*Annals of Botany* Page 1 of 9
doi:10.1093/aob/mcr024, available online at www.aob.oxfordjournals.org

ANNALS OF BOTANY
Founded 1887

PART OF A SPECIAL ISSUE ON EVOLUTION AND DEVELOPMENT

# Insights from ANA-grade angiosperms into the early evolution of *CUP-SHAPED COTYLEDON* genes

**Aurélie C. M. Vialette-Guiraud[1], Hélène Adam[2], Cédric Finet[1,†], Sophie Jasinski[1,‡], Stefan Jouannic[2] and Charles P. Scutt[1,*]**

[1]*Laboratoire de Reproduction et Développement des Plantes, UMR 5667, CNRS-INRA-Université de Lyon, Ecole Normale Supérieure de Lyon, 46, allée d'Italie, 69364 Lyon Cedex 07, France and* [2]*IRD, UMR DIAPC, IRD/CIRAD Palm Development Group, BP64501, 911, avenue Agropolis, F-34394 Montpellier Cedex 5, France*
** For correspondence. E-mail Charlie.Scutt@ens-lyon.fr*
[†]*Present address: Howard Hughes Medical Institute, University of Wisconsin-Madison, 1525 Linden Drive, Madison, WI 53706, USA.*
[‡]*Present address: Institut Jean-Pierre Bourgin, INRA, Route de Saint-Cyr, 78026 Versailles Cedex, France.*

Received: 9 September 2010 Returned for revision: 12 October 2010 Accepted: 22 November 2010

• *Background and Aims* The closely related *NAC* family genes *NO APICAL MERISTEM* (*NAM*) and *CUP-SHAPED COTYLEDON3* (*CUC3*) regulate the formation of boundaries within and between plant organs. *NAM* is post-transcriptionally regulated by *miR164*, whereas *CUC3* is not. To gain insight into the evolution of *NAM* and *CUC3* in the angiosperms, we analysed orthologous genes in early-diverging ANA-grade angiosperms and gymnosperms.
• *Methods* We obtained *NAM*- and *CUC3*-like sequences from diverse angiosperms and gymnosperms by a combination of reverse transcriptase PCR, cDNA library screening and database searching, and then investigated their phylogenetic relationships by performing maximum-likelihood reconstructions. We also studied the spatial expression patterns of *NAM*, *CUC3* and *MIR164* orthologues in female reproductive tissues of *Amborella trichopoda*, the probable sister to all other flowering plants.
• *Key Results* Separate *NAM* and *CUC3* orthologues were found in early-diverging angiosperms, but not in gymnosperms, which contained putative orthologues of the entire *NAM* + *CUC3* clade that possessed sites of regulation by *miR164*. Multiple paralogues of *NAM* or *CUC3* genes were noted in certain taxa, including Brassicaceae. Expression of *NAM*, *CUC3* and *MIR164* orthologues from *Am. trichopoda* was found to co-localize in ovules at the developmental boundary between the chalaza and nucellus.
• *Conclusions* The *NAM* and *CUC3* lineages were generated by duplication, and *CUC3* was subsequently lost regulation by *miR164*, prior to the last common ancestor of the extant angiosperms. However, the paralogous *NAM* clade genes *CUC1* and *CUC2* were generated by a more recent duplication, near the base of Brassicaceae. The function of *NAM* and *CUC3* in defining a developmental boundary in the ovule appears to have been conserved since the last common ancestor of the flowering plants, as does the post-transcriptional regulation in ovule tissues of *NAM* by *miR164*.

**Key words:** CUP-SHAPED COTYLEDON, CUC, NO APICAL MERISTEM, NAM, NAC, MIR164, *Amborella trichopoda*, *Cabomba aquatica*, *Ginkgo biloba*, angiosperm, gymnosperm.

## INTRODUCTION

The genes *NO APICAL MERISTEM* (*NAM*) in *Petunia hybrida* and *CUP-SHAPED COTYLEDON3* (*CUC3*) in *Arabidopsis thaliana* encode distinct members of the large NAC family of plant-specific transcription factors (Ooka *et al.*, 2003). NAC proteins contain a highly conserved, DNA-binding 'NAC' domain at their N terminus, and a more variable C-terminal region (Ernst *et al.*, 2004). *NAM*, *CUC3* and their respective orthologues in diverse angiosperms, which are here referred to generically as *CUC* genes, are expressed at, and control the formation of, tissue boundaries both within and between plant organs. Accordingly, the inactivation of *CUC* genes reveals a range of unique, redundant and partially overlapping phenotypes in processes including: shoot apical meristem (SAM) establishment, lateral meristem formation, lateral organ separation, leaf lobing and ovule development (Souer *et al.*, 1996; Aida *et al.*, 1997; Ishida *et al.*, 2000; Vroemen *et al.*, 2003; Hibara *et al.*, 2006; Nikovics *et al.*, 2006). In ovule development, *CUC* genes are expressed at and define the boundary between the nucellus and chalaza in several distantly related eudicots (Souer *et al.*, 1996; Ishida *et al.*, 2000; Weir *et al.*, 2004; Hibara *et al.*, 2006). The degree of genetic redundancy among *CUC* genes varies between species. For example, the inactivation of the single genes *NAM* in *P. hybrida* (Souer *et al.*, 1996) and *CUPULIFORMIS* (*CUP*) in *Antirrhinum majus* (Weir *et al.*, 2004) causes cotyledon fusion and prevents SAM formation, whereas in *Ar. thaliana* any two of the genes *CUC1*, *CUC2* and *CUC3* must be mutated to produce the equivalent phenotype (Aida *et al.*, 1997; Hibara *et al.*, 2006).

*Arabidopsis thaliana CUC1* and *CUC2* are post-transcriptionally regulated by *miR164* (Laufs *et al.*, 2004; Mallory *et al.*, 2004), and both *P. hybrida NAM* and *An. majus CUP* also possess target sites for this microRNA. This regulatory mechanism appears to control the balance between tissue separation and fusion in various different situations in the plant. For example, the post-transcriptional regulation of *CUC2* in *Ar. thaliana* is necessary for fusion to take place between the two carpels of the syncarpic gynoecium, while the balance of *CUC2* and *miR164* expression also controls the depth of leaf lobes (Nikovics *et al.*, 2006; Sieber *et al.*, 2007; Larue *et al.*, 2009). Interestingly, a similar mechanism controls leaf dissection and compound leaf formation in several distantly related eudicots in which these morphological traits were clearly shown to be of independent origin (Blein *et al.*, 2008; Berger *et al.*, 2009). The regulation of *CUC* genes by *miR164* thus seems to form a conserved genetic module that has been repeatedly recruited during angiosperm evolution to both leaf dissection and floral organ fusion.

To provide insights into the early evolution of *CUC* genes in the angiosperms, we studied their phylogenetic relationships, in both angiosperms and their sister group, the gymnosperms. To make conclusions on the *CUC* genes that were present in the last common ancestor of the extant angiosperms, we included in our analyses representatives of the ANA grade (ANA for Amborellales, Nymphaeales and Austrobaileyales), which, according to molecular phylogenetic analyses (Bremer *et al.*, 2009), comprises the first three extant lineages to have diverged from a common remaining lineage from which all other living angiosperms are descended. In particular, we focused on *Amborella trichopoda*, a shrub endemic to New Caledonia, which, as the only representative of Amborellales, is the probable sister to all other extant angiosperms. This work is complementary to an earlier study of *miR164* evolution, which identified the presence of at least two *MIR164* genes in the last common ancestor of the extant angiosperms (Jasinski *et al.*, 2010).

The present study investigated the origin, through duplication, of the angiosperm *NAM* and *CUC3* lineages, and the presence or absence of post-transcriptional regulation in these lineages in the last common ancestor of the extant angiosperms. Studies in *Am. trichopoda* have demonstrated *NAM* and *CUC3* orthologues to be expressed between the nucellus and chalaza in the ovule, as is also the case in eudicots, suggesting the conservation of *CUC* gene function in defining this tissue boundary throughout angiosperm evolution. As the *NAM* orthologue from *Am. trichopoda* appears to be regulated by *miR164*, we also investigated the expression of this microRNA in ovule tissues. The results suggest that the *NAM/miR164* genetic module has operated in ovule tissues since the last common ancestor of the extant angiosperms.

## MATERIALS AND METHODS

### *Plant material*

Material of *Amborella trichopoda* Baill. was field-collected at Col d'Amieu, New Caledonia, and that of *Ginkgo biloba* L. was obtained from the City of Lyon Botanic Garden. Plants of *Cabomba aquatica* Aublet were obtained from Anthias SA (Les Chères, Rhône, France) and grown to maturity in a small aquarium.

### *NAC gene identification*

RNA was extracted (Chang *et al.*, 1993) from plant reproductive tissues and reverse transcribed into cDNA using M-MuLV reverse transcriptase (Fermentas, Hanover, MD, USA). Conserved regions of *NAM* and *CUC3* homologues, within the 5′ *NAC* domain, were amplified by PCR using the partially degenerate primers: 5′-GGTTYCAYCCNACTGAYGARGAGCT and 5′-CTGCANATNACCCATTCYTCCTT. DNA fragments thus obtained were sequenced and used in initial phylogenetic analyses (data not shown) to determine their possible orthology to *P. hybrida NAM* and/or *Ar. thaliana CUC3*. Amplified fragments of possible *NAM* and/or *CUC3* orthologues were then used to obtain longer or full-length cDNAs by screening cDNA libraries, prepared as previously described (Fourquin *et al.*, 2005; Finet *et al.*, 2010). The sequence of *GbiNACa* from *G. biloba* was completed by RACE (rapid amplification of cDNA ends) PCR using a Marathon (Clontech, Mountain View, CA, USA) kit, while *CaqNACa* was initially identified from a *C. aquatica* flower expressed sequence tag database (C. Scutt and C. Finet, unpubl. data). Further possible *NAM* and *CUC3* orthologues, and closely related *NAC* genes from diverse angiosperms and gymnosperms, were identified by BLAST searching of publicly available databases.

### *Phylogenetic analyses*

Amino acid sequences from *NAC* genes were aligned using MUSCLE in the SEAVIEW program (Gouy *et al.*, 2010) and homologous sites for phylogenetic reconstructions were chosen manually. Maximum-likelihood phylogenies were then generated from these alignments in PhyML (Guindon *et al.*, 2009) using an LG evolutionary model. Amino acid alignments were also back-translated using TranAlign (http://emboss.sourceforge.net/), thus conserving homologous sites determined from amino acid alignments, and used to generate maximum-likelihood phylogenies in PhyML using a GTR evolutionary model. Statistical support for all phylogenetic analyses was provided by performing 1000 bootstrap replicates.

### In situ *hybridizations using polynucleotide probes*

Expression of *AtrNAM* and *AtrCUC3* in female reproductive tissues was detected using digoxygenein-labelled antisense riboprobes synthesized using T7 RNA polymerase (Promega, Madison, WI, USA) and DIG-RNA labelling mix (Roche, Basel, Switzerland) from full-length cDNAs amplified as PCR products to incorporate bacteriophage T7 promoter sequences. Negative and positive control hybridizations were also performed, using sense strand riboprobes and an antisense strand riboprobe of a *HISTONE4* gene from *Am. trichopoda*, respectively. *In situ* hybridizations were carried out as described by Nikovics *et al.* (2006), except that hybridizations were performed in solutions containing 50 % (v/v) formamide, 10 % (w/v) dextran sulphate, 3·5 mg $mL^{-1}$ tRNA, 2·5 × Denhardt's reagent, 0·3 M NaCl, 10 mM Tris/HCl, 10 mM sodium phosphate

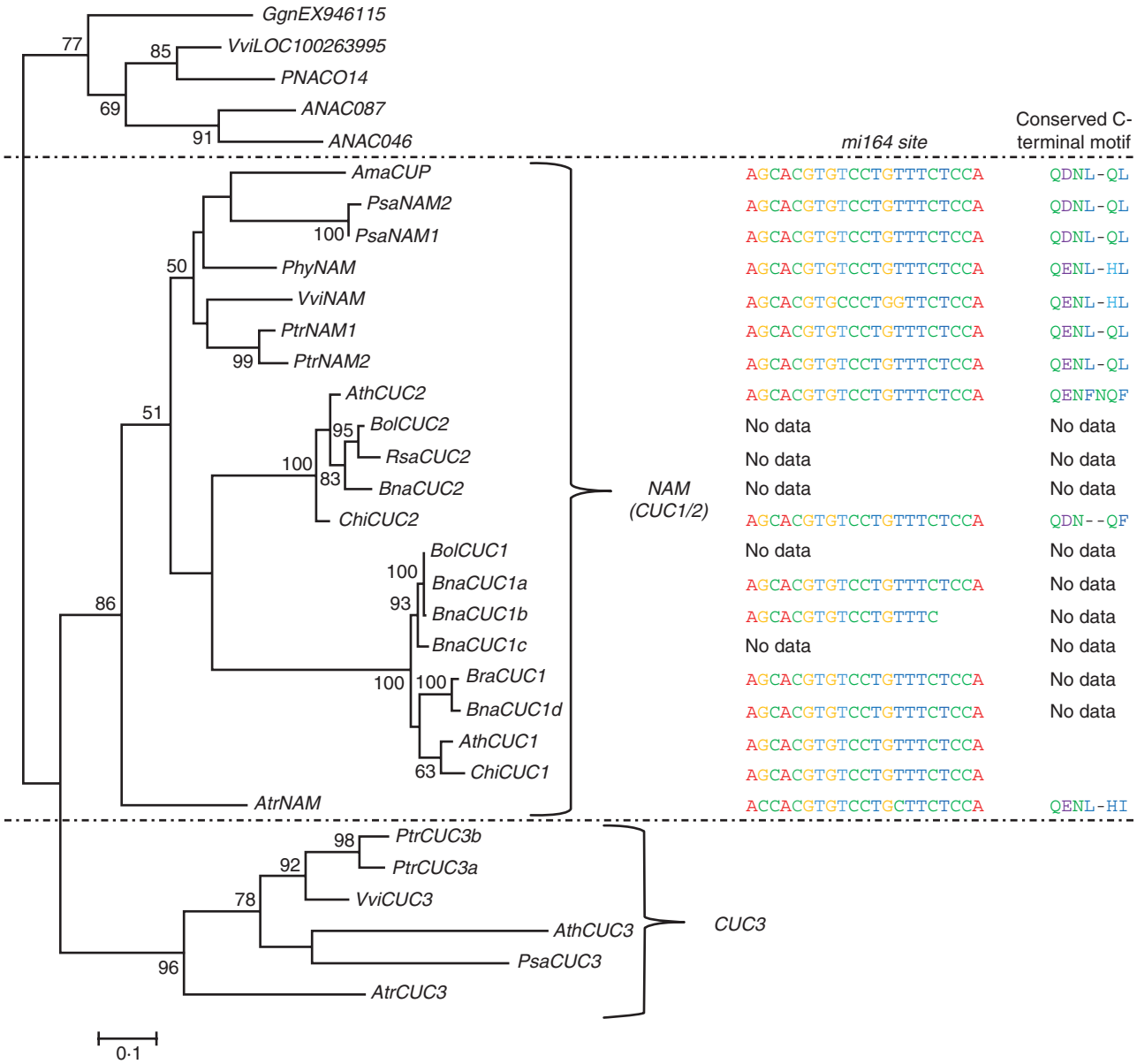


FIG. 1. A maximum-likelihood phylogenetic analysis of the *NAM* and *CUC3* clades of the *NAC* family. Gene name prefixes indicate the source organisms: *Ama*, *Antirrhinum majus*; *ANAC* or *Ath*, *Arabidopsis thaliana*; *Atr*, *Amborella trichopoda*; *Bna*, *Brassica napus*; *Bol*, *Brassica oleracea*; *Bra*, *Brassica rapa*; *Chi*, *Cardamine hirsuta*; *Ggn*, *Gnetum gnemon*; *Phy*, *Petunia hybrida*; *PNAC* or *Ptr*, *Populus trichocarpa*; *Psa*, *Pisum sativum*; *Rsa*, *Raphanus sativus*; *Vvi*, *Vitis vinifera*. Full sequence accession details are given in Table S1, and the nucleotide sites used to generate the phylogeny are shown in Fig. S1. Bootstrap support values ≥50 % are indicated at corresponding nodes, while sites of regulation by *miR164*, and a partially conserved C-terminal protein motif of unknown function, are shown to the right of the corresponding sequences.

and 5 mM EDTA (pH 6·8). Photographs were taken using an AxioCam MRc cooled camera (Zeiss, Jena, Germany) in conjunction with a Zeiss AxioImager M2 microscope, and images were processed with Photoshop software (Adobe, San Jose, CA, USA).

### In situ *hybridizations using oligonucleotide probes*

Expression of *miR164* in female reproductive tissues was detected using a 5′-digoxigenin-labelled miRCURY LNA probe (Exiqon, Vedbaek, Denmark) complementary to the predicted mature *Atr-miR164a* sequence from *Am. trichopoda* (Jasinski *et al.*, 2010). An LNA oligonucleotide complementary to murine *miR124* (Lagos-Quintana *et al.*, 2002), which has no predicted target sequences in known plant genomes, was used as a negative control. *In situ* hybridizations were carried out as described by Adam *et al.* (2007) at probe concentrations of 0·02 mM. Detection was performed using the Vector Blue Alkaline Phosphatase Substrate Kit III (Vector Laboratories, Burlingame, CA, USA). Photographs were taken using an Evolution MP 5·0 cooled camera (MediaCybernetics, Bethesda, MD, USA) in conjunction with a DMRB microscope (Leica, Solms, Germany), and images were processed with Photoshop software (Adobe).

## RESULTS

### *The* NAM *and* CUC3 *lineages separated prior to the radiation of the extant angiosperms*

To gain insight into the evolution of *CUC* genes in the angiosperms, we performed phylogenetic analyses using sequences from both angiosperms and gymnosperms, including *Am. trichopoda* (Amborellales, Amborellaceae) and *C. aquatica* (Nymphaeales, Cabombaceae), as representatives of the two earliest diverging flowering plant lineages (see Supplementary Data Table S1). We first performed analyses on the entire data set obtained (Fig. S1) to identify the relationship of major clades containing *CUC* genes.

Phylogenies obtained using both amino acid and nucleotide alignments (Fig. S2) yielded trees of similar overall topology in which a grade containing the genes *PNAC111*, *PNAC109*, *VviLOC100254997* and *ANAC074* grouped externally to all remaining sequences, followed by a clade containing the *Ar. thaliana* gene *NAC1*, which mediates auxin signalling in lateral root development (Xie *et al.*, 2000). Within the remaining clade of sequences analysed, sub-clades respectively containing *P. hybrida NAM* (*PhyNAM*) and *Ar. thaliana CUC3* (*AthCUC3*) occupied sister positions in both of the analyses performed. However, in the phylogeny produced using a nucleic acid alignment (Fig. S2B), a group of Poaceae monocot sequences grouped externally to the clade containing *PhyNAM*, as compared with its position in the phylogeny performed using an amino acid alignment (Fig. S2A). This effect may be due to highly biased codon usage in Poaceae. Another unstable feature of our analyses occurred in the position of *CaqNACa*, a possible *NAM* orthologue from *C. aquatica* (Nymphaeales). This sequence occupied a basal position within the *NAM* clade in a nucleic acid-based phylogeny (Fig. S2B), but appeared basal to the entire *NAM* + *CUC3* clade in an amino acid-based phylogeny (Fig. S2A). Following our initial studies, we performed a phylogenetic analysis using a sub-set of sequences centred on the sister clades containing *PhNAM* and *AthCUC3* to better resolve the nodes within these, but omitting the Poaceae and *CabNACa* sequences which had previously shown variable positions. This analysis (Fig. 1) indicated the internal structure of the *NAM* and *CUC3* clades to broadly recapitulate angiosperm phylogeny, with sequences from the ANA-grade angiosperm *Am. trichopoda* occupying basal positions in both clades. Thus, despite the apparent instability of certain sequences in our phylogenies, the presence of separate *NAM* and *CUC3* orthologues in diverse flowering plant groups, including the earliest diverging order Amborellales, strongly suggests these lineages to have separated through a gene duplication event prior to the last common ancestor of the extant angiosperms.

In phylogenetic analyses performed using a nucleic acid alignment from an extensive *NAC* family data set (Fig. S2B), a clade of four gymnosperm genes occupied a sister position to the combined angiosperm *NAM* + *CUC3* clade. Like NAM orthologues, these gymnosperm genes possess sites of regulation by *miR164* (Fig. S3). However, bootstrap support for the phylogenetic placement of these genes was not strong, and only one of them, *PtaEST1163314*, occupied the same position in phylogenetic analyses performed using an amino acid alignment (Fig. S2A). Although questions thus remain concerning the precise orthology of the *NAM/CUC3*-like genes identified here from gymnosperms, it appears significant that none of these genes grouped within the angiosperm *NAM* and *CUC3* clades in any of our phylogenetic analyses. These findings are consistent with the possibility that the *NAM* and *CUC3* lineages separated specifically in the angiosperm lineage, after its divergence from that of the extant gymnosperms. However, other interpretations of these data are possible. For example, separate gymnosperm orthologues of *NAM* and *CUC3* might have been missed in cloning procedures, or may have existed but have been lost from the gymnosperm clade subsequent to the separation of the angiosperm and gymnosperm lineages.

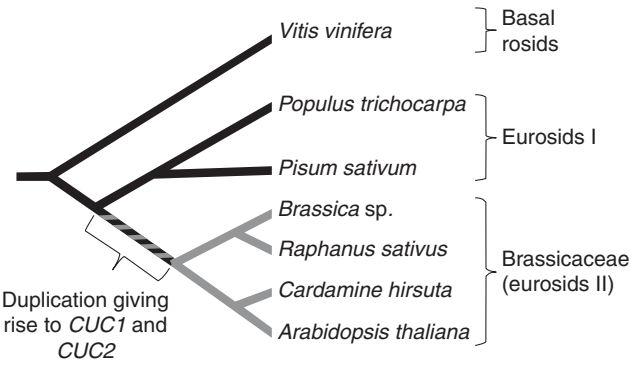


FIG. 2. Phylogenetic relationships of selected eurosid species showing the approximate position of the duplication that generated the *CUC1* and *CUC2* paralogues in Brassicaceae.

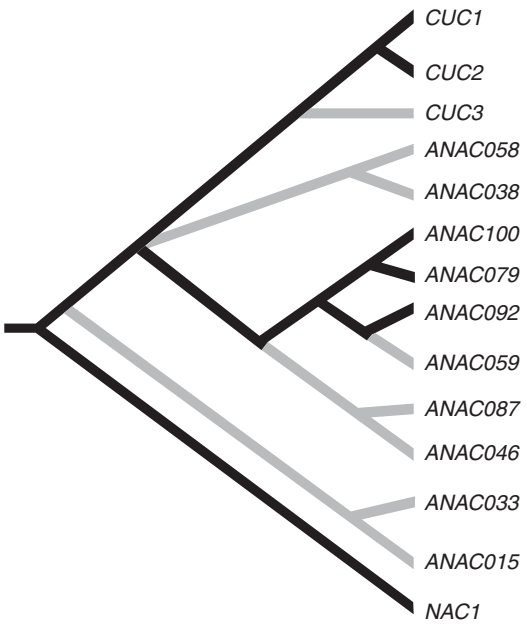


FIG. 3. Post-transcription regulation by *miR164* mapped onto a schematic phylogeny of *Ar. thaliana NAC* genes. The presence of sites of regulation by *miR164* that are known or predicted to be functional is indicated by black lines, while the loss of these is indicated by grey lines.

### *The* CUC1 *and* CUC2 *lineages separated by gene duplication near the base of Brassicaceae*

The largely redundant *Ar. thaliana* paralogues *CUC1* and *CUC2* are positioned in our phylogenetic analysis in each of

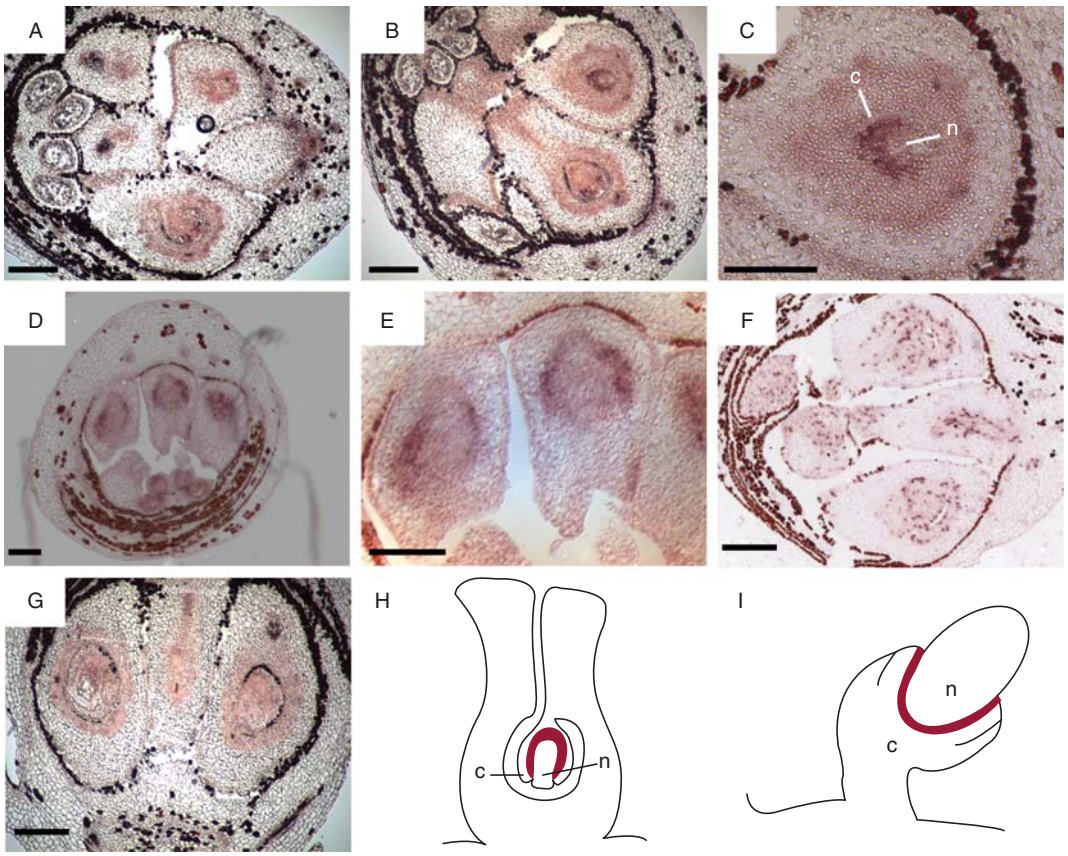


FIG. 4. Expression of *AtrNAM* and *AtrCUC3* in the *Am. trichopoda* gynoecium. *In situ* hybridizations to sections of female flower buds showing expression of *AtrCUC3* (A–C) and *AtrNAM* (D and E) in the zone separating the nucellus (n) and chalaza (c) of the ovule. Slight expression of these genes is also visible between the carpels. Control hybridizations are provided by a *HISTONE4* probe (F), which marks all rapidly dividing cells, and by a sense-strand *AtrCUC3* probe (G), which produces no significant hybridization signal. Diagrams illustrating the ovule arrangement in *Am. trichopoda* (H) and, for comparison, in *Ar. thaliana* (I) indicate in red the domain of *NAM* and *CUC3* expression between the nucellus and chalaza. Scale bars = 200 μm.

two sub-clades occupying sister positions within the main *NAM* clade of angiosperm NAC genes. Each of these sub-clades exclusively contains genes from Brassicaceae (Fig. 1), which groups within the eurosids II clade (Bremer *et al.*, 2009). The closest relatives of Brassicaceae sampled in this study are *Pisum sativum* and *Populus trichocarpa* from the eurosids I clade, which, like the other species external to Brassicaceae included in our analyses, do not appear to contain separate orthologues of *CUC1* and *CUC2*. We therefore conclude that the duplication that generated *CUC1* and *CUC2* occurred after the separation of eurosids I and II, and before the relatively early speciation event within Brassicaceae (Bailey *et al.*, 2006) that separated the common lineage of *Brassica* and *Raphanus* from that of *Arabidopsis* and *Cardamine*. These conclusions are summarized in Fig. 2. Further sampling within eurosids II could be used to localize this gene duplication event more precisely. At least one further gene duplication event within the *CUC1* clade appears to have occurred to generate two *CUC1* paralogues per diploid genome in *Brassica*, thus explaining the presence of four *CUC1* paralogues in the tetraploid genome of *Brassica napus* (Fig. 1). Our phylogenetic analyses also reveal the presence of recent gene duplications outside Brassicaceae in both the *NAM* and the *CUC3* clades, leading for example to the paralogues *PsaNAM1* and *PsaNAM2* in *Pisum sativum* and *PtrCUC3a* and *PtrCUC3b* in *Populus trichocarpa* (Fig. 1). Zimmermann and Werr (2005) noted the presence of a conserved motif, termed Motif 3, within the C-terminal region of certain *NAM*-like genes. Given the phylogeny of the *NAM* clade, as shown in Fig. 1, the absence of this motif in Brassicaceae CUC1 proteins (Fig. S4) suggests that sequences encoding it might have been lost from the *CUC1* lineage in a common ancestor of Brassicaceae.

### CUC3 *lost its regulation by* miR164 *before the radiation of the extant angiosperms*

All angiosperm *NAM* genes for which sequence data from the 3′-end of the coding region were available to this study were found to contain sites of potential regulation by *miR164* (Fig. 1), as were possible gymnosperm orthologues of the combined *NAM/CUC3* clade (Fig. S3). However, all *CUC3* orthologues analysed to date, including *AtrCUC3* from the ANA-grade angiosperm *Am. trichopoda*, lack sites of regulation by *miR164*. As all sites of regulation by *miR164* occur in the same position in their respective *NAC* genes, these sites appear to be homologous, and as *CUC3* groups within a clade of *NAC* genes that contain sites of regulation by *miR164*, we conclude the *CUC3* lineage to have previously possessed a site of regulation by *miR164*, but to have lost this site before the radiation of the extant

angiosperms. *miR164* is predicted to regulate seven genes in the *Ar. thaliana* genome (Gustafson *et al.*, 2005), of which *NAC1* is the most distantly related to the *Ar. thaliana* genes of the *NAM* clade, *CUC1* and *CUC2* (Fig. S2). In addition to *CUC3*, our phylogenetic analyses (Fig. S2) show seven further *Ar. thaliana* genes that are not predicted to be *miR164* targets, but which occupy phylogenetic positions that are intermediate between *NAC1* and the *NAM* clade. Of these genes, *ANAC015*, *ANAC038*, *ANAC046*, *ANAC058* and *ANAC087* show high numbers of mismatches to the *miR164* consensus sequence (13–17 of 21 nucleotides), suggesting ancient losses of this feature. *ANAC059*, by contrast, which is closely related to the predicted *miR164* target *ANAC092*, shows only eight mismatches to the *miR164* consensus sequence, suggesting the recent loss of post-transcriptional regulation in this gene. Mapping of predicted regulation by *miR164* onto a summarized phylogeny of *Ar. thaliana NAC* genes (based on Fig. S1A) suggests at least five independent losses of regulation by *miR164* have occurred (Fig. 3).

### *Expression patterns of* NAM, CUC3 *and* miR164 *orthologues suggest the conservation of function in ovule development since the last common ancestor of the extant flowering plants*

An interest in female reproductive development led us to study the expression patterns of the *NAM* and *CUC3* orthologues, *AtrNAM* and *AtrCUC3*, in female flowers of *Am. trichopoda.* Expression of these genes was found to co-localize in the peripheral region of the ovule, corresponding to the boundary of the nucellus and chalaza (Fig. 4). This expression pattern strongly resembles those of the orthologues of *AtrNAM* and *AtrCUC3* in *Ar. thaliana* (Ishida *et al.*, 2000; Hibara *et al.*, 2006) and those of *AtrNAM* in *P. hybrida* (Souer *et al.*, 1996) and *An. majus* (Weir *et al.*, 2004). These data suggest that *NAM* and *CUC3* define a developmental boundary in the *Am. trichopoda* ovule, as do the paralogous *NAM* clade genes *CUC1* and *CUC2* in *Ar. thaliana* (Ishida *et al.*, 2000). Such conservation of expression in a very precise tissue boundary strongly suggests the conservation of function in the studied plant lineages since the last common ancestor of the living flowering plants.

As *AtrNAM* possesses a site of regulation by *miR164*, we also performed *in situ* hybridizations using an LNA oligonucleotide probe corresponding to the known mature *miR164* sequence from *Am. trichopoda* (Jasinski *et al.*, 2010). These hybridizations (Fig. 5) demonstrated co-localization of *miR164* expression with that of its putative *NAM* target *AtrNAM* in ovule tissues, closely resembling the expression pattern of *miR164* in the eudicot *Nicotiana benthamiana* (Valoczi *et al.*, 2006). These results strongly suggest that the post-transcriptional regulation of *NAM* by *miR164* has been conserved in the ovule since the radiation of the extant angiosperms.

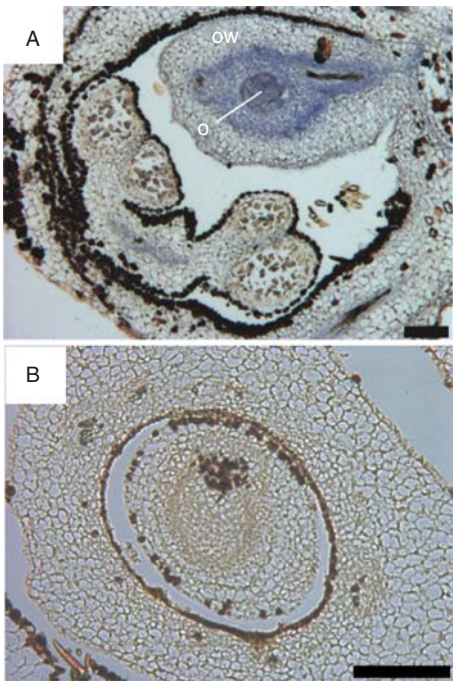


FIG. 5. Expression of *miR164* in the *Am. trichopoda* gynoecium. *In situ* hybridizations to sections of female flower buds showing expression of *Atr-miR164* (A) in the ovule, and lack of staining using a negative control probe complementary to murine *miR124* (B). Scale bars = 100 μm. Abbreviations: o, ovule; ow, ovary wall.

## DISCUSSION

### *A partial reconstruction of* NAM *and* CUC3 *evolution in the angiosperms*

We have analysed *NAC* family sequences from angiosperms and gymnosperms to clarify a number of points in the evolution of *CUC* genes in the flowering plants. We incorporated sequences from early-diverging angiosperms and gymnosperms in these analyses to infer the structure of the *NAC* family in the last common ancestor of the extant angiosperms. We also included all available *NAM* clade sequences from Brassicaceae to resolve questions relating to the evolutionary origin of the much studied paralogues *CUC1* and *CUC2* in *Ar. thaliana.*

The results of our analyses, which are summarized in Fig. 6, clearly indicate that both the *NAM* and *CUC3* lineages were present in the last common ancestor of the extant flowering plants. The apparent absence of separate orthologues of *NAM* and *CUC3* in gymnosperms suggests the duplication that generated these paralogous lineages occurred after the separation of the extant angiosperms and gymnosperms. Gymnosperm genes that group closely to the combined angiosperm *NAM* + *CUC3* clades in phylogenetic analyses have been identified in our analyses, although further work will be needed to adequately support the possible orthology of some or all of these to genes of the angiosperm *NAM* + *CUC3* clade.

Our analyses have also resolved a question relating to the evolutionary origin of *CUC1* and *CUC2* in *Ar. thaliana.* Earlier phylogenies of *NAM* genes, in which only *Ar. thaliana* was included to represent Brassicaceae, showed *CUC1* to group externally to a clade in which *Ar. thaliana CUC2* was present, together with *NAM* genes from distantly related eudicots such as *An. majus* and *P. hybrida* (Ooka *et al.*, 2003; Vroemen *et al.*, 2003; Weir *et al.*, 2004; Zimmermann and Werr, 2005; Blein *et al.*, 2008).

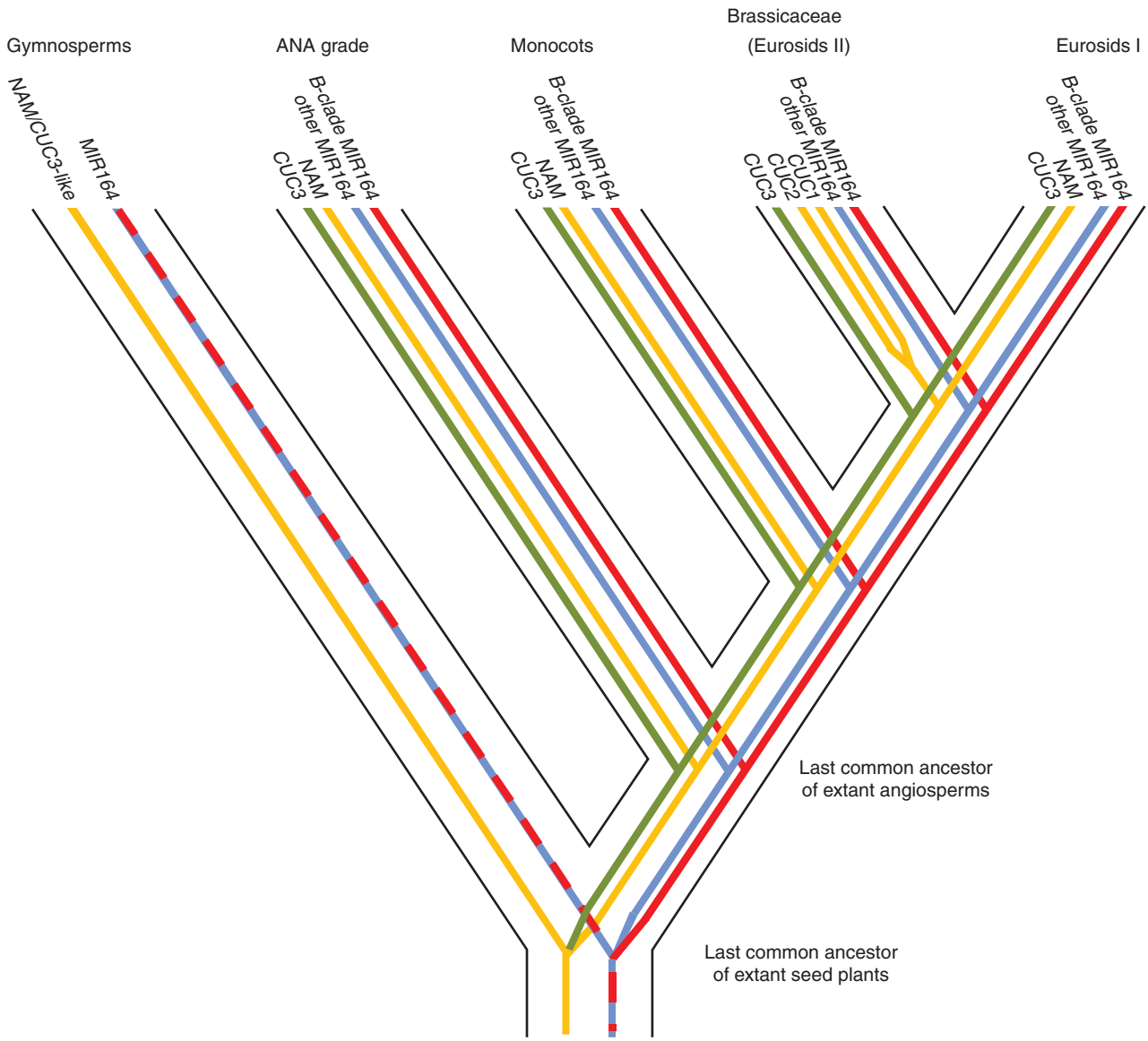


FIG. 6. Summary of *CUC* and *MIR164* gene evolution in the angiosperms. A duplication event prior to the radiation of the angiosperms generated the *NAM* and *CUC3* lineages, and a later duplication in eurosids II generated the *CUC1* and *CUC2* lineages present in Brassicaceae. At least two *MIR164* gene clades ('B-clade' and 'other') were present in the last common ancestor of the extant angiosperms (Jasinski *et al.*, 2010) and have persisted in both ANA-grade and more recently diverged angiosperm lineages.

However, our phylogenetic analysis (Fig. 1), which incorporates four genera that form two clades within Brassicaceae, better resolves the *NAM* clade, with the result that *CUC1* and *CUC2* appear as paralogous lineages that are unique to Brassicaceae, at least among the taxa included in our analysis (as summarized in Figs 2 and 6).

The basal positioning of *CUC1* observed in the phylogenies cited above suggests these analyses were distorted by long-branch attraction, possibly due to rapid evolution in the *CUC1* lineage. Such rapid evolution might suggest *CUC1* to have evolved specific functions through the positive selection of non-silent changes to its coding sequence. Specific functions of *CUC1* in cotyledon development, which might support this hypothesis, are revealed in plants in which *PIN-FORMED1* (*PIN1*) and *PINOID* (*PID*), which act to limit *CUC* expression in cotyledon tissues, are mutated. Accordingly, *cuc1 pin1* double mutants show enhanced defects in cotyledon separation compared with *cuc2 pin1* double mutants (Aida *et al.*, 2002), while *cuc1 pin1 pid* triple mutants show a greater recovery of the cotyledon development that is lost in *pin1 pid* double mutants compared with *cuc2 pin1 pid* triple mutants (Furutani *et al.*, 2004). However, these effects may reflect differential regulation of *CUC1* and *CUC2* in response to auxin, rather than differences in the biochemical properties of the proteins encoded by these genes. Thus, an alternative to positive selection operating on *CUC1* is relaxed selection, which may have allowed the coding sequence of this gene to change more rapidly than that of its paralogue *CUC2*. Indeed, both positive and relaxed selection may have operated on distinct domains of *CUC1*.

In addition to gene duplications in the *NAC* family, we have also addressed the question of the post-transcriptional regulation of *NAC* genes by *miR164*. We conclude that a loss of regulation by this microRNA has occurred in at least five *NAC* family lineages in the angiosperms, including that of *CUC3*, and further show that loss of post-transcriptional regulation in this lineage preceded the radiation of the extant angiosperms. By contrast, we have found no loss of regulation by *miR164* in any gene of the angiosperm *NAM* clade. It is interesting to note that the conservation of regulation by *miR164* in *NAM* genes since the last common ancestor of the extant angiosperms is paralleled by conservation in the structure of the *miR164* gene family (Fig. 6). Accordingly, at least two *miR164* genes appear to have been present in the last common ancestor of the extant flowering plants and these gene clades have been conserved to the present day in several distantly related angiosperm lineages, including those of the ANA grade (Jasinski *et al.*, 2010).

### *The function of* NAM *and* CUC3 *in early angiosperms*

We have focused here on the earliest diverging of all angiosperm lineages, uniquely represented by *Am. trichopoda*, to gain insight into the *CUC* genes present in the last common ancestor of the extant angiosperms. We show that expression of *NAM* and *CUC3* orthologues co-localizes with a developmental boundary that forms between the nucellus and chalaza of the *Am. trichopoda* ovule. This highly specific expression pattern appears to be entirely conserved between *Am. trichopoda* and diverse eudicots (Souer *et al.*, 1996; Ishida *et al.*, 2000; Weir *et al.*, 2004; Hibara *et al.*, 2006), suggesting its conservation since the last common ancestor of the extant angiosperms. We argue that conservation of such a highly specific expression pattern, in a developmental boundary that is also conserved between the species under comparison, implies the conservation of gene function. A full test of this hypothesis will, however, have to await the development of functional genetics approaches in *Am. trichopoda* or other ANA-grade angiosperms.

Our study has also demonstrated co-localization of *miR164* expression with that of its putative target *AtrNAM*, in *Am. trichopoda* ovule tissues, closely resembling the situation in eudicots (Valoczi *et al.*, 2006). No specific phenotype has yet been associated with the expression of *miR164* in ovules. However, expression of *miR164* in *Ar. thaliana* leaf and other tissues is known to modulate, rather than eliminate, the expression of its *CUC* gene targets (Laufs *et al.*, 2004; Nikovics *et al.*, 2006). Accordingly, the co-localization of *miR164* and *NAM* expression reported here suggests this fine-tuning mechanism also to operate in ovule tissues, and to have been conserved at this location since the last common ancestor of the extant flowering plants.

*NAM* genes and their *miR164* regulator are known to form a genetic module that has been recruited many times independently to processes such as compound leaf development and leaf lobe formation (Nikovics *et al.*, 2006; Blein *et al.*, 2008; Berger *et al.*, 2009). In addition, this module is known to be involved in the fusion of carpels into a syncarpic gynoecium, a process which has also evolved many times independently in the angiosperms (Armbruster *et al.*, 2002). For a genetic module to remain in a state of readiness to be recruited to such evolutionarily labile processes, it may also play some more constant developmental role that acts to maintain it over long evolutionary periods. Thus, the functions of *NAM*, *CUC3* and *miR164* in ovule development, which the results presented here suggest to have persisted throughout flowering plant evolution, in addition to basic functions in the SAM, may represent conserved roles from which these genes have been recruited to other more variable functions, such as leaf dissection, during angiosperm evolution.

## SUPPLEMENTARY DATA

Supplementary data are available online at www.aob.oxfordjournals.org and consist of the following. Fig. S1: Sequence alignments used in phylogenetic analyses. Fig. S2: Maximum-likelihood phylogenetic analyses of *CUC* genes and related sequences in angiosperms and gymnosperms. Fig. S3: Conservation of *miR164* target sites in *CUC*-like genes from angiosperms and gymnosperms. Fig. S4: Conserved domains in selected CUC proteins. Table S1: Accession numbers of sequences analysed.

## ACKNOWLEDGEMENTS

We thank Drs Marie Sémon and Pascale Chevret for advice on phylogenetic analyses, Dr Frédéric Pautz for access to the living collection of Lyon Botanic Garden, Drs Bruno Fogliani, Jérôme Munzinger, Marc Ducousso and Gildas Gateblé for help with field collection of *Am. trichopoda*, Dr Patrick Laufs for helpful discussions, and the technical staff of the RDP for their support. A.C.M.V-G. and C.F. received doctoral thesis funding from the French Ministry of Sciences and S.J. was supported by an ATER post of the ENS-Lyon. Our research group is supported by grant ANR-07-BLAN-0211-01 from the French National Research Agency.

## LITERATURE CITED

**Adam H, Jouannic S, Orieux Y, *et al.* 2007.** Functional characterization of MADS box genes involved in the determination of oil palm flower structure. *Journal of Experimental Botany* **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, Vernoux T, Furutani M, Traas J, Tasaka M. 2002.** Roles of *PIN-FORMED1* and *MONOPTEROS* in pattern formation of the apical region of the Arabidopsis embryo. *Development* **129**: 3965–3974.

**Armbruster WS, Debevec EM, Willson MF. 2002.** Evolution of syncarpy in angiosperms: theoretical and phylogenetic analyses of the effects of carpel fusion on offspring quantity and quality. *Journal of Evolutionary Biology* **15**: 657–672.

**Bailey CD, Koch MA, Mayer M, *et al.* 2006.** Toward a global phylogeny of the Brassicaceae. *Molecular Biology and Evolution* **23**: 2142–2160.

**Berger Y, Harpaz-Saad S, Brand A, *et al.* 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, *et al.* 2008.** A conserved molecular framework for compound leaf development. *Science* **322**: 1835–1839.

**Bremer B, Bremer K, Chase MW, *et al.* 2009.** An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* **161**: 105–121.

**Chang S, Puryear J, Cairney J. 1993.** A simple and efficient method for isolating RNA from pine tree. *Plant Molecular Biology Reporter* **11**: 114–117.

**Ernst HA, Olsen AN, Skriver K, Larsen S, Lo Leggio L. 2004.** Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors. *Embo Reports* **5**: 297–303.

**Finet C, Fourquin C, Vinauger M, *et al*. 2010.** Parallel structural evolution of Auxin Response Factors in the angiosperms. *The Plant Journal* **63**: 952–959.

**Fourquin C, Vinauger-Douard M, Fogliani B, Dumas C, Scutt CP. 2005.** Evidence that *CRABS CLAW* and *TOUSLED* have conserved their roles in carpel development since the ancestor of the extant angiosperms. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 4649–4654.

**Furutani M, Vernoux T, Traas J, Kato T, Tasaka M, Aida M. 2004.** *PIN-FORMED1* and *PINOID* regulate boundary formation and cotyledon development in Arabidopsis embryogenesis. *Development* **131**: 5021–5030.

**Gouy M, Guindon S, Gascuel O. 2010.** SeaView Version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* **27**: 221–224.

**Guindon S, Dufayard JF, Hordijk W, Lefort V, Gascuel O. 2009.** PhyML: fast and accurate phylogeny reconstruction by maximum likelihood. *Infection Genetics and Evolution* **9**: 384–385.

**Gustafson AM, Allen E, Givan S, Smith D, Carrington JC, Kasschau KD. 2005.** ASRP: the *Arabidopsis* Small RNA Project Database. *Nucleic Acids Research* **33**: D637–D640.

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

**Ishida T, Aida M, Takada S, Tasaka M. 2000.** Involvement of *CUP-SHAPED COTYLEDON* genes in gynoecium and ovule development in Arabidopsis thaliana. *Plant and Cell Physiology* **41**: 60–67.

**Jasinski S, Vialette-Guiraud ACM, Scutt CP. 2010.** The evolutionary-developmental analysis of plant microRNAs. *Philosophical Transactions of the Royal Society B-Biological Sciences* **365**: 469–476.

**Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T. 2002.** Identification of tissue-specific microRNAs from mouse. *Current Biology* **12**: 735–739.

**Larue CT, Wen J, Walker JC. 2009.** Genetic interactions between the *miRNA164-CUC2* regulatory module and *BREVIPEDICELLUS* in Arabidopsis developmental patterning. *Plant Signalling and Behaviour* **4**: 666–668.

**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.

**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. *Current Biology* **14**: 1035–1046.

**Nikovics K, Blein T, Peaucelle A, *et al*. 2006.** The balance between the *MIR164A* and *CUC2* genes controls leaf margin serration in *Arabidopsis*. *Plant Cell* **18**: 2929–2945.

**Ooka H, Satoh K, Doi K, *et al*. 2003.** Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. *DNA Research* **10**: 239–247.

**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, vanHouwelingen 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.

**Valoczi A, Varallyay E, Kauppinen S, Burgyan J, Havelda Z. 2006.** Spatio-temporal accumulation of microRNAs is highly coordinated in developing plant tissues. *Plant Journal* **47**: 140–151.

**Vroemen CW, Mordhorst AP, Albrecht C, Kwaaitaal M, 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 JP, Cook H, Causier B, Schwarz-Sommer Z, Davies B. 2004.** *CUPULIFORMIS* establishes lateral organ boundaries in *Antirrhinum*. *Development* **131**: 915–922.

**Xie Q, Frugis G, Colgan D, Chua N-H. 2000.** *Arabidopsis NAC1* transduces auxin signal downstream of *TIR1* to promote lateral root development. *Genes and Development* **14**: 3024–3036.

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