Insect Molecular Biology (2006) **15**(6), 743–748

Intact *mariner*-like element in tobacco budworm, *Heliothis virescens* (Lepidoptera: Noctuidae)

X. Ren, Y. Park* and T. A. Miller

*Department of Entomology, UC Riverside, Riverside, CA, USA; *Department of Entomology, Kansas State University, KS, USA*

Abstract

An intact *mariner*-like element was isolated from the genome of tobacco budworm, *Heliothis virescens*. This is the first report of an intact *mariner* element after the initial identification of *Mos1* from *Drosophila mauritiana*. The full-length *Hvmar1* has 30 bp inverted terminal repeats and a complete 1065 bp open reading frame (ORF) encoding the transposase with a 'D,D(34)D' motif in the catalytic domain. Polymerase chain reaction results show that at least one insertion of the *Hvmar1* element is conserved in this *Heliothis* strain. Phylogenetic analysis demonstrated that *Hvmar1* belongs to the *irritans* subfamily.

Keywords: *mariner*, transposable element, intact, tobacco budworm.

Introduction

The *mariner* family of transposons is one of the most diverse and widespread Class Π transposable elements (Robertson, 1993, 2002). Successful germ-line transformations of *Drosophila melanogaster* (Lidholm *et al*., 1993) and the yellow fever mosquito, *Aedes aegypti* (Coates *et al*., 1998) using *Mos1*, an active *mariner* element discovered from *D. mauritiana*, were achieved with entirely stable resultant strains. These achievements encouraged the use of *mariner* elements as genetic tools, and while they still have some disadvantages, our considerable understanding of their diversity and molecular evolution is a primary advantage of using *mariners* as genetic tools for insects.

Mariner elements are widespread in fungi, plants, invertebrates and vertebrates (Robertson, 1993). *mariner*-like elements (MLEs) include at least seven major subfamilies and 10 minor subfamilies based on sequence data previously published. Moreover, *mariner* elements from different subfamilies can coexist within the same genome (Robertson, 1993; Robertson & MacLeod, 1993). Typical full-length MLEs are about 1300 bp in length, with inverted terminal repeats (ITRs) of about 30 bp and a TA target duplication (Robertson, 1995). They encode a transposase distinguished by a characteristic 'D,D(34)D' motif in the catalytic domain (Hartl *et al*., 1997; Hartl, 2001; Coy & Tu, 2005), which has been shown to be necessary for transposition (Plasterk *et al*., 1999; Lohe *et al*., 1997). Most identified MLEs are nonfunctional (Robertson & lampe, 1995) due to multiple frameshifting indels and in-frame stop codons in their transposase genes. So far only two characterized *mariner* elements have been demonstrated to be autonomous. These include *Mos1*, originally discovered in the fruit fly *D. mauritiana* (Medhora *et al*., 1991) and *Himar1*, an artificially constructed consensus sequence of elements found in the horn fly, *Haematobia irritans* (Lampe *et al*., 1996).

Mariner elements have a wide phylogenetic distribution. However, the phylogenetic relationships of *mariner* elements are not in concord with those of their hosts. This suggests that horizontal transfer is prevalent in this group (Robertson *et al*., 1998). In this study, we report the identification of an intact *mariner*-like element in tobacco budworm, *Hethiothis virescens, Hvmar1*. This element belongs to the *irritans* subfamily mostly close to human *mariner* transposon (*Hsmar2*), which might suggest a horizontal transfer event.

Results

Isolation of an intact mariner element

An intact *mariner*-like element was isolated from tobacco budworm. Degenerate primers MAR-124F and MAR-276R, which were designed to bind to two stretches of conserved

doi: 10.1111/j.1365-2583.2006.00673.x

Received 2 February 2006; accepted after revision 19 April 2006. Correspondence: Xiaoxia Ren, Department of Molecular Microbiology and Immunology, Bloomberg School of Public Health, Johns Hopkins University 615 N. Wolfe St, Rm E5132, Baltimore, MD 21205, USA. Tel.: +1(410)502 3132; e-mail: xren@jhsph.edu

amino acid sequences within *mariner* transposase (Robertson, 1993; Robertson & MacLeod, 1993), were used to amplify possible fragments of *mariner* elements. The polymerase chain reaction (PCR) products were subsequently subcloned to the pGEM-T easy vector and sequenced. A BLASTN search showed that one of the fragments, which was 503 bp in length, was similar to other known *mariner* elements. Based on the sequence of this fragment, a pair of outward facing primers was designed for inverse PCR to obtain the full-length sequence of this *mariner* element. The obtained full-length sequence of this *mariner*-like element is shown in Fig. 1.

This element is 1269 bp in length and has an open reading frame of 1065 nucleotides encoding a transposase of 354 amino acid residues. This element also has a 'D,D(34)D' motif in the catalytic domain, 30 bp imperfect ITRs, which have only one nucleotide difference among 30 nucleotides, and a TA target duplication upon integration. The conceptually translated protein sequence has the highest amino acid sequence identity (36%) to the *Homo sapiens mariner* (*Hsmar2*) and 33% identity to the *Chrysoperla plorabunda mariner* (*CplmarCp3*). The putative amino acid sequence aligned with those of other *mariner* elements is shown in Fig. 2. The GenBank accession number for the *Hvmar1* is DQ174779.

Insertion site sequences and mobility

Using inverse PCR from *Eco*RI digestion, we obtained the flanking sequences of the element. BLAST analysis did not reveal any significant commonalities to other sequences in the database. Sequencing data showed that the complete EcoRI fragment is about 2.4 kb in length. To find out whether this insertion occurred in every individual, we designed a pair of inward facing primers, which are located in the 5′- and 3′-flanking genome sequences, to amplify the whole insert with an expected 1.7 kb band if the insertion is present. We randomly tested 14 individuals for this insertion, all of which were positive for the 1.7 kb amplicon, indicating that this insertion may be conserved in this *Heliothis* strain.

Estimation of the copy number

The copy numbers of MLEs are extremely variable in different species (Hartl *et al*., 1997). To quantify copy number of the *mariner* element in this *Heliothis* strain, we conducted a Southern hybridization using a 650 bp fragment, which includes the most conserved sequence of the *mariner* element, as a probe. Four individuals were tested. The results show that the copy number is less than 10 and each individual has different band pattern (Fig. 3). Among the bands, there are several found in every individual, one of these is the 2.4 kb band. The universal presence of this 2.4 kb band is consistent with previous inverse PCR results.

```
cgaaggtcactagataagttctgcaatacattcgttttattaatgg
agggtagttgaaaatatttttatcataagtcttagctaatgtcctc
acagttcatatgttacaagttgacataattccttttgctcagttaaa
taaaaaaaaattaaattaaaaa

  1 atgtcggagtggtctaatgacgagttgcgagttgtaatgcgctac
    M  S  E  W  S  N  D  E  L  R  V  V  M  R  Y
 46 aacttttgcgtaaattgtccatagatgagtgtgtgtctgaaatg
    N  F  L  R  K  L  S  I  D  E  C  V  S  E  M
 91 aaaactgttttgggagaaaagtgtcctcatcgtacaacagtttac
    K  T  V  L  G  E  K  C  P  H  R  T  T  V  Y
136 cgatggtatcgtgagtttgaacgtggtaactttaacgttaatgac
    R  W  Y  R  E  F  E  R  G  N  F  N  V  N  D
181 gcagcaagatcaggacgtcccgtagaggccactagctctgagaac
    A  A  R  S  G  R  P  V  E  A  T  S  S  E  N
226 gtggcaaaagtggaaaaattgttaaaggaagaccgacgtattacg
    V  A  K  V  E  K  L  L  K  E  D  R  R  I  T
271 tactcacagatagaggaggccctacagattagtgcaccgagcgta
    Y  S  Q  I  E  E  A  L  Q  I  S  A  P  S  V
316 cacaaaattcttcacgatatcctccgagtaaaaaaagtttgtacg
    H  K  I  L  H  D  I  L  R  V  K  K  V  C  T

          MAR-124F
361 ctgtgggtaccacatgaccttaaacctgaacaaaaggatacaaga
    L  W  V  P  H  D  L  K  P  E  Q  K  D  T  R
406 gtcaagtggtgtaaaaaaatgctgcaaatgtatgaaaatggtact
    V  K  W  C  K  K  M  L  Q  M  Y  E  N  G  T
451 tcaaataacatcaataatattgttaccggtgatgaaacatggtta
    S  N  N  I  N  N  I  V  T  G  D  E  T  W  L
                                    ▲
496 tattattttgatttaccatcgaaaaacaaaaacaaagtttggctc
    Y  Y  F  D  L  P  S  K  N  K  N  K  V  W  L
541 tttgaaaatgaacaaactccagtgcaggtgcgaaaatcacgttcg
    F  E  N  E  Q  T  P  V  Q  V  R  K  S  R  S
586 gtgaaaaaaaagatgatcgctgtatttttacaaggcggggtatt
    V  K  K  K  M  I  A  V  F  F  T  R  R  G  I
631 ttggaacgggtgttactggaaagtcagcgtactgtaacggcatcc
    L  E  R  V  L  L  E  S  Q  R  T  V  T  A  S
676 tggtacattaatgaatgtttgcctaaagtttttcaaaggcttcag
    W  Y  I  N  E  C  L  P  K  V  F  Q  R  L  Q
721 gaaattcgaccaaattcaagaatggacactccgcattttcatcat
    E  I  R  P  N  S  R  M  D  T  P  H  F  H  H
766 gacaacgcgcccgcgcatcgggcacgcgatactgtcgaatttttg
    D  N  A  P  A  H  R  A  R  D  T  V  E  F  L
    ▲                                       ←
811 aatagcagtggtgttagagtgctggatcatcctgcgtacagccca
    N  S  S  G  V  R  V  L  D  H  P  A  Y  S  P

       MAR-276R
    ←
856 gacctagcgccctgtgactttgctcttttcccaatcattaaggat
    D  L  A  P  C  D  F  A  L  F  P  I  I  K  D
                   ▲
901 cagctgaaaggcaggaagtttcaaatagaggtggagcttttgagc
    Q  L  K  G  R  K  F  Q  I  E  V  E  L  L  S
946 gcttgggaccaagcatgttcagaactatcagaagaaaagtggaag
    A  W  D  Q  A  C  S  E  L  S  E  E  K  W  K
991 gatatttttaatgattggtttttacgcatgacaaagtgtataaac
    D  I  F  N  D  W  F  L  R  M  T  K  C  I  N
1036 tgtaatggaaattatttcgaaagagtgtaa 1065
    C  N  G  N  Y  F  E  R  V  *

aaaataaagctgtattgcagaacttatctagtgacctttg
```

Figure 1. *Hvmar1* element nucleotide sequence and predicted amino acid translation of encoded transposase. The inverted terminal repeats are underlined. The positions of MAR-124F and MAR276R PCR primer sites are indicated by arrows. The aspartic acid residues of the 'D,D(34)D' catalytic domain are marked with solid triangles.

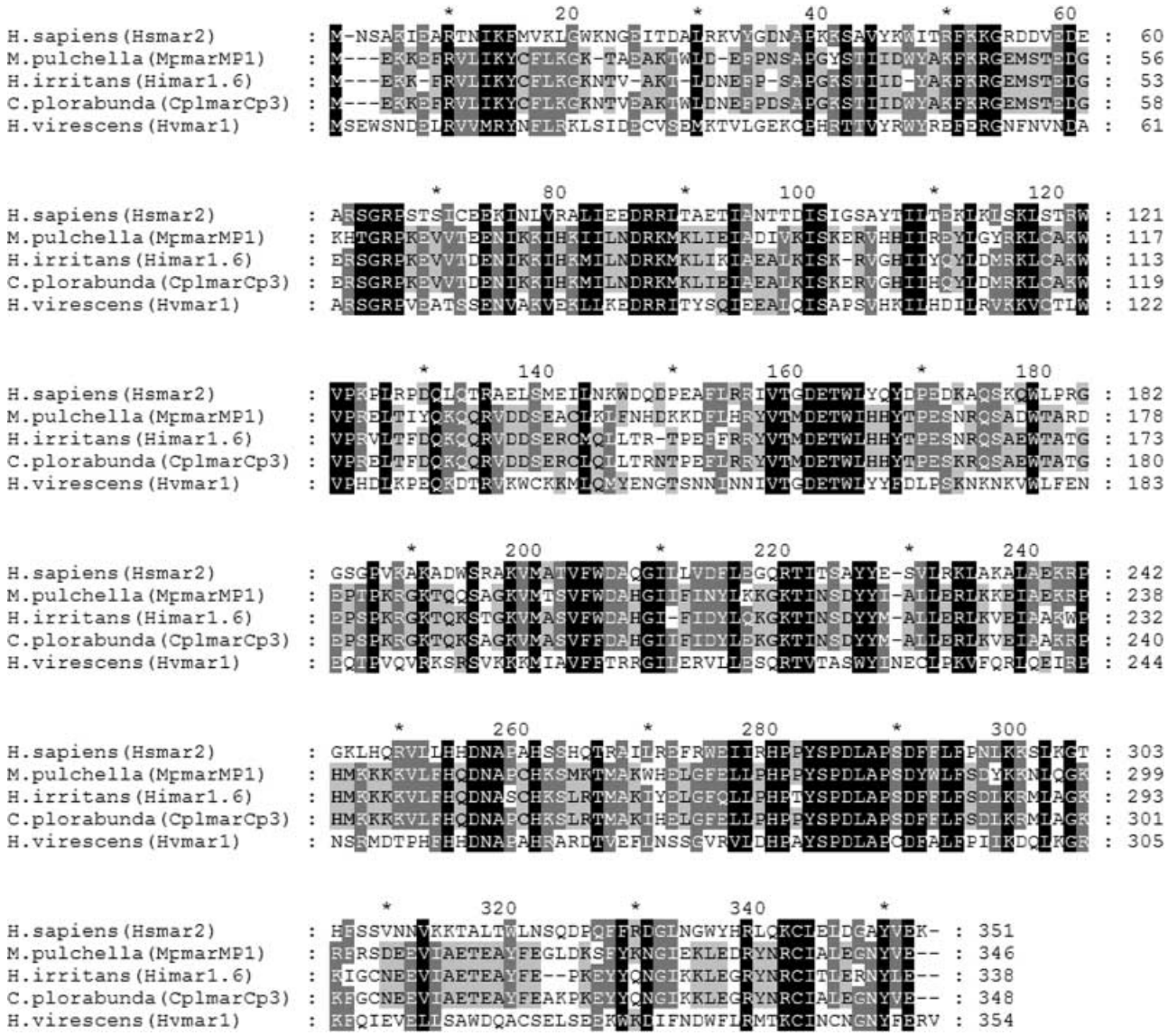

Figure 2. Amino acid sequence alignment of transposase coding regions from *irritans* subfamily of MLEs. Grey and black shaded regions highlight regions of amino acid similarity and identity, respectively. GenBank accession numbers for the sequences are: *Hsmar2* (AAC52011), *MpmarMP1* (U11649), *Himar1.6* (U11645), *CplmarCp3* (AAC46947) and *Hvmar1* (DQ174779).

Phylogenetic analysis

To investigate the relationship of *Hvmar1* to previously characterized full-length members of the *mariner* family, the transposase amino acid sequence was used to perform a phylogenetic analysis using MEGA3.1. Results of neighbour-joining analysis are summarized in Fig. 4. The Bootstrap support was evaluated with 1000 replications. The results clearly show that *Hvmar1* belong to the *irritans* subfamily with other members including *H. irritans* (*Himar1.6*), *C. plorabunda* (*CplmarCp3, Cplmar1, CplmarCp1, CplmarCp8, CplmarCp2*), *M. pulchella* (*MpmarMP1*) and *H. sapiens* (*Hsmar2*). Within the *irritans* subfamily, *Hvmar1* shows the nearest relationship with *Hsmar2*.

Comparison of inverted terminal repeats with those of other mariner elements

Hvmar1 has nearly perfect ITRs, which have only one nucleotide difference among 30 nucleotides. The alignment of the ITRs of *Hvmar1* with those of other members from the *irritans* subfamily shows that the ITRs of *Hvmar1* have the highest identity sequence with *Hsmar2*. This result is consistent with amino acid alignment result.

Discussion

This is the first report of a naturally occurring intact *mariner* element after the initial identification of *Mos1* from *D. mauritiana.*

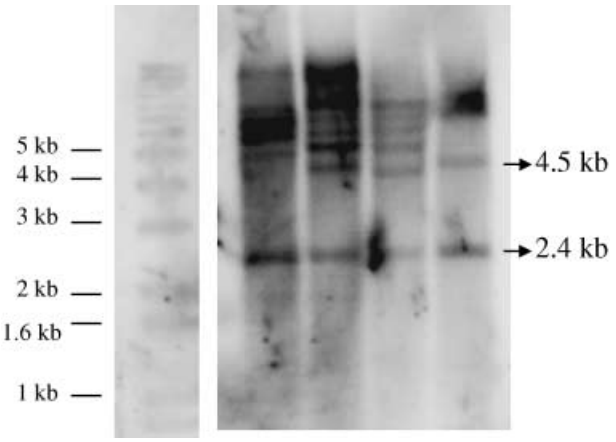

Figure 3. Southern blots of *Eco*RI-digested genomic DNA. The left single lane is 1 kb DNA ladder. The right four lanes are four individuals from the laboratory strain.

This intact *mariner* element is mostly closely related to the human *mariner* element (*Hsmar2*) according to the phylogenetic analysis. Even though the mechanism for the horizontal transfer still remains unknown, it is a reasonable explanation for the close relationship of *mariners* between a moth and the human being.

So far most MLEs examined are non-autonomous, and contain one or more mutations in the transposase coding region. The mutations caused by vertical inactivation accumulate in a time-proportional manner (Lohe *et al*., 1995), so time since horizontal transfer occurred could be deduced from the level of divergence of the various *mariner* copies with the host. The intact *mariner* in *Heliothis* has the complete ORF and the nearly complete ITRs, which leads to the hypothesis that this element might be a result of recent horizontal transfer.

The intact *Hvmar1* element may be autonomous due to its complete ORF and nearly perfect ITRs. A possible reason that this element is not mobile in this species, indicated by the stability of the 2.4 kb fragment, is that there might be a host factor that suppresses the activity in this host species. Theory predicts that hosts should evolve diverse regulatory mechanisms to suppress the activity of invading transposons. Indeed, there are many instances of such mechanisms, such as mutations of the *flamenco* gene allow activity of *gypsy* retrotransposons in *D. melanogaster* (Bucheton, 1995). Similar mechanisms are also known for transposons in various plants and fungi. Another possible reason for the stability of the intact *mariner* element in *Heliothis* is that this *mariner* element has its own built-in regulatory or repressive systems. This mechanism is well known for the *P* element in *D. melanogaster*, in which *P*-element transposition is repressed by *P* cytotype in *Drosophila P*-strain females (Roche *et al*., 1995).

In addition to the 2.4 kb band, more bands were detected by Southern blot, which could be due to the stringency used for the hybridization. These additional bands could be other more divergent element subfamilies in this species, which could explain why we could not identify more insertion sites other than the one of this intact element by using specific primers. Intrinsic feature of this element and its interaction with host factors could be the reason for the single copy in this species as discussed in above. Although the involvement of these mechanisms remains very speculative in the absence of functional data, the identification of this intact element gives us a good candidate to address these questions.

One important advantage of using *mariners* as genetic tools is our considerable understanding of their diversity and molecular evolution. Endogenous *mariners* can be easily identified using degenerate PCR primers designed from conserved regions of the transposase. So far at least seven major subfamilies and 10 minor subfamilies have been identified from different organisms. Such a large number of potentially useful transposons provide a great source to be exploited, if necessary, to introduce transgenes into a particular host insect.

Experimental procedures

Insect strain

Tobacco budworm pupae were received from the AgriPest Insects for Agricultural Research (http://www.intrex.net/agripest), and then maintained in laboratory for research purposes. All stages of tobacco budworms except larvae were kept in incubators with a 14 : 10 (L/D) light cycle, a temperature of 27–28 °C and 70–80% RH. Larvae were reared at RH between 30 and 40%. Dry tobacco budworm diet was purchased from Southland Products Inc. (201 Stuart Island Road, Lake Village, AR, USA).

Degenerate-primer polymerase chain reaction

The forward MAR124F (TGGGTNCCNCAYGARYT) and the reverse MAR276R (GGNGCNARRTCNGGNSWRTA) degenerate primers designed for amplifying the conserved regions of the *mariner* transposase gene (Robertson, 1993) were used to isolate endogenous *mariner* elements from *H. virescens* genomic DNA. Genomic DNA was isolated from the individual third-instar larvae using the QIAGEN DNeasy® Tissue Kit (27220 Turnberry Lane, Suite 200, Valencia, CA, 91355, USA). Genomic DNA samples were subjected to 40 amplification cycles under the following conditions (Hotstart PCR): 95 °C for 15 min, [94 °C for 1 min, 50 °C for 1 min, 72 °C for 1 min]: 40 cycles, 72 °C for 10 min. A 40 μl reaction included 20 μl HotStar Taq™ Master Mix (QIAGEN), 1 μM of each primer and 20–100 ng genomic DNA.

Inverse polymerase chain reaction

Two to four micrograms of genomic DNA from a single individual were digested overnight with *Eco*RI in a 20 μl volume. The restriction enzyme was inactivated at 65 °C for 15 min, and then the total volume was increased to 400 μl with the addition of 336 μl MiliQ H_2O, 40 μl 10 × ligation buffer and 4 μl T4 DNA ligase (Promega, 2800 Woods Hollow Rd, Madison, WI 53711, USA). The digested DNA was therefore at a concentration of 5–10 ng/ml, within the 1–10 ng/ml

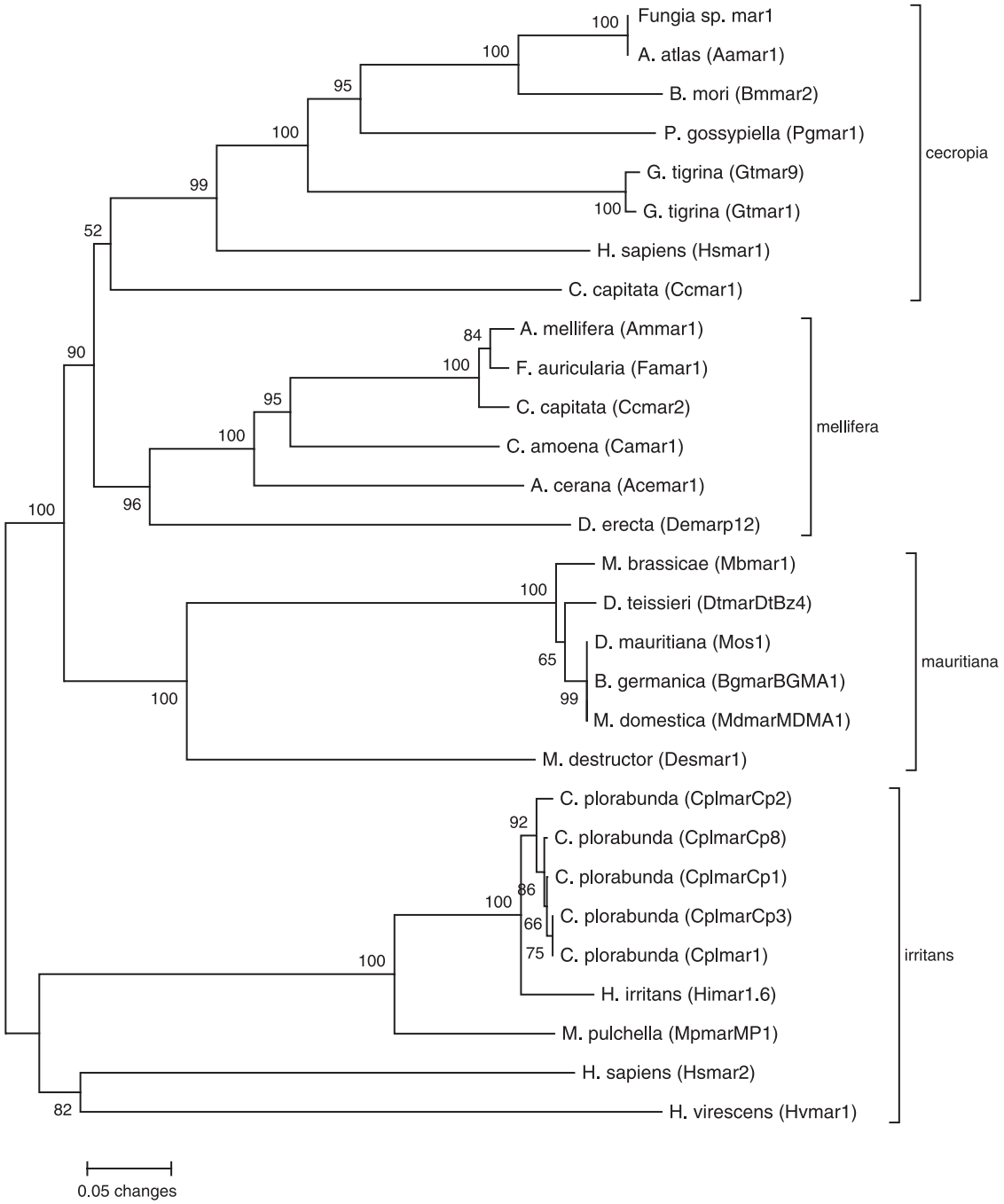

Figure 4. Phylogenetic relationships of *Hvmar1* with some other members of the *mariner* superfamily, based on predicted transposase amino acid sequences. Tree was constructed using neighbour-jointing method with model amino: p-distance. Numbers on the nodes are bootstrap support values in 1000 replications.

recommended for circularizing fragments during the ligation reaction. The ligation reactions were left at 4 °C overnight, then precipitated with 3 M NaAC (pH 5.2) and 2 × volumes of 100% ethanol and resuspended in 20 μl MiliQH2O.

A pair of outward-facing primers was designed for the inverse PCR: MARinvF (CTGTAACGGCATCCTGGTA) and MARinvR (CATGTTTCATCACCGGTAAC). Touchdown PCR was used for the amplifying and the PCR cycles were as follows: 94 °C for 3 min and followed by 10 cycles of 94 °C for 1 min, 62 °C for 1 min, 72 °C for 2 min with 1.0 °C lower per cycle, then followed by another 25 cycles of 94 °C for 1 min, 52 °C for 1 min, 72 °C for 2 min, finally 72 °C for 10 min.

Cloning and sequencing

All PCR products were subjected to gel electrophoresis in agarose with appropriate concentrations, purified using a MiniElute Gel Extraction Kit (QIAGEN), and cloned directly into the pGEM-T vector (Promega). Plasmids were isolated using QIAprep® Miniprep

Kit (QIAGEN). Sequencings of the clones were done by the DNA Sequencing Core Instrumentation Facility at the Institute for Integrative Genome Biology in the University of California, Riverside with SP6 or T7 primers.

Southern DNA hybridization

Approximately 5 μg genomic DNA samples from single individuals were digested to completion with *Eco*RI restriction enzyme, separated on 0.8% agarose gel, blotted to nylon membrane and immobilized by ultraviolet irradiation. Hybridization probe was generated from PCR amplification. The probe was then purified using MiniElute Gel Extraction Kit (QIAGEN) and radio-labelled with [^{32}P]-dCTP by random oligonucleotides priming (Roche Applied Science, 9115 Hague Road, Indianapolis, IN 46250, USA). Hybridization was performed at 42 °C overnight in ULTRAhyb™ Ultrasensitive Hybridization Buffer (Ambion, 2130 Woodward, Austin, TX 78744, USA). After hybridization, membrane was washed twice, 5 min each, in 2 × SSC, 0.1% SDS at 42 °C, followed by two washes in 2 × SSC, 0.1% SDS at 42 °C, 15 min each.

Acknowledgements

We would like to thank Dr Jason L Rasgon for reviewing the manuscript. This work was supported by Experiment Station Project CA-R-ENT-5887-H.

References

Bucheton, A. (1995) The relationship between the *flamenco* gene and *gypsy* in *Drosophila*: how to tame a retrovirus. *Trends Genet* **11**: 349–353.

Coates, C.J., Jasinskiene, N., Miyashiro, L. and James, A.A. (1998) *Mariner* transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. *Proc Natl Acad Sci USA* **95**: 3748–3751.

Coy, M.R. and Tu, Z. (2005) *Gambol* and *Tc1* are two distinct families of DD34E transposons: analysis of the *Anopheles gambiae* genome expands the diversity of the *IS630-Tc1-mariner* superfamily. *Insect Mol Biol* **14**: 537–546.

Hartl, D.L. (2001) Discovery of the transposable element *mariner*. *Genetics* **157**: 471–476.

Hartl, D.L., Lohe, A.R. and Lozovskaya, E.R. (1997) Modern thoughts on an ancyent marinere: function, evolution, regulation. *Annu Rev Genet* **31**: 337–358.

Lampe, D.J., Churchill, M.E.A. and Robertson, H.M. (1996) A purified *mariner* transposase is sufficient to mediate transposition *in vitro*. *EMBO J* **15**: 5470–5479.

Lidholm, D.A., Lohe, A.R. and Hartl, D.L. (1993) The transposable element *mariner* mediates germline transformation in *Drosophila melanogaster*. *Genetics* **134**: 859–868.

Lohe, A.R., Moriyama, E.N., Lidholm, D.A. and Hartl, D.L. (1995) Horizontal transmission, vertical inactivation, and stochastic loss of *mariner*-like transposable elements. *Mol Biol Evol* **12**: 62–72.

Lohe, A.R. and De Aguiar, D. and Hartl, D.L. (1997) Mutations in the *mariner* transposase, the 'D,D(35)E' consensus sequence is nonfunctional. *Proc Natl Acad Sci USA* **94**: 1293–1297.

Medhora, M., Maruyama, K. and Hartl, D.L. (1991) Molecular and functional analysis of the *mariner* mutator element *Mos1* in *Drosophila*. *Genetics* **142**: 907–914.

Plasterk, R.H.A., Izsvak, Z. and Ivics, Z. (1999) Resident aliens: the *Tc1/mariner* superfamily of transposable elements. *Trends Genet* **15**: 326–332.

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

Robertson, H.M. (1995) The *Tc1-mariner* superfamily of transposons in animals. *J Insect Physiol* **41**: 99–105.

Robertson, H.M. (2002) Evolution of DNA transposons in eukaryotes. In *Mobile DNA II* (Craig, N.L., Craigie, R., Gellert, M. and Lambowitz, A.M., eds), pp. 1093–1110. ASM Press, Washington DC.

Robertson, H.M. and Lampe, D.J. (1995) Distribution of transposable elements in arthropods. *Annu Rev Entomol* **40**: 333–357.

Robertson, H.M. and MacLeod, E.G. (1993) Five major subfamilies of *mariner* transposable elements in insects, including the Mediterranean fruit fly, and related arthropods. *Insect Mol Biol* **2**: 125–139.

Robertson, H.M., Soto-Adames, F.N., Walden, K.K.O., Avancini, R.M.P. and Lampe, D.J. (1998) The *mariner* transposons of animals: horizontally jumping genes. In *Horizontal Gene Transfer* (Syvanen, M. and Kado C.I., eds), pp. 268–284. Chapman & Hall, London.

Roche, S.E., Schiff, M. and Rio, D.C. (1995) *P* element repressor autoregulation involves germ-line transcriptional repression and reduction of third intron splicing. *Genes Dev* **9**: 1278–1288.

