Genome-Wide Analysis of the *Fusarium oxysporum mimp* Family of MITEs and Mobilization of Both Native and De Novo Created *mimps*

Mara Bergemann · Olivier Lespinet · Sarrah Ben M’Barek · Marie-Josée Daboussi · Marie Dufresne

Received: 23 May 2008 / Accepted: 28 August 2008 / Published online: 4 November 2008 © Springer Science+Business Media, LLC 2008

Abstract We have performed a genome-wide analysis of the *mimp* family of miniature inverted-repeat transposable elements, taking advantage of the recent release of the *F. oxysporum* genome sequence. Using different approaches, we detected 103 *mimp* elements, corresponding to 75 nonredundant copies, half of which are located on a single small chromosome. Phylogenetic analysis identified at least six subfamilies, all remarkably homogeneous in size and sequence. Based on high sequence identity in the terminal inverted repeats (TIRs), *mimp* elements were connected to different *impala* members. To gain insights into the mechanisms at the origin and amplification of *mimps*, we studied the potential of *impala* to cross-mobilize different *mimps*, native but also created de novo by inserting a short DNA segment between two TIRs. Our results show that TIR sequences are the main requirement for mobilization but that additional parameters in the internal region are likely to influence transposition efficiency. Finally, we show that integration site preference of native versus newly transposed *mimps* greatly varies in the host genomes used in this study.

Keywords Miniature inverted-repeat transposable element (MITE) · Mobilization · De novo origin · *Fusarium*

Introduction

Class II transposable elements (TEs), also called DNA transposons, have terminal inverted repeats (TIRs) at both ends and transpose using a “cut-and-paste” mechanism. This process involves cleavage at the ends of the transposon and strand transfer that allows excision of the element prior to its reinsertion at another locus in the genome. This mechanism is performed by a transposase encoded by autonomous elements which, after specific binding of the enzyme at the ends of the element, catalyzes the steps mentioned above. Class II TEs are either autonomous transposons encoding their own transposase or nonautonomous elements corresponding to mutation (mostly deletion) derivatives of the previous ones. Although not encoding for a functional transposase, most nonautonomous elements can still be mobilized by a transposase provided in trans, because of strong sequence similarity, in particular, in the TIRs region.

Miniature inverted-repeat transposable elements (MITEs) are short (<500-bp) elements, structurally similar to defective class II elements. In particular, they are characterized by the presence of TIRs. They also share other features that distinguish them from other nonautonomous DNA transposons such as a high copy number and size homogeneity (Feschotte et al. 2002). With the progress of large-scale sequencing, it has become apparent that
MITEs are important components of eukaryotic genomes. First identified in plants (Wessler et al. 1995), they have been subsequently found, generally in large copy numbers, in a wide range of animal and fungal genomes (for a review, see Feschotte et al. 2002). Similarity searches based on TIR sequences have allowed the identification of autonomous class II elements, potentially responsible for the mobilization of MITEs (Feschotte and Mouches 2000; Feschotte et al. 2003; Jiang et al. 2003; Kikuchi et al. 2003; Macas et al. 2005; Saito et al. 2005; Quesneville et al. 2006). However, only a few recent studies have demonstrated, either in vitro or in vivo (Yang et al. 2007; Dufresne et al. 2007), that class II transposases do have the potential to mobilize MITEs.

Recent genome-wide analyses have greatly improved our knowledge on MITE families, however, many questions remain about their origin and amplification. Some MITEs may have originated by deletion of autonomous class II elements (Feschotte and Mouches 2000; Jiang et al. 2003; Ramussen et al. 2004), as shown by extended regions of similarity to coexisting TEs. Other MITE families may rather originate de novo by fortuitous juxtaposition of sequences resembling the TIRs of autonomous elements (Feschotte et al. 2002; Quesneville et al. 2006).

In Fusarium oxysporum, numerous TEs belonging to all structural classes have been identified (Daboussi and Capy 2003). The impala family of DNA transposons belonging to the Tc1/mariner superfamly has been extensively studied in our laboratory (Hua-Van et al. 1998, 2001a, b). This family occurs at a low copy number in F. oxysporum and comprises at least five different subfamilies, which differ by a high level of nucleotide diversity (Hua-Van et al. 1998, 2001a). Inactive elements, representing the majority of impala copies, were identified in all subfamilies, in contrast with autonomous elements, encountered only in subfamilies E and D (Hua-Van et al. 1998, 2001a, b). Two short elements (~200 bp) displaying TIRs very similar to those of impalaE have been recognized in the genome of F. oxysporum (Hua-Van et al. 2000). These elements, called mimp1 and mimp2, exhibit features of MITEs. Recently, we demonstrated that mimp1 is mobilized through the action of the impalaE transposase (Dufresne et al. 2007).

In this paper, we extract mimp-like elements from the recently available genome sequence of F. oxysporum and classify them into different subfamilies. We analyze their chromosomal distribution and insertion sites relative to predicted genes. We also demonstrate the mobilization of a native mimp element belonging to another subfamily than mimp1. More interestingly, we show that a mimp element created in vitro can also be mobilized by the impalaE transposase.

Materials and Methods

Fungal Strains

The Fusarium oxysporum strains FOM24 (F. oxysporum f.sp. melonis) and FOL15 (F. oxysporum f.sp. lycopersici) were used in polymerase chain reaction (PCR) experiments either to identify novel mimp elements (both strains) or to generate the internal sequence of the de novo created mimpNia element (FOM24 strain). The F. graminearum nitrate reductase-deficient mutant Fg820nia5 was used as a recipient strain in all transformation experiments. Fg820nia5 and derivatives were grown on PDA plates. To obtain spores, flushes or 2-ml Eppendorf tubes containing 20 ml or 1 ml of liquid mung bean medium (Bai and Shaner 1996), respectively, were inoculated with a plug from an 8-day-old PDA plate and incubated with moderate shaking at 26°C for 2–5 days. For long-term storage, a plug of the culture was put on 500 µl PDA poured in a 2-ml Eppendorf tube and, after growth for 2–3 days at 26°C, was stored at 4°C.

Plasmids and Constructs

pHEO62 contains the open reading frame (ORF) encoding the impalaE transposase (Hua-Van et al. 1998), cloned between the gpdA promoter and the trpC terminator of Aspergillus nidulans. It also carries a cassette conferring hygromycin resistance from the pBC1004 plasmid (Carroll et al. 1994), which allows hygromycin selection.

pNm2Hr carries a mimp2 element into the first intron of the niaD gene and was constructed following the same steps as for the pNm1H18 construct (Dufresne et al. 2007). The mimp2 copy corresponds to the reamplification of the mimp2 element (AF076625) previously described (Hua-Van et al. 2000) using the HindImp primer (5'-GGCC AAGCTTACAGTGGGGTGCAATAAGTTG-3'). This primer anneals to the TIRs of mimp2 and contains a HindIII site (underlined) at the 5' end. PCR conditions were 1 min of denaturation at 95°C, followed by 30 cycles of 1 min 94°C, 1 min at 59°C, and 30 s at 72°C, then an elongation step of 10 min at 72°C. The PCR product was then digested with HindIII and inserted into the HindIII site of the first intron of niaD gene in plasmid pAN301 (Malandier et al. 1989), deleted of a dispensable Ndel fragment, resulting in the niaD::mimp2 construct, also called pNm2Hr.

The artificial mimp element was generated using PCR in a three-step procedure. The first step used specific primers to amplify the desired genomic region, which will correspond to the central region of the artificial mimp element. The two others used chimeric primers, allowing the obtainment of a full-length artificial mimp with TIRs in two
successive PCR rounds. First, part of the first exon of the *F. oxysporum* nia gene, showing <50% identity to the orthologous sequence in *F. graminearum*, was amplified using primers niafox871+ (5'-ATACCCACACGTCTTCTCTCCT-3') and niafox1027– (5'-AGGACAGCTTTGCGAG-3') from genomic DNA of the FOM24 strain. PCR conditions were 5 min at 94°C, followed by 35 cycles of 1 min at 94°C, 1 min at 52°C, and 30 s at 72°C, and a final extension step of 10 min at 72°C. The eluted 157-bp PCR product was then used as template for amplification with the chimeric primers mimp_nia5 (5'-AATAAGTTTGAGGCCATACCCACAG-3'; HindIII restriction site underlined) and mimp_nia3' (5'-CAATAAGTTTGAGACAGACACTTTG-3'; the region complementary to the primary PCR product is underlined). PCRam conditions were 5 min at 94°C, followed by 5 cycles of 1 min at 94°C, 1 min at 34°C, and 1 min at 72°C, then 30 cycles of 1 min at 94°C, 1 min at 62°C, and 1 min at 72°C, and a final extension step of 10 min at 72°C. The eluted PCR product was finally used as template in a final PCR with the HindImp primer (5'-GCCCTAAGCTTACA GTGGGCTGCAAATGTTTGTG-3'; HindIII restriction site underlined) under the following conditions: 5 min at 94°C, followed by 35 cycles of 1 min at 94°C, 1 min at 44°C, and 1 min at 72°C, and a final extension step of 10 min at 72°C. After all PCR steps, the eluted amplification product was cloned into the pGEM-T vector (Promega; Charbonnières-les-Bains, France). After sequencing analysis of the whole product, the plasmid was digested with the HindIII restriction enzyme (see HindImp primer) and inserted into the HindIII site of the first intron of niaD gene in plasmid pAN301 as described above for the niaD::mimp2. The resulting construct was named pNmNia.

**Transformation Experiments and Revertant Selection**

Protoplast transformation was performed as previously described (Dufresne et al. 2007). Plugs of each cotransformant were picked on plates of nitrate minimal agar medium. Revertants were easily detected as patches of aerial mycelium with a wild-type phenotype on a back- ground of sparse mycelium corresponding to a niaD mutant (see Fig. 4 in Dufresne et al. 2007).

**DNA Preparation and Southern Blot Analysis**

DNA extractions from *F. graminearum* strains were conducted as previously described (Dufresne et al. 2007). Ten micrograms of genomic DNA was digested with the appropriate restriction enzyme, separated electrophoretically on 0.7% agarose gels, and transferred on nylon membranes, using a vacuum blotter. DNA templates were 32P-labeled using the rediprime II kit (Amersham Biosciences). Hybridizations were conducted under standard conditions (Sambrook et al. 1989).

**PCR and Primer Sequences**

New mimps elements were amplified by PCR using the HindImp primer (see Plasmids and Constructs, under Materials and Methods section), under the same PCR conditions as described above, and genomic DNA of either the FOM24 or the FOL15 strain as template (50 ng per PCR reaction). Hybridization probes were obtained by PCR. The HindImp primer was used (PCR conditions: 1 min of denaturation at 95°C, followed by 30 cycles of 1 min at 94°C, 1 min at 59°C, and 30 s at 72°C, then an elongation step of 10 min at 72°C) to generate PCR products from different mimp elements. The 419-bp niaD probe was generated with primers niaDCG1 (5'-CAGCTACCTAGGTACAGACGACAC-3') and niaDCG2 (5'-CTGACGACGACGACGACGACAG-3'), using the pAN301 plasmid as a template. PCR conditions were as follows: 2 min at 94°C, then 30 cycles of 30 s at 94°C, 30 s at 59°C, and 2 min at 72°C.

Excision events were controlled by PCR on 50 ng of genomic DNA of putative revertants and of the corresponding transformants, using primers niaD144 (5'-GGT TATGGTGAGGACGCGG-3') and niaD54 (5'-AGT TGGGCTGCAAATGTTTGTG-3') under the following conditions: 4 min at 94°C, then 30 cycles of 1 min at 94°C, 1 min at 59°C, and 2 min at 72°C. The sizes of the expected PCR products are 717 bp for a transformant and 485 bp for a revertant.

Amplification of Sequences Flanking mimp Elements

*mimp2* flanking sequences were recovered using a modified TAIL-PCR approach as described for *mimp1* by Dufresne et al. (2007). The arbitrary degenerate (AD) primer AD2, 5'-AG(A/T)GNAG(A/T)ANCA(A/T)AGA-3', and the specific primers, m2Tail1 (5'-GAATCCAGGCTGAAGTACC TG-3'), and m2Tail2 (5'-CCCGATTGGGGTATTGCTAG-3'), were used with reduced-stringency and high-stringency annealing temperatures of 44°C and 63°C, respectively. To verify the duplication of a target site following transposition of *mimpNia*, an inverse PCR strategy using the *Hae*III or *Xba*I restriction enzymes was carried out, using primers mniaDiv5 (5'-CTTTCGCTTTCCCTCGGAGGAA G-3') and mniaDiv3 (5'-TCTCGCAGAGTGCTGGGTCC-3'). PCR conditions were 5 min at 94°C, followed by 35 cycles of 1 min at 94°C, 1 min at 60°C, and 4 min at 72°C, and a final extension step of 10 min at 72°C.
Cloning of PCR Products and DNA Sequencing

PCR products were directly cloned into the pGEM-T vector (Promega) using 3 μl of either rough or purified PCR products, following the manufacturer’s instructions. Sequencing of PCR products, either directly or cloned into the pGEM-T vector, was performed by Genome Express (Meylan, France) using an ABI Big Dye Terminator kit (Perkin Elmer) and the appropriate primer(s).

Computational Analysis

The design of optimal primers for PCR experiments was realized using the Primer3 program (http://primer3.sourceforge.net/). The quality of sequencing was visualized with the program Chromas Lite (http://www.technelysium.com.au/). The mfold algorithm (http://www.bioinfo.rpi.edu/applications/mfold/rna/form1.cgi) was used to visualize potential secondary structures formed by mimp-like elements as well as to determine their stability by calculating the ΔG0 value. Blast analyses on the F. oxysporum and F. graminearum genomes were performed on the Fusarium Group web site, available at the Broad Institute (http://www.broad.mit.edu/annotation/genome/fusarium_group/MultiHome.html). For the determination of a potential consensus sequence of mimp insertions, 10 bp upstream and downstream of the target site was analyzed with the SeqLogo program (Schneider and Stephens 1990).

MITE Search by Complete Genome Screening

To perform an exhaustive MITE search, we used an ad hoc Perl script based on the following algorithm (available upon request). The nucleic acid sequence of each target TIR of a sequence length l were compared with the sequence of all sliding windows of the same length extracted from the F. oxysporum genome. A nucleotide polymorphism of 25%, slightly higher than the maximum value observed for TIRs examined manually (5 mismatches over 27 bp of TIRs), was tolerated. Once a candidate 5’-TIR was identified, a potential 3’-TIR was searched in a 500-bp range, in the range of maximum sizes usually reported for MITEs, with the same cutoff value of 75% identity. All MITE candidates were checked manually.

Phylogenetic Analysis

Multiple sequence alignments of the mimp nucleotidic sequences were performed using the MAFFT program using the default parameters (http://align.bmr.kyushu-u.ac.jp/mafft/online/server/). Alignments were then visualized using the sequence editor available in the MEGA4.0 package (Tamura et al. 2007). This package was also used to generate distance matrices and phylogenetic trees. The evolutionary history was inferred using the neighbor-joining method (Saitou and Nei 1987).

Results

Extracting mimp-Like Elements from the F. oxysporum Genome

In a previous work two different mimp elements had been identified by sequencing of genomic regions of the F. oxysporum FOM24 strain (mimp1 and mimp2 [Hua-Van et al. 2000]; accession numbers AP076624 and AP076625, respectively). These elements share the same structure: 27-bp impalaE-like TIRs, a central region with no coding capacity and exhibiting no homology with known sequences or between each other, and an overall size of ~200–220 bp. To identify novel mimp-like elements, we conducted PCR experiments using a primer specific for part of the impalaE TIR sequence (HindImp; see Material and Methods section). After cloning of the rough PCR amplicon, 20 randomly chosen clones were analyzed by sequencing: 14 corresponded to mimp1-like elements, 3 to mimp2-like copies, and the 3 remaining sequences corresponded to two novel mimp elements called mimp3 (two clones) and mimp4 (one clone). mimp3 exhibits impalaE-TIRs (Fig. 1) as mimp1 and mimp2 and is 216 bp long. The mimp4 element, despite presenting a similar overall structure and size (220 bp), was shown, after BLASTN searches on the F. oxysporum genome sequence, to be characterized by the presence of impalaF-like TIRs (Fig. 1).

To determine the copy number of mimp elements in the genome of F. oxysporum, BLASTN searches were performed on the genome sequence recently available at the Broad Institute (http://www.broad.mit.edu/annotation/genome/fusarium_group/MultiHome.html) using the four mimp elements described above as templates. Following manual analysis of the BLASTN hits, 95 elements were identified. In the F. oxysporum genome, carrying several segmental duplications (L. J. Ma, Broad Institute; our laboratory, unpublished results), we removed redundant mimp copies. Among the 71 nonredundant elements, 58 could be assigned to one of the four known mimp subfamilies and 13 remained unclassified at this stage.

We wondered whether more divergent mimp elements could be missed by such a similarity-based search. We thus extended our analysis by developing a computer program based only on the search of similar TIR sequences (impalaE- or impalaF-TIRs) spaced by <500 bp (see Materials and Methods section). If other mimp subfamilies exist, this computer program should allow their detection without any
a priori knowledge of the internal nucleotide sequence. We searched the complete \textit{F. oxysporum} genome sequence and found eight additional \textit{mimp} copies corresponding to four nonredundant sequences. Overall, 75 independent \textit{mimp} copies were found in the whole \textit{F. oxysporum} genome sequence, including 17 new \textit{mimp} elements which could not be easily classified into any of the four known \textit{mimp} subfamilies because they harbored completely different internal sequences.

To investigate further whether these new elements could define novel subfamilies, a subset of sequences including three full-length copies of each of the four \textit{mimp} subfamilies and the 17 novel full-length \textit{mimp} copies, called \textit{mn}1 to \textit{mn}17 (for \textit{mimp}new), was aligned using the MAFFT algorithm and the alignments used in phylogenetic analyses to obtain a neighbor-joining tree (Fig. 2). Six different monophyletic groups supported by high bootstrap values (>90%) were defined. These six distinct \textit{mimp} subfamilies were called \textit{mimp}1 to \textit{mimp}6. Apart from isolated copies (\textit{mn}5, \textit{mn}15, and \textit{mn}16; see Supplementary Table 1) corresponding to highly divergent sequences, a seventh group of sequences appeared on the phylogenetic tree (copies \textit{mn}7, \textit{mn}14, \textit{mn}17, \textit{mn}13, and \textit{mn}4; see Supplementary Table 1). However, this group was not supported by >50% bootstrap values and was not considered significant. Considering the TIR sequences, three groups of \textit{mimp} elements can be distinguished in the \textit{F. oxysporum} genome. The first, comprised of the \textit{mimp}1, \textit{mimp}2, and \textit{mimp}3 subfamilies and most of the highly divergent \textit{mimp} elements, is characterized by the presence of \textit{impalaE} TIRs. The second exhibits \textit{impalaD} TIR sequences and comprises the \textit{mimp}6 subfamily as well as one divergent element, \textit{mn}16. Finally, \textit{mimp}4 and \textit{mimp}5 subfamilies constitute a third group, characterized by \textit{impalaF}-like TIRs.

Among the 75 identified \textit{mimp} elements, we found both complete copies, referred to as full-length elements (Table 1), with TIRs and a target site duplication (TSD) sequence and truncated versions with only one discernible TIR. In all \textit{mimp} subfamilies except \textit{mimp}5, consisting of only two complete copies, full-length elements are largely predominant (Table 1). Only the 64 full-length copies were considered for further analysis.
Nucleotide polymorphism was examined and showed that all mimp subfamilies are characterized by a high level of nucleotide identity (92–98%), even if the mimp4 subfamily exhibits a higher level of sequence variation (25%). Another striking characteristic of mimp elements is their remarkable size homogeneity (between 204 and 226 bp). Apart from mimp elements, our computer-based search identified two already known deleted versions of impalaD (367 and 471 bp, respectively [Hua-Van et al. 2001a; unpublished results]).

Mobilization of a Native and a De Novo Created mimp by the impalaE Transposase

We have previously demonstrated that the native mimp1 element could be mobilized by the impalaE transposase (Dufresne et al. 2007). Since sequence similarity between mimp1 and impalaE is restricted to the TIRs, mobilization of other elements sharing similar TIRs would be expected. However, cross-mobilization is not fully predictable since we do not know the role of sequences in the central region. To determine if similarity in the TIRs is the only requirement or if the central region contributes to the transposition process, we evaluated the cross-mobilization by the impalaE transposase of both a mimp2 element and a de novo created mimp, mimpNia, according to the same procedure as the one already developed for mimp1, in the heterologous F. graminearum genetic background (Dufresne et al. 2007).

Five cotransformants, carrying both the mimp2 element inserted into the niaD gene and the construct allowing expression of the source of impala transposase, were obtained and replicated each on three minimal medium plates (see Materials and Methods section) to test for mimp2 excision. Based on the phenotypic assay, colonies which grew aerial were considered as revertants resulting from excision of mimp2. All five cotransformants gave a high rate of excision (20 to >100 revertants per plate; Table 2). A set of 12 potential revertants was analyzed by Southern blot to confirm mimp2 excision and estimate reinsertion rates. The results indicated that all revertants correspond to an excision event (Fig. 3a, top). Sequencing of the empty sites revealed, in all cases, a footprint of five additional nucleotides, two corresponding to the duplication of the TA target site and three to one of the ends of the element, as already observed for impala (Hua-Van et al. 2001b) and mimp1 (Dufresne et al. 2007) (data not shown). Using the mimp2 element as a probe (Fig. 3a, bottom, and Table 2), reinsertion was observed in 50% (6/12) of the strains analyzed, a frequency which appears to be lower than that observed for mimp1 in the same genetic background (95% [Dufresne et al. 2007]).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Sequence, size variability, and location relative to the closest predicted genes of the six mimp families</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full-length</strong></td>
<td><strong>Truncated</strong></td>
</tr>
<tr>
<td><strong>ORF</strong></td>
<td><strong>&lt;500 bp</strong></td>
</tr>
<tr>
<td>mimp1</td>
<td>18</td>
</tr>
<tr>
<td>mimp2</td>
<td>8</td>
</tr>
<tr>
<td>mimp3</td>
<td>7</td>
</tr>
<tr>
<td>mimp4</td>
<td>18</td>
</tr>
<tr>
<td>mimp5</td>
<td>2</td>
</tr>
<tr>
<td>mimp6</td>
<td>4</td>
</tr>
<tr>
<td>Other mimp s</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>64</td>
</tr>
</tbody>
</table>

a To determine the distance of mimp reinsertion sites to genes, the annotated version of the F. oxysporum genome sequence was used; it is available at the Broad Institute web site (http://www.broad.mit.edu/annotation/genome/fusarium_group/MultiHome.html)

Table 2 | Sizes, predicted ΔG values, and transposition efficiencies of native mimp1 and mimp2 and the artificial mimpNia element |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>mimp element</td>
<td>Size (bp)</td>
<td>ΔG with TIRs (kcal/mol)</td>
<td>Excision a</td>
</tr>
<tr>
<td>mimp1</td>
<td>224</td>
<td>−80.50</td>
<td>20–100</td>
</tr>
<tr>
<td>Mimp2</td>
<td>213</td>
<td>−87.69</td>
<td>20–100</td>
</tr>
<tr>
<td>mimpNia</td>
<td>211</td>
<td>−59.07</td>
<td>1–2</td>
</tr>
</tbody>
</table>

a Number of Nia+ colonies detected per minimal medium agar plate (see Materials and Methods section)
b Determined by PCR using mimp1-specific primers (Dufresne et al. 2007)
c Determined by Southern blot analysis using a mimp2- or mimpNia-specific probe, respectively
To further test the impact of mimps central region, we constructed an artificial mimp element in vitro, exhibiting the following structure: typical 27-bp impalaE-like TIRs and a central region, giving an overall size of 211 bp. We selected a region in the first exon of the *F. oxysporum* *nia* gene showing low homology with the *F. graminearum* genome. The corresponding mimp element was called mimpNia. The size and ΔG₀ value for the mimpNia element compared with those of mimp1 and mimp2, with and without TIRs, are presented in Table 2. This element was then introduced into the first intron of the *A. nidulans* *niaD* gene leading to plasmid construct pNmNia (see Materials and Methods section). Among eight cotransformants carrying the pNmNia construct, two (N4 and N5) gave rise to only a few Nia⁺ colonies (one or two per plate; Table 2) within a 4- to 6-week incubation time. To evaluate the behavior of the artificial mimp, Southern blot experiments were conducted. A set of 11 revertants, resulting from the excision of mimpNia, was analyzed. Hybridizing bands observed using the *niaD* probe (Fig. 3b; top) showed that they all correspond to excision events. Sequencing of the corresponding empty sites showed that eight of them exhibited typical excision footprints as already found for native mimp1 and mimp2 elements (Fig. 4). In the three remaining cases, more or less extended deletions were observed (Fig. 4). These modifications (typical excision footprints or deletions) do not alter splicing of the *niaD* intron, explaining why a wild-type nitrate-reductase activity was restored in these strains. Concerning the fate of the excised copy, the use of the mimpNia element as a probe showed reinsertion in at least three cases (lanes 2, 3, and 11 in Fig. 3b, bottom, and Table 2).

Altogether, these results demonstrate that mimp2 and also the in vitro created mimpNia element can be transactivated by the impalaE transposase and transpose through a canonical “cut-and-paste” mechanism. However, looking in more detail at the different steps in the transposition process, some differences are observed. The mimp2 element differs from mimp1 only by a less frequent reinsertion of the excised copy (Table 2). In contrast, the overall transposition efficiency of the mimpNia element is highly reduced: excision events are rare and <30% of the excised copies reinsert in the genome (Table 2). These results suggest that if TIR sequences are the minimal requirement for mobilization by the impalaE transposase through a cut-and-paste mechanism, additional parameters are likely required for efficient transposition.

Genomic Organization and Insertion Sites of mimp Elements

Looking at the distribution of mimp elements on the 15 *F. oxysporum* chromosomes, a strong bias was revealed (Fig. 5a), with seven chromosomes containing no mimp element and one, chromosome 14, carrying nearly 50% of the copies (48/103; Fig. 5a). This biased distribution is not due to one specific mimp subfamily but is rather a tendency for all mimp elements (data not shown). No strict correlation between the distribution of mimp and that of impala elements could be established (Fig. 5a) since chromosomes were found carrying mimps and no impala (chromosome...
3), and vice versa (chromosomes 2 and 8). However, a higher density of both elements was noticed on chromosome 14 (Fig. 5a). To explain this bias, we examined the gene content of the 15 chromosomes. As shown in Fig. 5a, four chromosomes are characterized by a relative paucity of genes. Among them, chromosome 14 is the only one exhibiting such a high concentration of mimp elements. Looking in more detail at the location of mimp elements on chromosome 14, we noticed that insertion sites are not random. Half of the copies (22/48) cluster together, spaced by 40 bp to 22 kbp, most of them being located in a 100-kbp interval on supercontig 2.51 (Fig. 5b). No relationship with either gene density or TA content could be established to explain the concentration of mimp elements on this interval (Fig. 5b). One mimp4 element was found inserted within a mimp3 copy.

To analyze mimp2 insertional preference in the F. graminearum genome, we performed a systematic recovery of flanking sequences using tail PCR. After sequencing, BLASTN searches on the F. graminearum genome sequence were used to locate 42 independant reinsertion sites, according to the third release of the assembled and annotated genome, available in the Fusarium Group database at the Broad Institute (http://www.broad.mit.edu/annotation/genome/fusarium_group/MultiHome.html). Most insertions were unique (34/42), however, four were hit twice. In contrast to the observations for native mimps in the F. oxysporum genome, the 38 independent mimp2 reinsertion sites seem to occur rather randomly over the four F. graminearum chromosomes, even though a slight bias is revealed on chromosomes 2 and 4, with fewer insertions observed than expected on chromosome 2 and more on chromosome 4 (Table 3).

The analysis of 60 nt around the insertion site of both the 64 independent full-length mimp elements in F. oxysporum and the 38 independent mimp2 reinsertion sites in F. graminearum revealed that, in addition to the TA dinucleotide target site, no consensus sequence could be found at mimp insertion sites. However, mimp elements tend to target genomic regions enriched in A/T nucleotides. For native mimps, the AT content of insertion sites varies from 56.3% (mimp1 subfamily) to 65.5% (mimp5 subfamily), whereas the average value for the F. oxysporum genome is 51.6%, reaching 54.6% in noncoding regions (http://www.broad.mit.edu/annotation/genome/fusarium_group/MultiHome.html). Similar results were found when analyzing mimp2 reinsertion sites in the F. graminearum genome (data not shown).

**Table 3** Observed and expected numbers of mimp2 insertions on the four F. graminearum chromosomes: calculations were made assuming that all TA dinucleotides are equally accessible

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>543,481</td>
</tr>
<tr>
<td>2</td>
<td>431,023</td>
</tr>
<tr>
<td>3</td>
<td>360,659</td>
</tr>
<tr>
<td>4</td>
<td>372,165</td>
</tr>
<tr>
<td>5</td>
<td>1,707,328</td>
</tr>
</tbody>
</table>

**Fig. 4** Structure of mimpNia excision sites in F. graminearum revertants. The TA dinucleotide target site is indicated in boldface. Nucleotide insertions or deletions relative to the niaD sequence without the mimpNia element are framed. n: number of revertants exhibiting the indicated nucleotide sequence.
To determine whether the targeting of AT-rich regions correlates with a preferential location in noncoding regions, we analyzed the regions flanking the 64 full-length mimp elements and calculated the distance toward the closest predicted gene in the F. oxysporum genome. One and six-tenths percent of the copies lie within a predicted gene (one element in an intronic region), 18.7% lie <500 bp from an ORF, and 29.7% lie >500 and <1000 bp from an ORF (Table 1), while the average intergenic distance in the F. oxysporum genome is 1974 bp. The location of mimp endogenous elements varies from one subfamily to another. For example, whereas mimp4 elements are located predominantly far from genes (only 11% at <500 bp from an ORF; see Table 1), mimp1 elements are closer to genes (nearly 40% at <500 bp from an ORF; see Table 1). A similar analysis of mimp2 reinsertion sites in the cross-mobilization assay showed that more than 45% (18/38) were found to be within (3/38) or in the close vicinity of (<500 bp upstream or downstream [15/38]) predicted genes (Table 4), with a preference for 5′-regions (11 vs. 4 insertions in the 5′- and 3′-regions, respectively). These results are very similar to those observed for mimp1 in the same genetic background (Dufresne et al. 2007; unpublished results).

**Discussion**

In this study we took advantage of the recent availability of the F. oxysporum genome sequence to perform a genome-wide analysis of the mimp elements, a family of MITEs for which a few members had previously been identified in this fungal species (Hua-Van et al. 2000). Two approaches were developed. The first consisted in using sequences of two known mimp elements, mimp1 and mimp2 (Hua-Van et al. 2000), and of two novel mimp3 and mimp4 elements identified by a PCR approach (this study) as queries in standard BLASTN searches (Altschul et al. 1997). This analysis allowed the identification of both full-length and truncated copies. However, divergent mimp elements could have been missed in such a similarity-based search. We thus extended our analysis by developing a computer program based only on the presence of similar TIR sequences (75% nucleotide identity) spaced by <500 bp, to extract all mimp elements in the genome of F. oxysporum. Similar strategies have been conducted in other eukaryote genomes such as Arabidopsis thaliana to analyze the Emigrant family (Santiago et al. 2002) and Anopheles gambiae (Tu 2001).
Overall, our analysis identified 103 mimp copies. This number could appear to be unusually low in comparison with most MITE families in plant and animal species. However, considering the size of the F. oxysporum genome (60 Mb), it is in the same range as that observed for the Emigrant family of MITEs in Arabidopsis (Santiago et al. 2002). Forty-three mimp elements were found in two to five identical copies, being part of larger redundant genomic regions. A phylogenetic analysis conducted on a subset of mimp copies showed that mimp elements belong to at least six different subfamilies, called mimp1 to mimp6. Eight mimp copies are too divergent to be included in any of the six defined subfamilies. Nevertheless, mimp full-length elements share common characteristics: impala-like TIRs, which may, depending on the subfamily, be more similar to impalaE (mimp1, mimp2, mimp3), impalaD (mimp6), or impalaF (mimp4, mimp5) TIR sequences, and a very homogeneous size, ranging between 204 and 226 bp. Tourist and Stowaway, two large plant MITE families, also exhibited this remarkable length homogeneity (Bureau and Wessler 1992, 1994a). Since then, size homogeneity has been found to be one of the characteristics of MITEs that distinguish them from other nonautonomous elements (Wessler et al. 1995). mimp subfamilies exhibit different degrees of variability. Whereas mimp1 and mimp6 are very homogeneous (93% and 98% nucleotide identity, respectively), elements belonging to the mimp4 subfamily are more divergent from one another (74% nucleotide identity). This suggests that different amplification bursts, each starting from a distinct mimp element and occurring at different times, gave rise to these subfamilies. The absence of any active copy of impala, the mimp autonomous partner (Dufresne et al. 2007; this work), in this genome prevents further amplification of mimps at this stage (M. Dufresne and M. J. Daboussi, unpublished results). However, different situations might be encountered in other F. oxysporum strains in which potentially active members of the impala family are still present (Hua-Van et al. 2001a).

A strong bias of distribution of endogenous mimp elements was revealed, whatever mimp subfamily, when looking at their insertion sites along the 15 F. oxysporum chromosomes. Seven chromosomes do not contain any mimp element, whereas chromosome 14 carries 46.6% of the 103 identified mimp copies. The density of mimp elements on chromosome 14 is two to seven times higher than that observed for TEs belonging to other superfamilies (L. J. Ma, Broad Institute; our laboratory, unpublished results). Detailed analysis of the distribution on this chromosome revealed that mimp elements tend to group together, as nearly half of the copies are found adjacent, as close as 40 bp. Only a single case of a mimp insertion into another mimp was observed, a situation different from that reported for Tourist MITEs (Jiang and Wessler 2001). The enrichment in mimps on this chromosome, contrary to the observations on other chromosomes such as chromosomes 3 and 6, is not due to the occurrence of segmental duplications. TEs have been reported to accumulate in some genomic regions, such as heterochromatin. However, the concentration of nearly half of mimp copies in such a small region of the genome is outstanding. The only element for which such a peculiar distribution was observed is the Bar1 transposon of Drosophila. This element is present in D. melanogaster both as single-copy elements scattered in the euchromatin and as a discrete cluster in the heterochromatin (Marsano et al. 2003). The similar distribution of mimp and Bar1 elements could suggest that the F. oxysporum genomic region in which mimp elements are found concentrated may be in the process of heterochromatinization. Alternatively, this distribution may result from local hopping of mimp elements, a phenomenon already described for other transposons such as the P, Ac, and Sleeping Beauty elements (Mátés et al. 2007). However, such a behavior has never been reported for the master element impala in F. oxysporum, in which reinsertion has been shown to occur randomly (Migheli et al. 2000). Moreover, the mimp1 and mimp2 elements do not exhibit such an insertion bias in the heterologous species F. graminearum (Dufresne et al. 2007; this work).

Plant MITE families have often been found associated with genes (Bureau and Wessler 1992, 1994a, b; Bureau et al. 1996; Yang et al. 2001; Han and Korban 2007). The distribution of mimp elements belonging to the six different subfamilies with respect to predicted genes was examined. On average, mimp elements were found far from genes, with more than 75% of the insertions located >500 bp from the closest ORF (see Table 1). However, this inser- tional preference was far less pronounced for the mimp1 subfamily, in which nearly 40% of the insertions were found <500 bp from a predicted ORF. A general tendency could be that mimp elements close to genes have been eliminated during evolution. This phenomenon is expected to be less apparent for more recent elements, such as those belonging to the very homogeneous mimp1 subfamily. Our results are in contrast with the analysis conducted on the Emigrant family of MITEs in Arabidopsis, in which copies close to genes were shown to have been maintained during evolution (Santiago et al. 2002).

In analysis of the location of newly mobilized mimp2 elements toward predicted genes in F. graminearum, we obtained a rather different picture. Nearly half of the insertions were found <500 bp from a predicted ORF. This result may reflect the propensity of mimp2 to target genic regions and thus favor the hypothesis that mimp elements, including endogenous mimp2 elements, close to genes have been negatively selected during the evolution of the genome of the F. oxysporum sequenced strain. Alternatively, the
observed discrepancy between insertional preference of endogenous mimp2 elements in *F. oxysporum* and newly transposed mimp2 copies in *F. graminearum* could be due to structural differences between the two genomes. The genome of *F. graminearum* is characterized by a high gene density and a remarkable paucity in repeated sequences (Cuomo et al. 2007). In contrast, *F. oxysporum* genome exhibits a high expansion of repeated sequences together with the occurrence of several large segmental duplications.

In a previous study, we demonstrated that mimp1 could be mobilized through the action of the transposase of *impalaE*, a Tc1-like autonomous element (Dufresne et al. 2007). Genome-wide analysis of the mimp elements in *F. oxysporum* showed that a number of them have TIRs similar to those of the autonomous *impalaE* element. To further evaluate the potential of *impalaE* to mobilize mimp elements as well as to evaluate the impact of sequences of the central region of mimp elements, we used a second native mimp element, mimp2, and an in vitro created mimp element, mimpNia, both carrying *impalaE*-like TIR sequences, in transactivation experiments. We demonstrated that mimp2 and mimpNia are mobilized by the *impalaE* transposase, and transpose through a “cut-and-paste” mechanism, providing evidence that *impalaE* is the autonomous partner of different mimp subfamilies.

However, transposition efficiencies appear to be lower than that observed for the native mimp1 element, in particular, for the mimpNia element, for which a drastic reduction in both excision and reinsertion was observed. The TIR sequences of the three mimp elements tested so far in our transposition assay differ by only a single nucleotide from *impE* TIR sequences, and the mimp1 and mimpNia elements exhibit identical TIRs. Therefore, even if TIR sequences appear to be the main requirement for mimp mobilization, these results suggest that the internal sequence might influence transposition efficiency. Results of binding experiments between the *Lemil* transposase and *Emigrant* MITEs showed that subterminal repeated motifs are important for transposase interaction and, possibly, for mobilization of these elements (Loot et al. 2006). No such sequence feature was found in mimp1 and mimp2, the two native mimp elements for which mobilization has been demonstrated. Another parameter that could explain the difference in transposition efficiency is the lower potential of the element we created to form a stable secondary structure. However, this does not stand for mimp2, which exhibits as mimp1 a strong potential to form a stable secondary structure. Deciphering which parameters are crucial for efficient transposition will be investigated further. One way would be to construct mimp elements corresponding either to chimera or to mutated versions of native mimp elements and to compare their transposition efficiency. However, to ensure a proper quantitative comparison, this will require assignment of the different constructs to the same genomic locus to avoid some of the variations observed being due to the effect of donor loci.

The relationships between mimp elements and impala autonomous members will also be analyzed further. Our exhaustive analysis of mimp elements in *F. oxysporum* has shown that three groups of mimps are found carrying TIRs resembling those of *impalaE*, *impalaD*, or *impalaF*, respectively. The potential of autonomous elements other than *impalaE* to mobilize different mimps will be tested. Previous studies have shown that *impalaD*, the only other available autonomous copy of *impala*, can cross-mobilize defective *impalaE* elements (Hua-Van et al. 2001b). A similar result is expected for mimp elements exhibiting TIRs similar to *impalaE*. Testing whether cross-mobilization of mimps carrying more divergent *impalaF* TIRs can occur will be of particular interest.

Acknowledgments This work was supported by CNRS funding. We are grateful to Pierre Capy and Aurélie Hua-Van for critical reading of the manuscript.

References


Bureau TE, Wessler SR (1994a) Mobile inverted-repeat elements of the *Tourist* family are associated with the genes of many cereal grasses. Proc Natl Acad Sci USA 91:1411–1415

Bureau TE, Wessler SR (1994b) *Stowaway*: a new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. Plant Cell 6:907–916


Feschotte C, Mouches C (2000) Evidence that a family of miniature inverted-repeat transposable elements (MITEs) from the *Arabidopsis thaliana* genome has arisen from a pogo-like DNA transposon. Mol Biol Evol 17:730–737


Bureau TE, Wessler SR (1994a) Mobile inverted-repeat elements of the *Tourist* family are associated with the genes of many cereal grasses. Proc Natl Acad Sci USA 91:1411–1415

Bureau TE, Wessler SR (1994b) *Stowaway*: a new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. Plant Cell 6:907–916


Feschotte C, Mouches C (2000) Evidence that a family of miniature inverted-repeat transposable elements (MITEs) from the *Arabidopsis thaliana* genome has arisen from a pogo-like DNA transposon. Mol Biol Evol 17:730–737


