Polymorphism within and between populations of *Ceratitis capitata*: comparison between RAPD and multilocus enzyme electrophoresis data

L. BARUFFI, G. DAMIANI†, C. R. GUGLIELMINO‡, C. BANDI§, A. R. MALACRIDA & G. GASPERI*

Dipartimento di Biologia Animale, Laboratorio di Zoologia, Università di Pavia, Piazza Botta 9, I-27100 Pavia, †Istituto per la Difesa e la Valorizzazione del Germoplasma Animale, CNR, via Celoria 10, I-20133 Milano, ‡Dipartimento di Genetica e Microbiologia, Università di Pavia, Istituto di Genetica e Biologia Evoluzionistica, CNR, via Abbiategrasso 207, I-27100 Pavia and §Istituto di Patologia Generale Veterinaria, Università di Milano, via Celoria 10, I-20133 Milano, Italy

Random amplified polymorphic DNA (RAPD) analysis and multilocus enzyme electrophoresis (MLEE) were used to assess genetic variability in six wild populations and in five laboratory strains of *Ceratitis capitata*. The RAPD technique reveals larger amounts of genetic variation than the conventional MLEE, and can improve discrimination within and between populations and strains. In our experimental conditions, RAPD analysis with four different primers produces 174 polymorphic bands out of 176, while MLEE analysis at 26 enzyme loci scores 74 alleles. RAPD fingerprints are peculiar to African flies, while different laboratory strains have similar patterns, independently of their origins. The results obtained by these two methods are significantly correlated, and are in agreement with the general trend of decreasing variability from African populations towards the peripheral and laboratory ones. UPGMA dendrograms derived from MLEE (protein) and RAPD (DNA) data show that a major part of intraspecific variability involves the differentiation of central vs. peripheral populations.

Keywords: *Ceratitis capitata* (medfly), DNA fingerprinting, laboratory strains, multilocus enzyme electrophoresis, RAPD, wild populations.

Introduction

The economically important species *Ceratitis capitata* (medfly) is a polyphagous and multivoltine dipteran species which has recently spread from its supposed origin in sub-Saharan Africa to tropical and subtropical areas througout the world (Fletcher, 1989). The colonization history of this species has been documented and many aspects of its biology and ecology have been investigated (for a comprehensive review see Robinson & Hooper, 1989).

Multilocus enzyme electrophoresis (MLEE) has already been applied to *C. capitata*, providing a reliable tool for assessing the genetic makeup of populations (Milani *et al.*, 1989), to construct genetic maps of this species (Malacrida *et al.*, 1990) and to infer taxonomic

relationships between the medfly and several other tephritid species (Malacrida *et al.*, 1991). Allozymes have proved the existence of nonrandom distributions of genetic variability in geographical populations of *C. capitata*: spatial and temporal differentiations characterize this species after its spread from the putative source area to the periphery of its present geographical range (Gasperi *et al.*, 1991; Malacrida *et al.*, 1992). With respect to the colonization patterns, it has been suggested that the medfly populations, like those of *Drosophila melanogaster* (David & Capy, 1988), can be divided into three main categories: ancestral, ancient and new populations from sub-Saharan Africa, the Mediterranean basin, and the New World, respectively (Malacrida *et al.*, 1992).

DNA variability has been used in the study of populations of *C. capitata* to discriminate between medfly populations using mitochondrial DNA (mtDNA) variation (Frohlich *et al.*, 1993; McPheron *et al.*,

*Correspondence.

1994). Recently, techniques based on the polymerase chain reaction (PCR) using random primers have provided a powerful tool for producing a large amount of anonymous genetic markers called RAPD (random amplified polymorphic DNA) (Williams *et al.*, 1990; Welsh & McClelland, 1990). These techniques reveal larger amounts of variation than the conventional MLEE, and may improve discrimination both within and between populations from various species (Hadrys *et al.*, 1992; Hedrick, 1992; Black IV, 1993). RAPD analysis has already been applied to *C. capitata*, proving that this approach rapidly generates a large number of markers for genome mapping and for population genetics (Gasperi *et al.*, 1993; Haymer & McInnis, 1994).

In this paper, we compare genetic polymorphisms detected by RAPD and MLEE in several wild samples and laboratory strains of *C. capitata*. The aim of this work is to confirm and to extend the previously established knowledge on the genetic structure and differentiation of medfly populations (Gasperi *et al.*, 1991; Malacrida *et al.*, 1992). In *C. capitata*, RAPDs reveal a surprisingly high level of hidden genetic variability mainly in wild African populations, which decreases in the peripheral ones to levels similar to those observed in some laboratory colonies. These data are generally consistent with the results obtained by MLEE, although they can reveal different features of differentiation between medfly populations.

Materials and methods

Populations of C. capitata

Six natural different geographical populations (Kenya 1; Kenya 2; Morocco; Italy; Reunion; Australia) and five laboratory strains (Ispra; Israel; Metapa; M84; Hawaii) having different colonization histories were analysed. The wild samples were collected in 1992 as pupae and adult individuals were assayed after emergence. Kenya 1 and Kenya 2 samples were collected on coffee berries in two localities near Nairobi, Kenya, at Ruiru and Kabete, respectively. The Moroccan flies derive from the *Argania spinosa* forest on the western coast of that country. The Italian sample was found on figs near Milan, North Italy, in September 1992. The samples from Reunion and Australia (Perth) were collected from different fruits.

Hawaii, Israel and Ispra are laboratory strains, 35, 20 and 25 years old, respectively. They have been maintained in our laboratory since 1992, 1990 and 1979, respectively. The first two strains were established from flies collected in the corresponding countries whereas the Ispra strain (C.E.C. Joint

Research Centre, Ispra, Italy) was derived from flies from Sicily and Greece, and is regarded as a standard laboratory strain for our allozyme surveys. Metapa was obtained in 1990 from the MOSCAMED Biofactory, Tapachula, Chiapas, Mexico. M84 is a laboratory line derived from a single family produced by a backcross between the laboratory strains South Africa and the multimarked *dp*; *bwb*; *ye²*. This line was established to maintain the polymorphism at the following morphological and biochemical loci: *dp* (dark pupae), Got_2 (3rd chr.); *bwb* (brown body), *Zw*, *Fh* (5th chr.); *ye²* (yellow eye 2), Got_1 (6th chr.). All the laboratory strains are maintained at the Laboratory of Zoology, Department of Animal Biology, University of Pavia.

Multilocus enzyme electrophoresis (MLEE)

Single fly homogenization and electrophoretic procedures were performed according to a previous survey of medfly populations (Gasperi *et al.*, 1991). At least 30 individuals from each population sample were electrophoretically assayed using the following 26 enzyme loci: Hk_1 and Hk_2 (E.C. 2.7.1.1 hexokinase); *Pgm* (E.C. 2.7.5.1 phosphoglucomutase); *Pgi* (E.C. 5.3.1.9 phosphoglucose isomerase); *Pgk* (E.C. 2.7.2.3 phosphoglycerate kinase); *Zw* (E.C. 1.1.1.49 glucose 6-phosphate dehydrogenase); *Pgd* (E.C. 1.1.1.44 6-phosphogluconate dehydrogenase); *a-Gpdh* (E.C. 1.1.1.8 *a*-glycerophosphate dehydrogenase); *Idh* (E.C. 1.1.1.42 isocitrate dehydrogenase); *Fh* (E.C. 4.2.1.2 fumarate hydratase); Mdh_1 and Mdh_2 (E.C. 1.1.1.37 malate dehydrogenase); Ak_1 and Ak_2 (E.C. 2.7.4.3 adenilate kinase); Adh_2 (E.C. 1.1.1.1 alcohol dehydrogenase); *Aox* (E.C. 1.2.3.1 aldehyde oxidase); Est_1 and Est_2 (E.C. 3.1.1.– esterases); Got_1 and Got_2 (E.C. 2.6.1.1 glutamate oxaloacetate transaminase); *Gpt* (E.C. 2.6.1.2 glutamate pyruvate transaminase); *Had* (E.C. 1.1.1.35 hydroxyacid dehydrogenase); *Me* (E.C. 1.1.1.40 malic enzyme); *Mpi* (E.C. 5.3.1.8 mannose phosphate isomerase), $Acon_1$ and $Acon_2$ (E.C. 4.2.1.3 aconitate hydratase).

Random amplified polymorphic DNA (RAPD) analysis

DNA preparation from single flies was performed as described by Bender *et al.* (1983) with modifications. Single flies were ground in Eppendorf tubes, each containing 200 μL of 100 mM NaCl, 200 mM sucrose, 100 mM Tris-HCl (pH 9.1), 50 mM EDTA, 0.5 per cent SDS, and proteinase K to a final concentration of 100 μg/mL. After adding another 200 μL of grinding buffer, the homogenate was incubated at 65°C for 30 min. Then 8 M potassium acetate was added to a final concentration of 1 M, and the mixture was kept at 0°C

for 30 min. After centrifugation for 15 min at 15 800 g, the recovered supernatant was mixed with two volumes of absolute ethanol, left at room temperature for 5 min, and centrifuged again for 15 min at 15 800 g. The resulting precipitate was washed in 70 per cent ethanol and in absolute ethanol, dried and resuspended in 25 μL of TE (10 mM Tris-HCl, pH 8, 1 mM EDTA).

The DNA from five individuals of each population sample was amplified using the following oligonucleotide primers.

494N: 5'-AGCGCTGTGAGAAAGATGAAAGAT
D17: 5'-CGAAACAAGCGTGCATGAGCCCGAA
NP1: 5'-CGAAACAAGCGCTCATGAGCCCGAA
NP4: 5'-CTAATGCAGGAGTCGCATAAGGGAGA

Amplification reactions were performed in a 10 μL volume, containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 5 mM $MgCl_2$, 0.1 per cent (v/v) Triton X-100, 200 μM of each dNTP, 250 ng of each primer, 10–100 ng of DNA and 0.5 unit of *Taq* polymerase (Promega). The reaction mixtures were prepared at 0°C, primers were denatured and immediately chilled in wet ice before addition to the reaction mix. The tubes were than quickly transferred to the thermal cycler (Programmable Thermal Controller PTC100, MJ Research), preheated at 90°C. After incubation at 94°C for 90 s, reactions were cycled 45 times through the following temperature profile: 94°C for 30 s, 45°C for 1 min and 75°C for 2 min. The samples were finally incubated at 75°C for 10 min, at 60°C for 10 min and stored at − 20°C.

Aliquots of 5μL of each amplification were loaded onto a 3 per cent (w/v) agarose gel (1.5 per cent SeaKem GTG-FMC plus 1.5 per cent NuSieve GTG-FMC) with TBE buffer containing 0.5 mg/mL (w/v) of ethidium bromide (Sambrook *et al.*, 1989), and electrophoresis was carried out at 5 V/cm for 3 h.

Data analysis

Allozyme frequencies, mean heterozygosities (H), genetic distances (D) and *F*-statistics were computed on allozyme data, as described by Nei (1972) and Nei & Chesser (1983), using the BIOSYS-1 program of Swofford & Selander (1981).

The numerical taxonomy methods (Sneath & Sokal, 1973) and the NTSYS (Numerical Taxonomy System) program by James Rohlf (Applied Biostatistics, Setauket, New York) were used for analysing the PCR amplified polymorphisms. The dissimilarity index (d) was calculated by comparing banding patterns between pairs of individuals, according to the following formula:

$$d = N_{AB}/(N_A + N_B),$$

where N_A and N_B are the number of fragments in individuals A and B, respectively, and N_{AB} is the number of fragments that differ between the two individuals (Gilbert *et al.*, 1990; Yuhki & O'Brien, 1990). If we compare n individuals from the same population and average the $[n_1 \times (n_1 - 1)/2]$ dissimilarity values, we obtain an estimate of within-population dissimilarity; on the other hand, the dissimilarity value between different populations is given by the average of the $[n_1 \times n_2]$ dissimilarity values between the n_1 and n_2 individuals of the two populations. The computations are based on five individuals per population. In fact, running DNA samples from five individuals per population on large gels reduced band interpretation errors, often encountered in scoring RAPD band matches from individuals of different populations on separate gels. Furthermore, it is generally accepted that a high number of genetic markers may counterbalance the relatively low number of individuals and may therefore reduce the sampling variation of the dissimilarity estimates.

Trees showing relationships between individuals or populations were generated by the UPGMA method of cluster analysis on the values of the pairwise dissimilarity matrix.

To estimate the congruence between the two sets of data (MLEE and RAPD), the correlation between the distance matrices of allozymes and the dissimilarity matrices of RAPD was calculated with the nonparametric permutation test of Mantel (1967).

Results

MLEE analysis of C. capitata populations

Out of the 26 enzyme loci listed in Materials and methods, 22 were found to be polymorphic over the examined wild populations and laboratory strains of *C. capitata*. The four loci Ak_1, Adh_2, α-*Gpdh* and *Me* were uniformly monomorphic, and the total number of alleles scored by MLEE in the 11 population samples was 74. Allele frequencies generally follow Hardy–Weinberg expectation within each population. The amount of allozyme variability within each population is summarized in Table 1. In the wild population samples, statistics such as average number of alleles per locus (A), proportion of polymorphic loci (P) and average number of heterozygous individuals (H) seem to be related to the evolutionary history of the populations, given the general trend of decreasing genetic variation from the putative African source area (Kenya) toward the periphery of the species range, as already assessed in a previous survey (Gasperi *et al.*, 1991). The estimates observed for the wild peripheral

Table 1 Measures of genetic variation, at 26 biochemical loci, of six wild populations and five laboratory strains of *Ceratitis capitata*

Medfly samples	Mean no. of alleles per locus (A)	Proportion of polymorphic loci[a] (P)	Mean heterozygosity (H)
Wild populations			
Kenya 1	2.1 ($\pm$0.3)	61.5	0.154 ($\pm$0.043)
Kenya 2	2.3 ($\pm$0.3)	65.4	0.140 ($\pm$0.037)
Reunion	1.5 ($\pm$0.2)	30.8	0.091 ($\pm$0.039)
Morocco	1.5 ($\pm$0.2)	38.5	0.102 ($\pm$0.036)
Italy	1.2 ($\pm$0.1)	19.2	0.046 ($\pm$0.023)
Australia	1.1 ($\pm$0.1)	11.5	0.030 ($\pm$0.018)
Laboratory strains			
Hawaii	1.3 ($\pm$0.1)	19.2	0.062 ($\pm$0.028)
Israel	1.2 ($\pm$0.1)	19.2	0.050 ($\pm$0.023)
Ispra	1.2 ($\pm$0.1)	19.2	0.036 ($\pm$0.022)
Metapa	1.2 ($\pm$0.1)	11.5	0.037 ($\pm$0.021)
M84	1.3 ($\pm$0.1)	30.8	0.117 ($\pm$0.039)

Values in parentheses are standard errors.
[a]A locus is considered polymorphic if more than one allele was detected.

populations (Italy, Australia) are in the same range as those detected in the laboratory strains. Among the latter, M84 shows the highest values for variability, as expected, because of the artificial selection for polymorphism (see Materials and methods).

RAPD analysis of C. capitata *populations*

Genomic DNAs of individuals randomly sampled from the same populations of *C. capitata* already analysed by MLEE, were amplified by PCR using single arbitrary primers. RAPD patterns in individual flies show about 10–20 main bands, ranging from 50 to 2000 bp in size. The same patterns were obtained in independent experiments with the amount of template DNA varying between 500 and 50 ng. The reproducibility of these RAPD patterns is shown in Fig. 1. Our experimental conditions for generating RAPD markers are different from those reported in the standard protocols (Williams *et al.*, 1990; Welsh & McClelland, 1990). The reproducibility of this modified method has been already tested in other organisms (Bandi *et al.*, 1993). The Mendelian segregation of the DNA fragments as dominant markers was verified by single-pair matings and the attribution of several RAPD markers to the linkage groups of *C. capitata*, already constructed with biochemical and morphological loci (Malacrida *et al.*, 1990), are in a manuscript in preparation. Variation within and between populations is regarded as presence/absence of a given band, because, owing to

the dominance of RAPD markers, the products from a homozygous (presence/presence) cannot be distinguished from the heterozygous condition (presence/absence). Figure 2 shows an example of fragments amplified by a single primer (NP1) from five individual flies of each of the 11 populations analysed. Extremely variable banding patterns characterize the wild African flies, such as those from Kenya and Reunion, compared with those obtained from Morocco, Italy and Australia. These last patterns roughly resemble those obtained from the laboratory strains (Hawaii, Israel, Ispra, Metapa, M84) which generally show similar RAPD banding, irrespective of their origins and colonization histories. A similar trend is produced by the other primers used (494N, D17, NP4). We can exclude the possibility that the similarity of banding patterns in laboratory strains results from contamination, given that certain biochemical and morphological markers allow us to recognize many of these strains. Histograms showing the number of fragments generated by PCR in five individuals, and the trends mentioned above are given in Fig. 3 for each primer. Each primer generates a variable number of bands, depending on the population surveyed. Generally, the proportion of monomorphic bands exceeds that of polymorphic ones in the laboratory strains as well as in the wild peripheral populations (Italy and Australia). The African flies are clearly more polymorphic than the peripheral and laboratory populations. Out of the four primers used, NP4 appears to be
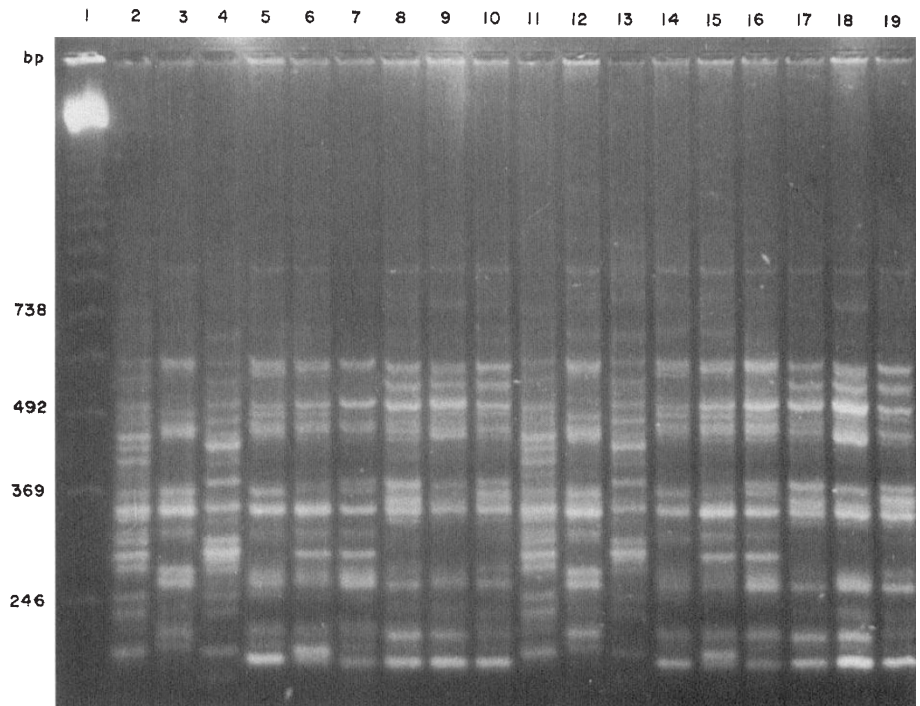
Fig. 1 Comparison of the genomic DNA amplification products obtained from single flies of different strains of *Ceratitis capitata* in two independent experiments. Dialysed template DNAs were used in lanes 2–10 whereas in lanes 11–19 the same DNAs were not treated. Molecular weight markers (base pairs, bp) are indicated in lane 1. Lanes 2–4 and 11–13: Kenya flies; lanes 5–7 and 14–16: Metapa flies; lanes 8–10 and 17–19: M84 flies. Amplification products were obtained using a NP4 primer of arbitrary nucleotide sequence and were resolved by electrophoresis in a 3 per cent (w/v) agarose gel which was stained with ethidium bromide.

the most effective in describing the cline of decreasing variability in wild populations and the difference in variability between wild and laboratory flies.

RAPD analysis using all four primers on the 11 populations allows the definition of 176 fragments, only two of which have been found to be monomorphic. However, bands reported as monomorphic may hide genetic heterogeneity, because of the dominance of RAPD markers (Williams *et al.*, 1990; our results, data not presented).

RAPD variation within and between populations

Table 2 shows the mean number of fragments per individual, the total number of fragments scored in each population and the dissimilarity index (d) estimates among individual flies in each population. The within-population d values range from 38.7 per cent to 15.8 per cent in wild flies, and from 13.1 per cent to 6.2 per cent in laboratory strains. The high value of d for Kenya and Reunion flies is consistent with the fact that individual fingerprints were observed in the samples from Kenya and Reunion, namely flies share on average a very low proportion of their RAPD bands.

A cluster analysis based on between-individual d values was performed to generate a tree showing the degree of differentiation between individuals of the 11 population samples (Fig. 4). All the flies cluster

together according to their populations. The individual pairwise comparisons indicate the suitability of RAPD analysis for medfly fingerprinting, given that very high values of individual dissimilarity (as in African flies) are associated with a clear-cut population grouping. The wild peripheral flies and the laboratory ones cluster in a single comprehensive group with a sharp distinction between the derived and the laboratory populations.

The between-population dissimilarity matrix (data not presented) shows that d values range from 0.645 (Kenya 2-Reunion) to 0.200 (Italy-Australia) for the wild populations, and from 0.285 (Hawaii-Metapa) to 0.194 (Ispra-Israel) for laboratory strains.

Congruence of RAPD and MLEE data

The degree of polymorphism detected in wild populations of *C. capitata* is shown in Fig. 5, which compares the mean proportion of polymorphic biochemical loci with the mean proportion of polymorphic DNA fragments. A greater level of polymorphism is revealed by RAPD (mean 0.66) compared with that exposed by MLEE (mean 0.38). The mean value of polymorphism detected by PCR is underestimated because of the dominance of RAPD markers. Both plots show a positive deviation from the mean percentage of polymorphism for the populations from Kenya (the supposed source area of the species) in comparison with the negative deviation for the ones derived from
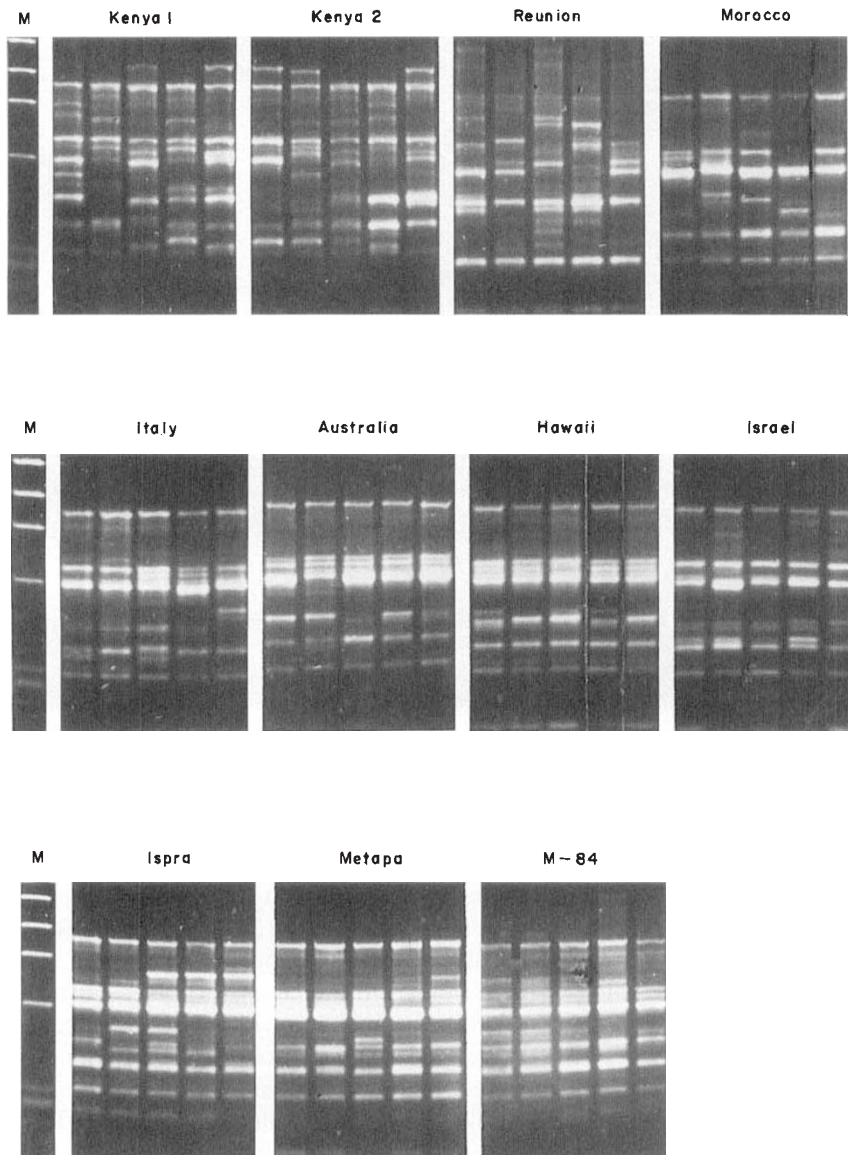
Fig. 2 RAPD patterns of five individuals of *Ceratitis capitata* randomly chosen from six wild populations (Kenya 1, Kenya 2, Reunion, Morocco, Italy, Australia) and from five laboratory strains (Hawaii, Israel, Ispra, Metapa, M84). The fragments were generated and resolved as described in Materials and methods, using NP1 as a single primer.

Italy and Australia (at the periphery of the species range). The deviation values for Morocco approach the mean percentage of polymorphism by the two methods. RAPD clearly highlights more variability than MLEE in the Reunion sample.

The dissimilarity index can be considered as a rough approximator of heterozygosity, based on the genetic model of Gilbert *et al.* (1990) and Lynch (1990) for DNA fingerprint evaluation. The correlation between the heterozygosity (MLEE) and the dissimilarity index (RAPD) over the 11 populations is 0.67, which is significant at the 5 per cent level by the standard test for significance of the correlation coefficient. In Fig. 6 we show the scatter of populations and the regression line of the dissimilarity indices on the heterozygosity

values. Most of the points are fairly close to the line; exceptions are the Reunion sample, which appears to be highly heterogeneous at the DNA level, and the laboratory strain M84, where a high heterozygosity value is expected, based on the polymorphism at biochemical loci in this strain.

Two UPGMA dendrograms showing relationships among the 11 medfly populations were constructed on average dissimilarity values for the RAPD data (Fig. 7a), and on Nei's genetic distances inferred from allozyme frequencies (Fig. 7b). In both the dendrograms, Reunion is the most differentiated, standing quite apart from all the other populations. Among the wild samples, those from Kenya cluster in a group very distinct from the peripheral populations, which group
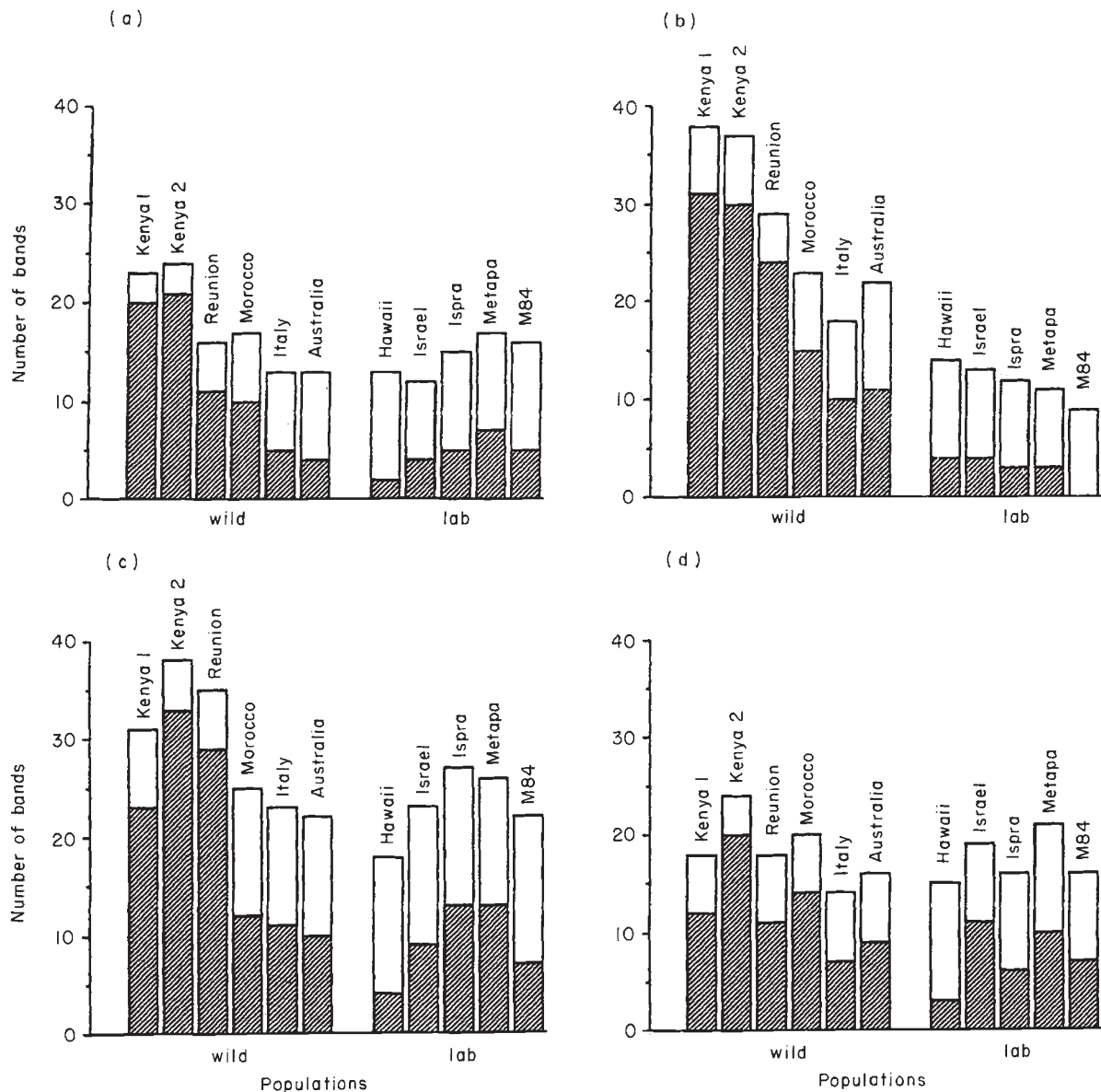
Fig. 3 Frequency distribution of polymorphic (▨) and monomorphic (□) DNA fragments in wild and laboratory populations of *Ceratitis capitata*, generated from each of the four primers: (a) NP1, (b) NP4, (c) 494N and (d) D17.

with the laboratory strains. The congruence between the RAPD and MLEE data has been evaluated by calculating the correlation coefficient, using the Mantel test, between the two sets of distances for the 55 pairwise comparisons among the 11 populations. The value of the correlation index is 0.698, which is statistically significant ($P < 0.05$) according to the Mantel test. However, the MLEE approach highlights the biochemical genetic differentiation of the laboratory strains Hawaii and M84. Table 3 presents the fixation index (F_{ST}) estimates in all the populations assayed, confirming the contribution of single biochemical loci to the differentiation of Reunion, Hawaii and M84. Private alleles at the loci Est_1 and Est_2 mainly contribute to the Hawaii differentiation whereas the differentiation of M84 derives from the contribution of the loci Got_2, Hk_1, Got_1 and Fh. The loci Est_1 and Est_2 also contribute to the Reunion differentiation, as previously assessed (Gasperi *et al.*, 1991).

Table 2 Summary of RAPD data obtained by amplification with four primers of individual flies from six wild populations and five laboratory strains of *Ceratitis capitata*

Medfly samples	Mean no. of fragments scored per individual	Total no. of fragments scored in populations	Dissimilarity index (D)
Wild populations			
Kenya 1	65.0	110	0.328
Kenya 2	66.2	123	0.387
Reunion	56.2	98	0.345
Morocco	58.2	85	0.228
Italy	51.6	68	0.171
Australia	55.2	73	0.158
Average ± SD	58.7 ± 5.7	92.8 ± 21.5	0.269 ± 0.10
Laboratory strains			
Hawaii	53.4	60	0.062
Israel	54.0	67	0.128
Ispra	56.8	70	0.122
Metapa	60.2	75	0.131
M84	54.2	63	0.091
Average ± SD	55.72 ± 2.8	67 ± 5.9	0.107 ± 0.03

Discussion

Allozymes and RAPD markers have been used to analyse the species *C. capitata*, and to relate intra-specific variability to the differentiation of its populations. RAPD markers detect arbitrary DNA sequences in the genome whereas allozyme markers allow one to infer variation at protein coding sequences. As both these kinds of markers are widely distributed over the genetic maps of *C. capitata* (Malacrida *et al.*, 1990; L. Baruffi *et al.*, unpublished data), they can be considered an unbiased sample of the medfly genome.

The evaluation of DNA polymorphism is based on the phenotypic rather than the genotypic level, unlike MLEE. It is well known that RAPD bands of equal size obtained with the same primer in different individuals do not necessary represent the same locus, and indeed the absence of an RAPD band may represent multiple allelic forms of a locus (Black IV, 1993). However, these difficulties were minimized by choosing RAPD markers, the majority of which follow Mendelian segregation (L. Baruffi *et al.*, unpubl. data). Therefore, we are confident that the polymorphic RAPD markers in the present study behave as individual loci having dominant and recessive alleles (DNA band presence or absence, respectively). RAPD data have been computed using the model of fingerprint analysis, making the assumption that RAPD markers are in Hardy–Weinberg equilibrium within and between loci (Gilbert *et al.*, 1990; Lynch, 1990). All the above assessments obey the criteria proposed by Clark &

Lanigan (1993) for estimating population genetic parameters using RAPD data.

The results obtained from the allozyme analysis are consistent with the genetic structure of *C. capitata* populations revealed by previous analyses. Genetic parameter estimates, such as the average number of alleles per locus, the proportion of polymorphic loci and the average number of heterozygous individuals, are correlated with the geographical distribution and the recent colonization history of medfly populations (Gasperi *et al.*, 1991; Malacrida *et al.*, 1992).

The RAPD analysis reveals a surprisingly high degree of polymorphism in *C. capitata*. The number of RAPD markers, identified in the medfly populations using four primers, exceeds about sevenfold the number of the biochemical loci analysed (176 vs. 26). In addition, the degree of polymorphism revealed by these RAPD markers is extensive: in the wild populations, the mean proportion of polymorphic DNA fragments is 0.66 whereas that for polymorphic biochemical loci is 0.38. Only two monomorphic RAPD fragments are conserved in all the individuals analysed. Several fragments are limited to some populations and/or individual flies. Unique bands characterize individual flies in the wild African populations (Kenya and Reunion), and the extremely variable patterns in these populations can represent DNA fingerprints. All these data demonstrate that there is a vast reservoir of unknown nuclear DNA polymorphism in the medfly. The RAPD approach adequately describes this hidden genomic variation and can be
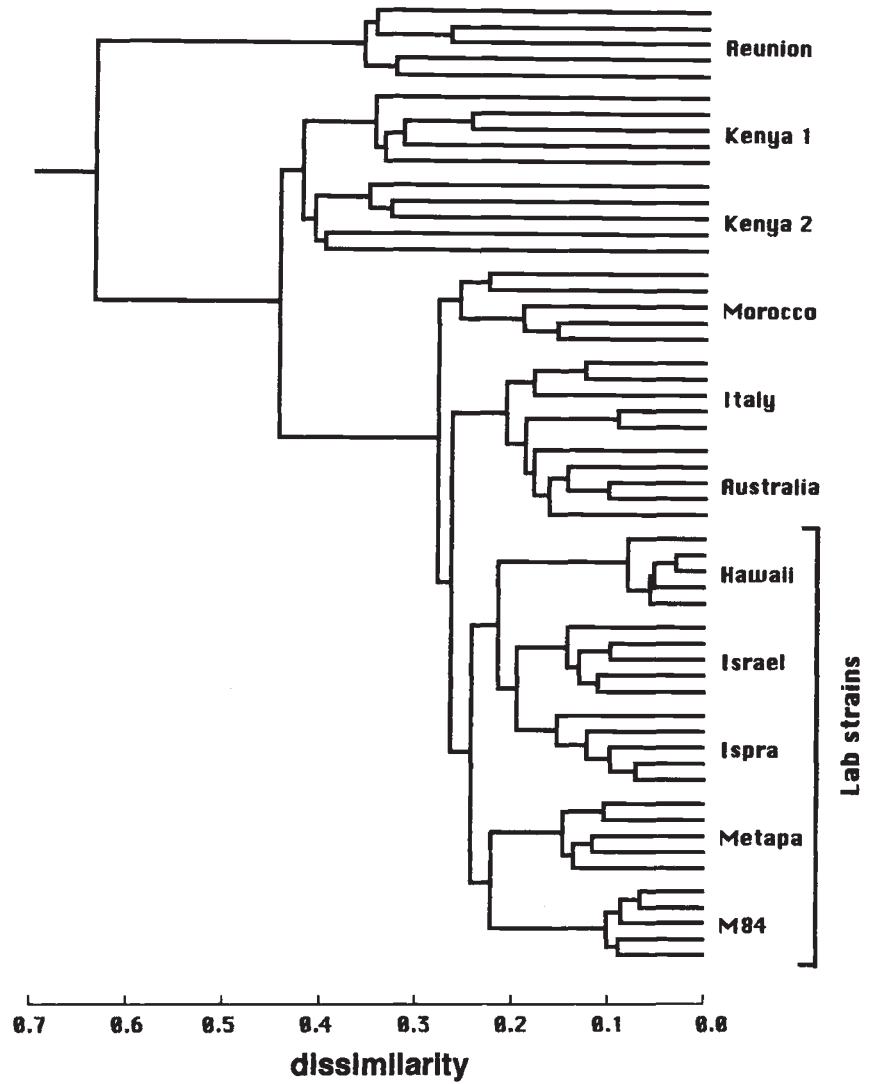
Fig. 4 Dendrogram representing individual phenetic relationships between randomly chosen flies from 11 populations of *Ceratitis capitata*. Dissimilarity values have been calculated on 176 RAPD markers amplified with the four primers used. Cluster analysis was performed using the UPGMA method.
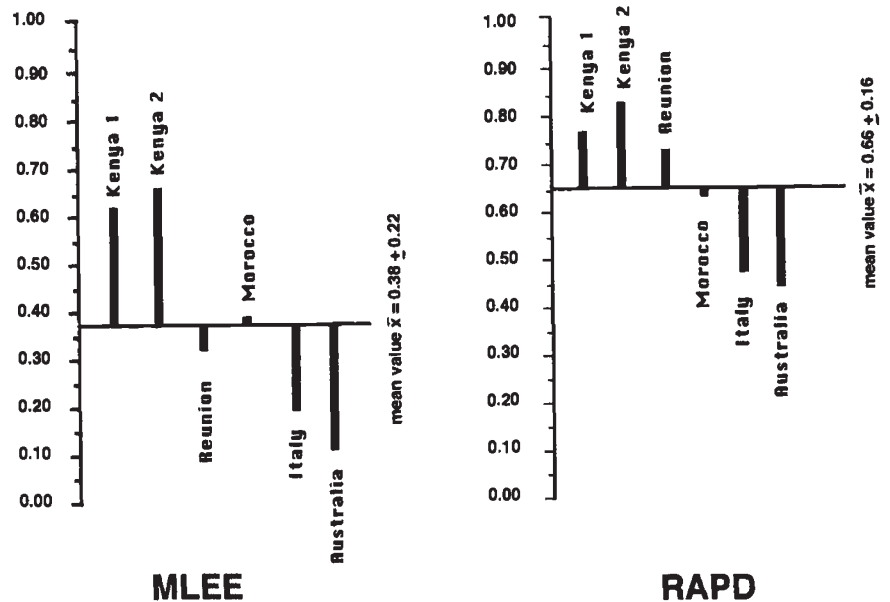

Fig. 5 Trend of polymorphism at the protein (MLEE) and the DNA (RAPD) levels in the six considered wild populations of *Ceratitis capitata*. $\bar{x}$, mean percentage of polymorphic biochemical loci, and of polymorphic DNA amplified fragments.
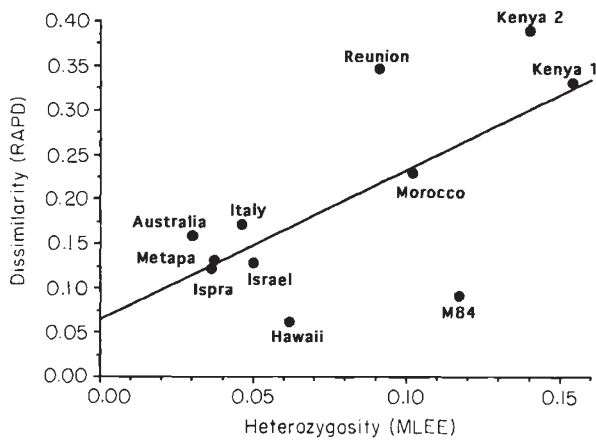
Fig. 6 Correlation between the dissimilarity values of RAPD markers and the average heterozygosities (H) at biochemical loci in 11 populations of *Ceratitis capitata* ($r = 0.67$, $P < 0.05$).

useful to address medfly intraspecific questions, concerning the genetic relatedness between individuals as well as among populations.

Of the six wild population samples considered, two are from an ancestral population (Kenya 1, Kenya 2) whereas four are from derived ones (Reunion, Morocco, Italy, Australia). The dissimilarity index estimates of intrapopulation variability obtained with RAPD show the trend of decreasing variability in medfly populations from central Africa (Kenya) towards the peripheral areas, as observed using the MLEE approach. The greater number of unique RAPD bands and biochemical alleles found in the Kenya population with respect to the derived ones, indicates that this population has the attributes of an ancestral tropical polyphagous population (Fletcher, 1989), which according to Nei's theory (Nei *et al.*, 1975) has maintained a large size over time. Therefore,
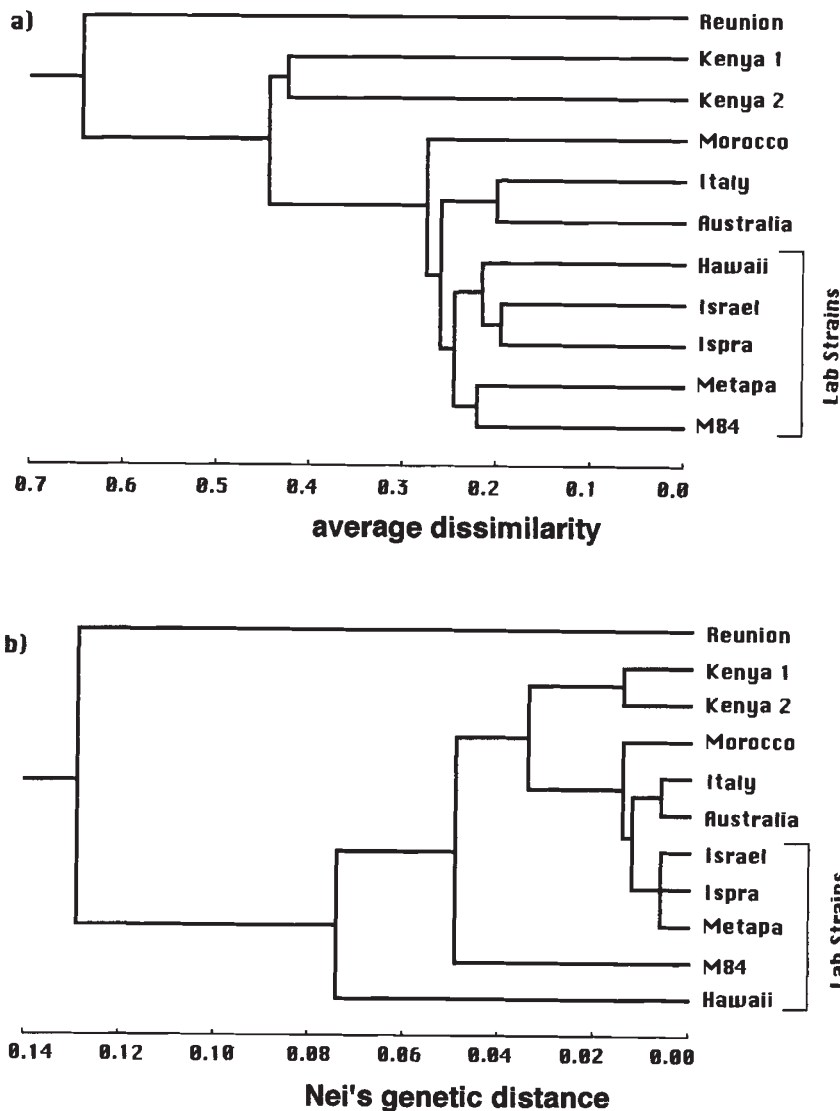

Fig. 7 Dendrograms of the 11 populations of *Ceratitis capitata* derived from RAPD dissimilarity values (a) and from genetic distances obtained by MLEE (b). The average dissimilarity values for RAPD data are based on 176 RAPD markers generated by four primers, and the Nei's genetic distances (D) for MLEE data are based on 26 biochemical loci. Cluster analysis was performed using the UPGMA method.

Table 3 Fixation index (F_{ST}) estimates in wild and laboratory populations of *Ceratitis capitata*, based on allelic frequencies at 22 polymorphic biochemical loci. Group A represents all the populations, excluding Reunion, Hawaii and M84

	F_{ST} values				
Locus	All populations	Group A	A + Reunion	A + Hawaii	A + M84
Est_1	0.654	0.191	0.595*	0.476*	0.189
Est_2	0.647	0.190	0.629*	0.510*	0.163
$Acon_1$	0.485	0.382	0.464	0.391	0.391
Got_2	0.370	0.040	0.133	0.041	0.442*
Hk_1	0.330	0.016	0.016	0.016	0.352*
Got_1	0.302	0.043	0.038	0.045	0.329*
Fh	0.255	0.092	0.095	0.095	0.244*
Mpi	0.236	0.181	0.215	0.183	0.203
Hk_2	0.179	0.178	0.184	0.178	0.171
Gpt	0.161	0.152	0.155	0.155	0.155
Zw	0.125	0.128	0.143	0.126	0.114
Pgm	0.118	0.106	0.111	0.111	0.111
Had	0.089	0.111	0.086	0.113	0.113
Aox	0.040	0.040	0.040	0.040	0.040
Idh	0.039	0.040	0.036	0.042	0.042
Pgd	0.030	0.030	0.030	0.030	0.030
Ak_2	0.022	0.022	0.022	0.022	0.022
Pgi	0.021	0.021	0.021	0.021	0.021
$Acon_2$	0.016	0.016	0.016	0.016	0.016
Pgk	0.015	0.015	0.015	0.015	0.015
Mdh_2	0.015	0.015	0.015	0.015	0.015
Mdh_1	0.013	0.013	0.013	0.013	0.013

*F_{ST} values were greatly increased by the addition of the single population.

the highest level of genetic variability found in Kenya supports the general opinion that this population is present in the source area of the species (Huettel *et al.*, 1980; Gasperi *et al.*, 1991). The dispersion itself might explain a reduction in variability, at the DNA and protein level, in the derived populations as a consequence of bottlenecks in population size (Nei *et al.*, 1975). The estimated correlation between RAPD and MLEE data suggests that the biochemical and molecular markers seem to be similarly affected by factors such as population size and drift during the colonization process. However, the amount of variability decreases fivefold from the ancestral to the peripheral populations when measured at the allozyme level (*H* ranging from 0.154 to 0.030) and only twofold when estimated at the DNA level (*d* ranging from 0.387 to 0.158). The peripheral populations appear, indeed, to lose less variation at the DNA level than at the protein level. The variation at the DNA level is detected at

random sequences of the genome, and consequently it may represent noncoding sequences. Different parts of the genome can evolve at different rates (Kimura, 1983; Nei, 1987). Therefore, it is possible that one or several regions amplified by PCR evolve at a higher rate compared with the enzymatic loci. Thus the variation gain can be explained as a result of the higher evolutionary rates of the DNA regions amplified, especially if they contain micro- and/or minisatellites. Variation detected at the protein level can be generated by processes different from those working on the entire genome. Although mainly drift may contribute to the decreasing variation at the anonymous DNA sequences (RAPD), directional selection may have reinforced the action of drift in the reduction of variability during the colonization processes at the protein level. A similar conclusion has been reported by Begun & Aquadro (1993) who pointed out that, in *Drosophila melanogaster*, the variation in restriction sites of

specific DNA loci and for allozymes probably experienced substantially different evolutionary dynamics.

It may be of some interest that the primer NP4 is the most efficient at highlighting the trend of decreasing variability from the ancestral to the derived populations. This primer may target some medfly chromosomal regions related to the colonization processes. That this primer detects a very low degree of polymorphism in laboratory strains supports this hypothesis. Fixation of specific markers in these strains may be the result of inbreeding from long-term laboratory maintenance.

Another genetic aspect of the spread of *C. capitata* is the differentiation of its populations. Differentiation processes can be inferred from the genetic distances at both DNA and protein levels. The trees of Figs 7a and b show that Nei's genetic distance and the average dissimilarity estimates between wild African, Mediterranean, and Australian samples both reflect the events during the colonization of the species. At both protein and DNA levels, a major part of the intraspecific variability involves the differentiation of central vs. peripheral populations. Reunion appears to be the most differentiated population in both the trees. However, the potential causes of this differentiation are different at the protein and DNA levels. At the protein level, the main contribution to the differentiation of this population is given by the presence of fixed alleles at certain biochemical loci (Est_1 and Est_2), which have been previously proposed as possible targets of natural selection in this population (Gasperi *et al.*, 1991). On the other hand, at the DNA level, the high dissimilarity index value found between individuals of this population (Table 2), and the high average dissimilarity between Reunion and the other populations, indicate that there is a tendency for Reunion to harbour genomic variability, providing further evidence of different evolutionary processes in this population in comparison with the others. It is possible that colonization of Reunion occurred at a different period compared with the other populations. Whether the differentiation history of Reunion differs because of isolation, as opposed to environmental variation, is an open question. The results obtained in Reunion at the protein level (allelic fixation) and those at the DNA level (high values of dissimilarity) reinforce the hypothesis that these two types of variation can be influenced by different evolutionary forces.

The laboratory strains show a certain degree of genetic similarity at the DNA level, despite their different origins; at the protein level they show either a high genetic similarity or a striking difference (Hawaii, M84). Their colonization histories, laboratory adapta-

tion effects and allelic fixation from random or non-random inbreeding, singly or in concert, have produced the observed results in these strains.

In conclusion, the contemporary use of RAPD and MLEE seems to give a very congruent picture of the population structure of *C. capitata*, for which drift and selection are the major causes of differentiation. On the other hand, when other genetic constraints interfere with drift, as in the laboratory strains, incongruence appears between the two methodological approaches.

References

BANDI, C., LA ROSA, G., COMINCINI, S., DAMIANI, G. AND POZIO, E. 1993. Random amplified polymorphic DNA technique for the identification of *Trichinella* species. *Parasitology*, **107**, 419–424.

BEGUN, D. J. AND AQUADRO, C. F. 1993. African and North American populations of *Drosophila melanogaster* are very different at the DNA level. *Nature*, **365**, 548–550.

BENDER, W., SPIERER, P. AND HOGNESS, D. S. 1983. Chromosomal walking and jumping to isolate DNA from *Ace* and *rosy* loci and the bithorax complex in *Drosophila melanogaster*. *J. Mol. Biol.*, **168**, 17–33.

BLACK IV, W. C. 1993. PCR with arbitrary primers: approach with care. *Insect Mol. Biol.*, **2**, 1–6.

CLARK, A. G. AND LANIGAN, C. M. S. 1993. Prospects for estimating nucleotide divergence with RAPDs. *Mol. Biol. Evol.*, **10**, 1096–1111.

DAVID, J. R. AND CAPY, P. 1988. Genetic variation of *Drosophila melanogaster* natural populations. *Trends Genet.*, **4**, 106–111.

FLETCHER, B. S. 1989. Movements of tephritid fruit flies. In: Robinson A. S. and Hooper G. H. S. (eds), *Fruit Flies,*

Their Biology, Natural Enemies and Control, vol. 3B, pp. 209–219. Elsevier, Amsterdam, The Netherlands.

FROHLICH, D. R., ROBINSON, A. S. AND WELLS, M. A. 1993. Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann) mitochondrial DNA: genes and secondary structures for six t-RNAs. *Insect Mol. Biol.,* **1**, 165–169.

GASPERI, G., GUGLIELMINO, C. R., MALACRIDA, A. R. AND MILANI, R. 1991. Genetic variability and gene flow in geographical populations of (medfly) *Ceratitis capitata* (Wied.). *Heredity,* **67**, 347–356.

GASPERI, G., MALACRIDA, A. R., GUGLIELMINO, C. R., BARUFFI, L., TORTI, C., MILANI, R., DAMIANI, G. AND BANDI, C. 1993. *Ceratitis capitata*: suitable markers for population genetics and genome analysis. In: *Management of Insect Pests: Nuclear and Related Molecular and Genetic Techniques,* pp. 251–256. International Atomic Energy Agency Press, Vienna, Austria.

GILBERT, D. A., LEHMAN, N., O'BRIEN, S. J. AND WAYNE, R. K. 1990. Genetic fingerprinting reflects population differentiation in the California Channel Island fox. *Nature,* **344**, 764–767.

HADRYS, H., BALICK, M. AND SCHIERWATER, B. 1992. Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. *Mol. Ecol.,* **1**, 55–63.

HAYMER, D. S. AND McINNIS, D. O. 1994. Resolution of populations of the Mediterranean fruit fly at the DNA level using random primers for the polymerase chain reaction. *Genome,* **37**, 244–248.

HEDRICK, P. 1992. Shooting the RAPDs. *Nature,* **355**, 679–680.

HUETTEL, M. D., FUERST, P. A., MARUYAMA, M. AND CHAKRABORTY, R. 1980. Genetic effects of multiple population bottlenecks in the Mediterranean fruit fly (*Ceratitis capitata*). *Genetics,* **94s**, 47–48 (abstract).

KIMURA, M. 1983. *The Neutral Theory of Molecular Evolution.* Cambridge University Press, Cambridge.

LYNCH, M. 1990. The similarity index and DNA fingerprinting. *Mol. Biol. Evol.,* **7**, 478–484.

McPHERON, B. A., GASPARICH, G. E., HAN, H.-Y., STECK, G. J. AND SHEPPARD, W. S. 1994. Mitochondrial DNA restriction map for the Mediterranean fruit fly, *Ceratitis capitata. Biochem. Genet.,* **32**, 25–33.

MALACRIDA, A. R., GASPERI, G., BARUFFI, L. AND MILANI, R. 1990. The contribution of formal genetic studies to the species characterization of *Ceratitis capitata* (Wied.). In: *Genetic Sexing of the Mediterranean Fruit Fly,* pp. 85–89. International Atomic Energy Agency Press, Vienna, Austria.

MALACRIDA, A. R., GUGLIELMINO, C. R., GASPERI, G., BARUFFI, L. VILLANI, P. C. AND MILANI, R. 1991. Genetical approach to systematics and phylogeny of Trypetinae (Diptera, Tephritidae). *Boll. Zool.,* **58**, 355–361.

MALACRIDA, A. R., GUGLIELMINO, C. R., GASPERI, G., BARUFFI, L. AND MILANI, R. 1992. Spatial and temporal differentiation in colonizing populations of *Ceratitis capitata. Heredity,* **69**, 101–111.

MANTEL, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.,* **27**, 209–220.

MILANI, R., GASPERI, G. AND MALACRIDA, A. R. 1989. Biochemical genetics [of *Ceratitis capitata*]. In: Robinson A. S. and Hooper G. H. S. (eds) *Fruit Flies: Their Biology, Natural Enemies and Control,* vol. 3B, pp. 33–56. Elsevier, Amsterdam, The Netherlands.

NEI, M. 1972. Genetic distance between populations. *Am. Nat.,* **106**, 283–292.

NEI, M. 1987. *Molecular Evolutionary Genetics.* Columbia University Press.

NEI, M. AND CHESSER, R. K. 1983. Estimation of fixation indices and gene diversities. *Ann. Hum. Genet.,* **47**, 253–259.

NEI, M., MARUYAMA, T. AND CHAKRABORTY, R. 1975. The bottleneck effect and genetic variability in populations. *Evolution,* **29**, 1–10.

ROBINSON A. S. AND HOOPER G. H. S. (EDS) 1989. *Fruit Flies, Their Biology, Natural Enemies and Control,* vol. 3 A and B. Elsevier, Amsterdam, The Netherlands.

SAMBROOK J., FRITSCH E. F. AND MANIATIS, T. 1989. *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, New York.

SNEATH, P. H. A. AND SOKAL, R. R. 1973. *Numerical Taxonomy.* W.H. Freeman and Company, San Francisco.

SWOFFORD, D. L. AND SELANDER, R. B. 1981. BIOSYS-1. *A Computer Program for the Analysis of Allelic Variations in Genetics.* University of Illinois, Urbana.

WELSH, J. AND McCLELLAND, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucl. Acids Res.,* **18**, 7213–7218.

WILLIAMS, J. G. K., KUBELIK, A. R., LIVAK, K. J., RAFALSKI, J. A. AND TINGEY, S. V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl. Acids Res.,* **18**, 6531–6535.

YUHKI, N. AND O'BRIEN, S. 1990. DNA variation of the mammalian major histocompatibility complex reflects genomic diversity and population history. *Proc. Natl. Acad. Sci. U.S.A.,* **87**, 836–840.