

Survey of transposable elements in sugarcane expressed sequence tags (ESTs)

Magdalena Rossi^{1,a}, Paula Gonçalves Araujo^{1,a}, and Marie-Anne Van Sluys^{1,*}

Abstract

The sugarcane expressed sequence tag (SUCEST) project has produced a large number of cDNA sequences from several plant tissues submitted or not to different conditions of stress. In this paper we report the result of a search for transposable elements (TEs) revealing a surprising amount of expressed TE homologues. Of the 260,781 sequences grouped in 81,223 fragment assembly program (Phrap) clusters, a total of 276 clones showed homology to previously reported TEs using a stringent cut-off value of e-50 or better. Homologous clones to *Copia/Ty1* and *Gypsy/Ty3* groups of long terminal repeat (LTR) retrotransposons were found but no non-LTR retroelements were identified. All major transposon families were represented in sugarcane including *Activator (Ac)*, *Mutator (MuDR)*, *Suppressor-mutator (En/Spm)* and *Mariner*. In order to compare the TE diversity in grasses genomes, we carried out a search for TEs described in sugarcane related species *O.sativa*, *Z. mays* and *S. bicolor*. We also present preliminary results showing the potential use of TEs insertion pattern polymorphism as molecular markers for cultivar identification.

INTRODUCTION

TEs are genetic units capable of movement within the genome. As a consequence of this mobility, these elements are mutagenic agents and their activity produces structural changes in single genes or overall genome followed by altered spatial and temporal patterns of gene expression and, ultimately, gene function. They are present in all living organisms and in some large genomes (*e.g.* maize) they represent over 50% of nuclear DNA, providing an enormous source of variability that can be used to create novel genes or modify genetic functions (Bennetzen, 2000). Sugarcane (*Saccharum officinarum*) has an even more complex genome than maize and no information is available concerning the analysis of TEs. These considerations point out the significance of TEs contributions in the sugarcane complex genome understanding.

According to their mode of transposition, TEs are classified into two major classes, RNA mediated transposable elements or retroelements (Class I) and DNA transposable elements or classical transposons (Class II). Retrotransposons transpose through an RNA intermediate molecule, which is reverse transcribed to generate a cDNA copy that will be integrated elsewhere in the genome. They are highly represented in plant genomes, where there are 4 of the 5 recognized classes in eukaryotic cells. These elements, when flanked with long terminal repeats (LTRs), are organized into two subclasses *copia* and *gypsy* differing in the position of the integrase domain within the polyprotein. Retroelements without LTRs are the long interspersed nuclear elements (LINEs) and the defective small interspersed nuclear elements (SINEs). LINEs are autonomous, encod-

ing for nucleocapsid (GAG) and reverse transcriptase (RT) proteins needed for transposition, while SINEs have only few hundred bases and need trans-acting RT polymerase and integrase to transpose. Finally, retroviruses, only found in animals, are derived from *gypsy* retrotransposons which have acquired an envelope protein that allows them to be infective (Xiong *et al.*, 1990).

Transposons move as DNA molecules and their activity was first identified and described in plants in the early work of McClintock (1946). This class comprises families of elements like *Ac* (Doring *et al.*, 1984), *En/Spm* (Gierl and Saedler, 1989) and *MuDR* (Barker *et al.*, 1984) of maize. The main characteristic of these elements is the presence of the transposase (Tpase) gene flanked by terminal inverted repeats (TIRs), this gene encodes the protein responsible for the excision/insertion of the transposon. DNA transposable elements with the same TIR are considered members of the same family (Bennetzen, 2000) and these TIRs may range from 11 bp for *Ac* to few hundred bases for *MuDR*.

There is another group of TE, miniature inverted-repeat transposable elements (MITEs) that, to date, hasn't been included in any of the previous described classes because their transposition mechanism is still unclear. They have short length (125-500 bp) and share features with both transposons and retrotransposons, because they have TIRs (10-15 bp) like transposons and a high copy number like retrotransposons (Zhang *et al.*, 2000). MITEs have been described as non-autonomous elements, and it was hypothesized that their either have an undiscovered autonomous element or that they use host functions to transpose like

¹Departamento de Botanica, IB-USP, Rua do Matão 277, 055080-090 São Paulo, SP, Brazil.

^aBoth these authors contributed equally to this paper.

*Send correspondence to Marie-Anne Van Sluys. E-mail: mavsluys@ib.usp.br.

DNA replication protein complex (J. Casacuberta pers. communication). Although the vast majority of MITEs are defective for transposase, recently, Tourist and Stow-away-like elements have been described containing putative transposases (Le *et al.*, 2000; Turcotte *et al.*, 2001).

Sugarcane is a highly allopolyploid derived from an interspecific hybridization between the wild species *Saccharum spontaneum* (2n = 36-128), contributing with the vegetative vigor and resistance traits, and the cultivated *Saccharum officinarum* (2n = 70-140) responsible for the high sugar content (Ming *et al.*, 1998). Thus, resulting in a large and very complex genome with a chromosome number between 100-130 (D'Hont *et al.*, 1996).

The sugarcane EST sequencing project (SUCEST) produced a large amount of partial cDNA sequences from different tissues submitted or not to different conditions of stress. SUCEST database represents an invaluable source of information for a little known and genetically complex species like sugarcane. After a thorough search of SUCEST database, we identified a variable spectrum of expressed TE. In this paper we present these data and analyze the polymorphic pattern of a particular clone of *Ac*-family homologue.

MATERIAL AND METHODS

SUCEST ESTs database screening

The SUCEST database contains sugarcane cDNA partial sequences of 37 libraries from different tissues, physiological and stressed conditions (Vettore *et al.*, 2001). Sequence comparisons against GenBank were available for keyword searching (Telles *et al.*, 2001). TEs survey was carried out on 260,781 sequences grouped in 81,223 fragment assembly program (Phrap) clusters (Telles and Silva, 2001). The keywords used were “transposable element”, “transposase”, “transposon” and “retrotransposon”. Due to the enormous amount of data and to avoid spurious matches a very stringent expectation cut-off value (e^{-50} or better) was used.

The protein and nucleotide public database of GenBank was also searched with the same keywords in order to detect all mobile genetic elements described in sugarcane related species *Zea mays*, *Oryza sativa* and *Sorghum bicolor*.

SUCEST clones were assigned a family according to the best sequence alignment against a fully characterized element using BLAST program (Altschul *et al.*, 1990).

Plant material

Sugarcane lines SP80-1043, SP89-1117, SP81-1763, SP84-2025, SP82-3349; and 4 individuals of C-905 population (SP80-1043 X SP3349 crossing) were used for polymerase chain reaction (PCR). Purified DNA was kindly

provided by Dr. Jorge A. Gonçalves da Silva (Copersucar Technology Center, Caixa Postal 162, 13400 Piracicaba, SP, Brazil).

PCR analysis

Two primers (52-2: 5'-GTGTCGCTTTAGTGCAA CAATGGA-3' and 52-4: 5'-TCCGGGAAGCGTGTT TGTGAGGG-3') were designed based on the full-length sequence of clone SCCCL4010E02. Each 50 μ L reaction contained 100-200 ng of genomic DNA, 10 pmoles of each primer, 0.2 mM dNTPs, 2.5 mM MgCl₂, 2.5 U Taq DNA polymerase (BRL, Bethesda, Md., USA). Cycling conditions were 94 °C for 5 min, followed by 35 cycles consisting of 94 °C for 45 s, 60 °C for 45 s, and 72 °C for 1 min, with a final extension cycle of 72 °C for 15 min. Aliquots were fractionated in a 0.8% agarose gel.

RESULTS

Expressed TEs in sugarcane

To analyze the number and types of expressed TEs present in SUCEST database, we used keyword search with an expectation cut-off value of e^{-50} or better identifying 276 TE homologous clones representing 117 Phrap clusters. A list of all the identified clusters can be found in SUCEST home page (<http://sucest.lad.dcc.unicamp.br/private/mining-reports/UT/UT-mining.htm>). Since we looked at elements that can present several polymorphic insertions along the genome, and a cluster may represent the grouping of mRNAs coming from different insertions, we decided to consider individual clones instead of clusters for our analyses.

The proportion of transposons (54%) and retrotransposons (46%) homologues was slightly different. Class I TEs were further classified into LTR, *Copia/Ty1* or *Gypsy/Ty3*, and non LTR retroelements. Table I presents the number of clones identified for each category. The amount of *Copia/Ty1* expressed elements was significantly higher than *Gypsy/Ty3* and no non-LTR retrotransposons were found.

Table I - Transposable element content of the SUCEST database.

Class	Type	Number of ESTs
I	LTR ^a	27
	Gypsy/Ty3	19
	Copia/Ty1	82
	Non LTR	-
II	Transposons	148
Total		276

^aUnclassified.

Clones were assigned to a family according to the best alignment against a fully characterized element using BLAST program which enable us to classify the 276 SUCEST clones in 21 different families (Figure 1). It is possible that some of the TEs described in this paper represent a new TE family specific from sugarcane but this will only be confirmed after a detailed analysis of the element complete sequences.

We found sugarcane cDNA clones homologous to various LTR retrotransposons, *Bs1* (Ji and Bennetzen, 1989), *Opie* (San Miguel and Bennetzen, 1998), *Hopschotch* (White *et al.*, 1994), *Sto4* (Marillonnet and Wessler, 1998), *Cinful* (San Miguel and Bennetzen, 1998) and *Centa* (Ananiev *et al.*, 1998), all from *Zea mays*; *MAGGY* (Farman *et al.*, 1996) from *Magnaporthe grisea*; *Retrofit* (Song *et al.*, 1997) from *Oryza longistaminata*; *RICCOPIA* (Voytas *et al.*, 1992) and *RIRE* (Kumekawa *et al.*, 1999) from *Oryza sativa*; *Tnt1* (Grandbastien *et al.*, 1989) from *Nicotiana tabacum*; *Retrosor1* (unpublished, AF098806) from *Sorghum bicolor*; and *deal* (Thomson *et al.*, 1998) from *Ananas comosus*.

Clones homologous to all major plant transposon families were identified, including *MuDr*, *Ac* and *En/Spm* originally described in *Zea mays*; *Soymar* (Jarvik *et al.*, 1998) from *Glicine max*; *Tag* (Liu and Crawford, 1998) from *Arabidopsis thaliana*; *Candystripe* (Chopra *et al.*, 1999) from *Sorghum bicolor*; *Tam* (Nacken *et al.*, 1991) from *Antirrhinum majus* and *Tip100* (Habu *et al.*, 1998) from *Ipomea purpurea*.

As expected, no MITEs were identified since these defective elements are not transcribed.

Comparison with related species

In order to compare TEs found in sugarcane with those reported for related species, a species specific keyword search was made in public GenBank for all mobile elements described for *Zea mays*, *Oryza sativa* and *Sorghum bicolor* obtaining a total of 5348, 18 and 543 entries respectively. All the entries were analyzed to eliminate redundancy and spurious matches and were subsequently organized according to the classes they belong. Table II was generated from this study plus sugarcane data, and represents all known TEs described, to date, for these 4 grass species. Of the total of 71 different TEs 60.6% belong to Class I, 23.9% to Class II and 15.5% representing MITEs.

Since plant TEs were discovered in maize, this species became the model for plant TEs study and consequently, the vast majority of mobile genetic elements were originally described in *Z. mays*. This fact is clearly reflected in Table II where 60.4% of retrotransposons, 64.7% of transposons and 63.6% of MITEs shown were first described in maize. Only 12 and 4 TEs were first characterized in rice and sorghum respectively while all the rice transposons were found by homology to previously described elements. Although retroelements are more represented, only *MuDR* and *Ac* families are present in all four species investigated.

Polymorphic insertion patterns of an *Ac* homologue Tase

To evaluate the polymorphism level between different sugarcane cultivars generated by any of the TE homologues described, we chose an *Ac*-like clone and a PCR based strategy. Clone SCCCCL4010E02 sequenced from a callus-derived library is similar to *Ac* Tase at e^{-114} value. Full length sequence of this clone resulted in a 1081 bp poliA^+ fragment homologous to the central part of the 3.5 kb *Ac* element mRNA (Kunze *et al.*, 1987). This result indicated that SCCCCL4010E02 may be either a partial cDNA clone where the 5' end is missing or an mRNA from a truncated element unit or both. A pair of primers with an expected PCR product of 517 bp for sugarcane *Ac* homologous clone was designed. No introns were expected within this PCR fragment according to the original *Ac* element sequence. Genomic PCRs were carried out using total genomic DNA isolated from different sugarcane cultivars. Each cultivar has its own fingerprint with some shared fragments. The 517 bp expected band is present in all cultivars but intensity varies. Presence-absence type of polymorphic bands (1.6 and 1.1 kb bands) indicate the presence of other insertions of this *Ac* homologous element in the genome with variable sizes of the amplified fragment (probable due to insertion within the 517 bp PCR product) or concatamers of the same element. Differences in intensity (0.9 and 0.5 kb bands) indicate the presence of variable numbers of copies of that particular locus in the different cultivars.

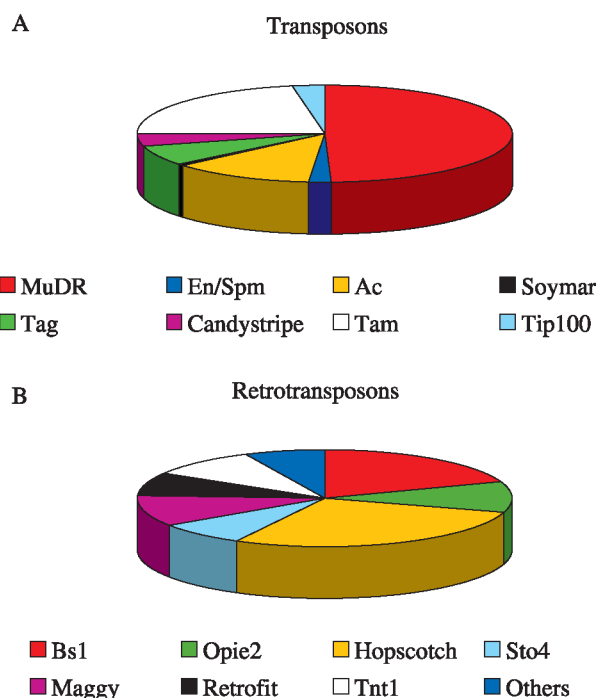


Figure 1 - Relative abundance of TE families in sugarcane. A: Transposons. B: Retrotransposons.

Table II - Transposable element spectrum in sugarcane and the related species rice, maize and sorghum.

Class	Element	Type	Source species	Also present in	Reference
	RICCOPIA	Ty1/copia	<i>Oryza.sativa</i>	Sugarcane	M94492
	Opie1-4	Ty1/copia	<i>Zea mays</i>	Rice, Sugarcane	T04111
	Fourf	Ty1/Copia	<i>Z.mays</i>	Sorghum	U68401
	Victim	Ty1/copia	<i>Z.mays</i>	Sorghum	U68410
	Hopscotch	Ty1/copia	<i>Z.mays</i>	Sugarcane	U12626
	Ji1-6	Ty1/copia	<i>Z.mays</i>	-	U68405
	PREM-2	Ty1/copia	<i>Z.mays</i>	Sorghum	T02724
	Tnt1	Ty1/copia	<i>Nicotiana tabacum</i>	Rice, Sugarcane	X13777
	Osr1	Ty1/copia	<i>O.sativa</i>	-	BAB03249
	Rrt	Ty1/copia	<i>O.sativa</i>	-	Z75540
	Stonor (Sto)	Ty1/copia	<i>Z.mays</i>	Rice, Sugarcane	T17429
	Tos1-20	Ty1/copia	<i>O.sativa</i>	-	D85879
	Retrofit	Ty1/copia	<i>Oryza longistaminata</i>	Sugarcane	AAB82754
	Romani	Ty3/gypsy	<i>Z.mays</i>	Rice	AJ002612
	Huck1-2	Ty3/gypsy	<i>Z.mays</i>	-	U68404
	Reina	Ty3/gypsy	<i>Z.mays</i>	-	U69258
	Cinful-1-2	Ty3/gypsy	<i>Z.mays</i>	Rice, Sugarcane	AF049110
	Cin1 (Cinful family)	Ty3/gypsy	<i>Z.mays</i>	-	T03247
	Zeon1 (Cinful family)	Ty3/gypsy	<i>Z.mays</i>	-	S58484
	Magellan	Ty3/gypsy	<i>Z.mays</i>	-	AF015269
	Grande1	Ty3/gypsy	<i>Zea diploperenis</i>	Rice	U68403
	MARE1	Ty3/gypsy	<i>Z.mays</i>	-	AB014749
	Tekay	Ty3/gypsy	<i>Z.mays</i>	-	AF050455
	Retrosor1	Ty3/gypsy	<i>Sorghum.bicolor</i>	Rice, Sugarcane	AF098806
	RIRE	Ty3/gypsy	<i>O.sativa</i>	Sugarcane	AB014755
	Retrosat	Ty3gypsy	<i>O.sativa</i>	-	AF111709
	Rle	Ty3/gypsy	<i>Z.mays</i>	-	AF057037
	MAGGY	Ty3gypsy	<i>Magnaporthe grisea</i>	Rice, Sugarcane	L35053
I (Retrotransposons)	dea1 (del family)	Ty3/gypsy	<i>Ananas comosus</i>	Sugarcane	Y12432
	del1-46	Ty3/gypsy	<i>Lilium henryi</i>	Rice, Sorghum	1510387A
	Levithan	LTR ^a	<i>S.bicolor</i>	Maize	U07816
	Bs1	LTR ^a	<i>Z.mays</i>	Sugarcane	X16080
	G	LTR ^a	<i>Z.mays</i>	-	Varagona <i>et al.</i> , 1992
	B5	LTR ^a	<i>Z.mays</i>	-	S46940
	CentA	LTR ^a	<i>Z. mays</i>	Sugarcane	AF078917
	P-SINE	SINE	<i>O.sativa</i>	-	D85067
	Colonist1-2	LINE	<i>Z.mays</i>	-	U90128
	Cin-4	LINE	<i>Z.mays</i>	-	Y00086
	RILN	LINE	<i>O.sativa</i>	-	AB016109
	MALN	LINE	<i>Z.mays</i>	-	AB016115
	pHind12-22	Unknown	<i>S.bicolor</i>	-	AF078902
	Take1-2	Unknown	<i>Z.mays</i>	-	U68406
	Milt	Unknown	<i>Z.mays</i>	Sorghum	U68407

Table II - (cont.)

Class	Element	Type	Source species	Also present in	Reference	
II (Transposons)	En/Spm	Transposon	<i>Z.mays</i>	Rice, Sugarcane	M25427	
	Tam (En/Spm family)	Transposon	<i>Anthirinus majus</i>	Rice, Sugarcane	X57297	
	Ac/Ds	Transposon	<i>Z.mays</i>	Rice, Sorghum, Sugarcane	X62655/X01380	
	Tip100 (Ac family)	Transposon	<i>Ipomoea purpurea</i>	Sugarcane	AB004906	
	MuDr	Transposon	<i>Z.mays</i>	Rice, Sorghum, Sugarcane	M76978	
	ILS1	Transposon	<i>Z.mays</i>	Sorghum	U07956	
	Ins2	Transposon	<i>Z.mays</i>	-	U32636	
	Bg	Transposon	<i>Z.mays</i>	-	X56877	
	Uq	Transposon	<i>Z.mays</i>	-	X59774	
	Soymar (Mariner family)	Transposon	<i>Glicine max</i>	Rice, Sugarcane	AF078934	
	Tnr	Transposon	<i>O.sativa</i>	-	D63711	
	Tz86	Transposon	<i>Z.mays</i>	-	AH00136	
	Jittery (Mu family)	Transposon	<i>Z.mays</i>	-	AF247646	
	Sleepy	Transposon	<i>Z. mays</i>	-	U28041	
	Candystripe	Transposon	<i>S. bicolor</i>	Sugarcane	AF206660	
	Shooter	Transposon	<i>Z.mays</i>	-	AAD55677	
	Tag	Transposon	<i>A.thaliana</i>	Sugarcane	L12220	
	MITEs ^b	Tourist	MITE	<i>Z.mays</i>	Rice, Sorghum	S48688
		Truncator	MITE	<i>O. longistaminata</i>	Rice	U72724
		L-K	MITE	<i>O.sativa</i>	-	AF172282
Wanderer		MITE	<i>O.sativa</i>	-	U34601	
Explorer		MITE	<i>O.sativa</i>	-	U70541	
Stowaway		MITE	<i>Z.mays</i>	-	U32636	
Heartbreaker		MITE	<i>Z.mays</i>	-	AF205927	
Ditto		MITE	<i>Z.mays</i>	-	AC079830	
Castaway		MITE	<i>Z.mays</i>	-	AC079830	
Gaifin		MITE	<i>Z.mays</i>	-	AC079830	
Crackle	MITE	<i>Z.mays</i>	-	AC079830		

^aUnclassified long terminal repeat retrotransposon.

These results strongly suggest the activity of an *Ac*-like transposable element during these cultivars establishment process (Figure 2A).

To confirm the specificity of the designed primers, some of the amplified bands were cloned, sequenced and analyzed against GenBank. Results showed that those bands were homologous to *Ac* element demonstrating that PCR patterns are indeed different insertions of *Ac* homologues (data not shown).

Figure 2B shows the amplification pattern for 4 individuals of SP80-1043 X SP82-3349 segregating progeny. There are no new polymorphic bands with respect to parental lines, meaning that no major rearrangements associated to the element occurred during the cross. On the other hand,

contributions of both parents to the *Ac*-like element genetic composition in the progeny can be observed as no single individual has exactly the same pattern as either of the parental lines.

DISCUSSION

Sugarcane revealed a surprising amount and diverse spectrum of expressed transposable elements. Considering that TEs are the most abundant non-genic DNA in plants, providing a rich source for genome evolution, and their special abundance in grass genomes (Bennetzen, 2000; Zhang and Peterson, 1999), it is absolutely necessary for sugarcane genomics understanding to evaluate transposable elements contribution.

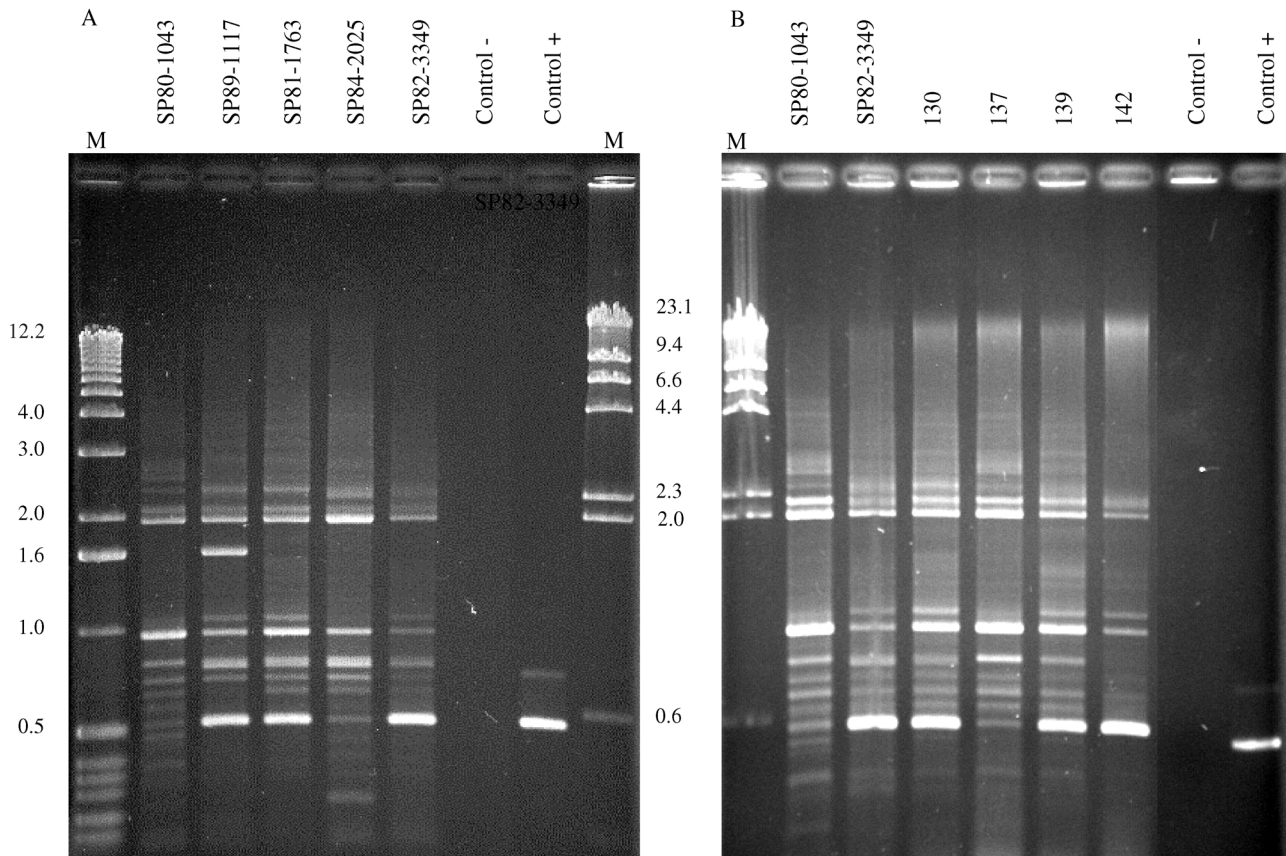


Figure 2 - *Ac* transposon homologue polymorphism. **A:** Genomic PCR pattern of sugarcane varieties SP80-1043, SP89-1117, SP81-1763, SP84-2025 and SP82-3349. **B:** Genomic PCR pattern of four individuals from SP80-1043 X SP82-3349 cross (130, 137, 139 and 142). M: molecular weight marker (molecular weights are in kb). The positive control (+) was done using DNA of SCCCL4010E02 clone as template while no template was used for the negative control (-).

Ubiquitous and abundant, retroelements constitute more than 50% of the maize genome and they could be equally or even more numerous in other plant species with large complex genomes like sugarcane (Bennetzen, 2000; Mao *et al.*, 2000). In contrast, retrotransposon expression has been demonstrated for only very few higher plants, probably due to the fine transcriptional control which makes difficult to find the temporal and spatial conditions for expression to occur (Takeda *et al.*, 1999; Grandbastien *et al.*, 1998), denoting an intelligent strategy to perpetuate them selves without compromising host genome. Based on this considerations, our results agree with the expectation about finding more expressed transposons (54%) than retrotransposon (46%).

How does sugarcane support such a large number of expressed TEs? Plants adapt to genome perturbation like changes in ploidy, over 50% of all plant species are polyploid or have undergone periods of polyploidy during their evolutionary history. Silenced transposable elements in a diploid may be activated in the new genetic environment of the polyploid since genetic redundancy in the polyploid may buffer the potential deleterious effects of transposition (Voytas and Naylor, 1998). On the other hand, plant ge-

nome expansion is an inevitable consequence of TEs insertion, the accumulation of retrotransposons blocks between genes is a major factor in the size difference between the maize genome and those of its smaller relatives (Federoff, 2000).

Genomic studies in rice by Turcotte *et al.* (2001) and Mao *et al.* (2000) report that *Gypsy/Ty3* retrotransposons are more abundant than *Copia/Ty1* elements. In contrast, our results show that, in sugarcane, the number of expressed *Copia/Ty1* elements was significantly higher than *Gypsy/Ty3*. We also found no non-LTR retrotransposons in our search of SUCEST database, which indicates that these TEs are poorly expressed in sugarcane, moreover, they showed to be few represented in grass genomes as presented in Table II and previously described in rice (Mao *et al.*, 2000).

In this paper we have reported 21 different expressing TEs in sugarcane. This finding strongly suggests that these elements are transpositionally active but it is important to point out that defective TE can also be expressed as seems to be the case for the *Ac* homologue analyzed here. If we consider that, prior to the SUCEST project, sugarcane was a largely unexplored species and our research concerned

transcriptionally active elements only, the diversity of sugarcane TEs presented in this paper is comparable to maize, where to date 39 different elements have been described (Table II).

Our novel approach enabled us to have a genomic overview of potentially active transposable elements in contrast to the classic strategies where TEs were first described at DNA level, and eventually, the transcription and transposition were studied.

In sugarcane, somaclonal variation is a common phenomena caused either by tissue culture procedures or vegetative propagation (Irvine *et al.*, 1991). In view of the fact that our results present a wide spectrum of potentially active previously unknown TEs in sugarcane, we can hypothesize that some of the somaclonal variation events reported in the literature can be a result of TEs activity.

Recently, reports were published evaluating the utility of TEs as source to develop molecular markers showing that TEs are highly polymorphic in terms of insertion sites and their internal structure (Ellis *et al.*, 1998). Furthermore, TE-derived markers efficiently distinguished rice cultivars (Fukuchi *et al.*, 1993). Our preliminary analyses with the *Ac* homologue revealed element transposition during cultivar establishment and showed that TE-mediated fingerprinting could become a powerful tool for sugarcane.

RESUMO

O projeto de ESTs (Expressed sequence tag) da cana-de-açúcar (SUCEST) produziu um grande número de seqüências de cDNA de diferentes tecidos submetidos ou não a diferentes condições de estresse. Neste trabalho nós analisamos os resultados de uma busca, por Elementos de Transposição (TEs), que apresentou um grande número de seqüências homólogas a TEs. Das 260.781 seqüências agrupadas em 81.223 *clusters* pelo programa Phrap (fragment assembly program), um total de 276 seqüências apresentaram homologia com TEs previamente descritos utilizando-se uma nota de corte estridente de 50 ou melhor. Clones homólogos aos grupos de retrotransposons com LTR (long terminal repeat) copia/Ty1 e Gypsy/ty3 foram achados porém nenhum retroelemento sem LTR foi identificado. A maioria das famílias de transposons estão representadas em cana-de-açúcar incluindo Activator (Ac), Mutator (MuDR), Supressor-mutator (En/Spm) e Mariner. Para podermos comparar a diversidade dos TEs nos genomas das gramíneas, nós buscamos os TEs em espécies relacionadas a cana-de-açúcar tais como O sativa, *Z. mays* e *S. bicolor*. Nós também apresentamos os resultados preliminares do potencial uso de TEs como marcadores moleculares na identificação de cultivares.

ACKNOWLEDGMENTS

This work undertaken in this project was supported by the Brazilian agency FAPESP. M. Rossi and P.G.

Araujo were recipients of FAPESP postdoctoral fellowships.

REFERENCES

- Ananiev, E.V., Phillips, R.L. and Rines, H.W.** (1998). Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions. *Proc Natl Acad Sci. U.S.A.* 27: 13073-13078.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J.** (1990). Basic alignment search tool. *J. Mol. Biol.* 215 (3): 403-410.
- Barker, R.F., Thompson, D.V., Talbot, D.R., Swanson, J. and Bennetzen, J.L.** (1984). Nucleotide sequence of the maize transposable element *Mu. Nucleic Acids Res* 12: 5955-5967.
- Bennetzen, J.L.** (2000). Transposable element contributions to plant genome evolution. *Plant. Mol. Biol* 42: 251-269.
- Chopra, S., Brendel, V., Zhang, J., Axtell, J.D. and Peterson, T.** (1999). Molecular characterization of a mutable pigmentation phenotype and isolation of the first active transposable element from *Sorghum bicolor*. *Proc. Natl. Acad. Sci. U.S.A.* 96: 15330-15335.
- D'Hont, A., Grivet, L., Feldmann, P., Rao, P.S., Berding, N. and Glaszmann, J.C.** (1996). Characterization of the double genomic structure of modern sugarcane cultivars, *Saccharum* spp., by molecular cytogenetics. *Mol. Gen. Genet.* 250: 405-413.
- Doring, H.P., Tillmann, E and Starlinger, P.** (1984). DNA sequence of the maize transposable element *Dissociation. Nature* 307: 127-130.
- Ellis, T.H.N., Poyser, S.J., Knox, M.R., Vershinin, A.V. and Ambrose, M.J.** (1998). Polymorphism of insertion sites of *Ty1-copia* class retrotransposons and its use for linkage and diversity analysis in pea. *Mol. Gen. Genet.* 260: 9-19.
- Farman, M.L., Tosa, Y., Nitta, N. and Leong, S.A.** (1996). MAGGY, a retrotransposon in the genome of the rice blast fungus *Magnaporthe grisea*. *Mol Gen Genet.* 251: 665-674.
- Fedoroff, N.** (2000). Transposons and genome evolution in plants. *Proc. Natl. Acad. Sci. U.S.A.* 97: 7002-7007.
- Fukuchi, A., Kikuchi, F., and Hirochika, H.** (1993). DNA fingerprinting of cultivated rice with rice retrotransposon probes. *Jpn. J. Genet.* 68: 195-204.
- Gierl, A., and Saedler, H.** (1989). The *En/Spm* transposable element of *Zea mays*. *Plant Mol Biol.* 13: 261-266.
- Grandbastien, M.A.** (1998). Activation of plant retrotransposons under stress conditions. *Trends in Plant Science* 3: 181-187.
- Grandbastien, M.A., Spielmann, A., and Caboche, M.** (1989). Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. *Nature* 337: 376-380.
- Irvine, J.E., Benda, G.T.A., Legendre, B.L., and Machado, G.R.** (1991). The frequency of marker changes in sugarcane regenerated from callus culture. II Evidence for vegetative and genetic transmission, epigenetic effects and chimeral disruption. *Plant Cell, Tissue and Organ Culture* 26: 115-125.
- Habu, Y., Hisatomi, Y. and Iida, S.** (1998). Molecular characterization of the mutable flaked allele for flower variegation in the common morning glory. *Plant J.* 16 (3): 371-376.
- Jarvik, T., and Lark, K.G.** (1998). Characterization of *Soymar1*, a mariner element in soybean. *Genetics* 149: 1569-1574.

- Jin, Y.K., and Bennetzen, J.L.** (1989). Structure and coding properties of *Bs1*, a maize retrovirus-like transposon. *Proc. Natl. Acad. Sci. U.S.A.* 86: 6235-6239.
- Kumekawa, N., Ohtsubo, H., Horiuchi, T., and Ohtsubo, E.** (1999). Identification and characterization of novel retrotransposons of the *gypsy* type in rice. *Mol. Gen. Genet.* 260: 593-602.
- Kunze, R., Stochaj, U., Laufs, J. and Stalinger, P.** (1987). Transcription of transposable element *Activator (Ac)* of *Z. mays*. *EMBO J.* 6: 1555-1563.
- Le, Q.H., Wright, S., Yu, Z. and Bureau, T.** (2000). Transposon diversity in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 97: 7376-7381.
- Liu, D., and Crawford, N.M.** (1998). Characterization of the putative transposase mRNA of *Tag1*, which is ubiquitously expressed in *Arabidopsis* and can be induced by *Agrobacterium*-mediated transformation with *dTag1* DNA. *Genetics* 149: 693-701.
- Mao, L., Wood, T.C., Yu, Y., Budiman, M.A., Tomkins, J., Woo, S., Sasinowski, M., Presting, G., Frisch, D., Goff, S., Dean, R.A. and Wing, R.A.** (2000). Rice transposable elements: A survey of 73,000 Sequence-Tagged-Connectors. *Genome Research* 10: 982-990.
- Marillonnet, S. and Wessler, S.R.** (1998). Extreme structural heterogeneity among the members of a maize retrotransposon family. *Genetics* 150: 1245-1256.
- McClintock, B.** (1946). Maize genetics. *Year Book Carnegie Inst. Washington* 45: 176-186.
- Ming, R., Liu, S.C., Lin, Y.R., da Silva, J., Wilson, W., Braga, D., van Deynze, A., Wenslaff, T.F., Wu, K.K., Moore, P.H., Burnquist, W., Sorrells, M.E., Irvine, J.E. and Paterson, A.H.** (1998). Detailed alignment of *Saccharum* and *Sorghum* chromosomes: comparative organization of closely related diploid and polyploid genomes. *Genetics* 150: 1663-1682.
- Nacken, W.K., Piotrowiak, R., Saedler, H. and Sommer, H.** (1991). The transposable element *Tam1* from *Antirrhinum majus* shows structural homology to the maize transposon *En/Spm* and has no sequence specificity of insertion. *Mol. Gen. Genet.* 228 (1-2): 201-208.
- San Miguel, P.J. and Bennetzen, J.L.** (1998). Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. *Annals of Botany* 82 (A): 37-44.
- Song, W.Y., Pi, L.Y., Wang, G.L., Gardner, J., Holsten, T. and Ronald, P.C.** (1997). Evolution of the rice *Xa21* disease resistance gene family. *Plant Cell* 9: 1279-1287.
- Takeda, S., Sugimoto, K., Otsuki, H. and Hirochika, H.** (1999). A 13 bp cis-regulatory element in the LTR promoter of the tobacco retrotransposon *Tto1* is involved in responsiveness to tissue culture, wounding, methyl jasmonate and fungal elicitors. *Plant Journal* 18: 383-393.
- Telles, G.P., Braga, M.D.V., Dias, Z., Lin, T.L., Quitzau, J.A.A., da Silva, F.R. and Meidanis, J.** (2001). Bioinformatics of the sugarcane EST project. *Genetics and Molecular Biology* 24 (1-4): 9-15.
- Telles, G.P. and da Silva, F.R.** (2001). Trimming and clustering sugarcane ESTs. *Genetics and Molecular Biology* 24 (1-4): 17-23.
- Thomson, K.G., Thomas, J.E. and Dietzgen, R.G.** (1998). Retrotransposon-like sequences integrated into the genome of pineapple, *Ananas comosus*. *Plant Mol Biol.* 38: 461-465.
- Turcotte, K., Srinivasan, S. and Bureau, T.** (2001). Survey of transposable elements from rice genomic sequences. *Plant Journal* 25: 169-179.
- Varagona, M.J., Purugganan, M. and Wessler, S.R.** (1992). Alternative splicing induced by insertion of retrotransposons into maize *waxy* gene. *Plant Cell* 4: 811-820.
- Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.** (2001). The libraries that made SUCEST. *Genetics and Molecular Biology* 24 (1-4): 1-7.
- Voytas, D.F. and Naylor, G.J.P.** (1998). Rapid flux in plant genomes. *Nature Genetics* 20: 6-7.
- Voytas, D.F., Cummings, M.P., Konieczny, A., Ausubel, F.M. and Rodermel, S.R.** (1992). *Copia*-like retrotransposons are ubiquitous among plants. *Proc. Natl. Acad. Sci. U.S.A.* 89: 7124-7128.
- 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 gene structure and expression. *Proc Natl Acad Sci U.S.A.* 91: 11792-11796.
- Xiong, Y. and Eickbush, T.H.** 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* 9: 3353-3362.
- Zhang, Q., Arbuckle, J. and Wessler, S.R.** (2000). Recent, extensive, and preferential insertion of members of the miniature inverted-repeat transposable element family heartbreaker into genic regions of maize. *Proc. Natl. Acad. Sci. U.S.A.* 97: 1160-1165.
- Zhang, J. and Peterson, T.** (1999). Genome rearrangements by nonlinear transposons in maize. *Genetics* 153: 1403-1410.