UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI MEDICINA ANIMALE, PRODUZIONI E SALUTE

Dipartimento di Agronomia Animali Alimenti Risorse Naturali e Ambiente

Tesi di laurea in Biotecnologie per l'Alimentazione

COMPUTATIONAL IDENTIFICATION OF CONSERVED MICRORNAS AND THEIR PUTATIVE TARGETS IN THE *HYPERICUM PERFORATUM* L. FLOWER TRANSCRIPTOME

Relatore
Prof. Gianni Barcaccia

Correlatore
Dott. Galla Giulio

Laureando:
Mirko Volpato
Matricola n. 1034525

ANNO ACCADEMICO 2012-2013

INDEX

1. ABSTRACT

MicroRNAs (miRNAs) have recently emerged as important regulators of gene expression in plants. Many miRNA families and their targets have been extensively studied in model species and major crops. We have characterized mature miRNAs along with their precursors and potential targets in *Hypericum* to generate a comprehensive list of conserved miRNA families and to investigate the regulatory role of selected miRNAs in biological processes that occur in the flower. St. John's wort (*H. perforatum* L., $2n=4x=32$), a medicinal plant that produces pharmaceutically important metabolites with therapeutic activities, was chosen because it is regarded as an attractive model system for the study of apomixis. A computational *in silico* prediction of structure, in combination with an *in vitro* validation, allowed us to identify 7 pre-miRNAs, including miR156, miR166, miR390, miR394, miR396 and miR414. We demonstrated that *H. perforatum* flowers share highly conserved miRNAs and that these miRNAs potentially target dozens of genes with a wide range of molecular functions, including metabolism, response to stress, flower development and plant reproduction. Our analysis paves the way towards identifying flower-specific miRNAs that may differentiate the sexual and apomictic reproductive pathways.

Keywords: miRNA, *Hypericum perforatum*, reproductive organs, apomixis.

RIASSUNTO

I microRNA (miRNA) sono stati recentemente indicati come importanti regolatori dell'espressione genica nelle piante. Molte famiglie di miRNA e i rispettivi geni target sono stati ampiamente studiati sia in specie modello sia in specie di interesse agro-alimentare. Nella presente ricerca i miRNA maturi e i loro potenziali target sono stati studiati in iperico (*Hypericum perforatum* L., 2*n*=4*x*=32) al fine di ottenere una lista completa delle famiglie di miRNA conservati in questa specie e di studiare il ruolo regolatore dei miRNA selezionati nei processi biologici che avvengono nel fiore. L'iperico è una pianta medicinale che produce metaboliti farmacologicamente importanti con attività terapeutiche. Questa specie è stata scelta perché è considerata un sistema modello interessante per lo studio dell'apomissia. Un iniziale lavoro di bioinformatica è stato condotto al fine di predire *in silico* la struttura dei microRNA e di validarne la presenza *in vitro*. Tale approccio sperimentale ha permesso di individuare 7 pre-miRNA, tra cui miR156, miR166, miR390, miR394, miR396 e miR414. Parallelamente, questo approccio ha dimostrato che miRNA altamente conservati sono espressi nei fiori di *H. perforatum* e che i target potenziali di questi miRNA sono decine di geni annotati secondo una vasta gamma di funzioni molecolari, tra cui lo sviluppo del fiore e il controllo della riproduzione. I dati prodotti costituiscono la base conoscitiva necessaria per l'individuazione di miRNA fiore-specifici che possono avere un ruolo chiave nella differenziazione dei processi riproduttivi sessuale e apomittico.

2 INTRODUCTION

2.1 Mechanisms of apomixis

Because apomictic reproduction entails the development of an embryo from a cell with a somatic chromosome number, there are several ways to produce embryos of apomictic origin. The simplest pathway avoids the production of an embryo sac, and the maternal embryo originates from one or more somatic cells of the ovule (Barcaccia and Albertini, 2013). Among the agriculturally important species, adventitious embryony (*i.e.*, sporophytic apomixis) has been noted in mango (*Mangifera indica*), several *Citrus* species, and orchids. The most comprehensive treatise on adventitious embryony has been published by Naumova (1992).

When the maternal embryo originates from a diploid egg cell differentiated in an unreduced embryo sac, the apomictic pathway is referred to as gametophytic apomixis (Nogler, 1984).

In gametophytic apomixis, the unreduced embryo sac may arise from a somatic nucellar cell that acquires the developmental program of a functional megaspore, a mechanism referred to as apospory. Alternatively, if the embryo sac forms from a megaspore mother cell with suppressed or modified meiosis, the pathway is referred to as diplospory. It is worth emphasizing that apomictic plants may or may not change meiosis itself, but in any case they do activate the gametic cell fate either in a somatic cell (apospory) or in an unreduced megaspore (diplospory) as surrogate for meiotic products (Barcaccia and Albertini, 2013; Hörandl and Hadacek, 2013). Once 2*n* female gametophytes and gametes are formed (apomeiosis), they subsequently undergo embryogenesis autonomously without fertilization by a male gamete (somatic parthenogenesis). Endosperm formation may be fertilization-independent (autonomous endosperm) or may require fertilization (pseudogamous endosperm). Among others, apospory has been reported in *Beta*, *Brachiaria*, *Cenchrus*, *Chloris*, *Compositae*, *Eriochloa*,

Heteropogon, *Hieracium*, *Hyparrhenia*, *Hypericum*, *Panicum*, *Paspalum*, *Pennisetum*, *Poaceae*, *Ranunculus*, *Sorghum*, *Themeda*, and *Urochloa,* whereas diplospory has been noted in *Agropyrum*, *Allium*, *Antennaria*, *Boechera* (formerly *Arabis*), *Datura*, *Eragrostis*, *Erigeron*, *Eupatorium*, *Ixeris*, *Parthenium*, *Paspalum*, *Poa*, *Taraxacum*, and *Tripsacum* (Barcaccia and Albertini, 2013).

2.2 Apomixis as a desiderable trait in modern agricolture

One of the greatest success stories in modern agriculture has been the tremendous yield increase achieved by coupling high-yield varieties with high-input agronomic systems, creating the so-called Green Revolution. Approximately one-third of the world's seed supply comes from the commercial seed market, another one-third is provided by publicly funded institutions and the seed saved by farmers accounts for the remainder (Barcaccia and Albertini, 2013).

Over the centuries, crop plants have followed the general pattern of introduction, selection, and hybridization. Crop introduction has been crucial for agriculture because many of the world's crops are produced outside their region of domestication. Once introgressed, selection and breeding strategies have led to the development of new cultivars with improved yield and adaptation. Plant breeders are working to extend the Green Revolution by intensifying selection, developing more hybrid varieties in more crops, and increasing the range of plant functions through mutation and transgenic breeding. Hence, plant breeding will continue to play a crucial role in crop improvement because the needs are many, the techniques are expanding, the new genetic combinations are limitless, and the successes of the past illuminate the potential of the future (Barcaccia and Albertini, 2013). In outcrossing species, alleles disseminate in the offspring; thus, the optimal genotype is lost together with the desired trait. Exact copies of a superior genotype can be made via vegetative propagation; however, this technique is usually not applicable to annual crops such as maize, rice and wheat. The fixation of a given genotype occurs naturally in species that exhibit an asexual type of seed production termed apomixis. This trait by itself is highly valuable for agriculture;

however, despite many efforts, it has not been possible to introduce apomixis into modern domesticated crop species (Barcaccia and Albertini, 2013).

As a reproductive strategy for cloning plants via seeds, apomixis is a highly desirable trait in modern agriculture. In fact, apomixis results in offspring that are exact genetic replices of the female parent because embryos are derived from the parthenogenic development of apomeiotic egg cells (for reviews on apomixis, see Bicknell and Koltunow 2004; Ozias-Akins 2006; Albertini et al. 2010; Pupilli and Barcaccia 2012; and Koltunow et al. 2013). From an evolutionary point of view, apomixis may be regarded as a consequence of sexual failure rather than as a recipe for clonal success (Silvertown 2008).

Introgression of apomixis from wild relatives into crop species and transformation of sexual genotypes into apomictically reproducing genotypes are long-held goals of plant breeding. Breeders believe that the introduction of apomixis into agronomically important crops will have revolutionary implications for agriculture. The potential benefits of harnessing apomixis are many and vary from full exploitation of heterosis by reseeding the best hybrids to clonal propagation of the superior genotypes in seed-propagated outcrossing crops. The impact of apomictic crops in agriculture would be massive in both developed and developing countries. Unfortunately, barring a few exceptions in some forage grasses and fruit trees, apomixis is not a common feature among crop species (Barcaccia and Albertini, 2013).

The fixation of hybrid vigor through apomixis is a desirable objective for breeders and farmers alike and is expected to have a revolutionary impact on food and agriculture production. The stabilization of heterozygous genotypes via apomixis would make breeding programs faster and cheaper. The impact of apomictic crops in agriculture would be comparable to, or even greater than, the impact of the Green Revolution, especially in Third World countries (Vielle-Calzada 1996; Pupilli and Barcaccia 2012). In fact, it has been estimated that the use of apomixis technology in the production of hybrid rice alone could provide benefits exceeding 1,800 million Euros per year (Spillane et al. 2004; Albertini et al. 2010).

Apomixis technology could also provide benefits for clonally propagated crops. Clonal crop yields are limited by pathogens (mainly viral and endophytic), which accumulate over successive rounds of vegetative propagation and seriously limit the yield and exchange of germplasms between countries. The use of apomixis technology in these crops would provide the additional option and benefit of propagation via clonal seeds, thus generating disease-free material that can be more easily stored and transported (Barcaccia and Albertini, 2013).The use of apomictic seed as an alternative to vegetative propagules would provide similar benefits (*e.g.*, lower costs and higher yields) over the current use of true seed of such crops. For example, apomixis technology could make true potato seeds a more attractive option for potato breeders and cultivators and would return benefits to growers of as much as 2.3 billion Euros per year (Spillane et al. 2004).

The development of apomixis technology in agriculture will require a deeper knowledge of the mechanisms regulating reproductive development in plants. Our molecular understanding of apomixis would be greatly increased if genes that are specifically or differentially expressed during the formation of the embryo and embryo sac could be identified. Over the last two decades, many scientists have speculated about the isolation of gene/s controlling key steps of the apomictic pathway, and many papers have postulated the production of engineered plants exhibiting apomictic-like phenotypes (Barcaccia and Albertini, 2013). In fact, none of the major crop plants have been bred for apomixis, and only some features of apomixis have been genetically engineered in model species. Consequently, even in the era of genomics, achieving an understanding of the genetic control and molecular regulation of apomixis appears much more complicated than expected. Large amounts of cytological and ecological information, along with genetic and molecular data, have been collected mainly from model species (*i.e.*, *Boechera holboellii*, *Hieracium* spp., *Hypericum perforatum*, *Paspalum* spp., *Poa pratensis*, *Ranunculus* spp., and *Taraxacum officinale*) and have often been tested in *Arabidopsis thaliana* (*Arabidopsis*) to elucidate the mechanisms of apomeiosis, parthenogenesis and apomixis (Barcaccia and Albertini, 2013). Several genes

involved in the formation of unreduced embryo sacs and egg cells, in addition to genes responsible for the autonomous development of the embryo and endosperm, have been cloned and characterized; however, none of these genes are capable of miming the apomictic pathway as a whole in crop plants (Barcaccia and Albertini, 2013). Hence, after two decades of substantial studies conducted in several laboratories and model plants, the asexual reproductive strategy termed "gametophytic apomixis" by Nogler (1984) still appears to be an unsolved puzzle. As a result, seed companies have lost interest in this research, and it has been difficult to acquire funds for conducting research on apomixis (Barcaccia and Albertini, 2013).

Currently, novel views and original concepts are emerging from the fog, including a link between apomixis and gene-specific silencing mechanisms (likely based on chromatin remodeling factors or trans-acting and heterochromatic interfering RNAs involved in both transcriptional and post-transcriptional gene regulation) and the parallel between the Y chromosome and apomixis-bearing chromosomes from the most primitive to the most advanced in evolutionary terms (comparative genomic analyses revealed common features such as few recombination events, accumulation of transposable elements and degeneration of genes). More recently, merging lines of evidence regarding the role of auxin in cell fate specification of the embryo sac and egg cell development have been reported in *Arabidopsis* (Barcaccia and Albertini, 2013).

2.3 *Hypericum perforatum* as a model species for apomixis researches

Although miRNAs have been studied in plants for years, no extensive study has yet been performed on *Hypericum* species. *Hypericum perforatum* L. is a medicinal plant belonging to the family *Hypericaceae*, whose members are widely distributed in a variety of habitats ranging from rocky and dry environments to moist grasslands and swamps (Nurk *et al*., 2012). Wild populations of *H. perforatum* are generally composed of tetraploid individuals (2*n*=2*x*=32), but

diploids and hexaploids have also been documented (Matzk *et al*., 2001; Robson, 2002). According to the polyploid nature of the plant, several studies have underlined a number of traits reflective of hybridity, including meiotic abnormalities and elevated pollen grain sterility (reviewed by Barcaccia *et al*., 2007).

Although *H. perforatum* is considered a weed in many countries (Buckley *et al*., 2003; Zouhar, 2004), during the last decade it was chosen as model species for investigating apomixis, a reproductive strategy rather common in polyploid plants that allows the progenies to inherit the whole maternal genome through seeds (Matzk *et al*., 2003; Barcaccia *et al*., 2006; Schallau *et al*., 2010; Galla *et al*., 2011). *H. perforatum* reproduces via aposporic apomixis, and hence a somatic cell of the ovule gives rise apomeiotically to a functional embryo sac in which the unreduced egg cell can develop autonomously via parthenogenesis into a maternally-derived embryo (Barcaccia *et al*., 2007; Koch *et al*., 2013).

2.4 MicroRNAs in plants

Plants regulate gene expression using many mechanisms to ensure normal development and reproduction, as well as to produce appropriate responses to biotic agents and environmental signals. One of these regulatory mechanisms involves small RNA (sRNA) molecules that act by silencing gene expression. Although miRNAs constitute only a small fraction of the sRNA population (Lu *et al*., 2005; Jones-Rhoades *et al*., 2006), the post-transcriptional regulation of genes guided by miRNAs is one of the most conserved and well-characterized gene regulatory mechanisms (Lewis *et al*., 2005; Jones-Rhoades *et al*., 2006; Voinnet, 2009). MicroRNAs are 21-24 nt non-coding RNA sequences derived from single-stranded RNA precursors that possess the ability to form intra-molecular complementary hairpin structures. This is a key feature that distinguishes miRNAs from other small RNAs, such as small interfering RNAs (siRNAs), which originate from double-stranded RNAs (dsRNAs) derived from the inter-molecular hybridization of two complementary RNA molecules.

Extending the definition proposed by Axtell (2013), miRNAs are single, short RNA molecules that originate from hairpin RNAs (hpRNAs), whose function is the down-regulation of mRNA species by direct annealing. It is generally accepted that miRNAs target RNA molecules that differ from those from which they are transcribed (Chen, 2005). In addition, increasing evidence shows that miRNAs negatively regulate their target genes, which function in a wide range of biological processes, including organogenesis, sexual reproduction and stress responses (Wu *et al*., 2006; Sunkar *et al*., 2007; Bowman *et al*., 2008) as well as molecular functions, such as the regulation of translational turnover and signal transduction (Brodersen *et al*., 2008; Axtell, 2013).

From the first miRNA identified in *C. elegans* twenty years ago (Lee *et al*., 1993), studies have focused their attention on the mechanisms of miRNA biogenesis and on their functional roles in both plants and animals (Pfeffer *et al*., 2004; Siomi and Siomi, 2010). The main components regulating the transcription and maturation processes of miRNAs are now known (Mallory and Vaucheret, 2006; Song *et al*., 2007; Faller and Guo 2008). The miRNAs are derived from a precursor sequence of approximately 70 or more nucleotides that commonly forms a 21 bp duplex with a conserved stem and variable loops, which is then excised to produce the mature miRNA. The 21 bp sequence matches one or more target sequences for cleavage (Jones-Rhoades *et al*., 2006; Bowman *et al*., 2008). All miRNAs were initially identified by direct cloning using bioinformatic prediction, or Sanger sequencing of relatively small-sized cDNA libraries (Llave *et al*., 2002; Sunkar *et al*., 2004). The application of deep sequencing by NGS technology has greatly facilitated the pace of miRNA identification in plants.

The conformation of RNA in the stem-loop structures that contain the miRNA:miRNA* duplex, and the high complementarity existing between a miRNA sequence and its target sequence, are biological aspects that have been extensively used as tools for the computational investigation of miRNAs and target genes. For example, comparative analyses revealed that some miRNA families are highly conserved among unrelated plant species while others have diverged and evolved,

generating abundant family- and species-specific miRNAs (Axtell *et al*., 2005; Jones-Rhoades *et al*., 2006; Cuperus *et al*., 2011). A recent analysis performed by Nozawa *et al*. (2012) revealed that nearly half of the miRNA genes in *Arabidopsis* have homologous miRNA genes in rice and vice versa. Surprisingly, not only were the miRNA sequences highly conserved, but also the miRNA/target relationships over long periods of plant evolution (Jones-Rhoades, 2012). These findings suggest that dynamic and evolving miRNA molecules may serve as a driving force for the selection of improved traits in plants (Zhu *et al*., 2012).

The number of miRNA genes in plant genomes is variable in different species and ranges from 72 found in papaya to as many as 378 in rice (Zhu *et al*., 2008; Nozawa *et al*., 2011). *A. thaliana*, which has the most highly annotated plant genome, contains approximately 160 miRNA genes grouped into 80 families, possibly reflecting the number of ancestors maintained by selection in this species (Rajagopalan *et al*., 2006). In addition to model species, miRNAs have been identified in many other crop plants, including poplar (Barakat *et al*., 2007), grape (Pantaleo *et al*., 2010), apple (Yu *et al*., 2011), peach (Zhu *et al*., 2012), tomato (Moxon *et al*., 2008), maize (Zhang *et al*., 2009), peanut (Zhao *et al*., 2010) and soybean (Song *et al*., 2011).

Currently, the most important worldwide Plant miRNA Database (http://bioinformatics.cau.edu.cn/PMRD/) contains nearly 11,000 miRNA sequences that were deduced either experimentally or computationally from 127 species. The distribution of miRNAs among species varies, with the most abundant contributions from *Arabidopsis thaliana* (1,530 sequences), *Oryza sativa* (2,773) and *Populus trichocarpa* (2,780), a reflection of the long-term role of model species as well as the availability of data derived from NGS approaches.

The crucial role of miRNAs in plant development is exemplified by the dramatic and pleiotropic developmental defects of mutants lacking single proteins participating in their biogenesis (reviewed by Mallory and Vaucheret, 2006). At cellular levels, miRNAs have been associated with cell proliferation and programmed cell death (Lynam-Lennon *et al*., 2009), of which deregulation is an

important trait in cancer progression in animals (Evan and Vousden, 2001). In plants, post-transcriptional activities of miRNAs are involved in the regulation of fundamentally important biological processes, such as plant development (*e.g.*, miR156/157, miR390/391), stress response (*e.g.*, miR395, miR398/399, miR408) or signaling pathways (*e.g.*, miR159, miR164, miR168) (Xie *et al.*, 2010). Among the targets whose functions are known and have been validated in plants, the proportion of proteins having transcription factor activity is relevant (Jones-Rhoades *et. al.*, 2006). Similarly, at least seven miRNA families are known to target gene products either in the detection of auxins (miR393 vs. TIR1) or in response to this hormone, such as the ARF proteins (miR160, miR167, miR172). As an example, if miR167 does not interact with its targets ARF6 and ARF8, which regulate gynoecium and stamen development in immature flowers, the result is the ectopic expression of these genes, which eventually affects ovule development and anther indehiscence (Wu *et al.*, 2006).

2.5 Objective of the reserch

The main goal of this work was to identify and completely catalogue conserved plant miRNAs in this species, including their precursors and targets, to shed light on the potential role of miRNAs in flower development and in the formation of reproductive tissues/organs and gametes in *H. perforatum.* Our analyses intend to pave the way towards discovering flower-specific miRNAs and predicting target genes that may differentiate and/or regulate the sexual and apomictic reproductive pathways.

3 MATERIALS AND METHODS

3.1 *H. perforatum* sequence datasets

The *H. perforatum* sequences used in this study for discovering conserved microRNAs and their targets have been recently produced, assembled and annotated by Galla *et al*. (2012). In brief, approximately 1.5 million reads were generated by 454 pyrosequencing, and the raw sequences were then processed using bioinformatics (Table 1). The assembled and annotated contigs/isotigs were stored in the *Hypericum* database (Hypdb) available for downloading at https://147.162.139.232/account/login/. Our flower EST database of the *H. perforatum* genotypes can be consulted and queried by researchers upon an expression of interest and after the formalization of a Material Transfer Agreement (MTA) to get authentication by username and password.

The cDNA libraries were obtained from flowers collected at developmental stages 1-12 (according to Smyth *et al*., 1990; and Galla *et al*., 2011) from two sexual and apomictic plants (HP13EU; HP36EU; HP38EU and HP1093US). These cDNA libraries were sequenced twice. Organ specific libraries were produced from plant accession HP4/13. The cDNAs were produced from the following flower parts: young buds (whole buds with a length < 3.0 mm, flower developmental stages 1 to 10), carpels (*i.e*., pistils), stamens (i.e., anthers) and sepals/petals with a bud length > 3 mm equal to flower developmental stages 11 to 14 (Galla *et al*., 2011; Galla *et al*., 2012).

Table 1. Biological materials used in this research. For each plant accession, the site of origin, ploidy level and degree of apomixis are reported.

Accession	Description	Origin	Ploidy	Apomixis	Reproductive behavior
13EU	Hybrid population	IPK-Gatersleben (D)	2n=4x	< 4%	Sexual
36EU	Hybrid population	IPK-Gatersleben (D)	2n=4x	< 4%	Sexual
Hp4/13	Wild population	UniPD-Cellarda BL (I)	2n=4x	24%	Facultative apomictic
39EU	Hybrid population	IPK-Gatersleben (D)	2n=4x	> 96%	Obligate apomictic
1973US	Wild population	UM-Tecumseh MI (USA)	2n=4x	> 95%	Obligate apomictic

3.2 Prediction of *H. perforatum* flower pre-miRNAs

The strategy that was followed for the prediction of flower pre-miRNAs is that proposed by Amiteye *et al*. (2011; 2013) with some modifications as reported in the experimental pipeline (Figure 1). The reference dataset of miRNA sequences was obtained from the plant microRNAs database PMRD, consisting of more than 10,000 sequences from approximately 130 plant species (http://bioinformatics.cau.edu.cn/PMRD/, Zhang *et al*., 2010). Computational investigations were based mainly on mature miRNA sequences.

Briefly, the miRNA database was queried by using a BLASTN strategy (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast±/) that was adapted for short sequences. The program used for all local BLAST searches was BLAST−2.2.26+. Algorithm parameters that produced the highest number of results were: −evalue 1, −word size 7, −gapopen 0, −gapextend 2, −penalty −1, −reward 1 and −max target seqs 20. The ESTs producing an alignment of 20-24 nucleotides in length, with three or less mismatches and no gaps were selected for further annotation steps. Sequence redundancy was removed with a self-BLASTN. Both tRNA and rRNA sequences were filtered with successive BLAST searches over the ribosomal RNAs database from Rfam (http://www.sanger.ac.uk/Software/Rfam/ that was previously purified of all miRNAs sequences) and the *Arabidopsis* transfer RNAs database (http://lowelab.ucsc.edu/GtRNAdb/Athal). Finally, a BLASTX search over the non-redundant protein database (ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/) was used to remove all of the most probable mRNA sequences.

Secondary structures of the putative pre-miRNA were generated using the Zuker folding algorithm implemented in MFOLD 3.2 (http://mfold.rna.albany.edu/?q=mfold; Zucker, 2003). The default parameters were used to predict the secondary structures of the selected sequences. All minimum folding free energies (MFEs) were expressed as negative kcal/mol. Adjusted MFE (AMFE) represented the MFE of 100 nucleotides and was calculated using the following formula: MFE/(length of RNA sequence) × 100. The minimal

folding free energy index (MFEI) was calculated using the following equation: MFEI = AMFE/(G+C)% from Zhang *et al*. (2008).

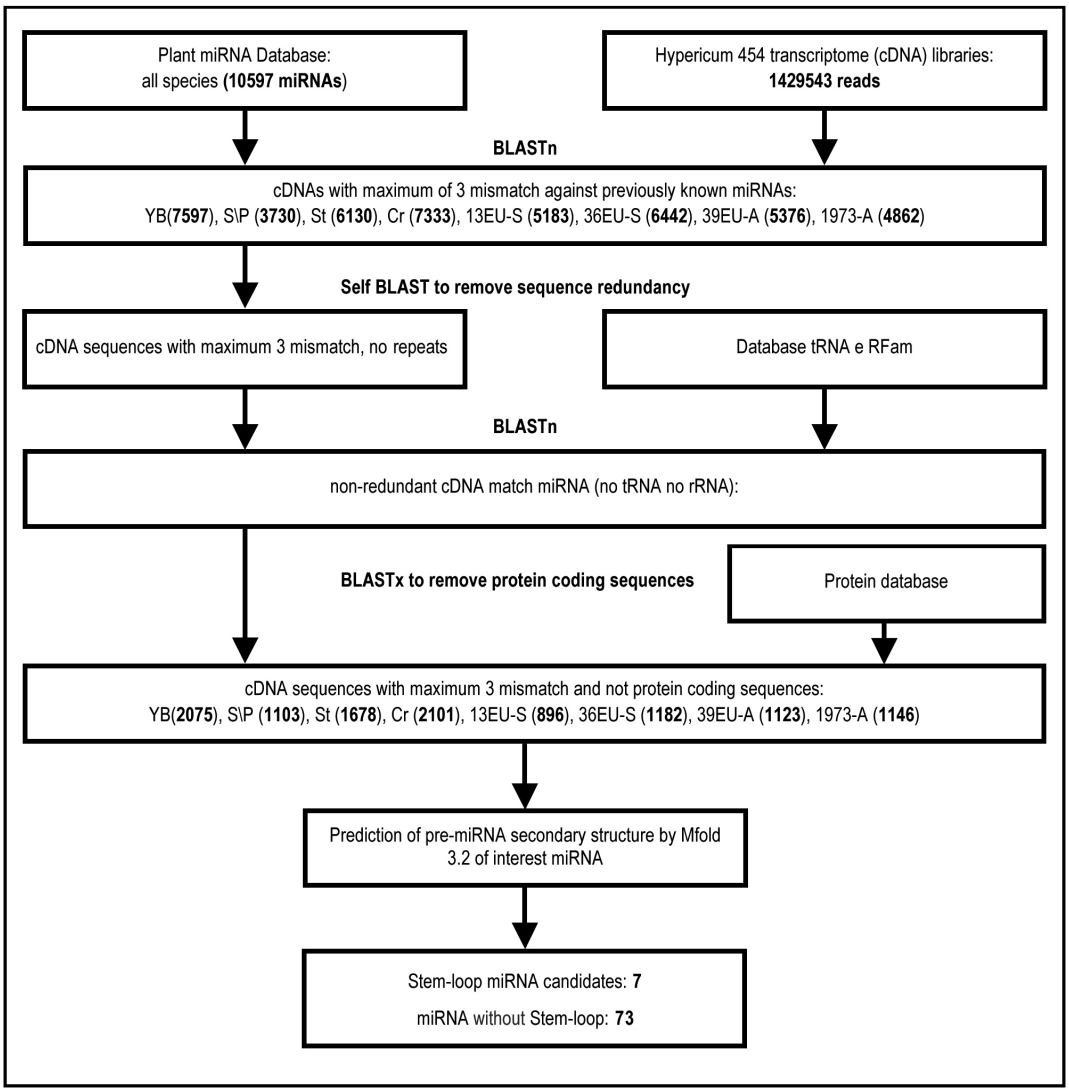

Figure 1. Pipeline for the bioinformatics search of conserved miRNAs in *Hypericum* species. YB, S/P, St, Cr, 13EU-S, 36EU-S, 39EU-A, 1973US-A indicate the names of the libraries (for each library, the number of sequences is indicated in brackets)

Stem-loops were selected as a candidate miRNA precursor when the RNA sequence could fold into an appropriate stem-loop hairpin secondary structure, and all but one of the following criteria were satisfied: i) the predicted mature miRNAs had no more than 3 nucleotide substitutions and no gaps compared with

the mature miRNA query; ii) there were no more than 6 mismatches between the predicted mature miRNA sequence and its opposite miRNA* sequence in the secondary structure; iii) no loops or breaks in the miRNA or miRNA* sequences; iv) the predicted secondary structure had a high MFEI and a negative MFE and v) the mature miRNA could be localized in one arm of the hairpin structure.

Multiple alignments of pre-miRNA sequences were generated using the algorithm MUSCLE implemented by the Geneious software v.3.6.1 (www.geneious.com). Pre-miRNA sequences were submitted to GenBank (http://www.ncbi.nlm.nih.gov/genbank/) with the accession numbers: KC884257-63.

3.3 Prediction of miRNA targets

The miRNA target sequences were predicted based on their complementarity with mature miRNA sequences. Protein-coding sequences were selected with BLASTX searches over the non-redundant protein database (ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/) using default algorithm parameters, with an E-value cutoff set to 1.0E−6. In this study, the following criteria were used for identifying potential miRNA targets: i) an alignment length longer that 18 nt; ii) no more than three mismatches between the mature miRNA and its potential target site; and iii) no gaps in complementary sites. Target miRNA sequences were stored in in the *Hypericum* database (Hypdb) and are available to researchers upon an expression of interest and after the formalization of a Material Transfer Agreement (MTA).

To annotate all putative targets, a BLASTX-based approach was used to compare the *H. perforatum* sequences to the nr database downloaded from NCBI (http://www.ncbi.nlm.nih.gov/). The GI identifiers of the best BLASTX, for all hits having an E-value ≤ 1E−09 and a degree of similarity ≥ 70%, were mapped in the UniprotKB protein database (http://www.uniprot.org/). Finally, the UniprotKB accessions were used to extract Gene Ontology terms for further functional annotations. The Blast2GO software v1.3.3 (http://www.blast2go.org/, Conesa *et*

al., 2005) was used to perform basic statistics on the GO annotations as reported by Botton *et al*. (2008) and Galla *et al*. (2009).

3.4 Prediction of mature miRNA from target sequences

BLASTN searches were used to query a PMRD database and predict putative miRNA sequences from the targets. Algorithm parameters were: –evalue 1, –word size 7, –gapopen 0, –gapextend 2, –penalty –1, -reward 1, –max target seqs 1 and –outfmt 5. The target regions with the highest complementarity to the miRNA were selected from the BLASTN alignments. The minimum requirements for an alignment to be considered were 18 nucleotides in length, with 3 or fewer mismatches and no gaps. Short miRNA target regions (SMTR) sharing high sequence complementarity with heterologous miRNAs were aligned with the algorithm MUSCLE implemented by Geneious software v.3.6.1 (www.geneious.com). Manual editing of the alignment was employed to minimize the effect of SNPs on short sequence alignments. The consensus sequence of each SMTR was extracted from each sequence alignment. The prediction of miRNAs sequences whose pre-miRNAs had been previously identified was used to test this method. The software DNASP v. 4 (Rozas *et al*., 2003) was used to generate haplotype sequences (likely attributable to miRNA family members).

3.5 Validation of *H. perforatum* flower pre-miRNAs

Total RNA was extracted from whole flowers and flower parts using the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich) following the manufacturer's protocol. Both anthers and pistils were collected separately at flower developmental stages 11-12a (later referred as An1 and Pi1) and 12b-14 (later referred as An2 and Pi2). Genomic DNA was removed by treatment with DNase I (Sigma-Aldrich) by following the manufacturer's protocol. cDNA synthesis was conducted using the SuperScript® III cDNA Synthesis Kit (Life Technologies) by

following manufacturer's protocol. A reaction mix without SuperScript® III was used as negative RT control.

For each pre-miRNA, the PCR and RT-PCR reactions were performed on a GeneAmp PCR System 9700 (Applied Biosystems) in a 20 µl volume containing 1× PCR buffer (100 mM Tris-HCl pH 9.0, 15 mM $MgCl_2$ and 500 mM KCl), 0.2 mM dNTPs, 0.2 µM of each primer and 0.5 U of *Taq* DNA polymerase (BIOline). All forward and reverse primers designed for each of the selected pre-miRNA sequences, along with their nucleotide sequences and corresponding references, are listed in Table S1 (see Supplementary materials). The *hgPP2*, encoding the *Hypericum* protein phosphatase 2A subunit A3, was used as housekeeping gene. The temperature conditions for the validations were as follows: 5 min at 94°C followed by 40 cycles of 30 sec at 94°C, 30 sec at 53°C to 55°C (depending on the specific pair of primers, see Table S1) and 30 sec at 72°C, followed by 7 min at 72°C followed by a 10°C hold. Genomic DNA and negative controls were used as reference standards. The PCR-derived fragments were resolved on 2% agarose/TAE gels and visualized under UV light using Sybr Safe DNA stain (Life Technologies). All amplification products were subjected to EXOI/FAP (Fermentas) treatment and then directly sequenced using an ABI3100 automated sequencer (Applied Biosystems).

Quantitative Real-Time PCR experiments were performed according to Galla *et al*. (2009), using a StepOne instrument (Applied Biosystems) equipped with a 96-well plate system. Samples were analyzed in three technical replicates. The amplifications were normalized using the $\Delta\Delta^{Ct}$ method with the *Hypericum hgPP2* gene as the housekeeping gene.

4 RESULTS

4.1 Identification of conserved miRNA families in *Hypericum perforatum* using bioinformatics analysis

The prediction of miRNAs was performed using the 454 raw reads to exclude the possibility that artifacts introduced by the assembly could have decreased the sensitivity and efficiency of the bioinformatic search. The computational surveys were performed as depicted in the experimental pipeline in Figure 1, using 10,597 mature microRNA sequences belonging to 127 plant species, and enabled the identification of 46,653 transcripts with sequence homology to mature miRNAs. Additional bioinformatic analyses were necessary to remove redundancy and select all sequences with high similarity scores to rRNAs and tRNAs, and with putative mRNAs. This procedure allowed us to identify 11,304 potential pre-miRNA sequences belonging to 1,542 families (Table 2). Therefore, approximately 6.4% of the ESTs analyzed here contained a region with a high sequence similarity to a known miRNA.

In this study, we focused our investigation on 36 families of miRNAs that were selected by nature of them having highest conservation (Zhang *et al*., 2006a) in addition to the biological processes in which they were believed to be involved (Sun Guiling, 2012). The computational prediction of the structures allowed us to identify seven pre-miRNAs belonging to six different families, including miR156, miR166, miR390, miR394, miR396 and miR414 (Table 3). The predicted hairpin structures of the selected miRNAs identified in *Hypericum* spp. are reported in Figure 2.

Table 2. Number of transcripts with sequence homology to a mature miRNA and descriptive statistics of libraries composition. For each plant accession, the table reports on plant organ considered for sequencing, and plant phenotypes. Number of reads procedures is also indicated.

Accession	Description	Phenotype	Reads	Putative miRNAs
Hp4/13	Pre-meiotic buds	Facultative apomictic	229311	2075
	Sepals and Petals		111665	1103
	Stamens		186567	1678
	Carpels		219166	2101
13EU	Whole flower	Sexual	162589	896
36EU		Sexual	201457	1182
39EU		Apomictic	161067	1123
1973US		Apomictic	150755	1146
Overall			177822	11,304

4.2 Characterization of the selected *Hypericum* miRNAs

The formation of secondary structures is not a unique feature of miRNA molecules (*e.g*. they are common in tRNAs and rRNAs), and thus the evaluation of the precursors was established on the basis of fundamental parameters required for the annotation of new miRNAs, including the minimum fold energy (MFE), the adjusted minimum energy fold (AMFE) and the minimum fold energy index (MFEI) (Zhang *et al*., 2006b). Zhang *et al*. (2006b) reported that the majority of identified miRNA precursors have a MFEI higher than 0.85, which is much higher than is usually scored by tRNAs (0.64), rRNAs (0.59) or mRNAs (0.62-0.64). Of the eight putative pre-miRNAs identified in *Hypericum*, the average MFEI value was 0.81±0.17, ranging from a minimum of 0.57 and a maximum of 1.03 (Table 3).

The pre-miRNA types miR156b, miR166 and miR414 were selected during the preliminary screen although their MFEI estimates were lower than 0.85 (Table 3). These precursors were still considered because they fulfilled at least five of the six established conditions (Zhang *et al*., 2008; Amiteye *et al*., 2011; 2013), including the number of nucleotide substitutions between the predicted miRNA and the mature miRNA used as a query (<4), the number of mismatches between the putative mature miRNA and its complement in the secondary structure (<5), and the absence of specific structures such as hairpins or loops within the mature

miRNA sequence (Figure 2). The content of the A + U nucleotides in the precursors varied from 46 to nearly 68% (Table 3), which was similar to the contents observed in other species (Amiteye *et al*., 2011, Pani *et al*., 2011). The majority of the precursors identified in the *Hypericum* samples contained more A + U than G + C nucleotides, as was expected from the available data (Zhang *et al*., 2006a; 2006b).

In these precursors, a different degree of identity was found between the sequence of the mature miRNA used as a query and the corresponding region of the *Hypericum* transcript. In addition, the mature miRNAs showed high similarity with the sequences of the putative orthologs (Table 3). The high extent of conservation was supported by three miRNA sequences that were identical to the reference sequence (*i.e*., miR156a, miR166 and miR396), whereas for the remaining miRNAs the number of nucleotide substitutions ranged between 1 and 3. The sequences most similar to the miRNAs of *Hypericum* were identified both in closely related species, such as *Populus* spp. and *Glycine max*, and in taxonomically distant species, such as *Arabidopsis* spp., *Vitis vinifera* and *Zea mays* (Table 3).

The length of the *Hypericum* miRNAs varied from 20 to 22 nucleotides, whereas the length of their precursors ranged between 143 and 237 nucleotides, with an average length of approximately 182 nucleotides (Table 3). The distribution of the precursor length and the mature miRNAs is consistent with the information available from other plant species (Zhang *et al*., 2006a; 2006b; 2008). As shown in Figure 2, the identified miRNAs were positioned predominantly at the 5' end of the stem-loop structures, with the exception of miR166 and miR414, for which the mature miRNAs were localized at the 3' end of the molecules (Figure 2).

Table 3. *Hypericum perforatum* miRNAs identified by comparative genomics and secondary structures.

miRNA family	Mature miRNAs sequence	Plant sp., NSs	ML	NN	%A	%C	%G	%T	%(G+C)
Hyp-miR156a	UUGACAGAAGAUAGAGAGCAC	Vvi	21	164	29,3	26,2	18,3	26,2	44,5
Hyp-miR156b	UUGACAGAAGAGAGAGAGCAA	Pts, C/A	21	216	33,3	19,4	18,1	29,2	37,5
Hyp-miR166	UCGGACCAGGCUUCAUUCCCC	Vvi	21	237	25,3	20,7	23,2	30,8	43,9
Hyp-miR390	AAGCUCAGGAGGGAGAGCACC	Gma, U/G	21	175	30,3	21,1	20,6	28,0	41,7
Hyp-miR394	UUGGCAUUCUGUCCACCUCCCU	Vvi, A/C	22	153	22,2	19,0	25,5	33,3	44,4
Hyp-miR396	UUCCACAGCUUUCUUGAACU	Ptc	20	143	28,0	14,0	18,2	39,9	32,2
Hyp-miR414	UCCUCCUCAUCCUCCUCGUC	Osa, A/C (3)	20	187	21,4	32,6	21,4	24,6	54,0

miRNA family	Mature miRNAs sequence	Plant sp., NSs	MFE	AMFE	MFEI	EST ID	Arm
Hyp-miR156a	UUGACAGAAGAUAGAGAGCAC	Vvi	64,4	39,27	0,88	KC884257	5'
Hyp-miR156b	UUGACAGAAGAGAGAGAGCAA	Pts, C/A	46,0	21,30	0,57	KC884258	5'
Hyp-miR166	UCGGACCAGGCUUCAUUCCCC	Vvi	73,3	30,93	0,70	KC884259	3'
Hyp-miR390	AAGCUCAGGAGGGAGAGCACC	Gma, U/G	60,5	34,57	0,83	KC884260	5'
Hyp-miR394	UUGGCAUUCUGUCCACCUCCCU	Vvi, A/C	66,5	43,46	0,98	KC884261	5'
Hyp-miR396	UUCCACAGCUUUCUUGAACU	Ptc	47,2	33,01	1,03	KC884262	5'
Hyp-miR414	UCCUCCUCAUCCUCCUCGUC	Osa, A/C (3)	68,0	36,36	0,67	KC884263	3'

Plant sp, NSs, Nucleotide substitutions between known plant query miRNAs and the corresponding miRNA in *Hypericum perforatum* species; ML, length of mature miRNAs; NN, Number of nucleotides hairpin length; MFE, minimum fold energy; AMFE, Adjusted minimum fold energy; MFEI, Minimum fold energy index; ARM, mature miRNA location in hairpin structure; EST ID, Identifier of the 454 transcripts from which miRNA was derived. Italicized, bold and underlined red letters show nucleotide substitutions in miRNAs of *Hypericum* species. Plant species: Aly, *Arabidopsis lyrata*; Gma, *Glycine max*; Osa, *Oryza sativa*; Ppt, *Physcomitrella patens*; Pts, *Populus tremuloides*; Ptc, *Populus trichocarpa*; Vvi, *Vitis vinifera*; Zma, *Zea mays*.

The sequences that passed the various filters and which were specified as putative precursors were then used to perform a multiple sequence alignment (). This alignment was useful to underline the regions corresponding to the miRNA and the miRNA*, as well as the loop sequence located between these two regions. The reliability of each position in the multiple alignments can be deduced from the sequence logo shown in Figure 3. Each of the multiple alignments was mainly conducted to further demonstrate the precursor existence and structure. This analysis confirmed the presence of two extremely conserved regions that corresponded to the miRNA and miRNA*, even in unrelated species, and an intermediate highly variable region that coincided with the sequence region that formed the loop in the secondary structure.

The most likely sequence of the mature miRNA, for which no secondary structure was found to support the identification of a pre-miRNA, was determined by performing multiple alignments among the sequences having a significant

match with a miRNA but lacking some of the requirements later used for the definition of a pre-miRNA. To select the most likely sequence corresponding to the mature miRNA, only the regions resulting from the BLASTN analysis of transcripts and miRNAs used as references were considered. The accuracy of the experimental pipeline was tested using the miRNA sequences extracted from the pre-miRNAs (Table 3 and Figure 4). All of the putative target sequences were grouped into haplotypes that summed up the variability of the nucleotide sequences for each of the multiple alignments, which enabled the determination of the most likely consensus sequence for each family. As seen from the web logo shown in Figure 4, all multiple alignments revealed a high degree of conservation of the different haplotypes. Of note, positions 10 and 11 were invariably preserved in the target sequences used for the construction of the haplotypes (Figure 4).

The comparison of the consensus sequences with the sequences from the mature miRNAs (see Supplementary material, Figure S1) demonstrated substantial homology, and hence confirmed the reliability of the method used for *a posteriori* prediction of the mature sequences. For example, the alignment of the miR414, which includes 31 haplotypes, was shown to differ from the sequence predicted from pre-miRNA for only two positions. The comparison between the two sequences belonging to the miR396 family, one deduced and the other predicted, suggested a potential interaction with the target that results in the formation of a hairpin in the region of sequence pairing, in addition to a canonical interaction in plants without the formation of a hairpin in the region of sequence pairing.

The same method applied on a wider scale enabled the identification of 72 additional consensus sequences of the miRNA motifs grouped into 36 distinct families, including those examined in this study (Table 4). The length of detected miRNA sequences ranged from 18 to 24 nucleotides, with an average value of 21 ± 1.3 (Table 4) and half of the sequences were 21 nucleotides long.

Multiple alignments resulted in a different degree of conservation among the sequences belonging to the same family, as indicated by the number and level of degenerate positions. Notably, 51% of these alignments generated a highly

conserved consensus sequence with no SNPs in the homologous regions. Among the fully conserved sequences, which were supported by multiple independent target mRNAs (contigs, Table 4), we recorded miR156, miR157, miR166 and miR396. In the remaining sequences, the degree of divergence was found to be variable, as expected in the case of multiple alignments of sequences that encode different gene products (Table 4).

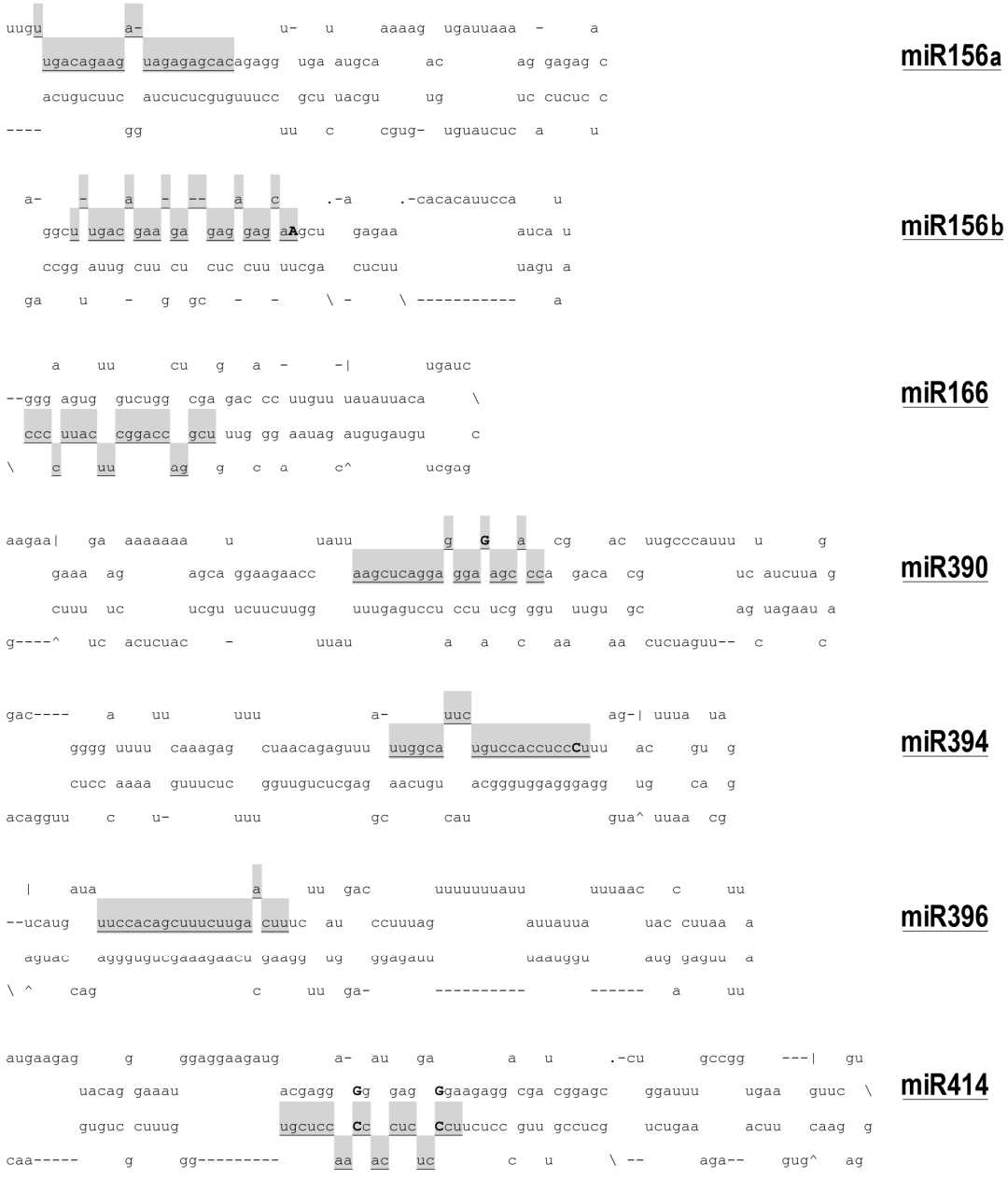

Figure 2. Predicted hairpin secondary structures of the selected *Hypericum* miRNAs identified in this study. Mature miRNA sequences are shaded and underlined. Nucleotide substitutions of conserved miRNAs in other plant species compared with the corresponding miRNAs in *Hypericum* species are shown as uppercase-bold (miRNA precursors may be slightly longer than the sequences shown in this figure).

The number of distinct target sequences, each belonging to a different transcript, was found to range from one single contig up to several dozen, as in the case of miR414. Overall, our prediction method of a given mature miRNA sequence on the basis of its respective target indicated that the degree of degeneration in a sequence may be correlated with the number of targets to which the microRNA is able to pair. As reported in Table 4, each family of miRNAs contains a variable number of members; the largest families were the miR159, miR169, miR395 and miR396, with four members each. Although the miR414 family was identified as containing the largest number of targets (580 contigs), these sequences assembled into a single 21 bp long SMTR that contained four variable positions (Table 4 and Figure 4). The extent and reliability of sequence conservation for this miRNA were further supported by the alignment of the SMTR sequence with the mature sequence deduced from the pre-miRNA (see Supplementary material, Figure S1).

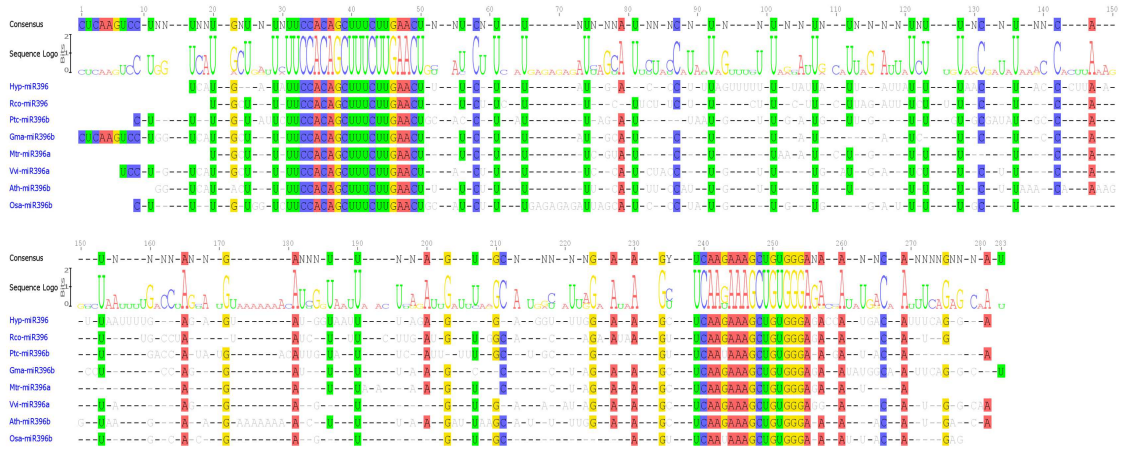

Figure 3. Multiple alignment of miR396 stem-loop sequences from different and unrelated plant species. Conserved nucleotides are shaded, while variable nucleotides are indicted with the white background. For each nucleotide position, the degree of conservation among the different haplotypes is graphically indicated by the web-logo. Rco, *Ricinus communis*; Ptc, *Populus trichocarpa*; Gma, *Glycine max*; Mtr, *Medicago truncatula*; Vvi, *Vitis vinifera*; Ath, *Arabidopsis thaliana*; Osa, *Oryza sativa*.

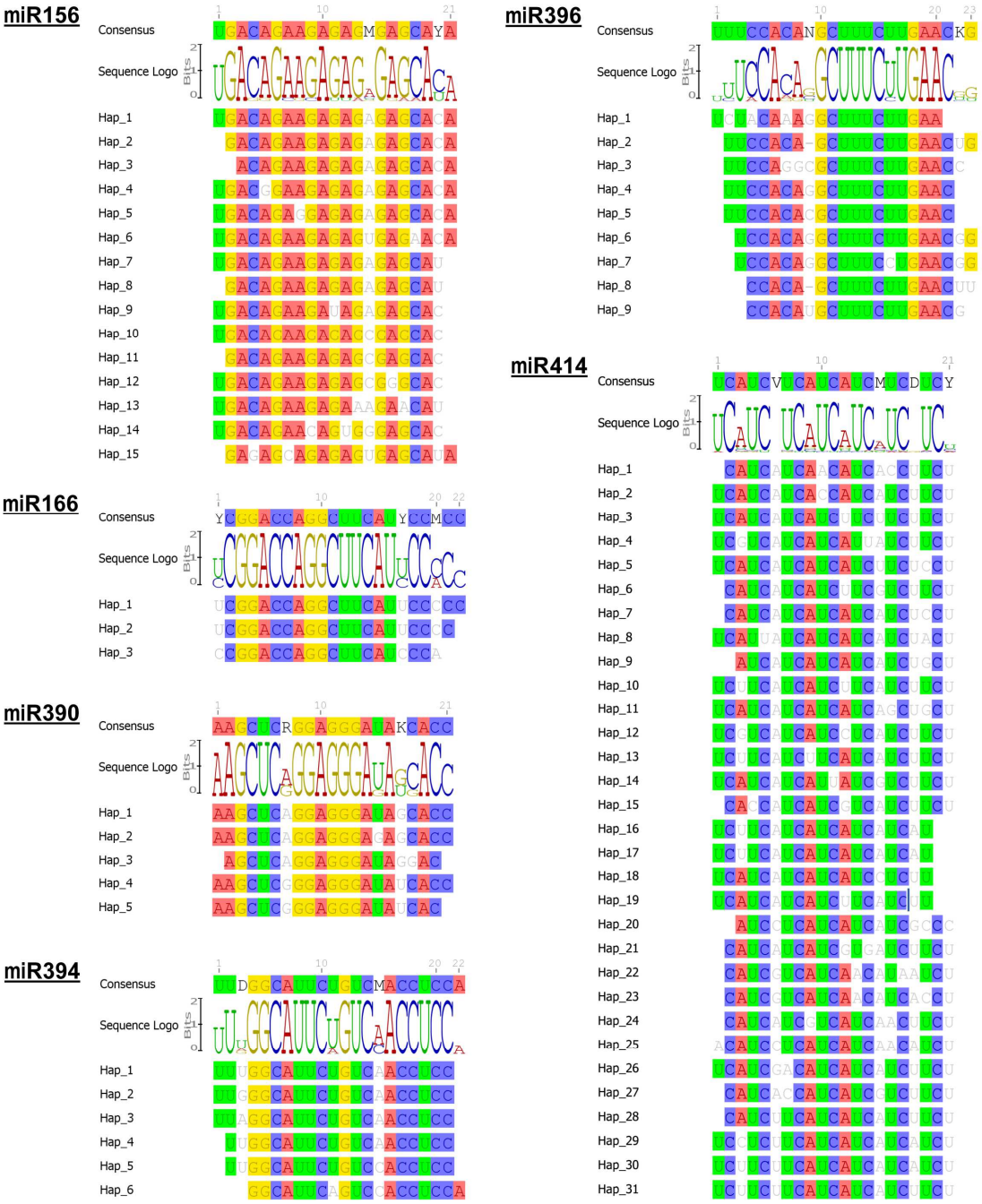

Figure 4. Multiple alignment of target sequences producing significant matches with a known miRNA sequence. The most likely sequence of the mature miRNA was deduced by the consensus sequence. For each nucleotide position, the degree of conservation among the different haplotypes is graphically indicated by the web-logo. For each sequence alignment, the name of the miRNA family is reported.

4.3 Identification of the putative miRNA targets in *Hypericum*

It is well known that the regulation of gene expression by miRNAs is achieved through its pairing with target sequences that normally include fewer than five mismatches (Axtell *et al*., 2013). This biological relationship was exploited for the detection of transcripts encoding putative targets. Overall, this search allowed us to annotate 170 unique sequence targets on the basis of their nucleotide complementarity with known miRNAs. These transcripts proved to belong to gene families with different biological functions. In particular, families that permitted the identification of the largest set of potential targets were the following: miR414, with 57 transcripts, and miR156/157 and miR172, with 17 and 12 transcripts, respectively.

Within the pool of targets, the vast majority (25%) proved to be transcription factors, whereas the others were associated with plant metabolism and response to environmental stress. Many miRNAs are related to the control of transcription factors that are involved in plant development at different levels. Among the most preserved miRNA families, such as the miR156/157, miR172, miR170/171, miR165/166, miR159/319, miR396, miR168, miR160 and miR390, some targets were of particular interest because they are implicated in the development of the flower, including the male and female reproductive organs.

The family miR156/157 controls the production of transcription factors belonging to the family of SQUAMOSA-like proteins, which are involved in defining the identity of the floral meristem.

Table 4. Putative mature miRNAs reported as consensus sequences obtained by multiple alignment of the different target sequences (ML: mature length). Degenerate nucleotide residuals are shown underlined in bold.

miRNA family	Mature sequence consensus	ML	EST	Contigs
Highly conserved miRNAs				
Hyp-miR156a	UUGACAGAAGAGAGAGAGCACA	22	361	25
Hyp-miR157a	GCUCUCUUUCCUUCUUCCA	19	26	6
Hyp-miR157b	CUCUCUA**S**CCUUCU**SY**C**RY**C*	20	11	3
Hyp-miR172a	UG**W**GAAUC**Y**UGAUGAUGCUGCA	22	76	23
Hyp-miR172b	AGAAU**S**UUGA**W**G**M**UG**H**UGCA	20	12	11
Hyp-miR172c	GGAAU**K**U**K**GAUGAUG**Y**UG**Y**AGCAG*	24	16	3
Hyp-miR170a	UCAUUGAGCCGUACCAAU*	18	9	2
Hyp-miR170b	GAUAUUGAUGUGGUUCAAUC*	20	2	1
Hyp-miR171a	GGUGGAGCAGCGCCAAUAUC*	20	13	1
Hyp-miR171b	UGAGCAGCUCCAAUAUCACAU*	21	2	2
Hyp-miR166a	UCGGACCAGGCUUCAUUCCCCC	22	7	7
Hyp-miR166b	AAUGUUGUCUGCCUCGAGG*	19	2	1
Hyp-miR166c	CCGGACCAGGCUUCAUCCCA*	20	2	2
Hyp-miR159a	CUUCCAUAUAUGGGGAGCUUC*	21	21	3
Hyp-miR159b	**W**UUGGA**KK**GAAGG**S**AGCUC**YH**	21	70	17
Hyp-miR159c	GAACUCCCUUACUCCAAAAC*	20	1	1
Hyp-miR159d	AGCUGCUUAGCUAUGGAUCCC*	21	1	1
Hyp-miR319	**Y**U**B**GGACUGAAGGGAGCUCACU	22	26	6
Hyp-miR396a	UUUCCACAGGCUUUCUUGAACGG	23	40	13
Hyp-miR396b	UCCCACAGCUUCACUGAACC*	20	11	1
Hyp-miR396c	AGUUC**M**AG**M**A**W**GU**S**CUUGG**W**AA*	22	7	3
Hyp-miR396d	UCUCCUC**N**GGCMUUCUUGAACUU*	23	3	2
Hyp-miR168a	GCUUGGUGC**W**GGUCG**R**GAAC*	20	11	1
Hyp-miR168b	GACCCCGCCUUGGGCCAAUUGAAU*	24	1	1
Hyp-miR160a	UGCCUGGCUCCCUGUAUGCC**W**	21	18	12
Hyp-miR160b	GUAUGAGGAGCCAUGCAUA*	19	1	1
Hyp-miR160c	GUACAGGGUAGUCAAGGAUG*	20	1	1
Hyp-miR390	AAGCUC**R**GGAGGGAUA**KS**ACC	21	7	6
Moderately conserved miRNAs				
Hyp-miR394	UUUGGCAUUCUGUC**M**ACCUCCA	22	17	7
Hyp-miR164a	UGGAGAAGCAGG**R**CAC**D**U**RMK**	21	49	19
Hyp-miR164b	UGGAGAAGCA**S**GG**S**AC**K**UG**M**U	21	20	6
Hyp-miR169a	**W**AGCCAAG**R**AUGA**M**UUGCC**K**G	21	66	16
Hyp-miR169b	CAAGUUGUCCUUCGGCUUCA*	20	21	2
Hyp-miR169c	UAGGCAAAAAUGGCUUGCCUA*	21	7	1
Hyp-miR169d	UGAGCCAAGUAAGGCUUGCC*	20	3	2
Hyp-miR167a	UGAAGCUGCCAG**S**CUGAUCUCA	22	133	15
Hyp-miR167b	AGGUCAUC**Y**U**R**CAGC**Y**UCAGU	21	8	5
Hyp-miR162a	UCGUUAAACCUUCGCAUCCAG	21	8	4
Hyp-miR162b	CGAUGAGUCUCUGCAUCCAG*	20	1	1
Hyp-miR398a	UGUGUUCUCAGGUCGCCCCUG	21	11	6
Hyp-miR398b	GGGUYGMCAUGAKRACAYAUG*	21	2	2

Table 4. Continues

miRNA family	Mature sequence consensus	ML	EST	Contigs
Hyp-miR414	UCAUC**V**UCAUCAUC**M**UC**D**UC**Y**	21	2346	580
Hyp-miR393a	UCAUGCUGUCUCUUUGAAUU*	20	1	1
Hyp-miR393b	UCCUAAGGGAUCUCCUUGAUCU*	22	1	1
Hyp-miR397a	UC**R**UUGAG**YR**C**M**GCGUUGA**Y**G	21	17	5
Hyp-miR397b	UU**Y**AU**Y**GACUGCAGUGUUUAUU*	22	9	2
Hyp-miR163a	UUGAAGAG**B**AC**Y**UGGAACUUCGAU	24	55	12
Hyp-miR163b	GAAGAAGAGUUGGAACUUA	19	106	6
Lowly conserved miRNAs				
Hyp-miR395a	UGAA**K**U**K**UU**WR**GGGGAACUC	20	30	14
Hyp-miR395b	UU**YY**CUUCAAG**M**ACUUCACGA	21	42	15
Hyp-miR395c	GAAGU**K**UUUGGGGGAUUCU	21	8	5
Hyp-miR395d	UUCCUUUCAAAC**M**CUUCACAU	22	4	4
Hyp-miR408a	CUGC**W**CUGCCUC**W**UCC**Y**UG**K**CU	22	18	12
Hyp-miR408b	CUGUGAACAGGCAGAGGAUG*	20	1	1
Hyp-miR399a	UGACAAAGGAGAUGUGCCCAG*	21	16	1
Hyp-miR399b	GGGCAA**WR**UC**W**C**Y**AUUGG**Y**AGA*	22	29	3
Hyp-miR161a	CCCGAUGUAGUGACUUUCAA*	20	8	3
Hyp-miR161b	UCAAGGUAUUGAAAGUGACUA*	21	8	3
Hyp-miR161c	UUAAAGGUGACUACAUCGGGG*	21	6	3
Hyp-miR173	UU**Y**GCUUG**S**AG**Y**GA**S**AAAUCAC	22	8	4
Hyp-miR474	CAAA**R**GU**K**G**Y**UGGGUUUGG**H**UGGG	24	46	19
Hyp-miR528a	UUGCAGGGACAGGGAGAGGA*	20	13	2
Hyp-miR528b	CUGUGGCUGCCUCUUCCAUU	20	39	7
Hyp-miR529	AGAAGAAGAGAGA**SRKS**A**S**AGC**Y**U	24	47	24
Hyp-miR534	UAUGUCCAUU**R**C**W**GUU**SY**AUAC	22	11	8
Hyp-miR824	UCUCAUCGAUGGUCUUGA*	18	1	1
Hyp-miR1028	UGACAUUGUAG**W**UCUA**Y**GU	19	32	4
Hyp-miR1219	UUUCCU**K**CCUCUCAC**W**AGCUU*	21	7	3
Hyp-miR1442	ACACCUCUAUUACUAUGAAU*	20	6	1
Hyp-miR1530	UUUUCACAUAAAUUAAAA**W**AU*	21	4	3

ML, length of mature miRNAs; EST, number of Expressed Sequence Tag used for alignment; Contigs, Number of assembled reads

Among the potential targets of the family miR172, several transcripts were identified by sequence homology with respect to genes encoding transcription factors that belong to the class APETALA-2. This type of target, implicated in the identity of floral organs and, in particular, in the development of the perianth (the sum of sepals and petals), is among the most significant for this family with an FDR equal to 1.70 E−02 (Table 5, see also Table S1 of Supplementary materials). The other targets identified for this family encode proteins involved in sugar metabolism, particularly the mannose, proteolysis and transcriptional regulation processes. Of the other conserved families, the miR159/319 family controls the production of transcription factors of the MYB type as well as the transcription factors involved in the response to heat stress.

The same approach confirmed that the miR396 family is likely to be responsible for the regulation of transcription factors termed UPA17 in *Hypericum* (see Table S1). This family is also involved in the control of transcription factors responsible for the development of pollen grains. Both miR168 and miR390 control targets associated with metabolism and ion binding. The miR160 family regulates the transcription of genes responsive to auxin, with functions that range from the development of sepals and roots, to the regulation of the cell cycle.

Among the less conserved miRNA families, selected for their possible role in the regulation of plant development, those involved in the regulation of transcription factors are the miR164 and miR163 families. In particular, the miR164 family controls a series of transcription factors, known as NAC-domain genes, which are involved in plant development at different levels, with particular reference to the flower. The miR163 family controls transcripts putatively related to transcription factors of the SUPERMAN type.

The family that presents the greatest number of transcripts and that also controls the largest number of functions is miR414, although it is most significantly associated with the enzyme cysteine peptidase as target (FDR = 1.00 E−02). Of particular interest are the miR167 and miR393 families, which control

transcription factors involved in the development of the flower and, in particular, the reproductive organs (Table 5).

Finally, four different miRNA families (miR529, miR534, miR824 and mir1219) led to the identification of targets whose sequences show no similarity to any protein in the GenBank databases.

The annotation of the target sequences for their putative molecular functions and biological processes underlined the presence of some predominant classes for each of the two GO vocabularies (Figure 5). For molecular function, the highest number of GO terms associated with the *Hypericum* sequences proved to be DNA binding, nucleotide binding and sequence-specific DNA binding and transcription factor activity (Figure 5, panel A). To a lesser extent, the *Hypericum* targets were annotated as chromatin binding, RNA binding and kinase activity. Moreover, the majority of GO annotations related to biological processes are related with responses to stress and endogenous stimuli, and signal transduction. A few sequences were also annotated as flower development, pollen and embryo development (Figure 5, panel B).

The statistical analysis of the relative abundance of ontological annotations assigned to the putative targets of each miRNA family identified 12 different families with terms associated with molecular function (17 GO terms) or to a biological process (30 GO terms) with FDR values lower than 0.05 (Table 5). For example, 7 out of 12 miRNA families are associated with the GO terms DNA binding or regulation of transcription. Similarly, GO terms related to hormone response (*e.g.*, auxin-mediated signaling pathway, response to salicylic acid and gibberellin acid stimuli) are associated with five different miRNA families: miR160, miR164, miR167, miR170/171 and miR393 (Table 5). The same analysis also showed how a large number of target sequences of miR159/319 and miR170/171 families are annotated with GO terms related to the response to heat, wounding and fungal attack. Finally, a significant number of target families regulated by miR160, miR167 and miR393 are annotated with GO terms associated with the development of the flower, the carpels or the stamens.

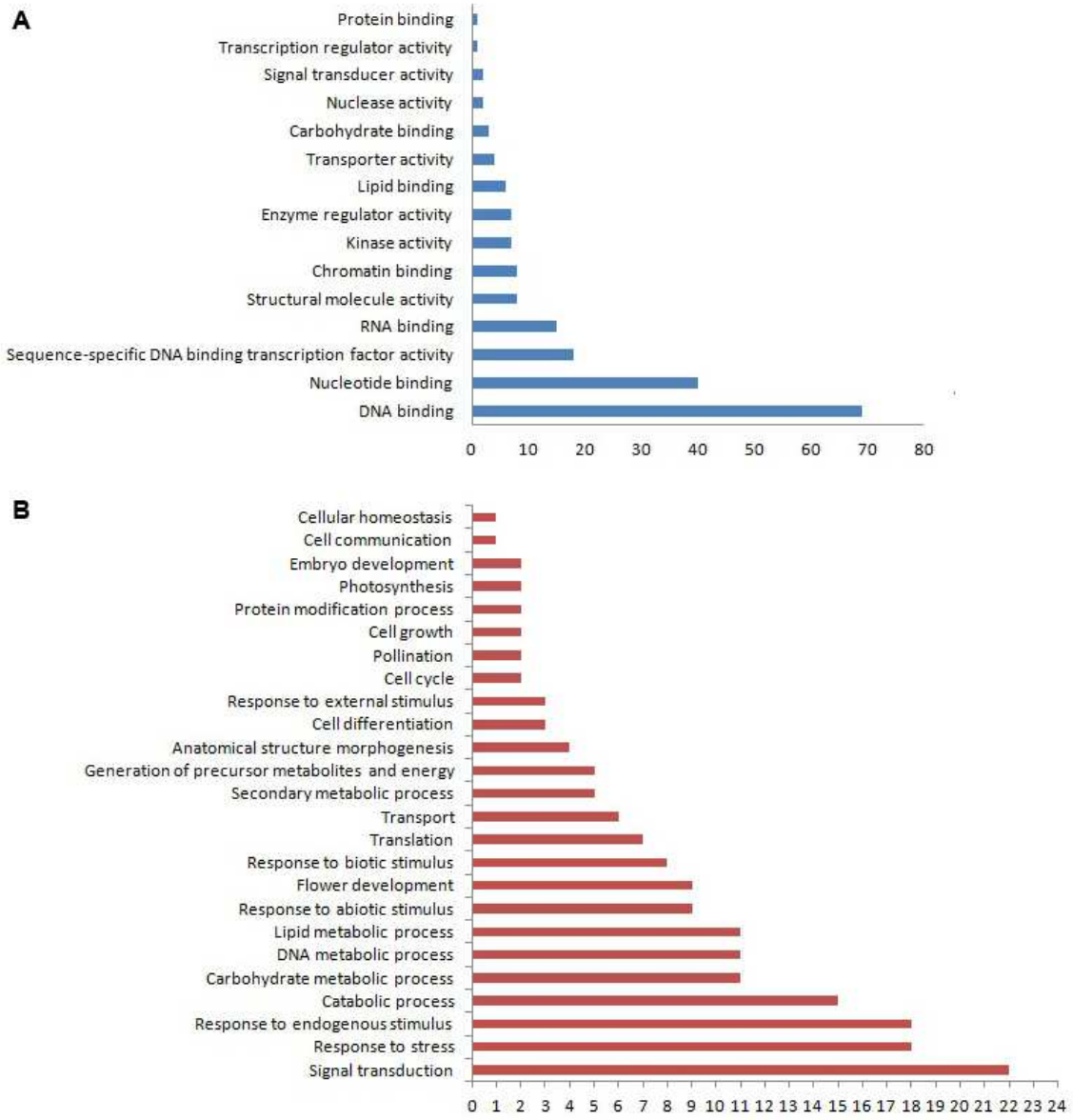

Figure 5. Annotation of the target sequences for their putative molecular functions and biological processes according to the two GO vocabularies. GO terms associated to molecular functions are shown above (panel A), whereas GO terms related to biological processes are shown below (panel B).

Table 5. Ontological terms significantly over-represented in each miRNA family. Only miRNA families with over-represented GO terms are shown. Target description, GO term and vocabulary are reported. For each enriched GO term, the False Discovery Rate (FDR) is also indicated. F: Molecular Function, B: Biological Process.

miRNA	Target description	GO term and vocabulary		FDR
miR156/157	Squamosa promoter-binding protein, putative	DNA binding	(F)	6,40E-16
miR159/319	Chaperone protein DnaJ	Response to heat	(P)	9,60E-04
		Protein folding	(P)	4,10E-02
	R2r3-myb transcription factor, putative	DNA replication	(P)	3,50E-02
	Carbonic anhydrase	Carbon utilization	(F)	3,80E-02
		Carbonate dehydratase activity	(P)	4,10E-02
miR160	Auxin response factor, putative	Auxin-mediated signaling pathway	(P)	3,00E-04
	Predicted protein	Regulation of transcription, DNA-dependent	(P)	3,00E-04
		Root cap development	(P)	3,00E-04
		DNA binding	(F)	6,00E-03
		Pattern specification process	(P)	6,00E-03
		Sepal development	(P)	2,20E-02
		Cell division	(P)	2,70E-02
miR164	Predicted protein/ NAC domain-containing protein 21/22 putative/ Transcriptional factor NAC35	DNA binding	(F)	1,60E-03
	NAC domain-containing protein 21/22, putative/ Transcriptional factor NAC35	Regulation of transcription, DNA-dependent	(P)	1,60E-03
	Predicted protein	Response to gibberellin stimulus	(P)	1,10E-02
		Response to salicylic acid stimulus	(P)	3,80E-02
miR167	Predicted protein	Auxin mediated signaling pathway	(P)	1,10E-04
		Flower development	(P)	1,20E-02
	Retrotransposon protein	Regulation of transcription, DNA-dependent	(P)	1,80E-02
miR169	Nuclear transcription factor Y subunit A-1, putative/Predicted protein	Sequence-specific DNA binding transcription factor activity	(F)	4,60E-09
	Myb3r3, putative/ Nuclear transcription factor Y subunit A-1 putative/ Predicted protein	DNA binding	(F)	1,50E-03
miR170/171	12-oxophytodienoate reductase 7	12-oxophytodienoate reductase activity	(F)	9,30E-08
		Jasmonic acid biosynthetic process	(P)	4,20E-06
		Response to ozone	(P)	4,70E-06
		FMN binding	(F)	5,70E-05
		Response to wounding	(P)	1,50E-04
		Response to fungus	(P)	1,50E-04
miR172	GDP-mannose 4,6-dehydratase	GDP-mannose 4,6-dehydratase activity	(F)	1,10E-02
	Putative polyprotein	2 iron, 2 sulfur cluster binding	(F)	1,70E-02
	GDP-mannose 4,6-dehydratase	GDP-mannose metabolic process	(P)	1,70E-02
	AP2 domain-containing transcription factor	Sequence-specific DNA binding transcription factor activity	(F)	1,70E-02
	AP2 domain-containing transcription factor	2-alkenal reductase [NAD(P)] activity	(F)	4,20E-02
miR393	F-box family protein/TIR1 protein putative	Inositol hexakisphosphate binding	(F)	9,60E-13
	F-box family protein	Auxin mediated signaling pathway	(P)	4,40E-07
		Cellular response to nitrate	(P)	1,00E-05
		Auxin binding	(F)	2,40E-05
		Response to molecule of bacterial origin	(P)	2,40E-05
		Pollen maturation	(P)	9,40E-05
		Primary root development	(P)	3,90E-04
		Lateral root development	(P)	1,40E-03
		Stamen development	(P)	2,70E-03
miR396	UPA17/Predicted protein	Regulation of transcription, DNA-dependent	(P)	1,90E-04
		Hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides	(F)	6,10E-03
miR397	Laccase, putative	L-ascorbate oxidase activity	(F)	1,90E-03
		Lignin catabolic process	(P)	1,90E-03
miR414	Cysteine proteases	Cysteine-type peptidase activity	(F)	1,00E-02

Validation of *H. perforatum* flower pre-miRNAs was carried out by RT-PCR. To rule out the possibility of RNA impurity due to genomic DNA contaminants, a negative RT sample was used as a negative control in all RT-PCR assays (data not shown). The amplification of each of the pre-miRNAs using flower cDNA as a template confirmed the expression of the precursors in the flower (Figure 6). For these pre-miRNAs, the PCR products were purified, sequenced and aligned to the pre-miRNA sequences to confirm the nature of the amplicons. In addition, preliminary investigations based on semi-quantitative RT-PCR assays showed that some precursors are preferentially expressed in specific flower tissues, whereas some others are uniformly expressed in all plant organs (Figure 7). In particular, we found that miR156b is strongly expressed in all plant organs, while miR390 and 396 are similarly expressed at much lower levels. Both miR166 and miR414 proved to be strongly expressed in all flower verticils (*i.e.*, anthers, pistils, sepals and petals), including young buds, while transcripts of these precursors were never detected neither in leaves nor in roots. Real-Time PCR analyses confirmed the expression levels and patterns of pre-miRNAs in flower verticils. Additionally, miR156a was found preferentially expressed in anthers at early developmental stages, in pistils at late developmental stages and also in pre-meiotic flower buds. The expression of miR394 in both anthers and pistils was much higher at early than late developmental stages (Figure 8). Finally, miR390 was shown uniformly expressed in all flower verticils, while miR396 was found equally expressed in pistils and differentially expressed in anthers at different developmental stages (see Figure 8).

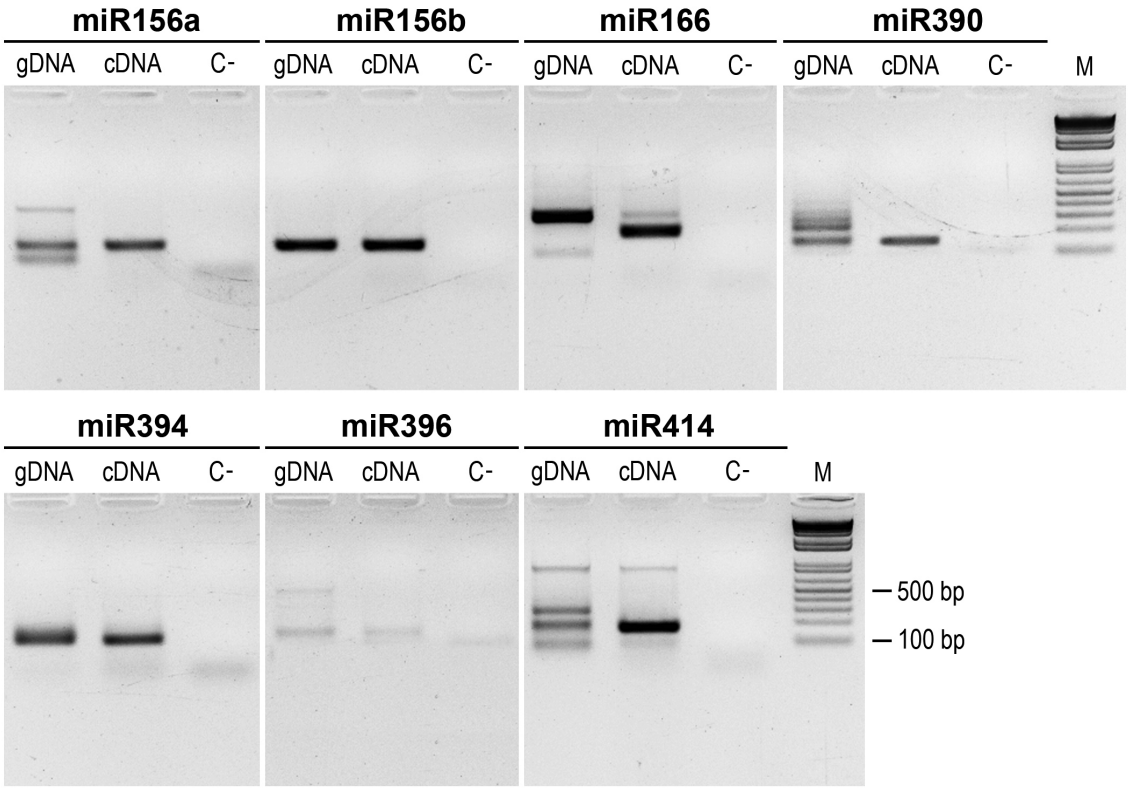

Figure 6. PCR amplification of miRNA stem-loop regions using genomic DNA (gDNA) and cDNAs from *H. perforatum* flowers as templates. Negative controls without template are also shown (C-). M indicates the 1 Kb Plus DNA ladder.

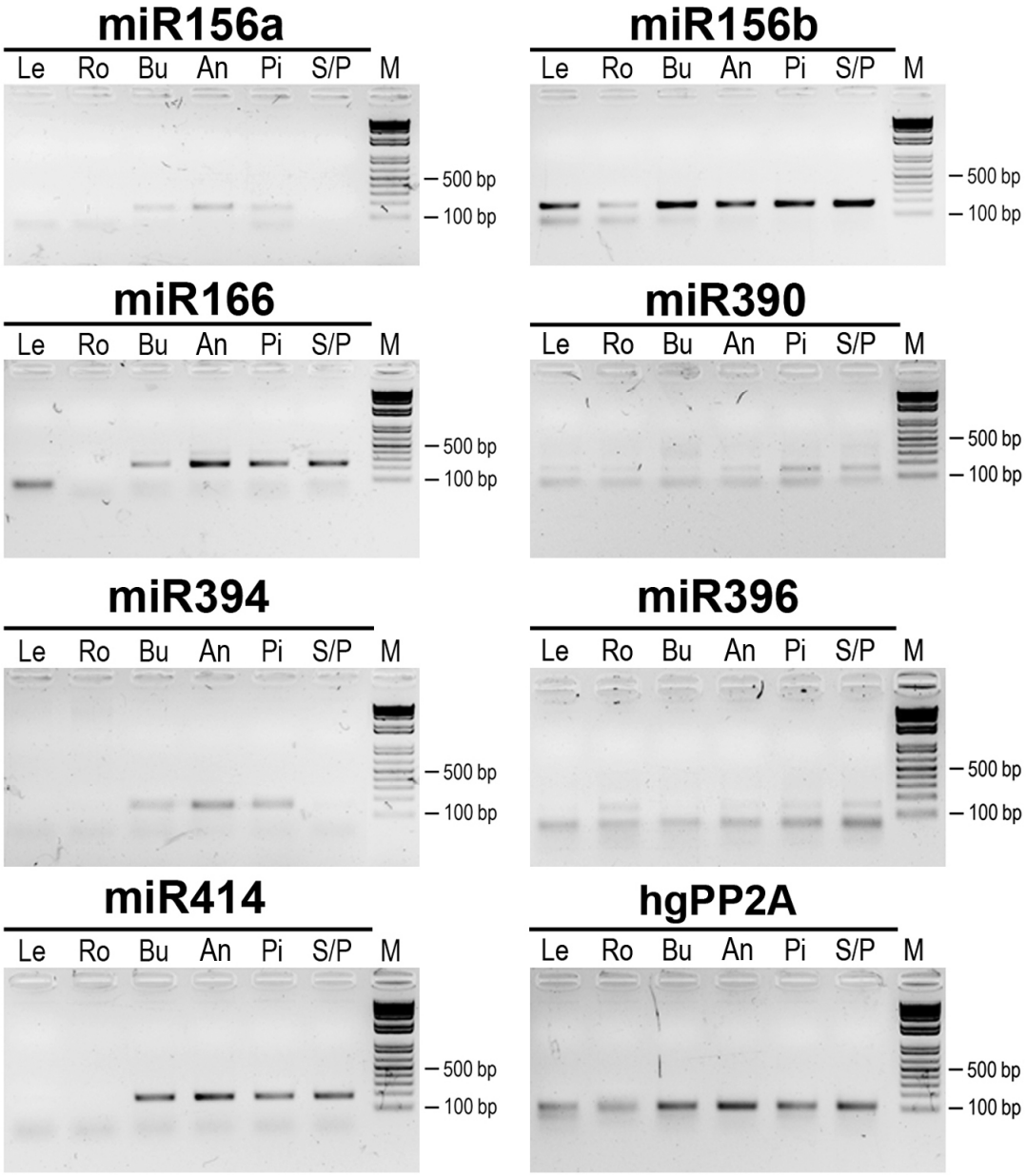

Figure 7. Presence of miRNA precursors in different plant tissues verified by semi-quantitative RT-PCR analysis. Specific primers for miR156a, miR156b, miR166, miR390, miR394, miR396, miR414 stem-loop regions were tested in leaves (Le), roots (Ro), pre-meiotic flower buds (Bu), anthers (An), pistils (Pi) as well as sepals and petals (S/P). M indicates the 1 Kb Plus DNA ladder. hgPP2 refers to the housekeeping gene PP2 encoding the *Hypericum* Protein Phosphatase 2A, subunit A3.

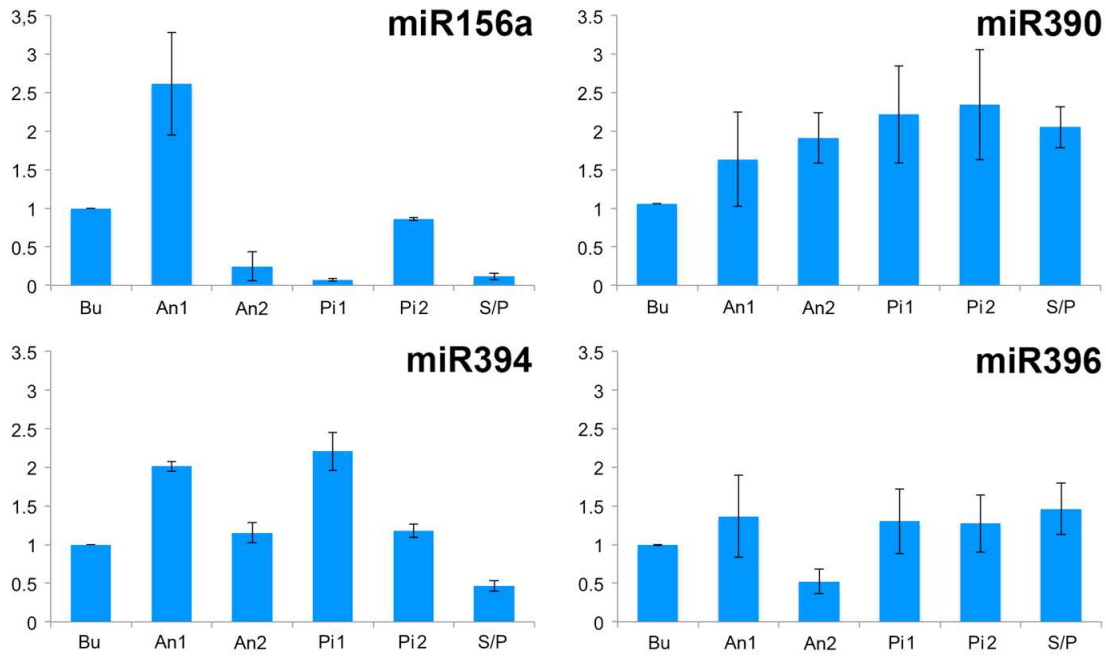

Figure 8. Expression levels and patterns of *Hypericum* pre-miRNAs assessed by Real-Time PCR. Bu: pre-meiotic flower buds; An: anthers; Pi: pistils; S/P: sepals and petals. Both anthers and pistils were investigated at two different flower developmental stages, such as 11-12a (An1/Pi1) and 12b-14 (An2/Pi2). Fold changes in mRNA expression were calculated relative to the control using the DDCt method. The presented values are the means of data collected for three biological replicates, each assayed using three technical replicates. Error bars indicate standard deviations.

5 DISCUSSION AND CONCLUSIONS

MicroRNAs (miRNAs) have recently emerged as important gene regulators in plants. This is the first study dealing with the identification and characterization of miRNAs and their target genes in St. John's wort (*H*. *perforatum* L., 2*n*=4*x*=32), a medicinal plant that produces pharmaceutically important metabolites with antidepressive, anticancer and antiviral activities. This species is regarded as a serious weed in many countries, and it is also considered an attractive model system for the study of apomixis. Apomixis is an asexual reproductive strategy for cloning plants through seeds. In apomixis, the offspring are exact genetic replicas of the parent because the embryos are derived from the parthenogenic development of apomeiotic egg cells (for recent reviews on apomixis, see Albertini *et al*., 2010; and Pupilli and Barcaccia, 2012).

It is hypothesized that apomixis, from an evolutionary perspective, may be considered a result of sexual failure rather than as a means of clonal success (Silvertown, 2008), it is also true that apomixis, as a biological process of seed formation, represents an altered form of sexuality rather than a new developmental program (Koltunow and Grossniklaus, 2003). The idea that apomixis is an altered form of sexuality which results from temporal and/or spatial alterations in the sexual seed formation program, suggests a link between apomixis and regulatory mechanisms acting at the post-transcriptional level (*e.g*., sRNAs). Indeed, this perspective would imply the presence of reproductive machinery largely shared amid sexual and apomictic plants, with regulators that differentially-modulate between the reproductive modes (*i.e*., likely based on chromatin re-modeling factors or transacting and heterochromatic interfering RNAs involved in both transcriptional and post-transcriptional gene regulation).

Small RNAs have been recently studied in different model systems, and it is now known that mutations in the molecular pathways that generate sRNAs may dramatically affect fertility (Van Ex *et al*., 2011; Tucker *et al*., 2012). Recent work demonstrated that strong mutant alleles of genes involved in the formation and activity of miRNAs such as AGO1, DCL1, HEN1 and HYL1 disrupt reproductive

development (reviewed by Van Ex *et al*., 2011). However, interpreting these phenotypes is frequently difficult because such mutations have ectopic effects and influence different aspects of plant development (Axtell *et al*., 2013). Recently, two members of the ARGONAUTE protein family, AGO5 and AGO9, which are involved in the regulatory pathway of sRNAs in plants, have been associated with cell specification and embryo sac development (Olmedo-Monfil *et al*., 2010; Tucker *et al*., 2012). The orthologue of AGO5 in rice was reported to be essential for the progression of pre-meiotic mitosis and meiosis (Nonomura *et al*., 2007), and the production of viable gametes without meiosis was reported in maize lacking the orthologue of AGO9 (Singh *et al*., 2011). Some of the genes isolated and characterized from sexual species may play a role in the framework of apomixis, and it may be possible that sRNAs act by silencing master genes directly involved in differentiating apomictic from sexual pathways.

In this study, we have conducted a comprehensive analysis of *Hypericum* miRNAs produced in flower organs, and have computationally predicted their putative targets. We focused our bioinformatics investigations on 36 families of miRNAs that were selected because they were among the most conserved families (Zhang *et al*., 2006a) and due to the biological processes in which they were involved (Guiling Sun, 2012). The prediction of the structures enabled us to identify seven pre-miRNAs belonging to six conserved miRNA families, including miR156, miR166, miR390, miR394, miR396, and miR414. In this study, some of the targets for *Hypericum* miRNAs have counterparts that were previously identified and validated in other species, such as *Arabidopsis thaliana* (Addo-Quaye *et al*., 2008), *Glycine max* (Song *et al*., 2011) and *Vitis vinifera* (Pantaleo *et al*., 2010), as well as *Populus trichocarpa*, *Medicago truncatula* and *Oryza sativa* (Guiling Sun, 2012).

Overall, the identification of the putative miRNA targets in *Hypericum* allowed us to annotate 170 sequences that are likely related to putative targets on the basis of their nucleotide complementarity with known miRNAs.

Several known targets of specific miRNAs, mainly transcription factors, are known to control different physiological processes and genetic programs associated with plant metabolism, flowering, hormone signaling and stress responses. This is a clearly emerging trend from our data, as some of the miRNA families in this study were characterized by over-representations of GO terms including as DNA-binding, response to hormones and response to stress and external stimuli. Additionally, the identification of ontological terms that are significantly enriched in one mRNA/target family means that the same processes of functions are prevalent in that mRNA/target family and are distinctive of the actions of that miRNA.

Large sets of potential targets were identified in the miR414, miR156/157 and miR172 families. In addition, the vast majority of targets were transcription factors associated with plant development at different levels. Among the most conserved miRNA families, such as miR156/157, miR172, miR170/171, miR165/166, miR159/319, miR396, miR168, miR160 and miR390, some targets were of particular interest because they were implicated in the development of the flower, including the male and female reproductive organs. Of particular interest was the miRNA167 family that controls patterns of auxin-responsive factors (*ARF6* and *ARF8*) in *Arabidopsis*, and regulates both female and male reproduction, with specific reference to ovules and anthers (Wu *et al*., 2006). The auxin-responsive factors are transcription factors that regulate the expression of auxin-responsive genes in both activation and repression modes (Guilfoyle and Hagen, 2007). The *ARF1* homolog of *Arabidopsis* was found expressed earlier in apomictic ovules compared to sexual ones in the grass *Paspalum simplex*, suggesting that the auxin response may affect the differentiation of aposporic initials from nucellar cells by repressing a class of auxin-responsive genes that maintain the undifferentiated state of nucellar cells once the megaspore mother cell is formed (Polegri *et al*., 2010). A role of the auxin-response machinery in apomictic reproduction has been hypothesized in *Hieracium* (Koltunow *et al*., 2001) and an auxin-responsive protein was found expressed in pistils of apomictic *Panicum maximum* (Yamada-

Akiyama *et al*., 2009). Finally, the developmental fate of non-reproductive cells has been switched in embryo sacs of *Arabidopsis* by manipulating auxin-response genes (Pagnussat *et al*., 2009).

In addition, the miR156/157 family controls transcription factors of the SQUAMOSA Promoter Binding Like (SPL) family, which is involved in the regulation of developmental timing, including the phase transition from juvenile to adult and from vegetative to reproductive, in concert with miR172 (Xie *et al.*, 2006; Gandikota *et al.*, 2007). In *Arabidopsis*, miR156 regulates the expression of miR172 by SPL9 that, redundantly with SPL10, directly promotes the transcription of miR172b (Chen *et al*., 2004). Among the potential targets of the miR172 family, several sequences revealed homology with genes encoding transcription factors that belong to the class APETALA-2 (AP2, class A gene). AP2, the progenitor of this transcription factor family, which also contains TOE1-3, SMZ and SNZ transcription factors, is known to be active in flower development. It is implicated in the identity of floral organs and acts in concert with AGAMOUS (AG, a class C gene), restricting each other's activities to their proper domains of action within the floral meristem.

Of the other conserved families, the miR159/319 family controls the production of transcription factors of the MYB type, a large family of proteins considered as key factors in regulatory networks controlling development, metabolism and responses to biotic and abiotic stresses (Palatnik, 2007). The miR396 family is responsible for the regulation of several transcription factors, including those which control conserved targets belonging to the growth-regulating factors (GRFs), which regulate cell proliferation in *Arabidopsis* leaves (Liu *et al.,* 2009; Rodriguez *et al.*, 2010). This miR396 family was also expressed in *Arabidopsis* inflorescences and pollen grains, and is likely to be involved in their maturation (Chambers and Shuai, 2009). The functional specialization of the miR396 regulatory network in plants was recently found to act through distinct microRNA-target gene interactions whose control may be biologically relevant (Debernardi *et al*., 2012). In *Hypericum* we also identified the precursor of miR394. While experimental evidence in *Arabidopsis* and rice indicated that miR394 in

involved in the regulation of the cell cycle, recent research in Japanese apricot (*Prunus mume*) revealed that miR394 is differentially expressed in perfect and imperfect flowers, leading to the conclusion that it may be associated with pistil development (Gao *et al*., 2012). Furthermore, Knauer *et al*. (2013) have recently reported that miR394 would play a central role in maintaining shoot meristem stem cell identity in *Arabidopsis* by repressing a specific F-box protein. In fact, the miR394 was identified as a mobile signal produced by the surface cell layer (*i.e*., the protoderm) that confers stem cell competence to the distal meristem (Knauer *et al*., 2013).

Among the less conserved miRNA families selected for their possible roles in the regulation of plant development, those involved in the regulation of transcription factors are the miR164 and miR163 families, which are involved in development of the plant at different levels, including the flower. In particular, the miR163 family is related to the transcription factors of SUPERMAN, a gene that in *Arabidopsis* plays a role in controlling the morphogenesis of flower organs, particularly the boundary between stamen and carpel development in the flower (Nakagawa *et al*., 2004). It is well known that this gene is expressed early in flower development in the stamen whorl adjacent to the carpel whorl, and additionally this gene interacts with the other genes of the ABC model of flower development in a variety of ways.

The miR414 family revealed the greatest number of transcripts and is also characterized as controlling the largest variety of functions, and has as many as 57 different targets, including some involved in the development of the flower and, in particular, the reproductive organs.

The miR390 family was expressed in the *Hypericum* flower transcriptome. In *Arabidopsis* miR390 targets the four known *TAS1-4* genes and triggers the entry of a specialized RNAi pathway that culminates with the production of 21 nucleotide trans-acting siRNAs (Chitwood *et al*., 2009). Mutants defective for one of these *TAS* genes exhibited aberrant flower morphologies and plant phenotypes with an accelerated juvenile-to-adult phase transition (Chitwood *et al*., 2009). The miRNA

390 has an elegant regulatory system in which the miRNA is broadly expressed and mobile throughout the shoot apical meristem (but not exclusively in the SAM), but its activity becomes restricted to specific cell layers by the expression of effector complexes that are spatially restricted (Chitwood *et al*., 2009). Unfortunately, the role of miR390 with respect to flower development has not yet been elucidated and further research will be necessary to understand if this short regulatory pathway will eventually interact or interfere with tissue-specific effectors and influence the formation of flower-specific components. We can speculate by stating that single mutants for each of the two genes, *RDR6* (an RNA-dependent RNA polymerase) and *SGS3* (an RNA-binding protein suppressor of gene silencing), which are responsible for two consecutive steps of the synthesis of the trans-acting siRNA from miRNA390/TAS precursors, resulted in phenotypes reminiscent of aposporic apomixis (Olmedo-Monfil *et al*., 2010). In addition, the importance of sRNAs in the determination of cell fate and gamete formation has been demonstrated by recent studies (Olmedo-Monfil *et al*., 2010; Singh *et al*., 2011; Tucker *et al*., 2012).

The statistical analysis of the relative abundance of annotations assigned to the putative targets of each miRNAs family allowed us to identify 12 families with GO terms associated with molecular functions and biological processes, such as DNA binding or regulation of RNA transcription. Several other GO terms were associated with the response to hormones (*e.g*., auxin-mediated signaling pathway, response to salicylic acid stimulus and response to gibberellin acid stimulus). Our data support the regulatory role of miR160, miR167 and miR393 in flower development and in the morphogenesis of carpels and stamens. These findings further confirm that the molecular machinery for the control of gene expression is frequently conserved among unrelated species, not only at the miRNA sequence levels but also in terms of miRNA/target relationships. Our research demonstrated that the mature sequences of known conserved miRNA could be deduced from the alignment of target sequences from the same species. In this light, predictive computational investigations could be implemented using the large accumulation

of biological and molecular data, particularly from NGS technologies, applied in whole-plant genomics and organ-specific transcriptomics. Thus the increasing availability of sequence data in public databases will promote the possibility of studying the spatial and temporal expression levels and patterns of conserved miRNAs.

In a near future, the availability of large transcriptome datasets will likely improve our computational ability to discover mature miRNA sequences and to predict their target genes. If this holds true, PCR-based methods would allow us cloning and sequencing the original miRNA precursors in a non-model species like *H. perforatum*. Gene expression studies will be feasible for a wide range of biological processes, including sexual and apomictic reproductive pathways. In particular, further investigations will help clarifying the possible role of miRNA167 (family that controls patterns of auxin-responsive factors and regulates plant reproduction), miR156a (preferentially expressed in anthers and pistils at different developmental stages), miR394 and miR396 (modulately expressed in anthers and pistils) during gametogenesis in plants.

In conclusion, we characterized miRNAs and their putative targets in *Hypericum* to provide a comprehensive list of conserved miRNA families and to reveal the potential role of their regulatory functions. We demonstrated that *H. perforatum* has both conserved and species-specific miRNAs and that these miRNAs potentially target dozens of genes with a wide range of biological functions in flower development and plant reproduction. This analysis paves the way towards identifying miRNAs specifically expressed in reproductive organs and that play a role in sexual and apomictic reproductive pathways.

6 ACKNOWLEDGEMENTS

This research was supported by the following grants: Research Project for Young Researchers of the University of Padova (year 2010), "Comparative and functional genomics for cloning and characterizing genes for apomixis" (code: GRIC101130/10), Principal investigator: Giulio Galla. Academic Research Project of the University of Padova (year 2012), "Transcriptomics of reproductive organs in model species for comparative analysis of the genetic-molecular factors characterizing sexual and apomictic processes" (code: CPDA128282/12), Principal investigator: Gianni Barcaccia.

Data presented and discussed in the present dissertation were recently published: see Galla G, Volpato M, Sharbel TF, Barcaccia G (2013) Computational identification of conserved microRNAs and their putative targets in *Hypericum perforatum* L. flower transcriptome. Plant Reprod, doi:10.1007S00497-013-0227-6

7 REFERENCES

Addo-Quaye C, Eshoo TW, Bartel DP, Axtell MJ (2008) Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome. Curr Biol 18(10): 758-762

Amiteye S, Corral J, Vogel H, Sharbel T (2011) Analysis of conserved microRNAs in floral tissues of sexual and apomictic Boechera species. BMC Genomics 12(1): 500

Albertini E, Barcaccia G (2007) Apomeiosis in *Poa pratensis* L. and *Medicago sativa* L. In: Horandl E, Grossniklaus U, Sharbel T and van Dijk P (eds) Apomixis: Evolution, Mechanisms and Perspectives. Gantner, Ruggel [Regnum Veg. 147], pp. 93-116

Albertini E, Barcaccia G, Mazzucato A, Sharbel TF, Falcinelli M (2010) Apomixis in the era of biotechnology. In: EC Pua and MR Davey (eds), Plant Developmental Biology - Biotechnological Perspectives, Volume 1. Springer, Heidelberg, Berlin (Germany), pp. 405-436

Axtell MJ, Bartel DP (2005) Antiquity of microRNAs and their targets in land plants. The Plant Cell 17(6): 1658-1673

Bowman JL, Axtell MJ (2008) Evolution of plant microRNAs and their targets. Trends Plant Sci 13(7): 343-349

Axtell MJ (2013) Classification and comparison of small RNAs from plants. Annual Rev Plant Biol 64:137-59.

Barakat A, Wall PK, DiLoreto S, W dePamphilis C, Carlson JE (2007) Conservation and divergence of microRNAs in Populus. BMC genomics 8(1): 481

Barcaccia G, Arzenton F, Sharbel TF, Varotto S, Parrini P, Lucchin M (2006) Genetic diversity and reproductive biology in ecotypes of the facultative apomict Hypericum perforatum L. Heredity 96(4): 322-334

Barcaccia G, Baumlein H, Sharbel TF (2007) Apomixis in St. John's wort: An overview and glimpse towards the future. In: Apomixis: Evolution, Mechanisms and Perspectives (Hörandle E., Grossniklaus U., Van Dijk P., Sharbel T.F eds.) Chapter XIV. Intl. Association of Plant Taxonomy - Koeltz Scientific Books, Vienna (Austria), pp. 259-280

Barcaccia G, Albertini E, (2013) Apomixis in plant reproduction: a novel perspective on an old dilemma. Plant Reprod doi:10.1007/s00497-013-0222-y

Bicknell RA, Koltunow AM (2004) Understanding apomixis: recent advances and remaining conundrums. Plant Cell 16: S228-S245.

Botton A, Galla G, Conesa A, Bachem C, Ramina A, Barcaccia G (2008) Large-scale Gene Ontology analysis of plant transcriptome-derived sequences retrieved by AFLP technology. BMC genomics 9(1): 347

Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O (2008) Widespread translational inhibition by plant miRNAs and siRNAs. Sci Signal 320(5880): 1185

Buckley YM, Briese DT, Rees M (2003) Demography and management of the invasive plant species Hypericum perforatum. I. Using multi-level mixed-effects models for characterizing growth, survival and fecundity in a long-term data set. J Appl Ecology 40(3): 481-493

Chambers C, Shuai B (2009) Profiling microRNA expression in Arabidopsis pollen using microRNA array and Real-Time PCR. BMC Plant Biology 9: 87

Chen X (2005) MicroRNA biogenesis and function in plants. FEBS Letters 579(26): 5923-5931

Chen X (2004) A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. Sci Signal 303(5666): 2022

Chitwood DH, Nogueira FT, Howell MD, Montgomery TA, Carrington JC,

Timmermans MC (2009) Pattern formation via small RNA mobility. Genes Dev 23(5): 549-554

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21(18): 3674-3676

Cuperus JT, Fahlgren N, Carrington JC (2011) Evolution and functional diversification of MIRNA genes. The Plant Cell 23(2): 431-442

Evan GI, Vousden KH (2001) Proliferation, cell cycle and apoptosis in cancer. Nature 411(6835): 342-348

Debernardi JM, Rodriguez RE, Mecchia MA, Palatnik JF (2012) Functional specialization of the plant miR396 regulatory network through distinct microRNA-target interactions. PLoS Genet 8(1): e1002419

Faller M, Guo F (2008) MicroRNA biogenesis: there's more than one way to skin a cat. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms 1779(11): 663-667

Galla G, Barcaccia G, Ramina A, Collani S, Alagna F, Baldoni L, Cultrera NGM, Martinelli F, Sebastiani L, Tonutti P (2009) Computational annotation of genes differentially expressed along olive fruit development. BMC Plant Biology 9: 128-144

Galla G, Barcaccia G, Schallau A, Puente Molins M, Baoumlein H, Sharbel TF (2011) The cytohistological basis of apospory in Hypericum perforatum L. Sex Plant Reprod 24(1): 47-61

Galla G, Sharbel TF, Barcaccia G (2012) De novo sequencing and annotation of the Hypericum perforatum flower transcriptome. Proc of the 56th Italian Society of Agricultural Genetics, 17-20 September, 2012, Perugia (Italy), p. 1.09

Gandikota M, Birkenbihl RP, Höhmann S, Cardon GH, Saedler H, Huijser P (2007)

The miRNA156/157 recognition element in the 3′ UTR of the Arabidopsis SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. The Plant J 49(4): 683-693

Gao Z, Shi T, Luo X, Zhang Z, Zhuang W, Wang L (2012) High-throughput sequencing of small RNAs and analysis of differentially expressed microRNAs associated with pistil development in Japanese apricot. BMC Genomics 13(1): 371

Guilfoyle TJ, Hagen G (2007) Auxin response factors. Curr Opin Plant Biol10(5): 453-60

Hörandl E and Hadacek F (2013) The oxidative damage initiation hypothesis for meiosis. Plant Reprod (special issue on Apomixis, in press).

Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAs and their regulatory roles in plants. Annu Rev Plant Biol 57: 19-53

Jones-Rhoades MW (2012) Conservation and divergence in plant microRNAs. Plant Mol Biol 80(1): 3-16

Knauer S, Holt AL, Rubio-Somoza I, Tucker EJ, Hinze A, Pisch M, Javelle M, Timmermans MC, Tucker MR, Laux T (2013). A protodermal miR394 signal defines a region of stem cell competence in the Arabidopsis shoot meristem. Developmental Cell 24: 125-132

Koch MA, Scheriau C, Betzin A, Hohmann N, Sharbel TF (2013) Evolution of cryptic gene pools in Hypericum perforatum: the influence of reproductive system and gene flow. Ann Bot 111(6): 1083-1094

Koltunow AM, Johnson SD, Lynch M, Yoshihara T, Costantino P (2001) Expression of rolB in apomictic Hieracium piloselloides Vill. causes ectopic meristems in planta and changes in ovule formation, where apomixis initiates at higher frequency. Planta 214(2):196-205

Koltunow AM, Grossniklaus U (2003) Apomixis: a developmental perspective. Annu Rev Plant Biol 54(1): 547-574

Koltunow AMG, Ozias-Akins P, Siddiqi I (2013) Apomixis. In: Becraft PW (ed) Seed Genomics, Wiley-Blackwell, Oxford, UK, pp 83-110.

Lee RC, Feinbaum RL, Ambros V (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75(5): 843-854

Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120(1): 15-20

Liu D, Song Y, Chen Z, Yu D (2009) Ectopic expression of miR396 suppresses GRF target gene expression and alters leaf growth in Arabidopsis. Physiol Plant 136(2): 223-236

Llave C, Kasschau KD, Rector MA, Carrington JC (2002) Endogenous and silencing-associated small RNAs in plants. The Plant Cell 14(7): 1605-1619

Lynam-Lennon N, Maher SG, Reynolds JV (2009) The roles of microRNA in cancer and apoptosis. Biol Rev Camb Philos Soc 84(1): 55-71

Lu C, Tej SS, Luo S, Haudenschild CD, Meyers BC, Green PJ (2005) Elucidation of the small RNA component of the transcriptome. Science 309(5740): 1567-1569

Mallory AC, Vaucheret H (2006) Functions of microRNAs and related small RNAs in plants. Nature Genetics 38: S31-S36

Matzk F, Meister A, Brutovska R, Schubert I (2001) Reconstruction of reproductive diversity in Hypericum perforatum L. opens novel strategies to manage apomixis. The Plant J 26(3): 275-282

Matzk F, Hammer K, Schubert I (2003) Coevolution of apomixis and genome size within the genus Hypericum. Sex Plant Reprod 16: 51-58

Moxon S, Jing R, Szittya G, Schwach F, Pilcher RLR, Moulton V, Dalmay T (2008) Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. Genome Res 18(10): 1602-1609

Naumova T (1992) Apomixis in Angiosperms: nucellar and integumentary embryony. CRC Press, Boca Raton, FL, USA.

Nakagawa H, Ferrario S, Angenent GC, Kobayashi A, Takatsuji H (2004) The petunia ortholog of Arabidopsis SUPERMAN plays a distinct role in floral organ morphogenesis. Plant Cell 16(4): 920-932

Nogler GA (1984) Gametophytic apomixis. In: Johri BM (ed) Embryology of angiosperms. Springer, Berlin Heidelberg New York, pp. 475-518.

Nonomura KI, Morohoshi A, Nakano M, Eiguchi M, Miyao A, Hirochika H, Kurata N (2007) A germ cell-specific gene of the ARGONAUTE family is essential for the progression of premeiotic mitosis and meiosis during sporogenesis in rice. The Plant Cell 19(8): 2583-2594

Nozawa M, Miura S, Nei M (2012) Origins and evolution of microRNA genes in plant species. Genome Biol Evol 4(3): 230-239

Nürk NM, Madriñán S, Carine MA, Chase MW, Blattner FR (2012) Molecular phylogenetics and morphological evolution of St. John's wort (Hypericum; Hypericaceae). Mol Phylogenet Evol 66(1): 1-16.

Olmedo-Monfil V, Durán-Figueroa N, Arteaga-Vázquez M, Demesa-Arévalo E, Autran D, Grimanelli D, Slotkin RK, Martienssen, RA, Vielle-Calzada JP (2010) Control of female gamete formation by a small RNA pathway in Arabidopsis. Nature 464(7288): 628-632

Ozias-Akins P (2006) Apomixis: Developmental characteristics and genetics. Crit. Rev. Plant Sci 25: 199-214.

Pantaleo V, Szittya G, Moxon S, Miozzi L, Moulton V, Dalmay T, Burgyan J (2010)

Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis. The Plant J 62(6): 960-976

Palatnik JF, Wollmann H, Schommer C, Schwab R, Boisbouvier J, Rodriguez R, Warthmann N, Allen E, Dezulian T, Huson D, Carrington JT, Weigel D (2007) Sequence and expression differences underlie functional specialization of Arabidopsis microRNAs miR159 and miR319. Dev Cell 13(1): 115-125

Pagnussat GC, Alandete-Saez M, Bowman JL, Sundaresan V (2009) Auxin-dependent patterning and gamete specification in the Arabidopsis female gametophyte. Science 324(5935): 1684-1689

Pani A, Mahapatra RK, Behera N, Naik PK (2011) Computational identification of sweet Wormwood (Artemisia annua) microRNA and their mRNA targets. Genomics Proteomics and Bioinformatics 9(6): 200-210

Pfeffer S, Zavolan M, Grasser FA, Chien M, Russo JJ, Ju J, John B, Enright AJ, Marks D, Sander C, Tuschl T (2004) Identification of virus-encoded microRNAs. Sci Signal 304(5671): 734

Polegri L, Calderini O, Arcioni S, Pupilli F (2010) Specific expression of apomixis-linked alleles revealed by comparative transcriptomic analysis of sexual and apomictic Paspalum simplex Morong flowers. J Exp Bot 61(6): 1869-1883

Pupilli F, Barcaccia G (2012) Cloning plants by seeds: Inheritance models and candidate genes to increase fundamental knowledge for engineering apomixis in sexual crops. J Biotechnol 159(4): 291-311

Rajagopalan R, Vaucheret H, Trejo J, Bartel DP (2006) A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. Genes Dev 20(24): 3407-3425

Silvertown J (2008) The evolutionary maintenance of sexual reproduction: evidence from the ecological distribution of asexual reproduction in clonal plants. Intl J Plant Sci 169: 157-168.

Siomi H, Siomi MC (2010) Posttranscriptional regulation of miRNA biogenesis in animals. Mol Cell 38: 323-332

Robson, NK (2002) Studies in the genus Hypericum L.(Guttiferae) 4 (2). Section 9. Hypericum sensu lato (part 2): subsection 1. Hypericum series 1. Hypericum. Bulletin of the Natural History Museum Botany, 32(2): 61-123

Rodriguez RE, Mecchia MA, Debernardi JM, Schommer C, Weigel D, Palatnik JF (2010) Control of cell proliferation in Arabidopsis thaliana by microRNA miR396. Development 137(1): 103-112

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19(18): 2496-2497

Schallau A, Arzenton F, Johnston AJ, Hahnel U, Koszegi D, Blattner FR, Altschmied L, Haberer G, Barcaccia G, Baümlein H (2010) Identification and genetic analysis of the APOSPORY locus in Hypericum perforatum L. Plant J 62: 772-784

Silvertown J (2008) The evolutionary maintenance of sexual reproduction: evidence from the ecological distribution of asexual reproduction in clonal plants. Intl J Plant Sci 169(1): 157-168

Singh M, Goel S, Meeley RB, Dantec C, Parrinello H, Michaud C, Leblanc O, Grimanelli D (2011) Production of viable gametes without meiosis in maize deficient for an ARGONAUTE protein. The Plant Cell 23(2): 443-458

Smyth DR, Bowman JL, Meyerowitz EM (1990) Early flower development in Arabidopsis. The Plant Cell Online 2(8): 755-767

Song L, Han MH, Lesicka J, Fedoroff N (2007) Arabidopsis primary microRNA processing proteins HYL1 and DCL1 define a nuclear body distinct from the Cajal body. Proc Natl Acad Sci USA 104(13): 5437-5442

Song QX, Liu YF, Hu XY, Zhang WK, Ma B, Chen SY, Zhang JS (2011) Identification of

miRNAs and their target genes in developing soybean seeds by deep sequencing. BMC Plant Biology 11(1): 5

Spillane C, Curtis MD, Grossniklaus U (2004) Apomixis technology development - virgin births in farmers' fields? Nat Biotech 22: 687-691.

Sunkar R, Chinnusamy V, Zhu J, Zhu JK (2007) Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. Trends Plant Sci 12(7): 301-309

Sunkar R, Zhu JK (2004) Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. Sci Signal 16(8): 2001

Sun G (2012) MicroRNAs and their diverse functions in plants. Plant Mol Biol 80(1): 17-36

Van Ex F, Jacob Y, Martienssen RA (2011) Multiple roles for small RNAs during plant reproduction. Curr Opin Plant Biol 14(5): 588-593

Tucker MR, Okada T, Hu Y, Scholefield A, Taylor JM, Koltunow AM (2012) Somatic small RNA pathways promote the mitotic events of megagametogenesis during female reproductive development in Arabidopsis. Development 139(8): 1399-1404

Vielle-Calzada JP, Crane CF, Stelly DM (1996) Apomixis: the asexual revolution. Science 274: 1322-1323.

Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. Cell 136(4): 669-687

Wu MF, Tian Q, Reed JW (2006) Arabidopsis microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. Development 133(21): 4211-4218

Xie K, Wu C, Xiong L (2006) Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and

microRNA156 in rice. Plant Phys 142(1): 280-293

Xie F, Frazier TP, Zhang B (2010) Identification and characterization of microRNAs and their targets in the bioenergy plant switchgrass (Panicum virgatum). Planta 232(2): 417-434

Yamada-Akiyama H, Akiyama Y, Ebina M, Xu Q, Tsuruta S, Yazaki J, Kishimoto N, Kikuchi S, Takahara M, Takamizo T, Sugita S, Nakagawa H (2009) Analysis of expressed sequence tags in apomictic guineagrass (Panicum maximum). J Plant Physiol 166(7): 750-761

Yu HP, Song CN, Jia QD, Wang C, Li F, Nicholas KK, Zhang XY, Fang JG (2011) Computational identification of microRNAs in apple expressed sequence tags and validation of their precise sequences by miR-RACE. Physiol. Plant 141(1): 56-70

Zhang B, Pan X, Cannon CH, Cobb GP, Anderson TA (2006a) Conservation and divergence of plant microRNA genes. The Plant J 46(2): 243-259

Zhang BH, Pan XP, Cox SB, Cobb GP, Anderson TA (2006b) Evidence that miRNAs are different from other RNAs. Cell Mol Life Sci 63(2): 246-254

Zhang B, Pan X, Stellwag EJ (2008) Identification of soybean microRNAs and their targets. Planta 229(1): 161-182

Zhang L, Chia JM, Kumari S, Stein JC, Liu Z, Narechania A, Maher CA, Guill K, McMullen MD, Ware D (2009) A genome-wide characterization of microRNA genes in maize. PLoS Genetics 5(11): e1000716

Zhang Z, et al. (2010) PMRD: plant microRNA database. Nucleic Acids Res 38(Suppl 1): D806-D813

Zhao CZ, Xia H, Frazier T, Yao YY, Bi YP, Li AQ, Li MJ, Li CS, Zhang BH, Wang XJ (2010) Deep sequencing identifies novel and conserved microRNAs in peanuts (Arachis hypogaea L.). BMC Plant Biology 10(1): 3

Zhu QH, Spriggs A, Matthew L, Fan L, Kennedy G, Gubler F, Helliwell C (2008) A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains. Genome Res 18(9): 1456-1465

Zhu H, Xia R, Zhao B, An YQ, Dardick CD, Callahan AM, Liu Z (2012) Unique expression, processing regulation, and regulatory network of peach (Prunus persica) miRNAs. BMC Plant Biology 12(1): 149

Zouhar K (2004) Hypericum perforatum. In: Fire Effects Information System [Online]. U.S. Department of Agriculture, Forest Service. Rocky Mountain Research Station, Fire Sciences Laboratory (Producer). Available at http://www.fs.fed.us/database/feis/

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31(13): 3406-3415

8 SUPPLEMENTARY

```
156 Consensus     UGACAGAAGAGAGMGAGCAYA
                  |||||||||| ||:||||||:
Hyp-miR156a       UUGACAGAAGAUAGAGAGCAC

156 Consensus     UGACAGAAGAGAGMGAGCAYA
                  |||||||||||||:|||||
Hyp-miR156b       UUGACAGAAGAGAGAGAGCAA

166 Consensus     YCGGACCAGGCUUCAUYCCMCC
                  :|||||||||||||||:||:|
Hyp-miR166        UCGGACCAGGCUUCAUUCCCC

390 Consensus     AAGCUCRGGAGGGAUAKCACC
                  ||||||:|||||||  |:|||||
Hyp-miR390        AAGCUCAGGAGGGAGAGCACC

394 Consensus     UUDGGCAUUCUGUCMACCUCCA
                   |:||||||||||||:||||||
Hyp-miR394         UUGGCAUUCUGUCCACCUCCCU

396 Consensus     UUUCCACABGCUUUCUUGAACKG
                   ||||||| ||||||||||||:
Hyp-miR396         UUCCACA-GCUUUCUUGAACU

414 Consensus     UCAUCVUCAUCAUCMUCDUCY
                  || ||:|||||| ||:||:||
Hyp-miR414        UCCUCCUCAUCCUCCUCGUC
```

Figure S1. Comparison of the consensus sequences with the sequences from the mature miRNAs which demonstrates substantial homology and confirms the reliability of the method used for *a posteriori* prediction of the mature sequences.

Table S1. List of forward and reverse primers designed for each of the selected pre-miRNA sequences. The name of primer corresponds to the miRNA family.

Primer name	Primer sequence (5'-3')	Length (nt)	Tm (°C)
Hyp-miR156a for	TTGAGAGGGAGAGGGAATTT	20	57.8
Hyp-miR156a rev	TTGAAGGTGATGACAGAAGC	20	56.3
Hyp-miR156b for	CCATGGTGCTAATCTTAATTCAT	23	57.3
Hyp-miR156b rev	CTGGCCATAACGAACGAC	18	57.6
Hyp-miR166 for	GTGTTGTCTGGCTCGAGGA	19	60.0
Hyp-miR166 rev	GAATAACGAGGATGCGAGAA	20	57.9
Hyp-miR390 for	AAAAAGCATGGAAGAACCTA	20	54.2
Hyp-miR390 rev	CGAAAAGAGTGAGATGAGCA	20	56.7
Hyp-miR394 for	TTTTTCAAAGAGTTTCTAACAGAGTTT	27	57.7
Hyp-miR394 rev	CCAAGAGGGTTTTACAAAGAGA	22	57.6
Hyp-miR396 for	TGATATTCCACAGCTTTCTTGA	22	57.5
Hyp-miR396 rev	AAGAAAGCAAAATCCTGAAATG	22	57.6
Hyp-miR414 for	AGAGTACAGGGAAATGGAGGA	21	57.7
Hyp-miR414 rev	CACAGCGAAACCCACGAG	18	61.5
Hyp-hgPP2A for	TCATCTGATTCTTCCCAGCA	20	63.4
Hyp-hgPP2A rev	GAAAGTAAATCAATCCCAATGACC	24	63.2

Ringraziamenti

Desidero ringraziare primariamente il Prof. Gianni Barcaccia per avermi permesso di svolgere il tirocinio nel suo laboratorio e per avermi così permesso di portare a termine il mio percorso di studi. Desidero inoltre ringraziare il Dr Giulio Galla per tutti gli insegnamenti , i consigli e le idee che mi hanno permesso di terminare con successo questo lavoro di tesi, e per il sostegno e la pazienza che ha avuto durante questo anno di lavoro.

Desidero ringraziare la mia famiglia, mamma, papà e mia sorella Alessia, per avermi sostenuto e essermi stato vicino ogni momento durante questi 5 anni di università.

Voglio infine ringraziare Vanessa per avermi sostenuto, aiutato ma soprattutto sopportato in quest'ultimo periodo pre-laurea.