

# Genetic variability and gene flow in geographical populations of *Ceratitis capitata* (Wied.) (medfly)

GIULIANO GASPERI, CARMELA R. GUGLIELMINO,\* ANNA R. MALACRIDA & RICCARDO MILANI

Dipartimento di Biologia Animale, Università di Pavia, Piazza Botta 9, 27100 Pavia, and \*Dipartimento di Genetica e Microbiologia, Università di Pavia ed Istituto di Genetica e Biologia Evoluzionistica del CNR, Pavia, Italy

Two African populations of *Ceratitis capitata* (Kenya and Réunion Isl.) and two Mediterranean ones (Sardinia and Procida Isl.) have been studied for genetic variability at 25 loci by electrophoresis. Wright's  $F_{ST}$ , Slatkin's  $Nm_*$  gene flow estimator, Nei's distance (D) together with measures of variability such as  $\bar{H}$ ,  $\bar{P}$ ,  $\bar{A}$  have been used to compare the population from Kenya with the other three. Parameters using gene frequencies ( $F_{ST}$ , D,  $Nm_*$ ) indicate the presence of substantial geographic heterogeneity, largely attributable to genetic drift and correlated with dispersion of the medfly from its source area (Subsaharan Africa) to the periphery. The Kenyan population has high genetic variability (assessed by  $\bar{H}$ ,  $\bar{P}$  and  $\bar{A}$ ), as might be expected given its native status. Significant gene flow estimates between Kenya and the derived Mediterranean populations supports the hypothesis of recent colonization. Part of the geographic heterogeneity is related to the presence of fixed alleles in the more differentiated Réunion population although it maintains the genetic attributes of the ancestral population. Selection or other forces may have played an important role in the differentiation of this population.

**Keywords:** *Ceratitis capitata*, electrophoretic markers, gene flow, genetic variability, geographical populations.

## Introduction

The Mediterranean fruit fly, *Ceratitis capitata* (Diptera, Tephritidae) originated in Africa and has a recent history of very rapid, extensive, and initially quite discontinuous colonization of tropical and subtropical countries, including those very remote from the source centre, such as South and Central America and Australia (Saul, 1986). Little is known about the habits of *C. capitata* in its original home but it has proved to be highly polyphagous in countries where it has become established, infesting a wide variety of commercial, ornamental and wild hosts. It is multivoltine, with the number of generations per year determined mainly by temperature (Fletcher, 1989). Considerable information is available on the biology, physiology, behaviour and ecology of the insect (for a comprehensive review see Robinson & Hooper, 1989). However, despite its economic importance, it is still poorly studied from the standpoint of population genetics. Knowledge of the genetic structure and

dynamics of natural populations affords opportunities to examine the genetic consequences of its spreading and the ecology genetics and evolution of the species (Milani *et al.*, 1989). Intraspecific genetic differentiation is expected to result from the colonization of *C. capitata*, despite the apparent morphological uniformity within this species (White, 1989). Previous studies have shown large genetic differences between introduced populations and African ones (Huettel *et al.*, 1980; Gasperi *et al.*, 1987; Loukas, 1989). Based on this variability the Mediterranean region has been considered to be the primary introduction area (Huettel *et al.*, 1980).

The present study concerns four populations from geographically and ecologically distinct areas, namely a native population (Kenya) and three derived ones (Réunion Island, to the east of Madagascar, and Sardinia and Procida islands in the Mediterranean basin). Data on the spatial distribution of allelic variation show that the pattern is strikingly heterogeneous for different loci, which suggests the possible involve-

ment of different types of evolutionary forces during the process of colonization. Such data provide information on the relative contributions of selection, gene flow, genetic drift and geographical isolation to genetic differentiation in this species.

## Materials and methods

### Populations of *Ceratitis capitata*

Two populations from the African region (Kenya and Réunion Island) and two populations from the Mediterranean basin (Sardinia, and Procida Island in the Bay of Naples, Italy) were studied according to sampling opportunities.

The Kenyan population was represented by two samples of pupae from coffee berries, collected in August and October 1984, respectively, in two farms near Nairobi. The population of Réunion Island was sampled twice, in October and December 1987, near St Denis from pupae collected on *Mimusops elengi* (Sapotaceae). The Procida population (eight samples) was sampled over seasons in 1985 (from February to October). The population from Sardinia was sampled once for each of these consecutive years (October 1985, September 1986, September 1987). Collections of wild flies from these Mediterranean populations were made by harvesting infested fruits (peach, apricot, fig, bitter orange) from the ground beneath trees and allowing the larvae to pupate in the laboratory. All the samples produced large number of flies, ranging from 156 to 751 (one of the two samples from Réunion yielded only 50 adult flies). Principal component analysis and other tests (Gasperi *et al.*, 1990) demonstrated that the separate collections within a location were substantially homogeneous.

### Electrophoretic procedures

At least 25 individuals were assayed for each population sample at the 25 enzyme loci listed in Table 1.

Adult medflies were homogenized in 100 mM Tris-HCl buffer pH 7.5, containing 0.47 mM  $\beta$ -mercaptoethanol, 1 mM  $\text{Na}_2$ -EDTA, 0.5 mM NAD and 0.5 mM NADP. Allozyme variability was assayed by electrophoresis on Cellogel (Cellulose acetate gel sheets from Chemetron, Milan, Italy). The buffer TEC (pH 7.5) was used for electrophoresis of the majority of enzymes (61.4 mM Tris, 4 mM EDTA and 13.6 mM citric acid). The buffer TBE (pH 7.2) was used (0.95 mM Tris, 4 mM EDTA and 310 mM boric acid) for electrophoresis of esterase isoenzymes. Staining for enzyme activities after electrophoresis was based on recipes of Meera Khan (1971) and Harris & Hopkinson (1976).

### Data analysis

The measurement and testing of heterogeneity among samples from different geographical areas were carried out in different ways.  $H_S$ ,  $H_T$ ,  $F_{IS}$ ,  $F_{IT}$  and  $F_{ST}$  statistics were computed by the method of Nei & Chesser (1983). The  $F_{ST}$  values were tested for departure from zero with the method of Workman & Niswander (1970). Genetic distances ( $D$ ) were calculated using Nei's method (1972) on the full dataset. Gene flow estimates ( $Nm_w$ ), were derived from  $F_{ST}$  values by the relationship  $F_{ST} = 1/(1 + 4Nm)$ , where  $m$  is the average rate of immigration in an 'island' model of population structure and  $N$  the local population size (Wright, 1931). Gene flow levels among populations were also estimated by Slatkin's methods (1981; 1985), based on the distribution of rare alleles ( $Nm_*$ ). Quantitative estimates of the amount of gene flow were calculated following Slatkin's (1985) formula:

$$Nm_* = \frac{\exp\{-[\ln\{\bar{p}(1)\} + 2.44]/0.505\}}{N/25}$$

where  $\bar{p}(1)$  is the average frequency of all alleles found in only one population, and  $N$  is the average number of individuals sampled per population.

## Results

The enzymes assayed in this survey are listed in Table 1. This table also includes data on the structure of the enzymes and on the number of alleles discovered. Most of the 25 enzymes studied directly control pathways of basic importance to energy metabolism, others are correlated with these pathways, while some, such as the esterases, are involved in specific, not always defined, metabolic functions.

Of the 25 loci, only 3 ( $Ak_1$ ,  $Aox$ ,  $\alpha$ - $Gpdh$ ) were found to be uniformly monomorphic. Out of the 22 polymorphic loci, 19 have been mapped to five of the six chromosomes of *C. capitata*. They appear to be widely distributed on the genetic map (Malacrida *et al.*, 1988), and thus can be considered to be a random set, suitable for the analysis of intraspecific variability. As in other organisms (Zouros, 1976), there is a correlation between the subunit structure of the enzymes and the extent of genetic polymorphism: heterozygosities for multimers are lower (dimeric:  $\bar{H} = 0.033$ ; tetrameric:  $\bar{H} = 0.002$ ) than heterozygosities for monomers ( $\bar{H} = 0.198$ ).

### Single-locus measures of population differentiation

The pattern of allelic differentiation within and between geographically distant populations of *Ceratitis*

**Table 1** Gene-enzyme systems analysed in field-collected samples of *Ceratitidis capitata*

Main metabolic pathway	Enzyme (E.C. number)	Enzyme symbol	Molecular subunit structure	Number of loci studied	Locus symbol	Total number of electrophoretic alleles per locus
Conversion of carbohydrate reserves to glucose 6-phosphate						
Hexokinase	(2.7.1.1)	HK-1	Monomeric	2	<i>Hk</i> <sub>1</sub>	2
		HK-2	Monomeric		<i>Hk</i> <sub>2</sub>	4
Phosphoglucumutase	(2.7.5.1)	PGM	Monomeric	1	<i>Pgm</i>	5
Glycolysis						
Phosphoglucose isomerase	(5.3.1.9)	PGI	Dimeric	1	<i>Pgi</i>	2
Phosphoglycerate kinase	(2.7.2.3)	PGK	Monomeric	1	<i>Pgk</i>	3
Pentose phosphate pathway						
Glucose 6-phosphate dehydrogenase	(1.1.1.49)	G6PD	Dimeric	1	<i>Zw</i>	2
6-Phosphogluconate dehydrogenase	(1.1.1.44)	6PGD	Dimeric	1	<i>Pgd</i>	3
$\alpha$ -Glycerophosphate shuttle						
$\alpha$ -Glycerophosphate dehydrogenase	(1.1.1.8)	$\alpha$ -GPDH	Dimeric	1	<i><math>\alpha</math>-Gpdh</i>	1
Krebs cycle						
Isocitrate dehydrogenase	(1.1.1.42)	IDH	Dimeric	1	<i>Idh</i>	3
Fumarate hydratase	(4.2.1.2)	FH	Tetrameric	1	<i>Fh</i>	2
Malate dehydrogenase	(1.1.1.37)	MDH-1	Dimeric	2	<i>Mdh</i> <sub>1</sub>	3
		MDH-2	Dimeric		<i>Mdh</i> <sub>2</sub>	2
Others						
Adenylate kinase	(2.7.4.3)	AK-1	Monomeric	2	<i>Ak</i> <sub>1</sub>	1
		AK-2	Monomeric		<i>Ak</i> <sub>2</sub>	2
Alcohol dehydrogenase	(1.1.1.1)	ADH-2	Dimeric	1	<i>Adh</i> <sub>2</sub>	2
Aldehyde oxidase	(1.2.3.1)	AOX	Dimeric	1	<i>Aox</i>	1
Esterases	(3.1.1.-)	EST-1	Monomeric	3	<i>Est</i> <sub>1</sub>	3
		EST-2	Monomeric		<i>Est</i> <sub>2</sub>	3
		EST-6	Monomeric		<i>Est</i> <sub>6</sub>	2
Glutamate oxalacetate transaminase	(2.6.1.1)	GOT-1	Dimeric	2	<i>Got</i> <sub>1</sub>	2
		GOT-2	Dimeric		<i>Got</i> <sub>2</sub>	2
Glutamate pyruvate transaminase	(2.6.1.2)	GPT	Dimeric	1	<i>Gpt</i>	2
Hydroxyacid dehydrogenase	(1.1.1.35)	HAD	Dimeric	1	<i>Had</i>	2
Malic enzyme	(1.1.1.40)	ME	Tetrameric	1	<i>Me</i>	2
Mannose phosphate isomerase	(5.3.1.8)	MPI	Monomeric	1	<i>Mpi</i>	7

*capitata* has been analysed in order to assess the degree of genetic divergence in relation to ecogeographical conditions. In Table 2, for each of the 22 polymorphic loci, the heterozygosity of the pooled four geographical populations was partitioned into two components within-population ( $H_S$ ) and between-populations ( $H_T$ ).  $H_S$  is the mean of the individual population heterozygosities at the locus and  $H_T$  is the total expected heterozygosity over all populations sampled, calculated from the total allele frequencies ( $P_i$ ) at the locus. The proportion of heterozygosity attributable to differentiation between populations is represented by the fixation index ( $F_{ST}$ ), given by  $F_{ST} = 1 - H_S/H_T$ . Loci are arranged in Table 2 in order of decreasing fixation

index. In the plot of Fig. 1, the lack of correlation between the average heterozygosity at these loci and their fixation indices is due to the presence of *Est*<sub>1</sub> and *Est*<sub>2</sub>. If these loci are removed from the dataset the correlation is highly significant ( $r = 0.66$ ,  $P < 0.01$ ).  $F_{ST}$  is significantly different from zero at 15 loci (Table 2). The average value of  $\bar{F}_{ST} = 0.123$  indicates that at least 12 per cent of the total variability in *C. capitata* is attributable to divergence between populations. An important feature of the results listed in Table 2 is the wide range of  $F_{ST}$  values among loci (from 0.685 at *Est*<sub>2</sub> to 0). As drift alone would affect all loci similarly, the observed  $F_{ST}$  heterogeneity may be caused by some form of selection. Part of the differentiation could

**Table 2** Comparison of polymorphism within and between geographical populations (Kenya, Réunion, Sardinia, Procida) of *C. capitata*

Locus	Number of alleles	Within-population heterozygosity $H_S$	Total heterozygosity $H_T$	Fixation index $F_{ST}$	Heterogeneity chi-squared (d.f.)
<i>Est</i> <sub>2</sub>	3	0.157	0.499	0.685	443.9 (6)***
<i>Est</i> <sub>1</sub>	3	0.208	0.537	0.613	522.3 (6)***
<i>Hk</i> <sub>2</sub>	4	0.477	0.611	0.219	279.6 (9)***
<i>Got</i> <sub>2</sub>	2	0.103	0.126	0.183	80.1 (3)***
<i>Hk</i> <sub>1</sub>	2	0.081	0.095	0.147	65.4 (3)***
<i>Ak</i> <sub>2</sub>	2	0.194	0.222	0.139	62.1 (3)***
<i>Had</i>	2	0.076	0.087	0.130	59.9 (3)***
<i>Est</i> <sub>6</sub>	2	0.395	0.452	0.126	32.2 (3)***
<i>Zw</i>	2	0.171	0.193	0.117	49.6 (3)***
<i>Mpi</i>	7	0.531	0.595	0.108	272.4 (18)***
<i>Pgm</i>	5	0.233	0.250	0.069	133.6 (12)***
<i>Pgk</i>	3	0.135	0.145	0.066	56.8 (6)***
<i>Gpt</i>	2	0.036	0.037	0.033	7.3 (2)*
<i>Pgd</i>	3	0.032	0.034	0.032	28.2 (6)***
<i>Got</i> <sub>1</sub>	2	0.029	0.030	0.032	13.4 (3)**
<i>Idh</i>	3	0.067	0.068	0.006	5.8 (6)
<i>Mdh</i> <sub>1</sub>	3	0.035	0.036	0.005	4.8 (6)
<i>Mdh</i> <sub>2</sub>	2	0.011	0.011	0.004	2.0 (3)
<i>Me</i>	2	0.009	0.009	0.002	0.0 (2)
<i>Pgi</i>	2	0.011	0.011	0.000	—
<i>Adh</i> <sub>2</sub>	2	0.005	0.005	0.000	—
<i>Fh</i>	2	0.006	0.006	0.000	—
Mean		$\bar{H}_S = 0.136$	$\bar{H}_T = 0.184$	$\bar{F}_{ST} = 0.123$	

\*\*\* $P < 0.001$ ; \*\* $P < 0.01$ ; \* $P < 0.05$ 

simply be due to drift and part could be maintained by some sort of selection that affects particular genes (see also Singh *et al.*, 1987).

Estimates of  $F_{IS}$ ,  $F_{IT}$  and  $F_{ST}$  for each of the polymorphic loci are partitioned according to linkage group (Table 3), because of the possibility that the linkage among loci might affect their pattern of variability.  $F_{IS}$  and  $F_{IT}$ , estimated as

$$1 - \frac{\text{observed heterozygosity}}{\text{expected heterozygosity}}$$

relate to the deviation from panmixia within individual populations and in the pooled population, respectively (Nei *et al.*, 1975). All linkage groups, except A (Chr. 4), appear to have the same values for these parameters and the average values of  $\bar{F}_{ST}$  range from 0.036 to 0.094. For linkage group A, the mean values of  $\bar{F}_{ST}$  is high (0.379) and the values of  $F_{IS}$  are negative for all four loci, which suggests a departure from panmixia for the local populations, with an excess of heterozygotes. However, for two of these loci, *Est*<sub>1</sub> and *Est*<sub>2</sub>,  $F_{IT}$

values, which are sensitive to high gene frequency differences among populations, have high positive estimates, which indicates an excess of homozygotes. Populations fixed for a single electromorph are considered to be undefined and eliminated from the computation of  $F_{IS}$ . The strong divergence between  $F_{IS}$  and  $F_{IT}$  values then depends on the presence of fixed alleles at the *Est*<sub>1</sub> and *Est*<sub>2</sub> loci. In this linkage group *Est*<sub>1</sub>, *Est*<sub>2</sub> and, to a lesser degree, *Hk*<sub>2</sub> contribute substantially to the geographical differentiation of *C. capitata*. Note that *Est*<sub>1</sub> and *Est*<sub>2</sub> are tightly linked and are located at  $26.8 \pm 0.03$  cM from *Hk*<sub>2</sub> (Gasperi *et al.*, 1986).

#### Genes and population differentiation

The contribution of single loci to population differentiation is presented in Fig. 2. If all populations are considered, the distribution of  $F_{ST}$  is mainly around 0.14 and only four loci show values around or greater than 0.20. These loci are *Est*<sub>2</sub>, *Est*<sub>1</sub>, *Hk*<sub>2</sub>, *Got*<sub>2</sub>. A large portion of the interpopulation differentiation is contributed by the Réunion population, as only *Hk*<sub>2</sub> shows



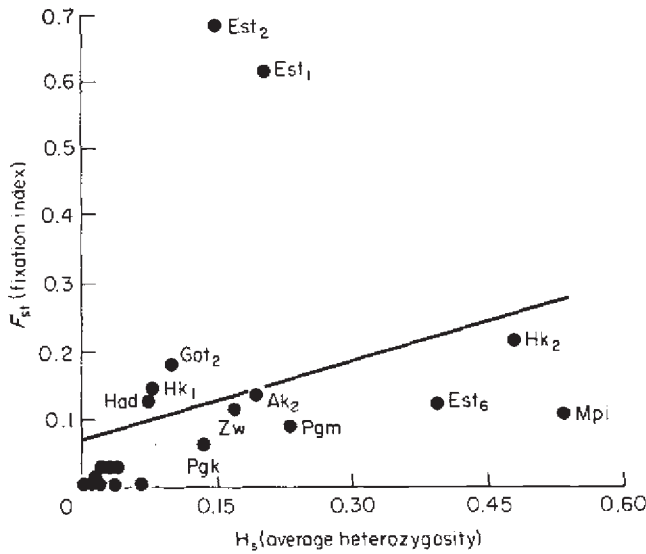


Fig. 1 The correlation between the fixation index ( $F_{ST}$ ) on average heterozygosity ( $H_s$ ) at allozyme loci in four geographical populations of *Ceratitidis capitata* ( $r=0.326$ ).

$F_{ST}$  values greater than 20 per cent if this population is omitted from the computation. Réunion seems to be significantly differentiated from all the other populations at the *Est* loci, due to the presence in this population of fixed alleles alternative to the most common in the other populations. *Hk<sub>2</sub>* ( $F_{ST} > 20$  per cent) contributes to the differentiation of the Kenya population. No significant changes are observed for the fixation index of the above mentioned loci when the other populations are omitted from the computation.

**Genetic structure of populations**

The levels of genetic variability for each population, considering all 25 loci, are summarized in Table 4. A general trend seems to characterize these populations of *C. capitata*, which are polymorphic to different degrees. In particular, the Mediterranean populations (Procida and Sardinia) are less polymorphic than subtropical (Réunion) and tropical ones (Kenya). This is supported by all three statistics:  $\bar{A}$  (average number of alleles per locus),  $\bar{P}$  (proportion of polymorphic loci) and  $\bar{H}$  (average number of heterozygous individuals). These findings are in agreement with the general rule of a decreasing trend of genetic variability away from the source area of a species towards the periphery of its geographical range. The Kenya population has the highest average number of alleles per locus ( $2.00 \pm 0.16$ ); the major contribution to this high value is due to the presence, in this population, of several low

Table 3  $F$ -statistics analysis for the enzyme loci located on various chromosomes, in geographical populations (Kenya, Réunion, Sardinia and Procida) of *C. capitata*

Locus	Number of alleles	$F_{IS}$	$F_{IT}$	$F_{ST}$	$\bar{F}_{ST}$
Linkage group A (Chr. 4)					
<i>Hk<sub>2</sub></i>	4	-0.195	0.066	0.219	
<i>Est<sub>1</sub></i>	3	-0.264	0.511	0.613	0.379
<i>Est<sub>2</sub></i>	3	-0.192	0.624	0.685	
<i>Pgi</i>	2	-0.087	-0.090	0.000	
Linkage group B (Chr. 5)					
<i>Zw</i>	2	-0.260	-0.112	0.117	
<i>Pgd</i>	3	-0.075	-0.040	0.032	
<i>Fh</i>	2	0.175	0.169	0.000	0.093
<i>Had</i>	2	0.078	0.197	0.130	
<i>Hk<sub>1</sub></i>	2	0.046	0.187	0.147	
<i>Gpt</i>	2	-0.014	0.019	0.033	
Linkage group C (Chr. 3)					
<i>Got<sub>2</sub></i>	2	-0.142	0.067	0.183	0.094
<i>Mdh<sub>1</sub></i>	3	-0.058	-0.052	0.005	
Linkage group D (Chr. 2)					
<i>Mpi</i>	7	0.281	0.359	0.108	
<i>Est<sub>6</sub></i>	2	-0.279	-0.118	0.126	
<i>Mdh<sub>2</sub></i>	2	-0.093	-0.089	0.004	0.059
<i>Adh<sub>2</sub></i>	2	0.099	0.093	0.000	
Linkage group E (Chr. 6)					
<i>Idh</i>	3	0.038	0.044	0.006	
<i>Pgm</i>	5	0.034	0.101	0.069	0.036
<i>Got<sub>1</sub></i>	2	-0.027	0.006	0.032	
Unmapped loci					
<i>Ak<sub>2</sub></i>	2	-0.133	0.024	0.139	
<i>Pgk</i>	3	-0.146	-0.070	0.066	0.069
<i>Me</i>	2	0.231	0.231	0.002	

frequency alleles. The number of rare alleles present only in this population (private alleles according to Slatkin, 1985) can be estimated to be about 25 per cent of the total number of alleles. The discovery of several low frequency alleles in Kenya supports the hypothesis that this population has maintained a large size (Nei *et al.*, 1975).

**Genetic distance**

When gene frequencies are taken into account to infer intraspecific Nei (1978) genetic distances (Table 5), the geographical trend of decreasing variability (previously observed in Table 4) is masked. In fact, while the smallest genetic distance is between the geographically close populations of Procida and Sardinia ( $D=0.011$ ),

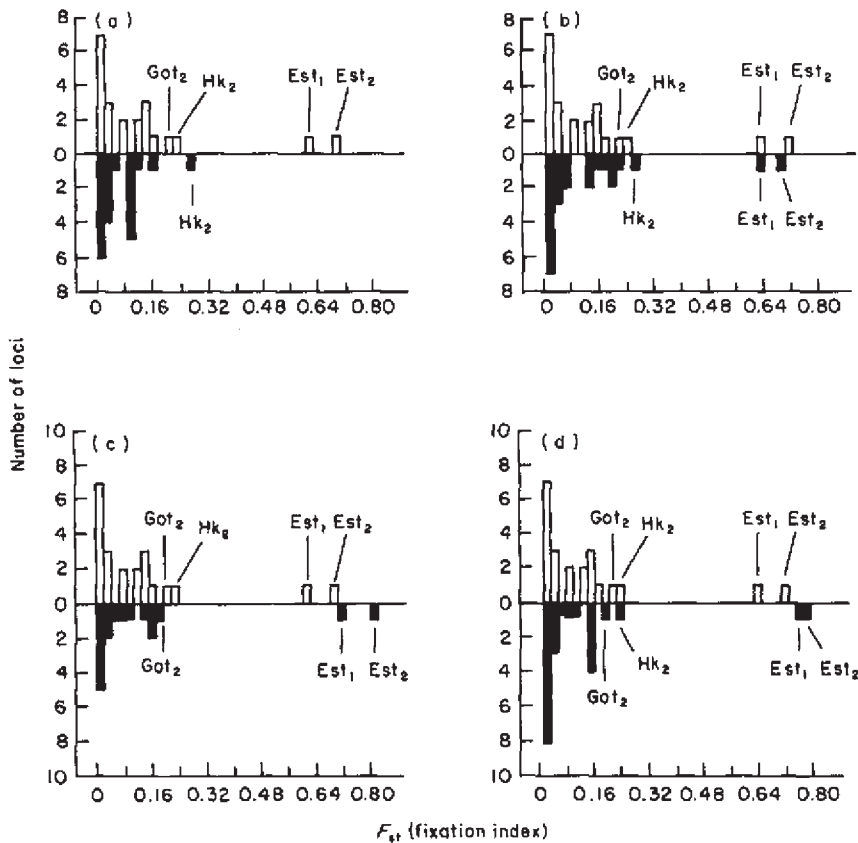


Fig. 2 The distribution of the fixation index ( $F_{ST}$ ) at polymorphic loci in four geographical populations of *Ceratitis capitata*. (□) All four populations; (■) excluding: (a) Réunion, (b) Procida, (c) Kenya, (d) Sardinia.

Table 4 Parameters of genetic variability in four populations of *C. capitata*

Geographical origin	$\bar{A} \pm \text{s.d.}$	$\bar{P} \pm \text{s.d.}$	$\bar{H} \pm \text{s.d.}$
Kenya	$2.00 \pm 0.16$	$0.568 \pm 0.08$	$0.173 \pm 0.01$
Réunion	$1.527 \pm 0.26$	$0.403 \pm 0.14$	$0.116 \pm 0.07$
Sardinia	$1.337 \pm 0.068$	$0.238 \pm 0.06$	$0.084 \pm 0.026$
Procida	$1.232 \pm 0.10$	$0.238 \pm 0.06$	$0.050 \pm 0.020$

the largest  $D$  values occur with all those comparisons to the Réunion population. Therefore, Réunion appears to be the most differentiated population. As expected from Nei's theory (1971; 1972), these high  $D$  values of the Réunion population are affected by the presence of the fixed alleles at  $Est_1$  and  $Est_2$  and the high frequencies of allele at  $Got_2$  unique to Réunion.

#### Gene flow

An attempt has also been made to infer the role of gene flow in relation to the genetic structure of *C. capitata*

during its dispersal. The gene flow is considered low and not significant if  $Nm < 1$ , i.e. if there is less than one immigrant per generation. The two methods (Wright, 1931; Slatkin, 1985) are expected to give similar estimates of the gene flow ( $Nm$ ). But a low value of gene flow ( $Nm_* = 0.289$ ) is obtained with Slatkin's formula, in contrast to Wright's formula ( $Nm_w = 1.782$ ). As pointed out by Slatkin (1987), his formula is particularly sensitive to the presence of private alleles with high frequency in a particular population; the inclusion of this population could mask the presence of substantial gene flow between the remaining populations. The analysis of different population subsets (Table 6) is therefore used to detect the population responsible for the lowering of this gene flow estimate. As expected, the Réunion population has this effect: by excluding it the highest and the only significant value of gene flow is observed. This further supports the position of Réunion as a geographical isolate, in which a bottleneck effect and/or selection may have contributed to population differentiation. Ignoring Réunion allows us to analyse the high values of gene flow among the remaining three populations. Excluding one of these three populations, very high  $Nm_*$  are obtained between Kenya and Sardinia and

**Table 5** Matrix of genetic distance coefficients (\*) between geographical populations of *C. capitata* (\*\*)

Populations	1	2	3	4
1 Procida	—			
2 Sardinia	0.011	—		
3 Réunion	0.101	0.102	—	
4 Kenya	0.026	0.028	0.107	—

\**D* values (unbiased distance; Nei, 1978).

\*\*Mean intraspecific distance  $\pm$  s.d. =  $0.0625 \pm 0.0452$  (range 0.107–0.011).

**Table 6** Number of private alleles (n.p.a.), average frequencies of private alleles  $\bar{p}(1)$ , and estimates of *Nm*, obtained from different subsets of populations of *C. capitata*

	Population excluded			
	Procida	Sardinia	Réunion	Kenya
n.p.a	20	21	16	14
$\bar{p}(1)$	0.167	0.153	0.073	0.227
<i>Nm</i> *	0.276	0.328	1.420	0.150

between Kenya and Procida, as seen in Fig. 3. The arrows indicate the most probable direction of the gene flow supported by the results obtained in this paper (i.e. that Kenya has the highest estimates of the genetic variability and the greatest number of private alleles). According to our results, the Kenya population has all of the properties of a native population and could belong to the centre of diffusion of this species. The gene flow pattern is in agreement with the geographical colonization of the medfly proposed by Hagen *et al.* (1981).

## Discussion

The results obtained in this study focus attention on the pattern and causes of geographical differentiation in natural populations of *C. capitata*.

Although only 25 loci have been assayed, these may be considered to be a random, although small, sample of the genome. These loci are widely distributed over the genetic map of *C. capitata* (Malacrida *et al.*, 1988); they code for a variety of enzymes with both singular and multiple physiological substrates, and for both monomeric and multimeric enzymes, which contribute differentially to overall heterozygosity (Kojima *et al.*, 1970; Gillespie & Langley, 1974; Zouros, 1976).

In the analysis of single locus differentiation some loci, as seen in Table 2, show a significant amount of genetic differentiation between populations, and others, while highly heterozygous (e.g. *Mpi*), are homogeneous between populations. This would suggest that genetic differentiation between populations is unrelated to genetic variation within populations. But the presence of at least three loci with very high  $H_T$  and high  $F_{ST}$  values suggests the presence of balancing and diversifying selection among populations (Singh *et al.*, 1987). If the differentiation indices are considered separately in relation to each linkage group (Table 3), their estimates provide grounds to believe that linkage constrains the pattern of variability. In fact, the mean fixation index  $F_{ST}$  is about 12 per cent and the highest contribution to this value is given by loci of linkage group A (Chr. 4). Segregation distortion for these loci may be one cause for the observed deviation from panmixia. Similar results have been reported for *Ragholetis completa* (Berlocher, 1984) in which segregation distortion was also proposed as one of the causes of deviation from panmixia. Sex ratio distorter factors ( $Sd_1$ ,  $Sd_2$ ) are present in *C. capitata* (Malacrida *et al.*, 1987; Milani *et al.*, 1989). These factors cause differential segregation to sons and daughters of associated genes, which creates a departure from random equilibrium of genotypic frequencies.  $Sd_1$ , the most common distorter factor, maps to linkage group A and is very tightly linked to the  $Est_1$  and  $Est_2$  loci, at  $26.8 \pm 0.03$  cM from  $Hk_2$  (Gasperi *et al.*, 1986). As pointed out above,  $Est_1$  and  $Est_2$  are the two loci which contribute most to the differentiation between populations, as a consequence of fixed alleles in Réunion. Note that in Réunion  $Sd_1$  is present (A. R. Malacrida, unpublished data). Even if the data do not allow a definitive conclusion a plausible hypothesis is that drift may play a more decisive role in this population, acting on an unbalanced genotypic background and causing differentiation in terms of fixed alleles. But selection or a bottleneck effect cannot be excluded.

When gene frequencies in *C. capitata* are used to estimate the parameters of variability or distances, the dynamic aspects of evolutionary processes, i.e. drift, selection, and gene flow, appear evident. Thus: (i)  $F_{ST}$  estimates reveal the presence of loci that are possible targets of natural selection. The population mainly involved in this process is Réunion. (ii) Genetic distances between Réunion and Kenya are very high, again as a consequence of the fixation of alleles in Réunion. (iii) Similarly, when gene flow is estimated by Slatkin's method, using the frequency of private alleles, the inclusion of the Réunion samples in the dataset considerably changes the gene flow estimates.

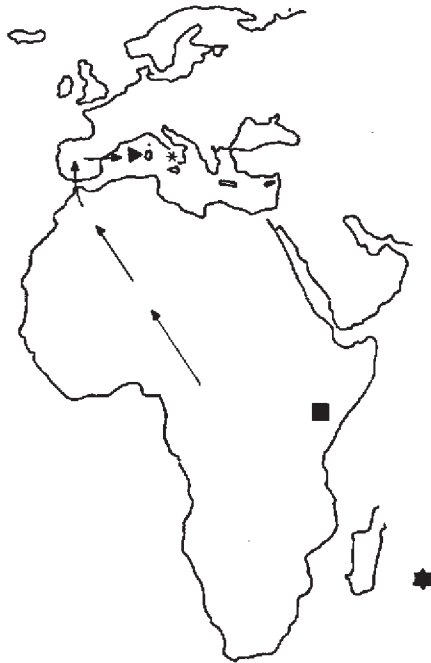
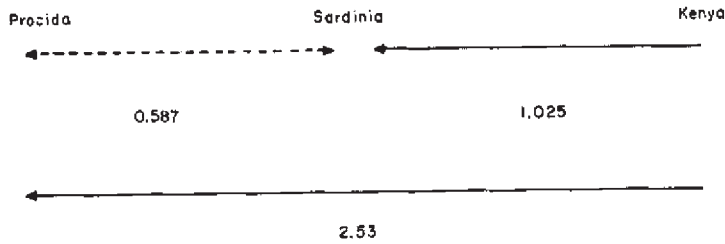


Fig. 3 Quantitative estimates of gene flow ( $Nm_*$ ) among (■) Kenya, (▶) Sardinia and (\*) Procida populations of *Ceratitis capitata*, and a possible route of medfly colonization from Africa to the Mediterranean basin (see also Hagen *et al.*, 1981). ★ = Réunion.

If the Réunion population or the loci of linkage group A are excluded from the computation, then the estimates of  $F_{ST}$  and gene flow more closely resemble those attributable only to forces such as drift and migration. In fact, the mean contribution of the remaining loci to the geographical differentiation is around 8 per cent. Wright's  $Nm_w$ , derived from this  $F_{ST}$  estimate ( $Nm_w = 2.9$ ), is significant as the value of  $Nm_*$  estimated by Slatkin's method, excluding Réunion ( $Nm_* = 1.42$ ); this estimate implies a consistent gene flow between Kenya and the Mediterranean populations.

On the other hand, estimates of variability not involving gene frequencies directly, such as the average number of alleles per locus ( $\bar{A}$ ), the proportion of polymorphic loci ( $\bar{P}$ ) and the average number of heterozygous individuals ( $\bar{H}$ ) seem more related to the populations' evolutionary history. Based on these 'static' parameters, it is observed that the Réunion samples are similar to the Kenyan samples than to the

distant Mediterranean populations. These statistics reveal a trend of decreasing variability between flies from the putative source area (Africa) and those from the periphery of the dispersion area. The African populations of *C. capitata* contain more alleles than the Mediterranean populations; on the other hand all the Mediterranean alleles were found in the two African populations. The greater number of alleles found in the Kenyan population can be explained by its larger population size as expected from a poliphagous tropical population (Fletcher, 1989). The dispersion itself might explain a reduction in variability in the derived Mediterranean populations, as a consequence of a bottleneck in population size (Nei, 1975). This low initial level of variability may be maintained over time because of different environmental constraints, e.g. seasonal changes. Seasonality may affect the demographic profile of a Mediterranean population. The genetic distances found between Kenyan and Mediterranean populations ( $D = 0.026$  and  $0.028$ ) are com-



patible with a relatively recent colonization. Hagen *et al.* (1981) indicate 1842 as the first record of medfly in the Mediterranean basin.

In conclusion, different evolutionary forces could have contributed, singly or in concert, to genetic changes during colonization by *C. capitata*. Selection and/or the bottleneck effect may have played an important role in the differentiation of the Réunion samples, which on the other hand maintain the genetic attributes of an ancestral population. Drift, which is correlated with geographical distance, seems to have played a major role in the dispersion processes of the two Mediterranean populations. On the other hand, gene flow, in terms of the number of immigrants per generation, is significant between the original and derived Mediterranean populations, which supports the hypothesis of a recent colonization, as Hagen *et al.* (1981) pointed out.

### Acknowledgements

We would like to thank M. Ashburner, F. Kafatos and E. Zouros for critical reading and comments on an earlier draft of this manuscript. We are indebted to the following people for providing the specimens studied: D. Lindquist and E. Busch-Petersen (IAEA, Vienna, Austria); T. Mukiyama (University of Nairobi, Kenya); S. Quilici (INRAT, Réunion, France); U. Cirio (ENEA, Rome, Italy) and G. Del Rio (University of Sassari, Italy).

This research was supported by the National Research Council of Italy, Special Project RAISA, Sub-project No. 2, Paper No. 40. Grants from MPI (funds 40 per cent from National Ministry of Education) and from IAEA (International Atomic Energy Agency, Vienna, Austria) also contributed to this work.

### References

- BERLOCHER, S. H. 1984. Genetic changes coinciding with colonization of California by the walnut husk fly, *Rhagoletis completa*. *Evolution*, **38**, 906–918.
- FLETCHER, B. S. 1989. Life history strategies of tephritid fruit flies. In: Robinson, A. S. and Hooper, G. (eds), *Fruit Flies: Their Biology, Natural Enemies and Control*, Vol. 3B. Elsevier Science Publishers, Amsterdam, pp. 195–208.
- GASPERI, G., MALACRIDA, A. R. AND MILANI, R. 1987. Protein variability and population genetics of *Ceratitidis capitata*. In: Economopoulos, A. P., (ed.), *Fruit Flies*. Elsevier Science Publishers, Amsterdam, pp. 149–157.
- GASPERI, G., MALACRIDA, A. R., GUGLIELMINO, C. R. AND MILANI, R. 1990. Electrophoretic multilocus analysis for the study of natural populations of the medfly *Ceratitidis capitata*. In: *Genetic Sexing of the Mediterranean Fruit Fly*, International Atomic Energy Agency, Vienna, Austria, pp. 90–94.
- GASPERI, G., MALACRIDA, A., TOSETTI, M. AND MILANI, R. 1986. Enzyme variability: a tool for investigating the genome organization and the population structure of *Ceratitidis capitata*. In: Cavalloro, R. (ed.), *Fruit Flies of Economic Importance 84*. A. A. Balkema, Rotterdam, pp. 153–161.
- GILLESPIE, J. H. AND LANGLEY, C. H. 1974. A general model to account for enzyme variation in natural populations. *Genetics*, **76**, 837–884.
- HAGEN, K. S., WILLIAM, W. W. AND TASSAN, R. L. 1981. Mediterranean fruit fly: the worst may be yet to come. *California Agriculture (University of California, Division of Agricultural Sciences, Reports on Progress in Research)*, **35**, 5–7.
- HARRIS, H. AND HOPKINSON, D. A. 1976. *Handbook of Enzyme Electrophoresis in Human Genetics*. North-Holland Publishing Company, Amsterdam.
- HUETTEL, M. D., FUERST, P. A., MARUYAMA, M. AND CHAKRABORTY, R. 1980. Genetic effects of multiple population bottlenecks in the Mediterranean fruit fly (*Ceratitidis capitata*). *Genetics*, **94s**, 47–48 (Abstract).
- KOJIMA, K., GILLESPIE, J. H. AND TOBARI, Y. N. 1970. A profile of *Drosophila* species enzymes assayed by electrophoresis. I. Number of alleles, heterozygosities, and linkage disequilibrium in glucose-metabolizing systems and some other enzymes. *Biochem. Genet.*, **4**, 627–637.
- LOUKAS, M. 1989. Population genetics studies of fruit flies of economic importance, especially medfly and olive fruit fly, using electrophoretic methods. In: H. D. Loxdale and J. Den Hollander (eds), *Electrophoretic Studies on Agricultural Pests*. Systematics Association Special Volume No. 39, Clarendon Press, Oxford, U.K., pp. 69–102.
- MALACRIDA, A. R., GASPERI, G., BARUFFI, L., BISCALDI, G. F. AND MILANI, R. 1988. Updating of the genetics of *Ceratitidis capitata* (Wied.). In: *Modern Insect Control: Nuclear Techniques and Biotechnology*. International Atomic Energy Agency, Vienna, Austria, pp. 221–227.
- MALACRIDA, A. R., GASPERI, G. AND MILANI, R. 1987. Genome organization of *Ceratitidis capitata*: linkage groups and evidence for sex-ratio distorters. In: A. P. Economopoulos (ed.), *Fruit Flies*. Elsevier Science Publishers, Amsterdam, pp. 169–174.
- MEERA KHAN, P. 1971. Enzyme electrophoresis on cellulose acetate gel: zymogram patterns in man-mouse and man-chinese hamster somatic cell hybrids. *Arch. Biochem. Biophys.*, **145**, 470–483.
- MILANI, R., GASPERI, G. AND MALACRIDA, A. 1989. Biochemical Genetics (of *Ceratitidis capitata*). In: Robinson, A. S. and Hooper, G. (eds), *Fruit Flies: Their Biology, Natural Enemies and Control*, Vol. 3B. Elsevier Science Publishers, Amsterdam, pp. 33–56.
- NEI, M. 1971. Interspecific gene differences and evolutionary time from electrophoretic data on protein identity. *Am. Nat.*, **105**, 385–398.
- 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.
- NEI, M. AND CHESSEY, 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. eds. 1989. *Fruit Flies: Their Biology, Natural Enemies and Control*, Vol. 3A, B. Elsevier Science Publishers, Amsterdam.
- SAUL, S. H. 1986. Genetics of the Mediterranean fruit fly (*Ceratitis capitata*) (Wiedemann). In: Russel, G. E. (ed.), *Agricultural Zoology Reviews*, Vol. 1. Intercept Ltd, Ponteland, Newcastle upon Tyne, U.K., pp. 73-108.
- SINGH, R. S. AND RHOMBERG, L. R. 1987. A comprehensive study of genetic variation in natural populations of *Drosophila melanogaster*. II. Estimates of heterozygosity and patterns of geographic differentiation. *Genetics*, **117**, 255-271.
- SLATKIN, M. 1981. Estimating levels of gene flow in natural populations. *Genetics*, **99**, 323-335.
- SLATKIN, M. 1985. Rare alleles as indicators of gene flow. *Evolution*, **39**, 53-65.
- SLATKIN, M. 1987. Gene flow and geographic structure of natural populations. *Science*, **236**, 787-792.
- WHITE, I. M. 1989. The state of fruit fly taxonomy and future research priorities. In: Cavalloro, R. (ed.), *Fruit Flies of Economic Importance* 87. A. A. Balkema, Rotterdam, pp. 543-552.
- WORKMAN, P. L. AND NISWANDER, J. D. 1970. Population studies on southwestern Indian tribes. II. Local genetic differentiation in the Papago. *Am. J. Hum. Genet.*, **22**, 24-49.
- WRIGHT, S. 1931. Evolution in Mendelian populations. *Genetics*, **16**, 97-159.
- WRIGHT, S. 1943. Isolation by distance. *Genetics*, **28**, 114-138.
- WRIGHT, S. 1951. The genetical structure of populations. *Ann. Eugen.*, **15**, 323-354.
- ZOUROS, E. 1976. Hybrid molecules and the superiority of the heterozygote. *Nature*, **262**, 227-229.