A COMPARATIVE ANALYSIS OF microRNAs IN *FLAVERIA* SPECIES OF C_3, C_3-C_4 INTERMEDIATE AND C_4 PHOTOSYNTHESIS

By

ROBERT POKOO

Bachelor of Science in Biochemistry
University of Cape Coast
Cape Coast, Ghana
2007

Master of Science in Biochemistry
Chungnam National University
Daejeon, South-Korea
2012

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2017

A COMPARATIVE ANALYSIS OF microRNAs IN *FLAVERIA* SPECIES OF C_3, C_3-C_4 INTERMEDIATE AND C_4 PHOTOSYNTHESIS

Dissertation Approved:

Dr. Ramanjulu Sunkar

Dissertation Adviser

Dr. Andrew Mort

Dr. Charles Chen

Dr. Gopal Kakani

ACKNOWLEDGEMENTS

Firstly, I wish to thank my advisor Dr. Ramanjulu Sunkar for his invaluable instructions and mentoring throughout this project. Secondly, my committee members, Dr. Andrew Mort, Dr. Gopal Kakani and Dr. Charles Chen for their positive criticisms and motivations throughout my period of research. Many thanks are due to Dr. Yun Zheng (Kunming University of Science and Technology, China) and Dr. Brian Couger (Oklahoma State University) for the help with analyses of the sequenced datasets, Dr. Udo Gowik (Heinrich-Heine-Universitaet Dusseldorf, Germany) and Dr. Rowan Sage (University of Toronto, Canada) for sharing the seeds of *Flaveria* spp. used in this study and Susanne von Caemmerer (Australian National University, Canberra, Australia) for the kind gift of antibody raised against Rubisco. Thanks are also due to Ms. Smitha Jose, Dr. Yong-Fang Li and Ms. Ferhunde Ayisin for their initial assistance with various molecular biology techniques.

Immense thanks also go out to my BMBGSA colleagues for their comradeship and support throughout my period of study. I had an enjoyable experience doing research and studying with such talented scientists such as Junho Cho, Freddy Saudale, Sikta Patnaik and Prabhat Kumar. Immense thanks go to my parents, Monica Pokoo and Robert Pokoo, my wife Afua Pokoo and son Aaron Afriyie Pokoo who have been my source of inspiration and motivators throughout this study, especially when things looked topsy-turvy.

Finally, I would like to thank the office staff of the Biochemistry & Molecular Biology Department and the entire faculty for the good relationship and assistance they always provided me especially Dr. Kevin Wilson and Dr. Jose´ Soulages for allowing me to use their laboratory and equipments when

needed. I have been able to get this far all because of such wonderful and dedicated people.

Name: ROBERT POKOO

Date of Degree: May, 2017

Title of Study: A COMPARATIVE ANALYSIS OF microRNAs IN *FLAVERIA* SPECIES OF C_3, C_3-C_4 INTERMEDIATES AND C_4 PHOTOSYNTHESIS

Major Field: BIOCHEMISTRY AND MOLECULAR BIOLOGY

MicroRNAs (miRNAs) are a group of small non-coding RNAs that negatively regulate expression of their target genes at the posttranscriptional level. The miRNA-mediated gene regulation has emerged as one of the critical modes of gene regulation important for almost all biological processes of a plant life cycle. However, the role of miRNAs in photosynthesis is unknown. Photosynthesis is a chief metabolic process through which plants synthesize carbohydrates in the presence of sunlight using atmospheric carbon dioxide and water. *Flaveria* genus belongs to the family *Asteraceae* and its species differ greatly in their mode of photosynthesis as different *Flaveria* spp. operate either C_3, C_3-C_4 intermediate, or C_4 photosynthesis. This makes *Flaveria* a valuable model system to investigate, at the molecular level, how C_4 type has evolved from the ancestral C_3 type. To address whether or not miRNAs have a role in this process, we have analyzed miRNAs in the leaves as well as from the isolated mesophyll and bundle sheath cells of leaves from *Flaveria robusta* (C_3), *Flaveria ramosissima* (C_3-C_4 intermediate) and *Flaveria bidentis* (C_4). The analyses revealed significant differences for the abundances of various miRNA families in *Flaveria robusta* (C_3), *Flaveria ramosissima* (C_3-C_4 intermediate) and *Flaveria bidentis* (C_4). To gain an insight into the miRNA targets in these plant species, degradome libraries were constructed and sequenced. This approach confirmed targets such as REVOLUTA (member of homeodomain-leucine zipper family) and Target of early activation tagged (EAT) for miR166 and miR172, respectively. Additionally, several other potential targets that could be regulated by miRNAs in the *Flaveria* spp. have been identified.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Photosynthesis is a chief metabolic process through which plants synthesize carbohydrates (stored as starch) in the presence of sunlight using atmospheric carbon dioxide and water. It is unequivocally regarded as one of the most important biochemical pathways because all life existence depends on it for acquisition of energy and survival (Ehleringer et al., 1993). Most plant scientists in recent times are also targeting improving on the photosynthetic pathway in crops that are widely propagated in order to ensure increasing the amount of food production to match the ever increasing human and animal population (Furbank, 2017).

Many evidences have so far suggested that photosynthesis has existed since the origination of life and this complex biochemical process has evolved over millions of years. The ability to photosynthesize is known to be widely distributed in cyanobacteria and green plants. Plants incorporate carbon dioxide into the Calvin cycle (dark stage of photosynthesis) and reduce it to produce starch at the end of this process (Ehleringer et al., 1993). Based on the first stable carbon compound formed after carbon dioxide is incorporated, plants have been classified into C_3 or C_4 photosynthesis. The C_3 plants are characterized by a photosynthetically active mesophyll cells where CO_2 fixation into the Calvin cycle occurs. Contrastingly, C_4 plants possess two photosynthetically active cell types namely, mesophyll and bundle sheath cells. CO_2 fixation occurs in the mesophyll cells whereas the Calvin cycle takes place in the bundle sheath cells. This distinctive feature in C_4 plants ensures a sequestration of a high concentration of CO_2 in the Calvin cycle containing bundle sheath cells required for photosynthesis.

As a result, C4 plants avoid photorespiration and so are photosynthetically more efficient than C3 plants (Ehleringer et al., 1997).

***Flaveria* as a model system to study the evolution of C4 from its ancestral C3 photosynthesis**

C4 photosynthesis was initially discovered in sugarcane (Hatch and Slack, 1966). However, due to the lack of a phylogenetically close C3 taxa for sugarcane, the study for potential regulators of the C4 pathway is limited in this plant species. *Flaveria* genus belongs to the family *Asteraceae* (Powell, 1978). The genus includes annuals, perennials and shrubs that are mostly found in the Fehuacan valley of Mexico, Southern United States, Asia (India), West Indies (Greater Antilles), Africa, Australia and South America. An interesting and very relevant feature is that *Flaveria* spp. differ greatly in their mode of photosynthesis, i.e., different *Flaveria* spp. operate either C3, or C3-C4 intermediate or C4 photosynthesis (Table. 1) (Mckown et al., 2005). It is also easier to biochemically purify/isolate their mesophyll and bundle sheath cells from these species (Kanai and Edwards, 1973). These attributes make *Flaveria* spp a valuable model system to investigate at the molecular level, how C4 photosynthesis evolved from its ancestral C3 type.

Table 1: Diversity of photosynthesis in *Flaveria* spp. (Modified from Gowik et al., 2011) (* denotes the species used in the current study).

***Flaveria* spp.**	**Photosynthetic type**
F. robusta	C3 photosynthesis*
F. ramosissima	C3-C4 intermediate*
F. sonorensis	C3-C4 intermediate
F. angustifolia	C3-C4 intermediate
F. anomala	C3-C4 intermediate
F. chloraefolia	C3-C4 intermediate
F. yucatan	C3-C4 intermediate
F. pubescens	C3-C4 intermediate
F. oppositifolia	C3-C4 intermediate
F. linearis	C3-C4 intermediate
F. floridana	C3-C4 intermediate
F.palmeri	C4-like photosynthesis

F. vaginata	C4-like photosynthesis
F. brownii	C4-like photosynthesis
F. kochiana	C4 photosynthesis
F. campestris	C4 photosynthesis
F. australasica	C4 photosynthesis
F. trinervia	C4 photosynthesis
F. haumanii	C4 photosynthesis
F. bidentis	C4 photosynthesis*

MicroRNAs

MicroRNAs (miRNAs) represent a class of endogenous small non-coding RNAs that regulate the expression of mRNAs that share sequence complementarity. MicroRNAs are at the heart of complex gene regulatory networks that control almost all growth and developmental processes of the plant life cycle including biotic and abiotic stress responses as well as nutrient deprived conditions (Jones-Rhoades et al., 2006; Sunkar et al., 2012). Several recent studies have focused on cataloguing miRNA populations from different plant species. Thus far, approximately 8,500 miRNAs from 73 different plant species have been deposited at the miRBase (www.miRBase.org) but none from the *Flaveria* spp.

Research Objectives

Although miRNAs have been associated with almost all biological processes of a plant life cycle, their roles in photosynthesis, the most important metabolic process, remain unclear. Because miRNAs represent major gene regulatory molecules controlling a variety of biological processes, we hypothesize that miRNAs could play important roles in photosynthesis, specifically in C3 and C4 modes of photosynthesis. This can be tested by identifying miRNAs expressed in leaves of C3 and C4 plants that are closely related. C4 photosynthesis operates in two distinct types of cells, namely, mesophyll and bundle sheath cells. This thesis catalogues and characterizes miRNAs expressed in not only entire leaves but also in mesophyll and bundle

sheath cells of C3, C3-C4 intermediate and C4 *Flaveria* spp. (Table 1). The outcome of this research contributes clues about whether miRNAs could function in photosynthesis.

CHAPTER II

LITERATURE REVIEW

2.1: Photosynthesis

The photosynthesis process is divided into two main types namely C_3 and C_4 based on the first stable carbon product formed after fixation of CO_2 from the atmosphere. Although atmospheric CO_2 is the general source of carbon to both pathways, a 3-Carbon compound (3-phosphoglycerate) is produced in the C_3 pathway, whereas a 4-Carbon compound (oxaloacetate) is synthesized in the C_4 pathway.

2.1.1: C_3 Photosynthesis

In C_3 photosynthesis (Fig. 2.1), atmospheric CO_2 diffuses through the stomata into mesophyll cells, the carboxylation site of photosynthesis (Ehleringer, 1997). In this step, the substrate Ribulose-1,5-Bisphosphate (RuBP) is carboxylated with CO_2 to generate 3-phosphoglycerate (3-carbon compound) and this reaction is catalyzed by Ribulose-1,5-Bisphosphate carboxylase/oxygenase (Rubisco).

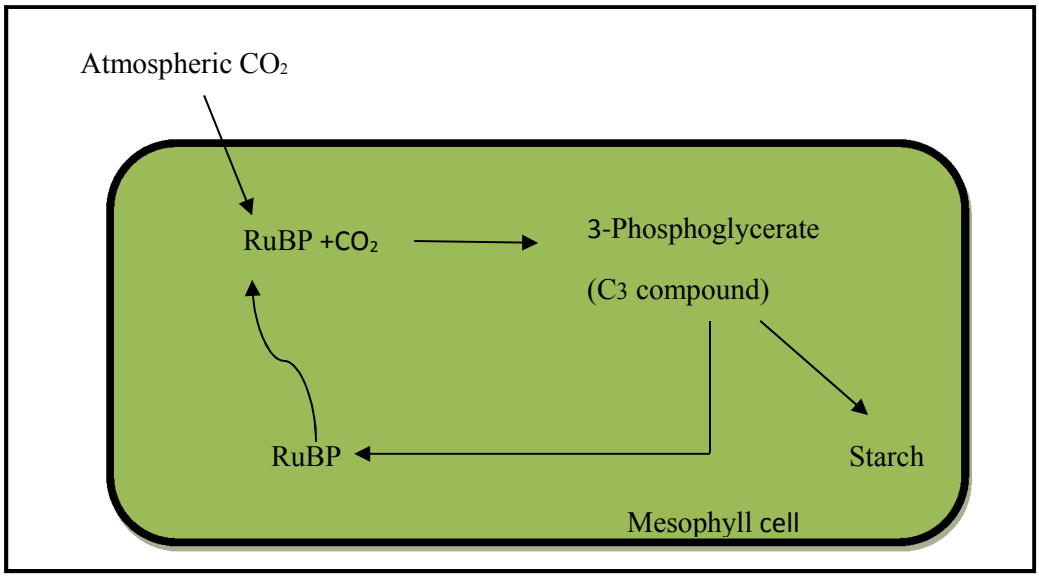


Figure 2.1**:** C_3 plants possess only photosynthetic mesophyll cells.

Plants that use C_3 photosynthesis are predominantly characterized by the presence of Rubisco in their mesophyll cells (Edwards and Walker, 1983). About 95% of land plants including the most important crop plants such as rice, wheat, potato and tomato use the C_3 pathway of carbon fixation (Bond et al., 2005). Due to the high demand of CO_2 for photosynthesis, C_3 plants open their stoma frequently, resulting in increased loss of water through transpiration. Moreover, C_3 plants often suffer from photorespiration due to the lack of an efficient CO_2 concentrating mechanism (Sharkey, 1988). Photorespiration is a process in which the Rubisco enzyme preferentially catalyzes the oxidation of Ribulose-1,5-Bisphosphate (RuBP) to 3-phosphoglycerate (3-PGA) and phosphoglycolate. Photorespiration is known to metabolize ATP during its cycle rather than producing it (Leegod et al., 2007).

2.1.2: C_4 Photosynthesis

Anatomically, the leaves of C_4 plants possess Kranz anatomy, i.e., wreath-like arrangement of mesophyll cells around bundle sheath cells. In C_4 photosynthesis, CO_2 first enters the mesophyll cells and is converted to bicarbonate by carbonic anhydrase. The enzyme PEPCase (Phosphoenolpyruvate Carboxylase) then catalyzes the combination of bicarbonate and

phosphoenolpyruvate to form oxaloacetate, a 4-Carbon compound. Oxaloacetate is then converted into malate for transportation from the mesophyll cells to bundle sheath cells where carbon fixation occurs (Fig.2.2). In the bundle sheath cells, malate is decarboxylated to regenerate pyruvate and CO_2. The CO_2 is channeled into the Calvin cycle for starch formation whereas pyruvate is transported back into the mesophyll cell to repeat the entire cycle of carbon fixation (Sage et al., 2006). In the leaf of a C4 plant, enlarged bundle sheath cells and reduced interveinal distance contribute immensely to an increase in flow of photosynthetic metabolites between cells. Besides, the Kranz anatomy facilitates a high CO_2 concentration at the active site of Rubisco (Dengler et al., 1999; Sage and Monson 1999; Still et al., 2003; Gowik et al., 2011).

About 3% of land plants such as maize, sorghum and sugarcane use the C4 photosynthetic pathway (Sage et al., 2009 and Zhu et al., 2008). C4 plants have low rates of water loss (transpiration) compared to C3 plants. The existence of an efficient CO_2 concentrating mechanism in C4 plants (which ensures a high concentration of CO_2 around the Rubisco enzyme) is known to increase their photosynthetic efficiency.

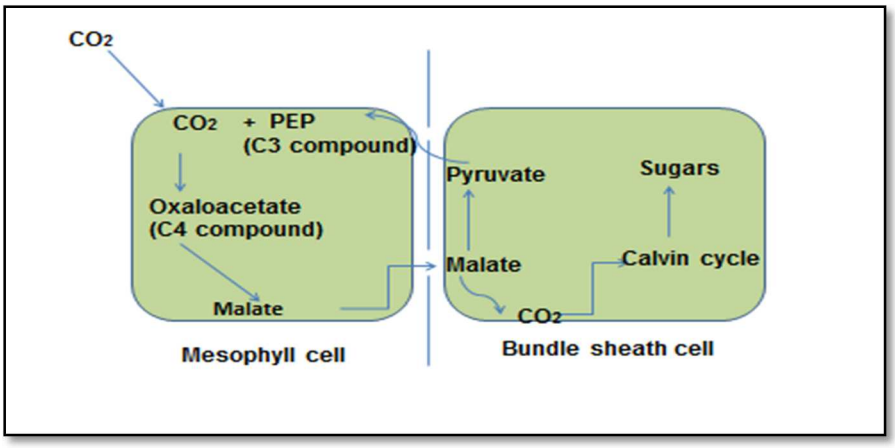


Figure 2.2: Cell-types and metabolites in C4 photosynthesis (PEP- Phosphoenolpyruvate)

2.1.3: C3-C4 intermediate Photosynthesis

C3-C4 intermediate photosynthesis is known to operate in *Panicum* spp., *Mollugo* spp., *Moricandia* spp., *Alternanthera* spp. and *Flaveria* spp. (Rawsthorne, 1992; Monson et al., 1989). It is referred to as such because its mechanisms involve 70% features of C3 and 30% of C4 photosynthesis. It also may represent a transition stage in the evolution of C4 photosynthesis from C3 type (Monson et al., 1984; Rawsthorne et al., 1992). C3-C4 intermediates are characterized by low rates of photorespiration when compared to C3 plants because they possess a photorespiratory CO_2 pump, also referred to as photorespiratory glycine shuttle that scavenges photorespiratory CO_2 into the bundle sheath cells (Ku et al., 1991; Mallman et al., 2014).

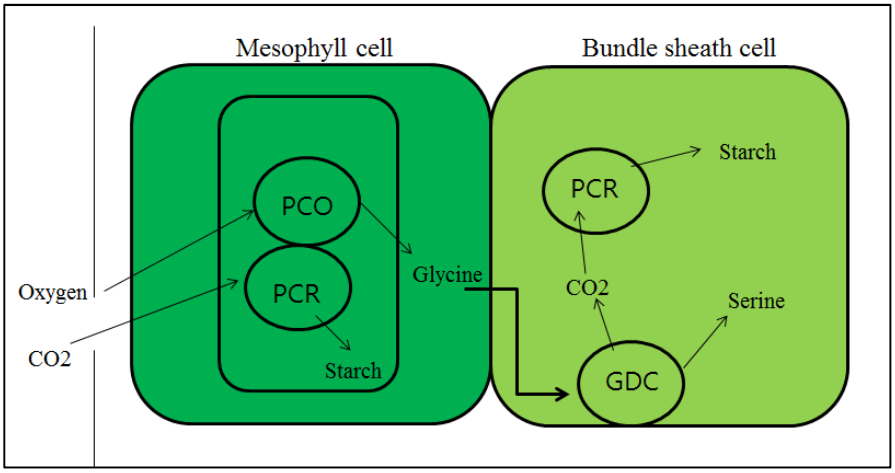


Figure 2.3: Figure 2.10. Schematic representation of the photorespiratory pump in C3-C4 Intermediate photosynthesis. Mesophyll and bundle sheath cells contain chloroplasts with functional Calvin cycle. Abbreviations: GDC-glycine decarboxylase; PCO-photosynthetic oxidative cycle; PCR- photosynthetic reductive cycle.

During photorespiration, the Rubisco enzyme preferentially oxidizes RuBP instead of carboxylating it. As a result, phosphoglycolate is produced and C3-C4 intermediate plants use their photorespiratory CO_2 pump to regenerate CO_2 from phosphoglycolate through a series of reactions, which occur in the bundle sheath cells. The photorespiratory CO_2 is generated from decarboxylation of glycine (product of photorespiration) in a reaction catalyzed by glycine

decarboxylase, which is restricted to the bundle sheath cells (Bauwe, 2011). Thus, although photorespiratory-derived glycine produced within the mesophyll cells but it is transferred to the bundle sheath cells for further metabolism, which ultimately results in the release of CO_2 and NH_3. By this pathway, C3-C4 intermediate plants ensure a high concentration of CO_2 around Rubisco enzyme within the bundle sheath cells (Gowik et al., 2011; Sage et al., 2014). In C3 plants, even though glycine decarboxylase enzyme is present in the mesophyll cells, the photorespiratory CO_2 pump does not exist. C4 plants do not have glycine decarboxylase enzyme; instead they possess the Kranz anatomy, which ensures high CO_2 concentration around Rubisco enzyme (Sage et al., 2014).

2.2: Evolution of C4 photosynthesis

The C4 cycle portrays a complex series of biochemical and anatomical modifications that contributes significantly to overcoming the deficiencies associated with the Rubisco enzyme (its oxygenase activity that competes with the carboxylase function thereby promoting photorespiration) and regulating the entire photosynthetic process (Mallman et al., 2014).

The C4 photosynthesis has independently evolved over 62 times in 19 different lineages from the ancestral C3 type (Hausler et al., 2002). It has also been postulated that these frequent evolutionary changes are because of the ease in the recruitment of C4 cycle genes over the period of evolution (Sage et al., 2012). Intriguingly, some genes essential for the C4 pathway are also expressed in the leaves of C3 plants but at relatively low levels. These include genes that encode pyruvate phosphate dikinase, phosphoenolpyruvate carboxylase, NADP-dependent malate dehydrogenase and phosphoenolpyruvate carboxykinase (Brautigam et al., 2011; Gowik et al., 2011; Bergh et al., 2014). The mechanisms by which these genes were recruited and are highly expressed in the C4 pathway are yet to be elucidated. Based on previous studies, it was postulated that the driving forces behind the evolutionary transformation of C3 to C4 photosynthetic pathway include photorespiration, drought, high temperatures and low concentration of CO_2, which may

contribute to differences in leaf anatomy and cell specific expression of photosynthetic enzymes (Sage, 2001; 2004; McKnown et al., 2005; Gowik et al., 2011). In addition, gene duplication, differential gene expression patterns within cell types (mesophyll and bundle sheath cells) and photorespiratory CO_2 pump have been suggested to be contributing factors that led to C4 photosynthesis evolution (Gowik et al., 2011).

2.2.1: Gene Duplication

Gene duplication has been proposed as one of the factors for C4 photosynthesis evolution (Bergh et al., 2014). This postulation suggests that during gene duplication, the original ancestral gene is maintained while the duplicate form of the gene acquires some other beneficial or additive role that were absent in the ancestral gene. This additive role of the duplicated gene could result in significant changes in the regulation of plant metabolism as well the phenotype (Monson et al., 1999; Monson, 2003). As C4 photosynthetic evolution progressed gradually from C3 pathway, it was hypothesized that C4 cycle genes namely carbonic anhydrase, phosphoenolpyruvate carboxylase, malate dehydrogenase, malate decarboxylase enzyme and phosphoenolpyruvate carboxykinase underwent modifications and were recruited into the C4 pathway (Huynen and Van Nimwegen, 1998; Harrison and Gerstein, 2002).

2.2.2: Localization and alteration in gene expression within the mesophyll and bundle sheath cells

In contrast to the above suggested gene duplication, localization and alterations in the expressions of specific genes in plant cell compartments are thought to play important roles in the evolution of C4 photosynthesis (Brautigam et al., 2011; Kulahoglu et al., 2014). Examples of enzymes that are expressed in specific plant cells within the leaves include carbonic anhydrase that catalyzes the conversion of CO_2 to bicarbonate, Ribulose-1,5-Bisphosphate carboxylase/oxygenase that catalyzes the carboxylation of Ribulose-1,5-Bisphosphate to 3-phosphoglycerate and pyruvate phosphate dikinase that catalyzes the regeneration of

phosphoenolpyruvate from pyruvate (Sheen 1991; Ludwig and Burnell, 1995). Even though gene duplication may play potential roles in the C4 evolution process, it does not necessarily contribute to cell specific expression of photosynthetic enzymes and proteins. It is important to note that for C4 photosynthesis to operate efficiently, an increase in the expression of C4 photosynthesis enzymes are required in both mesophyll and bundle sheath cells. In C3 photosynthesis however, few enzymes are required since the mesophyll cell is the only site of photosynthesis (Hibberd et al., 2010; Aubry et al., 2011).

A recent comparative leaf transcriptome analysis of C3 and C4 *Cleome* spp. has suggested that C4 photosynthesis evolution occurred due to an upregulation of genes that encode for C4 photosynthesis enzymes in C4 plants compared with C3 plants. The leaf transcriptome studies of *Cleome spinosa* (C3) and *Cleome gynandra* (C4) revealed enhanced expression of C4 photosynthesis transcripts in *C. gynandra* than in *C. spinosa*. The upregulated transcripts include carbonic anhydrase, phosphoenolpyruvate carboxylase, malate dehydrogenase, NADP-Malic enzyme, pyruvate carboxykinase, pyruvate orthophosphate dikinase, bile acid sodium symporter, pyruvate orthophosphate dikinase and ribulose bisphosphate carboxylase/oxygenase small-subunit (Brautigam et al., 2011). Furthermore, transcripts which encode GOLDEN2-LIKE transcription factors (associated with chloroplast positioning in cells and plasmodesmata conductance) in *Arabidopsis* were also more highly expressed in C4 plants than in C3 plants (Oikawa et al., 2003; Levy et al., 2007). Leaf transcriptome analysis of C3, C3-C4 intermediate and C4 *Flaveria* spp. (Gowik et al., 2011; Mallman et al., 2014) has also been studied and the results complement the findings in *Cleome* spp. (Table 2.1).

Generally, C4 photosynthesis involves the transport of large amounts of metabolites across the chloroplast but C3 pathway requires less. However, an interesting observation made from the transcriptome data of C4 and C3-C4 intermediate *Flaveria* spp. was the increase in expressions of the alanine aminotransferase gene in *F. ramosissima* (C3-C4 intermediate)

compared to *F. trinervia* (C4). Cytoplasmic and mitochondrial aspartate aminotransferase as well as mitochondrial NAD-malate dehydrogenase genes however, was significantly upregulated in *F. ramosissima* (C3-C4) than in *F. robusta* (C3) and *F. bidentis* C4. The upregulation of transcripts of regulatory proteins such as pyruvate orthophosphate dikinase regulatory protein and transporters of serine and glycine in *F. ramosissima* (C3-C4 intermediates) than in C3 and C4 *Flaveria* spp. confirms the proposition that C3-C4 intermediate photosynthesis involves the shuttling of more photosynthesis metabolites between plant cells (Gowik et al., 2011).

Table 2.1: The normalized abundances of transcripts associated with C4 cycle in *Flaveria* spp. (Gowik et al., 2011)

C4 Cycle transcripts	*F. trinervia* (C4)	*F. bidentis* (C4)	*F. ramosissima* (C3-C4)	*F. robusta* (C3)	*F. pringlei* (C3)
PPDK	18,213	17,580	2,330	101	140
PEPC	25,975	12,502	3,894	167	254
NADP-MDH	3,679	3,516	674	156	245
PEP-CK	24	38	9	31	9
PEPC-K	147	140	48	14	9
PPDK-RP	274	476	54	61	36
BASS 2	2,873	2,870	1,043	110	230
DiT 1	303	517	167	172	82

Key: PPDK=Pyruvate orthophosphate dikinase; PEPC=Phosphoenolpyruvate Carboxylase; NADP-MDH=NADP Malate dehydrogenase; PEP-CK= Phosphoenolpyruvate Carboxykinase; PEPC-K= Phosphoenolpyruvate Carboxylase; PPDK-RP= Pyruvate orthophosphate dikinase regulatory protein; BASS 2= Bile Acid Sodium Symporter; DiT 1= Dicarboxylate Transporter 1

2.2.3: Kranz Anatomy formation in C4 plants

Anatomical studies of C4 plants have revealed the presence of a wreath-like arrangement of bundle sheath and mesophyll cells around the vascular tissue (Kranz anatomy) and this is absent in C3 and C3-C4 intermediates (Ku et.al., 1991; Muhaidat et. al., 2007). This unique arrangement of cells in the leaf ensures the availability of increased CO_2 concentration within the bundle sheath cells, where Rubisco is located (Ludwig, 2013). Kranz anatomy formation has been proposed as the initial step towards C4 photosynthesis evolution. Recent studies have identified that the SCARECROW (SCR) transcription factor, which originally functions in root vein

formation, is also involved in bundle sheath cells proliferation (Bergh et al., 2014). This finding is a key step into identifying genes involved in regulating Kranz anatomy formation in C4 plants (Slewinski et al., 2012).

Comparative analysis of leaf transcriptome of maize (C4) and rice (C3) revealed differential expression of 425 genes between them. Of these, 71 genes have been predicted to encode transcription factors that could potentially regulate the Kranz anatomy (Wang et al., 2013). Putative negative regulators (members of transcription factor gene families such as *AP2*, *bZIP*, *bHLH*, *HB*, *MADS*, *NAC*, *SBP*, *MYB*, *CAMTA* and ARF) of Kranz anatomy were highly expressed in rice but repressed in maize (Wang et al., 2013). On the other hand, several transcription factor gene families such as *AP2*, *bHLH*, *C2H2*, *GATA*, *GRAS*, *GRF*, *HD-ZIP*, *MADS*, *MYB* and *SBP* factors that potentially function as positive regulators of Kranz anatomy were found to be highly expressed in maize (C4 plant) but at very low levels in rice (C3) (Wang et al., 2013; Slewinski et al., 2012). Intriguingly, several of the above mentioned transcription factor families are known to be regulated at the post-transcriptional level by conserved miRNAs in plants (Table 2.2).

The anatomy of leaves of C3-C4 intermediate species show bundle sheath cells surrounding vascular bundles but the mesophyll cells do not form a distinct wreath-like arrangement around the bundle sheath cells as observed in the leaves of C4 species. Rather, they are sparsely arranged just as in C3 species, which results in large interveinal distances between the mesophyll and bundle sheath cells (Brown and Hattersley, 1989). This arrangement facilitates loss of CO_2 during its transport from the mesophyll to bundle sheath cells. It is worth noting that *F. ramosissima* (C3-C4 intermediate) also exhibits an arrangement similar to the Kranz anatomy but the arrangement of bundle sheath and mesophyll cells are less distinctive (Bouton et al., 1986).

Table 2.2: Candidate transcription factors that positively regulate Kranz anatomy in maize (Wang et al., 2013)

Putative Positive Regulators	Transcription Factors
GRMZM2G146688	AP2-EREBP family
GRMZM2G151542	AP2-EREBP family
GRMZM2G021573	AP2-EREBP family
GRMZM2G399072	AP2-EREBP family
GRMZM2G121309	Aux/IAA family
GRMZM2G163975	bHLH FAMILY
GRMZM2G098988	bHLH FAMILY
GRMZM2G045883	bHLH FAMILY
GRMZM2G015666	bHLH FAMILY
GRMZM2GO82586	bHLH FAMILY
AC215201.3-FG008	bHLH FAMILY
GRMZM2G095899	bHLH FAMILY
GRMZM2G178182	bHLH FAMILY
GRMZM2G123900	C_2C_2-Dof OBP3-like
GRMZM2G318592	C_2H_2 family
GRMZM2G028046	C_2H_2 family MRPI-like
GRMZM2G136494	C_2H_2 family MRPI-like
GRMZM2G150011	C_2H_2 family DOT5-like
GRMZM2G002280	C_3HC_4 RING ZnF
GRMZM2G462623	DP-1 family
GRMZM2G140669	GATA ZnF family
GRMZM2G132794	GATA ZnF family (SHR)
GRMZM2G172657	GATA ZnF family (SHR)
GRMZM2G131516	GATA ZnF family (SCR1)

Putative Positive Regulators	Transcription factors
GRMZM2G119359	GRF
GRMZM5G850129	GRF
GRMZM5G893117	GRF
GRMZM2G178102	HD-ZIP III family
GRMZM2G098813	LFY family (ZFL 1)
GRMZM2G471089	MADS family
GRMZM2G171365	MADS family (ZmMADS1)
GRMZM2G469304	Putative Ternary complex factor MIP1
GRMZM2G039074	Myb family KAN-like
GRMZM2G374986	Myb family
GRMZM5G887276	Myb family
GRMZM2G111045	Myb family MIXTA-like
GRMZM2G040924	Myb family MIXTA-like
GRMZM2G312419	Myb family LOF-like
GRMZM2G131577	Basic TF (NAC domain)
GRMZM2G126018	SBP family
GRMZM2G061734	SBP family
GRMZM2G148467	SBP family
GRMZM2G097275	SBP family
GRMZM2G472945	TLP-family
GRMZM2G377217	WRKY family
GRMZM2G425236	ZnF-HD family
GRMZM2G417229	ZnF-HD family
GRMZM2G069365	ZnF-HD family

Table 2.3: Candidate transcription factors that negatively regulate Kranz anatomy in maize (Wang et al., 2013)

Putative Negative Regulators	Transcription Factors
GRMZM2G086573	AP2-EREBP family
GRMZM2G156006	AP2-EREBP family
GRMZM2G028980	ARF family
GRMZM2G085751	bHLH family
GRMZM2G064638	bHLH family
GRMZM2G045109	bHLH family
GRMZM2G180406	bHLH family
GRMZM2G137541	bHLH family
GRMZM2G077124	bZIP family
GRMZM2G052102	bZIP family
GRMZM2G000842	bZIP family
GRMZM2G176063	C2C2-Dof family
GRMZM2G140694	C2C2-Dof family
GRMZM2G171600	CAMTA family
GRMZM2G132367	HB family
GRMZM2G062244	HB family
GRMZM2G060544	LOB family
GRMZM2G005155	MADS family
GRMZM2G137510	MADS family
GRMZM2G181030	MYB-related family
GRMZM2G003715	NAC family
GRMZM2G065451	SBP family
GRMZM2G170034	LIM domain ZnF

2.2.4: Photorespiratory CO_2 pump formation

Another step towards the evolution of C4 photosynthesis is the formation of photorespiratory CO_2 pump or glycine shuttle (Bauwe et al., 2010). This shuttle is characterized by compartmentalization of glycine decarboxylase within the bundle sheath cells. Thus, the gene expression for this enzyme is restricted to the bundle sheath cells. Hence, products of photorespiration (glycine) are channeled from the mesophyll cells into the bundle sheath cells for decarboxylation to generate CO_2. Due to the activities of glycine shuttle, the carboxylation efficiency of Rubisco increases, while its oxygenase efficiency reduces. This ultimately increases

the efficiency of photosynthesis (Gowik et al., 2011; Bauwe et al., 2010). This unique feature is restricted to the C3-C4 intermediate plants.

2.3: Regulation of Photosynthesis

It is known that during the evolution of C4 photosynthesis cycle from the ancestral C3 type, significant changes at the molecular level did occur (Hausler et al., 2002). Moreover, C4 photosynthesis has independently evolved over 62 times in 19 different lineages. It has also been postulated that these frequent evolutionary changes are because of the ease in the recruitment of C4 cycle genes over the period of evolution (Sage et al., 2012). Genes required for regulating C4 photosynthesis cycle exist in ancestral C3 plants but they have been predicted to have different functions that include regulating some non-photosynthetic processes in C3 plants (Taylor et al., 2010). An example is the activity of pyruvate orthophosphate dikinase in nitrogen mobilization in leaves of C3 plants during senescence. In C4 plants however, pyruvate orthophosphate dikinase is utilized in the regeneration of phosphoenolpyruvate (Aoyagi et al., 1984). Another enzyme is β-carbonic anhydrase (an isoform of carbonic anhydrase) known to regulate lipid biosynthesis in cotton seeds and root nodules of legumes (Hoang et al., 1999; Hoang and Chapmann, 2002). In C4 photosynthesis however, carbonic anhydrase promotes the production and continuous transport of CO_2 to the carboxylation reaction site (Price et al., 1994; Aubry et al., 2011). Another regulatory enzyme is phosphoenolpyruvate carboxylase, which regulates nitrogen assimilation, the tricarboxylic acid cycle and biosynthesis of amino acids in C3 plants depending on growth stage of the plant development (Melzer et al., 1987; Guy et al., 1989). In the C4 photosynthesis cycle, phosphoenolpyruvate carboxylase is a key enzyme that catalyzes the production of oxaloacetate, which is the source of CO_2 for photosynthesis after decarboxylation (Aubry et al., 2011). Such genes need to be modified in order to increase their transcript abundance, alter their roles, and acquire cell-specific expression (Gowik et al., 2004). Even though, it has been suggested that no novel elements are required for utilization of the C4 cycle, altered gene

expression patterns are essential to the development, regulation of the Kranz anatomy formation (specific arrangement of mesophyll and bundle sheath cells), and cell-specific accumulation of enzymes (Sage, 1999).

2.4 The importance of transcription factors and untranslated regions (UTRs) in the transcripts in regulating photosynthetic genes

Thus far, our understanding of the role of transcription factors in C4 photosynthesis is limited. Studies in maize have suggested that the activity of transcription factors such as Golden2 and Golden-like1 transcription factors may be important in the evolution of C4 photosynthesis (Hall et al., 1998). This may be due to their ability to target high numbers of photosynthesis related genes in *Arabidopsis* (Waters et al., 2009).

Studies about photosynthesis related genes have underpinned the major role that untranslated regions (UTRs) of mRNAs play in cell specific patterns of gene expression (Ludwig, 2013). For instance, nucleotide sequences in the 5' and 3' UTR regions have been identified to be involved in regulation of carbonic anhydrase expression in mesophyll cells of C4 plants (Kajala et al., 2012). Similar observations have also been made in the specific expression of Rubisco in the bundle sheath cells and Pyruvate phosphate dikinase (PPDK) expression in the mesophyll cells of *F. bidentis* (C4). The UTR region of small subunit of Rubisco has been identified as a stability determinant involved in the increased accumulation of RBCS transcripts in bundle sheath cells of C4 plants (Patel et al., 2006; Ludwig et al., 2013). In addition, *MeI* (codes for NADP-ME) expression levels in *F. bidentis* (C4) are also known to be regulated by nucleotide sequences at the 3'UTR and 5'UTR regions of the gene (Lai et al., 2002; Ludwig et al., 2013).

2.5: Endogenous small RNAs (sRNAs)

Endogenous small non-coding RNAs of about 20-24 nucleotides have been identified to regulate gene expression in plants and animals (Ambros, 2004; Bartel, 2004). On the basis of the nature (single-stranded or double-stranded) of the precursor transcripts as well as processing steps involved in their biogenesis, small RNAs are broadly categorized into two classes, i.e., microRNAs (miRNAs), and small interfering RNAs (siRNAs). Similarities between miRNAs and siRNAs include their size and both are incorporated into RISC and act as specificity determinants. However, the fundamental difference between siRNAs and miRNAs is lies in their precursors from which they are generated. siRNAs are processed from long, double stranded RNAs whereas miRNAs are processed from single stranded RNA molecules that can adopt an imperfect stem-loop secondary structure (Lagos-Quintana et al., 2001; Lau et al., 2001).

Within the siRNAs, various sub-classes have been characterized. These include trans-acting short interfering RNA (tasi-RNAs - a class of siRNAs that are generated from a non-coding transcript that is cleaved due to miRNA targeting them - such cleaved transcripts are in turn converted into double stranded RNAs due to RNA dependent RNA Polymerase activity - which is again processed into 21-nt long siRNAs that in turn regulates the abundance of protein coding mRNAs by guiding a cleavage), phasiRNAs (Phased siRNAs - that are derived from the dsRNA resulting from miRNA targeting a protein coding mRNAs), natural antisense siRNAs (natsiRNAs - are derived from the dsRNA resulting from the expression of both sense and antisense strands of the same locus) and heterochromatic siRNAs (hc-siRNAs – derived from the transcripts originated from the heterochromatin) in plants (Brodersen and Voinnet, 2006; Sunkar and Zhu, 2007).

2.6: MicroRNAs

Plant miRNAs are of approximately 21 nucleotides long and form key components of a complex network of gene regulatory pathways (Jones-Rhoades et al, 2006). MicroRNAs were first identified in *Caenorhabditis elegans* in 1993 (Lee et al., 1993) and in *Arabidopsis* in 2002 (Llave et al., 2002). Since then approximately 36,000 miRNAs have been identified from diverse species of plants, animals, including viruses (www.miRBase.org).

2.7: MicroRNA biogenesis and their mode of action in plants

In plants, miRNAs are generated from the dicing of the primary miRNA transcripts that adopt a hairpin-like structure by the members of the DICER-LIKE (DCL) protein family (Xie et al., 2004; McHale et al., 2013).

Initially, RNA POLYMERASE II transcribes primary miRNA transcripts from the miRNA genes in a similar fashion as mRNA transcription. The resulting RNA molecule adopts a hairpin-like structure. Dicer-like 1 (DCL1) enzyme together with HYPONASTIC LEAVES (HYL1 which is a double stranded RNA binding protein), SERRATE (SE), a zinc-finger protein and DAWDLE (DDL-RNA binding protein) (Han et al., 2004; Vazquez et al., 2004; Lobbes et al., 2006; Yang et al., 2006) dices the stem-loop structure and releases miRNA (guide strand) and miRNA* (passenger strand) duplex. The terminal sugar at the 3' ends of miRNA's duplexes are methylated by HUA ENHANCER 1 (HEN 1), a methyl transferase (Yu et al., 2005). In the absence of methylation, the miRNA duplex is subjected to uridylation and can be degraded by small RNA degrading nucleases (Ramachandran et al., 2008). Additionally, methylation of miRNA duplexes is essential to avoid any modification of their 3'ends (Chen 2005; Yang et al., 2006). All these above mentioned steps involving miRNA biogenesis occurs in the nucleus (Fig. 2.4). The methylated miRNA-miRNA* duplex released in the nucleus are exported to the cytoplasm by a nuclear membrane-localized transport protein called HASTY (HST). In the cytosol, the guide miRNA is unwound from the passenger miRNA* and loaded onto Argonaute 1

(AGO1) of the RNA-Induced Silencing Complex (RISC). The miRNA* strand is then degraded and persists at low levels (Jones-Rhoades et al, 2006). The incorporated miRNA guides the RISC in recognizing target mRNA based on perfect or near perfect sequence complementarities between the nucleotides of the mature miRNA and the mRNA target. This results in mRNA cleavage and degradation or repression of protein synthesis (translation) or even both in plants (Tang et al., 2003; Bartel et al., 2004; Baulcombe et al., 2004; Aukerman and Sakai, 2003; Chen, 2004). miRNAs linked with developmental roles show regions of activity that apparently coincides entirely with the site of transcription of miRNA-coding genes (Parizotto et al., 2004). This characteristic of miRNAs makes them cell-type specificity determinants in plant leaves.

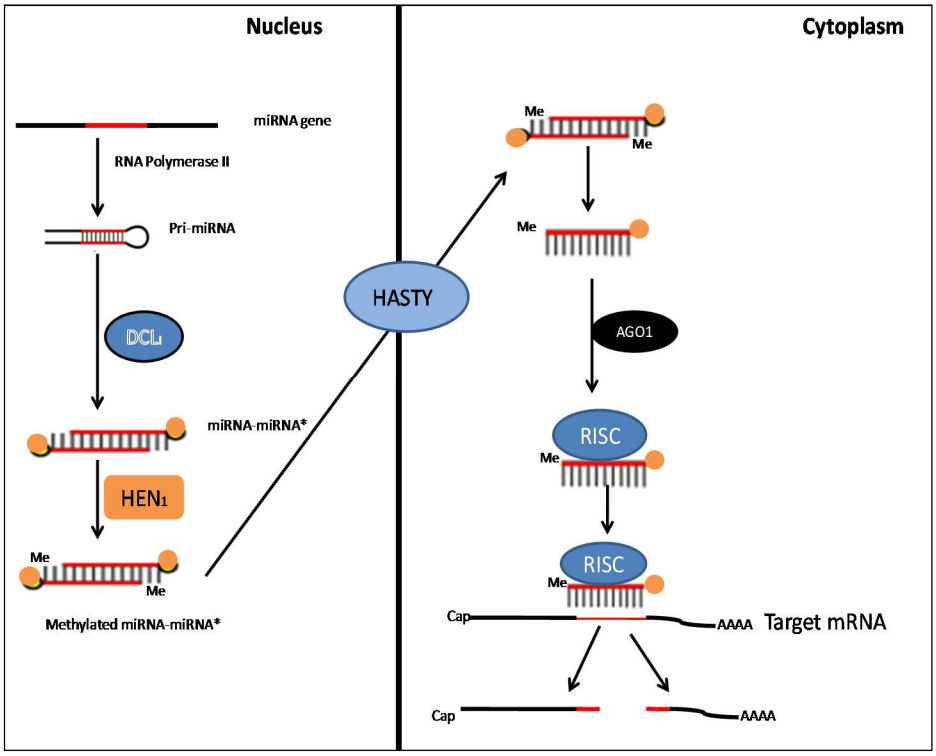


Figure 2.4: Plant miRNAs biogenesis and function (modified from Jones-Rhoades et al., 2006)

2.8: Conserved and non-conserved miRNAs in plants

Based on their conservation, miRNAs have been classified into conserved and non-conserved miRNAs. Until date, 22 miRNA families (Table 2.4) are known to be highly conserved between monocotyledonous and dicotyledonous plants (Jones-Rhoades et al., 2006). In addition to conserved miRNA families, recent advancements in sequencing technologies (that could sequence to a greater depth) have led to the discovery of lineage-specific and species-specific miRNA families in plants. For instance, miR403, a dicot-specific miRNA, has been found to be abundantly expressed in dicots but absent in monocots (Sunkar and Zhu, 2004; Sunkar and Jagadeeswaran, 2008). On the other hand, miR444 (Sunkar et al., 2005; Lu et al., 2008) has been identified in all monocots but not in dicots profiled to date. Some other lineage-specific miRNAs have been found to be conserved in closely related legumes, such as *Medicago truncatula*, chickpea, soybean and alfalfa but not in *Arabidopsis*, rice, or other non-leguminous plants (Jagadeeswaran et al., 2009). The prevalence of miRNAs specifically encoded in plant lineages and species portrays the existence of a complex network of post-transcriptional regulation that operates in plants. These lineage-specific and species-specific miRNAs may be involved in regulating metabolic pathways unique to a particular lineage or plant species.

Table 2.4: Highly conserved miRNA families and their target gene families in higher plants

miRNA family	Target gene family
miR156	SBP Factors
miR159	MYB Factors
miR160	Auxin Response Factors
miR162	DCL-1
miR164	NAC Factors
miR165/166	HD-ZIP Factors
miR167	Auxin Response Factors
miR168	Argonaute-1
miR169	NFY subunits or CAAT box Binding Factors
miR170/171	Scarecrow-like Factors
miR172	AP2-like Factors
miR319	TCP Factors
miR390/Tas3-siRNA	Auxin Response Factors
miR393	TIR1 (F-box proteins)
miR394	F-Box proteins
miR395	ATP Sulfurylases and a sulfate transporter
miR396	Growth Regulating Factors
miR397	Laccases
miR398	Cu/Zn SODs
miR399	E2 ligase and a phosphate transporter
miR408	Plantacyanin
miR444	MADS-box factors

Over the years, it has been well established that miRNAs are important regulators of gene expressions critical for apical meristem maintenance, leaf morphogenesis and development, flower formation and development, root growth and development, vegetative-to-reproductive phase transitions and myriad of other developmental processes (Aukerman et al., 2003; Palatnik et al., 2003; Chen et al., 2004; Schwab et al., 2005; Guo et al., 2005; Baker et al., 2005; Xie et al., 2006; Sieber et al., 2007; Wang et al., 2008; Carlesbecker et al., 2010; Schommer et al., 2012; Lima et al., 2012; Yang et al., 2013). Besides, miRNAs have been shown to alter their abundances in response to diverse biotic/abiotic stress conditions including nutrient deficiencies (Sunkar et al., 2007; 2012). Such an altered expression of miRNAs should affect the abundances of their target genes, which could contribute for plant fitness to survive the stressful conditions (Sunkar, 2010).

CHAPTER III

METHODOLOGY

Two approaches have been successfully used to identify and characterize miRNAs in plants. These include direct sequencing of small RNA libraries or bioinformatic prediction of miRNAs. The drawback of the bioinformatics approach is that it can only predict conserved miRNAs provided the sequenced genome information of the species in question is available. The leaves of *F. robusta* (C_3) only possess mesophyll cells, whereas *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4) have distinct bundle sheath cells in addition to the mesophyll cells. In this study, we sequenced small RNAs using a high throughput sequencing platform for identification of miRNAs not only in entire leaves but also in mesophyll and bundle sheath cells of three different *Flaveria* spp.

3.1: Plant material

To date 21 species of *Flaveria* have been identified (Table 1). The unique feature of this genus is that its' species are represented by C_3, C_3-C_4 intermediate and C_4 photosynthesis. Three *Flaveria* spp. namely *Flaveria robusta* (C_3), *Flaveria ramosissima* (C_3-C_4 intermediate) and *Flaveria bidentis* (C_4) were used in present investigation. These species were selected because they have been used for transcriptome and other molecular or biochemical studies (Gowik et al., 2011; Sage et al., 2012 and Mallman et al., 2014).

3.2: Growth Conditions

Seeds of the above named three *Flaveria* spp. were germinated in pots filled with 1:2:1 parts of sand, top soil and vermiculite, respectively, for 14 days at 26°C in a growth chamber. They were watered once every 3 days for four weeks, after which the seedlings were maintained in a growth chamber at 28/18°C with 16/8 h day/night cycle (day/night cycle) and 65% relative humidity. Light was provided by a combination of fluorescent and incandescent lamps.

3.3: Harvesting leaves from *Flaveria* spp.

The second and fourth pairs of matured, fully expanded leaves were harvested from four-month-old *F. robusta* (C_3), *F. ramosissima* (C_3-C_4 intermediate), and *F. bidentis* (C_4) in the presence of light. They were immediately flash frozen in liquid nitrogen and stored at -80°C until used for downstream analysis.

3.4: Isolation of total RNA from the leaves

Approximately 1g of frozen leaf tissue from each *Flaveria* spp. was ground into fine powder using a mortar and pestle. The powder was transferred to a 20ml centrifuge tube containing 10ml of Trizol and incubated at room temperature for 5 minutes. The tubes were then centrifuged at 13,000g for 5 minutes. The supernatant formed was decanted into a clean 50ml centrifuge tube and chloroform (1ml: 200µl) was added to it. The resultant mixture was vigorously vortexed twice and then centrifuged at 13,000g for 30 minutes. The aqueous supernatant layer formed was carefully pipetted into a new 50ml centrifuge tube and equal volumes of isopropanol were added to it. The supernatant was discarded and the RNA precipitate The mixture was incubated at room temperature for 10 minutes and centrifuged at 13,000g for 30 minutes. formed at the bottom of the tube was washed twice with 80% cold ethyl alcohol. The RNA pellet was then air-dried at room temperature whilst on ice for about 15 minutes and then dissolved in DEPC treated (RNase free) water. The quality of total RNAs was assessed by absorbance ratio (A260/A280) using a

Nanodrop ND-1000 spectrophotometer as well as by ribosomal RNA integrity using gel electrophoresis.

3.5: Small RNA library construction

The Truseq Small RNA Library Preparation protocol from Illumina was followed in constructing small RNA libraries for *Flaveria* spp. In short, approximately 10µg of total RNA sample was ligated with 3' and 5' illumina adapters, respectively. Using T4 RNA ligase, both ligation reactions (5' and 3' ligations) were performed at 37°C for 1 to 2 hours. This was followed by a reverse transcription (RT) to generate cDNA, which was amplified using 25 cycles by polymerase chain reaction (PCR). The amplified PCR product was size fractionated on 6% PAGE, purified and used for sequencing.

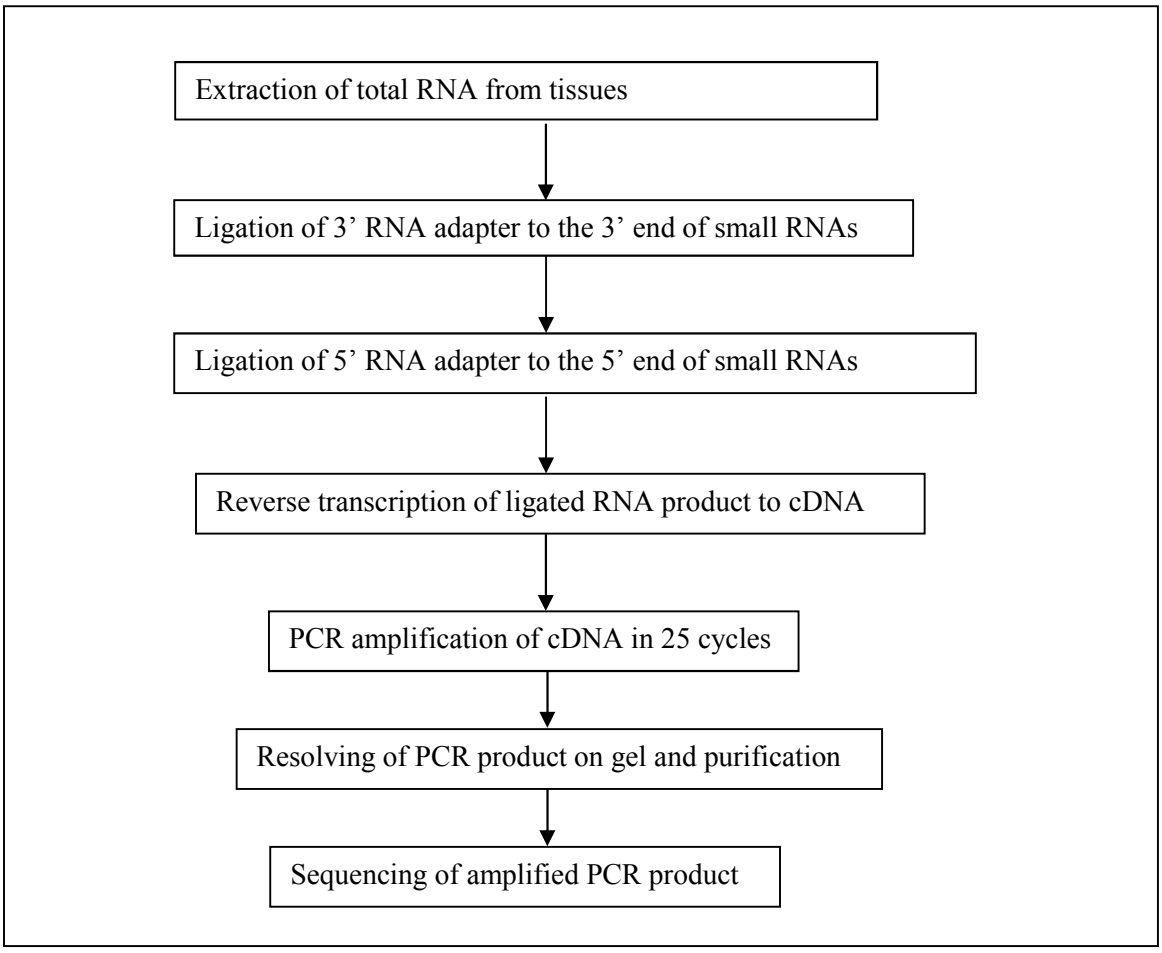


Figure 3.1: Basic outline of small RNA library construction.

3.6: Analysis of small RNA libraries

Firstly, the sequences in the small RNA libraries that match perfectly to 3' and 5' adapter sequences were removed because these represent adapter self-ligation products. Next, the 5' adapter was trimmed off and the small RNA was extracted from the reads. The small RNA reads that are shorter than 17 nucleotides and longer than 28 nucleotides were discarded. The remaining reads were BLAST searched against sequence entries in the European ribosomal RNA database to remove those sequences that match with rRNAs, tRNAs, snRNAs and snoRNAs. The remaining reads were mapped to conserved miRNA sequences downloaded from miRBase21 (www. miRBase.com) for identifying homologs of conserved miRNAs in *Flaveria* spp. In addition, the normalized frequencies for the miRNA families identified were recorded in Reads Per Ten million (RPTM). Normalized frequency (the number of reads for each miRNA family in the library divided by the total number of reads for that library and multiplied by 10,000,000 million. Figure 3.2 shows a summary of the steps that were involved in analyzing the sequence reads generated from small RNA libraries.

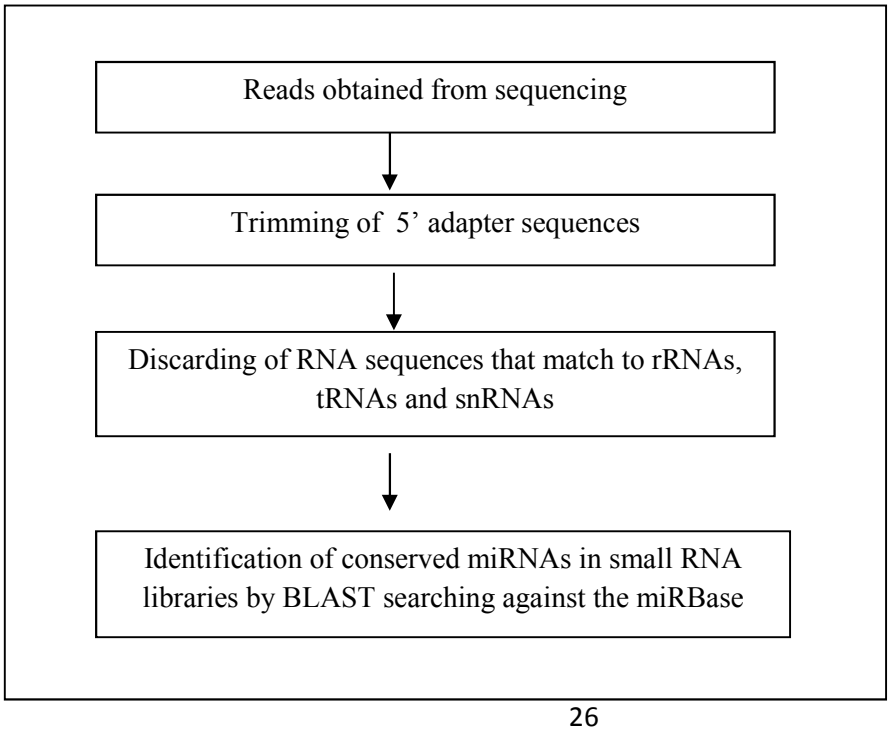


Figure 3.2: Basic outline of sequence analysis to identify conserved miRNAs.

3.7: Small RNA blot analysis for characterizing miRNA abundances in leaves

Small RNA blot analyses were performed to validate the results obtained from the high throughput sequencing of leaf small RNA libraries of *F. robusta* (C_3), *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4). In brief, 20µg of total RNA extracted from the leaves of each of the *Flaveria* spp. were resolved on a 15% denaturing polyacrylamide gel. The size fractionated RNAs were electrophoretically transferred from the gel onto Hybond-N^+ blotting membranes using a wet-blot transfer unit. The membranes were UV cross-linked and baked for 2 hours at 80°C in order to immobilize the RNAs. The membranes were then prehybridized in PerfectHYB$^+$ hybridization buffer (Sigma) at 37°C in an incubation chamber for at least 3 hours. This was followed by hybridization of the membranes using p32-labelled oligonucleotide probes. The probes represent DNA oligonucleotides whose sequences are complementary to miRNA sequences, which were end-labeled with γ-^{32}P-ATP in a reaction catalyzed by T4 polynucleotide kinase. After overnight hybridization, the membranes were washed with buffer containing 2xSSC and 0.1% SDS at 50°C and briefly air-dried. The membranes were exposed to a phosphorscreen for about 2 hours to overnight at room temperature. The phosphorscreen was scanned with a typhoon scanner to detect the signals derived as a result of probe hybridization with the miRNAs.

3.8: Enzymatic isolation of mesophyll cells from leaves of *Flaveria* spp.

The leaves of *F. robusta* (C_3) only possess photosynthetic mesophyll cells, whereas *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4) have distinct bundle sheath cells in addition to the mesophyll cells. In order to gain an insight into the differences in miRNA compositions within these named cells, mesophyll and bundle sheath cells were isolated from their leaves. The procedures described by Kanai and Edwards (1973), Lahiri et al., (2000) and Chang et al., (2012) were followed in isolating mesophyll cells from the leaves of *Flaveria* spp.

In brief, the harvested leaves were cut into 2mm segments with a sterilized razor blade and immersed into the digestion medium (1.5% cellulose Onozuka, 0.1% macerase Onozuka, 0.6M sorbitol, 20mM 2-[N-morpholino] ethane sulfonic acid [MES] pH 5.8, 100mM β-mercaptoethanol, 1mM calcium chloride and 10mM dithiothreitol) and incubated at room temperature for 3 hours. The mixture was filtered through a 500µm mesh and the mesophyll cells were extracted into the filtrate. The residues left on the mesh were washed gently with 0.6M sorbitol to release any additional mesophyll cells into the filtrate. The collected filtrate was further filtered through 80µm nylon mesh to separate mesophyll cells that could be entangled with bundle sheath strands. The resultant filtrate, which contains the mesophyll cells were collected and then centrifuged at 500g for 10 minutes. The supernatant was discarded and the precipitate formed represent pure mesophyll cells, which were suspended in Trizol for isolating total RNA.

3.9: Mechanical isolation of bundle sheath cells from the leaves of *Flaveria* spp.

The enzymatic digestion procedure for isolating mesophyll cells was known to cause damage to bundle sheath cells (Potter and Black, 1982). Therefore, a mechanical procedure described by Edwards and Black (1971) was followed to isolate bundle sheath cells from the leaves of C_3-C_4 intermediate and C_4 *Flaveria* spp. The leaves were cut into 2mm segments and immersed in isolation buffer (50mM Tris-HCL pH 8.0, 0.6M sorbitol, 5mM $MgCl_2$, and 100mM β-mercaptoethanol) and then homogenized using a Warring blender for 1 minute at high speed. This was followed with filtration of the homogenate through 500µm nylon mesh and discarding of the filtrate. The homogenization and filtration steps were repeated twice and the bundle sheath cells were again filtered through 80µm nylon mesh. The residues (bundle sheath cells) were then washed with cold isolation buffer to discard any attached mesophyll cells associated with the bundle sheath cells. This step was repeated three times to ensure purity of isolated bundle sheath cells. The residue (bundle sheath cells) was then suspended in Trizol for isolating total RNA.

3.10: Immunoblot analysis of Rubisco and PEPC for monitoring purity of isolated mesophyll and bundle sheath cells

It has been established that Rubisco enzyme operates within the bundle sheath cells of C_4 plants. On the other hand, it is abundant within the mesophyll cells of C_3 plants. It has also been confirmed that PEPC enzyme operates within the mesophyll cells of C_4 plants. In order to monitor the purity of isolated mesophyll and bundle sheath cells from the leaves of three *Flaveria* spp., the expressions of Rubisco and PEPC proteins (markers for specific cell types in C_3 and C_4 plant species) were investigated by using immunoblot assay. To achieve this, total proteins were extracted from the isolated mesophyll and bundle sheath cells using protein extraction buffer (0.7M sucrose, 0.5M Tris, 30mM HCl, 50mM EDTA, 0.1M KCl, 2% 2-mercaptoethanol and 2mM PMSF). The extraction process was similar for both cells and it involved homogenization of the cells by grinding followed by centrifugation. The supernatant formed was collected and Bradford's Assay was performed to determine total protein concentration. Approximately, 7µg of the total protein in the supernatant was mixed with 4X protein sample buffer (40% Glycerol, 240 mM Tris/HCl pH 6.8, 8% SDS, 0.04% bromophenol blue and 5% beta-mercaptoethanol) and loaded onto SDS-PAGE gels (consisting of 5% stacking gel and 12% separating gel for size fractionation). A protein marker was also loaded to assist in identifying proteins of interest based on their molecular weight. After resolving the proteins on the SDS-PAGE gel at 100V for an hour, one of the gels was stained with Commassie Blue reagent for about 2 hours and then destained using de-staining buffer (20% methanol and 10% glacial acetic acid) overnight at room temperature. For immunoblotting analysis, the resolved proteins were transferred onto a nitrocellulose membrane (Thermo Scientific) using an electroblotting semi dry transfer apparatus for 10 minutes. The membrane was stained using Ponceau reagent to check for successful transfer of proteins onto the nitrocellulose membrane. Then the membrane surface was blocked using 5% nonfat dry milk powder and 0.1% Tween 20 in Tris-buffered saline (TTBS) for an hour. This was followed by incubation of the membrane with polyclonal rabbit anti-Rubisco antibodies (1:5,000

dilution) (kind gift from Susanne von Caemmerer, Australian National University, Canberra) for overnight. The membrane was washed and incubated in goat anti-rabbit polyclonal antibody (secondary antibody), conjugated with Horse Radish Peroxidase (HRP) (1: 10,000 dilution) for 2 hours. The unbound secondary antibodies were washed off the membrane three times using TTBS. Using ECL developing reagent (Thermo Scientific), the blots were incubated for 5 minutes, exposed to X-ray film and developed using X-ray developer. The molecular weight of the detected Rubisco protein was estimated based on the molecular standards present in the protein ladder. The same procedures were used for PEPC (Agrisera antibodies) protein detection. However, due to low levels of PEPC in the mesophyll cells of *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4), a high concentration of polyclonal rabbit anti-PEPC antibodies (1: 2,000 dilution) and goat anti-rabbit polyclonal antibody conjugated with HRP (1: 5,000 dilution) were used.

3.11: Total RNA Isolation from the bundle sheath and mesophyll cells

The protocol described in section 3.7 was followed in isolating total RNAs from the isolated mesophyll and bundle sheath cells of *Flaveria* spp. Again, the quality of total RNA extracted independently from both cell types was assessed by absorbance ratio (A260/A280) using Nanodrop ND-1000 spectrophotometer as well as by ribosomal RNA integrity after size fractionation on an agarose gel. The RNA was used for constructing the small RNA libraries (section 3.5).

3.12: Construction of degradome libraries from the leaves of *Flaveria* spp.

A defining feature of miRNA-guided silencing in plants is that cleavage occurs precisely between the 10^{th} and 11^{th} nucleotide, counting from the 5' end of miRNA in the complementary region of the target transcripts. When miRNA guides cleavage on their target mRNA, two fragments namely the 5' (7-methylguanosine capped end) and 3'ends (polyadenylated end) of cleaved mRNAs are released. Degradome analysis is a high throughput sequencing-based

approach that sequences uncapped poly-adenylated transcripts including cleaved mRNA targets as a result of miRNA targeting them (Addo-Quaye et al., 2008; German et al., 2008; Li et al., 2010). Degradome sequencing was used to identify targets of miRNAs in *Flaveria* spp.

To construct degradome libraries from the leaves of *Flaveria* spp., total RNAs were first extracted as explained previously (section 3.7). The polyadenylated mRNA fragments were purified from total RNAs using the MicroPoly 'A' Purist kit (Ambion). Using T4 RNA ligase, the 5'end of the 3' cleaved fragment containing poly 'A' mRNA was ligated to an RNA adapter containing the *MmeI* restriction enzyme recognition site. Reverse transcription was carried out to generate the first strand cDNA using an oligo (dT) primer and reverse transcriptase. The cDNA formed, was amplified through 5 cycles of Polymerase Chain Reaction (PCR) and the product was digested with *MmeI* restriction enzyme. The 20-base pair long digested products were ligated to double stranded DNA adapter and amplified using 20 cycles of PCR. The amplified product was purified and sequenced. The schematic flow chart showing the steps followed for degradome libraries construction is summarized in Figure 3.3.

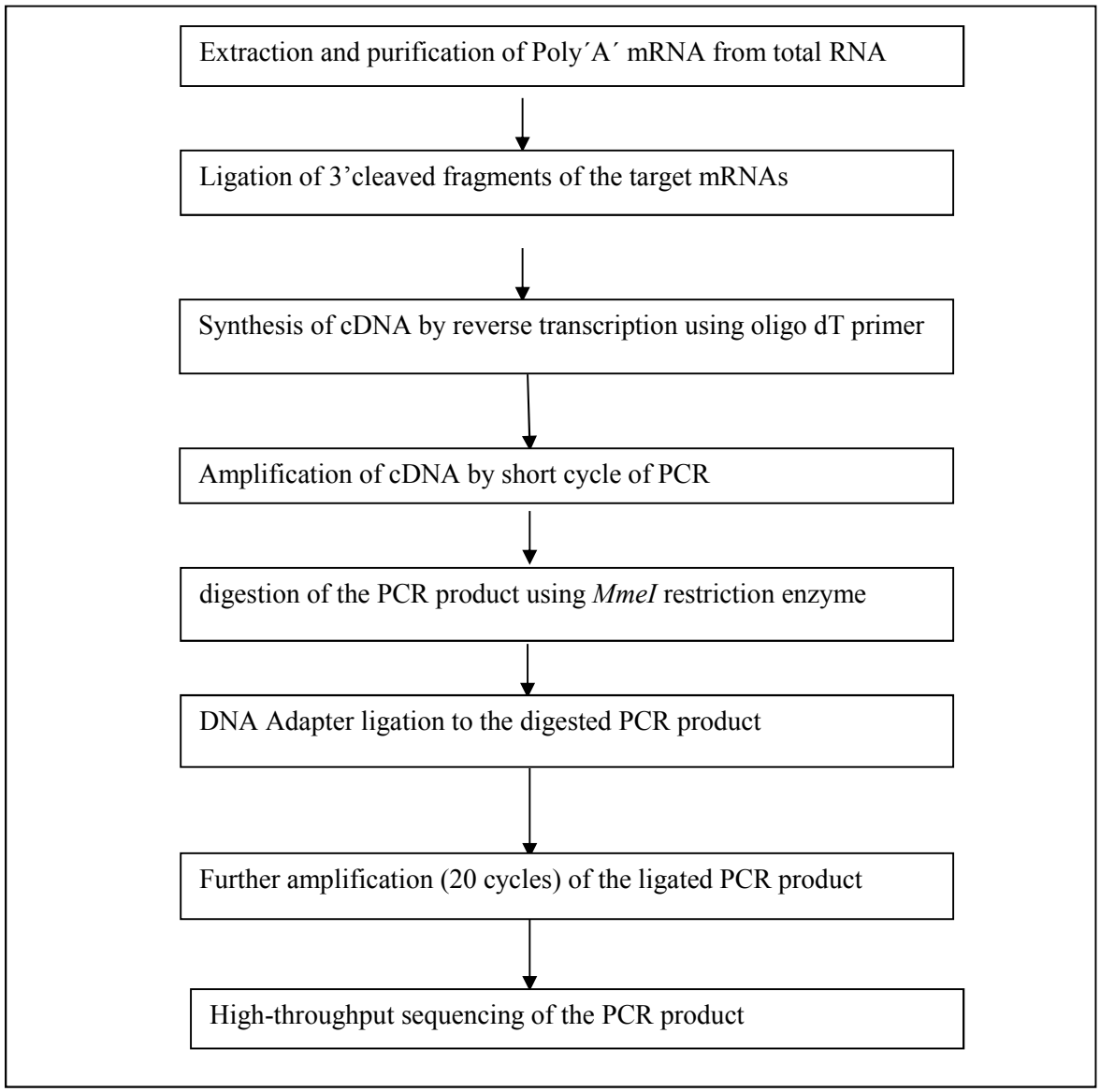


Figure 3.3: Basic outline of degradome library construction.

3.13: Computational analysis of degradome library sequencing reads

Computational algorithms such as CleaveLand pipeline and SeqTar (idm.fudan.edu.cn/zhengyun) have been developed for identifying miRNA guided cleavages on miRNA-targeted transcripts present in the sequenced degradome libraries. In this study, SeqTar computational pipeline was used in analyzing the degradome data of miRNA guided cleavages on targets in *Flaveria* spp. The SeqTar pipeline utilizes two features, i.e., (i) alignment of miRNAs and their corresponding target gene and (ii) evaluation of number of reads identified at the centre

of the miRNA: mRNA complementary region (Zheng et al., 2012). The target genes identified for each *Flaveria* spp. were limited to identifying miRNAs-target gene pairs containing less than four mismatches. The annotation for each identified target gene was obtained from the BLAST searches against the protein data base at the NCBI. Figure 3.4 shows a summary of the steps that were followed to identify the targets of miRNAs in each *Flaveria* spp.

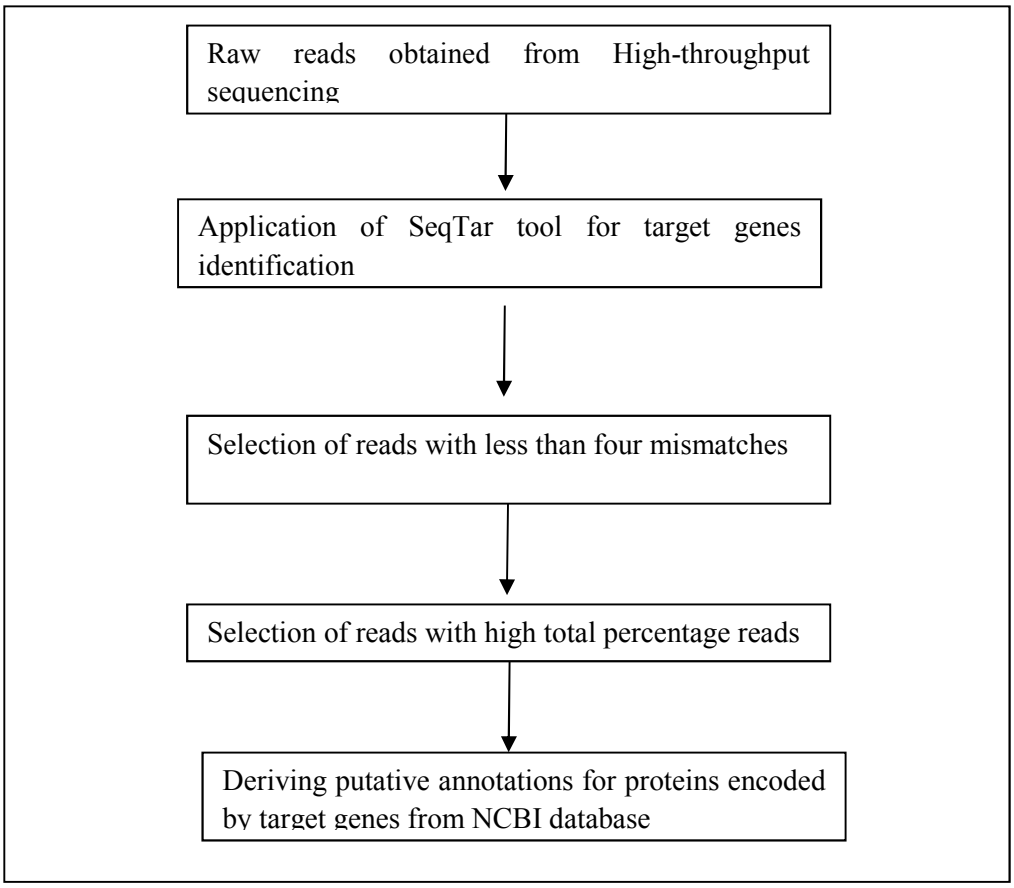


Figure 3.4: Flow chart for identifying miRNA targets in *Flaveria* spp.

CHAPTER IV

RESULTS

4.1: Identification of conserved miRNAs in *Flaveria* spp.

Conserved plant miRNAs can be predicted *in silico* by the computational approach (Zhang et al., 2006; Sunkar and Jagadeeswaran, 2008). This approach requires the availability of sequenced and assembled genome and/or significant numbers of Expressed Sequence Tags (ESTs). Cloning and high-throughput sequencing of small RNA libraries is another approach to identifying conserved and known miRNAs in plants (Lu et al., 2008; Sunkar et al., 2008). Currently for *Flaveria* spp., the genome has not been sequenced and the numbers of ESTs are also very limited. Due to these limitations, cloning and high-throughput sequencing of small RNA libraries is the best approach to identify miRNAs expressed in *Flaveria* spp. Therefore, small RNA libraries were constructed from the leaves of *F. robusta* (C_3), *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4) to identify miRNAs in the entire leaves. Additionally, to characterize miRNA component in the mesophyll and bundle sheath cells of C_3, C_3-C_4 and C_4 photosynthesis small RNA libraries were constructed from mesophyll and bundle sheath cells of *F. robusta* (C_3), *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4).

4.2: Analysis of Small RNA libraries generated from the leaves of *Flaveria* spp.

The small RNAs ranging in size between 18 - 28 nucleotides were used for the analysis. The summary of small RNA reads that can be mapped to various RNA categories such as mRNAs, ncRNAs, repeats and miRBase have been recorded (Table 4.1). There were also small RNAs that could not be mapped to annotated miRNAs or mRNAs or other categories of RNAs suggesting that they might represent endogenous siRNAs or degradation products from mRNAs. Mapping these reads require assembled transcripts or genomes. Within all small RNA libraries constructed from the leaves, the distribution of the small RNA sizes and their abundances were depicted in Figure 4.1. In all three *Flaveria* spp., two peaks, i.e., one at 21 nt and the other at 24 nt size classes were commonly observed for the total reads from the small RNA libraries (Fig 4.1). Such a bimodal distribution of small RNA sizes has been widely documented for plant small RNA populations. However, after the removal of redundant reads, the 24 nt size class of small RNAs was the highest in the unique small RNA reads for all three *Flaveria* spp. (Fig. 4.1) and such a pattern for 24 nt size class in the unique reads has also been well characterized in different plant species.

To identify homologs of conserved and known miRNAs in the small RNA libraries from the leaves, the unique reads were mapped to the miRBase21. This analysis revealed that the small RNA population from *Flaveria* spp. was represented by 42 known miRNA families. Of these, 39 were identified in *F. robusta* (C_3), 37 in *F. ramosissima* (C_3-C_4 intermediate) and 35 in *F. bidentis* (C_4). Within these known miRNA families, 21 families belonging to the highly conserved miRNA families were present in small RNA libraries of *F. robusta* (C_3) and *F. ramosissima* (C_3-C_4 Intermediate). In *F. bidentis* (C_4), 20 highly conserved miRNA families were recovered from the leaves. The commonly identified 20 conserved miRNA families in leaves of all three *Flaveria* spp. include miR156, miR159, miR160, miR162, miR164, miR165/166, miR167, miR168, miR169, miR171, miR172, miR319, miR390, miR393, miR394,

miR395, miR396, miR397, miR398, miR399. In addition to these, miR408 was found in only *F. robusta* (C_3) and *F. ramosissima* (C_3-C_4 intermediate) but not from *F. bidentis* (C_4). As expected, miR444, another miRNA family known to be conserved in monocots was not identified in any of the three *Flaveria* spp.

Most conserved miRNA families in the leaf small RNA libraries recorded higher abundances (Fig 4.2). Of these, miR166 was by far the most abundantly expressed miRNA family in the leaves of all three *Flaveria* spp. (Fig. 4.2A). The other abundantly expressed miRNA families (more than 5,000 RPTM) are represented by miR396, miR169, miR156, miR159, miR167, miR168 and miR403 (Fig 4.2). Other highly conserved miRNA families such as miR170/171 and miR398 recorded moderate abundances with their normalized frequencies ranging between 1,000 - 5,000 RPTM. Some of the highly conserved miRNA families however, recorded low (below 100 RPTM) expression levels as exemplified by miR394 and miR408 families (Fig. 4.1). Besides the above mentioned highly conserved miRNA families, several known miRNAs (known to be expressed in different plant species but not all plant species suggesting that these belong to less conserved miRNA families in plants) were also identified. These miRNA families include miR477, miR482, miR530, miR858, miR2111, miR2118, miR5139, and miR6173, whose abundances were less than 100 RPTM but more than 10 RPTM. A few other less conserved miRNA families were also recovered from *Flaveria* spp. whose abundances were extremely low (less than 10 RPTM) and these include miR894, miR1509, miR2643, miR4995, miR5072, miR5083, miR5162, miR5368, miR5770, miR6113, miR6478, miR8155 and miR8175.

An in depth analysis into the abundances of miRNA families in leaves revealed greater differences between *F. robusta* (C_3), *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4). miRNA families such as miR156, miR160, miR162, miR164 and miR390 recorded low abundances in *F. robusta* (C_3) but in moderate abundances both in *F. ramosissima* (C_3-C_4

intermediate) and *F. bidentis* (C4). miR397 recorded moderate frequencies in *F. robusta* (C3) but was low in *F. bidentis* (C4). Overall, several conserved miRNA families (miR156/157, miR159, miR162, miR167, miR168, miR169, miR170/171 and miR393) were more abundantly expressed in leaves of *F. bidentis* (C4) than in C3-C4 intermediate and C3 *Flaveria* spp.

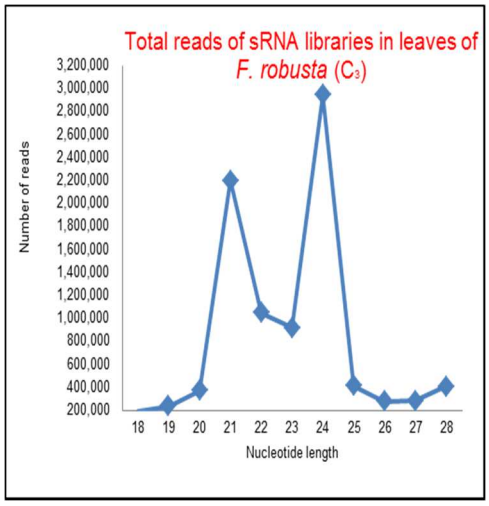
Most conserved miRNA families are represented by multiple loci/isoforms in plants. The abundances of miRNA isoforms within miRNA family varied greatly for all three *Flaveria* spp. Out of 129 individual miRNA homologs found in the library, miR166 family was represented by the most number (23) of isoforms in *Flaveria* spp. Within the miR166 family members, miR166a was the most abundantly expressed. Similarly, miR396 family was represented with 16 members, of which miR396n was the most abundantly expressed isoform. For both miR167 and miR398, ten isoforms were recovered from each family. Within the miR398 family, miR398f was the most abundant whereas in the miR167 family, miR167a recorded the highest abundance. miR156 was represented by 9 members, miR169 family by 8 members, miR159 family by 7 members, miR168 and miR171 family by 6 members, miR395 by 5 members, miR319 by 4 members, miR160 and miR393 by 3 members and lastly, miR162 with 2 members. The remaining miRNA families were represented by single isoforms in the *Flaveria* spp. (Table 4.6)

Besides the identification of known miRNAs in leaves, miRNA* for 6 known miRNAs were also identified (Table 4.2). While miR170*, miR393*, miR403* and miR530* were expressed in all three *Flaveria* spp., the rest of the miRNA* (miR397, miR408, miR530) were either expressed in only one or two of the three *Flaveria* spp.

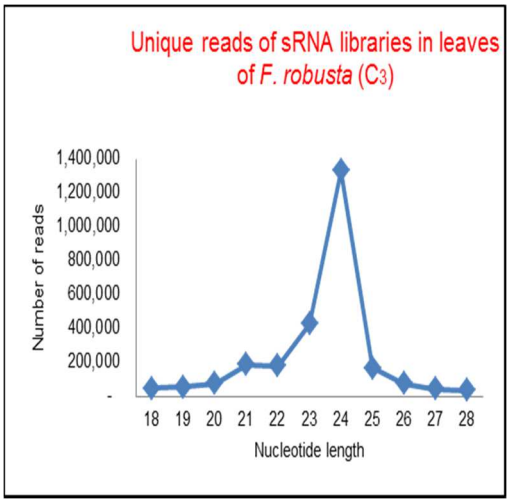
Table 4.1: A summary of total and unique reads distribution of small RNA libraries generated from the leaves of *Flaveria* spp.

	F. robusta (C_3)		*F. ramosissima* (C_3-C_4)		*F. bidentis* (C_4)	
RNA Category	reads	unique	reads	unique	reads	unique
Flaveria-mRNAs	7,319,605	536,281	4,877,428	417,041	6,583,769	545,597
ncRNAs	5,120,895	160,629	3,107,292	156,747	4,560,134	203,684
Pre-miRbase21	1,353,781	1,258	663,086	1,256	680,665	1,547
repeats	2,605,109	91,127	1,613,341	90,159	2,738,186	111,526
total	12,556,812	3,000,228	10,538,559	2,373,791	10,191,249	2,270,350

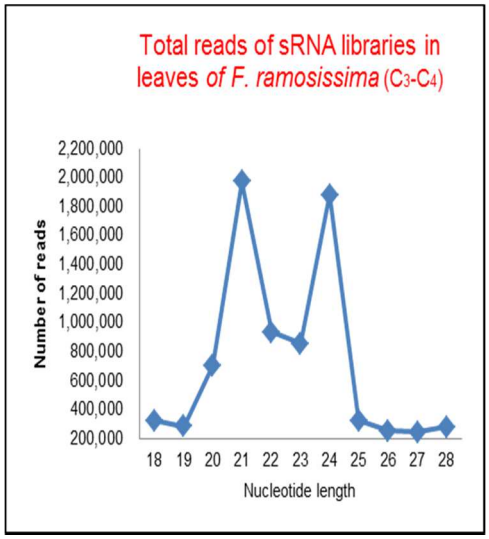
A)



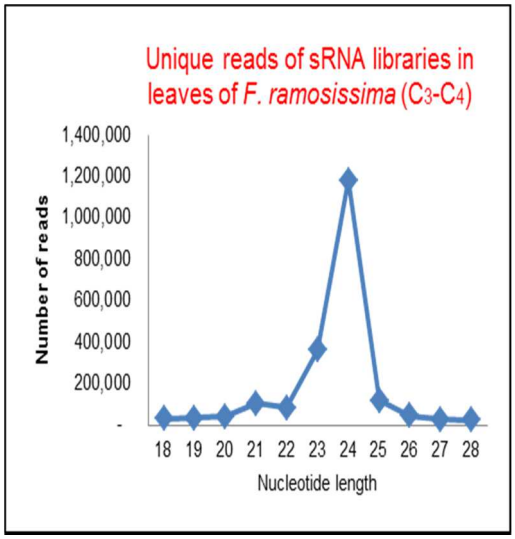
B)



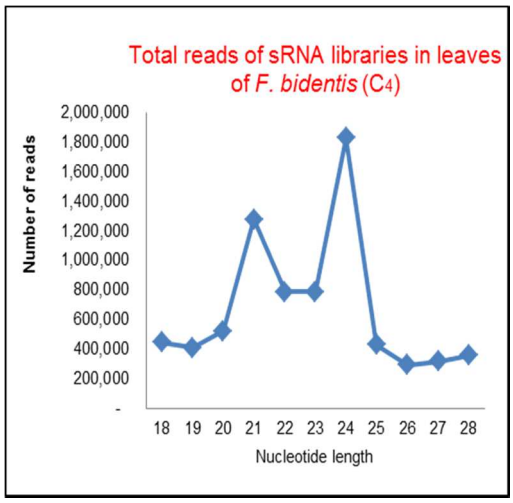
C)



D)



E)



F)

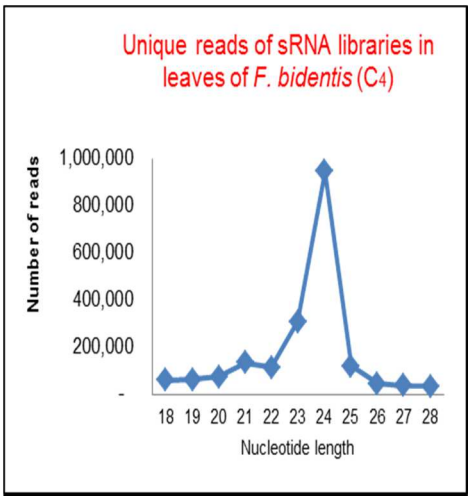
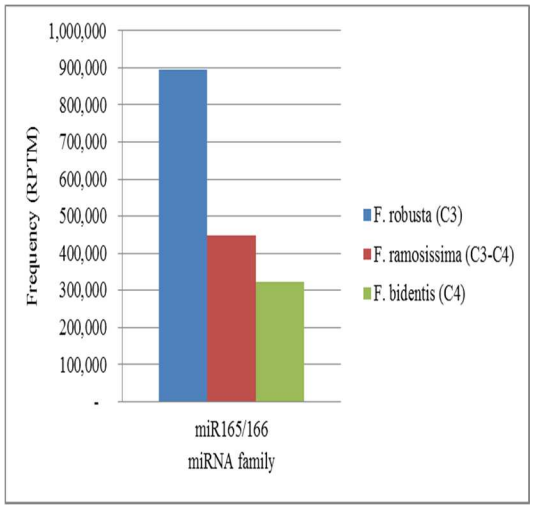
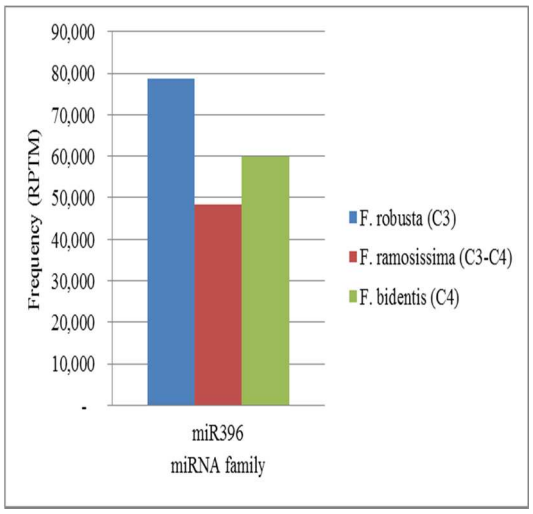


Figure 4.1 (A-F): Small RNA read length vs their abundances in the small RNA libraries from the leaves of *F. robusta* (C3), *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4).

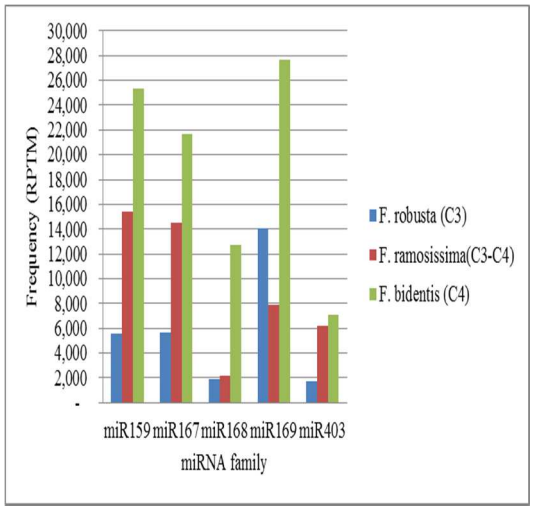
A)


1,000,000
900,000
800,000
700,000
600,000
500,000
400,000
300,000
200,000
100,000
-
Frequency (RPTM)
miR165/166
miRNA family
F. robusta (C3)
F. ramosissima (C3-C4)
F. bidentis (C4)

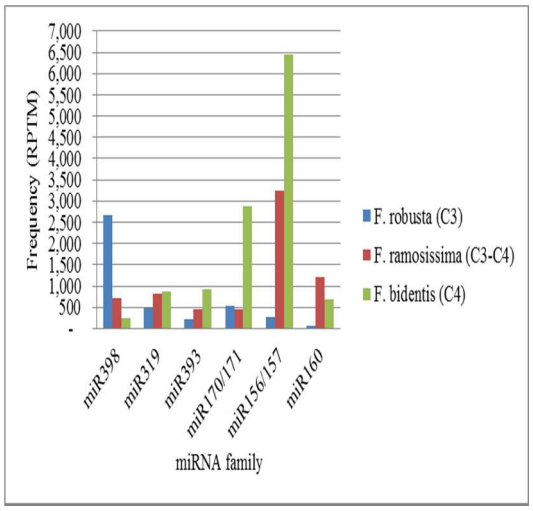
B)


90,000
80,000
70,000
60,000
50,000
40,000
30,000
20,000
10,000
-
Frequency (RPTM)
miR396
miRNA family
F. robusta (C3)
F. ramosissima (C3-C4)
F. bidentis (C4)

C)


30,000
28,000
26,000
24,000
22,000
20,000
18,000
16,000
14,000
12,000
10,000
8,000
6,000
4,000
2,000
-
Frequency (RPTM)
miR159 miR167 miR168 miR169 miR403
miRNA family
F. robusta (C3)
F. ramosissima(C3-C4)
F. bidentis (C4)

D)


7,000
6,500
6,000
5,500
5,000
4,500
4,000
3,500
3,000
2,500
2,000
1,500
1,000
500
-
Frequency (RPTM)
miR398
miR319
miR393
miR170/171
miR156/157
miR160
miRNA family
F. robusta (C3)
F. ramosissima (C3-C4)
F. bidentis (C4)

E)

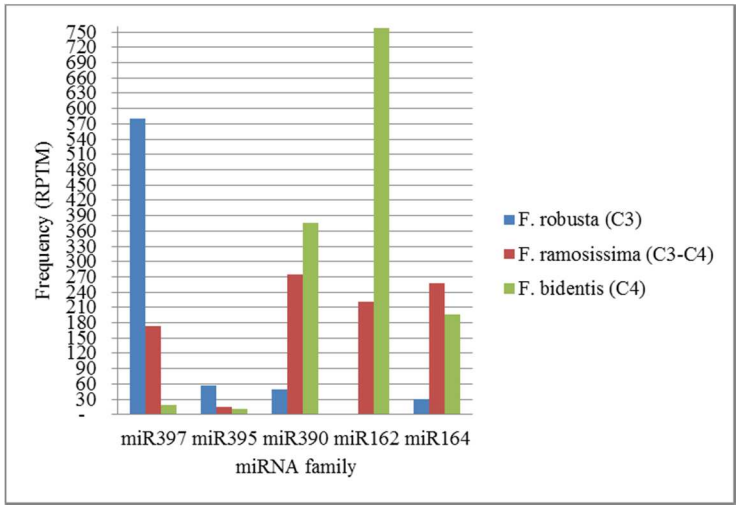


Figure 4.2 (A-E): The normalized frequencies of conserved miRNA families in the leaves of *F. robusta* (C3), *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4).

Table 4.2: A summary of miRNA* families and their normalized abundances in small RNA libraries from the leaves.

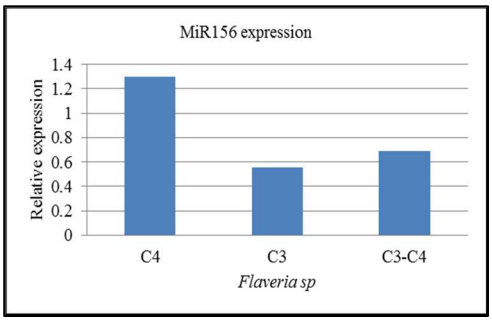
miRNA *	*F. robusta* (C3)	*F. ramosissima* (C3-C4)	*F. bidentis* (C4)
miR170	33	19	52
miR393	2	10	9
miR397	14	-	-
miR408	2	2	-
miR530	2	1	10
miR403	1,688	6,229	7,117

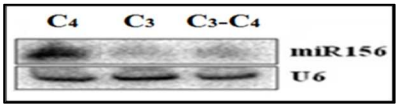
4.3: Expression analyses of conserved miRNAs in leaves of *Flaveria* spp. using small RNA blot analyses.

The sequencing and computational analysis of small RNA libraries revealed the identity of various miRNA families that are expressed in three *Flaveria* spp. The analyses also revealed the differences in abundances for various miRNA families between three *Flaveria* spp. In order to validate those differences independently, small RNA blot analyses was used. U6, a small non-coding nuclear RNA (snRNA) was used as a loading control to ensure uniformity in quantity of total RNAs that were resolved on the gel.

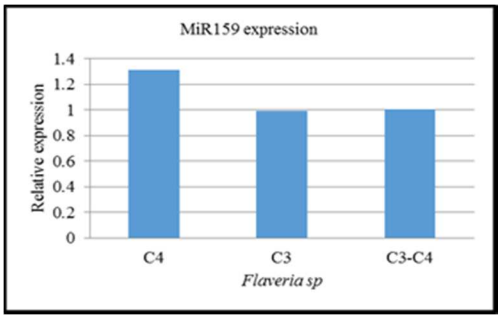
Small RNA blot analyses revealed that several of the most conserved miRNA families were differentially expressed between *F. robusta* (C_3), *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4) as observed in sequencing-based profiles. Based on their signal intensities, miR156, miR159, miR168, miR393 and miR395 were found to be more abundantly expressed in leaves of *F. bidentis* (C_4) than in *F. ramosissima* (C_3-C_4 intermediate) and *F. robusta* (C_3) (Fig.4.3). Some conserved miRNAs such as, miR162 and miR164 could not be detected in the leaves of all three *Flaveria* spp., which could be due to their extremely low abundances. Indeed, the sequencing-based profiles also suggest that these miRNAs have been expressed at very low levels. Although most of the differences in expression levels between the sequencing approach and small RNA blot analysis were correlated, there were instances in which the expression levels for miR159, miR160 and miR168 were only showed weak correlations between the two approaches, i.e., normalized frequencies derived in the sequencing approach and signal intensities observed in small RNA blot analysis. This discrepancy could potentially be attributed to biased ligation efficiencies of different miRNAs with the RNA adapters or other factors such as biased sequencing associated with specific miRNAs. In summary, results from the small RNA blot analysis of miRNAs in leaves of all three *Flaveria* spp. suggests an elevated expression of most conserved miRNAs in *F. bidentis* (C_4) than in *F. ramosissima* (C_3-C_4 intermediate) and *F. robusta* (C_3) and this trend was similar to what was observed in the sequencing-based profiles.

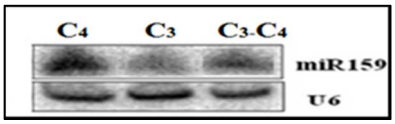
A)


MiR156 expression
Relative expression
1.4
1.2
1
0.8
0.6
0.4
0.2
0
C4
C3
C3-C4
Flaveria sp

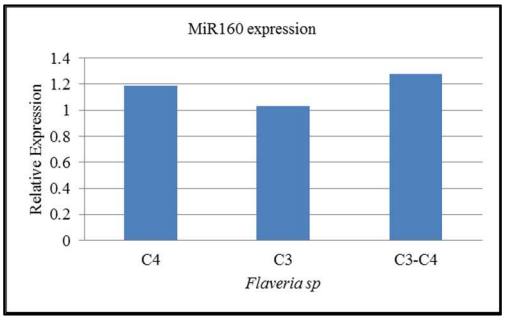

C4
C3
C3-C4
miR156
U6

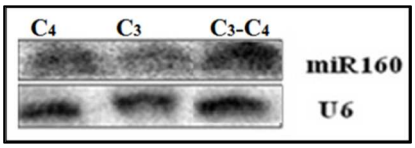
B)


MiR159 expression
Relative expression
1.4
1.2
1
0.8
0.6
0.4
0.2
0
C4
C3
C3-C4
Flaveria sp

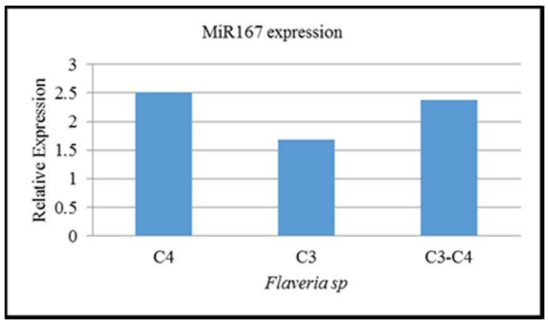

C4
C3
C3-C4
miR159
U6

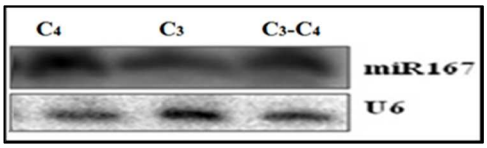
C)


MiR160 expression
Relative Expression
1.4
1.2
1
0.8
0.6
0.4
0.2
0
C4
C3
C3-C4
Flaveria sp

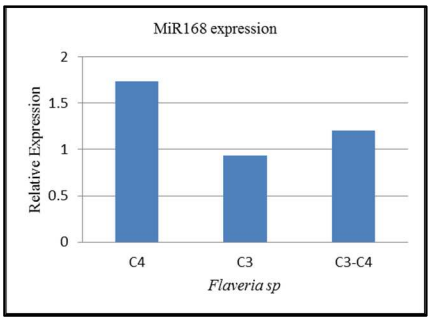

C4
C3
C3-C4
miR160
U6

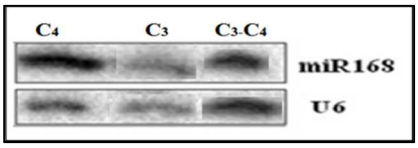
D)


MiR167 expression
Relative Expression
3
2.5
2
1.5
1
0.5
0
C4
C3
C3-C4
Flaveria sp

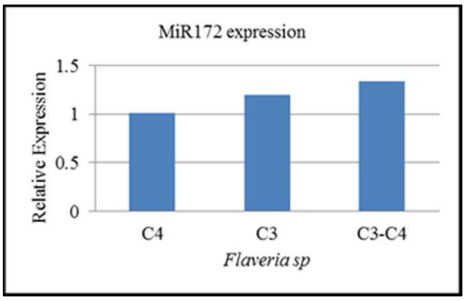

C4
C3
C3-C4
miR167
U6

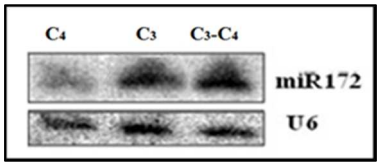
E)


MiR168 expression
Relative Expression
2
1.5
1
0.5
0
C4
C3
C3-C4
Flaveria sp

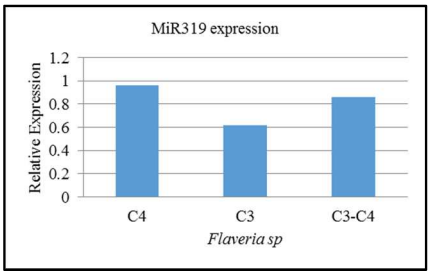

C4
C3
C3-C4
miR168
U6

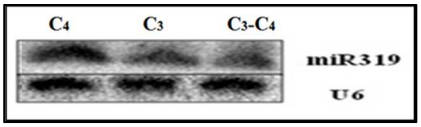
F)


MiR172 expression
Relative Expression
1.5
1
0.5
0
C4
C3
C3-C4
Flaveria sp

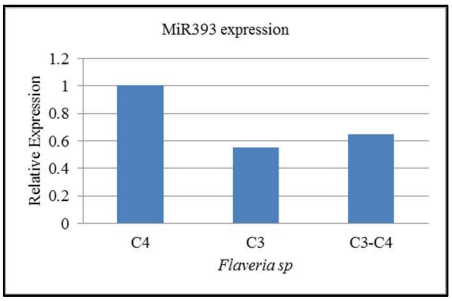

C4
C3
C3-C4
miR172
U6

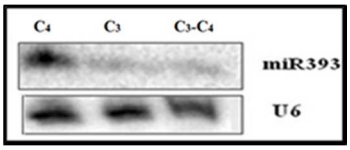
G)


MiR319 expression
Relative Expression
1.2
1
0.8
0.6
0.4
0.2
0
C4
C3
C3-C4
Flaveria sp

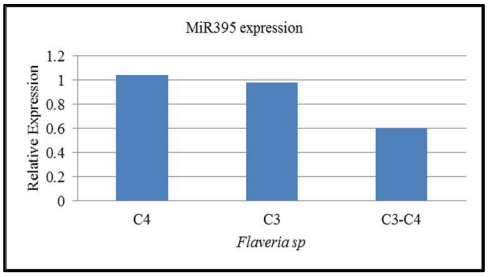

C4
C3
C3-C4
miR319
U6

H)


MiR393 expression
Relative Expression
1.2
1
0.8
0.6
0.4
0.2
0
C4
C3
C3-C4
Flaveria sp


C4
C3
C3-C4
miR393
U6

I)


MiR395 expression
Relative Expression
1.2
1
0.8
0.6
0.4
0.2
0
C4
C3
C3-C4
Flaveria sp

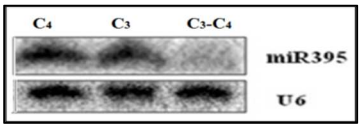

C4
C3
C3-C4
miR395
U6

Figure 4.3 (A-I): Distinct expression patterns of conserved miRNAs in *F. bidentis* (C_4), *F. robusta* (C_3) and *F. ramosissima* (C_3-C_4 intermediate) as determined by small RNA blot analyses and densitometry.

4.4: Analysis of small RNA libraries generated from mesophyll cells of *Flaveria spp*

To gain an insight into the miRNA composition and expression patterns of different miRNA families within the mesophyll and bundle sheath cells of *Flaveria* spp., small RNA libraries were constructed from total RNAs isolated from these cells. The purity of isolated mesophyll cells was assessed by determining Rubisco in the mesophyll cells of *F. robusta* (C_3) and in the bundle sheath cells of *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4), whereas PEPC in the mesophyll cells of *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4).

(I)

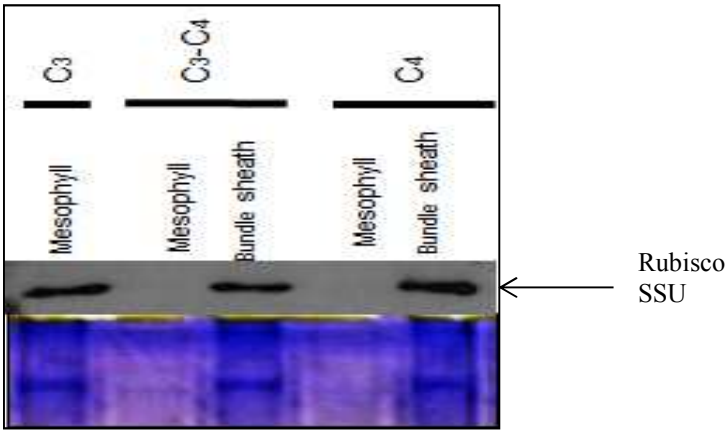


Fig. 4.4(I) Immunodetection of Rubisco small subunit in the mesophyll cells of *F. robusta* (C_3) and in the bundle sheath cells of *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4). The gel stained with Coomassie Blue was shown as loading control.

(II)

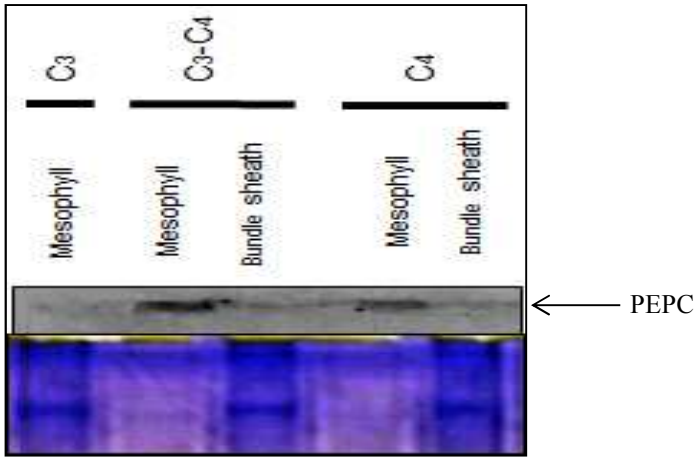


Fig. 4.4 (II) Immunodetection of PEPC in the mesophyll cells *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4). The gel stained with Coomassie Blue was shown as loading control.

Because the bundle sheath cells are not well developed in *F. robusta* (C3), miRNAs were only analyzed in the mesophyll cells in this species whereas in case of the *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4) both mesophyll and bundle sheath cells were analyzed. The read distributions of mesophyll small RNA libraries obtained after high-throughput sequencing are summarized in Table 4.3. The total reads were represented by 27,048,733, 30,692,393 and 36,402,516 from the *F. robusta* (C3), *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4), respectively. After removal of redundant reads, the unique reads were represented by 591,013 in *F. robusta* (C3), 922,852 in *F. ramosissima* (C3-C4 intermediate) and 1,343,281 in *F. bidentis* (C4). Of these, 717 unique reads from *F. robusta* (C3), 708 unique reads from *F. ramosissima* (C3-C4 intermediate) and 800 unique reads from *F. bidentis* (C4) were mapped to the pre-miRBase that has precursor sequences of miRNAs (Table 4.3). Further, these unique reads were matched with the miRBase21 for identification of conserved miRNA homologs. Similar to what was observed in the whole leaf small RNA libraries, only a small portion of the total unique reads could be mapped to the miRBase21.

The small RNAs ranged between 18-28 nucleotides from the mesophyll cells were used for further analysis. The nucleotide lengths of total and unique small RNA reads of mesophyll cell were similar to those observed in the libraries constructed from the whole leaves. Analysis of the total reads of mesophyll cell small RNA libraries of *F. robusta* (C3), revealed one peak at 21 nt size class. By contrast, the total small RNA reads from the mesophyll cells of *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4) revealed two peaks, one each at 22 nt and 26 nt. Small RNAs consisting of 21 nt size which is largely represented by the miRNAs was present in both total and unique reads but they were less abundant compared to the other size classes (Fig. 4.5C-F). Nucleotide length based analysis of unique reads generated from the mesophyll cells of all three *Flaveria* spp. revealed that the 24-nt long small RNAs were more abundant than the 21 nt long small RNAs. Generally, most 21 nt long small RNAs represent miRNAs, tasiRNAs or phasiRNAs, whereas most 24-26 nt long small RNAs represent heterochromatic small interfering RNAs (Lu et al., 2006).

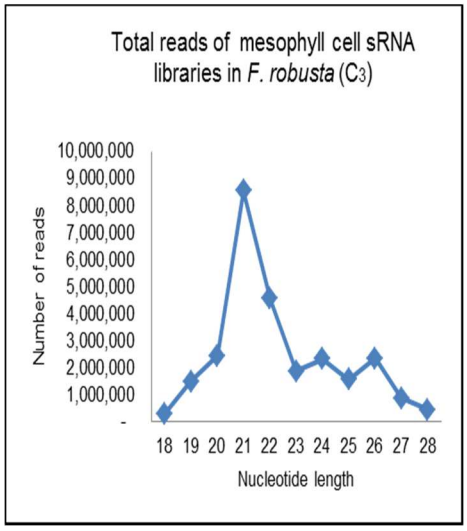
In order to identify homologs of conserved miRNAs in mesophyll cell small RNA libraries of all three *Flaveria* spp., the unique reads were mapped to miRBase21. This resulted in the identification of 32 conserved miRNA families. Of these, 20 miRNA families belong to the highly conserved miRNA families and the remaining were grouped as less conserved miRNA families. Unlike in the small RNA libraries from the entire leaves where qualitative differences were minimal between the three *Flaveria* spp., miRNA composition in the mesophyll cells is different. Of the 20 highly conserved miRNA families, three families (miR160, miR164 and miR390) were not found in mesophyll cells of *F. robusta* (C3). Similarly in case of *F. ramosissima* (C3-C4 intermediate), 16 of the 21 highly conserved miRNA families were detected but not miR160, miR164, miR169, miR390 and miR393 families. In the case of *F. bidentis* (C4), 20 highly conserved miRNA families were identified. Interestingly, miR394 family was not recovered from the mesophyll cells of all three *Flaveria* spp. Within the highly conserved

miRNA families, fourteen families that include miR156, miR159, miR162, miR165/166, miR167, miR168, miR170/171, miR319, miR395, miR396, miR397, miR398, miR399 and miR403 were commonly identified in the mesophyll cells of all three *Flaveria* spp. (Fig.4.6). Moreover, these fourteen miRNA families that were commonly detected in the mesophyll cells of three *Flaveria* spp. were also differentially expressed as some of these were more abundantly expressed in C3 while some others were more abundant in C4 *Flaveria* spp. For instance, miR156, miR159, miR166, miR167, miR168, miR319, miR396, miR397 and miR398 family frequencies were highly abundant in the mesophyll cells of *F. robusta* (C3) than in *F. ramosissima* (C3-C4 intermediate) or *F. bidentis* (C4). With the exception of miR170/171 family, which was more abundantly expressed in *F. bidentis* (C4), the abundances of most of these above mentioned miRNA families were relatively higher in the mesophyll cells of *F. ramosissima* (C3-C4 intermediate) than in *F. bidentis* (C4). Surprisingly, unlike what was observed in the small RNA libraries of whole leaves in which most conserved miRNAs were more abundantly expressed in *F. bidentis* (C4), several of the highly conserved miRNA families were expressed at relatively low levels in the mesophyll cells of *F. bidentis* (C4) (Fig4.6). In addition to the identification of conserved miRNAs, miRNA* for 6 miRNA families (miR162*, miR170*, miR393*, miR397*, miR403* and miR408*) were identified in the mesophyll cell small RNA libraries (Table 4.4).

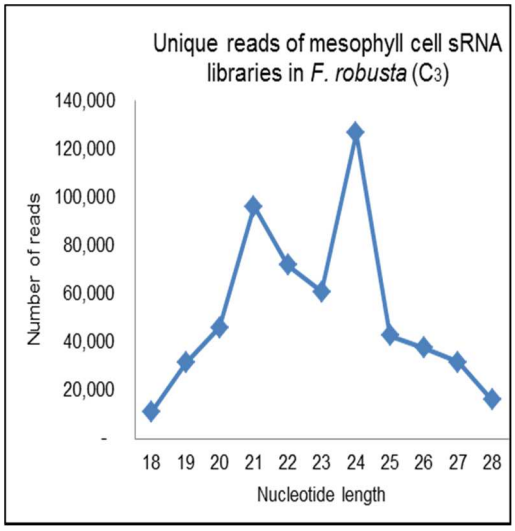
Table 4.3: A summary of total and unique reads distribution of small RNA libraries sequenced from the mesophyll cells of *Flaveria* spp.

Categories	*F. robusta* (C3)		*F. ramosissima* (C3-C4)		*F. bidentis* (C4)	
	reads	unique	reads	unique	reads	unique
Flaveria-mRNAs	24,823,680	279,961	28,929,872	586,897	35,293,910	1,004,559
ncRNAs	18,128,338	58,013	20,601,038	87,924	25,397,956	119,305
premiRbase21	4,518,324	717	709,823	708	523,949	800
repeats	6,366,347	29,770	5,356,715	41,329	11,043,014	58,472
total	27,048,733	591,013	30,692,393	922,852	36,402,516	1,343,281

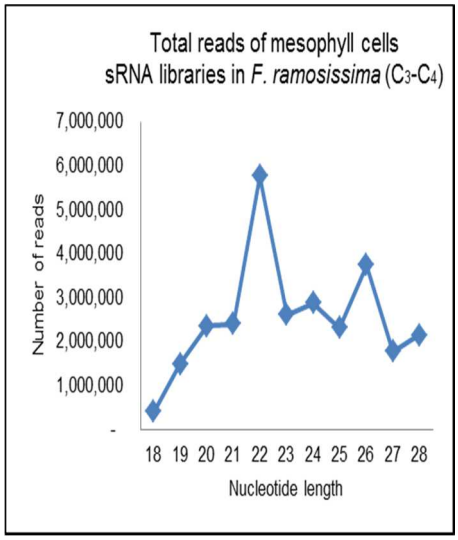
A)


Total reads of mesophyll cell sRNA libraries in F. robusta (C3)
Number of reads
10,000,000
9,000,000
8,000,000
7,000,000
6,000,000
5,000,000
4,000,000
3,000,000
2,000,000
1,000,000
-
18 19 20 21 22 23 24 25 26 27 28
Nucleotide length

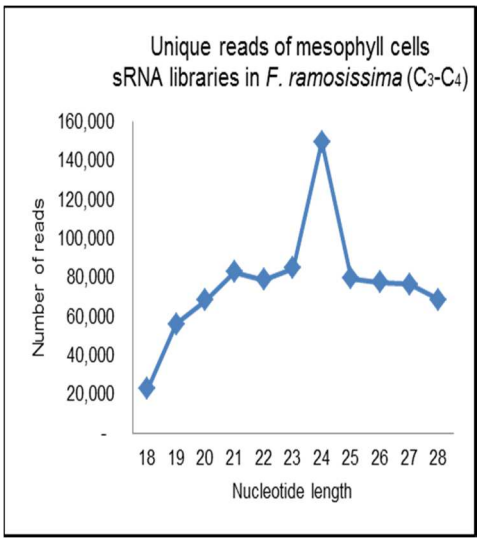
B)


Unique reads of mesophyll cell sRNA libraries in F. robusta (C3)
Number of reads
140,000
120,000
100,000
80,000
60,000
40,000
20,000
-
18 19 20 21 22 23 24 25 26 27 28
Nucleotide length

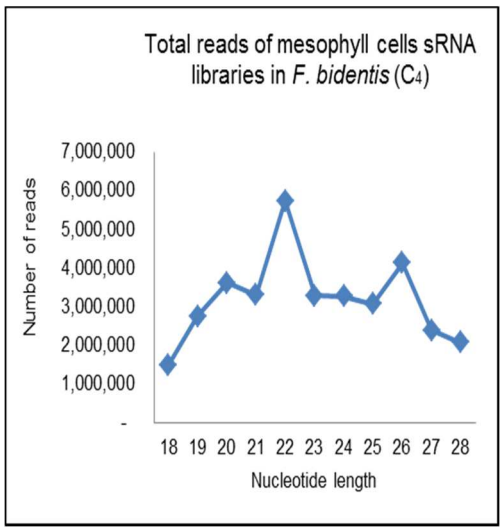
C)


Total reads of mesophyll cells sRNA libraries in F. ramosissima (C3-C4)
Number of reads
7,000,000
6,000,000
5,000,000
4,000,000
3,000,000
2,000,000
1,000,000
-
18 19 20 21 22 23 24 25 26 27 28
Nucleotide length

D)


Unique reads of mesophyll cells sRNA libraries in F. ramosissima (C3-C4)
Number of reads
160,000
140,000
120,000
100,000
80,000
60,000
40,000
20,000
-
18 19 20 21 22 23 24 25 26 27 28
Nucleotide length

E)



F)

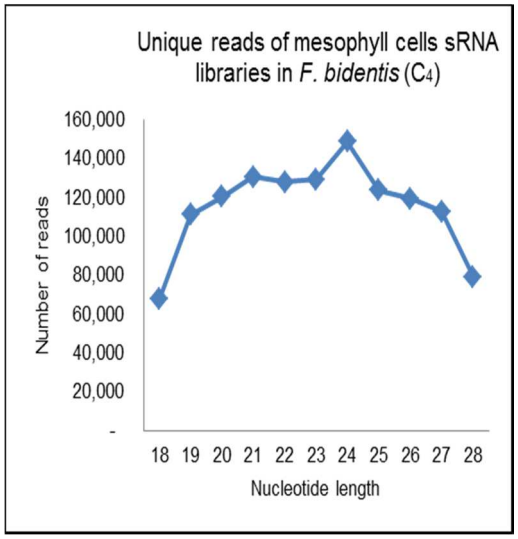
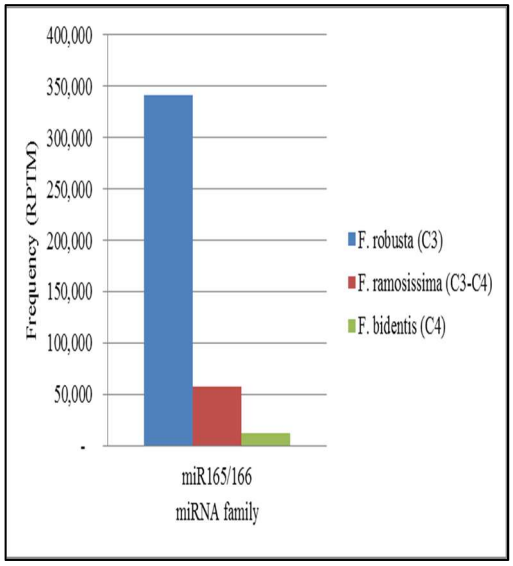
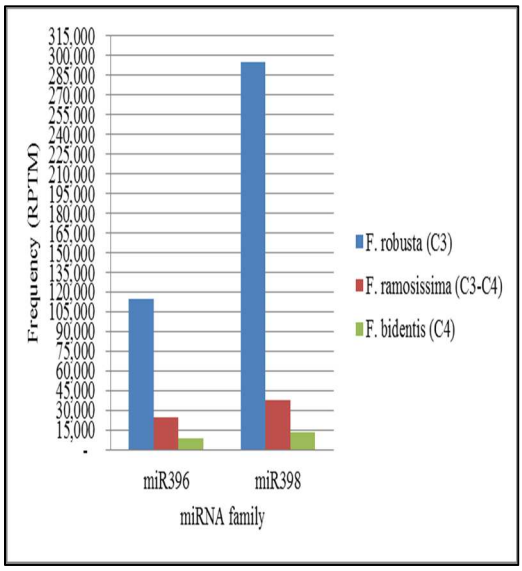


Figure 4.5 (A-F): Small RNA read length vs their abundances in the small RNA libraries from the mesophyll cells of *F. robusta* (C3), *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4).

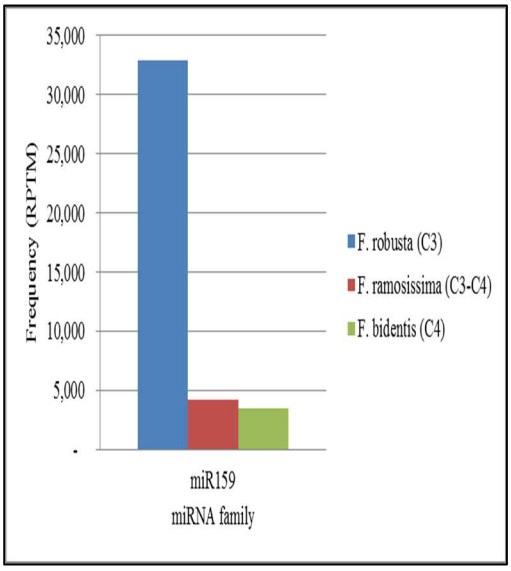
A)



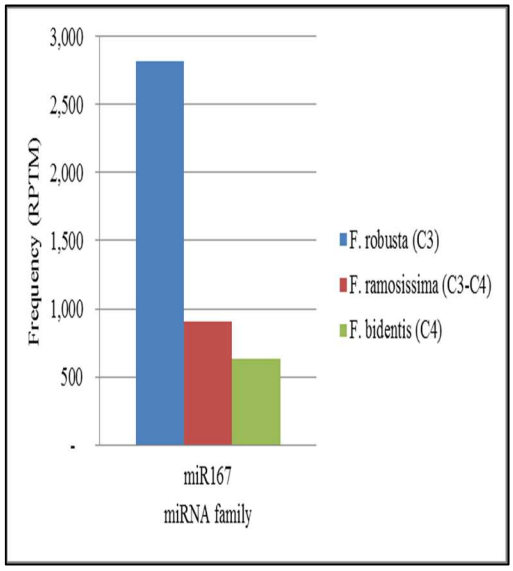
B)



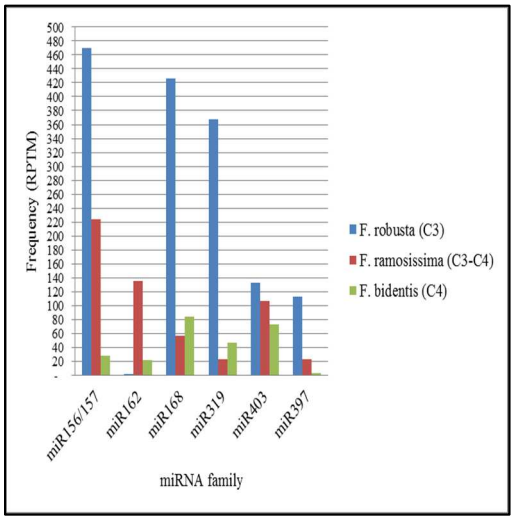
C)



D)



E)



F)

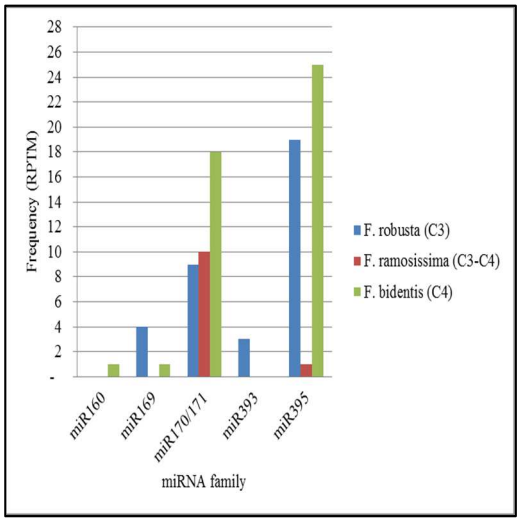


Figure 4.6 (A-F): Conserved miRNA families' distribution in mesophyll cell sRNA libraries of *F. robusta* (C3), *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4).

Table 4.4: Summary of normalized abundances of miRNA* families identified in the mesophyll cell small RNA libraries

miRNA *	*F. robusta* (C3)	*F. ramosissima* (C3-C4)	*F. bidentis* (C4)
miR162	-	9	3
miR170	3	-	-
miR393	1	-	-
miR397	11	2	-
miR408	6	-	-
miR403	133	107	73

4.5: Analysis of small RNA libraries generated from the bundle Sheath cells

In order to obtain insight into the miRNA composition within the bundle sheath cells of *Flaveria* spp., small RNA libraries were constructed from total RNAs isolated from these cells of *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4). These cells were not well-developed in *F. robusta* (C3), thus were not analyzed. The purity of isolated bundle sheath cells was assessed by determining Rubisco in the bundle sheath cells of *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4).

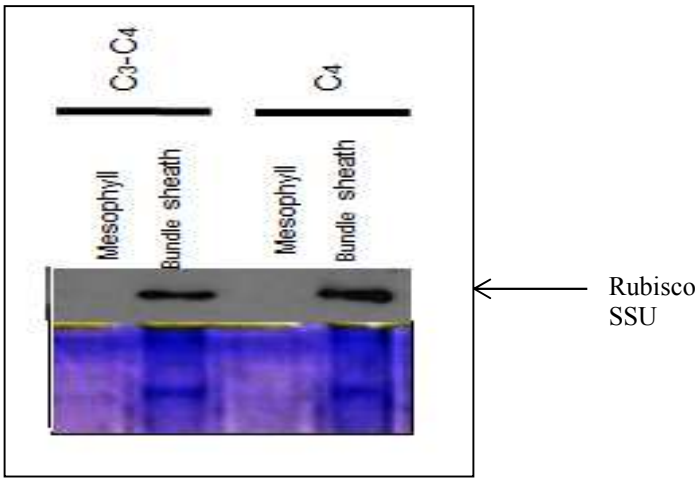


Fig. 4.7 Immunodetection of Rubisco in the bundle sheath cells of *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4). The gel stained with Coomassie Blue was shown as loading control.

The total and unique small RNA reads ranging in size 18 to 28 nucleotides were used for the analyses of small RNA population in the bundle sheath cells of *Flaveria* spp. Interestingly, in the case of both *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4), the sizes of total small RNA reads and their abundances revealed a highly conspicuous peak for 21 nt size class but only a minor peak for the 24 nt size class (Fig. 4.8). This feature suggests that the small RNA reads consisting of 21 nt were the most abundant in the total reads identified from the bundle sheath cells. In contrast, 24 nt size class was the most abundant and 21 nt size class were represented by a minor peak in the case of unique reads obtained from the bundle sheath cells. This suggests that only a small fraction may represent miRNAs in the small RNA population from the bundle sheath cells of *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4).

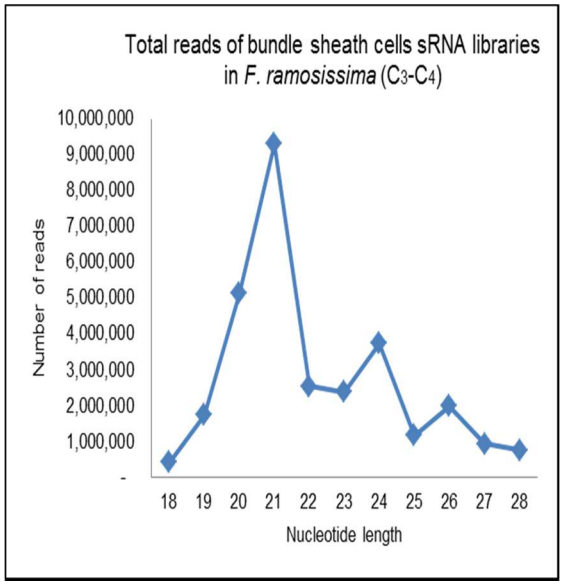
The read distribution of small RNAs to different RNA categories in the bundle sheath cells of two *Flaveria* spp. are summarized in Table 4.5. Similar to what was observed in the small RNAs from the leaves and mesophyll cells of all three *Flaveria* spp., only a small fraction of the entire unique reads could be mapped to the miRBase21. For example, out of the total unique reads of 2,764,943 in *F. ramosissima* (C_3-C_4 intermediate) and 762,544 in *F. bidentis* (C_4), only 1,461 and 829 unique reads, respectively, were mapped to miRBase21. The mapping of unique small RNA reads to the miRBase21 resulted in the identification of 34 known miRNA families. Of these, 22 families belong to highly conserved miRNA families and the remaining families belong to less conserved miRNA families. In the small RNA libraries of *F. ramosissima* (C_3-C_4 intermediate), 31 known miRNAs were identified and this include all 21 highly conserved miRNA families. In the case of *F. bidentis* (C_4), with the exception of miR160, 390 and miR408, the remaining 18 highly conserved miRNA families were identified. miR394 family was not recovered from the bundle sheath cells of either *F. ramosissima* (C_3-C_4 intermediate) or *F. bidentis* (C_4).

Within the highly conserved miRNA families, the normalized frequencies between different miRNA families greatly varied in the bundle sheath cells. For instance, miR166, miR159, miR396, miR398 and miR167 abundances were very high (above 47,708 RPTM) in bundle sheath cells of both *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4). On the other hand, several miRNA families (miR160, miR169, and miR390) are expressed at relatively low levels in bundle sheath cells of *F. ramosissima* (C3-C4 intermediate) or *F. bidentis* (C4) (Fig 4.9). Interestingly, miR160, miR164 and miR390 families were detected only in the bundle sheath cells of *F. ramosissima* (C3-C4 intermediate) but not in *F. bidentis* (C4). Most importantly, miRNA families (miR159, miR162, miR168, miR170/171, miR319, miR393, miR397 and miR403) were more abundantly expressed in the bundle sheath cells of *F. ramosissima* (C3-C4 intermediate) than in *F. bidentis* (C4). On the other hand, miR395 levels were more abundant in the bundle sheath cells of *F. bidentis* (C4) than in *F. ramosissima* (C3-C4 intermediate). Overall, the conserved miRNA families were more abundantly expressed than the less conserved miRNA families in the bundle sheath cell small RNA libraries. This is very similar to what was observed in the case of mesophyll cell and whole leaf small RNA libraries. In addition, miRNA* for 6 known miRNA families were identified (Table 4.6). These miRNA* were highly expressed in bundle sheath cells of *F. ramosissima* (C3-C4 intermediate) than in *F. bidentis* (C4).

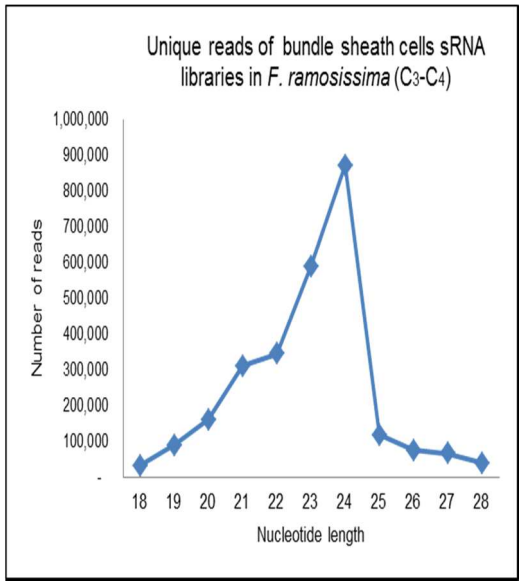
Table 4.5: A summary of total and unique reads distribution of small RNA libraries generated from the bundle sheath cells of *Flaveria* spp.

Categories	*F. ramosissima* (C3-C4 intermediate)-bundle sheath		*F. bidentis* (C4)-bundle sheath	
	reads	unique	reads	unique
Flaveria-mRNAs	22,790,661	584,040	28,384,655	192,921
ncRNAs	15,213,919	101,022	25,746,480	81,362
premiRbase21	5,555,240	1,461	11,798,661	829
repeats	6,969,173	35,488	10,913,000	40,769
total	30,553,548	2,764,943	32,936,658	762,544

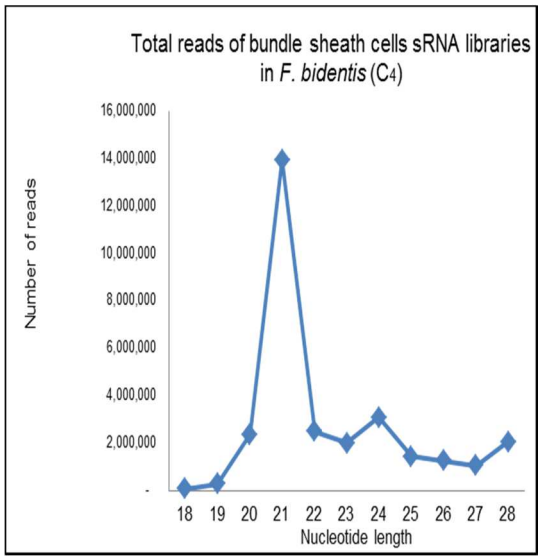
A)



B)



C)



D)

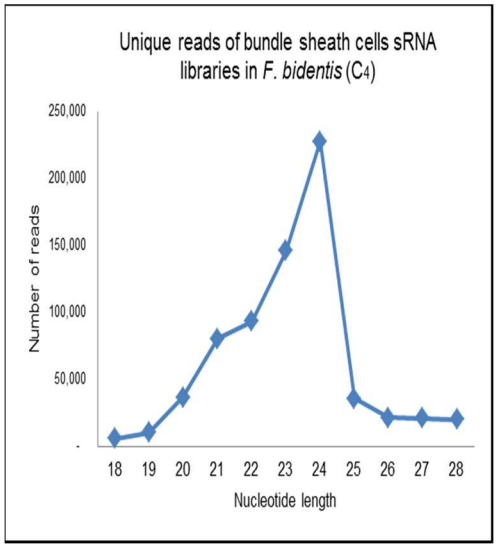
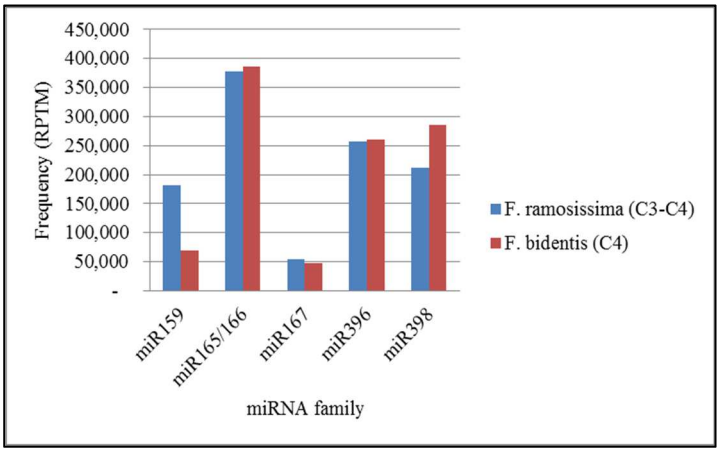
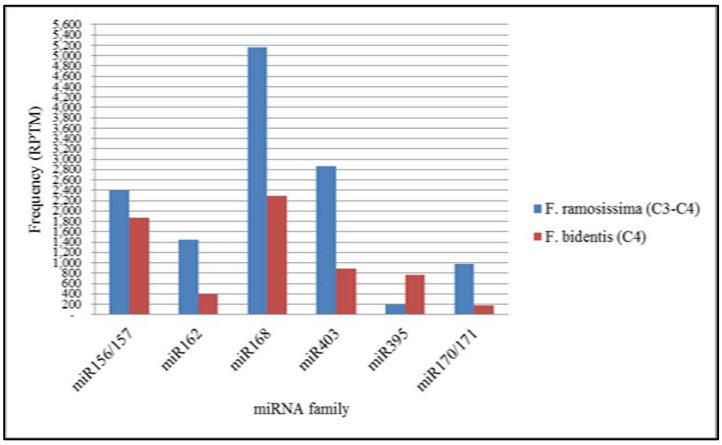


Figure 4.8 (A-D): Small RNA read length vs their abundances in the small RNA libraries from the bundle sheath cells of *F. robusta* (C3), *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4).

A)



B)



C)

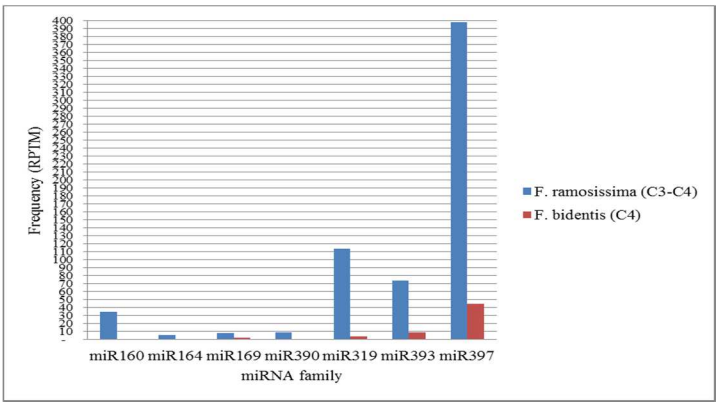


Figure 4.9 (A-C): Conserved miRNA families and their normalized frequencies in the bundle sheath cells of *F. robusta* (C3), *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4).

Table 4.6: A summary of normalized abundances of miRNA* families in the bundle sheath cells.

miRNA *	*F. ramosissima* (C_3-C_4)	*F. bidentis* (C_4)
miR162	58	8
miR170	-	7
miR393	7	-
miR397	4	-
miR408	3	-
miR403	2,865	882

Table 4.7: A summary of individual miRNAs within a family and their normalized abundances in leaves, mesophyll cells and bundle sheath cells of three *Flaveria* spp.

	C_3		C_3-C_4 Intermediate			C_4		
miRNA	*F.robusta*-leaves	*F.robusta*-mesophyll	*F.ramosissima*-leaves	*F.ramosissima*-mesophyll	*F.ramosissima*-bundle sheath	*F. bidentis*-leaves	*F. bidentis* -mesophyll	*F.bidentis*-bundle sheath
miR156a	34	65	603	42	912	1,218	7	549
miR156b	11	2	32	-	21	1,264	1	73
miR156c	111	111	1,573	64	1,006	2,599	9	696
miR156d	34	65	601	42	909	1,214	7	548
miR156e	43	18	383	21	204	549	3	228
miR156f	-	1	18	2	111	51	-	11
miR156g	34	78	603	43	945	1,220	7	553
miR156h	34	78	603	43	945	1,220	7	553
miR156i	108	338	1,231	137	1,049	1,996	15	864
miR157a	11	2	32	-	21	1,264	1	73
miR159a	5,429	32,475	15,153	4,088	169,649	24,640	3,258	68,439
miR159b	2,594	15,065	6,782	2,149	120,542	15,773	1,887	48,526
miR159c	6	52	16	6	1,043	81	7	140
miR159d	3,115	16,887	7,529	2,228	121,855	16,329	1,941	49,155
miR159e	2,634	15,361	6,791	2,155	120,599	15,796	1,894	48,608
miR159f	91	350	162	140	11,113	457	180	1,421
miR159g	25	110	96	17	1,660	227	27	313
miR160a	53	-	1,203	-	35	691	1	-
miR160b	56	-	1,206	-	35	696	1	-
miR160c	58	-	1,211	-	35	698	1	-
miR162a	-	-	39	9	58	157	3	8

miR162b	-	2	183	126	1,389	601	18	390
miR164a	31	-	257	-	6	196	-	-
miR165a	307	162	179	27	336	117	11	172
miR166a	887,122	334,320	421,139	55,125	351,774	304,746	10,878	353,947
miR166b	886,559	334,128	420,947	55,096	351,660	304,570	10,875	353,811
miR166c	2,026	771	1,687	326	12,889	3,695	132	4,248
miR166d	1,159	158	875	95	911	2,306	67	190
miR166e	2,673	886	2,250	375	13,294	4,074	147	4,564
miR166f	1,240	417	1,012	276	2,324	2,624	147	6,293
miR166d	158	23	275	5	189	278	2	45
miR166h	5,941	2,609	2,489	503	17,125	5,182	218	6,401
miR166i	142	159	96	23	149	61	4	152
miR166j	100	10	51	5	22	30	1	15
miR166k	5,036	5,854	24,795	2,614	12,369	13,439	857	28,271
miR166l	46	1	79	-	38	176	1	13
miR166m	10,797	12,860	5,442	1,731	26,183	22,065	1,905	61,732
miR166n	1,991	769	1,666	325	12,883	3,684	131	4,246
miR166o	45	626	104	111	804	263	122	1,773
miR166p	273	199	192	32	224	104	6	188
miR166q	2,025	771	1,687	326	12,888	3,695	132	4,248
miR166r	444	3	43	-	27	35	1	5
miR166s	1,236	417	1,007	276	2,323	2,622	147	6,292
miR166t	2,672	886	2,246	374	13,285	4,071	147	4,558
miR166u	417	334	302	56	345	404	9	349
miR166v	7,167	1,316	8,213	372	3,083	9,253	345	10,571
miR166w	8,487	11,813	4,440	1,537	23,294	18,344	1,661	57,506
miR167a	5,560	2,798	14,027	896	49,887	20,627	616	46,709
miR167b	124	26	524	13	4,996	1,074	12	999
miR167c	123	26	518	13	4,992	1,062	12	996
miR167d	163	381	337	307	37,390	926	118	35,547
miR167e	88	26	264	13	4,920	775	12	967
miR167f	132	389	380	308	37,427	968	121	35,560
miR167g	99	381	285	307	37,388	831	118	35,546
miR167h	87	24	257	12	4,919	767	12	962
miR167i	90	26	293	12	4,938	799	12	968
miR167j	90	26	294	12	4,938	801	12	968
miR168a	178	79	140	10	2,266	1,015	28	967
miR168b	1,695	342	2,002	46	2,886	11,695	55	1,320

miR168c	185	84	146	11	2,276	1,063	29	968
miR168d	136	54	97	6	252	292	3	91
miR168e	-	6	5	3	157	15	15	21
miR168f	11	-	24	-	99	82	-	5
miR169a	13,887	4	7,701	-	8	27,224	1	-
miR169b	13,886	4	7,696	-	7	27,222	1	-
miR169c	13,892	4	7,701	-	8	27,226	1	-
miR169d	169	-	98	-	-	394	-	-
miR169e	13,886	4	7,702	-	7	27,223	1	-
miR169f	13,894	4	7,718	-	8	27,249	1	2
miR169g	13,894	4	7,718	-	8	27,249	1	2
miR169h	169	-	98	-	-	394	-	-
miR171a	-	1	6	6	879	19	1	22
miR171b	533	1	440	1	2	883	-	12
miR171c	2	-	1	1	19	444	1	31
miR171d	50	1	175	-	1	183	-	10
miR171e	-	-	-	1	19	377	1	27
miR171f	2	7	5	2	85	1,522	16	114
miR172a	231	2	232	-	19	194	-	4
miR172b	232	2	234	-	20	203	-	4
miR319a	471	368	821	23	114	876	47	4
miR319b	143	268	419	17	83	493	41	3
miR319c	366	336	1,015	23	213	838	49	27
miR319d	146	268	421	17	83	493	41	3
miR3630	1	45	-	50	56	1	12	19
miR390a	49	-	274	-	9	376	-	-
miR393b	95	-	253	-	22	466	-	2
miR393c	38	-	121	-	28	311	-	5
miR393d	89	3	86	-	24	146	-	2
miR395a	-	-	-	-	1	14	3	239
miR395b	249	13	197	6	1,176	783	7	500
miR395c	249	13	197	6	1,184	783	7	500
miR395d	-	-	-	-	6	33	7	321
miR395e	56	19	15	1	198	12	25	765
miR396a	90	361	101	154	1,812	110	68	689
miR396b	4,339	1,098	431	79	3,218	545	11	274
miR396c	198	157	37	20	1,785	93	6	176
miR396d	539	28	249	3	264	316	4	43

miR396e	4	6	5	1	205	19	2	12
miR396f	17	17	4	2	44	8	-	7
miR396g	17	17	4	2	45	7	-	7
miR396h	2	15	9	3	22	103	2	27
miR396i	5,348	5,825	1,689	524	5,014	3,240	96	3,422
miR396j	192	156	36	20	1,784	92	6	176
miR396k	4	6	5	1	205	19	2	12
miR396l	4	27	9	9	147	14	2	43
miR396m	4,349	1,102	435	79	3,221	549	11	274
miR396n	77,991	114,689	47,898	24,486	255,357	59,562	8,334	260,492
miR396o	501	5,164	740	3,719	83,816	1,566	702	46,598
miR396p	55	37	9	3	75	12	-	4
miR397a	581	113	174	23	398	19	3	45
miR397b	573	113	157	22	390	19	3	45
miR398a	7	8	7	-	137	5	-	8
miR398b	7	8	7	-	137	5	-	8
miR398c	126	34	287	3	253	34	1	49
miR398d	2	31	-	6	73	-	8	57
miR398e	465	9	-	-	1	-	-	1
miR398f	1,750	293,307	416	37,897	210,126	196	13,643	284,868
miR398g	796	1,820	8	74	537	4	36	562
miR398h	7	8	7	-	137	5	-	9
miR398i	7	8	7	-	137	5	-	8
miR398j	30	469	6	66	485	3	34	530
miR403	1,688	133	6,229	107	2,865	7,117	73	882
miR6111	338	15	354	8	422	1,081	9	169
miR6114	357	6	6	-	1	156	-	48
miR6478	162	-	211	2	16	601	3	9
miR8175	24	-	54	-	9	521	-	-
miR858a	27	18	7	4	131	48	17	49
miR858b	37	26	8	6	187	70	20	56
miR858c	222	1,173	248	183	8,262	599	568	2,429

4.6: Degradome Library Construction and Sequencing Analysis

In higher plants, most miRNAs regulate the expression of their target genes by cleavage of mRNAs within the complementary region to the miRNAs (Llave et al., 2002). When miRNAs guides cleavage on their target mRNAs, 5' and 3' cleaved mRNA fragments will be released. RNA adapter was ligated to the 5'end of the 3' cleaved mRNA fragments (targets of miRNAs), amplified using few PCR cycles, digested and then DNA adapters were ligated and further amplified and the resultant PCR product was sequenced, a procedure known as 'degradome' or 'PARE (Parallel Analysis of RNA Ends)' analysis (Addo-Quaye et al., 2008; German et al., 2008). In order to identify potential targets for the miRNAs in the leaves of *F. robusta* (C_3), *F. ramosissima* (C_3-C_4 intermediate), and *F. bidentis* (C_4), degradome libraries were constructed, sequenced and the datasets were analyzed using the SeqTar pipeline (Zheng et al., 2012). Overall, 21,007,764, 18,407,012 and 20,519,940 reads from *F. robusta* (C_3), *F. ramosissima* (C_3-C_4 intermediate), and *F. bidentis* (C_4) respectively were sequenced from the degradome libraries. Upon further processing of the total reads, a total of 862,199, 1,106,857 and 1,173,019 unique reads were obtained for *F. robusta*, *F. ramosissima* and *F. bidentis,* respectively (Table 5.4). In all three degradome libraries, reads that were mapped to the non-coding RNAs and repeat sequences were highly represented. Only 851,256, 480,130 and 1,001,359 total reads represented by 239,564, 249,713 and 477,500 unique reads were mapped to the mRNAs from the *Flaveria* spp.

Table 4.8: Summary of reads distribution of degradome libraries constructed from the leaves of three *Flaveria* spp.

Category	*F. robusta* (C3)		*F. ramosissima* (C3-C4)			*F.bidentis* (C4)
	reads	unique	reads	unique	reads	unique
Flaveria-mRNAs	851,256	239,564	480,130	249,713	1,001,359	477,500
ncRNAs	527,466	17,894	183,063	16,681	293,935	18,831
premirbase21	16,540	258	29,466	582	36,003	719
repeats	424,089	10,118	91,076	9,266	198,523	10,426
Total reads	21,007,764	862,199	18,407,012	1,106,857	20,519,940	1,173,019

4.7: Identification of Targets of conserved miRNAs in *Flaveria* spp.

Until date, the CleaveLand pipeline has been the largely used computational pipeline for identifying plant miRNA targets present in degradome libraries (Addo-Quaye et al., 2009 and Zheng et al., 2012). However, the steps followed in CleaveLand method have been suggested to be too rigid and less flexible. This could result in omission of some potential miRNA targets (Zheng et al., 2012). On the other hand, a recently developed SeqTar algorithm is a more relaxed and effective pipeline for identifying cleaved miRNA targets with more mismatches from raw degradome data in plants. Therefore, SeqTar pipeline was used in identifying miRNA: mRNA pairs using the unique reads obtained from the degradome reads of each *Flaveria* spp. The analysis revealed only a small number of transcripts as targets for miRNAs in each of the *Flaveria* spp. investigated. In *F. robusta* (C3), 4 transcripts were identified as putative targets. These targets encode proteins namely REVOLUTA, member of homeodomain-leucine zipper family (HD-ZIP family), Target of early activation tagged (EAT) and Molecular chaperone DnaJ (Table 4.9). In *F. ramosissima* (C3-C4 intermediate), 3 transcripts were identified as targets (Protein ABIL4, hypothetical protein) and Choline/ethanolamine phosphotransferase 1) for miR159, miR160 and miR168 (Table 5.6). Lastly, in *F. bidentis* (C4), three transcripts, i.e., MYB domain transcription factor and Protein tyrosine phosphatase were identified as targets for miR398 and miR858 (Table 5.1).

Table 4.9: Identification of targets of conserved miRNAs in *F. robusta* (C3)

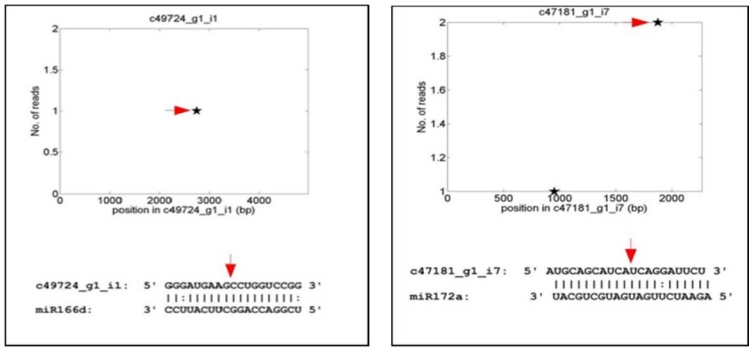
miRNAs	Total Mismatches	Total Reads	Percentage Reads	Reference Full Name	Encoded protein
miR166d	1	4	25	c49724	REVOLUTA, member of homeodomain-leucine zipper family
miR172a	0.5	3	66.67	c47181	Target of early activation tagged (EAT)
miR319a	2.5	2	50	c46580	Molecular chaperone DnaJ
miR319a	4	198	2.02	c32054	Unknown protein

Table 5.0: Identification of targets of conserved miRNAs in *F. ramosissima* (C3-C4 intermediate)

miRNAs	Total Mismatches	Total Reads	Percentage	Reference Full Name	Encoded protein protein
miR159a	4	19	10.53	TR4898	Protein ABIL4
miR160a	1	9	11.11	TR17740	hypothetical protein
miR168f	4	184	0.54	TR20673	Choline/ethanolamine phosphotransferase 1

Table 5.1: Identification of targets of conserved miRNAs in *F. bidentis* (C4)

miRNAs	Total_ Mismatches	Total Reads	Percentage	Reference Full Name	Encoded Protein
miR398c	2	9	11.11	c23379	Hypothetical protein
miR858b	3	6	16.67	c2950	MYB domain protein 111
miR858c	4	217	9.22	c9922	Protein tyrosine phosphatase



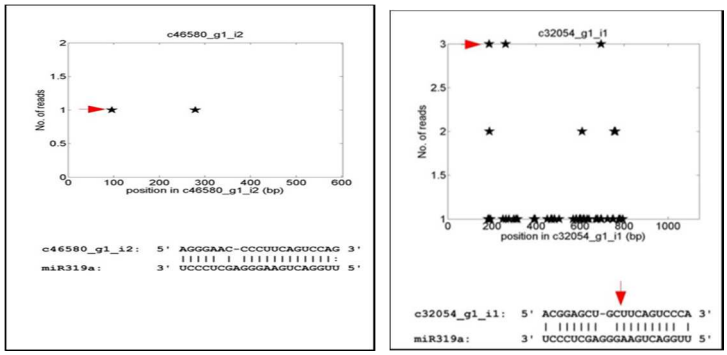


Figure 4.10: T-plots showing cleavage abundance of predicted targets of conserved miRNAs in *F. robusta* (C_3).

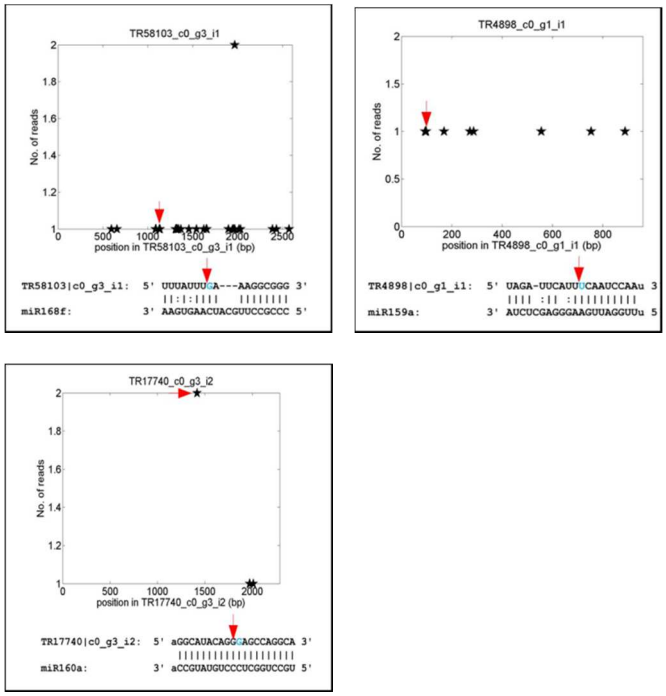


Figure 4.11: T-plots showing cleavage abundance of predicted targets of conserved miRNAs in *F. ramosissima* (C_3-C_4 intermediate)

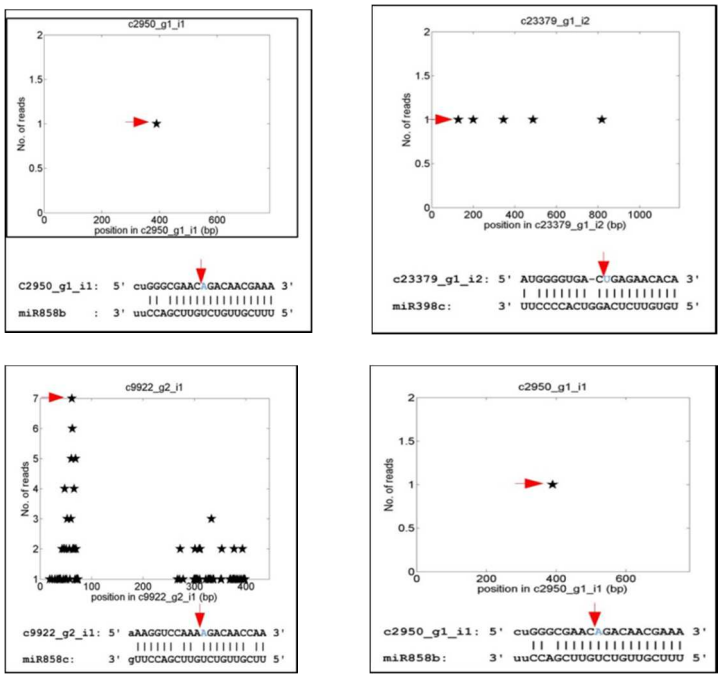


Figure 4.12: T-plots showing cleavage abundance of predicted targets of conserved miRNAs in *F. bidentis* (C_4).

CHAPTER V

DISCUSSION

Currently, about 3% of the world's total terrestrial plant species use the C_4 photosynthetic pathway for carbon fixation whereas, 95% of them use the C_3 pathway for carbon assimilation (Chang et al., 2012). More C_3 plants than C_4, represent most of the world's staple crops such as *Oryza sativa*, *Triticum aestivum* and *Glycine max* (Hibberd et al., 2008). Due to the inefficiency in photosynthesis and water use, most C_3 plants may be unable to thrive well under current global weather conditions including rising temperatures (Chang et al., 2012). This problem could at least partly be mitigated by effectively incorporating the desirable C_4 pathway traits and their regulatory networks into C_3 plants. To achieve this daunting task, we need a thorough understanding of gene regulatory networks that contribute for C_4 photosynthesis.

Flaveria has been widely used as a model plant system for investigating the evolution of C_4 photosynthesis from the ancestral C_3 type (Gowik et al., 2011). To obtain molecular insights especially in identifying genes associated with the evolution of C_4 photosynthesis from C_3, recent studies compared the transcriptome profiles of *F. robusta* (C_3), *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4). Such comparative analyses between C_3 and C_4 plants yielded major differences with respect to the abundances of various C_4 photosynthesis enzymes as well as diverse families of transcription factors (Gowik et al., 2011; Wang et al., 2013 and Mallman et al., 2014). However, the role of miRNAs in this important biochemical process is almost unknown. Thus far, the miRNA component has not been investigated in *Flaveria* spp.

In order to understand the roles of miRNAs in C3, C3-C4 intermediate and C4 photosynthesis, miRNA profiles were compared in *Flaveria spp.* that differ in their mode of photosynthesis. Anatomically, the C3 and C4 leaves were distinct represented by mesophyll and bundle sheath cells in C4 and only mesophyll cells in C3 plants. Therefore, the miRNA component in mesophyll and bundle sheath cells isolated from matured *F. robusta* (C3), *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4) were also investigated. In addition to this, three degradome libraries were constructed from the leaves of the above named species in order to identify the targets for the miRNAs in *Flaveria* spp.

5.1: Identification of conserved miRNAs and determining their differential abundances in the leaves of three *Flaveria* spp.

Extensive small RNA sequencing approaches from diverse plant species resulted in identification and deposition of vast number of miRNAs at the miRBase. At least 21 miRNA families have shown a high degree of conservation among land plants suggesting conserved biological functions for these miRNAs in all higher plants (Axtell et al., 2005; An et al., 2011).

The computational analysis of the small RNA reads from the leaves of three *Flaveria* spp. resulted in identification of homologs of 21 conserved miRNA families in *F. robusta* (C3) and *F. ramosissima* (C3-C4 intermediate). In *F. bidentis* (C4) however, 20 conserved miRNAs were identified. miR408, which was expressed in *F. robusta* (C3) and *F. ramosissima* (C3-C4 intermediate), was not recovered from *F. bidentis* (C4). It is worth mentioning that even in *F. robusta* (C3) and *F. ramosissima* (C3-C4 intermediate), miR408 levels were extremely low (3 or 2 RPTM). The non-recovery of miR408 could simply be due to its extreme low abundance in *F. bidentis* (C4).

Among all conserved miRNA families, miR166 family was the most abundantly expressed in all three *Flaveria* spp. investigated. However, miR166 family abundances greatly differed between the three *Flaveria* spp. In leaves of *F. robusta* (C3), its levels were almost twice

that of the *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4). Similar differences were also observed in the mesophyll cells in which miR166 family displayed greater abundances in *F. robusta* (C3) but much lower levels in *F. ramosissima (*C3-C4 intermediate) and *F. bidentis* (C4). The miR166 family abundances were very similar in the bundle sheath cells of *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4). The abundantly expressed miR166 in the leaves of all *Flaveria* spp. suggest an important role for this miRNA family. The miR166 family is known to target HD-ZIP transcription factors that in turn regulate leaf polarity and floral development in plants (Emery et al., 2003, Jung and Park, 2007; Nogueira et al., 2007). Thus far, miR166 has not been linked with photosynthesis either biochemically or anatomically. The present investigation has revealed clear differences with respect to the abundances of miR166 family not only in entire leaves but also in the mesophyll cells of C3 and C4 *Flaveria* spp. This suggests an important role for this miRNA family in either the origination or maintenance of anatomical or biochemical differences associated with the C3 and C4 photosynthesis. It is worth noting that the differences in miR166 abundances in bundle sheath cells of C3-C4 intermediate and C4 *Flaveria* spp., were only marginal.

In the leaves, miR396 was abundantly expressed in *F. robusta* (C3) compared with *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4). Similarly, in the mesophyll cells, miR396 levels were several fold greater in *F. robusta* (C3) compared with *F. ramosissima* (C3-C4 intermediate) and *F. bidentis*. However, its levels in bundle sheath cells did not differ between *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4). miR396 is known to regulate several Growth regulating factors in plants, which is important for leaf growth and development including adaxial-abaxial polarity formation during leaf morphogenesis (Liu et al., 2009; Rodriguez et al., 2010; Wang et al. 2011). Similar to the miR166, miR396 family has not been linked with photosynthesis, but it's differential expression in *Flaveria* spp., differing in their mode of photosynthesis, suggest a potential role in C3 and C4 photosynthesis.

miR159 and miR319 are two highly conserved miRNA families in land plants and have been shown to be involved in regulating leaf morphogenesis, plant growth and reproduction. miR159 represses MYB transcription factor genes, miR319 targets TCP transcriptional factors in plants (Jones-Rhoades et al., 2006). The activities of miR159 and its corresponding target genes have been shown to regulate seed germination, vegetative growth and reproductive organs development (Alonso-Peral et al, 2010; Allen et al., 2007 and Reyes et al., 2007) but miR319 and its target, TCP, regulate leaf morphology and senescence (Palatnaik et al., 2003 and Schommer et al., 2008). Interestingly, in the entire leaves both miR159 and miR319 levels were relatively high in *F. bidentis* (C4) compared with *F. robusta* (C3) *and F. ramosissima* (C3-C4 intermediate). By contrast in mesophyll cells, both miR159 and miR319 levels were several fold greater in *F. robusta* (C3) compared with *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4). Within the bundle sheath cells, miR159 and miR319 levels were greater in *F. ramosissima* (C3-C4 intermediate) compared with *F. bidentis* (C4). Such differential expression of both miR159 and miR319 that share very high sequence homologies suggests a putative role for both of these miRNA families in C3 and C4 photosynthesis.

Auxin signaling plays a central role in regulating diverse aspects of plant growth and development. Intriguingly, several key proteins (TIR1 and ARFs) involved in auxin signaling are regulated by miR160 (ARFs), miR167 (ARFs) and miR393 (TIR1) families as well as miR390-dependent TAS3 siRNAs (ARFs) (Jones-Rhoades et al., 2006). The profiling of miRNAs in leaves of three *Flaveria* spp. revealed distinct differences with respect to the expression levels of these miRNAs that fine tune auxin signaling in plants. For instance, miR167 levels in leaves were much greater in *F. bidentis* (C4) but intermediate in *F. ramosissima* (C3-C4 intermediate) and very low in *F. robusta* (C3). These differences were further confirmed using small RNA blot analysis in these three plant species. Interestingly, miR167 levels in the mesophyll cells were found to be opposite to what was observed in the entire leaves, as its levels were much greater in

F. robusta (C3), but intermediate in *F. ramosissima* (C3-C4 intermediate) and very low in *F. bidentis* (C4). Similarly, miR160 expression levels greatly differed; a very low level in leaves of *F. robusta* (C3), but much higher levels in both *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4). These differences were also confirmed in small RNA blot analysis. Surprisingly miR160 was almost not represented in three small RNA libraries constructed from the mesophyll cells. However, miR160 was recovered from the bundle sheath cells of both *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4) although its abundances were much higher in *F. ramosissima* (C3-C4 intermediate) compared with *F. bidentis* (C4). Also, miR393 differed greatly differed between *Flaveria* spp. miR393 levels that targets TIR1 (an auxin receptor) was found to be relatively high in leaves of *F. bidentis* (C4) and intermediate in *F. ramosissima* (C3-C4 intermediate) and low in *F. robusta* (C3). These differences in miR393 expression levels were also confirmed in small RNA blot analysis. miR393 reads were hardly recovered from the small RNA libraries of mesophyll cells of all three *Flaveria* spp. On the other hand, these were differentially expressed in bundle sheath cells, i.e., more abundant in *F. ramosissima* (C3-C4 intermediate) and extremely low in *F. bidentis* (C4). These results imply that miRNAs that fine tune auxin signaling differ greatly not only in the entire leaves but also in mesophyll and bundle sheath cells of three *Flaveria* spp. that differ in their mode of photosynthesis. These results support the notion that the miRNA-controlled auxin signaling is an important factor that could contribute to the differences in photosynthesis.

Small RNA analyses from the leaves revealed that the miR168 was highly expressed in *F. bidentis* (C4) compared with *F. ramosissima* (C3-C4 intermediate) and *F. robusta* (C3). Small RNA blot analysis also confirmed the greater abundances in *F. bidentis* (C4). By contrast, miR168 levels were high in the mesophyll cells of *F. robusta* (C3) than in *F. ramosissima* (C3-C4 intermediate) and *F. bidentis* (C4). miR168 is involved in regulating the expression of *AGO1*, which is critical for normal growth and development (Vaucheret et al., 2004). AGO1 is an

important component of the RNA-induced Silencing Complex into which mature miRNA is incorporated prior to target mRNA cleavage (Baldrich et al., 2014; Vaucheret, 2008 and Voinnet, 2009). The differential abundances of miR168 in total leaves, mesophyll and bundle sheath cells of *F. robusta* (C_3) than in *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4) suggests a role in photosynthesis perhaps by differentially regulating various miRNA targets in these *Flaveria* spp.

In general, miR395 abundances are known to be extremely low in several plants that have been grown on optimal nutrients. However, miR395 has been shown to be highly inducible during sulfate-deprivations (Jones-Rhoades et al., 2006; Sunkar et al., 2007). miR395 was differentially expressed in three *Flaveria* spp. In leaves, miR395 was found to be more abundant in *F. bidentis* (C_4) than in *F. ramosissima* (C_3-C_4 intermediate) and *F. robusta* (C_3) as determined by small RNA blot analysis. Interestingly, miR395 was also recovered from both the bundle sheath and mesophyll cells. It was more abundantly expressed in bundle sheath cells than in mesophyll cells. miR395 which targets a sulfate transporter and three ATP Sulfurylases, fine tunes sulfate uptake and assimilation in plants. Sulfur assimilation and metabolism which largely occurs in the chloroplast was thought to be associated with C_4 photosynthesis (Kopriva and Kopriva, 2005). Earlier comparative transcriptomic analyses between *Flaveria* spp. that utilizes C_3, C_4, and C_3–C_4 intermediate photosynthesis suggested a potential link between sulfate assimilation and C_4 photosynthesis (Weckopp and Kopriva, 2014). These analyses pointed out that there exists a greater need for reduced sulfur in C_4 photosynthetic plants (Weckopp and Kopriva, 2014). Although the results presented from the miRNA analysis here do not agree with this hypothesis, (a greater need for reduced sulfur in C_4 photosynthetic plants) the greater abundances of miR395 in *F. bidentis* (C_4) suggests it diminishes sulfur reduction rather than enhance it. In addition, the identified differential expression patterns of miR395 appear to have an important role in C_3 and C_4 photosynthesis.

5.2: Identification of targets of conserved miRNAs in *Flaveria* spp.

Following the analysis of the degradome library using SeqTar pipeline, only a small number of target transcripts were identified for some conserved and known miRNAs of *Flaveria* spp. Most of these predicted targets are non-conserved targets for the conserved miRNAs. This observation supports the suggestion that some conserved miRNAs may have non-conserved targets in distinct plant species (German et al., 2008 and Li et al., 2010). Many conserved miRNA targets could not be identified as miRNA targets in *Flaveria* spp., due to lack of a well-assembled *Flaveria* transcriptome.

5.3: Does miRNAs play a role in Photosynthesis?

C_4 photosynthesis is believed to have evolved independently from ancestral C_3 pants through intermediary forms (Sage et al., 2012; Schluter et al., 2016) and the evolution process is postulated to have involved modifications in leaf anatomy and changes in gene expression (Brautigam and Gowik, 2016). In C_4 photosynthetic species especially, the unique arrangement of mesophyll and bundle sheath cells in the leaves have been suggested to be crucial for photosynthesis (Sage and Monson, 1999). To date, the roles of miRNAs in photosynthesis are yet to be convincingly demonstrated. Although our study could not directly establish a link between miRNAs and the different forms of Photosynthesis, the differential expression of several conserved miRNA families, not only in leaves but also in mesophyll and bundle sheath cells, argues that such differential expressions could impact on the target gene expression differently. The identified differences with respect to the miRNAs require further analysis, but they lay a strong foundation to investigate the significance of miRNA-controlled gene regulation not only in C_3 and C_4 photosynthesis but also in the evolution of C_4 from C_3 photosynthesis. Because most conserved miRNAs regulate transcription factors, their differential regulation as determined by miRNAs in a cell-specific manner could differently impinge on the transcription factor controlled down-stream gene expression. This could in turn be important for the anatomical and biochemical

differences that exist between C_3 and C_4 photosynthetic plants and in the process, miRNAs might have played a significant role in the evolution of C_4 photosynthesis from the C_3 type.

CHAPTER VI

CONCLUSION

By constructing small RNA libraries and performing high throughput sequencing of small RNAs from leaves of three *Flaveria* spp., the following known miRNA families were identified. i.e., 39 in *F. robusta* (C_3), 37 in *F. ramosissima* (C_3-C_4 intermediate) and 35 in *F. bidentis* (C_4). From the differential expression patterns, miR156, miR159, miR162, miR167, miR168, miR169, miR170/171 and miR393 showed a more abundant expression in leaves of C_4 than in C_3 *Flaveria* spp. These sequencing-based differences in conserved miRNA profiles were largely confirmed using small RNA blot analysis. The miRNA analysis in the mesophyll cells revealed major differences between the three *Flaveria* spp. Although fourteen miRNA families were commonly detected in the mesophyll cells of the *Flaveria* spp., they were differently expressed. i.e., miR166, miR159, miR398, miR396 and miR319 family frequencies were more abundant in the mesophyll cells of *F. robusta* (C_3) than in *F. bidentis* (C_4). On the other hand, miR160, miR390 and miR393 families were more abundantly expressed in the bundle sheath cells of *F. ramosissima* (C_3-C_4 intermediate) but not in *F. bidentis* (C_4).

To date, no known targets of conserved miRNAs in *Flaveria* spp. have been identified. From the three degradome libraries constructed, only a small number of homologs of conserved miRNA targets were identified. Rather, non-conserved targets of conserved miRNAs were identified.

In the degradome libraries of *F. robusta* (C_3), REVOLUTA (member of homeodomain-leucine zipper family), Target of early activation tagged (EAT) and Molecular chaperone DnaJ were identified as potential targets of miR166, miR172 and miR319, respectively. In the libraries of *F. ramosissima* (C_3-C_4 intermediate), Protein ABIL4, hypothetical protein and Choline/ethanolamine and phosphotransferase 1 respectively were identified as potential targets of miR159, miR160 and miR168, respectively. Lastly in the libraries of *F. bidentis* (C_4), MYB domain protein, hypothetical protein 111 and Protein tyrosine phosphatase transcripts were identified as potential targets of miR398, miR858b and miR858c respectively.

Overall the results did not identify any potential target transcripts that can be directly associated with photosynthesis. However, due to the lack of a sufficient number of annotated transcripts from the *Flaveria* spp., we were prevented from identifying all miRNA targets. The missing link between the conserved miRNAs in *Flaveria* spp. and their roles in photosynthesis could be best elucidated with help of a well-assembled transcriptome database of *Flaveria* spp. This will assist in identifying all the potential miRNA targets that could function in photosynthesis regulation.

In conclusion, this study contributes to unearthing the identity of miRNAs in *Flaveria* spp. Most importantly, the expression profiles of conserved miRNAs differed greatly in leaves, in mesophyll cells and in bundle sheath cells of *F. robusta* (C_3), *F. ramosissima* (C_3-C_4 intermediate) and *F. bidentis* (C_4). Such differences in miRNA expression patterns in C_3, C_3-C_4 intermediate and C_4 *Flaveria* spp. makes them potential regulators of photosynthesis.

REFERENCES

Addo-Quaye, C., Eshoo, T. W., Bartel, D. P., & Axtell, M. J. (2008). Endogenous siRNA and miRNA targets identified by sequencing of the *Arabidopsis* degradome. *Current Biology*, *18*(10), 758-762.

Addo-Quaye, C., Miller, W., & Axtell, M. J. (2009). CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. *Bioinformatics*, *25*(1), 130-131.

Allen, R. S., Li, J., Stahle, M. I., Dubroué, A., Gubler, F., & Millar, A. A. (2007). Genetic analysis reveals functional redundancy and the major target genes of the *Arabidopsis* miR159 family. *Proceedings of the National Academy of Sciences*, *104*(41), 16371-16376.

Alonso-Peral, M. M., Li, J., Li, Y., Allen, R. S., Schnippenkoetter, W., Ohms, S., and Millar, A. A. (2010). The microRNA159-regulated GAMYB-like genes inhibit growth and promote programmed cell death in *Arabidopsis*. *Plant Physiology*, *154*(2), 757-771.

Ambros, V. (2004). The functions of animal microRNAs. *Nature*, *431*(7006), 350-355.

Aoyagi, K., and Bassham, J. A. (1984). Pyruvate orthophosphate dikinase of C3 seeds and leaves as compared to the enzyme from maize. *Plant physiology*, *75*(2), 387-392.

Aubry, S., Brown, N. J., and Hibberd, J. M. (2011). The role of proteins in C3 plants prior to their recruitment into the C4 pathway. *Journal of Experimental Botany*, *62*(9), 3049-3059

Aukerman, M. J., & Sakai, H. (2003). Regulation of flowering time and floral organ identity by a microRNA and its APETALA2-like target genes. *The Plant Cell*, *15*(11), 2730-2741.

Axtell, M. J., & Bartel, D. P. (2005). Antiquity of microRNAs and their targets in land plants. *The Plant Cell*, *17*(6), 1658-1673.

Baker, C. C., Sieber, P., Wellmer, F., & Meyerowitz, E. M. (2005). The early extra petals1 mutant uncovers a role for microRNA miR164c in regulating petal number in *Arabidopsis*. *Current Biology*, *15*(4), 303-315.

Baldrich, P., Kakar, K., Siré, C., Moreno, A. B., Berger, A., García-Chapa, M., & San Segundo, B. (2014). Small RNA profiling reveals regulation of *Arabidopsis* miR168 and heterochromatic siRNA415 in response to fungal elicitors. *BMC genomics*, *15*(1), 1083.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, *116*(2), 281-297.

Baulcombe, D. (2004). RNA silencing in plants. *Nature*, *431*(7006), 356-363.

Bauwe, H. (2011). Photorespiration: the bridge to C4 photosynthesis. In C4 photosynthesis and related CO_2 concentrating mechanisms. *Springer Netherlands*, (81-108).

Bauwe, H., Hagemann, M., and Fernie, A. R. (2010). Photorespiration: players, partners and origin. *Trends in plant science*, *15*(6), 330-336.

Bergh, V. D. E., Kulahoglu, C., Brautigam, A., Hibberd, J. M., Weber, A. P. M., Zhu, X. G., and Schranz, M. E. (2014). Gene and genome duplications and the origin of C4 photosynthesis: Birth of a trait in the *Cleomaceae*. *Current Plant Biology,* 1: 2–9.

Bond, W. J., Woodward, F. I., & Midgley, G. F. (2005). The global distribution of ecosystems in a world without fire. *New phytologist*, *165*(2), 525-538.

Bouton, J. H., Brown, R. H., Evans, P. T., and Jernstedt, J. A. (1986). Photosynthesis, Leaf Anatomy, and Morphology of Progeny from Hybrids between C3 and C3-C4 *Panicum Species*. *Plant Physiology*, *80*(2), 487-492.

Bräutigam, A., and Gowik, U. (2016). Photorespiration connects C3 and C4 photosynthesis. Journal of experimental botany, 56

Brautigam, A., Kajala, K., Wullenweber, J., Sommer, M., Gagneul, D., Weber, K. L., and Weber, A. P. (2011). An mRNA blueprint for C4 photosynthesis derived from comparative transcriptomics of closely related C3 and C4 species. *Plant Physiology*, *155*(1), 142-156.

Brodersen, P., and Voinnet, O. (2006). The diversity of RNA silencing pathways in plants. *TRENDS in Genetics*, *22*(5), 268-280.

Brown, R. H., and Hattersley, P. W. (1989). Leaf anatomy of C3-C4 species as related to evolution of C4 photosynthesis. *Plant Physiology*, *91*(4), 1543-1550.

Chang, Y. M., Liu, W. Y., Shih, A. C. C., Shen, M. N., Lu, C. H., Lu, M. Y. J., ... & Li, W. H. (2012). Characterizing regulatory and functional differentiation between maize mesophyll and bundle sheath cells by transcriptomic analysis. *Plant physiology*, *160*(1), 165-177.

Chen, X. (2004). A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development. *Science*, *303*(5666), 2022-2025.

Chen, X. (2005). MicroRNA biogenesis and function in plants. *FEBS letters*, *579*(26), 5923-5931.

Dengler, N. G., and Nelson, T. (1999). Leaf structure and development in C4 plants.133-172.

Dong, J., Kim, S. T., & Lord, E. M. (2005). Plantacyanin plays a role in reproduction in *Arabidopsis*. *Plant Physiology*, *138*(2), 778-789.

Drincovich, M. F., Casati, P., Andreo, C. S., Chessin, S. J., Franceschi, V. R., Edwards, G. E., and

Edwards, G. E., & Black, C. C. (1971). Isolation of mesophyll cells and bundle sheath cells from *Digitaria sanguinalis* (L.) leaves and a scanning microscopy study of the internal leaf cell morphology. *Plant physiology*, *47*(1), 149-156.

Edwards, G., & Walker, D. (1983). C3, C4*: mechanisms, and cellular and environmental regulation, of photosynthesis*. University of California Press.

Ehleringer, J. R., Cerling, T. E., and Helliker, B. R. (1997). C4 photosynthesis, atmospheric CO_2, and climate. *Oecologia*, *112*(3), 285-299.

Ehleringer, J. R., and Monson, R. K. (1993). Evolutionary and ecological aspects of photosynthetic pathway variation. *Annual Review of Ecology and Systematics*, 411-439.

Emery, J. F., Floyd, S. K., Alvarez, J., Eshed, Y., Hawker, N. P., Izhaki, A., and Bowman, J. L. (2003). Radial patterning of *Arabidopsis* shoots by class III HD-ZIP and KANADI genes. *Current Biology*, *13*(20), 1768-1774.

German M.A., Pillay M., Jeong D.H., et al. 2008. Global identification of microRNA–target RNA pairs by parallel analysis of RNA ends. *Nature Biotechnology* 26, 941–946

Gowik, U., Bräutigam, A., Weber, K. L., Weber, A. P., and Westhoff, P. (2011). Evolution of C4 photosynthesis in the genus *Flaveria*: how many and which genes does it take to make C4? *The Plant Cell*, *23*(6), 2087-2105.

Gowik, U., Burscheidt, J., Akyildiz, M., Schlue, U., Koczor, M., Streubel, M., and Westhoff, P. (2004). Cis-Regulatory elements for mesophyll-specific gene expression in the C4 plant *Flaveria trinervia*, the promoter of the C4 phosphoenolpyruvate carboxylase gene. *The Plant Cell*, *16*(5), 1077-1090.

Guo, H. S., Xie, Q., Fei, J. F., & Chua, N. H. (2005). MicroRNA directs mRNA cleavage of the transcription factor NAC1 to downregulate auxin signals for *Arabidopsis* lateral root development. *The Plant Cell*, *17*(5), 1376-1386.

Guy, R. D., Vanlerberghe, G. C., & Turpin, D. H. (1989). Significance of phosphoenolpyruvate carboxylase during ammonium assimilation Carbon isotope discrimination in photosynthesis and respiration by the N-limited green alga *Selenastrum minutum*. *Plant physiology*, *89*(4), 1150-1157.

Hatch, M. D., & Slack, C. R. (1966). Photosynthesis by sugar-cane leaves: a new carboxylation reaction and the pathway of sugar formation. *Biochemical Journal*, *101*(1), 103.

Hackenberg, M., Gustafson, P., Langridge, P., & Shi, B. J. (2015). Differential expression of microRNAs and other small RNAs in barley between water and drought conditions. *Plant biotechnology journal*, *13*(1), 2-13.

Hall, L. N., Rossini, L., Cribb, L., and Langdale, J. A. (1998). GOLDEN 2: a novel transcriptional regulator of cellular differentiation in the maize leaf. *The Plant Cell*, *10*(6), 925-936.

Han, M. H., Goud, S., Song, L., and Fedoroff, N. (2004). The *Arabidopsis* double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(4), 1093-1098.

Harrison, P. M., and Gerstein, M. (2002). Studying genomes through the aeons: protein families, pseudo genes and proteome evolution. *Journal of Molecular Biology*, *318*(5), 1155-1174.

Hausler, R. E., Hirsch, H. J., Kreuzaler, F., and Peterhänsel, C. (2002). Over expression of C4-cyc

le enzymes in transgenic C_3 plants: a biotechnological approach to improve C_3-photosynthesis. *Jo urnal of Experimental Botany*, *53*(369), 591-607.

Hibberd, J. M., and Covshoff, S. (2010). The regulation of gene expression required for C_4 photosynthesis. *Annual Review of Plant Biology*, *61*, 181-207.

Hibberd, J. M., Sheehy, J. E., & Langdale, J. A. (2008). Using C_4 photosynthesis to increase the yield of rice—rationale and feasibility. *Current opinion in plant biology*, *11*(2), 228-231.

Hoang, C. V., and Chapman, K. D. (2002). Regulation of carbonic anhydrase gene expression in cotyledons of cotton (Gossypium hirsutum L.) seedlings during post-germinative growth. *Plant molecular biology*, *49*(5), 449-458.

Hoang, C. V., Wessler, H. G., Local, A., Turley, R. B., Benjamin, R. C., and Chapman, K. D. (1999). Identification and expression of cotton (*Gossypium hirsutum L.*) plastidial carbonic anhydrase. *Plant and Cell Physiology*, *40*(12), 1262-1270.

Jagadeeswaran, G., Saini, A., & Sunkar, R. (2009). Biotic and abiotic stress down-regulate miR398 expression in *Arabidopsis*. *Planta*, *229*(4), 1009-1014.

Jones-Rhoades, M. W. (2012). Conservation and divergence in plant microRNAs. *Plant molecular biology*, *80*(1), 3-16.

Jones-Rhoades, M. W., Bartel, D. P., and Bartel, B. (2006). MicroRNAs and their regulatory role s in plants. *Annu. Rev. Plant Biol.*, *57*, 19-53.

Jung, J. H., & Park, C. M. (2007). MIR166/165 genes exhibit dynamic expression patterns in regulating shoot apical meristem and floral development in *Arabidopsis*. *Planta*, *225*(6), 1327-1338.

Kajala, K., Brown, N. J., Williams, B. P., Borrill, P., Taylor, L. E., and Hibberd, J. M. (2012). Multiple *Arabidopsis* genes primed for recruitment into C_4 photosynthesis. *The Plant Journal*, *69*(1), 47-56.

Kanai, R., & Edwards, G. E. (1973). Separation of mesophyll protoplasts and bundle sheath cells from maize leaves for photosynthetic studies. *Plant Physiology*, *51*(6), 1133-1137.

Kopriva, S., & Koprivova, A. (2005). Sulfate assimilation and glutathione synthesis in C_4 plants. *Photosynthesis research*, *86*(3), 363-372.

Ku, Maurice S.B, Jingrui Wu, Ziyu Dai, Rick A. Scott, Chun Chu, and Gerald E. Edwards (1991). Photosynthetic and photorespiratory characteristics of *Flaveria* spp. *Plant Physiology* 96, (2): 518-528.

Külahoglu, C., Denton, A. K., Sommer, M., Maß, J., Schliesky, S., Wrobel, T. J., and Weber, A. P. (2014). Comparative Transcriptome Atlases Reveal Altered Gene Expression Modules between Two Cleomaceae C_3 and C_4 Plant Species. *The Plant Cell Online*, *26*(8), 3243-3260

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. *Science*, *294*(5543), 853-858.

Lahiri, S. D., & Allison, L. A. (2000). Complementary Expression of Two Plastid-Localized ς-Like Factors in Maize. *Plant physiology*, *123*(3), 883-894.

Lai, L. B., Wang, L., and Nelson, T. M. (2002). Distinct but conserved functions for two chloroplastic NADP-malic enzyme isoforms in C3 and C4 *Flaveria* spp. *Plant Physiology*, *128*(1), 125-139.

Lau, N. C., Lim, L. P., Weinstein, E. G., & Bartel, D. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science*, *294*(5543), 858-862.

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

Leegod, R.C. (2007). A welcome diversion from photorespiration. *Nature Biotechnology*, *25*(5), 539-539.

Levy, A., Erlanger, M., Rosenthal, M., and Epel, B. L. (2007). A plasmodesmata-associated β-1, 3-glucanase in *Arabidopsis*. *The Plant Journal*, *49*(4), 669-682.

Li, Y. F., Zheng, Y., Addo-Quaye, C., Zhang, L., Saini, A., Jagadeeswaran, G., and Sunkar, R. (2010). Transcriptome-wide identification of microRNA targets in rice. *The Plant Journal*, *62*(5), 742-759.

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

Lima, J. C. D., Loss-Morais, G., & Margis, R. (2012). MicroRNAs play critical roles during plant development and in response to abiotic stresses. *Genetics and molecular biology*, *35*(4), 1069-1077.

Llave, C., Kasschau, K. D., Rector, M. A., and Carrington, J. C. (2002). Endogenous and silencing-associated small RNAs in plants. *The Plant Cell Online*, *14*(7), 1605-1619.

Lobbes, D., Rallapalli, G., Schmidt, D. D., Martin, C., and Clarke, J. (2006). SERRATE: a new player on the plant microRNA scene. *EMBO reports*, *7*(10), 1052-1058.

Lu, X. Y., and Huang, X. L. (2008). Plant miRNAs and abiotic stress responses. *Biochemical and biophysical research communications*, *368*(3), 458-462.

Ludwig, M. (2013). Evolution of the C4 photosynthetic pathway: events at the cellular and molecular levels. *Photosynthesis Research*, *117*(1-3), 147-161.

Ludwig, M., and Burnell, J. N. (1995). Molecular comparison of carbonic anhydrase from *Flaveria* spp. demonstrating different photosynthetic pathways. *Plant Molecular Biology*, *29*(2), 353-365.

Mallmann, J., Heckmann, D., Bräutigam, A., Lercher, M. J., Weber, A. P., Westhoff, P., and Gowik, U. (2014). The role of photorespiration during the evolution of C4 photosynthesis in the genus *Flaveria*. *elife*, *3*, e02478.

McHale, M., Eamens, A. L., Finnegan, E. J., and Waterhouse, P. M. (2013). A 22-nt artificial microRNA mediates widespread RNA silencing in *Arabidopsis*. *The Plant Journal*, *76*(3), 519-

529.

McKown A. D, Moncalvo J. M, Dengler N. G. 2005. Phylogeny of *Flaveria* (*Asteraceae*) and inference of C4 photosynthesis evolution. *American Journal of Botany* 92, 1911–1928.

Melzer, E., & O'Leary, M. H. (1987). Anapleurotic CO_2 fixation by phosphoenolpyruvate carboxylase in C3 plants. *Plant Physiology*, *84*(1), 58-60.

Monson, R. K. (2003). Gene duplication, neofunctionalization, and the evolution of C4 photosynthesis. *International Journal of Plant Sciences*, *164*(S3), S43-S54.

Monson, R. K. and Rowan, F. (1999). The origins of C4 genes and evolutionary pattern in the C4 metabolic phenotype. *C4 Plant Biology*, 377-410.

Monson, R. K., and Moore, B. D. (1989). On the significance of C3-C4 intermediate photosynthesis to the evolution of C4 photosynthesis. *Plant, Cell & Environment*, *12*(7), 689-699.

Monson, R. K., Edwards, G. E., and Ku, M. S. (1984). C3-C4 intermediate photosynthesis in plants. *BioScience*, *34*(9), 563-574.

Muhaidat, R., Sage, R. F., and Dengler, N. G. (2007). Diversity of Kranz anatomy and biochemistry in C4 eudicots. *American Journal of Botany*, *94*(3), 362-381.

Nogueira, F. T., Madi, S., Chitwood, D. H., Juarez, M. T., and Timmermans, M. C. (2007). Two small regulatory RNAs establish opposing fates of a developmental axis. *Genes & development*, *21*(7), 750-755.

Oikawa, K., Kasahara, M., Kiyosue, T., Kagawa, T., Suetsugu, N., Takahashi, F., and Wada, M. (2003). Chloroplast unusual positioning1 is essential for proper chloroplast positioning. *The Plant Cell Online*, *15*(12), 2805-2815.

Palatnik, J. F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J. C., & Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. *Nature*, *425*(6955), 257-263.

Palatnik, J. F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J. C., & Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. *Nature*, *425*(6955), 257-263.

Parizotto, E. A., Dunoyer, P., Rahm, N., Himber, C., and Voinnet, O. (2004). In vivo investigation of the transcription, processing, endonucleolytic activity, and functional relevance of the spatial distribution of a plant miRNA. *Genes & development*, *18*(18), 2237-2242.

Patel, M., Siegel, A. J., and Berry, J. O. (2006). Untranslated regions of FbRbcS1 mRNA mediate bundle sheath cell-specific gene expression in leaves of a C4 plant. *Journal of Biological Chemistry*, *281*(35), 25485-25491.

Potter, J. W., & Black, C. C. (1982). Differential Protein Composition and Gene Expression in Leaf Mesophyll Cells and Bundle Sheath Cells of the C4 Plant *Digitaria sanguinalis* (L.) Scop. *Plant physiology*, *70*(2), 590-597.

Powell, A.M. (1978). Systematics of *Flaveria* (*Flaveriinae-Asteraceae*. Annals of the Missouri Botanical Garden 65:590–636

Price, G. D., von Caemmerer, S., Evans, J. R., Yu, J. W., Lloyd, J., Oja, V., and Badger, M. R. (1994). Specific reduction of chloroplast carbonic anhydrase activity by antisense RNA in transgenic tobacco plants has a minor effect on photosynthetic CO_2 assimilation. *Planta*, *193*(3), 331-340.

Ramachandran, V., and Chen, X. (2008). Degradation of microRNAs by a family of exoribonucleases in *Arabidopsis*. *Science*, *321*(5895), 1490-1492.

Rawsthorne, S. (1992). C_3–C_4 intermediate photosynthesis: linking physiology to gene expression. *The Plant Journal*, *2*(3), 267-274.

Reyes, J. L., and Chua, N. H. (2007). ABA induction of miR159 controls transcript levels of two MYB factors during *Arabidopsis* seed germination. *The Plant Journal*, *49*(4), 592-606.

Rodriguez, R. E., Mecchia, M. A., Debernardi, J. M., Schommer, C., Weigel, D., & Palatnik, J. F. (2010). Control of cell proliferation in *Arabidopsis thaliana* by microRNA miR396. *Development*, *137*(1), 103-112.

Sage, R. F. (2001). Environmental and Evolutionary Preconditions for the Origin and Diversification of the C_4 Photosynthetic Syndrome. *Plant Biology*, *3*(3), 202-213.

Sage, R. F. (2004). The evolution of C_4 photosynthesis. *New phytologist*, *161*(2), 341-370.

Sage, R. F., & Monson, R. K. (1998). *C_4 plant biology*. Academic Press.

Sage, R. F., and McKown, A. D. (2006). Is C_4 photosynthesis less phenotypically plastic than C_3 photosynthesis? *Journal of Experimental Botany*, *57*(2), 303-317.

Sage, T. L., & Sage, R. F. (2009). The functional anatomy of rice leaves: implications for refixation of photorespiratory CO_2 and efforts to engineer C_4 photosynthesis into rice. *Plant and Cell Physiology*, *50*(4), 756-772.

Sage, R. F., Khoshravesh, R., and Sage, T. L. (2014). From proto-Kranz to C_4 Kranz: building the bridge to C_4 photosynthesis. *Journal of experimental botany*, 180.

Sage, R. F., Sage, T. L., and Kocacinar, F. (2012). Photorespiration and the evolution of C_4 photosynthesis. *Annual review of plant biology*, *63*, 19-47.

Sage, R. F. and R. K. Monson (1999). Why C_4 photosynthesis? *C_4 plant biology*, 3-16.

Schlüter, U., & Weber, A. P. (2016). The road to C_4 photosynthesis: evolution of a complex trait via intermediary states. *Plant and Cell Physiology*, *57*(5), 881-889.

Schommer, C., Palatnik, J. F., Aggarwal, P., Chételat, A., Cubas, P., Farmer, E. E., ... & Weigel, D. (2008). Control of jasmonate biosynthesis and senescence by miR319 targets. *PLoS Biol*, *6*(9), e230.

Schommer, C., Bresso, E. G., Spinelli, S. V., & Palatnik, J. F. (2012). Role of microRNA miR319 in plant development. In *MicroRNAs in Plant Development and Stress Responses* Springer Berlin Heidelberg. pp. 29-47.

Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M., & Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. *Developmental cell*, *8*(4), 517-527.

Sharkey, T. D. (1988). Estimating the rate of photorespiration in leaves. *Physiologia Plantarum*, *73*(1), 147-152.

Sheen, J. (1991). Molecular mechanisms underlying the differential expression of maize pyruvate, orthophosphate dikinase genes. *The Plant Cell Online*, *3*(3), 225-245.

Sieber, P., Wellmer, F., Gheyselinck, J., Riechmann, J. L., & Meyerowitz, E. M. (2007). Redundancy and specialization among plant microRNAs: role of the MIR164 family in developmental robustness. *Development*, *134*(6), 1051-1060.

Silva-Pérez, V., Furbank, R. T., Condon, A. G., & Evans, J. R. (2017). Biochemical model of C3 photosynthesis applied to wheat at different temperatures. *Plant, Cell & Environment*.

Slewinski, T. L., Anderson, A. A., Zhang, C., and Turgeon, R. (2012). Scarecrow plays a role in establishing Kranz anatomy in maize leaves. *Plant and Cell Physiology*, *53*(12), 2030-2037.

Still, C. J., Berry, J. A., Collatz, G. J., and DeFries, R. S. (2003). Global distribution of C3 and C4 vegetation: carbon cycle implications. *Global Biogeochemical Cycles*, *17*(1).

Sunkar, R. (2010). MicroRNAs with macro-effects on plant stress responses. In *Seminars in cell & developmental biology* (Vol. 21, No. 8, pp. 805-811). Academic Press.

Sunkar R., Zhu J.K. (2004). Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. Plant Cell *16*: 2001–2019.

Sunkar, R., & Jagadeeswaran, G. (2008). In silico identification of conserved microRNAs in large number of diverse plant species. *BMC plant biology*, *8*(1), 37.

Sunkar, R., Chinnusamy, V., Zhu, J., & Zhu, J. K. (2007). Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. *Trends in plant science*, *12*(7), 301-309.

Sunkar, R., Girke, T., Jain, P. K., & Zhu, J. K. (2005). Cloning and characterization of microRNAs from rice. *The Plant Cell*, *17*(5), 1397-1411.

Sunkar, R., Li, Y. F., and Jagadeeswaran, G. (2012). Functions of microRNAs in plant stress responses. *Trends in Plant Science*, *17*(4), 196-203.

Sunkar, R., Zhou, X., Zheng, Y., Zhang, W., & Zhu, J. K. (2008). Identification of novel and candidate miRNAs in rice by high throughput sequencing. *BMC plant biology*, *8*(1), 25.

Tang, G., Reinhart, B. J., Bartel, D. P., & Zamore, P. D. (2003). A biochemical framework for RNA silencing in plants. *Genes & development*, *17*(1), 49-63.

Taylor, L., Nunes-Nesi, A., Parsley, K., Leiss, A., Leach, G., Coates, S., and Hibberd, J. M. (2010). Cytosolic pyruvate, orthophosphate dikinase functions in nitrogen remobilization during leaf senescence and limits individual seed growth and nitrogen content. *The Plant Journal*, *62*(4), 641-652.

Vaucheret, H. (2008). Plant argonautes. *Trends in plant science*, *13*(7), 350-358.

Vaucheret, H., Vazquez, F., Crété, P., & Bartel, D. P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes & development*, *18*(10), 1187-1197.

Vazquez, F., Vaucheret, H., Rajagopalan, R., Lepers, C., Gasciolli, V., Mallory, A. C., and Crété, P. (2004). Endogenous trans-Acting siRNAs regulate the accumulation of *Arabidopsis* mRNAs. *Molecular Cell*, *16*(1), 69-79.

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

Wang, J. W., Schwab, R., Czech, B., Mica, E., & Weigel, D. (2008). Dual effects of miR156-targeted SPL genes and CYP78A5/KLUH on plastochron length and organ size in *Arabidopsis* thaliana. *The Plant Cell*, *20*(5), 1231-1243.

Wang, P., Kelly, S., Fouracre, J. P., and Langdale, J. A. (2013). Genome-wide transcript analysis of early maize leaf development reveals gene cohorts associated with the differentiation of C4 Kranz anatomy. *The Plant Journal*, *75*(4), 656-670.

Waters, M. T., Wang, P., Korkaric, M., Capper, R. G., Saunders, N. J., and Langdale, J. A. (2009). GLK transcription factors coordinate expression of the photosynthetic apparatus in *Arabidopsis*. *The Plant Cell*, *21*(4), 1109-1128.

Weckopp, S. C., & Kopriva, S. (2015). Are changes in sulfate assimilation pathway needed for evolution of C4 photosynthesis? *Frontiers in plant science*, *5*, 773.

Xie, K., Wu, C., & Xiong, L. (2006). Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and microRNA156 in rice. *Plant physiology*, *142*(1), 280-293.

Xie, Z., Johansen, L. K., Gustafson, A. M., Kasschau, K. D., Lellis, A. D., Zilberman, D., & Carrington, J. C. (2004). Genetic and functional diversification of small RNA pathways in plants. *PLoS biology*, *2*(5), e104.

Yang, L., Liu, Z., Lu, F., Dong, A., and Huang, H. (2006). SERRATE is a novel nuclear regulator in primary microRNA processing in *Arabidopsis*. *The Plant Journal*, *47*(6), 841-850.

Yang, X., Wang, L., Yuan, D., Lindsey, K., & Zhang, X. (2013). Small RNA and degradome sequencing reveal complex miRNA regulation during cotton somatic embryogenesis. *Journal of experimental botany*, ert013.

Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R. W. and Chen, X. (2005). Methylation as a crucial step in plant microRNA biogenesis. *Science*, *307*(5711), 932-935.

Zhang, B., Pan, X., Wang, Q., Cobb, G. P., & Anderson, T. A. (2006). Computational identification of microRNAs and their targets. *Computational biology and chemistry*, *30*(6), 395-407.

Zheng, Y., Li, Y. F., Sunkar, R., & Zhang, W. (2012). SeqTar: an effective method for identifying microRNA guided cleavage sites from degradome of polyadenylated transcripts in plants. *Nucleic acids research*, *40*(4), e28-e28.

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

VITA

Robert Pokoo

Candidate for the Degree of

Doctor of Philosophy

Thesis: A COMPARATIVE ANALYSIS OF MIRNAs IN *FLAVERIA* SPECIES OF C_3, C_3-C_4 INTERMEDIATE AND C_4 PHOTOSYNTHESIS

Major Field: Biochemistry and Molecular Biology

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Biochemistry and Molecular Biology at Oklahoma State University, Stillwater, Oklahoma in May, 2017.

Completed the requirements for the Master of Science in Biochemistry at Chungnam National University, Daejeon, Yuseong-Gu/South Korea in 2012.

Completed the requirements for the Bachelor of Science in Biochemistry at University of Cape Coast, Cape Coast/Ghana 2007.

Experience:

Professional Memberships:
2012-2017: Biochemistry and Molecular Biology Graduate Students' Association (Member, Faculty Representative)
2005-2007: Biochemistry Students' Association of Ghana (Member)