Molecular characterization of *mariner*-like elements in *Bruchus pisorum* and *Bruchus rufimanus* (Coleoptera: Bruchidae)

Salma Djebbi[1], Wiem Ben Amara[1], Hanem Makni[1,2], Mohamed Makni[1] and Maha Mezghani-Khemakhem[1,*]

[1] *Unité de Recherche Génomique des Insectes Ravageurs des Cultures d'intérêt agronomique (UR11ES10), Faculté des Sciences de Tunis, Université de Tunis El Manar*, 2092 El Manar, Tunis, Tunisia

[2] *Institut Supérieur de l'Animation pour la Jeunesse et la Culture (ISAJC), Université de Tunis*, Bir El Bey, Tunisia

*Corresponding author: mahakm@planet.tn

Abstract: *Mariner*-like elements (MLEs) are Class-II transposons that are widely present in diverse organisms and encode a D,D34D transposase motif. MLE sequences from two coleopteran species, *Bruchus pisorum* and *B. rufimanus* were obtained using the terminal-inverted repeats (TIRs) of *mariner* elements belonging to the *mauritiana* subfamily as primer. The characterized elements were between 1073 and 1302 bp in length and are likely to be inactive, based on the presence of multiple stop codons and/or frameshifts. A single consensus of MLE was detected in *B. pisorum* and was named *Bpmar1*. This element exhibited several conserved amino acid blocks as well as the specific D,D(34)D signature. As for *B. rufimanus*, two MLE consensuses, designated *Brmar1* and *Brmar2,* were isolated, both containing deletions overlapping the internal region of the transposase. Structural and phylogenetic analysis of these sequences suggested a relatively recent origin of *Bpmar1* versus a more ancient invasion of *Brmar1* and *Brmar2* in their respective host genomes. Given that MLEs are potential mediators of insect resistance and have been used as vectors to transfer genes into host genomes, the MLEs characterized in this study will have valuable implications for selecting appropriate transposable elements in transgenesis.

Key words: *Mariner*-like element; Coleoptera; *Bruchus pisorum*; *B. rufimanus; mauritiana* subfamily

INTRODUCTION

Transposable elements (TEs) are mobile genetic elements that have the ability to replicate and spread in host genomes. This fluidity leads to modifications of the gene structure and genome architecture [1]. Transposable elements have been traditionally classified into two classes, namely RNA-mediated (Class-I) and DNA-mediated (Class-II) elements, according to their transposition mode [2].

Mariner-like elements (MLEs) are Class-II transposons with 28-30bp inverted terminal repeats flanking a single open reading frame (ORF) coding for a transposase of approximately 350 amino acids. The MLEs' transposase contains two highly conserved motifs, WVPHEL and YSPDLAP, separated by approximately 150 amino acids, as well as a specific D,D(34)D signature motif [3]. Originally isolated from *Drosophila mauritiana* as an insertion in the white eye gene

[4,5], MLEs were subsequently identified in a wide range of animal [6] and plant genomes [7]. Currently, MLEs are clustered into five subfamilies: *cecropia, elegans/briggsae, irritans, mauritiana* and *mellifera/capitata* based on their phylogenetic relationships [8,9]. However, Rouault et al. [10] proposed a method based on hierarchical clustering and average linkage to sort 935 MLEs into 15 subfamilies, including the five major subfamilies previously described.

Most MLEs were found to be inactive as a consequence of the presence of stop codons and/or frameshifts within the coding region [11]. Even though some identified MLEs were found to encode full-length transposases in invertebrate and vertebrate species, only a few showed enzyme activity following genetic analysis for transposition [12]. To date, only three MLEs have been found to be naturally active in some insect species: *Mos1* from *Drosophila mauritiana* [13]; *Famar1,* isolated from the European earwig, *For-*

ficula auricularia [14] and *Mboumar-9*, identified in the ant *Messor bouvieri* [15].

MLEs' mobility made them become potential tools for transgenesis and mutagenesis in a wide variety of organisms, including insects [12]. So far, several species have been transformed using the functional *Mos1* element, and its ability to introduce genes has been well elucidated in *Aedes aegypti* [16,17]. Mathur et al. [18] have successfully transferred antipathogen effector molecules into the salivary glands of *Aedes aegypti* to block the dengue virus transmission using a *Mos1*-derived vector.

All of these transformations were based on exogenous transposable elements because the presence of endogenous elements related to those used in transformation vectors raises the problem of the potential cross-mobilization of the elements and the subsequent effects on the stability of the transformed systems [19].

Bruchus pisorum and *B. rufimanus* are important coleopteran storage insect pests that cause significant losses in pea and faba bean, respectively [20,21]. Insecticides have been commonly used to prevent grain losses; however, larval feeding within seeds limits the chemical insecticides' effects. Moreover, this chemical control is not cost-effective and is associated with concerns related to environmental pollution and food safety, which increases the need for alternative control approaches. One of the methods involves the use of TEs to integrate sterility genes in males reared in the laboratory before spreading them into natural populations to reduce their size [19]. Thus, the characterization of TEs in the coleopteran genomes could be used as a valuable biotechnological tool to promote such genetic control methods.

To date, among Coleopteran insects, only *mariner*-like elements belonging to the *irritans* subfamily have been described in *Agrilus planipennis* and more recently, partial MLEs were identified in *Coprophanaeus cyanescens*, *C. ensifer*, *Diabroctis mimas* and *Dichotomius schiffleri* (Coleoptera: Scarabaeidae) [12,22].

Given the importance of *Mos1* element as a genetic tool to transfer genes into host genomes, we were interested in investigating the presence of *mariner*-like elements of the *mauritiana* subfamily in *B. pisorum* and *B. rufimanus* genomes. Results of the study shed light on the genome structure of these pests and provide insight into the potential use of *Mos1* element to control them.

MATERIALS AND METHODS

PCR amplification

Genomic DNA was extracted from two individual insects of each *Bruchus* species using the Doyle & Doyle protocol [23]. To authenticate the studied *Bruchus* species, we sequenced their DNA barcode region (COI). The COI mitochondrial DNA region was amplified for all samples using the universal primers designed by Folmer et al. [24], LCO-1490: 5´TTTCTACAAATCATA-AAGATATTGG3´ and HCO-2198 5'TGATTTTTT-GGTCACCCTGAAGTTTA3´, and PCR conditions described by Mezghani et al. [25]. The full-length MLEs specific to the *mauritiana* subfamily were amplified by the degenerate *Mos1* primer 5'TAYCAGGRGTA-CAAGTAKGRAA3' described by Kharrat et al. [26].

Amplifications were performed in 25μl, using 25 ng of template DNA, 20 pMol of the degenerate primer, 0.2 mM of each dNTP, 1.5-2.5 mM $MgCl_2$ and 1 unit of Go Taq DNA Polymerase (Promega) in the provided buffer (5X). PCRs were performed using the following program: an initial denaturation at 94°C for 2 min, followed by 35 cycles at94°C for 30 s, 48°C for 30 s and 72°C for 1 min. A final extension step was carried out at 72°C for 10 min. The PCR products were separated on a 1% agarose gel, stained with ethidium bromide and visualized under UV light. PCR products were purified using spin columns (Wizard PCR Preps, Promega) and cloned into a pGEM-T Easy vector (Promega). Plasmids were extracted (Wizard Minipreps, Promega) and sequenced in both directions using the primers T7 and SP6 on an automated sequencer (ABI PRISM 3100 Genetic Analyzer, Applied Biosystems).

Sequence analysis

Mitochondrial COI sequences were submitted to Barcode of Life Data system version 3.0 BOLD to assign species names, then deposited in GenBank under accession numbers: KU982562-KU982565. For *mariner*-like elements, the homology analysis was

Table 1: Similarity index among the nucleotide sequences of the MLEs identified in two *Bruchus* species calculated using Bioedit software. Clones *Bpmar1.1-Bpmar1.6* were obtained from *B. pisorum* and *Brmar1.1-Brmar1.6* were identified in *B. rufimanus*

Clones	Bpmar1.2	Bpmar1.3	Bpmar1.4	Bpmar1.5	Bpmar1.6	Brmar1.1	Brmar1.2	Brmar1.3	Brmar1.4	Brmar1.5	Brmar1.6
Bpmar1.1	0.993	0.993	0.992	0.942	0.990	0.794	0.792	0.794	0.702	0.704	0.702
Bpmar1.2		0.993	0.990	0.943	0.990	0.793	0.793	0.793	0.701	0.701	0.702
Bpmar1.3			0.990	0.941	0.990	0.792	0.792	0.790	0.701	0.700	0.700
Bpmar1.4				0.942	0.990	0.793	0.792	0.792	0.700	0.702	0.702
Bpmar1.5					0.940	0.800	0.798	0.800	0.712	0.713	0.714
Bpmar1.6						0.790	0.790	0.789	0.700	0.701	0.701
Brmar1.1							0.992	0.991	0.662	0.665	0.666
Brmar1.2								0.994	0.664	0.663	0.664
Brmar1.3									0.661	0.665	0.665
Brmar1.4										0.992	0.993
Brmar1.5											0.993
Brmar1.6											-

performed using BLASTX in the NCBI server (www. ncbi.nlm.gov/cgibin/BLAST). The similarity index among nucleotide sequences was calculated with Bioedit [27]. Nucleotide sequences were translated into their presumed amino acid sequences by ExPASy (http://web.expasy.org/translate/), and HTH motifs within them were identified using the GYM 2.0 program [28]. Sequence alignments were performed using GeneDoc and a phylogenetic tree was constructed using MEGA software version 7.0.14 based on the ML method [29]. Bootstrap values for the branches were obtained with 1000 replications. The identified MLE sequences were deposited in the DNA Data Bank of Japan (DDBJ:http://www.ddbj.nig.ac.jp/) under accession numbers LC144637-LC144648.

RESULTS

Taxonomic identification of *Bruchus* samples

Each *Bruchus* sample was clearly identified at the species level as either *B. pisorum* or *B. rufimanus*, using the BOLD identification engine, with a maximum identity of 96-99% based on the COI sequence.

Amplification of full-length MLEs in *Bruchus pisorum* and *B. rufimanus*

PCR products between 1000 and 1300 bp were obtained from the two *Bruchus* species and three clones were sequenced from each individual. MLEs of *B. pisorum* were named *Bpmar1.1-Bpmar1.6* and those

of *B. rufimanus* were named *Brmar1.1-Brmar1.6,* following the nomenclature of Robertson and Asplund [30]. All elements had a total length ranging between 1073 and 1302bp.

MLE sequence analysis

To investigate the characteristics of the *mariner* elements in both *Bruchus* species, MLE nucleotide sequences obtained from all of the *Bruchus* clones were aligned and the similarity index was estimated (Table 1).

The results showed that MLE sequences obtained from *B. pisorum* were similar at both the intra- and interindividual levels, with identities higher than 94% (Table 1). Due to this high similarity, a single consensus sequence of 1302 bp, named *Bpmar1*, was constructed. However, MLE sequences generated from *B. rufimanus* differed according to the individuals from which they were derived, with a similarity index ranging from 66.1% to 99.4% (Table 1). Two consensus sequences, named *Brmar1* and *Brmar2* of 1221bp and 1073bp, respectively, were constructed.

Database searches in GenBank using BLASTX revealed that the three consensus sequences best matched the black garden ant *Lasius niger mariner* element, *Lnmar1*. *Bpmar1* shared 70% amino acid identity with *Lnmar1* (KMQ85296.1), while *Brmar1* and *Brmar2* shared only 40% and 67% identity, respectively.

Alignment of the consensus sequences with the naturally active *Mos1* element of *Drosophila mauritiana* (X78906) are shown in Fig. 1. The consensus se-

ITR-5' and other structural annotations (ITR-5', Met, HTH motif, W V P H E L, NLS bipartite, T G D E K W, NLS monopartite, H D N A, Y S P D L A P S D, Polyadenylation signal, ITR-3') appear above the alignment.

```
                    ITR-5'                                                                                          Met
Mos1    TACCAGGTGTACAAGTAGGGAATGTCGGTTCGAACATATA--G----------------ATGTCTCGCAAACGTAAATATTTATCGATTGTCATAAAAC   81
Bpmar1  ..T...G.........T....--C...AA.TG....C..TGATTCTCTAGCCGAACTAT.CA.T.A.T...AA.......------G.CTG.GCC.TT.   92
Brmar2  .............T....--C...AA.TG....C..TGATTCGCTATCCAAACTAT.C..TCA.T...AA.....G.------G.C.G.GCC.TT.   92
Brmar1  ................A..--C...A.TG....C...GATTCGCTAGCCAAACTAT.C.GTCA.T...AA.....G.------..C.A.CCC.TT.   91

Mos1    TTTGACCTTGTGAAGTGTCAACCTTGACTGTCGAACCACCA---TAGTTTGGCGCGAATTGAGCGTCATAATTGTTTACTCTCAGTGCAGTCAGTCATGTC  178
Bpmar1  ..CAGTA.CAGC.TC.......AA.A.T.TGTTT..A.A..ATCA...C.CATA.A-----A.AA.G.GG..T...T.G...------..A..A.....  180
Brmar2  G..TTTA.CAAC.TC......-AA.A.TC.GTA..TA.A..ATCA..A..CATA..T.-C.TA.AA.G..G..T...T.------------..A..A...  180
Brmar1  ...ATTA.CAAC.TC......AA.A.TCAGTTT.TA.A..ATCC.....CATA.A-----A.A.TG.GC..T...T.G...------..A..A.....  179

Mos1    GAGTTTCGTGCCGAATAAAGAGCAAACGCGGACAGTATTAATTTTCTGTTTTCATTTGAAGAAAACAGCTGCGGAATCGCACCGAATGCTTGTTGAAGCC  278
Bpmar1  ..A...T.....T.CA..CT.CG.TTT.......TC.C.TC....T....AC..A..A.............C.A....A..T...T.......G.....T  280
Brmar2  ..A...T.....T.CA..CTGCG.T.T.......TC....G....AC..A...........G..CA..A....A..T...T.......G.....TT  280
Brmar1  A.A.G.T.....T.CA..CT.CG.T.T.......TC...G..G....AC..A..........C..C..A....A..T...T.......G...T.T  279

Mos1    TTTGGCGAACAAGTACCAACTGTGAAAACGTGTGAACGGTGGTTTCAACGCTTCAAAAGTGGTGATTTTGACGTCGACGACAAAGAGCACGGAAAACCGC  378
Bpmar1  .ACA....G..T..TATTTGCAAA..C.CA...CT.TGAA.......A.TAAA.........C...........T..GA.AA.TG....A.GT....GG..A.  380
Brmar2  .AC.A...G..T.GT.TT.G.AAA..C.CA..TCT.TGAA......A.TAA--.........TA.TAGAA..G...CA.GT.....G..A.  377
Brmar1  .------------------------------------------------T.................T..G.AA.TG....A.GT.....G..A.  327

                                                                                HTH motif
Mos1    CAAAAAGGTACGAAGACGCCGAACTGCAAGCATTATTGGATGAAGACGATGCTCAAACGC AAAAACAACTCGCAGAGCAGTTGGAA GTAAGTCAACAA  476
Bpmar1  ....G.AA.TT.....T......T.AG..........A....TCCATGC.....A........G..T.....GC....A.T...G.TA......  478
Brmar2  .G..G-------------------------------------------------.  ----------------------------------------  382
Brmar1  A...T.AA.TT.....T......T.A........A.........TTCGTG......A.A....A....T......GCC...A..G..T.CA.....  427

                                                            W  V  P  H  E  L
Mos1    GCAGTTTCCAATCGCTTGCGAGAGATGGGAAAGATTCAGAAGGTCGGTAGATGGGTGCCACATGAGTTGAACGAGAGGCAGATGGAGAGGCGCA AAAAC  575
Bpmar1  A.GA..........A..A..T.CC.......G.....TA..ATG.A.A.AG.....T.......A....CT..A.......CAA..ACACA.A.....C.  577
Brmar2  ----------------------------------------------------------------------------------------------.  ----  382
Brmar1  ...A....T.....A...AT.TC........T...TTA..AT--.A.GA......T.........A....CT..A..A..CCAA..A.AC..G.G....C.  525

            NLS bipartite                                  T  G  D  E  K  W              NLS monopartite
Mos1    ACATGCGAAATTTTGCTTTCACGATACAAAAGGAAGTCGTTTTTGCATCGTATCGTTACTGGAGATGAAAAATGGATCTTTTTTGTTAATCCTAAACGTA  675
Bpmar1  .TT..T......C....CG.C..G........A..A...A............C..C.....G.....T.A...C.AG.....C..A.  677
Brmar2  ----------CC....CG.GA.........AA..A..A.A...A...G.A............CC.........T.A....AG....TA...TA.  470
Brmar1  .TT..T.....GC.....G.CA.G........A...AA.A...A.G...G.....T...C.......A..T.A...CCAGG..T.C.....A.  625

Mos1    AAAAGTCATACGTTGATCCTGGACAACCGGCCAC-----------ATCGACTGCTCGACCGAATCGCTTTGGCAAGAAGACGATGCTCTGTGTTTGGTG  763
Bpmar1  .....G.....TA....C..CA.T......T.A..ATCAAGGTCTAT......C..AAA...A......CA.A...T....A.........TC..  777
Brmar2  .....A....T......C..C..T.....TT.A.........C..AAA...A.......C.AA....C.A....G......C..  558
Brmar1  .....G-----------------......AT.A.G.......C..AAA.G.A...CA.A.........A...........TC..  696

Mos1    GGATCAGAGCGGTGTCAT-TTACTATGAGCTCTTGAAACCCGGCGAAACGGTGAATACGGCACGCTACCAACAACAATTGATCAATTTGAACCGTGCGCT  862
Bpmar1  C.....AC.T.....T....GA.....C..TC......G..T........T.C.GGTAA.............G.....GG........C..TT.  876
Brmar2  ......AC......CT..C.GG....A..TC......A......AA..T.C.GGT.AG.............G...GCAG..A.......A.ATT.  657
Brmar1  ......AC......T..C..GA.....C..TC......A.AT....AA..T.C.GGT.GG.A.........G.......GC.G........A...TT.  796

                                                            H  D  N  A
Mos1    TCAGAGAAAACGACCGGAATATCAAAAAAGACAACACAGGGTCATTTTTCTCCATGACAACGCTCCATCACATA-CGGCAAGAGCGGTTCGCGACACGTT  961
Bpmar1  GAGCCA......T..A..G.....GCT..G......AA..G.......TG......T.A.T...G.C.GAA...A.CA.AC...G..ATT.G.  975
Brmar2  G.GCCA......T..A.C.....GCT..G......AA..G......GTG......T..A......C.GAA....A.CA.AC...G..ATTAG.  756
Brmar1  G.GCCA........A....G..CT..G......AA..G-.....T.TT....A..T.TA......CAAAAA..A.CA.AC...G..ATT.G.  895

                                                    Y  S  P  D  L  A  P  S  D
Mos1    GGAAACACTCAATTGGGAAGTGCTTCCGCATGCGGCTTACTCACCAGACCTGGCCCCATCCGATTACCACCTATTCGCTTCGATGGGACACGCACTCGCT 1061
Bpmar1  A..TT.GTATGG.....A...CC..G..C.......C.T.GG..C..T....T.....A.......T.G.CT..AA.A......T.G..T..A 1075
Brmar2  A..TT.GTA.CGA.....CC..G.AC.......T....C..T....T.....AA.....T..T.G..T.A.............A..T..A  856
Brmar1  A..TCGGTA.GG.......CC..GT.C.......T.T...C..T....T........A..C.T..T.G..T..A.T.......A..G..T..A  995

Mos1    GAGCAGCGCTTCGAT--TCTTACGAAAGTGTGAAAAAATGGCTCGATGAATGGTTCGCCGCAAAAGACGATGAGTTCTACTGGCGTGGAATC-CACAAAT 1158
Bpmar1  .......A......T...GA.......G.G.....C.......A...AC.A...TT..A..A..T..T........ 1172
Brmar2  .......A...A....A...TC...A...A......G...T..A....C.--------..A..A.AC.A....T....A.A..T..T........  943
Brmar1  ..T..A.A.....A...AA.....T...A........GC.....G.....C......T.....A...AC.A...GT.T..AA..T..TC..A... 1095

                                        Polyadenylation signal
Mos1    TGCCCGAGAGATGGGAAAAATGTGTAGCTAGCGACGGCAAATACTTTGAATAAATGATTTTTTCTTTTTCCACAAAATTTAACGTGTTTTTTGATTTAAA 1258
Bpmar1  .A..AA.A...........G.A..T...T..T.TCT.........G.TTAA..G.A.AA.CG..TTCG.GC.A.A.......C.A...T..GC... 1271
Brmar2  ....A.GA....A.A.......T.TT...TC...........TTAA..G.A.AA.CG..TTCG..C.A.A..........A...T..GG... 1042
Brmar1  .A...A...A.C....AATG.....--.........TA...TC.....C....GCTTAA..G.A.AA..G..TTCG..C.A.A........A...T..GC... 1190

                    ITR-3'
Mos1    AAAAACGACATTTCATACTTGTACACCTGATA 1290
Bpmar1  ....TTCCTG..C............C....... 1303
Brmar2  ....-TCCAG..C.C.............G... 1073
Brmar1  ....-TACAG.................G... 1221
```

Fig. 1. Nucleotide sequence alignment of *Bpmar1, Brmar1* and *Brmar2* from *Bruchus pisorum* and *B. rufimanus*, respectively, with the *Mos1* element of *Drosophila mauritiana* (X78906).

Fig. 2. Schematic diagram showing the *Mos1* element of *Drosophila mauritiana* (X78906); *Bpmar1* from *B. pisorum* and the *Brmar1* and *Brmar2* transposons from *B. rufimanus*. Boxes represent canonical motifs of *mariner* transposase. Black triangles represent ITRs. MSS and MSN are start motifs, YFE and YLE are terminal motifs, Δ – Deletion, Ins – Insertion.

quence of the 28-bp inverted terminal repeats (ITRs) of *Bruchus* elements were similar to those of the *Mos1* element. The *Brmar1* 5'-ITR was the most similar to *Mos1*, exhibiting twenty-five exact matches. Lampe et al. [31] deduced the conservation of two motifs at positions 3-8 and 14-18 along the ITR sequence. These motifs are 5'AGGT(C/T/G)(T/G) 3' and 5'(T/A) A(A/G)(A/G)(T/G), exhibiting two highly conserved positions, 5 and 15, respectively. In comparison with the *Mos1* element, those motifs implicated in the protein-ITR interaction were found to be slightly modified in *Bpmar1* (three mismatches) but highly conserved in *Brmar1* (a unique mismatch) and identical in *Brmar2* (no mismatch) (Fig.1).

Putative MLE transposases deduced from consensus sequences showed that they are nonfunctional because they contain multiple stop codons and frameshifts. The obtained hypothetical transposases were aligned with the *Mos1* element to highlight their similarities (Fig. 2), and a schematic diagram of whole *mariner* elements was constructed (Fig. 2 A-D).

In comparison with *Mos1* (Fig. 2A), the *Bpmar1* element showed an insertion of 12bp and therefore, corresponds to a mutated transposase of 346 amino acid residues that exhibit many canonical motifs, several of which are highly conserved, such as the helix-turn-helix (HTH) and the WVPHEL. The signature sequence D,D(34)D, typical of MLEs, was also identified in *Bpmar1*. The first D was included in the motif

"TGDE"; the second D in "DDNA," slightly modified from the canonical "HDNA" motif; and the third D in "HWPDLAPSD," which is modified from the canonical motif YSPDLAPSD (Fig. 2B).

However, *mariner* elements of *B. rufimanus* were found to carry many deletions along their transposases (Supplementary Fig. S1). The *Brmar1* element displayed a transposase of 317 amino acids with only two conserved aspartic acid residues of the D,D(34)D motif. The first retained D was identified in the IGDE motif, which is altered from the canonical "TGDE," while the second D was missing due to mutations in the canonical "HDNA" motif. The third D was in the slightly modified YLPDLAPSD motif (Fig. 2C).

In silico translation of *Brmar2* generated a 271 amino acid transposase lacking the HTH and WVPHEL motifs due to an internal deletion spanning these regions. Only two conserved aspartic acid residues of the "D,D(34)D" motif were identified. The first retained D was identified in the TADE motif and the second D in the "DDNA" motif. The third D was missing due a mutation in the YSPDLAPSD motif, which is replaced by YSPDLAPSN (Fig. 2D).

Phylogenetic analyses

To investigate the phylogenetic relationships among *Bruchus* MLE consensus sequences and known full-length *mariner* elements, we used some of the published sequences representative of the five major *MLEs* subfamilies from GeneBank: *mauritiana, cecropia, mellifera, capitata* and *irritans,* as described by Robertson and MacLeod [8]. As expected, the phylogram constructed with the maximum likelihood method based on nucleotide sequences revealed five clades grouping the *mariner* transposase. The *Bruchus* MLE sequences were reliably assigned to the *mauritiana* subfamily with a bootstrap support of 99% (Fig. 3).

DISCUSSION

The present study is the first report of MLEs belonging to the *mauritiana* subfamily in monophagous coleopteran grain pests, *B. pisorum* and *B. rufimanus*. Using PCR followed by sequencing, we have identified two MLE consensuses, *Brmar1* and *Brmar2,* in the *B.*
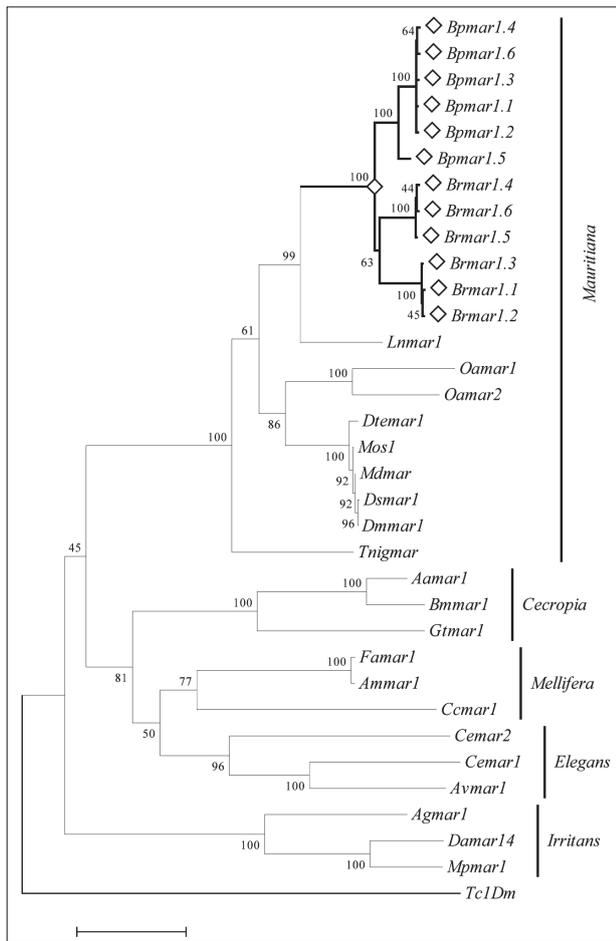
Fig. 3. Molecular phylogenetic analysis by the maximum likelihood method between *Bruchus* MLEs and other *mariner* elements based on the nucleotide sequences using Tc1 as the outgroup. The numbers on the nodes of the phylogram indicate bootstrap support (1000 replications). The accession numbers of the used elements are: X78906 (*Mos1: Drosophila mauritiana*), M14653.1 (*Dmmar1: Drosophila mauritiana)*, AF035567.1 (*Dtemar1: Drosophila teissieri)*, AF037060.1 (*Dsmar1: Drosophila simulans*), LBMM01014153.1 (*Lnmar1: Lasisus niger*), HE577149.1 (*Tnigmar: Tapinoma nigerrimum)* AF373028.1 (*Mdmar: Musca domestica)*, KC767725.1, KC767721.1 (*Oamar1, Oamar2: Oryctes agamemnon)* from the *mauritiana* subfamily, X71979.1 (*Gtmar1: Girardia tigrina*), AB006464.1 (*Aamar1: Attacus atlas*) D88671.1 (*Bmmar 1: Bombyx mori*) from the *cecropia* subfamily, AY155492.1 (*Famar1: Forficula auricularia*) and U40493.1 (*Ccmar1: Ceratitis capitata*) AY155490.1 (*Ammar1: Apis mellifera)* from the *mellifera* subfamily, ZC132.1 (*Cemar1: Caenorhabditis elegans*), DQ138248.1 (*Avmar1: Adineta vaga*), Y39A3A.1 (*Cemar2: Caenorhabditis elegans*) from the *elegans* subfamily, U11646.1 (*Damar14: Drosophila ananassae*), U11649.1 (*Mpmar1:Mantispa pulchella* clone) and GQ398105.1 (*Agmar1: Agrilus planipennis*) from the *irritans* subfamily. Tc1-like element sequences from *Drosophila melanogaster* (*Tc1Dm*: S60466) were used as the out-group to root the tree.

rufimanus genome and a single consensus element, *Bpmar1*, in *B. pisorum*.

Analysis of these three MLEs revealed, like most other *mariner* elements, that they are defective copies containing deletions, insertions, multiple stop codons and frameshifts throughout the sequences [11]. This indicates that they have accumulated mutations during the process of vertical transmission [32].

The *Bpmar1 mariner* element identified in the *B. pisorum* genome exhibited low levels of sequence divergence between the clones, suggesting a more recent origin of the corresponding *mariner* element copies in the *B. pisorum* genome. In contrast, the *Brmar1* and *Brmar2* elements showed high nucleotide sequence diversity, suggesting that either they have inhabited the *B. rufimanus* genome for a long time, or two distinct MLE lineages have invaded the *B. rufimanus* genome and evolved differently.

Moreover, both *mariner* elements showed internal deletions in the 5' region that contains the coding region of the MLEs. Such deletions have been observed in several elements belonging to *IS630-Tc1-mariner* [26,33,34]. Hua-Van. et al. [33] suggested that the low polymorphism in the region containing the two functional domains (DDD and NLS) could be due to the existence of selective pressures on these crucial domains, which might be explained by the recent activity of the corresponding elements.

Different mechanisms could generate deletions in MLEs, such as slippage during replication or ectopic recombination [35]. The origin of deletions might also be related to an active regulatory mechanism that is involved in the inactivation of full-length MLE copies [36] or to the transposition mechanisms of Class-II elements. In fact, MLEs transpose by a cut-and-paste model, which requires the repair of the gap left by the excision [37]. When the host DNA repair machinery is not efficient, internal deletions could appear [35]. Indeed, many other Class-II elements have shown internal deletions as a consequence of element transposition, such as the piggyBac-like elements of *Aphis gossypii* [38] and the *P* and *hobo* elements of *D. melanogaster* [39]. Plasmid-based mobility assays using *Mos1* have been carried out in several insects, such as *Aedes aegypti*, *Drosophila melanogaster*, *Lucilia cuprina* and *Bactrocera tryoni* [40,41]. Green et al. [42] reported lower frequen-

cy of transposition events in *B. tryoni* compared to *D. melanogaster*, which was explained by the presence of endogenous MLE copies in the *B. tryoni* genome that interfered, even in their defective form, with the activity of the exogenous *Mos1* transposase along with the associated regulatory system. Indeed, the presence of endogenously similar MLEs might affect the efficiency of the transgenesis causing instability of the transposase-mediated insertions that resulted from cross-mobilization events. Moreover, the authors concluded that both recombination between different MLEs and the potential for cross-mobilization could be influenced by the level of similarity between regions of the encoded transposases and the nucleotide sequence of the elements [42]. In addition, Bigot et al. [9] demonstrated the presence of motifs within the ITRs of *MLEs* that interact with the transposase. In this study, comparison of *Bruchus MLE* ITRs with those of *Mos1* showed the conservation of these motifs implicated in transposase fixation. Therefore, the presence of these endogenous MLEs similar to *Mos1* elements may interfere with the efficiency of the transgenesis vector system.

To summarize, when considering germline transformation vector technology to control the impact of these serious storage insect pests on legume crops, factors regulating or repressing transposable elements within their corresponding host genome need to be investigated. In the present study, we identified *mariner* elements of the *mauritiana* subfamily in *B. pisorum* and *B. rufimanus*. The presence of these endogenous elements might interfere with the efficiency of the transformation system based on *Mos1*. Therefore, the discovery of additional transposon-based vectors and engineered elements may advance the conception of highly effective systems to conduct transgenesis within these two storage insect pests.

Authors' contribution: Salma Djebbi performed the experiments; Salma Djebbi, Wiem Ben Amara and Maha Mezghani-Khemakhem analyzed the data; Hanem Makni and Mohamed Makni helped in designing the study; Salma Djebbi and Maha Mezghani-Khemakhem wrote the paper.

Conflict of interest disclosure: The authors declare that there is no conflict of interests regarding the publication of this paper.

REFERENCES

1. Doolittle WF, Sapienza C. Selfish genes, the phenotype paradigm and genome evolution. Nature.1980;(17):284:601-3.
2. Finnegan DJ. Transposable elements. Curr Opin Genet Dev. 1992;2(6):861-7.
3. Robertson HM. The *mariner* transposable element is widespread in insects. Nature. 1993;362(6417):241-5.
4. Haymer DS, Marsh JL. Germ line and somatic instability of a white mutation in *Drosophila mauritiana* due to a transposable genetic element. Dev Genet. 1986;6(4):281-91.
5. Jacobson JW, Medhora MM, Hartl DL. Molecular structure of a somatically unstable transposable element in *Drosophila*. Proc Natl Acad Sci USA. 1986;83(22):8684-8.
6. Hartl DL, Lozovskaya ER, Nurminsky DI, Lohe AR. What restricts the activity of *mariner*-like transposable elements. Trends Genet. 1997;13(5):197-201.
7. Feschotte C, Jiang N, Wessler SR. Plant transposable elements: where genetics meets genomics. Nat Rev Genet. 2002;3(5):329-41.
8. Robertson HM, MacLeod EG. Five major subfamilies of mariner transposable elements in insects, including the Mediterranean fruit fly, and related arthropods. Insect Mol Biol. 1993;2(3):125-39.
9. Bigot Y, Brillet B, Auge-Gouillou C. Conservation of Palindromic and Mirror Motifs within Inverted Terminal Repeats of *mariner*-like Elements. J Mol Biol. 2005;351(1):108-16.
10. Rouault JD, Casse N, Chenais B, Hua-Van A, Filee J, Capy P. Automatic classification within families of transposable elements: application to the *mariner* Family. Gene. 2009;448(2):227-32.
11. Maruyama K, Schoor KD, Hartl DL. Identification of nucleotide substitutions necessary for trans-activation of mariner transposable elements in *Drosophila*: analysis of naturally occurring elements. Genetics. 1991;128(4):777-84.
12. Delaurière L, Chénais B, Hardivillier Y, Gauvry L, Casse N. *Mariner* transposons as genetic tools in vertebrate cells. Genetica. 2009;137(1):9-17.
13. Medhora M, Maruyama K, Hartl DL. Molecular and functional analysis of the mariner mutator element Mos1 in *Drosophila*.Genetics. 1991;128(2):311-8.
14. Barry EG, Witherspoon DJ, Lampe DJ. A bacterial genetic screen identifies functional coding sequences of the insect mariner transposable element *Famar1* amplified from the genome of the earwig, *Forficula auricularia*. Genetics. 2004;166(2):823-33.
15. Munoz-Lopez M, Siddique A, Bischerour J, Lorite P, Chalmers R, Palomeque T. Transposition of Mboumar-9: identification of a new naturally active *mariner*-family transposon. J Mol Biol. 2008;382(3):567-72.
16. Coates CJ, Jasinskiene N, Miyashiro L, James AA. Mariner transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. Proc Natl Acad Sci U S A. 1998;95(7):3748-51.
17. Atkinson PW, Pinkerton AC, O'Brochta DA. Genetic transformation systems in insects. Annu Rev Entomol. 2001;46:317-46.
18. Mathur G, Sanchez-Vargas I, Alvarez D, Olson KE, Marinotti O, James AA. Transgene-mediated suppression of dengue

viruses in the salivary glands of the yellow fever mosquito, *Aedes aegypti*. Insect Mol Biol. 2010;19(6):753-63.

19. Ashburner M, Hoy MA, Peloquin JJ. Prospects for the genetic transformation of arthropods. Insect Mol Biol. 1998;7(3):201-13.

20. Smith AM. Pea Weevil (*Bruchus Pisorum L.*) and Crop Loss - Implications for Management. In: Fujii K, Gatehouse AMR, Johnson CD, Mitchel R, Yoshida T, editors. Bruchids and Legumes: Economics, Ecology and Coevolution: Proceedings of the Second International Symposium on Bruchids and Legumes (ISBL-2);1989Sep 6-9; Okayama,Japan. Dordrecht: Springer Netherlands; 1990. p. 105-14.

21. Clement SL, Wightman JA, Hardie DC, Bailey P, Baker G, McDonald G. Opportunities for integrated management of insect pests of grain legumes. In: Knight R, editor. Linking Research and Marketing Opportunities for Pulses in the 21st Century: Proceedings of the Third International Food Legumes Research Conference. Dordrecht: Springer Netherlands; 2000. p. 467-80. (Current plant science and biotechnology in agriculture; vol. 34).

22. Oliveira SG, Cabral-de-Mello DC, Moura RC, Martins C. Chromosomal organization and evolutionary history of *Mariner* transposable elements in *Scarabaeinae* coleopterans. Mol Cytogenet. 2013;6(1):54.

23. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaves tissue. Phytochem Bull. 1987;19:11-5.

24. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol. 1994;3(5):294-9.

25. Mezghani-Khemakhem M, Ben Lazhar W, Bouktila D, Ben Slimen H, Makni H, Makni M. A rapid diagnostic technique of *Bactrocera cucurbitae* and *Bactrocera zonata* (Diptera: Tephritidae) for quarantine application. Pest Manag Sci. 2013;69(6):744-6.

26. Kharrat I, Mezghani M, Casse N, Denis F, Caruso A, Makni H, et al. Characterization of *mariner*-like transposons of the *mauritiana* Subfamily in seven tree aphid species. Genetica. 2015;143(1):63-72.

27. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser. 1999;41:95-98.

28. Narasimhan G, Bu C, Gao Y, Wang X, Xu N, Mathee K. Mining protein sequences for motifs. J Comp Biol. 2002;9(5):707-20.

29. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol.2016;33:1870-4.

30. Robertson HM, Asplund ML. *Bmmar1*: a basal lineage of the *mariner* family of transposable elements in the silkworm moth, *Bombyx mori*. Insect Biochem Mol Biol. 1996;26(8-9):945-54.

31. Lampe DJ, Walden KK, Robertson HM. 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. 2001;18(6):954-61.

32. Capy P, David JR, Hartl DL. Evolution of the transposable element mariner in the *Drosophila melanogaster* species group. Genetica. 1992;86(1-3):37-46.

33. Hua-Van A, Hericourt F, Capy P, Daboussi MJ, Langin T. Three highly divergent subfamilies of the *impala* transposable element coexist in the genome of the fungus *Fusarium oxysporum*. Mol Gen Genet. 1998;259(4):354-62.

34. Torti C, Gomulski LM, Malacrida AR, Capy P, Gasperi G. Genetic and molecular investigations on the endogenous mobile elements of non-drosophilidfruitflies. Genetica. 1997;100(1-3):119-29.

35. Brunet F, Giraud T, Godin F, Capy P. Do deletions of Mos1-like elements occur randomly in the Drosophilidae family? J Mol Evol. 2002;54(2):227-34.

36. Capy P, Bazin C, Higuet D, Langin T. Dynamic and Evolution of Transposable Elements. Austin, Texas, USA: R.G. Landes Company; 1997.197 p.

37. Engels WR, Johnson-Schlitz DM, Eggleston WB, Sved J. High-frequency *P* element loss in *Drosophila* is homolog dependent. Cell. 1990;62(3):515-25.

38. Luo GH, Wu M, Wang XF, Zhang W, Han ZJ. A new active *piggyBac-like* element in *Aphis gossypii*.Insect Sci. 2011;18(6):652-62.

39. Daniels SB, Chovnick A, Boussy IA. Distribution of *hobo* transposable elements in the genus *Drosophila*. Mol Biol Evol. 1990;7(6):589-606.

40. Coates CJ, Turney CL, Frommer M, O'Brochta DA, Warren WD, Atkinson PW. The transposable element mariner can excise in non-drosophilid insects. Mol Gen Genet. 1995;249(2):246-52.

41. Coates CJ, Turney CL, Frommer M, O'Brochta DA, Atkinson PW. Interplasmid transposition of the mariner transposable element in non-drosophilid insects.Molecular & General Genetics. 1997;253(6):728-33.

42. Green CL, Frommer M. The genome of the Queensland fruit fly *Bactrocera tryoni* contains multiple representatives of the *mariner* family of transposable elements. Insect Mol Biol. 2001;10(4):371-86.

Supplementary Data

Supplementary Fig.S1. Alignment of the *in silico* translation of the consensus *Bpmar1* transposase sequences of *Bruchus pisorum* and *Brmar1* and *Brmar2* of *B. rufimanus* with the Mos1 element of *Drosophila mauritiana* (X78906). (-) indicates gaps introduced to maintain the alignment; (*) indicates stop codons. Black and gray boxes are for identical and similar amino acids.
Fig. S1. can be assessed from: http://serbiosoc.org.rs/sup/Suppl.Fig.S1.docx