1. INTRODUCTION

1.1 Background

MicroRNAs (miRNAs) are a class of 21–24 nucleotide endogenous noncoding small RNAs, transcribed from MIR genes, that act as negative regulators of gene expression in eukaryotes. Plant miRNAs negatively regulate endogenous target genes whose transcripts are recognized based on their close or complete complementarity. This regulation occurs through cleavage or translational suppression of the target transcript (Mlotshwa et al., 2006).

miRNAs play critical roles in development and environmental responses of plants, including regulation of leaf and flower morphology, floral organ identity and flowering time. Plant miRNAs are also engaged in self-regulation of the miRNA pathway (Yang et al., 2007). Many plant miRNAs have perfect or near-perfect complementarity to their target transcripts and usually have only one single site inside their target mRNA (Rhoades et al., 2002). Plant miRNAs arise from the genome, are distinct from other annotated genes and are evolutionary conserved (Reinhart et al., 2002; Bartel, 2004). Up to date, many miRNA genes, either conserved or tissue-specific have been identified in various organisms especially those with completed genome sequences. Recently miRbase (release 15.0), the registry that coordinates naming of miRNAs released 14,197 entries representing hairpin precursor miRNAs, expressing 15,632 mature miRNA products, in 133 species (miRbase, 2010).

In plants, miR172 is one of the special miRNAs due to its ability to use both methods of target regulation. Like most of the animal miRNAs, miR172, downregulates its target genes mainly by translational repression rather than cleavage of the target gene, although cleavage products as by-products of overlapping between the two regulation pathways also is possible (Aukerman and Sakai, 2003; Chen, 2004).

Studies have demonstrated that miR172 family is engaged in flowering time control (Aukerman and Sakai, 2003; Chen, 2004; Mlotshwa et al., 2006) and developmental phase transition. As it was confirmed in *Arabidopsis thaliana*, overexpression of miR172 caused early flowering and some disordering in floral organ identity, while overexpression of miR172 in *Zea mays* leads to a delayed phase change from the vegetative to the reproductive stage (Lauter et al., 2005).

The targets of miR172 are members of the APETALA2 (AP2) transcription factor gene family such as AP2, TOE1, TOE2, TOE3, SNZ and SMZ which are known in *Arabidopsis thaliana* and glossy15, Indeterminate spikelet1 and Sister of Indeterminate spikelet1 which are identified in *Zea mays* (Park et al., 2002; Aukerman and Sakai, 2003; Chuck et al., 2008).

The discoveries of miRNAs in plants and the growing evidence of their involvement in a variety of functional roles have produced a great deal of excitement in plant biology (Yang et al., 2007). Among miRNA registry entries (release 15.0), 27 species belonged to eudicot plants and just 6 species belonged to monocot plants, while this number increased up to 75 species for animals (miRBase, 2010). Hence still more effort is needed to identify novel miRNAs and predict their targets in other monocot and dicot plants.

Oil palm (*Elaeis guineensis*) is a monocot and perennial crop and belongs to the *Elaeis* genus. Oil palm is a distinctive crop because it produces two types of oil: kernel oil which is valued usually in the oleochemical industry and the mesocarp oil which is mostly used for its edible assets. About 90% of Palm oil is used for edible products, while the rest is used for non-edible purposes (Rival, 2007). Malaysia is one of the largest palm oil producers and exporters in the world. Malaysia and Indonesia together contirbute 85% of the crude palm oil production in the world. It was reported by the Malaysian Palm Oil Board (MPOB) that Malaysian palm oil export was 1.21 million

tonnes in the month of December 2009. In addition to oils and fats production, there is also a special attention regarding oil palm for renewable energy (Sumathi et al., 2008).

The main difficulty towards miRNA and its target gene study in oil palm is due to its unestablished genome in any public database, however there are currently promising projects towards genomic and transcriptome aspects of oil palm by MPOB, Sime Darby and Asiatic Centre for Genome Technology Sdn Bhd (ACGT). In the near future, their information disclosure would improve oil palm projects in wider aspects.

1.2 Importance of the Study

Although several EST of oil palm are available in the National Centre for Biotechnology Information (NCBI) database, so far, few miRNA and no confirmed targets have been reported in any of its species. Since oil palm is an important plantation crop, this study aimed to identify and characterize miR172 and its potential targets, using both computational and molecular methods.

As the oil palm fruit is the source of the main economic output for this crop, by identifying and isolating the miR172 and its potential targets, new insights associated with floral development which is a critical process related to productivity of this crop might be provided.

1.3 Research Objectives

This study aims to:

(i) Identify microRNA172 (miR172) gene sequences from oil palm (*Elaeis guineensis Jacq*).

(ii) Determine the expression pattern of miR172 in different developmental stages of oil palm.

(iii) Identify potential transcripts of miR172 target genes in various tissues from oil palm.

(iv) Determine the expression pattern of candidate target mRNA in different developmental stages in oil palm.

2. LITERATURE REVIEW

2.1 Oil Palm

Oil palm is one of the most important perennial and long-lived crops, sometimes living more than 100 years (Dransfield and Uhl, 1998). A typical oil palm tree growing in a plantation in Malaysia is shown in Figure 2.1. The oil palm has a single stem and, as is common among the majority of palm species, bears a single vegetative shoot apical meristem (SAM). The SAM is located at the center of the leaf crown, is sustained through the lifetime of the plant and under good climatic conditions, it constantly will be active. Oil palm is widely grown in tropical areas and after soya is the most important world vegetable oil producing plant. Even though it has economical importance, only limited studies of its inflorescence and floral development, have been reported (Adam et al., 2005).

2.1.1 Scientific Classification of Oil Palm

Palms are woody monoecious plants which belong to the order *Arecales* and the family *Arecaceae* (*Palmae*) (Price et al., 2007; Jones, 1995). The oil palm is a member of the *Arecoideae* subfamily, *Cocoeae* tribe and *Elaeidinae* subtribe (Corley and Tinker, 2003). The subtribe *Elaeidinae* consists of two genera, i.e. *Elaeis* and *Barcella* (Hahn, 2002). Currently, the genus *Barcella* does not have any commercial application. The genus *Elaeis* has two species; the African oil palm, *E. guineensis Jacq*., and the Latin American oil palm, *E. oleifera Cortez* (Corley and Tinker, 2003).

Figure 2.1 Oil Palm Plant (MPOC, 2010)

2.1.2 *Elaeis guineensis*

E. guineensis Jacq. has no subspecies and contains 16 pairs of chromosomes (Madon and Clyde, 1995). On the basis of the length of *E. guineensis* chromosomes, three groups were created (Madon and Clyde 1995). The estimated size of the haploid *E. guineensis* genome is 3.4×10^9 bp (Price et al., 2007).

2.1.3 Flowers

Without any doubt, studies on inflorescence and flower development are key factors for finding out the evolutionary relationships in the palm family and considering the other monocots, even in angiosperms as a whole (Adam et al., 2005). Moreover, oil palm is important because of its capability for producing two types of oils (palm kernel oil and mesocarp oil) and also other by-products.

It takes almost 2 to 3 years until unfolded leaflets at the middle of the oil palm crown are developed from initial leaves. Separate female and male flowers emerge in the axils of vegetative leaves. The crown protects the immature inflorescence. Furthermore, the growing axes of the inflorescence are protected by a thick and woody pprophyll and a peduncular bract (Figure 2.2) (Adam et al., 2005).

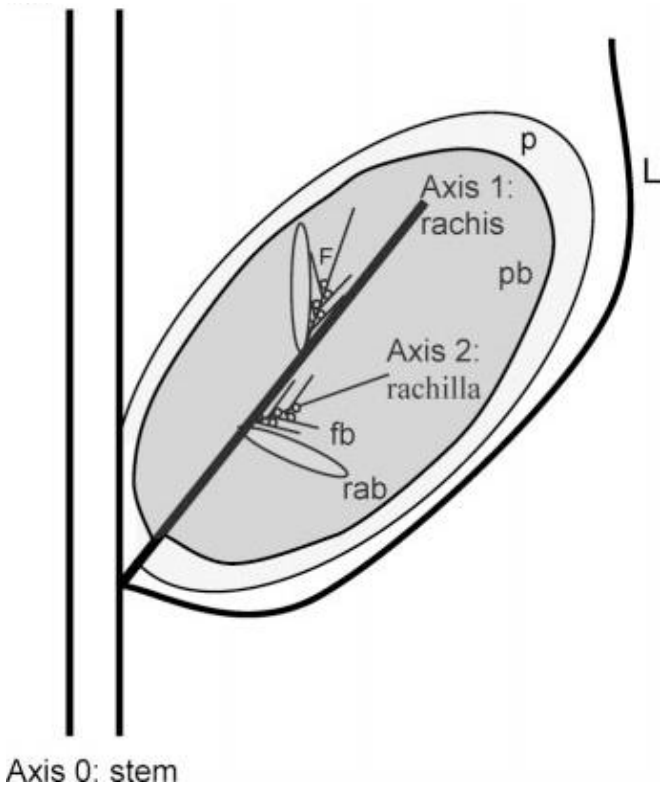


Figure 2.2 Features of Reproductive Development in Oil Palm. (A) Oil palm inflorescence structure and axis organization. The inflorescence unit axis (axis1) is in the axil of a leaf. *Abbreviations:* F, flower; fb, flower bract; L, leaf; p, prophyll; pb, peduncular bract; rab, rachilla bract. (Reproduced from Adam et al., 2005)

It takes approximately three years for mature male or female inflorescences to develop while just about two inflorescences are initiated monthly during this time. The oil palm alternatively produces male and female flowers in a period of about 6 months, so it is an outcrossing plant. Studies on the inflorescences demonstrated that the primordium of the flower is capable of producing both male and female organs, while one of them mostly remains undeveloped (Hartley, 1988). The male inflorescence is made of rachillae with 400-1,500 staminate flowers on each of them (Figure 2.3 C-D). The female inflorescence is an indeterminate cluster that includes a different number of rachillae which carry 5-30 floral triads, each of them made of a pistillate flower (female inflorescence) joined with two nonfunctional (staminate flowers) male inflorescences (Adam et al., 2005; Price et al., 2007). Each oil palm seed arises from only one carpel while each female flower contains a tricarpellate ovary (Price et al., 2007). Figure 2.3 A-B, shows the details of female inflorescences. The final development of the flower occurs in the pollen sac with the formation of microsporocytes after this no further development happens and prior to maturation of pistillate flowers, abscission occurs. In the perianth, the petaloid appearance of sepals and petals, especially in pistillate flowers, are similar. In some rare cases at the transition between the male and female cycles, mixed sex inflorescences are produced (Biradar, 1978). The sex of the inflorescence is affected by age, genetic factors, and especially external conditions (Corley, 1976). The cycle transition toward male inflorescences can be influenced by negative environmental conditions such as water stress while favorable environmental situation may affect female inflorescence cycle transition (Price et al., 2007).

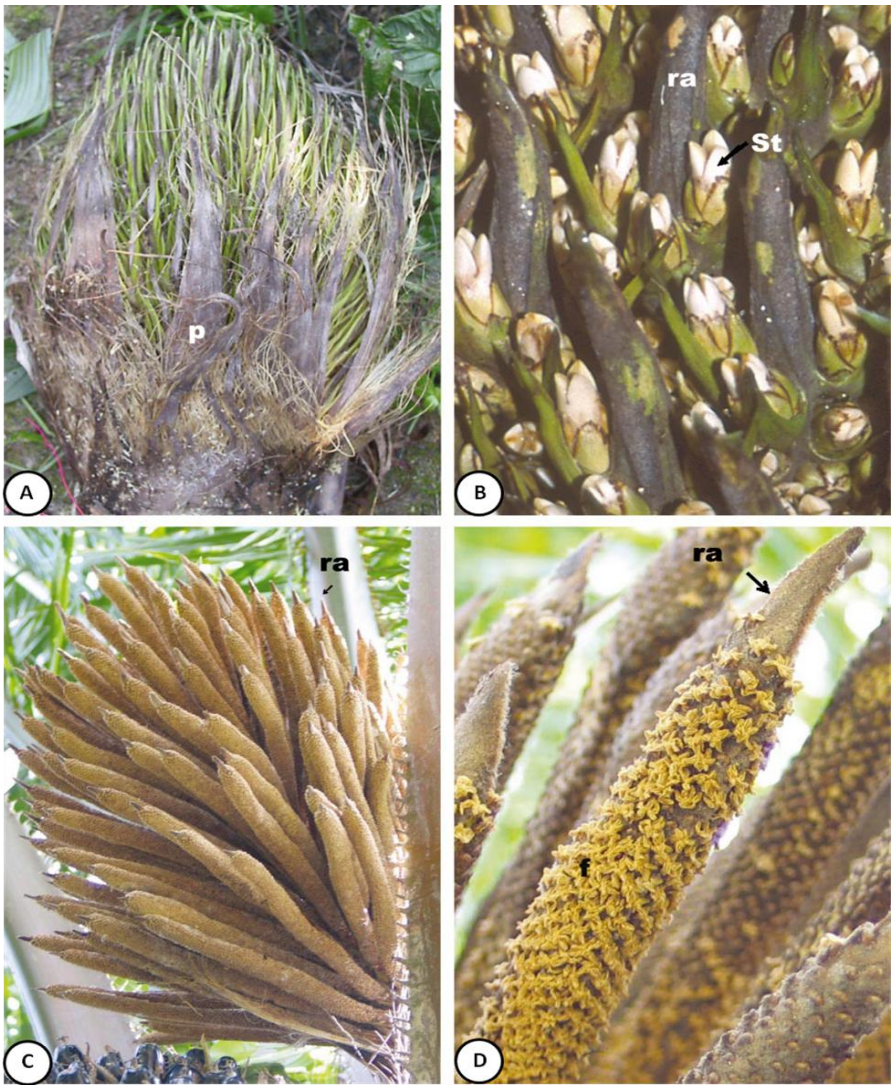


Figure 2.3 Macroscopic View of Mature Female and Male Inflorescences of Oil Palm. (A) Female inflorescence at anthesis after opening of the peduncular bract and prophyll. (B) Detail of female rachillae carrying pistillate flowers at maturity with exposed stigmas. (C) Male inflorescence at anthesis. (D) Detail of male rachilla and flowers at anthesis. *Abbreviations:* f, flower; p, prophyll; ra, rachilla; St, stigma. (Reproduced from Adam et al., 2005)

In palm species, through a single shoot, two major and distinct forms of flowering behavior may be observed (Tomlinson, 1990). One of them, that happens in about 5% of all palm species is Hapaxanthy, which refers to a single shoot wherein, there is a quick floral transition. In this type, after the vegetative phase has been stopped, floral branches (inflorescences) grow in larger numbers. Therefore, the flowering phase is short.

The other form, Pleonanthy, which occurs in about 95% of palms, refers to a single shoot in which inflorescences emerge in the axils of vegetative leaves and as the palm continues vegetative development, the inflorescences keep on being produced. In this type the flowering phase is extensive and undetermined. The oil palm belongs to the second type. In early stages, soon after germination of oil palm, the inflorescences are initiated in the axil of the leaves and this is maintained throughout the lifetime of the plant. The expansion of the inflorescences happens over a two year period (Adam et al., 2005).

2.1.4 Fruit

In fruits, seeds (kernel) and the pulp (mesocarp) are very rich in oil (Figure 2.4). The most important differences inside the fruit arise from the thickness of the shell (endocarp).

Dura, Pisifer and Tenera are three fruit forms that were found in 1941 to be due to a single gene (Price et al., 2007; Beinaret and Vanderweyen, 1941). This gene is important, since only intermediate type plants, *Tenera*, are developed for commercial usage. *Dura* with the genotype of *ShSh*, has a thick shell (endocarp) while *Pisifera* of genotype *shsh,* just has a fibrous ring (mesocarp) and no endocarp is developed. In *Tenera,* the intermediate form, with *Shsh* genotype, the endocarp is thinner than *Dura*, while it also has the fibrous ring. Although the endocarp thickness differs noticeably in

the *Dura* and *Tenera* fruits, sometimes they overlap, thus the crucial factor for differentiating these two fruit forms is due to the fibrous ring that is present in the *Tenera*. Since the crude palm oil (CPO) is delivered from this fibrous ring (mesocarp), it is essential in the palm oil industry.

Initially it was supposed that oil palm was wind pollinated, because of its plentiful pollen and reduced flower form (Price et al., 2007). The *Elaiedobius kamerunicu*, a pollinating insect which was introduced from Africa into Malaysia in 1981 (Corley and Tinker, 2003), proved that insect pollination essentially is engaged in the fruit formation. This can mainly happen in wet circumstances (Hartley, 1988).

Figure 2.4 Oil Palm Fruit (USDA, 2009)

2.1.5 Economic Importance of Oil Palm

In many south east Asian developing countries, especially in Malaysia and Indonesia, oil palm has a critical role in the economy. The main importance of the oil palm is because of its fruit pulp (mesocarp) oil and seed kernel oil. These two oils have significant differences in their fatty acid composition (Corley and Tinker, 2003). Between 40-70% of the palm oil content is derived from mesocarp. The major oil palm oil is extracted from palm kernel (seed). The content of oil varies from 43–51% per seed (Price et al., 2007). Compared with all other oil crops, oil palm is the most productive and has the most flexibility in usage compared to other vegetable oils (Edem 2002). Palm oil is a main competitor of soybean oil (USDA-FAS, 2006) and it has become a main source for edible oil in the world. Around 90% of palm oil is for edible usage, while the rest is applied for nonfood applications like oleochemical and soap production (Price et al., 2007). As reported by Global Oils & Fats Business Magazine (GOFBM), 2009, Malaysia produced 17.7 million tonnes of palm oil on 4.5 million hectares of land in 2008 (Malaysian Palm Oil Industry Performance, 2008). After Indonesia, Malaysia is the largest exporter of palm oil in the world (Salleh et al., 2010). Oleochemicals such as soaps and glycerol are used for non-edible application. In addition, diesel fuel can be replaced by oil palm fatty acid methyl esters, as a renewable energy source which is called Biodiesel (Pioch and Vaitilingom, 2005; Price et al., 2007). It has been suggested that palm biodiesel is the most feasible in comparison with other vegetable biodiesels such as rapeseed, soy and brassica (Lin et al., 2006). Other waste products which result from oil processing including empty fruit bunches and Palm oil mill effluent (POME) are used in oil palm plantations as organic fertilizer, and palm kernel shell and fiber are applied as fuel and make palm oil mills energy self sufficiency (Yusoff, 2006). Besides these, oil palm can also be used as a source for producing Biodegradable plastics such as polyhydroxybutyrate (BHP) (Masani Mat

Yunus et al., 2001). Figure 2.5 illustrates different usages of various parts of oil palm plants.

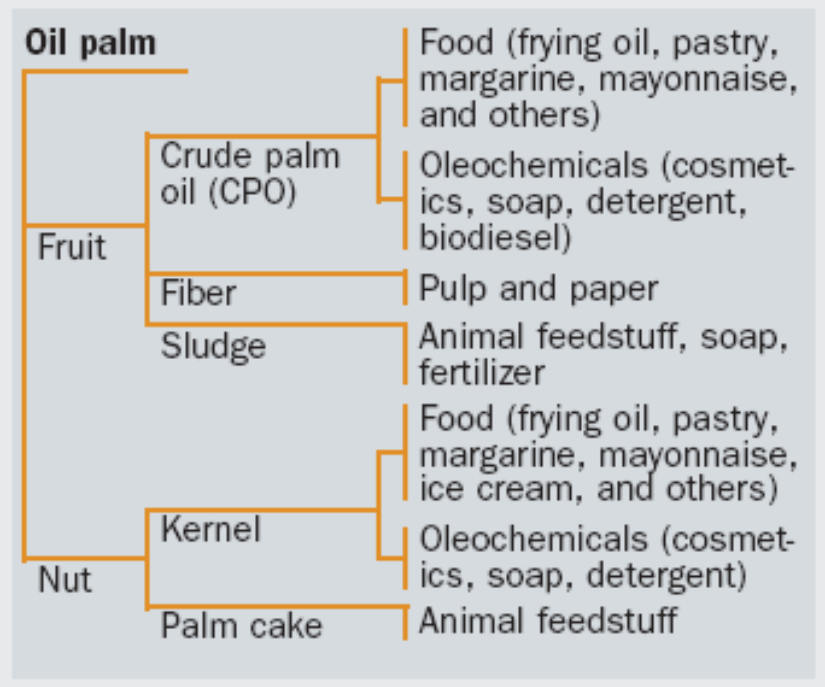


Figure 2.5 Uses of Oil Palm (Fassler, 2006)

2.2 Introduction to microRNAs

MicroRNAs (miRNAs) are a class of small non-coding RNAs (sRNAs) which are 20-22 nucleotides in length (Chellappan et al., 2010). miRNAs are fundamental, sequence-specific regulatory elements (Voinnet, 2009), encoded in the genome of plants, animals, algae, some other unicellular organisms, and many DNA viruses (Fahlgren et al., 2010). miRNAs regulate gene expression at the post-transcriptional level in a sequence specific manner. They generally bind to the 3' UTR of their target mRNAs with near perfect complementary and suppress protein production by cleaving the mRNA and/or translational silencing (Cannell et al., 2008). Several of these mRNA targets are transcription factors engaged in developmental control (Axtell and Bowman, 2008). Many miRNA families are conserved across most important lineages of plants, such as mosses, gymnosperms, monocots and dicots (Zhang et al., 2006a), but many are present only in closely related species or are species specific (Fahlgren et al., 2010). The evolutionary conservation in plant miRNAs and their targets goes back over 400 million years, when seed plants (spermatophytes) and bryophytes had their last common ancestor (Li and Mao, 2007).

The regulatory role of miRNA in developmental timing of *C. elegans* was first recognized by Ambros and colleagues in 1993; the *lin-4* and *let-7* miRNAs suppressed the genes engaged in the transition of larval development at different stages (Lee et al., 1993; Reinhart et al., 2000; Willmann and Poethig, 2007). Later in 2001, when many other small RNAs were found in *C. elegans*, the term 'microRNA' widely came to be used (Lau et al., 2001). By 2008 there were over 8,000 miRNAs that had been identified in more than 30 different species (Griffiths-Jones et al., 2008). Several experimental approaches highlighted their incidence even in a unicellular alga besides animals and plants (Zhao et al., 2007). They are involved in various functions such as developmental control (Qiang, et al., 2007), cell proliferation and apoptosis processes

which are important in cancer formation (Zhang et al., 2007a) , stress resistance and fat metabolism (Ambros, 2003), differentiation and viral infection (Wang and Wang, 2006).

2.2.1 Function of microRNAs in Plants

Evidence showed that plant miRNAs regulate their target genes negatively, in a wide range of developmental processes, such as organ morphogenesis and floral differentiation and development (Chen, 2004; Zhang et al., 2006b; Qiang et al., 2007), meristem cell identity (Qiang et al., 2007), boundary formation/organ separation and floral organ identity (Chen, 2005). They are also engaged in differentiation (Rhoades et al., 2002), auxin responses (Wang et al., 2005), stress condition responses, leaf polarity (Li and Mao, 2007), phosphate starvation (Fujii et al., 2005), sex determination and organ polarity (Ding et al., 2009). More studies on miRNAs possibly will enlighten many other roles (Willmann and Poethig, 2007).

Numerous plant miRNA targets are members of regulatory protein families such as transcription factors, implying that miRNAs are master regulators (Jones-Rhoades et al., 2006). These transcription factors which mainly have regulatory roles in plant development, contain a major category of miRNA target genes. As suggested by Chen (2005), miRNAs help direct an extensive range of cell division and cell fate decisions among plants (Chen, 2005).

2.2.2 Biogenesis of miRNAs

Primary microRNAs (pri-miRNAs) that may be over one kb in length are generated by RNA polymerase II (Xie et al., 2005). Pri-miRNA also undergoes common splicing, polyadenylation and capping (Kurihara & Watanabe, 2004; Xie et al., 2005). Later, pri-miRNA are processed to precursor microRNA (pre-miRNA) which are about 60-90bp

(Krol et al., 2004). This pre-miRNA acquires the typical stem-loop structure which contains the mature microRNA in one of its arms.

miRNA transcripts are processed differently in animal and plants. In animals, the cleavage of pri-miRNA into pre-miRNA, in the nucleus, is mediated by the RNase III activities of Drosha (Lee et al., 2003). Later this pre-miRNA is exported to the cytoplasm and processed by Dicer to create miRNA-miRNA* duplexes. miRNA* stands for the antisense strand (Billy et al., 2001; Provost et al., 2002). The nucleocytoplasmic export of the pre-miRNAs is an essential step in miRNA biogenesis (Lee et al., 2002; Lund et al., 2004); while the last cleavage step, by Dicer, results in two nucleotide 3' overhangs on either side of the duplex (Elbashir et al., 2001).

In plants, the pri-miRNA is also processed to create the mature miRNA in two steps, but since plants do not have any Drosha homolog, as an alternative, Dicer Like-1 (DCL1), a homolog to animal Dicer and an RNase III enzyme, is essential for both steps of miRNA maturation (Xie et al., 2004) which occur in the nucleus (Papp et al., 2003; Vazquez et al., 2004). HYL1 which is a dsRNA-binding domain protein partner and SE (Serrate) help to release a miRNA/miRNA* duplex (Chapman and Carrington, 2007).

HYL1 and HEN1 are two other proteins which are needed in biogenesis of plant miRNAs. HEN1contains a dsRNA-binding motif and a C-terminal methyltransferase domain (Chen, 2005) and causes miRNA duplex methylation (Chapman and Carrington, 2007). The changes that are made by HEN1 on miRNA, lead to protection of the miRNA from additional degradation or alteration, or even may smooth the progress of its assembly into the RNA-induced silencing complex (RISC) (Du and Zamore, 2005), which eventually causes the target RNA degradation or repression.

In plants, the resulting miRNA/miRNA* duplex is exported from the nucleus by HASTY which is the plant ortholog of animal Exportin 5. Afterward, in the cytoplasm

the miRNA/miRNA* duplex is assembled into the RNA-induced silencing complex (RISC) (Park et al., 2005). Figure 2.6 shows the biogenesis pathway for plant and animal miRNAs.

miRNAs in plants have a methyl group on the ribose of the last nucleotide while animal miRNAs end with free 2' and 3' hydroxyl groups on the ribose of the last nucleotide (Du and Zamore, 2005).

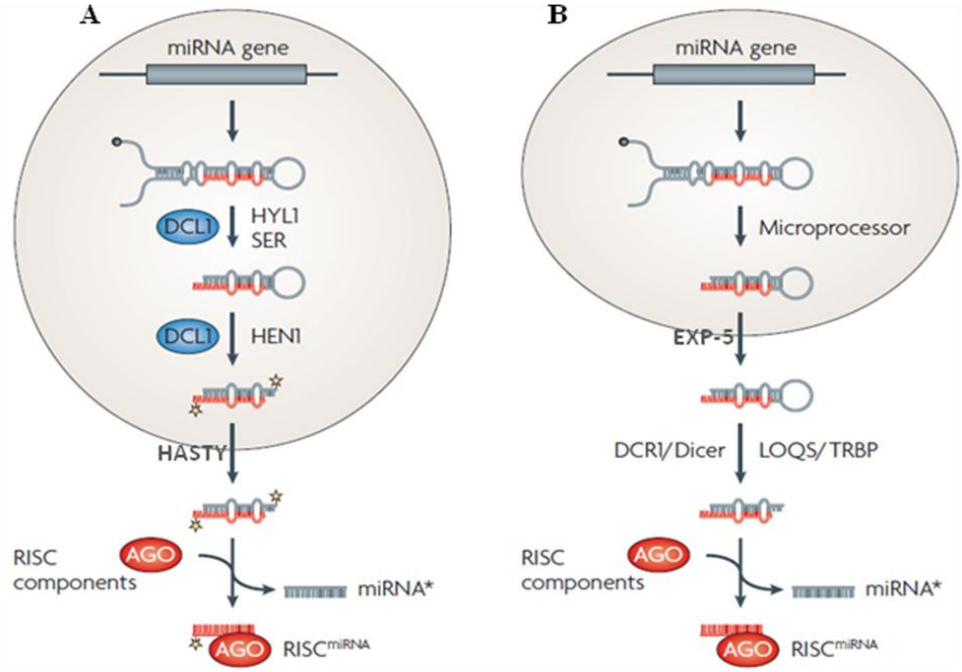


Figure 2.6 miRNA Biogenesis in Plants and Animals. (A) Transcripts from *Arabidopsis thaliana* microRNA (miRNA) genes adopt an imperfect foldback structure, and are processed by DICER-LIKE 1 (DCL1) in concert with HYPONASTIC LEAVES 1 (HYL1) and SERRATE (SE) to release a miRNA–miRNA* (miRNA* represents the antisense strand) duplex. This duplex is methylated by the methyltransferase HUA ENHANCER 1 (HEN1) and exported to the cytoplasm where the mature miRNA strand associates with an Argonaute (AGO) protein, most commonly AGO1, to form the RNA-induced silencing complex (RISC). (B)Transcripts from animal miRNA genes adopt an imperfect foldback structure and undergo sequential processing by the microprocessor complex (Drosha and Pasha in flies; Drosha and DiGeorge syndrome region gene 8 protein (DCGR8) in mammals) to form precursor miRNA (pre-miRNA) intermediates. Pre-miRNAs are exported to the cytoplasm and processed by Dicer in concert with a specialized RNA-binding protein, such as Loquacious (LOQS) in flies or trans-activation responsive RNA-binding protein (TRBP) in humans. (Reproduced from Chapman and Carrington, 2007)

2.2.3 miRNA-mediated Suppression Mechanism in Plants

In a variety of eukaryotes, the endonuclease activity of the RNA-Induced Silencing Complex (RISC), is formed by the incorporation of the mature strand of the miRNA into a complex of ribonucleotide proteins (RNPs), resulting in cleavage of target RNA molecules guided by complementary miRNAs or siRNAs (Hammond et al., 2000; Zamore, 2002). In both plants and animals, RISC plays a role in gene silencing mediated by small RNAs (Du and Zamore, 2005). The primary proteins in this complex are members of the Argonaute (AGO) family, whose members all contain a central PAZ domain and a carboxy terminal PIWI domain. Structural studies demonstrate that the binding site for PIWI domain is located at the 5' end of miRNAs, while the 3' end of miRNAs is the site of PAZ domain binding (Parker et al., 2005). The cleavage of targets needs the base pairing of the 5' end of the miRNA with its cognizant target mRNA and occurs at the 10th or 11th nucleotide from the 5' end of the miRNA (Xie et al., 2003; Jones-Rhoades and Bartel, 2004). The cleaved products include 3' hydroxyl and 5' phosphate groups (Jover-Gil et al., 2005). In a study in Arabidopsis loss-of-function ago1 mutants, mRNA target genes were upregulated, while miRNA levels were not reduced. This suggests that AGO1 works downstream of DCL1, HYL1 and HEN1 proteins (Kidner and Martienssen, 2004; Vaucheret et al., 2004). Also it has been shown that numerous conserved amino acid residues in AGO1, seem to be essential for the endonuclease activity of other AGO proteins (Liu et al., 2004; Song et al., 2004). These facts implied that AGO1 formed part of the plant RISC in the miRNA pathway.

In plants, miRNAs show a near-perfect complementary to their targets, resulting in a mechanism to cleave the target mRNAs, which is one of the forms of miRNA-mediated regulation (Jones-Rhoades et al. 2006), while most of the animal miRNA targets are

regulated by repressing protein translation (Humphreys et al., 2005; Pillai et al., 2005) (Figure 2.7).

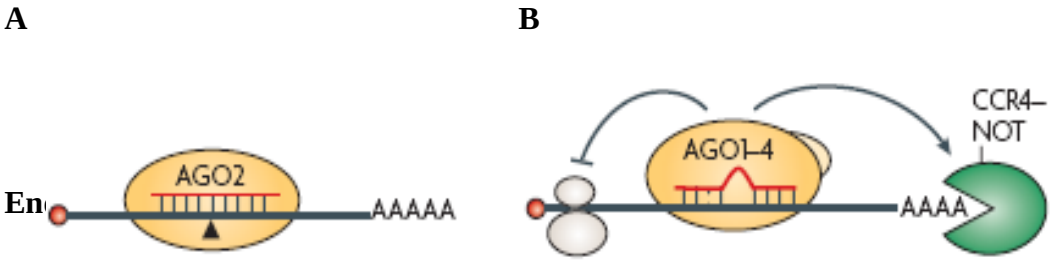


Figure 2.7 The Actions of miRNAs in Target Regulation. (A) Messenger RNA cleavage specified by a miRNAs. Black arrowhead indicates site of cleavage. (B) Translational repression specified by miRNAs. (Reproduced from Filipowicz et al., 2008)

2.2.4 miRNA Precursors (Pre-miRNA)

Plant miRNA precursors form stem-loop (hairpin) structures (Wang et al., 2007), with a huge variation in their lengths and secondary structures (Axtell and Bowman, 2008). These pre-miRNA secondary structures may have lower minimal free folding energy than other types of RNA such as rRNA and tRNA; e.g. this value for 5SrRNA and tRNA is -34.60 kcal/mol and 24.15 kcal/mol respectively (Zhang et al., 2006c).

In animals the secondary structure of miRNA precursors' hairpin is strictly defined. There are very slight differences in the position of miRNA:miRNA* duplexes or even in the size of their loops concerning the secondary structure, while in plants, there are major variations in their lengths and also secondary structures of miRNA precursors.

It is probable that these differences cause various molecular mechanisms, which are required for the distinguishing and processing the miRNA precursors (Krol et al., 2004).

2.2.5 Origin and Evolution of Plant miRNA

The loci that encode plant miRNAs are called the MIR genes, which are obviously distinctive from previously annotated genes (Bartel and Bartel, 2003). Different from a majority of metazoan MIR genes, which are mostly found within introns/exons, most plant MIR genes are intergenic (Voinnet, 2009).

The majority of Arabidopsis miRNA loci are independent transcription units which are expressed under the control of their own promoters (Xie et al., 2005), whilst moss miRNAs mostly overlap protein-coding genes (Axtel et al., 2007). At least five non-conserved miRNAs in Arabidopsis are transcribed from introns of protein-coding genes (Rajagopalan et al., 2006), and a number of miRNAs in mammals (Rodriguez et al., 2004) and fruitflies (Ruby et al., 2007) are encoded in intronic regions. Plant MIR

genes are rarely arranged in tandem, even though clustering does not seem unusual in some plants such as soybean (Voinnet, 2009).

One of the significant features of miRNAs is their evolutionary conservation. In a study by Reinhart et al., (2002), from nineteen miRNAs from Arabidopsis, eight of them had a minimum perfect homology in the location of the hairpin precursor compared to rice, and the rest, had paralogy with some mismatches (≤ 3). Mature plant miRNAs are expected to be encoded in the 5'or 3' hairpin arm; nevertheless, when multiple MIR genes encode an miRNA, in all members of the gene family, the miRNA is always encoded in the same arm of the hairpin (Bartel and Bartel, 2003). The conservation in sequence and structure hints that numerous plant miRNAs have had key roles, since about 250 million years ago prior to when monocots and dicots diverged (Meyerowitz, 2002; Bartel and Bartel, 2003; Voinnet, 2009). Conservation of several miRNA families in moss indicates their very ancient origin (Garcia, 2008), while none of the miRNAs recognized in the Chlamydomonas, single-cell algae, appear to be conserved in multicellular plants (Voinnet, 2009). The evidence proved the conservation of expressed miRNAs among the angiosperms. An miRNA gene in one species may be present in another as a homolog or ortholog (Zhang et al., 2006a).

In spite of the small sizes of miRNAs, the evolutionary history of miRNA gene families seems to be similar to their protein-coding counterparts (Li and Mao, 2007). Many conserved plant miRNA genes are the results of extensive genome duplications and rearrangements; therefore these MIR genes usually have multiple loci and are generally highly expressed (Voinnet, 2009). In plants, there are also another two classes of low and moderately conserved miRNA families, which seem to have important roles in plant development and species-specific characteristics (Zhang el al., 2006a). In Arabidopsis, their predicted targets contain many broader ranges of proteins than those regulated by conserved miRNAs (Voinnet, 2009).

A possible theory for the evolutionary origin of plant miRNAs was derived from the study of the sequence similarity between current miRNA genes and their targets (Axtell and Bowman, 2008). This theory, which is called the inverted duplication hypothesis, suggests that the origin of miRNA genes is from inverted duplications of their target genes (Willmann and Poethig, 2007). In this scenario the hairpin transcripts which are substrates for DCL activities, are the products of the transcription of the inverted duplication (Axtell and Bowman, 2008). Plant genomes rather than animal genomes have larger but fewer miRNA gene families (Li and Mao, 2007). Plants and animals are suggested to have different miRNA evolutionary mechanisms (Allen et al., 2004; Chen and Rajewsky, 2007). Plant miRNAs have perfect or nearly perfect complementary with their target mRNAs and their binding sites mostly are located within the coding region of their target genes. In Plants, target mRNAs usually have an evolutionarily relation to the miRNAs which regulate them. Also the number of miRNA targets is limited (Axtell and Bowman, 2008).

2.2.6 Post-transcriptional Gene Regulation

Gene expression regulation is a complex process which needs the involvement of several factors at various steps (Orphanides and Reinberg, 2002).

Pathways which use double-stranded RNAs (dsRNAs) to regulate genes are called RNA interference (RNAi) or Post-Transcriptional Gene Silencing (PTGS) (Cerutti and Casas-Mollano, 2006). RNAi was found to be present in many eukaryotic organisms and is implicated in numerous gene-silencing phenomena (Murchison and Hannon, 2004). These dsRNAs can be endogenous microRNAs (miRNAs) and small interfering RNAs (siRNAs) or exogenous siRNAs. Four classes of endogenous siRNAs have been described from a variety of organisms. These are: repeat-associated siRNAs (rasiRNAs), trans-acting siRNAs (tasiRNAs), Piwi-interacting (pi)RNAs, small-scan

(scn)RNAs (Rana, 2007) and heterochromatic siRNAs (hc-siRNAs) (Chellappan et al., 2010).

2.2.6.1 siRNAs versus miRNAs

The small interfering RNAs (siRNAs) and microRNAs (miRNAs) have been distinguished in numerous eukaryotes (Grivna et al., 2006). In plants, animals and also their related viruses, miRNAs widely have been recognized while siRNAs seem to be a part of a basic mechanism that also functions in unicellular eukaryotes and probably in prokaryotes (Bartel and Bartel, 2003). The processing of both of these types of small RNAs is based on the cleavage of precursor double stranded RNAs by Dicers, an RNase III-type enzyme (Grivna et al., 2006). Although miRNAs and siRNAs share similar functions and pathways (Bartel and Bartel, 2003), they still have some major differences.

The length of miRNAs is 20-22nt which is shorter than siRNAs (23-27nt) (Chellappan et al., 2010). siRNAs derive from long dsRNAs, while miRNAs are processed from hairpin precursors of single RNA transcripts. A mature miRNA is made from one of the miRNA precursor arms, while in siRNAs, various siRNAs may be made from both siRNA precursor strands (Bartel and Bartel, 2003). Another key difference between miRNAs and siRNAs arises from their target genes. The target genes of miRNAs are other than the genes from which miRNAs were generated; so miRNAs indicate hetero silencing, while endogenous siRNAs normally are auto silencing, meaning that they regulate the target DNA from which they originated (Bartel, 2004; Chen, 2005).

These two classes of small RNAs usually differ in their origins. miRNAs originate from genomic loci, MIR genes, which are different from other known genes, while siRNAs derive from transposons, mRNAs, heterochromatic DNA or viruses. So, all miRNAs are endogenous small RNAs while siRNAs can be exogenous or endogenous

(Rana, 2007). Moreover, recently in a study by Chellappan and colleagues (2010), a novel class of MIR genes was found in Arabidopsis which produces both 20-22 nt and 23-27 nt small RNAs from the same loci. Biogenesis analysis demonstrated that the 20-22 nt species are miRNAs that are dependent on DCL1 while they are independent from RNA Dependent RNA Polymerase (RDR) or RNA polymerase IV (Pol IV). On the other hand, the 23-27 nt species are siRNAs which are dependent on DCL3, RDR2 and Pol IV, these are elements of a classic Heterochromatic siRNAs (hc-siRNAs) pathway. hc-siRNAs protect genome integrity by promoting heterochromatin formation via DNA methylation and/or histone modifications. The biogenesis of hc-siRNAs is dependent on DCL3, RDR2 and Pol IV (Chellappan et al., 2010). Figure 2.8 (A, B) illustrates the common origin and biogenesis pathways of miRNAs and siRNAs, while Figure 2.9 demonstrates the model for the MIR genes of dual function.

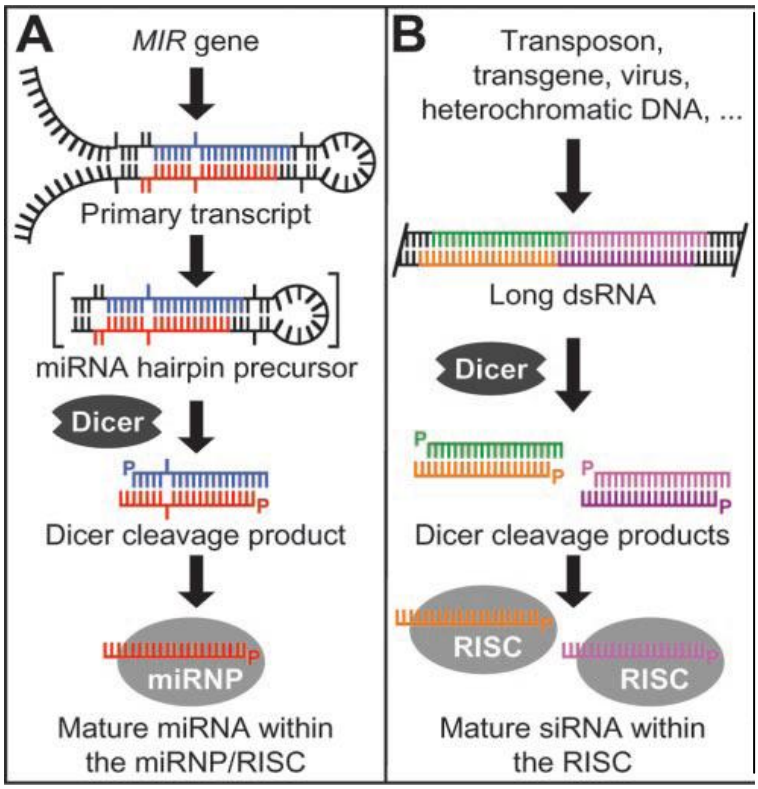


Figure 2.8 Biogenesis of miRNAs and siRNAs. (A) The portion of the primary transcript that contains the miRNA sequence (red) resides on one arm of a predicted stem-loop precursor structure. The transcription start and stop sites for miRNA primary transcripts have not yet been defined. In animals, the hairpin precursor (in brackets) is processed from the primary transcript, but such intermediates have not been detected in plants. Either the primary transcript or this processed hairpin is cleaved by Dicer to yield paired approximately 21-nt RNAs with 2-nt 3' overhangs, 5' phosphates, and 3' hydroxyls. One strand of this short-lived double-stranded intermediate accumulates as the mature miRNA (in red), which acts as a guide RNA within the miRNP/RISC complex. (B) Long dsRNA is processed into many different siRNA species. siRNAs from both strands of the precursor accumulate within RISC complexes. (Reproduced from Bartel and Bartel, 2003)

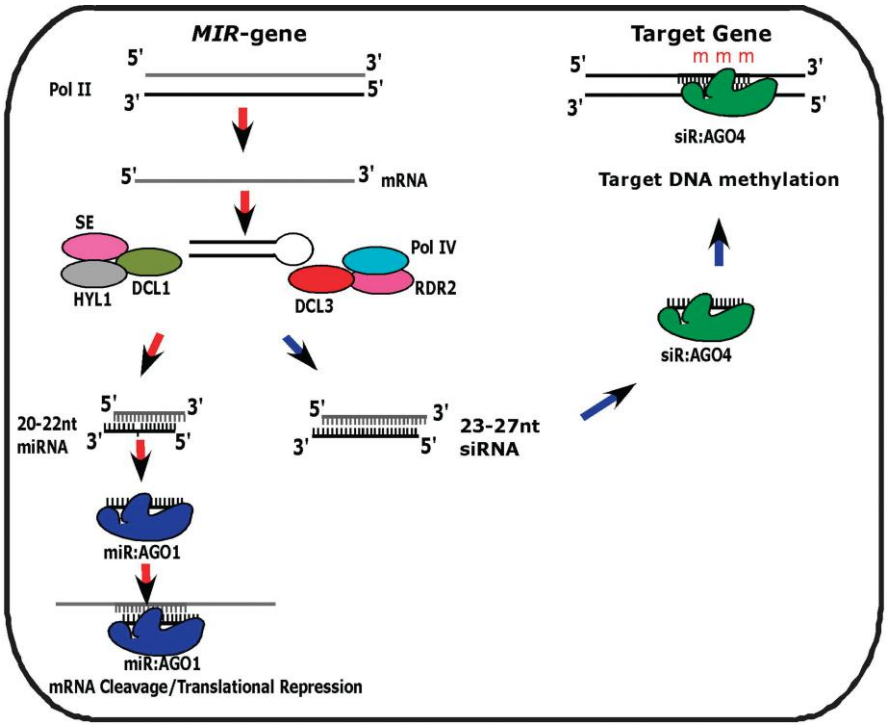


Figure 2.9 A Model for the MIR Genes of Dual Function. The miRNA precursors generate two species by following two pathways: (i) 20- to 22-nt miRNAs are processed predominantly by DCL1/HYL1/SE pathway from pri-miRNA transcripts, and (ii) siRNAs of 23- to 27-nt species are generated by the activities of Pol IV/RDR2/DCL3 pathway. Impairment of any protein in the pathway would affect the biogenesis and function of the particular species of sRNAs. The miRNAs are associated with AGO1 and mediated target mRNA cleavage or translational repression, whereas the siRNAs are associated with AGO4 and direct de novo DNA methylation at target loci. (Reproduced from Chellappan et al., 2010)

2.2.7 Regulation of miRNA Genes in Plants

Studies on the functions of the miRNAs suggested that the expression and activity of miRNA genes are specifically regulated in plants (Chen, 2005). This regulation of plant miRNA can occur in many ways, including transcriptional control and regulation at the level of miRNA processing (Voinnet, 2009), such as pri-miRNA processing by DCL1, miRNA methylation by HEN1, miRNA loading into RISC, miRNA export to the cytoplasm and possible transport of the miRNA inside or outside of cells or tissues (Yoo et al., 2004).

miRNA gene expression also may be under direct or indirect effects of their own or other miRNAs. For instance, miR162-mediated dicer-like1 transcript cleavage (Chen, 2009) and miR168-mediated AGO1 transcript cleavage (Vaucheret, 2009), both act on the general pathways of miRNA biogenesis, forming two feedback loops that result in an miRNA regulating its own expression. Also miR156 regulates the expression of miR172 via its SPL9 target mRNAs which directly promotes the transcription of miR172b in Arabidopsis (Wu et al., 2009).

So far, the methods of regulation of miRNAs are not well-known and it seems that in the future, probably, one of the challenges in microRNA biology would be, understanding possible miRNA regulation mechanisms.

2.2.8 Plant miRNA and Target mRNAs

The first miRNAs discovered in plants were found to be engaged in developmental control by regulating their related genes or transcription factors (Willmann and Poethig, 2007).

miRNAs negatively regulate their target genes by either one of two mechanisms: processing by cleavage (usually in plants) or translational inhibition (in animals) of the target mRNAs (Chen, 2005). Animal miRNAs often bind to target mRNAs through

imperfect complementary at multiple binding sites in the 3' UTR of the target mRNAs (Olsen and Ambros, 1999; Bartel, 2004), blocking ribosome movement along the mRNA, and repress gene expression (Carrington and Ambros, 2003), while plants have similar mechanisms in gene expression regulation. miRNAs in plants have a high degree of sequence complementary to their target mRNAs (Reinhart et al., 2002; Allen et al., 2005), most target mRNAs usually contain a single miRNA complementary site located in the coding region (Rhoades et al.,2002; Bartel, 2004), and most perfectly matched miRNAs normally cleave the target mRNAs (Bartel, 2004). Distinct from animal miRNA targets, the complementary sites in plants can be anywhere along the target mRNA rather than in the 3' UTR. Studies on Arabidopsis and rice indicated the nearly perfect complementary between miRNAs and their target mRNAs (Rhoades et al., 2002); also validation of targets for miRNAs in plants is done broadly using the cleavage of the target in the middle of the miRNA complementary site. Similar to animal miRNAs, and different from most plants, miRNA172 (He and Hannon, 2004; Chen, 2004) and miR398 (Dugas and Bartel, 2008) are plant miRNAs known to repress the translation of their mRNA targets. Nevertheless, another study also indicated that miR172 is also capable to guide cleavage of its target mRNAs (Schwab et al., 2005). As it was also mentioned by Zhang and colleagues (2006b), that this can support the idea that, in plants, miRNAs may be involved in more complicated mechanisms to control gene expression than in animals (Zhang et al., 2006b).

2.2.8.1 miR172 and its APETALA2-like Targets

Studies on *Arabidopsis thaliana*, a flowering plant, showed that miRNA172 regulates the expression of APETALA2 (AP2)-like genes (Glazinska et al., 2009).

The AP2/ERF (Ethylene Responsive Element Binding Factor) protein family is defined by DNA-binding domain that contains sixty conserved amino acids (Shigyo et al., 2006; Canella et al., 2010). The AP2/ERF multigene family includes the genes with one or two AP2 domains: the ERF (EREBP) subfamily has one AP2 domain and the AP2 subfamily contains two AP2 domains. This multigene family is responsible for coding putative transcription factors (Nakano et al., 2006; Hinz et al., 2010).

In order to elucidate the AP2 subfamily's molecular evolution, Shigyo and colleagues (2006) isolated and sequenced AP2-like genes with two domains from *Cycas revoluta*, *Ginkgo biloba* and *Gnetum parvifolium* which all are gymnosperms. The results showed that among gymnosperm AP2 homologues, the miR172 target site is considerably conserved. This indicates that the mechanism of gene regulation by miR172 has been conserved from the time that gymnosperm and flowering plant ancestries had become diverged, more than three hundred million years ago (Shigyo et al., 2006).

Several studies confirmed that miR172 is engaged in developmental transition (such as transition from juvenile phase to adult), floral organ identity and flowering time (Aukerman and Sakai, 2003; Chen, 2005; Wu et al., 2009; Zhu et al., 2009; Glazinska et al., 2009).

Flowering organs are set off in spirals from the floral meristem. By the combined activities of three major classes of genes, known as A, B and C floral Homeotic genes, the characteristics of the floral organ primordia was identified (Jack, 2004). Class A gene, APETALA2 (AP2), has its transcript present within the flowering meristem. Loss-of-function mutations in the class A (AP2) and C (AGAMOUS (AG)), genes

indicated that these two genes function antagonistically to hamper each other's activity within the floral meristem (Chen, 2005).

Aukerman and Sakai (2003) demonstrated that miR172 overexpression causes early flowering and disturbs the flowering organ characteristics in Arabidopsis. In general, miR172 is expressed in a temporal manner and it was suggested that it is engaged in controlling of flowering time. A subfamily of the APETALA2 transcription factor family, (AP2-like), are the regulatory targets of miR172 (Aukerman and Sakai, 2003; Chen, 2004). As shown in studies by Chen (2004; 2005), miR172 overexpression from the 35S promoter results in decreasing the levels of AP2 protein and floral homeotic phenotypes similar to ap2 loss-of-function mutants and leads to early flowering. Overexpressing of a miR172-resistant type of AP2 caused a replacement of reproductive organs by perianth organs. Moreover, some intensive abnormalities in the interior two spirals of Arabidopsis flowers had happened, while an AP2 promoter miR172-resistant type was expressed (Chen, 2004; Chen, 2005). These observations can show the key role of miR172 in regulation of AP2 gene.

In addition to AP2 family members, miR172 regulates several AP2-like genes including TOE1, TOE2, TOE3, SMZ and SNZ in Arabidopsis (Mathieu et al., 2009). The mutant toe1 takes the plant to flowering stage a little earlier while the toe2 mutant does not have any effect on flowering time. But compared with wild type, a double mutant of toe1 and toe2 flowered much earlier, indicating that TOE1 and TOE2 are additional suppressors of the vegetative to reproductive phase change (Chen, 2005), while the overexpression of TOE1 (Aukerman and Sakai, 2003), SMZ or SNZ (Schmid et al., 2003) causes delay in flowering.

miR172 overexpression which accelerates flowering (Jung et al., 2007; Wu et al., 2009), together with the results of other studies (Chen, 2004; Chen et al., 2005), indicate that AP2, TOE1 and TOE2 are the genes through which, miR172 regulates the

vegetative to reproductive phase change. However, studies also showed that the level of TOE1 mRNA was not reduced by the overexpression of miR172 in Arabidopsis (Aukerman and Sakai, 2003; Chen, 2005), while overexpression of miR172 lead to a decline in TOE2 mRNA; indicating that the method of TOE1 regulation is translational inhibition, whereas the regulation mechanism for TOE2 is mRNA target degradation. These two results suggest that the miR172 uses different mechanisms to regulate its different targets (Schmid et al., 2003; Chen, 2005).

Glossy15 and indeterminate spikelet1 were recognized as AP2-like genes in maize (Aukerman and Sakai, 2003). Glossy15 regulates the transition from juvenile to adult leaves in maize (Lauter et al., 2005). The juvenile leaves are different from adult ones in several epidermal features. The RNA of glossy15 is only found in juvenile leaves of maize and this Ap2-like gene, induces juvenile leaf identity (Kerstetter and Poethig, 1998; Chen, 2005). miR172 can be detected in adult leaves while glossy15 mRNA is absent. This makes it possible that miR172 acts on glossy15 mRNA allowing vegetative phase transition (Chen, 2005). miR172 sequence complementarity to the glossy15 mRNA and the miR172-mediated cleavage products could be detected in vivo; these two results together were taken to imply that glossy15 is a miR172 target gene (Lauter et al., 2005).

MiR172 has also been shown to target AP2-like family members in several plant species, such as, wheat (Yao et al., 2007), tomato (Itaya et al., 2008), soybean (Subramanian et al., 2008), rice (Zhu et al., 2009) and *Ipomoea nil* (Glazińska et al., 2009).

2.2.9 Other miRNAs Engaged in Flowering and Developmental Transition

Beside miR172 and its target AP2 which are engaged in regulation of flowering, miR156 is another miRNA where overexpression in both maize and Arabidopsis delays the expression of juvenile vegetative characteristics and prolongs flowering time (Chuck et al., 2007; Wu et al., 2009); therefore miR156 is essential for the phase transition and regulates the juvenile to adult phase change. miR172 acts downstream of miR156 to promote adult epidermal identity and these two miRNAs have opposite expression patterns (Wu et al., 2009). In Arabidopsis, out of sixteen, ten members of the SPL transcription factors have been identified as targets for this miRNA (Wu and Poethig, 2006).

miR159 is another miRNA which is engaged in regulation of flowering time. In Arabidopsis, under short day conditions, miR159 overexpression causes late flowering but this will not happen under long day conditions. miR159 targets MYB transcription factors such as GAMYB (Schwab et al., 2005; Wang et al., 2007). miR319/Jaw is another miRNA which is engaged in regulation of both leaf development and flowering (Palatnik et al., 2003). The targets of miR319/Jaw family are a group of TCP transcription factors (Wang et al., 2007). Overexpression of miR319/Jaw showed phenotypes of crinkly leaves and also delayed in flowering in Arabidopsis. It also reduced the levels of five TCP target genes (Palatnik et al., 2003; Chen, 2005). Baker and colleagues (2005) showed that, miR164 targets transcription factors CUC1 and CUC2 genes. In Arabidopsis, miR164c is involved in control of petal number in a scarce manner. The analysis of overexpressed miR164a or miR164b or expressing CUC1/CUC2 miRNA-resistant demonstrated that miRNA regulation of CUC1 and CUC2 is necessary for normal flower development (Baker et al., 2005; Wang et al., 2007).

3. MATERIALS AND METHODS

3.1 Collection and Characterization of Oil Palm Tissue Samples

Inflorescence tissues at different developmental stages were collected from a single oil palm (*Elaeis guineensis*), from Sime Darby Plantations, East Plantation, Carey Island, Selangor, Malaysia. The sampled tree was grown from seed harvested in 1977 from a Deli Dura (mother palm) x BM119 (AVROS) Pisifera from the series, GH200 (deriving from SP540 Indonesian stock) pollen. The different developmental stages of oil palm flowers used in this study are defined in Table 3.1. In oil palm plants, the youngest expanding leaf was numbered as leaf 0. The leaf produced immediately before leaf 0 is thus leaf F+1 and the leaf produced subsequently is leaf F-1 (this leaf is not yet expanded). The reproductive development in oil palm is illustrated in Figure 3.1. Samples were labeled and immediately placed in liquid nitrogen for transfer from the field to the laboratory at University of Malaya (UM) where they were placed in a -80°C freezer for storage and prior to RNA extraction. Eight month old oil palm plants from the same genetic line were grown at the UM in separate polybags under 12 hours daylight at 23-25 °C with watering once each day. Shoot apical meristem which is located at the base of leaves in each plant, root and leaf samples were collected from these plants at the 10, 11 and 12 month old stages.

Table 3.1 Developmental Stages of Oil Palm Flowers

Tissue Name	Frond Number*	Description
-Inflorescence	-27 to -1	Immature inflorescences positioned at the base of the fronds inside of the trunk
+Inflorescence	+1 to +14	Immature inflorescences positioned at the base of the fronds outside of the trunk
Emerged Flower	+16	Inflorescences which have emerged from the leaf bract but are still not distinguishable as male or female
Mature Flower	+22 and higher	Inflorescences which are distinguishable as male or female.

***** Positive numbers indicate flowers on the outside of the palm trunk with higher numbers indicating the most mature inflorescences.

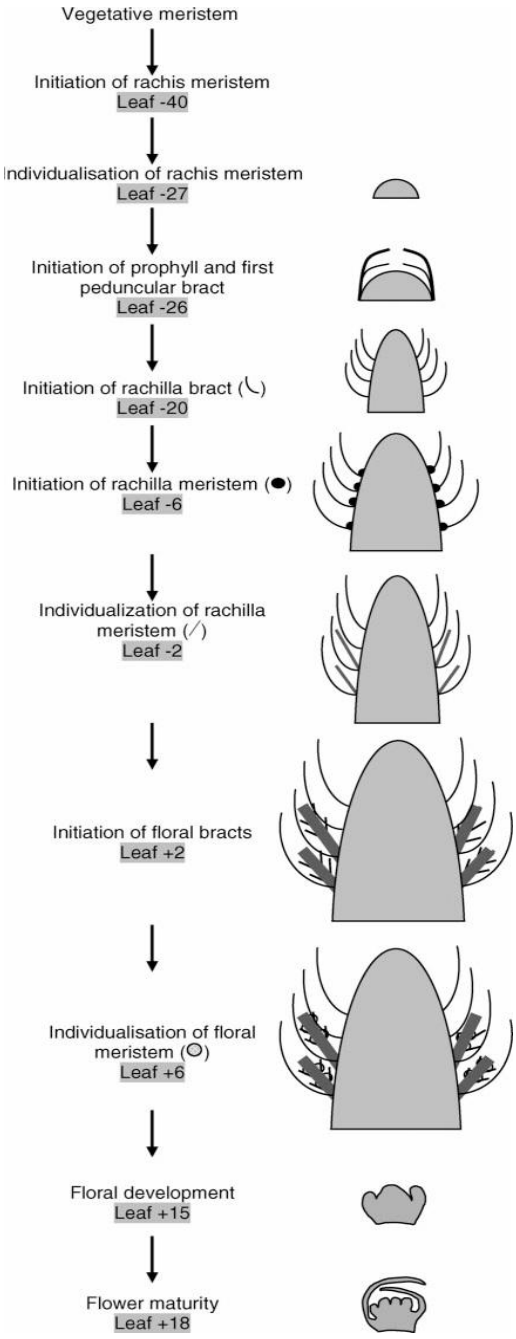


Figure 3.1 Reproductive Development in Oil Palm Chronology of the principal phases of inflorescence and flower development in oil palm. Note that this is a generalized scheme applying to inflorescences of both sexes. (Reproduced from Adam et al., 2005)

3.2 Plant Total RNA Extraction

Total RNA was extracted from samples using a modified method based on Kiefer et al., 2000. One hundred to one hundred and fifty mg of each oil palm tissue: Shoot, root, leaf (six, eight, ten, eleven and twelve month old), immature flowers (+ and – inflorescence), emerged inflorescence and mature flowers were ground separately using cold mortar and pestles under liquid nitrogen. The powdered tissues were then placed in a 2 ml microcentrifuge tube and 1ml of pre-warmed (65°C) extraction buffer (100mM Tris-HCI, 25mM EDTA, 2M NaCI, 2% CTAB w/v, 2% PVP w/v, 2% mercaptoethanol) was added to the tube. After that, an equal volume of chloroform/isoamyalcohol (24:1 v/v) was added and the tube was vortexed at 1,800 rpm at room temperature for 5 minutes. The tube was then centrifuged for 15 minutes at 4°C, 10,600 x g and the supernatant transferred to a new 2ml microcentrifuge tube. Five hundred µl of chloroform/isoamyalcohol (24:1; v/v) was then added to the supernatant. The tube was vortexed at room temperature followed by centrifugation for 15 minutes at 4°C, 10,600 x g. The supernatant was transferred to a 2ml microcentrifuge tube and five hundred µl of Phenol: Chloroform: Isoamylalcohol (PCI) (24:1:25; v/v/v) was added to it. The tube was again vortexed at room temperature followed by centrifugation for 15 minutes at 4°C, 10,600 x g. Another time the supernatant transferred to a new 2ml microcentrifuge tube. Five hundred µl of chloroform/isoamyalcohol (24:1; v/v) was then added to the supernatant. The tube was again vortexed at room temperature followed by centrifugation for 15 minutes at 4°C, 10,600 x g. The supernatant was transferred to a 1.5ml microcentrifuge tube and 3 volumes of cold absolute ethanol (4°C) and 0.1 volume of NaOAc (Sodium Acetate) was added to it. The tube was incubated at -80°C overnight, followed by centrifugation for 30 minutes, 4°C at 17,900 x g. The pellet was then washed with cold 70% ethanol

and centrifuged for 5min, 4°C at 17,900 x g. The supernatant then was discarded and the pellet was resuspended in 50µl of DEPC treated dH2O.

3.3 Agarose Gel Electrophoresis

In order to check the RNA extraction quality, a 1% agarose gel was run. In a flask 20mg of agarose powder was weighed and then 20 ml 1X TBE (Tris/Borate/EDTA) was added to it. The flask then transferred to a microwave at 20% power until all agarose was dissolved (usually 45 seconds). The liquid then was cooled under cold running water for 10-15 seconds or was allowed to become cool by letting it sit at RT until it was not too hot to hold. Ethidium bromide (1µl per 20 ml) was added to the liquid agar in the flask. Then, the gel was poured into the gel mold held in a cassette, with the comb in place. The gel was allowed to set for about 20 minutes. After that, the comb was removed and the gel in the gel mold was transferred to a tank containing of 1X TBE buffer in a way that the gel was completely submerged. Then on a sheet of parafilm 1µl of RNA was mixed with 1µl of 2X RNA loading dye and then was loaded into each well of the gel. In order to estimate the RNA band sites 1µl of a 50bp ladder was loaded. Then the tank was connected to the power supply and the gel was electrophoresed at 120V for 40 minutes. After that the RNA bands were visualized using a gel imaging system.

3.4 DNase Treatment

In a 0.5 ml tube, 5µl of DNase buffer (40mM Tris-HCI, 60mM MgCI2, pH7.5) and 5µl (5U) of DNase I were added to 5µl of RNA. The mixture was then incubated in a 37°C water bath for 30 minutes. Later, 5µl of RQ1 DNase Stop Solution was added and incubated in 65°C water bath for 10 minutes. This DNase treated RNA was used for one step RT-PCR and also for quantitative reverse transcription PCR.

3.5 Gel Extraction of Small RNA

Since DNase treatment only removes DNA contamination, small RNA gel extraction was performed to enrich for miRNA. A 2ml tube was filled with 0.4M NaCl and placed in an ice box. Extracted RNA from the previous step was run on a 2% agarose gel. According to the ladder the RNA which was less than 250bp located in the lower part of the gel, was cut under the UV in a dark room and the fragments were transferred to the 2ml tube of 0.4 M NaCl. After this step the ice box was placed on a belly dancer, with the speed of 3 for four hours. Then the tube was stored at 4°C overnight. After that the liquid content of the tube was transferred into a new tube, then 3 volumes of cold absolute ethanol (4°C) and 0.1 volume of NaOAc (Sodium Acetate) were added to it. The tube was incubated at -80°C overnight, followed by centrifugation for 30 minutes, 4°C at 17,900 x g. The pellet was then washed with cold 70% ethanol and centrifuged for 5min, 4°C at 17,900 x g. The supernatant then was discarded and the pellet was resuspended in 50µl of DEPC treated dH2O. The small RNAs were kept in -80°C until use.

3.6 Primer Design

3.6.1 Potential Target RNA

The PDA primer design software (http://dbb.nhri.org.tw/primer/index.html) was used to design the forward and reverse primers for the 662bp potential target RNA which was obtained from EST of oil palm and later will be explained in section 3.14. The sequences of the primers were: F: 5'- GTA CAT ATA TCT TGG GCT ATT CG -3' (upstream), R: 5'- GGA GCC GGT AGG CTG AAC TGG -3' (downstream).

Since in quantitative reverse transcription PCR a shorter amplification product is important, the primers were designed to amplify a 105bp fragment while the miR172

complementary site also was located within this amplicone. The sequences of the primers were: F: 5'-GCG TGG CAG ATG CAT GGC -3' (upstream), R: 5'- AGG CAG CGA GGA AGG ATG G -3' (downstream).

3.6.2 miR172

The miR172 sequences from three monocots plants (*Zea mays*, *Oryza sativa* and *Sorghum bicolour*) were obtained from miRbase, version 14, (http://microrna.sanger.ac.uk/). Forward and reverse primers were designed by aligning the precursors of the three miR172s using ClustalW (Figure 3.2). The primer sequences were: F: 5'- GTG CAG CAT CAT CAA GAT TCA -3' (upstream), R: 5'- ATG CAG CAT CAT CAA GAT TCT -3' (downstream).

```
osa-MIR172a     GUGUUUGCGGGCGUGGCAUCAUCAAGAUUCACAUCCAUG----CAUAUAUCACAAGACGC 56
zma-MIR172a     ----------GCGUGGCAUCAUCAAGAUUCACAACC-------CAUCAAUC-CGAACCAC 42
sbi-MIR172a     ----------GUGCAGCAUCAUCAAGAUUCACAUCCAGCUCAUCCUCGGUGAUAUGCUAU 50
                          * *  ******************** **       * *   *

osa-MIR172a     ACAUAUACAU-----CCGAUUUGGC----UGAGAAUCUUGAUGAUGCUGCAUCCGCAGAC 107
zma-MIR172a     UGAUUUGGAA-----UGCAUGYA------UGAGAAUCUUGAUGAUGCUGCAU-------- 83
sbi-MIR172a     AUACAUAAAUAUAUGCGUAUGUGAUGAGAUGAGAAUCUUGAUGAUGCUGCAU-------- 102
                  *  *  *          **          ************************

osa-MIR172a     AA 109
zma-MIR172a     --
sbi-MIR172a     --
```

Figure 3.2 Multiple Alignment for miR172 Primer Design. Three different precursor miR172s were aligned using ClustalW. osa (*Oryza sativa*), zma (*Zea mays*) and sbi (*Sorghum bicolour*).

3.7 One Step Reverse Transcription PCR

One-step reverse transcription PCR was carried out using the Access RT-PCR System, (Promega, USA), to detect miR172 and candidate target RNA, using the specific designed primer sets. For both RT-PCR reactions (miR172 and candidate target) the RNAs were used from five different oil palm tissues (shoot, root, leaf, mature and immature flowers). The reaction mixture was as shown in Table 3.2.

Table 3.2 Reaction Components for One Step RT-PCR

Reaction Component	Volume (µl)	Final Concentration
AMV/Tfl 5X Reaction buffer	5	1X
10mM dNTP	0.5	0.2mM
miR172 Forward primer (10µM)	2.5	1µM
miR172 Reverse primer (10µM)	2.5	1µM
25 mM MgSo4	1.5	1.5mM
5U/µl AMV Reverse transcriptase	0.5	0.1U/µl
5U/µl *Tfl* DNA polymerase	0.5	0.1U/µl
RNA template	5	10pg-1µg
Nuclease free water	7	-
Total	**25**	-

The thermal cycling profile was as follows: reverse transcription (45°C for 45 minutes), inactivation of AMV reverse transcriptase (94°C for 2 minutes), denaturation of RNA/cDNA primer, (94°C for 2 minutes), denaturation (94°C for 30 seconds), annealing (59°C for 1 minute), extension (68°C for 2 minute), final extension (68°C for 7 minute), hold (4°C) with the denaturation, annealing and extension steps repeated for 35 cycles. The PCR products were separated on a 1.0% agarose gel run at 120V for 25 minutes. The PCR products were analyzed under UV light after ethidium bromide staining (0.5 µg/ml).

3.8 Gel Purification of PCR Products

Fifty µl of the PCR products were separated on 2% agarose gel and then the expected bands were purified using QIAquick Gel Extraction Kit (Qiagen, Germany). The DNA fragments were excised from the agarose gel with a clean, sharp scalpel. The gel slices were weighed in a 1.5 ml microcentrifuge tube. Then 3 volumes of buffer QG to 1 volume of gel was added (100 mg ~100). The tubes then were incubated at 50°C for 10 minutes, to help dissolve the gel, every 2-3 minutes the tubes were vortex during the incubation. After the gel slices were dissolved completely, 1 gel volume of Isopropanol was added to the each tube. One QIAquick spin column was placed in a provided 2ml collection tube. To bind DNA, the sample was applied to the QIAquick column and was centrifuged at 17,900 x g for 1 minute and the flow through discarded. Next, 750µl of Washing Buffer (buffer PE) was added to the column and centrifuged at 17,900 x g for 1 minute. The supernatant was discarded and the column centrifuged again at 17,900 x g for 1 minute to dry the membrane. The DNA was then eluted by adding 30µl of the Elution Buffer directly to the centre of the column, letting it stand for 1 minute and then centrifuging for 1 minute at 17,900 x g.

3.9 Cloning of PCR Products

PCR products were cloned using *E.coli* competent cell prior to sequencing in the following three steps:

3.9.1 Ligation

The pGEM®-T Easy Vector from pGEM®-T and pGEM®-T Easy Vector Systems (Promega, USA) was briefly centrifuged to collect contents at the bottom of the tube and the 2X Rapid Ligation Buffer was Vortexed vigorously. The ligation reactions were prepared for each samples separately in different 0.5 ml tubes as described in

Table 3.3. The ligation reactions, then, were mixed by pipetting and left for incubation overnight at 4°C.

Table 3.3 Ligation of DNA in the pGEM®-T Easy Vector

Reagents	**Volume/ Standard Reaction (µl)**
10X Rapid Ligation Buffer	1
pGEM®-T Easy Vector (50ng)	1
PCR product (100ng)	6
T4 DNA Ligase (3 Weiss units/ µl)	1
Nuclease free water	1
Total	**10**

3.9.2 *E.coli* Transformation

In a 250 ml bottle 7 g of LB-agar was dissolved in 200 ml dH2O by heating, and then it was autoclaved in 121°C for 15 minutes with pressure of 103 kPa. After cooling down, 1ml of IPTG (0.1M) (Isopropyl β-D-thiogalactopyranoside), 320 µl of X-gal (5%) and 100 µl ampicillin (100 mg) were added to it and after a gentle agitation 30-35 ml of medium was poured into 85 mm petri dishes. The agar was left to harden and put at 37°C for 1 hour before use.

The ligation reactions were centrifuged briefly and 5 µl of each ligation reaction was added to a sterile 1.5 ml microcentrifuge tube on ice. The high-efficiency *JM109* Competent Cells from Promega pGEM®-T and pGEM®-T Easy Vector Systems were placed in an ice bath until just thawed (5 minutes), then the tubes were flicked gently. Hundred µl of cells carefully were transferred to the ligation reaction tubes prepared in the previous step. Gently the tubes were flicked, and incubated on ice for 20 minutes. The cells were put for 45–50 seconds in a water bath at exactly 42°C to be heat shocked. Then without any shaking, immediately were returned to ice for 2 minutes.

After that 900 µl of room temperature LB Broth media was added to the ligation reaction transformations and incubated for 1.5 hours at 37°C with shaking (~150rpm). For the negative control, 100 µl competent cells and 900 µl LB Broth were placed into a 1.5ml tube and incubated for 1.5 hours at 37°C. Afterward 200 µl of each transformation was cultured on duplicate LB/ampicillin/IPTG/X-gal plates. The plates were sealed with parafilm and were incubated at 37°C overnight (16-24 hours).

3.9.3 Colony PCR

Well-isolated white colonies were picked using a flamed and cooled bacterial loop, and transferred to a new LB-agar plate which was called Library, containing the ampicillin for overnight culture in 37°C and plasmid prep. Part of each colony was transferred to 30 µl of sterile water, which later was used as the DNA template for colony PCR. The colony PCR was done using GoTaq® Flexi DNA Polymerase (Promega, USA) as described in Table 3.4.

Table 3.4 Colony PCR using GoTaq® Flexi DNA Polymerase

Reaction Components	Volume (µl)	Final Concentration
5X Reaction buffer	2.5	1X
dNTP mix (10mM of eachdNTP)	0.25	0.2mM each
GoTaq® DNA polymerase (5u/ µl)	0.1	0.025U/µl
25 mM MgCl2	0.75	1µM
10mM M13 Forward primer	1.25	1.5mM
10mM M13 Reverse primer	1.25	1mM
DNA template	3	-
Nuclease free water	3.4	-
Total	**12.5**	-

The thermal cycling profile was as follows: initial denaturation, (94°C for 2 minutes), denaturation (94°C for 1 minute), annealing (59°C for 1 minute), extension (72°C for 1 minute), final extension (72°C for 5 minutes) and hold (4°C) with the denaturation, annealing and extension step repeated for 35 cycles. The completed PCR products then were analyzed by 1% agarose gel electrophoresis for the appropriate size product, which indicated if the correct insert was present in the clone.

3.9.4 Plasmid Isolation

The appropriate colonies were sub-cultured in different universal bottles containing 10 ml of LB/ampicillin (100 μg/ml) at 37°C shaking water bath (~150 rpm) overnight. Afterwards, the 10 ml of the LB/ampicillin/*JM109* cells culture were transferred to 15ml centrifuge tubes and centrifuged for 2,656 x g for 15 minutes. The supernatants were discarded and the pellets dried. Two hundred μl of solution I (50 mM glucose, 25 mM Tris-HCI, pH8 and 10 mM EDTA) was added to the each pellet, mixed and vortexed vigorously to dissolve the pellets. The mixtures then were transferred to new 1.5ml microcentrifuge tubes. Next, 200μl of freshly prepared solution 2 (1% SDS and 0.2M NaOH) was added and then left in room temperature for 4 minutes. Later, 200μl of solution 3 (60 ml of 5M potassium acetate, 11.5 ml of glacial acetic acid and 28.5 dH2O) was added and the mixtures were left on ice for 15 minutes. The mixtures were centrifuged for 10 minutes at 17,382 x g and the supernatant transferred to new 1.5 ml microcentrifuge tubes. After that, 5μl of 50mg/ml RNase A was added and incubated in a 37°C water bath for 3 hours. After RNase A incubation, 600μl of phenol was added, and the mixtures vortexed and centrifuged at 17,382 x g for 3 minutes. The supernatants were removed and placed in new microcentrifuge tubes. Next, 600μl of chloroform was added, the mixtures vortexed and centrifuged again at 17,382 x g for 3 minutes. The aqueous layers were removed and 0.1 volume of 3M sodium acetate and

2.5 volume of absolute ethanol were added to the each tube. The mixtures were then placed on ice for 20 minutes. Later, it was centrifuged at 17,382 x g for 15 minutes. The supernatants were discarded and 1ml of 70% ethanol was added to each one and centrifuged for 15 minutes at 17,382 x g. The supernatants were discarded by pipetting and the pellet was vacuum dried. Lastly, 30µl of dH2O was added to the each pellet and the samples were left overnight at 4°C (method modified from Sambrook et al., 2001).

3.10 DNA Sequencing

One µl big dye, 7µl Buffer, 2µl plasmid template, 1µl M13 forward primer (-20) and 9µl of dH2O was added into a clean PCR tube. The contents of the tube were then mixed well and centrifuged briefly. The tube was then placed into the thermo cycler-PTC 100 (MJ Reserch Inc., USA) with the following thermal cycler conditions used: 96°C for 1 minute, 96°C for 10 seconds, 50°C for 5 seconds, 60°C for 4 minutes, Hold for 4°C with denaturing, annealing and extension steps repeated for 24 cycles.

After that, the contents of the tube were again vortexed and centrifuged briefly. Five µl of 125 mM EDTA and 60µl of 100% ethanol was added and mixed by inverting. The tube was incubated at room temperature for 15 minutes and then centrifuged at maximum speed for 15 minutes. The supernatant was discarded and 60µl of 70% ethanol was added. The tube was later centrifuged at full speed for 10 minutes at 4ºC. The supernatant was discarded and the tube centrifuged for another 1 minute at full speed. Later, the pellet was precipitated and vacuum dried. Next, 10µl of Hi-Di-Formamide (Applied Biosystems, USA) was added to resuspend the pellet and the mixture denatured at 95ºC for 5 minutes. Lastly, the tube was inserted into the 3130 xl DNA analyzer (Applied Biosystems, USA) for sequencing.

Later, the output sequences were analysed with Chromas Software (Technelysium Pty Ltd, Australia) to remove the vector sequences, and obtain the desired sequences.

3.11 mFold and miRbase Analysis

The sequencing results from previous step were subjected to mFold software at the http://mfold.bioinfo.rpi.edu website (Zuker, 2003). mFold is a tool which was developed for predicting the secondary hairpin structure of RNA and DNA, mostly by using thermodynamic methods. The output sequences were analyzed for their ability to form miRNA precursors based on the following criteria for a potential miRNA: (a) The RNA sequences should be able to fold into a stem-loop structure, in a way that mature miRNA locate through one of the stem-loop arms containing 30-70% A+U. (b) The free folding energy of potential miRNA should be less than other types of RNA. (c) The mature miRNA:miRNA* sequences must have less than six mismatches, while maximum of three nucleotide should be in any loop or bulge located in the mature miRNA:RNA* (Zhang et al., 2006c; Qiu et al., 2007). Accordingly by using mFold and considering above criteria, after filtration of the unwanted sequences, the potential miR172 precursors were selected.

Afterward the output sequences were analyzed by blast against miRbase (http://microrna.sanger.ac.uk) in order to compare potential oil palm miR172 sequence with those of other miR172 family members.

3.12 Nested PCR of Six Putative Precursor egu-miR172 Paralogs

Nested PCR was performed using *Tfl* DNA Polymerase reagents (Promega, USA). DNA template (section 3.8) was used for all of the six reactions. The six sets of interior nested PCR primers were designed based on the sequences of putative precursor egu-miR172 paralogs obtained in section 3.11. The primers were as shown in Table 3.5.

Table 3.5 Six Primer Sets for Nested PCR

Primers	Forward and Reverse Sequences (5'-3')
Set 1	F: CATCTGTTAATCACTGCC R: CATCCTTTGTCCCTGTAGAC
Set 2	F: CATCCTGGTTCTGAATTAAG R: CATCCATCGCCCATCTCTGC
Set 3	F: CATCCATCGCCCATCTCTGC R: CATCCTGGTTCTGAATTAAGAG
Set 4	F: CATCGTTCCCATCTCTGGG R: AAGTCAAGAGCCTCAAGAAG
Set 5	F: CAGCTTCTGATACCAATAAAC R: AGGGTTGCAGGGACAATTAC
Set 6	F: GCCUUUGACUUGUCGUAGCGAG R: ACGCTATCGAGACGGAAATTCC

Each set of primers was used with oil palm tissues (shoot, root, leaf, immature and mature flowers). The reaction mixtures are shown in Table 3.6.

Table 3.6 Reaction Components for Nested PCR

Reaction Component	Volume (µl)	Final Concentration
Tfl 5X Reaction buffer	5	1X
10mM dNTP	0.5	0.2mM
25 mM MgSo4	1.5	1µM
miR172 Forward primer (10µM)	2.5	1µM
miR172 Reverse primer (10µM)	2.5	1.5mM
5U/µl *Tfl* DNA polymerase	0.5	0.1U/µl
DNA template	1	10pg-1µg
Nuclease free water	11.5	-
Total	**25**	-

The thermal cycle condition were: initial denaturation (95°C for 2 minutes), denaturation (95°C for 45 seconds), annealing (60°C for 1 minute), extension (68°C for 2 minutes), final extension (68°C for 5 minutes), hold (4°C) with the denaturation, annealing and extension step repeated for 34 cycles. The PCR products were separated on a 1.0% agarose gel run at 120V for 30 minutes. The PCR products were analyzed under UV light after ethidium bromide staining.

3.13 Quantitative Reverse Transcription PCR Analysis

3.13.1 SYBR-Green

SYBR-Green chemistry was used for expression analysis of six potential pre-miR172 using primers mentioned in Table 3.5 as well as candidate AP2-like target using primers already mentioned in section 3.6.1 in different developmental stages of oil palm.

3.13.1.1 Selection of an Endogenous Control

SYBR-Green quantitative reverse transcription PCR was performed using cDNAs from five different oil palm tissues, i.e. shoot, root, leaf, mature and immature flowers together with four different endogenous control gene primer sets as mentioned in Table 3.7. For each sample, three replicates were prepared. The preparation of reaction mixtures for both Reverse transcription and PCR amplification are described in Tables 3.8 and 3.9.

Table 3.7 Endogenous Controls Forward and Reverse Primers

No.	Endogenous Controls	Forward and Reverse Sequences (5'-3')	Amplicon length (bp)
1	β-actin (EU284857.1)	F: GCCAAGAGCAGCTCCTCTGTAG R: GGAATTGTAGGTAGTTTCAGG	130
2	18S rRNA (EL693697.1)	F: CGAGACCTCAGCCTGCTAAC R: AACCTTGTTACGACTTCTCC	450
3	eEF1-α (AK073620; Jain et al.,2006)	F: TTTCACTCTTGGTGTGAAGCAGAT R: GACTTCCTTCACGATTTCATCGTAA	103
4	UBQ5 (AK061988; Jain et al., 2006)	F: ACCACTTCGACCGCCACTACT R: ACGCCTAAGCCTGCTGGTT	69

In order to determine the most stably expressed reference genes in a given set of tissues (Vandesompele et al., 2002) the geNorm VBA applet for Microsoft Excel was performed. geNorm calculates the gene expression stability measure M. the most sTable gene should have the smallest M value (M<1.5).

For start up the geNorm applet, Excel was opened and macro's were enabled when prompted. Then the raw expression data matrix (not yet normalized expression levels) was loaded as the input in a way that the first column contained the sample names and the first row contained the gene names. The first cell or the first row and column should be empty.

After that the Calculate button was pressed. The M values of the least and most stable genes were shown in red and green respectively.

3.13.1.2 Reverse Transcription

Reverse transcription was done using the High Capacity cDNA Reverse transcription kit (Applied Biosystems, USA). The reaction mixture was prepared as shown in Table 3.8.

Table 3.8 Reaction Components for Reverse Transcription

Reaction Components	Volume (µl)	Final Concentration
10X RT Buffer 2 1X	2	1X
25X dNTP mix (100mM)	0.8	1X or 4mM
10X RT Random primers	2	1X
50U/ml MultiScribeTM Reverse transcriptase	1	50U
40U/ml Ambion RNase Inhibitor	0.5	1U/µl
Nuclease free water	-	-
RNA template	10	2mg
Total	**20**	

The thermal cycling profile was as following: 25°C 10 min 37°C 120 min 85°C 5 seconds and hold 4°C.

3.13.1.3 PCR Amplification

Quantitative reverse transcription PCR was carried out using Power SYBR- Green PCR Master Mix kit (Applied Biosystems, USA). The PCR Master Mix was prepared as shown in Table 3.9.

Table 3.9 Reaction Components for Quantitative Reverse Transcription PCR

Reaction Component	Volume (µl)	Final Concentration
2XSYBR GreenPower Master Mix	12.5	1X
Forward primer(10µM)	1.25	500nM
Reverse primer(10µM)	1.25	500nM
cDNA	2	10-100ng
Nuclease free water	8	-
Total	**25**	**-**

The PCR Master mixes were prepared using both β-actin endogenous control strips (the most sTable control gene) as well as pre-miRNA and potential target RNA primers for the sample strips. The master mix containing β-actin primers was added into the endogenous control strips while the master mix containing the pre-miRNA primers were added into the sample strips. cDNA and nuclease free water was then added to the specific strips. For each sample at least three replicates were performed. Then the optical 8 tubes strips was loaded into the Applied Biosystem 7500 Real-Time PCR system and the SDS 1.3.1 (Sequence Detection Software) used to create a relative quantification (ddCt) plate. The run was then started. The amplification program was: 95°C for 10 minutes, 40 cycles of 95°C for 15 seconds and finally 60°C for 1 minute.

After the running of the program, the SDS software was again used to generate a dissociation curve and later a relative quantification (ddCt) study was created to analyze the results. The SDS software uses the equation 2-UUCT to calculate the expression level of the pre-miRNA relative to the endogenous control.

3.13.2 Taqman Assay

Taqman assay was used for analysis of mature miR172 expression in different developmental stages of oil palm, using ath-172a Taqman Assay Kit (Applied Biosystems, USA).

3.13.2.1 MicroRNA Reverse Transcription

TaqMan MicroRNA Reverse Transcription kit (Applied Biosystems, USA) was used for microRNA reverse transcription. The reaction mixture was prepared as shown in Table 3.10.

Table 3.10 Reaction Components for Reverse Transcription

Component	Master Mix (Volume/15µl Reaction)
100mM dNTPs (with dTTP)	0.15
MultiScribe Reverse Transcriptase, 50U/µl	1
10X Reverse Transcription Buffer	1.5
RNase inhibitor, 20U/µl	0.19
Nuclease free water	4.16
Primer (10 µM)	3µl
RNA template (1µg)	5µl
Total	**15µl**

The kit components were thawed on ice. This procedure assumed that miRNA from total RNA samples were quantifying. The tubes were mixed gently and centrifuged to bring the solution to the bottom of the tube. The following thermal cycling profile was used: 16°C for 30 minutes, 42°C for 30 minutes, 85°C for 5 seconds and hold for 4°C.

3.13.2.2 PCR Amplification

PCR Amplification was done using TaqMan MicroRNA assay Kit, ath-172a, (Applied Biosystems,USA). The reaction mixture was prepared as shown in Table 3.11.

Table 3.11 Reaction Components for Real-Time PCR

Component	Master Mix (Volume/15µl Reactions)
Taqman MicroRNA Assay (20X)	1
Product from RT reaction (Minimum 1:15 Dilution)	1.33
Taqman 2X Universal PCR Master Mix. No AmpErase UNG	10
Nuclease free water	7.67
Total	**20**

The thermal cycling profile used was as followed: 95°C for 10 minutes, 40 cycles of 95°C for 15 seconds and finally 60°C for 1minute.

3.14 Computational Target Findings

3.14.1 Potential Target for miR172 in EST of Oil Palm

RNAhybrid software (http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/) (Rehmsmeier et al., 2004), was used to determine the potential AP2-like targets in oil palm. RNAhybrid is an algorithm constructed to find the lowest free energy hybridization between two RNA molecules. This sofware performs in silico hybridizations between a miRNA and a possible target mRNA in a way that optimizes the free energy of the hybridization. It also provides graphics showing the duplex between miRNA and predicted target. The EST of oil palm together with mature miR172 was subjected to RNAhybrid software to find out the complementary site and also the mfe; this energy

should be at least 72% of a perfect match calculated with the same miRNA, and the actual value should be -30 kcal/mol or lower.

Screening criteria for candidate selection were: perfect base pairing of the duplex from nucleotide 8 to 12 counting from the 5'end of the miRNA while loops with a maximum length of one nucleotide in either strands; bulges with no more than one nucleotide in size; and unpaired end overhangs not longer than two nucleotides were allowed (Alves-Junior et al., 2009). The same procedure was done for AP2/miR172 from Arabidopsis to compare the results.

3.14.2 BLAST Analysis

In order to confirm the RNAhybrid results, Homology search using miR172 and EST of oil palm in NCBI through (http://www.ncbi.nlm.nih.gov/) website was done. Blastx was performed using Non-redundant protein sequences database, in order to find any homology between the candidate target protein with other AP2 families.

3.15 5' Rapid Amplification of cDNA Ends (RACE)

RNA ligase mediated 5' RACE was performed to detect miR172 cleavage products by using the GeneRacer Kit (Invitrogen, USA) on RNA sample isolated from mature flower. The dephosphorylation and decapping steps were omitted, so only 5' ends of truncated transcripts could be ligated to the GeneRacer RNA Oligo.

3.15.1 Ligating the RNA Oligo to Truncated mRNA

This oligo is specifically designed to optimize ligation to truncated and decapped mRNA. In particular, it has: (a) Minimal secondary structure to provide a free 3′ end for efficient ligation (b) Adenines at the 3′ end to increase ligation efficiency (Uhlenbeck and Gumport, 1982).

Seven µl of total RNA was transferred to a tube containing the pre-aliquoted, lyophilized GeneRacer™ RNA Oligo (0.25 µg). Pipetting was done several times in order to mix and resuspend RNA Oligo. Brief centrifugation was to collect the fluid in the bottom of the tube. Afterward, to relax the RNA secondary structure, incubation was done at 65°C for 5 minutes. After this step, the total volume of this solution may decrease by 1 µl due to evaporation. Later the tube was place on ice to chill (~2 minutes) followed by brief centrifugation. The following reagents were added to the tube, and mixed gently by pipetting: one µl of 10X Ligase Buffer, 1 µl of 10 mM ATP, 1 µl RNaseOut™ (40 U/µl), 1 µl of T4 RNA ligase (5 U/µl) and the tube was incubated at 37°C for 1 hour. Next, it was again centrifuged briefly and placed on ice.

3.15.2 RNA Precipitation

After ligation, 90 µl DEPC water and 100 µl phenol:chloroform was added to the tube and vortexed vigorously for 30 seconds. After that it was centrifuged at 10,600 x g in a microcentrifuge for 5 minutes at room temperature. The aqueous (top) phase was transferred to a new microcentrifuge tube (~100 µl) and 2 µl 10 mg/ml mussel glycogen, 10 µl 3 M sodium acetate, pH 5.2, was added to the tube and mixed well. Then, 220 µl 95% ethanol was added and the tube was vortexed briefly. Later, the tube was left, at -80°C overnight. After that, to pellet RNA, centrifugation at 10,600 x g for 20 minutes at +4°C was performed. The position of the pellet was noted and supernatant was removed by pipet. Next, 500 µl 70% ethanol was added to the tube,

inverted several times, and vortexed briefly. Centrifugation was done at maximum speed for 2 minutes at +4°C. The ethanol was removed using a pipet without disturbing the pellet. Centrifugation was repeated to collect remaining ethanol. The ethanol was carefully, removed by pipet and the pellet was air-dried for 2 minutes at room temperature. Finally, the pellet was resuspended in 10 µl DEPC water.

3.15.3 Reverse Transcribing mRNA

A SuperScript™ III RT Reaction was used in order to reverse transcribe the RNA. The ollowing reagents were added to the 10 µl of ligated RNA obtained from previous step: one µl GeneRacer™ Oligo dT Primer 5′-GCTGTCAACGATACGCTACGTAACGGCATGACAGTG$(T)_{24}$-3′, 1 µl of dNTP Mix and 1 µl of sterile, distilled water. The tube was incubated at 65°C for 5 minutes to remove any RNA secondary structure and chilled on ice for at least 1 minute and centrifuged briefly. Afterward, the following reagents were added to the 13-µl ligated RNA and primer mixture: 4 µl of 5X First Strand Buffer, 1 µl 0.1 M DTT, 1 µl of RNaseOut™ (40 U/µl), 1 µl SuperScript™ III RT (200 U/µl). At this stage the total volume was 20 µl. Then the reaction was mixed well by pipetting gently. Brief centrifugation was performed and the tube was incubated at 50°C for 30-60 minutes. The RT reaction was then deactivated at 70°C for 15 minutes, chilled on ice for 2 minutes and centrifuged briefly at 10,600 x g in a microcentrifuge. Later, one µl of RNase H (2 U) was added to the reaction mix and was incubated at 37°C for 20 minutes. The reaction mix was then centrifuged briefly and used immediately for PCR amplification.

3.15.4 Amplifying cDNA Ends

At this stage, the reaction was set up to amplify the 5′ end of the gene of interest by using Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen, USA). The gene specific primer (GSP) was designed from the end of the AP2-like target mRNA in oil palm as follows: 5'- GATGTCGGTGGTCTAGAGCATAC -3'. The GeneRacer™ 5′ Primer sequence was: 5′- CGACTGGAGCACGAGGACACTGA -3′. The reaction mixture was prepared as shown in Table 3.12.

Table 3.12 Reaction Components for Amplifying cDNA Ends

Component	5′ RACE (Volume/50µl Reaction)
GeneRacer™ 5′ Primer, 10 µ M	3 µl
Reverse GSP, 10 µM	1 µl
RT template (1µg)	2 µl
10X High Fidelity PCR Buffer	5 µl
dNTP Solution (10 mM each)	1 µl
Platinum® Taq DNA Polymerase High Fidelity, 5U/µl	0.5 µl
MgSO4, 50 mM	2µl
Sterile Water	37.5 µl
Total	**50µl**

The thermal cycling profile used for amplifying the 5’ RACE products was as shown in Table 3.13.

Table 3.13 Thermal Cycling Profile for Amplifying cDNA Ends

Temperature	Time	Cycle
94°C	2 minutes	1
94°C	30 seconds	5
72°C	1 min/ 1 kb DNA	
94°C	30 seconds	5
70°C	1 min/ 1 kb DNA	
94°C	30 seconds	
60°C	30 seconds	20-25
72°C	1 min/ 1 kb DNA	
72°C	10 minutes	1

4. RESULTS AND DISCUSSION

4.1 Confirmation of Putative miR172 Gene Expression in Oil Palm Tissues

Total RNA extraction using a CTAB modified method, RNeasy Mini Kit (Qiagen, Germany) and a Trizol method was carried out on five different oil palm tissues. Only the CTAB modified method gave good quality RNA, so this was found to be the most suitable method for RNA isolation from oil palm.

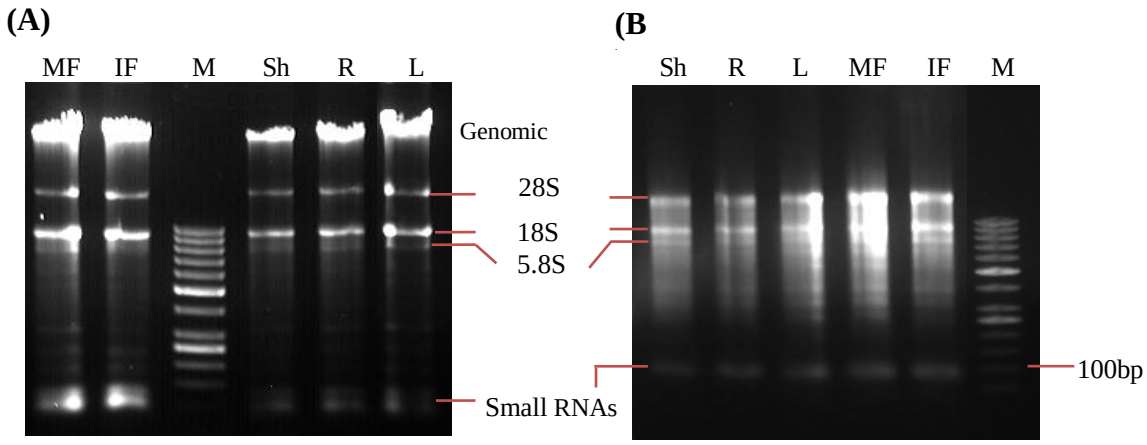


Figure 4.1 RNA Extraction using CTAB Modified Method. (A) Before DNase treatment (B) After DNase treatment. MF: Mature Flower, IF: Immature Flower, Sh: Shoot, R: Root, L: Leaf and M: 50bp Marker.

It is worth noting that there was also one main modification of the CTAB method for oil palm tissues whereby two steps of chloroform:Isoamylalchohol (CI) extraction followed by one step of Phenol:Chloroform:Isoamylalcohol (PCI) extraction were used respectively to eliminate the proteins and polysaccharide contamination from the extract; Moreover, one extra step of CI extraction was included to remove any phenol residues to avoid interference during PCR. This was compared to the original CTAB method (Kiefer et al, 2000) where only two steps of CI extraction were suggested, as mentioned in section 3.2. These extra extraction steps helped to increase the purity of

samples with high polysaccharide contamination, especially from leaf and shoot of oil palm.

The white bulks at the top of the RNA samples as observed in gel electrophoresis (Figure 4.1), show genomic DNA contamination, indicating that this modified CTAB RNA extraction method also extracted large amounts of genomic DNA other than the desired RNA, whilst with the Trizol method no DNA is detectable on the agarose gel. Thus to eliminate the DNA contamination DNase treatment was performed and during this procedure some of the RNA also was lost as expected because of the degradation which usually happens during multiple incubation steps (Figure 4.1 B).

Although the main concern in this study is the small RNAs which can be seen at the bottom of the RNA in the Figure 4.1, the key indicator for confirming that the RNA is intact and can be used for further steps, is the distinct and recognizable bands of 28S and 18S ribosomal RNAs, while in better extracted RNAs even 5.8S rRNA also is detectable. In some cases 5S rRNA also can be seen but mostly because of slight difference in size cannot be differentiated from 5.8S rRNA.

In the next step, extracted RNAs from five different oil palm tissues were subjected to RT-PCR using miR172 specific primers. The presence of miR172 in all five oil palm tissues was confirmed by obtaining the amplicons of the desired size of approximately 200bp on a 1% agarose gel (Figure 4.2). Later these bands were cloned and the presence of the putative miR172 gene transcript was verified via DNA sequencing. The sequences are shown in Appendix 1.

As will be shown later in Figure 4.3, miR172s from *Oryza sativa*, *Sorghum bicolor* and *Populus tricocarpa* were of a similar size as the potential egu-pre-miR172s (miRbase, 2010). It was reported by Wang and colleagues that miRNA precursors in plants generally range from ~50 up to 700 nt in size, while more than 80% are larger than 100 nt (Wang et al., 2007). The forward and reverse primers designed for amplifying

miR172 in oil palm were actually mature miR172 and miR172* sequences located at the two arms of the precursor miR172.

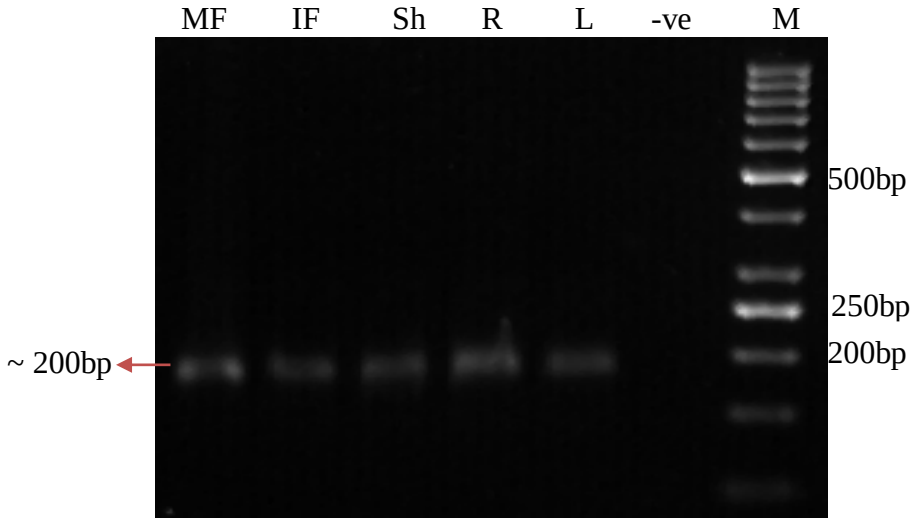


Figure 4.2 Agarose Gel of RT-PCR Products. Reverse transcription PCR of (MF) Mature flower (IF) Immature flower (Sh) Shoot (R) Root (L) Leaf RNA and miR172 primer pairs. (-ve) Negative control and (M) 50bp DNA Marker.

4.2 Potential egu-miR172 Paralogs

Through sequencing of 100 colonies, eight unique candidate sequences were identified and of these, six were predicted to form strong hairpin structures with the mature miR172 sequence on the 3' arm of the helix in all cases (Table 4.1).

All the six candidate paralogs contained mature 21nt miR172 sequences with perfect identity to mature miR172 from other plant families present in miRBase. Two isoforms of the mature miR172 were each represented by three candidates each differing only in the 5' and 3' terminal nucleotides, with mature egu-miR172a, b and c (5' AGAAUCUUGAUGAUGCCGCAU 3') sharing perfect identity with mature ath-miR172a and b, osa-miR172a and d and ptc-miR172f, whilst mature egu-miR172d, e and f (5' UGAAUCUUGAUGAUGCUGCAC 3') were identical to the mature osa-miR172c and sbi-miR172e. Each of the predicted precursors forms a relatively long stem loop structure. In different families, mature miRNAs can be located at either arm of the stem-loop hairpin structure; some at the 5′ end of the miRNA precursor sequences, others at the 3′ end (Zhang et al., 2007c). In this study, the mature sequences from all of the six potential egu-miR172s were located at the 3' arm of the hairpin structures, excluding any bulge or loop through its 21nt sequence. This is consistent with previous studies reporting each miRNA family has a typical location in which the mature miRNAs are always encoded in the same arm of the hairpin (Bartel and Bartel, 2003; Jones-Rhoades, 2010).

Generally, sequences of precursor miRNAs can be distinguished from those of endogenous siRNAs on the basis of the ability of the miRNA-surrounding sequences to adopt a hairpin structure (Wang et al., 2007).

Table 4.1 Predicted Secondary Structures of egu-miR172 Paralogs

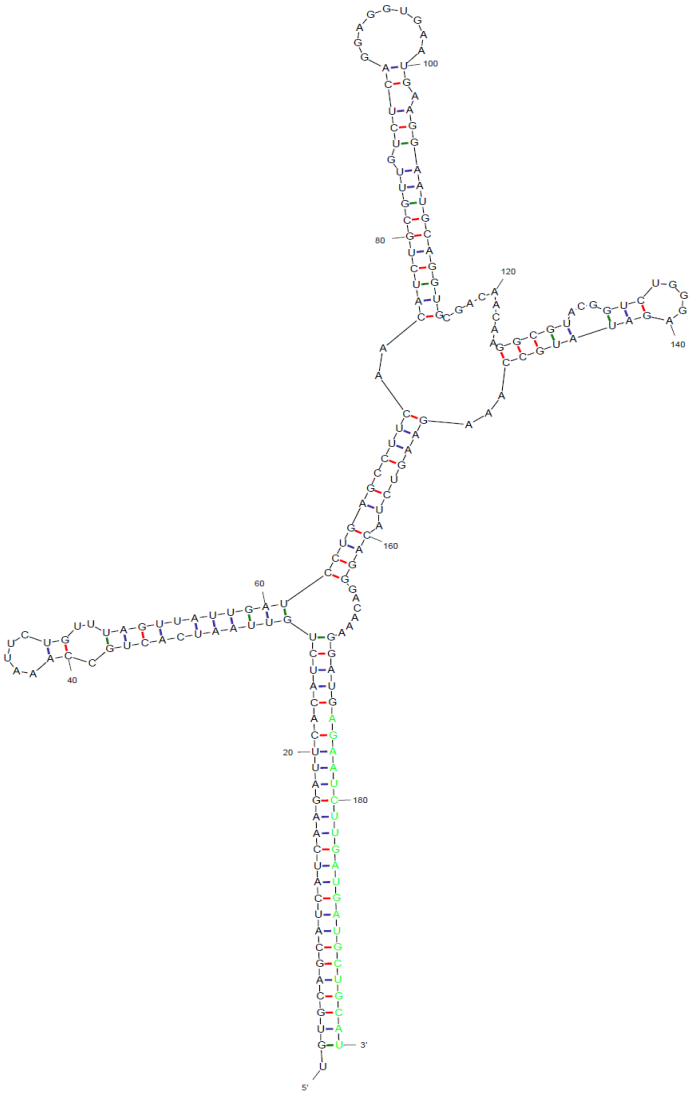
Sequence:	egu-miR172 a: 5'- AGAAUCUUGAUGAUGCCGCAU -3' (21nt)
Minimal Folding Energy (MFE):	-64.57 Kcal/mol
Secondary Structure:	

Table 4.1 (*continues*)

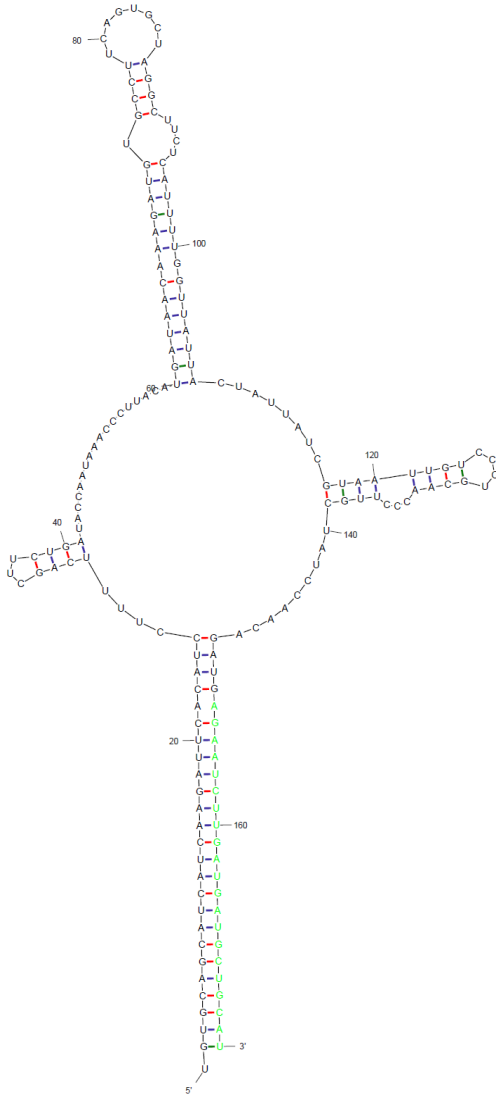
Sequence:	egu-miR172 b: 5’- AGAAUCUUGAUGAUGCCGCAU -3’ (21nt)
Minimal Folding Energy (MFE):	-50.54 Kcal/mol
Secondary Structure:	

Table 4.1 (*continues*)

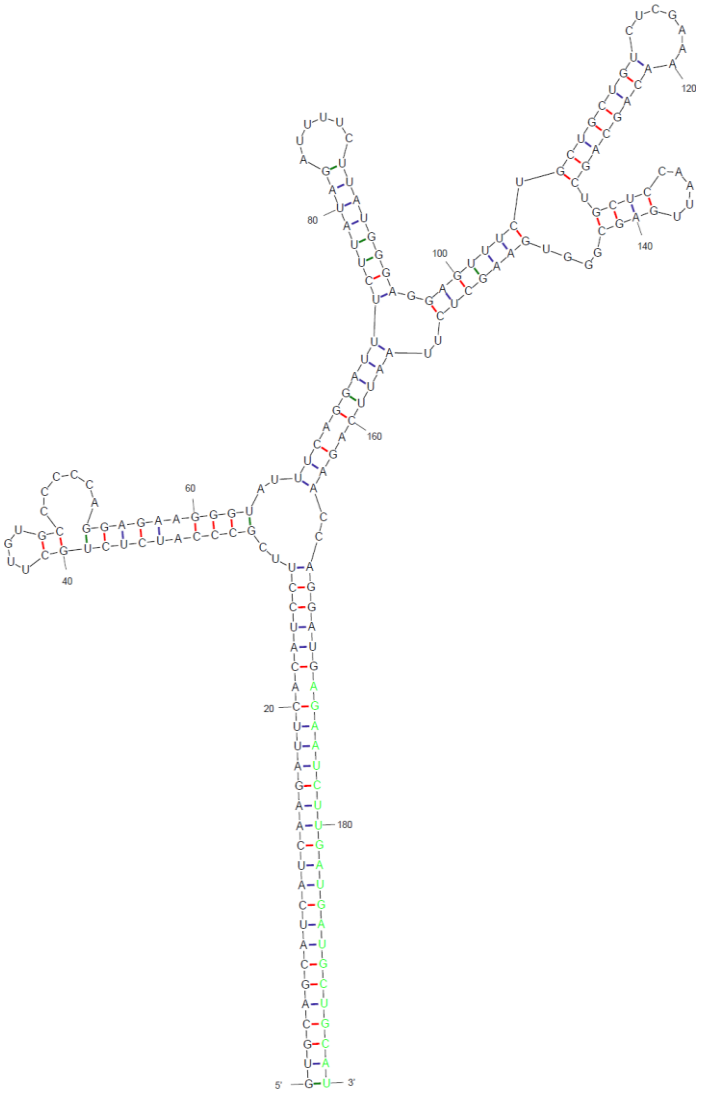
Sequence:	egu-miR172 c: 5’- AGAAUCUUGAUGAUGCCGCAU -3’ (21nt)
Minimal Folding Energy (MFE):	-81.10 Kcal/mol
Secondary Structure:	

Table 4.1 (*continues*)

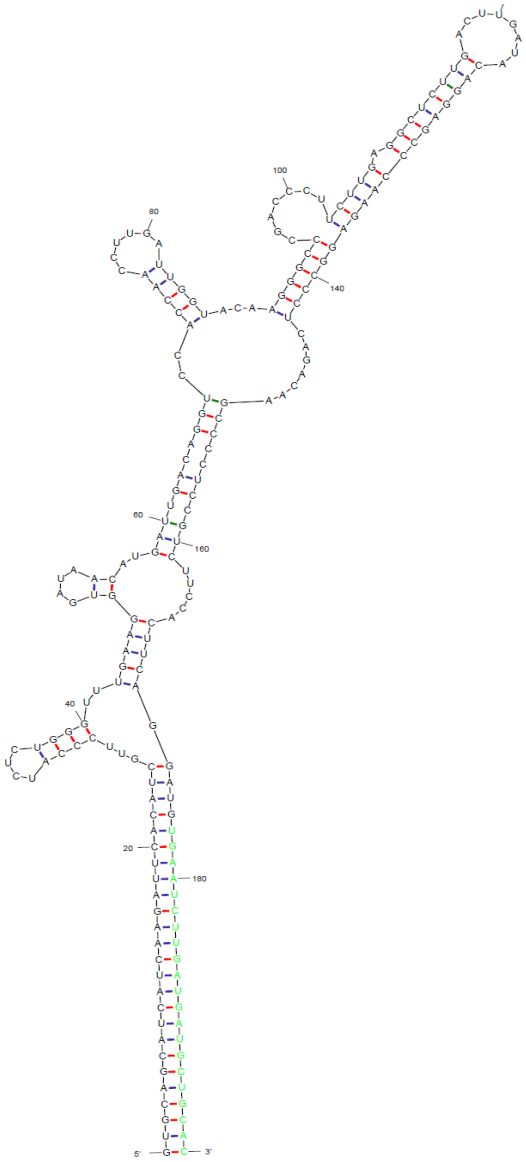
Sequence:	egu-miR172 d: 5’- UGAAUCUUGAUGAUGCCGCAC -3’ (21nt)
Minimal Folding Energy (MFE):	-84.37 Kcal/mol
Secondary Structure:	

Table 4.1 (*continues*)

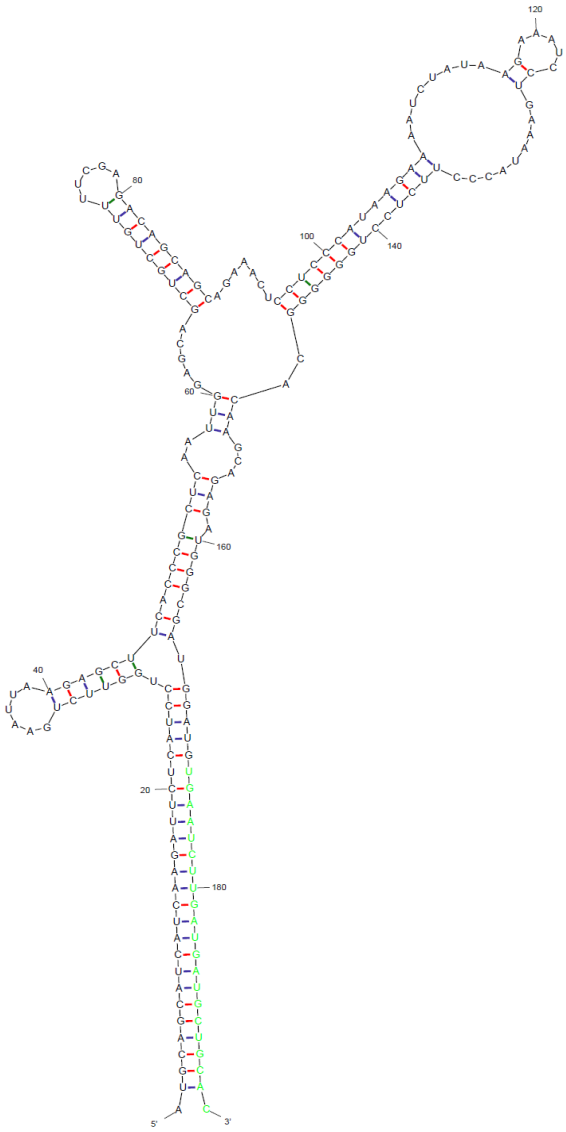
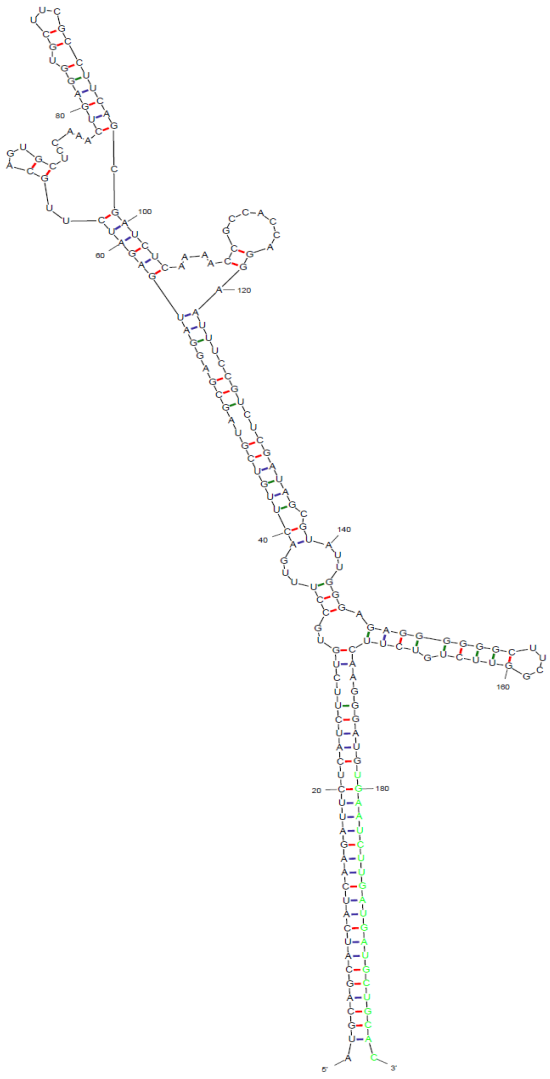
Sequence:	egu-miR172 e: 5’- UGAAUCUUGAUGAUGCCGCAC -3’ (21nt)
Minimal Folding Energy (MFE):	-81.17 Kcal/mol
Secondary Structure:	

Table 4.1 (*continues*)

Sequence:	egu-miR172 f: 5’- UGAAUCUUGAUGAUGCCGCAC -3’ (21nt)
Minimal Folding Energy (MFE):	-73.45 Kcal/mol
Secondary Structure:	

Predicted secondary structures and free folding energies of egu-miR172 precursors determined using mFold software. The green nucleotides indicate mature egu-miR172 sequences within the 3’ arm of a hair-pin loop.

In this study, six criteria which were mentioned by Zhang and colleagues (2006c) were used to confirm the pre-miR172 structures. From the results shown for the predicted secondary structure in the Table 4.1, RNA sequences were considered to be miRNA candidates only if they fitted all of the following criteria: (1) predicted mature miR172s had no more than four nucleotides different from other mature miR172s obtained from miRbase; (2) the RNA sequences could fold into appropriate stem-loop hairpin secondary structures; (3) the mature miRNA sat in one arm of the hairpin structure; (4) there were not more than six mismatches between the predicted mature miRNA sequences and their opposite miRNA* sequences in the secondary structures; (5) no loop or break in the miRNA or miRNA* sequences were observed; (6) predicted secondary structure had higher negative free folding energy (MFE).

According to the results, the length of the cloned sequences for the six potential egu-miR172 paralogs varied between 172-199 nucleotides. Consistent with the miR172 sequences from 20 species deposited in miRbase (2010), some miR172s have short stem-loop sequences while a few are longer than these six predicted egu-miR172s. These six sequences are among the longer miR172 stem loops compared to the majority of miR172 precursors reported to date (miRBase, release 15.0, 2010). So far, the longest miR172 sequence, 277bp, belongs to pre-ata-miR172 (*Aegilops taushii*) (Dryanova et al., 2008) while the shortest one, 74bp, belongs to pre-zma-miR172d (*Zea mays*) (miRbase, release 15.0, 2010).

Generally different sizes of miRNA precursors cause small differences in their secondary structures. The longer precursors may contain other functional elements and present specific roles for miRNA biogenesis or gene expression regulations (Zhang et al., 2006a).

The minimal free folding energy for the six sequences was between -50.54 Kcal/mol for egu-miR172e to -84.37 Kcal/mol for egu-miR172d. Although this minimal free

folding energy usually is lower than other types of RNAs, still this energy depends on the length of the RNA. Longer sequences have lower free folding energy (Zhang et al., 2006c); as examples the predicted free folding energy values are -95.60 Kcal/mol and and -34.80 Kcal/mol for pre-ata-miR172 and pre-zma-miR172d respectively (miRbase, release 15.0, 2010; mFold, 2010). Table 4.1 shows the mature egu-miR172 sequences and their predicted secondary structures.

Plant miRNAs are less conserved than animal miRNAs. Usually, only mature miRNAs are conserved in plants instead of miRNA precursors that are more highly conserved between animals (Bartel, 2004). In this study, ClustalW alignment of the six potential egu-miR172 paralogs with their best matching orthologs from miRBase, indicated that although mature and "star" sequences of miR172 are highly conserved between plant species, the other parts of the pre-miR172 precursors are less conserved (Figure 4.3). ClustalW alignment of three paralogs with the same isoforms also showed low similarity outside of the mature and miRNA* regions (Appendix 3). Previously a similar result was reported for miR156 in cotton (Zhang et al, 2007b).

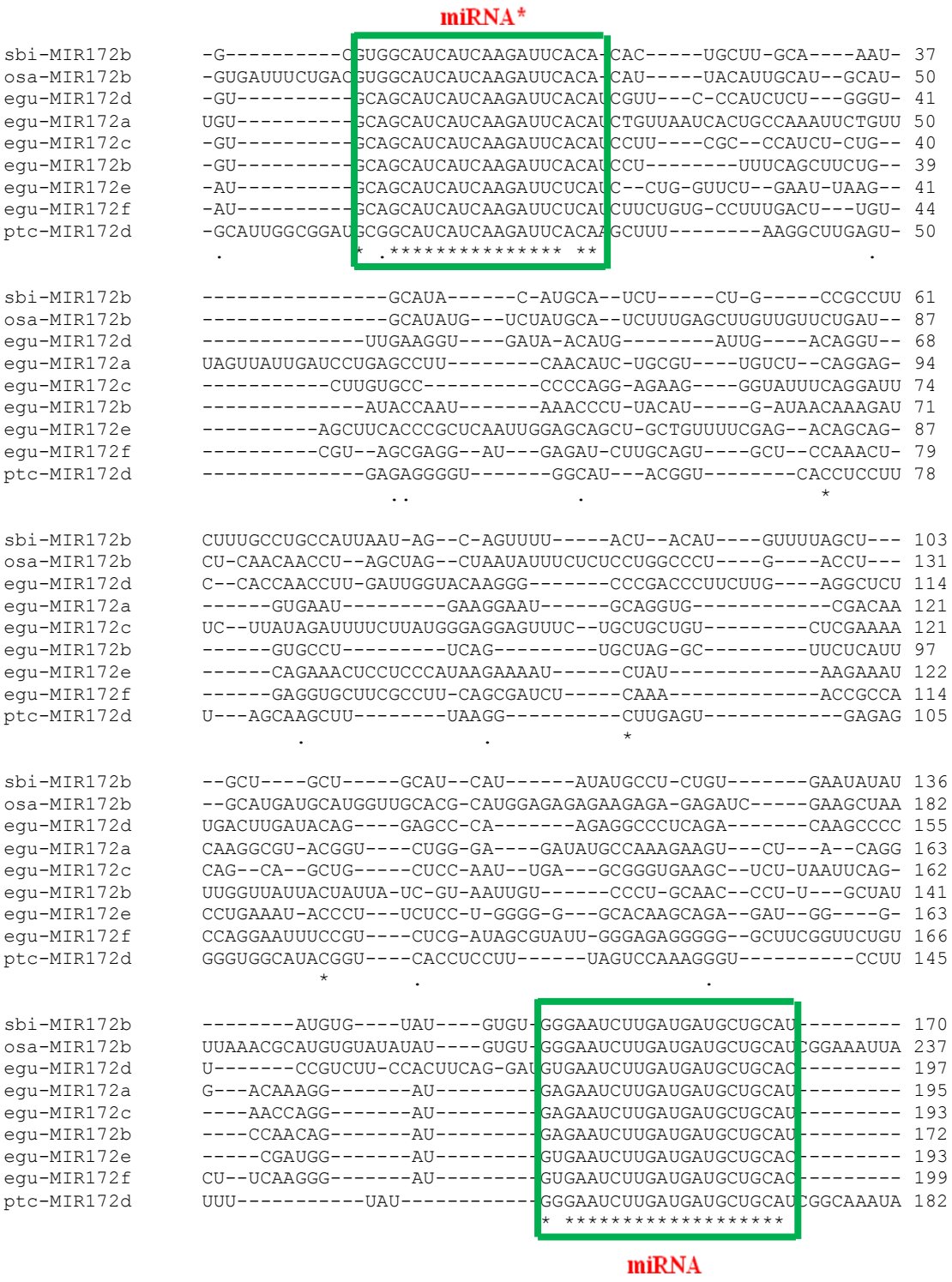


Figure 4.3 Multiple Sequence Alignment of egu-miR172 Paralogs with Their Orthologs from Other Plant Species. Precursor sequences of oil palm miR172 paralogs (egu-MIR172a-f) aligned to miRBase entries osa-MIR172b (*Oryza sativa*), sbi-miR172b (*Sorghum bicolor*) and ptc-miR172d (*Populus trichocarpa*). Boxes indicate the mature miR172 and miR172* sequences.

miR172 seems to be universally expressed among various angiosperms. A subset of this microRNA family is even more ancient, since it was also present in gymnosperms, lycopods and bryophytes (Axtell and Bowman, 2008) which, shows the conservation of this gene among plant species. Furthermore, it is also common within a species that minimum conservation among miRNA family members occurs outside the hairpin loop regions. For example a study on pre-miR166 family members in maize showed that stem conservation was observed between miR166b and miR166d hairpins (Maher et al., 2004). A similar result has been observed in this study among egu-miR172 members.

Small patches of relatively conserved regions observed amongst members of the same miRNA gene families indicate that different types of one miRNA gene may be derived from the same ancestral copy, as was reported for miR395 members in rice (Li and Mao, 2007). The level of identity among miRNA family members shows the evolutionary timescale of their genes. It was reported by Guddeti and colleagues that a cascade of sequence similarity among the miR395 family members specified duplication events which occurred at different times during evolution. Several of them appeared to be rather recent, as they were quite identical (Guddeti et al., 2005). Numerous conserved plant microRNA genes are thought to have resulted from extensive genome duplications and rearrangements, which are considered as a normal basis for microRNA gene family formation, and hence mostly include multiple loci (Voinnet, 2009). miRNA genes of the same family are frequently spread within the genome, suggesting that considerable shuffling have been experienced by plant genomes from the time that the amplification of these families started (Li and Mao, 2007). miRNA gene families may locate on the same chromosome or different ones and they may have different sizes. For example, two MIR168 loci, coding miR168a and b, have been recognized in the genomes of Arabidopsis and poplar. miR168a and

miR168b are located on chromosome 4 and 5 in Arabidopsis respectively, whilst they are located on linkage group III (LGI) and scaffold 86 in poplar. It was suggested that independent duplication events may have resulted in the MIR168 family (Reinhart et al., 2002; Gazzani et al., 2009).

In rice, miR169n and miR169o were recognized as high salinity-responsive miRNAs and their genes were located within an miRNA cluster with a distance of 3,707 base pairs. The high degree of conservation and close phylogenic distance of pre-miR-169n and pre-miR-169o demonstrated that both of them originated from a very recent tandem duplication event (Zhao et al., 2009).

Studies on the miR172 family and its targets showed that this miRNA has a specific function in wheat and includes an additional target interaction with the mRNA for succinyl-CoA ligase, which implies that the function and expression of this miRNA have drifted through long evolutionary timescales (Yin and Shen, 2010).

Since the genome information and loci of miR172 in oil palm is not yet published, it is not yet possible to comment on the evolutionary event in this plant; however considering other studies (Axtell and Bowman, 2008; Yin and Shen, 2010) together with low similarity among the six different paralogs obtained in this study, which is limited to the miRNA and miRNA* regions (Figure 4.3), it seems likely that the evolution of this miRNA goes back to a long time ago as was reported in wheat (Yin and Shen, 2010).

4.3 Expression Analysis of miR172 in Oil Palm using Nested PCR

With the aim of finding whether each of the egu-miR172 precursors were expressed in all five oil palm tissues or if they were tissue specific, six sets of interior primers were used for nested PCR on oil palm shoot, root, leaf, mature and immature flowers. The nested PCR results showed amplified fragments of approximately 100 to 150bp for all samples on a 1% agarose gel (Figure 4.4). the size of amplicon for primer set one was 153bp while primer sets three and five resulted in amplicons with a size of 151bp. the size of amplicon using primer set four was 100bp while the lengths of amplified fragments using primer sets two and six were respectively 106bp and 108bp (<110bp). All of the PCR products were of the expected size, based on the sequence of each cloned paralog, which indicated that all of the six egu-miR172 isoforms were expressed in all vegetative and reproductive tissues in oil palm (Figure 4.4).

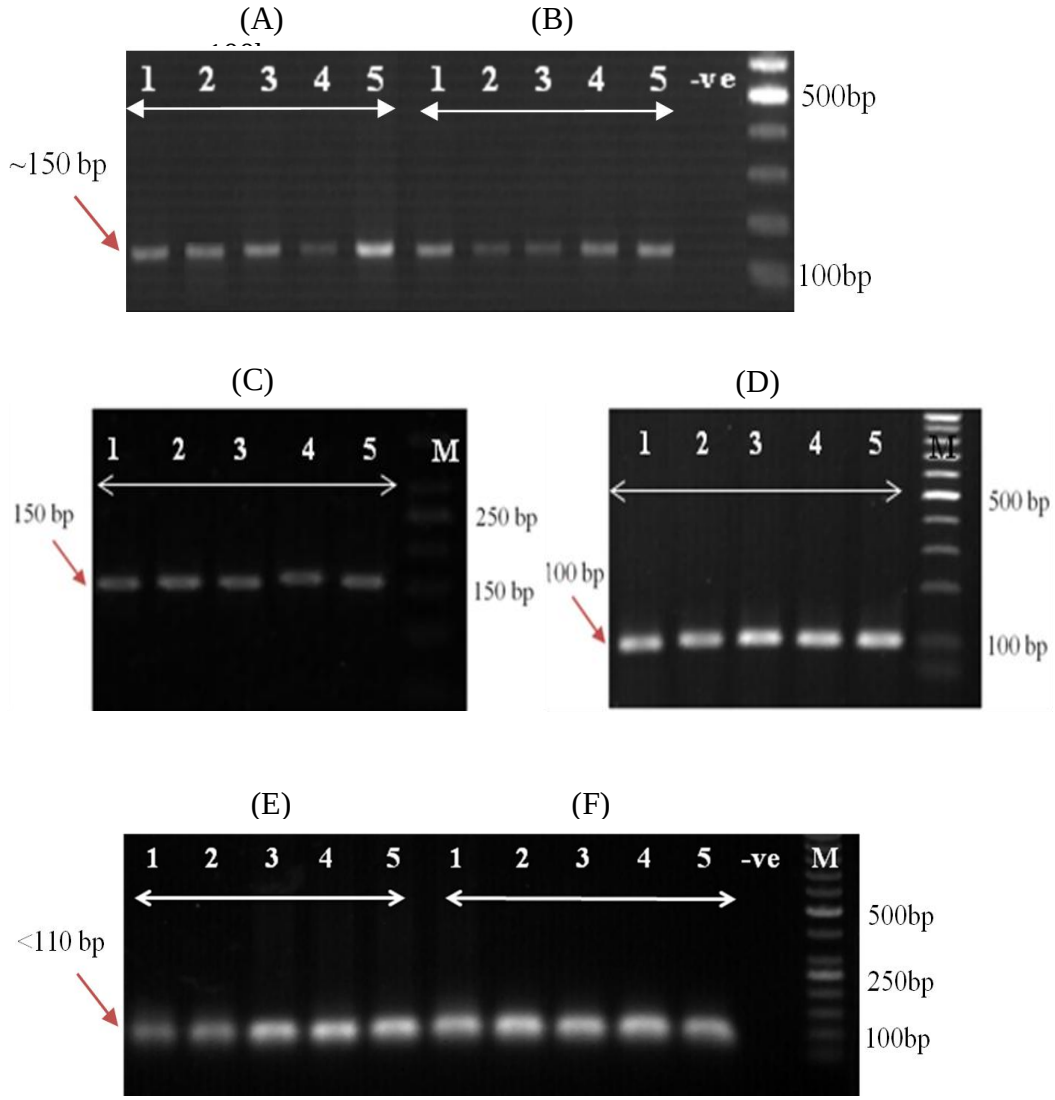


Figure 4.4 Nested PCR Amplification of egu-pre-miR172. 1: root, 2: leaf, 3: shoot, 4: mature flower, 5: immature flower using six nested primer sets to amplify six egu-pre-miR172 isoforms. -ve: Negative control and M: 50bp DNA Marker. Primers used were: (A) egu-pre-miR172a (B) egu-pre-miR172e (C) egu-pre-miR172c (D) egu-pre-miR172d (E) egu-pre-miR172b and (F) egu-pre-miR172f.

4.4 Quantitative Reverse Transcription PCR Analysis of miR172 Precursors

4.4.1 Selection of Internal Control

In order to select an appropriate endogenous control gene for quantitative reverse transcription PCR with SYBR green, cDNA from different oil palm tissues was amplified using β-actin, eEF1-α, UBQ5 and 18S rRNA primers to select a set of housekeeping gene primers that had the most constant expression in the all of the oil palm tissues. Housekeeping genes are essential for basic cell survival and the transcription of these genes in the cells is quite stable in different tissues (Pfaffl, 2001). The geNorm VBA applet for Microsoft Excel (Windows version), verifies the most stably expressed reference genes from a set of tested genes in a given cDNA sample panel (Vandesompele et al., 2002). The gene expression stability measure, M, for reference genes were calculated using this software. This M value should be less than 1.5 and among the controls, those with smallest M value, show the most stable and constant genes which will be the most suitable endogenous controls in quantitative reverse transcription PCR. Results from geNorm analysis showed that in the oil palm samples used in this study, both the β-actin and 18S rRNA were suitable endogenous controls for quantitative real time PCR. Since the M value of β-actin was the lowest, it was selected as the endogenous control for quantitative reverse transcription PCR using SYBR-Green (Table 4.2).

Table 4.2 Average Expression Stability Values (*M*) of Four Control Genes

Data	β-actin	18S rRNA	eEF1-α	UBQ5
Shoot	2.94E+01	2.70E+01	2.30E+01	3.40E+01
Root	3.00E+01	2.90E+01	3.50E+01	2.10E+01
Leaf	3.22E+01	3.10E+01	3.10E+01	2.70E+01
Mature Flower	3.05E+01	3.12E+01	2.50E+01	3.24E+01
Immature Flower	3.10E+01	3.20E+01	3.00E+01	3.10E+01
M < 1.5	**0.196**	**0.198**	**0.323**	**0.366**

4.4.2 Expression of miR172 Precursors in Oil Palm

To study the expression pattern of the six miR172 precursors in oil palm tissues, quantitative reverse transcription PCR analysis relative to β-actin was carried out using SYBR green, which results in the detection of all double stranded DNA. A dissociation assay was generated immediately after the relative quantification run to determine whether non-specific amplification products were present. The dissociation curve showed a single peak for all the six egu-miR172 precursors and β-actin, representing that all the six pairs of nested miR172 and internal control primers used in this study were specific (Appendix 4).

Among the six predicted miR172 paralogs, egu-miR172a, b, c and d had the highest accumulation in mature flowers with significant difference compared to other tissues. Although egu-miR172e and f were the most highly expressed in immature flowers and this expression was significantly higher than vegetative tissues, there was no significant difference against mature flower. The lowest expression for all the six paralogs was observed in the vegetative tissues (root, shoot and leaf) as shown in Figure 4.5.

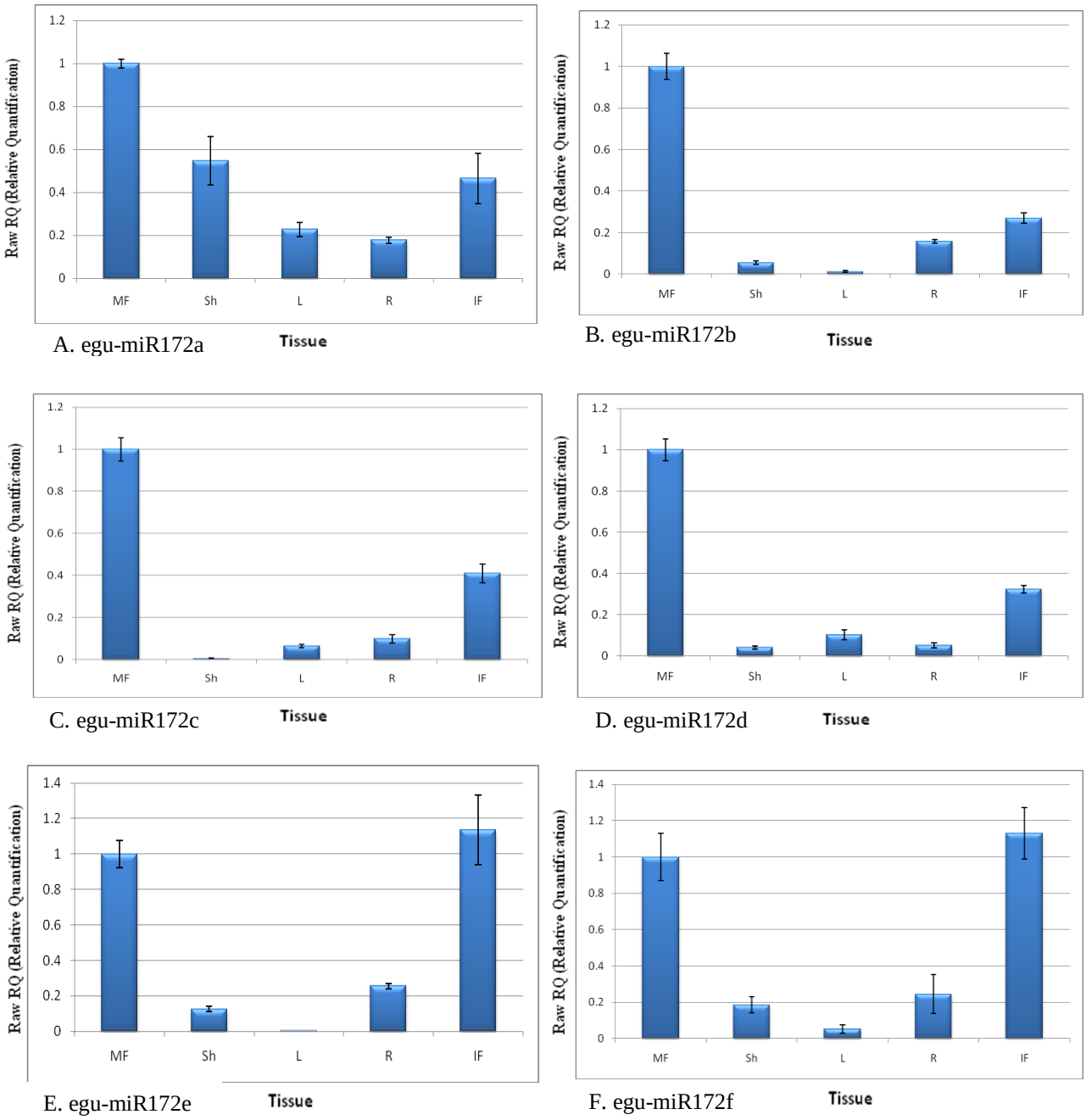


Figure 4.5 Quantitative Reverse Transcription PCR Expression Profiles of egu-miR172 Paralogs. Five oil palm tissues were examined; MF: mature flower; Sh: shoot; L: leaf; R: root; IF: immature flower. Relative expression levels are to a beta-actin internal control and have been normalized to the mature flower (MF) expression level, which is set as 1. Each bar represents the mean concentration level (RQvalue) from three replicates and the vertical lines indicate standard error of the mean concentration level. Note: paralogs a, b and c are one isoform (isoform 1) whilst d, e and f are the other isoform (isoform 2).

All the six egu-miR172 paralogs showed broadly similar expression profiles in the five tissues examined, with relatively high levels of expression in floral tissues (Figure 4.5). A comparison of the various paralogs revealed differences in relative expression in immature flowers, showing a roughly four-fold difference between egu-miR172b and egu-miR172e. There were also similar differences in expression between various paralogs for the non-reproductive tissues most notably for shoot (egu-miR172a vs egu-miR172c) and leaf (egu-miR172a vs egu-miR172e), suggesting that different paralogs are tissue specific and may have different regulatory roles.

In a wide range of organisms, many miRNAs are differentially expressed at various developmental stages and in different tissues and cell types (Wang et al., 2007; Matts et al., 2010). For instance, several studies showed that miR165/166 is limited to the abaxial side of leaves in maize and Arabidopsis (Chen, 2005; Liu et al., 2009) and the level of miR156 accumulated in both the upper and the lower sets of seedling leaves but was almost undetectable in similar sets of mature switchgrass plant leaves (Matts et al., 2010). miR172 showed higher expression in later stage vegetative tissues and developing young panicles in rice, while its accumulation was below the detection limit in 10 days after fertilization in grains (Zhu et al., 2009). It was also revealed that, in Arabidopsis, the level of miR159 is elevated by gibberellin (GA) (Liu and Chen, 2009) and that miR164 levels change in response to auxin. This miRNA regulates auxin signaling to reduce lateral root production in Arabidopsis (Guo et al., 2005; Simon et al., 2009). The expression pattern of miR164c was investigated using a GUS reporter gene from the MIR164c promoter in wild-type plants. GUS reporter gene is a tool to measure a gene activity in transgenic plants. During plant development the promoter of miR164c causes active expression of the GUS reporter gene in Arabidopsis (Baker et al., 2005), signifying that miRNA gene transcriptions are regulated.

Little is known whether the expression or function of the members within the same miRNA family is different in plants (Jiang et al., 2006); however, there are examples of specific expression differences in tissue/ cell type, developmental stages or environmental conditions among miRNA individuals within same gene families. For instance, in maize, the expression of all eight members of miR396 family is generally higher in the juvenile root and seedling but only particular family members (miR396a, miR396b and miR396g) are highly expressed in pollen (Zhang et al., 2009). Another study by Mica and colleague (2009) on grapevine (*Vitis vinifera*), showed that, miR171e transcripts were detected only in callus, miR171f was only transcribed in stem while miR171g was observed in callus and root. A similar condition was observed for some other families including: miR166 and miR167 (Mica et al., 2009). Furthermore, miR319c, e and f are all expressed in stem, whilst miR319c and g are expressed in callus in grapevine. These data suggested that tissue specific expression of different precursors within single miRNA families is widespread in this plant (Mica et al., 2009). Out of seventeen miR169 deposited in miRbase, three miR169 members (miR169g and miR169n and miR169o) were identified as high salinity-responsive miRNAs in rice (Zhao et al., 2009). In addition, overexpressed-miR172b delayed the transition from spikelet to floral meristem, and caused significant floret defects and reduced fertility in rice (Zhu et al., 2009).

The data presented here, could not exactly distinguish these possibilities for oil palm but did suggest that miR172 may be one of these tissue specific miRNAs and it is likely to be involved in floral organ identity and flowering time in oil palm.

4.5 miR172 Target Prediction

Two mature miR172 isoforms each represented by three candidates (egu-miR172a, b and c) (isoform1) and (egu-miR172d, e and f) (isoform2) were compared against the *Elaeis guineensis* EST database using RNAhybrid software (http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/submission.html). Prediction of miRNA targets was limited to EST sequences in the GenBank database as the complete genome of oil palm has not been published at this time. Nine EST sequences were identified as potential miR172 target transcripts (Table 4.3), two of which (GenBank: EL692343.1 and EY408858.1), are 99% identical. These two EST share 46-58% translated amino acid sequence identity with AP2 and AP2-like transcription factors from several plant species. ClustalW alignment showed that EY408858.1 is a truncated version of EST EL692343.1 with two nucleotide differences outside of the predicted miR172 target region (Appendix 5).

Table 4.3 Predicted miR172: mRNA Target Hybrids

Predicted Hybrid		mfe (kcal/mol)	Length of EST (bp)	Target position in EST	Predicted function of EST (%aa identity)
EY401361.1 Eug-miR172a,b,c	5' A G3' GCAGCA CAUCAGGAUUU CGUCGU GUAGUUCUAAG 3' UA A A5'	-30.5	523	423	Mitochondrial phosphate transporter Protein from several plants (~65%)
EL692343.1 Eug-miR172a,b,c	5'C C 3' UGCAGCAUCAUCAGGAUUCU ACGUCGUAGUAGUUCUAAGA 3'U 5'	-38.8	662	484	AP2 and AP2 like transcription factors from several plant species (46-58%)
EL683787.1 Eug-miR172a,b,c	5' A C 3' GCAG AUCAUCAGGGUUCU CGUC UAGUAGUUCUAAGA 3' UA G 5'	-32.4	848	594	No significant match to any known protein
EY401361.1 Eug-miR172a,b,c	5' A G3' GCAGCA CAUCAGGAUUU CGUCGU GUAGUUCUAAG 3' UA A A5'	-30.5	523	423	Mitochondrial phosphate transporter Protein from several plants (~65%)

Table 4.3 ***(Continued)***

Predicted Hybrid		mfe (kcal/mol)	Length of EST (bp)	Target position in EST	Predicted function of EST (%aa identity)
EL692343.1 Eug-miR172d,e,f	5'C C 3' UGCAGCAUCAUCAGGAUUCU ACGUCGUAGUAGUUCUAAGA 3'U 5'	-37.7	662	484	AP2 and AP2 like transcription factors from several plant species (46-58%)
EL689180.1 Eug-miR172d,e,f	5' C G A 3' GUGCGGCGUCGUCG GA UCA CACGUCGUAGUAGU CU AGU 3' U A 5'	-31.0	857	552	Hypothetical protein from different plants
EL683764.1 Eug-miR172d,e,f	5' C U 3' UGCAGCAUCAUCAGGAUUC ACGUCGUAGUAGUUCUAAG 3' C U 5'	-37.7	130	6	No significant match to any known protein
EY399139.1 Eug-miR172d,e,f	5' U 3' UGCAGCAUCAUCA GUUCA ACGUCGUAGUAGU UAAGU 3' C UC 5'	-31.1	144	126	Conserved hypothetical protein [*Ixodes Scapularis*]

Table 4.3 ***(Continued)***

Predicted Hybrid		mfe (kcal/mol)	Length of EST (bp)	Target position in EST	Predicted function of EST (%aa identity)
EY401258.1 Eug-miR172d,e,f	5'A A A 3' GCAGUAUCAUCAAGGU CA CGUCGUAGUAGUUCUA GU 3' CA A 5'	-31.9	661	613	Hypothetical protein from different plants
EY409361.1 Eug-miR172d,e,f	5' A G 3' GCAGCA CAUCAGGAUUU CGUCGU GUAGUUCUAAG 3' UA A A 5'	-30.4	523	423	Mitochondrial phosphate travsporter protein from several plants (~65%)

Potential hybrid structures of miR172:miR172 targets from oil palm EST predicted with RNAhybrid. Alignment with EST EY4088581 not shown as this EST is a truncated version of EST EL692343.1 with a single nucleotide difference which is outside of the predicted miR172 target region. Predicted functions of EST are based on BLASTX identities.

A homology search using mature miR172 sequence (isoform 1) against ESTs of oil palm was conducted using BLAST program in National Center for Biotechnology Services (http://www.ncbi.nlm.nih.gov). It also verified that the identity between miR172 (Query) and the complementary site of the EST EL692343.1 (Subject) was 20/21nt (95%) (Figure 4.6).

The deduced amino acid sequence of the EST EL692343.1 was compared against the non-redundant protein sequence (nr) database using blastx, which confirmed significant alignments with the AP2 super family (Appendix 6) and showed close relation to other AP2-like genes and the highest to HAP2 from *Hyacinthus orientalis* (Appendix 7).

(A)

>gi|161970769|gb|EL692343.1|EL692343 OPR02383 Elaeis guineensis root Elaeis guineensis cDNA clone OPVRXX_1913A.seq, mRNA sequence

5'- ATACAATACTTGCGTGCAGGTACATATATCTTGGGCTATTCGACAGCGAAATAGAAGCTGCAAGGG
CTTATGACAAGGCAGCTATAAAATGTAATGGAAGGGATGCTGTTACCAATTTTGAGCCCAGCACCTATGA
AGGAGAACTGCTTACTGAGGCTAATAGTGAAGCAACTGGCCATGATGTTGATCTGAACTTGAGGATTTCT
CAACCTGTTGCCCATAGTCCAAAGAAGGATCACAATTCAATAGGCATCCAATTCCACTATGGCTTGCTTG
AATCTTCGGATGCCAAGAAAGTAACGATTGACAGCACGTCTTCCCAATTGGCTGGTCAGCCGCATCATGT
ATGGACTGCTCAACGCCCAGCTCTCTTTCCTACCATTGAGGAAGGAGCAAGAGAGAAGAGGCTTGAGGTT
GGTTCTCAAGCTCTACCCGCCTGNGCGTGGCAGATGCATGGCCCTACTCCATTGCCGCTGTTCTCTT***CTG***
CAGCATCATCAGGATTCTCAACTACTGCTGTGACATCTGCCCATCCTTCCTCGCTGCCTCCTCCTTCGGC
AACCCTGCATTCCCAGTTCAGCCTACCGGCTCCATCCAATTTTCGATTCAGGAGCTGAAATCCATGTGCT
TGGACCAGTGTATGTATGCTCTAGACCACCGACATC -3'

(B)

> gb|EL692343.1| OPR02383 Elaeis guineensis root Elaeis guineensis cDNA clone
OPVRXX_1913A.seq, mRNA sequence.
Length=662

```
 Score = 32.2 bits (16),  Expect = 0.027
 Identities = 19/20 (95%), Gaps = 0/20 (0%)
 Strand=Plus/Minus

Query  1    AGAATCTTGATGATGCTGCA  20
            |||||| ||||||||||||
Sbjct  504  AGAATCCTGATGATGCTGCA  485
```

Figure 4.6 Candidate miR172 Target from Oil Palm. (A) Homology search using mature miR172 against EST of oil palm through NCBI, resulted in 662bp, partial sequence which has a complementarity site with the mature miR172 (bold underlined bases). (B) The sequence comparison conducted through nucleotide database search using BLAST program in NCBI.

The predicted miR172: EL692343.1 hybrid had the lowest minimal free energy (mfe) value among the predicted target hybrids (-38.8 kcal/mol) and showed complete identity to the ath-miR172:AP2 hybrid of *Arabidopsis thaliana* reported by Park et al., 2002; both hybrids having just one unmatched nucleotide (Figure 4.7).

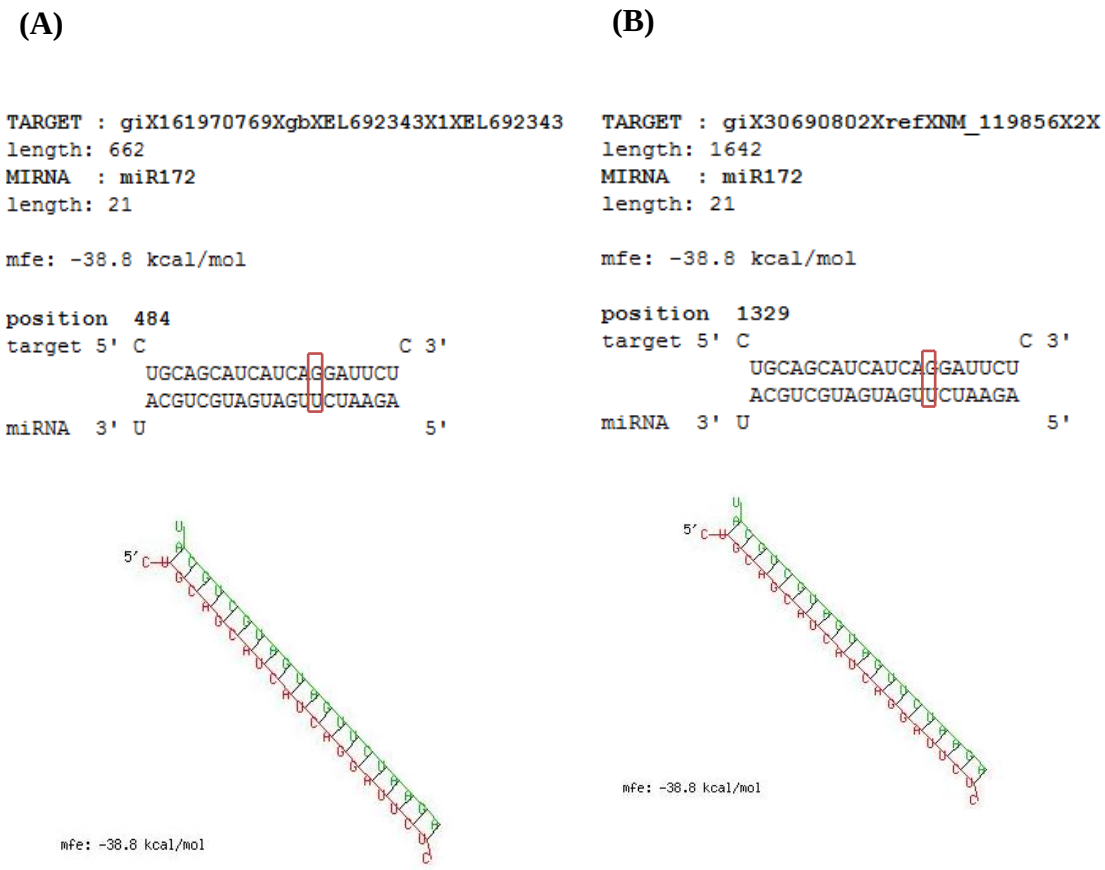


Figure 4.7 The Predicted miR172: AP2-like Targets Hybrid. (A) Shows the complementary site between miR172 and its target candidate in oil palm. (B) The complementary site of miR172 and its AP2 target in *Arabidopsis thaliana*.

Several AP2-like target genes have been recognized in different plants. In Arabidopsis, six AP2-like transcription factors act as floral repressors. This clade of proteins includes APETALA 2 (AP2), TOE1, TOE2, and TOE3, SMZ and its paralog SNZ (Mathieu et al., 2009), while in maize glossy15 was recognized as AP2-like genes (Aukerman and Sakai, 2003). The rice miR172 family contains four members (miR172a-d), which are predicted to target five AP2-like genes (Os03g60430, Os04g55560, Os05g03040, Os06g43220 and SNB) (Zhu et al., 2009). miR172 also targets AP2-like family members in several plant species, such as, wheat (Yao et al., 2007), tomato (Itaya et al., 2008), soybean (Subramanian et al., 2008), and Ipomoea nil (Glazińska et al., 2009).

The results presented here, confirmed that the EST EL692343.1 is a good candidate as an AP2-like family member, but beside this, there are also other candidates predicted among oil palm ESTs by RNAhybrid software (Table 4.3), which are likely to be other miR172 targets, although their predicted functions (from the EST annotation) were not of AP2-like family members as they were only annotated as hypothetical proteins or sequences with no significant matches to any known sequence, which may point out them as other putative miR172 targets which have not yet been recognized.

From the supplementary data provided by Alves-Junior et al., (2009), fifteen possible targets with different predicted functions were recognized through RNAhybrid software and using the Arabidopsis ath-miR172a sequence. As an example, a target EST (AT4G24630.1) annotated as a zinc ion binding protein was among these fifteen predicted targets. In comparison, RNAhybrid analysis of the rice genome using zma-miR172b resulted in 72 potential targets. Among them there were various functional descriptions, such as hypothetical proteins, nodulin-like protein, no specific function, etc. http://bibiserv.techfak.uni-bielefeld.de/cgi-bin/rnahybrid_tdb_mirnas.cgi (Alves-Junior et al., 2009).

4.6 Experimental Validation of Predicted AP-like Target

For experimental validation of the predicted oil palm AP2-like target (EST, EL692343.1) reverse transcription PCR was performed on shoot, root, leaf, mature and immature inflorescences of oil palm. The reported length of this EST was 662bp (EST, EL692343.1), while the length of the new amplified fragment using a specific primer set was 571bp (~90bp less than the sequence reported for EST EL692343.1). Reverse transcription-PCR using the target specific primers showed that the EL692343.1 mRNA was indeed present in all five tissues of oil palm. A band of the expected size (< 600bp) was seen on a 1% agarose gel (Figure 4.8).

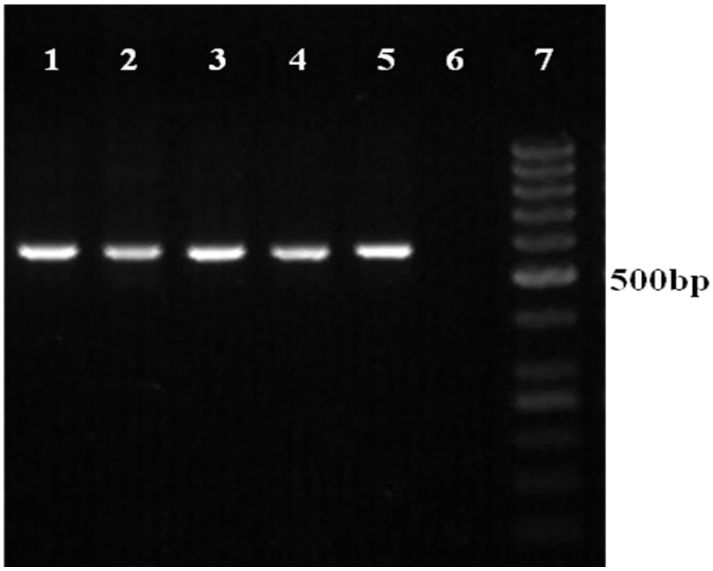


Figure 4.8 Agarose Gel (1%) of RT-PCR Products. Reverse transcription-PCR of (1) Mature flower (2) Immature flowers (3) Shoot (4) Root (5) Leaf (6) Negative control (7) 50bp ladder

Later for confirmation of the correct amplified fragment, the sequences were subjected to sequencing and then the results were aligned with the oil palm EST EL692343.1. The alignment showed almost full complementary between the RT-PCR product and the potential target RNA fragment from oil palm, with just 3 mismatches (99.5% identity) (Figure 4.9).

```
CLUSTAL W (1.83) multiple sequence alignment

Query                                 ------------------CGTACATATATCTTGGGCTATTCGACAGCGAA 32
gi|161970769|gb|EL692343.1|EL6        ATACAATACTTGCGTGCAGGTACATATATCTTGGGCTATTCGACAGCGAA 50
                                                        ******************************

Query                                 ATAGAAGCTGCAAGGGCTTATGACAAGGCAGCTATAAAATGTAATGGAAG 82
gi|161970769|gb|EL692343.1|EL6        ATAGAAGCTGCAAGGGCTTATGACAAGGCAGCTATAAAATGTAATGGAAG 100
                                      **************************************************

Query                                 GGATGCTGTTACCAATTTTGAGCCCAGCACCTATGAAGGAGAACTGCTTA 132
gi|161970769|gb|EL692343.1|EL6        GGATGCTGTTACCAATTTTGAGCCCAGCACCTATGAAGGAGAACTGCTTA 150
                                      **************************************************

Query                                 CTGAGGCTAATAGTGAAGCAACTGGCCATGATGTTGATCTGAACTTGAGG 182
gi|161970769|gb|EL692343.1|EL6        CTGAGGCTAATAGTGAAGCAACTGGCCATGATGTTGATCTGAACTTGAGG 200
                                      **************************************************

Query                                 ATTTCTCAACCTGTTGCCCATAGTCCAAAGAAGGATCACAATTCAATAGG 232
gi|161970769|gb|EL692343.1|EL6        ATTTCTCAACCTGTTGCCCATAGTCCAAAGAAGGATCACAATTCAATAGG 250
                                      **************************************************

Query                                 CATCCAATTCCACTATGGCTTGCTTGAATCTTCTGATGCCAAGAAAGTAA 282
gi|161970769|gb|EL692343.1|EL6        CATCCAATTCCACTATGGCTTGCTTGAATCTTCGGATGCCAAGAAAGTAA 300
                                      ********************************** ****************

Query                                 CGATTGACAGCACGTCTTCCCAATTGGCTGGTCAGCCGCATCATGTATGG 332
gi|161970769|gb|EL692343.1|EL6        CGATTGACAGCACGTCTTCCCAATTGGCTGGTCAGCCGCATCATGTATGG 350
                                      **************************************************

Query                                 ACTGCTCAACGCCCAGCTCTCTTTCCTACCATTGAGGAAGGAGCAAGAGA 382
gi|161970769|gb|EL692343.1|EL6        ACTGCTCAACGCCCAGCTCTCTTTCCTACCATTGAGGAAGGAGCAAGAGA 400
                                      **************************************************

Query                                 GAAGAGGCTTGAGGTTGGTTCTCAAGCTCTACCCGCCTGGGCGTGGCAGA 432
gi|161970769|gb|EL692343.1|EL6        GAAGAGGCTTGAGGTTGGTTCTCAAGCTCTACCCGCCTGNGCGTGGCAGA 450
                                      ************************************ **********

Query                                 TGCATGGCCCTACTCCATTGCCGCTGTTCTCTTCTGCAGCATCATCAGGA 482
gi|161970769|gb|EL692343.1|EL6        TGCATGGCCCTACTCCATTGCCGCTGTTCTCTTCTGCAGCATCATCAGGA 500
                                      **************************************************

Query                                 TTCTCAACTACTGCTGTTACATCTGCCCATCCTTCCTCGCTGCCTCCTCC 532
gi|161970769|gb|EL692343.1|EL6        TTCTCAACTACTGCTGTGACATCTGCCCATCCTTCCTCGCTGCCTCCTCC 550
                                      ***************** ********************************

Query                                 TTCGGCAACCCTGCATTCCCAGTTCAGCCTACCGGCTCC----------- 571
gi|161970769|gb|EL692343.1|EL6        TTCGGCAACCCTGCATTCCCAGTTCAGCCTACCGGCTCCATCCAATTTTC 600
                                      ****************************************

Query                                 ---------------------------------------------------
gi|161970769|gb|EL692343.1|EL6        GATTCAGGAGCTGAAATCCATGTGCTTGGACCAGTGTATGTATGCTCTAG 650

Query                                 ------------
gi|161970769|gb|EL692343.1|EL6        ACCACCGACATC 662
```

4.9 Sequence Alignment of 571bp RT-PCR Product (Query) with EST EL692343.1. ClustalW alignment of EL692343.1 and the RT-PCR amplified product from oil palm.

4.7 Expression Profiles of miR172 and its Potential Target Gene

To determine transcript expression patterns of miR172 and its predicted target during development of oil palm, expression of miR172 and its potential AP2-like target mRNA were analyzed using quantitative reverse transcription PCR which is a highly sensitive type of PCR (Dugas and Bartel, 2008).

The results presented here showed that in oil palm, the expression of miR172 depends on the age of the plant (Figure 4.10 A and B). This is consistent with previous studies in Arabidopsis which demonstrated that the accumulation of miR172 is depended on the age of the plant (Jung et al., 2007; Wu et al., 2009). Also, in *Ipomoea nil*, miR172 levels increased with the age of the seedling (Glazin´ska et al., 2009).

The expression of miR172 varied considerably between organs and developmental stages in oil palm. Mature miR172 accumulation increased significantly in vegetative tissues (leaves, roots and shoots) but not in eleven and twelve month old shoots as plants grew. Maximum accumulation was observed respectively in 12 month old roots and leaves (Figure 4.10A).

Generally, similar expression patterns of miR172 have also been observed in vegetative tissues of Arabidopsis, maize and rice (Aukerman and Sakai, 2003; Lauter et al., 2005; Zhu et al., 2009). In reproductive tissues, the abundance of miR172 increased gradually during floral organ development showing a big jump in expression in the most mature flowers, while the lowest expression was observed in the most undeveloped flowers ,the "-inflorescences" (Figure 4.10B). Higher expression of miR172 in later stage vegetative tissues and developed floral tissues is consistent with the proposed role of miR172 in regulating of floral organ identity and flowering development in plants (Aukerman and Sakai, 2003; Chuck et al., 2009; Wu et al., 2009).

The abundance of intact transcripts of miR172 target genes was analyzed by quantitative reverse transcription PCR using primer pairs spanning the miR172

cleavage sites. Expression of potential AP2-like target RNA was highest in ten month old shoots and it was also high in emerged flowers and twelve month old roots (Figure 4.10C, D). This gene had a very low expression level in mature flower (Figure 4.10D).

Expression of the mature miR172a sequence showed an inverse expression pattern to that of the predicted oil palm AP2-like target sequence in oil palm eleven month shoot, leaf and floral tissues at various ages, most notably for the mature flower (Figure 4.10A, B, C and D). The Taqman probe used to detect mature miR172a was designed for ath-miR172a, which has an identical mature miRNA sequence to the egu-miR172a, b and c (isoform 1), so will detect only this isoform i.e. mature miR172a, but not mature miR172b (isoform 2), arising from the paralogs egu-miR172d, e and f.

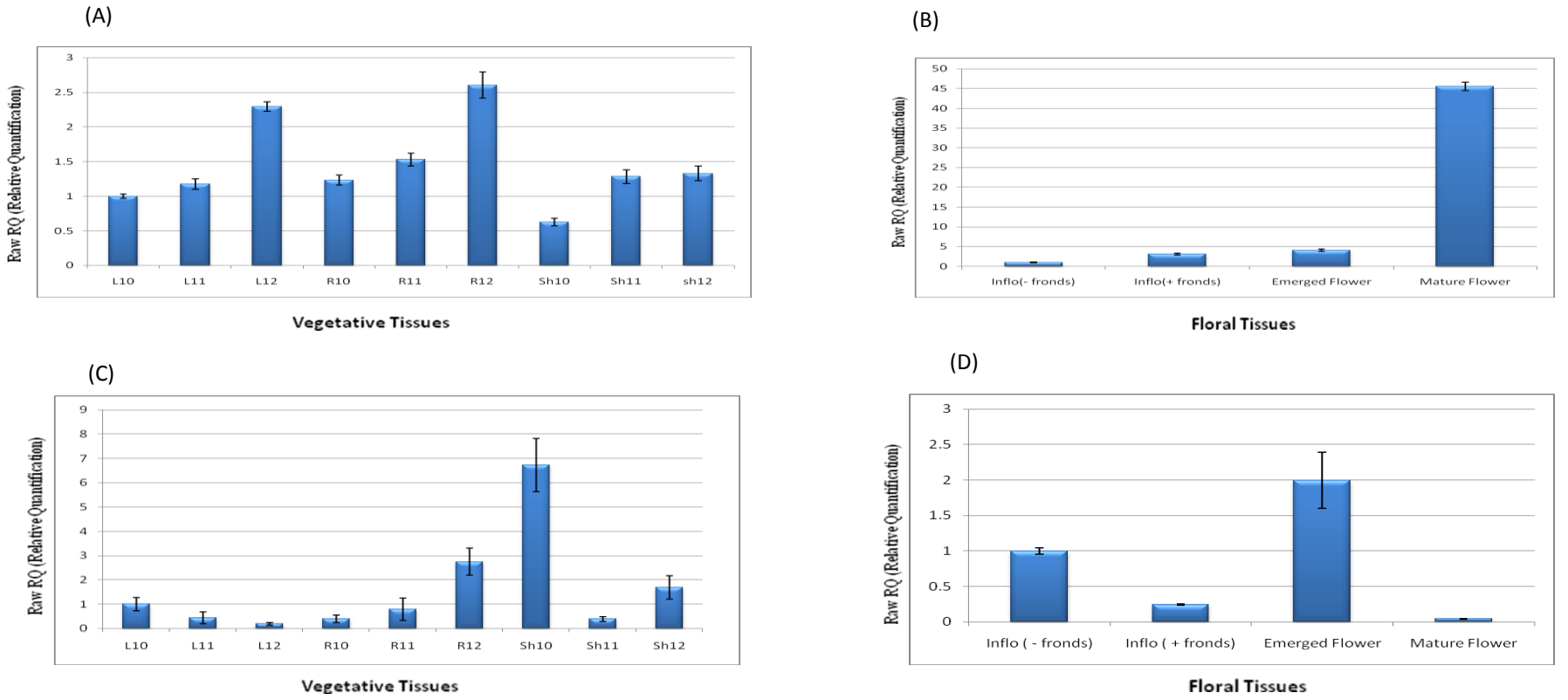
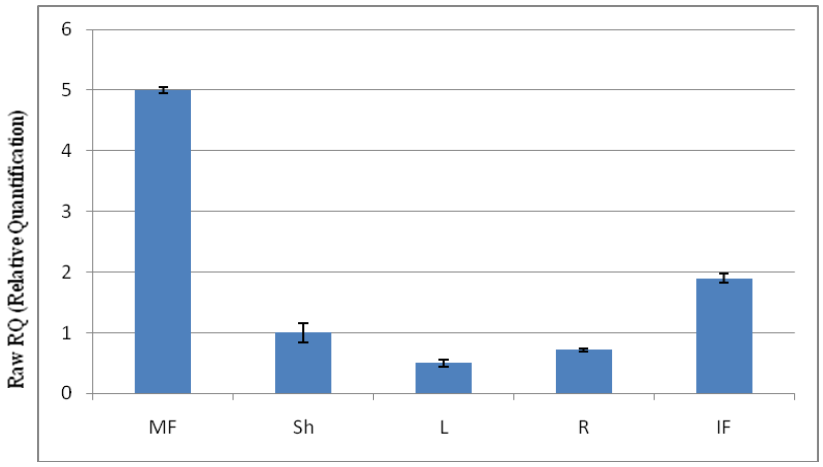


Figure 4. 10 Quantitative Reverse Transcription PCR of Mature miR172 and AP2-like Target from Oil Palm. (A) Relative expression of mature miR172 in leaf (L), root (R) and shoot (Sh) with 10-month-old leaf as the calibrator (value set at 1.0). (B) Relative expression of mature miR172 in flower with -Inflorescence as the calibrator (value set at 1.0). (C) Relative expression of potential miR172 target (EST EL692343.1) in leaf (L), root (R) and shoot (Sh) with 10-month-old leaf as the calibrator (value set at 1.0). (D) Relative expression of potential miR172 target (EST EL692343.1) in flower with "–Inflorescence" as the calibrator (value set at 1.0).

A comparison of the expression patterns of the different egu-miR172 precursors to that of the mature miR172a (isoform1), showed a similar profile for the mature miR172 and the three precursors sharing sequence identity in their mature region (i.e. egu-miR172a, b and c) (isoform1), with significantly higher expression in mature flower compared to immature flower (Figure 4.11). Considering the expression pattern of potential AP2-like gene, this suggests that for the mature miR172a (isoform 1), there may be significant cleavage of the AP2-like target. Two of the paralogs (egu-miR172e and egu-miR172f) representing the mature miR172b (isoform 2) sequence, showed similar expression levels in immature and mature flowers (Figure 4.5; E and F), suggesting that this isoform may be involved in regulation via a translational silencing mechanism in addition to cleavage, as reported for Arabidopsis (Aukerman and Sakai, 2003; Chen, 2004).

(A)



(B)

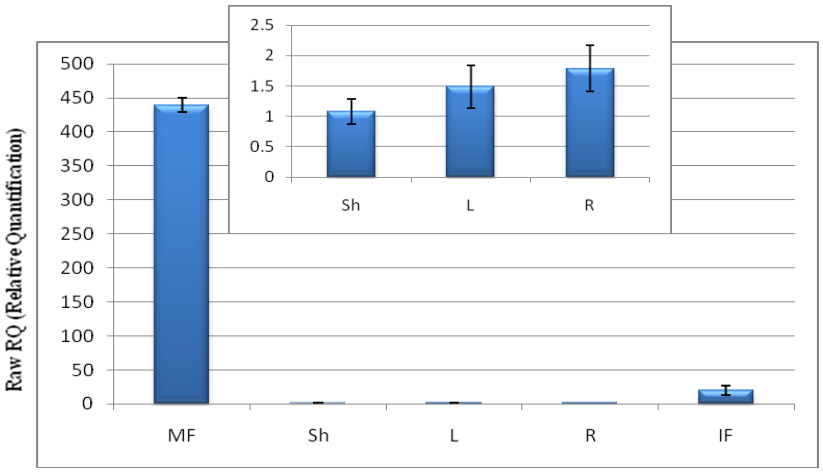
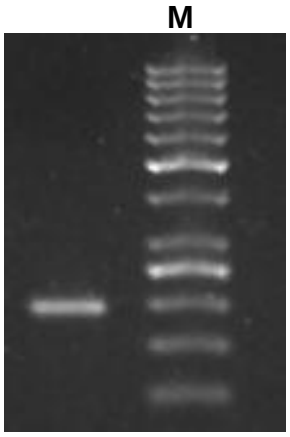


Figure 4.11 Expression Comparison of egu-miR172a Paralogs and Isoforms (A) three precursors sharing sequence identity in their mature region (egu-miR172a, b and c) (B) mature egu-pre-miR172a. MF: mature flower; Sh: shoot; L: leaf; R: root; IF: immature flower, with shoot (Sh) as the calibrator (value set at 1.0).

In this study, to confirm the presence of cleavage products, the potential AP2-like target mRNA from mature flower, was subjected to target cleavage analysis, using 5' RACE (rapid amplification of 5' complementary DNA ends). As was previously mentioned by Kasschau and colleagues (2003), miRNA guided cleavage phenomena have two diagnostic aspects. First, cleavage occurs close to the middle of the miRNA complementarity region, and next, the 5' end of the 3' cleavage product is capable to be ligated to an RNA adaptor. Accordingly, miRNA-guided cleavage phenomena can be identified by RNA ligase-mediated in 5' RACE followed by sequence analysis of the cloned PCR products (Kasschau et al., 2003). A gene-specific 5' RACE primer was designed and the presence of the expected product was confirmed by obtaining the amplicons of the desired size of approximately 200bp on a 1% agarose gel (Figure 4.12A). This product was cloned, and the sequence of twelve inserts was determined for the reaction. 100% of the 5' ends of inserts terminated at a position corresponding to the middle of the region of complementarity with the miR172 (between nucleotides 10 and 11 from the 5' end of the miRNA sequence) (Figure 4.12B).

(A)



(B)

Figure 4.12 5' Race (A) Agarose gel showing the 5' RACE product (~200bp) for miR172 target. Lane M, 100 bp DNA ladder. (B) Represent the target mRNA cleavage sites. The egu-miR172 sequences and partial sequence of the potential AP2-like target are shown in bold. The underlined lowercase letters indicate the mismatches in the complementary site. The numbers of 5' RACE clones sequenced that correspond to the cleavage product is indicated by vertical arrowhead.

Although it was thought that miR172 translational repression is the predominant mode of regulation in plants (Zhu et al., 2009), in several studies, miR172-guided cleavage products have been detected for AP2-like genes in Arabidopsis (Aukerman and Sakai, 2003; Kasschau et al., 2003; Schwad et al., 2005), maize (Lauter et al., 2005; Chuck et al., 2007) and rice (Zhu et al., 2009), which are consistent with the 5' RACE results, presented in this study. As examples, in rice, to discover if all of the five AP2-like targets are cleaved by miR172, 5' RACE analysis was performed by Zhu and colleagues (2009), using RNAs from shoots, grains and booting panicles (BP). Cleavage of Os04g55560 target was identified in a mixed sample of shoot and grain in addition to BP; cleavage product was barely detected in the mixed sample for Os06g43220; and cleavage of SNB was only detected in BP. No cleavage was detected for Os03g60430 and Os05g03040 in any of the samples analyzed. All the cleavage sites were located at the middle of the complementary region (between nucleotides 10-11 from the 5' end of the miR172 sequence) (Zhu et al., 2009).

In another study conducted by Aukerman and Sakai (2003), 5' RACE was used to detect cleavage products for the AP2-like target genes in both wild-type and miR172-overexpressing plants (*eat-D*), which indicated that the cleavage site for miR172 in Arabidopsis is, mostly, located between nucleotides 10-11 from 5' end of this miRNA. The numbers of 5' RACE clones sequenced that correspond to this cleavage site was: 19/22 sequences for AP2, 4/4 sequences for TOE1 and 17/19 sequences for TOE2 (Aukerman and Sakai, 2003). In addition, Kasschau and colleagues (2003), by performing 5' RACE proved that, at least 75% of the 5' ends of inserts terminated at the same position at the middle of the complementary site. This is a common aspect of a RISC-like processing event and strongly suggests that at least part of each predicted target mRNA is cleaved by a process dependent on the indicated miRNAs (Kasschau et al., 2003).

Aukerman and Sakai (2003), reported that no cleavage products were detected on RNA gel blots which shows that the detected cleavage products by 5' RACE, are a very small part of the total AP2-like transcript population. This is while the overexpression of miR172 did not change the steady state levels of the *AP2*-like target genes. Beside this, AP2 protein was measured by immunoblot analysis (using an antibody specific for AP2) and the results showed that, compared to the wild type, the AP2 protein declined considerably in the miR172a overexpression line, while the AP2 transcript accumulated to normal levels in the same transgenic line. These data suggest that miR172a negatively regulates AP2-like genes by translational repression and not by target mRNA cleavage and that the small amount of RNA cleavage observed, possibly is the outcome of overlap between these two regulatory pathways (Aukerman and Sakai, 2003). To study miRNA-guided cleavage, Schwab and colleagues (2005) developed a transient assay with *Nicotiana benthamiana* as host and *Agrobacterium tumefaciens* as delivery vehicle. Overexpressed miR172a caused a reduction in the levels of full-length TOE2 transcript and led to accumulation of a shorter mRNA indicating the TOE2 cleavage products. miR172a showed no effect on an miRNA-resistant version of TOE2, and, additionally, even in the absence of exogenous miR172a its accumulation increased to higher levels. These experiments together, revealed that miR172a guides TOE2 cleavage, which leads to much reduction in steady-state levels of TOE2 in overexpressed-miR172a (this was not observed by the Aukerman research group earlier in 2003) (Schwab et al., 2005). In rice, as mentioned earlier, also miR172-mediated cleavage for three AP2-like genes (SNB, Os04g55560 and Os06g43220) was observed. This is while the abundance of SNB mRNA was not reduced in overexpressing miR172b plants (Zhu et al., 2009). Although significant increases in the cleavage products of AP2 and TOE1 also was detected in Arabidopsis, still no appropriate decline in steady-state levels of TOE1 or AP2 was observed. A comparison between the

AP2 transcript levels in wild-type and in transgenic plants that overexpressed an miRNA-resistant form of AP2 (mAP2) was done. Results showed a significant reduction in the levels of endogenous, wild-type AP2 transcript in mAP2 overexpression. This result confirmed that AP2-like targets are under direct or indirect feedback regulation by their own products. It was concluded that miR172 both causes mRNA target cleavage and translational suppression and the resulted reduction in protein accumulation causes increase in the target's transcription (Schwab et al., 2005).

Studies on transgenic oil palm are impractical because of the long life cycle which causes serious limitation in the field of miRNA studies, for example it is not feasible to study mutated miRNA or miRNA target sequences within oil palm for proof of function. However, taken together with the evidence from previous studies in other plants (Aukerman and Sakai, 2003; Kasschau et al., 2003; Chen, 2004; Schwab et al., 2005; Zhu et al., 2009), the unchanged or increased abundance of the potential oil palm AP2-like target in some tissues in older oil palm plants may be an indicator that miR172 acts as a translational suppressor in these tissues in oil palm; whereas the cleavage product detected in mature flowers by using 5' RACE, together with the decreased accumulation of the potential AP2-like target in some other tissues in older stages, such as leaf (Figure 4.10), indicate that this miRNA also regulates its potential AP2-like target through cleavage in the mentioned oil palm tissues.

In this study six paralogs and two isoforms of miR172 together with one AP2-like target were recognized, considering the number of miR172s in miRbase and also AP2-like genes from other species, it is likely that there may be other members of the egu-miR172 and AP2-like target gene family which have not yet been identified, which might affect the function of these known egu-miR172s. For instance, so far, the miR172 gene family from *Populus trichocarpa* (ptc) possesses the largest number of individuals in miRbase which is eight pre-miR172 paralogs and four mature-miR172

isoforms (from its genome analysis); while *Glycine max* (gma) and *Sorghum bicolor* (sbi) are each reported to have six paralogs and four mature isoforms (miRbase, 2010), which suggests the possibility that there may be other members for MIR172 in oil palm as well.

It is probable that one miRNA isoform e.g. egu-miR172a (isoform 1) regulates one or two targets, while egu-miR172b (isoform 2) regulates other AP2-like targets in oil palm. It is also likely that other isoforms (which have not been recognized yet) have overlapping roles with the function of these two identified isoforms (isoforms 1 and 2); for instance, miR172e targets IDS in maize (Chuck et al., 2007b) while other miR172 members target glossy15 in this plant (Lauter et al., 2005).

It is also likely that one miRNA isoform may be under direct or indirect regulation by other miRNAs or their targets, as was described in Arabidopsis where miR156 regulates expression of miR172 via one of its targets named SPL9 (Wu et al., 2009). Moreover, one isoform may be engaged in floral development while the other may be involved in vegetative phase change and maturity. As an example, in maize, miR172e (ts4) mutants showed a branching deficiency in which the distal portion of the inflorescence, compared to the wild type, was highly branched. miR172e did not show any effect on vegetative phase change; this is while the other miR172 family members were engaged in the vegetative phase transition in this plant (Lauter et al., 2005; Chuck et al., 2007). This result in which miR172e mutant (loss-of-function) in maize caused an increase in inflorescence meristem branching and expanded the tassel's carpels, also may suggest that in some cases, just one of the miR172 members plays a functional role in plants, signifying that miR172e has a specific role in this plant (Chuck et al., 2007; Zhu et al., 2009).

Another possibility is that two miRNAs regulate one target, using two different methods, or that one miRNA regulates two different AP2-like members simultaneously or at different developmental stages of the plant. This could be a result of spatiotemporal expression differences between individual members of the miR172 family, or their targets. For example, the phenotypes resulting from overexpression of miR172b represses SNB and at least one of the other miR172 targets in rice. These results were obtained while overexpression of miR172b showed same phenotypes of snb and also other additional developmental defects which were not been observed in snb, which implied that in overexpressing miR172b plants, SNB and at least one of the other AP2-like target genes are downregulated, demonstrating that other AP2-like family members also are involved in rice floret development (Zhu et al., 2009).

As reported earlier for other plant species such as Arabidopsis and Rice (Aukerman and Sakai, 2003; Zhu et al., 2009), it is also likely that in oil palm, the AP2-like transcript be under negative feedback regulation through its protein products. This phenomenon causes difficulties to make very strong conclusions on the major regulation mechanism of egu-miR172 and its exact role in oil palm; however, the results presented in this study, suggest that at least two identified egu-miR172 isoforms regulate one potential AP2-like target gene, using both cleavage and translational suppression, in different tissues in oil palm. To confirm which of these two regulatory mechanisms plays the major role in oil palm, performing proteomic studies such as measuring the AP2-like protein levels at different developmental stages of oil palm might be helpful.

Access to the oil palm genome as well as finding other probable egu-miR172s and their targets, which can later be accompanied by investigation of their expression and interactions, can be future steps to improve these findings and possibilities.

5. CONCLUSION AND FURTHUR STUDIES

In this study six paralogs and two isoforms of egu-miR172, as well as, a potential AP2-like target gene were identified in five different tissues in oil palm. The expression pattern of the potential egu-miR172s and target gene were studied in vegetative and floral tissues.

The data demonstrated that the egu-miR172a and the potential AP2-like gene had overlapping expression patterns and their expression did not show an obvious negative correlation in all of the oil palm tissues. Also, there was not a uniform decrease in the expression of the potential AP2-like target mRNA as the egu-miR172a expression increased in various tissues in different developmental stages, with the exception of leaf. The results also demonstrated that the expression of different pre-egu-miR172 paralogs is tissue specific and for at least one mature egu-miR172 isoform in oil palm, there may be significant cleavage of the AP2-like target. Although this study supports that miR172 regulates at least one of its AP2-like target genes in oil palm by using both degradation and translational suppression, it is not possible to confirm the major regulation mechanism of egu-miR172 and its exact role in this plant.

miR172 has been reported to be involved in floral organ identity and flowering development in several plant species; higher expression of egu-miR172 in later stage vegetative and floral tissues in oil palm is consistent with this proposed role for miR172. Further studies are required to prove this idea. Artificial miR172 may be used to suppress gene expression for knockdown of AP2-like target genes and to study the gene function as was already proposed by Zhang and colleague for Arabidopsis (Zhang et al., 2006b). Gain-of-function and loss-of-function analyses in transgenic plants is used to study both RNA and protein expression e.g. using RNA gel blots and immunoblots of protein can be good indicators to discover the major regulation

mechanism of miR172. However, these methods are quite time consuming in oil palm, due to the long life cycle of this crop and the time taken for transition from its vegetative to reproductive phase. Availability of the oil palm genome sequence is anticipated in the near future and this should at least enable the identification of other miRNA family members (if they exist) and potential targets of miR172 from oil palm.

This study has layered the fundamentals for identifying egu-miR172, potential miR172 targets and their functional analysis in *Elaeis guineensis.* Elucidation of the presence and roles of egu-miR172 paralogs may contribute to sequence based selection in oil palm improvement programs.