Mol Breeding (2010) 26:627–637
DOI 10.1007/s11032-010-9399-z

Allele-specific expression of a weeping lovegrass gene from the lignin biosynthetic pathway, caffeoyl-coenzyme A 3-*O*-methyltransferase

Marina Lucía Díaz · Ingrid Garbus · Viviana Echenique

Received: 31 October 2008 / Accepted: 25 January 2010 / Published online: 24 February 2010

Abstract *Eragrostis curvula* is an important forage grass in marginal regions for its capacity to grow and develop in sandy, not very fertile soils and for its drought tolerance. However, its widespread use for animal production is limited at present since it has low forage quality. In forage species, lignin content has been recognized as one of the main factors that affects this parameter. The *O*-methylation at the C3 position of the phenolic ring of caffeic acid constitutes a key step in the lignin biosynthetic pathway. The enzyme caffeoyl-CoA 3-*O*-methyltransferase (CCoAOMT) catalyzes such methylation and is thus considered an interesting target for molecular breeding programs. Here we report the isolation of the full-length *CCoAOMT* cDNA from the *E. curvula* inflorescences library. Primers based on this sequence led to the amplification of seven unigenes from genomic DNA from cvs. Tanganyika, Don Pablo and Kromdraai. Since the major differences were located in the intron regions, these seven sequences resulted in four possible allelic forms. We further evaluated the allelic expression per tissue in cv. Tanganyika, the most variable genotype. The four alleles predicted by the genomic sequences were found to be expressed. Three of them were common to inflorescences, roots and leaves, while the other one seemed to be specific for inflorescences. The statistical analysis showed that the expression was not organ-independent. The information reported here constitutes a valuable tool for *E. curvula* breeding programs, aiming to alter lignin biosynthesis to improve forage quality without causing undesirable effects.

Keywords *CCoAOMT* gene · Lignin · Weeping lovegrass · Allele expression

M. L. Díaz · I. Garbus
Departamento de Biología, Bioquímica y Farmacia, Universidad Nacional del Sur, San Juan 670, 8000 Bahía Blanca, Argentina

I. Garbus · V. Echenique
Centro de Recursos Naturales Renovables de la Zona Semiárida (CERZOS), CONICET, Camino de La Carrindanga Km 7, 8000 Bahía Blanca, Argentina

V. Echenique (✉)
Departamento de Agronomía, Universidad Nacional del Sur, San Andrés 800, 8000 Bahía Blanca, Argentina
e-mail: echeniq@criba.edu.ar

Introduction

Weeping lovegrass, *Eragrostis curvula* (Shrad.) Nees is a member of the Poaceae, subfamily Chloridoidea. The genus *Eragrostis* is composed of more than 250 species, characterized by a basic number of $x = 10$ chromosomes (Streetman 1963). Cultivars are highly polyploid and reproduce by apomixis. It is a valuable forage resource in marginal regions, able to grow in sandy, not very fertile soils, with high tolerance to drought and extreme pH conditions. Like other forage species with summer growth, it tends to have higher

fiber and lower protein contents when compared to those of temperate climates. This results in lower digestibility and lower animal production. For these reasons, forage quality is one of the main aspects to be considered in an *E. curvula* breeding program.

Lignification of plant cell wall has been identified as the major factor responsible for lowering digestibility of forage tissues as they mature (Buxton and Russell 1988; Vogel and Jung 2001). Lignin is the major structural component of secondarily thickened plant cell walls and provides mechanical support to plant tissues. The metabolic routes and enzymes involved in lignin biosynthesis have largely been described, and representative genes have been cloned from several plant species (De Melis et al. 1999; Grimming and Matern 1997; Heath et al. 1998; Larsen 2004; Osakabe et al. 1995; Tsuruta et al. 2007) and genetically manipulated in model species like tobacco (Pinçon et al. 2001) and poplar (Lu et al. 2004; Meyermans et al. 2000) and in forages like alfalfa (Guo et al. 2001; Marita et al. 2003).

The methylation catalyzed by caffeic acid 3-*O*-methyltransferase (COMT) and caffeoyl-CoA 3-*O*-methyltransferase (CCoAOMT) constitutes key steps of lignin biosynthesis (Do et al. 2007; Goujon et al. 2003; Inoue et al. 1998; Rastogi and Dwivedi 2008; Ye et al. 2001; Zhong et al. 1998). It has been reported that a CCoAOMT transgenic alfalfa exhibited a dramatic decrease in lignin content, although lignin composition remained unaltered (Marita et al. 2003). In *Arabidopsis CCoAOMT 1* mutants, the lignin content decrease was accompanied by a change in lignin quality. Authors also reported that expression was organ-specific, being restricted to vascular tissues, except in seedlings where it seems to be constitutively expressed (Do et al. 2007).

In several plant species, *CCoAOMT*s constitute a gene family. Three genes were identified in rice (Zhao et al. 2004), four were cloned in tobacco (Martz et al. 1998) and five to ten in *Zinnia* (Ye et al. 1994). However, in other species the situation is less complex, such as in parsley and alfalfa where only one gene was reported (Grimming and Matern 1997; Inoue et al. 1998) or poplar, with two genes (Chen et al. 1998).

The potential of *E. curvula* as a valuable forage resource reveals the importance of directing breeding efforts to improve its digestibility. The characterization of *E. curvula* genes involved in lignification opens up the possibility of improving forage digestibility by genetically modifying lignin biosynthesis. In the present paper, *E. curvula CCoAOMT* cDNA isolated from a cDNA inflorescence library was used as an efficient tool for the extended characterization of the corresponding genes in three *E. curvula* cultivars, which allowed an organ-specific expression analysis.

Materials and methods

cDNA libraries' construction and characterization

Details of libraries' construction and characterization were reported by Cervigni et al. (2008). Briefly, total RNA from cv. Tanganyika ($2n = 4x = 40$) immature inflorescences was extracted using the RNeasy total RNA isolation kit (Promega). cDNAs were obtained using the SMART PCR synthesis kit (Clontech), cloned into the pGEM®-T Easy Vector(Promega) and used for transformation of XL10-Gold Ultracompetent *E. coli* cells (Stratagene). Each insert was characterized by a single read using a MegaBACE 4000 sequencer. BLASTX algorithm (Altschul et al. 1994) against the SwissProt/Trembl databases of the sequences from library EC02 (from inflorescences) was used for the detection of ESTs related to the lignin biosynthetic pathway.

Sequencing and analysis of variation in the CCoAOMT gene of *E. curvula*

Genomic DNA was isolated from leaves of 50-day-old plants from cultivars Tanganyika, Don Pablo INTA ($2n = 7x = 70$) and Kromdraai ($2n = 6x = 60$) using a CTAB protocol (CIMMYT 2005). Primers were designed based on the identified *CCoAOMT* sequence. The sequence and positions of the primers (CCoAOMT 3 and CCoAOMT 4) along the cDNA are indicated in Fig. 1. PCRs were carried out in 25 µl reaction mixture containing 5 µl Taq DNA polymerase buffer (5×), 1 µl of each primer (10 µM), 2 µl dNTP (2.5 mM), 1.25 U of Taq DNA polymerase (Promega) and 150 ng of genomic DNA. After an initial denaturation at 95°C for 3 min, the reaction mixture was subjected to amplification in a thermal cycler (My Cycler™, BIORAD) for 35 cycles consisting of 1 min at 95°C, 30 s at 56°C and 1 min at 72°C, followed by 10 min incubation at 72°C.

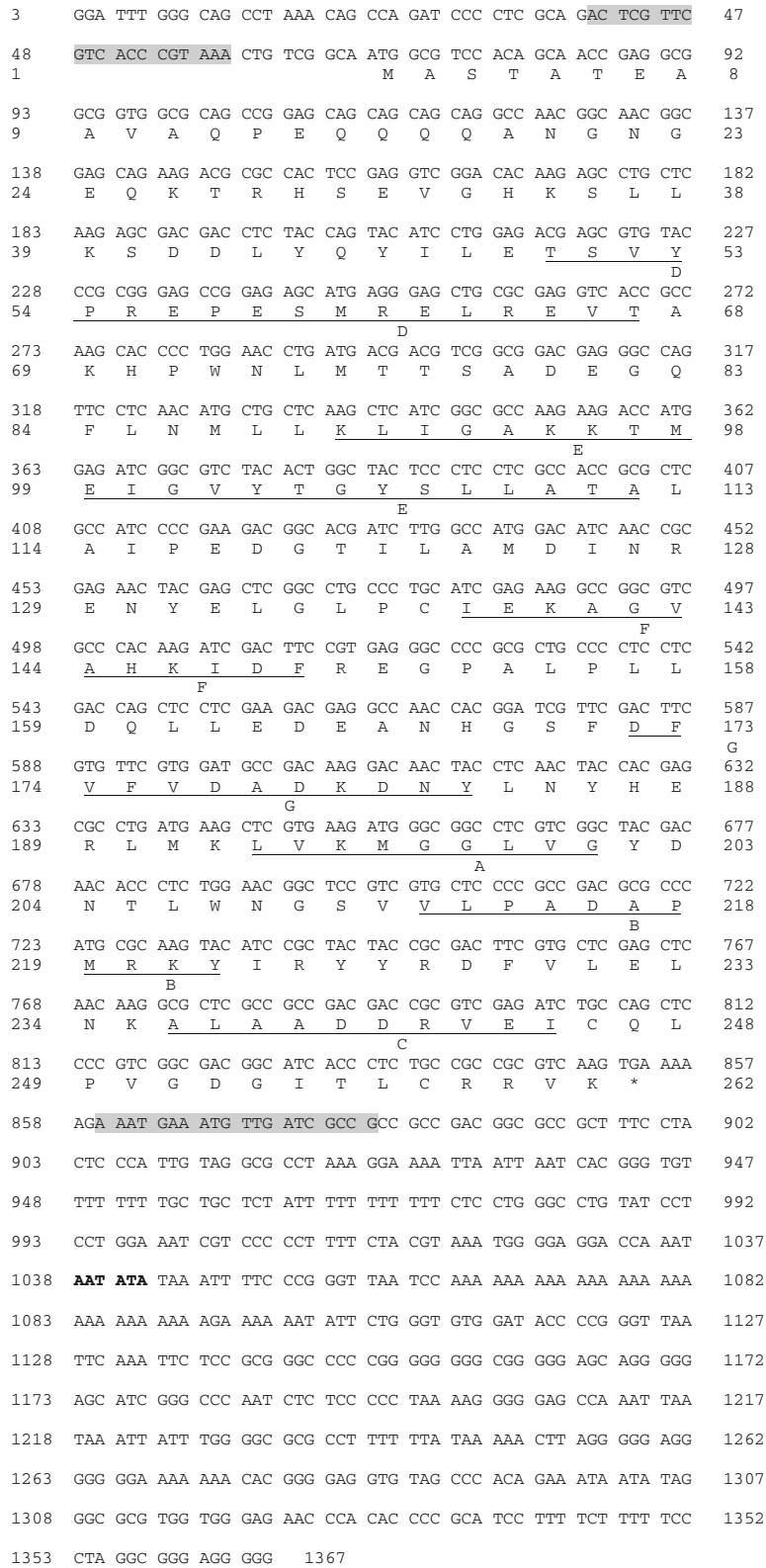

Fig. 1 Nucleotide and deduced amino acid sequence of *Eragrostis curvula CCoAOMT* cDNA. The nucleotide sequence is numbered from the 5′ end of the cDNA. The first methionine of the open reading frame is designated as the first amino acid of the putative polypeptide. Underlined sequences A, B and C represent putative SAM-binding motifs; D, E, F, G and H represent additional CCoAOMT signature motifs (Joshi and Chiang 1998). Polyadenylation signal is indicated in *bold*. Darkened sequences represent the sequences of primers CCoAOMT 3 and 4

The amplification products were cloned using the pGEM®-T Easy Vector kit (Promega) and 11, 16 and 18 clones from the cvs Tanganyika, Kromdraai and Don Pablo, respectively, were sequenced. Contig assembly was done using the CAP3 software (Huang and Madan 1999), with a high stringency level (sequence identity $P = 90$) and overlapping lengths of $O = 40$. The SSR Discovery program (Robinson et al. 2004) was used for the identification of SSR markers. For the SSR selection, a minimum of 5 repeats for di-, 4 for tri- and 5 for tetranucleotide arrays were considered.

Southern hybridization

Twenty micrograms of genomic DNA from cv. Tanganyika were digested separately with *Eco*RI, *Hind*III and *Apa*I. The restriction fragments were separated by electrophoresis on a 0.8% agarose gel and blotted onto a nylon membrane by capillarity transfer. The DNA was fixed onto the membrane by 1 min UV irradiation followed by 2 h incubation at 80°C. The filter was prehybridized for 2 h at 63°C in prehybridizing solution (5× SSC, 0.02% SDS, 0.01% *S*-lauryl-Sarcosine) and 0.2% of blocking solution (0.1 M maleic acid, 0.15 M NaCl, 0.2% casein p/v; pH 7.5) and subsequently hybridized overnight at 63°C with the dig-dUTP labeled probe (5-25 ng/ml). Prior to autoradiography, it was washed four times for 15 min. Two washes were performed with 2× SSC, 0.1% SDS at room temperature and the others at 60°C using a 0.25× SSC and 0.1% SDS.

Expression studies

RT-PCR and CCoAOMT expression profile

Total RNA was isolated from different tissues of cv. Tanganyika plants (leaves and roots of 12-day-old plants and immature inflorescences of mature plants growing in the glasshouse) using the SV Total RNA Isolation System Kit (Promega). The RNA was resuspended in DEPC-treated water and the concentration in each sample was determined by the absorbance at 260 nm. For all the samples, reverse transcription was performed with 3 µg of total RNA using the First-Strand cDNA Synthesis Kit (Amersham Biosciences) with oligo *Not* I-d$(T)_{18}$ according to the manufacturer's instructions.

PCR amplification of the cDNA was performed using two specific primer pairs based on the *E. curvula CCoAOMT* cDNA sequence: CCoAOMT 1 (5′-AAGAGCGACGACCTCTACCA-3′) and CCoA OMT 2 (5′-AACACGAGTCGAACGATCC-3′). The PCR program consisted of a denaturation step at 95°C for 3 min, followed by 25 cycles of 1 min at 95°C, 30 s at 59°C and 1 min at 72°C, followed by an extra extension step of 10 min at 72°C. As an internal control, the expression of the actin gene from *E. curvula* (EH189703.1) was used. The primers Act1 (5′-GGATCGTTCGACTTCGTGTT-3′) and Act2 (5′-CACATCTGCTGGAATGTGCT-3′) were designed for this purpose. cDNA synthesis and PCR reactions were repeated twice.

RT-PCR and organ-specific CCoAOMT allele identification

In order to search for organ-specific allele expression, one of the replicates of the single strand DNA obtained in the previous section was amplified by PCR using the primers CCoAOMT 3 and CCoAOMT 4 (Fig. 1). The PCR program consisted of a denaturation step at 95°C for 3 min, followed by 30 cycles of 1 min at 95°C, 30 s at 56°C and 1 min at 72°C, followed by an extra extension step of 10 min at 72°C.

Competent DH5α *E. coli* bacteria were transformed with the amplification products cloned into pGEM®-T Easy Vector (Promega) and plated in X-Gal/IPTG LB-agar with ampicillin (100 mg/L). Plasmids from white colonies were purified (Wizard Plus SV Minipreps DNA Purification System, Promega) and sequenced. A total of 21 sequences from inflorescences, 25 from leaves and 22 from roots were obtained. For each tissue, white colonies were selected from plates with more than 100 eligible colonies. The sequence alignment was performed with the Clustal W program. Data were statistically analyzed using the InfoStat (2006) software.

Phylogenetic trees construction

The multiple sequences alignments were performed with the Clustal W program using the Neighbor-Joining (NJ) method (Saitou and Nei 1987) with MEGA4 software (Tamura et al. 2007). Sites containing alignment gaps were removed before the final

analysis (complete-deletion option). Consensus trees and confidence values for the nodes were calculated using 1,000 bootstraps. Trees were drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic trees.

Results

Analysis of weeping lovegrass CCoAOMT cDNA

The BLASTX algorithm allowed the detection in the library of one EST, the EC02_d_2916.ab1 (GenBank accession no. EH189157.1), corresponding to the *E. curvula CCoAOMT* gene. The clone was completely sequenced, and the full-length cDNA was obtained. It consisted of 1,367 bp with an open reading frame (ORF) of 786 bp, a 5′ noncoding region of 68 bp and a 3′ noncoding region of 513 bp, including a poly(A) tail (Fig. 1). The GC content of the coding region was 65.26%, which is similar to the observed values in other monocot species, like rice, maize and *Brachypodium*. Contrastingly, GC content reported for dicots species ranges between 40.81 and 51.83% (Campell and Gowri 1990). The polypeptide predicted by the coding region consisted of 261 amino acids, with an estimated molecular mass of 29.3 kDa.

Comparison of the *E. curvula* CCoAOMT cDNA with other methyltransferases: identification of conserved motifs and evolutionary relationships

The comparison between weeping lovegrass *CCoAOMT* cDNA allowed the construction of a phylogenetic tree that showed the expected division between the two most important taxonomic groups, monocots and dicots (Fig. 2). Group I included dicots and the *CCoAOMT* from a gymnosperm, and Group II included only monocots sequences. This suggests that the *CCoAOMT* gene might remain conserved during the evolution of the monocots. Contrastingly, two of the three rice *CCoAOMT* genes proved to be species-specific (Zhao et al. 2004).

The proposed structure for CCoAOMT protein consists of three putative SAM-binding motifs (A, B and C) and five other regions (D, E, F, G and H), that may serve as CCoAOMT signatures in plants (Joshi and Chiang 1998). Accordingly, the CCoAOMT

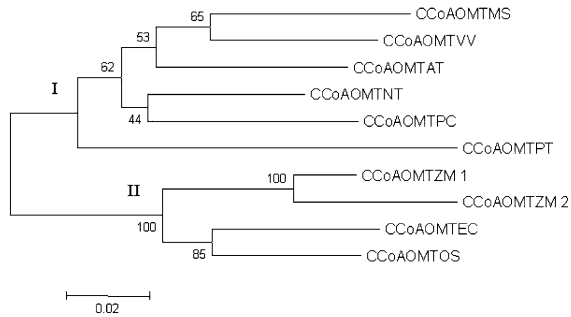

Fig. 2 Evolutionary relationship of the 10 CCoAOMT amino acid sequences aligned in Fig. 2, using the Neighbor-Joining method (Saitou and Nei 1987). The optimal tree with the sum of branch length = 0.63762624 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling 1965) and are in the units of the number of amino acid substitutions per site. There were a total of 240 positions in the final dataset

protein sequence predicted here revealed the presence of the eight mentioned motifs (A–H) with conserved spatial relationships (Fig. 1). Moreover, the *E. curvula* CCoAOMT protein sequence exhibits the complete conservation of the key residues identified previously by X-ray crystallography in alfalfa by Ferrer et al. (2005), such as Lys21 and Arg206, involved in the interaction with the CoA, and Met61, Asp165, Trp193, Tyr208 and Tyr212, related to substrate recognition.

Analysis of the variation in the CCoAOMT gene of *E. curvula*

A single band of ~1,300 bp was amplified by PCR from cvs. Tanganyika, Don Pablo and Kromdraai with the primer pair CCoAOMT 3/CCoAOMT 4, based on the cDNA sequence (Fig. 1).

The genomic sequence of the gene is arranged into five exons separated by four introns of 128, 104, 85 and 92 bp (Fig. 3). Comparison of the three genotypes at the predicted amino acid sequences level showed differences at seven positions, being only two non-equivalent substitutions. However, these substitutions did not involve key residues of the protein.

The 45 sequences were analyzed with CAP3 software, resulting in 42 assembled in 4 contigs and

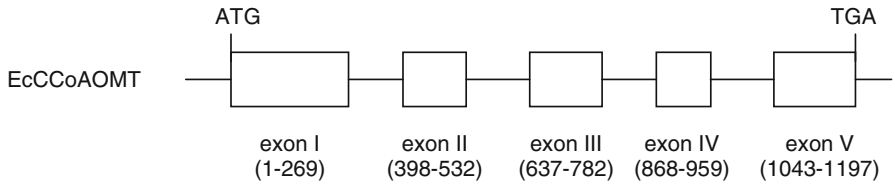

Fig. 3 Schematic representation of the structure of the *E. curvula CCoAOMT* (*EcCCoAOMT*) gene. Five exons with their initial and final nucleotides are shown

3 sequences remaining as singletons (Table 1). The major differences that govern this segregation reside in the intron regions. In contig #4, assembled with 25 sequences, a microsatellite in the first exon was identified. It consisted of a trinucleotide (CAG) repetition at position 70 and led to an additional amino acid residue, a glutamine (Gln). The low number of sequences assembled in the other three contigs hampered a statistically valid analysis of the occurrence of microsatellites.

Determination of the copy number of CCoAOMT gene by Southern hybridization

Cultivar Tanganyika genomic DNA was digested separately with the restriction enzymes *Apa*I, *Eco*RI and *Hind*III. Clone EC02_d_1042 (GenBank accession no. EH187122) was used as probe. It consisted of a partial sequence of 1,070 bp of the *E. curvula CCoAOMT* cDNA, having a restriction site for the enzyme *Apa*I and no internal sites for *Eco*RI and *Hind*III. The complete cDNA has an additional *Apa*I restriction site near the end of the cDNA clone, at 1,147 bp. The most plausible interpretation of the restriction profiles observed with *Hind*III and *Apa* I (Fig. 4) suggested the existence of at least two copies of the *CCoAOMT* gene, with the *Hind*III site positioned between both copies of the gene and thus leading to two bands. If there was only one copy per homologue chromosome, the presence of two *Apa*I sites (Fig. 4) would result in two bands, since the last portion of the gene would not be detected by the probe. The highest band corresponds to the first portion of the gene and the other one, a ~600 bp band, is framed by the *Apa*I sites. The existence of two copies gave a ~600 bp band and two other bands, each one corresponding to the first portion of the gene. Finally, the absence of *Eco*RI restriction sites between both copies of the gene, evidenced by the presence of a single band, suggests that both copies are relatively

Table 1 *E. curvula CCoAOMT* contigs and singletons assembled using the CAP3 software

Assembled unigenes	Sequences included	Sequences origin
1	8	Kromdraai (5)
		Tanganyika (3)
2	3	Tanganyika (2)
		Kromdraai (1)
3	7	Don Pablo (7)
4	25	Tanganyika (4)
		Don Pablo (11)
		Kromdraai (10)
5	1	Tanganyika
6	1	Tanganyika
7	1	Tanganyika

Assembled unigenes (contigs or singletons) are numbered consecutively from 1 to 7 in the first column. The second column shows the quantity of sequences included, whereas the last one shows the cultivar from which the sequences were obtained with the individual contribution in brackets

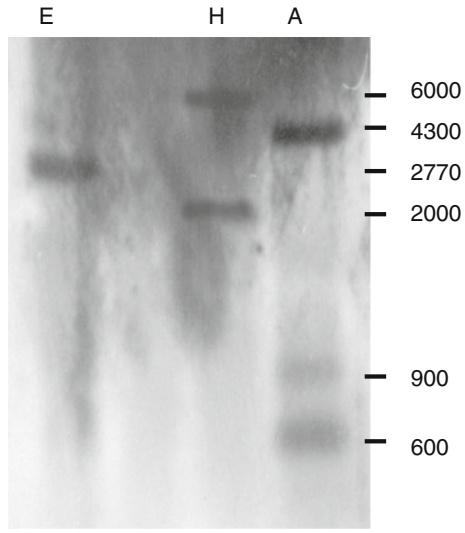

Fig. 4 Southern hybridization of *E. curvula* cv. Tanganyika genomic DNA digested with *Eco*RI (E), *Hind*III (H) and *Apa*I (A), separated on a 0.8% agarose gel, and blotted onto nitrocellulose filters. The estimated band sizes are indicated in bp. The 1,070 bp *CCoAOMT* clone (H7) was used as hybridization probe

close. The additional bands may not reflect the presence of different allelic forms, considering that they showed the same restriction pattern.

Expression of CCoAOMT gene in different tissues

The expression pattern of the *CCoAOMT* gene in different tissues was explored by RT-PCR, using the primer pair CCoAOMT 1/CCoAOMT 2. This study was carried out in cv. Tanganyika plants because it proved to be the most variable cultivar, as judged by the number of unigenes revealed from its genomic sequences. The *CCoAOMT* gene was expressed in inflorescences, leaves and roots of *E. curvula* plants, the former showing the strongest expression signal (Fig. 5a). A possible source of genomic DNA contamination was discarded, since a single band of 1,100 bp would be expected to be amplified from genomic DNA (Fig 5b).

The amino acid sequence deduced from cv. Tanganyika's unigenes (genomic sequences) (see Table 1) revealed the existence of six polymorphisms, arranged in a way that leads to the identification of four alleles (Figs. 5, 6, 7). The number of expressed alleles is lower than expected from the observed unigenes since, as was mentioned in the previous section, the major polymorphisms were located in the introns. We further analyzed the expression of these alleles in different tissues. Thus, one of the cDNA replicates per tissue of cv. Tanganyika was subjected to PCR amplification with the primer pair CCoAOMT 3/CCoAOMT 4. The amplification products were further cloned and sequenced, obtaining 21, 25 and 22 sequences from inflorescences, leaves and roots, respectively.

The translated sequences of the 68 cDNA clones revealed the expression of the four alleles predicted by the genomic sequences, named *CCoAOMT1.1*, *CCoAOMT1.2, CCoAOMT2.1* and *CCoAOMT2.2*. The statistical analysis of the contribution of the individual alleles to the different organ expression demonstrated that such expression was not organ-independent ($\chi^2 = 27.36$, 6 *df*, $P = 0.0001$) (Table 2). Since the *CCoAOMT2.1* allele was absent in leaves and roots, it was removed from the further analysis that showed differences in the expression of the remaining alleles ($\chi^2 = 10.28$, 4 *df*, $P = 0.0359$).

Discussion

In this study we reported the characterization and expression pattern of the *E. curvula CCoAOMT* gene. The cDNA identified in the library and the deduced amino acid sequence showed a high homology with other *CCoAOMTs* from different plant species, demonstrating at the same time that this gene is highly conserved among species and that the clone effectively codes for this enzyme. The alignment with the corresponding sequence from maize (Q9XGD6 and Q9XGD5.1) and bamboo (ABO26812.1) showed a high homology level (90–95%, *e*-value $\leq$ 4e-126) with the *E. curvula* predicted protein sequence. Furthermore, the conserved motifs and their characteristic spatial relationships for plant *S*-adenosyl-L-methionine-dependent methyltransferases (Joshi and Chiang 1998) were identified. Finally, the residues involved in the active site and other key sites of alfalfa caffeoyl coenzyme A 3-*O*-methyltransferase (Ferrer et al. 2005), essential for SAM-binding pocket formation (Ferrer et al. 2005; Lepelley et al. 2007), were identified in the present sequence.

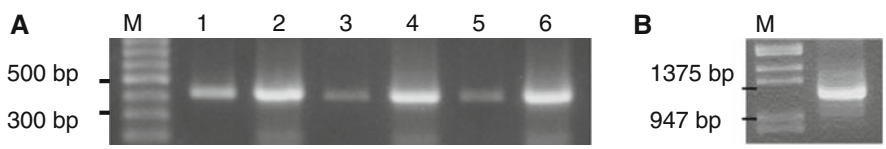

Fig. 5 **a** RT-PCR of mRNA from different tissues of *Eragrostis curvula*. *Lanes 1, 3 and 5* correspond to RT-PCR products amplified with specific CCoAOMT 1 and 2 primers from immature inflorescences, leaves and roots of 12-day-old plants, respectively. *Lanes 2, 4 and 6* correspond to RT-PCR products amplified with specific *E. curvula* actin primers (Act1 and 2) from immature inflorescences, leaves and roots of 12-day-old plants, respectively. *Lane M* contained a 100-bp ladder (Promega). **b** PCR product amplified from *E. curvula* genomic DNA using primers CCoAOMT 1 and 2. *Lane M* contained a *λ*/*Eco*RI *Hind*III ladder (Promega)

```
EcCCoAOMT1.1 MASTATEAAVAQPEQQQQ-ANGNGEQKTRHSEVGHKSLLKSDDLYQYILETSVYPREPES 59
EcCCoAOMT1.2 MASTATEAAVAQPEQQQQQANGNGEQKTRHSEVGHKSLLKSDDLYQYILETSVYPREPES 60
EcCCoAOMT2.1 MASTATEAAVAQPEQQQQ-ANGNGEQKTRHSEVGHKSLLKSDDLYQYILETSVYPREPES 59
EcCCoAOMT2.2 MASTATEAAVAQPEQQQQQANGNGEQKTRHSEVGHKSLLKSDDLYQYILETSVYPREPES 60
             ****************** ****************************************

EcCCoAOMT1.1 MRELREVTAKHPWNLMTTSADEGQFLNMLLKLIGAKKTMEIGVYTGYSLLATALAIPEDG 119
EcCCoAOMT1.2 MKELREVTAKHPWNLMTTSADEGQFLNMLLKLIGAKKTMEIGVYTGYSLLATALAIPEDG 120
EcCCoAOMT2.1 MRELREVTAKHPWNLMTTSADEGQFLNMLLKLIGAKKTMEIGVYTGYSLLATALAIPEDG 119
EcCCoAOMT2.2 MKELREVTAKHPWNLMTTSADEGQFLNMLLKLIGAKKTMEIGVYTGYSLLATALAIPEDG 120
             *:************************************************************

EcCCoAOMT1.1 TILAMDINRENYELGLPCIEKAGVAHKIDFREGPALPLLDQLLEDEANHGSFDFVFVDAD 179
EcCCoAOMT1.2 TILAMDINRENYELGLPCIEKAGVAHKIDFREGPALPLLDQLLEDEANHGSFDFVFVDAD 180
EcCCoAOMT2.1 TILAMDINRENYELGLPCIEKAGVAHKIDFREGPALPLLDQLLEDEANHGSFDFVFVDAD 179
EcCCoAOMT2.2 TILAMDINRENYELGLPCIEKAGVAHKIDFREGPALPLLDQLLEDEANHGSFDFVFVDAD 180
             ************************************************************

EcCCoAOMT1.1 KDNYLNYHERLMKLVKMGGLVGYDNTLWNGSVVLPADAPMRKYIRYYRDFVLELNKALAA 239
EcCCoAOMT1.2 KDNYLNYHERLMKLVKMGGLVGYDNTLWNGSVVLPADAPMRKYIRYYRDFVLELNKALAA 240
EcCCoAOMT2.1 KDNYLNYHERLMKLVKMGGVVGYDNTLWNGSVVLPDDAPMRKYIRYYRDFVIVLNKALAA 239
EcCCoAOMT2.2 KDNYLNYHERLMKLVKMGGVVGYDNTLWNGSVVLPDDAPMRKYIRYYRDFVIVLNKALAA 240
             ******************:*************** ***************: *******

EcCCoAOMT1.1 DDRVEICQLPVGDGITLCRRVK 261
EcCCoAOMT1.2 DDRVEICQLPVGDGITLCRRVK 262
EcCCoAOMT2.1 DDRVEICQLPVGDGITLCRRVK 261
EcCCoAOMT2.2 DDRVEICQLPVGDGITLCRRVK 262
             **********************
```

Fig. 6 Multiple alignment of the translated amino acid sequences of *E. curvula* CCoAOMT as deduced from four alleles. The analysis revealed six polymorphisms: at position 18 an additional glutamine (Q) is observed and at positions 61, 199, and 230 there are functionally conserved differences (:). Amino acids at positions 214 and 231 are not equivalent

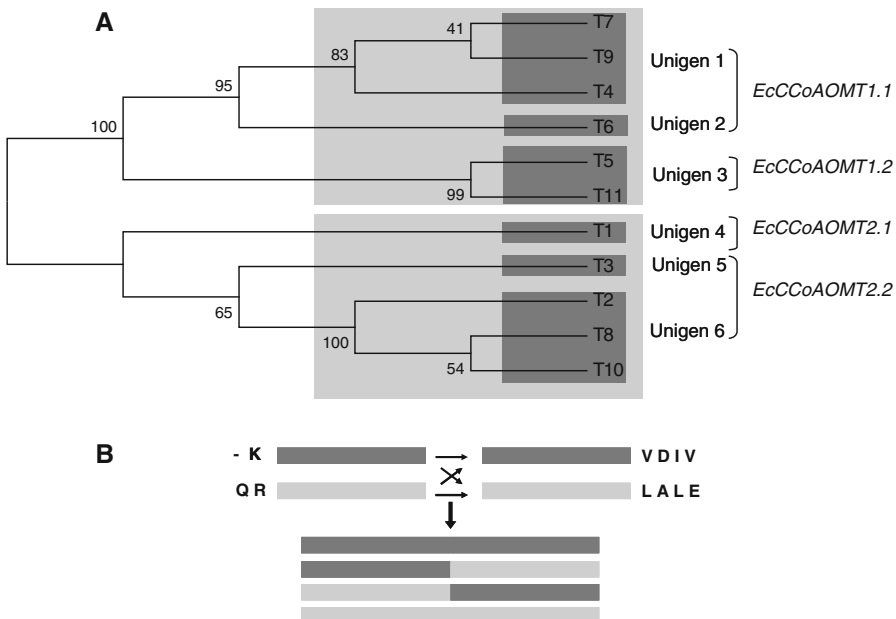

Fig. 7 **a** Evolutionary relationship of the 11 sequences obtained from cv. Tanganyika *E. curvula* (T1–11) (linearized). The optimal tree with the sum of branch length = 0.02966018 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) are shown next to the branches (Felsenstein 1985). The phylogenetic tree was linearized assuming equal evolutionary rates in all lineages (Takezaki et al. 2004). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. There were a total of 1202 positions in the final dataset. **b** Schematic representation of the origin of *E. curvula CCoAOMT* alleles. *Arrows* indicate the possible recombination events. Amino acid polymorphisms are shown in the original sequences. The minus sign (−) indicates the absence of a glutamine residue in the sequence

Table 2 Alleles of *E. curvula* cv. Tanganyika *CCoAOMT* (*EcCCoAOMT*) genes expressed in three different organs

Allele	Organ			Total
	I	L	R	
EcCCoAOMT1.1	2	9	3	14
EcCCoAOMT1.2	9	8	6	23
EcCCoAOMT2.1	7	0	0	7
EcCCoAOMT2.2	3	8	13	24
Total	21	25	22	68

The first column numbers the identified alleles. The following columns indicate the quantity of sequences for each allele identified in each tissue, i.e., inflorescence (I), leaf (L) or root (R). The last row shows the total sequences analyzed by tissue whereas the last column indicates the frequency of individual alleles in the total analyzed clones

The copy number of *CCoAOMT* gene exhibits a great variation among species, even among those related, like maize and rice. As inferred from the Southern hybridization analysis, two copies of the *CCoAOMT* gene were identified in *E. curvula* cv. Tanganyika (4x) (Fig. 4).

Based on the analysis of 45 genomic sequences obtained from the three different cultivars included in this study, it was demonstrated that *E. curvula CCoAOMT* genes are organized in five exons and four short introns (Fig. 3), in concordance with the maize *CCoAOMT1* gene intron–exon arrangement (Guillet-Claude et al. 2004). In contrast, three introns were identified in rice *CCoAOMT* (AY644636) and maize *CCoAOMT2* (AY279035.1) genes (Guillet-Claude et al. 2004). The analysis performed with the CAP3 software of the genomic DNA sequences of the three cultivars led to the discovery of seven different alleles (Table 1).

An unexpected variability was detected in cv. Tanganyika (4x) genomic sequences, assembling three contigs, whereas three sequences remained as singletons. A phylogenetic tree performed with the genomic sequences obtained from this cultivar revealed the existence of two ancestral *EcCCoAOMT* genes, *EcCCoAOMT1* and *EcCCoAOMT2,* as predicted by the Southern hybridization, each one leading to three different sequences (unigenes) (Fig 7a). The predicted proteins coded by these genomic sequences would result in four allelic forms, called *EcCCoAOMT1.1*, *EcCCoAOMT1.2*, *EcCCoAOMT2.1* and *EcCCoAOMT2.2*, identical to those obtained from the expression studies (Fig 7a, b; Table 2). The most plausible explanation of such findings could be that a duplication in the *EcCCoAOMT* gene occurred at the diploid stage of the species by unequal crossing-over or rearrangement of this genomic region. The observed allelic variants resulted from recombination events among the gene copies (Fig. 7b) and then, probably in tetraploid apomictics, insertion/deletions in the intron region resulted in the six reported unigenes.

We further evaluated the organ distribution of the *EcCCoAOMT* expression. The selection of cv. Tanganyika for these studies was based on the fact that it is the best candidate to be genetically manipulated, as justified by many criteria. First, its genetic variability, since it was the cultivar that showed the highest number of unigenes, possessing six of the seven identified alleles. Second, Tanganyika, as well as Don Pablo, is a fully apomictic cultivar, particularly interesting for in vitro manipulations. Third, cv. Don Pablo (7x) is highly polyploid, having also a high content of silica that could alter its digestibility and also palatability (unpublished data).

The RT-PCR analysis demonstrated that *EcCCoAOMT* gene(s) is/are being expressed in the tissues under consideration, i.e., inflorescences, leaves and roots (Fig. 5a). The detailed analysis of the expressed sequences revealed that: (a) the four predicted alleles were expressed, (b) the expression of allele *EcCCoAOMT2.1* was restricted to inflorescences, and (c) alleles *EcCCoAOMT1.1, EcCCoAOMT1.2* and *EcCCoAOMT2.2* were expressed in inflorescences, leaves and roots. However, statistical analysis of the expressed sequences revealed that the expression pattern was not independent of the considered tissue. The expression of more than one allelic form in all the analyzed tissues is an advantage since: (a) the *EcCCoAOMT* genes' downregulation strategies could be allele-specific, improving forage quality without the drastic effect that could result from a single allele being suppressed: this could lead to reduced pathogen resistance, increased winter mortality and reduced biomass, among other undesirable effects; and (b) the fact that allele expression is unbalanced between tissues offers the additional benefit of directing the suppression in a way that predominantly affects one tissue over the others. Considering the alleles expressed in leaves at equal level, *EcCCoAOMT2.2* is not a good candidate to be

downregulated considering that it is highly expressed in roots. The manipulation of *EcCCoAOMT1.2* may cause an undesirable alteration in inflorescences, essential for breeding programs, even taking into account that panicle production is detrimental to digestibility. The allele *EcCCoAOMT1.1* constitutes a potential candidate considering the high expression level shown in leaves compared to other organs. We strongly believe that the idea presented is very relevant for forage breeding purposes.

Acknowledgments We warmly acknowledge Natalia Moirano for performing the Southern Blotting, and Dr Alicia Carrera for helpful discussion. This work was funded by the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT, PICT 14624 and PAV 137) and Secretaría de Ciencia y Técnica (SECyT -UNS, PGI 24/A133), Argentina.

References

Altschul SF, Boguski MS, Gish W et al (1994) Issues in searching molecular sequence databases. Nat Genet 6:119–129

Buxton DR, Russell JR (1988) Lignin constituents and cell wall digestibility of grass and legume stems. Crop Sci 28:553–558

Campell W, Gowri G (1990) Codon usage in higher plants, green algae and cyanobacteria. Plant Physiol 92:1–11

Cervigni GD, Paniego N, Díaz M et al (2008) 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

Chen C, Meyermans H, Doorsselaere J et al (1998) A gene encoding caffeoyl coenzyme A 3-*O*-methyltransferase (CCoAOMT) from *Populus trichocarpa* (accession no. AJ223621) (PGR 98-104). Plant Physiol 117:719

CIMMYT (2005) Laboratory protocols: CIMMYT applied molecular genetics laboratory, 3rd edn. CIMMYT, Mexico, DF

De Melis LE, Whiteman PH, Stevenson TW (1999) Isolation and characterisation of a cDNA clone encoding cinnamyl alcohol dehydrogenase in *Eucalyptus globulus* Labill. Plant Sci 143:173–182

Do C, Pollet B, Théyenin J et al (2007) Both Caffeoyl Coenzyme A 3-*O*-methyltransferase 1 and caffeic acid *O*-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in *Arabidopsis*. Planta 226:1117–1129

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Ferrer JL, Zubieta C, Dixon RA et al (2005) Crystal structures of alfalfa caffeoyl CoA 3-*O*-methyltransferase. Plant Physiol 137:1009–1017

Goujon T, Sibout R, Eudes A et al (2003) Genes involved in the biosynthesis of lignin precursors in *Arabidopsis thaliana*. Plant Physiol Biochem 41:677–687

Grimming B, Matern U (1997) Structure of the parsley caffeoyl-CoA *O*-methyltransferase gene, harbouring a novel elicitor responsive cis-acting element. Plant Mol Biol 33:323–341

Guillet-Claude C, Birolleau-Touchard C, Manicacci D et al (2004) Genetic diversity associated with variation in silage corn digestibility for three *O*-methyltransferase genes involved in lignin biosynthesis. Theor Appl Genet 110:126–135

Guo D, Chen F, Inoue K et al (2001) Downregulation of caffeic acid 3-*O*-methyltransferase and caffeolyl-CoA 3-*O*-methyltransferase in transgenic alfalfa: impacts on lignin structure and implications for the biosynthesis of G and S lignin. Plant Cell 13:73–88

Heath R, Huxley H, Stone B et al (1998) cDNA cloning and differential expression of three caffeic acid *O*-methyltransferase homologues from Perennial Ryegrass (*Lolium perenne*). J Plant Physiol 152:649–657

Huang X, Madan A (1999) CAP3: a DNA sequence assembly program. Genome Res 9:868–877

InfoStat (2006) Grupo InfoStat, FCA. Universidad Nacional de Córdoba, Argentina

Inoue K, Sewalt WJH, Balance GM et al (1998) Developmental expression and substrate specificities of alfalfa caffeic acid 3-*O*-methyltransferase and caffeoyl coenzyme A 3-*O*-methyltransferase in relation to lignification. Plant Physiol 117:761–770

Joshi CP, Chiang VL (1998) Conserved sequences motifs in plant S-adenosyl-L-methionine-dependent methyltransferases. Plant Mol Biol 37:663–674

Larsen K (2004) Cloning characterization of a ryegrass (*Lolium perenne*) gene encoding cinnamoyl-CoA reductase (CCR). Plant Sci 166:569–581

Lepelley M, Cheminade G, Tremillon N et al (2007) Chlorogenic acid synthesis in coffee: an analysis of CGA content and real-time RT-PCR expression of HCT, HQT, C3H1, and CCoAOMT1 genes during grain development in *C. canephora*. Plant Sci 172:978–996

Lu J, Zhao H, Wei J et al (2004) Lignin reduction in transgenic poplars by expressing antisense CCoAOMT gene. Prog Nat Sci 14:1060–1063

Marita J, Ralph J, Hatfield R et al (2003) Structural and compositional modifications in lignin of transgenic alfalfa down-regulation in caffeic acid 3-*O*-methyltransferase and caffeoyl coenzyme A 3-*O*-methyltransferase. Phytochemistry 62:53–65

Martz F, Maury S, Pinçon G et al (1998) cDNA cloning, substrate specificity and expression study of tobacco caffeoyl-CoA 3-*O*-methyltransferase, a lignin biosynthetic enzyme. Plant Mol Biol 36:427–437

Meyermans H, Morreel K, Lapierre C (2000) Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl -coenzyme A *O*-methyltransferase, an enzyme involved in lignin biosynthesis. J Biol Chem 47:36899–36909

Osakabe Y, Ohtsubo Y, Kawai S (1995) Structure and tissue-specific expression of genes for phenylalanine ammonia-lyase from a hybrid aspen, *Populus kitakamiensis*. Plant Sci 105:217–226

Pinçon G, Maury S, Hoffmann L et al (2001) Repression of *O*-methyltransferase genes in transgenic tobacco affects

lignin synthesis and plant growth. Phytochemistry 57:1167–1176

Rastogi S, Dwivedi U (2008) Manipulation of lignin in plants with special reference to *O*-methyltransferase. Plant Sci 174:264–277

Robinson AJ, Love CG, Batley J et al (2004) Simple sequence repeat marker loci discovery using SSR primer. Bioinformatics 20:1475–1476

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Streetman L (1963) Reproduction of the lovegrasses, the genus Eragrostis: Eragrostis chloromelas Stend., E. curvua (Schrad.) Nees., E. leshmanniana Nees. and E. superba Peyr. Wrightia. Am J Bot 3:41–51

Takezaki N, Rzhetsky A, Nei M (2004) Phylogenetic test of the molecular clock and linearized trees. Mol Biol Evol 12:823–833

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Tsuruta S, Ebina M, Nakagawa H et al (2007) Isolation and characterization of cDNA encoding cinnamyl alcohol dehudrogenase (CAD) in sorghum (*Sorghum bicolor* (L.) Moench). Grassland Sci 53:103–109

Vogel K, Jung H (2001) Genetic modification of herbaceous plants for feed and fuel. Crit Rev Plant Sci 20:15–49

Ye Z, Kneusel R, Matern U et al (1994) An alternative methylation pathway in lignin biosynthesis in *Zinnia*. Plant Cell 6:1427–1439

Ye ZH, Zhong R, Morrison WH et al (2001) Caffeoyl coenzyme A *O*-methyltransferase and lignin biosynthesis. Phytochemistry 57:1177–1185

Zhao H, Sheng Q, Lü S (2004) Characterization of three rice CCoAOMT genes. Chin Sci Bull 49:1602–1606

Zhong R, Morrison H, Negrel J et al (1998) Dual methylation pathways in lignin biosynthesis. Plant Cell 10:2033–2045

Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ (eds) Evolving genes and proteins. Academic Press, New York, pp 97–166