Divergent Expression Patterns of *miR164* and *CUP-SHAPED COTYLEDON* Genes in Palms and Other Monocots: Implication for the Evolution of Meristem Function in Angiosperms

Hélène Adam,[1] Mélanie Marguerettaz,†,[1] Rashad Qadri,[1] Bernard Adroher,[2] Frédérique Richaud,[1] Myriam Collin,[1] Anne-Céline Thuillet,[3] Yves Vigouroux,[3] Patrick Laufs,[2] James W. Tregear,[1] and Stefan Jouannic*,[1]

[1]IRD, UMR DIAPC, IRD/CIRAD Palm Development Group, Montpellier Cedex, France

[2]Institut Jean-Pierre Bourgin, UMR1318 INRA-AgroParisTech, INRA Centre de Versailles-Grignon, Versailles Cedex, France

[3]IRD, UMR DIAPC, Dynadiv Group, Montpellier Cedex, France

†Present address: UMR BGPI-CIRAD Campus International de Baillarguet, 34 398 Montpellier Cedex 5, France

***Corresponding author**: E-mail: stephane.jouannic@ird.fr.

Associate editor: Neelima Sinha

Research article

Abstract

In order to understand how the morphology of plant species has diversified over time, it is necessary to decipher how the underlying developmental programs have evolved. The regulatory network controlling shoot meristem activity is likely to have played an important role in morphological diversification and useful insights can be gained by comparing monocots and eudicots. These two distinct monophyletic groups of angiosperms diverged 130 Ma and are characterized by important differences in their morphology. Several studies of eudicot species have revealed a conserved role for *NAM* and *CUC3* genes in meristem functioning and pattern formation through the definition of morphogenetic boundaries during development. In this study, we show that *NAM*- and *CUC3*-related genes are conserved in palms and grasses, their diversification having predated the radiation of monocots and eudicots. Moreover, the *NAM-miR164* posttranscriptional regulatory module is also conserved in palm species. However, in contrast to the *CUC3*-related genes, which share a similar expression pattern between the two angiosperm groups, the expression domain of the *NAM-miR164* module differs between monocot and eudicot species. In our studies of spatial expression patterns, we compared existing eudicot data with novel results from our work using two palm species (date palm and oil palm) and two members of the Poaceae (rice and millet). In addition to contrasting results obtained at the gene expression level, major differences were also observed between eudicot and monocot *NAM*-related genes in the occurrence of putative *cis*-regulatory elements in their promoter sequences. Overall, our results suggest that although NAM- and CUC3-related proteins are functionally equivalent between monocots and eudicots, evolutionary radiation has resulted in heterotopy through alterations in the expression domain of the *NAM-miR164* regulatory module.

Key words: evo-devo, meristem, microRNA, palms, CUC, miR164.

Introduction

Angiosperm species display an enormous diversity in their architecture and in the structure of their individual organs. One of the most important evolutionary questions is to understand how this morphological diversity arises during evolution. Plant form diversity has resulted in part from variations in the organogenic activity of meristems (apical and/or axillary) during the vegetative and reproductive phases of development. In order to understand how angiosperm morphology has radiated during evolution, it is therefore interesting to investigate how meristem function has diverged between different groups. One important morphological dichotomy observed during angiosperm evolution is the split between the monocots and eudicots lineages, which diverged 130 Ma.

In angiosperms, the shoot meristem is typically a dome-shaped structure with a characteristic inner organization of tissues, containing a group of undifferentiated and dividing cells (Kwiatkowska 2008). Its activity is tightly controlled in order to carry out two processes: first, its self-maintenance; and second, the formation of lateral organs. The formation of an organ primordium without disrupting the integrity of the meristem requires the creation of boundaries that separate the primordium from surrounding tissues, in order to delineate territories, which will undergo different cell fates.

Several genes involved in the definition of the morphogenetic boundaries have been identified in a number of species (Aida and Tasaka 2006). One such class identified includes the *Petunia hybrida* gene *NO APICAL MERISTEM* (*PhNAM*; Souer et al. 1996), the *Arabidopsis thaliana* genes *CUP-SHAPED COTYLEDON 1*, 2, and 3 (*AtCUC1*, *AtCUC2*, *AtCUC3*; Aida et al. 1997; Takada et al. 2001; Vroemen et al. 2003) and the *Antirrhinum majus* gene *CUPULIFORMIS* (*AmCUP*) (Weir et al. 2004). All these genes encode

related members of the NAC family (for NAM, ATAF1/2, CUC), which constitute one of the largest plant-specific families of transcription factors (105 genes in *A. thaliana* and 75 in rice; Ooka et al. 2003).

In eudicot species, *NAM/CUC3* genes are known to be involved in shoot apical meristem (SAM) establishment during embryogenesis, in axillary meristem establishment at later stages, and also in the establishment and maintenance of meristem-organ (M-O) and organ-organ (O-O) boundaries during plant development as a whole (Souer et al. 1996; Aida et al. 1997, 1999; Takada et al. 2001; Vroemen et al. 2003; Weir et al. 2004; Hibara et al. 2006; Keller et al. 2006; Blein et al. 2008; Raman et al. 2008; Berger et al. 2009). They are also involved in lateral organ enlargement (Larue et al. 2009), in leaf shape patterning in both simple and compound-leaved species (Nikovics et al. 2006; Blein et al. 2008; Berger et al. 2009), and in the postmeristematic maintenance of phyllotaxy through the inhibition of cell division and elongation in their expression domain, thereby leading to structural separation (Peaucelle et al. 2007). The two classes of *CUC* gene are expressed in narrow domains and in the M-O and O-O boundaries, and their contribution may differ between species and/or tissues, as illustrated by *AtCUC1* and *AtCUC2* versus *AtCUC3* involvement in *A. thaliana*, indicating both partial redundancy and subfunctionalization of the *NAM* and *CUC3* genes (Hibara et al. 2006; Nikovics et al. 2006).

Several studies have shown that some *NAM*-related genes are posttranscriptionally regulated by a microRNA, *miR164*, which limits boundary expansion by cleavage of their mRNAs (Laufs et al. 2004; Mallory et al. 2004; Baker et al. 2005; Nikovics et al. 2006; Peaucelle et al. 2007; Sieber et al. 2007). In *A. thaliana*, *miR164* acts by regulating transcript abundance in a pre-existing transcriptional pattern, leading to a dampening of expression and therefore the reduction or stabilization of target gene expression levels (Baker et al. 2005; Nikovics et al. 2006; Sieber et al. 2007; Raman et al. 2008). This suggests that the spatial regulation of *NAM* genes depends to a large extent on initial transcriptional activity; however, nothing is known as yet of the factors directly involved in this process. A different regulatory process occurs in tomato, in which the *GOBLET* gene, the tomato *NAM* ortholog, and *miR164* are expressed in complementary domains in the SAM and also in the developing leaf primordium during leaflet formation (Berger et al. 2009). This suggests that the regulatory mechanism of the *NAM-miR164* module might vary between species, from the dampening system in *A. thaliana*, to an exclusive system as in tomato (Voinnet 2009). However, at this point, it cannot be concluded which regulatory system is more typical in eudicots.

In *Zea mays* and *Oryza sativa* (Poales, Poaceae), the *CUC3*-related genes present an expression pattern similar to that observed in eudicots (i.e., M-O and O-O boundaries) (Hibara and Nagato 2005; Zimmerman and Werr 2005; Hibara and Nagato 2006; Oikawa and Kyozuka 2009). In contrast, the expression patterns of the *NAM*-related genes from these two species show some divergence compared with eudicots (Hibara and Nagato 2005; Zimmerman and Werr 2005; Hibara and Nagato 2006; Oikawa and Kyozuka 2009). During maize reproductive development, the *NAM*-related genes are expressed within the spikelet and floral meristems and not in the M-O and floral O-O boundaries (Zimmerman and Werr 2005). These results suggest that the *NAM/CUC3* genes from maize and rice may be partially divergent in their function compared with the situation in eudicots. It is not clear, however, whether this is a general characteristic of monocots or if it is limited to these grass species. Furthermore, no data are yet available on *miR164*-dependent regulation of *NAM*-related genes in maize and rice; thus, we do not yet know whether the *NAM* expression domains seen in these two species are associated with a similarly divergent *miR164* expression pattern.

In order to broaden current knowledge of *NAM/CUC3* genes and *miR164* function in monocotyledons, we chose to characterize these genes in an order distinct from the Poales: the Arecales (Arecaceae or palm family). The palms are a group of plants of particular interest because they have certain distinctive features compared with other monocot groups (Tomlinson 2006; Jouannic et al. 2007). In this study, we focused mostly on the reproductive development of oil palm (*Elaeis guineensis*), a monoecious species producing male and female inflorescences at different times in an alternating cycle, and date palm (*Phoenix dactylifera*), a dioecious species. Both produce highly branched inflorescences consisting of a central rachis bearing rachillae, which carry the unisexual flowers. Male and female inflorescences are strongly dimorphic, the former being characterized by a higher density of rachillae and flowers due to the larger number of axillary and floral meristems initiated (Adam et al. 2005; Daher et al. 2010).

We describe here the isolation of *NAM*- and *CUC3*-related genes from oil palm and date palm, on the basis of a phylogenetic sequence analysis. We characterized their expression patterns during development and investigated the biological activities of their encoded proteins. We established that the *NAM-miR164* posttranscriptional regulatory module is present in palms. To broaden our studies to other monocot phyla, we included two additional monocot species from the Poales in our expression studies, namely pearl millet (*Pennisetum glaucum*) and rice (*O. sativa*). In contrast to the situation for *CUC3*-related genes, the *NAM-miR164* module displays a divergent spatial expression pattern in the monocot meristems when compared with eudicots. We provide evidence suggesting that this may be due in part to the divergence of *cis*-regulatory sequences in the monocot and eudicot gene promoters since these two plant groups separated during evolution.

Materials and Methods

Plant Material

Oil palm (*E. guineensis*) vegetative shoot apex-containing samples and leaf samples were harvested from 6-month-old plants

(C1001 genotype) grown in the glasshouse in Montpellier (France). Oil palm tissues enriched in rachis, rachilla, and floral meristems were excised from inflorescences collected from the axils of leaves ranging from the leaf number F−20 to F−10, F−6 to F+2, and F+6 to F+7, respectively, on 20-year-old plants at the INRAB plantation (Pobé, Benin). The youngest expanding leaf was numbered as leaf F0. The leaf produced immediately before leaf F0 is thus leaf F+1 and the leaf produced subsequently is leaf F−1 (this leaf is not yet expanded). For a histological description of the corresponding stages, see Adam et al. (2005). In vitro somatic embryos from oil palm were produced from embryonic cell suspensions as described in Aberlenc-Bertossi et al. (1999).

Date palm (*p. dactylifera*) inflorescences were collected from both male and female plants in Sanremo (Italy). Pearl millet (*p. glaucum*; ligui ecotype) and rice (*O. sativa*; Nipponbare ecotype) inflorescences were harvested from 2-month-old plants grown in the glasshouse in Montpellier (France).

Arabidopsis thaliana plants of the ecotype Columbia (Col) were supplied by the Nottingham Arabidopsis Stock Centre (see http://nasc.nott.ac.uk and the *cuc2-1* mutant line (*cuc2-1 Ler* allele back-crossed in Col background) was kindly provided by M. Aida.

Complementary DNA and Gene Cloning

Oil palm and date palm *NAM/CUC3*-related cDNA fragments were amplified by polymerase chain reaction (PCR) from phage DNA of male inflorescence cDNA (complementary DNA) libraries constructed in the λZAPII vector (Stratagene) using the degenerate primers CUCS1 and CUCAS1 (see supplementary table S1, Supplementary Material online) designed against the NAC domain. After cloning and sequencing, one of the PCR products of interest was used as a probe to screen the oil palm male inflorescence cDNA library, using standard conditions (Sambrook et al. 1989). The phagemids were then isolated according to the manufacturer's instructions (Stratagene). Date palm full-length *NAM/CUC3*-related cDNAs were obtained by 5′ and 3′ RACE (rapid amplification of cDNA ends) PCRs using male inflorescence mRNA-derived reverse transcription (RT) obtained with the SMARTer RACE cDNA Amplification kit (Clontech).

The corresponding genes were isolated according to a nested PCR–based cloning approach using total DNA from a male date palm and an oil palm genomic DNA library (FN4 genotype) in conjunction with the Genome walker kit (Clontech) according to the manufacturer's instructions. The primers used for the promoter, the terminator, and the coding region of the corresponding genes are listed in supplementary table S1 (Supplementary Material online).

All the PCRs were performed using the Advantage 2 PCR Kit (Clontech). The PCR products were cloned using a pGEM-T easy ligation kit (Promega, France) with *Escherichia coli* JM109 competent cells (Promega) and sequenced by Beckman Coulter Genomics (UK).

Computational and Database Analysis

Similarity searches were performed using the BlastP and BlastX programs with default parameters in conjunction with protein sequence databases and TBlastN with nucleotide sequence collections, expressed sequence tags and high throughput genomic sequence databases provided by the NCBI server (http://www.ncbi.nlm.nih.gov/blast/). Multiple sequence alignments were performed using ClustalW with default parameters (http://www.ebi.ac.uk/clustalw/). Phylogeny and molecular evolutionary analyses were based on neighbor joining (NJ), maximum parsimony (MP), and maximum likelihood (ML) methods using Seaview program version 4 (http://pbil.univ-lyon1.fr/software/seaview.html; Gouy et al. 2010). ML phylogenies were generated using LG (after authors) evolutionary model, a model of amino acid replacement model proved to be better than previously proposed model (Le and Gascuel, 2008). Statistical support of all phylogenetic analyses was provided by performing 100 bootstrap replicates. Accession numbers of all the sequences used for the NAM/CUC3 relationship analysis are listed in supplementary table S2 (Supplementary Material online).

Promoter sequence comparisons of the *NAM*- and *CUC3*-related genes from the different species were performed using mVISTA web server (http://genome.lbl.gov/vista/index.shtml; Frazer et al. 2004).

RLM-RACE PCRs

To determine the transcription initiation sites of *EgNAM1* and *EgCUC3*, nested RACE PCRs were conducted using total RNA obtained from male inflorescences in conjunction with a GeneRacer kit (Invitrogen) according to the manufacturer's instructions. EgNAM1AS3 and EgNAM1AS4 primers, and EgCUC3AS2 and EgCUC3AS3 primers were used for *EgNAM1* and *EgCUC3*, respectively (see supplementary table S1, Supplementary Material online). To detect putative truncated mRNAs from *EgNAM1* at the *miR164* site, nested RACE PCRs were conducted using RNA ligase-mediated (RLM)-based RTs without CIP and TAP treatment (GeneRacer kit; Invitrogene) in conjunction with the EgNAM1AS7 and EgNAM1AS8 primers (see supplementary table S1, Supplementary Material online). All the PCRs were performed using the Advantage 2 PCR Kit (Clontech). PCR products were cloned using pGEM-T easy ligation kit (Promega) with *E. coli* JM109 competent cells (Promega) and sequenced by Beckman Coulter Genomics (UK).

RNA Extraction and Expression Analysis

Total RNAs (mRNAs and smallRNAs) were extracted using the RNeasy Plant Mini Kit with the RLT and RWT buffers (Qiagen, France). DNAse treatments were performed using the RNase-free DNase set (Qiagen). The mRNA-derived RTs were obtained using the ImProm-II Reverse Transcription System (Promega) from 1 μg of total RNA according to the manufacturer's instructions. The miRNA-derived RTs were obtained using a miR-X miRNA First-Strand Synthesis Kit (Clontech) from 1 μg of total RNA according to the

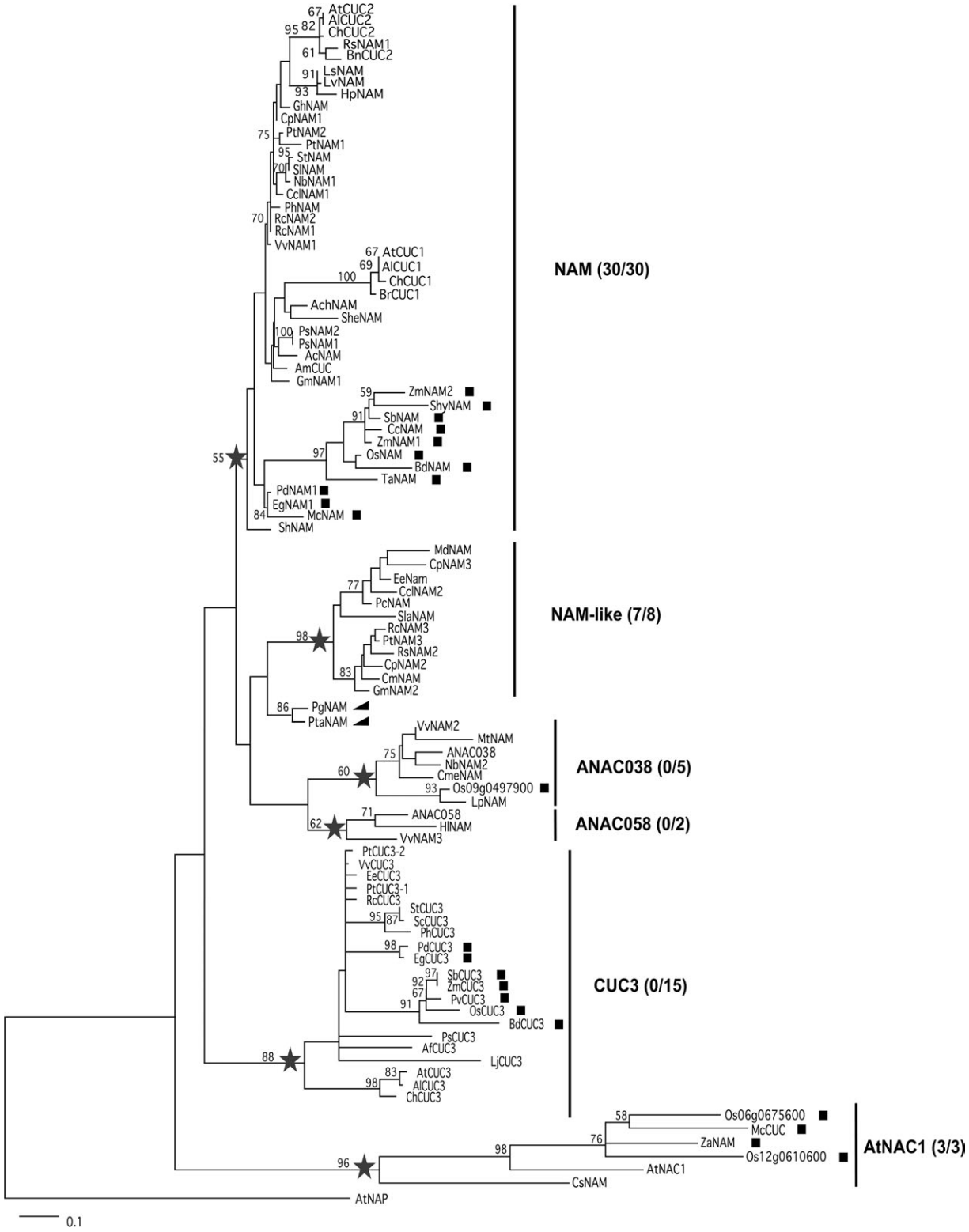

Fig. 1. Phylogenetic analysis of the NAM/CUC3 proteins. The tree is a bootstrap consensus from maximum likelihood analysis performed on the basis of NAC domain amino acid alignment, using LG evolutionary model. Bootstrap confidence values were obtained by 100 replicates and only bootstrap values greater than 50% are indicated. Black squares indicate sequences from monocot species and black triangles indicate sequences from gymnosperms. Asterisks indicate the nodes of the clades defined here. Values in brackets to the right of clade names indicate the number of full-length sequences possessing the *miR164*-binding site and the number of full-length sequences, respectively, for the corresponding clade. A list of the protein sequence accession numbers used is detailed in supplementary table S2 (Supplementary Material online). Species name abbreviations are as follows: *Actinidia chinensis*, Ach; *Antirrhinum majus*, Am; *Arabidopsis lyrata*, Al; *Arabidopsis thaliana*, At; *Aquilegia coerulea*, Ac; *Aquilegia formosa*, Af; *Brachypodium distachyon*, Bd; *Brassica napus*, Bn; *Brassica rapa*, Br; *Cardamine hirsuta*, Ch; *Carica papaya*, Cp; *Castanea mollissima*, Cm; *Cenchrus ciliaris*, Cc; *Citrus clementina*, Ccl; *Crocus sativus*, Cs; *Cucumis melo*, Cme; *Elaeis guineensis*, Eg; *Euphorbia esula*, Ee; *Glycine max*, Gm; *Gossypium*

manufacturer's instructions. RT synthesis was performed twice for each RNA sample and tested independently for PCR amplifications to avoid artifacts related to the quality of RT synthesis.

Semiquantitative RT-PCRs using mRNA-derived RTs as templates were performed using GoTaq polymerase (Promega) and the following conditions: 3 min 94 °C (30 s 94 °C; 30 s 55 °C; 30 s 72 °C) for 25–35 cycles, 10 min 72 °C. Nonquantitative RT-PCRs using miRNA-derived RTs were performed using the LightCycler 480 SYBR Green I Master kit (Roche, France) according to the manufacturer's instructions and the following conditions: 10 s 95 °C, (10 s 95 °C, 15 s 60 °C) 40 cycles. The primer pairs EgNAM1S1–EgNAM1AS5, EgCUC3S5–EgCUC3AS4, EgKNOX1S4–EgKNOX1AS3, and EgEF1aS2–EgEF1aAS2 were used for specific amplification of *EgNAM1*, *EgCUC3*, *EgKNOX1* (accession number DQ890420) and *EgEF1-α1* (accession number AY550990) transcripts, respectively. The primer pair EgNAM1S10–EgNAM1AS5 was used for the specific amplification of the 5′-untranslated region (UTR) and coding region of *EgNAM1*. MIR164A and MIR164C primers were used for specific *miR164* amplification in conjunction with the mRQ 3′ primer from the miR-X miRNA First-Strand Synthesis Kit (Clontech). PCR amplifications of *EgEF1-α1* and *U6* reverse transcripts were used as RT controls. Amplified cDNAs were cloned and sequenced to check the specificity of the amplification products. The sequences of all the primers used for RT-PCRs are listed in supplementary table S1 (Supplementary Material online).

RNA In Situ Hybridization

To obtain DNA templates for the RNA probe synthesis, PCR amplifications were performed with gene-specific antisense primers tailed with a T7 RNA polymerase binding site. PCRs were performed with the EgNAM1S9–EgNAM1AST7, EgCUC3S3–EgCUC3AST7, PdNAM1S2–pDNAM1AS2T7, and EgH4S1–EgH4AST7 primer pairs for *EgNAM1*-, *EgCUC3*-, *PdNAM1*-, and *EgH4*-specific probes, respectively (see supplementary table S1, Supplementary Material online for primer sequences). The resulting DNA fragments were used directly as templates for synthesizing antisense riboprobes incorporating UTP–digoxigenin (Roche) as the label in conjunction with a T7 Maxi Script kit (Ambion). The amplification products were tested by Southern hybridization to evaluate the specificity of the probe. For *miR164* detection, 0.02 μM of a 5′ digoxigenin–labeled miRCURY LNA microRNA detection probe complementary to *ath-miR164a* (Exiqon, Denmark) was used. In situ hybridization experiments were carried out as described by Adam et al. (2007). An LNA complementary to mouse *miR124* (5′-TTAAGGCACGCGGTGAATGCCA-3′) with no predicted target sequences in plants was used as a negative control. No signal was observed in the different plant tissues tested for the in situ hybridization experiments.

Detection was performed using the Vector Blue Alkaline Phosphatase Substrate Kit III (Vector Laboratories). Photos were taken using an Evolution MP 5.0 cooled camera (MediaCybernetics) in conjunction with a Leica DMRB microscope and images were processed with Photoshop software (Adobe, France).

Plant Transformation

The complete open reading frames (ORF) of *EgNAM1* and *EgCUC3* were amplified by PCR using the Advantage 2 PCR Kit (Clontech) with the EgNAM1S5–EgNAM1AS1 and EgCUC3S4–EgCUC3AS1 primer pairs for *EgNAM1* and *EgCUC3*, respectively (see supplementary table S1, Supplementary Material online for the primer sequences) and cloned using the pENTR-D-TOPO cloning kit (Invitrogen). The corresponding ORFs were transferred to the pMDC32 binary vector (Curtis and Grossniklaus 2003) using the Gateway LR Clonase enzyme mix (Invitrogen) according to the manufacturer's instructions. The resulting constructs were named 35S:*EgNAM1* and 35S:*EgCUC3*, respectively. In order to obtain the *CUC2* promoter fusion constructs, the *EgNAM1* ORF was cloned in the pCR2.1-TOPO vector (Invitrogen). The resulting plasmid was digested by *Eco*RI, and the *Eco*RI–*Eco*RI insert fragment was ligated to the pGREEN0129 binary vector (Hellens et al. 2000) containing a 3.7-Kb-long fragment, the *CUC2* promoter region (*pCUC2*; Nikovics et al. 2006). The resulting construct was named pCUC2:*EgNAM1*. The *pCUC2*:*CUC2* construct was done as previously described (Nikovics et al. 2006).

Wild-type *A. thaliana* plants of the ecotype Columbia and *cuc2-1 A. thaliana* mutant plants were transformed using the floral dip method (Clough and Bent 1998), with the 35S and the pCUC2 constructs, respectively. Transformants that survived in the medium containing hygromycin were checked for transgene insertions by PCR. Selected plants were transferred to the glasshouse and grown under standard conditions for flowering. Plant phenotypes were observed in the T0 and T1 generations for the overexpression lines The floral phenotype of transformed *cuc2-1* plants was calculated as the average of the sepal fusion score (Laufs et al. 2004) of ten flowers of the main stem of T1 plants (the untransformed *cuc2-1* line has a score of 6.2 ± 0.7 [standard deviation]). For the scoring of the leaf phenotype of transformed *cuc2-1* lines, 20 T2 plants of each line were grown alongside with four standard lines showing increasing dissection levels. At bolting, a serration score ranging from 1 to 5 was given to each line by comparing it with a smooth line: *cuc2-1*, serration score = 1; normal: wild type, serration score = 2; moderate increase of dissection: *mir164a-4*, serration score=3; intermediate increase of

←
hirsitum, Gh; *Humulus lupulus*, Hl; *Helianthus paradoxus*, Hp; *Lactuca sativa*, Ls; *Lactuca virosa*, Lv; *Lolium perenne*, Lp; *Lotus japonicus*, Lj; *Malus domestica*, Md; *Medicago trunculata*, Mt; *Musa cachaco*, Mc; *Nicotiana benthamania*, Nb; *Oryza sativa*, Os; *Paullinia cupana*, Pc; *Panicum virgatum*, Pv; *Petunia hybrida*, Ph; *Phoenix dactylifera*, Pd; *Picea glauca*, Pg; *Pinus taeda*, Pta; *Pisum sativum*, Ps; *Populus trichocarpa*, Pt; *Raphanus sativus*, Rs; *Ricinus communis*, Rc; *Saccharum hybrid*, Shy; *Saruma henryi*, Sh; *Silene latifolia*, Sla; *Solanum chacoense*, Sc; *Solanum lycopersicum*, Sl; *Solanum tuberosum*, St; *Sorghum bicolor*, Sb; *Striga hermonthica*, She; *Triticum aestivum*, Ta; *Vitis vivnifera*, Vv; *Zantedeschia aethiopica*, Za; *Zea mays*, Zm.

dissection: *CUC2g-m4*, serration score = 4; stronger increase of dissection, serration score = 5.

Results

Palm *NAM/CUC3* Homologs and Related Genes from Other Plant Species

The use of degenerate PCR primers designed on the basis of highly conserved amino acid motifs within the NAC domain (FRFHPTDE and MKKTLVFY) allowed the isolation of *NAM/CUC3*-related cDNAs, namely *EgNAM1* and *EgCUC3* from oil palm (GenBank accession numbers HM622270 and HM622271, respectively) and *PdNAM1* and *PdCUC3* from date palm (GenBank accession numbers HM622272 and HM622273, respectively). Blast analyses revealed that the deduced EgNAM1 and PdNAM1 polypeptides were related to the proteins AtCUC1 and AtCUC2 from *A. thaliana*, whereas EgCUC3 and PdCUC3 were more related to the *A. thaliana* AtCUC3 protein. The NAC domain, which is subdivided into five subdomains (A to E; Ooka et al. 2003), is highly conserved (similarities ranging from 92% to 99%) between the different NAM- and CUC3-related proteins (supplementary fig. S1, Supplementary Material online).

Detailed phylogenetic analyses were performed in order to examine the relationship between the palm NAM/CUC3 sequences and those of related proteins (fig. 1). Our analyses were based on the protein sequences of the NAC domain available in public databases from the various established genome sequences (ESTs and whole genomes available in June 2010). Several algorithms (NJ, MP, and ML) were used and all three methods produced a similar tree topology. The ML tree topology is shown in figure 1. Several noticeable clades are indicated with a star, notably in the cases of the NAM and the CUC3 clades. As expected, EgNAM1 and PdNAM1 belong to the NAM clade, whereas EgCUC3 and PdCUC3 belong to the CUC3 clade. Within the CUC3 and NAM clades, the sequences identified in monocotyledonous plants were grouped into two subbranches, distinct from the eudicot sequences, with a cleavage between Poales and non-Poales sequences. Within the NAM and CUC3 clades, the sequences from eudicot species were resolved in several subclades.

Within the CUC3 clade, each species studied to date possesses only a single gene of this type with the exception of *Populus trichocarpa*. In contrast, several duplication events appear to have occurred in the NAM clade, as suggested by the presence of paralogs in species from the orders Brassicales (Malvids-Rosids I) and Rosids II (Fabids). In the monocot NAM subbranches, a single *NAM*-related gene is present per species with the exception of *Z. mays*, which possesses two closely related genes. In date palm, only one *NAM*-related gene has been found in the available genome sequence data (see http://qatar-weill.cornell.edu/research/datepalmGenome/).

The NAM-like clade contains sequences specifically from eudicot species but not from *A. thaliana*. NAM-related sequences from *Picea* and *Pinus* species from gymnosperms are branched at the base of the NAM-like clade. As regards the three other identified clades, the ANAC038 and AtNAC1 groups contained both eudicot and monocot species, in contrast to the ANAC058 clade that comprised only eudicot species.

In parallel to the NAC domain based relationship analysis, members of the NAM and CUC3 clades were found to share distinct features in their C-terminal region outside the NAC domain. The C-terminal domain of the CUC3-related proteins is poorly conserved, but the NAM-related proteins share distinct conserved domains as already reported (Zimmermann and Werr, 2005; see supplementary fig. S1, Supplementary Material online). These motifs are conserved in both eudicots and monocots (comprising Poales and the two palm species), with the exception of motif I that is highly conserved in the eudicot and palm proteins but not in the Poales proteins (supplementary fig. S1, Supplementary Material online).

No *miR164*-binding site was found in the full-length gene sequences corresponding to the CUC3, ANAC058, and ANAC038 clades (fig. 1). In contrast, a *miR164*-binding site was found in available full-length gene sequences corresponding to the NAM, AtNAC1, and NAM-like clades (fig. 1). The *miR164*-binding site alignment revealed an overall well-conserved sequence with a few variations, allowing the definition of different groups (fig. 2*A*). Eudicot sequences from the NAM clade belong mainly to groups 1 and 2 and monocot species to groups 3 (Poales) and 4 (palms). The NAM-related sequence from the gymnosperm species *Picea glauca* possesses a *miR164*-binding site identical to that observed in the two palm species (group 4). In accessions from the NAM-like clade, sequences displaying divergences compared with the *miR164*-binding site from NAM sequences were identified (fig 2*B*). In *AtNAC1*, the *miR164*-binding site showed only one additional nucleotide change compared with the *miR164*-binding site identified in *LsNAM*. Interestingly, all the monocot sequences of the NAM clade, along with the entire NAM-like clade, AtNAC1, and the NAM-related sequence from *P. glauca*, share a C nucleotide at position 14, in contrast to the other eudicot species. Moreover, all the aforementioned sequences, plus the sequences from group 2, share a C nucleotide at position 9. These two substitutions lead to a better pairing between the mature *miR164* and its targets. In parallel, various paralogs of the mature *miR164*, characterized by single nucleotide substitutions, have been observed in individual species (fig. 2*B*). The set of paralogous mature *miR164* may differ between species. However, these sequence variations do not correlate with sequence modifications in the binding site of the target mRNA sequences (fig. 2). Finally, it was observed that the sequence of mature *Ath-miR164a/b* was conserved in all the species for which data were available from both eudicots and monocots. This mature *miR164* sequence may correspond to the ancestral form (fig. 2*B*).

Genomic Structure of *NAM*- and *CUC3*-related Genes

Genomic sequencing revealed that *EgNAM1* and *EgCUC3* were both organized into three exons and two introns

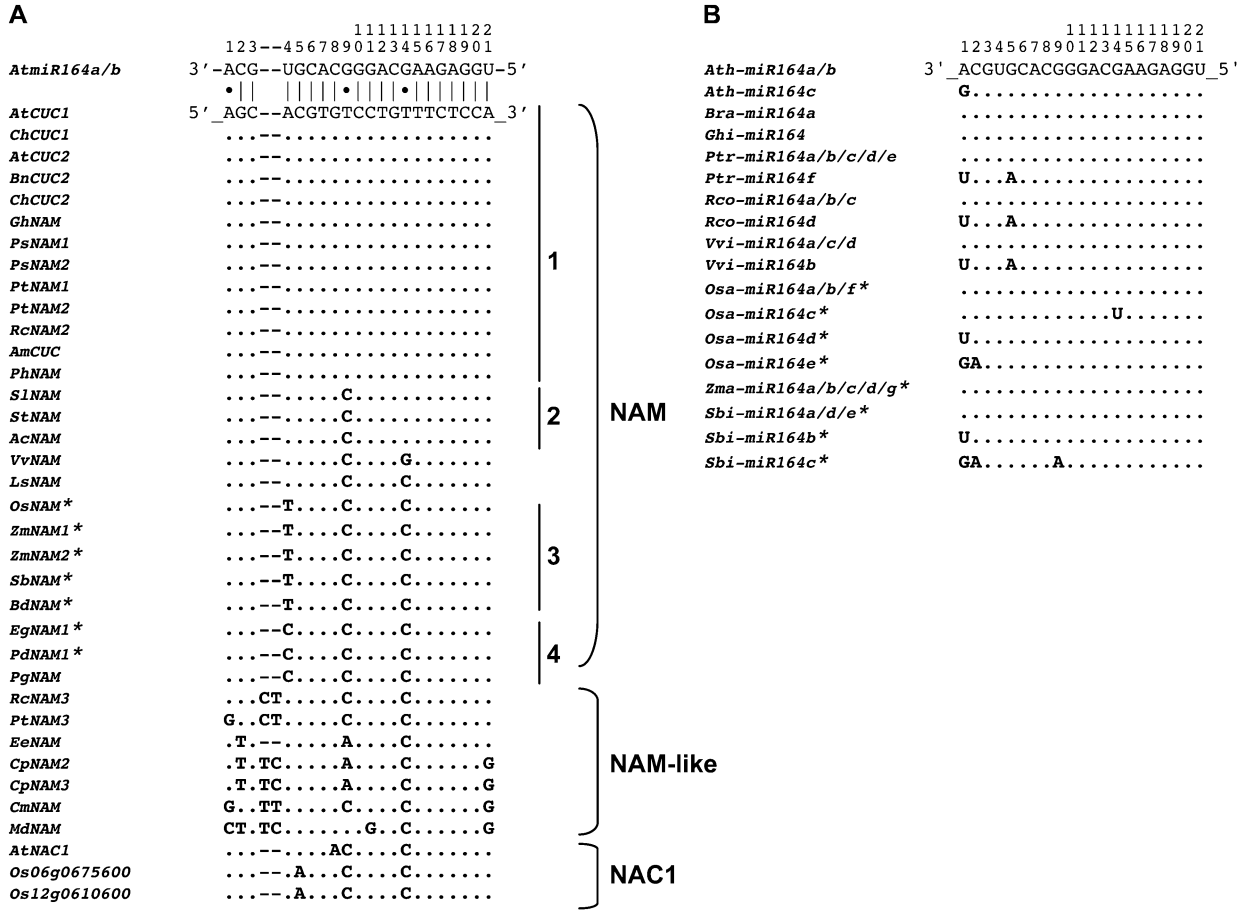

Fig. 2. Alignments of *miR164* binding site and mature *miR164* sequences. (*A*) Alignment of the *miR164*-binding sites present in sequences of the NAM, NAM-like, and AtNAC1 clades. The alignment of the *miR164*-binding site sequences with *Ath-miR164* from *Arabidopsis thaliana* is shown. Position numbers are indicated. Dots in the *miR164*-binding site sequences indicate identical positions to *AtCUC2* sequence. The different sequence groups are indicated on the right. Asterisks indicate sequences from monocot species. (*B*) Alignment of mature *miR164* sequences available in miRBASE. Dots indicate identical position to *Ath-miR164a and Ath-miR164b* sequences from *A. thaliana*. Asterisks indicate sequences from monocot species. See legend of figure 1 for the species name abbreviations.

within the coding region (fig. 3*A*). A similar organization was observed for the coding region of the *PdNAM1* and *PdCUC3* genes (fig. 3*A*). Sequencing of RLM-based 5′ RACE PCR products (using RTs from inflorescences) revealed that transcription initiation sites of *EgCUC3* and *EgNAM1* were located about 69 bp and 518 bp upstream of the translation initiation codon, respectively (supplementary fig. S1, Supplementary Material online). It also confirmed the presence of a 392 bp intron in the 5′ UTR of *EgNAM1*, leading to a transcription unit organized into four exons and three introns for this gene (fig. 3*A*).

A comparative analysis was carried out of the genomic structure of *NAM*- and *CUC3*-related genes from palms and several other species for which full-length cDNA or EST sequences were available, with the exception of *OsNAM* and *SbCUC3* for which no transcript sequences have been reported in the public databases (fig. 3*A*). The general structure of the *NAM*-related genes was found to be conserved, with one intron at a common position in the NAC domain coding region and a second intron present in the C-terminal coding region. The only deviation observed from this structure was the rice *OsNAM* gene, which has lost the intron in the NAC domain coding region. Further investigation of the noncoding sequences, by the comparison of genomic and full-length cDNA accessions, revealed the presence of an additional 5′ UTR intron in the maize *ZmNAM1* and *Sorghum bicolor SbNAM* genes (cDNA accession numbers: EU970442 and XM_002441933, respectively). It was not possible to confirm the presence of a 5′-UTR intron in the *ZmNAM2*, *OsNAM*, and *BdNAM* genes from maize, rice, and *Brachypodium distachyum*, respectively, as no full-length cDNA or EST sequences were available in the public databases. Nevertheless, a comparison between these genes with *ZmNAM1* and *SbNAM* revealed sequence conservation in all cases in and around the positions corresponding to the splice sites. In contrast, no intron was observed in the 5′ UTR of the *A. thaliana AtCUC1* and *AtCUC2* genes, nor in the tomato *SlNAM* gene. These findings suggest that the 5′ UTR intron is a characteristic of *NAM*-related genes from palms and grasses and is not shared by eudicot genes.

A detailed analysis of the 5′ upstream region of the *NAM*-related genes revealed four conserved noncoding

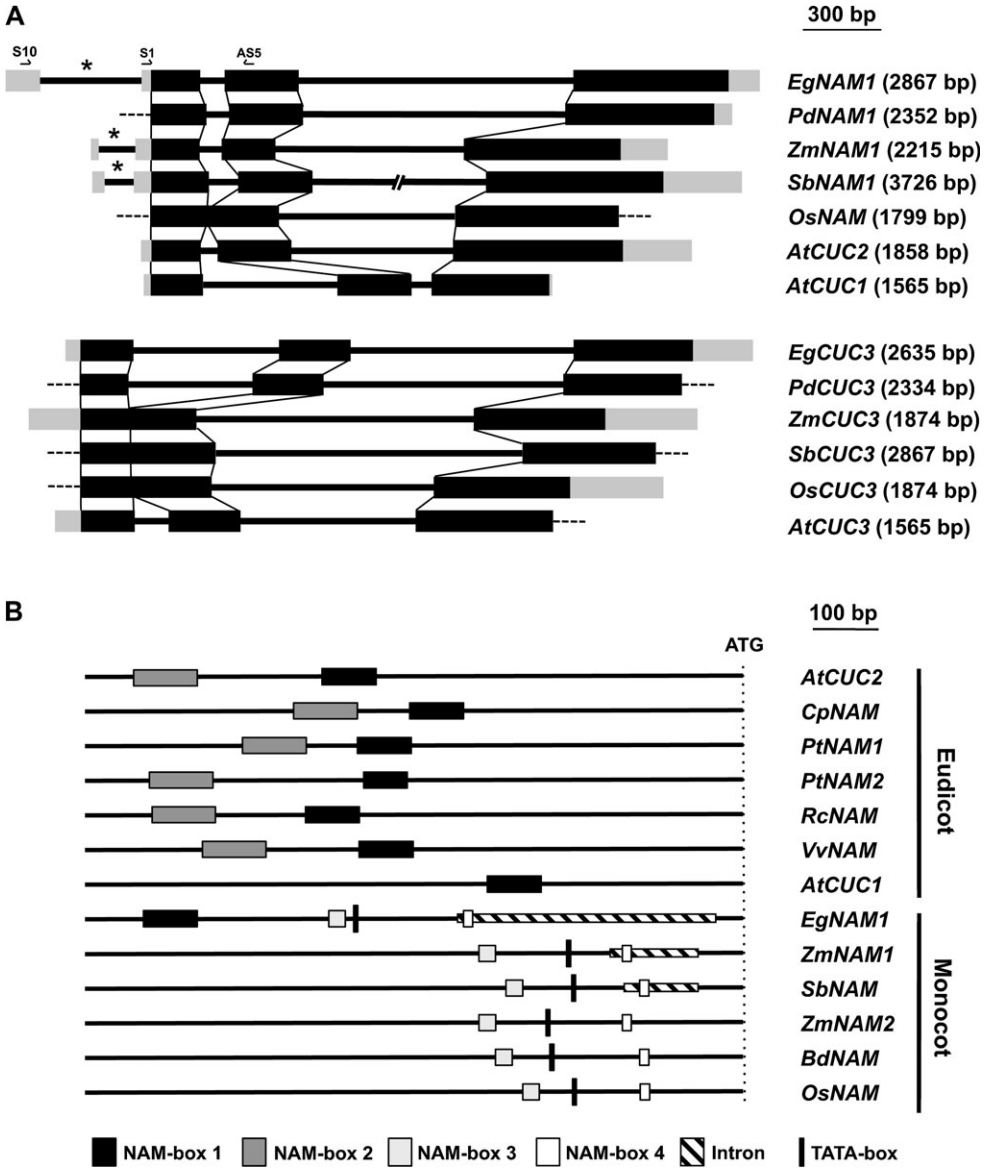

FIG. 3. Genomic structure of *NAM*/*CUC3*-related genes and CNSs in the 5′ flanking regions of *NAM*-related genes. (*A*) Genomic structure of *NAM*- and *CUC3*-related genes in *Elaeis guineensis* and *Phoenix dactilyfera* in comparison with *Zea mays*, *Oryza sativa* and *Arabidopsis thaliana*. Black lines correspond to intron sequences, light and dark gray boxes correspond to untranslated and translated exon sequences, respectively. Asterisks indicate the presence of a 5′-UTR intron. The positions of the primers used for the *EgNAM1* RT-PCR–based expression analysis are indicated by arrows (S10 = EgNAM1S10 primer; S1 = EgNAM1S1 primer; AS5 = EgNAM1AS5 primer). (*B*) The position of the four CNSs identified in the 5′ upstream regions from the translation start codon of *NAM*-related genes are illustrated. The alignments of the corresponding sequences are shown in supplementary fig. S1 (Supplementary Material online).

sequences (CNSs), which were named NAM-box 1, NAM-box 2, NAM-box 3, and NAM-box 4. Their relative positions with respect to the translation initiation codon are shown in figure 3*B* (see supplementary fig. S1, Supplementary Material online for the sequence alignments). NAM-box 1 (ca. 130 bp long) was conserved in all the eudicot genes investigated and in the *EgNAM1* gene from oil palm. NAM-box 2 (ca. 110 bp long) is conserved specifically in the eudicot species, with the exception of *AtCUC1*. NAM-box 3 (ca. 39 bp long) and NAM-box 4 (ca. 30 bp long) were conserved specifically in the monocot genes investigated, upstream of the putative TATA-box and in the 5′ half of the 5′ UTR intron region, respectively. A BlastN search against the complete genomic sequences from maize, rice, *Sorghum* and *Brachypodium* did not reveal any NAM-box 1 or 2 sequences in the more distant 5′ flanking regions of their *NAM* genes, nor anywhere else in these genomes. In the same way, no NAM-box 3 or 4 related sequences were identified in the *Arabidopsis*, *Populus*, *Carica*, or *Vitis* genomic sequences. None of these sequences share similarities with any known regulatory domains. Taken together, these data indicate a wide divergence between eudicots and monocots with respect to the organization of their *NAM*-related gene 5′ flanking regions.

As regards the *CUC3*-related genes, our comparisons showed that their overall exon–intron organization is similar

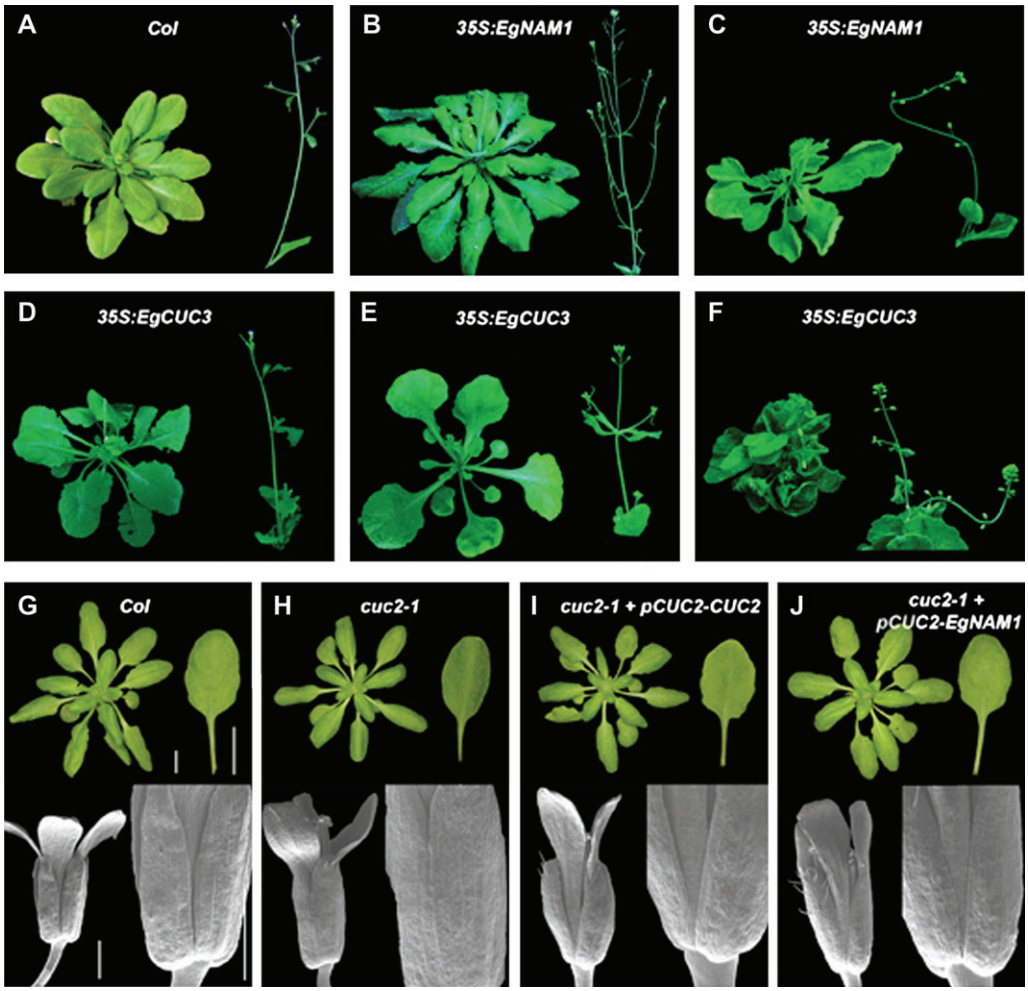

Fig. 4. Functional analysis of *EgNAM1* and *EgCUC3* genes in *Arabidopsis thaliana*. (*A*–*F*) Leaf and inflorescence phenotypes of independent *A. thaliana Col0* plants transformed with the 35S-*EgNAM1* (two independent lines) and 35S-*EgCUC3* constructs (three independent lines). (*G*–*J*) Leaf and flower phenotypes of independent *A. thaliana cuc2-1* mutant plants transformed with the *pCUC2–CUC2* and *pCUC2–EgNAM1* constructs. Scale bars: 1 cm (rosette and leaf); 1 mm (flower and sepals).

to that of the *NAM*-related genes, with three exons, one intron at a common position in the NAC domain coding region, and a second intron in the C-terminal coding region. The exceptions were the *ZmCUC3*, *OsCUC3*, and *SbCUC3* genes from maize, rice, and *Sorghum*, respectively, for which the intron in the NAC domain coding region was absent, suggesting a loss of this intron in the Poales. No 5′-UTR introns were observed in either the monocot or the eudicot *CUC3*-related genes, nor were any CNSs detected in their 5′ flanking regions. Interestingly, an intron is present at a conserved position in the NAC domain coding region of both *CUC3*- and *NAM*-related genes (fig. 3*A*) and also in some other NAC-related clades (data not shown), suggesting an ancient origin for this feature within the NAC gene family.

Are EgNAM1 and EgCUC3 the Functional Equivalents of the AtCUC2 and AtCUC3 Proteins?

To investigate the biological functions of the proteins encoded by the oil palm *EgNAM1* and *EgCUC3* genes, *A. thaliana* transgenic plants ectopically expressing each of them under the control of the CaMV 35S promoter were generated (fig. 4*A*–*F*). In both cases, alterations to leaf shape and inflorescence architecture were observed in comparison with wild-type plants. In less severely affected lines, plants were affected only in their leaf shape with a highly pronounced serration of the leaf edges (fig. 4*B* and *D*). In more severely affected lines, the overall leaf shape was altered to produce larger leaves and petioles and the leaves were often crenated. Plants of this type also displayed a modified inflorescence structure with alterations to phyllotaxy and branching (fig. 4*C*, *E*, and *F*). These phenotypic alterations have also been observed in plants overexpressing the *AtCUC2* gene (Laufs et al. 2004), in plants expressing a *miR164*-resistant form of *AtCUC2*, CUC2g-m4 (Nikovics et al. 2006; Peaucelle et al. 2007), and in plants expressing *AtCUC3* under the control of the *AtCUC2* promoter (B. Adroher and P. Laufs, unpublished data).

A complementation assay of the *A. thaliana cuc2-1* mutant with *EgNAM1* under the control of *AtCUC2* promoter confirmed that *EgNAM1* is functionally equivalent to *AtCUC2*. This was witnessed by the complementation of the

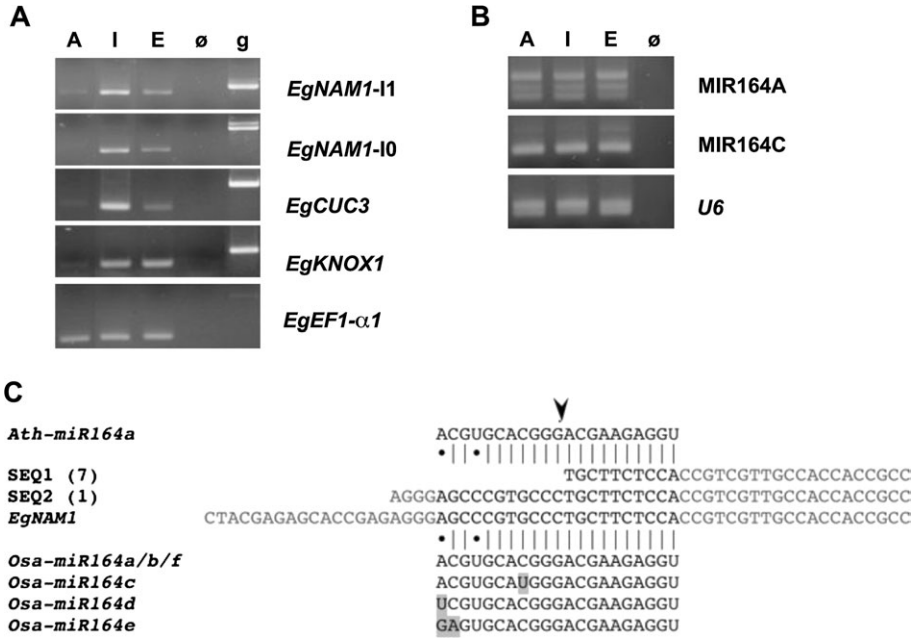

Fig. 5. RT-PCR expression profiles of *EgNAM1*, *EgCUC3*, and *Egu-miR164* in different plant organs and detection of truncated *EgNAM1* transcripts at the *miR164* recognition site. (*A*) *EgNAM1*, *EgCUC3*, and *EgKNOX1* expression profiles in different tissues and organs. *A*: apex; *I*: male inflorescence; *E*: somatic embryo; ø: no RT control; *g*: genomic DNA control. (*B*) *miR164* expression profiles in different tissues and organs using MIR164A and MIR164C primers (see supplementary table S1, Supplementary Material online). See (*A*) for the legend. *EgEF1-α1* and *U6* were used as RT controls for the mRNA- and the miRNA-derived RTs, respectively. The primer pairs EgNAM1S1–EgNAM1AS5 and EgNAM1S10–EgNAM1AS5 were used for EgNAM1-I1 and EgNAM1-I0 RT-PCRs, respectively (see supplementary table S1, Supplementary Material online, for primer sequences; see figure 3*A* for the position of the primers). (*C*) The sequences of the truncated *EgNAM1* transcripts detected by RLM-RACE PCR from inflorescence RTs were aligned to the mature *miR164* sequences from *Arabidopsisc thaliana* (*Ath-miR164a*) and rice (*Osa-MIR164a/b/f, c, d* and *e*). The arrowhead indicates the *miR164*-directed cleavage site as detected in other species. The number of clones obtained is indicated in brackets at the corresponding sequence.

cuc2-1 leaf and flower defects, that is, restoration of leaf serration and elimination of sepal fusion to the same level as that produced with *AtCUC2* (fig. 4*G–I*; see supplementary fig. S2, Supplementary Material online for quantitative data).

These data collectively indicate that the EgNAM1 and EgCUC3 proteins are functionally interchangeable with the AtCUC2 and AtCUC3 proteins. Moreover, on the basis of the complementation assay, it can be concluded that the *EgNAM1* gene can be regulated by *Ath-miR164* in the *Arabidopsis cuc2-1* mutant, as no phenotypes similar to those associated with the *miR164*-resistant form of *AtCUC2* were observed (i.e., CUC2g-m4; Nikovics et al. 2006; Peaucelle et al. 2007).

Characterization of the *EgNAM1*/*EgCUC3*/*Egu-miR164* Expression Patterns and Demonstration of *miR164*-dependent Cleavage of *EgNAM1*

The expression patterns of *EgNAM1* and *EgCUC3* were investigated by RT-PCR using various different tissues and organs of oil palm and primers located at the 5′ end of the transcripts in order to detect full-length ones (fig. 5). Both *EgNAM1* and *EgCUC3* were expressed in inflorescence (both male and female) and embryos in a similar fashion to the *EgKNOX1* gene (fig. 5*A*; Jouannic et al. 2007). In SAM-containing tissue, both *EgNAM1* and *EgCUC3* transcripts were detected, but at a very low level. The splicing of the 5′-UTR intron of *EgNAM1* was confirmed by the use of a second pair of primers (see EgNAM1-I1 vs. EgNAM1-I0 in fig. 5*A*; see the position of the primers in fig. 3*A*).

In parallel, the accumulation of mature *miR164* was investigated in the same tissues using nonquantitative RT-PCR. As the sequence of the palm *miR164s* is unknown, we used two different primers sets, MIR164A designed against *Osa-miR164a*, *b*, or *f* forms and MIR164C designed against *Osa-miR164c* form. Several bands were obtained with the primer MIR164A. However, DNA sequencing revealed that only the smallest PCR product corresponded exactly to the expected sequence. Mature *miR164* transcripts were detected in all the organs and tissues tested (fig. 5*B*). Thus, *miR164* is conserved in oil palm and accumulates in a range of different tissues in which *EgNAM1* transcripts are also found.

To assess whether *miR164* was able to mediate the cleavage of *EgNAM1* transcripts in oil palm tissues, the presence of truncated *EgNAM1* transcripts starting at the predicted *miR164*-directed cleavage site was determined by RML-based 5′ RACE PCR using RTs synthesized from inflorescence RNAs (fig. 5*C*). DNA sequencing of eight cloned PCR products revealed that seven of them started at the predicted *miR164*-directed cleavage site. Thus, our results confirmed the existence of *miR164*-dependent cleavage of *EgNAM1* transcripts in inflorescences

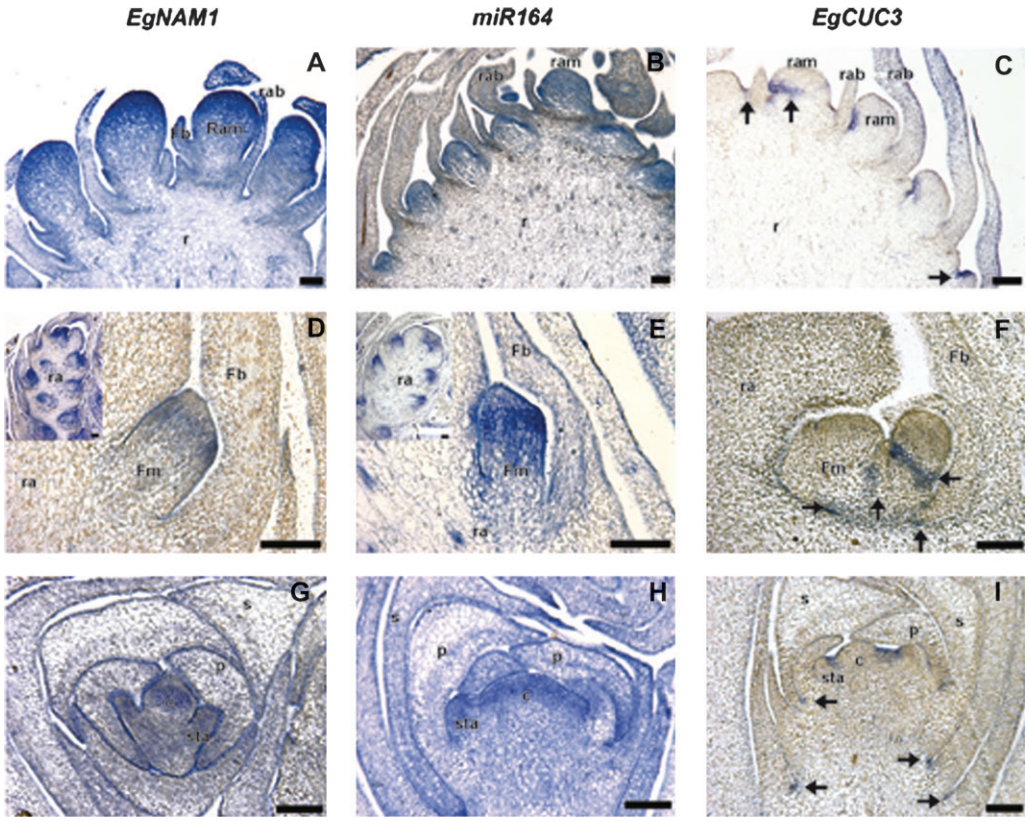

Fig. 6. In situ hybridization analysis of the expression of *EgNAM1* (*A*, *D*, *G*), *Egu-miR164* (*B*, *E*, *H*) and *EgCUC3* (*C*, *F*, *I*) in the oil palm female inflorescence. (*A*, *B*, *C*) Female inflorescence rachilla meristems initiated from the rachis at the axil of rachilla bracts. (*D*, *E*, *F*) Female inflorescence floral triad meristem initiated along the rachilla at the axil of floral bracts. (*G*, *H*, *I*) Pistillate flowers with developed sepals and petals and reproductive organ primordia. Arrows indicate *EgCUC3*-associated signals. c, carpel; Fb, floral bract; Fm, floral meristem; p, petal; r, rachis; ra, rachilla; rab, rachilla bract; rap, rachilla primordium; ram, rachilla meristem; sta, staminode; s, sepal; bars = 100 μm.

and the conservation of the *NAM-miR164* regulatory module in a monocot species.

Expression Patterns of *EgNAM1*, *Egu-miR164*, and *EgCUC3* in Reproductive Organs

Oil palm inflorescences are compound structures consisting of a central rachis that gives rise to branches known as rachillae. The oil palm produces functionally unisexual male and female inflorescences in alternation. During reproductive development, the rachis produces a series of rachilla bracts that subtend in their axils the meristems from which rachillae will be produced. The density of floral meristems produced on the rachilla is related to the sex of the inflorescence; thus, male rachillae can be seen to display a larger number of floral bracts than the future female rachillae (Adam et al. 2005). While the male inflorescence produces individual functional male flowers, the rachillae of the female inflorescence bear triads of flowers, each consisting of a functional pistillate flower flanked by two abortive staminate flowers.

EgNAM1, *Egu-miR164*, and *EgCUC3* transcript localization was studied in longitudinal sections of female inflorescences at different stages of development: at rachilla meristem initiation (fig. 6*A*–*C*), in the floral triad meristem (fig. 6*D*–*F*), and in the pistillate flower (fig. 6*G*–*I*). In the rachilla meristem, a signal corresponding to *EgNAM1* transcripts was detected in the apical cells of rachillae and at a lower level in the cells of rachilla bracts (fig. 6*A*). To follow in situ *miR164* expression, we used a probe complementary to *Osa-miR164a/b/f*, which recognizes at least a subset of the palm *miR164* pool (fig. 5*B*). Mature transcripts of *Egu-miR164* were also detected in the apical cells of rachilla meristems but not in the rachilla bracts (fig. 6*B*). At the same developmental stage, *EgCUC3* transcripts were detected specifically in the boundary cells between the rachilla bracts and the rachilla meristem (fig. 6*C*). At the floral triad meristem stage, *EgNAM1* and *Egu-miR164* transcripts were detected in the same cells of floral meristems. Transcripts were located in the cells surrounding the floral meristem (i.e., in a ringlike zone) and at a lower level in the apical tip of the meristem (fig. 6*D*–*E*). In contrast, *EgCUC3* transcripts were detected in the cells separating adjacent organs: between the rachillae and the floral triad meristems and in the zone dividing the three floral triad meristems (fig. 6*F*). At the flower development stage in the female inflorescence, both *EgNAM1* and *Egu-miR164* transcripts were accumulated in the cells of the reproductive organ primordial, but *EgNAM1* transcripts were not detected in the L1 cell layer of these organs (fig. 6*G*–*H*). At the same developmental stage, a signal corresponding to *EgCUC3* transcripts was

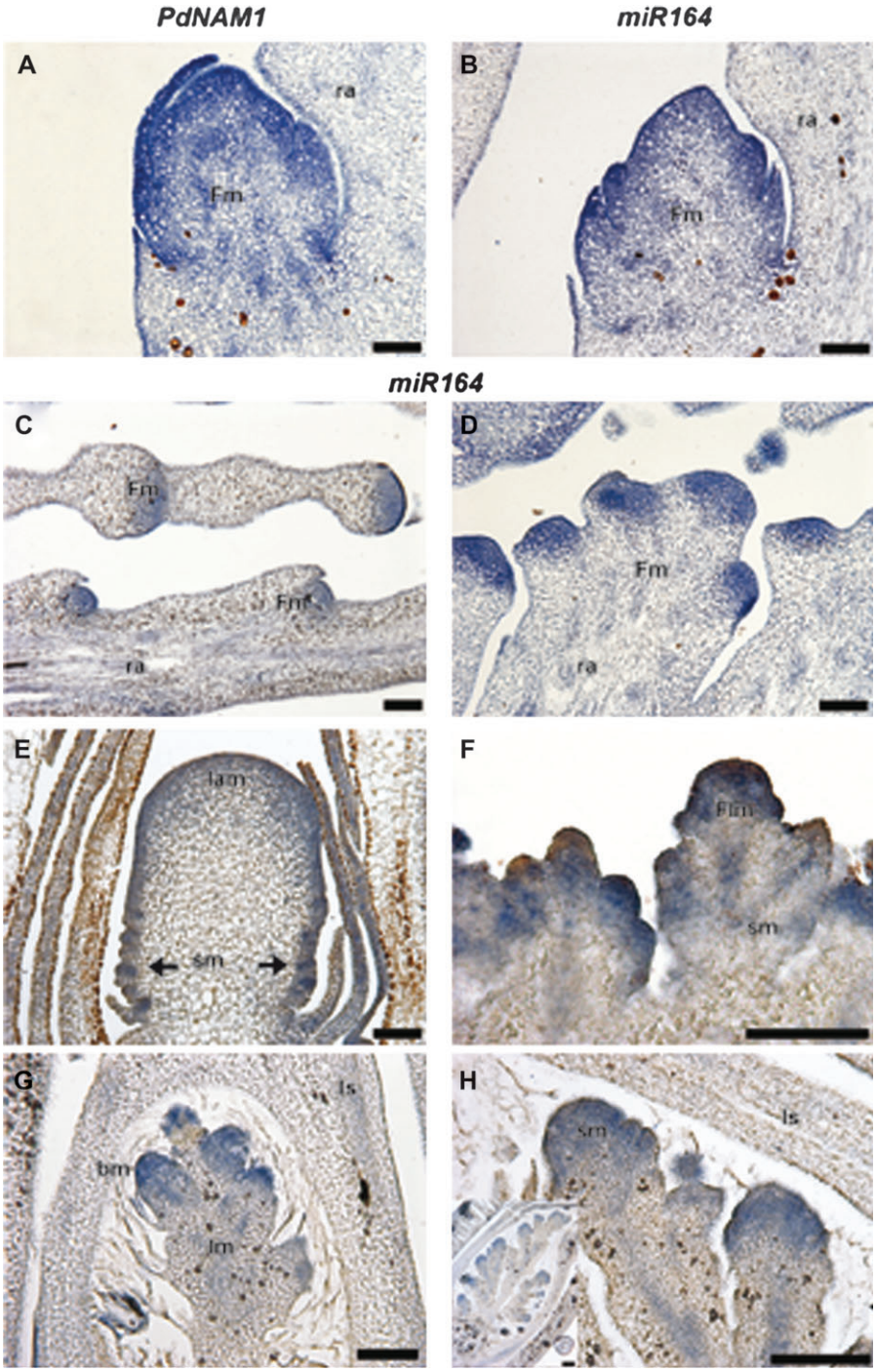

Fig. 7. In situ hybridization analysis of the expression of *PdNAM1* (*A*) and *miR164* (*B* to *H*) in the inflorescence meristems of other monocot species. (*A*, *B*) Date palm pistillate flower showing intiation of floral organs. (*C*) Date palm female rachilla bearing floral meristem. (*D*) Date palm male rachilla bearing floral meristem. (*E*, *F*) Millet reproductive meristems; *E*, millet inflorescence meristem, spm are initiated at the base of inflorescence, and initiate later the spikelet meristem, *F*, millet spikelet meristem with floral meristems. (*G*, *H*) Rice inflorescence meristem, Im initiated branch meristem which later bears the spikelet meristem. The SM initiates sterile glumes, followed by lemmas, containing florets in their axils. Iam, Inflorescence apeical meristem; Im, inflorescence meristem; Flm, Floret meristem; Fm, floral meristem; bm, branch meristem; Is, leaf sheath; sm, spikelet meristem; spm, spikelet pair meristem; ra, rachilla; bars = 100 μm.

detected in the cells separating adjacent organs (i.e., the whorl–whorl boundary) and in the residual floral meristem of carpel primordia (fig. 6*I*).

Studies of the localization of *EgNAM1* and *EgCUC3* transcripts and mature *Egu-miR164* were also performed, using equivalent developmental stages, on the male inflorescence. In this case, the patterns of expression observed were similar to those seen for the female inflorescence (supplementary fig. S3, Supplementary Material online).

The above data collectively indicate that in both female and male inflorescences and flowers, *EgNAM1* and *Egu-*

miR164 transcripts are colocalized in meristematic zones. A notable difference is seen between the indeterminate rachilla meristem, in which *EgNAM1* and *Egu-miR164* were detected in the apical tip of the meristem, and the determinate floral meristems in which transcripts were detected mainly in a ringlike zone below the apex. In contrast to *EgNAM1* and *Egu-miR164* transcripts, mRNAs of *EgCUC3* showed a distinct spatial pattern of accumulation in meristem–organ boundaries (in the case of rachillae and floral meristems) and whorl boundaries (in the case of the developing flower).

Expression Patterns of *miR164* and *NAM*-related Genes in Reproductive Organs of Other Monocot Species

To assess whether the expression pattern of the *NAM-miR164* module observed in oil palm was conserved in another palm species, the expression of *PdNAM1* and of *miR164* orthologs was studied in reproductive tissues of date palm (fig. 7*A* and *B*). The signals visualized on longitudinal sections of the pistillate floral meristem showed that *PdNAM1* and *miR164* transcripts were present in the same cells, in the floral meristem, and in the developing sepal and petal primordia (fig. 7*A* and *B*). *MiR164* transcripts were also detected at earlier stages of floral meristem development in both the male and female inflorescence (fig. 7*C* and *D*). These accumulation patterns were similar to those observed in inflorescences of oil palm. Although date palm is strictly dioecious, as opposed to the temporally dioecious oil palm, no distinction between male and female inflorescences was observed in our studies, these data again mirroring those obtained with oil palm.

To determine whether the expression pattern of *miR164* was also conserved in reproductive organs of other monocot species, we selected two species from the family Poales: pearl millet (*P. glaucum*) and rice (*O. sativa*) (fig. 7*E*–*H*), which both produce hermaphrodite flowers. In pearl millet, a signal corresponding to *miR164* was detected in the cells of spikelet meristems and in the apical region of the inflorescence meristem (fig. 7*E*). Later during development, *miR164* transcripts were also detected in the cells of the floral meristem but were apparently excluded from the L1 cell layers (fig. 7*F*). In rice inflorescences, *miR164* transcripts were observed to accumulate in a similar way: in the cells of the lateral meristems, in the spikelet meristems, and in the floral meristem.

These data indicate that the similar *miR164* expression patterns observed in two different palm species belonging to distinct subfamilies (Arecoideae and Coryphoideae) are also conserved in two species of the Poales, again belonging to distinct subfamilies (Panicoideae and Ehrhartoideae).

Discussion

NAM/CUC3 Gene Relationship and Evolution of the NAM-miR164 Regulatory Module

Our phylogenetic reconstruction based on NAC domain sequences identified three main groups of *NAM*- and *CUC3*-related genes: the NAM, the NAM-like, and the CUC3 clades. Genes belonging to the NAM and CUC3 clades are present in both eudicots (core and basal eudicots) and monocots (Arecales, Poales, Zingiberales), indicating that these clades predate eudicot–monocot separation (130 Ma; Moore et al. 2007). Within the NAM clade, several duplication events have occurred, notably in the Brassicales (*CUC1* vs. *CUC2* genes) and also in *Populus*, pea, and maize. In contrast, the NAM-like clade appears to have originated specifically in eudicots from an ancestral eudicot *NAM*/*NAM-like* gene, which may have been lost in the Brassicales. The weak support along the backbone of the phylogenetic tree, especially the branching of the gymnosperm sequences, unables us to conclude about the evolutionary diversification of the different clades. In order to reconstruct more precisely how *NAM/CUC3* evolution occurred within the plant kingdom, it will be necessary to assemble more data from other monocot families, basal angiosperms, gymnosperms, and lower branching tracheophyte groups.

The *miR164*-binding site is globally well conserved in all the full-length sequences belonging to the NAM and AtNAC1 clades along with the NAM-related sequence from the gymnosperm species *P. glauca*. The fact that the *miR164* recognition site is conserved in terms of both transcribed mRNA and encoded protein sequences in *NAM*-related genes and other NAC domain encoding genes targeted by *miR164* adds further weight to the hypothesis of a common origin for the *miR164*-targeted genes rather than an evolutionary convergence. A few modifications of the *miR164* recognition site sequences were observable in relation to the phylogenetic relationships of the plant species (i.e., eudicot vs. monocots; Poales vs. Arecales). However, these modifications were not associated with changes to the mature *miR164* sequence in these species. This may be because *miR164* tolerates mismatches at the recognition site for its pairing to the target mRNA, leading to possible polymorphisms in the recognition site. A degenerate *miR164* recognition site was observed in full-length sequences belonging to the NAM-like clade. Unfortunately, it cannot be concluded whether these sites are still recognized by *miR164*. Interestingly, no polymorphism was observed in any of the seven positions adjoining the 3′ end of the binding site (with the exception of a few NAM-like sequences), suggesting that the nucleotide sequence of this region is crucial to the activity of *miR164* or its cofactors. This is corroborated by studies of the effects of mutations in this region on the regulation of the target mRNA, as illustrated by the goblet mutant from tomato (Berger et al. 2009), the *A. thaliana cuc2-1D* mutant (Larue et al. 2009), or the *miR164*-resistant form of *AtCUC2* (Laufs et al. 2004).

MiR164-related sequences have been identified in a wide range of vascular plants, including core eudicots, monocots, basal angiosperm species such as *Amborella trichopoda* and *Cabomba aquatica* (Jasinski et al. 2010), and the gymnosperm species *P. glauca* and *Pinus resinosa* (Axtell and Bartel 2005; Sunkar and Jagadeeswaran 2008). In contrast, no *miR164*-related sequences were detected in the sequenced

genome of the bryophyte *Physcomitrella patens*. Based on the pattern of conservation observed within the vascular plant lineages for *miR164* and its recognition site in target mRNAs, it seems probable that the *miR164* regulatory module pre-existed gymnosperm/angiosperm separation but was acquired after bryophyte/gymnosperm separation. This *miR164* ancestral gene may have arisen from inverted duplications of an ancestral *NAC* gene (Axtell and Bowman 2008).

Functional Evolution of the NAM-miR164 Module in Monocots

No data have been published to date on the conservation of the *NAM-miR164* regulatory module and *CUC3* genes in monocots from a functional point of view. Indirect information on their function can be deduced from heterologous functional complementation (our study) and expression patterns (Hibara and Nagato 2005; Zimmermann and Werr 2005; Hibara and Nagato 2006; Oikawa and Kyozuka 2009; our study). The ability of the oil palm EgNAM1 and EgCUC3 proteins to interact with the other effectors of the AtCUC2 or AtCUC3 regulatory networks in *A. thaliana* (e.g., DNA binding sites and protein partners) imply, at least in part, a conservation of the *NAM* and *CUC3* regulatory networks between plant families/orders, which diverged 130 Ma (Moore et al. 2007). Furthermore, our expression study suggests that the *NAM-* and *CUC3*-related genes do not act redundantly in the functioning of meristems in monocots, in contrast to the situation observed in eudicots.

The conservation of *CUC3*-related gene expression patterns observed in eudicots, palms, and grasses suggests a similar function throughout angiosperms in the control of M-O and O-O boundaries and also in the regulation of SAM and axillary meristem establishment. However, functional analysis will be required in order to confirm this hypothesis.

The detection of *EgNAM1* transcripts truncated at the *miR164*-directed cleavage site indicated that the *NAM-miR164* module is functionally conserved in the Arecales. Moreover, a recent study revealed the presence of *OsNAM* transcripts cleaved at the *miR164* recognition site in the rice mRNA degradome (Li et al. 2010). It seems reasonable to suppose that this is also the case with other monocot groups. Further evidence for functional conservation of the *NAM-miR164* module comes from the fact that the oil palm *EgNAM1* gene is able to complement the *A. thaliana cuc2* mutant, which implies that there is binding of the *A. thaliana miR164* to the *EgNAM1* mRNA. The task will now be to determine its function in monocot meristems, as it does not appear to be involved in the fine definition of postembryonic M-O and O-O boundaries (Zimmerman and Werr, 2005; our study). Indeed, *NAM*-related transcripts from palms and grasses are accumulated within the reproductive meristems proper rather than being confined to their boundaries. The colocalization of *NAM*-related transcripts with *miR164* inside the reproductive meristems in palms and the similar transcriptional activity of *MIR164* inside the reproductive meristems of rice and pearl millet suggest that the "dampening" system of regulation of target mRNAs by *miR164* occurs in these monocot species as observed in *A. thaliana*. It has to be kept in mind that the probe used for in situ hybridizations was designed against *Ath-MIR164a*, which is also conserved in grasses (fig. 2*B*). Consequently, the signal detected may be specific to the *miR164a* form only, and it cannot be excluded that other putative forms of *miR164* are expressed in other domains. Alternatively, it is possible that different forms of *miR164* were being detected at the same time in these experiments. However, despite the divergence of transcript localization between monocots and eudicots, it will be important to determine whether the NAM-related proteins colocalize with their transcripts within meristems and in which way these proteins may still be involved (or not) in boundary definition in monocots.

Importantly, the divergence of expression patterns of *NAM*-related genes in eudicots and monocots may be due to divergence of their *cis*-regulatory sequences, leading to a divergence in their transcriptional regulatory networks. Indeed, major differences between eudicot and monocot *NAM*-related genes are observable within their 5′ flanking regions, notably with respect to the presence or absence of several different CNSs and an intron in the 5′ UTRs. Evolutionarily conserved motifs have been identified in the promoter regions of a range of genes affecting plant development, examples including the *LATE ELONGATED HYPOCOTYL* and *SHOOTMERISTEMLESS* genes from *A. thaliana* (Uchida et al. 2007; Spensley et al. 2009). However, in contrast to the latter two classes, *NAM*-related genes display a clear dichotomy between currently identified monocot and eudicot genes. Moreover, 5′-UTR introns are known to be involved in the transcriptional regulation of certain genes through the action of specific DNA sequences on the level and/or specificity of their expression (see Kooiker et al. 2005, Chung et al. 2006, and Karthikeyan et al. 2009 for examples). In future studies, it will be of great interest to elucidate the regulatory role of these CNSs in both the monocot and eudicot contexts. Since *NAM*-related transcripts and *miR164* both show a divergent expression pattern in monocots compared with eudicots, it can be inferred that there has been a coevolution of their transcriptional regulatory network. One explanation is the modification of the activity or the expression pattern of common regulatory factor(s) between the two genes during angiosperm evolution.

In the context of evolutionary developmental biology, the role of *NAM/CUC3* genes is of great interest due to their involvement in the determination of plant architecture, especially through their participation in meristem functioning and boundary definition. The description of *NAM-miR164* module and *CUC3* genes in a wider panel of angiosperm species has pointed to major differences between monocots and eudicots in the way these regulatory mechanisms have been recruited to regulate development.

Supplementary Material

Supplementary tables S1 and S2 and supplementary figs. S1–S3 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

Research for this work was carried out at the IRD (Montpellier) and the Institut Jean-Pierre Bourgin (Versailles). We thank Louise Audebert for technical help with in situ hybridizations. We thank the generous support of colleagues at INRAB (Pobé Station, Benin) and FELDA (Malaysia), for the provision of oil palm material. Alain Rival (CIRAD, UMR DIAPC), David Cros and Bruno Nouy (CIRAD, PalmElit, France) are thanked for logistical help and harvesting. We are very grateful to Claudio Littardi and Robert Castellana (CSRP Sanremo, Italy) for help and access to date palm material. Finally, we wish to thank M. Aida for providing the backcrossed *A. thaliana cuc2-1* mutant. This work was supported by institutional funding from IRD and CIRAD. R. Qadri was the recipient of a Higher Education Commission fellowship from Pakistan.

References

Aberlenc-Bertossi F, Noirot M, Duval Y. 1999. BA enhances the germination of oil palm somatic embryos derived from embryogenic suspension cultures. *Plant Cell Tiss Org Cult.* 56:53–57.

Adam H, Jouannic S, Escoute J, Verdeil JL, Duval Y, Tregear JW. 2005. Reproductive developmental complexity in the African oil palm (*Elaeis guineensis*). *Am J Bot.* 92:1836–1852.

Adam H, Jouannic S, Orieux Y, Morcillo F, Richaud F, Duval Y, Tregear JW. 2007. Functional characterization of MADS box genes involved in the determination of oil palm flower structure. *J Exp Bot.* 58:1245–1259.

Aida M, Ishida T, Fukaki H, Fujisawa H, Tasaka M. 1997. Genes involved in organ separation in Arabidopsis: an analysis of the *cup-shaped cotyledon* mutant. *Plant Cell.* 9:841–857.

Aida M, Ishida T, Tasaka M. 1999. Shoot apical meristem and cotyledon formation during Arabidopsis embryogenesis: interaction among the *CUP-SHAPED COTYLEDON* and *SHOOT MERISTEMLESS* genes. *Development* 126:1563–1570.

Aida M, Tasaka M. 2006. Morphogenesis and patterning at the organ boundaries in the higher plant shoot apex. *Plant Mol Biol.* 60:915–928.

Axtell MJ, Bartel DP. 2005. Antiquity of microRNAs and their targets in land plants. *Plant Cell.* 17:1658–1673.

Axtell MJ, Bowman JL. 2008. Evolution of plant microRNAs and their targets. *Trends Plant Sci.* 13:343–349.

Baker CC, Sieber P, Wellmer F, Meyerowitz EM. 2005. The early extra petals1 mutant uncovers a role for microRNA *miR164c* in regulating petal number in Arabidopsis. *Curr Biol.* 15:303–315.

Berger Y, Harpaz-Saad S, Brabd A, Melnik H, Sirding N, Alvarez JP, Zinder M, Samach A, Eshed Y, Ori N. 2009. The NAC-domain transcription factor GOBLET specifies leaflet boundaries in compound tomato leaves. *Development* 136:823–832.

Blein T, Pulido A, Vialette-Guiraud A, Nikovics K, Morin H, Hay A, Johansen IE, Tsiantis M, Laufs P. 2008. A conserved molecular framework for compound leaf development *Science* 322:1835–1839.

Chung BY, Simons C, Firth AE, Brown CM, Hellens RP. 2006. Effect of 5'UTR introns on gene expression in *Arabidopsis thaliana*. *BMC Genomics*. 7:120.

Clough SJ, Bent AF. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.* 16:735–743.

Curtis M, Grossniklaus U. 2003. A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.* 133:462–469.

Daher A, Adam H, Chabrillange N, Collin M, Mohamed N, Tregear J, Aberlenc-Bertossi F. 2010. Cell cycle arrest characterizes the transition from a bisexual floral bud to a unisexual flower in *Phoenix dactylifera*. *Ann Bot.* 106:255–266.

Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004. VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* 32:W273–W279.

Gouy M, Guindon S, Gascuel O. 2010. Seaview version 4: a multiplateforme graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol.* 27:221–224.

Hellens RP, Edwards EA, Leyland NR, Bean S, Mullineaux PM. 2000. pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Mol Biol.* 42:819–832.

Hibara K, Nagato Y. 2005. Petunia NO APICAL MERISTEM ortholog, *OsNAM*, is expressed in the embryonic SAM and organ boundary in rice. *Rice Genet Newsletter*. 22:79–81.

Hibara K, Nagato Y. 2006. *OsNAM* and *OsCUC3* are expressed specically in organ boundaries. *Rice Genet Newsletter*. 23:96–97.

Hibara K, Karim MR, Takada S, Taoka K, Furutani M, Aida M, Tasaka M. 2006. Arabidopsis *CUP-SHAPED COTYLEDON3* regulates postembryonic shoot meristem and organ boundary formation. *Plant Cell.* 18:2946–2957.

Jasinski S, Vialette-Guiraud A, Scutt CP. 2010. The evolutionary-developmental analysis of plant microRNAs. *Phil Trans R Soc.* 365:469–476.

Jouannic S, Collin M, Vidal B, Verdeil JL, Tregear JW. 2007. A class I *KNOX* gene from the palm species *Elaeis guineensis* (Arecaceae) is associated with meristem function and a distinct mode of leaf dissection. *New Phytol.* 174:551–568.

Karthikeyan AS, Ballachanda DN, Raghothama KG. 2009. Promoter deletion analysis elucidates the role of cis elements and 5'UTR intron in spatiotemporal regulation of AtPht1;4 expression in Arabidopsis. *Physiol Plant.* 136:10–18.

Keller T, Abbott J, Moritz T, Doerner P. 2006. Arabidopsis *REGULATOR OF AXILLARY MERISTEMS1* controls a leaf axil stem cell niche and modulates vegetative development. *Plant Cell.* 18:598–611.

Kooiker M, Airoldi CA, Losa A, Manzotti PS, Finzi L, Kater MM, Colombo L. 2005. *BASIC PENTACYSTEINE1*, a GA binding protein that induces conformational changes in the regulatory region of the homeotic Arabidopsis gene *SEEDSTICK*. *Plant Cell.* 17:722–729.

Kwiatkowska D. 2008. Flowering and apical meristem growth dynamics. *J Exp Bot.* 59:187–201.

Larue CT, Wen J, Walker JC. 2009. A microRNA-transcription factor module regulates lateral organ size and patterning in Arabidopsis. *Plant J.* 58:450–463.

Laufs P, Peaucelle A, Morin H, Traas J. 2004. MicroRNA regulation of the *CUC* genes is required for boundary size control in Arabidopsis meristems. *Development* 131:4311–4322.

Le SQ, Gascuel O. 2008. An improved general amino acid replacement matrix. *Mol Biol Evol.* 25:1307–1320.

Li Y, Zheng Y, Addo-Quaye C, Zhang L, Saini A, Jagadeeswaran G, Axtell M, Zhang W, Sunkar R. 2010. Transcriptome-wide identification of microRNA targets in rice. *Plant J.* 62:742–759.

Mallory AC, Dugas DV, Bartel DP, Bartel B. 2004. MicroRNA regulation of NAC-domain targets is required for proper formation and separation of adjacent embryonic, vegetative, and floral organs. *Curr Biol.* 14:1035–1046.

Moore MJ, Bell CD, Soltis PS, Soltis DE. 2007. Using plastid genome-scale date to resolve enigmatic relationships among basal angiosperms. *Proc Natl Acad Sci USA*. 104:19363–19368.

Nikovics K, Blein T, Peaucelle A, Ishida T, Morin H, Aida M, Laufs P. 2006. The balance between the *MIR164A* and *CUC2* genes controls leaf margin serration in Arabidopsis. *Plant Cell*. 18:2929–2945.

Oikawa T, Kyozuka J. 2009. Two-Step regulation of LAX PANICLE1 protein accumulation in axillary meristem formation in rice. *Plant Cell*. 21:1095–1108.

Ooka H, Satoh K, Doi K, et al. (16 co-authors). 2003. Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. *DNA Res*. 10:239–247.

Peaucelle A, Morin H, Traas J, Laufs P. 2007. Plants expressing a *miR164*-resistant *CUC2* gene reveal the importance of post-meristematic maintenance of phyllotaxy in Arabidopsis. *Development* 134:1045–1050.

Raman S, Greb T, Peaucelle A, Blein T, Laufs P, Theres K. 2008. Interplay of *miR164*, *CUP-SHAPED COTYLEDON* genes and LATERAL SUPPRESSOR controls axillary meristem formation in *Arabidopsis thaliana*. *Plant J*. 55:65–76.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

Sieber P, Wellmer F, Gheyselinck J, Riechmann JL, Meyerowitz EM. 2007. Redundancy and specialization among plant microRNAs: role of the *MIR164* family in developmental robustness. *Development* 134:1051–1060.

Souer E, van Houwelingen A, Kloos D, Mol J, Koes R. 1996. The no apical meristem gene of Petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. *Cell* 85:159–170.

Spensley M, Kim JY, Picot E, Reid J, Ott S, Helliwell C, Carré IA. 2009. Evolutionarily conserved regulatory motifs in the promoter of the *Arabidopsis* clock gene *LATE ELONGATED HYPOCOTYL*. *Plant Cell*. 21:2606–2623.

Sunkar R, Jagadeeswaran G. 2008. In silico identification of conserved microRNAs in large number of diverse plant species. *BMC Plant Biol*. 8:37.

Takada S, Hibara K, Ishida T, Tasaka M. 2001. The *CUP-SHAPED COTYLEDON1* gene of *Arabidopsis* regulates shoot apical meristem formation. *Development* 128:1127–1135.

Tomlinson PB. 2006. The uniqueness of palms. *Bot J Linn Soc*. 151:5–14.

Uchida N, Townsley B, Chung KH, Sinha N. 2007. Regulation of SHOOT MERISTEMLESS genes via an upstream-conserved noncoding sequences coordiantes leaf development. *Proc Natl Acad Sci USA*. 104:15953–159583.

Voinnet O. 2009. Origin, biogenesis, and activity of plant microRNAs. *Cell* 136:669–687.

Vroemen CW, Mordhorst AP, Albrecht C, Kwaaitaal MA, de Vries SC. 2003. The *CUP-SHAPED COTYLEDON3* gene is required for boundary and shoot meristem formation in *Arabidopsis*. *Plant Cell*. 15:1563–1577.

Weir I, Lu J, Cook H, Causier B, Schwarz-Sommer Z, Davies B. 2004. *CUPULIFORMIS* establishes lateral organ boundaries in *Antirrhinum*. *Development* 131:915–922.

Zimmermann R, Werr W. 2005. Pattern formation in the monocot embryo as revealed by *NAM* and *CUC3* orthologs from *Zea mays* L. *Plant Mol Biol*. 58:669–685.