Computational prediction and qPCR validation of miRNA in Ceratopteris richardii

Yevgeniy Kharonov

Baruch College

Follow this and additional works at: https://academicworks.cuny.edu/bb_etds

Recommended Citation
Kharonov, Yevgeniy, "Computational prediction and qPCR validation of miRNA in Ceratopteris richardii" (2013). CUNY Academic Works.
https://academicworks.cuny.edu/bb_etds/31
Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

YEVGENIY KHARONOV

Submitted to the Committee on Undergraduate Honors of Baruch Honors College
of the City University of New York
in partial fulfillment of the requirement for the degree of
Bachelor of Arts in *ad hoc* Biology with Honors

Submitted Spring Semester 2013

Approved by the Department of Natural Sciences:

________________________________
Dr. Edward B. Tucker
Mentor and faculty sponsor

________________________________  ______________________________
Dr. Jason Munshi-South    Dr. Seymour Schulman
Committee member            Committee member
# TABLE OF CONTENTS

**EXECUTIVE SUMMARY**

**ABSTRACT**

## 1.0 INTRODUCTION

1.1 **CERATOPTERIS RICHARDII, A MODEL ORGANISM**

1.2 **THE LIFE CYCLE OF CERATOPTERIS RICHARDII**

1.3 **THE CELL WALL & PLASMOLYSIS**

1.4 **REGULATORY RNA**

1.5 **MICRORNA & THEIR ROLE IN PLANT STRESS RESPONSE**

1.6 **A BRIEF CHRONICLE OF BIOINFORMATICS**

1.7 **QUANTITATIVE REAL-TIME POLYMERASE CHAIN REACTION**

1.8 **FOCUS**

## 2.0 METHODS

2.1 **MICRORNA PREDICTION**

2.2 **CERATOPTERIS RICHARDII CULTURING**

2.3 **PLASMOLYSIS**

2.4 **RNA EXTRACTION**

2.5 **POLY(A) TAILING AND CDNA SYNTHESIS**

2.6 **QPCR PROTOCOL**

## 3.0 RESULTS

3.1 **PREDICTIONS**

3.2 **EFFECT OF PLASMOLYSIS**

3.3 **QPCR**

## 4.0 DISCUSSION

4.1 **CONSERVED PRE-MICRORNA**

4.2 **CONSERVED MICRORNA**

4.3 **FALSE POSITIVES & NEGATIVES**

4.4 **PLASMOLYSIS REGULATES MICRORNA EXPRESSION**

## 5.0 CONCLUSIONS

**ACKNOWLEDGEMENTS**

**REFERENCES**
Executive Summary

With its active molten core, thick stormy atmosphere, and huge quantities of water, Earth, is anything but an inert planet. Earth partakes in wonderful and awe-inspiring transformations that occur on astonishing timescales. To be successful on Earth, life had to develop a way to convey experience from the environment to the deepest level of the cell, the genetic code. Life’s ability to adapt is contained in a hereditary, self-replicating molecule called deoxyribonucleic acid (DNA). The information contained in certain segments of DNA called genes is the blueprint for organismal development. But how does life quickly convey information from the environment to its genetic code?

In the last two decades it has been shown that vast networks of complex gene regulatory mechanisms exist in all forms of life. Imagine the fate of a cell as a destination and gene regulation as the GPS. As the GPS unit, gene regulation, allows organisms to alter the levels of proteins expressed at any given time in response to the surroundings. It is important to alter protein expression levels because proteins carry out the majority of the work inside of cells. This GPS allows the organism to navigate through various degrees of commitment on a developmental path.

In this study we focused on an emerging mechanism of gene regulation known as microRNA (miRNA) interference. MiRNA are short segments of RNA that originate from non-protein-coding genes in DNA and associate with specific proteins to form a molecular machine called the miRNA-induced silencing complex (RISC). In plants, microRNA direct the binding of RISC to messenger RNA (mRNA), causing their degradation. The point of interference follows the transcription of DNA into RNA but precedes the translation of RNA into protein and is
therefore called post-transcriptional gene regulation. This mechanism of post-transcriptional gene regulation has drastic effects on cell commitment and fate.

We observed that the fern (*Ceratopteris richardii*), when exposed to dehydration (a type of abiotic stress), would develop aberrantly and display uncharacterized deformities in leaf, stem, root, and sexual organ structure and function. We believe that miRNA are somehow involved in a type of genetic reprogramming that occurs after dehydration. This reprogramming informs the fern of the abiotic stress induced and guides the organism toward a new developmental path for survival. The goal of the project is to characterize miRNA in our organism and determine their effects.

The genome for our model organism has not been sequenced yet and so a full description of the genetic code is not available, making it difficult to determine which miRNAs are expressed. In order to predict the miRNA in our model we used complex computational methods to access repositories of genes known as Expressed Sequence Tags (ESTs). From ESTs we were able to predict about 100 miRNA that we believe exist in this fern. We were also able to computationally predict miRNA target interactions. In other words, we predicted miRNA, their mRNA targets, and the effect of rising or diminishing levels of miRNA on the fern cell. Furthermore, we experimentally measured the expression levels of several miRNAs using a quantitative technique called real-time polymerase chain reaction.

We now believe that the stress of dehydration on these ferns caused the loss of cell-cell adhesions that mediate communication between cells of the fern. With the loss of this communication, the now individual cells had to reorganize the information they expressed to ensure survival. We believe the miRNA predicted in this paper are powerful contributors to the
Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

reprogramming observed in protein expression. The role of these miRNA and their targets are discussed and a foundation is set for future work in this area.

Revealing the role of miRNA in plants as well as other life forms (including human) will drastically alter the way in which farmers and even healthcare practitioners carry out their work. In the field of agriculture, specific miRNA molecules can be used to perform RNA interference and make crops less susceptible to environmental factors (biotic and abiotic stress), have a different nutritional value, or alter taste and shelf life. In medicine, miRNA will provide a form of patient customized therapy that can potentially be used to treat and possibly cure life threatening or therapeutically limited diseases.
Abstract

MicroRNAs were predicted in the Expressed Sequence Tags available for the fern *Ceratopteris richardii*. 9 unique pre-miRNAs were predicted from a total of 327,172 EST alignments. 89 unique mature miRNAs were predicted from a total of 76,940 EST alignments. The role of several miRNAs in response to plasmolysis was investigated. 8 and 12-day-old *C. richardii* gametophytes were treated with 1.0 Molar mannitol for 30 minutes followed by rehydration in dH2O. Mannitol/dH2O treatment induced plasmolysis that may have caused the breakdown of cell-cell adhesions leading to an aberrant morphology termed the "secondary gametophyte." RNA isolation was performed pre-plasmolysis, 24, 48, and 72 hours post-plasmolysis.

MiR106b-5p a regulator of cyclin-dependent kinase inhibitor 1A (p21/Cip1) was up-regulated post-plasmolysis and may be responsible for the release of cell cycle arrest typical in developing gametophytes. MiR181c was up-regulated and is thought to inhibit several key cell cycle proteins, including transportin1 a nuclear transporter, fidgetin a protein thought to sever microtubule and spindle formation, as well as DDX3X an ATP-dependent RNA helicase important for regulation of apoptotic pathway and G1/S transition of mitotic cell cycle. MiR181c may play a global role in gametophyte reprogramming post-plasmolysis. Cre-miR914 was detected and although it has no known targets, its detection contributes depth to the extent of which miRNA are conserved from alga to plant.
CHAPTER 1
INTRODUCTION

1.1  *Ceratopteris richardii, a model organism*

One of the major obstacles to the scientific analysis of life systems is to find a suitable model organism. An ideal organism is one that enables observation, experimentation and data collection over an extended period of time, is easy to culture and maintain, is not complex and grows quickly. The underlying physiological and biochemical mechanisms contributing to the development of the homosporous fern *Ceratopteris richardii* have been well characterized over the past century. *C. richardii* has been utilized as a plant model system for pedagogical application in classrooms all over the world.

Freshman biology courses typically provide each student with an individual plate to culture *C. richardii* spores and periodically observe gametophyte development over the semester. Typically gametophytes reach sexual maturity between 2-3 weeks when swimming sperm cells secreted from the antheridia make their way to the eggs in the archegonia. Within 4-5 weeks young embryos can be observed and many sporophytes with well-developed roots may be transferred to terraria after approximately 6 weeks. Students are provided with a rare opportunity to scrutinize basic biological events such as germination, gametophyte development, spermatogenesis, fertilization, embryogenesis, organogenesis and sporophyte growth (Renzaglia, et al 1995). The utility of employing model organisms for basic science research is to be able to find some translational value between the model system and other more complex systems, without introducing confounding factors or unnecessary complexity. The major developmental events in *C. richardii* are depicted in *Figure 1*. Luckily for us, there are still many mysteries left to unravel in this classic model organism.
Figure 1 (Hickok LG and Warne TR, 1998)

Schematic of the major developmental events in life cycle of the fern *C. richardii*.

© American Institute of Biological Sciences 1998.
1.2 *The life cycle of Ceratopteris richardii*

The normal life cycle of our fern is important to consider when gauging the effects of our experimental stress on the organism. Typically, single celled haploid spores (average diameter of 120 μm), which if dry can remain dormant for years; develop rapidly upon exposure to suitably moist conditions. Spore imbibition requires approximately 24 hours and is followed by the initiation of spore germination, which is a light-dependent process that is normally controlled by phytochrome and other photoreceptors (Cooke et al., 1987). Sexual type is determined between 3-5 days, although distinct phenotypic differences are not evident until 7 days (Banks et al., 1993). The presence of both sexual types within a population is caused by asynchrony in spore germination. Gametophytes that germinate early and develop first become hermaphroditic, whereas later-developing gametophytes become male (Hickok et al., 1998). Sexual determination is regulated by a pheromone called antheridiogen that is secreted by hermaphrodites and directs later-germinating gametophytes so that they become male (Alifarag, 2012).

Antheridiogen has not been isolated and structurally identified, however several studies indicate that inhibitors of gibberellin biosynthesis have demonstrated that this compound may share a biosynthetic relationship to gibberellin (Wynne et al., 1998; Rademacher, 2000; Srivastava, 2002). Alifarag (2012) demonstrated that gametophytes sown on a media that absorbed antheridiogen failed to undergo sexual differentiation. Hermaphrodites and males are clearly distinguishable after 10 days, at which point sexual maturity has been reached. Males lack a meristem (plant stem cell tissue) and show determinate growth, whereas hermaphrodites possess a meristem (notch meristem) and grow indeterminately.
Hermaphrodites are obliquely heart-shaped, 2-3 mm in diameter, and possess both antheridia and archegonia. Archegonia are located directly behind the notch meristem, and the few antheridia of hermaphrodites are located initially on the margins and subsequently throughout the body of the gametophyte. By contrast, males are thumb shaped, approximately 1 mm long, and are covered by many small round antheridia on their surface. Each antheridium consists of spermatogenous cells surrounded by a cup-shaped basal cell, a doughnut-shaped ring cell, and a disc-shaped cap cell (Raghavan, 1989).

Fertilization occurs if adequate moisture is present to allow sperm from the antheridia to swim along a chemotactic gradient towards the egg contained in the archegonium. Fertilization of the egg results in a diploid zygote, which quickly develops into a young embryo. Three days after fertilization, young embryos are easily observed. After a successful fertilization, the notch meristem of the hermaphrodite ceases to divide and dedifferentiates into vegetative tissue (Banks, 1997). Embryo development progresses rapidly during which point the gametophyte gradually senesces and dies and the sporophyte continuously forms initial leaves and roots. Young sporophytes can be grown under a variety of conditions (greenhouse, in mini-terraria, hydroponically in fresh water). Sporophytes reach maturity when their leaves contain fertile sporangium in which four large spore mother cells undergo meiosis to produce a total of 16 haploid spores. Viable spores can be obtained from mature sporophytes at 90 days (Hickok and Warne, 1998).
1.3 The cell wall and plasmolysis

To fully appreciate the effect of dehydration on a plant, we need to recognize the function of an organelle that is absent in other forms of life: the plant cell wall. Plant cell walls form load-bearing structures in which polymers of polysaccharides (mostly cellulose microfibrils) are cross-linked with hemicelluloses, including mannans, xylans, mixed-linkage glucans, and xyloglucans (Fry, 2004). This network of sugars is further embedded in a matrix of protein-linked polysaccharides including homogalactuonan, rhamnogalacturonan-I, rhamogalacturonan-II, and xylogalacturonan (Caffal and Mohnen, 2009). Primary cell walls appear first during plant development and establish the foundation for cell shape and allow the cells to resist the tensile forces exerted by turgor pressure and are capable of controlled expansion to enable cell growth (Scheller and Ulvskov, 2010). In non-growing plant tissues, some cells are typically surrounded by “secondary walls” whose function is to resist compressive force (Hepler et al., 1970).

Plasmolysis is a term used to describe the action of osmotic pressure on the cell contents or protoplasm. Healthy plant cells are described as being turgid, a state in which the cell and the vacuole is full of water with the cell membrane right up against the cell wall. The turgid state confers a degree of rigidity to the cells as a unit and is a desirable state because it increases tensile strength, allowing the plant to grow towards the sunlight and against gravity (Campbell et al, 2008). When a plant cell is placed into a hypertonic solution (a solution in which the concentration of solute is higher than the concentration of solute in the cell sap) water rushes out of the cell and causes the cell membrane to recede from the cell wall as displayed in Figure 2.

Furthermore, it is known that the cell wall contains special tubes called plasmodesmata through which individual plant cells achieve cytoplasmic continuity with neighboring cells (Burch-Smith and Zambryski, 2012).
A normal gametophyte (left) and a gametophyte plasmolyzed in 1.0 M-mannitol solution (right) viewed under 100x magnification. In the normal gametophyte the protoplast is right against the cell wall. In the plasmolyzed gametophyte the protoplast has receded away from the cell wall, potentially rupturing inter-cellular connections formed by the plasmodesmata. The loss of this cell-cell adhesion may inform the individual cells to begin altering their gene expression or perish. Through mechanisms of gene regulation, each individual cell alters its developmental fate.
1.4 Regulatory ribonucleic acids

The “central dogma” of molecular biology describes the flow of information within a biological system from DNA to RNA to protein (Crick, 1970). Since Crick’s postulation there have been a bevy of new discoveries that indicate that the flow of information is exponentially more complicated than imagined. The discovery of RNA interference (RNAi) by Andrew Fire, Craig Mello, and colleagues (Fire et al., 1998), for which they were awarded the 2006 Nobel Prize in Physiology and Medicine, redefined how scientists approached the long-believed concepts about RNA. Their discoveries led to the so-called “Renaissance of the Regulatory RNAs.” This new breed of RNA was dubbed “Non-coding RNAs (ncRNAs)” and originates from various types of regulatory DNA, previously termed “junk DNA;” as researchers at the time were unsure of the function of the majority of DNA in humans and other species. Several classes of ncRNAs have emerged in recent years through novel computational and experimental strategies undertaken to identify ncRNA candidates in various model organisms from *Escherichia coli* to human.

The biological functions of these newly characterized ncRNAs are still relatively obscure and researchers are employing novel high-through-put methods such as real-time PCR and next generation sequencing for analysis. These experiments will help researchers better understand the biological significance of ncRNAs in cellular processes as well as the potential therapeutic applications for ncRNAs as biomarkers for diagnosis, prognosis, and prediction of therapeutic outcome (Ghosh and Mallick, 2012). *Figure 3* summarizes the flow of information in Crick’s central dogma as well as some members of the expanding noncoding RNA landscape.
Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

**Figure 3** (Ghosh and Mallick, 2012)

The expanding noncoding RNA landscape.
1.5 MicroRNA and their role in plant stress response

It has been shown that miRNA genes predate the divergence of monocot and dicot plants with many miRNA genes conserved in rice and Arabidopsis (Floyd and Bowman, 2004). The miRNA miR-166 has binding sites for class III homeodomain-leucine zipper (HD-ZIP III) genes that regulate apical embryo patterning, embryonic shoot meristem formation, organ polarity, vascular development, and meristem function (Prigge et al., 2004) from bryophytes to seed plants, indicating that the postranscriptional regulation mediated by miR-166 is a mechanism that can be traced back more than 400 million years (Floyd and Bowman, 2004).

In plants, RNA silencing plays key roles in regulation of biotic stresses such as bacterial, pathogenic, viral, fungal, insect, nematode, etc. and abiotic stresses such as drought, soil salinity, extreme temperatures, heavy metals, etc. (Chao et al., 2005; Si et al., 2009). RNA silencing encompasses a wide variety of mechanisms that typically depend upon a guide RNA molecule that can potentially affect all the different levels of gene expression (e.g. transcription, RNA processing, translation, and postranscriptional modifications).

MicroRNAs (miRNAs) are important for regulating gene expression in eukaryotes and act as a negative regulator of its target mRNA. MiRNA genes exist as independent transcriptional units and are transcribed by RNA polymerase II into long primary transcripts (pri-miRNAs) (Kim, 2006). In plants, the pri-miRNAs precursor is processed by ribonuclease III enzyme Dicer-like1 (DCL1) (Jones-Rhoades, et al., 2006; Chen, 2005) and other proteins including the zinc-finger domain protein Serrate (SE) (Lobbes et al, 2006) and a double-stranded RNA (dsRNA) binding-domain proteins, Hyponastic Leaves1 (HYL1) (Han et al, 2004; Wu et al, 2007) to release the mature miRNA from the nucleus. Hua Enhancer (HEN1) methylates plant miRNAs at their 3’ ends in order to confer stability and prevent exonuclease degradation (Yu et
MiRNAs are exported from the nucleus as a duplex of miRNA/anti-miRNA by the enzyme HASTY1 (an orthologue of the animal protein exportin 5) and subsequently loaded onto Argonaute 1 (AGO1) (Baumberger and Baulcombe, 2005) to form the miRNA induced silencing complex (miRISC), where the miRNA guides the cleavage or translational repression of its target mRNA by base-pairing with the target (Bartel, 2004) and the anti-miRNA is degraded by Small RNA Degrading Nuclease 1 (SDN1) an exoribonuclease that degrades single-stranded small RNAs in vitro and limits the accumulation of small RNAs in vivo (Ramchandran and Chen, 2008). Figure 4 details the biogenesis of plant miRNAs as known thus far.

Plant miRNAs differ from animal miRNAs in their biogenesis with the localization of the pri-miRNA processing (nucleus versus nucleus then cytoplasm) and the enzymes involved (Drosha and Dicer in animals and DCL1 in plants) (Yang, 2007). Plant miRNAs have more variable pre-miRNAs with larger stem-loop structures. Mature plant miRNAs pair to their target sites with near-perfect complementarity. In animals, miRNAs typically recognize several target sequences in the 3’ untranslated region (UTR) of mRNAs and cause translation inhibition, while in plant miRNAs typically target a single site with high specificity in the coding region and guide the mRNA to cleavage (Brodersen and Voinnet, 2006; Mallory and Vaucheret, 2010).
Figure 4 (Diana V Dugas and Bonnie Bartel, 2004) ©Elsevier Ltd.

A simplified schematic of miRNA biogenesis in plants. The transcription of microRNA containing genes gives rise to pre-microRNA. Transcripts are processed by the DCL1 ribonuclease, assisted by the HEN1 and HYL1 proteins, into paired dsRNA of about 21 nucleotides. This dsRNA is shuttled out of the nucleus by HASTY1 (not shown) after which the strands are separated by an RNA helicase. The 5’- 3’ miRNA strand is associated with catalytic proteins, forming RISC. The 3’- 5’ anti-miRNA is degraded by SDN1 (not shown). RISC targets mRNA for cleavage or translation repression.
Systematic and integrative genomic approaches have been applied to understand the roles of miRNA in plant nutrient acquisition, assimilation, metabolism, and adaptations in response to stressors (abiotic and biotic). Several key miRNAs have been implicated in plant homeostasis to ensure normal growth, development and the continuation of the life cycle.

The miRNA miR-399 is important in phosphate metabolism (phosphorus is one of the most influential macronutrients involved in phosphorylation reactions, energy delivery, synthesis of phospholipids and nucleic acids) by its action on the mRNA of the PHO2 gene (coding for an E2 ubiquitin conjugase related enzyme). Plants can readily access inorganic phosphate (P$_i$) from the soil (Raghothama, 1999). The expression of miR-399 as well as the ncRNA Induced by Phosphate Starvation 1 (IPS1) are induced during phosphate starvation and appear to be activated by the PHR1 transcription factor through recognition of the P1BS motif in miR-399 promoters (Fujii et al., 2005; Bari et al., 2006; Franco-Zorrilla et al., 2007; Wang et al., 2009; Pacak et al., 2010; Huang et al., 2011). Inhibition of PHO2 expression by miR-399 is counterbalanced by the short ncRNA IPS1, thus indicating that the miR-399-PHO2-IPS1 circuit plays an important component of P$_i$-deficiency signaling pathway.

The miRNA miR-395 plays an important role in sulfur homeostasis. Sulfur in its reduced form is found mainly in amino acids, peptides and proteins, in iron-sulfur clusters, in lipoic acid and in other co-factors, and in its oxidized form as sulfonate group in modified proteins, polysaccharides and lipids (Rausch and Wachter, 2005; Kopriva, 2006). The major source of sulfur for plants is inorganic sulfate taken up by the roots and re-distributed by sulfate transporters and then assimilated into cysteine, methionine, and glutathiones and thionin (Shukla et al., 2008; Tan et al., 2010). In Arabidopsis, miR-395 was shown to regulate one of the three known sulfate transporters SULTR2;1 and three of the known ATP Sulfurylasese genes APS1,
ASP3, and APS4 which function in the first step of sulfate assimilation (Jones-Rhoades and Bartel, 2004; Allen et al., 2005; Sunkar et al., 2007; Huang et al., 2010; Liang and Yu, 2010). Under sulfate deprivation the expression of sulfate transporters in the root were induced by the protein sulfur limitation1 (SLIM1), leading to an increase in the uptake of sulfate while at the same time inducing the accumulation of miR-395 to restrict the expression of SULTR2;1 to the xylem parenchyma, which leads to enhanced translocation of sulfate from the roots to the shoots and prevents its transport in the phloem from the shoots to the roots (Kawashima et al., 2009, 2011).

Copper (Cu) is crucial as a cofactor to proteins involved in electron transfer reactions and in plant protection against reactive oxygen species (ROS) (Burkhead et al, 2009). Under copper deficient stress conditions miR-397, miR-398, miR-408, and miR-857 were up-regulated and led to simultaneous down-regulation of nonessential Cu-containing proteins and consequently save copper for essential functions (Puig et al, 2007; Yamasaki et al., 2007; Beauclair et al., 2010; Abdel-Ghany and Pilon, 2008; Bonnet et al., 2004; Sunkar, 2010).

MiR-398 has also been implicated in response to oxidative stress or the accumulation of ROS such as superoxide radicals (O$_2^-$), hydrogen peroxide (H$_2$O$_2$) and hydroxyl radicals (OH-) (Mittler 2002; Bartels and Sunkar, 2005). ROS cause damage to nucleic acids proteins and membrane lipids and plants have enzymatic (superoxide dismutases, catalases, peroxidases) and non-enzymatic (carotenoids, xanthophylls, glutathione, tocopherol, ascorbate, etc.) scavenging mechanisms to neutralize oxidative stress. Superoxide dimutases (SODs) play a fundamental role in the stress response by converting highly toxic O$_2^-$ into less toxic hydrogen peroxide (Mittler, 2002; Sunkar et al., 2006). In Arabidopsis two SODs (Cu/Zn CSD1 and CSD2) are regulated by miR-398-directed cleavage of their mRNAs (Yamasaki et al., 2007). Under both abiotic and
Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

Biotic stress conditions the expression of miR-398 is down-regulated transcriptionally and leads to increased accumulation of CSD1 and CSD2 mRNAs and increased tolerance to oxidative stress (Sunkar et al., 2006).

Drought is one of the main environmental factors influencing the yield and distribution of agricultural crops. Numerous genes with altered expression during drought stress have been identified through genome wide gene expression and proteomics to further the understanding of plant tolerance to water limitations (Bartels and Sunkar, 2005; Shinozaki and Yamaguchi-Shinozaki, 2007). Recent studies have also shown that the expression of miRNAs is altered in response to drought stress in many plant species such as cowpea (Barrera-Figueroa et al., 2011), tobacco (Frazier et al., 2011), *Triticum dicoccoides* (Kantar et al., 2011), soybean (Kulcheski et al., 2011), and *Phaseolus vulgaris* (Arenas-Huertero et al., 2009).

In *Arabidopsis* the accumulation of miR-393, miR-319, and miR-397 is increased in response to dehydrations (Sunkar and Zhu, 2004). Several other miRNAs have been revealed to play a role in drought stress in *Arabidopsis* including miR-157, miR-167, miR-168, miR-171, miR-408, miR-393, miR-396 (Liu et al., 2008). In rice, MIR-169g, MIR-171a and MIR393 are up-regulated in response to dehydration (Zhao et al., 2007; Jian et al., 2010; Zhou et al., 2010). In *Medicago truncatula* miR-398a/b increases in accumulation following water deficit conditions (Trindade et al., 2010). In *Populus trichocarpa* miR-1446a-e, miR-1444a, miR-1447 and miR-1450 were found to be significantly down-regulated while the expression level of miR-1711a, miR-482.2, miR-530a, miR-827a, miR-1445a, and miR-1448 were only slightly decreased during drought stress (Lu et al. 2008). miR-393 was consistently up-regulated during drought stress in rice, *Arabidopsis*, *M. truncatula*, and *P. vulgaris*. However, the effect on the expression
of its target genes TIR1/AFB2 Auxin Receptors during drought is unkown (Zhao et al., 2007; Liu et al., 2008; Arenas-Huertero et al., 2009; Jian et al., 2010; Trindade et al., 2010).

Low temperatures trigger physiological disorder and structural lesions in plants and depend on the duration of exposure to low temperatures (acute or chronic) and on the rate of temperature decrease. Thermal shocks generated through sudden temperature decrease are more harmful to plant cells. Several miRNAs have been revealed to be up-regulated under cold conditions in Arabidopsis including miR-172, miR-393, miR-397b, miR-402, miR-319c, miR-165/166, miR-169, miR-396, miR-397, miR-408, miR-156/157, miR-159/319 (Sunkar and Zhu, 2004; Zhou et al., 2008). MiR-319c was reported in the case of cold stress, but not under other stresses such as dehydration or high salt concentration. MiR-319 regulates the mRNA levels of TCP transcription factors (a family of transcription factors that play various roles in plant development) (Palatnik et al., 2007). Although it is not clear what the role of these genes is in response to cold, some speculate that this regulation is part of a complex response leading to synthesis of protective molecules and to a strong reduction of plant growth. In rice, 18 cold-responsive miRNAs have been identified (miR-156k, miR-166k, miR-166m, miR-167a/b/c, miR-168b, miR-169e, miR-169f, miR-169h, miR-171a, miR-535, miR-319a/b, miR-1884b, miR-444a.1, miR-1850, miR-1868, miR-1320, miR-1435, and miR-1876) (Lv et al., 2010). Mir-169 expression may serve to inhibit cell wall loosening and thus contribute with other miRNAs like miR-393 to repress plant growth under low temperatures (Zhou et al., 2008).

High salt concentrations in soil limits the plant’s ability to uptake water and hinders worldwide agricultural production. Salt stress in plants elicits similar metabolic and cellular processes as drought stress (Munns, 2005). However, genes and pathways of plants are differentially regulated between salt and drought stress (Bartels and Sunkar, 2005; Golldack et
Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

al., 2011). In rice, three salt-inducible members of the miR-169 family (miR-169g/n/o) as well as miR-393 have been identified (Zhao et al., 2009; Gao et al., 2011). These miRNAs selectively cleaved the transcript of the NF-YA gene, the subunit of the NF-Y complex, an evolutionarily conserved transcription factor present in a wide range of organisms from yeast to humans (Kabe et al, 2005). Recent microarray experiments using salt-tolerant and salt-sensitive lines of *Zea mays* identified 98 miRNAs from 27 families showing significant changes in their expression after salt treatment (Ding et al., 2009). In *Arabidopsis*, miR-168 was among the salt stress-responsive miRNA expressed and correspondingly AGO1 mRNA levels were decreased (Ding et al, 2009). AGO1 is the main effector protein required for miRNA function in plants and it will be important to determine whether alteration of miR-168 function leads to general changes in miRNA functions.

Sun UV radiation is another major stressor for plants. Most UV-B (280-320 nm) is absorbed by the ozone layer; UV-A (320-400 nm) penetrates the atmosphere to reach the surface of earth. The depletion of the stratospheric ozone layer is known to negatively affect plant growth and development (due to elevated UV-B) that can accelerate the generation and accumulation of ROS (McKenzie et al., 2007, Zhou et al., 2007). 24 putative light-response miRNAs falling into 11 families were predicted in *Arabidopsis* (miR-156, miR-159, miR-160, miR-165/166, miR-167, miR-169, miR-170, miR-172, miR-393, miR-398, and miR-401) (Zhou et al., 2007). Recently, 24 UV-B stress-responsive miRNAs were identified in *Populus tremula* (Jia et al., 2009). Several miRNAs such as miR-168, miR-395, and miR-398 are involved in response to other stresses and might regulate important homeostatic processes including miRNA-AGO1 system, energy related metabolism, and ROS inactivation. The role of these various
miRNAs in response to UV remains to be determined and it will be important to resolve whether UV response regulatory mechanisms are conserved between different plant species.

MiRNAs are involved in nutrient homeostasis and other abiotic stresses in plants, yet they also play an important role in defending against viruses, bacteria, fungi, insects, and nematodes. RNA silencing in defense against viruses was unraveled several years ago, while the involvement of miRNA regulation in protection against bacterial pathogens has only recently emerged (Navarro et al., 2006). Perception of flagellin (the principal substituent of bacterial flagellum) is crucial for plant resistance to *Pseudomonas syringae* bacterium (Gómez-Gómez and Boller, 2002). The miRNA miR-393 seems to be important for the process of flagellin recognition (Navarro et al., 2006). A group of bacteria-regulated miRNAs that target genes encoding proteins of the auxin, abscisic acid, and jasmonic acid (plant hormones) biosynthetic and/or signalizing pathways were identified and miR-160, miR-167, miR-393, miR-159 were shown to be down-regulated and their respective targets accumulated upon *Pseudomonas* infection (Zhang et al., 2011). These results indicate that miRNA may be important for plant defense signaling by regulating and fine-tuning multiple plant hormone pathways. MiRNAs responsive to plant infection by viruses were identified in many plants species including *Brassica rapa* (He et al, 2008), rice (Du et al., 2011), *Arabidopsis* (Bazzini et al., 2009; Blevins et al., 2011; Hu et al., 2011) and tomato (Lang et al.,2011). It is believed that plants use the general RNA silencing machinery to degrade viral RNAs or target viral DNAs for methylation (Pantaleo, 2011; Hohn and Vazquez, 2011).

It is clear that miRNAs are heavily involved in plant responses to biotic and abiotic stresses. *Table 1* summarizes the miRNAs discussed above and their expression levels. The role these various miRNAs play is an extremely interesting area of research and holds promise for
future understanding of the diversity and specificities of plant stress responses. Further investigation may yield the production of on-demand plants that are bred or genetically engineered to resist or quickly acclimate to stressors, increasing global food yield and efficiency of our resources. At the same time miRNA mediated pathways in plants hint at the importance of RNA silencing in all eukaryotes and will further the understanding of earth’s earliest life forms and the small hereditary molecules that allowed them to exist.
### Table 1: Summary of miRNA expression changes in different plant species.

<table>
<thead>
<tr>
<th>miRNA</th>
<th>Sulfur Starvation</th>
<th>Copper Starvation</th>
<th>Oxidative Stress</th>
<th>Cold Stress</th>
<th>Salt</th>
<th>Drought</th>
<th>UV-B</th>
<th>Biotic Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>395</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>398</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>408</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>397</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>393</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>319</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>165</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>166</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>169</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>172</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>396</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>156</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>157</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>164</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>394</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>167</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>168</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>171</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>159</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>160</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
</tbody>
</table>

**RED** – Predicted in *C. richardii* ESTs
1.6  

A brief chronicle of bioinformatics

We have come a long way since the days of Jensen and Evans who positioned a single amino acid (a terminal phenylalanine) in insulin (Jensen and Evans, 1935), Sanger who elucidated its complete sequence, the first of any protein (Sanger et al., 1955), Franklin and Gosling’s foundational work towards the elucidation of the structure of DNA (Franklin and Gosling, 1953) and Watson and Crick who opportunistically interpreted their data (Watson and Crick, 1953). The era of single amino acid or nucleic acid elucidation is long gone and it is evident bioinformatics plays a central role in modern experimental and theoretical biology. While many important databases and institutions have been necessarily omitted, it is clear that the ones discussed have a global reach and provide freely available easy-to-use software tools with which to analyze the data and draw meaningfull conclusions. The future of bioinformatics is broad and the disciplines borders are becoming all-encompassing with the majority of emerging research being driven almost entirely through computational predictions and high-throughput methods of experimental validation.

Both the term and discipline of bioinformatics emerged in the 1970’s when the Dutch theoretical biologist Paulien Hogeweg described her main field of research as bioinformatics, and established a bioinformatics group at the University of Utretch (Hogeweg, 1978; Howeseg and Hesper, 1978). The discipline itself ultimately grew out of the need for researchers to access and analyze biomedical data, which with the advent of modern biotechnology and the spur of computer processing, was accumulating rapidly. The enormous technological leaps that yielded unprecedented amounts of biological sequence data led to wide-scale applications of computational resources and algorithms for the analysis, manipulation, and storage of these growing quantities of data. The recruitment of computers was the key to systematically analyze
Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

and store the accumulating sequence and structure data. Several organizations and infrastructural initiatives have emerged over the last 50 years and their growth into vast repositories of biological and chemical data has changed the way research is conducted today.

1.7 *Real-time quantitative PCR*

While bioinformatics is imperative to investigating the role of miRNAs in life systems, the computational predictions made still need to be experimentally validated to gain merit. In 1987 Nobel Prize laureate Kary Mullis patented his method (Mullis, 1985) for the amplification of nucleotides or the polymerase chain reaction (PCR). Mullis’ discovery forever changed molecular biology. Some have compared the advent of PCR to that of the Internet,

> “Both inventions have emerged in the last 20 years to the point where it is difficult to imagine life without them. Both have grown far beyond the confines of their original simple design and have created opportunities unimaginable before their invention. Both have also spawned a whole new vocabulary and professionals literate in that vocabulary.” (Bartlett and Stirling, 2003).

In the most basic sense, PCR is a chemical reaction that amplifies a target DNA sequence. The reaction requires several reagents including the DNA template to be amplified, target specific oligonucleotides called DNA primers (required for initiation of DNA synthesis), certain ions (magnesium, manganese, potassium) necessary for DNA synthesis usually that incorporated into a buffer solution, and deoxynucleoside triphosphates (dNTPs) that will get incorporated into the new DNA. A special enzyme called Taq polymerase (DNA polymerase) from the organism *Thermus aquaticus* (a species of bacterium that thrives at 70 °C) enables the reaction to proceed through what is known as thermal cycling. Initially the reaction mixture is kept cool while reagents are loaded into PCR tubes or wells. Once the wells are loaded and
Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

capped, they can be placed into the thermal cycler or PCR machine that heats and cools the wells to specific temperatures to allow for annealing of primers to the target sequence, synthesis of the complimentary sequences by Taq polymerase, followed by denaturing the newly synthesized DNAs to allow the process to start over again. As this is a chemical reaction in a closed system, eventually the limiting reagents will be consumed. However, in the first 25-30 cycles exponential amplification is observed (assuming 100% reaction efficiency) and minute quantities of DNA can be doubled with each cycle to yield millions or billions of copies. Eventually the reaction begins to level off as Taq polymerase loses activity and reagents are consumed until a plateau is reached and no more product accumulates. PCR product can then be analyzed by gel electrophoresis and compared to a DNA ladder (DNA sequences of different lengths) to determine if the amplified gene is the right number of base pairs (if it is amplified at all).

The original PCR reaction has been morphed into many other techniques that rely on this basic principle yet alter it in some way to achieve more specific expression profiling for various downstream applications. Real-time quantitative PCR (qPCR) is a technique that uses fluorescent dyes (fluorophore) linked to the PCR product that can be detected by a camera in the thermal cycler to measure the amount of product in each well (Heid et al., 1996). A common fluorophore is SYBR Green I which is a dye that has a high affinity for double-stranded DNA (dsDNA) (Simpson et al., 2000). The amount of fluorescence is indicative of the amount of dsDNA present in each well, and the only dsDNA present is the one that was amplified by the target primers. Theoretically a well with more starting DNA target should have a higher fluorescence within fewer cycles than a well with less starting DNA target. Several mathematical models have emerged and allow for relative, absolute, and normalized quantification of qPCR product.
1.8 Focus

Gametophytes of the organism *Ceratopteris richardii* displayed characteristics of the loss of cell-cell communication following treatment with hypertonic solution. The cells exhibited an aberrant morphology as a consequence of plasmolysis and displayed characteristics indicative of the loss of cell-cell adhesions and the formation of the “secondary gametophyte.” MiRNAs were thought to play a role in this reprogramming and in this thesis we used complex computational methods to predict miRNAs in the Expressed Sequence Tags (ESTs; partial cDNA sequences of expressed genes cloned into a plasmid) (Adams et al., 1991; Matukumalli et al., 2004) available for the organism *C. richardii*. The aim of the project was to determine which miRNAs are expressed in *C. richardii* and validate them with qPCR. We were also interested in the expression levels of miRNAs implicated in human cancer progression (oncomiRs) as the reprogramming observed in *C. richardii* after plasmolysis is similar to that observed in human cancer cells prior to and during metastasis (Kharonov and Tucker, unpublished data).

Furthermore we were interested in the expression levels of several miRNAs that are known to be regulators of key cell cycle proteins that may be contributing to the development of the secondary gametophytes observed post-plasmolysis. This study will determine putative miRNAs that are conserved in *C. richardii* and will provide insights into their biogenesis, nucleotide sequence, and target interactions.
CHAPTER 2
METHODS

2.1 MicroRNA prediction

Computational predictions were carried out using a local instance of the Galaxy server (https://main.g2.bx.psu.edu/) with National Center for Biotechnology Information (NCBI) tools. A total of 10,102 mature miRNAs and 10,058 pre-miRNAs were obtained from miRbase volume 18 (http://www.mirbase.org; Griffiths-Jones, 2004). A total of 6,225 ESTs were obtained from NCBI (http://www.ncbi.nlm.nih.gov/nucest). MiRNAs and pre-miRNAs were matched against the ESTs with the Basic Local Alignment Search Tool Nucleotide (BLASTn 2.2.26; NCBI). BLASTn parameters were set to: expect 1000; the number of description and alignments was set to 500. The word-size between a query and a database sequence was set to 7. EST sequences that were closely matched (mismatch <4), had a high percent identity (>85%), few gaps (<2) and with an Expect value less than or equal to e-6 and e-12 for miRNA and pre-miRNA, respectively, were selected. Predictions that passed selection criteria were BLAST Translated (BLASTx 2.2.26) against all known plant proteins on UniProt (http://www.uniprot.org; UniProt Consortium, 2011). Non-coding ESTs were selected with an Expect value less than e-5 and percent identity less than 25. Secondary structures can be predicted from microRNAs within non-coding ESTs using default RNA 3.0 folding parameters for Mfold on the DINA-Melt Server (Markham and Zuker, 2005), but were not collected. Predicted Target genes for the putative miRNAs were identified using default parameters on the psRNATarget server (Dai et al, 2010). Validated miRNA targets were retrieved from DIANA-microT web server (http://www.diana.cslab.ece.ntua.gr/microT.com; Maragkakis et al., 2009). Figure 5 is a schematic of the described workflow.
Figure 5

A simplified workflow of miRNA prediction from Expressed Sequence Tags.
2.2  *Ceratopteris richardii* culturing

Wild type *C. richardii* spores (RNWT1) were obtained from Carolina Biological Supply (North Carolina) and sown on macronutrient medium (Basile, 1978). Spores were incubated at 22 C° with 24hr light and adequate moisture. Spores were allowed to develop into 8 and 12-day-old gametophytes before plasmolysis.

2.3  *Plasmolysis*

Gametophytes were plasmolyzed in 200μL of 1.0 M mannitol solution in 1.5mL tubes for 30 minutes after which they rehydrated with 500μL of distilled water and transferred back into a sterile petri dish with macronutrient agar. Rehydrated gametophytes were placed back into the incubation rack and allowed to continue their development for 24, 48, and 72 hour periods before they were sacrificed.

2.4  RNA extraction

The RNAqueous®-4PCR Kit (Ambion® Life Technologies) was used to extract and purify total RNA from control, 24, 48, and 72 hour post-plasmolysis cells from both 8 and 12 day-old gametophytes. Cells were transferred into 350μL of cold Lysis/Binding Solution in 1.5mL tubes and disrupted for 10 minutes with a micro-pestle and vigorous vortexing. 350μL of 64% ethanol was added to the lysate and gently but thoroughly mixed in by vortexing. The 700μL lysate/ethanol mixture was transferred into a filter cartridge inside a collection tube and centrifuged for 30 seconds at 10,000 rpm. The flow through was discarded and 700μL of Wash Solution 1 was added to the filter cartridge and centrifuged for 15 seconds at 10,000 rpm. The procedure was repeated twice with 500μL of Wash Solution 2 and centrifuged for 15 seconds at
10,000 rpm. The filter cartridge was replaced into the collection tube and centrifuged for an addition 30 seconds at 10,000 rpm to remove residual wash fluid. The used collection tube with the flow through was discarded and the filter cartridge was placed into a fresh collection tube. 100μL of 80°C Elution Solution was added to the filter cartridge in aliquots of 60μL and 40μL and centrifuged for 30 seconds after the addition of each aliquot at 10,000 rpm to collect the RNA from the filter. The filter cartridge was discarded and the RNA was treated with 1μL of DNase I with 10μL of 10x DNase I Buffer and incubated at 37°C for 30 minutes to remove DNA contamination. 11μL of DNase Inactivation Reagent were added to the RNA and mixed gently at room temperature for 2 minutes. The tube was centrifuged for 1 minute at 10,000 rpm to pellet the DNase Inactivation Reagent and the supernatant was transferred into a sterile 1.5mL tube. RNA quality and quantity was measured with a NanoDrop® 1000 (Thermo Scientific). The RNA was stored at -20°C.

### 2.5 Poly(A) Tailing and cDNA synthesis

The NCode™ EXPRESS SYBR® GreenER™ miRNA qRT-PCR Kit Universal (Invitrogen™, Life Technologies) was used to synthesize cDNA from the RNA extract. The following components were combined in a tube on ice: 4μL 5X Reaction Mix, 2μL 10X SuperScript® Enzyme Mix, 2μL of RNA, and 12μL of DEPC-treated DNase/RNase free water. The tube was capped, gently vortexed, and centrifuged for 15 seconds at 10,000 rpm to collect the contents. The tube was then incubated at 37°C for 60 minutes. The reaction was terminated at 95°C for 5 minutes. The cDNA was stored at -20°C.
2.6 *qPCR protocol*

The NCode™ EXPRESS SYBR® GreenER™ miRNA qRT-PCR Kit Universal (Invitrogen™, Life Technologies) was used for qPCR with EXPRESS SYBR Green. The following components were added to each well: 10μL of EXPRESS SYBR® GreenER™ qPCR SuperMix Universal, 0.4μL of 10µM miRNA-specific forward primer (200nM final), 0.4μL of 10µM Universal qPCR Primer (200nM final), 2μL of cDNA (20μg/μL), and 7.2μL of DEPC-treated DNase/RNase free water. The total volume for each reaction well was 20μL. The reference gene GAPDH was amplified with its own mRNA forward specific primer with all other components equal. The 96-well plates were sealed with translucent film and placed inside the MyiQ™2 Two-Color Real-Time PCR Detection System (Bio-Rad). The cycling program was set to 50°C for 2 minutes (UDG incubation), 95°C for 2 minutes, 45 cycles of 95°C for 15 seconds and 60°C for 1 minutes (with a gradient of +12°C), melting curve analysis from 60°C-95°C, and a hold step at 4°C. Data was analyzed according to established mathematical model (Pfaffi, 2001).
CHAPTER 3
RESULTS

3.1 Predictions

Nine unique pre-miRNAs were predicted from a total of 327,172 EST alignments. Eighty-nine unique mature miRNAs were predicted from a total of 76,940 EST alignments. Pre-miRNAs had an Expect value less than e-12 while mature miRNAs had an Expect value less than e-6. Table 2 summarizes pre-miRNA predictions and Table 3 summarizes mature miRNA predictions.
Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

<table>
<thead>
<tr>
<th>Query id</th>
<th>Subject id</th>
<th>% Identity</th>
<th>A. Length</th>
<th>M. Match</th>
<th>G. Opening</th>
<th>Q. Start</th>
<th>Q. End</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>55452087 at-MIR172</td>
<td>100</td>
<td>41</td>
<td>0</td>
<td>0</td>
<td>293</td>
<td>333</td>
<td>4.00E-18</td>
</tr>
<tr>
<td>gi</td>
<td>55452087 sb-MIR396c</td>
<td>100</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>286</td>
<td>323</td>
<td>1.00E-16</td>
</tr>
<tr>
<td>gi</td>
<td>20046501 tae-MIR160</td>
<td>100</td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>182</td>
<td>213</td>
<td>2.00E-13</td>
</tr>
<tr>
<td>gi</td>
<td>55451392 mmu-mir-466q</td>
<td>97.37</td>
<td>38</td>
<td>1</td>
<td>0</td>
<td>204</td>
<td>241</td>
<td>1.00E-15</td>
</tr>
<tr>
<td>gi</td>
<td>20046793 smo-MIR1082a</td>
<td>97.3</td>
<td>37</td>
<td>1</td>
<td>0</td>
<td>112</td>
<td>148</td>
<td>2.00E-14</td>
</tr>
<tr>
<td>gi</td>
<td>9960040 sha-mir-716b</td>
<td>96.61</td>
<td>59</td>
<td>2</td>
<td>0</td>
<td>723</td>
<td>781</td>
<td>3.00E-25</td>
</tr>
<tr>
<td>gi</td>
<td>9959882 bna-MIR166b</td>
<td>94.87</td>
<td>39</td>
<td>2</td>
<td>0</td>
<td>573</td>
<td>611</td>
<td>1.00E-14</td>
</tr>
<tr>
<td>gi</td>
<td>55452655 ath-MIR5645d</td>
<td>94.74</td>
<td>38</td>
<td>2</td>
<td>0</td>
<td>376</td>
<td>413</td>
<td>2.00E-13</td>
</tr>
<tr>
<td>gi</td>
<td>55450631 aly-MIR408</td>
<td>90.7</td>
<td>43</td>
<td>3</td>
<td>1</td>
<td>250</td>
<td>292</td>
<td>9.00E-13</td>
</tr>
</tbody>
</table>

**Table 2:** 9 unique pre-microRNA predictions and their EST origins.

Query id: The accession number for NCBI EST.

Subject id: The accession number for mirBase.

% Identity: The percent of matches that are identical between the query and the subject.

A. Length: The total length of the alignment in nucleotides.

M. Match: The number of mismatches between the query and the subject in nucleotides.

G. Opening: The number of gap openings in the alignment.

Q. Start: The number of the first nucleotide of the query in the alignment.

S. Start: The number of the first nucleotide of the subject in the alignment.

E-value: The Expect value establishes a statistical significance threshold for reporting database sequence matches. The lower the E-value, the less likely it is for the alignment to occur by chance.
### Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

<table>
<thead>
<tr>
<th>Query id</th>
<th>Subject id</th>
<th>% Identity</th>
<th>A. Length</th>
<th>M. Match</th>
<th>G. Opening</th>
<th>Q. Start</th>
<th>Q. End</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>9959029</td>
<td>mmu-miR-466i-5p</td>
<td>100</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>150</td>
<td>167</td>
</tr>
<tr>
<td>gi</td>
<td>9959029</td>
<td>mmu-miR-466i-5p</td>
<td>100</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>150</td>
<td>167</td>
</tr>
<tr>
<td>gi</td>
<td>9959029</td>
<td>mmu-miR-466i-3p</td>
<td>100</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>232</td>
<td>249</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>mmu-miR-466j</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>85</td>
<td>100</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>mmu-miR-669m-5p</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>85</td>
<td>100</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>mmu-miR-466m-5p</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>85</td>
<td>100</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>oan-miR-1419d</td>
<td>100</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>399</td>
<td>415</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>ath-miR5658</td>
<td>100</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>125</td>
<td>141</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>ath-miR414</td>
<td>100</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>313</td>
<td>329</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>ath-miR853</td>
<td>100</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>437</td>
<td>453</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>osa-miR808</td>
<td>100</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>44</td>
<td>60</td>
</tr>
<tr>
<td>gi</td>
<td>9958847</td>
<td>ath-miR5021</td>
<td>100</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>196</td>
<td>212</td>
</tr>
<tr>
<td>gi</td>
<td>9958690</td>
<td>bmo-miR-3233</td>
<td>100</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>560</td>
<td>576</td>
</tr>
<tr>
<td>gi</td>
<td>9958615</td>
<td>bta-miR-2293</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>126</td>
<td>141</td>
</tr>
<tr>
<td>gi</td>
<td>9958611</td>
<td>bta-miR-2365</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>168</td>
<td>183</td>
</tr>
<tr>
<td>gi</td>
<td>9960933</td>
<td>mmu-miR-669f-5p</td>
<td>100</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>86</td>
<td>100</td>
</tr>
<tr>
<td>gi</td>
<td>9960550</td>
<td>cel-miR-5550-5p</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>210</td>
<td>225</td>
</tr>
<tr>
<td>gi</td>
<td>9960041</td>
<td>hsa-miR-105-3p</td>
<td>95.24</td>
<td>21</td>
<td>0</td>
<td>1</td>
<td>515</td>
<td>535</td>
</tr>
<tr>
<td>gi</td>
<td>9959990</td>
<td>bta-miR-424</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>266</td>
<td>281</td>
</tr>
<tr>
<td>gi</td>
<td>9959949</td>
<td>cel-miR-1819-3p</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>74</td>
<td>89</td>
</tr>
<tr>
<td>gi</td>
<td>9959846</td>
<td>hsa-miR-3688-5p</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>255</td>
<td>270</td>
</tr>
<tr>
<td>gi</td>
<td>9959710</td>
<td>ath-miR5021</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>157</td>
<td>172</td>
</tr>
</tbody>
</table>
Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

<table>
<thead>
<tr>
<th>ID</th>
<th>MiRNA</th>
<th>%</th>
<th>Fold</th>
<th>Prim</th>
<th>Amp</th>
<th>Size 1</th>
<th>Size 2</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>9959624</td>
<td>mml-miR-944</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>540</td>
<td>555</td>
</tr>
<tr>
<td>gi</td>
<td>9959616</td>
<td>181b-3p hsa-miR-424-5p</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>482</td>
<td>497</td>
</tr>
<tr>
<td>gi</td>
<td>9958966</td>
<td>mml-miR-424 ptr-miR-424</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>266</td>
<td>281</td>
</tr>
<tr>
<td>gi</td>
<td>9958966</td>
<td>eca-miR-424 ssc-miR-424 ppy-miR-424</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>266</td>
<td>281</td>
</tr>
<tr>
<td>gi</td>
<td>9958966</td>
<td>mtr-miR2590 g</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>320</td>
<td>335</td>
</tr>
<tr>
<td>gi</td>
<td>9958886</td>
<td>gga-miR-1758 aly-miR838 mmu-miR541-3p</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>255</td>
<td>270</td>
</tr>
<tr>
<td>gi</td>
<td>9958837</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>449</td>
<td>464</td>
<td>8.00E-06</td>
</tr>
<tr>
<td>gi</td>
<td>9958837</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>197</td>
<td>212</td>
<td>8.00E-06</td>
</tr>
<tr>
<td>gi</td>
<td>9958705</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>573</td>
<td>588</td>
<td>8.00E-06</td>
</tr>
<tr>
<td>gi</td>
<td>9958668</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>518</td>
<td>533</td>
<td>8.00E-06</td>
</tr>
<tr>
<td>gi</td>
<td>9958663</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>411</td>
<td>426</td>
<td>8.00E-06</td>
</tr>
<tr>
<td>gi</td>
<td>995868</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>113</td>
<td>128</td>
<td>8.00E-06</td>
</tr>
<tr>
<td>gi</td>
<td>9958431</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>248</td>
<td>263</td>
<td>9.00E-06</td>
</tr>
<tr>
<td>gi</td>
<td>9958421</td>
<td>100</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>525</td>
<td>540</td>
<td>9.00E-06</td>
</tr>
<tr>
<td>gi</td>
<td>9959758</td>
<td>95.24</td>
<td>21</td>
<td>1</td>
<td>0</td>
<td>229</td>
<td>249</td>
<td>4.00E-07</td>
</tr>
<tr>
<td>gi</td>
<td>9959075</td>
<td>mmu-miR-1187</td>
<td>95.24</td>
<td>21</td>
<td>1</td>
<td>0</td>
<td>224</td>
<td>244</td>
</tr>
</tbody>
</table>

38
<table>
<thead>
<tr>
<th>ID</th>
<th>Accession</th>
<th>Rank</th>
<th>miR Name</th>
<th>miRNA Name</th>
<th>Value</th>
<th>Error</th>
<th>Log10(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>9958977</td>
<td>osa-miR169i.2</td>
<td>95</td>
<td>20</td>
<td>1</td>
<td>0</td>
<td>2024</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>hsa-miR-372</td>
<td>95</td>
<td>20</td>
<td>1</td>
<td>0</td>
<td>649</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>mml-miR-372</td>
<td>95</td>
<td>20</td>
<td>1</td>
<td>0</td>
<td>649</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>ptr-miR-372</td>
<td>95</td>
<td>20</td>
<td>1</td>
<td>0</td>
<td>649</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>ppy-miR-372</td>
<td>95</td>
<td>20</td>
<td>1</td>
<td>0</td>
<td>649</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>rno-miR-466b-1*</td>
<td>95</td>
<td>20</td>
<td>1</td>
<td>0</td>
<td>228</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>tca-miR-219-3p</td>
<td>95</td>
<td>20</td>
<td>1</td>
<td>0</td>
<td>569</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>bmo-miR-3329</td>
<td>95</td>
<td>20</td>
<td>1</td>
<td>0</td>
<td>532</td>
</tr>
<tr>
<td>gi</td>
<td>9958977</td>
<td>sme-miR-2201*</td>
<td>91.67</td>
<td>24</td>
<td>1</td>
<td>1</td>
<td>676</td>
</tr>
<tr>
<td>gi</td>
<td>9958690</td>
<td>bta-miR-2285b</td>
<td>94.44</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td>43</td>
</tr>
<tr>
<td>gi</td>
<td>9960893</td>
<td>mmu-miR-1187</td>
<td>94.74</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>78</td>
</tr>
<tr>
<td>gi</td>
<td>9960893</td>
<td>ath-miR5658</td>
<td>94.74</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>222</td>
</tr>
<tr>
<td>gi</td>
<td>9960696</td>
<td>gma-miR4415a-5p</td>
<td>94.74</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>339</td>
</tr>
<tr>
<td>gi</td>
<td>9960346</td>
<td>bmo-miR-3314</td>
<td>94.74</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>86</td>
</tr>
<tr>
<td>gi</td>
<td>9959196</td>
<td>pma-miR-190b*</td>
<td>94.74</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>177</td>
</tr>
<tr>
<td>gi</td>
<td>9959001</td>
<td>ppt-miR902b-5p</td>
<td>94.74</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>705</td>
</tr>
<tr>
<td>gi</td>
<td>9958989</td>
<td>ppt-miR902c-5p</td>
<td>94.74</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>705</td>
</tr>
<tr>
<td>gi</td>
<td>9958966</td>
<td>ppt-miR902d-5p</td>
<td>94.74</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>705</td>
</tr>
<tr>
<td>gi</td>
<td>9958966</td>
<td>ppt-miR902e-5p</td>
<td>94.74</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>705</td>
</tr>
<tr>
<td>gi</td>
<td>9958966</td>
<td>ppt-miR902f-5p</td>
<td>94.74</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>705</td>
</tr>
<tr>
<td>gi</td>
<td>9958966</td>
<td>ppt-miR902g-5p</td>
<td>94.74</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>705</td>
</tr>
</tbody>
</table>
### Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

<table>
<thead>
<tr>
<th>ID</th>
<th>miRNA</th>
<th>EST 1</th>
<th>EST 2</th>
<th>EST 3</th>
<th>EST 4</th>
<th>EST 5</th>
<th>EST 6</th>
<th>EST 7</th>
<th>EST 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi9958966</td>
<td>ebv-miR-BART8</td>
<td>94.74</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>502</td>
<td>520</td>
<td>8.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9958476</td>
<td>dme-miR-6-2-5p</td>
<td>94.74</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>606</td>
<td>624</td>
<td>9.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9958444</td>
<td>dsi-miR-6*</td>
<td>94.74</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>606</td>
<td>624</td>
<td>9.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9959758</td>
<td>hsa-miR-2916</td>
<td>91.3</td>
<td>23</td>
<td>2</td>
<td>0</td>
<td>146</td>
<td>168</td>
<td>5.00E-07</td>
<td></td>
</tr>
<tr>
<td>gi9959523</td>
<td>mmu-miR-574-5p</td>
<td>91.3</td>
<td>23</td>
<td>2</td>
<td>0</td>
<td>220</td>
<td>242</td>
<td>5.00E-07</td>
<td></td>
</tr>
<tr>
<td>gi9959250</td>
<td>ath-miR414</td>
<td>91.3</td>
<td>23</td>
<td>2</td>
<td>0</td>
<td>220</td>
<td>242</td>
<td>5.00E-07</td>
<td></td>
</tr>
<tr>
<td>gi9958977</td>
<td>mmu-miR-466f-3p</td>
<td>90.48</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>197</td>
<td>217</td>
<td>5.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9958574</td>
<td>ath-miR414</td>
<td>90.48</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>292</td>
<td>312</td>
<td>6.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9960078</td>
<td>gga-miR-1588</td>
<td>90.48</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>454</td>
<td>474</td>
<td>7.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9960442</td>
<td>ath-miR5658</td>
<td>90.48</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>59</td>
<td>79</td>
<td>7.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9960442</td>
<td>sme-miR-748*</td>
<td>90.48</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>274</td>
<td>294</td>
<td>7.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9960401</td>
<td>dps-miR-2525</td>
<td>90.48</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>37</td>
<td>57</td>
<td>7.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9959523</td>
<td>sme-miR-2166</td>
<td>90.48</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>297</td>
<td>317</td>
<td>8.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9959523</td>
<td>sme-miR-5388</td>
<td>90.48</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>454</td>
<td>474</td>
<td>8.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9959410</td>
<td>mmu-miR-1187</td>
<td>90.48</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>253</td>
<td>273</td>
<td>8.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9959401</td>
<td>oan-miR-1421ab</td>
<td>90.48</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>150</td>
<td>170</td>
<td>8.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9959304</td>
<td>cin-miR-4123-5p</td>
<td>90.48</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>176</td>
<td>196</td>
<td>8.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9959232</td>
<td>osa-miR5144-3p</td>
<td>90.48</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>59</td>
<td>79</td>
<td>8.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9959232</td>
<td>bta-miR-2416</td>
<td>90.48</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>281</td>
<td>301</td>
<td>9.00E-06</td>
<td></td>
</tr>
<tr>
<td>gi9958495</td>
<td>sme-miR-133a*</td>
<td>90.48</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>370</td>
<td>390</td>
<td>9.00E-06</td>
<td></td>
</tr>
</tbody>
</table>

**Table 3:** 89 microRNA predictions and their EST origins.
3.2 Effect of plasmolysis

The gametophyte that developed after plasmolysis was different than anything described in the development for *C. richardii*. Thirty days post-plasmolysis the gametophyte, lacking sexually differentiated organs (antheridia or archegonia), developed a lettuce leaf-like morphology. This abnormally formed gametophyte had gametophyte-like projections emanating from its base in all directions. *Figure 6* exhibits the difference between a typical gametophyte and one that has been plasmolyzed.

*Figure 6*

The normal gametophyte (left) is an obliquely heart-shaped hermaphrodite (10 to 11-days-old), 2-3 mm in diameter, possessing both antheridia and archegonia. The secondary gametophyte (right) was plasmolzyed at 10 days, rehydrated, and allowed to continue growing for an addition 20 days. The secondary gametophyte (30 to 31-days old) is vastly different from anything that has been observed in ferns to this day. After plasmolysis each cell was separated from adjacent cells. Gene regulation may have turned on a pathway that promotes the formation of a new gametophyte from a single cell. The result is a lettuce-leaf-like secondary gametophyte with primary gametophyte projections emanating from its base in all directions.
3.3 \textit{qPCR}

Quantitative real-time PCR was performed on both 8 and 12-day-old gametophytes, tracking gene expression through four conditions (control, 24, 48, and 72 hours post-plasmolysis) using miRNA specific forward primers found in Table 4. Both relative and normalized expression values were calculated according to methods detailed by Pfaffi, however only relative expression values were graphed. The results for 8-day-old gametophytes are displayed in Table 5 and the results for 12-day-old gametophytes are displayed in Table 6. Figure 6 summarizes the results for 8-day-old and 12-day-old gametophytes, respectively.

Normalized expression was calculated with the \( \Delta \Delta C_t \) method using

\[
E^{-2(\Delta C_t \text{ Reference} - \Delta C_t \text{ Unknown})}
\]

where \( E \) or Efficiency was set to 2 and \( \Delta C_t \) is the difference in cycle threshold between treatments. The normalized expression is therefore a way to quantize the expression level of an unknown against a gene that is expressed at a constant level throughout the treatment.

Relative expression was calculated with the \( \Delta C_t \) method using

\[
E^{-2(\text{Unknown Ct} - \text{Reference Ct})}
\]

where \( E \) was set to 2 and the difference between the Ct of the unknown and the reference.
Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

**Table 4:** Forward specific primers.

<table>
<thead>
<tr>
<th>Forward Primer</th>
<th>Sequence</th>
<th>Tm °C</th>
<th>Anhyd. Mol. Wgt.</th>
<th>Amount of Oligo</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH</td>
<td>5’-ACCACAGTCCATGCCATC-3’</td>
<td>58.65</td>
<td>5,991.0</td>
<td>1626.7pm</td>
</tr>
<tr>
<td>cre-MIR914</td>
<td>5’-GCC GGC GGC GCC GGA ATC-3’</td>
<td>68.40</td>
<td>5,526.6</td>
<td>24.0nm</td>
</tr>
<tr>
<td>cre-MIR-914</td>
<td>5’-GCG GGC GGC CGG ACG TTA CGG C-3’</td>
<td>72.00</td>
<td>6,818.4</td>
<td>30.5nm</td>
</tr>
<tr>
<td>miR-106b</td>
<td>5’-GCC GCG GTA AAG TGC TGA CAG TG-3’</td>
<td>62.80</td>
<td>7,169.7</td>
<td>30.1nm</td>
</tr>
<tr>
<td>MIR-106b</td>
<td>5’-GCG GGC GCC TGC CGG GCC TAA AG-3’</td>
<td>71.50</td>
<td>7,131.6</td>
<td>27.4nm</td>
</tr>
<tr>
<td>miR-181c</td>
<td>5’-GCC GCG GAA CAT TCA ACC TGT C-3’</td>
<td>61.20</td>
<td>6,720.4</td>
<td>32.9nm</td>
</tr>
<tr>
<td>MIR-181c</td>
<td>5’-GCC GCG GCG GAA AAT TTG CCA AG-3’</td>
<td>64.50</td>
<td>7,138.7</td>
<td>29.4nm</td>
</tr>
</tbody>
</table>

MIR – pre-microRNA

miR – microRNA

Tm °C – melting temperature

Anhyd. Mol. Wgt. – Anhydrous molecular weight

Amount of oligo- in 10 pico moles or nano moles per 10μg
### Table 5: qPCR results for 8-day-old gametophytes.

C – control

24h, 48h, 72h – hours post-plasmolysis
### Table 6: qPCR results for 12-day-old gametophytes.

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Replicate #</th>
<th>Mean Ct</th>
<th>Normalized Expression (ΔΔCt)</th>
<th>Relative Expression (ΔCt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH C</td>
<td>3</td>
<td>21.71</td>
<td>1.00E+00</td>
<td>1.00E+00</td>
</tr>
<tr>
<td>GAPDH 24h</td>
<td>3</td>
<td>22.4</td>
<td>1.00E+00</td>
<td>6.20E-01</td>
</tr>
<tr>
<td>GAPDH 48h</td>
<td>3</td>
<td>20.92</td>
<td>1.00E+00</td>
<td>1.73E+00</td>
</tr>
<tr>
<td>GAPDH 72h</td>
<td>3</td>
<td>21.08</td>
<td>1.00E+00</td>
<td>1.55E+00</td>
</tr>
<tr>
<td>miR-106b C</td>
<td>1</td>
<td>27.34</td>
<td>2.02E-02</td>
<td>2.02E-02</td>
</tr>
<tr>
<td>miR-106b 24h</td>
<td>1</td>
<td>29.87</td>
<td>-5.17E-01</td>
<td>5.64E-03</td>
</tr>
<tr>
<td>miR-106b 48h</td>
<td>1</td>
<td>22.58</td>
<td>2.79E+01</td>
<td>3.16E-01</td>
</tr>
<tr>
<td>miR-106b 72h</td>
<td>1</td>
<td>22.28</td>
<td>3.40E+01</td>
<td>4.35E-01</td>
</tr>
<tr>
<td>pre-miR-106b C</td>
<td>1</td>
<td>27.18</td>
<td>2.26E-02</td>
<td>2.26E-02</td>
</tr>
<tr>
<td>pre-miR-106b 24h</td>
<td>1</td>
<td>22.55</td>
<td>2.41E+01</td>
<td>9.01E-01</td>
</tr>
<tr>
<td>pre-miR-106b 48h</td>
<td>1</td>
<td>25.17</td>
<td>4.82E-00</td>
<td>5.26E-02</td>
</tr>
<tr>
<td>pre-miR-106b 72h</td>
<td>2</td>
<td>34.22</td>
<td>6.38E-01</td>
<td>1.11E-04</td>
</tr>
<tr>
<td>miR-181c C</td>
<td>1</td>
<td>25.74</td>
<td>6.12E-02</td>
<td>6.12E-02</td>
</tr>
<tr>
<td>miR-181c 24h</td>
<td>1</td>
<td>22.71</td>
<td>7.48E+00</td>
<td>8.07E-01</td>
</tr>
<tr>
<td>miR-181c 48h</td>
<td>1</td>
<td>23.8</td>
<td>4.63E+00</td>
<td>1.36E-01</td>
</tr>
<tr>
<td>miR-181c 72h</td>
<td>1</td>
<td>23.2</td>
<td>6.45E+00</td>
<td>2.30E-01</td>
</tr>
<tr>
<td>pre-miR-181c C</td>
<td>3</td>
<td>16.14</td>
<td>4.75E+01</td>
<td>4.75E+01</td>
</tr>
<tr>
<td>pre-miR-181c 24h</td>
<td>3</td>
<td>12.5</td>
<td>1.18E+01</td>
<td>9.55E+02</td>
</tr>
<tr>
<td>pre-miR-181c 48h</td>
<td>3</td>
<td>13.02</td>
<td>9.48E+00</td>
<td>2.39E-02</td>
</tr>
<tr>
<td>pre-miR-181c 72h</td>
<td>3</td>
<td>12.4</td>
<td>1.40E+01</td>
<td>4.10E+02</td>
</tr>
<tr>
<td>cre-miR-914C</td>
<td>3</td>
<td>17.08</td>
<td>2.48E+01</td>
<td>2.48E+01</td>
</tr>
<tr>
<td>cre-miR-914 24h</td>
<td>3</td>
<td>16.38</td>
<td>9.35E-01</td>
<td>6.49E+01</td>
</tr>
<tr>
<td>cre-miR-914 48h</td>
<td>3</td>
<td>16.75</td>
<td>2.05E+00</td>
<td>1.80E+01</td>
</tr>
<tr>
<td>cre-miR-914 72h</td>
<td>3</td>
<td>16.14</td>
<td>2.55E+00</td>
<td>3.07E+01</td>
</tr>
<tr>
<td>pre-cre-miR914 C</td>
<td>3</td>
<td>19.41</td>
<td>4.92E+00</td>
<td>4.92E+00</td>
</tr>
<tr>
<td>pre-cre-mir914 24h</td>
<td>3</td>
<td>20.74</td>
<td>-2.92E-01</td>
<td>3.16E+00</td>
</tr>
<tr>
<td>pre-cre-mir914 48h</td>
<td>3</td>
<td>20.37</td>
<td>1.30E+00</td>
<td>1.46E+00</td>
</tr>
<tr>
<td>pre-cre-mir914 72h</td>
<td>3</td>
<td>19.91</td>
<td>1.34E+00</td>
<td>2.25E+00</td>
</tr>
</tbody>
</table>

C – control

24h, 48h, 72h – hours post-plasmolysis.
Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

Figure 6: Relative level of microRNA expression after plasmolysis in 8-day-old (A-F) and 12-day-old (G-L) gametophytes over a 72 hour period.

**A-B:** Accumulation of miR and pre-miR106b.

**C-D:** Accumulation of miR181c and steady expression of pre-miR181c over 72 hours

**E-F:** Slight decrease of miR914 and steady pre-miR914 over 72 hours

**G-H:** Accumulation of miR106b after 48 hours, accumulation of pre-miR106b after 24 hours.

**I-J:** Accumulation of miR181c after 24 hours, colossal accumulation of pre-miR181c after 24 hours.

**K-J:** Large accumulation of miR-914 and slight loss of pre-miR914
4.1 *Conserved pre-miRNA*

EST alignments and target predictions suggest that miRNA are highly conserved between organisms within the plant kingdom and may have an orthologous function in the animal kingdom. The effector targets may have a conserved sequence as indicated by BLAST alignment and target predictions from psRNA target server and the DIANA server. These predictions indicate that miRNA have diverse functions in both plants and animals. The translational value from a plant to an animal may be limited, however it is believed that eukaryotes’ share a common ancestor, one that may have had these miRNA genes serving their own respective functions. 20 miRNAs were listed in *Table 1* with their expression in response to various stresses. Of these 20 miRNAs, 7 (miRNA-408, miRNA-166, miRNA-169, miRNA-172, miRNA-396, miRNA-156, miRNA-160) are predicted to exist in *C. richardii* from EST alignment. These miRNAs and others that are not found in the literature are discussed here.

Pre-miRNA-172 from *Aegilops tauschi* (Tausch’s goatgrass), a grass that is related to wheat, targets the APETALA2 (AP2) and AP2-like gene glossy 15 (gl15). AP2 and ethylene-responsive element binding proteins (EREBPs) are the prototypic members of a family of transcription factors unique to plants, which contain the AP2 DNA-binding domain. AP2/REBP genes form a large multigene family and have been shown to have a variety of functions throughout the plant life cycle as key regulators of several developmental processes (floral organ identity determination, control of leaf epidermal cell identity, response to various types of abiotic and biotic stress) (Riechmann and Meyerowitz, 1998; Gil-Humanes et al., 2009). Pre-miRNA-172 also regulates gl15 an AP2 like gene that is important for juvenile-to-adult transition and
flowering timing. MiRNA-172 has been shown to regulate gl15 expression. Increasing gl15 activity in transgenic maize not only increases the number of leaves expressing juvenile traits, but also delays the onset of reproductive development. MiRNA-172 promotes the transition to the adult phase by down-regulation of gl15 (Lauter et al, 2005). In *C. richardii* miR-172 may also regulate the expression of AP2 and AP2-like genes having a similar role in the plant life cycle and mediating reproductive maturation timing and response to stress.

Pre-miRNA-396c from *Sorghum bicolor* (sorghum grass) is thought to regulate the Growth-regulating factor (GRF) gene, which regulates calcium signaling and cell proliferation in plants. MiRNA-396c may regulate GRF and GRF-interacting factors (GIF) that have been shown to interact with components of a chromatin-remodeling complex in yeast and as a regulator of leaf growth and morphology in *Arabidopsis* (Treich et al. 1995; Kim JH and Kende H, 2004) and corn (Zhang et al., 2008) and may play similar roles in *C. richardii*. Pre-miRNA-396c from *Citrus sinensis* (orange) was also predicted reinforcing the impression that miRNA-396c regulates GRF by mRNA cleavage in plants. However, the orange variant of pre-miRNA-396c had 3 mismatched base pairs with the EST alignment, making it a weaker prediction than the sorghum grass variant.

Pre-miRNA-160 from *Triticum aestivum* (bread wheat) is a regulator of the Auxin Response Factor (ARF), a transcriptional regulator of auxin (hormone) expression. ARF and indole-3-acetic acid proteins regulate the expression levels of auxin and play a global role in various plant hormone signaling pathways (jasmonic acid, salicyclic acid) that regulate a variety of physiological processes including apical dominance, tropic responses, lateral root formation, vascular differentiation, embryo patterning, and shoot elongation (Staswick et al., 2002;
Okushima et al., 2005). Pre-miRNA-160 regulation of ARF may play a similar role in *C. richardii* hormonal regulation of growth.

Pre-miRNA-466q from *mus musculus* (house mouse) lacks validated target predictions in plants. However, in animals it is predicted to target the 6-phosphofructo-2-kinase/fructose-2 (Pfkfb2) gene, the butyrophilin-related 1 (Butr1) gene, the sine oculis-related homeobox 6 (Six6 in *Drosophila*) gene, and 8 as of yet uncharacterized ESTs and cDNAs. While Butr1 and Six6 are uncharacterized in plants, Pfkfb2 is known to be involved in both the synthesis and degradation of fructose-2,6-bisphosphate, a regulatory molecule that controls glycolysis in eukaryotes. Several orthologs have been characterized in *Arabidopsis* and rice (Theologis et al., 2000; Ohyanagi et al., 2006). The encoded protein has a 6-phosphofructo-2-kinase activity that catalyzes the synthesis of fructose-2,6-bisphosphate, and a fructose-2,6-biphosphatase activity that catalyzes the degradation of fructose-2,6-bisphosphate. Essentially regulating a key step in energy metabolism and developmental timing.

Pre- miRNA-1082a from *Selaginella moellendorffii*, an ancient vascular plant (~400 million years old), is predicted to bind to an unreviewed predicted protein A9T955 in *Physcomitrela patens* (Rensing et al, 2008). Computational resources predict that this gene codes for a membrane protein with D-arabinono-1,4-lactone oxidase activity, UDP-N-acetylmuramate dehydrogenase activity, and flavin adenine dinucleotide binding activity (UniProt). The role and existence of this protein in *C. richardii* is unknown, however if computational predictions are accurate a similar protein should exists considering the close evolutionary lineage between moss and fern.
Pre-miRNA-716b from *Sarcophilus harrisii* (Tasmanian devil) has no validated targets and target prediction servers yield weak results for *Arabidopsis* and no result for *Physcomitrella*. The role of pre-miRNA-716b in *C. richardii* and other eukaryotes remains elusive and microarrays of global mRNA expression are a probable way of target elucidation.

Pre-miRNA-166b from *Brassica napus* (rapeseed) has been shown to target the class III homeodomain-leucine zipper (HD-ZIPIII) genes in *Arabidopsis*. HD-ZIPIII gene is conserved in plants and thought to be critical in early development as leaf primorida become polarized along their adaxial/abaxial axes. The establishment of the ad/abaxial polarity is fundamental to several aspects of shoot development and leaf vascularity (Canales et al., 2005, Carraro et al., 2006). In the absence of HD-ZIPIII protein, lower abaxial leaf fates develop and meristems fail to form (Talbert et al., 1995; Prigge et al., 2004). The functions of HD-ZIPIII in lower plants are yet to be determined, however it is hypothesized that this gene plays a similar role in leaf polarity, shoot development, vascularity, and plant growth (Tang et al; 2003; Ochando et al., 2006). Pre-miRNA-166b in *C. richardii* may regulate HD-ZIPIII expression by mRNA cleavage.

Pre-miRNA-5645d from *Arabidopsis thaliana* (mouse-ear crest) is predicted to target AT5G30247.1 in *Arabidopsis thaliana* (psRNA Target Server), which is predicted to be an ortholog of Q9C87Q7 a putative protein also known as Athila open reading frame 1 (ORF1) (UniProt). The Athila ORF1 gene family is a group of *Arabidopsis* retrotransposons belonging to the Gy3/gypsy family of the long terminal repeat (LTR) class of eukaryotic retrotransposons (Pelssier et al., 1995; Wright et al., 1998; Marin et al., 2000). The role of these retrotransposon genes, if any, is uncharacterized in *C. richardii*. However, retrotransposons are particularly abundant in plants and in some species they compose between 50-80% of the entire genome and may have various regulatory functions (Staton et al., 2009).
Pre-miRNA-408 from *Arabidopsis lyrata* is a weaker prediction than the ones discussed thus far. All the pre-miRNA predictions so far (excluding pre-miRNA-396 orange variant) had between 0-2 basepair mismatches between the EST sequence and miRNA sequence alignment. Pre-miRNA-408 has 3 base pair mismatches, which presents a problem for target prediction. As discussed earlier, plant miRNAs bind to their mRNA targets with almost 100% complementarity. The presence of mismatches makes target predictions weaker and the alignments may have totally unique targets. Pre-miRNA-408 has no validated targets, but is predicted to bind putatative protein AT1G17180.1 (psRNA target server) in *Arabidopsis thaliana*, a homolog of Q9SHH7 (UniProt) also known as Glutathione S-transferase U25 in *Arabidopsis thaliana*. This protein may be involved in the conjugation of reduced glutathione to a wide number of exogenous and endogenous hydrophobic electrophiles and have a detoxification role against certain toxins (Wagner et al., 2002).

In summary, predicted pre-miRNAs have very low E-values (less than e-12) which indicates that these alignments are not chance alignments. Target prediction reveals that pre-miRNAs 172, 396c, and 160 target genes that regulate plant hormones and growth factors necessary for development, cell proliferation, and maturation in plants. Pre-miR466q and 1082a target genes necessary for cellular metabolism. Pre-miRNAs 166b and 5645d target genes necessary for chromatin remodeling and retrotransposon activity, respectively. Pre-miRNA 408 targets a gene necessary for dealing with oxidative stress and toxins in plants. Finally pre-miRNA-716b had no known targets and needs further characterization.
4.2 *Conserved mature miRNA*

Mature miRNA predictions were more numerous than pre-miRNA predictions, yet had an e-value cut off twice as high (e-6), doubling the likelihood that the alignment occurred by chance. Mature miRNA sequence alignments are also shorter than pre-mature miRNA alignments and have a higher chance of being randomly aligned to repeated sequences in ESTs. For these reasons only miRNAs from plants with 0 base pair mismatches between the EST and miRNA alignment were assessed for target interactions. The remaining predictions may yet be valid (including the many animal miRNA predictions), however they will not be discussed to avoid potentially false predictions with misleading target interactions.

Predicted miRNA-5658, miRNA-414, and miRNA-853 in *Arabidopsis thaliana* are all aligned with the same EST (gi|9958977) about a 100 base pairs apart from each other. MiRNA-5156 had only one significant target (AT2G15840.1; psRNA Target Server) with an Expect value of 2.0. AT2G15840.1 is a pseudogene or hypothetical protein predicted to exist in *Arabidopsis*. MiRNA-169i.2 had no significant target prediction. MiRNA-858 had no known targets on either the psRNA target server or the DIANA server.

MiRNA-5658 is predicted to interact with NF-YA5 a nuclear transcription factor shown to be regulated by miRNA-166 during plant stress response. MiRNA-5658 is also predicted to target the Botrytis-induced kinase (BAK)-1-interacting-receptor-like kinase 1 (BIR1), a gene thought to initiate plant innate immunity by responding to bacterial protein flagellin 22 (flg22) (Cinchilla et al., 2007; Lu et al., 2010). BAK1 also functions as a negative regulator in a brassinosteroid-independent cell-death pathway (Kemmerling et al., 2007; He et al., 2008). BIR1 was recently identified as a regulator of multiple resistance signaling pathways (Gao et al., 2011).
and the loss of function of BIR1 induces cell death and constitutive defense responses in *Arabidopsis thaliana* (Wang et al., 2011).

MiRNA-414 has a diverse range of targets that includes 3 proteins from the previously discussed AP2 transcriptional factor family that regulate auxin expression and control growth and proliferation. Nucleosome assembly protein 1 (NAP1), a protein conserved from yeast to human facilitates the *in vitro* assembly of nucleosomes as a histone chaperone (Dong et al., 2005), is another important target for chromosome function and remodeling. WRKY58 is a DNA binding protein part of the WRKY superfamily of transcription factors in plants that regulate plant growth, development and responses to abiotic and biotic stress (Wang et al., 2006). YAO transducing/WD40 a nucleolar protein in *Arabidopsis* is required for the correct positioning of the first zygotic division plane making it critical for gametogenesis and embryogenesis as well as rRNA processing in plants. YAO is expressed ubiquitously with high level of expression in tissues under active cell divisions, including embryo sacs, pollen, embryos, endosperms, and root tips (Li et al., 2010) making it an important miRNA-414 target. MiRNA-414 has several other targets including an F-box/RNI-like protein, a glutaredoxin family protein, a pentatricopeptide repeat family, a calmodulin-binding protein, the wall associated kinase-like 4 protein, the DUO3 homeodomain-like protein, a C2H2 type zinc finger transcription factor and several ribosomal proteins of the RPL5B family, as well as several other uncharacterized predicted proteins. This diverse range of targets make miRNA-414 critical to regulating all aspects of plant life from early development, to growth and proliferation, signal transduction, metabolism and ion balance, and even DNA/histone modification and constituent RNA expression.

Five other plant miRNA were predicted in EST alignment miRNA-5021 from *Arabidopsis*, miRNA-2590g from *Medicago truncatula*, miRNA-838 from *Arabidopsis lyrata*,
miRNA-5156 and miR-169i.2 in rice. MiRNA-5021 targets the cytoskeletal protein-like protein thought to be involved in cytoskeletal reorganization (Q84JE5; UniProt). Alpha-L-fucosidase I(ATFUC1) is an enzyme that is able to modify fucosylated xyloglucan oligosaccharides in plant cell walls having an indirect effect on controlling the rate of cell expansion and distribution of cell wall signaling molecules (de la Torre et al., 2002). MiRNA-5021 also targets the RING/U-box superfamily protein a component of ubiquitin-activating enzymes (E1, E2, E3) thought to be important for protein degradation and has been linked to important aspects of plant growth and development ranging from embryogenesis to senescence (Gray et al., 2002). WRKY14 part of the WRKY superfamily is another important transcription factor that is regulated by miRNA-5021. Mitogen-activated protein kinase kinase kinase 21 (MAPKKK21), cyclin A2 (CYCA2), S-adenosyl-L-methionine-dependent methyltransferase, ADP-ribosylation factor C1 (ARFC1), and ammonium transporter 2 (ATAMT2), as well as several uncharacterized predicted proteins are other predicted targets of miRNA. MiRNA-5021 also seems to play a diverse function in plants from cell cycle regulation, protein degradation, cell growth and expansion, and dealing with abiotic and biotic stress.

MiRNA 2590g targets several genes; although most do not have a low enough Expect value from target prediction to be considered significant. The only gene that is potentially a good target prediction is the transposable element gene also known as mutator-like transposase (E-value 2.0), thought to function as a zinc ion binding protein (Q9LRP4; UniProt).

MiRNA-838 also has many predicted targets but only two has low enough Expect values to be considered significant. Two proteins (AT4G39420.2 and ATG45030.1; psRNA Target Server) had a low enough Expect value 2.0. AT4G39420.2 is an uncharacterized protein that contains a zinc-binding domain and needs further characterization. ATG45030.1 is an O-
fucosyltransferase family protein that is believed to have transferase activity and functions to transfer glycosyl groups of plant cell wall polymers (xyloglucans, rhamnogalacturonan, arabinogalactan, N-glycans) that regulate cell wall integrity and extensibility in growing tissues (Wu et al., 2010). It is important to note that miRNA-838, unlike the other miRNA predictions, is the only miRNA predicted to function as a translational inhibitor as opposed to cleavage of its mRNA targets.

These miRNA predictions reveal the diverse roles that miRNA regulation has in the plant life cycle including cellular metabolism, developmental timing, alteration of chromosome structure, and response to biotic and environmental stress among many other more specific roles discussed. Conservation of miRNAs within plants and between animals reinforces the theory that miRNAs are an ancient regulatory mechanism, evolving long before plants and animals diverged.

4.3 False positives and negatives

The existence of false positives and false negatives is a challenge in many studies. Comparing EST sequences to previously known Arabidopsis and rice miRNA sequences dramatically reduced the number of false positives. ESTs were blasted against all known proteins in plants to remove ESTs that potentially code for protein instead of regulatory RNA. According to the characteristic that plant miRNAs are highly conserved and only a few nucleotides change between plant species, ESTs with only 0-3 mismatched nucleotides were considered as potential miRNA candidates. As discussed earlier, only EST alignments with very low e-values (e-12 in pre-miRNA and e-6 in miRNA) were considered. Combining these strategies reduced the
number of alignments from 327,172 to 9 for pre-miRNA and 76,940 to 89 for mature miRNA, increasing the confidence levels of miRNAs identified in this study.

False negatives are bound to exist, as this was a limited study of roughly 6000 ESTs in *C. richardii*. Similar studies conducted in cotton and rice used over 150,000 ESTs. There remain many gaps in miRNA identification for *C. richardii* and the sequencing of this organism’s genome would contribute greatly to our understanding of miRNA regulation in early land plants. However, the overlap of miRNA predicted in this study with those known to be important in regulating stress in plants (*Table 1*), certainly displays that significant findings can be made from a relatively small number of ESTs.

### 4.4 Plasmolysis regulates microRNA expression

Plasmolysis of *C. richardii* gametophytes significantly altered gametophyte morphology, leading to the secondary gametophyte (*Figure 6*). Plasmolysis, as evidenced by qPCR, altered the expression levels of 3 miRNA genes: miRNA-106b, miRNA-181c, and cre-miRNA-914.

Human miRNA-106b-5p is thought to target cyclin-dependant kinase inhibitor 1A (P21/CIP1; DIANA server) a profoundly important regulator of cell cycle that binds to and inhibits the activity of cyclin-CDK2 or CDK4 complexes, and thus functions as a regulator of cell cycle progression at G1. P21 is thought to be tightly controlled by the tumor suppressor protein P53 and is thought to mediate a P53-dependent cell cycle G1 phase arrest in response to a variety of stress stimuli. P21 also interacts with proliferating cell nuclear antigen (PCNA), a DNA polymerase accessory factor that plays a role in S phase DNA replication and DNA damage repair (Becker et al., 2006).
Initially it was predicted that miRNA-106b would demonstrate an increase in expression level since cells of the secondary gametophyte were undergoing rapid proliferation and mitosis. An increase in miRNA-106b expression should lead to inhibition of P21 by mRNA cleavage or translational repression, allowing for CDK2/CDK4 complexes to drive the cell cycle through G1 and into S, G2 and mitosis. The relative expression levels of mature miRNA-106b in both 8-day-old and 12-day gametophytes, quadrupled post-plasmolysis. Pre-miRNA-106b was greatly elevated (8x) in 24-hours post-plasmolysis after which it was almost undetectable in 48 to 72 hours in 12-day-old gametophytes. Mature miRNA-106b was almost undetectable in control and 24-hours after plasmolysis and dramatically increased in 48 and 72-hours; after pre-miRNA-106b had accumulated in 24 hours.

MiRNA-181c is thought to regulate fidgeting (FIGN) a member of the AAA protein superfamily with an important role in microtubule severing and depolymerizing the enzyme used to regulate mitotic spindle architecture, dynamics, and anaphase progression. Human fidgeting is thought to actively inhibit microtubule growth from and attachment to centrosomes (Mukherjee et al., 2012). MiRNA-181c also targets transportin1 (TNPO1) a protein that recognizes nuclear import signals on proteins that are different from the classical basic nuclear localization signals. This transportin-mediated nuclear import pathway is highly conserved between man, yeast, and plants (Ziemienowicz et al., 2003) allowing for efficient and bi-directional traffic of many different proteins, RNAs and ribonucleoprotein complexes. MiRNA-181c is also thought to target ATP-dependent RNA helicase (DDX3X/Dead box protein 3), a multifunctional RNA helicase involved in several steps of gene expression, such as transcription, mRNA maturation, mRNA export and translation. The molecular functions of this RNA helicase involve RNA secondary structure unwinding, cellular response to osmotic stress, chromosome segregation,
negative regulation of apoptotic pathway, negative regulation of cell growth, mature ribosome assembly, positive regulation of G1/S transition of mitotic cell cycle, among many other biological processes (UniProt).

Pre-miRNA-181c was up-regulated in both 8-day-old and 12-day-old C. richardii post-plasmolysis. In 12-day-old ferns pre-miRNA-181c showed the greatest up-regulation with a relative expression level 1000x higher in 24-hour cells versus control cells. Expression levels remained very high, about 500x in both 48 and 72-hour cells. Mature miRNA-181c increased accordingly after 24-hours, and slowly trailed off after 48 and 72-hours. In 8-day-old ferns, pre-miR-181c was expressed at a much lower level before and post plasmolysis, as was mature miR-181c. However, miRNA-181c showed up-regulation in both 8 and 12-day-old ferns.

Cre-miRNA-914 has no known targets in any organism. The reason it was incorporated into this study was to demonstrate that miRNA are highly conserved between eukaryotes. This miRNA was detected in Chlamydomonas reinhardtii (Zhao et al., 2007). Detection by qPCR in 8-day-old ferns showed that the mature miRNA-914 was expressed at a relatively high level in the control and was down-regulated slightly post-plasmolysis, while pre-miRNA-914 was expressed at an almost constant level. In 12-day-old ferns the expression of mature miRNA-914 tripled 24-hours post-plasmolysis and then declined back to control levels after 48 and 72-hours, while the pre-miRNA-914 was detected to be down-regulated or modified into mature miRNA-914 post-plasmolysis.
CHAPTER 5
CONCLUSIONS

These findings indicate that the miRNA detected may play a diverse set of roles in post-plasmolyzed gametophytes. 8-day-old gametophytes had a much lower expression level miRNA-106b; indicating that P21 activity is important in very young gametophytes to prevent the progression of the cell cycle. After plasmolysis of 8-day-old ferns, miRNA-106b levels increased and may have led to the inhibition of P21 and the loss of inhibition of CDK2/CDK4 complexes, essentially releasing cell cycle arrest at a time when gametophytes are not supposed to proliferate and divide. The up-regulation of miRNA-181c may inhibit FIGN-like proteins in *C. richardii*, preventing the inhibition of microtubule organization and mitotic progression. MiRNA-181c may also inhibit TPO1, altering the transport of macromolecules between the cytosol and the nucleus. MiRNA-181c may also inhibit DDX3X, potentially affecting all downstream functions of this ATP-dependent RNA helicase. Cre-miR-914 has an unknown role, however its detection adds depth to the extent of which miRNA are conserved in evolution.

Activation and regulation of gene expression at different levels (translation, transcription, postranscriptional) is important to all life. Certain small RNA molecules are important components of regulatory circuits and are involved in wide-ranging aspects of gene regulation. *C. richardii* proved to be a useful model organism and EST analysis has yielded deep insights into 98 predicted miRNA regulatory pathways. The miRNA predicted and those validated by qPCR post-plasmolysis serve to expand our understanding of the effects of hydrostatic stress on plant cells. RNA interference and the establishment of miRNA-specific-genetically modified crops may be used to increase crop yields, enhance plant adaptation to stress, and secure global food resources.
Computational prediction and qPCR validation of miRNA in Ceratopteris richardii

ACKNOWLEDGEMENTS

Firstly, I would like to thank you, the reader, for taking the time to assess this work. This study stemmed from an independent study project in 2011 in which Dr. Edward Tucker introduced me to the role of microRNAs in humans and their detection by microarray in plants. I was very intrigued by these small molecules and could not help but ponder on the intricacy of Mother Nature and the complex ways in which life has emerged on our planet.

My mentor, Dr. Tucker, fully believed in my intellectual abilities, and fostered my creativity and originality in this project. Dr. Tucker trusted in my maturity, integrity, perseverance, and dependability; allowing me to independently plan and conduct all aspects of the experiment. I was given a lot of freedom and took the initiative to complete this project to the best of my ability. I was able to apply analytical and problem solving skills to get the results and in the process hone my written and oral expression skills. I could not have done this project without the continued support of Dr. Tucker, who not only encouraged my intellectual pursuits, but also acquired the funding necessary for our project.

During my freshman year I was lucky enough to have Dr. Jason Munshi-South as my professor for a course I did not select, but loved from the start. Environmental 1020 was the only class in my Learning Community Block that actually challenged me intellectually and fascinated me. Dr. Munshi-South was a great professor and I thank him for getting me onto the Natural Science track. I wholly acknowledge Dr. Munshi-South for his continued support throughout this project. Dr. Munshi-South gave me access to his laboratory and allowed me to use his supercomputer to perform the bioinformatic predictions and analysis. He also provided me with a space to conduct the wet-lab aspects of the experiment. Most importantly Dr. Munshi-South set me up on his quantitative real-time PCR machine, which was imperative for the progression of
Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

this study. Dr. Munshi-South was a great help in both the experimental aspect of this project and as one of the Natural Science Department committee readers. His notes and revisions to this paper are greatly appreciated.

In 2012 I took Developmental Biology, taught by Dr. Seymour Schulman. Dr. Schulman is always on top of the latest discoveries in genetics and his interest in epigenetics overlapped with mine in regulatory RNAs. I thank Dr. Schulman for his continued support and persistent inquiries on the status of the project. I am happy to say that, although these topics are inherently dull, we have never had a dull discussion. Dr. Schulman agreed to be one of the Natural Science Department committee members and I cannot thank him enough for help with the revisions. I also thank Dr. Schulman for heading my Natural Science Department recommendation letter for medical school as well as helping me revise my personal statement.

I thank the Natural Science Department at Baruch College and all of the wonderful professors who have given me a great experience. Dr. John Wahlert the department chair of the department, who in addition to approving my honors proposal, taught great biology and anatomy courses. Dr. Emil Gernert for teaching the most rigorous biology and human physiology courses at Baruch College. I also thank Dr. Gernert for giving me the opportunity to tutor biology throughout the years, as well as his support of both the Biomedical Society and my role this year as the Vice President. I thank Dr. Helene Eisenman for helping me with the honors prospectus and for her advice on qPCR. I thank Dr. Robert Butler, Dr. Joel Brind, Dr. Edyta Greer, Dr. Ramzi Khuri, Dr. Chandra Kulatileke, Dr. Keith Ramig, and Dr. David Szalda for offering me their knowledge of biology, biochemistry, general chemistry, organic chemistry, and physics. I never imagined that I would have access to such world-class and highly accomplished educators and researchers; who genuinely care for and revel in the success of their students. The Natural
Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

Science Department has been paramount to my development as an individual and has given me the tools to succeed as a researcher.

I am extremely grateful to my family, especially my mother Maria Kharonova, for putting me through college and believing in my abilities. My brothers Rudy Kelner and Gregory Kharonov for always being great role models and helping me develop into the person I am today. I also thank my lovely girlfriend Julia Gertopoky for her love and support throughout this project. A very big ‘thank you’ to Maxat Askarbekov for his help with the bioinformatics and miRNA annotation, as well as his help in the lab. Thank you to the soon to be doctor, Stephen Harris, you are a spreadsheet genius. I could not have filtered the predictions without Dr. Harris’ help. Last but not least, I’d like to thank my good friend Christopher Cosgriff for his work on the galaxy server, setting up NCBI tools, and spending many arduously long evenings side by side with me, writing our theses.
Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*

References


Athila, a new retroelement from *Arabidopsis thaliana*. Plant Mol Biol 29:441-452.


Computational prediction and qPCR validation of miRNA in Ceratopteris richardii


Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*


Computational prediction and qPCR validation of miRNA in *Ceratopteris richardii*


Talbert PB, Alder HT, Parks DW, Comai L (1995). The REVOLUTA gene is necessary for apical meristem development and for limiting cell divisions in the leaves and stems of *Arabidopsis thaliana*. Dev 121:2723-2735.


