ELSEVIER

Detection of a *mariner*-like element and a miniature inverted-repeat transposable element (MITE) associated with the heterochromatin from ants of the genus *Messor* and their possible involvement for satellite DNA evolution

Teresa Palomeque *, José Antonio Carrillo, Martín Muñoz-López, Pedro Lorite

Departamento de Biología Experimental. Área de Genética. Universidad de Jaén. 23071, Jaén, Spain

Abstract

The satellite DNA of ants *Messor bouvieri*, *M. barbarus* and *M. structor*, studied in a previous paper, is organized as tandemly repeated 79-bp monomers in the three species showing high sequence similarity. In the present paper, a *mariner*-like element (*Mboumar*) and a new MITE (miniature inverted-repeat transposable element) called IRE-130, inserted into satellite DNA from *M. bouvieri*, are analyzed. The study of *Mboumar* element, of its transcription and the putative transposase that it would encode, suggests that it could be an active element. *Mboumar* elements inserted into IRE-130 elements have also been detected. It is the first time, to our knowledge, that a MITE has been described in Hymenoptera and it is also the first time that a *mariner*-like element inserted into a MITE has been detected. A *mariner*-like element, inserted into satellite DNA from *M. structor* and in *M. barbarus*, also has been found. The results seem to indicate that transposition events have participated in the satellite DNA mobilization and evolution.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Ants; *Mariner*-like element; MITE; Satellite DNA evolution

1. Introduction

Transposable elements comprise a major fraction of the eukaryotic genome. *Mariner-Tc1*, a diverse and widespread superfamily of class II transposable elements (transposable elements that move via a DNA intermediate), includes *mariner*-like elements, Tc1 and Tc1-like elements (Robertson, 1996). *Mos*1 was the first *mariner* transposable element isolated from *Drosophila mauritiana* (reviewed in Hartl, 2001). *Mos*1 is 1286 bp in length contains imperfect inverted terminal repeats (ITRs) and encodes a transposase of 345 amino acids. Numerous

studies on this transposable element and other *mariner*-like transposable elements (MLEs) have been made as well as on transposition mechanism (Lampe et al., 2001; Krieger and Ross, 2003; Augé-Gouillou et al., 2005a,b; Bigot et al., 2005, among others).

Miniature inverted-repeat transposable elements (MITEs) have been identified in many animals and plants (Feschotte et al., 2002; Casacuberta and Santiago, 2003). MITEs are also members of the class II transposons (reviewed in Kidwell, 2002). They are characterized by a short length, have no coding potential, are AT-rich and may have a stable secondary structure. They are often present in a high copy number per family and genome and they show individual target-sequence preferences. Generally, MITEs have terminal inverted repeat or subterminal inverted repeat. For example, *Micron* elements lack the above-mentioned repetitions and contain several internal repeated sequences and subterminal inverted repeats (Akagi et al., 2001).

Abbreviations: bp, base pairs; HTH, helix-turn-helix; IRE, interspersed repetitive element; ITR, inverted terminal repeat; kb, kilobase(s); M., *Messor*; MITE, miniature inverted-repeat transposable element; MLE, *mariner*-like transposable elements; NLS, nuclear location signal; stDNA, satellite DNA; TLE, Tc1 transposable element; TSD, target-site duplication.

* Corresponding author. Tel.: +34 953 212156; fax: +34 953 211875.
  E-mail address: tpalome@ujaen.es (T. Palomeque).

Various authors have suggested that MITEs could be cross-mobilized by transposases encoded in *trans* by autonomous elements. Supporting this suggestion, sequence similarity has been detected between the ITRs and target-site duplications (TSDs) of MITEs and the autonomous element in the same genome. Concretely, *Tourist*-like MITEs show sequence similarity with the ITRs of PIF/Harbinger superfamily (Zhang et al., 2004) and the ITR and TSD of *Stowaway*-like show sequence similarities with the ITR and TSD of Tc1/*mariner* element (Feschotte et al., 2002). Recently, it has been demonstrated that OSMAR5 (a *mariner*-like transposase) interacts in vitro with the inverted terminal repeat of certain *Stowaway* MITEs (Feschote et al., 2005).

Heterochromatin, well represented in eukaryotic genomes, is rich in satellite DNA (stDNA) and transposable elements, as documented for species as diverse as *Drosophila melanogaster*, *Arabidopsis* spp. and *Zea mays*. The origin of many stDNA families is not definitively known, but a few have been generated from transposable elements or have a major component that is related to part of a mobile element (reviewed in Kidwell, 2002).

MLEs have been described in several species of ants (Hymenoptera, Formicidae). In specific, these elements have been found in *Tapinoma sessile* and *Crematogaster cerasi* (Robertson and MacLeod, 1993), *Myrmica ruginodis* (Bigot et al., 1994) and in several species belonging to the genus *Solenopsis* (Krieger and Ross, 2003). Two LTR-retrotransposons have also been found in *Solenopsis invicta* genome (Krieger and Ross, 2003).

The stDNA of ants *Messor bouvieri*, *M. barbarus* and *M. structor* is organized as tandemly repeated 79-bp monomers in all species, showing high similarity and all belonging to the same stDNA family. A 25-bp DNA fragment was found in the 5′ end of a clone of stDNA in the ant *M. bouvieri* considered to be unrelated flanking DNA (Lorite et al., 2002). The analysis of new stDNA clones revealed other clones with a similar DNA sequence. In addition, these 25-bp fragments are also present in their reverse complementary form in the 3′ ends. These new studies also have determined that these fragments are part of the inverted ends of a new MITE inserted into the stDNA. A presumably active *mariner*-like element was also found, sometimes inserted into the stDNA repeats or into the MITE element. *Mariner*-like elements are also present in stDNA from *M. structor* and in *M. barbarus*. The possible role of these elements in the evolution of stDNA is analyzed.

2. Materials and methods

2.1. Materials, DNA extraction, PCR amplification and cloning

Ants of *M. bouvieri* were collected at the dam of Giribaile (population GI-1) and dam of Rules (population RU-2) from the provinces of Jaén and Granada (Spain), respectively. Different individuals from each population (A, A$_1$ and B, B$_1$, respectively) were used for genomic-DNA extraction.

A 25-bp DNA fragment was found in the 5′ end of the MEBO-128 clone of stDNA in the ant *M. bouvieri* (Lorite et al., 2002). The 25-bp fragments are also present in their reverse complementary form in the 3′ ends. Using this 25-bp sequence, we designed a single primer to amplify the DNA inserted in the stDNA (MEBOTRA: 5′-AGTCAGAAATGACACCTC-GATC). PCR amplifications were performed using the following cycling profile: 25 cycles at 95 °C (20 s), 50.6 °C (30 s), 68 °C (2 min), with a final elongation step of 68 °C for 7 min. Reactions were set up in a 50-μl mixture containing 100 ng of genomic DNA, 5% DMSO, 0.5 mM dNTPs, 50 pmol of the primer and 1 U of Expand Long Template polymerase (Roche).

Two other pairs of primers were designed, the first pair based on the sequence of the *Mboumar mariner* element isolated, one at each end of the *mariner* and directed outwards (MAR-1: 5′-CAGGCACAAAACTCGACAT and MAR-4: 5′-CCATTC-TAAGTAACGTGT). The second pair was based on the sequence of the stDNA of this species (MEBOSAT-1: 5′-CGCAAGTAAATCTTAGTTTTCT and MEBOSAT-2: 5′-AGAAAACTAAGATTTACTTGCG). PCR amplifications were performed using the following cycling profile: 35 cycles at 95 °C (20 s), 52 °C (1 min), 68 °C (1.5 min), with a final elongation step of 68 °C for 3 min. Reactions were set up in a 50-μl mixture containing 100 ng of genomic DNA, 5% DMSO, 0.5 mM dNTPs, 40 pmol of each primer and 1 U of Expand Long Template polymerase (Roche).

Two different PCRs were performed to verify whether transposable elements are present in genomic DNA from *M. structor* and *M. barbarus*, using a primer derived from the sequence data obtained from the inverted terminal repeats of *mariner*-like elements studied in the present paper (ITR-MAR: 5′-CCAGGTGTGTCGGTAATTCCTTTCCGG) and the primer MEBOTRA mentioned above. Finally, a PCR was performed to determine whether the *mariner*-like element is also inserted directly into its stDNAs, using a primer based on the stDNA (MEBOSAT-1) and another based on the *mariner* sequence (MAR-1). PCR amplifications were performed in all the cases with the same cycling profile described for the PCRs, also using MEBOTRA as a primer.

PCR products were analyzed by electrophoresis in agarose gels, eluted from agarose gel and inserted into the pGEM-T Easy vector (Promega).

2.2. Sequence analysis

Recombinant plasmids were directly sequenced on both strands by the dideoxy sequencing method. Multiple-sequence alignment was performed using the CLUSTALW program. Sequence comparison, the search for ORFs and other sequence analyses were performed using the programs available on-line in the Resources Centre INFOBIOGEN (http://www.infobiogen.fr). Sequence divergences were calculated as the average of nucleotide substitution per site between sequences (D_{xy} value from DnaSP program). The NSP@Network Protein Sequence Analysis program was used for the prediction of helix-turn-helix motifs. The putative TATA box and polyadenylation signals were determined using the programs HCtata (Hamming-Clustering Method for TATA Signal Prediction in Eukaryotic Genes) and HC polya (Hamming Clustering Poly-A Prediction in

Eukaryotic Genes). The phosphorylation target sites at casein kinase II were determined using the Motif Finder program (http://motif.genome.jp/). The MITEs were analyzed using the MAK program (http://perl.idmb.tamu.edu/mak.htm). The theoretical secondary structure of these elements was studied using the FoldDNA program (http://www.bioinfo.rpi.edu/applications/mfold/fold/dna/form1cgi).

2.3. Amplification of mariner-flanking sequences

Two different techniques based on PCR have been used for the identification of *mariner*-flanking sequences, inverse PCR and partially unspecific PCR.

For inverse PCR, genomic DNA was digested with *Eco*RI, since there are no targets for this enzyme in the *mariner* elements isolated in *M. bouvieri*. Digested DNA fragments were circularized with T4 ligase. This DNA was used as template DNA in PCR experiments using the primers MAR-1 and MAR-4 described above. PCR amplification was carried out using the following cycling profile: 35 cycles at 95 °C (20 s), 53 °C (30 s), 68 °C (2 min), with a final elongation step of 68 °C for 7 min. Reactions were set up in a 50-µl mixture containing 100 ng of genomic DNA, 5% DMSO, 0.5 mM dNTPs, 40 pmol of each primer and 1 U of Taq polymerase (Bioline). PCR products were cloned as described above (*Mboumar*-INV-PCR clones).

Partially unspecific PCR is based in the technique of Karlyshev et al. (2000). In this PCR experiments, undigested genomic DNA and the primers MAR-1 and MAR-4 were used. PCR amplification was carried out using the same conditions that inverse PCR although with lower annealing temperature (40–45 °C). Under these conditions, unspecific amplification fragments are obtained but some of them a formed by an end of the *mariner* sequence and the flanking DNA. The obtained PCR fragments were cloned (*Mboumar*-PCR clones). *Mboumar*-9 *mariner* sequence was digoxygenin-labeled by random priming with the DIG system (Roche) and used as hybridization probes for plasmids screening in order to test those clones that include the sequence of one of the end of the *mariner* element. The conditions used were hybridization at 65 °C and final wash in 2×SSC at 65 °C.

2.4. RNA purification and RT-PCR

Total RNA was isolated from worker adults (150 mg) using the Eppendorf RNA extraction kit. RNA was suspended in 50 µl of RNAse-free water. Traces of DNA were eliminated by treatment with 10 U of pancreatic RNAse-free DNAseI (Invitrogen) for 1 h at 37 °C.

A sample of 5 µl of the RNA obtained was treated for 1 h at 37 °C with 10 U DNase-free RNase H (Invitrogen). The sample was precipitated and suspended in 100 µl of water. This sample was used as a control for PCRs. A second sample (RT+) was treated as follows: 5 µl of the RNA obtained were mixed with 50 pmol of the oligo (dT)$_{20}$ and with 1 µl dNTP mix (19 mM). Sterile and distilled water was added to a final volume of 13 µl. The mixture was heated to 65 °C for 5 min and incubated on ice for 1 min. Afterwards, 200 U of SuperScript III reverse

transcriptase (200 U/µl) (Invitrogen), 40 U of RNase inhibitor (Amershan), 1 µl of DTT 0.1 M and the corresponding buffer were added to a final volume of 20 µl. The mixture was incubated to 25 °C for 5 min, to 50 °C for 1 h and to 70 °C for 15 min, after which the rest of the RNA was eliminated by treatment with 2 U of DNAse-free RNAase (Invitrogen) at 37 °C for 20 min. This sample containing cDNA was precipitated and suspended in 100 µl of water. A third sample of 5 µl, not reverse transcribed (RT−), was suspended in 100 µl of water and used as control.

PCRs were performed using 5 µl of each of the above samples. For PCR amplification, the primers *trans*-1 (5′-GCTAGTAGAGACTTACGGTGAGC) and *trans*-2 (5′-GTCCATCCCTCAGACAATTTATG) were used. These primers amplify a region of 900 bp of the ORF of the *mariner* transposase. PCRs were performed in 50 µl using 30 pmol of each primer and 3 U of Taq polymerase (Bioline). The PCR program used was 2 min at 92 °C and 40 cycles of 20 s at 92 °C, 1 min at 50 °C, 40 s at 72 °C for 30 cycles. The PCR products were inserted into the pGEMT Easy vector (Promega).

3. Results and discussion

3.1. Isolation of the IRE-130 and mariner elements

The stDNA of ants *M. bouvieri*, *M. barbarus* and *M. structor* has previously been studied and found to be organized as tandemly repeated 79-bp monomers. A 25-bp DNA fragment was found in the 5′ end of a clone of stDNA in the ant *M. bouvieri* (Lorite et al., 2002). New analyses of the stDNA showed the existence of other clones with this 25-bp sequence at either the 5′ or 3′ end. All these sequences interrupted the stDNA at the same point—the trinucleotide TAA (positions 23–25, Fig. 1a). They were very similar to each other, with the 5′ and 3′ copies being inverted with respect to each other. Based on these sequences, we designed a single primer (MEBOTRA, Fig. 1b) in order to characterize the sequences between the inverted repeats. The results of the four PCRs (A, A-1, B, B-1) with the primer MEBOTRA, using genomic DNA from two *M. bouvieri* populations (GI-1 and RU-2), showed of a variable number of amplification products, ranging from less than 500 bp to several kilobases (Fig. 1 in the Appendix). In the present work, we studied the fragments with sizes of 0.5 kb (clone MEBOTRA-500-2 from PCR A), 1 kb (clones MEBOTRA-1-1, -2, -3, -4 and -5 from PCR-B) and 1.4 kb (clones MEBOTRA-1.4-5, -6, -9 and -19 from PCR-A and MEBOTRA-1.4-B6 from PCR-B). We studied the fragments with the sizes indicated, because they correspond to the most prominent and/or coincident amplification bands in all the PCRs performed. The name of the clones includes the size of corresponding fragments and number of the clone (Table 1 in the Appendix). The letter corresponding to the PCR is included only in the last clone (GenBank accession nos. AJ780971–AJ780975, AJ132589, AJ84417 and from AJ781763–AJ781772).

In the sequence of the 0.5-kb fragment (clone MEBOTRA-500-2), three regions can be differentiated (Fig. 2a).The lateral are formed by sequences of about 130 bp, which are very
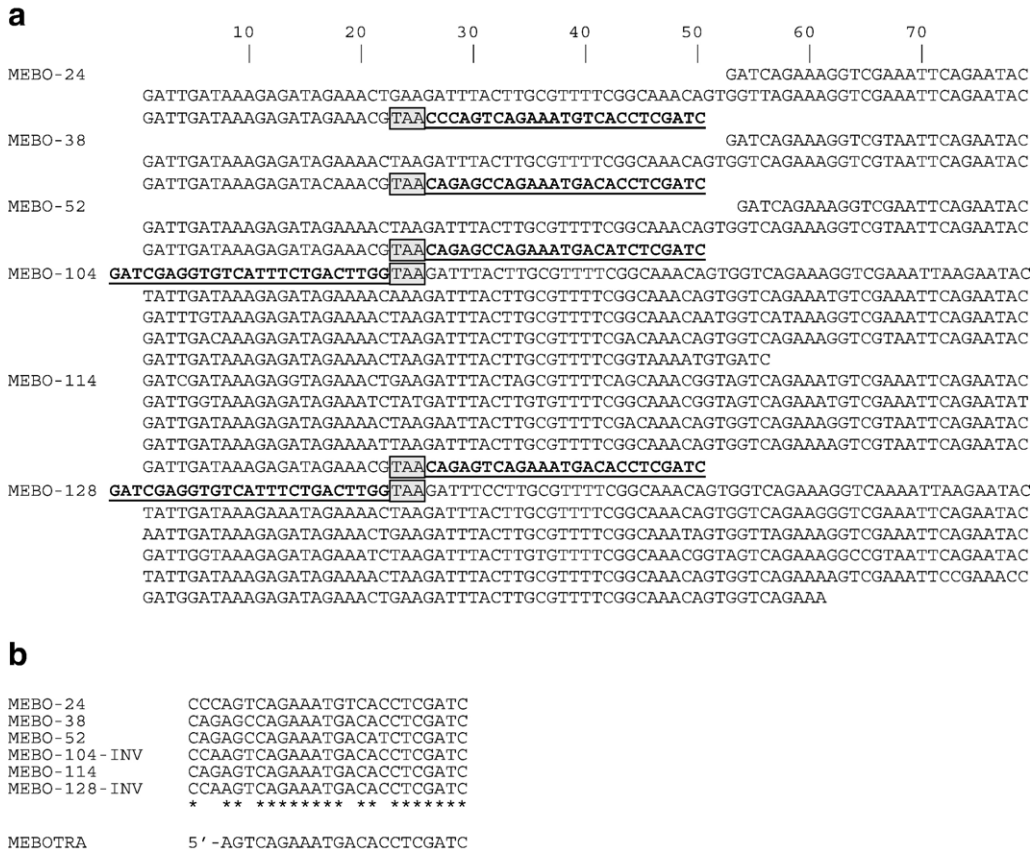
a

```
                10        20        30        40        50        60        70
                 |         |         |         |         |         |         |
MEBO-24                                                    GATCAGAAAGGTCGAAATTCAGAATAC
           GATTGATAAAGAGATAGAAACTGAAGATTTACTTGCGTTTTCGGCAAACAGTGGTTAGAAAGGTCGAAATTCAGAATAC
           GATTGATAAAGAGATAGAAACGTAACCCAGTCAGAAATGTCACCTCGATC
MEBO-38                                                    GATCAGAAAGGTCGTAATTCAGAATAC
           GATTGATAAAGAGATAGAAAACTAAGATTTACTTGCGTTTTCGGCAAACAGTGGTCAGAAAGGTCGTAATTCAGAATAC
           GATTGATAAAGAGATACAAACGTAACAGAGCCAGAAATGACACCTCGATC
MEBO-52                                                    GATCAGAAAGGTCGAATTCAGAATAC
           GATTGATAAAGAGATAGAAAACTAAGATTTACTTGCGTTTTCGGCAAACAGTGGTCAGAAAGGTCGTAATTCAGAATAC
           GATTGATAAAGAGATAGAAACGTAACAGAGCCAGAAATGACATCTCGATC
MEBO-104   GATCGAGGTGTCATTTCTGACTTGGTAAGATTTACTTGCGTTTTCGGCAAACAGTGGTCAGAAAGGTCGAAATTAAGAATAC
           TATTGATAAAGAGATAGAAACAAAGATTTACTTGCGTTTTCGGCAAACAGTGGTCAGAAATGTCGAAATTCAGAATAC
           GATTTGTAAAGAGATAGAAAACTAAGATTTACTTGCGTTTTCGGCAAACAATGGTCATAAAGGTCGAAATTCAGAATAC
           GATTGACAAAGAGATAGAAAACTAAGATTTACTTGCGTTTTCGACAAACAGTGGTCAGAAAGGTCGTAATTCAGAATAC
           GATTGATAAAGAGATAGAAAACTAAGATTTACTTGCGTTTTCGGTAAAATGTGATC
MEBO-114   GATCGATAAAGAGGTAGAAACTGAAGATTTACTAGCGTTTTCAGCAAACGGTAGTCAGAAATGTCGAAATTCAGAATAC
           GATTGGTAAAGAGATAGAAATCTATGATTTACTTGTGTTTTCGGCAAACGGTAGTCAGAAATGTCGAAATTCAGAATAT
           GATTGATAAAGAGATAGAAAACTAAGAATTACTTGCGTTTTCGACAAACAGTGGTCAGAAAGGTCGTAATTCAGAATAC
           GATTGATAAAGAGATAGAAATTAAGATTTACTTGCGTTTTCGGCAAACAGTGGTCAGAAAAGTCGTAATTCAGAATAC
           GATTGATAAAGAGATAGAAACGTAACAGAGTCAGAAATGACACCTCGATC
MEBO-128   GATCGAGGTGTCATTTCTGACTTGGTAAGATTTCCTTGCGTTTTCGGCAAACAGTGGTCAGAAAGGTCAAAATTAAGAATAC
           TATTGATAAAGAAATAGAAACTAAGATTTACTTGCGTTTTCGGCAAACAGTGGTCAGAAGGGTCGAAATTCAGAATAC
           AATTGATAAAGAGATAGAAACTGAAGATTTACTTGCGTTTTCGGCAAATAGTGGTTAGAAAGGTCGAAATTCAGAATAC
           GATTGGTAAAGAGATAGAAATCTAAGATTTACTTGTGTTTTCGGCAAACGGTAGTCAGAAAGGCCGTAATTCAGAATAC
           TATTGATAAAGAGATAGAAAACTAAGATTTACTTGCGTTTTCGGCAAACAGTGGTCAGAAAAGTCGAAATTCCGAAACC
           GATGGATAAAGAGATAGAAACTGAAGATTTACTTGCGTTTTCGGCAAACAGTGGTCAGAAA
```

b

```
MEBO-24       CCCAGTCAGAAATGTCACCTCGATC
MEBO-38       CAGAGCCAGAAATGACACCTCGATC
MEBO-52       CAGAGCCAGAAATGACATCTCGATC
MEBO-104-INV  CCAAGTCAGAAATGACACCTCGATC
MEBO-114      CAGAGTCAGAAATGACACCTCGATC
MEBO-128-INV  CCAAGTCAGAAATGACACCTCGATC
              *   ** ******** ** *******

MEBOTRA       5'-AGTCAGAAATGACACCTCGATC
```

Fig. 1. (a) Multiple-sequence alignments of clones from stDNA of *Messor bouvieri*, showing in all clones a 25-bp sequence (underlined) at the 5′ end or at the 3′ end interrupting to the stDNA at the trinucleotide TAA (in boxes). The 25-bp sequences were very similar to each other with the 5′ and 3′ copies being inverted with respect to each other. (b) Alignment of the 25-bp sequences from the clones represented in (a) and MEBOTRA primer based on these sequences. The 25-bp sequences corresponding to MEBO-104 and MEBO-128 are inverted in this figure.

a

A	B	C
IRE-130	stDNA	IRE-130

b

```
          5' AGTCAGA-AATGACACCTCGATC 3'
500-2A    ---AGTCAGA-AATGACACCTCGATCGACGTCGAATCGACGTCGATTCAATGTCGAAAATAACATGTTGA
500-2C    CCCAGTAATTTAATGACACCTCGATCGACGTCGAATCGACGTCGATTCAATGTCGAAAATAACATGTTGA
          *** *  *****************************************************

                        3' GATCGAGGTGTCATTTCTGACT 5'
500-2A    CATAGAATCTACGTCGATTAGACGTCGATTAGACGTCGAACGAGGTATCATTTCTGACTGGG
500-2C    CATAGAATCGACGTCGATTAGACGTCGATTAGACATCGATCGAGGTGTCATTTCTGACT---
          ******** ********************** **** ****** ************
```

c

```
MEBO     GATTGATAAAGAGATAGAAAACTAAGATTTACTTGCGTTTTCGGCAAACAGTGGTCAGAAAGGTCGAAATTCAGAATAC
500-2B        (500-2-A).....TAAGATTTACTTGGGTTTTCGGCAAACAGTAGTCAGAAAGGCGGTAATGCAGAGTAC
         GATTTATAAAGAGATAGAAAGCTAAGATTTACTTGCGTTTTACGGCAAACAGTGGTCAGAAAGGTCGTAATTCAGAATAC
         GATTGATAAAGAGAATAGAAAACTAAGATTTACTTGCGTTTTCGACAATCAGTGGTCAGAAAGGTCGTAATTCAGAATAA
         AATTGATAAAGAGATAGAAAACTAA....(500-2-B)
```
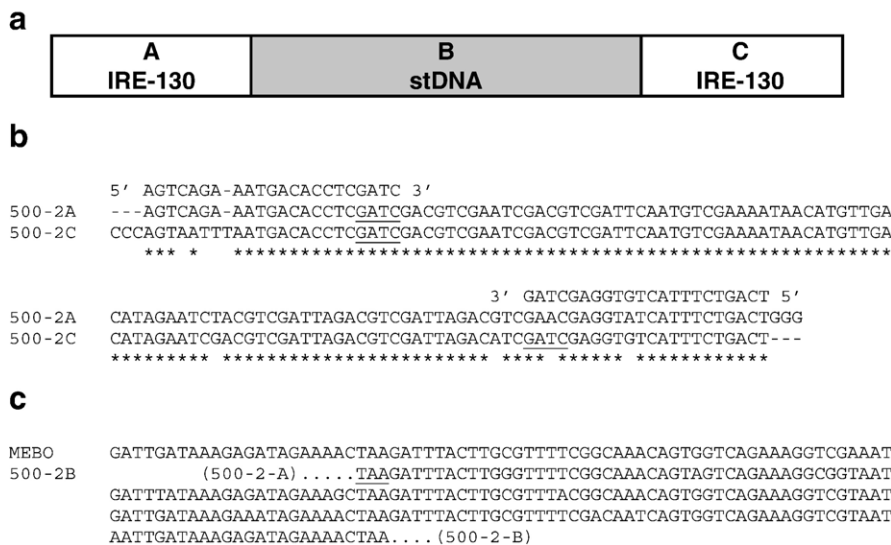
Fig. 2. (a) Schematic representation of the composition of MEBOTRA-500-2 clone, showing three differentiated regions (A, B and C). The lateral zones are formed by sequences of about 130 bp or IRE-130 (interspersed repetitive element of 130 bp). In the middle appear three complete monomers of stDNA. (b) Sequence alignments of the IRE-130 sequences (MEBOTRA-500-2A and MEBOTRA-500-2C). The MEBOTRA primers are also shown. The target sequences for *Sau*3A are underlined. (c) Sequence alignments of monomer consensus sequence of stDNA from *M. bouvieri* (Lorite et al., 2002) and the three monomers of the stDNA (500-2B), showing that the lateral monomers are interrupted by the IRE-130 element. The duplicates of the three nucleotides (TAA) are underlined.

similar (more than 90%) and have a palindrome arrangement (MEBOTRA-500-2A and -2C in Fig. 2b). In the middle of the 0.5-kb fragment, there were three monomers of the stDNA (500-2B in Fig. 2c). The lateral monomers were interrupted by these 130-bp sequences at the same site, the trinucleotide TAA, as in the isolated clones of stDNA (Fig. 1a). These results suggest that the 130-bp sequences are dispersed in the stDNA of this species, being inserted consistently at the same position. As a result of the integration, a duplication of three nucleotides takes place (TAA) (Fig. 2c). We call the interrupted stDNA sequence IRE-130 (interspersed repetitive element of 130 bp). IRE-130 has two targets for Sau3A at its end (Fig. 2b), resulting after the digestion of genomic DNA with this enzyme, in the 25-bp fragments associated to the stDNA (Lorite et al., 2002). The complete results of study of the IRE-130 sequence are analyzed in the following section (see Section 3.2).

Five nucleic-acid sequences with 1-kb sizes were analyzed. The analyses of the sequences of the MEBOTRA-1-1, -2 and -3 clones (Fig. 3a and b) reveal that they were formed by IRE-130 sequences that flank 11 monomers of stDNA. These clones showed a complete IRE-130 at one of their ends; this sequence interrupted a stDNA monomer repeat at the level of the TAA of positions 23–25, as it did previously for other IRE-130 repeats. The sequences corresponding to these IRE-130 elements are shown in Fig. 4 and we call them MEBOTRA-1-1B, -2A and -3B, respectively. These clones have, at the other end (Fig. 3a and b), an incomplete IRE-130 element interrupted at a TA dinucleotide (positions 72–73), the target site of the mariner-like element studied here for its insertion in IRE-130, which we will analyze below. In the clones MEBOTRA-1-1 and -3, the incomplete IRE-130 is followed by stDNA, this beginning in the TA of positions 46–47. Notably, at positions 46–47 of stDNA is the dinucleotide AA in all the monomers studied in this work and in Lorite et al. (2002). Nevertheless, it is possible that a mutation gave rise to the dinucleotide TA. The incomplete IRE-130 of the clone MEBOTRA-1-2 is followed by stDNA repeats but in the TA at positions 77–78. The positions 77–78 of stDNA are one of the insertion sites of the mariner-like element, as we will also analyze next.

Clones MEBOTRA-1-4 and -5 (with 1-kb sizes) are formed by an element IRE-130 with an inserted sequence of 851 bp (Fig. 4). As a consequence of the insertion, a duplication of

14 bp in element IRE-130 took place, including the dinucleotide TA used by the mariner element as a target site (Fig. 4). The 851-bp sequences of both clones were practically identical, with 99.53% of homology (GenBank accession no. AJ781767), and they were not significantly similar to previously described sequences deposited in databases.

The five sequences with 1.4-kb sizes (MEBOTRA-1.4-5, -6, -9, -19 and -B6 clones) are generated by the insertion of a sequence of about 1.3 kb in an IRE-130 element (Fig. 4). The sizes of the inserted sequences ranged between 1275 and 1286 bp. Database comparisons indicated that these sequences have similarities with MLEs, especially those isolated in D. mauritiana or D. simulans (70% and 69.4%, respectively) (Jacobson et al., 1986; Capy et al., 1990). Following the nomenclature of Robertson and Asplund (1996), we call these sequences Mboumar (M. bouvieri mariner-like elements).

The Mboumar sequences were inserted into the IRE-130 element at the same nucleotide position in all the cases, specifically at the TA dinucleotide of the 72–73 positions, generating its duplication. The integration of the Mboumar-9, besides prompting of the duplication typical of TA, triggered the duplication of a 13 bp at the 3′ end of the insertion site. Fig. 4 shows these results and the consensus sequence of the IRE-130 elements analyzed. Altogether, these results could indicate that IRE-130 elements carry a hot-spot for Mboumar insertion. Peculiarly, the 851-bp long sequences were inserted almost at the same location as the Mboumar element and they prompted the duplication of 14 bp. These data could suggest a possible relationship of the sequence of 851 bp with transposable elements.

3.2. Sequence analysis of IRE-130 elements

Several features of the IRE-130 element suggest that it is a new MITE. The IRE-130 element is small and highly conserved (Fig. 4), with an average number of nucleotide substitutions per site in relation to the consensus sequence (Dxy (JC)) of 0.02 ± 0.01. It has been suggested that the high degree of conservation of nucleotide sequence indicates a recent spread of an element within a genome (Akagi et al., 2001). The consensus sequence is almost a perfect palindrome and it has content in A+T of 53.85%, and the theoretical secondary structure is
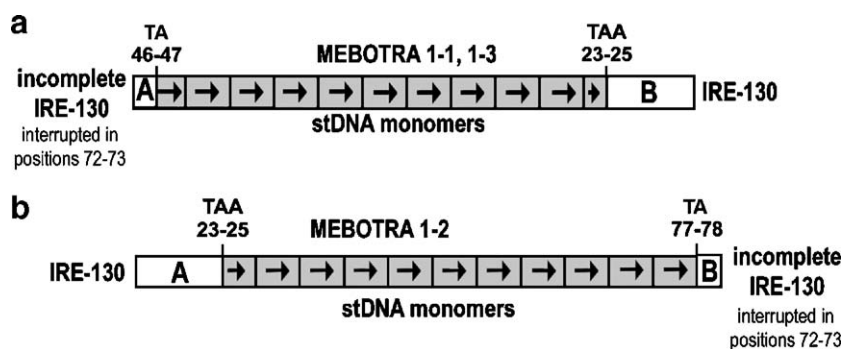

Fig. 3. Schematic representation of the clones MEBOTRA-1-1, MEBOTRA-1-3 (a) and MEBOTRA-1-2 (b). These are formed by 11 monomers of stDNA and an IRE-130 element (MEBOTRA-1-1B, MEBOTRA-1-3B and MEBOTRA-1-2A, respectively) and an incomplete IRE-130 element (A and B, respectively). The dinucleotide TA and the trinucleotide TAA as well as the corresponding numbers indicate the insertion site of IRE-130 in stDNA.
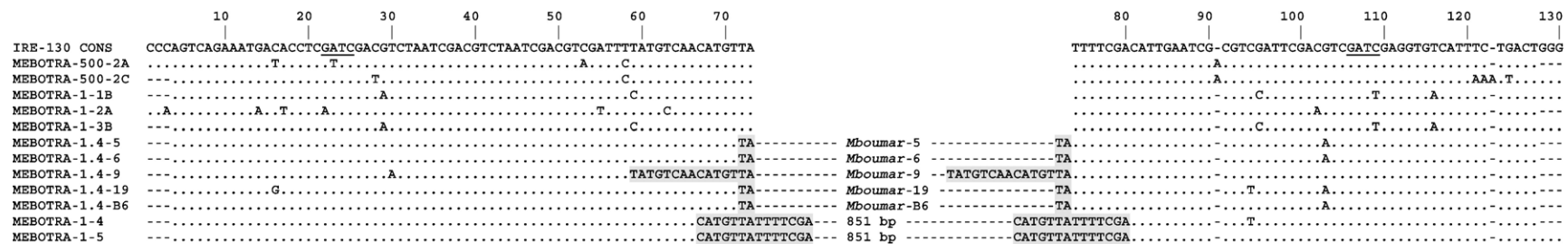
```
                    10        20        30        40        50        60        70                                80        90       100       110       120       130
                    |         |         |         |         |         |         |                                 |         |         |         |         |         |
IRE-130 CONS    CCCAGTCAGAAATGACACCTCGATCGACGTCTAATCGACGTCTAATCGACGTCGATTTTATGTCAACATGTTA                    TTTTCGACATTGAATCG-CGTCGATTCGACGTCGATCGAGGTGTCATTTC-TGACTGGG
MEBOTRA-500-2A  ...................T.....T.........................................A....C..............            ...................A..............................-.........---
MEBOTRA-500-2C  ---...................T.......................................C..............              ...................A..........................AAA.T.....
MEBOTRA-1-1B    ---..................A................................................C..............      .................-....C...............T......A......-.......
MEBOTRA-1-2A    ..A.........A..T....A.......................................T......C..........           ...............-.......A....................................---
MEBOTRA-1-3B    ---..................A................................C..............                    .................-....C...............T......A......-.......
MEBOTRA-1.4-5   ---..................................................................TA---------- Mboumar-5 --------------TA...............................A..................---
MEBOTRA-1.4-6   ---..................................................................TA---------- Mboumar-6 --------------TA...............................A..................---
MEBOTRA-1.4-9   ---..................A...............................TATGTCAACATGTTA---------- Mboumar-9 --TATGTCAACATGTTA.................-.............................---
MEBOTRA-1.4-19  ---..........G.......................................................TA---------- Mboumar-19 -------------TA...............-....T......A..................---
MEBOTRA-1.4-B6  ---..................................................................TA---------- Mboumar-B6 --------------TA...............................A..................---
MEBOTRA-1-4     ---...........................................CATGTTATTTTCGA--- 851 bp -----------CATGTTATTTTCGA.............-....T....................-.......
MEBOTRA-1-5     ---...........................................CATGTTATTTTCGA--- 851 bp ------------CATGTTATTTTCGA.............-........................-.......
```

Fig. 4. Sequence alignments of the all the IRE-130 sequences studied in the different clones and consensus sequence derived from them. The target sequences for *Sau*3A are underlined. Schematic representation of the composition of the clones MEBOTRA-1.4-5, -6, -9, -19 and -B6 clones. The composition of the clones MEBOTRA-1-4 and -5 is also shown. The duplicated TA and other duplicated nucleotide generated by the insertion of the sequences (*Mboumar* and 851-bp sequences) in IRE-130 are shaded.

highly stable (dG=−65.2 to temperature of 25 °C) (Fig. 2a and b in the Appendix). The IRE-130 is a non-autonomous element, since it lacks ORFs.

The sequence of IRE-130 elements does not resemble any sequence deposited in the sequence databases. No relationship with other MITEs known was found using the MAK program, which was developed specifically to locate known MITE family members and to identify related TPase-encoding elements present in the same database.

Most MITES, at least in plants, can be included in the *Tourist*-like or *Stowaway*-like superfamilies, with target-site duplication TAA and TA, respectively. *Tourist*-like MITEs are widespread in plants and animals, whereas *Stowaway*-like MITEs are widespread only in plants (Jiang et al., 2004). The TSD of the IRE-130 element seems to be TAA, as *Tourist*-like MITEs. In the present study, we found IRE-130 elements inserted into stDNA consistently at the same position of the momomers, i.e. the trinucleotide TAA (positions 23–25). As a result of the integration, three nucleotides were duplicated (Figs. 1a, 2c and 3). In addition, the number of analyzed clones with IRE-130 elements inserted into stDNA is quite high. Consequently, at least for these cases, the TSD seems to be TAA. Nevertheless, it would be also possible that the insertion of IRE-130 in the trinucleotide TAA from stDNA represented a single insertion event evolutionarily propagated by subsequent stDNA expansion instead of several independent insertions at the same target. We also found incomplete IRE-130 elements, IRE-130 elements with a MLE or with an 851-bp sequence inserted. In these copies of the IRE-130 element, we do not have information on its flanking sequences and possible TSD.

Most MITEs are considered to be associated with noncoding regions of plant genes, most likely because of a strong insertion preference for low-copy sequences (Jiang et al., 2004). Nevertheless, MITEs with other locations also have been described. For example, *Micron*, a MITE of the rice genome, has as target-site microsatellite loci (Akagi et al., 2001).

As commented above (see Section 1), various authors have suggested that MITES could be cross-mobilized by transposases encoded by autonomous elements such as MLEs. This relationship is based especially on the sequence similarities between inverted terminal repeats of both types of transposable elements. Nevertheless, the sequence-comparison programs (see Section 2.2) show that it does not have sequence similarities between IRE-130 elements and ITRs of the MLEs studied here. Therefore, like other MITEs, the origin and mode of amplification of the IRE-130 element remains unknown.

3.3. Sequence analysis of the mariner element and results of RT-PCR

DNA sequences of the MLEs isolated from *M. bouvieri* show a high degree of conservation, with a homology over 99%. All the clones have 1287 bp except *Mboumar*-5, which has 1286 bp, presenting two point deletions and one insertion. The clone *Mboumar*-9 presents in the 5′-ITR a T, instead of an A, at the nucleotide position 1203 (Fig. 3 in the Appendix). These ITRs of 32 bp show 60% and 63% identity, respectively, with the ITRs from *Mos1* and *Himar1 mariner* transposons, the known active *mariners* most studied. *Mos1* and *Himar1* show 33% identity as well (Lampe et al., 2001). The sequence and the length of the ITRs vary among members of the MLEs, but the terminal nucleotides located just beside the duplicated TA dinucleotide at the insertion site conform to the consensus 5′-YYAGRT-3′ (Langin et al., 1995). Lampe et al. (2001) have performed a sequence-logo analysis of the ITRs of different subfamilies of *mariner* elements. In agreement with these analyses, two regions (positions 3–8 and 14–18) appear to be the most conserved. Conservation of these positions suggests that the transposase might be making base-specific contacts within these regions of the ITRs. Indeed, the most conserved positions in each region (positions 5 and 15) are one helical turn away from each other, suggesting that transposase might be making contact with the DNA on one face of the ITR at two locations of the major groove. The sequence logo of these regions is AGGTBK and WARRK, respectively. All these regions are conserved in the ITR from *Mboumar* (Fig. 4 in the Appendix).

The sequences from *Mboumar*-6, -9, -19 and -B6 clones present ORFs that could encode a putative protein with 345 amino acids (GenBank accession no. AJ781769–AJ781772). However, the existence of deletions and insertions in *Mboumar*-5 originate several stop codons, so that it is a defective copy. These putative proteins showed a high sequence identity (Fig. 5 in the Appendix). Database searches revealed that these proteins shared close protein-sequence identity (67–68%) with the transposase of the *Mos1* from *D. mauritiana*, the first *mariner* element to be described (Jacobson et al., 1986). Fig. 5 shows the sequence and putative protein of *Mboumar*-9.

The *Mos1* transposase has two domains: an N-terminal ITR-binding domain and a C-terminal catalytic domain. The existence of one or two helix-turn-helix (HTH) motifs in this N-terminal domain is discussed. Plasterk et al. (1999) proposed that TLE transposases and probably also MLE transposases have two HTH motifs. Zhang et al. (2001) have reported that a putative helix-turn-helix motif between residues 88 and 108 is required by the DNA-binding activity of *Mos1* transposase. We used a helix-turn-helix motif-prediction program to analyze the protein from *Mos1* and the putative protein studied here. We found that the scores (6.96 for *Mos1* and 5.39 for all putative transposases studied) suggest roughly 100% probability that these proteins contain a HTH motif at 87–108 positions (Fig. 5). It should be emphasized that the transposases studied here differ with respect to *Mos1* in 8 amino-acid positions; nevertheless, they can form a similar motif.

A bipartite nuclear location signal (NLS), RRK and RYKAK (131–133 and 142–146 positions, respectively) was also found to be present in *Mos1*. The middle R of the RRK region is essential for transposase activity (Lohe et al., 1997). The middle R of the RRK region is conserved (Fig. 5). The transposase NLS is flanked by phosphorylation target sites of casein kinase II. Phosphorylation of these sites is a potential checkpoint in the regulation of transposition (Plasterk et al., 1999). *Mos1* has three potential target sites, two of these flanking the NLS (24–27, 154–157 and 170–173 positions, respectively) and three of

```
CCAGGTGTGTCGGTAATTCCTTTCCGGTTTTTCCGGCAGATGTCACTAGCCATAAGTATG   60   TATA box

AAATGTTATGATTTGATACATATGTCATTTTATTCTACTGACATTAACCTTAAAACTACA  120

CAAGTTACGTTCCGCCAAAATAACAGCGTTATAGATTTATAATTTTTTGAAAATGTCGAG  180
                                                        M  S  S    3

TTTTGTGCCTGAAAACGTGCATTTGCGGCACGCATTACTTTTCTTGTTTCATCAAAAGAA  240
 F  V  P  E  N  V  H  L  R  H  A  L  L  F  L  F  H  Q  K  K    23

AAGAGCCGCTGAAAGTCATCGTTTGCTAGTAGAGACTTACGGTGAGCATGCTCCAACCAT  300
 R  A  A  E  S  H  R  L  L  V  E  T  Y  G  E  H  A  P  T  I    43

AAGAACATGTGAAACGTGGTTTCGACAATTCAAATGTGGTGATTTCAACGTTCAAGACAA  360
 R  T  C  E  T  W  F  R  Q  F  K  C  G  D  F  N  V  Q  D  K    63

AGAACGTCCTGGTAGGCCGAAAACGTTTGAAGACGCGGAATTGCAGGAGTTATTGGATGA  420
 E  R  P  G  R  P  K  T  F  E  D  A  E  L  Q  E  L  L  D  E    83

AGACTCAACACAAACTCAAAAACAATTAGCAGAAAAGTTGAATGTGAGCCGAGTAGCAAT  480
 D  S  T  Q  T  Q  K  Q  L  A  E  K  L  N  V  S  R  V  A  I   103  HTH motif (87-108)

TTGTGAACGATTACAAGCGATGGGAAAGATCCAAAAGATGGGAAGATGGGTGCCACATGA  540
 C  E  R  L  Q  A  M  G  K  I  Q  K  M  G  R  W  V  P  H  E   123  conserved aminoacidic
                                                                    motif WVPHEL (119-124)

ATTGAATGACAGGCAAATGGAAAATCGAAAAATCGTCAGTGAAATGCTGCTTCAAAGGTA  600
 L  N  D  R  Q  M  E  N  R  K  I  V  S  E  M  L  L  Q  R  Y   143  bipartite nuclear
                                                                    localization signals
                                                                    (131-133, 142-146)

TGAAAGGAAGTCATTTCTGCATCGAATAGTAACGGGTGATGAAAAATGGATTTATTTCGA  660
 E  R  K  S  F  L  H  R  I  V  T  G  D  E  K  W  I  Y  F  E   163

AAATCCCAAGCGAAAAAAATCATGGTTATCACCTGGTGAAGCCGGCCCATCGACTGCAAG  720
 N  P  K  R  K  K  S  W  L  S  P  G  E  A  G  P  S  T  A  R   183

GCCAAATCGCTTTGGTCGCAAGACCATGCTCTGTGTCGGTGGGACCAGATTGGTGTCGT  780
 P  N  R  F  G  R  K  T  M  L  C  V  W  W  D  Q  I  G  V  V   203

ATATTATGAACTGTTGAAACCTGGAGAAACAGTTAATACTGATCGCTACCGACAACAAAT  840
 Y  Y  E  L  L  K  P  G  E  T  V  N  T  D  R  Y  R  Q  Q  M   223

GATCAATTTGAACTGTGCATTGATCGAAAAACGTCCACAATACGCTCAAAGACATGATAA  900
 I  N  L  N  C  A  L  I  E  K  R  P  Q  Y  A  Q  R  H  D  K   243

AGTGATTTTGCAACATGACAACGCGCCGTCTCATACAGCAAAACCAGTCAAAGAAATGTT  960
 V  I  L  Q  H  D  N  A  P  S  H  T  A  K  P  V  K  E  M  L   263

GAAATCACTTGGATGGGAAGTTTTATCGCACCCGCCGTACTCTCCAGACCTGGCTCCATC 1020
 K  S  L  G  W  E  V  L  S  H  P  P  Y  S  P  D  L  A  P  S   283  conserved aminoacidic
                                                                    motif YSPDLAP (176-183)

CGACTACCACCTTTTCGCATCGATGGGGCACGCGCTTGCAGAGCAGCACTTCGCCGATTT 1080
 D  Y  H  L  F  A  S  M  G  H  A  L  A  E  Q  H  F  A  D  F   303

CGAAGAAGTTAAAAAATGGCTCGATGAATGGTTTAGCTCAAAGGAGAAACTGTTCTTTTG 1140
 E  E  V  K  K  W  L  D  E  W  F  S  S  K  E  K  L  F  F  W   323

GAATGGCATTCATAAATTGTCTGAGAGATGGACAAAATGTATAGAATCAAATGGCCAATA 1200
 N  G  I  H  K  L  S  E  R  W  T  K  C  I  E  S  N  G  Q  Y   343

CTTTGAATAAATAAATTTGAACTATCCATTCTAAGTAACGTGTTTTCTTTAACGAAAAAA 1260  poli-adenilation
 F  E  *                                                      345  signal

CCGGAAAAGAATTACCGACACTCCTGG                                  1287
```

Fig. 5. Sequence of the *Mboumar*-9 element. The ITRs are underlined. The putative TATA box and polyadenylation signal are doubly underlined. The putative protein is also shown. The numbers on the right are, respectively, the number of nucleotides and amino acids. The HTH motif and other important protein regions are shown on the amino-acid sequence. A key signature of the *mariners*, the amino acids of the catalytic D,D(34)D motif are boxed.

the putative transposases studied have five potential target sites, three of these flanking the NLS (35–38, 71–74, 154–157, 173–176 and 315–318 positions, respectively).

The C-terminal catalytic domain in active transposases of the *mariner* groups invariably contains the conserved D,D(34)D motif (Robertson, 1996; Lohe et al., 1997). The transposases studied also present this conserved motif (Fig. 5). As an exception in *Mboumar*-B6 element, the first aspartic residue was replaced by a glycine, resulting in a motive G,D(34)D (Fig. 5 in the Appendix). A few MLEs are considered active in vivo and are believed to encode transposase with two highly conserved amino-acid motifs, WVPHEL and YSPDLAP(I/S/T) separated by about 150 amino-acid residues (Robertson and

MacLeod, 1993; Augé-Gouillou et al., 2005b). Both motifs are conserved in the transposases studied (Fig. 5).

In addition, an examination of the 5′ and 3′ untranslated region with the appropriate computer program (see Section 2.2) revealed a putative TATA box at position 57 and a polyadenylation signal at position 1206 (Fig. 5).

The transcription of the *M. bouvieri* MLEs was checked by reverse transcription with oligonucleotides dT on RNA extracted from workers. The reverse transcripts were amplified by PCR with two primers designed to amplify a region of approximately 900 bp from the sequence of the *mariner* element isolated. PCRs were performed on three samples: one treated with RNase (negative control), another one with reverse
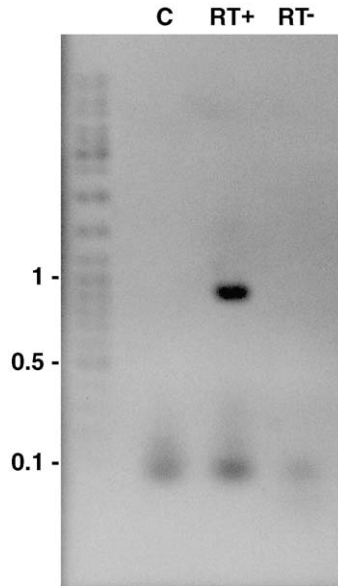
Fig. 6. A 1% agarose gel showing the amplification bands obtained by RT-PCR. Lane 1: sample treated with RNase (negative control). Lane 2: sample with reverse transcriptase (RT⁺), showing an amplification band of about 900 bp. Lane 3: sample without reverse transcriptase (RT⁻), also used as control and without amplification band.

transcriptase (RT+) and finally another sample without reverse transcriptase (RT−) also used as control (see Section 2.4) (Fig. 6, lanes 1, 2 and 3, respectively). A single band appeared in the second sample (Fig. 6, lane 2). Amplification bands on lanes 1 and 3 did not appear. The result for the negative control demonstrated that there was no DNA in RNA extract. The result for the sample without reverse transcriptase (RT−) also demonstrated the absence of DNA. The data sequenced (RT-PCR-900-1) from the PCR products showed a sequence with 99.5% identity with the corresponding region of the *mariner*

element (GenBank accession no. AJ784418 and Fig. 6 and Table 1 in the Appendix).

Altogether, the data presented here strongly suggest that *Mboumar* is an active transposon. However, the MLEs from the ant *M. ruginodis* and the bumble bee *Bombus terrestris* do not have an open reading frame similar to that of the active element *Mos1*, although they conserve some of the characteristic amino-acid motifs of active transposases (Bigot et al., 1994). The putative transposase from *S. invicta* (*Sinvmar1*) contains five frameshifts and two stop codons, indicating that it is a defective copy (Krieger and Ross, 2003). *Acmar1*, a MLE from *Apis cerana japonica* (Hymenoptera, Apocrita) contained a sequence with the motif D,D(34)D but with a deletion of five amino-acid residues in the frame within the motif as compared to the corresponding sequence of *D. mauritiana Mos1* (Sumitani et al., 2002). All the characteristics of the *Mboumar* elements analyzed appear to indicate that it is an active element; nevertheless, this needs to be confirmed by in vivo transposition assays or in vitro excision assays.

3.4. Study of the insertion sites of mariner elements in M. bouvieri and of its presence in other species of the Messor genus

The results so far presented do not enable us to determine whether *Mboumar* are directly inserted into the stDNA, or whether the presence of IRE-130 is necessary. To determine the possible insertion of *mariner* elements directly in the stDNA, we carried out PCRs using primers based on the sequence of both types of DNA. Two primers based on the stDNA (MEBOSAT-1 and -2) and two based on the *mariner* sequence (MAR-1 and -4) were designed with this purpose (Fig. 7a). PCRs were performed using the four-primer combinations. We identified several amplification bands ranging between 100 bp and 500 bp.
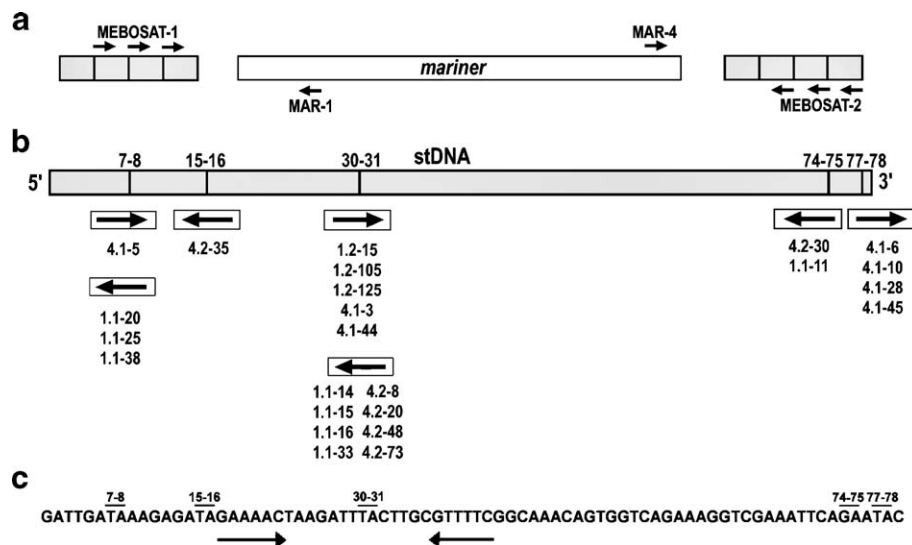

Fig. 7. (a) Schematic representation of the location of the primers designed to determine the insertion site of *Mboumar* elements in the stDNA. (b) Schematic representation of the insertion site of the *Mboumar* elements on the monomer stDNA. The arrows indicate the direction of the insertion of this element. The numbers indicate the corresponding clones. (c) Consensus-sequence monomer stDNA from *M. bouvieri* showing the position of target site of the *Mboumar* elements. The arrows under the consensus sequence indicate the inverted repeat of 6 bp.

A total of 24 clones were selected for further studies. Eight of these clones were obtained using the primers MAR-1/MEBOSAT-1 (MEBOMARSAT-1.1 clones), three with MAR-1/MEBOSAT-2 (MEBOMARSAT-1.2 clones), seven with the primers MAR-4/MEBOSAT-1(MEBOMARSAT-4.1 clones) and six with MAR-4/MEBOSAT-2 (MEBOMARSAT-4.2 clones) (GenBank accession no. are from AJ783743–AJ783753 and from AJ784404–AJ784416; Figs. 7, 8, 9 and 10, and Table 1 in the Appendix). All clones presented an end of *Mboumar* and several monomers of stDNA. The IRE-130 element was not detected in any of the clones, indicating that the *mariner* element can integrate itself directly into the stDNA without requiring the presence of the IRE-130 element.

Mboumar elements insert into the stDNA in either direction but invariably at the level of a dinucleotide TA. We detected five TA dinucleotides in the stDNA in which *Mboumar* was inserted (Fig. 7b and c). In the stDNA consensus sequence, one of these positions (74–75) was GA (Fig. 7c). The variability of stDNA by mutation could cause the change of the first G to T, giving a TA that could act as a target by the *mariner* element integration. Zhang et al. (2001) have studied mutant transposases that allow the insertion at dinucleotides other than TA, including sequences with GC base pairs. Therefore, it is also possible that a *mariner* element could be inserted into a dinucleotide GA.

With greater frequency the *Mboumar* integrates at positions 30–31 (Fig. 7b) located in the middle of a degenerate palindrome sequence of 26 bp with an inverted repeat of 6bp at its ends (Fig. 7c). Bigot et al. (1994) reported that Hymenopteran MLEs lie at a specific insertion site. In agreement with this contention, these elements would be inserted into the host genomes, in the middle of a degenerated 30-bp palindrome, which is itself located in a 85-bp conserved region with a purine-rich tail at one of these ends. The *Mboumar* element is inserted at different positions into the stDNA but only a degenerated palindrome exists at one of these positions. On the other hand, the *Mboumar* elements are inserted into different sequences, as, for example, into IRE-130 sequences. Nevertheless, the IRE-130 consensus sequence is almost a perfect palindrome, as analyzed above (see Section 3.2). Consequently, it is possible that target-site selection is determined primarily at the level of DNA structure and not by specific base-pair interactions, as suggested Vigdal et al. (2002).

We used inverse PCR and partially unspecific PCR to determine other DNA flanking the *Mboumar* elements. By inverse PCR two sequences of 115 and 354 bp containing a *mariner* element were isolated (*Mbmar*-INV-PCR-L8,-16; GenBank accession no. AJ784419 and AM158970) (Fig. 11 in the Appendix). In one of the clones (*Mbmar*-INV-PCR-L8), *Mboumar* had lost its 5′-ITR and part of its 5′-UTR, possibly as a consequence of an aberrant integration. By partially unspecific PCR, several sequences flanking the 5′ or 3′ end of the *Mboumar* were also isolated (*Mbmar*-PCR clones, GenBank accession no. from AM158971–AM158978) (Figs. 12 and 13 in the Appendix). Neither of the cloned fragments was found to present the IRE-130 elements or stDNA. The flanking DNA

sequences isolated are not significantly similar to previously described sequences deposited in databases. The present results indicate that *Mboumar* elements exist outside of the stDNA array.

To verify whether IRE-130 and *Mboumar* elements are present in *M. barbarus* and *M. structor*, we performed PCRs using primers based on the ITRs from the *Mboumar* elements studied in the present work and also the MEBOTRA primer. Amplification bands of about 1.3 kb were found using the first primer and about 1.4 kb using the second primer. This result suggests the existence of a *mariner*-like element (1.3 kb) that could be associated with an element similar to IRE-130 (1.4 kb) in these species of the genus *Messor* (Fig. 14 in the Appendix).

In order to determine whether the *mariner*-like element is also inserted directly into the stDNA from *M. barbarus* and *M. structor*, we carried out PCRs using a primer based on the stDNA (MEBOSAT-1) and another based on the *mariner* sequence (MAR-1). All sequenced clones (MEBAMARSAT and MESTMARSAT, respectively) showed an end of a *mariner*-like element and several monomers of stDNA (Fig. 15 and Table 1 in the Appendix).

3.5. Potential impact of Mboumar mariner-like and IRE-130 elements on satellite DNA evolution

In this study, we found IRE-130 elements and *Mboumar* elements inserted into stDNA and, IRE-130 elements with an inserted *Mboumar* element at the same nucleotide position in all the cases (see Section 3.1). To our knowledge, this is the first time that a *mariner* element inserted into a MITE has been detected. On the contrary, MITEs inserted within fixed copies of non-autonomous *P* elements have been reported (Holyoake and Kidwell, 2003). Nevertheless, the MEBOTRA-1-1, -2 and -3 clones are perhaps of most interest in relation to the evolution of the stDNA.

MEBOTRA-1-1, -2 and -3 clones (Fig. 3) have an incomplete IRE-130 element interrupted at a TA dinucleotide (positions 72–73), target site of the *Mboumar* element for its insertion in IRE-130. We propose that this peculiar configuration of sequences around of the dinucleotide TA could be the evolutionary vestige where two *mariner* elements had been inserted. The incomplete IRE-130 of the clone MEBOTRA-1-2 is followed by stDNA repeats in the TA at positions 77–78. This position is also one of the target sites of a *mariner*-like element in stDNA. We suggest that the sequence of the MEBOTRA-1-2 clone could have originated in the following way: evolutionarily, two *mariner* elements could have existed, one inserted directly into the stDNA (positions 77–78) and the other inserted in an IRE-130 element, as shown in Fig. 8. At a certain time in evolution, an excision process could have occurred, involving the 5′-ITR of one *mariner* and the 3′-ITR of the other, acting as a single transposable element and dragging both elements and the stDNA sequences that could exist between the two *mariners* (Fig. 8). In the clones MEBOTRA-1-1 and -3, the incomplete IRE-130 is followed by stDNA, this beginning in the TA of
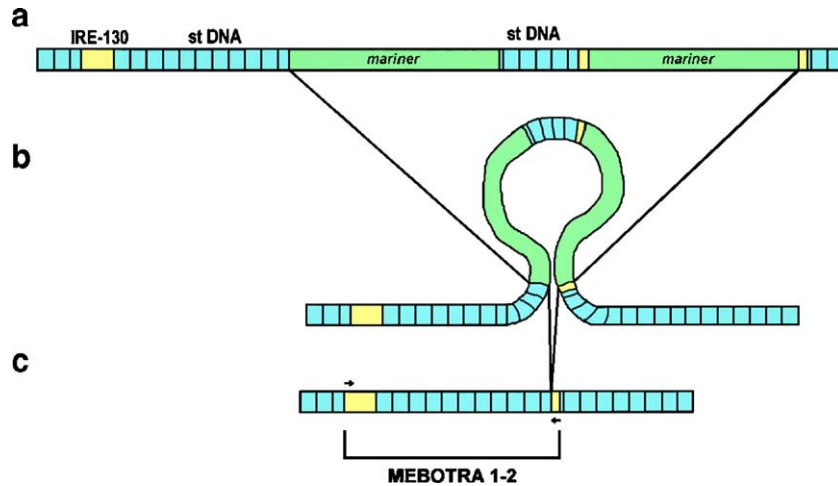
Fig. 8. Possible evolutionary origin of the MEBOTRA-1-2 clone. (a) Schematic representation of a hypothetical segment of stDNA with an IRE-130 element and a *mariner* element inserted in positions 23–25 and 77–78 respectively (left in the scheme) and with other *mariner* element inserted in positions 72–73 of an IRE-130 element (right in the scheme). (b) Representation of an excision process involving the 5′-ITR of one *mariner* and the 3′-ITR of the other, acting as a single transposable element. (c) Representation of the hypothetical segment of DNA originated by the excision process represented that could include the sequence corresponding to the clone MEBOTRA-1-2. As the MEBOTRA-1-2 clone, this region is formed by monomers of stDNA and two IRE-130 elements, one of them complete (inserted into 23–25 positions from the stDNA) and another incomplete (inserted in 77–78 positions from the stDNA). The incomplete IRE-130 element would be interrupted at a TA dinucleotide (positions 72–73), target site of the *mariner*-like element studied here for its insertion into IRE-130.

positions 46–47. Consequently, a similar excision process, although with a *mariner* element inserted into the stDNA but in positions 46–47, could have given rise the sequences of these two clones. Positions 46–47 could represent a new insertion site of the *mariner* in the stDNA. The stDNA located between the two *mariners* could even have been at a different location from the chromosome and/or in a different chromosome. The joint split of two *mariner* elements may have been favored by mutations in the ITRs, or by anomalous insertions that caused a deletion at the level of some of the ITRs, as seems to have occurred in the clone *Mbmar*-INV-PCR-L8 (see Section 3.3), in which a deletion of 39 bp has completely eliminated the ITR in one of the ends of *mariner* elements.

As indicated above (see Section 1), the stDNA from *M. barbarus*, *M. bouvieri* and *M. structor* shows high interspecific sequence conservation, just as happens in other ant species (Lorite et al., 2004). The existence of MLEs elements has also been shown in stDNA from *M. barbarus* and *M. structor*. Probably, IRE-130 is also present in these species. Several hypotheses have been advanced to explain the sequence conservation of a stDNA family. In agreement with Dover (2002), the sequence conservation could be achieved as the result of some highly efficient homogenization mechanism, including gene conversion and unequal crossing-over, transposition, slippage replication and RNA-mediated exchanges. The possible genomic mobility of stDNA repeats through transposable elements could be an important molecular mechanism in the conservation of these sequences.

Thompson-Stewart et al. (1994) reported that the subtelomeric repeats were mobilized, frequently altering the number of repeats, upon the excision of the transposable *P* element from tandemly repeated sequences. Pons (2004) has suggested that transposition could be one of the putative molecular mechanisms involved in the strong amplification and/or homogeniza-

tion of the *Pst*I stDNA family (with motifs similar to those found in transposable element) from the tenebrionid *Misolampus goudoti*. Similarly, the mobilization of the repeat could contribute to the conservation and expansion of the stDNA from the genus *Messor*.

4. Conclusions

In summary, we conclude that:

(1) Two transposable elements, concretely *Mboumar* (MLE) and IRE-130 (MITE) are inserted into stDNA from *M. bouvieri*.
(2) The *Mboumar* element has also been detected to be inserted into IRE-130 element and consistently in the same target site. On the contrary, *Mboumar* is inserted into different target sites from stDNA.
(3) The study of the *Mboumar* element, of its transcription and the putative transposase that it would encode, suggests that it could be an active element.
(4) The *mariner*-like element is also present in stDNA from *M. barbarus* and *M. structor*.
(5) The study of certain clones analyzed in this work suggests that the possible genomic mobility of stDNA repeats through transposable elements could be an important molecular mechanism in the conservation of stDNA. The stDNA from three species of the *Messor* genus show a high similarity and all belong to the same family of satellite DNA.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2005.11.032.

References

Akagi, H., Yokozeki, K., Inagaki, A., Mori, K., Fujimura, T., 2001. *Micron*, a microsatellite-targeting transposable element in the rice genome. Mol. Genet. Genomics 266, 471–480.

Augé-Gouillou, C., Brillet, B., Hamelin, M.H., Bigot, Y., 2005a. Assembly of the *mariner* Mos1 synaptic complex. Mol. Cell. Biol. 25, 2861–2870.

Augé-Gouillou, C., Brillet, B., Germon, S., Hamelin, M.H., Bigot, Y., 2005b. *Mariner* Mos 1 transposase dimerizes prior to ITR binding. J. Mol. Biol. 351, 117–130.

Bigot, Y., Hamelin, M.H., Capy, P., Periquet, G., 1994. *Mariner*-like elements in hymenopteran species, insertion site and distribution. Proc. Natl. Acad. Sci. U. S. A. 91, 3408–3412.

Bigot, Y., Brillet, B., Augé-Gouillou, A., 2005. Conservation of palindromic and mirror motifs within inverted terminal repeats of mariner-like elements. J. Mol. Biol. 351, 108–116.

Capy, P., Chakrani, F., Lemeunier, F., Hartl, D.L., David, J.R., 1990. Active mariner transposable elements are widespread in natural populations of *Drosophila simulans*. Proc. R. Soc. Lond., B Biol. Sci. 242, 57–60.

Casacuberta, J.M., Santiago, N., 2003. Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes. Gene 311, 1–11.

Dover, G., 2002. Molecular drive. Trends Genet. 18, 587–589.

Feschotte, C., Zhang, X., Wessler, S.R., 2002. Miniature inverted repeat transposable elements and their relationship to established DNA transposons. In: Craig, N.L., Craigie, R., Geller, M., Lambowith, A.M. (Eds.), Mobile DNA II. ASM Press, Washinton, DC, pp. 1147–1158.

Feschote, C., Osterlund, M.T., Peeler, R., Wessler, S.R., 2005. DNA-binding specifity of rice *mariner*-like transposases and interactions with *Stowaway* MITES. Nucleic Acids Res. 33, 2153–2165.

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

Holyoake, A.J., Kidwell, M.G., 2003. Vege and Mar, two novel hTA MITE families from *Drosophila willistoni*. Mol. Biol. Evol. 20, 163–167.

Jacobson, J.W., Medhora, M.M., Hartl, D.L., 1986. Molecular structure of a somatically unstable transposable element in *Drosophila*. Proc. Natl. Acad. Sci. U. S. A. 83, 8684–8688.

Jiang, N., Feschotte, C., Zhang, X., Wessler, S.R., 2004. Using rice to understand the origin and amplification of miniature inverted repeat transposable elements (MITEs). Curr. Opin. Plant Biol. 7, 115–119.

Karlyshev, A.V., Pallen, M.J., Wren, B.W., 2000. Single-primer PCR procedure for rapid identification of transposon insertion sites. BioTechniques 28, 1078–1082.

Kidwell, M.G., 2002. Transposable elements and the evolution of genome size in eukaryotes. Genetica 115, 49–63.

Krieger, M.J.B., Ross, K.G., 2003. Molecular evolutionary analyses of mariners and other transposable elements in fire ants (Hymenoptera, Formicidae). Insect Mol. Biol. 12, 155–165.

Lampe, D.J., Walden, K.K.O., Robertson, H.M., 2001. Loss of transposase-DNA interaction may underlie the divergence of *mariner* family transposable elements and the ability of more than one mariner to occupy the same genome. Mol. Biol. Evol. 18, 954–961.

Langin, T., Capy, P., Daboussi, M.J., 1995. The transposable element Impala, a fungal member of the Tc1-*mariner* superfamily. Mol. Gen. Genet. 246, 19–28.

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

Lorite, P., Carrillo, J.A., Tinaut, A., Palomeque, T., 2002. Comparative study of satellite DNA in ants of the *Messor* genus (Hymenoptera, Formicidae). Gene 297, 113–122.

Lorite, P., Carrillo, J.A., Tinaut, A., Palomeque, T., 2004. Evolutionary dynamics of satellite DNA in species of the genus *Formica* (Hymenoptera, Formicidae). Gene 332, 159–168.

Plasterk, R.H., Izsvák, Z., Ivics, Z., 1999. Resident aliens, the Tc1/*mariner* superfamily of transposable elements. Trends Genet. 15, 326–332.

Pons, J., 2004. Cloning and characterization of transposable-like repeat in the heterochromatin of the darkling beetle *Misolampus goudoti*. Genome 47, 769–774.

Robertson, H.M., 1996. Members of the pogo superfamily of DNA-mediated transposons in the human genome. Mol. Gen. Genet. 252, 761–766.

Robertson, H.M., Asplund, K.L., 1996. *Bmmar*1, a basal lineage of the *mariner* family of transposable elements in the silkworm moth, *Bombyx mori*. Insect Biochem. Mol. Biol. 26, 945–954.

Robertson, H.M., 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.

Sumitani, M., Lee, J.M., Hatakeyama, M., Oishi, K., 2002. Cloning and characterization of *Acmar*1, a *mariner*-like element in the Asiatic honey bee, *Apis cerana japonica* (Hymenoptera, Apocrita). Arch. Insect Biochem. Physiol. 50, 183–190.

Thompson-Stewart, D., Karpen, G.H., Spradling, A.C., 1994. A transposable element can drive the concerted evolution of tandemly repetitious DNA. Proc. Natl. Acad. Sci. U. S. A. 91, 9042–9046.

Vigdal, T.J., Kaufman, C.D., Izsvák, Z., Voytas, D.F., Ivics, Z., 2002. Common physical properties of DNA affecting target site selection of sleeping beauty and other Tc1/*mariner* transposable elements. J. Mol. Biol. 323, 441–452.

Zhang, L., Dawson, A., Finnegan, D.J., 2001. DNA-binding activity and subunit interaction of the *mariner* transposase. Nucleic Acids Res. 29, 3566–3575.

Zhang, X., Jiang, N., Feschotte, C., Wessier, S.R., 2004. *PIF*- and *Pong*-like transposable elements, distribution, evolution and relationship with *Tourist*-like miniature inverted-repeat transposable elements. Genetics 166, 971–986.