DOI: 10.1007/s10535-015-0569-z BIOLOGIA PLANTARUM 60 (1): 55-67, 2016

Repetitive sequences in *Eragrostis curvula* cDNA EST libraries obtained from genotypes with different ploidy

J. ROMERO[1], J.P. SELVA[1,2], S. PESSINO[3], V. ECHENIQUE[1,4], and I. GARBUS[1,5]*

Centro de Recursos Naturales Renovables de la Zona Semiárida, CONICET, 8000, Bahía Blanca, Argentina[1]
Departamento de Biología Bioquímica y Farmacia, Universidad Nacional del Sur, 8000, Bahía Blanca, Argentina[2]
Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, S2125ZAA, Zavalla, Argentina[3]
Departamento de Agronomía, Universidad Nacional del Sur, 8000, Bahía Blanca, Argentina[4]
Departamento de Ciencias de la Salud, Universidad Nacional del Sur, 8000, Bahía Blanca, Argentina[5]

Abstract

Eragrostis curvula (Schrad) Nees (weeping lovegrass) represents important cultivated forage in semiarid regions, and the most useful cultivars are tetraploid and reproduce by pseudogamous diplosporous apomixis. We previously produced a series of genetically related *E. curvula* lines that provide a suitable system for the identification of gene(s) involved in diplosporous apomixis and ploidy, including a natural apomictic tetraploid (T), a diploid sexual line (D), and a tetraploid sexual plant (C). A collection of expressed sequence tags (ESTs) was generated from cDNA libraries obtained from panicles of the D, T, and C, and leaves of the T. The present study aimed to analyze the repetitive content of these four cDNA libraries and further identify and characterize transposable element (TE)-related ESTs. Repetitive sequences were identified through the interface *RepeatMasker* (*RM*) using the database *Repbase Update* and further classification of TEs was performed manually from the *RM* output. The different contribution of ESTs with identity to TEs among libraries was further evaluated, and such differences were validated through RT-qPCR. We found that the percentage of repetitive content in the leaf cDNA library was almost double than in inflorescence libraries, with retrotransposons contributing mostly in all libraries. The expression of TE-related ESTs was compared in cDNA samples extracted from D, T, and C leaves or inflorescences revealing that seven mRNAs containing *MuDR*-like DNA transposons, *Gypsy*-like, and *Copia*-like retrotransposons were differentially represented according to tissue, reproductive mode, or ploidy. The euploid series of *Eragrostis curvula* is a useful model to the study of epigenomic changes produced after changes in ploidy. The present work constitutes the first detailed report on repetitive sequences of *Eragrostis curvula* at the transcriptome level.

Additional key words: diplosporous apomixis, transposable elements, weeping lovegrass.

Introduction

The term C-value paradox was coined by Thomas (1971) to describe lack of correlation between the structural complexity of an organism and its genome size attributed to differences in the content of repetitive sequences among species. Several examples of such variation exist in the plant kingdom with repetitive sequences accounting for 10 % (Initiative 2000), 35 % (Project IRGSP 2005), 40 % (Zhou and Xu 2009), 61 % (Paterson *et al.* 2009), and 80 % (Chandler and Brendel 2002, Messing *et al.* 2004, Rabinowicz and Bennetzen 2006) of the genomes of *Arabidopsis thaliana*, *Oryza sativa*, *Populus trichocarpa*, *Sorghum bicolor*, and *Zea mays*, respectively, and showing an apparent correlation with their genome sizes of 135, 382, 403, 697, and 2 500 Mb,

Submitted 24 July 2014, *last revision* 15 July 2015, *accepted* 2 September 2015.
Abbreviations: C - tetraploid sexual plant; D - diploid sexual line; EST - expressed sequence tag; LTR - long terminal repeat; RM - repeat masker; SINE - short interspersed nuclear elements; T - natural apomictic tetraploid plant; TE - transposable element.
Acknowledgements: This work was funded by the CONICET (PIP 112-200801-01517 and PIP-112-201101-00767 to V.E), SECyT UNS (PGI 24/A168 to V.E.) and the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT, PICT 2013-2029 to I.G.).
* Corresponding author; fax: (+54) 0291 4862882, e-mail: igarbus@criba.edu.ar

respectively. Thus, the number and complexity of repetitive elements varies between species, and those with larger genomes generally have more repetitive elements (Kidwell 2002). Repetitive sequences comprise two main classes, transposable elements (TEs) and tandem repeats. Based on a transposition mechanism, TEs can be further subdivided into Class I retrotransposons that move *via* so-called "copy and paste" mechanisms using RNA intermediates, and Class II DNA transposons that move *via* "cut and paste" through DNA intermediates (Wicker *et al.* 2007). On the other hand, tandem repeats encompass any sequence found in consecutive copies along a DNA strand, organized in tandem arrangements of the monomeric unit, and typically are located in centromeres, telomeres, and heterochromatic regions of many eukaryotes (for review see Plohl 2008).

Due to their mobility through a process known as transposition, TEs have a potential to alter a genome by bringing about growth in genome size at the same time as they can induce mutations, disrupt genes, mediate chromosomal rearrangement, and transport genes to new genomic locations when a captured gene and the TE are copied together to a new location (reviewed in Lisch 2013). In order to prevent their potentially damaging actions, cells have developed mechanisms for TE control, including silencing by DNA methylation, RNA interference, or histone modifications (for review see Blumenstiel 2011, Saze and Kakutani 2011, Lisch 2013). However, under certain circumstances, some TEs escape this cellular control and transcribe and sometimes transpose either under normal conditions or in response to biotic or abiotic stresses (Pouteau *et al.* 1991, Hirochika 1993, Mhiri *et al.* 1997, Muthukumar and Bennetzen 2004, Gómez *et al.* 2006, Ramallo *et al.* 2008, Ueki and Nishii 2008, Cheng *et al.* 2009, Picault *et al.* 2009). Although data from EST libraries in grasses show that most TEs are poorly transcribed, EST-based analyses have demonstrated the presence of TE transcripts in several organs and cell types (Vicient *et al.* 2001, Echenique *et al.* 2002, Vicient and Schulman 2002, Kashkush *et al.* 2003, Vicient 2010).

Polyploidization is considered to be major factor in the evolutionary history of angiosperms and one of the most important contributor mechanisms to speciation and adaptation in the plant kingdom (Wendel 2000). For instance, *Lonicera japonica* tetraploids show to be more resistant to heat stress than diploids (Li *et al.* 2011). Polyploidization may involve major genetic and epigenetic alterations of progenitor genomes (for review see Ma and Gustafson 2005). Newly synthesized polyploids provide useful systems to reveal such alterations occurring immediately following a modification at the ploidy level. Gene expression repatterning associated with genetic and epigenetic modifications was reported for allopolyploid wheat (Kashkush *et al.* 2002, 2003), *Arabidopsis* (Comai *et al.* 2000, Wang *et al.* 2004), autopolyploid *Paspalum* (Martelotto *et al.* 2005), and cotton (Adams *et al.* 2003).

Eragrostis curvula (Schrad.) Nees (weeping lovegrass) is member of the *Poaceae* family, subfamily *Chloridoideae* (Watson and Dallwitz 1992), native to southern Africa and cultivated in semiarid regions. The most useful cultivars for forage are tetraploids (2n = 4x = 40) that reproduce by pseudogamous diplosporous apomixis (Crane 2001, Voigt *et al.* 2004). In this asexual reproduction through seeds (Asker and Jerling 1992), the megasporocyte undergoes two rounds of mitotic division to form a non-reduced tetranucleate embryo sac with an egg, two synergids, and one polar nucleus (Meier *et al.* 2011). Besides *Eragrostis*, diplosporous apomixis has been described in *Agropyrum*, *Boechera*, *Paspalum*, *Poa*, *Tripsacum* among other species (Nogler 1984, Barcaccia and Albertini 2013). Our laboratory previously produced a near-isogenic series of genetically related *E. curvula* lines consisting of: *1*) a natural apomictic tetraploid (cv. Tanganyika, the T; 2n = 4x = 40), *2*) a diploid sexual plant (cv. Victoria, the D; 2n = 2x = 20) obtained from the tissue culture of the T, and *3*) a tetraploid sexual plant (cv. Bahiense, the C; 2n = 4x = 40) derived from the D by colchicine treatment (Cardone *et al.* 2006). These plants provide a suitable system for the identification of gene(s) involved in diplosporous apomixis, ploidy level-regulated expression, and forage quality candidate variation using a transcriptomic approach, as well as for the development of potential markers for genetic mapping. A collection of ESTs was generated from cDNA libraries prepared from panicles of near-isogenic lines with different ploidy and reproductive modes and from leaves of 12-d-old line T plants (Cervigni *et al.* 2008a). The present study analyzes the repetitive content of these four cDNA libraries, identifies and characterizes the TE sequences represented in the ESTs, and further investigates their possible expression. The euploid series of *E. curvula* constitutes an excellent model to study epigenomic changes produced immediately after changes in ploidy. This is the first analysis of the *E. curvula* repetitive sequences at the transcriptome level.

Materials and methods

Libraries of ESTs: Four cDNA libraries described in Cervigni *et al.* (2008a) were analyzed in the present study. Three of them, Ec01, Ec02, and Ec04, were constructed from panicles in the premeiotic developmental stage from the sexual diploid (D, 2n = 2x = 20), tetraploid apomictic Tanganyika (T, 2n = 4x = 40) and tetraploid sexual plant obtained through colchicine treatment (C, 2n = 4x = 40), respectively. The fourth

library, Ec03, was constructed from leaves of 12-d-old Tanganyika plants. The EST sequences are available in GenBank (EH183417 to EH195711).

Identification of repetitive elements: The program *RepeatMasker* (*RM*, http://www.repeatmasker.org) was utilized for identification of repetitive sequences among *E. curvula* cDNA libraries deposited at the National Center for Biotechnology information (*NCBI*), with similarity searches at the Genetic Information Research Institute (*GIRI*) repeat element database without filtering. *Fast Alignment* (*FASTA*)-formatted EST accessions from the four libraries (Ec01: EH183417 - EH187066; Ec02: EH187067 - EH190710; Ec03: EH190711 - EH192319; and Ec04: EH192320 - EH195711) were used as program input. *RepeatMasker* returned three output files (masked query sequence(s), an annotated list of the masked sequences, and a summary of the repeat content of the analyzed sequences) that were used to generate the statistics using Aho, Weinberger, and Kernighan (awk) scripts (Table 1 Suppl.). Sequence comparisons of signatures for each family of TEs were performed using the alignment program *Crossmatch* (*open version 3.3.0*) and the Smith-Waterman-Gotoh algorithm in conjunction with the script *MaskerAid* (Bedell *et al.* 2000). A repeat database used was *Repbase Update* (*RU*) *v. 18.07*; (*GIRI*, http://www.girinst.org/repbase/index.html) containing prototypical sequences representing repetitive DNA from different species of the *Poaceae* family. Most prototypical sequences in *RU* are consensus sequences of large families and subfamilies of repeats. Smaller families are represented by sequence examples. *Repbase Update* includes many families of repeats unreported elsewhere. Classification of TEs was performed manually from the *RM* output. The *BLASTN* algorithm (*v. 2.2.25*) was used to search for the identity of the sequences from the four libraries with retroelements deposited at *GIRI*, utilized as local database, following the hierarchy suggested by Wicker *et al.* (2007): class, subclass, superfamily. In order to classify recognizable repeat families, subclasses were further used to divide DNA transposons into terminal inverted repeats and non-terminal inverted repeats , whereas retrotransposons were classified as long terminal repeats (LTR) and non-long terminal repeats (non LTR). The rationale used to homology analysis was based on the fact that the EST sequences were compared to the available databases, where no *Eragrostis* sequences are included. Thus, we assume that when the ESTs showed homologies with TEs and not with other sequences, even if Wicker's criterion was not fulfilled in most cases, they can be considered valid matches. Following this reasoning, the sequences that aligned with repetitive elements over at least 40 bp with an e-value $\leq e^{-3}$ were included in this analysis. The criterion was then restricted in order to select a subset of ESTs that could result more reliable to a deeper analysis. To make statistical comparisons among the contribution of each element to each library, 95 % confidence intervals for each percentage were calculated through the method of Newcombe (1998; http://vassarstats.net/prop1.html).

Different contribution of ESTs with identity to TEs among libraries: The evaluation of the different contribution of ESTs with identity to TEs among libraries was performed following a more stringent criterion in order to filter redundant data and maximize the quality of the results. Thus, the analysis was carried out including the ESTs that simultaneously met the criteria of having an alignment $\geq$ 80 bp with the proposed TE and an e-value $\leq e^{-10}$. The initial comparison of the differentially expressed TE sequences was performed manually. The existence of protein-coding-like sequences within the TE-related ESTs was explored through a *BLASTX* analysis performed against the non-redundant protein database using the *BLAST2GO* platform (Conesa *et al.* 2005).

Validation assays using PCR: Total RNA was isolated from inflorescences and leaves of plants C, D, and T grown under the same experimental conditions using a spin or vacuum total RNA isolation kit (*Promega*, Madison, USA) following the manufacturer's instructions. An RNA concentration in each sample was determined by measuring absorbance at 260 nm. For all samples, reverse transcription was performed with 0.5 µg of the total RNA using an *ImProm-II*™ reverse transcription system (*Promega*) with the Oligo (dT)15 primer according to the manufacturer's instructions. Specific primers were designed through the integrated DNA technology (*IDT*) webpage tools (http://www.idtdna.com/Scitools/Applications/RealTimePCR/) and synthesized by *IDT* after virtually checking their specificity.

Amplification of the cDNA was performed using PCR and specific primer pairs matching specific regions of the corresponding TE-related EST sequence that occurred in a single library (Table 1 Suppl.). The primers were first demonstrated to amplify products from DNA extracted from the plants D, T, and C (data not shown). The PCR program consisted of a denaturation step at 94 °C for 2 min followed by 40 cycles at 94 °C for 15 s, at 55 °C for 20 s, and at 72 °C for 30 s, followed by an extra extension step at 72 °C for 5 min. The expression of the *ubiquitin conjugating enzyme* gene, reported to be optimal reference gene in the apomictic species *Brachiaria brizantha* (Duarte Silveira *et al.* 2009), was used as internal control. Primers EC_UBICEqPCR_F (AAGGAGCTCAAGGACCTGCAGAAA) and EC_UBICE qPCR_R (TCACTAAGAACACACCAC CGGCAT) were previously designed based on the EH186329.1 sequence from *E. curvula*.

Real time PCR reactions were prepared in a final reaction volume of 20 mm^3 including 50 pmol of forward

and reverse primers, 5 mm^3 of cDNA diluted 100-fold, and 10 mm^3 of *Real Mix* (*Biodynamics*, Buenos Aires, Argentina). Amplification was carried out in a *Rotor Gene 6000* (*Corbett Research*, Sydney, Australia). Cycling consisted of 94 °C for 2 min followed by 35 cycles at 94 °C for 15 s, at 55 °C for 20 s, and at 72 °C for 30 s. Finally, a melting curve was constructed (5 s for each step ramping from 73 °C to 95 °C, increasing the temperature by 0.2 °C after step 2). The $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen 2001) was used to normalize and calibrate transcript values relative to the *ubiquitin conjugating enzyme* expression. Amplification efficiency was equivalent for the samples and corresponding internal control. Two negative controls, one without reverse transcriptase and the non-template control, were included in the analyses. The reactions were conducted in three technical replicates of at least two biological samples of RNA isolated from the D, T, and C plants. Differences between mean values were evaluated by unpaired Student's *t*-test. Values $P < 0.05$ were considered significant. The cycle threshold (CT) indicates the fractional cycle at which the amplified target reaches its threshold. The CT was determined from the exponential phase of the PCR by the iQ5 real time detection system software (*Bio-Rad*, Hercules, CA, USA). The ΔCT value for a sample was calculated by subtracting the CT of each gene from that of the internal reference gene. The ΔΔCT of a gene was calculated by subtracting the ΔCT of each sample from the ΔCT of the control genotype (Tanganyika).

Results

The program *RM* was used for the identification of repetitive sequences among *E. curvula* cDNA libraries, and they were deposited at the *NCBI* (EH183417 to EH195711). The libraries were generated without any normalization and/or enrichment steps in order to evaluate a different expression among lines with different ploidy levels and reproductive modes (Cervigni *et al.* 2008a). Masked and unmasked regions were computed by *RM* without filtering using *FASTA* formatted sequences from the libraries Ec01 - Ec04 as inputs. The percentages of masked bases over the total length were ~10 % for the three inflorescence libraries (Ec01, Ec02, and Ec04) and ~18 % for the leaf library (Ec03) (Table 1). In the four analyzed cDNA libraries, the most frequently found repetitions were retroelements followed by small RNAs mainly composed by rRNAs (Table 1). However, the contribution of such repetitive elements was greater in Ec03 than in the inflorescence libraries, thus accounting for the differences in the total repetitive content among the libraries constructed from different tissues (Table 1). Concerning the difference in composition of a small RNA fraction between libraries even when no validation analysis was conducted, three rRNAs were differentially found in the inflorescence libraries, whereas one was suggestive of being preferentially found in leaves. Simple repeats, low complexity, and unclassified sequences represented small proportions of the masked sequence in the libraries, whereas satellites were identified only in libraries Ec01 and Ec03 (Table 1).

The *BLASTN* algorithm was used to search for the identity of the sequences from the four libraries to retroelements and DNA elements from the *GIRI* database of repeat elements (an e-value $\leq e^{-3}$). Classification of recognizable repeat families was performed following the recommendations reported in Wicker *et al.* (2007). Between 1.5 and 2 % of the ESTs of each library showed a significant sequence similarity to TE families included in the *GIRI* database (Table 2). The number of sequences with similarity to RNA transposons in the four EST libraries was 4 to 6 times greater than that of sequence similarity to DNA transposons (Table 2). In the

Table 1. Repetitive elements identified in *Eragrostis curvula* cDNA libraries. The element class is indicated in the first column. Information about each library (Ec01, Ec02, Ec03, and Ec04) is provided in three columns, indicating the number of elements (#), the length of the sequence occupied by these elements (bp), and the percentage of the library that transposable element represents [%]. The total length of each library (bp) is indicated under the library name.

Repetitive element type	Ec01 (3082956 bp)			Ec02 (3256100 bp)			Ec03 (961037 bp)			Ec04 (2810498 bp)		
	#	[bp]	[%]	#	[bp]	[%]	#	[bp]	[%]	#	[bp]	[%]
Retroelements	578	114850	3.73	633	119483	3.67	329	71422	7.43	450	70667	2.51
DNA transposons	39	4272	0.14	75	7696	0.24	26	3850	0.40	50	5166	0.18
Unclassified	9	1329	0.04	10	1609	0.05	30	7103	0.74	3	350	0.01
Small RNA	161	72920	2.37	152	86039	2.64	159	75676	7.87	160	75049	2.80
Satellites	4	464	0.02	0	0	0	1	98	0.01	0	0	0
Simple repeats	1023	44242	1.44	1083	48641	1.49	216	7734	0.80	1569	78331	2.79
Low complexity	340	42194	1.37	328	63855	1.96	95	4627	0.48	236	51093	1.82
Total	2154	280271	9.10	2281	327323	10.05	856	170510	17.73	2468	280656	10.11

Table 2. The distribution of transposable element (TE) superfamilies identified in *Eragrostis curvula* cDNA expressed sequence tag (EST) libraries. For each library (Ec01, Ec02, Ec03, and Ec04), the number of TE-related ESTs for each family (n), the percentage of the library that it represents [%], and the lower and upper limits of the 95 % confidence intervals (C.I.) are shown.

TE class	Family		Ec01			Ec02			Ec03			Ec04		
			n	[%]	C.I.	n	[%]	C.I.	n	[%]	C.I.	n	[%]	C.I.
DNA transpo-sons	non TIR	*Helitron*	5	0.14	0.06-0.32	3	0.08	0.03-0.24	0	0.00	0.00-0.24	2	0.06	0.02-0.22
	TIR	*Cacta*	1	0.03	0.01-0.16	2	0.05	0.01-0.19	1	0.06	0.01-0.35	3	0.09	0.03-0.26
		Harbinger	2	0.06	0.01-0.19	2	0.05	0.01-0.19	0	0.00	0.00-0.24	0	0.00	0.00-0.11
		Hat	0	0.00	0.00-0.11	1	0.03	0.01-0.16	1	0.06	0.01-0.35	1	0.03	0.01-0.17
		MuDR	4	0.11	0.04-0.28	1	0.03	0.01-0.16	2	0.12	0.03-0.45	3	0.09	0.03-0.26
	total		12	0.33	0.19-0.58	9	0.25	0.13-0.47	4	0.25	0.10-0.64	9	0.27	0.14-0.51
Retro-transpo-sons	non LTR	*LINE*	0	0.00	0.00-0.11	0	0.00	0.00-0.11	1	0.06	0.01-0.35	0	0.00	0.00-0.11
		SINE	0	0.00	0.00-0.11	3	0.08	0.03-0.24	4	0.25	0.10-0.64	0	0.00	0.00-0.11
	LTR	*Copia*	14	0.38	0.23-0.64	15	0.41	0.41-0.68	16	0.99	0.61-1.60	7	0.21	0.10-0.43
		Gypsy	34	0.93	0.67-1.30	26	0.71	0.48-1.04	5	0.31	0.13-0.72	53	1.56	1.19-2.04
	Total		48	1.32	1.00-1.74	44	1.21	0.90-1.62	26	1.62	1.11-2.36	60	1.77	1.38-2.27
Total TEs			60	1.64	1.28-2.11	53	1.45	1.11-1.89	30	1.87	1.31-2.64	69	2.03	1.59-2.58

Table 3. The gene ontology analysis of transposable element (TE)-related expressed sequence tags (ESTs). The TEs identified in the ESTs, the number of the ESTs that contain each TE, and the *BLAST2GO* annotation of the ESTs are shown in columns 1 - 3. Columns 4 - 7 indicate TE presence by library (Ec01 - Ec04). The accession numbers of the ESTs that contain each TE are provided in Table 2 Suppl.

Identified TE	#ESTs	*BLAST2GO* annotation	Ec01	Ec02	Ec03	Ec04
CASTAWAY_*Harbinger_Oryza*	2	N.A.	0	2	0	0
*Copia*10-ZM_I_*Copia_Zea*	2	retrotransposon unclassified	2	0	0	0
Copia-18_BD-I_*Copia_Brachypodium*	25	cell wall-associated hydrolase	3	9	12	0
		transcription factor btf3	1	0	0	0
Copia-28-ZM_I_*Copia_Zea*	2	retrotransposon ty1-*Copia* subclass	0	0	0	2
Copia-33_BD-I_*Copia_Brachypodium*	3	hypothetical protein	0	0	0	3
Copia-3-ZM_I_*Copia_Zea*	1	retrotransposon unclassified	0	0	1	0
CRMA1_LTR_*Gypsy_Oryza*	1	retrotransposon ty3-*Gypsy* sub-class	0	1	0	0
DNA-8-1N_SBi_DNA	2	N.A.	0	0	2	0
*ENSPM*1_ZM_*EnSpm_Zea*	1	af464738_3 transposase	1	0	0	0
EnSpm-3_TA_*EnSpm_Triticum*	1	rna polymerase beta subunit	0	0	1	0
EnSpm-37_SBi_*EnSpm_Sorghum*	2	TPA: hypothetical protein *ZEA*MMB73_586644	0	0	0	2
F524_SINE_*Oryza*	2	N.A.	0	1	0	0
		N.A.	0	0	1	0
Gypsy-12_SB-I_*Gypsy_Sorghum*	1	retrotransposon ty3-*Gypsy* sub-class	0	0	1	0
Gypsy-124N_SBi-I_*Gypsy_Sorghum*	1	dna-directed rna polymerase e subunit 1-like	1	0	0	0
Gypsy-28_SB-LTR_*Gypsy_Sorghum*	2	retrotransposon ty3-*Gypsy* sub-class	0	2	0	0
Gypsy-36-ZM_I_*Gypsy_Zea*	1	wd repeat-containing protein 20-like	0	1	0	0
Gypsy-9_BD-I_*Gypsy_Brachypodium*	1	Os09g0491900	0	1	0	0
HARB-7_SBi_*Harbinger_Sorghum*	1	transposon en spm sub-class	1	0	0	0
Helitron-3_ZM_*Helitron_Zea*	1	uncharacterized loc101210396	0	0	0	1
Helitron-4_ZM_*Helitron_Zea*	3	mutant low phytic acid protein 1	3	0	0	0
Helitron-N3_ZM_*Helitron_Zea*	2	rrna 2-o-methyltransferase fibrillarin 1-like	0	1	0	0
		mediator of rna polymerase ii transcription subunit 36a-like	0	1	0	0
MuDR-12_SBi_*MuDR_Sorghum*	1	rna polymerase beta subunit	0	1	0	0
MuDR-13_SBi_*MuDR_Sorghum*	1	N.A.	0	0	1	0
MuDR-9_ZM_*MuDR_Zea*	4	prp18 domain containing protein	4	0	0	0
MUTRIM1_*MuDR_Triticum*	1	transposon expressed	0	0	1	0
SINE2-1_TAe_SINE2/tRNA_*Triticum*	1	N.A.	0	1	0	0
Total	65		16	21	20	8

inflorescence libraries Ec01 and Ec04, the number of TE-related ESTs for *Gypsy* LTR retrotransposons evaluated through a 95 % confidence interval superposition was higher than the sum of the other retrotransposon classes, whereas in Ec02 such a comparison did not show significant differences. On the contrary, the *Copia* LTR retrotransposons were the most abundant retroelements (Table 2) in leaf library Ec03.

The existence of protein-coding-like sequences within the TE-related ESTs analysis was restricted to the subset of ESTs with a *BLASTN* e-value < e^{-10} and alignment over 80 bp. Such criteria resulted in 65 TE-related ESTs which represented 0.53 % of the ESTs from the four libraries. A total of 57 (87.7 %) ESTs showed a significant similarity to at least one expressed protein (Table 3 and Table 2 Suppl.). Among those, 18.5 % coded for transposon-related proteins, 58.4 % for annotated proteins not related with transposons, and 10.8 % for hypothetical proteins. Notably, ~29 % of sequences in the "non TE-related" group coded for proteins involved in RNA metabolism (Table 3).

Seven superfamilies were represented among the TE-related ESTs including *Harbinger*, *MuDR*, *Helitron*, and *Cacta/EnSpm* type transposons, and LTR-*Copia*, LTR-*Gypsy*, and short interspersed nuclear elements (*SINE*) type retrotransposons. In most cases, *BLASTX* analysis of the TE-related ESTs identified proteins and/or protein functions associated with these superfamilies (Table 3). Exceptions to this rule were the Copia-18 family- and Helitron-N3-containing ESTs. Almost all of the Copia-18-containing ESTs showed identity to a cell wall-associated hydrolase, whereas one of them showed identity to transcription factor btf3. Of the Helitron-N3-containing ESTs, one showed identity to rRNA 2-*o*-methyltransferase fibrillarin 1-like, whereas another was similar to a mediator of RNA polymerase II transcription subunit 36a-like.

Since the *E. curvula* ESTs were obtained from libraries constructed with RNA extracted from genotypes differing in reproductive mode (sexual *vs*. apomictic), ploidy (diploid *vs*. tetraploid), and/or tissue (leaf *vs.* inflorescence), subsequent analysis focused on a different representation of sequences from different TE families using a "virtual northern" strategy based on the presence of the ESTs in each library.

The analysis of the occurrence of individual TE sequences among the libraries showed that 39 ESTs (60 %) were present in a single library (Table 3). Thus,

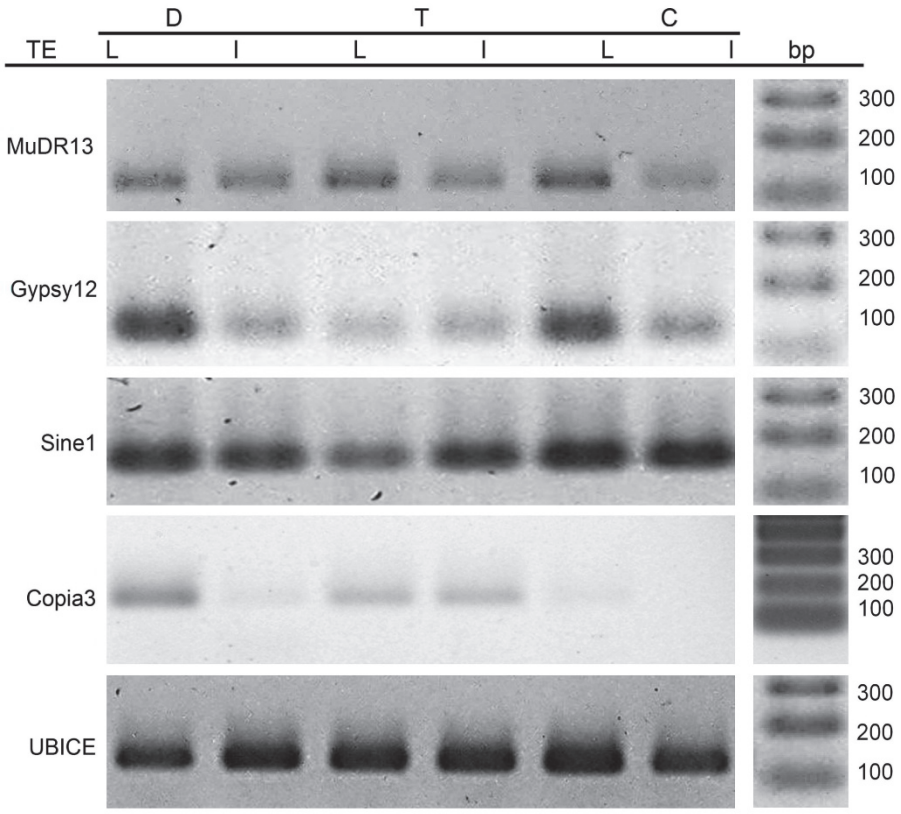

Fig. 1. The RT-PCR analysis of expression of candidate transposable elements (TEs) in different tissues. The primer pair used, designed based on a TE similar to expressed sequence tags, is shown *on the left* of the agarose gel fragments, whereas molecular mass markers are shown *on the right*. All of the analyzed TEs were identified in the leaf library (Ec03). The RT-PCR was performed on the cDNA from leaves (L) or inflorescences (I) of diploid (D), tetraploid (T), and induced tetraploid (C) genotypes. *Ubiquitin* (UBICE) was used as housekeeping gene.

even assuming that the major limitation of our experimental system was a low genome coverage of the four libraries (Cervigni *et al.* 2008a), it was hypothesized that the different representation of the TE sequences in the libraries could reflect in some cases that those elements were differentially represented among the tissues used to prepare the libraries.

We first experimentally checked by RT-PCR the differential expression of 13 TE-related ESTs **(**Figs. 1,3) using RNA samples extracted from leaves and inflorescence from the D, T, and C. Extra RT-qPCR analysis were performed (Figs. 2 and 4) to clarify the expression profiles of some TE-related ESTs. Primers based on the TE-related ESTs that were identified in a single library were designed (Table 1 Suppl.), except for the primers based on ESTs that showed identity with F524_*SINE*_Oryza sativa. Two primer pairs were designed for such TE-related ESTs since the ESTs EH190997 from Ec03 and EH188637 from Ec02 (Table 1 Suppl.) showed no similarity between them suggesting that they may correspond to different mRNAs carrying the same TE, or alternatively, to different regions of the expressed gene.

The tissue specificity of the mRNA expression was tested using the primer pairs specific for the ESTs found in the leaf library (Ec03) that showed identity to the

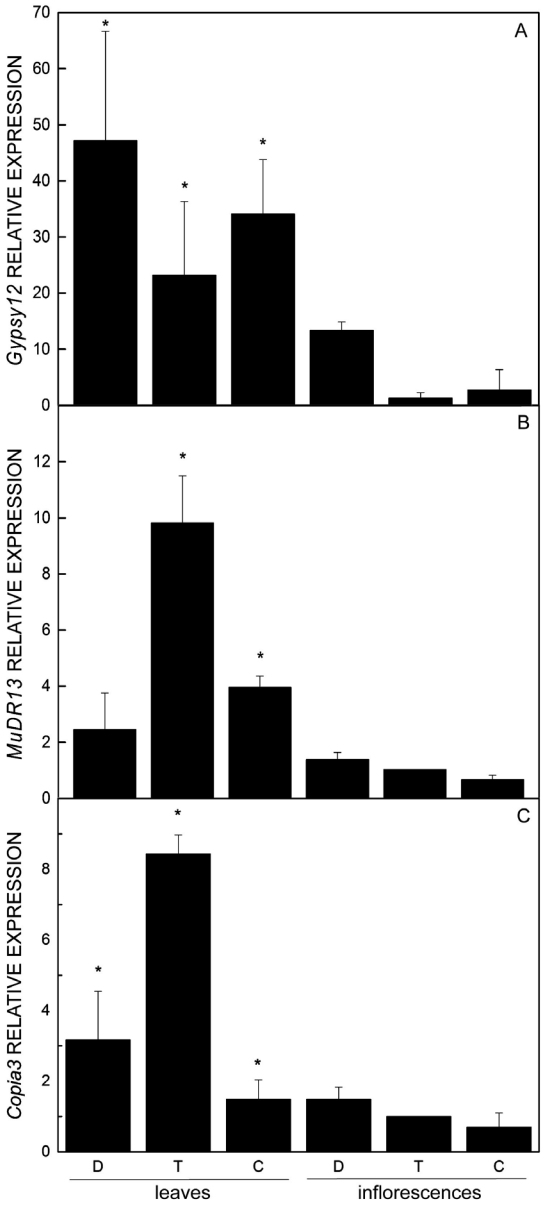

Fig. 2. The differential expression of transposable element-containing mRNAs. The results of RT-qPCR with specific primers *Gypsy-12*, *MuDR-13*, and *Copia-3* performed on the cDNA from leaves (L) or inflorescences (I) of diploid (D), tetraploid (T), and induced tetraploid (C) genotypes. * - significant differences between L and I for each genotype ($P < 0.05$).

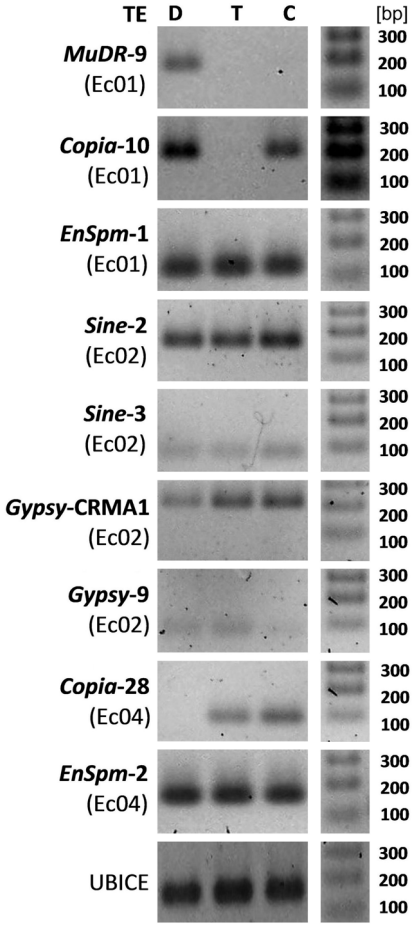

Fig. 3. The RT-PCR analysis of the expression of candidate transposable elements. The primer pair used, designed based on a TE similar to expressed sequence tags, is shown *on the left* of the agarose gel fragments with the library in which it was identified *in silico* shown *in brackets*. Molecular mass markers are shown *on the right*. The RT-PCR was performed on the cDNA from inflorescences of diploid (D), tetraploid (T), and induced tetraploid (C) genotypes. *Ubiquitin* (UBICE) was used as housekeeping gene.

TEs F524_*SINE* (*O. sativa*), *MuDR*-13 (*S. bicolor*), *Gypsy*-12 (*S. bicolor*), and *Copia*-3 (*Z. mays*) (Fig. 1). For the three genotypes, the expression of mRNA with identity to *Gypsy*-12, *MuDR*-13, and *Copia-3* was statistically greater in leaves than in inflorescences (Fig. 2).

Then, to evaluate different expression according to the ploidy level, primer pairs based on the ESTs from the library Ec01 that contain sequences similar to TEs *MuDR*-9 (*Z. mays*), *Copia*-10 (*Z. mays*), and *En/Spm*-1 (*Z. mays*) were used (Fig. 3). The mRNA containing a sequence similar to *MuDR*-9 was differentially expressed in the diploid genotype, whereas a sequence similar to *Copia*-10, predicted *in silico* to be diploid-specific, showed specificity to the sexual genotypes (D and C; Fig. 3). A sequence similar to *EnSpm*-1 was highly expressed in the three genotypes (Fig. 3).

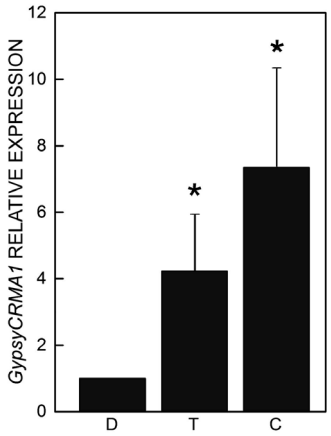

Fig. 4. The differential expression of *Gypsy*CRMA1-like containing mRNAs. The results of RT-qPCR using primers with specificity to *GypsyCRMA1* performed on the cDNA obtained from diploid (D), tetraploid (T), and induced tetraploid (C) genotype inflorescences. * - significant differences between the T and C as compared to the D ($P < 0.01$).

The use of primers based on TEs SINE2-1 (*T. aestivum*), *Gypsy* CRMA1 (*O. sativa*), and *Gypsy*-9 (*B. distachyon*), identified in the Ec02 library and thus being specific candidates for the apomictic genotype T, revealed no specific expression through RT-PCR analysis (Fig. 3), whereas the use of qPCR allowed the confirmation of different expression of mRNA containing a sequence similar to *Gypsy*CRMA1 in the tetraploid genotypes (Fig. 4).

We also used primers based on TEs *Copia*-28 (*Z. mays*) and *EnSpm*-37 (B. *distachyon*) identified in the Ec04 library showing a *Copia*-28-like sequence differentially represented in the tetraploid genotypes, whereas a high expression was found for a sequence similar to *EnSpm*-1 in the three genotypes (Fig. 3).

Inconsistencies observed in the TE-containing ESTs between the *in silico* estimates and experimental analysis through RT-PCR and qRT-PCR can be attributed to both a low TE-related EST number considered and a difference in precision of both RT-PCR methods. Except for *MuDR*-9, *Copia*-10, and *Copia*-28 that gave clear RT-PCR profiles, RT-qPCR was needed to decide about the different expression of the *Gypsy*-12, *MuDr*-13, *Copia*-3, and *GypsyCRMA1* TEs.

Overall, seven mRNAs corresponding to TE-related ESTs were found to be differentially expressed in samples originated from different tissues, ploidy, or reproductive mode in the *E. curvula* lines. The *BLASTX* sequence analysis established that five of these ESTs coded for TEs belonging only to the retrotransposon class I. The two other ESTs harbored sequence identity with *MuDR* sequences and contained segments with identity to a pre-mRNA-splicing factor, and the other showed identity to a putative ACT-domain containing protein kinase family protein (Table 4).

Discussion

In the present work, the repeatome of *E. curvula* cDNA leaf and inflorescence libraries was characterized. The four pooled EST libraries included in this study represent 22 % of the *E. curvula* genome, whereas the genome coverage of each library represents 5 - 10 % of the genome with no differences among libraries (Cervigni *et al.* 2008a). Moreover, similar redundancies, *i.e.*, an average EST count per gene, were estimated for the four libraries (from 1.19 to 1.34) indicating that a similar number of new genes would be captured from each library if new sequencing runs were conducted (Cervigni *et al.* 2008a).

Thus, the percentage of repetitive sequences in the leaf library (Ec03) was almost double regarding those of the three inflorescence libraries (Ec01, Ec02, and Ec04), which can be attributed to a greater representation of retroelements and small RNA transcripts in leaves (Table 1). In agreement with this, a higher expression of TEs in leaves compared to other tissues has been reported in wheat ESTs (Echenique *et al*. 2002). However, a lower TE expression in leaves compared to reproductive tissues has been reported in maize (Vicient 2010).

Identification and quantification of retroelements in transcriptomes could be taken as potential indicator of their transcriptional activity. A more detailed insight was provided though *BLAST* analysis of the repeatome against the *GIRI* database, thus enabling identification and classification of TEs. This strategy has an additional advantage of allowing a manual setting of parameters including the identity and alignment length. In the four libraries, the RNA retrotransposons were four to six times more expressed than the DNA transposons. It is

interesting to note that the *Gypsy*-like retrotransposons were the most frequently found superfamily in the ESTs from the inflorescence libraries (Ec01, Ec02, and Ec04), whereas in the leaf library, the most represented were the *Copia*-like elements constituting the first report of a higher expression of *Copia*-like TEs in somatic tissues compared to *Gypsy*-like elements. The further analysis of the *Gypsy*-like elements shows that this was not result of a high abundance of one or a few families, but nearly of 25 *Gypsy* families in the EST databases. Generally, most *Gypsy* families have a higher transcription compared with other elements (Vicient *et al.* 2010). In agreement with this finding, we found that a *Gypsy*12-related transcript was more expressed in leaves than in inflorescences.

Table 4. The detailed analysis of transposable element (TE)-related expressed sequence tags (ESTs) found to be differentially expressed. The TE-related ESTs and sequence lengths are shown in the 1st and 2nd columns. When more than one EST showed identity to the TE, a contig was assembled with the sequences. For each EST or contig, the associated TE, the *BLASTX* analysis, the number of identities, and the alignment range are indicated in the following columns. Q refers to the EST/contig probed; S refers to the sequence in the database with identity to the probe.

Sequence	Length [bp]	Associated TE	*BLASTX*	Identity	Positives	Alignment ranges	E-value
EH191456	760	*Gypsy-12*	AAP53838: retrotransposon protein, putative, Ty3-gypsy subclass (*Oryza sativa* japonica group); length: 1470 aa	142/165 (86 %)	151/165 (92 %)	Q: 535 - 41 S: 539 - 703	$6e^{-91}$
				24/41 (59 %)	30/41 (73 %)	Q: 638 - 516 S: 505 -545	$2e^{-27}$
EH192111	208	*MuDR-13*	EMS64397: serine/threonine-protein kinase HT1 [*Triticum urartu*]; length: 560 aa	27/32 (84 %)	30/32 (93 %)	Q: 85 – 180 S: 192 - 223	$7e^{-09}$
EH190917	819	*Copia-3*	ABA96411: retrotransposon protein, putative [*Oryza sativa* japonica group]; length: 256 aa	131/174 (75 %)	146/174 (83 %)	Q: 125 - 643 S: 82 - 225	$4e^{-86}$
Contig_3 (EH186963, EH183711, EH184869, EH185446)	1125	*MuDR-9*	pre-mRNA-splicing factor 18-like [*Setaria italica*] (XP_004958093); length: 441 aa	255/287 (89 %)	265/287 (92 %)	Q: 104 - 961 S: 155 - 441	$9e^{-150}$
EH189614	1285	*Gypsy-CRMA1*	AAX96752: retrotransposon protein, putative, Ty3-gypsy subclass (*Oryza sativa* japonica group); length: 1222 aa	42/73 (58 %)	50/73 (68 %)	Q: 366 -584 S: 1119 - 1109	$3e^{-16}$
				16/33 (48 %)	20/33 (61 %)	Q: 583 - 681 S: 1190 -1222	$3e^{-16}$
Contig_2 (EH194637, EH192502)	590	*Copia-28*	ABG00000: retrotransposon protein, putative, Ty1-copia subclass [*Oryza sativa* japonica group]; length: 2074 aa	90/132 (68 %)	100/132 (75 %)	Q: 5 - 397 S: 1379 - 1510	$3e^{-54}$
Contig_1 (EH186646, EH183457)	712	*Copia-10*	BF95666: retrotransposon protein, putative, Ty1-copia subclass, expressed [*Oryza sativa* japonica group]; length: 976 aa	48/63 (76 %)	53/63 (85 %)	Q: 358 - 546- S: 795 - 857	$2e^{-27}$
				20/24 (83 %)	21/24 (87 %)	Q: 548 - 619 S: 858 - 881	$2e^{-27}$
				10/21 (48 %)	15/21 (71 %)	Q: 615 - 677 S: 880 - 881	$3e^{-27}$

Although no estimates for repetitive sequence content in the *E. curvula* genome have been reported, the percentage of expressed sequences is expected to be only a small portion of the genome content, mainly due to efficient mechanisms that impede the expression of mobile elements (Blumenstiel 2011, Saze and Kakutani 2011, Lisch 2013). Here, it was found that 1.5 to 2 % of the ESTs of each library showed a significant sequence similarity to TE families in the *GIRI* database (Table 2). Similar values were obtained for maize and flax (Vicient 2010, González and Deyholos 2012), species with a different repetitive content, thus suggesting that TE transcriptional activity is maintained at low levels independently of the repetitive content in the species. The retrotransposon-related sequences content for libraries Ec02 and Ec04 has been previously estimated to be 0.7 and 1.5 % (an e-value $\leq e^{-8}$), respectively (Zappacosta *et al*. 2014). Here, as expected, a less stringent condition (Table 2, an e-value $\leq e^{-3}$) leads to a higher number of retrotransposon-related sequences, resulting to be 1.21 and 1.77 %, respectively.

Gene ontology analysis through the *BLAST2GO* interface of the 65 EST sequences selected according to more stringent parameters reveals that 18.5 % of the

sequences matched TEs. Combined *BLAST* analysis against non redundant and *GIRI* databases reveals that 12 ESTs were exclusively composed of TE sequences, whereas 38 were chimeric, composed of TE-gene sequences, and remaining 15 sequences were not available or were hypothetical proteins. Interestingly, 24 copies out of the 38 TE-gene chimeras found correspond to chimeras displaying sequence homologies with both the TE *Copia*-18_BD-I_*Copia_Brachypodium* and the gene of cell wall-associated hydrolase sequences. Biological bases for those increased activities of the *Copia*-18_BD-I_*Copia_Brachypodium* and the gene for a cell wall-associated hydrolase remains unknown. A future work will be conducted to analyze the expression of such transcripts in inflorescences and leaves of *E. curvul*a in order to investigate a possible regulatory link between the expression of the element and the homologues to the captured sequence. The other *Copia*- containing ESTs were not chimeric. *Copia*-like elements have been identified as having an insertion preference near genes in maize, whereas *Gypsy*-like elements have been observed to preferentially insert into other repetitive elements (Bennetzen 1996). Moreover, such an overrepresentation of *Copia*-like LTR retrotransposons in chimeric gene-TE ESTs has been previously reported in *Arabidopsis* concomitantly with a significant bias against *Gypsy*-like LTR retrotransposons (Lockton and Gaut 2009). On the other hand, 6 out of the 65 ESTs contained the full TE sequence inserted. All the *MuDR*-like containing ESTs were chimeric, whereas no gene-*SINE* chimeras were detected. Contributions of TEs to individual genes through the creation of introns, exons, or chimeric genes are not rare events, and a substantial proportion of genes in angiosperms harbor TEs (Oliver *et al.* 2013). It has been suggested that TE insertion events led to the formation of TE-gene chimeras (Lockton and Gaut 2009). Near one sixth of rice genes are associated with retrotransposons, with insertions either in the gene itself or within putative promoter regions, supporting the idea that TEs partake in the regulation of host genes (Krom *et al.* 2008).

Different EST based analyses demonstrate the presence of TE transcripts in several organs and cell types (Vicient *et al.* 2001, Vicient and Schulman 2002, Kashkush *et al.* 2003, De Araujo *et al.* 2005). Here, we established that three sequences, carrying *Gypsy-12*-like, *Copia-3*-like, and *MuDr-13*-like TE elements, were more highly expressed in leaves than in inflorescences of the three analyzed genotypes. In accordance with this, the percentage of ESTs similar to retrotransposons is higher in cDNA libraries from wheat leaf tissues than in those from other wheat tissues (Echenique *et al.* 2002).

It is generally accepted that apomictic individuals show a variable trait expressivity (Rebozzio *et al.* 2011). There is growing evidence for a relationship between apomixis and epigenetic mechanisms including different expression of transposable elements (Ochogavía *et al.* 2011, Okada *et al.* 2011). Differential display analysis of expression performed on the *E. curvula* 'back-and-forth' plant series revealed that one sequence associated with the sexual genotype and three with the apomictic one showed a significant similarity to TEs (Selva *et al.* 2012). Here, a *Copia*-10_like element was expressed in the inflorescences of the sexual genotypes but not in the apomictic ones, thus suggesting that TE activation in the sexual genotypes could be associated to mechanisms related to the expression of sexuality.

The effect of ploidy on TE expression in *E. curvula* inflorescences was also analyzed. The tetraploids T and C are autopolyploid (Cardone *et al*. 2006, Diaz *et al*. 2010). Genomic polymorphism arising during autotetra-ploidization was reported in the synthetic autotetraploid *Paspalum notatum* (Martelotto *et al.* 2005), where repetitive elements can play an important role in the increment of variability and adaptation of polyploid accessions (Rodriguez *et al.* 2012). Moreover, in the *E. curvula* 'back-and-forth' ploidy-altered system analyzed here, rapid epigenetic and expression changes associated with successive modifications of the ploidy level were shown (Mecchia *et al.* 2007, Cervigni *et al.* 2008b, Selva *et al.* 2012). None of these studies focused on expression of TEs in autopolyploids. Here, we identified three TEs that were differentially expressed according to ploidy. A *Copia-28-ZM_I_*-like and a *GypsyCRMA1*-like TEs were selectively identified in the tetraploid genotypes, whereas a *MuDR*-like TE was found in the diploid genotype, thus suggesting that polyploidization in *E. curvula* could lead to transposable element activation or deactivation. It is interesting to note that the plant D (2n = 2x = 20) and plant C (2n = 4x = 40) have recently suffered from genomic modifications leading to changes in ploidy (Cardone *et al.* 2006), a processes that involve large genome changes that can disrupt the mechanisms of TE control. The *BLASTX* analysis revealed the chimeric nature of the *MuDR*-containing ESTs showing similarity to a predicted pre-mRNA-splicing factor 18-like (XP_004958093). It has been previously reported that *MuDR* insertion in the first exon of a gene encoding a splicing factor has pleiotropic effects (Chung *et al.* 2007) caused by inefficient rRNA processing in endosperm as well as by up-regulation of ribosomal protein genes (Chung *et al.* 2009). Upregulation of two ribosomal genes, one in the diploid genotype and the other in the tetraploid ones, was also observed in the euploid series (Cervigni *et al.* 2008b, Selva *et al* 2012).

Thus, even when the little number of sequences of the EST libraries may make questionable the development of an accurate virtual northern strategy based on the abundance of ESTs, such libraries allowed us to postulate the existence of different ESTs based on their individual occurrence among the libraries Ec01 - Ec04. As expected, some of these predictions could not be demonstrated at biological samples since they showed to be expressed

similarly in most samples. However, the followed rationale allowed the experimental demonstration of seven EST-related sequences to be differently expressed among tissues and/or genotypes. It is expectable that increasing library sizes would have revealed a stronger correlation between different virtual and biological expression.

In the present work, we characterized the *E. curvula* repeatome in the genotypes of different ploidy level and/or reproductive mode. The repetitive sequences represented from 1.5 to 2 % of the total ESTs. Different expression according to tissue, ploidy, and reproductive mode could be demonstrated for three *Copia*-like, two *Gypsy*-like, and two *MuDR*-like TEs. Our results support the idea that TE activity is modulated during ploidy and reproductive mode changes in the *Poaceae*.

References

Adams, K.L., Cronn, R.C., Percifield, R., Wendel, J.F.: Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. - Proc. nat. Acad. Sci. USA **100**: 4649-4654, 2003.

Arabidopsis Genome Initiative: Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. - Nature **408**: 796-815, 2000.

Asker, S.E., Jerling, L.: Apomixis in Plants. - CRC Press, Boca Raton 1992.

Barcaccia, G., Albertini, E.: Apomixis in plant reproduction: a novel perspective on an old dilemma. - Plant Reprod. **26**: 159-179, 2013.

Bedell, J.A., Korf, I., Gish, W.: MaskerAid: a performance enhancement to RepeatMasker. - Bioinformatics **16**: 1040-1041, 2000.

Bennetzen, J.L.: The contributions of retroelements to plant genome organization, function and evolution. - Trends Microbiol. **4**: 347-353, 1996.

Blumenstiel, J.P.: Evolutionary dynamics of transposable elements in a small RNA world. - Trends Genet. **27**: 23-31, 2011.

Cardone, S., Polci, P., Selva, J.P., Mecchia, M., Pessino, S., Hermann, P., Cambi, V., Voigt, P., Spangenberg, G. Echenique, V.: Novel genotypes of the subtropical grass *Eragrostis curvula* for the study of apomixis (diplospory). - Euphytica **151**: 263-272, 2006.

Cervigni, G.D.L., Paniego, N., Díaz, M., Selva, J.P., Zappacosta, D., Zanazzi, D., Landerreche, I., Martelotto, L., Felitti, S., Pessino, S., Spangenberg, G., Echenique, V.: Expressed sequence tag analysis and development of gene associated markers in a near-isogenic plant system of *Eragrostis curvula*. - Plant mol. Biol. **67**: 1-10, 2008a.

Cervigni, G.D.L., Paniego, N., Pessino, S., Selva, J.P., Díaz, M., Spangenberg, G., Echenique, V.: Gene expression in diplosporous and sexual *Eragrostis curvula* genotypes with differing ploidy levels. - Plant. mol. Biol. **67**: 11-23, 2008b.

Chandler, V.L., Brendel, V.: The maize genome sequencing project. - Plant Physiol. **130**: 1594-1597, 2002.

Cheng, X., Zhang, D., Cheng, Z., Keller, B., Ling, H.Q.: A new family of *Ty1-copia*_like retrotransposons originated in the tomato genome by a recent horizontal transfer event. - Genetics **181**: 1183-1193, 2009.

Chung, T., Kim, C.S., Nguyen, H.N., Meeley, R.B., Larkins, B.A.: The maize *zmsmu*2 gene encodes a putative RNA-splicing factor that affects protein synthesis and RNA processing during endosperm development. - Plant Physiol. **144**: 821-835, 2007.

Chung, T., Wang, D., Kim, C.S., Yadegari, R., Larkins, B.A.: Plant *SMU*-1 and *SMU*-2 homologues regulate pre-mRNA splicing and multiple aspects of development. - Plant Physiol. **151**: 1498-1512, 2009.

Comai, L., Tyagi, A.P., Winter, K., Holmes-Davis, R., Reynolds, S.H., Stevens, .Y, Byers, B.: Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. - Plant Cell **12**: 1551-1567, 2000.

Conesa, A., Götzm, S., Garcia-Gomez, J.M., Terol, J., Talon, M., Robles, M.: Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. - Bioinformatics **21**: 3674-3676, 2005.

Crane, C.F.: Classification of apomictic mechanisms. - In: Savidan, Y., Carman, J.G., Dresselhaus, T. (ed.): The Flowering of Apomixis: from Mechanisms to Genetic Engineering. Pp. 24-43. CIMMYT, Mexico 2001.

De Araujo, P.G., Rossi, M., De Jesus, E.M., Saccaro, N.L., Kajihara, D., Massa, R., De Felix, J.M., Drummond, R.D., Falco, M.C., Chabregas, S.M., Ulian, E.C., Menossi, M., Van Sluys, M.A.V.: Transcriptionally active transposable elements in recent hybrid sugarcane. - Plant J. **44**: 707-717, 2005.

Diaz, .M.L, Garbus, I., Echenique, V.: Allele-specific expression of a weeping lovegrass gene from the lignin biosynthetic pathway, *caffeoyl-coenzyme A 3-O-methyl-transferase*. - Mol. Breed. **26**: 627-637, 2010.

Duarte Silveira, E., Alves-Ferreira, M., Arrais Guimaraes, L., Rodrigues da Silva, F., Tavares de Campos Carneiro, V.: Selection of reference genes for quantitative real-time PCR expression studies in the apomictic and sexual grass *Brachiaria brizantha*. - BMC Plant Biol. **9**: 84, 2009.

Echenique, V., Stamova, B., Wolters, P., Lazo, G., Carollo, L., Dubcovsky, J.: Frequencies of *Ty1- Copia* and *Ty3- Gypsy* retroelements within the *Triticeae* EST databases. - Theor. appl. Genet. **104**: 840-844, 2002.

Gómez, E., Schulman, A.H., Martínez-Izquierdo, J.A., Vicient, C.M.: Integrase diversity and transcription of the maize retrotransposon *Grande*. - Genome **49**: 558-562, 2006.

González, L.G., Deyholos, M.K.: Identification, characterization and distribution of transposable elements in the flax (*Linum usitatissimum* L.) genome. - BMC Genomics **13**: 644, 2012.

Hirochika, H.: Activation of tobacco retrotransposons during tissue culture. - EMBO J. **12**: 2521-2528, 1993.

Kashkush, K., Feldman, M., Levy A.: Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. - Genetics **160**: 1651-1659, 2002.

Kashkush, K., Feldman, M., Levy, A.A.: Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. - Nat. Genet. **33**: 102-106, 2003.

Kidwell, M.G.: Transposable elements and the evolution of genome size in eukaryotes. - Genetica **115**: 49-63, 2002.

Krom, N., Recla, J., Ramakrishna, W.: Analysis of genes associated with retrotransposons in the rice genome. - Genetica **134**: 297-310, 2008.

Li, W.D., Hu, X., Liu, J.K., Jiang, G.M., Li, O., Xing, D.: Chromosome doubling can increase heat tolerance in *Lonicera japonica* as indicated by chlorophyll fluorescence imaging. - Biol. Plant. **55**: 279-284, 2011.

Lisch, D.: How important are transposons for plant evolution? - Nat. Rev. Genet. **14**: 49-61, 2013.

Livak, K.J., Schmittgen, T.D.: Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. - Methods **25**: 402-408, 2001.

Lockton, S., Gaut, B.S.: The contribution of transposable elements to expressed coding sequence in *Arabidopsis thaliana*. - J. mol. Evol. **68**: 80-89, 2009.

Ma, X.-F., Gustafson, J.P.: Genome evolution of allopolyploids: a process of cytological and genetic diploidization. Cytogenet. Genome Res. **109**: 236-249, 2005.

Martelotto, L.G., Ortiz, J.P.A., Stein, J., Espinoza, F., Quarin, C.L., Pessino, S.C.: A comprehensive analysis of gene expression alterations in a newly synthesized *Paspalum notatum* autotetraploidy. - Plant Sci. **169**: 211-220, 2005.

Mecchia, M.A., Ochogavía, A., Selva, J.P., Laspina, N., Felitti, S., Martelotto, L.G., Spangenberg, G., Echenique, V., Pessino, S.C.: Genome polymorphisms and gene differential expression in a 'back-andforth' ploidy-altered series of weeping lovegrass (*Eragrostis curvula*). - J. Plant Physiol. **164**: 1051-1061, 2007.

Meier, M., Zappacosta, D., Selva, J.P., Pessino, S., Echenique, V.: Evaluation of different methods for assessing the reproductive mode of weeping lovegrass plants, *Eragrostis curvula* (Schrad.) Nees. - Aust. J. Bot. **59**: 253-261, 2011.

Messing, J., Bharti, A.K., Karlowski, W.M., Gundlach, H., Kim, H.R., Yu, Y., Wei, F., Fuks, G., Soderlund, C.A., Mayer, K.F., Wing, R.A.: Sequence composition and genome organization of maize. - Proc. nat. Acad. Sci. USA **101**: 14349-14354, 2004.

Mhiri, C., Morel, J.B., Vernhettes, S., Casacuberta, J.M., Lucas, H., Grandbastien, M.A.: The promoter of the tobacco *Tnt*1 retrotransposon is induced by wounding and by abiotic stress. - Plant mol. Biol. **33**: 257-266, 1997.

Muthukumar, B., Bennetzen, J.L.: Isolation and characterization of genomic and transcribed retrotransposon sequences from sorghum. - Mol. Genet. Genom. **271**: 308-316, 2004.

Newcombe, R.G.: Two-sided confidence intervals for the single proportion: comparison of seven methods. - Statistics Medicine **17**: 857-872, 1998

Nogler, G.A., Gametophytic apomixis. - In: Johri, B.M. (ed.): Embryology of Angiosperms. Pp. 475-518. Springer-Verlag, Berlin 1984.

Ochogavia, A.C., Seijo, J.G., Gonzalez, A.M., Podio, M., Duarte Silveira, E., Machado Lacerda, A.L., Tavares de Campos Carneiro, V., Ortiz, J.P., Pessino, S.C.: Characterization of retrotransposon sequences expressed in inflorescences of apomictic and sexual *Paspalum notatum* plants. - Sex. Plant Reprod. **24**: 231-246, 2011.

Okada, T., Ito, K., Johnson, S.D., Oelkers, K., Suzuki, G., Houben, A., Mukai, Y., Koltunow, A.M.: Chromosomes carrying meiotic avoidance loci in three apomictic eudicot *Hieracium* subgenus *Pilosella* species share structural features with two monocot apomicts. - Plant Physiol. **157**: 1327-1341, 2011.

Oliver, K.R., McComb, J.A., Greene, W.K.: Transposable elements: powerful contributors to angiosperm evolution and diversity. - Genome Biol. Evol. **5**: 1886-1901, 2013.

Paterson, A.H., Bowers, J.E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T., Poliakov, A., Schmutz, J., Spannagl, M., Tang, H., Wang, X., Wicker, T., Bharti, A.K., Chapman, J., Feltus, F.A., Gowik, U., Grigoriev, I.V., Lyons, E., Maher, C.A., Martis, M., Narechania, A., Otillar, R.P., Penning, B.W., Salamov, A.A., Wang, Y., Zhang, L., Carpita, N.C., Freeling, M., Gingle, A.R., Hash, C.T., Keller, B., Klein, P., Kresovich, S., McCann, M.C., Ming, R., Peterson, D.G., Mehboob-ur-Rahman, Ware, D., Westhoff, P., Mayer, K.F., Messing, J., Rokhsar, D.S.: The *Sorghum bicolor* genome and the diversification of grasses. - Nature **457**: 551-556, 2009.

Picault, N., Chaparro, C., Piegu, B., Stenger, W., Formey, D., Llauro, C., Descombin, J., Sabot, F., Lasserre, E., Meynard, D., Guiderdoni, E., Panaud, O.: Identification of an active LTR retrotransposon in rice. - Plant J. **58**: 754-765, 2009.

Plohl, M., LuchettI, A., Mestrović, N., Mantovani, B.: Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. - Gene **409**: 72-82, 2008.

Pouteau, S., Huttner, E., Grandbastien, M.A., Caboche, M.: Specific expression of the tobacco *Tnt1* retrotransposon in protoplasts. - EMBO J. **10**: 1911-1918, 1991.

Project IRGSP: The map-based sequence of the rice genome. - Nature **436**: 793-800, 2005.

Quarin, C.L., Espinoza, F., Martínez, E.J., Pessino, S.C., Bovo, O.A.: A rise of ploidy level induces the expression of apomixis in *Paspalum notatum*. - Sex. Plant Reprod. **13**: 243-249, 2001.

Rabinowicz, P.D., Bennetzen, J.L.: The maize genome as a model for efficient sequence analysis of large plant genomes. - Curr. Opin. Plant Biol. **9**: 146-156, 2006.

Ramallo, E., Kalendar, R., Schulman, A.H., Martínez-Izquierdo, J.A.: *Reme*1, a *Copia* retrotransposon in melon, is transcriptionally induced by UV light. - Plant mol. Biol. **66**: 137-150, 2008.

Rebozzio R., Sartor M., Quarin C., Espinoza F.: Residual sexuality and its seasonal variation in natural apomictic *Paspalum notatum*. - Biol. Plant. **55**: 391-395, 2011.

Rodriguez, M.P., Cervigni, G.D.L., Quarin, C.L. Ortiz, J.P.A.: Frequencies and variation in cytosine methylation patterns in diploid and tetraploid cytotypes of *Paspalum notatum*. - Biol. Plant. **56**: 276-282, 2012.

Saze, H., Kakutani, T.: Differentiation of epigenetic modifications between transposons and genes. - Curr. Opin. Plant Biol. **14**: 81-87, 2011.

Selva, J.P., Pessino, S., Meier, M., Echenique, V.: Identification of candidate genes related to polyploidy and/or apomixis in *Eragrostis curvula*. - Amer. J. Plant Sci. **3**: 403-416, 2012.

Thomas, C.A., Jr.: The genetic organization of chromosomes. - Annu. Rev. Genet. **5**: 237-256, 1971.

Ueki, N., Nishii, I.: Idaten is a new cold-inducible transposon of *Volvox carteri* that can be used for tagging developmentally important genes. - Genetics **180**: 1343-1353, 2008.

Vicient, C.: Transcriptional activity of transposable elements in maize. - BMC Genomics **11**: 601, 2010.

Vicient, C.M., Jääskeläinen, M.J., Kalendar, R., Schulman, A.H.: Active retrotransposons are a common feature of grass genomes. - Plant Physiol. **125**: 1283-1292, 2001.

Vicient, C.M., Schulman, A.H.: *Copia*-like retrotransposons in

the rice genome: few and assorted. - Genome Lett. **1**: 35-47, 2002.

Voigt, P., Rethman, N., Poverene, M.: Warm-Season (C4) Grasses. - American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison 2004.

Wang, J., Tian, L., Madlung, A., Lee, H.S., Chen, M., Lee, J.J., Watson, B., Kagochi, T., Comai, L., Chen, Z.J.: Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. - Genetics **167**: 1961-1973, 2004.

Watson, L., Dallwitz, M.J.: The Grass Genera of the World. - CAB International, Wallingford 1992.

Wendel, J.F.: Genome evolution in polyploids. - Plant mol. Biol **42**: 225-249, 2000.

Wicker, T., Sabot F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., San Miguel, P., Schulman, A.H.: A unified classification system for eukaryotic transposable elements. - Nat. Rev. Genet. **8**: 973-982, 2007.

Zappacosta, D.C., Ochogavía, A.C., Rodrigo, J.M., Romero, J.R., Meier, M.S., Garbus, I., Pessino S.C., Echenique, V.C.: Increased apomixis expression concurrent with genetic and epigenetic variation in a newly synthesized *Eragrostis curvula* polyploid. - Sci. Rep. **4**: 4423**,** 2014.

Zhou, F., Xu. Y.: RepPop: a database for repetitive elements in *Populus trichocarpa.* - BMC Genomics **10**: 14, 2009.