J Mol Evol (1998) 46:288–298

JOURNAL OF MOLECULAR EVOLUTION

Characterization and Evolution of *mariner* Elements from Closely Related Species of Fruit Flies (Diptera: Tephritidae)

Cristina Torti,[1] Ludvik M. Gomulski,[1] Anna R. Malacrida,[1] Pierre Capy,[2] Giuliano Gasperi[3,*]

[1] Department of Animal Biology, University of Pavia, Piazza Botta 9, I-27100 Pavia, Italy
[2] Laboratoire de Populations, Génétique et Evolution, CNRS, 91198 Gif-sur-Yvette Cedex, France
[3] Institute of Zoology, University of Sassari, via Muroni 25, I-07100 Sassari, Sardinia, Italy

Received: 25 April 1997 / Accepted: 31 July 1997

Abstract. *Mariner* elements were amplified using the polymerase chain reaction from two species of tephritid flies, *Ceratitis rosa* and *Trirhithrum coffeae.* The sequences were ~1.3 kb in length. None of these elements appeared to be functional, as in every case the open reading frame (ORF) was disrupted by the presence of frameshifts or stop codons. These elements, *Crmar1* and *Tcmar1,* are very similar to the *Ccmar1* element previously amplified from the closely related tephritid species *C. capitata* and are members of the mellifera subfamily of *mariner* elements. The phylogeny and pattern of divergence of these elements were examined in relation to the phylogeny of the host species. It is highly probable that the elements were present in the ancestral lineage prior to the divergence of the three species. The copy numbers of the elements within each species are very different, ranging from about 10 in *T. coffeae* to 5,000 in *C. rosa.* The possible mechanisms which determine the copy number of an element in the host genome are discussed.

Key words: *Ceratitis capitata — Ceratitis rosa — Trirhithrum coffeae — mariner —* Transposon

Correspondence to: G. Gasperi
* *Present address:* Department of Animal Biology, University of Pavia, Piazza Botta 9, I-27100 Pavia, Italy; e-mail gasperi@ipv36.unipv.it
Sequence Availability: The DNA sequences of clones *Crmar1.2, Crmar1.4, Crmar1.8, Crmar1.16, Tcmar1.51,* and *Tcmar1.81* have been deposited in the GenBank database and have the accession numbers U88159–U88164, respectively

Introduction

Mariner-like elements (MLEs) have an exceptionally wide distribution in insect species (Robertson and Lampe 1995a; Robertson and McLeod 1993). They can be classified into several distinct subfamilies according to sequence similarities (Robertson 1993). The major subfamilies all appear to be widely distributed among species and any particular species may contain *mariner* elements from different subfamilies (Robertson 1993). An example is provided by the tephritid fruit fly *Ceratitis capitata,* in which members of at least seven subfamilies of *mariner* have been identified (Robertson et al. 1997). Horizontal transmission of *mariner* elements has been proposed as an explanation for the widespread distribution of these elements in different distantly related insects, including species of different orders (Robertson and Lampe 1995b). MLEs have therefore attracted interest because of their potential use for genetic manipulation of insect species with special emphasis on insects of economic importance (Kidwell 1993). The vast majority of MLEs are not functional as they contain multiple inactivation mutations (Capy et al. 1992; Maruyama and Hartl 1991; Robertson and McLeod 1993). The only *mariner* elements demonstrated to be autonomous are all closely related to the functional element *Mos1* (Medhora et al. 1991) and are present only in *Drosophila mauritiana* and *D. simulans* (Capy et al. 1992). The *Mos1* element is also capable of transposition in other nonhost insect species (Lindhom et al. 1993; Lohe and Hartl 1996a) including nondrosophilids (Coates et al. 1995). Analyses of the distribution of *mariner* elements within

the *D. melanogaster* subgroup have made important contributions to our understanding of the evolutionary dynamics of *mariner* elements. Although *mariner* elements are widespread, the distribution pattern and abundance of a particular type of MLE may be quite different in closely related species. It has been proposed that the two main modes of evolution of these elements are horizontal transmission and vertical inactivation, both of which are limited by processes resulting in stochastic loss (Capy et al. 1993; Lohe et al. 1995).

We are studying the population biology and molecular evolution of *mariner* elements in related species of Tephritidae flies. A putative functional full-length *mariner* element, *Ccmar1,* was recently discovered in the medfly *C. capitata* (Gomulski et al. 1997). *Ccmar1* is a member of the mellifera *mariner* subfamily, but it is different from the mellifera sequences previously amplified from the medfly genome by Robertson and McLeod (1993). It represents a peripheral lineage of the mellifera subfamily. Within *C. capitata,* the *Ccmar1* element appears to have been evolving neutrally. It is present at high copy number in the genomes of both ancestral and derived populations as elements homogeneous in size, with very few deleted copies. Stochastic reduction in the copy number has not been observed during the very recent colonization process of the medfly (Gomulski et al. in press). Here we report that this element is also present in two other tephritid species *Ceratitis* (*Pterandrus*) *rosa* and *Trirhithrum coffeae,* closely related to *C. capitata,* within the Ceratitini tribe of the Ceratitinae subfamily (White and Elson-Harris 1992). These three tephritid species share the same source area, southeast Africa, where they live in sympatry, but they have different life histories and different geographic dispersion patterns (Fletcher 1989; White and Elson-Harris 1992). Sympatric native populations of *C. capitata, C. rosa,* and *T. coffeae* exhibit different degrees of genetic variability, which appears to be correlated to the zoogeography and to the biological traits of these three species (Malacrida et al. 1996).

Sequence comparisons and determination of the copy numbers of *Ccmar1*-related *mariner* elements in *C. capitata, C. rosa,* and *T. coffeae* are presented. The evolutionary biology and dynamics of these *mariner* elements within and between the host species are discussed.

Materials and Methods

Species Samples. Three wild samples of *C. rosa,* one wild sample of *T. coffeae,* and one wild sample of *C. capitata* were analyzed. The three samples of *C. rosa* were from Réunion (collected in St. Denis in 1994), from Mauritius (in 1993), and from Kenya (1995), respectively. The samples of *T. coffeae* and *C. capitata* were from Kenya. All the Kenyan samples of these three species were from sympatric populations and were collected together in Ruiru, from coffee berries in 1995.

PCR Amplification, Cloning, and Sequencing of mariner *Elements.* The *mariner* elements were amplified and sequenced from one individual chosen at random from each of the samples. DNA preparations from single flies were performed as described by Baruffi et al. (1995). Following treatment with RNase A, the DNA was extracted with phenol/chloroform, precipitated with ethanol, and resuspended in TE (10 mM Tris-HCl, pH 8.0, 1 mM EDTA). The *mariner* elements from *T. coffeae* and *C. rosa* were amplified using a primer corresponding to the complete 30 bp ITR of the *postdoc* element found in the Adh_l gene of *C. capitata* (Brogna et al. 1994; Gomulski et al. 1997): 5′ TTGGATGAGTGCATAAGTTCGTGCCCGATT 3′ and a shorter primer corresponding to the external 21 bp of the ITR, respectively. PCR with these primers enabled us to amplify entire *mariner*-like elements. The amplification conditions were those of Gomulski et al. (1997). PCR amplification products were resolved by electrophoresis on 1.2% agarose gels (SeaKem GTG, FMC) with TBE buffer containing 0.5 mg/ml of ethidium bromide. PCR products of ~1.3 kb were eluted from the gel using the QIAquick Gel Extraction Kit (Qiagen). The purified DNA fragments were ligated into the *Sac* II site of the pGEM-T vector (Promega). Positive colonies were selected and the size of the insert checked by PCR. Sequencing of the PCR products was performed using an Applied Biosystems model 373A DNA sequencing system and the PRISM Ready Reaction DyeDeoxy Terminator Cycle Sequencing Kit. Sequencing of the insert was achieved using T-7 and Sp-6 specific primers and nine additional internal primers.

Ceratitis capitata mariner *Clones.* The sequences of the following *mariner* clones from *C. capitata* (Gomulski et al. 1997) were used for comparative purposes: *Ccmar1.4* (U76903), *Ccmar1.7* (U76904), *Ccmar1.13* (U76905), and *Ccmar1.18* (U40493). The clone *Ccmar1.18* which is a complete full-length element, was used for hybridization analysis.

Hybridization and Dot-Blot Analyses. Genomic DNA (500 ng) from individual flies of *C. capitata, C. rosa,* and *T. coffeae,* collected in Kenya, was digested with *Xba* I and *Sal* I, which do not cleave within the *Ccmar1* element. Electrophoresis, Southern blot, and hybridization with labeled clone *Ccmar1.18* insert were as described in Gomulski et al. (1997). The copy numbers of the elements in *C. capitata, C. rosa,* and *T. coffeae* were determined by dot-blot analysis. Genomic DNAs of wild single flies from the three species were prepared as previously described. The concentration of the DNA was determined using a Beckman DU650 spectrophotometer. Serial dilutions of the genomic DNA (from 500 to 20 ng) from three individuals of each species and two replicates of the *mariner* clone *Ccmar1.18* insert DNA (from 50 ng to 10 pg) were manually spotted onto a positively charged membrane using the protocol in Ausubel et al. (1994). The filter was then probed with digoxigenin-labeled clone *Ccmar1.18* insert DNA at 68°C as previously described. The intensities of the signals on the exposed X-ray film were evaluated using a Molecular Dynamics Personal Densitometer. The procedure was repeated several times with different dilution series and exposure times to confirm the copy number estimates. Comparison of the linear portions of the regression lines of DNA concentration against optical density for the control *Ccmar1.18* insert DNA and the genomic DNA samples of the three species permitted the estimation of the portion of each genome which consists of *mariner* elements. The copy number in each species could then be calculated as the estimates of the genome sizes and the size of the *mariner* elements are known.

Phylogenetic Analysis. Sequence comparisons of the *C. rosa* and *T. coffeae* elements with other known sequences, including those from *C. capitata,* were performed using the BLAST family of programs from NCBI (Altschul et al. 1990), version 1.4.9MP. Sequences were aligned initially using Clustal W 1.6 (Thompson et al. 1994) and then adjusted manually to optimize the alignment. Phylogenetic trees were inferred using maximum parsimony analysis with the heuristic search option together with TBR branch swapping (PAUP version 3.1.1, Swofford 1993). Gaps were treated as missing characters. In the case of

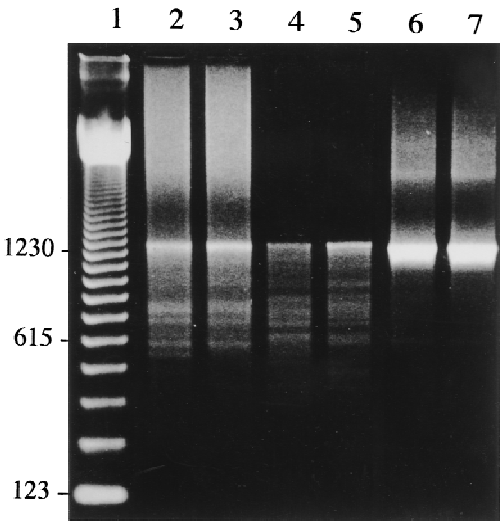

Fig. 1. PCR amplification of *mariner* elements using the ITR primers of *Ccmar1* from *C. capitata* (*lanes 2, 3*), *T. coffeae* (*lanes 4, 5*), and *C. rosa* (*lanes 6, 7*) genomic DNA. *Lane 1* contains a 123-bp molecular weight ladder (Gibco-BRL).

amino acid sequence data, stop codons and frame shifts were treated as extra characters. To investigate the phylogenetic relationships at the nucleotide level within tephritid *mariner* elements, 1,000 bootstrap replications were performed. The molecular evolution of the coding region of the sequenced elements was examined using the MEGA computer program (Kumar et al. 1993).

Results

Sequence Analysis

PCR amplification on the total genomic DNA from *C. rosa* and *T. coffeae* resulted in 1.3-kb products, the same size as the *C. capitata* full-length *mariner Ccmar1* (Gomulski et al. 1997), and in a number of bands corresponding to deleted elements. Figure 1 shows the amplification from individuals of the three species samples which are sympatric in Kenya; it is evident that the intensities of the 1.3-kb fragments are different among the considered species. The 1.3-kb PCR products from *C. rosa* and *T. coffeae* were cloned and six clones were sequenced: one from *C. rosa* from Kenya (clone 8), two from *C. rosa* from Mauritius (clones 2 and 4), one from *C. rosa* from Réunion (clone 16), and two from *T. coffeae* (clones 51 and 81), both from Kenya. According to the nomenclature proposed by Robertson and Asplund (1996) the elements from *C. rosa* and *T. coffeae* will be named *Crmar1* (*C. rosa mariner 1*) and *Tcmar1* (*T. coffeae mariner 1*), respectively. The sequences of the four clones *Crmar1.2, Crmar1.4, Crmar1.16,* and *Crmar1.8* were found to be 1,281, 1,266, 1,270, and 1,275 bp in length, respectively. The two clones *Tcmar1.51* and *Tcmar1.81* were found to be 1,289 and 1,277 bp in length, respectively.

The consensus sequence of the four *C. rosa* clones and the conceptual translation of the encoded putative transposase are shown in Fig. 2. Two ambiguities remain: one at position 31 (Y) where two of the clones contain a cytosine base, and the other two a thymine base; the second at position 230 (R), in the coding region, where two clones had an adenine, while the other two had a guanine base. These alternative bases represent nonsynonymous substitutions, coding, respectively, for isoleucine (Ile) and valine (Val) amino acids. Another ambiguity, at position 1,226, was resolved using another partially sequenced cloned. The complete consensus sequence, *Crmar1,* is 1,284 bp in length. The conceptual open reading frame (ORF), encoding 342 amino acids, is not continuous, as it contains three frameshifts at positions 552, 598, and 937. Comparison with the *Ccmar1* consensus sequence (Gomulski et al. 1997) indicates that these frameshifts are due to small insertions at positions 552 and 937 and a deletion at position 598.

The two sequences from *T. coffeae* appear to be somewhat divergent. *Tcmar1.51* appears to be very heavily mutated in comparison with the other sequences from *T. coffeae, C. rosa,* and *C. capitata.* This, together with the low number of clones sequenced from *T. coffeae,* did not permit the determination of a consensus sequence for this species.

All six sequences contain the expected features of *mariner* transposable elements: They are flanked by two short inverted terminal repeats and they contain relatively long 5′ and short 3′ noncoding regions. *Crmar1* and *Tcmar1.51* also contain the D,D,34,D catalytic domain motif proposed by Doak et al. (1994) and Robertson (1995), while in the *Tcmar1.81* sequence the third D is replaced by a glycine (G) (see Fig. 3). Other features present in the *Crmar1* and *Tcmar1* clones are putative TATA boxes and polyadenylation signals.

Sequence Comparisons of Ccmar1, Crmar1, *and* Tcmar1

The six elements from *C. rosa* and *T. coffeae* were compared with the individual clones from *C. capitata* previously described by Gomulski et al. (1997). All the sequences were very similar at the nucleotide level. The percentage nucleotide identity between the clones of each element ranged from 94.9 to 95.9% in *C. capitata,* from 93.2 to 96.3% in *C. rosa,* and 90.9% in *T. coffeae.* The lower level of identity found in *T. coffeae* is due to clone *Tcmar1.51,* which appears to be very heavily mutated in comparison with *Tcmar1.81* and with respect to the clones from the other species.

In the interspecific comparisons the *Ccmar1* and *Crmar1* clones show identities of between 93.6 and 96.5%, very similar to the within-species values. Interspecific comparisons with the *Tcmar1.81* are in the same range, 94.5–95.3%, when compared to *Ccmar1,* and 94.4–96.4% when compared to *Crmar1.* The compari-

```
ttggatgagtgcataagttcgtgcccgattycgctggatgccgtacgaatcgtttgagag   60
tggcgctcgtgaaagaaatatatacatattatcatatcgttggaaaggtgacagtccgaa  120
ctgtaaatcaagcataaaataacttcattttgatttgagttgcttagtgaaaagacgcat  180
ggaagATGAACAACGAAAAAGATCATATGCGTCATATTATGTTATACGAATTCCRCAAAG  240
      M  N  N  E  K  D  H  M  R  H  I  M  L  Y  E  F  R  K    18

GAAAAACAGTGGGCGCTGCAACTAAAGATATTCGCGAAGTTTATTTGGACCGTGCTCCAG  300
G  K  T  V  G  A  A  T  K  D  I  R  E  V  Y  L  D  R  A  P    38

CACTCCGCACAGTAAAGAAATGGTTCGCGAAATTTCGTTCTGGAGATTTTAACCTCGAAG  360
A  L  R  T  V  K  K  W  F  A  K  F  R  S  G  D  F  N  L  E    58

ATCAACCTCGCAGTGGACGGCCTTCTGAGCTTGATGACGATGTTCTAAGGACTTTAGTTG  420
D  Q  P  R  S  G  R  P  S  E  L  D  D  D  V  L  R  T  L  V    78

CGAATAACTCACGTATTTCGACGGAAGAGGTTGCCAGTGAATTGAACGTCAACAAATCAA  480
A  N  N  S  R  I  S  T  E  E  V  A  S  E  L  N  V  N  K  S    98

CTGCGTTTCGTCGTTTAAAAAAGGTTGGGTACACTTTGAAGCTCGATACATGGGTGCCAC  540
T  A  F  R  R  L  K  K  V  G  Y  T  L  K  L  D  T  W  V  P   118

ATCAGTTGAGTGAAAAAGAACAAAGTGGACCGTATGTCAACAGCAATTTCTTTGCTTCGA  600
H  Q  L  S  E  K  #  N  K  V  D  R  M  S  T  A  I  S  L  L  R   138

CGGATCAAAACGAACCTTTTTTGGATCGGCTCGTGACTGGTGATGAAAAATGGGTCCTGT  660
 R  #  Q  N  E  P  F  L  D  R  L  V  T  G  D  E  K  W  V  L    158
                                          *

ACAACAATGTTCAACGCAAAAGAACATGGAAACAGGCACACGAAGGGGCGGAACCGATGT  720
Y  N  N  V  Q  R  K  R  T  W  K  Q  A  H  E  G  A  E  P  M   178

CGAAAGGTGGATTGCATCCGATGAAGGTACTGCTGTGCATTTGGTGGGATATCCGAGGCG  780
S  K  G  G  L  H  P  M  K  V  L  L  C  I  W  W  D  I  R  G   198

TGATCTATTTTGAGCTCTTGCCAGCTGGAGAAACGATCACTGCCAACAAGTATTGTCAGC  840
V  I  Y  F  E  L  L  P  A  G  E  T  I  T  A  N  K  Y  C  Q   218

AATTGGTCGAATTGAAGAAAGCAATTGATGAAAAACGTCCGATTTTGGCCAATCGCAAAG  900
Q  L  V  E  L  K  K  A  I  D  E  K  R  P  I  L  A  N  R  K   238

GAGTTCTTTTCCATCATGACAACGCCAGGCCACATGTTGCCAAAACCGACCCTGGCCAAA  960
G  V  L  F  H  H  D  N  A  R  P  H  V  A  #  K  P  T  L  A  K   258
                  *

CTGAAGGAGATGAATTGGGAAATCATGCCGCATCCCCCATATTCACCCGACATTGCACCT 1020
 L  K  E  M  N  W  E  I  M  P  H  P  P  Y  S  P  D  I  A  P   278

TCTGATTATCATTTGTTTCGATCGCTGCAGAACAATTTGAATGGAAAAAAATTTAAAAAT 1080
 S  D  Y  H  L  F  R  S  L  Q  N  N  L  N  G  K  K  F  K  N   298
    *

GTGGAAGACGTCAAAAACCACCTTGACACCTTTTTCAACGAGAAACCGCGCGATTTCTAT 1140
 V  E  D  V  K  N  H  L  D  T  F  F  N  E  K  P  R  D  F  Y   318

GAATCAGGCATCCGTAAATTGGTTGAACGTTGGGAGTGGATTGCCGAACATGATGGCGAA 1200
 E  S  G  I  R  K  L  V  E  R  W  E  W  I  A  E  H  D  G  E   338

TACATAATTGATtaataaaagcgctttcttccaaaaatttcaattttagtttgcacttga 1260
 Y  I  I  D  .                                                 342

aatcgaacttatgcactcatccaa                                      1284
```

Fig. 2. DNA sequence of the *Crmar1* element from *C. rosa* (consensus of four copies) and the corresponding conceptual amino acid translation. The ITRs are *underlined*, the putative TATA box and polyadenylation signal are shown in *bold*, and the D,D,34,D catalytic domain residues are indicated by *asterisks*. The termination codon is indicated by a dot.

sons with *Tcmar1.51* are lower, 90.1–91.0% with *Ccmar1* and 90.4–92.1% with *Crmar1*. The clone *Tcmar1.81* is more closely related to the *Ccmar1* and *Crmar1* clones than to the *Tcmar1.51* element. Among the considered *mariner* elements, *Tcmar1.51* is the most differentiated, while *Crmar1.16* (from Réunion) shares the highest identities when compared with all the other clones (95.2–96.5%).

Using the shorter, 21-bp, primer it was possible to detect differences in the internal part of the ITRs of the *C. rosa* clones. Both the ITRs of *Crmar1.4* appear to have a deletion of six bases (from position 22 to position 27) and a mismatch between the most internal base of the two ITRs. The other three *Crmar1* sequences all possess the same deletion as *Crmar1.4* at the 3′ ITR of the element; however, in the case of *Crmar1.16*, the deletion is of five rather than six bases. These three clones share the same 30-bp 5′ ITR as found in *postdoc*, *Ccmar1*, and in the *Tcmar1* clones.

An alignment of the conceptual amino acid sequences of the consensus *Ccmar1*, *Crmar1* and the individual *Tcmar1* elements is shown in Fig. 3. In all the elements sequenced from *C. rosa* and *T. coffeae*, the ORFs are not continuous, due to the presence of insertions and deletions which result in frameshifts and thus truncated transposases. The conceptual translations of clones *Crmar1.2*, *Crmar1.4*, *Crmar1.8*, and *Crmar1.16* are 331, 344, 332, and 337 amino acids long, respectively. The *Tcmar1.51* and *Tcmar1.81* sequences code for 337 and 333 amino acids, respectively.

The highest amino acid identity/similarity was observed in the comparison between *Ccmar1* and *Crmar1*

```
Ccmar1     MDNEKDHMR-HIMLYEFRKGKTVGAA-TKDIR-EVYLDRAP-ALR-TVKK   50
Crmar1     MNNEKDHMR-HIMLYEFRKGKTVGAA-TKDIR-EVYLDRAP-ALR-TVKK
Tcmar1.81  MNNEKDHMR-HIMLY#FRKGKTVGA-#TKDIR-EVYLDRAP-ALR#IVKK
Tcmar1.51  MINDKDHMP#HIMSYEFRKGK#AGAA#SKDIR#EVLIWSVPPALP-HRKV
           * *+****  *** * *****  **  +**** ** +   * **    *

Ccmar1     W-FAKFRSGDFNLEDQPR--------SGRPSELDDDVLRTLVANNSRIST  100
Crmar1     W-FAKFRSGDFNLEDQPR--------SGRPSELDDDVLRTLVANNSRIST
Tcmar1.81  W#FAKFRSGDFNLEDQPR--------SGRPSELDDDVLRTLVANNSRIST
Tcmar1.51  N-WLPLKFSLFEISS-PRRITPPSVVRSLPSELDDDVLRTLVANNSRIST
                 +   * +   **         + ********************

Ccmar1     EEVASELNVNKSTAFR-RLKKVGYTLKLDTWVPHQLSEK-NKVDRMSTAI  150
Crmar1     EEVASELNVNKSTAFR-RLKKVGYTLKLDTWVPHQLSEK#NKVDRMSTAI
Tcmar1.81  E----#LNVNKSTAFR-RLKKVGYTLKLDTWVPHQLSEK-EQSGPYVNSK
Tcmar1.51  EEVASELNVNKSTAFR#RLKKVGYTLKLDTWMPHQLSEK-NKV#DSVTAV
           *     ********** ************** *******  +      +

Ccmar1     SLLRRI--KNEPFLDRLVTGDEKWILYNNVQRKRTWK-QAHEGA-EPMSK  200
Crmar1     SLLRR--#QNEPFLDRLVTGDEKWVLYNNVQRKRTWK-QAHEGA-EPMSK
Tcmar1.81  FLLLRTGSKTKPFLDRLVTGDEKWVPYNNVQRKRT+K-QAHEGA-EPMSK
Tcmar1.51  SALRRII--NDPSWERLVT-DEKWVLYNMSTQK--NM.QAHEGA#EPMSR
             * *      *  +**** **** + **   +*     ****** ****+

Ccmar1     GGLHPMKVLLCIWWDIRGVIYFELLPAGETITANKYCQQLVELKKAIDEK  250
Crmar1     GGLHPMKVLLCIWWDIRGVIYFELLPAGETITANKYCQQLVELKKAIDEK
Tcmar1.81  GGLHPMKVLLCIWWDIRGVIYFELLPAGETITANKYCQQLVELKKAIDET
Tcmar1.51  GGLHQMKVLLCIRWDIRRVIF#---AAGETITANKYCQQLAN.RKSIDEK
           **** *******+**** **       **************   +*+***

Ccmar1     RPILANRKGVLFHHD-NARPHVAKPTLAK-LKEMNW-EIMPH-PPYSPDI  300
Crmar1     RPILANRKGVLFHHD-NARPHVA#PTLAK-LKEMNW-EIMPH-PPYSPDI
Tcmar1.81  RPILANRKGVLFHHD-NARRHVAKPTLAK#LKEMNW#EIMPH#PPYSP#I
Tcmar1.51  RPIVP-IAGSLFHHDD#ARPTCCEPTLAK-TEEVNW-EIMPH-PPYSPDI
           ***+    * *****  **     *****   *+** ***** ***** *

Ccmar1     APSDYHLFR-SLQNNLNGKKFKNVEDVKNHLDTFFNEKPRDFYESGIRKL  350
Crmar1     APSDYHLFR-SLQNNLNGKKFKNVEDVKNHLDTFFNEKPRDFYESGIRKL
Tcmar1.81  GPFGYHLFR#PLLNNLNGK#FKNVEDVKNHLDTFF#EKPRDFYESGIRKL
Tcmar1.51  APSDYHLFR-SLQNNLNGL-FKNVEDV.NHLDTFFNEKPRDFYGSGIRKL
           +*  *****  * *****  ******* ******* ******* ******

Ccmar1     VERWEWIAEHDGEYIID.                                  367
Crmar1     VERWEWIAEHDGEYIID.
Tcmar1.81  VERWERIAEHDGEYIID.
Tcmar1.51  VERWEWIAEHDGEYIID.
           *****+***********
```

Fig. 3. Alignment of the *Ccmar1* putative transposase with conceptual *mariner* transposases from the consensus sequence of *Crmar1* and the *Tcmar1.51* and *Tcmar1.81* clones. *Dots* and # signs in the sequences indicate stop codons and framshifts, respectively. The D,D,34,D motif is shown in *bold.* The symbols * and + below the sequences indicate identities and conservative replacements, respectively.

(99.1/99.7%, respectively). Similar identities/similarities were found between *Tcmar1.81* and *Ccmar1* (93/94.5%) and between *Tcmar1.81* and *Crmar1* (92.4/93.9%). The two *Tcmar1* clones showed 70% amino acid identity and 75% similarity. The lowest identities/similarities were observed in the interspecific comparisons with clone *Tcmar1.51:* 77.2/83.3% and 67.2/82.3% identity/similarity with *Ccmar1* and *Crmar1,* respectively.

Pattern of Evolution

The mode of evolution of *mariner* elements within a genome can be deduced by comparing the frequencies of synonymous (silent) and nonsynonymous (replacement) nucleotide substitutions in the transposase coding regions of several copies.

The Jukes-Cantor (Jukes and Cantor 1969) corrected proportions of the synonymous (K_S) and nonsynonymous (K_N) substitutions in the transposase coding regions of the clones were computed according to the method of Nei and Gojobori (1986) using the computer program MEGA (Kumar et al. 1993). For interspecific comparisons the two individual sequences of *Tcmar1* and the two consensus sequences *Ccmar1* and *Crmar1* were used. Table 1 shows the values of K_S and K_N calculated in intraspecific and interspecific comparisons of the elements. The ratios of synonymous and nonsynonymous substitutions in the clones of *Crmar1* (K_S/K_N = 1.50) and *Tcmar1* (K_S/K_N = 1.49) are close to 1, which suggests that these elements are evolving neutrally. A similar result was observed for the *Ccmar1* elements (K_S/K_N = 1.19, Gomulski et al. 1997). In the interspecific comparisons the value of K_S/K_N was 1.81 for the *Ccmar1/Crmar1* comparison, 1.08 for the *Ccmar1/Tcmar1* comparison, and 1.25 between *Crmar1* and *Tcmar1.* To assess the significance of these values, G-tests of independence with Williams' correction for continuity (G_{adj}) were performed (Sokal and Rohlf

Table 1. Proportions of synonymous (K_S) and nonsynonymous (K_N) changes (with Jukes-Cantor correction) between the elements from *C. capitata, C. rosa,* and *T. coffeae*[a]

Comparisons	K_S	K_N	K_S/K_N	G_{adj}[b]	*P*
Within species					
Ccmar1	0.050	0.042	1.19	0.10	n.s.
Crmar1	0.087	0.058	1.50	1.19	n.s.
Tcmar1	0.122	0.082	1.49	0.43	n.s.
Between species					
Ccmar1/Crmar1	0.049	0.027	1.81	0.14	n.s.
Ccmar1/Tcmar1	0.050	0.046	1.08	0.04	n.s.
Crmar1/Tcmar1	0.084	0.067	1.25	0.02	n.s.

[a] For the between species comparisons the consensus sequences from *Crmar1* and *Ccmar1* were used. The consensus sequence of *Ccmar1* is taken from Gomulski et al. (1997)

[b] G-test of independence with Williams' correction

1981). Each nucleotide substitution was classified according to its effect on the amino acid sequence (replacement silent) and according to its status in the different clones (fixed or polymorphic). In every comparison no significant level of departure from parity between synonymous and nonsynonymous changes was observed (Table 1). It appears, therefore, that these sequences have been evolving neutrally in the three species.

Phylogeny of mariner *Elements of Tephritid Species*

The sequences of the four clones from *C. rosa* and the two from *T. coffeae* were compared with the four clones from *C. capitata* (Gomulski et al. 1997). Maximum parsimony analysis yielded nine trees of equal length, 946, with consistency indices of 0.924. The tree shown in Fig. 4 represents the consensus tree of 1,000 bootstrap replicates of the original data set. The tree was rooted using *Dugesia tigrina, Dtmar1* as an outgroup. The placement of *Dtmar1* is uncertain as it can cluster in the mellifera, in the cecropia, or at the base of the mauritiana subfamily of *mariners* (Robertson et al. 1997).

In the tree, the *Ccmar1* clones are grouped together while the *Crmar1* and *Tcmar1* sequences do not form separate clusters. With reference to the *Ccmar1* clones, the *Tcmar1.51* and *Crmar1.4* elements are the most differentiated, with branch lengths of 78 and 49, respectively. The least differentiated sequence is *Crmar1.16* (branch length 10). The low divergence between the 10 elements at each node reflects the high level of nucleotide identity already observed. Phylogenetic investigations using *Mos1* from the mauritiana subfamily of *mariner* elements, *Hsmar1* (U52077, Robertson et al. unpublished) from the cecropia subfamily, and *Gpmar1* (U13808, Blanchetot and Gooding 1995) from the mellifera subfamily as outgroups gave very similar tree topologies.

Comparisons With Other mariner *Elements*

The BLASTIN algorithm was used to search the DNA sequence databases. The *Crmar1* consensus and the *Tcmar1* clone sequences showed high identities (~93.1%) with the *C. capitata Ccmar1.18 mariner* element (Gomulski et al. 1997; U40493), 65.9% with *Forficula auricularia mariner* clone 5.9 (Robertson 1993; L10479), and 60–61% with the two *Apis mellifera mariner* clones 4.2 (Robertson 1993; L10430) and TnM1a (Ebert et al. 1995; U19902), all of which are members of the mellifera subfamily of *mariner* elements.

Also at the amino acid level, using the central transposase sequences, all three elements, *Crmar1, Tcmar1.51,* and *Tcmar1.81,* showed the highest identities/similarities with members of the mellifera *mariner* subfamily.

Phylogenetic Relationships with Other mellifera Subfamily Elements

Using only the internal conserved region of the transposase, we extended these results to a detailed phylogenetic analysis including the *Ccmar1, Crmar1,* and *Tcmar1* individual sequences and 54 other *mariner* elements from the mellifera subfamily (Robertson et al. 1997) (Fig. 5). The latter include elements from the medfly (clones 25.3, 25.4, 25.12) and from another tephritid, the Caribbean fruit fly *Anastrepha suspensa* (clone 32.2). The data were analyzed by the maximum parsimony method using the *Bombyx mori mar1* from the mori subfamily as an outgroup (Robertson et al. 1997). The *Bmmar1* sequence was chosen as the outgroup as this element has been shown to represent a basal lineage of the *mariner* family (Robertson and Asplund 1996). Due to the large size of the data set, the heuristic algorithm was used together with TBR branch swapping. An arbitrary representative of more than 9,800 equally parsimonious trees of length 1,807 and consistency index of 0.573 is shown in Fig. 5. The *Ccmar1, Crmar1,* and *Tcmar1* clones form a tight cluster, reflecting the high amino acid identity of these sequences on the periphery of the mellifera subfamily. The *Crmar1* and *Tcmar1* sequences do not form clusters according to the hosts' classification: *Tcmar1.81* is grouped with *Crmar1.4* and *Crmar1.16,* while *Tcmar1.51* is grouped with Crmar1.8 and *Crmar1.2.*

Our clones do not appear to be closely related to the other *mariner* elements from tephritids included in the tree. The highest identities were observed with medfly 25.3 (34.7–48.7%), medfly 25.12 (36.3–48.2%), and the Caribbean fruit fly 32.2 (35.2–47.9%). In each comparison, again, the highest identities were with *Ccmar1* and *Crmar1* and the lowest with *Tcmar1.51.*

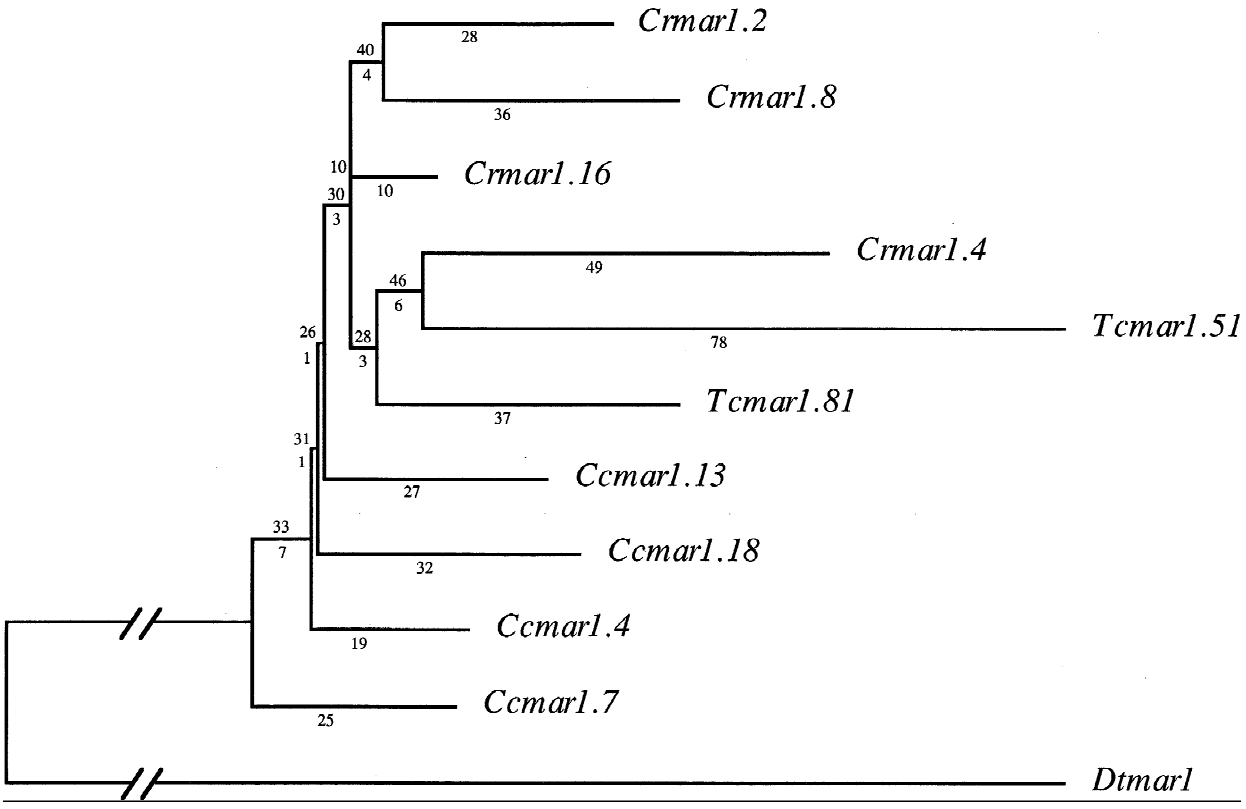

Fig. 4. Phylogeny of the elements based on their nucleotide sequences. Consensus tree of nine equally parsimonious trees of length 946, with consistency indices of 0.924. Bootstrap values above the branches and nucleotide changes below the branches.

Copy Number per Genome

Genomic DNAs from wild sympatric specimens of *C. capitata, C. rosa,* and *T. coffeae* from Kenya were used. Southern hybridization of the labeled *Ccmar1.18* probe to *C. rosa* and *T. coffeae* genomic DNA resulted in extensive hybridization to *C. rosa* and no detectable hybridization to *T. coffeae* DNA. The *C. rosa* samples produced a smear suggesting a very high copy number of *mariner* elements distributed randomly in the genome. The hybridization signal with the *C. rosa* DNA was many times more intense compared with *C. capitata* DNA, suggesting that *C. rosa* contains more copies of the element.

To estimate the copy number of these elements in the genome of *C. rosa* and *T. coffeae,* a dot blot of genomic DNA from single flies was probed at high stringency with labeled *Ccmar1.18* probe. The intensity of the hybridization signal on the exposed X-ray film was evaluated densitometrically. We estimate that the *mariner* elements make up 1.12% of the genome of *C. rosa.* The genome size of this species is approximately 5.8×10^8 bp (A. R. Malacrida in preparation); therefore, 6.5×10^6 bp consists of *mariner* sequences. Dividing by the length of an intact element (~1,285 bp) yields a copy number of approximately 5,000. For *T. coffeae* we estimated the elements make up 0.002% of the genome. Given a genome size of 5.3×10^8 bp (A. R. Malacrida in preparation) the copy number is estimated to be about 10.

Discussion

Using a PCR primer specific to the ITR sequence of a deleted *C. capitata mariner* element, we isolated four copies of a related *mariner* element from *C. rosa* (*Crmar1*) and two from *T. coffeae* (*Tcmar1*). When compared to the four *Ccmar1* elements previously isolated from *C. capitata* using the same PCR primer (Gomulski et al. 1997), the nucleotide sequence identities of the *Crmar1* and *Tcmar1* elements are very high. The *Crmar1* and *Tcmar1* elements share the same basis structure previously identified in *Ccmar1:* They are about 1.3 kb in length and contain the remnants of a putative, though discontinuous, ORF. All but *Tcmar1.81* contain conserved features typical of *mariner* elements, such as the D,D,34,D motif (Doak et al. 1994; Robertson 1995). When compared to the sequences available in the databases, all the elements showed highest similarity to members of the mellifera subfamily of *mariner* elements, particularly when only the central coding regions were considered. Gomulski et al. (1997) previously demonstrated that the *Ccmar1* elements formed a cluster at the

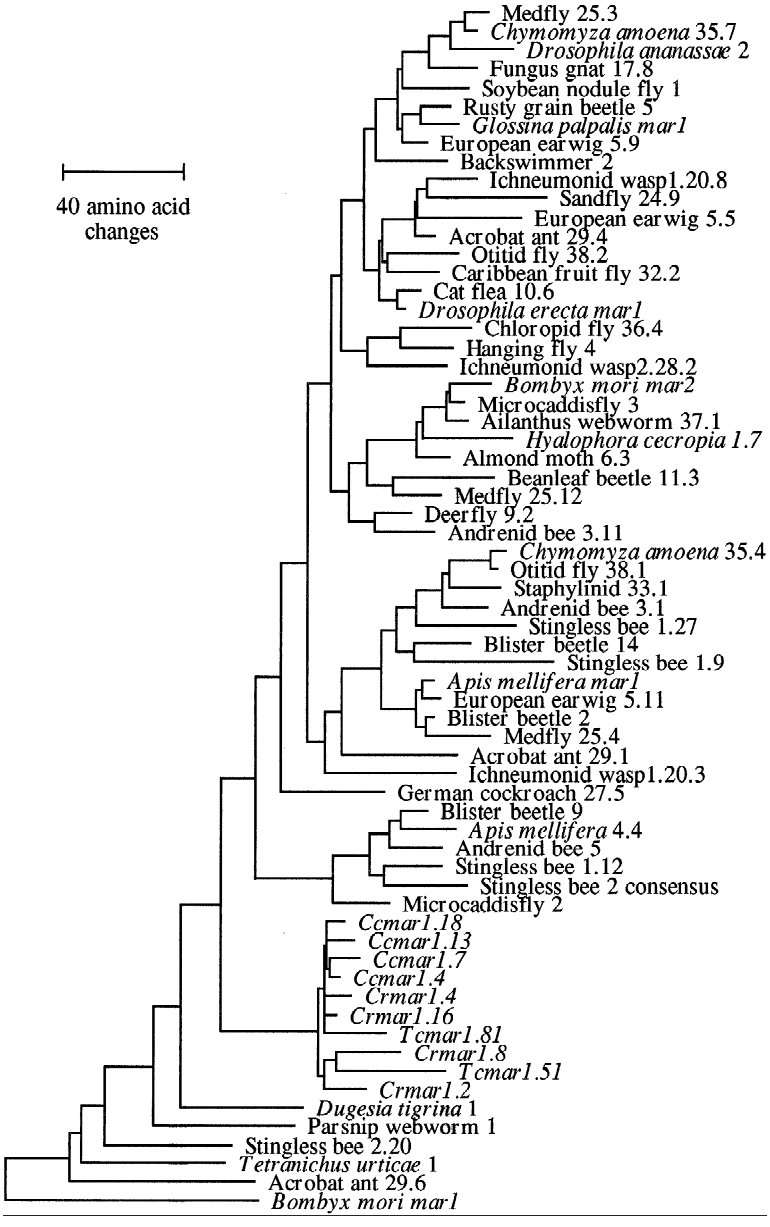

Fig. 5. A representative of the more than 9,800 equally parsimonious trees of 58 *mariner* transposase amino acid sequences, rooted with the *Bombyx mori Bmmar1* sequence. All the sequences are members of the mellifera subfamily after Robertson's classification. The tree has 1,807 steps, with a consistency index of 0.573.

periphery of the mellifera subfamily. We refined this phylogenetic analysis by comparing the *Ccmar1, Crmar1,* and *Tcmar1* elements with 54 members of the mellifera subfamily (Robertson et al. 1997), many of which were not previously available. The placement of our elements, in a tight cluster, near the periphery of the subfamily is confirmed.

Differentiation of Ccmar1, Crmar1, *and* Tcmar1

The proportion of elements examined in relation to the total number present was very different in the three species; two out of 10 in *T. coffeae,* four out of 5,000 in *C. rosa,* and four out of 500 in *C. capitata* (Gomulski et al. 1997). These differences in the relative number of the elements examined could have influenced the estimated level of diversity within and between the *Ccmar1, Crmar1,* and *Tcmar1* elements. However, our isolation strategy was designed to identify *Ccmar1*-like elements as it selected only those elements which shared common ITR sequences. All of the 10 examined sequences, although from different species, share more than 90% DNA identity. On this basis we can hypothesize that the related elements *Ccmar1, Crmar1,* and *Tcmar1* might not be highly divergent at the intraspecific level, and consequently our data may adequately represent the true diversity of these elements. These elements showed no close relationships with any previously published medfly and tephritid elements (Robertson et al. 1997).

In the phylogenetic analysis, based on the nucleotide sequences of the 10 clones of *Ccmar1, Crmar1,* and *Tcmar1,* the *Ccmar1* clones form a distinct group, while the *Crmar1* and *Tcmar1* elements do not form separate clusters according to the host species classification. The *Tcmar1.81* element is grouped within the *Crmar1* elements, reflecting the lack of divergence between this element and the *Crmar1* clones. No congruence between the placement of the *Crmar1* elements and the geographic origin of the hosts was evident. Colonization of Mauritius and Réunion by *C. rosa* occurred only in the 1950s (Orian and Moutia 1960), far too recently for any significant genetic differentiation to have occurred. Using allozyme polymorphism data, Malacrida et al. (1996) estimated a genetic distance (Nei 1972) of only 0.048 between Kenyan and Réunion populations of this species.

The *Crmar1* clones all contain a five- or six-base deletion in the 3′ ITR. The 5′ ITR is identical to that of *Ccmar1* and *Tcmar1,* except in one clone, *Crmar1.4,* in which both ITRs share the same deletion. These observations suggest that the inverted terminal repeats of the *Crmar1* clones could be shorter, at 21 bp, than those of *Ccmar1* and *Tcmar1.* An alternative interpretation could be that the shorter 3′ ITRs are merely due to a deletion. However, the fact that the same, or very similar, deletion is present in all four clones sequenced is puzzling. A possible explanation is that the deletion occurred prior to or early in the rapid increase in copy number of the element. This is feasible only if such a deletion did not impair the element's ability to increase its copy number. If the increase in copy number occurred prior to the divergence of the host species the same deletions should also be present in copies of the element in the other species; however, the use of the longer PCR primer did not permit their amplification and hence detection. Such differences could not be detected using the full-length primer, as annealing of the primer would have been impaired; however, it would be interesting to know if such a deletion is present in *mariner* elements from *C. capitata* and *T. coffeae,* isolated with the shorter primer.

Evolutionary Ages of Ccmar1, Crmar1, *and* Tcmar1

The host species, *C. capitata, C. rosa,* and *T. coffeae,* appear to be a relatively young group; they last shared a common ancestor about 2 Mya (Malacrida et al. 1996 and our unpublished data). Within these species, the copies of *Ccmar1, Crmar1* and *Tcmar1* elements appear to be evolving independently of each other. This pattern is particularly evident for the two copies from *T. coffeae,* each of which has independently accumulated frameshifts and stop codons. That *Ccmar1, Crmar1,* and *Tcmar1* sequences are evolving neutrally is supported by the finding that the number of nucleotide substitutions is distributed randomly through synonymous and nonsynonymous positions.

It is difficult to estimate the time of divergence of these elements within and between the host species because the rates of molecular evolution of these species are not known. However, if we accept that these copies are evolving at neutral rates, then we can apply the rate of 1% nucleotide divergence per Myr calculated for drosophilids (Werman et al. 1990; Powell et al. 1993). At this rate, the *C. rosa* clones, which differ, on average, by 5.2%, have been in the genome for 2.6 Myr. This is comparable to the value of 2.3 Myr calculated for *Ccmar1* (Gomulski et al. 1997). The two sequences from *T. coffeae,* which show a DNA divergence of 8.6% seem to have diverged over a period of 4.3 Myr. The *Tcmar1.51* clone shows 9.4% and 9% average DNA divergence from the *Ccmar1* and *Crmar1* clone sequences, respectively, suggesting that the divergence times of these elements are 4.7 and 4.5 Myr, respectively. *Tcmar1.81* shows 5.1% and 4.8% divergence from the *Ccmar1* and *Crmar1* clone sequences, indicating divergence times of 2.55 and 2.4 Myr, respectively. Finally, the *Ccmar1* and *Crmar1* clone sequences appear to have diverged ~2.6 Mya, as they show average nucleotide divergences of 5.2%. The very similar evolutionary times of *Ccmar1, Crmar1,* and *Tcmar1.81* elements indicate a recent divergence from an ancestral element. Assuming that there is no very high undetected differentiation between all the copies of these elements in the three species, the simple hypothesis is that this ancestral element was present in the lineage prior to the divergence of the three species. Under this assumption, the vertical transfer into *T. coffeae,* or its ancestor, appears to have been the most ancient. Based on the degree of divergence estimated between *Tcmar1* copies, *Tcmar1.51* appears to be the older element in the *T. coffeae* lineage. This interpretation of the evolutionary history of *Ccmar1, Crmar1,* and *Tcmar1* elements seems congruent with the phylogeny of the host species. In fact, within the Ceratitini tribe, *T. coffeae* appears the most differentiated species with respect to *C. rosa* and to *C. capitata* (White and Elson-Harris 1992; Willhoeft and Franz 1996).

Differences in Copy Number

The abundance of the related elements *Ccmar1, Crmar1,* and *Tcmar1* differs dramatically in the three species: 500 copies were found in *C. capitata* (Gomulski et al. 1997), 5,000 copies in the closely related *C. rosa,* but only about 10 in the somewhat distantly related *T. coffeae.*

One of the four clones of *C. capitata* (*Ccmar1.18*) has a potentially active transposase (Gomulski et al. 1997); the related elements sequenced from *C. rosa* and *T. coffeae* are almost certainly nonfunctional. However, we

cannot exclude the possibility that active copies are present in these species. Clearly the high copy numbers in *C. rosa* and to a lesser degree in *C. capitata* are the result of very high transposition rates. Whether the large variation in copy number of *Ccmar1, Crmar1,* and *Tcmar1* is related to differences in the regulation of transposition in the different host species remains to be investigated. However, Robertson and Lampe (1995b) and Lampe et al. (1996) found no evidence for species-specific host factors for transposition of the *Himar1* element, a member of the irritans subfamily.

A self-regulation mechanism of transposition (OPI) which implies interaction with the host genome has been proposed by Lohe and Hartl (1996b) on the basis of studies of the *Mos1* element. To apply this mechanism of regulation to *Ccmar1, Crmar1,* and *Tcmar1,* in order to explain their differences in copy number, we must accept that the host species will have a very different optimal copy number at which the rate of transposition per element is maximized. This would imply that the closely related elements *Ccmar1, Crmar1,* and *Tcmar1* have very different levels of expression in the three host genomes. However, given the extremely wide range of variation in copy number observed (5,000–10) it is improbable that self-regulation is in effect and hence the diversity in copy number does not reflect a state of equilibrium but merely represents different stages in the evolution of the elements. Moreover, in the absence of regulation, the different copy number estimates do not seem to be at equilibrium, as we would have expected a positive correlation between copy number and host genome size (Lohe and Hartl 1996b). The genomes of the three species are fairly similar (from 5.3×10^8 in *T. coffeae,* 5.4×10^8 in *C. capitata,* to 5.8×10^8 in *C. rosa*) whereas the copy numbers differ by a factor of 500 in the three species.

Under the vertical transmission hypothesis, following the initial expansion of the copy number, the ultimate fate of most of the copies of *Ccmar1, Crmar1,* and *Tcmar1* would be the loss of transposase activity due to vertical inactivation (Lohe et al. 1995). Most of the sequenced clones of our elements were clearly subjected to vertical inactivation. The elimination of inactive elements from a species is a stochastic process and it is unpredictable and largely dependent on genetic drift (Lohe et al. 1995). However, element loss need not be purely a random process, as the probability of loss in different host species may be a function of the evolutionary history of the species themselves and of subtle biological differences between these species. Nothing is known about the history of *C. capitata, C. rosa,* and *T. coffeae* over evolutionary time. From its home range in southeast Africa, in the last 100 years, *C. capitata* succeeded in reaching an almost cosmopolitan geographical distribution; *C. rosa* is widespread in Africa, where it is the major pest of fruits, and since the late 1950s it has spread out of Africa where in some places it has displaced *C. capitata* (Mukiama 1985). *C. capitata* and *C. rosa* are extremely polyphagous and have attributes of *r*-strategist species (Fletcher 1989). *T. coffeae* is considered an endemic species of western Africa (White and Elson-Harris 1992); it is monophagous, being totally dependent on coffee (Greathead 1972). In their home ranges *C. capitata, C. rosa,* and *T. coffeae* live in sympatry on coffee, where they compete for the host. High levels of genetic variability have been detected in the native populations of *C. rosa* and *C. capitata,* whereas sympatric populations of *T. coffeae* show very low levels of variability. These data suggest that the effective sizes of natural populations of *T. coffeae* are very small compared to those of *C. rosa* and *C. capitata* (Baruffi et al. 1995; Malacrida et al. 1996). These data are in agreement with sympatric population size estimates of these three species during the coffee season in Kenya (Mukiama 1985). The population sizes of *C. capitata* and *C. rosa* appear to be very similar, whereas those of *T. coffeae* appear to be about twentyfold smaller. *T. coffeae* experiences a series of severe population crashes when the coffee resource is unavailable.

Given the above, it appears that genetic drift affects the genetic variability of the three host species to different extents. We can assume that genetic drift would have accelerated the loss of *Tcmar1* copies in *T. coffeae.* We propose that the vertical transmission of the element into *T. coffeae* is an old event with respect to that of *C. capitata* and *C. rosa.* The difference in the time of acquisition of the elements could account for the copy number diversity in the three species, particularly the very low copy number estimate in *T. coffeae.* The more recent transmission of *Ccmar1* and *Crmar1,* associated with their host biology and the high gene flow estimates within the host species (Gasperi et al. 1991; Malacrida et al. 1996 and our unpublished data), could explain the high copy number of these elements.

Acknowledgments. We thank Hugh Robertson for sending us the alignment of the amino acid sequences used in Fig. 5 and for his helpful suggestions. We also thank F. Godin, F. Chartois, and B. Denis for their generous advice on techniques. We also thank the referees for helpful suggestions. This collaborative research was supported by the Galileo Project 1996 (Italian-French cooperation). A grant from the European Communities Commission (IC18-CT96-0100) contributed to this work.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (eds) (1994) Current protocols in molecular biology. Green Publishing Associates and John Wiley and Sons, New York

Baruffi L, Damiani G, Guglielmino CR, Bandi C, Malacrida AR, Gasperi G (1995) Polymorphism within and between populations of *Ceratitis capitata:* comparison between RAPD and multilocus enzyme electrophoresis data. Heredity 74:425–437

Blanchetot A, Gooding RH (1995) Identification of *mariner* element from the tsetse fly *Glossina palpalis palpalis.* Insect Mol Biol 4:89–96

Brogna S, Bourtzis K, Babaratsas A, Zacharopoulou A, Savakis C (1994) The *Adh* locus of *Ceratitis capitata.* Fourth International Symposium on Fruit Flies of Economic Importance, Sand Key, Florida, June 5–10, 1994

Capy P, Koga A, David JR, Hartl DL (1992) Sequence analysis of active *mariner* elements in natural populations of *Drosophila simulans.* Genetics 130:499–506

Capy P, David JR, Hartl DL (1993) Evolution of the transposable element *mariner* in the *Drosophila melanogaster* species group. In: McDonald JF (ed) Transposable elements and evolution. Kluwer Academic, Dordrecht, The Netherlands, pp 290–299

Coates CJ, Turney CL, Frommer M, O'Brochta DA, Warren WD, Atkinson PW (1995) The transposable element *mariner* can excise in non-drosophilid insects. Mol Gen Genet 249:246–252

Doak TG, Doerder FP, Jahn CL, Herrick G (1994) A proposed subfamily of transposase-related genes: new members in transposase-like elements of ciliated protozoa and a common ''D35E'' motif. Proc Natl Acad Sci USA 91:942–946

Ebert PR, Hileman IV JP, Nguyen HT (1995) Primary sequence, copy number, and distribution of *mariner* transposons in the honey bee. Insect Mol Biol 4:69–78

Fletcher BS (1989) Life history strategies of tephritid fruit flies. In: Robinson AS, Hooper GHS (eds) Fruit flies: their biology, natural enemies and control, vol 3B. Elsevier, Amsterdam, The Netherlands, pp 195–208

Gasperi G, Guglielmino CR, Malacrida AR, Milani R (1991) Genetic variability and gene flow in geographical populations of *Ceratitis capitata.* Heredity 67:347–356

Gomulski LM, Torti C, Malacrida AR, Gasperi G (1997) *Ccmar1,* a full length *mariner* element from the Mediterranean fruit fly, *Ceratitis capitata.* Insect Mol Biol 6:241–253

Greathead DJ (1972) Notes on coffee fruit flies and their parasites at Kowanda (Uganda). Tech Bull Commonwealth Institute Biol Control 15:11–18

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Monro HN (ed) Mammalian protein metabolism III. Academic Press, New York, pp 21–132

Kidwell MG (1993) Voyage of an ancient *mariner.* Nature 362:202

Kumar S, Tamura K, Nei M (1993) MEGA: molecular evolutionary genetics analysis, version 1.0.1. Pennsylvania State University

Lampe DJ, Churchill MEA, Robertson HM (1996) A purified *mariner* transposase is sufficient to mediate transposition in vitro. EMBO J:5470–5479

Lindholm DA, Lohe AR, Hartl DL (1993) The transposable element *mariner* mediates germline transformation in *Drosophila melanogaster.* Genetics 134:859–868

Lohe AR, Hartl DL (1996a) Germline transformation of *Drosophila virilis* with the transposable element *mariner.* Genetics 143:365–374

Lohe AR, Hartl DL (1996b) Autoregulation of *mariner* transposase activity by overporduction and dominant-negative complementation. Mol Biol Evol 13:549–555

Lohe AR, Moriyama EN, Lidholm DA, Hartl DL (1995) Horizontal transmission, vertical inactivation, and stochastic loss of *mariner*-like transposable elements. Mol Biol Evol 12:62–72

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

Maruyama K, Hartl DL (1991) Evolution of the transposable element *mariner* in *Drosophila* species. Genetics 128:319–329

Medhora M, Maruyama K, Hartl DL (1991) Molecular and functional analysis of the *mariner* mutator element *Mos1* in *Drosophila.* Genetics 128:311–318

Mukiama TK (1985) East African strains of *Ceratitis capitata.* Report of the Joint FAO/IAEA Research Coordination Meeting on the Development of Sexing Mechanism in Fruit flies through manipulation of radiation. International Atomic Energy Agency, Vienna

Nei M (1972) Genetic distance between populations. Am Nat 106:283–292

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418–426

Orian AJC, Moutia LA (1960) Fruit flies (Trypetidae) of economic importance in Mauritius. Rev Agricole Sucrière Ile Maurice 39: 142–150

Powell JR, Caccone A, Gleason JM, Nigro L (1993) Rates of DNA evolution in *Drosophila* depend on function and developmental stage of expression. Genetics 133:291–298

Robertson HM (1993) The *mariner* transposable element is widespread in insects. Nature 362:241–245

Robertson HM (1995) The *Tc1-mariner* superfamily of transposons in Animals. J Insect Physiol 41:99–105

Robertson HM, Asplund ML (1996) *Bmmar1:* a basal lineage of the *mariner* family of transposable elements in the silkworm moth, *Bombyx mori.* Insect Biochem Mol Biol 26:945–954

Robertson HM, Lampe DJ (1995a) Distribution of transposable elements in arthropods. Annu. Rev Entomol 40:333–357

Robertson HM, Lampe DJ (1995b) Recent horizontal transfer of a *mariner* transposable element among and between Diptera and Neuroptera. Mol Biol Evol 12:850–862

Robertson HM, MacLeod EG (1993) Five major subfamilies of *mariner* transposable elements in insects, including the Mediterranean fruitfly, and related arthropods. Insect Mol Biol 2:125–139

Robertson HM, Soto-Adames FN, Walden KKO, Avancini RMP, Lampe DJ (1997) The *mariner* transposons of animals: horizontally jumping genes. In: Syvanen M, Kado C (eds) Horizontal Gene Transfer. Chapman and Hall, London

Sokal RR, Rohlf FJ (1981) Biometry. 2nd ed. Freeman WH, San Francisco

Swofford DL (1993) PAUP: phylogenetic analysis using parsimony, version 3.1.1. Computer program distributed by the Illinois Natural History Survey, Champaign, IL

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Werman SD, Davidson EH, Britten RL (1990) Rapid evolution in a fraction of the *Drosophila* nuclear genome. J Mol Evol 30:281–289

White JM, Elson-Harris M (1992) Fruit flies of economic significance: their identification and bionomics. C.A.B. International, Wellingford, Oxfordshire

Willhoeft U, Franz G (1996) Comparison of the mitotic karyotypes of *Ceratitis capitata, C. rosa* and *Trirhithrum coffeae* Bezzi (Diptera, Tephritidae) by C-banding and FISH. Genome 39:884–889