The Plant Journal (1998) **16**(1), 79–85

SHORT COMMUNICATION

Presence of miniature inverted-repeat transposable elements (MITEs) in the genome of *Arabidopsis thaliana*: characterisation of the *Emigrant* family of elements

Elena Casacuberta, Josep M. Casacuberta, Pere Puigdomènech* and Amparo Monfort
Departament de Genètica Molecular, Centre d'Investigació i Desenvolupament (CSIC), Jordi Girona 18, 08034 Barcelona, Spain

Summary

Although the genome of *Arabidopsis thaliana* has a small amount of repetitive DNA, it contains representatives of most classes of mobile elements. However, to date, no miniature inverted-repeat transposable element (MITE) has been described in this plant. Here, we describe a new family of repeated sequences that we have named *Emigrant*, which are dispersed in the genome of *Arabidopsis* and fulfil all the requirements of MITEs. These sequences are short, AT-rich, have terminal inverted repeats (TIRs), and do not seem to have any coding capacity. Evidence for the mobility of *Emigrant* elements has been obtained from the absence of one of these elements in a specific *Arabidopsis* ecotype. *Emigrant* is also present in the genome of different *Brassicae* and its TIRs are 74% identical to those of *Wujin* elements, a recently described family of MITEs from the yellow fever mosquito *Aedes aegypti*.

Introduction

Transposable elements have been divided into two classes according to their mode of propagation. Class I elements, also known as retrotransposons, transpose via an RNA intermediate, while class II elements transpose by a DNA–DNA mechanism. Elements of both classes have been described in plants and both seem to be widely distributed, although retrotransposons are by far the most abundant (Grandbastien, 1992; Saedler and Gierl, 1996).

In the last few years a new class of transposable elements, called MITEs (miniature inverted-repeat transposable elements), has been described in plants (Wessler *et al.*, 1995). These elements share features of both class I and class II elements and, therefore, remain unclassified (Wessler *et al.*, 1995). The different MITEs described so far share structural, but not sequence, similarity. They are short A/T rich DNA sequences, have no coding capacity, have potential to form DNA secondary structure, and are flanked by inverted repeat sequences (Wessler *et al.*, 1995). Elements from the same family share similar inverted-repeat sequences, but only elements belonging to the same subfamily have internal sequence similarities (Bureau and Wessler, 1994a; Río *et al.*, 1996).

While MITEs were first described in plants, other short interspersed repeated elements having some characteristics of MITEs are present in animals, e.g. *Xenopus laevis* (Morgan and Middleton, 1990; Ünsal and Morgan, 1995) and humans (Morgan, 1995; Smith and Riggs, 1996). Recently, different families of mobile elements with all the characteristics of MITEs have been described in the yellow fever mosquito *Aedes aegypti* (Tu, 1997).

Arabidopsis thaliana has one of the smallest known genome among higher plants (Leutwiler *et al.*, 1984), and contains a very low amount of interspersed repetitive DNA that constitute only approximately 4% of its sequence (Meyerowitz, 1994). Nevertheless, mobile elements of class I (Konieczny *et al.*, 1991; Pélissier *et al.*, 1995; Voytas and Ausubel, 1988; Voytas *et al.*, 1990; Wright *et al.*, 1996) and class II (Frank *et al.*, 1997; Tsay *et al.*, 1993) have been described in this plant. On the other hand, although attempts to identify MITE elements in the genome of *Arabidopsis* have been made (Bureau *et al.*, 1996), until now no elements of this type have been identified. We describe here a new family of short repetitive elements from *Arabidopsis* that we have named *Emigrant*. The characteristics of *Emigrant* (*Emi*) elements are consistent with their being the first family of MITE elements described in *Arabidopsis*.

Results and discussion

Emigrant *is a new family of MITEs from* Arabidopsis thaliana

During the characterisation of the *Arabidopsis thaliana* chromosome IV genomic sequences obtained in our laboratory (within the framework of the European Arabidopsis Genome Project), a short sequence was found to have a high level of sequence similarity with four sequences

Received 28 May 1998; revised 28 July 1998; accepted 5 August 1998.
*For correspondence (fax +343 204 5904; e-mail pprgmp@cid.csic.es).

Table 1. Characteristics of *Emigrant* elements

	Chromosome	A+T (%)	Similarity to consensus (%)	Closest ORF prediction	ΔG (kCal mol^{-1})	Size (bp)	Acc. Number
Emi 1	I	71.8	86	n.d.	n.d.	235	AC000103
Emi 2	I	81.1	89	n.d.	–56.2	529	AC000107
Emi 3	I	83.3	81	n.d.	–87.9	604	AC000098
Emi 4	IV	75.0	95	1.2kb	n.d.	372	T19P19*
Emi 5	II	81.1	88	1.1kb	–58.3	533	u78721
Emi6	II	81.8	88	n.d.	–59.9	547	AC002505
Emi7	II	80.8	87	n.d.	–55.1	518	AC003673
Emi8	II	82.2	82	n.d.	–39.6	461	AC003673
Emi9	IV	81.3	86	1.5kb	–57.1	540	Z97344
Emi10	II	82.5	76	n.d.	–49.3	529	AC002521
Emi11	IV	82.6	84	1.5kb	–59.9	518	AF13294
Emi12	IV	81.6	89	8kb	–75.7	545	Z97337
Emi13	IV	81.4	90	400pb	–41.6	437	Z97337
Emi14	II	81.3	87	1kb	–69.5	530	AC003000

ΔG° was not determined (n.d.) when only partial sequence was available. ORF predictions are not available for sequences flanking *Emi1*, *Emi2*, *Emi3*, *Emi6*, *Emi7*, *Emi8*, and *Emi10*. Thus, the closest ORF prediction is not shown (n.d.).
*The name of the BAC clone containing this element is given.

dispersed in the *Arabidopsis* genome. A careful search in databases, using these five sequences as a query, revealed that the genome of *Arabidopsis* contains at least 14 short sequences displaying a high degree of sequence similarity (see Table 1). Other sequences showing a more limited degree of sequence similarity are also present in the genome of this plant (not shown). These sequences are found in different locations on different chromosomes of *Arabidopsis* (see Table 1). We have named this new family of repetitive sequences *Emigrant* (*Emi*).

We have not found any sequence similarity between *Emigrant* sequences and any other repetitive sequence described to date. Nevertheless, *Emigrant* has some of the features of transposable elements. The localization of the 14 *Emi* elements in three different chromosomes (see Table 1), as well as the result of the Southern blot hybridizations (not shown), suggest that this element is dispersed in the *Arabidopsis* genome. The ends of *Emigrant* elements are inverted repeated sequences of 24 nt (see Figure 1). Terminal inverted repeats (TIRs) are characteristics of class II transposons, and also of the new class of short repeated elements known as MITEs. Like MITEs, *Emigrant* elements do not seem to have any coding capacity, they are AT-rich and have the potential to form stable secondary structures with ΔG° values comparable to those reported for other families of MITEs (Bureau and Wessler, 1992; Bureau and Wessler, 1994a,b) (see Table 1). In addition, *Emi* elements are flanked by the dinucleotide TA which could represent target site duplication generated upon insertion (see Figure 1), and coincides with the TA(A) target site duplications of MITEs (Bureau and Wessler, 1992; Bureau and Wessler, 1994a,b; Bureau *et al.*, 1996; Río *et al.*, 1996; Tenzen *et al.*, 1994). Since the TIR sequences of *Emigrant* elements do not have any significant homology with those of other plant MITE families described (see Figure 2), we propose that the *Emigrant* elements are a new family of MITEs.

While the Arabidopsis genome contains class I (Konieczny *et al.*, 1991; Pélissier *et al.*, 1995; Voytas and Ausubel, 1988; Voytas *et al.*, 1990; Wright *et al.*, 1996) and class II (Klimyuk and Jones, 1997; Tsay *et al.*, 1993) transposons, no MITEs have been described in this plant until now. A short transposon-like element, *Tat1* (Peleman *et al.*, 1991), that could resemble this type of element, has been previously described in *Arabidopsis*. However, *Tat1* does not seem to have a target site preference, in contrast to the preference for TA(A) of the different families of MITEs described. In addition, the duplications generated by *Tat1* insertion are of 5 nt, which is longer than the two or three nucleotides of typical MITEs (Bureau and Wessler, 1992; Bureau and Wessler, 1994a,b; Bureau *et al.*, 1996; Río *et al.*, 1996; Tenzen *et al.*, 1994). Moreover, recent evidence suggests that the already described *Tat1* sequences are solo-LTR derivatives of a LTR-retrotransposon of the Ty3-*gypsy* family (Wright and Voytas, 1998). Thus, *Emigrant* is the first family of MITEs described in *Arabidopsis*.

Evidence for mobility of Emigrant *elements*

We have studied the possible mobility of *Emigrant* elements by looking for their presence at particular sites among different *Arabidopsis* ecotypes. PCR amplification of seven regions that contain an *Emigrant* element in *Columbia* ecotype revealed the polymorphic presence of two of them among the four *Arabidopsis* ecotypes studied here. Figure 3 presents the analysis of one of these polymorphic regions.

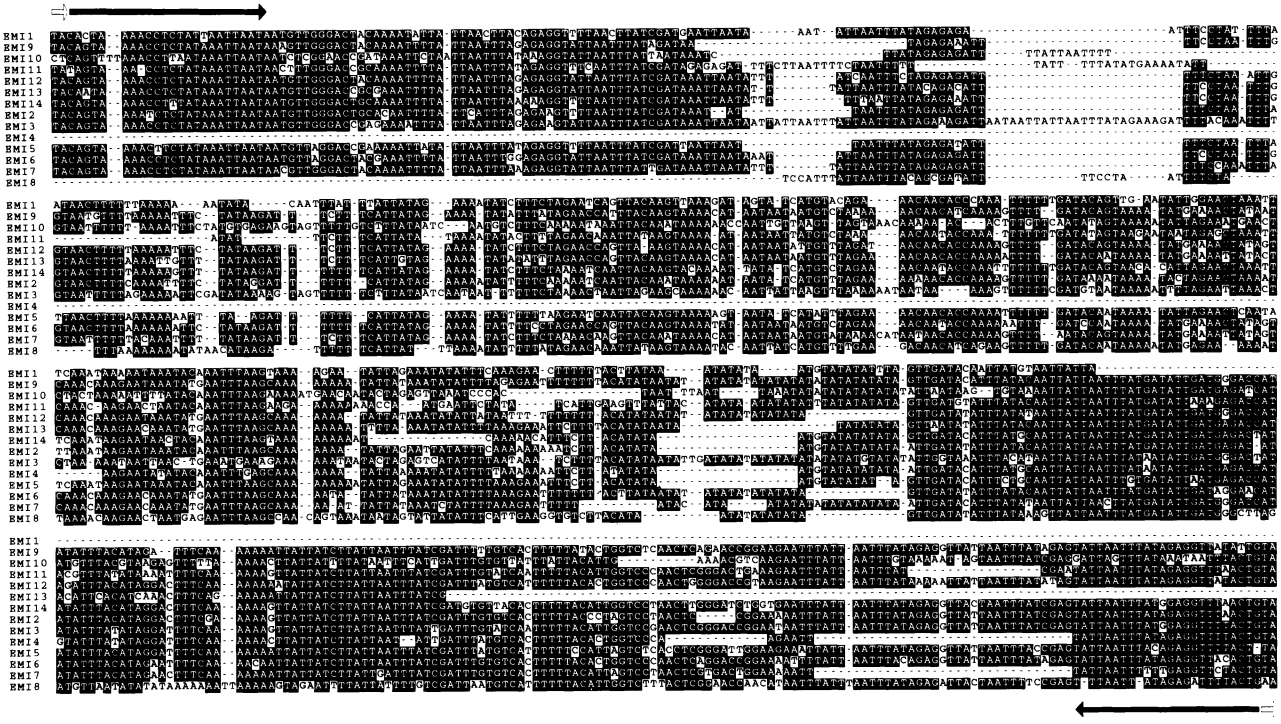

Figure 1. Multiple sequence alignment of *Emigrant* elements.
The 14 *Emigrant* sequences were aligned using the CLUSTAL V program (UWGCG). Conserved nucleotides are shown by white letters on a black background. The gaps are shown by dotted lines. The inverted-repeated sequences are indicated by black arrows, and the target site duplications by open arrows. *Emi1, Emi4* and *Emi13* are truncated copies.

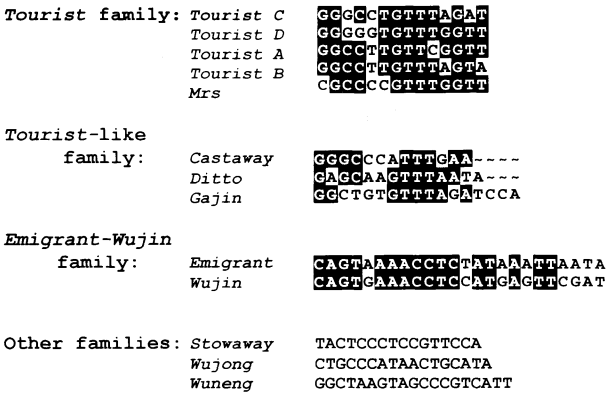

Figure 2. Comparison of the TIR sequences of the different families of MITEs.
TIR sequences of the *Tourist* family of elements are shown according to Bureau and Wesler (1992) (*Tourist A)*, to Bureau and Wesler (1994a) (*Tourist B, C and D*), and to Río *et al.* (1996) (*Mrs*); those of *Tourist*-like family according to Bureau *et al.* (1996). The rest of the TIR sequences are shown according to Bureau and Wesler, (1994b) (*Stowaway*) and to Tu (1997) (*Wujin, Wujong* and *Wuneng*).

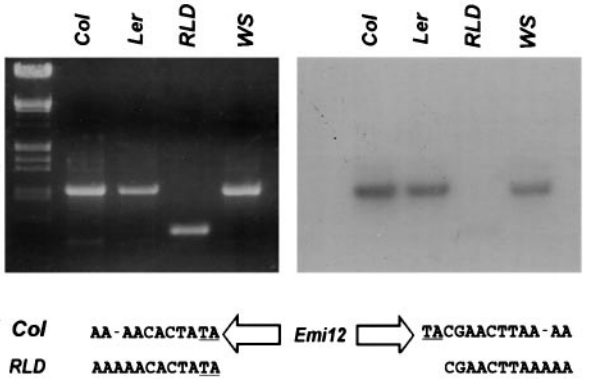

Figure 3. Polymorphic presence of *Emi12* in different *Arabidopsis* ecotypes Etidium bromide staining (left box) and hybridization with an *Emi* specific probe (right box) of the PCR products amplified with oligonucleotydes flanking the *Columbia Emi12* element from DNA obtained from *Columbia (Col), Landsberg erecta (Ler), WS* and *RLD Arabidopsis* ecotypes. The sequence flanking the *Emi12* element in *Columbia (Col)* ecotype and the sequence of the corresponding empty site in *RLD* ecotype are shown below. The TA target site duplicated after insertion is underlined.

This result suggests that *Emigrant* elements have actively transposed since the divergence of these ecotypes. The comparison between the sequences containing an *Emigrant* insertion with the corresponding empty sites has allowed us to confirm that *Emi* elements generate a duplication of the dinucleotide TA upon insertion (see Figure 3), which coincides with the consensus TA(A) target site duplication of previously described MITE elements (Bureau and Wessler, 1992; Bureau and Wessler, 1994a,b; Bureau *et al.*, 1996; Río *et al.*, 1996; Tenzen *et al.*, 1994).

Emigrant *is not associated with genes in* Arabidopsis

The presence of 14 highly homogeneous *Emi* elements over the 15 Mbp of genomic *Arabidopsis* DNA, available

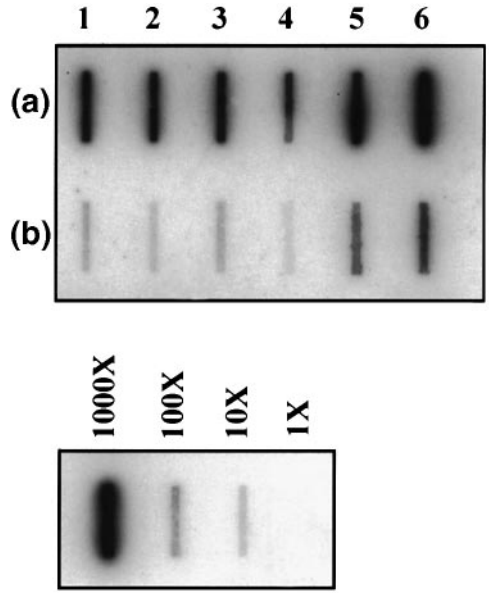

Figure 4. Distribution and copy number of *Emigrant* elements
(a) 1 μg and (b) 0.1 μg of total genomic DNA from *Columbia* (1), *Landsberg erecta* (2), *RLD* (3), *WS* (4), *Arabidopsis* ecotypes, and 4 μg (a) and 0.4 μg (b) of total genomic DNA from *Brassica napus* (5) and *Brassica juncea* (6) were hybridized with an *Emi4* probe. The hybridization obtained with different amounts of DNA of a clone containing the *Emi4* element corresponding to 1000, 100, 10 and 1 copies of the *Emigrant* element are shown for quantification.

through the databases since December 31st 1997, suggests that if *Emi* elements were homogeneously distributed, *Arabidopsis* should contain around 150 of these highly conserved sequences within the 145 Mbp of its genome. However, the number of *Emi*-related sequences should be higher, as sequences having a more limited degree of similarity have also been detected in these searches (not shown). In order to determine the number and distribution of *Emi*-related sequences, we have analysed by slot-blot hybridizations their presence in different *Arabidopsis* ecotypes and *Brassicae*. The results present in Figure 4 show that all the different *Arabidopsis* genomes analysed contain between 500 and 1000 *Emi*-related sequences, and that this element is also present in other *Brassicae*. *Emigrant* is thus less abundant than other MITEs, which can be present at more than 10 000 copies per genome (Wessler *et al.*, 1995).

MITEs, as well as retrotransposons, have frequently been found to be associated with genes in plants (Wessler *et al.*, 1995; White *et al.*, 1994). However, genomic sequencing projects have shown that the organisation of the genome of *Arabidopsis* may be different in some aspects to that of other plant genomes. Indeed, retro-elements seem to be dispersed in the genome of *Arabidopsis* (Bevan *et al.*, 1998) in clear contrast to the pattern of retro-elements in larger genomes such as maize, where retrotransposons form nested structures of multiple elements comprising at least 50% of the nuclear DNA of the plant (SanMiguel *et al.*, 1996). The *Emi* elements described here lie in non-coding regions, and only one of them has an open reading frame prediction within 1 kb upstream or downstream (see Table 1). It would seem that, in contrast to other MITEs (Wessler *et al.*, 1995), *Emi* elements are present in low copy number in the genome of *Arabidopsis* and are not frequently associated with genes. As *Emigrant* elements are longer than other previously described families of MITEs, their insertion within transcribed regions is more likely to interfere with gene expression. This could be a possible explanation for its particular pattern of insertion. Nevertheless, if other MITEs exist in Arabidopsis, they probably share this characteristic with *Emi* elements, as recent computer-based searches that have detected 37 MITE sequences within rice genes have failed to detect these elements in the close vicinity of *Arabidopsis* genes, although there are four times as many *Arabidopsis* gene sequences than rice genes in the GenBank and EMBL databases (Bureau *et al.*, 1996). Because of the close association of MITEs with plant genes, it has been suggested that these elements could have been involved in the evolution of genes in plants (Bureau *et al.*, 1996; Wessler *et al.*, 1995). Nevertheless, the results presented here show that while MITEs are present in *Arabidopsis*, its impact on the evolution of gene regulation in this species has been less important than in other species, such as maize or rice. On the other hand, it has also been suggested that the association of MITEs with genes could be a consequence of their already unknown mechanism of transposition. If MITEs transpose by an RNA intermediate, their presence within transcribed regions could facilitate mobilisation (Río *et al.*, 1996). Alternatively, the association of MITEs with coding sequences could reflect a preference of this type of element for integration in transcribed sequences. The existence of MITEs not associated with genes in *Arabidopsis* suggests that this association is not essential for the transposition of these elements, although we cannot rule out the possibility that the elements described here were generated from other active *Emi* elements lying in the close vicinity of a gene.

The inverted repeats of Emigrant *are similar to those of* Wujin *from the yellow fever mosquito*

Within the 23 nt of *Emigrant* TIR sequences, 17 are identical to those of *Wujin* (see Figure 2), a recently described MITE in the yellow fever mosquito *Aedes aegypti* (Tu 1997). The sequence of the TIRs, as well as the size and sometimes the sequence of the target site duplication generated upon integration, are believed to be specific for each family of transposable element that share integration machinery. There is no sequence similarity between *Emigrant* and *Wujin* elements except in their TIRs. This is a similar situation to that found for the different subfamilies of plant

Tourist elements, which have 65–85% identity in their TIRs and little, or no, similarity in their internal sequences (Bureau and Wessler, 1994a; Río *et al.*, 1996). Therefore, *Emigrant* and *Wujin* are probably two different subfamilies of the same MITE family of elements, and constitute the first example of a MITE family present in two species that belong to different phylogenetic kingdoms. It is tempting to present this as an example of horizontal transfer between plant and animal genomes. Horizontal transmission events have been repeatedly proposed to explain the wide distribution of other mobile elements, such as copia-like retrotransposons, between very distant species (see Flavell *et al.*, 1994). Nevertheless, when an extensive sampling of elements from related species is performed, the results obtained are consistent with a vertical transmission-based evolution of these elements (VanderWiel *et al.*, 1993; Vernhettes *et al.*, 1998). If MITEs are transmitted mainly vertically, as retrotransposons seem to be, the presence of the same MITE family in the genomes of *Arabidopsis* and the yellow fever mosquito would indicate an ancient association of MITEs with the eukaryote genome. Alternatively, it could be an indication of a convergent evolution of the TIRs of both elements due to constraints imposed by the use of a conserved cellular machinery for their mobility.

Conclusion

The genome of *Arabidopsis thaliana* contains a very low amount of interspersed repetitive DNA (Meyerowitz, 1994). Nevertheless, it contains representatives of most classes of transposable elements. Indeed, more than 10 different families of LTR-retrotransposons of the Ty1-*copia* family (Konieczny *et al.*, 1991; Voytas and Ausubel, 1988; Voytas *et al.*, 1990), two different LTR-retrotransposons of the Ty3-*gypsy* family (Pélissier *et al.*, 1995; Wright and Voytas, 1998), and 17 different families of non-LTR retrotransposons (Wright *et al.*, 1996), as well as one class II transposon (Frank *et al.*, 1997; Tsay *et al.*, 1993) have been characterised in *Arabidopsis*. What makes the *Arabidopsis* genome different is that most of the characterised *Arabidopsis* retroelements are of low copy number, ranging from one to no more than sevem copies (Konieczny *et al.*, 1991; Wright *et al.*, 1996), with the only exception of the *Athila* retrotransposon which is present in 150 copies in *Arabidopsis* mostly associated with its major satellite (Pélissier *et al.*, 1996).

Until now, no MITEs have been described in *Arabidopsis*. The characterisation of the *Emigrant* family of elements shows that, as for the other families of transposable elements, the genome of *Arabidopsis* does contain MITEs. Nevertheless, *Emigrant* is present at a lower copy number than typical MITEs in other plant genomes. *Emigrant* elements may have been abundant in an ancestor of *Arabidopsis*, being mostly lost since then, as suggested for retrotransposons (Wright *et al.*, 1996). Alternatively, MITEs could have been unsuccessful in proliferating after being introduced in *Arabidopsis*. If *Emigrant* elements, in contrast to the previously described families of MITEs, avoid transcribed regions, it will perhaps be difficult for these elements to find targets due to the high gene density genome of *Arabidopsis*. In any case, our results show that, as for the other classes of mobile elements, MITEs are not as abundant in *Arabidopsis* as in other plant genomes. This suggests a general rule restricting mobile elements to a low copy number in *Arabidopsis*, which seems to control their activity more strictly. Subtle differences in the host DNA repair machineries of maize and *Arabidopsis* have recently been suggested to explain differences in the footprints generated after *Ac* excision in these two plants (Rinehart *et al.*, 1997). Thus, the constraints of the *Arabidopsis* genome to mobile element proliferation, in comparison to other plant genomes, could be a consequence of differences in the general cellular mechanisms responsible for genome dynamics and integrity.

Experimental procedures

DNA sequencing and computer analyses

The nucleotide sequence of *Emi 4* was determined by the dideoxynucleotide chain termination method using an automatic fluorescence sequencer (ABI377 Perkin-Elmer). Sequence similarity searches were made using FASTA and Blast programs of UWGCG, software package (Genetics Computer Group, Madison, WI, USA) against the AT Data Base which contains the last submission of the Arabidopsis Genomic project (http://genome-www.stanford.edu/Arabidopsis). Multiple alignments of sequences were performed using CLUSTAL V and Boxshade (UWGCG) programs. $\Delta G°$ values were calculated using the MFOLD program of the UWGCG package. The consensus *Emigrant* sequence used to calculate the percentage of sequence similarity of the different copies was constructed after CLUSTAL V prediction. The coding sequence predictions were made with the programme Genefinder (Green and Hillier, in preparation). The ORF predictions of clone 19P19 were made using BLAST analysis and the NetPlantGene Program (Hebsgaard *et al.*, 1996).

Slot blot analysis

DNA from four different *Arabidopsis* ecotypes (*Columbia, Landsberg erecta, RLD* and *WS*) and two *Brassica* species (*Brassica napus* and *Brassica juncea*) was obtained by standard procedures (Dellaporta *et al.*, 1984). One μg and 0.1 μg of total genomic DNA of each *Arabidopsis* ecotype, and 4 μg and 0.4 μg of total genomic DNA of *Brassica* species was denatured and applied to a Nytran membrane (Schleicher and Schuell). One ng, 0.1 ng, 10 pg and 1 pg of a plasmid which contains the *Emi4* element, corresponding to 1000, 100, 10 and 1 copies of the *Emigrant* element were also applied to the membrane. After neutralisation and fixation, the membrane was hybridized and washed at low stringency (20 mM Na_2HPO_4 pH:7.2, 1% SDS, 1 mM EDTA, at 37°C) with a probe corresponding to the *Emi4* element.

PCR amplifications

PCR amplifications were performed by standard procedures with oligonucleotides corresponding to sequences flanking the *Emi*12 element in *Columbia* ecotype (5′-GAGAGCTTTAGAGTGTCAT-ACC-3′ and 5′-GCGCCATGGAGGATACTCTTC-3′). PCR products were run in an agarose gel and transferred to a nylon membrane (Schleicher and Schuell) by standard procedures. The membrane was hybridized with an *Emigrant* specific probe and washed at medium stringency (20 mM Na_2PO_4 pH:7.2, 1% SDS, 1 mM EDTA at 50°C).

Acknowledgements

We acknowledge the support of the European Genome Project and Plan Nacional de Investigación Científica y Técnica (grant BIO97–1419-CE). This work has been carried out within the framework of the Centre de Referència de Biotecnologia de Catalunya.

References

Bevan, M., Bancroft, I., Bent, E. *et al.* The E.U. Arabidopsis Genome Project (1998) Analyses of 1.9 Mb.of contiguous sequence from chromosome IV of Arabidopsis thaliana. *Nature*, **391**, 485–488.

Bureau, T.E., Ronald, P.C. and Wessler, S.R. (1996) A computer based systematic survey reveals the predominance of small inverted-repeat elements in wild-type genes. *Proc. Natl Acad. Sci. USA*, **93**, 8524–8529.

Bureau, T.E. and Wessler, S.R. (1992) *Tourist*: A large family of small inverted repeat elements frequently associated with Maize genes. *Plant Cell*, **4**, 1283–1294.

Bureau, T.E. and Wessler, S.R. (1994a) Mobile inverted-repeat elements of the *Tourist* family are associated with the genes of many cereal grasses. *Proc. Natl Acad. Sci. USA*, **9**, 1411–1415.

Bureau, T.E. and Wessler, S.R. (1994b) *Stowaway*: A new family of inverted -repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. *Plant Cell*, **6**, 907–916.

Dellaporta, S.L., Wood, J. and Hicks, J.B. (1984) *Molecular Biology of Plants: A Laboratory Course Manual.* Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Flavell, A.J., Pearce, S.R. and Kumar, A. (1994) Plant transposable elements and the genome. *Curr. Opin. Genet. Devel.* **4**, 838–844.

Frank, M.J., Liu, D., Tsay, Y.F., Ustach, C. and Crawford, N.M. (1997) *Tag1* is an autonomous transposable element that shows somatic excision in both Arabidopsis and tobacco. *Plant Cell*, **9**, 1745–1756.

Grandbastien, M.A. (1992) Retroelements in higher plants. *Trends. Genet.* **8**, 103–108.

Hebsgaard, S.M., Korning, P.G., Tolstrup, N., Engelbrecht, J., Rouzé, P. and Brunak, S. (1996) Splice prediction in *Arabidopsis thaliana* pre-mRNA by combining local and global sequence information. *Nucl Acids Res.* **17**, 3439–3452.

Klimyuk, V.I. and Jones, J.D.G. (1997) *AtDMC1*, the *Arabidopsis* homologue of yeast *DMC1* gene: characterisation, transposon-induced allelic variation and meiosis-associated expression. *Plant J.* **11**, 1–14.

Konieczny, A., Voytas, D.F., Cummings, M.P. and Ausubel, F.M. (1991) A superfamily of *Arabidopsis thaliana* retrotransposons. *Genetics*, **127**, 801–809.

Leutwiler, L.S., Hough-Evans, B.R. and Meyerowitz, E.M. (1984) The DNA of *Arabidopsis thaliana. Mol. Gen. Genet.* **194**, 15–23.

Meyerowitz, E.M. (1994) Structure and organization of the *Arabidopsis thaliana* nuclear genome. In: *Arabidopsis.* (Meyerowitz, E.M. and Somerville, C., eds). Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Morgan, G.T. (1995) Identification in the human genome of mobile elements spread by DNA-mediated transposition. *J. Mol. Biol.* **254**, 1–5.

Morgan, G.T. and Middleton, K.M. (1990) Short interspersed repeats from Xenopus that contain multiple octamer motifs are related to known transposable elements. *Nucl. Acids Res.* **18**, 5781–5786.

Peleman, J., Cottyn, B., Van Camp, W., Van Montagu, M. and Inze, D. (1991) Transient occurrence of extrachromosomal DNA of an *Arabidopsis thaliana* transposon-like element, *Tat1. Proc. Natl Acad. Sci. USA*, **88**, 3618–3622.

Pélissier, T., Tutois, S., Deragon, J.M., Tourmentes, S., Genestier, S. and Picard, G. (1995) *Athila*, a new retroelement from *Arabidopsis thaliana. Plant. Mol. Biol.* **29**, 441–451.

Pélissier, T., Tutois, S., Tourmentes, S., Deragon, J.M. and Picard, G. (1996) DNA regions flanking the major *Arabidopsis thaliana* satellite are principally enriched in *Athila* retroelement sequences. *Genetica*, **97**, 141–151.

Rinehart, T.A., Dean, C. and Weil, C.F. (1997) Comparative analysis of non-random DNA repair following *Ac* transposon excision in maize and *Arabidopsis. Plant J.* **12**, 1419–1427.

Río, A., Puigdomènech, P. and Casacuberta, J.M. (1996) *Mrs*, a new subfamily of *Tourist* transposable elements. *Plant. Mol. Biol.* **32**, 1221–1226.

Saedler, H. and Gierl, A. (1996) Transposable elements. In *Current Topics in Microbiology and Immunology.* Berlin: Springer-Verlag.

SanMiguel, P., Tikonov, A., Jin, Y.-K. *et al.* (1996) Nested retrotransposons in the intergenic regions of the Maize genome. *Science*, **274**, 765–767.

Smith, F.A. and Riggs, A.D. (1996) *Tiggers* and other DNA transposon fossils in the human genome. *Proc. Natl Acad. Sci. USA,* **93**, 1443–1448.

Tenzen, T., Matsuda, Y., Ohtsubo, H. and Ohtsubo, E. (1994) Transposition of *Tnr1* in rice genomes to 5′-PuTAPy-3′sites, duplicating the TA sequence. *Mol. Gen. Genet.* **245**, 441–448.

Tsay, Y.-F., Frank, M.J., Page, T., Dean, C. and Crawford, N.M. (1993) Identification of a mobile endogenous transposon in *Arabidopsis thaliana. Nature*, **260**, 342–344.

Tu, Z. (1997) Three novel families of miniature inverted-repeat transposable elements are associated with genes of the yellow fever mosquito, *Aedes Aegypti. Proc. Natl Acad. Sci. USA*, **94**, 7475–7480.

Ünsal, K. and Morgan, G.T. (1995) A novel group of families of short interspersed repetitive elements (SINEs) in Xenopus: evidence for a specific target site for DNA-mediated transposition of inverted-repeat SINEs. *J. Mol. Biol.* **248**, 812–823.

VanderWiel, P.L., Voytas, D.F. and Wendel, J.F. (1993) *Copia*-like retrotransposable element evolution in diploid and polyploid cotton (*Gossypium L.*). *J. Mol. Evol.* **36**, 429–447.

Vernhettes, S., Grandbastien, M.-A. and Casacuberta, J.M. (1998) The evolutionary analysis of the Tnt1 retrotransposon in *Nicotiana* species reveals the high variability of its regulatory sequences. *Mol. Biol. Evol.* **15**, 827–836.

Voytas, D.F. and Ausubel, F.M. (1988) A *copia*-like transposable element family in Arabidopsis thaliana. *Nature*, **336**, 242–244.

Voytas, D.F., Konieczny, A., Cummings, M.P. and Ausubel, F.M. (1990) The structure, distribution and evolution of the *Ta1* retrotransposable element family of *Arabidopsis thaliana. Genetics*, **126**, 713–721.

Wessler, S., Bureau, T.E. and White, S.E. (1995) LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. *Curr. Op. Genet. Dev.* **5**, 814–821.

White, S.E., Habera, L.F. and Wessler, S.R. (1994) Retrotransposons in the flanking regions of normal plant genes: a role for *copia*-like elements in the evolution of the gene structure and expression. *Proc. Natl Acad. Sci. USA,* **91**, 11792–11796.

Wright, D.A., Ke, N., Smalle, J., Hauge, B.M., Goodman, H.M. and Voytas, D.F. (1996) Multiple non-LTR retrotransposons in the genome of Arabidopsis thaliana. *Genetics*, **142**, 569–578.

Wright, D.A. and Voytas, D.F. (1998) Potential retroviruses in plants: Tat*1* is related to a group of *Arabidopsis thaliana*.Ty3/*gypsy* retrotransposons that encode envelope-like proteins. *Genetics*, **149**, 703–715.