Identification and frequency of transposable elements in *Eucalyptus*

Mauricio Bacci Jr.¹, Rafael B.S. Soares¹, Eloíza Tajara², Guilherme Ambar², Carlos N. Fischer³, Ivan R. Guilherme³, Eduardo P. Costa³ and Vitor F.O. Miranda⁴

¹Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Centro de Estudos de Insetos Sociais, Rio Claro, SP, Brazil.
²Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas, São José do Rio Preto, SP, Brazil.
³Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Geociências e Ciências Exatas, Departamento de Estatística, Matemática Aplicada e Computação, Rio Claro, SP, Brazil.
⁴Universidade de Mogi das Cruzes, Herbarium Mogiense, Mogi da Cruzes, SP, Brazil.

Abstract

Transposable elements (TE) are major components of eukaryotic genomes and involved in cell regulation and organism evolution. We have analyzed 123,889 expressed sequence tags of the *Eucalyptus* Genome Project database and found 124 sequences representing 76 TE in 9 groups, of which copia, MuDR and FAR1 groups were the most abundant. The low amount of sequences of TE may reflect the high efficiency of repression of these elements, a process that is called TE silencing. Frequency of groups of TE in *Eucalyptus* libraries which were prepared with different tissues or physiologic conditions from seedlings or adult plants indicated that developing plants experience the expression of a much wider spectrum of TE groups than that seen in adult plants. These are preliminary results that identify the most relevant TE groups involved with *Eucalyptus* development, which is important for industrial wood production.

Key words: *Eucalyptus*, transposable element.

Received: May 28, 2004; Accepted: May 30, 2005.

Transposable elements (TE) are present in most of the eukaryotic cells. They represent up to 40% of the human genome size (Yoder et al., 1997; Smit, 1999) and 50-90% of the genome size of important agricultural plants, such as maize, wheat and barley (Flavel, 1986; SanMiguel et al., 1996; Shirasu et al., 2000). They are tightly related to chromosome structure and evolution and thus, ultimately, to organism evolution (McDonald, 1995; Britten, 1996; Kidwell and Lisch, 1997; Fedoroff, 2000).

Based on the mode by which TE can move from one location to another in the genome, they are divided into two classes. Class I TE are transcribed to RNA intermediates, reverse-transcribed and integrated in a new genome site. Because of this activity, members of this class are also called retroelements, which are further subdivided into retrotransposons (*e.g.* copia-like and gypsy-like) that have long terminal repeats (LTRs), as well as the so-called non-LTR retroelements (*e.g.* long (LINE) and short (SINE) interspersed nuclear elements). Class II TE are those presenting terminal inverted repeats (TIR) and capable of moving from one site to another in the genome through a ‘copy-and-paste’ process that involves the action of transposases, which interact with TIR (Berg and Howe, 1989). Class II is further subdivided accordingly to TE structure and sequence or features of the target duplication site generated upon insertion (Capy et al., 1996). Examples of groups from Class II are CACTA, which is flanked by inverted repeats that terminate in a conserved CACTA motif (Wicker et al., 2003); MuDR, which codes for a transposase (Benito and Walbot, 1997) or other genes (Lisch et al., 1999) whose function remains unknown and control the expression of TE in the *MU* system, which is the most active and mutagen transposable element described in plants (Rossi et al., 2004); hAT, which is widespread among fungi, plants and animals (Rubin et al., 2001); and *IS* (insertion sequence), which is commonly found in bacteria (Schnetz and Rak, 1995).

Knowledge of TE has advanced significantly with the sequencing of many genomes, which has led to the accumulation of a large number of sequences and resulted in the
The discovery of new TE such as those in the Jittery group, which are homologous to FAR1 and FHY3 genes (FAR1 family) and present regulatory functions (Hudson et al., 2003). Regulatory function of genes in the FAR1 family is proposed to involve a mechanism that is similar to that occurring in the linkage of transposase to the TIR of a Class II TE in the MuDR group (Hudson et al., 2003). Therefore, many of the genes which are currently classified in distinct groups of TE are likely to be involved in cell regulatory processes, so that by following TE expression patterns in different physiologic conditions one could identify putative regulatory genes.

In the present investigation, we have searched for TE in different tissues and physiological conditions in Eucalyptus, which is an important source of wood for industrial purposes. To accomplish that, we identified TE in libraries of the Eucalyptus Transcriptome Project (FORESTS, Table 1) through the utilization of a keyword search in the FORESTS database, which was carried out with keywords 'transposon,' 'transposase' or names of each group of TE. Only the retrieved EST (EST = expressed sequence tags) showing e-values ≤ 10⁻⁵ were considered. An additional search for sequences of TE in the FORESTS database was done through a blastn (Basic Local Alignment Search Tool, nucleotide-nucleotide) utilizing the query sequences from the GenBank, representing every group of TE. Then, sequences of TE that were retrieved from the FORESTS database were classified based on their similarity to previously described sequences in the GenBank. To achieve that, each Eucalyptus transposable element sequence was utilized as a query in a blastx (translated queries against the GenBank protein database) analysis and protein sequences showing e-values ≤ 10⁻⁵ and scores ≥ 80 were considered.

This procedure for identification and classification of TE resulted in 124 EST in 16 libraries of the FORESTS database, and these EST were grouped in 76 clusters (Table 2). Many of these clusters (57) were found as singletons, which may correspond to rarely expressed genes, although it is likely that some of these singletons represented different regions of a single gene, considering the high gene size of TE (up to 9 kb (Capy et al., 1996)) compared to sizes of EST that were generated in the FORESTS project (~ 800 bp). The remaining 19 identified clusters were composed of 2 to 9 EST each (Table 2), and 16 of these clusters were composed of EST from more than one library, which indicates that some of the identified TE are expressed in more than a single plant tissue.

Most of the identified TE were in Class I (retroelements) and belonged to the FAR1 (29.8%) or copia (22.6%) groups while other retroelements, such as LINE and gypsy, were poorly represented (4.0 and 2.4%, respectively). Within the Class II TE, MuDR (16.9%) and hAT (12.1%) groups prevailed, followed by CACTA (4.8%), non-classified LTR (4.0%) and IS (3.2%) groups (Table 2). These results are in agreement with the high amount of expressed MuDR found in sugarcane (Rossi et al., 1996).

Table 1 - EST libraries from Eucalyptus.

| Library | Source | Species | EST* |
|---------|--------|---------|------|
| BK1     | Bark, heartwood, softwood and medulla of an eight-year old tree | E. grandis | 1,052 |
| CL1     | Calli formed in the dark | E. grandis | 9,998 |
| CL2     | Calli formed in the presence of light | E. grandis | 2,535 |
| FB1     | Inflorescence and fruit | E. grandis | 12,275 |
| LV1     | Leaf from seedlings | E. grandis | 2,048 |
| LV2     | Leaf from disease-sensitive, phosphate and borate deficient tree | E. grandis | 7,532 |
| LV3     | Leaf colonized by Thyrinteina | E. grandis | 4,341 |
| RT3     | Root of seedlings | E. grandis | 13,252 |
| RT6     | Root of frost-resistant or susceptible trees | E. grandis | 6,877 |
| SL1     | Seedlings cultivated in the dark and exposed to light | E. grandis | 6,182 |
| SL4     | Seedlings cultivated in the dark | E. globulus | 6,718 |
| SL5     | Seedlings cultivated in the dark | E. saligna | 7,165 |
| SL6     | Seedlings cultivated in the dark | E. urophylla | 1,217 |
| SL7     | Seedlings cultivated in the dark | E. grandis | 4,120 |
| SL8     | Seedlings cultivated in the dark | E. calmodulensis | 2,035 |
| ST2     | Stalk of seedlings susceptible to water deficit (0.6-2.0 kb insert) | E. grandis | 11,032 |
| ST6     | Stalk of seedlings susceptible to water deficit (0.8-3.0 kb insert) | E. grandis | 12,558 |
| ST7     | Stalk of frost-resistant or susceptible tress | E. grandis | 2,728 |
| WD2     | Wood | E. grandis | 10,224 |

*Number of sequenced EST in each library.
Table 2 - Number of TE EST in different libraries of *Eucalyptus* and their corresponding clusters.

| TE and cluster* | Library / Class I | Total |
|-----------------|-------------------|-------|
|                 | BK1   | CL1   | CL2   | FB1   | LV2   | LV3   | RT3   | RT6   | SL1   | SL4   | SL5   | SL7   | SL8   | ST2   | ST6   | WD2   |
| copia           |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| 1               | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 2               | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 2     | -     | 2     | 2     |
| 3               | -     | -     | -     | -     | -     | -     | -     | -     | 1     | -     | -     | -     | -     | -     | -     | -     |
| 4               | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 5               | -     | -     | 1     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 6               | -     | -     | 1     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 7               | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     | -     | -     | 1     |
| 8               | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     | 1     |
| 9               | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 10              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     | -     | -     | -     | -     | 2     |
| 11              | -     | -     | 3     | 3     | 1     | -     | -     | -     | -     | -     | -     | 2     | -     | -     | 9     |
| 12              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 2     | 2     |
| 13              | -     | -     | 2     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 2     | -     | 1     | 5     |
| Subtotal        | 0     | 0     | 1     | 6     | 3     | 0     | 3     | 1     | 2     | 0     | 0     | 3     | 2     | 3     | 1     | 3     |
| gypsy           |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| 14              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 15              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 16              | -     | -     | -     | 1     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| Subtotal        | 0     | 0     | 0     | 1     | 1     | 0     | 0     | 0     | 0     | 0     | 0     | 1     | 0     | 0     | 0     | 0     |
| LINE            |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| 17              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     | -     | -     | 1     |
| 18              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 19              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     | -     | -     | 1     |
| 20              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     | 1     |
| 21              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| Subtotal        | 0     | 0     | 0     | 0     | 0     | 1     | 0     | 0     | 0     | 0     | 0     | 0     | 2     | 1     | 1     | 5     |
| nc              |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| 22              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 23              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 24              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 25              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     | -     |
| 26              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| Subtotal        | 0     | 0     | 0     | 0     | 0     | 1     | 0     | 0     | 0     | 0     | 0     | 0     | 2     | 1     | 1     | 5     |
| MuDR            |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| 27              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     | 1     |
| 28              | -     | -     | -     | -     | -     | -     | -     | -     | 2     | -     | -     | -     | -     | -     | -     | 1     |
| 29              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     | -     | -     | 1     |
| 30              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 31              | -     | -     | 1     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 32              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 33              | -     | -     | -     | -     | -     | -     | -     | -     | 1     | -     | -     | -     | -     | -     | -     | 1     |
| 34              | -     | -     | -     | -     | -     | 1     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 35              | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     |
| 36              | -     | -     | 1     | -     | -     | -     | -     | -     | -     | -     | -     | -     | 1     | 2     | -     | 1     |
| 37              | -     | -     | -     | -     | -     | -     | -     | -     | 1     | -     | -     | -     | -     | -     | -     | 1     |
| Subtotal        | 0     | 5     | 0     | 1     | 0     | 0     | 2     | 1     | 0     | 4     | 1     | 0     | 0     | 1     | 4     | 2     | 21    |
| FARI            |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| 38              | -     | -     | -     | -     | -     | -     | -     | -     | 1     | -     | -     | -     | -     | -     | -     | 3     |
| 39              | -     | -     | -     | -     | -     | -     | 3     | -     | -     | -     | -     | -     | 1     | -     | 4     | 4     |
| 40              | -     | -     | -     | -     | -     | -     | -     | -     | 1     | -     | -     | -     | -     | -     | -     | 1     |
Table 2 (cont.)

| TE and cluster* | Library / Class I | Total |
|-----------------|------------------|-------|
|                 | BK1  | CL1 | CL2 | FB1 | LV2 | LV3 | RT3 | RT6 | SL1 | SL4 | SL5 | SL7 | SL8 | ST2 | ST6 | WD2 |
| 41              | -    | -   | -   | -   | -   | -   | -   | -   | 1   | 1   | -   | -   | -   | -   | 2   |
| 42              | -    | -   | -   | 1   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   |
| 43              | -    | -   | -   | 1   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   |
| 44              | -    | -   | -   | -   | -   | -   | -   | 1   | -   | -   | -   | -   | -   | -   | 1   |
| 45              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   | -   | -   | 1   |
| 46              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   |
| 47              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   |
| 48              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   |
| 49              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   | -   | -   |
| 50              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   | -   |
| 51              | -    | -   | -   | -   | -   | -   | 1   | -   | -   | -   | -   | -   | -   | -   | 1   |
| 52              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   | -   |
| 53              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   |
| 54              | -    | -   | -   | -   | -   | -   | 1   | -   | -   | -   | -   | -   | -   | -   |
| 55              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   | -   | -   | 1   |
| 56              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   | -   | 1   |
| 57              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   | -   |
| 58              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   | -   | -   |
| 59              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   | -   | -   |
| 60              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 2   | -   |
| 61              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   | -   | -   |
| 62              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 3   | -   |
| 63              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 2   | -   |
| **Subtotal**    |       | 2   | 1   | 0   | 3   | 5   | 0   | 11  | 1   | 2   | 1   | 2   | 2   | 1   | 3   |
| **hAT**         | 64    | -   | 1   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   |
| 65              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   | -   |
| 66              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   |
| 67              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 3   | -   | -   |
| 68              | -    | 4   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| 69              | -    | -   | 1   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| 70              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   | -   | -   | -   |
| 71              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| 72              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| **Subtotal**    |       | 0   | 6   | 0   | 0   | 0   | 0   | 1   | 2   | 0   | 0   | 0   | 0   | 0   | 15  |
| **IS**          | 73    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 3   |
| 74              | -    | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   |
| **Subtotal**    |       | 0   | 0   | 0   | 0   | 0   | 1   | 0   | 2   | 0   | 0   | 0   | 0   | 0   | 0   |
| **CACTA**       | 75    | -   | 1   | -   | -   | -   | -   | -   | -   | -   | -   | 1   | -   | -   | 5   |
| 76              | -    | -   | -   | 1   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 1   |
| **Subtotal**    |       | 0   | 1   | 0   | 1   | 2   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   |
| **Total**       |       | 2   | 13  | 1   | 12  | 12  | 1   | 22  | 3   | 4   | 8   | 5   | 8   | 4   | 9   | 11  |

*Clusters are divided in groups of TE. Cluster numbers are correlated to FORESTS codes as described in http://omega.rc.unesp.br/transposable/tabela.php.
+non-classified LTR
However, to our knowledge, the presence of TE in the IS group has never been reported in plants. Elements of this group are quite common in bacteria, where they act as enhancers (e.g. Schnetz and Rak, 1995). It remains to be seen whether elements in the IS group, which were found in libraries from roots and leaves, have a regulatory function in Eucalyptus.

In order to better compare the relative amounts of distinct groups of TE in Eucalyptus, we have calculated the TE frequency F by the equation $F = (n / N) \times 10^3$, where n is the number of EST in a given group of TE in a given library (values in Table 2) and N is the number of sequenced EST in this library (i.e. each of the 19 values in Table 1). The frequency FG was calculated for groups of TE, according to the equation $FG = \frac{1}{N} \sum F_i$, where $F_i$ represents each of the F values within a given group of TE. FG values indicated FAR1 as the most frequent group, followed by copia, MuDr, hAT and CACTA groups, non-classified LTR and gypsy, LINE and IS groups (Figure 1).

To compare the relative amounts of TE in different libraries, we have calculated FL, using $FL = \frac{1}{N} \sum F_i$, representing each of the F values within a given library. Our results are shown in Figure 2, which also shows the relative contribution of each of the groups of TE to FL values. These values vary from zero to 1.96, with a mean value of 0.92 (calculated for the FL values represented in Figure 2 plus the zero FL value of LV1, SL6 and ST7 libraries). This finding indicates that FORESTS libraries contained, on average, close to 1 transposable element per 1,000 EST, suggesting that a comparable expression rate may occur in Eucalyptus. This average value is very low, considering that TE usually represent 50-90% of plant genomes (Flavel, 1986; SanMiguel et al., 1996). This finding indicates that only a small fraction of TE that can be expressed in Eucalyptus is efficiently transcribed. Inhibition of expression of TE has been called silencing (Okamoto and Hirochika, 2001) and found to be widespread within many organisms such as maize (Fedoroff and Chandler, 1994; Rudenko et al., 2003), Arabidopsis (Hirochika et al., 2000; Steimer et al., 2000) or Drosophila melanogaster (Jensen et al., 1999; Malinsky et al., 2000).

It is unclear whether the calculated frequencies are related to the expression levels of TE; however, the FL values that we found are indicative of a highly variable expression pattern in distinct Eucalyptus tissues or even within single tissues submitted to distinct physiological conditions. For instance, CL1 and CL2 libraries, which were obtained from E. grandis calli, in the dark or in the presence of light, respectively, presented a dramatic difference in FL. CL1 presented a high FL value and contained four different groups of TE (especially MuDR and hAT, yet small amounts of FAR1 and CACTA) and CL2 showed a low FL value and contained TE only in the copia group.

In addition, libraries from seedlings (SL1, SL4, SL5, SL7 and SL8) contained almost all groups of Eucalyptus TE identified in our investigation, except members of LINE and IS groups. This pattern strongly contrasts that found in the BK1 library, which was obtained from eight-year-old trees and contained only FAR1 representatives. Similarly, we found a greater variety and frequency of groups of TE in the RT3 library (which was made from seedlings and contained almost all groups of TE that were identified in Eucalyptus, except CACTA and gypsy) than those found in the RT6 library (which was made from trees and contained only small amounts of copia and MuDr representatives). Finally, libraries from seedling stalk (ST2 and ST6) also had a great variety of groups of TE, although no transposable element was detected in any library from adult tree stalk (ST7). Taken together, these findings suggest that developing plants experience the expression of a much wider spectrum of TE than that seen in adult plants.
Our current mining efforts have identified the clusters and by consequence the FORESTS clones which contain genes of TE that may be differentially expressed in distinct tissues or physiological conditions in *Eucalyptus*. Starting from this preliminary information, further studies on the expression of these genes can be carried out in order to identify the most relevant TE involved in plant development, which is important for wood production on *Eucalyptus* plantations.

Acknowledgements

We would like to thank Centro Virtual Omega / PROPE-UNESP for financing the computers utilized in this annotation work. We would also like to thank the FORESTS Consortium, which provided us with access to its database.

References

Benito MI and Walbot V (1997) Characterization of the maize *Mutator* transposable element MURA transposase as a DNA-binding protein. Mol Cell Biol 17:5161-5175.

Berg DE and Howe MM (1989) Mobile DNA. American Society for Microbiology, Washington, DC, 972 p.

Britten, RJ (1996) DNA sequence insertion and evolutionary variation in gene regulation. Proc Natl Ac Sci USA 93:9374-9377.

Capy P, Vitalis R, Langin T, Higuet D and Bazin C (1996) Relationships between transposable elements based upon the integrase-transposase domains: Is there a common ancestor? J Mol Biol 42:359-368.

Fedoroff N (2000) Transposons and genome evolution in plants. Proc Natl Ac Sci USA 97:7002-7007.

Fedoroff N and Chandler V (1994) Inactivation of maize transposable elements. In: Paszkowisky J (ed) Homologous Recombination and Gene Silencing in Plants. Kluwer Academic Publisher, Dordrecht, pp 349-368.

Flavel RB (1986) Repetitive DNA and chromosome evolution in plants. Philos Trans R Soc Lond B Biol Sci 312:227-242.

Hirochika H, Okamoto H and Kakutani T (2000) Silencing of retrotransposons in *Arabidopsis* and reactivation by the *ddm1* mutation. Plant Cell 12:357-369.

Hudson ME, Lisich DR and Quail PH (2003) The *PHY3* and *FAR1* genes encode transposase-related proteins involved in regulation of gene expression by the phytochrome A-signaling pathway. Plant J 34:453-471.

Jensen S, Gassama MP and Heidmann T (1999) Taming of transposable elements by homology-dependent gene silencing. Nat Gen 21:209-212.

Kidwell MG and Lisch DR (1997) Transposable elements as sources of variation in animals and plants. Proc Natl Ac Sci USA 94:7704-7711.

Lisch D, Girard L, Donlin M and Freeling M (1999) Functional analysis of deletion derivatives of the maize transposon MuDR delineates roles for the MURA and MURB proteins. Genetics 151:331-341.

Malinsky S, Bucheton A and Busseau J (2000) New insights on homology-dependent silencing of I factor activity by transgenes containing ORF1 in *Drosophila melanogaster*. Genetics 156:1147-1155.

McDonald JF (1995) Transposable elements: Possible catalysis of organismic evolution. Trends Ecol Evol 10:123-126.

Okamoto H and Hirochika H (2001) Silencing of transposable elements in plants. Trends Plant Sci 6:527-534.

Rossi M, Araújo PG and Van Sluys M (2001) Survey of transposable elements in sugarcane expressed sequence tags (ESTs). Gen Mol Biol 24:147-154.

Rossi M, Araujo PG, De Jesus EM, Varani, AM and Sluys MA (2004) Comparative analysis of Mutator-like transposases in sugarcane. Mol Genet Genomics 272:194-203.

Rubin E, Lithwick G and Levy A (2001) Structure and evolution of the hAT transposon superfamly. Genetics 158:949-957.

Rudenko GN, Ono A and Walbot V (2003) Initiation of silencing of maize *MuDR/Mu* transposable elements. Plant J 33:1013-1025.

SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z and Bennetzen JL (1996) Nested retrotransposons in the intergenic regions of the maize genome. Science 274:765-768.

Schnetz K and Rak B (1995) ISS1: A mobile enhancer of transcription in *Escherichia coli*. Proc Natl Ac Sci USA 89:1244-1248.

Shirasu K, Schulman AH, Lahaye T and Schulze-Lefert P (2000) A continuous 66 kb barley DNA sequence provides evidence for reversible genome expansion. Genome Res 10:908-915.

Smit AF (1999) Interspersed repeats and other mementos of transposable elements in mammalian genome. Curr Opin Genet Dev 9:657-663.

Steimer A, Amedeo P, Afsar K, Franz P, Scheid OM and Paszkowski J (2000) Endogenous targets of transcriptional gene silencing in *Arabidopsis*. Plant Cell 12:1165-1178.

Wicker T, Guyot R, Yahiaoui N and Keller B (2003) *CACTA* Transposons in Triticeae. A diverse family of high-copy repetitive elements. Plant Physiol 132:52-63.

Yoder JA, Walsh C and Boster TH (1997) Cytosine methylation and the ecology of intragenomic parasites. Trends Genet 13:335-340.

Associate Editor: Marie Anne Van Sluys