Journal of Experimental Botany, Vol. 57, No. 11, pp. 2601–2612, 2006
doi:10.1093/jxb/erl013 Advance Access publication 4 July, 2006
This paper is available online free of all access charges (see http://jxb.oxfordjournals.org/open_access.html for further details)

Journal of Experimental Botany

RESEARCH PAPER

Characterization of five microRNA families in maize

Erica Mica, Luca Gianfranceschi and Mario Enrico Pè*

Department of Biomolecular Sciences and Biotechnology, Università degli Studi di Milano, Italy

Received 5 January 2006; Accepted 12 April 2006

Abstract

In recent years, microRNAs (miRNAs) have polarized the interest of the scientific community as a new category of gene expression regulators, present in both plants and animals. Plant miRNAs are involved in processes such as plant development, organ identity, and stress response. Nonetheless, knowledge of their functions is still incomplete, and it is conceivable that further new processes in which they are involved will be discovered. For these reasons, structural and functional characterization of MIR genes, that are also in crop species such as *Zea mays* L., becomes instrumental in addressing genetic and molecular mechanisms controlling phenotype determination and phenotypic adaptation to growing conditions. The present study contributes to the characterization of five miRNA families in maize, from the determination of their expression pattern in different maize tissues and genotypes, to the identification of putative targets by bioinformatic means and subsequent experimental validation of three targets by modified 5′ RACE experiments. Furthermore, 30 different MIR genes belonging to these five miRNA families were analysed by their attribution to maize chromosomes using oat–maize addition lines and by investigating their phylogenetic relationship with genes from other cereals. In particular, sequence homology was determined by the reciprocal best BLAST hit approach, to define groups of homologous genes between maize, rice, and sorghum.

Key words: Expression analysis, maize, microRNAs, MIR genes, orthologous genes, sorghum.

Introduction

Non-coding RNAs represent a large group of molecules in a eukaryotic cell (Meyers *et al.*, 2004). Among these molecules, in recent years, microRNAs (miRNAs) have polarized the interest of the scientific community as new categories of gene expression regulators (Carrington and Ambros, 2003; Hake, 2003); they are short RNAs, 21–24 nucleotides in length, which play an important role in post-transcriptional gene regulation in animals and plants, where several examples of miRNA-mediated gene regulation have been described (Bartel, 2004; Kidner and Martienssen, 2005; Zhang *et al*., 2005).

miRNAs are encoded by MIR genes that resides in distinct genomic regions. A single miRNA can be produced by the processing of one to several longer primary transcripts. Primary transcripts fold into secondary stem–loop structures which are processed in a two-step manner, through RNase III-like enzymes such as Drosha (Lee *et al.*, 2003), Dicer, and Dicer-like enzyme (Bernstein *et al.*, 2001; Kurihara and Watanabe, 2004). The resulting single-stranded miRNA is loaded in a ribonucleoproteic complex, called RISC (Hammond *et al.*, 2000).

In plants, miRNAs regulate the transcript level mostly by promoting, upon base pairing, the degradation of their target mRNA molecules. This function is performed by an enzyme with a 'slicer activity' (Liu *et al.*, 2004) that belongs to the RISC complex. However, in plants too there are a few examples of translational control performed by miRNAs (Aukerman and Sakai, 2003; Chen, 2003).

miRNAs in plants were first described very recently (Reinhart *et al.*, 2002; Llave *et al.*, 2002*a*; Bartel and Bartel, 2003). Searching for these molecules was accomplished by direct cloning together with genetic approaches and bioinformatic analyses, resulting in the identification, mainly in *Arabidopsis thaliana* and *Oriza sativa* (rice), of

* To whom correspondence should be addressed at: Via Celoria, 26, 20133 Milano, Italy. E-mail: enrico.pe@unimi.it
Abbreviations: AZM, assembled *Zea mays*; EST, expressed sequence tag; LW-RNA, low molecular weight RNA; miRNA, microRNA; OMA lines, oat–maize addition lines; pre-miRNA, microRNA precursor; RACE, rapid amplification of cDNA ends; RBH, reciprocal best BLAST hit; siRNA: small interfering RNA; TC, tentative contig; T_d, dissociation temperature.

several dozen miRNAs, their corresponding precursors, and MIR genes (Palatnik *et al*., 2003; Rhoades and Bartel, 2004; Sunkar and Zhu, 2004; Wang *et al.*, 2004; Sunkar *et al*., 2005).

More recently, other plant species such as *Glycine max*, *Medicago truncatula*, *Saccharum officinalis*, *Sorghum bicolor*, and *Zea mays* (Maher *et al*., 2004; Bedell *et al*., 2005) were also investigated and new miRNA families have been deposited in the RNA Registry (www.sanger.ac.uk/Software/Rfam/).

Computational analysis based on sequence similarity proved to be a reliable and successful way to identify target genes, since the number of mismatches allowed between the small RNA and its target in plants is low. Identified target genes of plant miRNAs are often transcription factors, involved in organ morphogenesis and plant development (Rhoades *et al.*, 2002; Bonnet *et al.*, 2004; Mallory *et al.*, 2004; Vaucheret *et al.*, 2004; Guo *et al.*, 2005; Lauter *et al.*, 2005).

For example, miR165/miR166 is involved in the determination of the adaxial/abaxial pattern in developing leaves (Kidner and Martienssen, 2004), miR172 governs floral organ development (Aukerman and Sakai, 2003), and miR-JAW regulates the level of *TCP*-family transcripts regulating leaf development (Palatnik *et al.*, 2003). It is remarkable that homologous miRNAs in different species have similar targets and conserved regulatory roles: miR165/miR166, for instance, regulates the expression pattern of the HD-ZIP III gene family in *Arabidopsis* and maize (Juarez *et al.*, 2004). In addition, it has recently been shown in *Arabidopsis*, rice, and *Populus trichocarpa* that some miRNAs are stress regulated and could be involved in cell responses to abiotic stresses such as salinity, cold, and dehydration (Rhoades and Bartel, 2004; Sunkar and Zhu, 2004; Lu *et al*., 2005). It is believed, however, that our current knowledge on plant miRNAs represents only a snapshot of the functions that miRNAs might perform in a plant cell: given their role as regulators of transcription factors and, more generally speaking, as transcription level regulators, it is conceivable that new processes in which miRNAs are involved will be discovered. This thesis is also supported by the indication that miRNAs can represent up to 1% of all predicted genes in animals and in plants (Bartel, 2004).

For the reasons mentioned above, structural and functional characterization of MIR genes, also in crop species such as *Z. mays* L. (maize), becomes instrumental in addressing genetic and molecular mechanisms controlling phenotype determination and phenotypic adaptation to growing conditions, on which yield potential, yield stability, and yield quality depend. In maize, where the phenomenon of heterosis finds its best examples, these phenotypic characteristics are strongly influenced by heterozygosity. With the long-term goal of exploring the possibility of the involvement of miRNAs in the heterotic phenomenon, some putative MIR genes coding for five miRNAs in maize were studied here. To explore a possible function for these miRNAs, the expression pattern of mature miRNAs was analysed by RNA gel blot analysis of RNA purified from different maize tissues in different genotypes: two inbred lines and their heterotic F_1 hybrid. To investigate miRNA involvement in plant architecture further, putative targets for all these five miRNAs were also identified, and some of these were experimentally confirmed. Finally, MIR gene sequences were assigned to maize chromosomes utilizing oat–maize addition lines (OMA lines) (Kynast *et al.*, 2001; Okagaki *et al.*, 2001) to analyse the distribution of these gene families over the maize genome.

Materials and methods

Plant material

Plant material was collected from maize inbred lines B73, H99, and their F_1 hybrid B73×H99. For each genotype, the three tissues examined were derived from at least 10 independent plants. Seedlings were obtained from plants grown in a growth chamber; in particular, seeds were sown in 4 cm diameter well plateaux on an inert substrate (Agriperlite, BPB Italia SpA, Italy) and incubated in a growth chamber with a 14 h light/10 h dark photoperiod at 26 °C. Plant material was collected from each individual plant when the third adult leaf apex became visible. The term seedling refers to any vegetative tissues above the first adult leaf, which was excluded. Immature ears were harvested from plants cultivated in the open field, selecting those whose silks reached no more than two-thirds of the ear length; silks and ear apexes were discarded. Developing kernels were collected from plants cultivated in the open field, 10 d after controlled pollination (sibbing for inbred lines and cross-pollination with B73 as female parent for the hybrid).

All plant material was immediately frozen in liquid nitrogen and stored at −80 °C until used.

RNA extraction and RNA gel blot

Total RNA was extracted with TRizol® reagent (Invitrogen), as described in the user's manual, and subsequently low molecular weight RNA (LW-RNA) was isolated using the DNA/RNA midi kit (Qiagen) as described by Carrington *et al*. (2002). A 20 μg aliquot of LW-RNA was loaded on a 10% polyacrylamide denaturing gel, electroblotted on an N^+ nylon membrane (Amersham). Filters were hybridized with oligo probes, end-labelled with [γ-^{32}P]ATP, using T4 polynucleotide kinase (Roche). Blots were prehybridized for 30 min at 65 °C using a solution containing 6× SSPE, 5× Denhardt's, 0.5% SDS, and salmon sperm DNA (200 μg ml^{-1}). Subsequently, blots were hybridized for 4 h at a temperature 15 °C below the oligonucleotide dissociation temperature (T_d) [T_d (°C)=4(C+G)+2(A+T)]. Membranes were washed three times with 6× SSPE at room temperature and once at a temperature 10 °C below the T_d. After air-drying, the hybridization signal was detected using a phospho-imager (Typhoon 8600™, Amersham). Computer-generated images were analysed using Image Quant software (Amersham), in order to quantify the optical density for each spot. These measures were normalized against the corresponding optical density for loaded RNA. The values were exported in an Excel file and used to generate histograms, where the highest signal was arbitrarily set as 100. Background and values not significantly different from the background were set as zero.

miRNA target gene prediction and 5' RACE

Target gene identification have been conducted analysing public maize expressed sequence tag (EST) databases (http://www.ncbi.nlm.nih.gov), and The Institute for Genomic Research (http://www.tigr.org) databases using the BLASTN program (Altschul *et al.*, 1990). When necessary, parameters were set to allow the program to deal with short nucleic acid sequences.

In addition, a new software, specifically designed for plant genomes, was utilized to find miRNA targets in maize. This software is available at the miRU web server http://bioinfo3.noble.org/miRU.htm (Zhang, 2005).

To validate target gene prediction, a rapid amplification of 5′ cDNA ends (5′ RACE) assay was performed, using the First-Choice RLM-RACE kit (Ambion). Total RNA was extracted with TRizol® reagent (Invitrogen) from tissues where abundant miRNA expression was detected, and poly(A)$^+$ mRNA was purified with an mRNA Purification kit (Amersham); poly(A)$^+$ mRNA was directly ligated, without other enzymatic pretreatment, to the 45 nucleotide adaptor from the kit. Subsequent steps were according to the manufacturer's instructions. For each annotated tentative contig (TC), two reverse primers were designed, with the respective forward primer used to check the presence of the specific cDNA molecules, and to assess that the working conditions were correct (supplementary Table S1 with gene-specific primers is available at JXB online); nested PCRs were performed with reverse primers and adaptor-specific primers. PCR products were subsequently sequenced.

Chromosome mapping

Lyophilized green tissues of OMA lines were kindly provided by Professor Ronald Phillips of the Department of Agronomy and Plant Genetics, University of Minnesota. DNA was extracted with the CTAB method and a PCR (35 cycles, 57 °C as annealing temperature, 50 ng of genomic DNA from OMA lines per reaction tube) was performed with pairs of primers specifically designed for this protocol (see supplementary Table S2 at JXB online). PCR products, ~200 nucleotides long, were run on a 1% agarose gel and stained with ethidium bromide.

Phylogenetic analysis

MIR gene sequences from maize, rice, and sorghum were collected, and the reciprocal best BLAST hit (RBH) method was performed running a BLASTN algorithm and selecting the first three best matches, always setting the lowest acceptable E-value limit to 1E-9, 1E-7, and 1E-06 when blasting on sorghum, rice, and maize databases, respectively. The data sources for the genomic sequences used in this analysis are the following: TIGR database (http://maize.tigr.org/release4.0/assembly.shtml) for maize, the NCBI database (www.ncbi.nlm.nih.gov) for rice, and the GSS division of GenBank (www.ncbi.nlm.nih.gov) for sorghum.

Results

Research effort was focused on the following five miRNA families: miR156, miR160, miR166, miR167, and miR169, for which, upon searching the Rfam database (www.sanger.ac.uk/Software/Rfam), 30 different pre-miRNAs were identified.

Similarity between maize, rice, and sorghum MIR genes

Data on these five miRNA families in maize were mostly derived by bioinformatics analyses, which hypothesized pre-miRNA molecules as putative precursors from an *in silico* prediction.

Since phylogenetic conservation of the 21 nucleotide maize miRNA sequence and its predicted precursor secondary structure is considered an important validation for MIR genes (Ambros *et al.*, 2003), the relationship between maize, rice, and sorghum MIR genes was addressed.

The very limited sequence conservation among different precursors of the same miRNA both within and among species did not allow a direct investigation on the phylogenetic relationship existing between MIR genes. Conventional PHYLIP analysis was in fact unable to estimate phylogenetic distances existing between sequences in these three evolutionarily related species. Using the evidence that loop sequence is less conserved than stem sequence, phylogenetic analysis was restricted to the sequence of the region comprising the miRNA and the anti-miRNA (~60–80 nucleotides). Again, even if a higher degree of similarity was found in these short sequences, phylogenetic analysis was still meaningless. However, the application of the RBH method provided a useful hint regarding putative homology among a number of maize, rice, and sorghum MIR genes. According to this method, two sequences A and B are reciprocal best hits if sequence B is the best hit from the organism of sequence B when sequence A is the query sequence, and vice versa. Several databases that collect orthologous sequences, such as cluster of orthologous groups (COG), use this method both in prokaryotes and in eukaryotes as the basis for several algorithms (Li *et al.*, 2003). The method was applied to find RBHs between maize and sorghum and between maize and rice, in order to define groups of orthologous genes. The results of this analysis are summarized in Tables 1 and 2, where it can be observed that many maize precursors have a putative orthologue in both rice and sorghum. Twenty-three maize genes were found to have a rice orthologue, and for 14 of them there is also a sorghum orthologue. For zma-MIR166f and zma-MIR166h, only the sorghum orthologue was found. In addition, only one putative common orthologue in sbi-MIR166b was found for these two genes, and three other sorghum genes have a double orthologue in maize (sbi-MIR169a, sbi-MIR166a, and sbi-MIR167a). The results are only partial, due to the largely incomplete sorghum database; however, interesting clues can be derived from the maize–rice comparison. For example, two maize genes—zma-MIR156b and zma-MIR156c—positioned in tandem configuration, identified two rice genes also in tandem configuration 250 bp apart. In addition, in six other cases, two maize precursors correspond to a single rice precursor, according to the tetraploid origin of maize (Gaut and Doebley, 1997; Swigonova *et al.*, 2004).

Chromosome mapping

As a first step for mapping MIR genes in maize, OMA lines were used to place MIR sequences on individual maize chromosomes (Okagaki *et al.*, 2001). Specific primers were

Table 1. *Reciprocal best BLAST hit between maize and rice sequences*

Maize MIR genes		Best BLAST hit on rice			Rice MIR genes		Best BLAST hit on maize		
Name	Corresponding accession no.[a]	Rice accession no.[b]	E-value	Bit score	Name	Corresponding accession no.[b]	Maize accession no.[a]	E-value	Bit score
zma-MIR156c	AZM4_138317	AP002836 AP002816	4E-13	78	osa-MIR156b	AP002836, AP002816[c]	AZM4_138317	2.6 E-14	452
zma-MIR156b	AZM4_138317	AP002836 AP002816	8E-23	109	osa-MIR156c	AP002836, AP002816[c]	AZM4_138317	2.5 E-06	280
zma-MIR156e	AZM4_13978	AL606594 AL442115	8E-24	113	osa-MIR156e	AL442115 AL606594	AZM4_13978 AZM4_39102	5.8 E-13 5.1 E-12	426, 367
zma-MIR156h	AZM4_39102	AL606594 AL442115	8E-27	123					
zma-MIR156i	AZM4_69013	AP005286	4E-40	167	osa-MIR156i	AP005286	AZM4_69013 AZM4_13978	1.2 E-12 8.4 E-11	422, 379
zma-MIR156d	AZM4_114270	AP004087	3E-26	121	osa-MIR156d	AP004087	AZM4_114270	5.8 E-15	473
zma-MIR160b	AZM4_32674	AP004989 AP004863	2E-24 2E-14	115, 82	osa-MIR160a	AP004863	AZM4_32674 AZM4_105618	2.0 E-08 3.4 E-08	330, 312
zma-MIR160a	AZM4_105618	AP004989	6E-07	56	osa-MIR160b	AP004989	AZM4_32674 AZM4_105618	1.9 E-12 9.9 E-11	419, 361
zma-MIR160d	AZM4_135674	AC084320 AC091123	1E-10	70	osa-MIR160d	AC084320, AC091123	AZM4_135674 AZM4_15265	0.0076 0.0100	205, 203
zma-MIR160e	AZM4_15265	AC084320 AC091123	7E-08	61					
zma-MIR160c	AZM4_108039	AP004591	2E-13	78	osa-MIR160c	AP004591	AZM4_108039	3.2 E-09	343
zma-MIR166e	AZM4_1320 OGUER01TV	AP004868 AP004839	9E-14	82	osa-MIR166d	AP004868 AP004839	AZM4_70223 OGUER01TV AZM4_1320	3.9 E-11 1.8 E-10 8.7 E-10	389 359
zma-MIR166c	AZM4_70223	AP004868 AP004839	1E-13	80					
zma-MIR166b	AZM4_26017	AE017112 AC025296	1E-15	86	osa-MIR166a	AC025296	AZM4_26017 AZM4_35871	8.9 E-12 1.4 E-10	401, 377
zma-MIR166d	AZM4_35871	AE017112 AC025296	3E-16	88					
zma-MIR166i	AZM4_51640	AC027660 AE017094	2E-08	62	osa-MIR166f	AC027660 AC025044	AZM4_51640	5.1 E-08	320
zma-MIR166g	AZM4_91065	AC104428	9E-10	66	osa-MIR166c	AC104428	AZM4_91065	9.0 E-09	216
zma-MIR167b	AZM4_2152 PUGHC59TB	AC092558	4E-13	78	osa-MIR167b	AC092558	PUGHC59TB AZM4_2152	5.2 E-10 2.8 E-09	226
zma-MIR167a	AZM4_102805	AL928781	6E-12	74	osa-MIR167a	AL928781	AZM4_102805	7.0 E-09	205
zma-MIR167c	AZM4_32681	AC092076	8E-14	80	osa-MIR167c	AC092076	AZM4_3101 AZM4_32681	9.2 E-14 0.00015	269, 245
zma-MIR167d	AZM4_3101	AC092076	7E-18	94					
zma-MIR169a	AZM4_40748	AP002817	4E-10	68	osa-MIR169	AP002817	AZM4_40748 AZM4_59852	5.0 E-13 6.0 E-06	
zma-MIR169b	AZM4_59852	AP002817	7E-09	64					

[a] Genomic contigs from methylation-filtered and highCot maize genomic clones database of TIGR (http://maize.tigr.org/release4.0/assembly.shtml).
[b] Rice genomic bacterial artificial chromosome (BAC) clones.
[c] These two BAC clones overlap in the MIR genes region.

Table 2. *Reciprocal best BLAST between maize and sorghum sequences*

Maize MIR genes		Best BLAST hit on sorghum			Sorghum MIR genes		Best BLAST hit on maize		
Name	Corresponding accession no.[a]	Sorghum accession no.[b]	E-value	Bit score	Name	Corresponding accession no.[b]	Maize accession no.[a]	E-value	Bit score
zma-MIR156d	AZM4_114270	CW199989 CW316697	1E-51	204	sbi-MIR156a	CW199989 CW468726 CW460595 CW316697	AZM4_114270 AZM4_80916	2.5 E-22 7.8 E-11	635 382
zma-MIR160b	AZM4_32674	CL185832 CL185833 CW079321	4E-61 4E-030	236 133	sbi-MIR160b	CL185832 CL185833	AZM4_32674 AZM4_105618	8.7 E-23 9.0 E-10	647 343
zma-MIR160d	AZM4_135674	CW228260 CL174305	1E-011 2E-010	72 68	sbi-MIR160d	CL174305	AZM4_135674	1.5 E-10	373
zma-MIR160c	AZM4_108039	CW378233 CW378234	2E-015	84	sbi-MIR160c	CW378233 CW378234	AZM4_108039	5.3E-14	446
zma-MIR160e	AZM4_15265	CW315135 CL151764	1E-019	99	sbi-MIR160e	CW315135	AZM4_105590 AZM4_135674 AZM4_15265	1.5 E-07 8.6 E-07 6.4 E-0	309 291 273
zma-MIR166h	AZM4_124162	CW037452 CW046554 CW178934 CW284000	2E-015	84	sbi-MIR166b	CW178934 CW046554 CW037452 CW284002 CW284000	AZM4_124162 AZM4_53768	3.3 E-17 2.4 E-12	524 414
zma-MIR166f	AZM4_53768	CW037452 CW046554 CW178934 CW284000	2E-009	64					
zma-MIR166e	AZM4_1320 OGUER01TV	CL183678 CW320001 CW478952	3E-48	193	sbi-MIR166d	CW320001 CL183678 CW478952	OGUER01TV AZM4_1320 AZM4_70223	2.4 E-14 1.1 E-13 4.3 E-12	445 410
zma-MIR166c	AZM4_70223	CL183678 CW320001 CW478952	7E-41	168					
zma-MIR166b	AZM4_26017	CL172555 CW110153 CW170188	1E-020 4E-017	101 90	sbi-MIR166a	CL172555 CW170188 CW110153 CW110154	AZM4_26017 AZM4_35871 AZM4_37606	7.7 E-19 2.3 E-16 1.1 E-13	555 504 440
zma-MIR166d	AZM4_35871	CL172555 CW110153 CW170188	1E-016	88					
zma-MIR167b	AZM4_2152 PUGHC59TB	CL172387 CW240599 CW434124	1E-24, 4E-015	115 84	sbi-MIR167b	CL172387 CW434124	PUGHC59TB AZM4_2152	9.5 E-09 2.1 E-08	328
zma-MIR167c	AZM4_32681	CL173383 CL173384 CW135847 CW135848 CW255985 CW255986	9E-49	194	sbi-MIR167a	CL173383 CW255985 CW135847 CW135848 CW255986 CL173384	AZM4_3101 AZM4_32681	1.1 E-20 1.8 E-19	595 574
zma-MIR167d	AZM4_3101	CL173383 CL173384 CW135847 CW135848 CW255986	4E-055	216					
zma-MIR169a	AZM4_40748	CW428429	8E-22	105	sbi-MIR169a	CW428429	AZM4_40748 AZM4_59852	1.1 E-18 1.1 E-09	553 356
zma-MIR169b	AZM4_59852	CW428429	2E-23	111					

[a] Genomic contigs from methylation-filtered and highCot maize genomic clones database of TIGR (http://maize.tigr.org/release4.0/assembly.shtml).
[b] Genomic contig from methylation-filtered sorghum genomic clones of the GSS division of GenBank (www.ncbi.nlm.nih.gov).

designed for each miRNA precursor and standard PCR was performed on the genomic DNA of OMA lines (see Fig. 1, for one example).

This approach allowed 23 MIR genes to be assigned to a chromosome, as summarized in Table 3, whereas seven sequences were not amplified on OMA lines.

Tissue and genotypic expression profile

The expression of the five miRNAs was tested by RNA gel blot analysis on samples from seedlings, developing ear, and developing kernels (10 d after pollination) derived from two inbred lines and their corresponding F_1 hybrid, as shown in Fig. 2 (upper part of each panel). The intensity of each spot has been subsequently measured and normalized, as described in the Materials and methods, to create the histograms shown in the lower part of the panels in Fig. 2.

All five miRNAs are present in the RNA fractions, with a size of ~20 nucleotides, showing tissue and genotype specificity of the expression pattern. With the exception of miR156, which is expressed mainly in seedlings, the other miRNAs are present in two or all the analysed tissues; for example, miR160 and miR169 are mainly expressed in seedlings and developing ears, and miR167 is present in developing kernels and seedlings.

Genotypic specificity is also evident, indicating that different genotypes express the same miRNA in the same tissues, but at different levels. miR166, for instance, shows higher expression in kernels from inbred line H99 than B73 and the B73×H99 F_1 hybrid. Furthermore, miR167 is expressed at similar levels in seedlings and kernels in both inbred lines, but shows a different pattern in the F_1 hybrid; similarly, miR169 has a similar expression pattern in the B73 inbred line and F_1 hybrid, which is different in the H99 inbred line.

Target gene identification

Given the high homology between miRNAs and their targets in plants, the maize EST database was searched for homology to the five miRNAs sequences using a BLASTN algorithm.

In 41 ESTs, regions complementary to the five miRNA sequences were identified, in which fewer than three mismatches were present. This search was refined comparing the 41 ESTs with the TIGR database (http://www.tigr.org/tdb/tgi/plant.shtml) in which most of the maize ESTs are assembled into TCs. Twelve TCs and a few singletons have been found; for some of these TCs, a function has been annotated. The same procedure was applied to the sorghum data set. Six TCs resulted as putative targets for the five miRNAs analysed.

Very recently, the miRU software was developed (Zhang, 2005) allowing an automated on-line search of miRNA targets to be performed. This was searched for both maize and sorghum miRNA targets, using rice as the reference organism for gene function conservation. Targets with a function not conserved between the two plant species are defined as possible false positives. A large list of putative target genes was thus produced, containing all the putative target genes previously identified by means of a manual search. In Table 4, conclusive results are shown. It appears that the putative functions of target TCs are the same as those reported for *Arabidopsis* and rice (Bartel and Bartel, 2003).

To verify the nature of the predicted miRNA targets, a modified 5′ RACE experiment was set up, as described in the Materials and methods. This is one of the most common and widely used methods in the literature (Llave *et al*., 2002*b*; Kasschau *et al*., 2003; Rhoades and Bartel, 2004; Lauter *et al*., 2005; Lu *et al*., 2005) to support bioinformatics data. The 5′ RACE analysis is based on the evidence that miRNA-cleaved transcripts have a phosphorylated 5′ end, which can be directly ligated to the RNA adaptor, and that maps to the 10th nucleotide of the coupled miRNA. Given the clear tissue-specific pattern of expression of miR156 (Fig. 2), these analyses were performed on a few of its putative targets, using the RNA extracted from B73 seedlings, where miR156 is more abundantly expressed. TC294022 and TC280157 were confirmed as real miR156 targets, since the 5′ end of the RACE product maps to the 9th and 10th nucleotide, respectively, of the coupled miRNA (Fig. 3A, B). These TCs are similar to two rice proteins coding for squamosa promoter-binding proteins. Similarly, TC306179 was confirmed as a real target for miR169 in B73 seedlings. In fact, the 5’ end of the RACE product maps to the 10th nucleotide of the coupled miRNA (Fig. 3C). TC306179 codes for a protein highly homologous to a CCAAT-binding transcription factor.

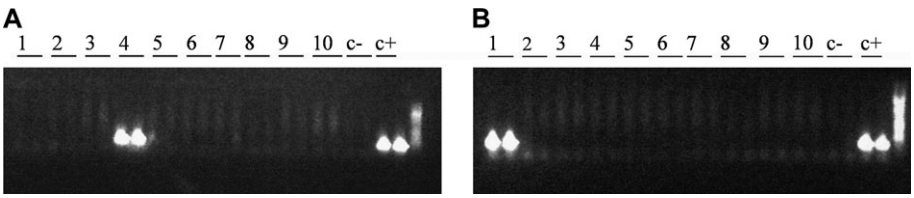

Fig. 1. PCR amplification on OMA lines of maize putative MIR genes. PCR amplifications were conducted in duplicate on genomic DNA of OMA lines and were loaded onto an agarose gel, and stained with ethidium bromide. The 10 different addition lines (one for each maize chromosome) are loaded from left to right (1=OMA1; 2=OMA2, etc.). c−= negative control, oat genomic DNA from genotype SunII; c+= positive control, maize genomic DNA for genotype Seneca 60. (A) Amplification for zma-MIR160a. (B) zma-MIR167d.

Table 3. *Chromosomal localization of maize MIR genes*

The chromosome localization for each MIR gene is shown as found from amplification on OMA DNA panels as detailed in the text.

MIR name	Chromosome
zma-MIR156a	4
zma-MIR156b	3
zma-MIR156c	3
zma-MIR156d	5
zma-MIR156e	2
zma-MIR156f	2
zma-MIR156g	7
zma-MIR156h	n.a.
zma-MIR156i	n.a.
zma-MIR160a	4
zma-MIR160b	n.a.
zma-MIR160c	6
zma-MIR160d	7
zma-MIR160e	7
zma-MIR166b	1
zma-MIR166c	n.a.
zma-MIR166d	5
zma-MIR166e	5
zma-MIR166f	n.a.
zma-MIR166g	1
zma-MIR166h	1
zma-MIR166i	1
zma-MIR166j	n.a.
zma-MIR166k	5
zma-MIR167a	n.a.
zma-MIR167b	5
zma-MIR167c	5
zma-MIR167d	1
zma-MIR169a	3
zma-MIR169b	8

n.a.= not assigned.

Discussion

Evolutionary conservation of MIR genes between maize and rice

miRNAs have recently emerged as important players in plant development and in participating in the regulation of plant response to stresses (Carrington and Ambros, 2003; Hake, 2003; Sunkar and Zhu, 2004; Kidner and Martienssen, 2005).

Most current knowledge on plant miRNAs and their corresponding MIR genes derives directly from studies on the model plant *Arabidopsis* and from a bioinformatics scan on the rice genome. Therefore, for a better understanding of phenotypic features crucial for the economic success of crop species, it is of relevance to identify and determine the function of miRNAs in other crops. In this respect, maize and sorghum are of double interest, since their evolutionary closeness to rice could prove important for addressing the function of miRNAs and their evolution in grasses, not yet investigated by other groups.

It was challenging to establish a degree of correlation between miRNA 'families' among these three grasses, because only the rice genome has been completely sequenced and it is known that grasses exhibit a high rate of duplication events (Gaut, 2001). In addition, miRNAs do not show clear patterns of sequence conservation, making it very difficult to define the evolutionary history of these sequences.

It is possible that some MIR genes might have a common ancestor, but many might have evolved recently by duplication and/or translocation events or, as recently hypothesized, might have been generated *ex novo* from target duplication (Allen *et al.*, 2004).

Despite all the limitations of the analysis conducted here, the finding that at least some precursors of the same miRNA are very similar among species is of relevance. In the literature, there are no reports on the analysis of evolutionary relationship among MIR genes in rice, maize, and sorghum. For the features of the sequences analysed, the RBH proved the most reliable and useful method. E-values together with bit scores reported for the BLAST analysis corroborate the conclusion. In three cases (Table 1), blasting on maize database E-values was higher, even if the alignment and bit score are acceptable. This depends on the presence in those sequences of regions rich in short repeats.

The data summarized in Tables 1 and 2 show that there are groups of homologous genes among the five different miRNAs 'families'. A number of maize MIR genes found an orthologue in the rice genome, and a smaller number found an orthologue in the sorghum genome, despite the closer relationship between sorghum and maize. This is obviously due to the largely incomplete sorghum data set.

As expected from the allotetraploid origin of maize, two different maize MIR genes often correspond to a unique rice or sorghum gene. There is one particular example of a duplication event which occurred prior to speciation: MIR156b and MIR156c in tandem configuration are present both in maize and in rice at a comparable distance (100 and 250 nucleotides). This observation is in agreement with the work of Guddeti *et al*. (2005) assessing that many rice MIR genes are organized in clusters.

Another interesting result is that relating to zma-MIR166f and zma-MIR166h, which seem to have an orthologues only in sorghum but not in rice. This group might have been generated after divergence between rice and the progenitor of both maize and sorghum. A more specific algorithm should be developed and further analysis should be performed to define the nature of these events precisely.

Expression profile of maize miRNAs

RNA blot analysis clearly demonstrates that 20–21 nucleotide long RNA molecules, possibly derived from one or more of the putative pre-miRNAs identified, are expressed in various maize tissues. These analyses indicate that there are differences in the amount of the various miRNAs in the different tissues tested, also depending on the genotype utilized. For instance, miR156 gave a positive signal only

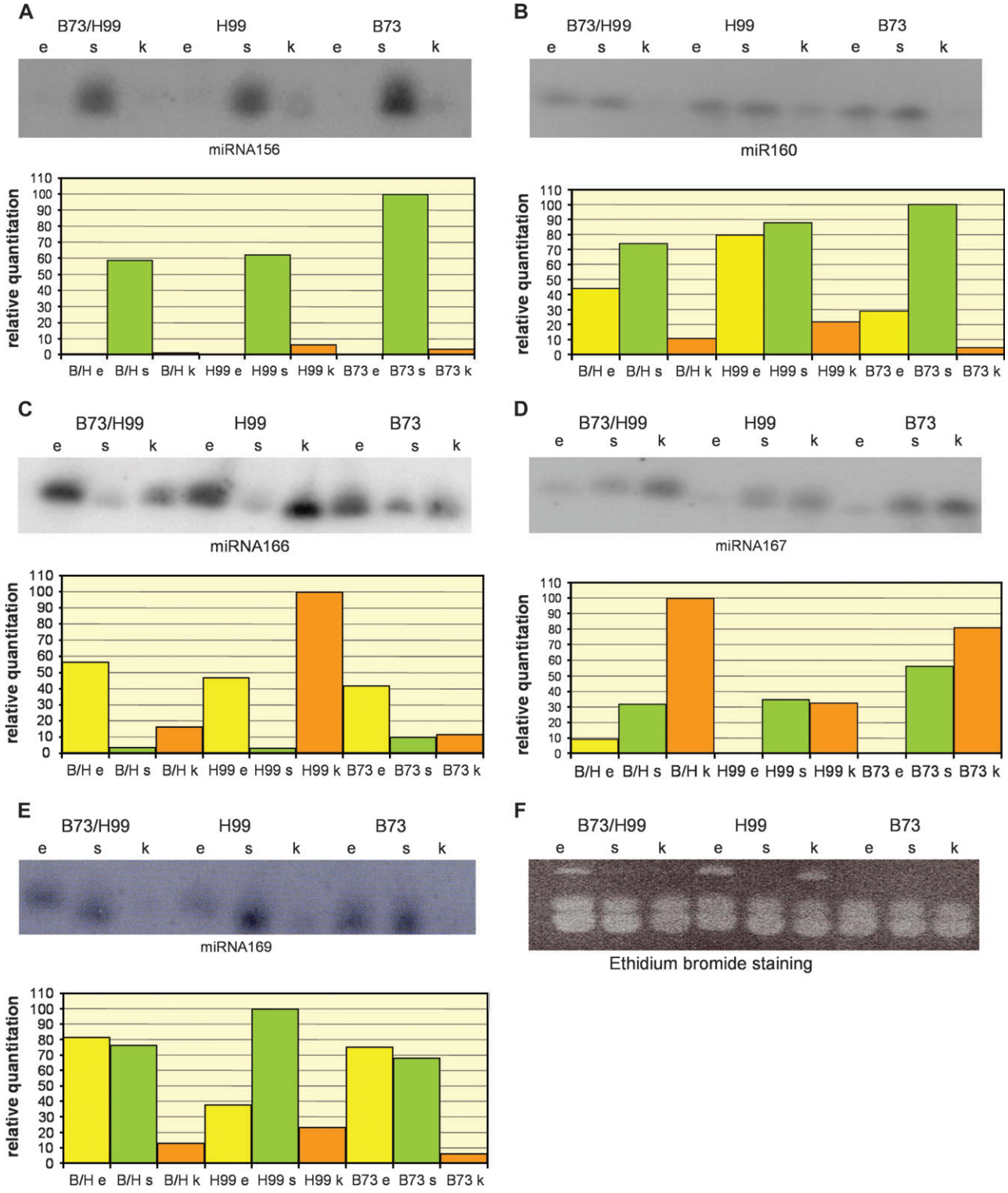

Fig. 2. Expression analysis of five maize miRNAs. (A–E) Upper panels: RNA gel blots of RNA, enriched for the low molecular weight fraction, isolated from different tissues and different genotypes were probed with end-labelled oligonucleotides specific for miR156(A), miR160(B), miR166(C), miR167(D), and miR169 (E). Lower panels: histograms showing the relative quantization of RNA gel blots, calculated as described in the Materials and methods. E, developing ears (yellow); s, seedlings (green); k, kernels at 10 d after pollination (orange), from H99 and B73 parental lines and their F_1 hybrid B73×H99. (F) RNA loaded onto the polyacrylamide gel and visualized by ethidium bromide staining.

Table 4. *Putative miRNA target genes in maize*

miRNA	Target function in rice	Maize TC putative target	Conserved with rice	Sorghum TC putative target	Conserved with rice
156	Squamosa promoter-binding protein (SBP)-like protein	TC284975		CD230908	
		TC282875	yes	TC108912	
		TC305894	yes	AW747167	yes
		TC302436		TC98039	
		TC305612	yes	TC107024	
		TC294022	yes	TC100722	
		TC280157	yes	CF071767	
		TC280160	yes	CN152532	
		TC294023		TC103749	
		TC304431		BG158029	
		TC311582		TC95574	
		TC295816			
		TC305845			
		TC304104			
160	Auxin response factor	TC295067		BM325690	
		CF648203		CF481512	
		TC310566		TC93604	
				TC97122	
166	HD-zip transcription factor	TC286620	yes	TC110352	yes
		TC281649		TC111000	
167	Auxin response factor	TC310360		TC109994	
		TC310360	yes	TC94769	
		CF630597		TC98839	
		TC309752			
		CF0055534			
169	CCAAT-binding factor HAP-2-like proteins	CD960569		CD206089	
		TC281368		TC108495	
		TC306902		TC104953	yes
		TC288454	yes	BM328363	
		TC307223	yes	CD423929	
		TC295929	yes		
		TC300826	yes		
		TC295655	yes		
		TC306179	yes		

in seedlings, whereas the other four miRNAs are expressed, although at different levels, in two or three tissues. As for genotypic differences, it appears that the F_1 hybrid shows a signal intensity similar to that of the parental inbred line with the strongest signal (see miR167 in 10 kernels 10 d after pollination and miR169 in developing ears), although this is not the case for miR166 expression in kernels 10 d after pollination.

miRNA target prediction

To assess and define a putative function of miRNA molecules in the grasses, and specifically in maize, a further step is represented by target identification.

The most efficient tool available is certainly the bioinformatics approach that in plants is facilitated by the high degree of homology between miRNA and target sequences (Rhoades *et al*., 2002; Bartel and Bartel, 2003). In fact this was also the case in our analysis where a long list of putative targets was obtained. Our analysis reveals, as expected, that many of the predicted targets, both in maize and in sorghum, have a conserved function with other plant miRNA targets. Nonetheless, some TCs with unknown function or with a function distinct from *Arabidopsis* or rice genes are also present, and they could be targets involved in processes that are species specific, or tissues specific (Lu *et al*., 2005).

Three target genes were experimentally validated, two for miR156 (TC294022 and TC280157) and one (TC306179) for miR169. The validation was obtained by performing the modified 5′ RACE protocol on mRNA extracted from seedlings, where it was previously demonstrated that both miR156 and miR169 are abundant. miR156 targets code for proteins similar to two rice proteins (Q6H508 and Q6Z461) homologous to squamosa promoter-binding protein, which is a plant-specific family of transcript factors involved in early flower development. This evidence is in agreement with the expression pattern of miR156 that is present at a very low level, if any, in developing ears (Fig. 2). Also miR169, which is expressed mainly in seedlings and in immature ears, has a validated target. TC306179 codes for a protein highly homologous to a rice protein (Q851D5) that is the B subunit of the CCAAT. This protein belong to a family common to many

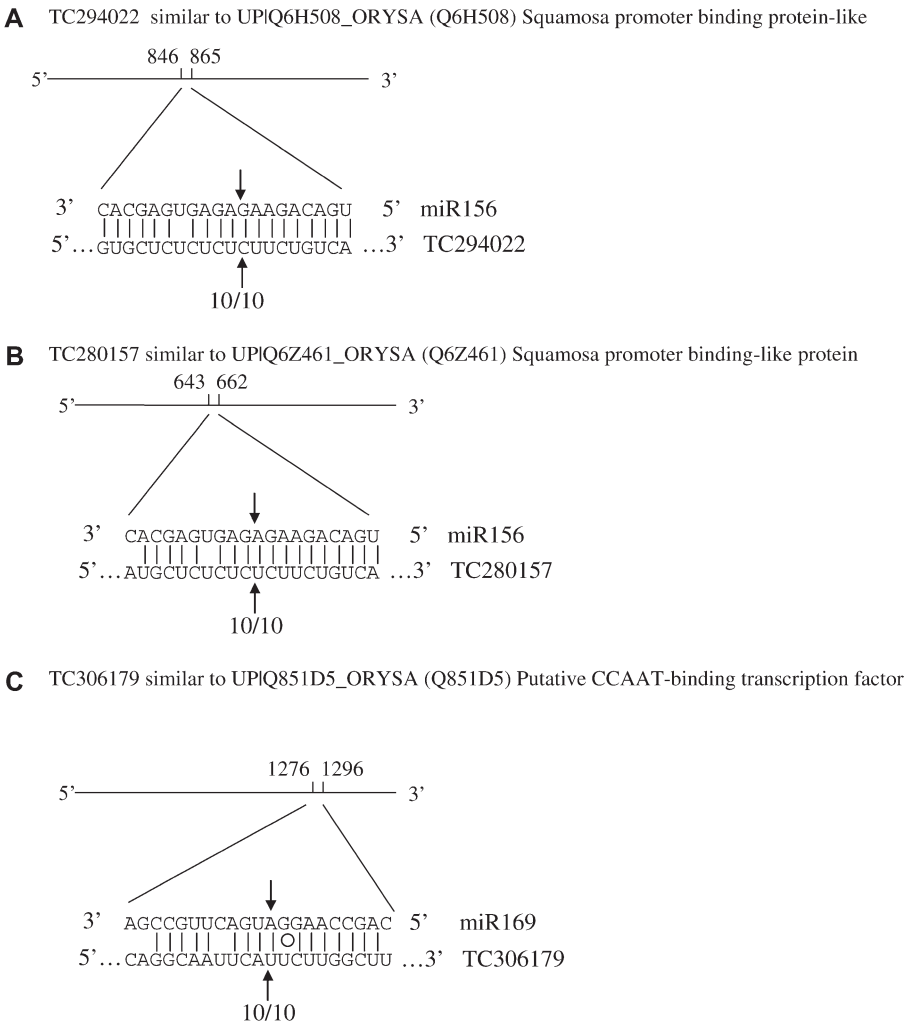

Fig. 3. Experimental validation of putative miRNA targets. The results of modified 5′ RACE are shown. The upper line represents TC; 5′–3′ orientation and the position of the miRNA complementary sites are indicated. The miRNA sequence and its complementary target sequence are reported in the expanded region. Watson–Crick pairing (vertical dashes) and G:U wobble pairing (circles) are indicated. Black arrows indicate the 5′ termini of cleaved mRNA, isolated from the tissues, with the frequency of clones shown. (A, B) Two miR156targets: TC294022 and TC280157; (C) TC306179, the miR169 target.

eukaryotes (human, mammals, yeast, nematodes, fungi, and green plants). In *Arabidopsis*, many of the CCAAT-binding transcription factors such as leafy cotyledon 1 and 2 (LEC1 and LEC2) are involved in embryogenesis induction and embryo development. It is interesting to notice that in maize, miR169 has a lower expression in kernels at 10 d after pollination (Fig. 2) where CCAAT-binding factor should be actively present. Obviously the present data have to be taken solely as suggestions and intriguing clues about a possible role for these miRNA families in maize. Further analysis should be performed to define and investigate these aspects.

Four additional TCs, considered as putative targets of miR156 and miR169, were also tested. Since no results were obtained, despite always using the same tissue as poly(A)$^+$ RNA source, no conclusions can be drawn regarding target validation.

Given the present limited knowledge on MIR gene function in maize, it is premature to speculate about a possible role for miRNAs in the phenomenon of heterosis. However, reports in the literature suggest that quantity, timing, and quality of gene expression could account for phenotypic differences between inbred lines and their corresponding F_1 hybrid (Romagnoli *et al*., 1990; Leonardi *et al*., 1991; Osborn *et al*., 2003; Song and Messing, 2003). Since a single miRNA could be produced by post-transcriptional processing of different primary transcripts, it would be of interest to verify if different primary transcripts are used in different tissues and/or in different growing conditions. In addition, *cis* regulation of allele-specific expression seems to be quite a common phenomenon in maize hybrids (Guo *et al*., 2004). It is not unreasonable to hypothesize that a similar mechanism is also acting on MIR genes in heterozygous plants.

Acknowledgements

We thank Dr D Horner, for useful discussion on the analysis of homologies between sequences, Dr I Salomoni for her help on OMA lines, and Dr G Pea for helpful discussion and advice. A special thanks to Professor R Phillips of the University of Minnesota for providing us with OMA lines. This work was supported by MIUR Cofin Program 2003 'Molecular and quantitative analysis of heterosis in maize'.

References

Allen E, Xie Z, Gustafson AM, Sung GH, Spatafora JW, Carrington JC. 2004. Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*. *Nature Genetics* **36,** 282–1290.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215,** 403–410.

Ambros V, Bartel B, Bartel DP, *et al.* 2003. A uniform system for microRNA annotation. *RNA* **9,** 277–279.

Aukerman MJ, Sakai H. 2003. Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-like target genes. *The Plant Cell* **15,** 2730–2741.

Bartel B, Bartel, DP. 2003. MicroRNAs: at the root of plant development? *Plant Physiology* **132,** 709–717.

Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116,** 281–297.

Bedell JA, Budiman MA, Nunberg A, *et al.* 2005. Sorghum genome sequencing by methylation filtration. *PLoS Biology* **3(1),** e13.

Bernstein E, Caudy AA, Hammond SM, Hannon GJ. 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409,** 363–366.

Bonnet E, Wuyts J, Rouzè P, Van de Peer Y. 2004. Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes. *Proceedings of the National Academy of Sciences, USA* **101,** 11511–11516.

Carrington JC, Ambros V. 2003. Role of microRNAs in plant and animal development. *Science* **301,** 336–338.

Chen X. 2003. A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. *Science* **303,** 2022–2025.

Gaut BS. 2001. Patterns of chromosomal duplication in maize and their implications for comparative maps of the grasses. *Genome Research* **11,** 55–66.

Gaut BS, Doebley JF. 1997. DNA sequence evidence for the segmental allotetraploid origin of maize. *Proceedings of the National Academy of Sciences, USA* **94,** 6809–6814.

Guddeti S, Zhang DC, Li AL, Leseberg CH, Kang H, Li XG, Zhai WX, Johns MA, Mao L. 2005. Molecular evolution of the rice miR395 gene family. *Cell Research* **15,** 631–638.

Guo HS, Xie Q, Fei JF, Chua NH. 2005. MicroRNA directs mRNA cleavage of the transcription factor NAC1 to downregulate auxin signals for arabidopsis lateral root development. *The Plant Cell* **17,** 1376–1386.

Guo M, Rupe MA, Zinselmeier C, Habben J, Bowen BA, Smith OS. 2004. Allelic variation of gene expression in maize hybrids. *The Plant Cell* **16,** 1707–1716.

Hake S. 2003. MicroRNAs: a role in plant development. *Current Biology* **13,** 851–852.

Hammond SC, Bernstein E, Beach D, Hannon GJ. 2000. An RNA-directed nuclease mediates posttranscriptional gene silencing in *Drosophila* cells. *Nature* **404,** 293–296.

Juarez MT, Kui JS, Thomas J, Heller BA, Timmermans MCP. 2004. MicroRNA-mediated repression of rolled leaf1 specifies maize leaf polarity. *Nature* **428,** 84–88.

Kasschau KD, Xie Z, Allen E, Llave C, Chapman EJ, Krizan KA, Carrington JC. 2003. P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA function. *Developmental Cell* **4,** 205–217.

Kidner CA, Martienssen RA. 2004. Spatially restricted microRNA direct leaf polarity through ARGONAUTE1. *Nature* **428,** 81–84.

Kidner CA, Martienssen RA. 2005. The developmental role of microRNA in plants. *Current Opinion in Plant Biology* **8,** 38–44.

Kurihara Y, Watanabe Y. 2004. Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. *Proceedings of the National Academy of Sciences, USA* **101,** 12753–12758.

Kynast RG, Riera-Lizarazu O, Vales MI, *et al.* 2001. A complete set of maize individual chromosome additions to the oat genome. *Plant Physiology* **125,** 1216–1227.

Lauter N, Kampani A, Carlson S, Goebel M, Moose SP. 2005. MicroRNA172 down-regulates glossy15 to promote vegetative phase change in maize. *Proceedings of the National Academy of Sciences, USA* **102,** 9412–9417.

Lee Y, Ahn C, Han J, *et al.* 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425,** 415–419.

Leonardi A, Damerval C, Herbert Y, Gallais A, DeVienne D. 1991. Association of protein amount polymorphisms (PAP) among maize lines with performances of their hybrids. *Theoretical and Applied Genetics* **82,** 552–560.

Li L, Christian J, Stoeckert J, Roos DS. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Research* **13,** 2178–2189.

Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song J, Hammond SM, Joshua-Tor L, Hannon GJ. 2004. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305,** 1437–1441.

Llave C, Kasschau KD, Rector MA, Carrington JC 2002*a*. Endogenous and silencing-associated small RNAs in plants. *The Plant Cell* **14,** 1605–1619.

Llave C, Xie Z, Kasschau KD, Carrington JC. 2002*b*. Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. *Science* **297,** 2053–2056.

Lu S, Sun YH, Shi R, Clark C, Li L, Chiang VL. 2005. Novel and mechanical stress-responsive microRNAs in *Populus trichocarpa* that are absent from Arabidopsis. *The Plant Cell* **17,** 2186–2203.

Maher C, Timmermans M, Stein L, Ware D. 2004. Identifying microRNA in plant genomes. *Proceedings of the 2004 IEEE Computational System Bioinformatics Conference*, 16–19 August 2004, Stanford, USA. IEEE Computer Society, 718–723.

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.

Meyers BC, Vu TH, Tej SS, Ghazal H, Matvienko M, Agrawal V, Ning J, Haudenschild CD. 2004. Analysis of the transcriptional complexity of *Arabidopsis thaliana* by massively parallel signature sequencing. *Nature Biotechnology* **22,** 1006–1011.

Okagaki RJ, Kynast RG, Livingston SM, Russel CD, Rines HW, Phillips RL. 2001. Mapping maize sequences to chromosomes using oat–maize chromosome addition materials. *Plant Physiology* **125,** 1228–1235.

Osborn TC, Pires JC, Birchler JA, *et al.* 2003. Understanding mechanisms of novel gene expression in polyploids. *Trends in Genetics* **19,** 141–147.

Palatnik JF, Allen E, Wu X, Schommer C, Schwab R, Carrington JC, Weigel D. 2003. Control of leaf morphogenesis by microRNAs. *Nature* **425,** 257–263.

Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP. 2002. MicroRNA in plants. *Genes and Development* **16,** 1616–1626.

Rhoades MW, Bartel DP. 2004. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Molecular Cell* **14,** 787–799.

Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP. 2002. Prediction of plant microRNA target. *Cell* **110,** 513–520.

Romagnoli S, Maddaloni M, Livini C, Motto M. 1990. Relationship between gene expression and hybrid vigor in primary root tips of young maize (*Zea mays L.*) plantlets. *Theoretical and Applied Genetics* **80,** 769–775.

Song R, Messing J. 2003. Gene expression of a gene family in maize based on non-collinear haplotypes. *Proceedings of the National Academy of Sciences, USA* **100,** 9055–9060.

Sunkar R, Zhu JK. 2004. Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. *The Plant Cell* **16,** 2001–2019.

Sunkar R, Girke T, Jain PK, Zhu JK. 2005. Cloning and characterization of microRNA from rice. *The Plant Cell* **17,** 1397–1411.

Swigonova Z, Lai J, Ma J, Ramakrishna W, Llaca V, Bennetzen JL, Messing J. 2004. Close split of sorghum and maize genome progenitors. *Genome Research* **14,** 1916–1923.

Vaucheret H, Vazquez F, Crete P, Bartel DP. 2004. The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes and Development* **18,** 1187–1197.

Wang J, Zhou H, Chen Y, Luo Q, Qu L. 2004. Identification of 20 microRNAs from *Oryza sativa*. *Nucleic Acids Research* **32,** 1688–1695.

Zhang BH, Pan XP, Wang QL, Cobb GP, Anderson TA. 2005. Identification and characterization of new plant microRNAs using EST analysis. *Cell Research* **15,** 336–360.

Zhang Y. 2005. miRU: an automated plant miRNA target prediction server. *Nucleic Acids Research* **33** (web server issue), W701–W704.