CrossMark

ORIGINAL PAPER

Spatial and temporal expression of cytosine-5 DNA methyltransferase and DNA demethylase gene families of the *Ricinus communis* during seed development and drought stress

Drosou Victoria[1] · Kapazoglou Aliki[1] · Koidou Venetia[1] · Merkouropoulos Georgios[1] · Hilioti Zoe[1]

Abstract Cytosine methylation plays a key role in genome stability and gene expression regulation. The dynamic and reversible nature of cytosine methylation is orchestrated by the coordinated and opposing action of two different gene families: the cytosine-5 DNA methyltransferases (C5-MTases) and the DNA demethylases. Here, we identified and characterized eight C5-MTases and three DNA demethylase family members in *Ricinus communis* (castor plant), an oilseed species of the Euphorbiaceae family. Bioinformatic analysis of the protein primary structure revealed the presence of conserved functional domains defining each family. The amino acid identities between C5-MTases and DNA demethylases from castor plant and their protein homologues from *Jatropha curcas* and *Manihot esculenta*, were 65 and 85%, respectively. Intriguingly, protein localization predictions suggested multiple targeting to nucleus, mitochondria and chloroplasts for some methyltransferases. Tissue and developmental gene expression analysis, with emphasis in seed maturation stages, showed that all family members varied widely in their expression across tissues suggesting distinct biological roles. In silico promoter analysis revealed several *cis* acting elements associated with drought responses indicating an epigenetic mechanism of regulation for the drought response by means of a methylation/demethylation switch. Consequent quantitative PCR analysis revealed that although C5-MTase genes were differentially expressed under drought stress, all the demethylase genes analyzed in this study (*RcDME, RcDML-3* and *RcROS1*) were significantly upregulated suggesting fine-tuning of DNA methylation and demethylation events. The detection and organization of putative transposable elements (TEs) along promoter regions of most of the genes indicates a possible role in transcriptional control. The current results set the foundation for functional studies of the C5-MTase and DNA demethylase gene families in castor plant.

✉ Hilioti Zoe
   zhilioti@certh.gr

   Drosou Victoria
   vdrossou@hotmail.com

   Kapazoglou Aliki
   akapazoglou@gmail.com

   Koidou Venetia
   vkoidoup@gmail.com

   Merkouropoulos Georgios
   georgios.merkouropoulos@gmail.com

[1] Institute of Applied Biosciences (INAB), CERTH, Thermi, Thessaloniki 57001, Greece

Introduction

DNA methylation is a conserved epigenetic mechanism in eukaryotes by which C5-MTases catalyze the transfer of a methyl group to the fifth carbon of a cytosine base of DNA to form 5-methylcytosine while DNA demethylases remove the 5-methylcytosine by a base excision mechanism. In plants, DNA methylation occurs in all three sequence contexts: the symmetric CG and CHG contexts (where H denotes C, A or T) and the asymmetric CHH context (Chan et al. 2005). The genome-wide levels of CG, CHG and CHH methylation are 24, 6.7 and 1.7% in *Arabidopsis thaliana*

🖄 Springer

(Cokus et al. 2008). DNA methylation in plants has been implicated in defending genome stability by repressing the activity of invading and mobile DNA elements such as transgenes, transposons, retroelements and viruses, as well as by regulating gene expression during development and in response to stress (Lisch 2013; Zhang et al. 2010).

The distribution and abundance of DNA methylation in plant genomes are regulated by the action of two different and antagonizing classes of enzymes: cytosine-5 DNA methyltransferases (C5-MTases) and DNA demethylases. In plants, C5-MTases are classified into four main families: methyltransferases (METs) (orthologous to mammalian *Dnmt1*), chromomethyltransferases (CMTs), de novo domains-rearranged methyltransferases (DRMs) (orthologous to *Dnmt3a/Dnmt3b*) and the DNA methyltransferase homologue 2 (*Dnmt2*). DNA METHYLTRANSFERASE1 (MET1) family establishes and maintains CG methylation (Chan et al. 2005), whereas CHROMOMETHYLASES CMT2 and CMT3 in cooperation with DOMAINS REARRANGED METHYLASE DRM2 shape the non-CG methylation pattern of the plant genome (Stroud et al. 2014; Lindroth et al. 2001), in coordination with H3K9 methylation, methylate CHH and CNG sites, respectively (Du et al. 2012; Matzke and Mosher 2014; Stroud et al. 2014). *DRM2* regulates non-CG methylation and is being guided to specific targets by an RNA-directed DNA methylation pathway that involves 24-nt small interfering RNAs (Law and Jacobsen 2010).

Four DNA demethylase genes have been identified in *A. thaliana; DEMETER (DME), REPRESSOR OF SILENCING 1 (ROS1), DEMETER LIKE 2 (DML2)* and *DEMETER LIKE 3 (DML3)*. These DNA glycosylase enzymes remove 5-methylcytosine from its targets through a base excision repair mechanism and replace it with an unmethylated cytosine (Morales-Ruiz et al. 2006; Gong et al. 2002; Gehring et al. 2006). DME was shown to be expressed mostly in central cells of female gametophytes where it was involved in maternal allele demethylation and gene imprinting in the endosperm (Choi et al. 2002; Kinoshita et al. 2004). DME also plays a role in the genomic demethylation of vegetative nuclei in developing pollens (Schoft et al. 2011). In other species of agronomical relevance such as barley, wheat and tomato, *DME* transcripts were found present in later stages of seed development and in response to abiotic stress (Kapazoglou et al. 2013a; Wen et al. 2012; Liu et al. 2015) ROS1, DML2 and DML3 are expressed in vegetative tissues where they prevent the DNA hypermethylation at genomic regions in close proximity to TEs (Yamamuro et al. 2014; Penterman et al. 2007). It has been proposed that ROS1 may be targeted to specific sequences by REPRESSOR OF SILENCING 3 (ROS3), a protein that binds small single-stranded RNAs suggesting an RNA directed DNA demethylating mechanism (Zheng et al. 2008). Mutations in the DNA demethylases genes cause increased DNA methylation in all sequence contexts at specific genomic loci (Law and Jacobsen 2010; Ortega-Galisteo et al. 2008).

Euphorbiaceae is a large family of the order of Malpighiales, composed of over 6300 species. It includes crops of high economic importance with numerous industrial applications such as castor plant (*R. communis*), cassava (*Manihot esculenta*), rubber tree (*Hevea brasiliensis*) and physic nut *(Jatropha curcas)*. *Ricinus communis*, one of the most important non-edible oilseed crops, produces seeds that accumulate ricinoleic acid (Kapazoglou et al. 2013b; Severino et al. 2012), along with a great number of additional metabolites (Merkouropoulos et al. 2016) that could be used in chemical industry. In the context of the challenges faced by breeders and agriculturalists worldwide to develop new plant varieties, the understanding of the phenotypic, genetic and epigenetic diversity of *R. communis* is important. Previous investigations in this species have shown low genetic diversity among castor plant germplasm (Allan et al. 2008; Foster et al. 2010) making epigenetic screening approaches necessary in the quest of new diversity. To this end studies in related species such as *Jatropha* (Yi et al. 2010), cassava (Xia et al. 2014) and rubber tree (Uthup et al. 2011) detected significant epigenetic diversity. The epigenetic state of the genome may largely affect plant physiology influencing a range of agronomical characteristics such as yield and environmental adaptation (Kapazoglou et al. 2012; Tsaftaris et al. 2012). Thus, the study of epigenetic marks and the respective epigenetic modifiers could unravel valuable epigenetic diversity for breeding purposes in this important industrial crop.

Members of C5-MTases and DNA demethylases have been identified in several crop plants such as tomato (Cao et al. 2014; Liu et al. 2015), legumes (Garg et al. 2014), rice (Ahmad et al. 2014; Sharma et al. 2009) and maize (Qian et al. 2014), as well as in the early land plant *Physcomitrella patens* (Malik et al. 2012). Continuing our involvement on the study of *R. communis* (Merkouropoulos et al. 2016), we sought to identify and characterize the *R. communis* C5-MTase and DNA demethylase gene families to better understand the molecular mechanisms of epigenetic regulation. The phylogenetic relationship among various types of C5-MTases and DNA demethylases has been inferred. Gene expression analysis of C5-MTases and DNA demethylases in various tissues during development and also upon the application of stress was performed in order to reveal putative functions of these enzymes.

Materials and methods

Identification of *R. communis* C5-MTase and DNA demethylase genes

The protein sequences of ten C5-MTases and four DNA demethylases in *A. thaliana* were downloaded from the National Center for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/) and were used to search for *R. communis* related sequences using the Blastp tool of the Phytozome database (http://www.phytozome.net/). The open reading frame (ORF) length, number of exons and the protein length were retrieved from the Phytozome database. The protein molecular weight (Mw) and isoelectric point (pI) were determined using the Expasy software (http://web.expasy.org/compute_pi/). The protein domains were characterized using the Smart *(http://smart.embl-heidelberg.de/) and Prosite tools (http://prosite.expasy.org/), while motifs were detected using MEME (version 4.9.0) (http://meme-suite.org/).*

Phylogenetic analysis, gene structure analysis and protein localization predictions

Phylogenetic relationships were determined using the neighbor-joining algorithm in the MEGA 6.0 (Tamura et al. 2013). Bootstrap analysis was performed using 5000 replicates with the pairwise deletion option. Amino acid % identity and protein alignments were determined using Protein blast at NCBI (http://blast.ncbi.nlm.nih.gov) and ClustalX. The gene structure analysis was determined using GSDS (http://gsds.cbi.pku.edu.cn/) by performing the comparison of the genomic sequences and coding sequences (CDS). The SWISS-MODEL structure homology-modelling server (https://swissmodel.expasy.org) was used to determine the protein three dimensional models and to generate a PDB file for each protein. The structure-based phylogeny was inferred by using root mean square deviation (RMSD) metric of structure homology in the VMD software. The localization prediction programs LOCALIZER and cNLS Mapper available on the Internet (see http://localizer.csiro.au, http://nls-mapper.iab.keio.ac.jp), were selected to provide localization predictions for cellular compartments.

Plant materials

Ricinus communis plants were grown in the experimental field of the Institute of Applied Biosciences (INAB), Thessaloniki, Greece. For gene expression studies, roots, leaves, stems, cotyledons, shoot apical meristem (SAM) and seeds at three developmental stages (S1-approximately 30 days after fertilization-DAF; S2 approximately 40–50 DAF; and S3-approximately 50–60 DAF) were collected. The criteria adopted to discriminate the seed stages were based on morphologic characteristics of the seed such as size and colour (Greenwood and Bewley 1982). Seeds < 1 cm in length and not fully expanded were assigned as stage S1, fully expanded and non-pigmented seeds were defined as stage S2 seeds, whereas pigmented seeds were grouped in stage S3. Tissues were quickly frozen in liquid nitrogen and stored at −80 °C until further use.

Drought stress experiment

Drought experiment was carried out in a growth chamber under 16/8 h light/dark cycle and 60% humidity. Healthy plants growing in individual pots in an automated growth chamber were chosen for the drought treatments. All plants were at the same developmental stage (emergence of the third true leaf). Individual plants were transplanted in large pots with each pot containing four plants. The pots were well watered in order to ensure successful plant establishment. Thereafter, water was provided every 5 days only to two of the four pots leaving the other two pots under a water deficit regime. Soil moisture sensors were placed in each pot to monitor changes in soil water content. Three weeks later, leaf samples were collected from the control and drought-stressed plants, quickly frozen in liquid nitrogen and stored at −80 °C. Each sample was taken from the lateral lobes of the third leaf. The mean soil moisture in the pots with the control plants was maintained at 70% compared to 41% in the pots with the stressed plants.

Expression analysis

Total RNA was extracted from different tissues and drought-stressed leaves using the NucleoSpin RNA Plant kit (Macherey–Nagel Co., Duren, Germany) according to the instructions of the manufacturer. Genomic DNA was removed from RNA preparations by digestion with DNase I (Macherey–Nagel Co., Duren, Germany) according to the manufacturer's protocol. RNA quantity and quality was confirmed by spectrophotometry (Thermo Fisher Scientific, model: NanoDrop™ 1000) and gel electrophoresis. First strand cDNA synthesis was performed using the SensiFAST™ cDNA Synthesis Kit (Bioline Reagents Ltd.) according to the specifications of the manufacturer. The cDNA products were diluted (1:5) and stored at −80 °C. Semi-qualitative RT-PCR reactions were performed in 20 μL volume using 1 μL of the diluted cDNA as template and KAPA Taq DNA Polymerase (Kapa Biosystems, Woburn, MA, USA) employing the following protocol: (i) 94 °C for 3 min, (ii) 25 to 35 cycles at 94 °C for 30 s, 52–60 °C for 30 s, and (iii) 72 °C for 30 s. The *R. communis NDUB8* (*NADH Dehydrogenase [Ubiquinone] 1 Beta Subcomplex Subunit 8*) gene was used as endogenous reference control. For real-time PCR, each

sample reaction was set up in a PCR reaction mix (20 μl), containing 1× buffer, 0.2 mM dNTPs, 0.2 mM forward and reverse primers, 1.5 mM Syto®9, 0.5U Kapa Taq DNA polymerase and 1 μl of the 1:5 diluted cDNA. Reactions were performed with a Corbett Rotor Gene 6000 Thermocycler (Corbett Research, Sydney, Australia). Two biological repeats were used and three technical replicates were performed for each one. General thermocycler conditions were: (i) 95 °C for 3 min, (ii) 45 cycles at 95 °C for 20 s, 53 °C for *RcDME, RcDRM3, RcDML-3, RcROS1, RcCMT2*, or 57 °C for *RcMET1, RcMET2, RcCMT1, RcDRM1, RcDRM1, RcDRM2* for 20 s, 72 °C for 20 s, and (iii) at 72 °C for 7 min. Relative quantification was performed using *NDUB8* as the reference gene. The data was analyzed using the REST software (Pfaffl et al. 2002). All primers used in expression analysis correspond to non-conserved regions (Supplementary Table S1).

In silico promoter analysis of C5-MTase and DNA demethylase genes

To identify potential *cis*-acting regulatory elements present in the promoter regions of the C5-MTase and DNA demethylase genes, nucleotide sequences of 1000 bp upstream regions from the translational start codon (ATG) were retrieved from the Phytozome database. In silico promoter analysis was carried out using the PLACE database (http://www.dna.affrc.go.jp/PLACE/signalscan.html). Detection of TEs was carried out utilizing the CENSOR software tool in the 2000 bp region located upstream of the ATG translation start codon (http://www.girinst.org/censor/).

Results

Identification and structural analysis of *R. communis* C5-MTases and DNA demethylases

A search for the identification of *R. communis* homologous genes to *A. thaliana* C5-MTases and DNA demethylases was performed using *AtDME, AtMET1, AtDRM1, AtCMT1* and *AtDnmt2* as queries. In total, eight C5-MTases and three DNA demethylases were identified in *R. communis* with e-values ranging from $5.2e^{-87}$ for *RcDRM3* to $8.1e^{-166}$ for *RcDML3* and $4.6e^{-171}$ for *RcDRM1*. As shown in Supplementary Table S2, the length of the C5-MTases open reading frame varied from 4755 bp for *RcMET1−2 to* 1215 bp for *RcDnmt2*, while for DNA demethylases the open reading frame varied from 5631 bp for *RcDME* to 4905 bp for *RcROS1*. The exon–intron structure of the eight C5-MTase and three DNA demethylase genes was studied (Supplementary Fig. S1). As for C5-MTases, *RcMET1-1* and *RcMET1-2* genes had similar structure with the exception of two small

inserted introns at the 5′ and 3′ end of *RcMET1-2*. The gene structure of the other members of the C5-MTase family was more variable. As for the *R. communis* DNA demethylases, all three showed comparable structure with 19 exons. Overall, the gene structure analysis revealed that the exon–intron structure considerably varies in the C5-MTase family in contrast to the consistent organization in the DNA demethylase family.

The proteins encoded varied from 404 to 1584 amino acids for C5-MTases and 1634 to 1876 amino acids for DNA demethylases. The isoelectric point varied from 5.24 to 9.20 for C5-MTases, and from 6.88 to 6.90 for DNA demethylases.

Structural analysis revealed that the eight C5-MTases are divided into four groups on the basis of the diverse structural features comprising their primary structure: RcMET1-1 and RcMET1-2 comprise the METs group, RcCMT-1 and RcCMT2 comprise the CMT group, RcDRM1, RcDRM2 and RcDRM3 comprise the DRM group, whereas RcDnmt2 is the sole member in the fourth group. All members of the C5-MTase family possess the carboxyl terminal catalytic domain with the conserved motifs I, IX, VI, VIII, IV and X aligned in a specific order, while the amino terminal domain is more variable having diverse structural features among different members of the family (Fig. 1). *Ricinus communis* contains a single putative *RcDnmt2* gene that codes for a small protein similar to prokaryotic MTases that possess only the MTase domain and lacks the amino terminal domain. Two RcMETs were identified harbouring two bromo adjacent homology (BAH) domains and two replication foci domain (RFD). Two CMTs were also identified in *R. communis* characterized by the presence of a single BAH domain and a conserved chromodomain inserted between motifs I and IV in the C5-MTase catalytic domain. In their N-terminus, the three DRM proteins identified, *RcDRM1, RcDRM2* and *RcDRM3*, contain a ubiquitin associated (UBA) domain, which is absent from the rest of the C5-MTase family members, while there is a circular permutation of motifs in the cytosine methyltransferase domain with motifs VI through X preceding the motifs I-IV. *RcDRM3* lacks the motif IV, which contains the invariant prolylcysteinyl doublet that has been identified as the functional active site of all known C5-MTases (Pavlopoulou and Kossida 2007; Cao et al. 2000). *RcDRM3* also lacks glutamic acid in motif IX and glycine in motif X similar to *Arabidopsis* and maize DRM3, pointing to a catalytically mutated DRM. *RcDRM3* possesses two UBA domains, although the second UBA domain showed substitution of the conserved glycine residue (MGF/MGY) for asparagine (Supplementary Fig. S2), which was predicted to abolish proper folding of UBA domain (Mueller and Feigon 2002).

Ricinus communis DNA demethylases have a common structure (Fig. 1). They harbour a DNA glycosylase
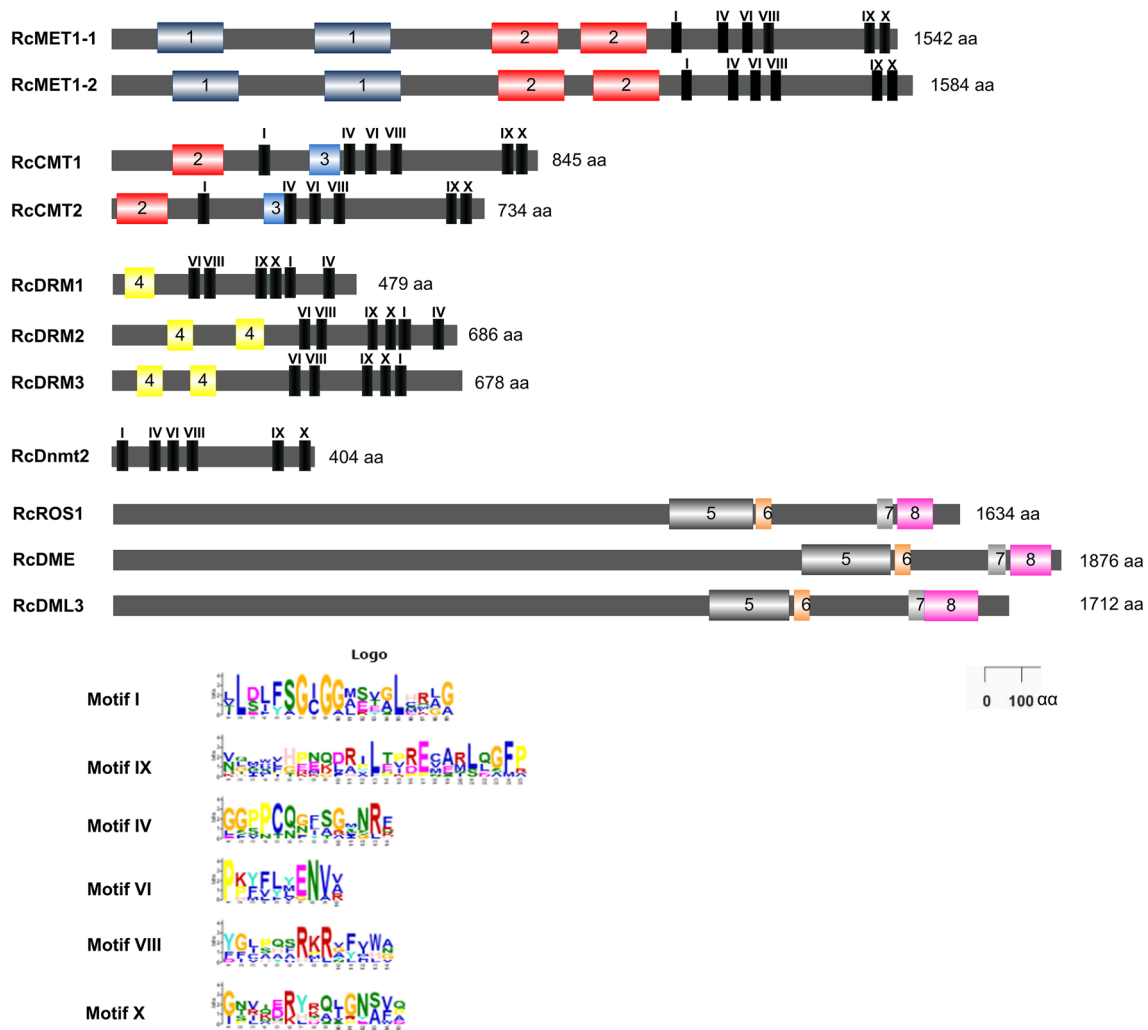
Fig. 1 Schematic representation of domains and motifs found in *R. communis* C5-MTase and DNA demethylase proteins. Each domain is represented by a colored box. 1: DNMT1-RFD domain; 2: BAH domain; 3: chromodomain; 4: ubiquitin-associated domain; 5: HhH- GPD domain; 6: FES domain; 7: Perm-CXXC domain; 8: RRM-DME domain. The numbers on the right denote the length of each protein, scale 100 amino acids

domain that includes the helix–hairpin–helix motif and a glycine–proline rich region flanked by a conserved aspartate (HhH-GPD) and the EndIII 4Fe-4Se domain. At the C-terminus of DNA demethylases there is a Perm-CXXC domain which is followed by the RRM DME domain that is present in the family of DEMETER proteins and facilitates the interaction of the catalytic domain with ssDNA or regulatory RNA (Iyer et al. 2011).

Phylogenetic, structural and subcellular localization analysis of *R. communis* C5-MTase and DNA demethylase genes

To examine the phylogenetic relationships among the members of the C5-MTase and DNA demethylase families identified in *R. communis*, two phylogenetic trees were constructed (one for each family) based on the alignments of full-length protein homologue sequences of maize, rice, tomato, *Arabidopsis*, cassava, *Jatropha* and soybean (Fig. 2a, b). As for C5-MTases, the generated phylogenetic tree formed four distinct clades; one for each of the four subfamilies (METs, CTMs, DRMs, and Dnmt2) (Fig. 2a). This classification was consistent with the domain analysis described in Fig. 1. On the C5-MTase tree, the *R. communis* proteins appeared to be very closely related to the protein homologues of the *Jatropha* and cassava plants, which also belong to the Euphorbiaceae family along with the *R. communis* (Fig. 2a). Specifically, RcMET1-1 displays the highest similarity having 85 and 82% identity to MeMET1-1 and JcMET1-1, respectively, followed by RcDRM2 (82 and 83% identity with MeDRM2 and JcDRM2, respectively) (Supplementary Table S3). RcCMT1 shows high homology to

Fig. 2 Phylogenetic analysis of *R. communis* C5-MTases and DNA demethylases. The amino acid sequences of *Arabidopsis*, tomato, *Glycine max*, maize, rice, *Manihot esculenta, Jatropha curcas* and *R. communis* C5-MTases (**a**) and DNA demethylases (**b**) were aligned with ClustalX, and the phylogenetic trees were constructed using the neighbour-joining method in MEGA 6.0 software. The numbers at the nodes represent bootstrap values from 5000 replicates. The sequences used and their accession numbers are shown in Supplementary Table S5
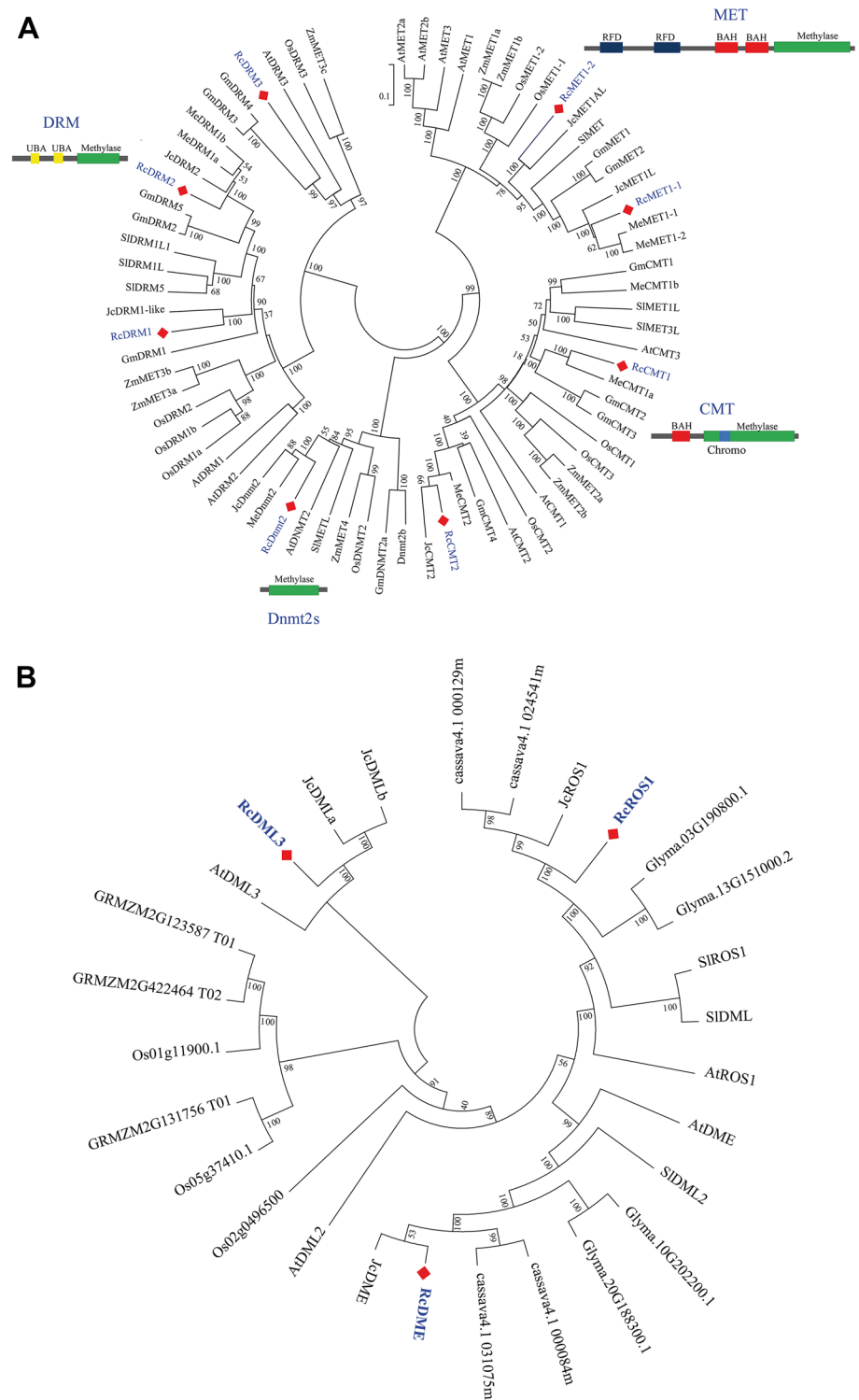
MeCMT1a, while RcCMT2 clusters together with MeCMT2 and JcCMT2. Likewise, RcDRM1 clusters together with the JcDRM1-like, RcDRM2 is found in the same clade with the MeDRM1a, MeDRM1b and JcDRM2, while DRM3 proteins cluster together in a distinct conserved clade. RcDnmt2 protein clusters together with both the JcDnmt2 and MeDnmt2

(Fig. 2a). Similarly, *R. communis* DNA demethylases have higher homology to orthologues from *Jatropha* and cassava (Fig. 2b). Overall protein identities ranged from 55 to 85% (Supplementary Table S3).

A bioinformatics analysis was performed in order to predict the subcellular localization of the 11 proteins. The

results suggested that all of cytosine-DNA modifying proteins carry nuclear localization signals (NLS), but two members (DRM1 and DRM3) of the DRM family were also predicted to carry mitochondria transit peptide (probability 0.79 for DRM1) and chloroplast transit peptide (probability 0.93 for DRM3; probability 0.8 for DRM1) (Fig. 3a).

The suggestion that DRM family proteins might be dual targeted to mitochondria and chloroplasts to regulate gene expression through DNA methylation modification in these tiny genomes led us to examine how the protein structures of this family evolved. The three-dimensional protein structures were predicted and the structural relationship of the proteins was determined based on the RMSD metric of structure

homology. The structure-based phylogeny of the family shows that DRM1 is the ancestral protein while DRM2 and DRM3 protein structures are closely (50% sequence identity) related (Fig. 3b).
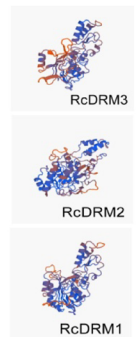
Spatial expression of C5-MTase and DNA demethylase genes during development

To gain insight into the function of C5-MTase and DNA demethylase genes in *R. communis*, their spatio-temporal expression patterns were analysed using semi-quantitative RT-PCR on RNA isolated from seeds (three developmental

Fig. 3 a Subcellular localization predictions of *R. communis* epigenetic modifiers. Nuclear localization signals (shown) were detected in all proteins while two members of the DRM family were predicted to localize also in bioenergetic organelles namely chloroplasts and mitochondria. **b** Structural information in RcDRM protein family evolutionary inference. Structure-based phylogenetic tree using three-dimensional protein structures predicted by SWISSMODEL and the structural relationship based on the root mean square deviation (RMSD) metric (values in Angstroms) of structure homology
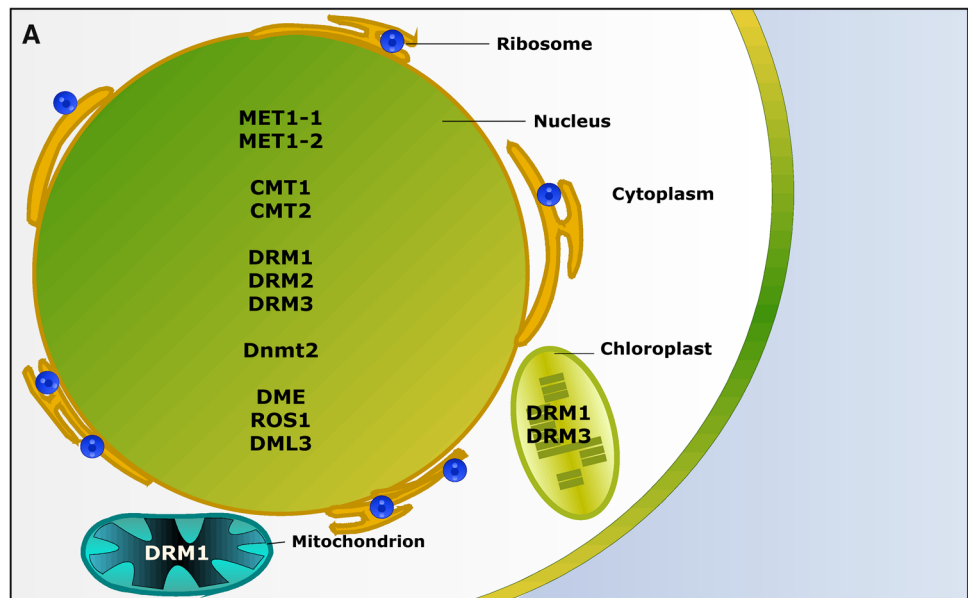
MET1-1:KRIR, RKKK, RKRK, KKNKGKPKSSVSKANKEQINNIKEKKRSSSQKNEEPAGSKKRPKA
MET1-2:KKKR, RKKK, KRGLSKNVENVYSSRKMTK
CMT1:RKYGTTPQSDFQKYIRLKK
CMT2:EQGKSKRPYARLWWDETVPTVVTYPSIHSQAM
DRM1:PCKKRKL, RKALEDCDGEPTLRAKKY, KKRKLHQNRSSKSQESKRK, KRKLHQNRSSKSQESKRKAK
DRM2:KKRNYAEYDMWKKKKQMK
DRM3:HRGKRPKQEY
Dnmt2:RQGLQKHSGDARAFSFLKILELIPHTKQPPH, RQGLQKHSGDARAFSFLKILELIPHTKQPPHM, EHISLRQRLSISFVGSMCMFDDAPLTVCRLAI
DME:PKRRK, PKQKRKY, KKRKPRP, KRRKHRPKVIVEGKPKKT, RRKHRPKVIVEGKPKKTPK
ROS1:KRRS, KKRSR, PTGKRKY, RKKRTKK, RRKKHRPKVITEGKPRPRK, RKKHRPKVITEGKPRPRKP
DML3:KKKK, KKKR, KRSGLVPGETRKTTK, KEYKRKRMRRKPKSK, KRKNVTKTQPKEKEKSK, RKRRVSSSLEEKEKRKNV, KRSRNGIDVDQRPLKRTKI, KKHRHRPKVAGAGRKRRVS

stages: S1, S2 and S3), root, stem, leaves, cotyledons and SAM (Fig. 4).

In seeds, six genes (*RcMET1-2, RcCMT2, RcDRM2, RcDRM3, RcDnmt2* and *RcDME*) were found to be expressed in all three seed stages analyzed and their expression level mainly reduced during seed development. Two genes (*RcDML3* and *RcROS1*) were found to be expressed at stages S1 and S2, whereas expression of another three genes (*RcMET1-1, RcDRM1* and *RcCMT1*) was detected either in the very early stage of the seed development (S1) or during seed development (S2). Through vegetative development, expression of eight genes (*RcMET1-2, RcCMT2, RcDRM2, RcDRM3, RcDnmt2, RcDML3, RcDME* and *RcROS1*) was detected in all the tissues examined although at different expression levels and with some noticeable expression peaks as in the case of *RcDRM3* in SAM and *RcDnmt2* in cotyledons. Expression of the remaining three genes (*RcMET1-1,*

RcCMT1 and *RcDRM1*) was not detected in the root, while their expression in the shoot was either not detectable or basal. Taken together, these results show that expression of the C5-MTases and DNA demethylases have distinct differences revealing a variable level of tissue specificity.

In silico promoter analysis of C5-MTase and DNA demethylase genes

A profile of the *cis*-acting regulatory elements present in the promoters is important for understanding the mechanism that regulates gene expression during plant development and under different environmental conditions. The 5′ upstream regions of C5-MTase and DNA demethylase genes were analyzed using the PLACE database for putative regulatory *cis*-acting elements (Supplementary Table S4). Totally, eight types of stress-related *cis*-elements were detected, including

Fig. 4 **a** Expression profiles of *R. communis* C5-MTase and DNA demethylase genes in various tissues. Semi-quantitative RT-PCR analyses on RNA isolated from *R. communis* seeds of stages S1, S2 and S3 as well as root (R), stem (S), leaves (L), cotyledons (C) and shoot apical meristem (SAM) were performed. *NDUB8* was used as an internal control. **b** Table summarizes the gene expression data from Fig. 4a based on band detection after DNA electrophoresis
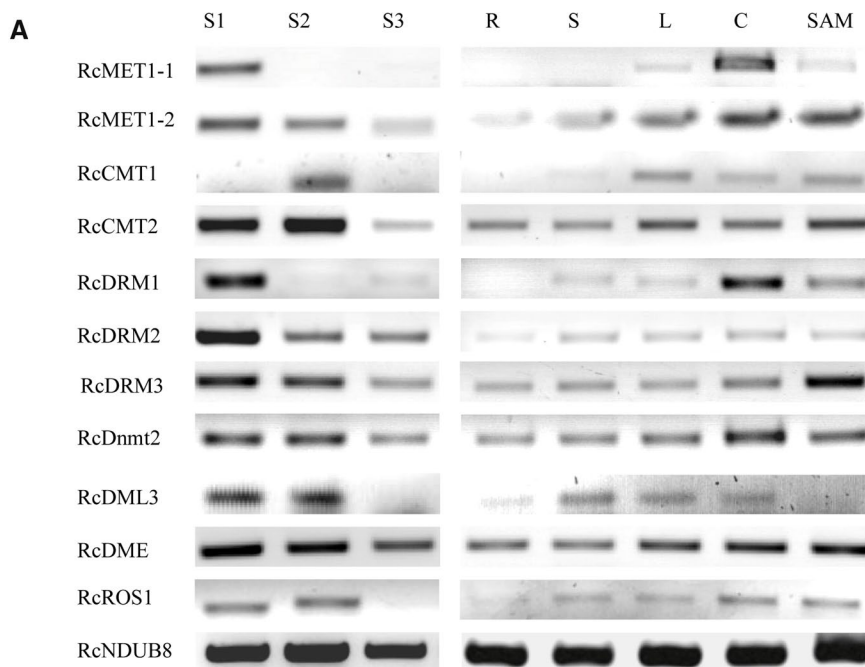
Gene	Seed stage 1	Seed stage 2	Seed stage 3	Root	Stem	Leaf	Cotyledon	Shoot apical meristem
RcMET1-1	◉					○	◉	○
RcMET1-2	◉	◉	○	○	○	○	◉	◉
RcCMT1		◉				○	○	○
RcCMT2	◉	◉	○	○	○	○	○	○
RcDRM1	◉		○		○	○	◉	○
RcDRM2	◉	◉	◉	○	○	○	○	○
RcDRM3	◉	◉	◉	○	○	○	○	◉
RcDnmt2	◉	◉	◉	○	○	○	◉	○
RcDML3	◉	◉		○	○	○	○	
RcDME	◉	◉	◉	◉	◉	◉	◉	◉
RcROS1	◉	◉		○	○	○	○	○
RcNDUB8	◉	◉	◉	◉	◉	◉	◉	◉

◉ Enriched   ○ Expressed

the dehydration-responsive elements (DRE), MYB binding sites involved in drought-inducibility (MBS), heat shock elements (HSE), low temperature-responsive elements (LTR), defence and stress-responsive elements (TC-rich repeats), EIRE (elicitor-responsive element), fungal elicitor responsive elements (BOX-W1) and an essential element for the anaerobic induction [ARE (anaerobic-responsive element)]. In addition, elements that are associated with plant responses to hormones were found such as abscisic acid [ABRE (ABA responsive element)], gibberellins (GARE-motif, TATC-box, P-box), ethylene [ERE (ethylene-responsive element)], methyl jasmonate (MeJA) (CGTCA-motif and TGACG-motif), auxin (TGA-element; AuxRR-core) and salicylic acid (SA) (TCA-element). Two seed-specific expression *cis*-motifs (Skn-1 motif, GCN4 motif) were conserved in the promoter regions of some C5-MTase and DNA demethylase genes. A great number of light responsive *cis*-acting regulatory elements, such as Box-4, Box-I, G-box, were found in all the promoter regions of the C5-MTase and DNA demethylase genes. Mapping the main *cis*-acting elements within the 1000 bp sequences upstream of the predicted start codon of the genes (Supplementary Fig. S3) allowed the emergence of several patterns: (i) each gene family member has a subset of unique motifs not present in other members suggestive of differential responsiveness and regulation of transcription, (ii) the CCAAT-box motif is localized to positions approximaly −1000 upstream of the transcriptional start in *MET1-1* and *CMT2* promoters, (iii) all DRM genes contained the seed-specific Skn1 motif in their promoters, (iv) unlike DRM2 and DRM3 genes, DRM1 lacks LTR retrotransposon and DNA transposon elements, (v) DRM1 and DRM3 posses abiotic stress-related motifs including ABRE and ERE. Collectively, the results underscore the diverse roles of C5-MTase and DNA demethylase genes in castor plant development and responses to biotic and abiotic stresses.

Drought-specific expression of C5-MTase and DNA demethylase genes

The expanding cultivation of castor plant will inevitably be challenged by the climate change to dry conditions worldwide. Since our in silico analysis located putative drought-related elements on the C5-MTase and DNA demethylase gene promoters, we examined the induction of the genes under water deficit conditions by subjecting castor plants to drought stress for a 3-week period. Under such stressful conditions, the phenotypes of the plants were severely affected. Specifically, the leaf surface area, stem height and internode length were severely reduced, in striking contrast to control plants (Fig. 5a). At a molecular level, prolonged stress signals may affect plant morphology through a global transcription reprogramming orchestrated by epigenetic modifiers of chromatin (Fig. 5b). To examine

the effects of drought stress on *R. communis* epigenetic modifiers such as C5-MTase and DNA demethylase genes, quantitative RT-PCR analysis was performed using leaves from *R. communis* that had been subjected to drought (Fig. 5c). Specifically, *RcMET1-1* showed no significant change in expression in leaves under drought stress conditions, whereas significant induction in transcript levels of approximately threefold, was observed for the *RcMET1-2* gene. A severe reduction in transcript levels was evidenced for *RcCMT1* and *RcCMT2* of about 3.4-fold and 3.6-fold, respectively. *RcDRM2* showed a small reduction by 1.4-fold whereas *RcDRM1* and *RcDRM3* do not exhibit significant changes in gene expression upon drought. *RcDnmt2* also displays a significant reduction in transcript levels of about twofold in response to drought. With regard to the DNA demethylase genes, all three *RcDME, RcDML3* and *RcROS1* demethylase genes showed significantly marked induction after drought of approximately threefold, 3.5-fold and fourfold, respectively (Fig. 5d). These results demonstrated that *R. communis* C5-MTase and DNA demethylase genes were differentially expressed under drought stress treatment pointing to distinct roles regarding abiotic stress responses.

Analysis of transposable elements within the promoters of C5-MTase and DNA demethylase genes

TEs are repeated DNA sequences that have a profound effect on genomic structure, evolution and gene regulation (Bennetzen and Wang 2014). In order to gain insight into the distribution of TEs along the C5-MTase and DNA demethylase promoters, we used bioinformatic tools for identification and characterization of TEs in the 2000 bp region located upstream of the ATG translation start codon. We found that there is a widespread contribution of TE sequences to C5-MTase promoters (Fig. 6; Supplementary Table S6). The TEs length ranged from 37 to 1183 nucleotides with the LTR retrotransposons being the most highly represented. TEs are abundant within the promoters of the C5-MTase genes with the exception of the *RcDnmt2* that lacks any TE element. Interestingly, the largest portion of the *RcCMT2* promoter is occupied by a large copy TE element, whereas *RcDRM1* promoter lacks any TE element near the ATG. As for DNA demethylases, all three promoters are mainly devoid of LTR transposons and the identified TEs correspond to DNA transposons and non-LTR retrotransposons. A detailed analysis of TEs in relation to *cis*-elements present in the promoter region of 1000 bp upstream of the predicted start codon of the genes (Supplementary Fig. S3) provides additional clues about possible positional effects of TEs on *cis*-elements.
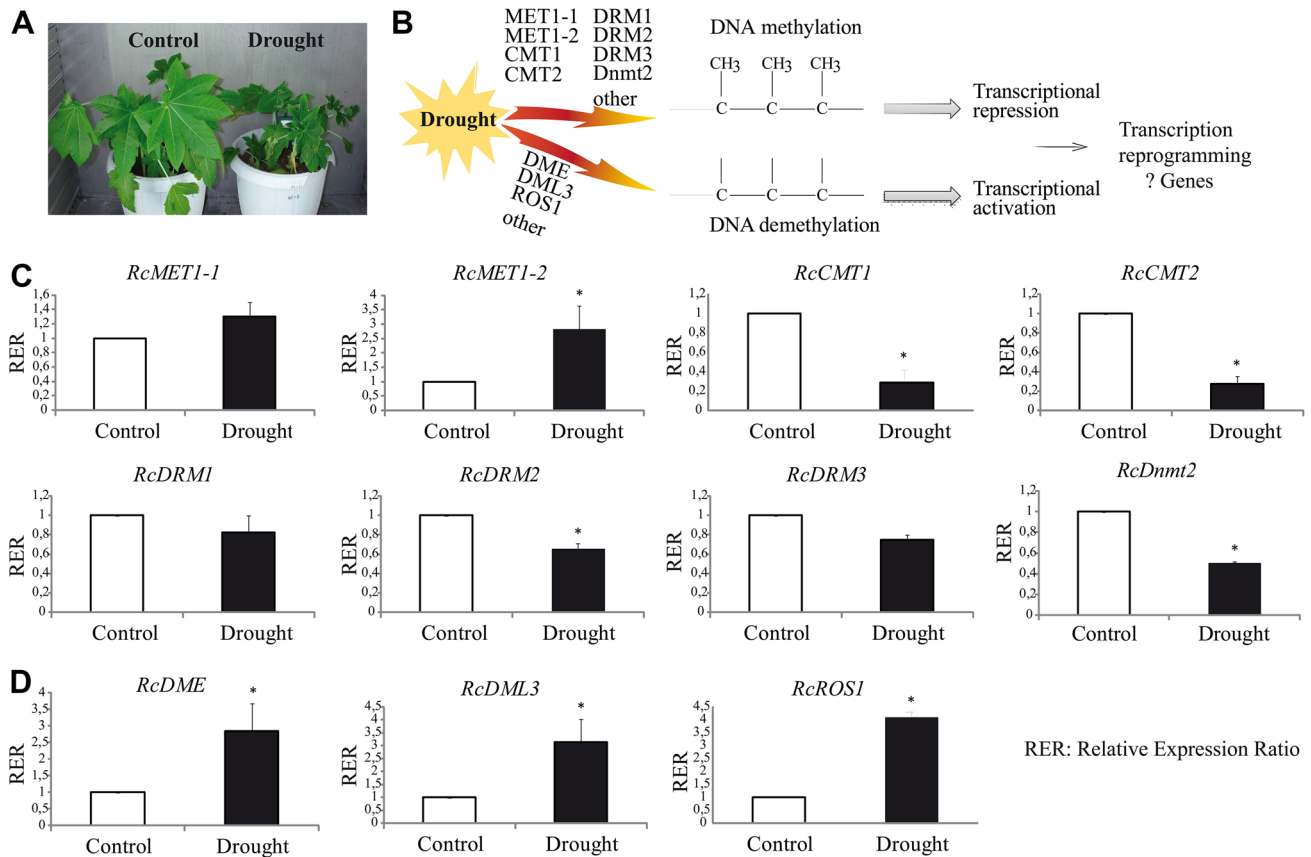
Fig. 5 a Visual comparison of morphological changes of control (left) and drought stressed (right) *R. communis* plants. **b** A schematic view of how epigenetic modifiers such as DNA methyltransferases and DNA demethylases may integrate drought stress signals to regulate global transcription reprogramming through DNA methylation/demethylation by their opposing actions. **c** Drought-induced gene expression analysis of *R. communis* C5-MTases. **d** Drought-induced gene expression analysis of DNA demethylases. Quantitative RT-PCR analyses with RNA isolated from leaves of castor plants sub-jected to drought stress. *RcNDUB8* was used as the endogenous control. White bars: control untreated plants; black bars: drought-treated plants. Expression values were normalized to those of *RcNDUB8*. Relative expression ratio of each sample was compared to the control group which was assigned the value of 1. Data represent mean values from two independent experiments with standard deviations. Values significantly different (P < 0.05) from the untreated plants are marked with an asterisk (*)

Discussion

There is an increasing interest concerning epigenetic variation in plants and its effect on quantitative traits and phenotypes as it would constitute ultimately a molecular tool for the selection of appropriate genotypes for crop improvement strategies. Epigenetic variants among individuals of the same species can be used in breeding programs. Epigenetic states can be regulated by C5-MTases and DNA demethylases which in turn are regulated by developmental and environmental cues.

The current results show that eight C5-MTases are present in the *R. communis* genome, which is fewer than those in *Arabidopsis* (10), rice (10) and soybean (13), and equal to the maize counterparts (8). Based on C5-MTase domain conservation, *R. communis* C5-MTases are classified into four distinct classes, similar to *Arabidopsis*, legumes and cereals. Specifically, castor plant has two members of the MET subfamily, two members of the CMT subfamily, three DRMs and one Dnmt2. In castor plant, three DNA demethylase genes were identified, while in *Arabidopsis* and tomato there are four and three DNA demethylase encoding genes, respectively. Furthermore, phylogenetic analysis indicated that castor plant C5-MTases and DNA demethylases are more closely related to *Jatropha* and cassava (amino acid identity ranges from 55–85%) consistent with the evolutionary relationships among these species which all belong to the Euphorbiaceae family. Subcellular localization predictions of the 11 *R. communis* epigenetic modifiers revealed that while all proteins contain NLS, two members of the DRM family appear to localise also in bioenergetic organelles such as mitochondria and chloroplasts. Cytosine methylation of the mitochondrial and chloroplastic genomes has remained largely overlooked since they represent tiny

Fig. 6 Map of TEs of *R. communis* C5-MTase and DNA demethylase gene promoters. TEs were detected in the promoter regions of castor C5-MTase and DNA demethylase genes. Details about the TEs are included in Supplementary Table S6
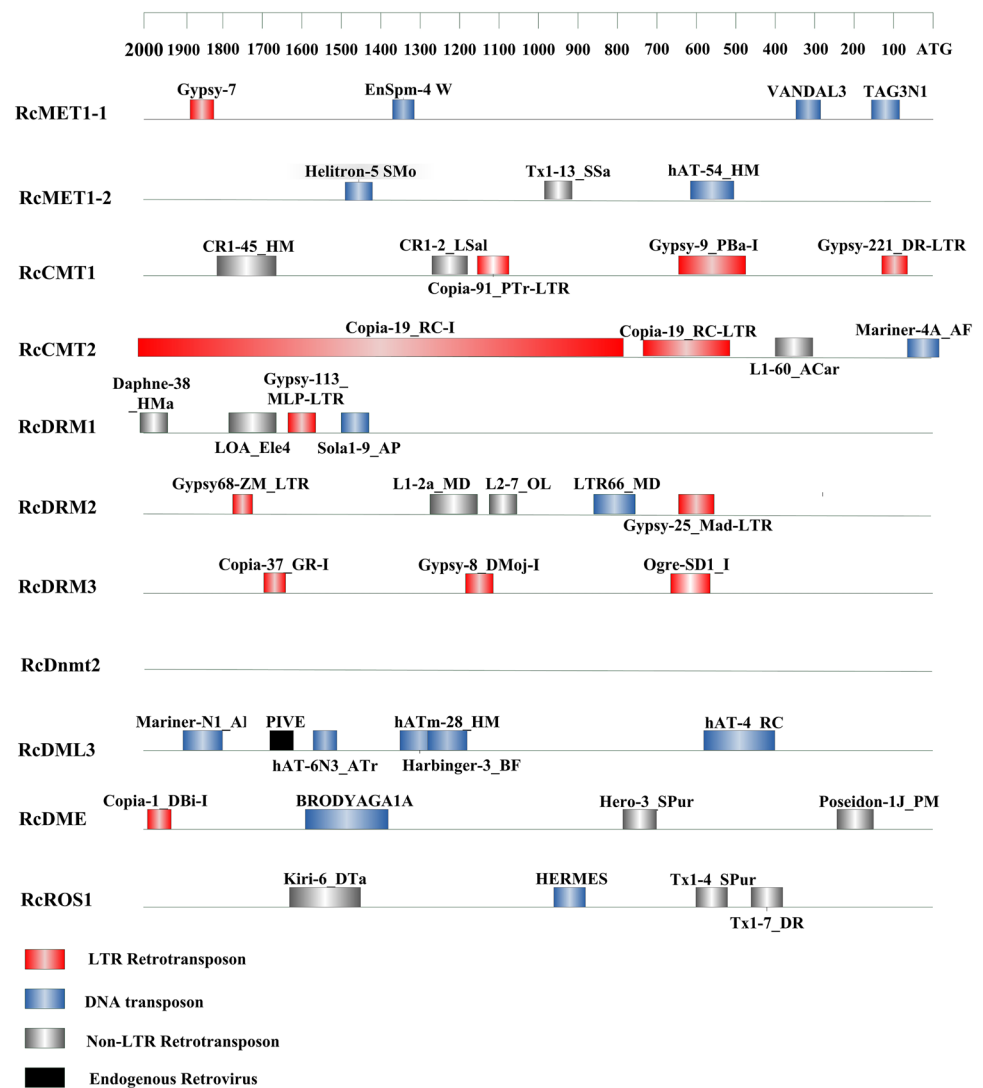
genomes. The current results raise the possibility that nuclear-encoded DRM proteins may have a second home in mitochondria or chloroplasts to modify gene expression in these organelles. Dual localization may enable some of these proteins to act as direct regulators for the coordinated expression of the mitochondrial, chloroplastic and nuclear genomes in response to environmental and developmental cues. Considering the DRM protein sequence and structure diversity found in *R. communis*, establishing the subcellular location of each protein will advance our understanding of their biological function. In the same context, a recent study (He et al. 2017) found that several methylated DNA fragments isolated from polymorphic methylated loci of castor plant accessions mapped in distinct regions of nuclear, mitochondrial and chloroplastic genomes.

During vegetative development most castor plant DNA C5-MTases were predominately expressed in tissues with actively dividing cells such as shoot apical meristem (SAM) and young cotyledons. This is in agreement with previous reports in *Arabidopsis*, legumes and cereals where METs and CMTs were induced in the SAM and young leaves (Garg et al. 2014; Qian et al. 2014; Sharma et al. 2009). In addition, *R. communis* DNA C5-MTases showed pronounced accumulation in S1 stage of seed development, where active endosperm cell proliferation and development is taking place (Baldoni et al. 2010). S1 is followed by the maturation phases S2, where protein and lipid accumulation begins, and S3, where protein and lipid storage bodies are fully formed (Baldoni et al. 2010). Presumably, during S1 of actively dividing cells and early endosperm development, specific gene expression programs have to be switched on and off and DNA methylation rearrangements take place along the genome. DNA C5-MTases are required to perform maintenance and establishment of DNA methylation marks at wider genomic areas or specific loci. *RcMET1-1* and *RcDRM1* show specific transcript accumulation in S1 seed, whereas they are nearly undetectable in the other two seed stages. Specific expression of *RcMET1-1* and *RcDRM1*

in S1 seeds implies specific functional roles of the corresponding enzymes at this particular seed stage. In addition elevated expression of *RcMET1-2, RcCMT2, RcDRM2* and *RcDRM3* may also suggest a functional role for these DNA C5-MTases in early seed development. Pronounced induction of *RcCMT2* at S2 of seed development may reflect the need for CHG maintenance of DNA methylation at chromatin regions associated with gene expression programs that promote the onset of the maturation phase. Furthermore, *RcMET1-1 and RcDRM1* are induced only in S1, whereas *RcMET1-2* and *RcDRM2* and *RcDRM*3 are present in all three stages of seed development implying functional diversification of gene members of the same family. Differential induction of DNA C5-MTase genes has also been evidenced during embryo and early seed development in *Arabidopsis*, legumes and cereals (Sharma et al. 2009; Garg et al. 2014; Qian et al. 2014). Downregulation of all *R. communis* DNA C5-MTases in seeds of S3 suggests a general decline in cell division and replication activities as well as gene expression activities during the later stages of seed development.

To determine factors that may regulate transcriptionally the specific epigenetic genes in castor plant, we analyzed their promoter sequences for putative *cis* elements within the 1 kb upstream of the ATG. Promoter bioinformatic analysis revealed seven types of stress-related *cis*-elements distributed along the promoters of the C5-MTase and DNA demethylase genes. Putative drought-related elements were located on the promoters prompting us to conduct a detailed investigation aiming to understanding the effect of drought stress on transcript regulation of the two family members. This investigation is relevant to plant challenges faced under current climate change which is characterized by long periods of drought stress in agricultural areas, a decrease in water availability and eventually loss of crop productivity.

The current results revealed differential regulation of transcript abundance of C5-MTase and DNA demethylase genes under drought stress conditions. Specifically, a general reduction of gene expression was observed for the DNA methyltransferases with the exception of *RcMET1-2*. Notably, a marked induction of three to four-fold was observed for all three DNA demethylases, *RcDME, RcDML3* and *RcROS*. This may imply a need for the erasure of methylation marks presumably from regulatory regions of drought-responsive transcription factors and other drought-associated genes, in order to drive drought-specific expression of important factors involved in survival under such abiotic stress conditions. Additionally, a reduction in DNA methylation is likely to favor activation of the transcription of transposons which is often accompanying stress conditions, further altering the global transcription pattern (Sanchez and Paszkowski 2014). The current results are in general agreement with previous studies in other species showing that C5-MTase

and DNA demethylase genes change expression patterns when responding to abiotic stresses (Garg et al. 2014; Qian et al. 2014).

C5-MTase and DNA demethylase promoters in castor plant are surrounded by different sets of TEs. Plant genomes contain a large number of TEs and it is now clear that TEs affect the transcription of nearby genes. A study in maize showed that the closer a TE sequence is to the transcriptional start site, the stronger the effect it has on the expression of the gene (Eichten et al. 2012). It has been proposed that regulation of TEs is mediated by DNA methylation, which can spread to flanking DNA and affect the expression of nearby genes (Gehring et al. 2009; Diez et al. 2014; Hollister and Gaut 2009). In *Arabidopsis*, methylation changes around promoter TEs regulate the expression of stress responsive genes (Le et al. 2014). It is striking that castor DNA demethylase promoters are devoid of LTRs and are enriched with DNA retrotransposons, whereas in castor genome the length covered by LTR elements accounts for about one-third of all repeats and DNA TEs constitute less than 2% (Chan et al. 2010). The different number, types and distribution of TEs in the promoter sequences of *R. communis* C5-MTase and DNA demethylase genes implies functional relationships between TEs and differential transcriptional regulation.

Conclusions

Gene expression patterns of high-order regulators of epigenetic state, namely C5-MTases and DNA demethylases, were studied during the seed maturation process and drought stress in *R. communis*. The primary structure of the eight C5-MTases and three DNA demethylases showed high-conservation in signature motifs found in homologues of other species. At early stages of seed development, nearly all genes were transcriptionally activated whereas at late stages their transcription ranged between low and undetectable levels. Under drought stress, the expression of C5-MTases varied with *CMT*s and *Dnmt2* being significantly downregulated while all DNA demethylases were remarkably upregulated suggesting a potential role in stress-related mechanisms. The gene expression profiling of the two families revealed under developmental stages and drought stress provides novel information that can facilitate functional future studies of DNA methylation in the growth and development of oilseed species.

Compliance with ethical standards

References

Ahmad F, Huang X, Lan HX, Huma T, Bao YM, Huang J, Zhang HS (2014) Comprehensive gene expression analysis of the DNA (cytosine-5) methyltransferase family in rice (*Oryza sativa* L.). Genet Mol Res 13(3):5159–5172

Allan G, Williams A, Rabinowicz PD, Chan AP, Ravel J, Keim P (2008) Worldwide genotyping of castor bean germplasm (*Ricinus communis* L.) using AFLPs and SSRs. Genet Resour Crop Evol 55(3):365–378

Baldoni AB, Araújo ACG, de Carvalho MH, Gomes ACMM, Aragao FJL (2010) Immunolocalization of ricin accumulation during castor bean (*Ricinus communis* L.) seed development. Int J Plant Biol 1(2):12

Bennetzen JL, Wang H (2014) The contributions of transposable elements to the structure, function, and evolution of plant genomes. Annu Rev Plant Biol 65:505–530

Cao X, Springer NM, Muszynski MG, Phillips RL, Kaeppler S, Jacobsen SE (2000) Conserved plant genes with similarity to mammalian de novo DNA methyltransferases. Proc Natl Acad Sci 97(9):4979–4984

Cao D, Ju Z, Gao C, Mei X, Fu D, Zhu H, Luo Y, Zhu B (2014) Genome-wide identification of cytosine-5 DNA methyltransferases and demethylases in *Solanum lycopersicum*. Gene 550(2):230–237

Chan SWL, Henderson IR, Jacobsen SE (2005) Gardening the genome: DNA methylation in *Arabidopsis thaliana*. Nat Rev Genet 6(5):351–360

Chan AP, Crabtree J, Zhao Q, Lorenzi H, Orvis J, Puiu D, Melake-Berhan A, Jones KM, Redman J, Chen G (2010) Draft genome sequence of the oilseed species *Ricinus communis*. Nature Biotechnol 28(9):951–956

Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL (2002) DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. Cell 110(1):33–42

Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE (2008) Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. Nature 452(7184):215–219

Diez CM, Roessler K, Gaut BS (2014) Epigenetics and plant genome evolution. Curr Opin Pl Biol 18:1–8

Du J, Zhong X, Bernatavichute YV, Stroud H, Feng S, Caro E, Vashisht AA, Terragni J, Chin HG, Tu A (2012) Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. Cell 151(1):167–180

Eichten SR, Ellis NA, Makarevitch I, Yeh C-T, Gent JI, Guo L, McGinnis KM, Zhang X, Schnable PS, Vaughn MW (2012) Spreading of heterochromatin is limited to specific families of maize retrotransposons. PLoS Genet 8(12):e1003127

Foster JT, Allan GJ, Chan AP, Rabinowicz PD, Ravel J, Jackson PJ, Keim P (2010) Single nucleotide polymorphisms for assessing genetic diversity in castor bean (*Ricinus communis*). BMC Plant Biol 10(1):1

Garg R, Kumari R, Tiwari S, Goyal S (2014) Genomic survey, gene expression analysis and structural modeling suggest diverse roles of DNA methyltransferases in legumes. PloS ONE 9(2):e88947

Gehring M, Huh JH, Hsieh T-F, Penterman J, Choi Y, Harada JJ, Goldberg RB, Fischer RL (2006) DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. Cell 124(3):495–506

Gehring M, Bubb KL, Henikoff S (2009) Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science 324(5933):1447–1451

Gong Z, Morales-Ruiz T, Ariza RR, Roldán-Arjona T, David L, Zhu J-K (2002) ROS1, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase. Cell 111(6):803–814

Greenwood J, Bewley J (1982) Seed development in *Ricinus communis* (castor bean). I. Descriptive morphology. Can J Bot 60(9):1751–1760

He S, Xu W, Li F, Wang Y, Liu A (2017) Intraspecific DNA methylation polymorphism in the non-edible oilseed plant castor bean. Plant Diversity (In Press)

Hollister JD, Gaut BS (2009) Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome Res 19(8):1419–1428

Iyer LM, Zhang D, Rogozin IB, Aravind L (2011) Evolution of the deaminase fold and multiple origins of eukaryotic editing and mutagenic nucleic acid deaminases from bacterial toxin systems. Nucleic Acids Res 39(22):9473–9497

Kapazoglou A, Engineer C, Drosou V, Kalloniati C, Tani E, Tsaballa A, Kouri ED, Ganopoulos I, Flemetakis E, Tsaftaris AS (2012) The study of two barley type I-like MADS-box genes as potential targets of epigenetic regulation during seed development. BMC Plant Biol 12(1):1

Kapazoglou A, Drosou V, Argiriou A, Tsaftaris AS (2013a) The study of a barley epigenetic regulator, HvDME, in seed development and under drought. BMC Plant Biol 13(1):172

Kapazoglou A, Drosou V, Nitsos C, Bossis I, Tsaftaris A, Triantafyllidis K, Hilioti Z (2013b) Biofuels get in the fast lane: developments in plant feedstock production and processing. Adv Crop Sci Tech 1(4):117–132 b)

Kinoshita T, Miura A, Choi Y, Kinoshita Y, Cao X, Jacobsen SE, Fischer RL, Kakutani T (2004) One-way control of FWA imprinting in *Arabidopsis* endosperm by DNA methylation. Science 303(5657):521–523

Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet 11(3):204–220

Le T-N, Schumann U, Smith NA, Tiwari S, Au PCK, Zhu Q-H, Taylor JM, Kazan K, Llewellyn DJ, Zhang R (2014) DNA demethylases target promoter transposable elements to positively regulate stress responsive genes in *Arabidopsis*. Genome Biol 15(9):1

Lindroth AM, Cao X, Jackson JP, Zilberman D, McCallum CM, Henikoff S, Jacobsen SE (2001) Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science 292(5524):2077–2080

Lisch D (2013) How important are transposons for plant evolution? Nat Rev Genet 14(1):49–61

Liu R, How-Kit A, Stammitti L, Teyssier E, Rolin D, Mortain-Bertrand A, Halle S, Liu M, Kong J, Wu C (2015) A DEMETER-like DNA demethylase governs tomato fruit ripening. Proc Natl Acad Sci 112(34):10804–10809

Malik G, Dangwal M, Kapoor S, Kapoor M (2012) Role of DNA methylation in growth and differentiation in *Physcomitrella patens* and characterization of cytosine DNA methyltransferases. The FEBS J 279(21):4081–4094

Matzke MA, Mosher RA (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity (vol 15, 394, 2014). Nat Rev Genet 15 (8)

Merkouropoulos G, Kapazoglou A, Drosou V, Jacobs E, Krolzig A, Papadopoulos C, Hilioti Z (2016) Dwarf hybrids of the bioenergy crop *Ricinus communis* suitable for mechanized harvesting reveal differences in morpho-physiological characteristics and seed metabolic profiles. Euphytica 210(2):207–219

Morales-Ruiz T, Ortega-Galisteo AP, Ponferrada-Marin MI, Martinez-Macias MI, Ariza RR, Roldan-Arjona T (2006) DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. Proc Natl Acad Sci 103(18):6853–6858

Mueller TD, Feigon J (2002) Solution structures of UBA domains reveal a conserved hydrophobic surface for protein-protein interactions. J Mol Biol 319(5):1243–1255

Ortega-Galisteo AP, Morales-Ruiz T, Ariza RR, Roldán-Arjona T (2008) *Arabidopsis* DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks. Plant Mol Biol 67(6):671–681

Pavlopoulou A, Kossida S (2007) Plant cytosine-5 DNA methyltransferases: structure, function, and molecular evolution. Genomics 90(4):530–541

Penterman J, Uzawa R, Fischer RL (2007) Genetic interactions between DNA demethylation and methylation in *Arabidopsis*. Plant Physiol 145(4):1549–1557

Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 30(9):e36-e36

Qian Y, Xi Y, Cheng B, Zhu S (2014) Genome-wide identification and expression profiling of DNA methyltransferase gene family in maize. Plant Cell Rep 33(10):1661–1672

Sanchez DH, Paszkowski J (2014) Heat-induced release of epigenetic silencing reveals the concealed role of an imprinted plant gene. PLoS Genet 10(11):e1004806

Schoft VK, Chumak N, Choi Y, Hannon M, Garcia-Aguilar M, Machlicova A, Slusarz L, Mosiolek M, Park J-S, Park GT, Fischer RL, Tamaru H (2011) Function of the DEMETER DNA glycosylase in the *Arabidopsis thaliana* male gametophyte. Proc Natl Acad Sci 108(19):8042–8047. doi:10.1073/pnas.1105117108

Severino LS, Auld DL, Baldanzi M, Cândido MJD, Chen G, Crosby W, Tan D, He X, Lakshmamma P, Lavanya C (2012) A review on the challenges for increased production of castor. Agron J 104(4):853–880

Sharma R, Mohan Singh RK, Malik G, Deveshwar P, Tyagi AK, Kapoor S, Kapoor M (2009) Rice cytosine DNA methyltransferases-gene expression profiling during reproductive development and abiotic stress. FEBS J 276(21):6301–6311

Stroud H, Do T, Du J, Zhong X, Feng S, Johnson L, Patel DJ, Jacobsen SE (2014) Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. Nature Struct Mol Biol 21(1):64–72

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30(12):2725–2729

Tsaftaris AS, Kapazoglou A, Darzentas N (2012) Epigenetics, epigenomics, and implications in plant breeding. Plant Biotechnology and Agriculture: Prospects for the 21st Century. Academic press, Cambridge, pp 207–226

Uthup TK, Ravindran M, Bini K, Thakurdas S (2011) Divergent DNA methylation patterns associated with abiotic stress in *Hevea brasiliensis*. Mol Plant 4(6):996–1013

Wen S, Wen N, Pang J, Langen G, Brew-Appiah RAT, Mejias JH, Osorio C, Yang M, Gemini R, Moehs CP, Zemetra RS, Kogel K-H, Liu B, Wang X, von Wettstein D, Rustgi S (2012) Structural genes of wheat and barley 5-methylcytosine DNA glycosylases and their potential applications for human health. Proc Natl Acad Sci 109:20543–20548

Xia Z, Zou M, Zhang S, Feng B, Wang W (2014) AFSM sequencing approach: a simple and rapid method for genome-wide SNP and methylation site discovery and genetic mapping. Sci Rep 4

Yamamuro C, Miki D, Zheng Z, Ma J, Wang J, Yang Z, Dong J, Zhu J-K (2014) Overproduction of stomatal lineage cells in Arabidopsis mutants defective in active DNA demethylation. Nat Commun 5(5):4062

Yi C, Zhang S, Liu X, Bui HTN, Hong Y (2010) Does epigenetic polymorphism contribute to phenotypic variances in *Jatropha curcas* L.? BMC Plant Biol 10(1):1

Zhang M, Kimatu JN, Xu K, Liu B (2010) DNA cytosine methylation in plant development. J Genet Genomics 37(1):1–12

Zheng X, Pontes O, Zhu J, Miki D, Zhang F, Li W-X, Iida K, Kapoor A, Pikaard CS, Zhu J-K (2008) ROS3 is an RNA-binding protein required for DNA demethylation in *Arabidopsis*. Nature 455(7217):1259–1262