Recent amplification of miniature inverted-repeat transposable elements in the vector mosquito *Culex pipiens*: characterization of the *Mimo* family

Cédric Feschotte, Claude Mouchès *

Laboratoire Ecologie Moléculaire and Faculté Sciences et Techniques Côte-Basque, Université de Pau et des Pays de l’Adour, BP 1155, 64 013 Pau, France

Received 10 February 2000; accepted 12 April 2000

Received by W.-H. Li

Abstract

We describe a new family of repetitive elements, named *Mimo*, from the mosquito *Culex pipiens*. Structural characteristics of these elements fit well with those of miniature inverted-repeat transposable elements (MITEs), which are ubiquitous and highly abundant in plant genomes. The occurrence of *Mimo* in *C. pipiens* provides new evidence that MITEs are not restricted to plant genomes, but may be widespread in arthropods as well. The copy number of *Mimo* elements in *C. pipiens* (~1000 copies in a 340 Mb genome) supports the hypothesis that there is a positive correlation between genome size and the magnitude of MITE proliferation. In contrast to most MITE families described so far, members of the *Mimo*-family share a high sequence conservation, which may reflect a recent amplification history in this species. In addition, we found that *Mimo* elements are a frequent nest for other MITE-like elements, suggesting that multiple and successive MITE transposition events have occurred very recently in the *C. pipiens* genome. Despite evidence for recent mobility of these MITEs, no element has been found to encode a protein; therefore, we do not know how they have transposed and have spread in the genome. However, some sequence similarities in terminal inverted-repeats suggest a possible filiation of some of these mosquito MITEs with *pogo*-like DNA transposons. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Genome evolution; MITEs; *pogo*; Repetitive DNA; Transposons

1. Introduction

Recently, numerous families of short (100–500 bp) transposable elements with terminal inverted repeats have been found in high copy numbers in a wide range of plant genomes (Bureau and Wessler, 1992, 1994a; Bureau et al., 1996; Pozuelo-Romero et al., 1996; Casacuberta et al., 1998; Song et al., 1998; Charrier et al., 1999). Based on common structural features, these elements have all been designed as MITEs (miniature inverted-repeat transposable elements). Some mobile elements similar in structure to the plant MITEs have been described in fungi (Yeadon and Catcheside, 1995) and in some vertebrates (Unsal and Morgan, 1995; Smit, 1996; Izhâk et al., 1999). In invertebrates, typical MITEs have been characterized from only two insect species: the yellow fever mosquito, *Aedes aegypti* (Tu, 1997), and the beetle, *Tenebrio molitor* (Braquart et al., 1999), whereas none have been described from the more extensively studied genome of *Drosophila melanogaster*. This distribution suggests that there might be a correlation between the level of MITE proliferation and the size and organization of the host genome; the larger it is, the higher the MITE copy number could be (Tu, 1997; Casacuberta et al., 1998; Braquart et al., 1999; Charrier et al., 1999).

To date, the mechanism of transposition of MITEs remains unknown. Several of their characteristics, including TIRs and short (2–3 bp) specific duplications generated upon insertion, suggest that MITEs may have spread by DNA transposition. This hypothesis is further
supported by the recent evidence that one plant MITE family has originated from a larger element encoding a transposase (Feschotte and Mouchès, 2000). However, so far, no excision event has been reported for a MITE. Furthermore, it is unlikely that DNA transposition, which by itself is a non-replicative process, is responsible for such a high level of proliferation in genomes (Izsvák et al., 1999; Feschotte and Mouchès, 2000).

Here, we report that the *C. pipiens* mosquito, an insect with a relatively large genome and a high amount of repetitive DNA, does contain MITEs. Structural analysis of one family, Mimo, indicates that it has a recent amplification origin and may have originated from a pogo-like DNA transposon, as previously demonstrated for a plant MITE family (Feschotte and Mouchès, 2000).

### 2. Material and methods

#### 2.1. Mosquito strains and genomic DNAs

The first Mimo-Cp1 copy was identified in a λ clone previously isolated from a genomic library of the Tem-R strain of *C. pipiens* (California, USA). All other copies are from the Ravenna strain of *C. pipiens* (Italy). For Southern experiments, we also used genomic DNAs from the following insects: *C. pipiens* Willow (California), *C. pipiens* Pro-R (California), *C. pipiens* cells (Taiwan), *C. hortensis* (collected in the field, South of France), *Aedes aegypti* (collected in the field, South of France), *A. triseriatus* cells (Trois Rivieres, Canada), *Anopheles stephensi* (obtained from the MNHN Paris), *Drosophila melanogaster* (Canton), and *Musca domestica* (collected in the field, South of France).

#### 2.2. Southern blot analysis

DNA was prepared from adult mosquitoes, as described previously (Mouchès et al., 1986). Aliquots of 10 μg of genomic DNA were digested to completion with EcoRI restriction endonuclease. Resulting fragments were separated on 1% agarose gels, transferred to a nylon membrane (Amersham) and hybridized at high stringency (65°C) with radiolabelled probes. Other procedures were as previously described (Mouchès et al., 1990). Probes were obtained by PCR amplification from plasmids carrying a copy of the element using a degenerate primer corresponding to the TIR (see below), gel-purified and labelled with α²P-dCTP by random priming (Pharmacia).

#### 2.3. Isolation of Mimo copies from a *C. pipiens* genomic library and estimation of copy number

A library was prepared by complete EcoRI digestion of genomic DNA from the Ravenna strain of *C. pipiens* and ligated into a λ-gt11 cloning vector (Stratagene). About 20,000 recombinant phages were plated and screened using a Mimo-Cp1 probe obtained by PCR amplification of the original Mimo element found in CM-gag4, a copy of the retroposon CM-gag (Bensaad-Merchermek et al., 1997) (see Section 3.1). Prehybridization, hybridization and washing were carried out at 65°C, as previously described (Mouchès et al., 1990). After a first round of screening, a large number of positives were obtained. Several positive plaques were plugged in SM buffer and amplified, and each was used as a template for PCR amplification with universal primers for the arms of λ-gt11 vector. PCR parameters were the same as those described below, except that the annealing temperature was reduced to 54°C and the elongation time increased to 2 min 30 s.

PCR products containing Mimo-Cp1 copies were subcloned into pcRII-TOPO plasmid vectors (Invitrogen). Copy number for Mimo-Cp1 elements was estimated based on the ratio of positive phage plaques to the total number of plaques screened, taking into account the haploid genome size of *C. pipiens* of 840 Mb (Black and Rai, 1988) and an average genomic library insert size of 4 kb.

#### 2.4. Additional Mimo copies isolated by PCR amplification

PCR reactions were performed according to standard procedures using a single oligonucleotide, 5′-CART-AGTTGTTCCGTAACTKG-3′, which is complementary to TIRs of the Mimo-Cp1 element. Cycling conditions were an initial 5 min at 94°C followed by 30 cycles of 30 s at 94°C, 90 s at 55°C, 60 s at 72°C and a final 10 min elongation at 72°C. To assess Mimo length variation, additional PCR reactions were performed with increased elongation steps up to 2 min. PCR products were analysed on agarose gels and cloned directly into pcRII-TOPO vectors (Invitrogen). Positive clones were identified by Southern hybridization with the Mimo-Cp1 probe.

#### 2.5. Sequence analysis

Sequencing was done by the Eurogentec sequencing department with synthetic primers, using an ABI-377 automatic sequencer. Most of the sequence analysis was done using tools available at the Infobiogen WWW server (http://www.infobiogen.fr). Database searches
were performed with BLASTN (Altschul et al., 1990), using default parameters. Multiple sequence alignments were constructed by CLUSTAL W version 1.7 (Thompson et al., 1994) using default parameters. Pairwise alignments of nucleotide sequences were done using the ALIGN program (Myers and Miller, 1988) of the FASTA package. The ability of elements to form secondary structures was estimated by the DNA FOLD program (Zuker, 1994) available through the server http://mfold.wustl.edu/, and free energies were determined according to SantaLucia (1998). Sequences reported in this article appear in GenBank under the Accession Nos AF217611 to AF217616.

3. Results

3.1. Mimo, a family of highly reiterated elements in the genome of C. pipiens

The first Mimo element was discovered during the analysis of a genomic clone from the Tem-R strain of C. pipiens, which contains a copy of the retroposon CM-gag (Bensaadi-Merchermek et al., 1997). Sequence alignment of this copy with the canonical CM-gag1 element (Accession No. AF030588) shows that it is interrupted by a 348 bp insertion at position 1678 (Fig. 1). The presence of almost perfect TIRs and flanking 2 bp direct repeats suggests that this insertion may be a transposable element. A computer-assisted search in current DNA databases using this insertion sequence as a query failed to detect any similarity with previously described sequences. However, it shares a strong structural similarity with the newly described class of short transposons called MITEs (see Section 3.2). We have named this insertion sequence Mimo-Cp1 for MITE mosquito.

Mimo-Cp1 was used as a probe to investigate the presence of related sequences in several dipterous species. Southern hybridizations with various genomic DNA digests reveal a long continuous smear for all tested strains of C. pipiens (data not shown). This demonstrates that Mimo-Cp1 belongs to a family of highly reiterated elements in this species. No hybridization signal was detected in other insects, including Aedes and Anopheles mosquito species. All C. pipiens strains contain about the same copy number of Mimo elements, but some variations in the banding pattern suggest that several locations are polymorphic between strains. The copy number of the Mimo MITEs in the C. pipiens genome was estimated by screening a genomic library from the Ravenna strain with Mimo-Cp1 as a probe. Based on the ratio of positive plaques to the total number of plaques screened and assuming a haploid genome size for C. pipiens of 540 Mb (Black and Rai, 1988), the copy number of Mimo elements is ~1000 per haploid genome. This is probably an underestimation since several Mimo copies can be present in the same phage clone.

3.2. Mimo is a novel family of MITEs with a high sequence conservation

The variability of Mimo family was analysed by characterization of several copies. Using Mimo-Cp1 as a probe, we screened a genomic library from the Ravenna strain of C. pipiens. Two positive phage clones were randomly chosen and further characterized. Each genomic clone contained a Mimo copy with both structural and sequence similarity to the original Mimo-Cp1 element (Fig. 2).

Furthermore, we used PCR amplification of genomic DNA with a degenerate primer for the TIRs to assess length heterogeneity among the Mimo copies. Numerous PCR fragments, ranging from approximately 300 to 400 bp, were obtained from Ravenna genomic DNA and subsequently cloned. Twelve of them were randomly chosen for further characterization. All clones strongly hybridize with the Mimo-Cp1 probe under high-stringency conditions, which further indicates that there is a high sequence homogeneity between Mimo elements. Ten of the 12 copies are also homogeneous in size, ranging from 310 to 380 bp. The shortest of these clones (Mimo-Cp6, 311 bp) contains several large deletions in the central region of the element (Fig. 2) but a similar substitution rate compared to other Mimo copies (data not shown). Two PCR clones were significantly longer, with 520 and 707 bp, respectively. Sequence analysis for these two clones reveals that they contain ‘regular’ Mimo copies (Mimo-Cp4 and Mimo-Cp5 in Fig. 2) interrupted by other mobile elements (see Section 3.3). The pairwise
sequence identity between Mimo copies is high, ranging from 74.3 to 95.8%.

A consensus sequence (Fig. 2) was calculated from the six Mimo copies (i.e., three elements obtained from genomic clones and three copies isolated by PCR amplification). Mimo elements are A+T-rich (~60%) and have no coding capacity. They possess 23 bp TIRs and additional short subterminal inverted repeats, so they could form stable secondary structures (ΔG values in the range of −53.5 to −68.6 kcal/mol). Analysis of the borders of the three Mimo copies isolated from the genomic libraries reveals that they are flanked by TA duplications (Fig. 2). This indicates that Mimo elements may insert preferentially at TA sites in the C. pipiens genome and that the target sequence is probably duplicated upon insertion. Since these features fit well with those defining the plant MITE families (Wessler et al., 1995), we conclude that Mimo elements belong to a novel family of highly conserved MITEs.

3.3. Some Mimo elements contain other MITE-like elements

Two of the 12 Mimo copies randomly chosen among PCR clones obtained from C. pipiens genomic DNA are interrupted by other putative transposable elements. Insertions occurred at two distinct positions in these copies (indicated in Fig. 2) but, interestingly, both have been inserted in Mimo elements after a CCA site. Comparison of Mimo-Cp4 and Mimo-Cp5 with other Mimo elements (Fig. 3) suggests that both insertion events were followed by a duplication of the dinucleotide CA. Southern experiments using each of these insertion sequences as a probe against C. pipiens genomic DNA show that these elements are highly reiterated in its genome (data not shown).

The first of these elements, named Mint1 (miniature nested transposon), was recognized as a 141 bp insertion sequence in Mimo-Cp4. Unlike classical MITEs, Mint1
do not have recognizable TIRs but possesses nearly perfect (2 mismatches out of 37 bp) subterminal inverted-repeats (SIRs) so that it can be folded into a stem-loop structure (Fig. 3A). The absence of TIRs has recently been reported for two miniature transposon-like elements of the rice genome, Pop and Crackl (Song et al., 1998), that also have SIRs and the potential to form secondary structures. Although we assume that additional Mint sequences are needed to confirm this unusual structure, Mint1 probably belongs to this new type of MITEs. Mint-Cp1 is interrupted by another putative transposon that we have named Nemo1 (for nested mosquito MITE). This element is 324 bp long and contains imperfect 25 bp TIRs (Fig. 3B). Like Mimo and Mint1, Nemo1 is A+T-rich and does not appear to encode a protein. Because it has several SIR motifs, it could presumably form stable secondary structures (AG = −53.3 kcal/mol). Based on these features, it is likely that Nemo1 is also a member of a new MITE family.

Mimo, Mint and Nemo elements have no significant sequence identity to each other or to any other known transposable elements. However, some striking similarities, in both size and sequence, exist between TIRs and target site duplications of Mimo and Nemo and those of several MITEs and Class II transposons (Fig. 4). This indicates a possible filiation of these C. pipiens MITEs with DNA transposons (see Section 4.4).

---

**Fig. 3.** MITE-like elements nested in Mimo elements. (A) Polymorphism corresponding to Mint1 insertion into Mimo-Cp4 and sequence of Mint1 MITE-like element. Subterminal inverted-repeats are boxed. (B) Polymorphism corresponding to Nemo1 insertion into Mimo-Cp5 and sequence of Nemo1 MITE-like element. TIRs are boxed. In each multiple alignment, dots denote identity, and dashed lines indicate gaps corresponding to the element insertion and one copy of the presumed target site. Putative target site duplications are underlined.

---

**Fig. 4.** Homologies in TIRs and target site duplications between some of the mosquito MITEs and DNA transposons of the pogo family. Alignment of the TIRs sequences was constructed by eye. Mimo, Nemo1, Wujin (Tu, 1997) and MERs (II) (Smit and Riggs, 1996) can all be associated with a pogo-like Class II element (PLE). According to TIR similarities, Mimo is closer to pogo, Nemo and Wujin are closer to Lemi1, and MERs(II) are closer to Tiggers. For each group, conserved bases are highlighted by white letters on a black background. TIRs, target site duplication (TSD) sequences and length of Wujin and Lemi/Emigrant and Mimo families are deduced from their consensus sequences (Tu, 1997; Casacuberta et al., 1998; Feschotte and Mouche’s, 2000 and this work). Since only one copy of Nemo1 is available, and since it possesses imperfect TIRs, we have aligned the 5 ’ TIR of Nemo1, but we have indicated mismatches in lowercase letters. MERs(II) represents a general consensus based on a simple majority rule that we have determined for the second group of human MERs, i.e., MER20, MER28, MER2, MER44, MER6, MER66 and MER7 (Smit and Riggs, 1996). N indicates a highly variable nucleotide. Other information on human transposons is from Smit and Riggs (1998), and data on pogo elements are from Tudor et al. (1992).
4. Discussion

4.1. Occurrence of the Mimo family of MITEs in the C. pipiens genome

We have shown that Mimo is a family of repetitive elements from the mosquito C. pipiens with a high structural and sequence homogeneity. Mimo sequences do not share a sequence similarity with any transposable element described to date, but they do possess all the characteristics that define the novel class of transposons called MITEs (Wessler et al., 1995). These include a small and homogeneous size, the presence of TIRs and the absence of coding capacity. Moreover, Mimo elements can be potentially folded into stable secondary structures, another frequent characteristic of MITEs. It is noteworthy that, to date, very few MITE families have been described in animals, while it is argued that MITEs are the most prevalent type of transposon associated with flowering plant genes (Wessler, 1998). To our knowledge, C. pipiens is only the third invertebrate species for which a highly reiterated MITE family is described. This provides new evidence that MITEs are not restricted to plant genomes, but occurred in arthropods also.

We estimate the copy number of Mimo elements to be ~1000 per haploid genome. This is a lower copy number than those generally reported for other MITE families. For example, the Tourist family is present at 10^5 copies in the maize genome (Bureau and Wessler, 1992), and MITE families described from the yellow fever mosquito, Aedes aegypti, are reiterated at ~2000 copies per genome (Tu, 1997). The DEC element is also present at a high copy number in the beetle Tenebrio molitor (~3500 copies, Braquart et al., 1999). This has led some authors to propose that the proliferation of MITEs may be associated with large and complex genomes, in both plant and animal kingdoms (Bureau and Wessler, 1992; Tu, 1997; Braquart et al., 1999; Charrier et al., 1999). In this case, the copy number of Mimo elements is in the range of those from the yellow fever mosquito, since the genome of C. pipiens is approximately 1.5-fold smaller than that of A. aegypti (540 against 810 Mb). Studies using reassociation kinetics have shown that Aedes and Culex, both members of the Culicines tribe, possess complex genomes with a short-interspersion pattern of repetitive elements (Black and Rai, 1988). It was also demonstrated that there has been a general increase in genome size during the evolution of Culicines mosquitoes, with all classes of repetitive DNA increasing linearly in amount with total genome size (Black and Rai, 1988). Our results are consistent with this previous work, showing that the genome of C. pipiens, like those of A. aegypti, does contain MITEs but in lower amounts. This difference could reflect a higher level of proliferation of MITEs, and probably other classes of transposable elements, in Aedes than in Culex and/or differing rates of ‘junk’ DNA elimination between these mosquito species, as recently shown between Drosophila fruit-fly and Laupala crickets (Petrov et al., 2000).

4.2. Evidence for a recent amplification history of Mimo

The first member of the Mimo family was identified as an insertion sequence in a copy of another repetitive element, the CM-gag retroposon. This insertion event, as well as a putative target site duplication, provides evidence for past mobility of this Mimo copy. Furthermore, the CM-gag retroposons are highly conserved, indicating a recent amplification of these retroelements (Bensaadi-Merchermek et al., 1997). The CM-gag4 copy also shares a high sequence similarity with other CM-gag elements (from 86 to 90%), which suggests a very recent insertion event of Mimo-Cp1 in CM-gag4. In addition, we show that Mimo elements isolated from C. pipiens are highly conserved in both terminal and internal sequences, even between two distinct geographical isolates and also when secondary insertions of mobile elements have occurred within them. This is in contrast with most MITE families described earlier. Generally, members of the same family share a good conservation in structure and in terminal sequences but are generally more divergent in internal sequences, with a sequence variation up to 30% being frequent between two members of the same family (Bureau and Wessler, 1994a,b; Pozueta-Romero et al., 1996; Rio et al., 1996; Smit and Riggs, 1996; Charrier et al., 1999; Izsâék et al., 1999). We conclude, therefore, that amplification of Mimo elements is relatively recent in the C. pipiens genome. It is possible that some Mimo elements may still be able to transpose.

4.3. MITEs may be a nest for other MITEs

Among the 12 Mimo copies randomly isolated by PCR using the TIR as a primer, we report that two copies contain nested transposable elements, named Mtm1 and Nemo1. Although we acknowledge that this approach is not quantitative, the ratio of secondary insertions in Mimo copies (16.6%) could be higher because the PCR cycling conditions that we used (1 min extension) have probably enhanced amplification of the shorter Mimo elements rather than larger copies. Nested MITEs in other MITEs were also reported in Sorghum bicolor where Stowaway-Sb1 was discovered as an insertion in Tourist-Sh5 (Bureau and Wessler, 1994a) and in A. aegypti with the Wujin-Aa1 element being found in the Wukong-Aa5 copy (Tu, 1997). Thus, it is possible that MITEs could be a preferential insertion site for other MITEs. Nested insertions in Mimo elements further support that recent and multiple transposition
bursts of various MITE-like elements occurred in the genome of *C. pipiens*.

### 4.4. Mimo and Nemo elements could have originated from pogo-like transposons

The transposition mechanism of MITEs has not yet been elucidated. Due to several structural similarities with DNA transposons as well as some sequence similarity in TIRs, it has been suggested that MITEs could transpose via a DNA intermediate by using transposase activity encoded elsewhere in the genome by another element (Bureau and Wessler, 1994a; Unsal and Morgan, 1995; Smit and Riggs, 1996; Teschendorf and Mouchès, 2000). This suggests that a similar mechanism for the origin of MITEs exists in distinct eukaryote genomes. In this regard, it is striking that *Mimo* and *Nemo1* as well as *Wujin*, a MITE family from the *A. aegypti* genome (Tu, 1997), display significant TIRs similarities with pogo-like transposons (Fig. 4). In addition, insertion of these elements is probably followed by a 2 bp target site duplication that appears to be specific (TA) for *Mimo* and *Wujin* elements. This feature is reminiscent of the TA target site preference of pogo-like elements and other members of the Tc1/mariner superfamiliy (van Luenen et al., 1994; Hartl et al., 1997; Plasterk et al., 1999). It is therefore tempting to speculate that *Mimo* and *Nemo1*, from *C. pipiens*, as well as the *Wujin* family from *A. aegypti*, are pogo-like transposon derivatives. In this case, pogo-like elements may have resided, at least at an ancient time, in their genomes. We now have to investigate the presence of such elements in several mosquito genomes to extend the DNA transposon origin of MITEs to these families.

### Acknowledgements

We thank P. Blanchard, Y. Gibert and S. Karama for technical assistance. We are grateful to F. Brunet and J. O’Connor for critical reading of the manuscript. C.F. was supported by a grant from the Ministère de l’Éducation Nationale, de la Recherche et de la Technologie à Université Paris 6.

### References


