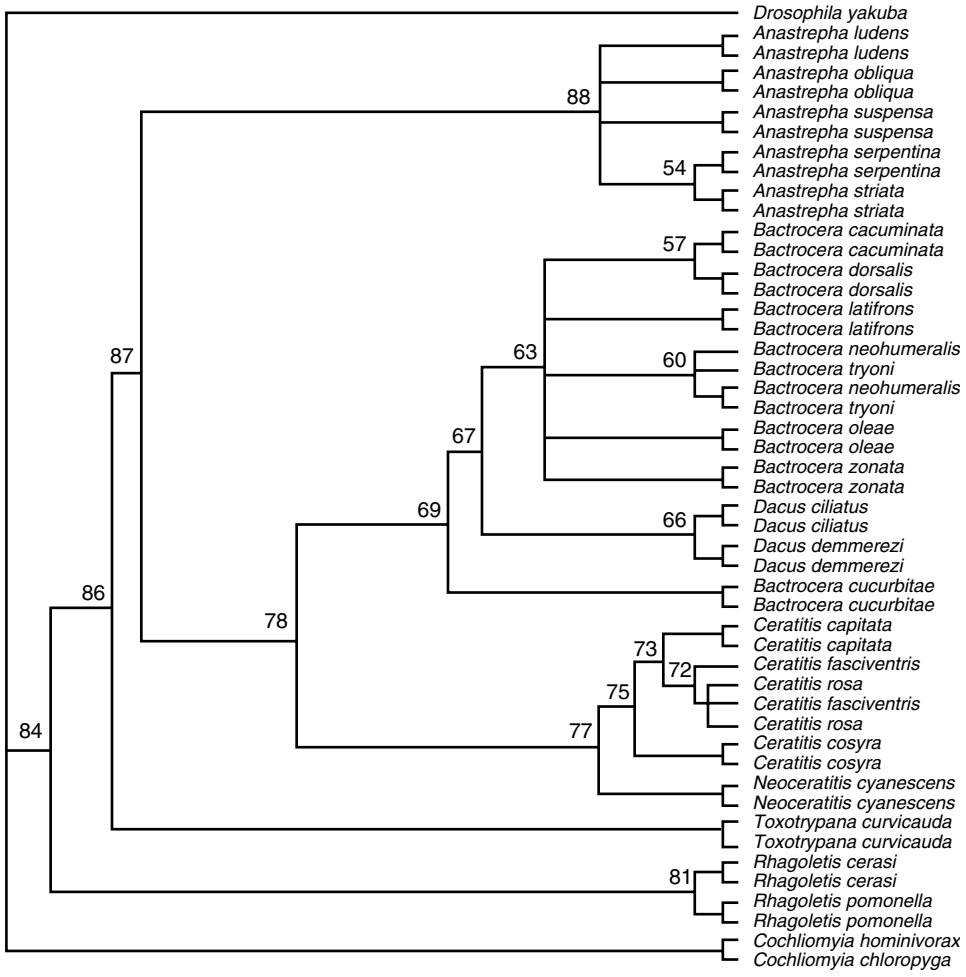

Fig. 1. Strict consensus tree of the 10 most parsimonious trees (TL = 454, CI = 0.4890). Bootstrap values for 1000 replicates are given at the nodes.

Ceratitis and *Neoceratitis*). The second large assemblage is the subfamily Trypetinae. Within this, the genera analysed form part of the tribes Toxotrypanini (*Anastrepha* and *Toxotrypana*) and Carpomyini (*Rhagoletis*).

Sequence analysis

The A : T : G : C ratio for the mtDNA sequence analysed was 45 : 35 : 9 : 11, i.e. A + T = 80%. This A-T bias in the mtDNA has been observed not only in the family Tephritidae (80.8% in Han & MacPheron, 1997; 80% in Han & McPheron, 2000), but also in other dipteran taxa (Simon *et al.*, 1994). This proportion might indicate a common ancestry, or be due to a consistent asymmetrical substitutional bias. As a possible basis for selection in favour of A + T nucleotides in *Drosophila* mtDNAs, it has been suggested that once DNA has become rich in A + T during long-term evolution (for reasons that are obscure), the enzymes responsible for transcription and replication of this DNA become adapted to function optimally with it, and therefore less optimally with more G + C-rich DNA (Wolstenholme & Clary, 1985).

The average ts : tv ratio in the studied sequence was 2.48. As for other dipteran taxa, the family Tephritidae has a high ts : tv ratio that tends to become smaller as distances increase and transitions become saturated (Han & McPheron, 1997).

The length of the sequence ranged from 286 to 301 bp. All the indels were situated in regions that do not affect gene function, namely the intergenic region, the 3′-end of the genes coding for proteins, and regions of $tRNA^{Ser}$ other than the anticodon loop or paired regions of its secondary structure. With respect to substitutions, the protein-coding genes were more variable at the third position of the codons, which experience less strong selection (55.8% of substitutions in *cytb* and 61.7% in *ND1*). In the region corresponding to the tRNA gene, 20 substitutions were seen, but only two of these were located in base-pair regions in the stems that might affect the secondary structure of the tRNA. The anticodon loop and the base-paired region that precedes it were strictly conserved. In summary, both indels and substitutions occurred mainly in regions not expected to suffer strong selective pressure.

The uncorrected sequence divergence among the taxa in this study ranged from 0% (between *Bactrocera tryoni* and *B. neohumeralis*, the members of the *tryoni* species complex, and *Ceratitis fasciventris* and *C. rosa* (only recently recognized as different species due to their great similarity))

https://doi.org/10.1079/BER2006425 Published online by Cambridge University Press

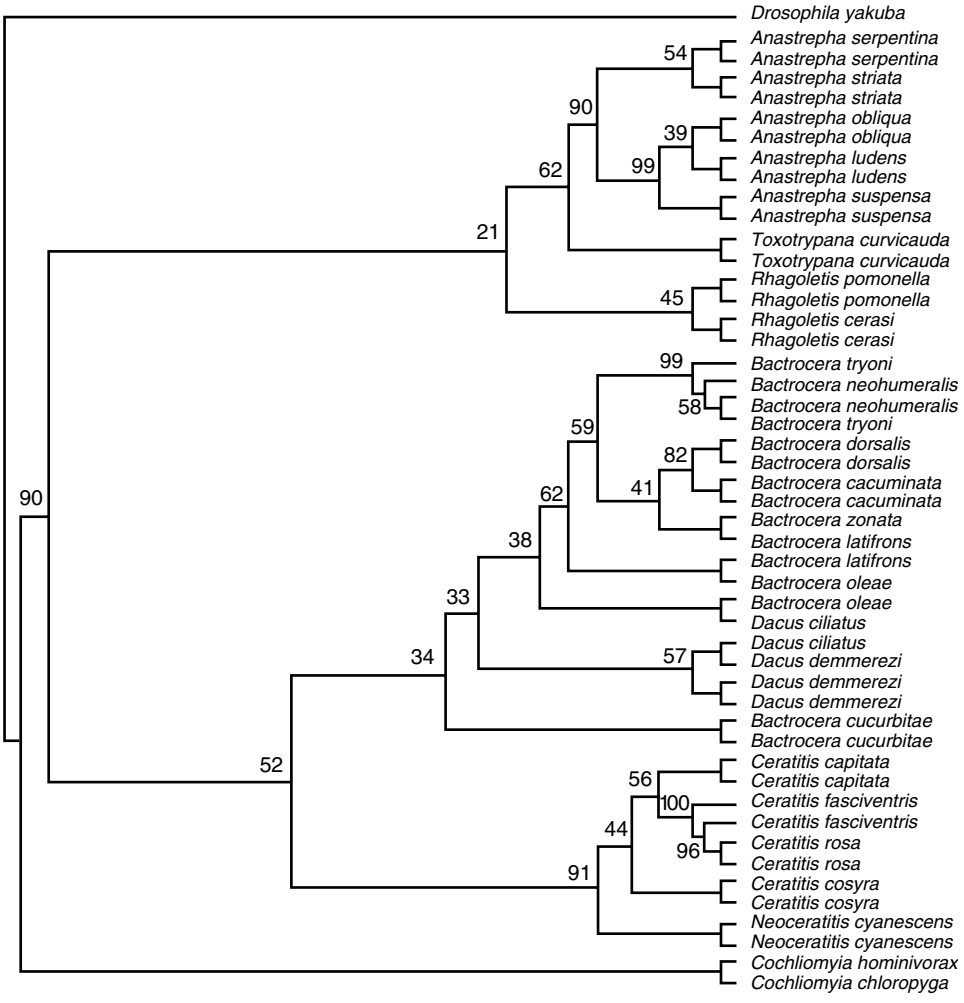

Fig. 2. Majority rule consensus tree based on neighbour joining trees. Bootstrap values for 1000 replicates are given at the nodes.

to 18% (between *B. cacuminata* and *Rhagoletis cerasi*), with an average of 12%. These distances are quite similar to those obtained by Han & McPheron (1997) for the same family when analysing another mtDNA region (3–18%, average 11%).

Sequence divergences among members of the same genus were lower than 10.2%, except for the genus *Bactrocera*. In the latter, the distance between *B. cucurbitae* and the other species were as high as 13.4% in some cases. The differences among *B. cucurbitae* and the other species of the genus are discussed below.

As expected, the distances became larger as higher taxonomic groups were compared.

Taxonomic identification

Several substitutions were seen in *tRNA*^{Ser} (table 2) and in the amino acid Cytb and ND1 sequences (table 3), which were common to all members of the same taxonomic groups. These might be considered autopomorphies of these taxonomic groups and be used to identify them. In the family Tephritidae, molecular methods may be particularly useful since they allow the identification of individuals independent of their developmental stage, whereas most of the taxonomic information available is based only on adult characteristics. Different techniques have been used for the identification of tephritids including amplification fragment length polymorphism (AFLP) (Kakouli-Duarte *et al*., 2001), polymerase chain reaction–restriction fragment length polymorphism (PCR–RFLP) (Salazar *et al*., 2002), PCR, sequencing (Douglas & Haymer, 2001) and random amplified polymorphic DNA (RAPD) (Fernandez *et al*., 2001). However, the literature contains no information on substitutions in tRNA gene sequences nor on any amino acid sequence of tephritids that might be used to identify the different taxa. The characteristic substitutions observed in the present study could therefore act as a new source of taxonomic information.

Phylogenetic relationships

According to previous classifications, two large clusters corresponding to the subfamilies Dacinae and Trypetinae should have been expected. The genera belonging to the

https://doi.org/10.1079/BER2006425 Published online by Cambridge University Press

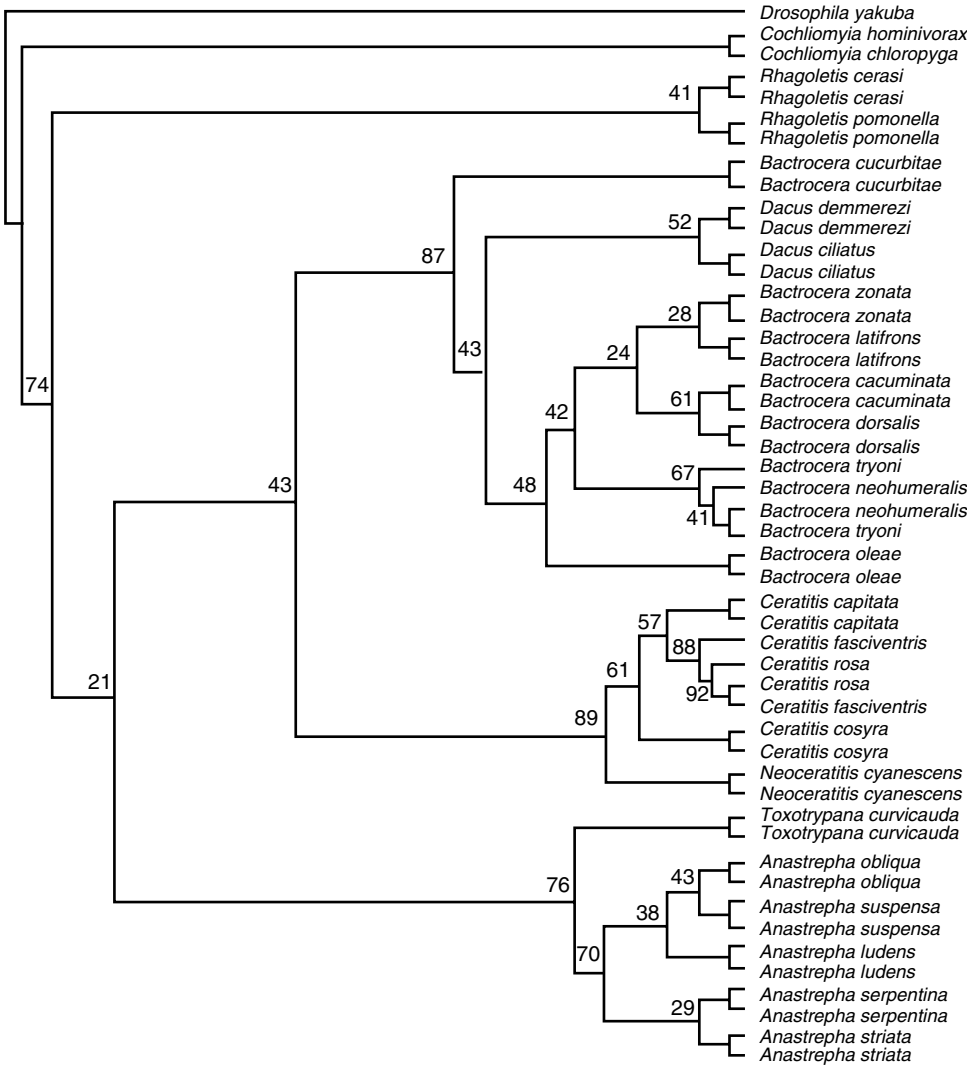

Fig. 3. Majority rule consensus tree based on maximum likelihood trees. Bootstrap values for 1000 replicates are given at the nodes.

subfamily Dacinae (*Dacus*, *Bactrocera*, *Ceratitis* and *Neoceratitis*) were indeed grouped together. However, the assemblage formed by the three species belonging to the subfamily Trypetinae was only present in the neighbour joining tree (NJ, fig. 2), and even then with low bootstrap support. The genus *Rhagoletis* was not grouped with the other two genera of the subfamily studied (*Anastrepha* and *Toxotrypana*), either in the maximum parsimony tree (MP, fig. 1), or in the maximum likelihood tree (ML, fig. 3). It is generally accepted that the subfamily Trypetinae may be a polyphyletic or paraphyletic cluster that remains after the exclusion of members of the other two subfamilies of Tephritidae, i.e. Dacinae and Tephritinae (Korneyev, 2000). The present results are congruent with the supposed non-monophyletic character of this subfamily.

Thus, two groups were considered to exist within the subfamily Trypetinae, namely the tribes Carpomyini (genus *Rhagoletis*) and Toxotrypanini (genera *Anastrepha* and *Toxotrypana*). In the tribe Carpomyini, the two species of the genus *Rhagoletis* were grouped together in all three trees.

The monophyletic nature of the tribe Toxotrypanini and the close relationship between the genera *Anastrepha* and *Toxotrypana* were clear, with high bootstrap support in the NJ and ML trees (although not so obvious in the MP tree). This relationship is also supported by two substitutions in the amino acid sequence of ND1 (table 2) that affected only the *Anastrepha* and *Toxotrypana* species. These might be considered synapomorphies of these taxa.

McPheron *et al*. (2000) and Norrbom *et al*. (2000) suggested that *Anastrepha* might be paraphyletic (if *Toxotrypana* is not included). The present analysis, however, supports the supposed monophyly of *Anastrepha*; high associated bootstrap values were seen in all trees (76–90). In addition, there were three substitutions, one in Cytb and two in ND1 (table 2), present in this genus that were absent in *Toxotrypana*. These might be considered autopomorphies.

Within *Anastrepha*, 18 species groups have been recognized, including the *fraterculus* group (which embraces the species *A. ludens* (Loew), *A. obliqua* (Macquart) and *A. suspensa* (Loew)), and the *serpentina* and *striata* groups.

https://doi.org/10.1079/BER2006425 Published online by Cambridge University Press

There is good morphological evidence (hypothesized synapomorphies) for these three groups (Norrbom *et al*., 2000). The results of the analysis of *rRNA 16S* (McPheron *et al*., 2000) also suggests the monophyletic nature of the *fraterculus* group of species, although the evidence is not conclusive. In the present study, a cluster with the three species representing this group was seen in all the trees. This assemblage was supported by significant bootstrap values. Also, hypothetical autopomorphies were found in the amino acid sequence of ND1, and these are shared by all the members of the group, thus supporting its monophyletic nature. The other two species of the genus *Anastrepha*, i.e. *A. serpentina* (Wiedemann) and *A. striata* (Schiner), were grouped together in all the trees, even though the bootstrap support was somewhat weak. They also shared a possible synapomorphy in ND1. Norrbom *et al*. (2000) suggested the existence of a quite close relationship between the *serpentina* and *striata* species groups. The present results also point in this direction.

The other subfamily in the present sample was the subfamily Dacinae. This is very probably monophyletic, and its early separation is supported by the presence of several plesiomorphies absent in the other subfamilies (Korneyev, 2000). The topology of all the trees is congruent with the monophyly of the Dacinae, even though the bootstrap values are not too high. However, there is a widely held notion supported by simulation tests (Kim, 1993; Avise, 1994) that agreement among trees estimated by different methods lends greater credibility to phylogenetic estimates. Different studies based both on morphological (Foote *et al*., 1993; Korneyev, 2000) and molecular data (Han & McPheron, 1994, 1997, 2000) agree, as does the present analysis, with the supposed monophyletic nature of this subfamily. Similarly, the monophyly of the tribes Dacini and Ceratitini, established by morphological studies, can be inferred from all the trees. The monophyly of these groups has high bootstrap support in all three. In addition, in the amino acid sequence of ND1 (table 3) there were substitutions unique for each of these tribes (two for Ceratitidini and three for Dacini). These might be considered autopomorphies and thus support a monophyletic nature for the Dacinae.

Tribe Dacini includes four genera, but most of the species belong to *Bactrocera* and *Dacus*. The genus *Dacus* was represented in the present study by two species, *Dacus (Didacus) ciliatus* and *Dacus (Dacus) demmerezi*. These two species form an assemblage with quite high bootstrap support in all trees.

Within the genus *Bactrocera* there are 30 recognized subgenera grouped into two major groups of subgenera, *Zeugodacus* and *Bactrocera* (Drew & Hancock, 2000). The monophyly of these subgenera and groups of subgenera is quite questionable (White, 2000). Accordingly, the species in the present study belong to three subgenera; subgenus *Zeugodacus*, which is part of the *Zeugodacus* group of subgenera (*B. cucurbitae*), and the subgenera *Daculus* (*B. oleae* (Rossi)) and *Bactrocera* (*B. cacuminata*, *B. dorsalis*, *B. latifrons* (Hendel), *B. neohumeralis*, *B. tryoni* and *B. zonata* (Saunders)), which are part of the *Bactrocera* group of subgenera. There is also morphological and behavioural evidence to support the closer relationship of the *Zeugodacus* group of subgenera to *Dacus*, than to the *Bactrocera* group of subgenera (I.M. White, personal communication). The *Bactrocera* group is considered monophyletic by Drew & Hancock (2000) and White (2000), but Muraji & Nakahara (2001) and Smith *et al*. (2003) call this into question. In the present study, all the species belonging to this group formed a cluster in all the trees. Though bootstrap support was not too high in the MP and ML trees, the coincidence of these lends confidence to the idea that the *Bactrocera* group is monophyletic. The monophyly of the subgenus *Bactrocera*, supported by the study of Jammnongluk *et al*. (2003a), was not clear in the present study nor that performed by Muraji & Nakahara (2001). What was clear in the present study was the sister group relationship between the subgenera *Bactrocera* and *Daculus*, as suggested by Smith *et al*. (2003).

The division of the *Bactrocera* and *Dacus* genera was not clear in any of the trees. *Bactrocera cucurbitae* split from the other species of the genus *Bactrocera*, laying nearer to the genus *Dacus*. As reported by White & Elson-Harris (1992), *Zeugodacus* species have a pattern of host relationships that differentiate them from the other *Bactrocera*. Many *Dacus* spp. and species of the *Zeugodacus* group are associated with Cucurbitaceae, a host relationship almost unknown in the *Bactrocera* group of subgenera. In addition, species attracted to male lures in *Dacus* and the *Zeugodacus* group are normally attracted to cue lure (there are just two apparent exceptions, both unconfirmed); whereas an attraction to methyl eugenol is common in the *Bactrocera* group. This separation from the other *Bactrocera* species has been observed in other studies (Malacrida *et al*., 1996).

The analysis of this region of mtDNA did not distinguish between the species *B. tryoni* and *B. neohumeralis*, both members of the *dorsalis* species complex. Using four different nuclear and mitochondrial markers, Morrow *et al*. (2000) also observed that the genetic difference between these two species was very small. This could be due to recent speciation or to rapid separation.

Finally, *B. cacuminata* and *B. dorsalis*, from the *dorsalis* species complex (White & Elson-Harris, 1992), formed an assemblage with high bootstrap support in all trees. A possible synapomorphy in the ND1 sequence confirms the close relationship between them.

The tribe Ceratitidini was represented in the present sample by two genera, *Ceratitis* and *Neoceratitis*. The monophyletic character of the subfamily and the close relationship between both genera was inferred from all trees, and had high bootstrap support. The genus *Ceratitis* also formed a monophyletic assemblage in all the trees, with high bootstrap values in the MP and ML trees. Here again, a substitution was seen in the Cytb sequence unique for the members of these taxa, thus corroborating the monophyletic nature of the group.

The genus *Ceratitis*, as it stands today, is a composite of six subgenera. The exact boundaries of these subgenera, and decisions on which species belong to which subgenus, are subject to continuous change (De Meyer, 2000, 2001a). The species in the present sample were classified in the subgenera *Ceratitis* (*C. capitata*), *Pterandrus* (*C. rosa* and *C. fasciventris*) and *Ceratalapsis* (*C. cosyra*). In all the trees produced, a cluster with strong bootstrap support was formed by the species of the subgenera *Pterandrus* and *Ceratitis*. This indicates the close relationship between them. Finally, the species *C. rosa* and *C. fasciventris* could only be distinguished in the NJ tree. Until recently the latter was considered a variety of *C. rosa*, but the most recent studies (De Meyer, 2001b; Baliraine *et al*., 2003, 2004) consider it a different species.

In summary, despite the number of studies that have been performed on this species-rich and economically

https://doi.org/10.1079/BER2006425 Published online by Cambridge University Press

important family, some controversy in the classification of the family Tephritidae still exists. This is probably due to the large size of the group, the fact that the taxonomic studies undertaken have largely been regional, and that most have been based on morphological traits that intergrade between higher taxonomic groups. In the present study, 23 species of tephritid fruit flies were included in the phylogenetic analysis, but more must be analysed and data from longer sequences are needed. Nonetheless, the results show that this region of mtDNA can help resolve the relationships of the Tephritidae at various taxonomic levels and that it will allow the evolutionary history of the family to be explored. Molecular data allows us to approach phylogenetic questions from a new direction, and perhaps provide a tool for testing existing classifications. A more exhaustive knowledge of the phylogeny of this family should throw light on the evolution of traits related to the pest status of this group and may provide information useful in the development of control programmes.

Acknowledgements

This work was funded by the EC Project FAIR3 CT96-1972 'Development of standardized molecular techniques for the identification of insect quarantine pests'. The authors are very grateful to Dr C. Fleming for the primers, and to Drs A. Alemany, C. Fleming, J. Frey, A. García-Ortiz, B. Gikaria, T. Grové, A. Handler, D.S. Haymer, E. Hernández-Ortiz, E. Jang, F. Jerónimo, P. Lawrence, J.M. Llorens, S. Lux, D. Moreno, N. Papadopoulos, S. Quilici, M. Robson, T. Rodríguez, M. Sala, S. Seewooruthun and N. Zenz for the biological samples. The authors also wish to thank to the referees for their valuable comments.

References

Avise, J.C. (1994) *Molecular markers, natural history and evolution*. New York, Chapman and Hall.

Baliraine, F.N., Bonizzoni, M., Osir, E.O., Lux, S.A., Mulaa, F.J., Zheng, L., Gomulski, L.M., Gasperi, G. & Malacrida, A.R. (2003) Comparative analysis of microsatellite loci in four fruit fly species of the genus *Ceratitis* (Diptera: Tephritidae). *Bulletin of Entomological Research* **93**, 1–10.

Baliraine, F.N., Bonizzoni, M., Guglielmino, C.R., Osir, E.O., Lux, S.A., Mulaa, F.J., Gomulski, L.M., Zheng, L., Quilici, S., Gasperi, G. & Malacrida, A.R. (2004) Population genetics of the invasive African fruit fly species, *Ceratitis rosa* and *Ceratitis fasciventris* (Diptera: Tephritidae). *Molecular Ecology* **13**, 683–695.

Clary, D.O. & Wolstenholme, D.R. (1985) The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization and genetic code. *Journal of Molecular Evolution* **22**, 252–271.

Clary, D.O., Wahleitner, J.A. & Wolstenholme, D.R. (1983) Transfer RNA genes in *Drosophila* mitochondrial DNA: related 5′ flanking sequences and comparisons to mammalian mitochondrial tRNA genes. *Nucleic Acids Research* **11**, 2411–2455.

De Meyer, M. (2000) Phylogeny of the genus *Ceratitis* (Dacinae: Ceratitidini). pp. 409–428 *in* Aluja, M. & Norrbom, A.L. (*Eds*) *Fruit flies (Tephritidae): phylogeny and evolution of behaviour*. Boca Raton, Florida, CRC Press.

De Meyer, M. (2001a) Distribution patterns and host-plant relationships within the genus *Ceratitis* MacLeay (Diptera: Tephritidae) in Africa. *Cimbebasia* **17**, 219–228.

De Meyer, M. (2001b) On the identity of the Natal fruit fly, *Ceratitis rosa* (Diptera: Tephritidae). *Bulletin d l'Institut Royal des Sciences Naturelles de Belgique, Entomologie* **71**, 55–62.

Douglas, L.J. & Haymer, D.S. (2001) Ribosomal ITS1 polymorphisms in *Ceratitis capitata* and *Ceratitis rosa* (Diptera: Tephritidae). *Annals of the Entomological Society of America* **94**, 726–731.

Drew, R.A.I. & Hancock, D.L. (2000) Phylogeny of the tribe Dacini (Dacinae) based on morphological, distributional and biological data. pp. 491–504 *in* Aluja, M. & Norrbom, A.L. (*Eds*) *Fruit flies (Tephritidae): phylogeny and evolution of behaviour*. Boca Raton, Florida, CRC Press.

Felsenstein, J. (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* **17**, 368–376.

Felsenstein, J. (2001) *PHYLIP (Phylogeny Inference Package)*, version 3.6a2. Department of Genetics, University of Washington, Seattle.

Fernández, P., Segura, D. & Ochando, M.D. (2001) Taxonomy and genetic differentiation of *Ceratitis capitata* (Wiedemann, 1824) (Diptera, Tephritidae) and *Bactrocera oleae* (Gmelin, 1790) (Diptera, Tephritidae) using molecular markers (RAPD–PCR). *Boletin de la Real Sociedad Española de Historia Natural (Sec. Biol.)* **96**, 321–329 (in Spanish, English abstract).

Fitch, W.M. (1971) Distinguishing homologous from analogous proteins. *Systematic Zoology* **20**, 406–416.

Foote, R.H., Blanc, F.L. & Norrbom, A.L. (1993) *Handbook of the fruit flies (Diptera: Tephritidae) of America North of Mexico*. Ithaca, New York, Comstock Publishing Associates.

Han, H.Y. (2000) Molecular phylogenetic study of the tribe Trypetini (Diptera: Tephritidae), using mitochondrial 16S ribosomal DNA sequences. *Biochemical Systematics and Ecology* **28**, 501–513.

Han, H.Y. & McPheron, B.A. (1994) Phylogenetic study of selected tephritid flies (Insecta: Diptera: Tephritidae) using partial sequences of the nuclear 18S ribosomal DNA. *Biochemical Systematics and Ecology* **22**, 447–457.

Han, H.Y. & McPheron, B.A. (1997) Molecular phylogenetic study of Tephritidae (Insecta: Diptera) using partial sequences of the mitochondrial 16S ribosomal DNA. *Molecular Phylogenetics and Evolution* **7**, 17–32.

Han, H.Y. & McPheron, B.A. (2000) Nucleotide sequence data as a tool to test phylogenetic relationships among higher groups of Tephritidae: a case of study using mitochondrial ribosomal DNA. pp. 115–132 *in* Aluja, M. & Norrbom, A.L. (*Eds*) *Fruit flies (Tephritidae): phylogeny and evolution of behaviour*. Boca Raton, Florida, CRC Press.

Jamnongluk, W., Baimai, V. & Kittayapong, P. (2003a) Molecular evolution of tephritid fruit flies in the genus *Bactrocera* based on the cytochrome oxidase I gene. *Genetica* **119**, 19–25.

Jamnongluk, W., Baimai, V. & Kittayapong, P. (2003b) Molecular phylogeny of tephritid fruit flies in the *Bactrocera tau* complex using the mitochodrial COI sequences. *Genome* **46**, 112–118.

Kakouli-Duarte, T., Casey, D.G. & Burnell, A.M. (2001) Development of a diagnostic DNA probe for the fruit flies *Ceratitis capitata* and *Ceratitis rosa*. *Journal of Economic Entomology* **94**, 989–997.

https://doi.org/10.1079/BER2006425 Published online by Cambridge University Press

Kim, J. (1993) Improving the accuracy of phylogenetic estimation by combining different methods. *Systematic Biology* **42**, 331–340.

Korneyev, V.A. (2000) Phylogenetic relationships among higher groups of Tephritidae. pp. 73–113 *in* Aluja, M. & Norrbom, A.L. (*Eds*) *Fruit flies (Tephritidae): phylogeny and evolution of behaviour*. Boca Raton, Florida, CRC Press.

Kumar, S., Tamura, K., Jacobsen, I.B. & Nei, M. (2001) *MEGA (Molecular Evolutionary Genetics Analysis*), versión 2.1. Arizona State University, Tempe.

Malacrida, A.R., Guglielmino, C.R., D'Adamo, P., Torti, C., Marinoni, F. & Gasperi, G. (1996) Allozyme divergence and phylogenetic relationships among species of tephritid flies. *Heredity* **76**, 592–602.

McPheron, B.A. & Han, H.Y. (1997) Phylogenetic analysis of North American *Rhagoletis* (Diptera: Tephritidae) and related genera using mitochondrial DNA sequence data. *Molecular Phylogenetics and Evolution* **7**, 1–16.

McPheron, B.A., Han, H.Y., Silva, J.G. & Norrbom, A.L. (2000) Phylogeny of the genera *Anastrepha* and *Toxotrypana* (Trypetinae: Toxotrypanini) based upon 16S rRNA mitochondrial DNA sequences. pp. 343–361 *in* Aluja, M. & Norrbom, A.L. (*Eds*) *Fruit flies (Tephritidae): phylogeny and evolution of behaviour*. Boca Raton, Florida, CRC Press.

Morrow, J., Scott, L., Congdon, B., Yeates, D., Fromer, M. & Sved, J. (2000) Close genetic similarity between two sympatric species of tephritid fruit flies reproductively isolated by mating time. *Evolution: International Journal of Organic Evolution* **54**, 899–910.

Muraji, M. & Nakahara, S. (2001) Phylogenetic relationships among fruit flies, *Bactrocera* (Diptera, Tephritidae), based on the mitochondrial rDNA sequences. *Insect Molecular Biology* **10**, 549–559.

Norrbom, A.L., Zucchi, R.A. & Hernández-Ortiz, V. (2000) Phylogeny of the genera *Anastrepha* and *Toxotrypana* based on morphology. pp. 299–342 *in* Aluja, M. & Norrbom, A.L. (*Eds*) *Fruit flies (Tephritidae): phylogeny and evolution of behaviour*. Boca Raton, Florida, CRC Press.

Reyes, A., Linacero, R. & Ochando, M.D. (1997) Molecular genetics and integrated control: a universal genomic DNA integrated method for PCR, RAPD, restriction and Southern analysis. *IOBC wprs Bulletin* **20**, 274–284.

Saitou, N. & Nei, M. (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406–425.

Salazar, M., Theoduloz, C., Vega, A., Pobrete, F., González, E., Badilla, R.N. & Meza-Basso, L. (2002) PCR–RFLP identification of endemic Chilean species of *Rhagoletis* (Diptera: Tephritidae) attacking Solanaceae. *Bulletin of Entomological Research* **92**, 337–341.

Simon, C., Frati, F., Beckenback, A., Crespi, B., Liu, H. & Flookal, P. (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* **87**, 651–701.

Smith, P.T., Kambhampati, S. & Armstrong, K.A. (2003) Phylogenetic relationships among *Bactrocera* species (Diptera: Tephritidae) inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution* **26**, 8–17.

Swofford, D.L. (2001) *PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods*), version 4.0b8. Sunderland, Massachusetts, Sinauer.

Swofford, D.L., Olsen, G.J., Waddell, P.J. & Hillis, D.M. (1996) Phylogenetic inference. pp. 407–514 *in* Hillis, D.M. & Moritz, C. (*Eds*) *Molecular systematics*. Sunderland, Massachusetts, Sinauer Associates.

Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673–4680.

White, I.M. (2000) Morphological features of the tribe Dacini (Dacinae): their significance to behavior and classification. pp. 505–533 *in* Aluja, M. & Norrbom, A.L. (*Eds*) *Fruit flies (Tephritidae): phylogeny and evolution of behaviour*. Boca Raton, Florida, CRC Press.

White, I.M. & Elson-Harris, M. (1992) *Fruit flies of economic significance; their identification and bionomics*. Wallingford, Oxon, CAB International.

Wolstenholme, D.R. & Clary, D.O. (1985) Sequence evolution of *Drosophila* mitochondrial DNA. *Genetics* **109**, 725–744.

(Accepted 17 January 2006)
© CAB International, 2006

https://doi.org/10.1079/BER2006425 Published online by Cambridge University Press