

## The evolutionary-developmental analysis of plant microRNAs

Sophie Jasinski, Aurélie C. M. Vialette-Guiraud and Charles P. Scutt

*Phil. Trans. R. Soc. B* 2010 **365**, 469-476

doi: 10.1098/rstb.2009.0246

---

### References

[This article cites 42 articles, 22 of which can be accessed free](#)

<http://rstb.royalsocietypublishing.org/content/365/1539/469.full.html#ref-list-1>

### Rapid response

[Respond to this article](#)

<http://rstb.royalsocietypublishing.org/letters/submit/royptb;365/1539/469>

### Subject collections

Articles on similar topics can be found in the following collections

[molecular biology](#) (196 articles)

[developmental biology](#) (128 articles)

[evolution](#) (1350 articles)

### Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

---

To subscribe to *Phil. Trans. R. Soc. B* go to: <http://rstb.royalsocietypublishing.org/subscriptions>

---

# The evolutionary-developmental analysis of plant microRNAs

Sophie Jasinski<sup>†</sup>, Aurélie C. M. Vialette-Guiraud<sup>†</sup>  
and Charles P. Scutt<sup>\*</sup>

*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*

MicroRNAs (miRNAs) control many important aspects of plant development, suggesting these molecules may also have played key roles in the evolution of developmental processes in plants. However, evolutionary-developmental (evo-devo) studies of miRNAs have been held back by technical difficulties in gene identification. To help solve this problem, we have developed a two-step procedure for the efficient identification of miRNA genes in any plant species. As a test case, we have studied the evolution of the *MIR164* family in the angiosperms. We have identified novel *MIR164* genes in three species occupying key phylogenetic positions and used these, together with published sequence data, to partially reconstruct the evolution of the *MIR164* family since the last common ancestor of the extant flowering plants. We use our evolutionary reconstruction to discuss potential roles for *MIR164* genes in the evolution of leaf shape and carpel closure in the angiosperms. The techniques we describe may be applied to any miRNA family and should thus enable plant evo-devo to begin to investigate the contributions miRNAs have made to the evolution of plant development.

**Keywords:** microRNA; miR164; angiosperm; flower; carpel; leaf dissection

## 1. INTRODUCTION

Evolutionary-developmental (evo-devo) biology aims to explain how developmental processes have evolved (Raff 2000). Evo-devo analyses frequently begin from an understanding of gene function in model species and a robust species phylogeny in which those models are placed. Orthologous genes may then be studied in species occupying key phylogenetic positions to deduce the molecular changes that were responsible for evolutionary events of interest. To date, plant evo-devo study has mainly focused on genes encoding developmental regulators such as transcription factors and signal transduction components. However, it is now clear that many plant developmental processes are regulated by microRNAs (miRNAs; Mallory & Vaucheret 2006), suggesting these molecules also may have played important roles in the evolution of plant development.

miRNAs are single-stranded RNA molecules of approximately 21 nt that negatively regulate gene expression by hybridizing to complementary target sites in specific messenger RNAs (mRNAs; Filipowicz *et al.* 2008). In plants, miRNAs typically show very high similarity to the complement of their target sites in mRNAs, and act by directing the cleavage of these

(Jones-Rhoades *et al.* 2006). miRNAs are transcribed from nuclear genes to generate primary transcripts which are then processed via intermediate forms termed pri- and pre-miRNAs. In these intermediate molecules, an miR site of approximately 21 nt forms a duplex with a partially complementary miR\* site. After further processing, the miR–miR\* duplex separates and the miR site, corresponding to the mature miRNA, becomes integrated into an RNA-induced silencing complex, which is then able to target mRNAs in the cytoplasm.

The *MIR164* family in the model angiosperm *Arabidopsis thaliana* consists of three genes, *Ath-MIR164a*, *b* and *c*, which are potentially capable of targeting seven members (Gustafson *et al.* 2005) of the *NAC* transcription factor family. Among these targets, the developmental roles of *CUP-SHAPED COTYLEDON1* and 2 (*CUC1/2*; Aida *et al.* 1997; Laufs *et al.* 2004; Peaucelle *et al.* 2007; Raman *et al.* 2008), *NAC1* (Xie *et al.* 2000) and *ORE1* (Kim *et al.* 2009) are known. The expression domains of *Ath-MIR164a*, *b* and *c* partially overlap (Sieber *et al.* 2007), which may explain the combination of redundant and non-redundant phenotypes observed in *mir164* mutants. For example, only *ath-mir164a* shows an increase in the depth of leaf sinuses (Nikovics *et al.* 2006), whereas both *ath-mir164a* and *ath-mir164b* show an increase in lateral root branching (Guo *et al.* 2005). Other *mir164* mutants show additive phenotypes. For example, *ath-miR164c* shows a slight defect in carpel fusion (Baker *et al.* 2005), whereas the *ath-mir164abc* triple mutant shows complete carpel separation (Sieber *et al.* 2007). The distinct

\* Author for correspondence ([charlie.scutt@ens-lyon.fr](mailto:charlie.scutt@ens-lyon.fr)).

<sup>†</sup>The first two authors contributed equally to this work.

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rstb.2009.0246> or via <http://rstb.royalsocietypublishing.org>.

One contribution of 16 to a Discussion Meeting Issue 'Darwin and the evolution of flowers'.

functions of *MIR164* family members in *A. thaliana* illustrate the need for evo-devo studies of miRNA genes in plants: the processes of gene duplication, gene loss, sub-functionalization and neo-functionalization, which are known to have shaped families of protein-coding genes (Ohno 1970), have clearly also acted on miRNA families, with potentially important consequences for the evolution of plant development.

Most of the plant miRNA genes currently known were identified by the *in silico* searching of partial or complete genome sequences (Bonnet *et al.* 2004; Jones-Rhoades & Bartel 2004; Adai *et al.* 2005; Li & Zhang 2005; Li *et al.* 2005). To study the evolutionary importance of miRNAs, methods are required for the identification of miRNA genes in plants occupying key phylogenetic positions whose genomes have not necessarily been sequenced. The high-throughput sequencing of small RNAs (Barakat *et al.* 2007; Buhtz *et al.* 2008) can be used to identify mature miRNAs in any species. However, the sequences obtained by this method are too short for phylogenetic analysis, and are therefore of limited value to evo-devo studies.

To help resolve this problem, we present here methods for the efficient identification of miRNA genes in any plant species. To achieve this, we have devised a two-step method, which is specifically applicable to plants due to the generally high conservation of miR sites within any given miRNA gene family throughout the plant kingdom. We furthermore show that standard methods of phylogenetic analysis can be applied to miRNA genes, permitting the reconstruction of miRNA family evolution over considerable phylogenetic distances. As a test case, we have analysed the evolution of the *MIR164* family in the angiosperms, or flowering plants. To draw conclusions on the structure of the *MIR164* family in the last common ancestor of the extant angiosperms, we incorporated in our studies two representatives of the ANA grade (ANA for Amborellales, Nymphaeales and Austrobaileyales), which represents the three earliest-diverging lineages of extant angiosperm (Stevens 2001). The ANA grade species used in this study were *Amborella trichopoda*, which is a shrub endemic to the rainforests of New Caledonia and the only known member of Amborellales, and *Cabomba aquatica*, which is a small aquatic plant native to northern South America (Orgaard 1991) and a convenient representative of Nymphaeales. We also analysed the *MIR164* family in *Aquilegia vulgaris*, which occupies a key phylogenetic position as a member of the most basally diverging eudicot order, Ranunculales (Stevens 2001).

Our phylogenetic analyses have enabled us to conclude that at least two *MIR164* genes were present in the last common ancestor of the extant angiosperms. Furthermore, gene expression analyses across a wide taxonomic range indicate these putative ancestral *MIR164* genes to have probably conserved characteristics of their expression patterns throughout angiosperm evolution. We use the results of our analyses to propose testable hypotheses on the roles of *miR164* in the evolution of leaf shape and carpel closure in the angiosperms. The methods we have developed should be applicable to any miRNA family and should

thus enable studies of the roles that these important regulatory molecules have played in plant evolution.

## 2. RESULTS

### (a) *A two-step procedure for the identification of miRNA genes in plants*

We have developed a two-step procedure for the identification of miRNA genes of known families in any plant species, as summarized in the electronic supplementary material, figure S1. The first step of this procedure consists of the polymerase chain reaction (PCR) suppression–amplification (Broude *et al.* 2001) of miRNA gene fragments from adaptor-ligated plant genomic DNA using an miRNA-specific PCR primer in combination with an adaptor primer that is common to all DNA molecules. Such amplifications typically yield mixtures of PCR products containing a proportion of miRNA gene fragments. In the second step of our procedure, a database is generated from sequences obtained by PCR suppression–amplification, and then searched for the miR\* sites present in miRNA gene fragments. To quantify the efficiency of the PCR suppression–amplification step of our procedure, we used this to re-identify known *MIR164* genes from *A. thaliana* and *Zea mays*. PCR products generated from genomic DNA of these species were ligated into plasmid vectors, cloned in *Escherichia coli*, and screened by colony hybridization using *MIR164* gene-specific radiolabelled probes. All six *MIR164* probes used, *Ath-MIR164a*, *b* and *c* from *A. thaliana*, and *Zma-MIR164a*, *b* and *d* (Griffiths-Jones *et al.* 2008) from *Z. mays*, revealed positively hybridizing bacterial colonies at frequencies of between 0.1 and 8 per cent (see the electronic supplementary material, table S1). This test thus established that our PCR procedure had exhaustively amplified the *MIR164* genes present in *A. thaliana*, in which this family had been completely characterized, and also amplified all three of the *MIR164* family members searched for in *Z. mays*. Despite the success of this initial test, it should not however be assumed that the miRNA gene identification method detailed here will prove to be exhaustive in all cases.

Following the above test, we employed our full, two-step procedure to identify novel *MIR164* genes from *A. trichopoda*, *C. aquatica* and *A. vulgaris*. The PCR products inserted into approximately 220 recombinant plasmids, generated from each of these species as described above, were sequenced from one extremity. A DNA database was compiled from sequences of each species under study, and then searched using an optimized BLAST protocol. This procedure resulted in the identification of three putative *MIR164* genes from each of *A. vulgaris* (*Avu-MIR164a–c*) and *C. aquatica* (*Caq-MIR164a–c*), and one such gene from *A. trichopoda* (*Atr-MIR164a*). The putative miRNA genes identified were present in populations of sequenced plasmids at frequencies of between 0.4 and 2 per cent (see the electronic supplementary material, table S1), in broad agreement with our earlier trial runs using *A. thaliana* and *Z. mays*. Further rounds of genomic PCR were then used to complete the sequences of the putative *MIR164* genes identified.



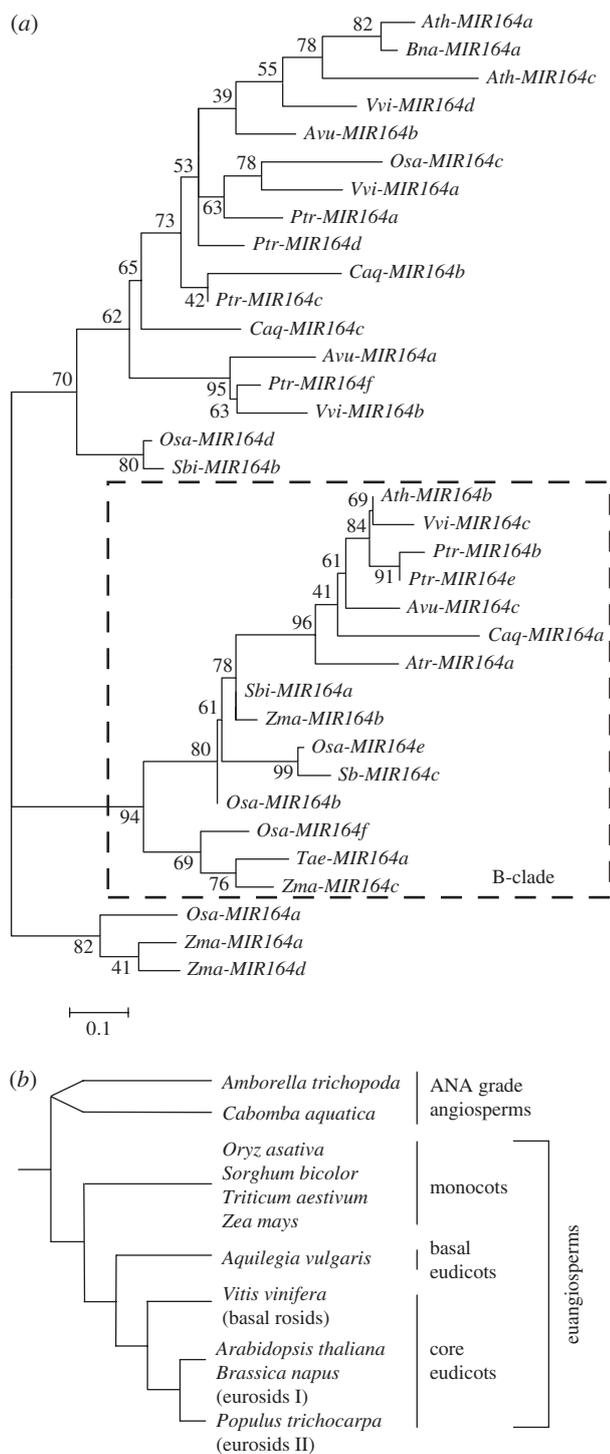


Figure 2. A phylogeny of the *MIR164* family in the angiosperms. (a) A maximum likelihood phylogeny of the *MIR164* family generated using the HKY model. Bootstrap supports are indicated. Accession numbers for the genes included are given in the electronic supplementary material, table S2. (b) The consensus view of angiosperm phylogeny (Stevens 2001), showing the relationships of species sampled in the present study.

family. Similar effects have been noted in organismal phylogenies constructed using large datasets in which only Poaceae were used to represent the monocots, as discussed by Soltis *et al.* (2004). The clear identification of both B-clade and other *MIR164* genes in species whose lineages diverged very early in flowering plant

evolution strongly suggests the presence of at least two *MIR164* genes in the last common ancestor of the extant flowering plants.

Several other nodes of our phylogeny also appear well supported by high bootstrap values. Such nodes frequently link pairs of genes from closely related species (e.g. *Ath-MIR164a* and *Bna-MIR164a* from Brassicaceae, and *Osa-MIR164d* and *Sbi-MIR164b* from Poaceae), suggesting these genes to be orthologues that were generated by recent speciation events. Other high levels of bootstrap support link multiple *MIR164* genes from a single species (e.g. *Ptr-MIR164b* and *Ptr-MIR164e*, from *P. trichocarpa*), suggesting these genes to be paralogues that were generated by recent duplication events.

### (c) Expression patterns of *MIR164* genes in the angiosperms

We used semi-quantitative reverse transcriptase-PCR (RT-PCR) to analyse the expression patterns of *MIR164* genes in *A. thaliana*, *C. aquatica*, *A. vulgaris* and *A. trichopoda* (figure 3). B-clade genes of the *MIR164* family were expressed at similar levels in the leaves, flowers and roots of all these species (figure 3a–d), except for a somewhat lower level of expression in the roots of *A. trichopoda* (figure 3d). These data suggest the conservation of similar levels of B-clade gene expression in leaves, flowers and roots of the lineages tested since the initial radiation of the angiosperms. However, further studies on dissected plant tissues, performed using more quantitative methods of analyses, will be required to precisely define the degree of conservation of gene expression pattern between the lineages compared.

A pronounced foliar dimorphism is present in *C. aquatica*, which possesses floating leaves with entire margins that help to carry its inflorescences to the water's surface, in addition to submerged leaves that are very highly dissected, with the effect of reducing drag in water currents. The B-clade gene *Caq-MIR164a* was more highly expressed in submerged than in floating leaves of *C. aquatica* (figure 3c), while *Caq-MIR164b*, which grouped externally to the B-clade in phylogenetic analyses (figure 2a), was expressed specifically in submerged leaves (figure 3c). Given the known role of *miR164* in determining leaf morphology in *A. thaliana* (Nikovics *et al.* 2006), these data suggest *miR164* may also be involved in the mechanism generating high levels of dissection in the submerged leaves of *C. aquatica*.

Alone among the sequences studied here, *Caq-MIR164c* from *C. aquatica* showed no measurable expression (figure 3c). Demonstrable expression has been proposed as a prerequisite for the identification of novel miRNA genes (Ambros *et al.* 2003). However, *Caq-MIR164c* may be considered as a *MIR164* gene as it occupies a well-supported position in phylogenetic reconstructions of the *MIR164* family (figure 2a). This gene may have a very specific expression profile, which was not revealed in the analyses presented here, or may alternatively represent a recently generated pseudogene which is no longer expressed.

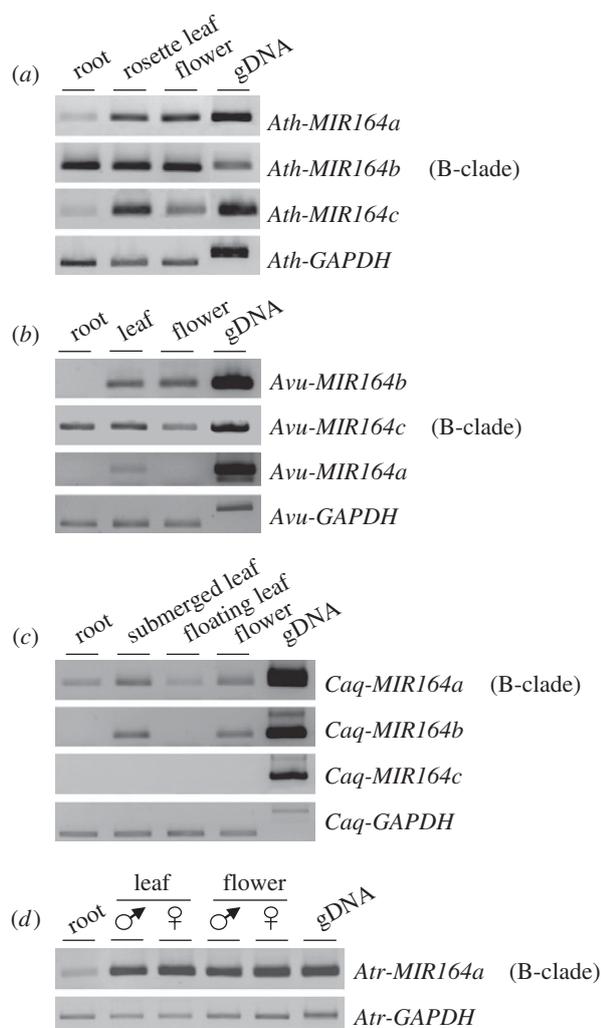


Figure 3. Semi-quantitative RT-PCR analysis of *MIR164* gene expression. Amplifications of *miR164* and *GAPDH* transcripts were simultaneously performed on first-strand cDNA prepared from the tissues indicated of: (a) *A. thaliana*, (b) *A. vulgaris*, (c) *C. aquatica* and (d) *A. trichopoda*. Genes corresponding to the B-clade, as identified in phylogenetic analyses, are indicated. Positive control amplifications from genomic DNA are also shown. Negative control amplifications, from RNA samples untreated with reverse transcriptase (not shown), yielded no PCR products in all cases.

### 3. DISCUSSION

#### (a) Novel methods have permitted a partial reconstruction of *MIR164* evolution in the angiosperms

We have devised and tested a combination of methods for the identification and phylogenetic analysis of miRNA genes in plants which should facilitate studies of the roles these genes have played in plant evolution. We have demonstrated the efficiency of these procedures by partially reconstructing the evolution of the *MIR164* family since the last common ancestor of the extant angiosperms. Our phylogenetic analysis thus considerably extends the taxonomic range of previous studies of the evolution of miRNA families (Mica *et al.* 2006; Zhang *et al.* 2006a). According to the results of our analyses, at least two lineages of *MIR164* genes were present in the last common ancestor of the extant angiosperms (figure 4). We therefore

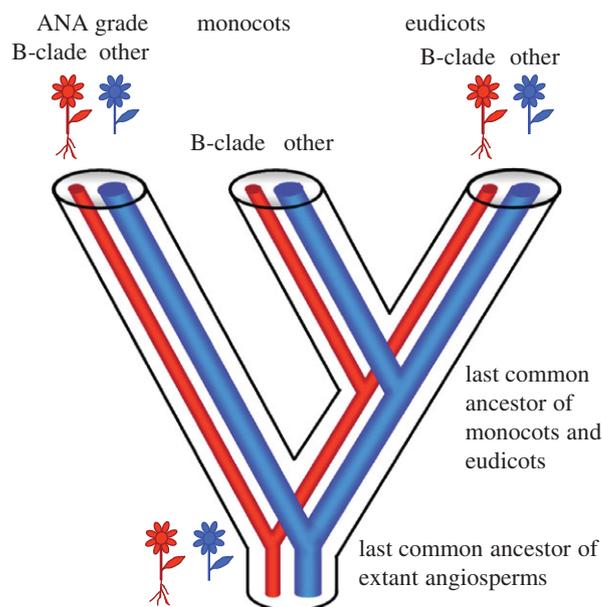


Figure 4. *MIR164* evolution in the angiosperms. B-clade (red) and one or more other (blue) *MIR164* genes were present in the last common ancestor of the extant angiosperms. Genes of both lineages show broadly conserved expression patterns between ANA grade angiosperms (*A. trichopoda* and *C. aquatica*) and eudicots (*A. thaliana* and *A. vulgaris*), as represented by plant organ symbols (flowers, leaves and roots).

conclude that one or more duplications in the *MIR164* family occurred before the radiation of the angiosperms to generate a B-lineage, in addition to at least one other *MIR164* lineage.

Our expression analyses (figure 3) have indicated B-clade *MIR164* genes to be expressed at a relatively high level in roots, whereas the other *MIR164* genes studies were less expressed in roots than in other tissues. The broad similarity of B-clade gene expression between ANA grade and other angiosperms suggest this pattern to be conserved in the lineages studied since the last common ancestor of the extant angiosperms (figure 4). Similarly, the *MIR164* genes external to the B-clade, which may belong to one or more *MIR164* lineages, appear in general to have conserved higher expression in leaves and flowers than in roots from an early stage in angiosperm evolution (figure 4).

#### (b) The roles of *MIR164* genes in angiosperm evolution

Leaf lobing and dissection are highly variable traits within the angiosperms, which have clearly arisen many times independently (Bharathan *et al.* 2002). Such polyphyletic origins for a given morphological trait might imply distinct underlying molecular mechanisms. However, the involvement of *CUC* orthologues in leaf lobing and dissection has recently been demonstrated in several distantly related eudicots, suggesting the independent recruitment of a pre-existing genetic mechanism to these processes in distinct plant lineages (Blein *et al.* 2008; Berger *et al.* 2009). In *A. thaliana*, the effect of *CUC2* on leaf lobing is known to be modulated by *Ath-MIR164a* (Nikovics *et al.* 2006), suggesting not only the parallel

recruitment of *CUC* genes to affect leaf morphology, but also of their *miR164* regulator. In the present study, we have identified a marked difference in the levels of *miR164* between the two leaf forms of *C. aquatica*, which differ considerably in their degree of dissection. We found *miR164* to be more highly expressed in the dissected submerged leaves of this species, which may reflect a need in these organs for a greater modulation of *NAC* gene activity, related to the production of a highly dissected leaf lamina. Hence, the preliminary results obtained here are consistent with a proposed role for *miR164* and its *NAC* family targets in leaf dissection in an ANA grade angiosperm. A conserved role in leaf dissection for the *NAC* gene/*miR164* expression balance between species as widely diverged as *A. thaliana* and *C. aquatica*, in which this characteristic is certainly of independent origin, may indicate these genes to function as a genetic module that has been recruited to modify leaf morphology in distinct lineages from an early stage in angiosperm evolution. Such a module might for example be conserved in all angiosperms, where it would participate in basic morphological processes in meristems, etc. and be 'switched on and off' in leaves over the course of evolution to effect relatively ephemeral changes to the lobing or dissection of these organs.

The major defining feature of the angiosperms is the carpel, the specialized female reproductive organ that encloses the ovules. In gymnosperms, which form the sister group to the angiosperms, ovules typically occur as naked structures, borne in the axils of open sporophylls. The process of fusion between carpels is known to be controlled in *A. thaliana* by the relative expression of *miR164* and *CUC2* (Baker *et al.* 2005; Nikovics *et al.* 2006; Sieber *et al.* 2007). Further studies may now be undertaken to determine whether a similar mechanism, involving the *MIR164* genes identified in the present work, might be responsible for the closure of the simple carpels present in *A. trichopoda* and *C. aquatica*. Evidence for the involvement of the *miR164/CUC* expression balance in carpel closure in multiple ANA grade lineages would suggest a potential mechanism for the initial evolution of the closed carpel in early flowering plants.

#### 4. EXPERIMENTAL PROCEDURES

##### (a) *Plant material*

Plants and tissues of *A. trichopoda* and *A. vulgaris* were field-collected from Col d'Amieu (New Caledonia) and from Corveissiat (Ain, France), respectively. Seed of the Columbia-0 ecotype of *A. thaliana* (accession no. N1092) was obtained from the Nottingham Arabidopsis Stock Center. Plants of the above species were grown in peat-based compost under greenhouse conditions. Plants of *C. aquatica* were obtained from Anthias S.A. (Les Chères, France) and grown in a freshwater aquarium.

##### (b) *PCR suppression–amplification of miRNA gene fragments*

Genomic DNA was extracted from leaf tissues of *A. trichopoda*, *A. vulgaris*, *C. aquatica* and *A. thaliana*

using a Nucleon PhytoPure kit (Amersham), incorporating an RNAase digestion step. Genomic DNA of *Z. mays* inbred line A188 (Gerdes & Tracy 1993) was kindly provided by Dr Peter Rogowsky (RDP, Lyon). To amplify miRNA gene sequences, PCR-suppression amplifications were performed, based on the method of Broude *et al.* (2001). Aliquots of genomic DNA (10 µg) from all species studied were digested with *DraI* or *HaeIII*. Double-stranded DNA adaptors (Broude *et al.* 2001) were ligated to the resulting cleaved DNA fragments, and unincorporated adaptors were then removed using NucleoSpin PCR clean-up columns (Macherey-Nagel). PCR amplifications were performed from aliquots (25 ng) of adaptor-ligated DNA, using a 21-nt primer corresponding to the consensus *miR164* sequence from *A. thaliana* (figure 1) and the A-adaptor primer (Broude *et al.* 2001). Amplifications were also performed using a 19-mer *miR164* primer, lacking the two bases at the 3'-extremity of the consensus *miR164* sequence. PCR amplifications were performed using 0.02 units µl<sup>-1</sup> of EX-TAQ Polymerase (Takara), 0.2 µM of each primer, 0.2 mM dNTPs and reaction buffer (1×), as supplied by the manufacturer, and continued for 30 thermal cycles, each consisting of a denaturation step of 94°C for 30 s, a thermal gradient annealing step of 45–65°C for 30 s, and an elongation step of 72°C for 90 s. Aliquots of the PCR products generated were analysed on 1 per cent agarose gels. DNA samples for further study were selected from amplifications performed at the highest annealing temperature that yielded products in each case. Aliquots (1 µl) of the DNA samples selected were ligated into the *pCRII* vector (Invitrogen) and the recombinant plasmids generated were used to transform *E. coli* DH5α competent cells and plated on ampicillin-containing media.

To estimate the proportion of *MIR164* gene fragments in PCR products amplified from *A. thaliana* and *Z. mays* genomic DNA, 520 and 1040 ampicillin-resistant colonies generated, respectively, from these two species were transferred onto Hybond-N (Amersham) membranes for colony hybridizations using radio-labelled, gene-specific probes of *Ath-MIR164a*, *b* and *c* (for *A. thaliana* genes), or *Zma-MIR164a*, *b* and *d* (for *Z. mays* genes). To identify novel *MIR164* genes from *A. trichopoda*, *C. aquatica* and *A. vulgaris*, approximately 220 antibiotic-resistant bacterial colonies from each species, generated as described above, were sent to a commercial service provider for plasmid purification and DNA sequencing using the M13 forward sequencing primer. PCR products derived from DNA cleaved by *DraI* and *HaeIII*, and from amplifications involving full-length and shortened *miR164* primers, were selected for sequencing in approximately equal proportions.

##### (c) *Database construction and BLAST searching for miR\* sites*

DNA sequences of PCR products, generated as described above, were imported into CONTIG EXPRESS (Invitrogen). Cloning vector regions were removed, and the resulting sequences were assembled into

contigs. For contigs containing a *miR164* primer at the 5'-end, only the longest contributing sequence was retained for further analysis, whereas for contigs beginning with an adaptor primer, the whole contig was saved, reversed and complemented. The resulting sequences from each plant species were exported as a FASTA file and aligned over their terminal *miR164* primer sequences, where present, using MUSCLE (Edgar 2004), which allowed these primer sites to be removed easily from the aligned sequences. The resulting primer-less sequences were then compiled into stand-alone DNA database using FORMATDB. The DNA databases generated in this way were searched by BLAST (Altschul *et al.* 1997) for internal miR\* sites using the mature *miR164* consensus sequence (figure 1). BLAST options were used to specify searching of the bottom strand only (S2), with a word-size of 6 (W6), no filtering (FF), giving a score of 4 for an exact match (r4), and a penalty of 5 for each mismatch (q-5). BLAST hits were used to search the available plant nucleotide databases to eliminate any false positives that were homologous to known genes. The remaining sequences, representing putative *MIR164* genes, were retained for further study.

#### (d) PCR amplification across miR sites

To determine the sequences of mature miR sites present in the putative *MIR164* genes identified, and of the regions immediately adjacent to these, a second round of PCR suppression–amplification was undertaken. These amplifications were performed on further aliquots of adaptor-ligated genomic DNA by the sequential use of two miRNA gene-specific primers derived from the novel sequences obtained, in combination with the A-adaptor primer (Broude *et al.* 2001). The resulting PCR products were cloned in plasmid vectors and sequenced, and the novel data obtained were used to complete the sequences of the putative *MIR164* genes identified.

#### (e) RNA secondary structure prediction and stability calculations

Secondary structures of putative *pre-miR164* molecules were predicted in MFOLD 3.2 (Zuker 2003) using default parameters. MFEI for the structures generated were calculated as described by Zhang *et al.* (2006b).

#### (f) Phylogenetic analyses of miRNA genes

Sequence alignments were performed using M-Coffee (Wallace *et al.* 2006) using 'regular' parameters and homologous sites were chosen for ML phylogenetic analyses on the basis of their clear alignment. The full alignment is shown for clarity in the electronic supplementary material, figure S2, though only the sites chosen for phylogenetic analyses in this figure may be considered as homologous. Phylogenetic analyses were performed using TREEFINDER v. May 2006 (Jobb *et al.* 2004), initially assuming the HKY (Hasegawa *et al.* 1985) model of DNA substitution. Congruent results were also obtained assuming the GTR (Rodriguez *et al.* 1990) and TrNef+ $\Gamma$  (Tamura & Nei 1993) models. In each analysis 1000 bootstrap replicates were performed.

#### (g) Semi-quantitative reverse transcriptase-PCR analyses

Plant tissues were ground under liquid nitrogen, together with 0.1 g of polyvinylpyrrolidone (mw 40 kDa) per gram of tissue, and RNA was extracted from these tissue powders using an RNA-Easy Kit (Qiagen). Samples (0.5  $\mu$ g) of total RNA were then treated with RNase-free DNase (Ambion), and each divided into two aliquots. These aliquots were processed for first-strand cDNA synthesis, respectively with and without the addition of RevertAid M-MuLV RT (Fermentas), using oligo-dT primers and other reaction components as described by the manufacturer. Appropriate volumes of diluted cDNA samples, and equal volumes of similarly diluted negative control samples (not treated with RT), were used as templates in PCR amplifications using *MIR164* gene-specific and *GAPDH* primers. Full details of the primer sequences and thermal cycle conditions used are given in the electronic supplementary material, table S3.

We thank Patrick Laufs for helpful discussions and for supplying *Arabidopsis* miR gene-specific probes, Manolo Gouy and Dominique Mouchiroud for advice on phylogenetic analysis, Guy Perrière for advice on database methods, and the technical staff of the RDP Laboratory for their support. S.J. was supported by an ATER post of the ENS-Lyon. A.V. receives a doctoral thesis grant from the French Ministry of Research.

#### REFERENCES

- Adai, A., Johnson, C., Mlotshwa, S., Archer-Evans, S., Manocha, V., Vance, V. & Sundaresan, V. 2005 Computational prediction of miRNAs in *Arabidopsis thaliana*. *Genome Res.* **15**, 78–91. (doi:10.1101/gr.2908205)
- 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. (doi:10.1105/tpc.9.6.841)
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W. & Lipman, D. J. 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* **25**, 3389–3402. (doi:10.1093/nar/25.17.3389)
- Ambros, V. *et al.* 2003 A uniform system for microRNA annotation. *RNA—Publ. RNA Soc.* **9**, 277–279.
- Baker, C. C., Sieber, P., Wellmer, F. & Meyerowitz, E. M. 2005 The early extra petals1 mutant uncovers a role for microRNA miR164c in regulating petal number in *Arabidopsis*. *Curr. Biol.* **15**, 303–315. (doi:10.1016/j.cub.2005.02.017)
- Barakat, A., Wall, K., Leebens-Mack, J., Wang, Y. J., Carlson, J. E. & dePamphilis, C. W. 2007 Large-scale identification of microRNAs from a basal eudicot (*Eschscholzia californica*) and conservation in flowering plants. *Plant J.* **51**, 991–1003. (doi:10.1111/j.1365-3113X.2007.03197.x)
- Berger, Y. *et al.* 2009 The NAC-domain transcription factor GOBLET specifies leaflet boundaries in compound tomato leaves. *Development* **136**, 823–832. (doi:10.1242/dev.031625)
- Bharathan, G., Goliber, T. E., Moore, C., Kessler, S., Pham, T. & Sinha, N. R. 2002 Homologies in leaf form inferred from KNOXI gene expression during development. *Science* **296**, 1858–1860. (doi:10.1126/science.1070343)
- Blein, T., Pulido, A., Vialette-Guiraud, A., Nikovics, K., Morin, H., Hay, A., Johansen, I. E., Tsiantis, M. &

- Laufs, P. 2008 A conserved molecular framework for compound leaf development. *Science* **322**, 1835–1839. (doi:10.1126/science.1166168)
- Bonnet, E., Wuyts, J., Rouze, P. & Van de Peer, Y. 2004 Detection of 91 potential in plant conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes. *Proc. Natl Acad. Sci. USA* **101**, 11 511–11 516. (doi:10.1073/pnas.0404025101)
- Broude, N. E., Zhang, L. G., Woodward, K., Englert, D. & Cantor, C. R. 2001 Multiplex allele-specific target amplification based on PCR suppression. *Proc. Natl Acad. Sci. USA* **98**, 206–211. (doi:10.1073/pnas.98.1.206)
- Buhtz, A., Springer, F., Chappell, L., Baulcombe, D. C. & Kehr, J. 2008 Identification and characterization of small RNAs from the phloem of *Brassica napus*. *Plant J.* **53**, 739–749. (doi:10.1111/j.1365-313X.2007.03368.x)
- Edgar, R. C. 2004 MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 1–19.
- Filipowicz, W., Bhattacharyya, S. N. & Sonenberg, N. 2008 Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet.* **9**, 102–114.
- Gerdes, J. T. & Tracy, W. F. 1993 Pedigree diversity within the Lancaster surecrop heterotic group of maize. *Crop Sci.* **33**, 334–337.
- Griffiths-Jones, S., Saini, H. K., van Dongen, S. & Enright, A. J. 2008 miRBase: tools for microRNA genomics. *Nucl. Acids Res.* **36**, D154–D158. (doi:10.1093/nar/gkm952)
- Guo, H. S., Xie, Q., Fei, J. F. & Chua, N. H. 2005 MicroRNA directs mRNA cleavage of the transcription factor NAC1 to downregulate auxin signals for *Arabidopsis* lateral root development. *Plant Cell* **17**, 1376–1386. (doi:10.1105/tpc.105.030841)
- Gustafson, A. M., Allen, E., Givan, S., Smith, D., Carrington, J. C. & Kasschau, K. D. 2005 ASRP: the *Arabidopsis* small RNA project database. *Nucl. Acids Res.* **33**, D637–D640. (doi:10.1093/nar/gki127)
- Hasegawa, M., Kishino, H. & Yano, T. A. 1985 Dating of the human ape splitting by a molecular clock of mitochondrial-DNA. *J. Mol. Evol.* **22**, 160–174. (doi:10.1007/BF02101694)
- Jobb, G., von Haeseler, A. & Strimmer, K. 2004 TREEFIN-DER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evol. Biol.* **4**, 18. (doi:10.1186/1471-2148-4-18)
- Jones-Rhoades, M. W. & Bartel, D. P. 2004 Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol. Cell* **14**, 787–799. (doi:10.1016/j.molcel.2004.05.027)
- Jones-Rhoades, M. W., Bartel, D. P. & Bartel, B. 2006 MicroRNAs and their regulatory roles in plants. *Annu. Rev. Plant Biol.* **57**, 19–53. (doi:10.1146/annurev.arplant.57.032905.105218)
- Kim, J. H., Woo, H. R., Kim, J., Lim, P. O., Lee, I. C., Choi, S. H., Hwang, D. & Nam, H. G. 2009 Trifurcate feed-forward regulation of age-dependent cell death involving miR164 in *Arabidopsis*. *Science* **323**, 1053–1057. (doi:10.1126/science.1166386)
- 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. (doi:10.1242/dev.01320)
- Li, X. & Zhang, Y. Z. 2005 Computational detection of microRNAs targeting transcription factor genes in *Arabidopsis thaliana*. *Comput. Biol. Chem.* **29**, 360–367. (doi:10.1016/j.compbiolchem.2005.08.005)
- Li, Y., Li, W. & Jin, Y. X. 2005 Computational identification of novel family members of microRNA genes in *Arabidopsis thaliana* and *Oryza sativa*. *Acta Biochim. Biophys. Sinica* **37**, 75–87. (doi:10.1111/j.1745-7270.2005.00012.x)
- Mallory, A. C. & Vaucheret, H. 2006 Functions of microRNAs and related small RNAs in plants. *Nat. Genet.* **38**, S31–S36. (doi:10.1038/ng1791)
- Mica, E., Gianfranceschi, L. & Pe, M. E. 2006 Characterization of five microRNA families in maize. *J. Exp. Bot.* **57**, 2601–2612. (doi:10.1093/jxb/erl013)
- 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. (doi:10.1105/tpc.106.045617)
- Ohno, S. 1970 *Evolution by gene duplication*. Berlin, Germany: Springer.
- Orgaard, M. 1991 The genus *Cabomba* (Cabombaceae)—a taxonomic study. *Nordic J. Bot.* **11**, 179–203. (doi:10.1111/j.1756-1051.1991.tb01819.x)
- Peaucelle, A., Morin, H., Traas, J. & Laufs, P. 2007 Plants expressing a miR164-resistant CUC2 gene reveal the importance of post-meristematic maintenance of phyllo-taxy in *Arabidopsis*. *Development* **134**, 1045–1050. (doi:10.1242/dev.02774)
- Raff, R. A. 2000 Evo-devo: the evolution of a new discipline. *Nat. Rev. Genet.* **1**, 74–79. (doi:10.1038/35049594)
- 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. (doi:10.1111/j.1365-313X.2008.03483.x)
- Rodriguez, F., Oliver, J. L., Marin, A. & Medina, J. R. 1990 The general stochastic-model of nucleotide substitution. *J. Theor. Biol.* **142**, 485–501. (doi:10.1016/S0022-5193(05)80104-3)
- Sieber, P., Wellmer, F., Gheyselinck, J., Riechmann, J. L. & Meyerowitz, E. M. 2007 Redundancy and specialization among plant microRNAs: role of the MIR164 family in developmental robustness. *Development* **134**, 1051–1060. (doi:10.1242/dev.02817)
- Soltis, D. E. *et al.* 2004 Genome-scale data, angiosperm relationships, and ‘ending incongruence’: a cautionary tale in phylogenetics. *Trends Plant Sci.* **9**, 477–483. (doi:10.1016/j.tplants.2004.08.008)
- Stevens, P. F. 2001 Angiosperm phylogeny website, v. 7. <http://www.mobot.org/MOBOT/research/APweb/>.
- Tamura, K. & Nei, M. 1993 Estimation of the number of nucleotide substitutions in the control region of mitochondrial-DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**, 512–526.
- Wallace, I. M., O’Sullivan, O., Higgins, D. G. & Notredame, C. 2006 M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucl. Acids Res.* **34**, 1692–1699. (doi:10.1093/nar/gkl091)
- Xie, Q., Frugis, G., Colgan, D. & Chua, N. H. 2000 *Arabidopsis* NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. *Genes Dev.* **14**, 3024–3036. (doi:10.1101/gad.852200)
- Zhang, B. H., Pan, X. P., Cannon, C. H., Cobb, G. P. & Anderson, T. A. 2006a Conservation and divergence of plant microRNA genes. *Plant J.* **46**, 243–259. (doi:10.1111/j.1365-313X.2006.02697.x)
- Zhang, B. H., Pan, X. P., Cox, S. B., Cobb, G. P. & Anderson, T. A. 2006b Evidence that miRNAs are different from other RNAs. *Cell. Mol. Life Sci.* **63**, 246–254. (doi:10.1007/s00018-005-5467-7)
- Zuker, M. 2003 Mfold web server for nucleic acid folding and hybridization prediction. *Nucl. Acids Res.* **31**, 3406–3415. (doi:10.1093/nar/gkg595)