SMALL RNAS IN GOSSYPIUM AND THEIR ROLES IN THE RESPONSE TO HEAT STRESS

by

Hongtao Hu

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
Dec 14, 2013

Small RNAs, heat stress, microRNAs

Approved by

Robert D. Locy, Chair, Professor of Biological Sciences
Narendra K. Singh, Professor of Biological Sciences
David B. Weaver, Professor of Agronomy and Soils
Zhanjiang (John) Liu, Professor of Fisheries and Allied Aquacultures
Aaron M. Rashotte, Associate Professor of Biological Sciences

Abstract

Small regulatory RNAs, which usually are 20 to 24 nt in length, play crucial roles in plant growth, development, and stress response. Based on their origin, biogenesis and function, plant small RNAs can be classified into two major classes, hairpin RNAs (hpRNAs) and short interfering siRNAs (siRNAs). hpRNAs and siRNAs can be subdivided into two and three subfamilies, respectively. hpRNAs, including microRNAs (miRNAs) and other hairpin RNAs (ohpRNAs), are produced from hairpin-shaped precursors. siRNAs are produced from double-stranded precursors and requires RNA dependent Polymerase. siRNAs can be categorized into three subfamilies, repeat-associated siRNAs (ra-siRNAs), pha-siRNAs, and *cis* natural antisense RNAs (*cis*-nat siRNAs). In general, miRNAs regulate gene expression at transcription level, while siRNAs can do either postranscriptionally or transcriptionally. In addition, ra-siRNAs are considered to regulate gene expression by DNA methylation.

Gossypium plants, such as *G. hirsutum* and *G. barbadense*, are economically important plants, and provide natural fiber for textile industry. In addition, they also are important sources for proteins and seed oils. Heat stress caused by elevated high temperature is an important factor that affects diverse physiological processes of plants, resulting serious loss in yield. The mechanism of heat response has been extensively studied in anatomical, biochemical, and gene levels. But till one decade year ago, gene expression regulated by small RNAs is shown to be important for animal and plant stress response.

Growing studies have shown that small RNAs, primarily miRNAs and siRNAs, are involved in heat stress response in plants. In *Arabidopsis*, *Oryza*, or *Medicago*, the small RNA loci have been thoroughly annotated, which identified a large number of small RNAs that regulates stress response. However, little is known about *Gossypium* small RNAs and their regulatory roles in heat stress response.

To annotate *Gossypium* small RNAs and study their roles, we constructed and sequenced forty four small RNAs libraries from eight different tissues from two different *Gossypium* species, *G. raimondii* and *G. hirsutum*. The samples were collected from heat-stressed or non-heat-stressed seedlings grown in growth chamber, or plants grown in fields under high temperature summer. Two deep sequencing methods, ABI Solid and Illumina, were employed to sequence the small RNA libraries. This generated approximately 790 million sequencing reads, which enabled us to annotate a large number of miRNAs, ohpRNAs, ra-siRNAs, pha-siRNAs and *cis*-nat siRNAs. Expression analysis revealed that a large number of small RNAs were significantly regulated between heat-stressed and non-heat-stressed tissues, or between heat-susceptible and heat-tolerant genotypes. Some small RNAs were differentially expressed among different tissues. This suggested that small RNAs might be involved in the stress response of *Gossypium* plants to heat stress and might play crucial roles in *Gossypium* heat stress tolerance.

This work in this study developed a systematical method in analyzing the small RNA transcriptomes and in annotating the small RNA loci in the genome. This would facilitate the study in the small RNAs and their regulatory mechanism in plant heat tolerance.

Acknowledgments

The author deeply appreciates his adviser, Dr. Robert Locy, for his wisdom, training, patience and encouragement during the whole period of the author's study in Auburn University. Dr. Locy gave his numerous support and time that are very vital for the author to complete his research. Dr. Locy not only teaches the author how to do science, but also encourages the author to use the correct and logical methods in conducting his studies. This is valuable treasure for the author and will affect his entire career.

The author would like to thank his committee, Drs. Narendra Singh, David Weaver, Zhanjiang Liu, and Aaron Rashotte. In the past five years, they help the author overcome many problems occurring in his studies and provide numerous insightful suggestions for improving his studies. The author would thank Dr. Eric Peatman for serving as outsider reader for the dissertation and for giving constructive suggestion for improving the dissertation.

The author would like to express his appreciation to Mrs. Chia-chen Weng, Mr. Shankar Pant, Dr. Roelof Sikkens, and Ms. Ruijuan Li, for their collaboration, support, and assistance in lab techniques and field managements.

The author also wishes to thank Auburn University, the College of Sciences and Mathematics, and the Department of Biological Sciences for their financial support of both teaching and research assistantships throughout the period of this dissertation.

Also research funding and support from the Alabama Cotton Commission and the Alabama Agricultural Experimental Station making this work possible is greatly acknowledged.

The author also acknowledges the collaboration of the Hudson Alpha Institute of Biotechnology, Huntsville, AL for their assistance in DNA sequencing. The special assistance provided by the Alabama Super Computer Authority, Dr. David Young specifically, is also greatly acknowledged.

Many thanks go to Dr. Scott Santos and Dr. Leslie Goertzen for their constructive suggestions for completing the analysis.

The author would like to thank his parents for their selfless support, which is very important for the author to complete his study. The author also wants to thank his wife and his son for their patience, understanding and supporting.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

DCL	Dicer-Like Protein
microRNA	miRNA
siRNA	short interfering RNA
ra-siRNAs	repeat-associated siRNAs or heterochromatic siRNAs
hc-siRNAs	heterochromatic siRNAs
ta-siRNAs	trans-acting siRNA
cis-nat siRNA	*cis* natural antisense siRNA
miRNA*	miRNA star, the pattern strand of miRNA
sec-siRNAs	secondary siRNAs
AGO	Argonaute protein
ohpRNAs	other hairpin RNAs
RDP	RNA dependent RNA polymerase

I. LITERATURE REVIEW

1. Plant Responses to Heat Stress

Heat stress, a major abiotic stress, is often defined as high temperature above the optimum that causes plant damage and alters growth and development (Wahid et al. 2007). High temperature typically affects vegetative growth and reproductive development often causing reproductive failure and severe yield losses (Barnabas et al. 2008; Snider et al. 2009) by creating a wide range of adverse impacts on physiology, metabolism, and development (Cheikh and Jones 1994; Aloni et al. 2001; Salvucci and Crafts-Brandner 2004; Zhao et al. 2005; Qin et al. 2008; Ginzberg et al. 2009; Goswami et al. 2010). These inhibitory effects are mediated by disturbing cellular homeostasis (Wang et al. 2004; Kotak et al. 2007a; Martiniere et al. 2011), by direct effects on the electron transport chain (Oukarroum et al. 2012), or through effects on Ribulose-1,5-bisphosphate carboxylase oxygenase (Rubisco) activity (Law and Crafts-Brandner 1999; Crafts-Brandner and Law 2000), and by repressing respiration (Lin and Markhart 1990).

1.1. Heat Stress as a Signal

Heat stress response in plants involves a complex series of events, but the focus of heat stress work has traditionally centered on the heat-shock response and the regulation of heat shock and other proteins during and after the perception of heat stress. Heat stress responses could be divided into different phases consisting of perception of heat, signal transduction, and alterations of gene expression (Verslues and Zhu 2005).

Two basic mechanisms for the sensing of heat by cells have been proposed. These include sensing of heat signal by membranes, particularly the plasma membrane, based on changes in membrane fluidity, and in photosynthetically active tissues sensing of heat by the photosystems of the chloroplast typically described as induced photoinhibition.

The plasma membrane is considered among the initial responses involved in sensing stress (Hofmann 2009). The fluctuation of temperature disrupts the integrity of biological membranes, and leads to the changes in the components of fatty acid (Saidi et al. 2009; Martiniere et al. 2011). These changes in the lipid composition of the membrane play a role in heat perception and initiation of stress signal transduction (Vigh et al. 1993; Falcone et al. 2004; Los and Murata 2004; Saidi et al. 2010). Membrane fluidity changes regulate calcium-dependent heat-signaling pathways (Saidi et al. 2010). These changes in membrane fluid and components are required for the generation and transmission of stress signals to activate HSP (heat shock protein) gene via HSF1 (heat stress transcription factor) (Vigh et al. 2007). Heat stress signal is perceived by the sensors embedded in the membrane and triggers the influx of Ca^{2+} from extracellular to cytoplasm. The elevated Ca^{2+} activates the receptors of Ca^{2+} binding proteins and transduces the signal to down-stream effectors, including MAP kinase, heat stress and factors, etc. Consequently, the heat-stress responsive genes are activated. In this model, G proteins are proposed to play an important role in activation of the signalling pathway and transduction of the signal through GTP-coupled phosphorylation. In addition, ROS is considered a crucial secondary messenger that transfers the heat stress signal to downstream effectors, such as Calmodulins. But it is not clear how the Ca^{2+} is released under stress condition, because plants lack the receptors of InP_3 that is found in animals.

During photosynthesis, electrons stripped from water encounter PSII, the most heat-labile component of the photosynthetic apparatus (Law and Crafts-Brandner 1999; Sharkey 2005). The PSII core complex is destabilized by high temperature, causing sensitivity of PSII to heat (De Las Rivas and Barber 1997). Manganese ions (Mn^{2+}) are critical to the oxygen evolving apparatus that adjoins PSII, and Mn^{2+} is released from the PSII core complex by high temperature, causing a loss of oxygen-evolving activity and leading to ROS production (Nash et al. 1985; Yamane et al. 1998). ROS, induced by high temperature oxidizes PS II core reaction center D1 protein (De Ronde et al. 2004; Wang et al. 2010a). Further photosynthetic inhibition is mediated by the effect of ROS on thiol-containing enzymes in the Calvin cycle leading to photoinhibition of carbon fixation (Kaiser 1979; Nishiyama et al. 2001; Takahashi et al. 2002). In addition, Rubisco appears to be a limiting factor that inhibits photosynthesis upon heat stress (Salvucci and Crafts-Brandner 2004).

1.2. Transduction of the heat signal and gene expression

Perception of increase in temperature is a cellular signal which can then be transduced through a series of second messengers (Mahajan and Tuteja 2005; Verslues and Zhu 2005). Ca^{2+} and ROS constitute the most important intracellular signaling molecules that participate in the regulation and integration of diverse cellular functions (Mazars et al. 2010). In additional, there are significant interactions among different pathways creating a complex regulatory network that may be far more important than individual effects.

1.2.1. Calcium signaling

Ca^{2+} is a critical component in heat stress signaling pathways (Kleinhenz and Palta 2002; Galon et al. 2010; Reddy et al. 2011; Wu and Jinn 2012). Ca^{2+} sensors are also called calcium-binding

proteins that sense the change in cytosolic Ca^{2+} concentration (Ranty et al. 2006; Boonburapong and Buaboocha 2007), and transduces the signal to downstream factors, including phosphorylation cascades activating the transcription of stress-responsive genes (Szymanski et al. 1996; Choi et al. 2005; Mahajan and Tuteja 2005; Kushwaha et al. 2008; Sinha et al. 2011; Kim 2013). The majority of these Ca^{2+} binding proteins possess one or more highly conserved helix-loop-helix motifs known as EF-hands that specifically bind Calcium ions (Day et al. 2002). The EF-hand Ca^{2+} binding proteins can be broadly grouped into three major families: calmodulin (CaM) and calmodulin-like proteins (CMLs) (Luan et al. 2002), calcium-dependent protein kinases (CDPKs, Harper et al. 2004; Harper and Harmon 2005; Reddy et al. 2011; Kim 2013), and calcineurin B-like proteins (CBLs) with the associated CBL-interacting protein kinases (CIPKs, Luan 2009 (Kim 2013). These sensor complexes detect calcium resulting in transduction of the signal to downstream effectors (Liu et al. 2003; Kim et al. 2009).

Ca^{2+}-regulated gene expression has been implicated in plant stress responses, as the key secondary messenger of stress signaling (Kim et al. 2009; Tan et al. 2011; Wu and Jinn 2012; Cheval et al. 2013). Ca^{2+}/Calmodulin also mediates the phosphorylation of NADPH oxidase (NOX), which is critical for the generation of ROS (Pandey et al. 2011).

1.2.2. *ROS Signaling*

Reactive oxygen species (ROS) play a dual role in stress physiology, i.e. they can be both responsible for stress damage, and they can serve as local or systemic signals producing secondary stress responses. Several molecular ROS species including superoxide (O_2^-), hydrogen peroxide (H_2O_2), and hydroxyl radical ($OH^\bullet$), impair organelle function (Foyer and Noctor 2005; Triantaphylides et al. 2008) and induce programmed cell death (Rhoads et al. 2006;

Zhang et al. 2009). ROS are continually and unavoidably produced by the electron transport chain during aerobic respiration in the mitochondria (Apel and Hirt 2004) and as a result of the electron transport process in chloroplasts during photosynthesis (Turrens 2003; Apel and Hirt 2004; Asada 2006). Despite their destructive activity, ROS may function as second messenger involved in cell rescue and defense signaling pathways (Desikan et al. 2001; Yan et al. 2007; Sharma et al. 2012).

Plants code multiple antioxidant proteins, such as SODs, FeSOD, MnSOD, and Cu/ZnSOD, which are capable of scavenging of ROS and alleviating oxidative damage (Miller et al. 2007; Gill and Tuteja 2010). SODs are considered to be the first line of defense system against ROS and catalyze the dismutation of superoxide (O_2^-) into H_2O_2 (Alscher et al. 2002). In the later reduction reactions, H_2O_2 and other products are reduced by a series of enzymes, including APX (ascorbate peroxidase), Catalase (CAT), MDAR (malondialdehyde reductase), GPX (Glutathione peroxidase), et al. (Apel and Hirt 2004).

Besides SODs, several other proteins, including APX (ascorbate peroxidase) and peroxiredoxin, play crucial roles in plant response to oxidative stress (Yoshimura et al. 2000; Suzuki and Miter 2006; Kim et al. 2010). The transgenic Arabidopsis plants overexpressing heat shock transcription factor 3 (HSF3) demonstrate basal thermotolerance, and an isoform of ascorbate peroxidase (APX), APX^S, is consistently expressed in transgenic plants and strongly induced in non-transgenic plants upon heat stress, which suggests APX might play a regulatory role in heat stress response (Panchuk et al. 2002).

TFs may be critical components in the ROS signaling pathway (Desikan et al. 2001; Huang et al. 2010; Bassel et al. 2011; Pitzschke and Hirt 2006). Transgenic *Arabidopsis* plants

overexpressing *CBF2* (C-repeat/ dehydration-responsive-element binding factor genes), a member of CBF transcription factors, have enhanced tolerance to hydrogen peroxide stress, and produce less ROS (Sharabi-Schwager et al. 2011). However, overexpressing *CBF2* has no impact on the expression of antioxidant enzyme genes and oxidative stress responsive genes, including SOS, CAT, APX, and DPX et al. (Sharabi-Schwager et al. 2011).

1.2.3. Inositol-tris-phosphate Signaling Pathway

Inositol-1,4,5-phosphate (IP_3) is an important secondary messenger that is involved in a wide range of cellular process (DeWald et al. 2001; Walker et al. 2002; Tuteja and Sopory 2008a; Reddy et al. 2011). IP_3 is produced through the hydrolysis of phosphatidylinositol 4,5-bisphophate by phospholipase C (PLC) (Walker et al. 2002). In animals, IP_3 regulates the release of Ca^{2+} from the Ca^{2+} stores acting on receptors located on the membrane (Richardson and Taylor 1993, Mikoshiba 2007). However, due to the absence of IP3 receptors in plants, the mechanism of IP3 is not understood (Krinke et al. 2007, Kudla et al. 2010). Alternatively, IP_3 might be converted into IP_6 (Inositol hexakisphosphate or $InsP_6$) that induces Ca^{2+} influx from vacuole (Lemtiri-Chlieh et al. 2003).

IP3 also participates directly in a variety of physiological processes (Di Paolo and De Camilli 2006). Consequently, IP3 rapidly accumulates in the cells in plants and animals likely by activation PLC under heat stress (Kiang and McClain 1993; Liu et al. 2006). The accumulation of IP3 induces the expression of heat shock protein (HSP) gene, and inhibition of PCL down-regulates the expression of HSP genes (Liu et al. 2006).

1.2.4. Hormone-regulated gene expression under stress

Heat stress has been shown to change hormone homeostasis, including content, biosynthesis, and compartmentalization (Maestri et al. 2002, Qu et al. 2013). Some hormones, including Abscisic acid (ABA), ethylene, and Salicylic acid (SA), have been implicated in signaling pathways in response to environmental stimuli including heat stress (Gong et al. 1998; Borsani et al. 2001; Larkindale and Knight 2002; Cao et al. 2006), thereby mediating heat-stress-dependent changes in gene expression (Verslues and Zhu 2005; Nakashima et al. 2009). Also the application of plant hormones can alleviate the stress injury and induces stress responsive gene expression (Dive et al. 2010; Ghanashyam and Jain, 2009).

ABA regulates plant growth and development as well as adaption to stresses including heat stress (Larkindale and Knight 2002; Tuteja 2007; Cutler et al. 2010; Romero et al. 2012).). The roles of ABA in regulating the heat stress response have been studied through the screening of ABA-deficient and ABA-insensitive mutants (Larkindale et al. 2005; North et al. 2007). The accumulation of ABA and perception of ABA induces intermediate signaling events that control downstream stress responses, including sugar-sensing pathways and ROS signaling pathways (Verslues and Zhu 2005).

H_2O_2 has been demonstrated to be a crucial signaling molecule in stress signal transduction and is required for the expression of heat shock genes (Volkov et al. 2006; Qu et al. 2013). ABA is able to induce the production of H_2O_2 and enhance the activities of antioxidant enzymes localized to both chloroplast and cytoplasm, while an inhibitor of NADPH oxidase almost completely represses the increase in the activities of those antioxidant enzymes (Hu et al. 2005). In addition, ABA-induced H_2O_2 is accumulated in only apoplast, and the greatest

accumulation occurs in the walls of mesophyll cell facing large intercellular spaces (Hu et al. 2005).

Of nine heat shock genes strongly induced by heat stress, two genes, OsHSP71.1 is induced by ABA and OsHSP24.1 is repressed (Zou et al. 2009). It appears that ABA is involved in plant heat stress response probably via H_2O_2 signaling pathway in which some specific heat shock protein genes are induced or repressed for adaption to heat stress

There are ABA receptors both localized in membranes and free in the cytosol that can function in ABA singling pathways (Guo et al. 2011), and externally-applied ABA can induce plant heat stress tolerance (Gong et al. 1998). In addition, $CaCl_2$ pre-treatment is able to enhance ABA-induced heat tolerance. This observation is consistent with a role for Ca^{2+}-signaling in ABA-induced heat stress tolerance (Gong et al. 1998).

Ethylene as a phytohormone affects myriad development processes (Cao et al. 2008; Diaz and Alvarez-Buylla 2009; Zhang et al. 2013) and adaptive responses to various stresses (Guo and Ecker 2004; Cao et al. 2007; Firon et al. 2012). Heat stress induces the production of ethylene that regulates the heat stress induced kernel abortion and suppression of grain maturation (Hays et al. 2007).

Ethylene signaling participates in various stress responses via regulating the expression of ethylene receptors (ETR) localized to the ER (Chen et al. 2002; Grefen et al. 2008; Zhong et al. 2008; Mayerhofer et al. 2012) and that regulate downstream signal transduction possibly involved in the ROS signaling pathway (Youm et al. 2008; Tian et al. 2011; Xu et al. 2011). Calmodulin plays an important role in plant by directly binding to the ethylene receptors, implicating a crucial role of Ca^{2+} in ethylene signaling pathway (Yang and Poovaiah 2000; Nie

et al. 2012). Overexpression of ERF/AP2 (ethylene response factor/APETALA2) transcription factor enhances tolerance to oxidative stress caused by heavy metal and heat through modulation of antioxidant, such as SOD, APOX (ascorbate peroxidase), and GR (glutathione reductase) (Tang et al. 2005), and heat shock protein 70 (HSP70) (Xu et al. 2011).

Salicylic acid (SA) is a plant phenolic hormone that functions as a plant growth regulator and regulates an array of physiological processes, including a variety of stress responses (Kang et al. 2012; Lan et al. 2013; Shakirova et al. 2013). Growing evidence demonstrates that SA regulates plant stress responses, including oxidative stress caused by heat (Wang et al. 2010a; Noriega et al. 2012; Mejia-Teniente et al. 2013). Although the mechanism of SA in a heat stress signaling pathway is not clear, a number of studies have revealed that SA is involved in the plant heat stress response (Larkindale and Knight 2002; Wang et al. 2010a).

These studies show that SA can induce plant tolerance to oxidative stress by improving the activities of antioxidant enzymes, such as peroxidase (POD), CAT, SOD, GR, and APX (Al-Whaibi et al. 2012; Mejia-Teniente et al. 2013). SA can activate the lipid signaling enzymes via increase the activity of phospholipase D and NADPH oxidase that induces the products of ROS (Kalachova et al. 2013). These suggest that a key role of SA is to induce ROS signaling that can activate the down stress gene expression.

SA activates the expression of PR protein genes (Potlakayala et al. 2007; Zhang et al. 2010), and a critical component of the SA signaling is the non-expression of pathogenesis related protein, NPR1, that is translocated into nucleus and interacts with TGA transcriptional factors (bZIP type), which are required for the expression of PR1 gene in response to SA levels (Rochon et al. 2006; Wu et al. 2012). In addition, it has been shown that the application of exogenous SA

is able to activate phospholipase D (PLD) and subsequently to induce the accumulation of phosphatidic acid (PA), which in turn results in the rapid increase of ROS generated by NADPH oxidase (Kalachova et al. 2013).

Studies have shown that cross-talks between phytohormones and other molecules might be present in plant stress response. For example, ROS and reactive nitrogen species, nitric oxide (NO), are key molecules regulating ABA-induced signaling (Zhao et al. 2001; Desikan et al. 2004). These observations suggest that plant hormones, like ABA, regulate plant stress response through a complicated networks and other studies have shown that at least ethylene- and salicylic acid- signaling networks are involved in heat stress signaling (Larkindale et al. 2005).

1.3. Heat Shock and Heat-induced Changes in Gene Expression.

Plants, like many other organisms, have developed various adaptive mechanisms to cope with stress (Floris et al. 2009; Wu and Jinn 2012; Kimura et al. 2013). These mechanisms invariably involve changes in the expression of genes (Kotak et al. 2007b; Wu and Jinn 2012; Mackova et al. 2013). Although the level of protein expression largely relies on regulation at the transcriptional level, studies have revealed that post-transcriptional regulation of gene expression plays crucial roles in plant responses to environmental stresses (Wang et al. 2010a; Li et al. 2011; Zhu et al. 2011; Khraiwesh et al. 2012).

Heat stress changes the pattern of gene expression leading to the up-regulation of heat shock proteins (HSPs) and the down-regulation of many housekeeping proteins (Volkov et al. 2003; Goswami et al. 2010). The HSPs are molecular chaperones that are required for the development of thermotolerance in plants and other organisms (Hong and Vierling 2001). These proteins are highly conserved across different species and are produced in response to a brief

sublethal high temperature exposure followed by exposure to an otherwise lethal temperature in virtually all organisms (Bondino et al. 2012). Exposure to high temperature results in protein denaturation with the HSP proteins serving as chaperones to prevent protein denaturation. HSPs play a crucial role in protein folding by maintaining protein conformation and function (Hilton et al. 2013).

Stress granules are stress-induced messenger riboucleoprotein (mRNP) particles that consist of the small ribosomal subunit, translation initiation factors, poly(A)-binding protein (PABP), and other RNA-binding proteins (Anderson and Kedersha 2006; Mazroui et al. 2006; Uniacke and Zerges 2008; Buchan and Parker 2009). These granules function to sequester, silence and/or degrade RNA transcripts as part of a mechanism that adapts patterns of local RNA translation to facilitate the stress response (Buchan and Parker 2009; Adeli 2011; Wolozin 2012). Housekeeping transcripts are stored in stress granules whereas stress-responsive gene transcripts are excluded from these particles in heat- stressed cells (Kedersha and Anderson 2002). In plants, stress granules can be induced in chloroplast by associating with the large subunit of Rubisco in response to oxidative stress under high light (Uniacke and Zerges 2008). The stress granules are speculated to protect the mRNAs from heat-induced damage, though the mechanism of formation of stress granules remains elusive (Hu et al. 2010; Onomoto et al. 2012).

2. Small Regulatory RNAs: An Overview.

The *lin*-4 mutation in *Caenorhabditis elegans* was first shown to regulate developmental timing, through inhibiting the translation of the *Lin*-14 gene (Lee et al. 1993). This was the initial study that uncovered an evolutionarily conserved mechanism of regulating gene expression by non-coding small RNAs. The small regulatory RNA was named micro-RNA (miRNA). The

miRNAs are 20-24nt non-coding RNAs that post-transcriptionally regulate gene silencing in eukaryotes (Bartel 2004). In the past decade, thousands of miRNAs and a variety of other classes of small regulatory RNAs have been identified from various animal and plant sources (Kozomara and Griffiths-Jones 2011) and their functions have been explored in considerable details (Fehniger et al. 2010; Zhang et al. 2011b; Wu et al. 2012).

In higher plants, miRNAs regulate a variety of physiological processes, including vegetative growth, flower timing, and auxin signaling (Wu et al. 2006; Nag et al. 2009). Recent studies have shown that among the myriad of plant small regulatory RNAs, miRNAs play a critical role in the regulatory responses to various environmental stimuli (Trindade et al. 2010; Li et al. 2011; Zhu et al. 2011; Khraiwesh et al. 2012). Our current knowledge of small regulatory RNAs provide diverse classes and complex mechanisms of biogenesis and mode of action among these regulatory RNAs. Following is a brief overview of the small regulatory RNAs..

Small regulatory RNAs (srRNAs), are short non-protein coding RNAs that regulate gene expression by either transcriptional gene silencing (TGS) or posttranscriptional gene silencing (PTGS) (Khraiwesh et al. 2012). Based on their characteristics, origin and biogenesis, small RNAs are grouped into three major classes: hairpin RNAs (hpRNAs), short interfering (siRNA), and piwi-interacting siRNAs (piRNAs) (Axtell 2013). The hpRNAs can be divided into miRNAs and a variety of other hpRNAs or non-mRNA hpRNAs), and siRNA can be divided into heterochromatic siRNAs (hc-siRNAs), natural-antisense transcript siRNAs (nat-siRNA), and secondary siRNAs such as trans-acting siRNAs (ta-siRNA) (Axtell 2013). The piRNAs are a class of special small RNAs that are expressed in germ line cells and have only been found in animals (Sarot et al. 2004; Vagin et al. 2006; Axtell 2013).

The hpRNAs are produced from single-stranded hairpin precursors cleaved by DCL proteins (Axtell 2013). The miRNAs are produced by precise excision primarily by DCL1 that yield one or a few functional small RNAs, whereas non-miRNA hpRNAs are processed from hairpins by imprecise cleavage of DCL2-4 that produce small RNA population with size distribution of 21 to 24nt (Axtell 2013). Typically, precursors of non-miRNA hpRNAs are much larger than typical miRNAs (Axtell 2013).

In plants, canonical miRNAs are ~21nt in length, while a small subset of miRNAs are longer or shorter (Axtell et al. 2011). Generally, canonical miRNAs are produced by DCL1, and other classes are produced by different DLCs (DCL2-4) and associated with different AGO proteins (Axtell 2013). Canonical miRNAs usually repress gene expression either by inhibition of target gene translation or by guiding their cleavage (Chorostecki et al. 2012), while the 24nt long miRNAs can direct the methylation of target genes (Wu et al. 2010). Some canonical miRNAs can trigger the biogenesis of so-called secondary siRNAs, such as ta-siRNAs (Zhai et al. 2011; Xia et al. 2013).

Non-miRNA hpRNAs are initially observed in plants and produced from long inverted repeats (IRs) (Henderson et al. 2006). The hpRNAs can be processed by different DCLs depending upon their respective RNA polymerase and corresponding promoters (Wang et al. 2008). The 21nt small RNAs are typically produced by 35 S Pol line, whereas 24-and/or 22-nt small RNAs are produced by Pol III line (Wang et al. 2008). The 24nt small RNAs can direct mRNA degradation, and both 24nt and other small RNAs can induce histone cytosine methylation (Wang et al. 2008). These observations suggest that non-miRNA hpRNAs are regulatory small RNAs and that their processing depends upon respective promoters (Wang et al. 2008).

In animals, some non-miRNAs hpRNAs are produced from the loop of miRNA hairpins and can be recruited into Argonaute complexes to direct target repression (Okamura et al. 2013). The load of loop small RNAs into AGO complexes are highly regulated and conserved in flies and mammals (Okamura et al. 2013). Thus, hairpins can be processed by DCLs in at least two distinct manners, precision and imprecision. But the mechanism controlling the precise or imprecise cleavage of hairpins is not well understood.

2.1 miRNAs and siRNAs

Two major classes of small RNAs are the miRNAs and the siRNAs that are produced by a ribonuclease III-like protein referred to as Dicer that cleaves sRNAs from a double-stranded region of RNA (Bartel 2004; Khraiwesh et al. 2012). Although miRNAs and siRNAs are similar in size (21-24 nt) and composition, they differ by their origin and biogenesis (Bartel 2004; Chen 2009).

The endogenous siRNAs are more diverse and are 21-24nt in length. These are often derived from perfectly double-stranded precursors produced by the RNA polymerase IV or V and RNA-dependent RNA polymerase (RDR2) (Mosher et al. 2009; Haag et al. 2012; Zhang et al. 2012). The siRNAs represent largest family of small RNAs typically found in most eukaryotes, and can be classified into several functional groups such as heterochromatic siRNAs (hc-siRNAs) also known as repeat associated siRNAs (ra-siRNAs), tans-acting RNAs (ta-siRNAs), natural antisense siRNAs (nat-siRNAs), and long siRNA (lsiRNA). The hc-siRNAs, also referred to as ra-siRNAs, are typically 23nt-24nt in length, and are produced from repeat-regions of the chromosome (Pikaard et al. 2008). These siRNAs are specifically involved in transcriptional gene silencing by sequence specific DNA methylation and are particularly

important in maintaining silencing of transposable element sequences in the heterochromatic regions of chromosomes utilizing DNA-methylation (Blevins et al. 2009; Zhai et al. 2011). The 24-nucleotide hc-siRNAs or ra-siRNAs are produced by the PolIV/RDR2/DCL3 pathway and function in RNA-directed DNA methylation and histone modifications (Matzke and Birchler 2005; Zaratiegui et al. 2007; Matzke et al. 2009; Wierzbicki et al. 2009). The ta-siRNAs are generally 21nt in length, and are produced by *TAS* genes (Yoshikawa et al. 2005). The *TAS* genes are only found in plants and are not known to encode any proteins (Yoshikawa et al. 2005). The processing of *TAS* transcripts is trigged by certain miRNAs, and requires RNA-dependent RNA polymerase (RdRP) (Yoshikawa et al. 2005). The ta-siRNAs require DCL4, RDR6, and a miRNA to mediated cleavage to initiate processing (Howell et al. 2007).

The nat-siRNAs are likely the most abundant group of siRNAs (Borsani et al. 2005; Katiyar-Agarwal et al. 2007; Zubko and Meyer 2007; Jin 2008; Ron et al. 2010; Zhang et al. 2012), and are produced from many diverse locations in the genome utilizing the RNAi pathway described above. Functionally, these RNAs are more likely involved in biotic and abiotic stress in plants (Borsani et al. 2005; Katiyar-Agarwal et al. 2006; Jin et al. 2008; Zhang et al. 2012). Some nat-siRNAs accumulate in specific developmental stages (Zubko and Meyer 2007; Ron et al. 2010). The biogenesis of salt- and bacterium-induced nat-siRNAs appear to require DCL1 and/or DCL2, RDR6, and PolIV in Arabidopsis (Borsani et al. 2005; Katiyar-Agarwal et al. 2006). Moreover, the expression of ARIADNE14 is de-repressed in dcl1, hen1, hyl1, sde4, rdr2 and sgs3, suggesting that the nat-siRNAs generated from the ARIADNE14 /KOKOPELLI overlapping pair is dependent on DCL1, HEN1, HYL1, RDR2, SGS3 and PolIV (Ron et al. 2010). The nat-siRNAs have been found in budding yeast and *Schistosoma japonicum* (Drinnenberg et al. 2009; Cai et al. 2011). Despite these reports, there remains a need for

conclusive evidence for the occurrence of nat-siRNAs in plants (Henz et al. 2007), because their features are not well understood.

Table 1. A summary of miRNAs and siRNAs in plants

Small RNAs	**Length (nt)**	**Enzyme for transcription**	**Enzyme for cleavage**	**Precursors**
miRNAs	21-24 (majority 21)	Pol II	DCL1	Hairpin-like precursors
hpRNAs	21, 22 or 24	Pol II or Pol III	DCL2-4	Hairpin-like precursors
ta-siRNAs	21	-	DCL4, SGS3 and RDP6	*TAS* gene
hc-siRNAs	23-24	Pol IV	RDP2, and DCL3	Repeat sequences
nat-siRNAs	24	Pol IV	DCL1, DCL2, or DCL3, and RDP6	Hybridization of complementary and independently transcribed RNAs

On the other hand, miRNAs are typically 21nt-22nt in length, and regulate gene expression post-transcriptionally (Bartel 2004; Khraiwesh et al. 2012). The miRNAs are generated from the double stranded stem of a hairpin loop structure that is transcribed as a single strand RNA using RNA polymerase II (Bartel 2004; Axtell et al. 2011).

2.2 Biogenesis of plant miRNAs.

The primary miRNAs (pri-miRNAs) are transcribed as long RNA precursor by RNA polymerase II from miRNA genes that are located in the intergenic regions between protein-coding genes, as shown in Figure 2 (Chen 2009). The pri-miRNAs are post-transcriptionally modified with 5'-cap and 3'-polyadenylate tail in a manner similar to mRNA (Chen 2009). The pri-miRNAs are processed by dicer protein to produce precursor miRNA (pre-miRNAs). Pre-miRNAs are capable of self-folding into hairpins, which is a key characteristic for bioinformatic identification of miRNAs (Fang and Spector 2007; Meyers et al. 2008). Pre-miRNAs are then cleaved by DCLs into miRNA-miRNA* (miRNA*=miRNA star, the complementary strand to the mature miRNAs) duplexes, with the assistance of two proteins: HYL1 (a double-stranded RNA binding protein) (Han et al. 2004) and SE (a zinc finger protein) (Yang et al. 2006). The miRNA-miRNA* duplexes are stabilized by 2-base methylated overhangs at 3' ends, which are generated the HEN1 protein (Yu et al. 2005). With the assistance of HSATY, miRNA-miRNA* duplexes are exported into cytoplasm (Park et al. 2005; Chen 2009). In the cytoplasm, mature miRNAs are loaded into AGO protein complex to form an RNA Induced Silencing Complex (RISC) that guides the degradation or translation inhibition of target mRNAs (Baumberger and Baulcombe 2005). Sequence partners of mature miRNAs, miRNA*s, are proposed to be degraded without further function (Brodersen and Voinnet 2009).

In Arabidopsis, there are four Dicer-like (DCL) proteins (named for the genes from which they derive, i.e. DCL1, DCL2, DCL3, and DCL4). These DCL protein homologs are believed to be responsible for miRNA biogenesis, but different DCLs produce different length miRNAs (Axtell et al. 2011). DCL1 is responsible for the biogenesis of 21nt miRNAs, while DCL2, DCL3, and DCL4 generate 22nt, 24nt, and 21nt miRNAs, respectively. Different length miRNAs appear to be associated with different AGO clades. For example, 21nt miRNAs are

associated with AGO1/10/7, while 24nt miRNAs is loaded into AGO4/6/9 complex (Axtell et al. 2011; Manavella et al. 2011). Some studies have shown that the first base of miRNAs may be crucial for miRNA loading to AGO proteins (Wang et al. 2010; Manavella et al. 2011).

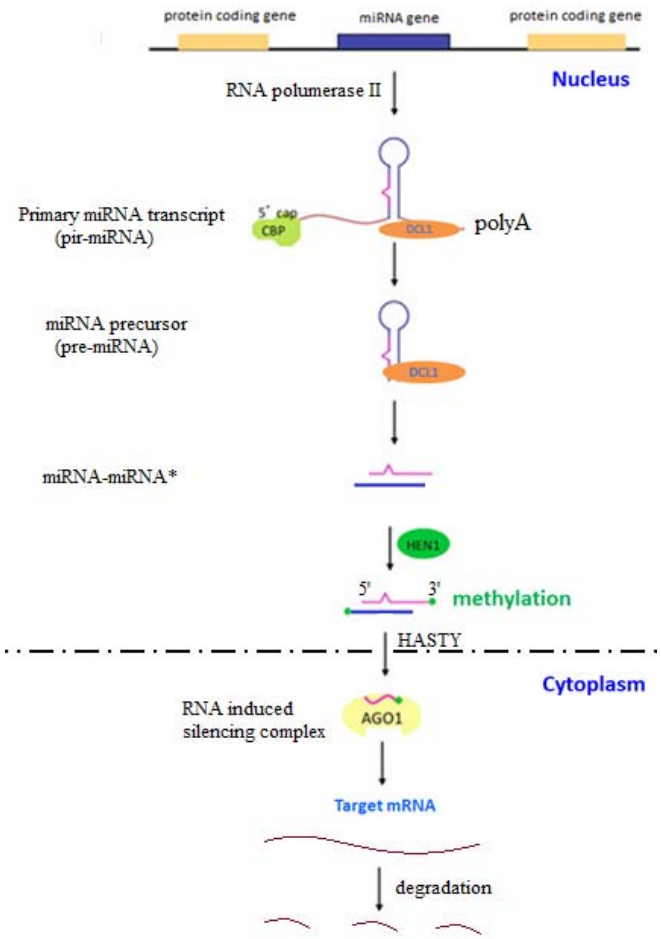


Figure 1. Biogenesis of plant miRNAs, modified from Chen (2009)

The primary miRNA transcripts (pri-miRNAs) are transcribed by RNA polymerase II and modified by adding 5' cap and 3' poly adenine (A). Pri-miRNAs are cleaved into pre-miRNAs (miRNA precursors) and then mina-miRNA* duplexes characterized with two base overhangs. The duplexes are methylated by HEN1 at the 3' ends, and then exported into cytoplasm by HASTY (HST). The mature miRNAs are incorporated into RNA-induced silencing complex with the core of AGO (Argonaute), which functions to slice the target mRNAs.

Although plant and animal miRNAs share some similarities in functions, these are distinct in their origins and biogenesis pathways (Axtell et al. 2011). The biogenesis of plant miRNAs is completed in nucleus. First, plant pri-miRNAs are transcribed by RNA polymerase II, and processed into pre-miRNAs by Dicer-like proteins (primarily DCL1) and consequently miRNA-miRNA* duplexes are produced (Axtell et al. 2011). In contrast, the maturation of animal miRNAs is primarily completed in cytoplasm. Most of animal pri-miRNAs are transcribed by RNA polymerase II (Kim et al. 2009), though some of animal pri-miRNAs are products of RNA polymerase III (Borchert et al. 2006). The pre-miRNAs of animals are exported into cytoplasm and subsequently processed to miRNA-miRNA duplexes by Dicer (Kim et al. 2009). In addition, plant miRNA-miRNA* complexes are methylated at 3'-ends for preventing degradation (Yu et al. 2005), whereas most products of animal miRNAs are not (Axtell et al. 2011). Also plant miRNA precursors vary in length and minimum free energy (MFE) compared to animal pre-miRNAs (Thakur et al. 2011; Xuan et al. 2011). For example, pre-miRNAs in plants vary from 60 to several hundred nucleotides, whereas most of animal pre-miRNAs are 70-80nt (Zhang et al. 2006a; Thakur et al. 2011; Yang and Li 2011).

2.3 Recognition of miRNA-target

The miRNAs regulate gene expression either by inhibition of translation or by degradation of mRNAs, and the specificity of this regulation is determined by the complementarity between miRNAs and their targets (Bartel 2004). When miRNAs show perfect or near-perfect complementary match to their target mRNA, the target mRNAs are cleaved and degraded (Brodersen and Voinnet 2009). Alternatively, if the mRNA does not have a perfect match to miRNA, translational inhibition rather than cleavage can occur (Zeng et al. 2003; Bartel 2004; Dai et al. 2011). In some cases, translational repression can also occur even with near-

perfect complementarity between the target and the miRNA (Gandikota et al. 2007). For example, a mutation in miR398 does not affect the accumulation target genes (*CSD1* and *CSD2*) (Dugas and Bartel 2008). Alternatively, translational repression can also occur even with near-perfect complementarity between targets and miRNAs (Gandikota et al. 2007). Target sites of plant miRNAs are predominantly found in coding regions, but can be located in UTRs regions, as well (Brodersen and Voinnet 2009).

In both animals and plants, miRNAs recognize their targets through "seed" sequences (Wang et al. 2010). The core of the seed sequences resides between nucleotides 2-7 measured from the 5'-end of the guide strand, which sometimes can include the nucleotide in position 8 (Wang et al. 2010). However, in some cases a single nucleotide change beyond the seed with central pairing can be important for miRNA-target recognition (Wang et al. 2010). It has also been observed that both an unpaired "bulge" and GU-wobble are present in plant miRNA/target-interaction (Jones-Rhoades and Bartel 2004).

2.4 Regulatory Roles of miRNAs in Plants.

Plant Growth and Development. The functions of plant miRNAs have been intensively reviewed (Bartel 2004; Yang et al. 2007; Chen 2009; Cuperus et al. 2011). In plants, miRNAs are involved in a variety of physiological processes, including growth and development, auxin homeostasis, and biotic and abiotic stress responses (Abdel-Ghany and Pilon 2008; Meng et al. 2010; Zhu et al. 2011; Yu et al. 2012). The functions of plant miRNAs have been determined using transgenic plants overexpressing miRNA genes (Chuck et al. 2011) or using gain-of-function mutants in which miRNA-resistant mutant genes are expressed (Jung et al. 2009). The targets of miR165/166 are PHABULOSA (PHB) and PHAVOLUTA (PHV), two genes that

control leaf polarity (Mallory et al. 2004). When a single mismatch is added to the complementary region between miR165 and the PHB mRNA, transgenic plants have radicalized leaves with ectopic patches of adaxial tissue on the abaxial surface, and ectopic meristems (Mallory et al. 2004).

The first identified miRNA in Arabidopsis, *MIR-JAW* encoded by *jaw*-D, guides messenger RNA cleavage of several *TCP* genes (*TCP*1-4) controlling leaf development (Palatnik et al. 2003). In the dominant mutant of jaw-*D* in which *MIR-JAW* is overexpressed, the expression of *TCP* genes is reduced (Palatnik et al. 2003) . As a results, the leaves are not flat, but are epinastic and "crinkly" similar to the *CIN* gene mutation of *Antirrhinum majus* (Palatnik et al. 2003). The constitutive expression of *TCP2* and *TCP4* partially rescues the leaf defects of *jaw-D* (Palatnik et al. 2003). In addition , a mutant of *TCP2* is able to rescue the leaf shape and curvature defects of jaw-D more completely than overexpression of *TCP2* (Palatnik et al. 2003), and a TCP4 overexpression mutant has longer hypocotyles and smaller leaves than wild type, which is similar to plants with *TCP4* mutation (Palatnik et al. 2003).

Jaw-D is currently known as miR319a that regulates the development of sepals through *TCP4* as a target (Nag et al. 2009). A miR319a mutant, miR319a^{129}, displays aberrant development of petals, while wild-type miR319a is able to rescue this defect. GUS activity tests show that there is no functional-overlapping among miR319a, miR319b, and miR319c. The mutant of miR319a129 (G→A at the 12th base) causes miR319a^{129} to lose targeting ability to *TCP4* mRNA. However, introduction of a mutated recognition site for miR319a^{129} into TCP4 mRNA suppresses the phenotypic defect. Therefore, miR319a functions to regulate the development of petals by controlling posttranscriptional expression of the *TCP4* gene (Nag et al. 2009). In addition, miR319 also participates in plant stress response, such as salt and drought, et

al. (Zhou et al. 2013). Overexpressing osa-miR319a (rice miR319a) in Creeping bentgrass (*Agrostis stolonifera* 'Penn A-4') alters morphological changes and exhibits enhanced tolerance to salt and drought via regulating at least putative target genes belonging to TCP family, AsPCE5-8 and AsTCP14 (Zhou et al. 2013).

The miR172 regulates the expression of the APETALA2 (AP2) transcription factor gene in Arabidopsis by regressing AP2 translation (Aukerman and Sakai 2003; Wollmann et al. 2010; Zhu and Helliwell 2011). The AP2 family belongs to class A gene in the ABC model of floral organ identities. The overexpression of EAT encoding miR172a-2 causes extremely early flowering and floral organ identity defects (Aukerman and Sakai 2003). In addition, miR172 has also acquired specialized species-specific functions in other aspects of plant development such as cleistogamy and tubeerization (Zhu and Helliwell 2011).

Auxin, regulates almostmany different aspects of plant growth and development (Perrot-Rechenmann 2010). At least three distinct miRNAs, miR393, miR160 and miR167, have been demonstrated to be involved in auxin responses via regulating either TAAR (TRANSPORT INHIBITOR RESPONSE1 / AUXIN SIGNALING F-BOX1 AUXIN RECEPTOR) genes (encoding a crucial proteins for perceiving the auxin signal) or auxin response transcription factors (ARFs that transcriptionally activate the expression of auxin response genes) (Si-Ammour et al. 2011; Windels and Vazquez 2011). miR393 regulates the expression TAAR genes via producing secondary siRNAs (so-called siTAARs) that mediate expression of TAAR genes (Si-Ammour et al. 2011). miR160 and miR167 target ARFs and consequently control the expression of auxin response genes (Mallory et al. 2005; Wu et al. 2006).

Stress-responsive miRNAs. Plant responses to abiotic stresses are regulated by a complicated network of various factors. Recent studies have demonstrated that post-transcriptional regulation of gene expression by miRNAs or siRNAs is crucial for environmental stress adaption/ tolerance of plants (Gao et al. 2010; Giacomelli et al. 2012; Khraiwesh et al. 2012).

Nutrient Deficiency and Toxicity Stresses. miRNAs are responsive to various nutrient deficiency, including phosphate deficiency (Chiou et al. 2006) and sulfate deficiency (Jones-Rhoades and Bartel 2004). The miR399, the first miRNA demonstrated to respond phosphate deficiency via regulating the expression of the PHO gene (Aung et al. 2006; Chiou et al. 2006; Valdes-Lopez et al. 2008; Kuo and Chiou 2011). In Arabidopsis, the PHO gene (PHOSPHATE), also known as UBC24, encodes E2 enzyme that functions in Pi translocation and remobilization (Aung et al. 2006). Upon Pi deficiency, the transcripts of the PHO gene are cleaved under the direction of miR399, which is crucial for Pi hemostasis (Bari et al. 2006; Chiou and Lin 2011). This process also is regulated by transcription factors, such as PHR1 (phosphate response gene) (Valdes-Lopez et al. 2008). These studies suggest that plants possess a possible Pi signaling pathway consisting of PHO, PHR1, and miR399 (Bari et al. 2006; Wu and Wang 2008; Yang and Finnegan 2010) .

The miR395 regulates the plant response to sulfate deficiency among other things (Jones-Rhoades and Bartel 2004). Under sulfate starvation, miR395 is up-regulated to represses the expression of ATP sulfurylase genes (Jones-Rhoades and Bartel 2004; Liang et al. 2010; Matthewman et al. 2012). In addition, miR395 might be involved in the allocation of sulfate, because the shoots in the transgenic plants overexpressing miR395 still display sulfate deficiency even though the over-accumulation of sulfate is found in the shoot (Liang et al. 2010).

Matthewman et al. (Matthewman et al. 2012) show that miR395 participates not only in the uptake of sulfate, but also in the assimilation of sulfate.

A number of miRNAs are involved in plant response to the copper availability (Abdel-Ghany and Pilon 2008; Lu et al. 2011). Copper is an essential micronutrient for all living organisms (Puig et al. 2007). In higher plants, two major copper proteins are plastocynain participating in photosynthetic electron transport in the thylakoid lumen of chloroplasts (Yamasaki et al. 2008), and Cu/Zn superoxide dismutase (Cu/Zn CSD) localized to the cytoplasm, the chloroplast stroma and peroxisome (Kliebenstein et al. 1998). Under copper deficiency, expression of Cu/Zn CSD is inhibited through the cleavage directed by miR398 (Yamasaki et al. 2007). In the copper-replete plants, almost no miR498 is observed, while miRNA498 is highly accumulated when copper is deficient or limited (Abdel-Ghany and Pilon 2008). In addition, miR397 and miR857 are strongly expressed in stem and root tissues in response to low levels of copper (Abdel-Ghany and Pilon 2008). Lu et al. report that at least four miRNAs of *Populus trichocarpa*, miR397, miR398, miR408 and miR144, are responsive to Cu-deficiency (Lu et al. 2011).

Oxidative stress. The miR398 is a conserved miRNA family present in most seed plants (Cuperus et al. 2011) that is responsive to and appears to regulate plant oxidative stress (Sunkar et al. 2006; Zhu et al. 2011). miR398 targets two closely related Cu/Zn superoxide dismutase genes (cytosolic CSD1 and chloroplastic CSD2) (Sunkar et al. 2006)(ref). As discussed above in section 1, Cu/Zn SODs are crucial enzymes that are responsible for oxidative stress responses. The expression of CSDs is regulated by miR398-directed cleavage of the CSD (Sunkar et al. 2006; Guan et al. 2013)(ref). Overexpression of miR398 resulted inresistant form of CSD2

accumulate more CSD2 than plants overexpressing nonmutant CSD2 and consequently show greater tolerance to oxidative stress (Sunkar et al. 2006).

Lyer et al. (Iyer et al. 2012), demonstrated that 22 miRNAs are differentially expressed early in response to ozone (a model abiotic elicitor of reactive oxygen species) in Arabidopsis. They identified several stress responsive cis-elements that were enriched in the promoters of ozone responsive genes through in silico promoter analysis of miRNA gene.

miRNAs involved in Heat Stress. In wheat, the expression levels of 24 and 12 miRNAs are altered in response to powdery mildew infection and heat stress, respectively (Xin et al. 2010). Through NGS sequencing, Chen, et al., identified 52 miRNAs from 15 families are responsive to heat stress in *Populus*, and most of them are down-regulated (Chen et al. 2012). Five miRNA families, including miR398, are responsive to heat stress in cabbage (*Brassica rapa*) (Yu et al. 2012). Under heat stress, miR398 is heat-inhibitive and regulates Brac*CSD1*, and miR156 is heat-induced and its putative target Brac*SPL2* is down-regulated (Yu et al. 2012). Based these studies, some miRNAs, such as, miR398, are implicated in plant responses to heat. One of the possible mechanisms is to regulate the expression of antioxidant proteins, such as SODs. But it may not be the only pathway for regulating plant heat tolerance. Further studies are needed to understand the regulatory roles of miRNAs in plant heat tolerances. The data to date only demonstrate that changes in certain miRNAs correlate with heat stress, and do not unequivocally demonstrate a link between heat stress tolerance and the expression of any miRNA.

The roles of various miRNAs in plant response to various stresses have been demonstrated, but most of these studies are limited to a few model species, such as, Arabidopsis,

or rice. Our current knowledge is rather limited about regulatory roles of miRNAs in non-model plants, such as, cotton and other economically important plants species to potentially take advantage of this new and emerging tool in crop improvement.

2.5 Identification and nomenclature of miRNAs in plants

Deep sequencing or next generation sequencing of RNA has vastly revolutionized our ability to identify and analyze RNAs of various length (Ansorge 2009; Malone and Oliver 2011; Williamson et al. 2013). This technology allows us to simultaneously produce billions of sequences, which is highly efficient and low cost for DNA or RNA analysis (Zhou et al. 2010).

Typically, miRNAs can be divided into conserved and non-conserved miRNAs, according to their conservation in plants (Cuperus et al. 2011). Alternatively, miRNAs that have been reported or identified in specific species are called known miRNAs. But some miRNAs with identical sequences but generated from different genomic loci are assigned the same number with sequential alphabetical suffixes, such as, a, b and c, etc. according to the order of discovery (Meyers et al. 2008). Due to historical reasons, some miRNAs, such as miR156/miR157, miR165/166, and miR170/171, are assigned different numbers, but in reality they should belong to the same families (Meyers et al. 2008). The strand complementary to miRNAs are called miRNA*s (miRNA stars), but recently miRBase has changed the star (*) and instead assigns these by their arms, such as 3p (3' end) or 5p (5' end). Novel miRNAs represents three types of miRNAs, undiscovered in plants, undiscovered in specific species, or discovered in new genomic loci. But novel miRNAs can be conserved miRNAs, or identical to known miRNAs in sequence.

Identification of known and conserved miRNAs. Identification of known miRNAs is quite straight forward and can be completed by mapping to miRNA loci previously reported in the genomes (Meyers et al. 2008; Mackowiak 2011) using open-source mapping software, such as Bowtie (Langmead 2010) and Blastn (Boratyn et al. 2013). Alternatively, a subset of highly conserved miRNAs in plants (Cuperus et al. 2011) can be identified by sequence comparison with conserved miRNAs found in other plant species (Zhang et al. 2005; Sunkar and Jagadeeswaran 2008; Zhang et al. 2011a).

The miRBase, a collective database of miRNAs (http://www.mirbase.org), has become be an important source of miRNAs for both animals and plants (Kozomara and Griffiths-Jones 2011). But a large proportion of miRNAs registered in miRBase appear to be not authentic (Jeong et al. 2011). It has been suggested that the miRNA list in miRBase should be carefully refined (Meng et al. 2012). Thus, sequence similarity or homology is insufficient for annotation of authentic miRNAs (Meyers et al. 2008).

Identification of miRNA*s. The miRNA*s are complementary sequence of miRNAs, and the base-pairing between miRNAs and miRNA*s is crucial for identification of plant miRNAs (Meyers et al. 2008). miRNA-miRNA* complexes produced by DCL proteins are characterized as two base overhangs at 3' ends (Axtell et al. 2011), which is the major basis identifying miRNA*s. We have developed a Perl script that can identify corresponding miRNA* sequences, which facilitates the identification of miRNAs.

Identification of novel miRNAs. Identification of novel miRNAs are required to meet the criteria of plant miRNAs (Meyers et al. 2008). Computational programs or scripts are designed to search potential hairpins that can be qualified as miRNA precursors (Jones-Rhoades and

Bartel 2004; Friedlander et al. 2008). In animals, many computational methods have been developed, such as miRDeep (Friedlander et al. 2008) and MIReNA (Mathelier and Carbone 2010). But these methods are designed specifically for animal miRNAs, such as human and dog (Friedlander et al. 2008), and cannot be directly applied in plants, because of distinct characteristics of plant miRNAs (Thakur et al. 2011; Yang and Li 2011). In plants, algorithms, such as, MIRCheck (Jones-Rhoades and Bartel 2004) and miRDeep-P (Yang and Li 2011), have been developed and applied in analyzing miRNA transcriptome data, as follows.

The MIRCheck algorithm is the first algorithm designed specifically for plant miRNAs (Jones-Rhoades and Bartel 2004). It is a Perl script that computationally identifies 20mers within the potential to encode plant miRNAs (Jones-Rhoades and Bartel 2004). MIRCheck extracts the imperfectly inverted repeats and evaluates the secondary structure by examining the base-pairing between miRNA and miRNA* (Jones-Rhoades and Bartel 2004). This algorithm has been applied in Arabidopsis, rice, and other plants (Jones-Rhoades and Bartel 2004; Ma et al. 2010; Schreiber et al. 2011).

The miRDeep-P is a recently developed algorithm that is modified from miRDeep, based on the characteristics of plant miRNAs (Yang and Li 2011). The algorithm of miRDeep uses a probabilistic model of miRNA biogenesis to score compatibility of the position and frequency of sequenced RNA with the secondary structure of the miRNA precursor, based on data generated from human and dog (Friedlander et al. 2008). Due to lower minimum free energy of plant miRNAs that leads to lower miRDeep scores, many plant miRNAs cannot pass the scoring system of miRDeep. To use miRDeep in plants, Yang and Li reduced the scoring threshold of miRDeep scoring system for allowing plant miRNAs (Yang and Li 2011). This algorithm has

been used in Arabidopsis, potato, and peach, et al. (Yang et al. 2011; Colaiacovo et al. 2012; Zhang et al. 2013).

The miREAP (http://sourceforge.net/projects/mireap/) is another algorithm that has been used for miRNA. This algorithm is developed to detect and score Dicer hairpin products (Jeong et al. 2011). A crucial aspect of miREAP is to evaluate secondary structures by examining the base-paring of miRNA:miRNA* (Jeong et al. 2011). This algorithm has been used in *Medicago*, rice, cassava, and cotton, et al. (Jeong et al. 2011; Perez-Quintero et al. 2012; Yin et al. 2012).

Li et al. investigated the performance of eight computational programs including miRDeep and miREAP, and found that miREAP is better than miRDeep in efficiency and effectiveness (Li et al. 2012). In our preliminary test, we found that miREAP had a relatively better performance than miRDeep-P (unpublished data). But miREAP produces a large number of false positives. For example, miREAP predicts 2,420 putative miRNAs from 14,469 unique sequences, and consequently only 106 are identified to authentic miRNAs (Zhai et al. 2011).

Although some computational methods have been developed for the identification of plant miRNAs, there are a large proportion of reported miRNAs that are not authentic (Jeong et al. 2011; Meng et al. 2012).

The majority of sequences in plant small RNA libraries are siRNAs, with miRNAs representing only a very small proportion of total sequence (Xie et al. 2005; Kasschau et al. 2007; Zhai et al. 2011). The RDR2 (RNA-dependent RNA polymerase 2) is a crucial protein involved in the biogenesis of 24nt siRNAs (Xie et al. 2004). In wild type, 24nt siRNAs are the most abundant class, but dramatically reduced in *rdr2* mutants (Kasschau et al. 2007). In addition, so-called hairpin RNAs (non-miRNAs) and siRNAs often complicate the identification

of miRNAs (Jeong et al. 2011; Axtell 2013). For example, ~150 out of ~400 reported rice miRNAs in miRBase (a database of miRNAs, http://www.mirbase.org) are characterized as siRNAs and unable to assign as authentic miRNAs (Jeong et al. 2011).

Computational method for identification of novel miRNAs largely relies on the analysis of secondary structures of precursor RNAs. Thus, accuracy of secondary structures of precursors is crucial for authentic miRNAs (Zhai et al. 2011). RNAfold is very useful in analyzing secondary structures of RNAs, including miRNA precursors (Friedlander et al. 2008; Wan et al. 2012), but might produce less accurate secondary structures (Sato et al. 2009). This might lead to some false positives (Jeong et al. 2011).

Solution to reduce false positives: miRNAs are proposed to be characterized as strand bias and abundance bias, because miRNAs are produced from single stranded precursors (Jeong et al. 2011). Stand bias is a parameter that indicates the proportion of small RNAs matching the sense strand of miRNA precursors, and abundance bias indicates that the proportion of the most small RNA (miRNA candidate) and abundant (isomer of miRNA candidate) in the total of all small RNAs matching the corresponding precursors. Jeong et al. (2011) reported that 0.9 and 0.6 could be the threshold of strand bias and abundance bias, respectively. The hpRNAs are produced from imprecise excision probably by DCL proteins (Axtell 2013), and can be speculated to lack strand bias and abundance bias. In addition, conserved miRNAs have lower minimum free energy, and some energetic-related parameters, such as MFEI (minimum free energy index), and are able to distinguish miRNAs from other non-protein coding RNAs (Zhang et al. 2006b). These indicators have been applied in recent studies of miRNAs (Wang et al. 2011; Yang et al. 2012).

Finally, use of CentroidFold for the analysis, produce more accurate secondary structure results (Jeong et al. 2011; Zhai et al. 2011). By examining Arabidopsis miRNAs registered in miRBase, some dead miRNAs that are to be removed from miRBase cannot form hairpins using CentroidFold (unpublished data). Although a large number of miRNAs have been identified primarily through deep sequencing, several pieces of evidences show that a relatively large proportion of those are not authentic (Jeong et al. 2011). This suggests that new, operational criteria should be proposed to uniform the annotation of plant miRNAs. Use of CentroidFold software might help improve the true positive and reduce false positives (Jeong et al. 2011; Zhai et al. 2011).

3.0. Summary and Proposal

Small regulatory RNAs from plants can be classified into two distinct families, short interfering RNAs (siRNAs) and hairpin RNAs (hpRNAs), on the basis of their origin and function (Axtell 2013). siRNAs arise from linear double stranded precursors, while hpRNAs are produced from single-stranded, hairpin-shaped precursor (Axtell 2013).

Based on their biogenesis and function, siRNAs can be divided into three major subgroups, repeat associated siRNAs (ra-siRNA) or heterochromatic siRNAs (hc-siRNAs), trans-acting siRNAs (ta-siRNAs), and natural antisense siRNAs (NAT-siRNAs) (Jamalkandi and Masoudi-Nejad 2009; Allen and Howell 2010). hc-siRNAs typically are 24 nt small RNAs that are produced from repetitive sequences by the RNA polymerase IV/RDP2/DCL3 pathway (Chapman and Carrington 2007; Xie and Qi 2008). ta-siRNAs are a class of secondary siRNAs that are produced from so-called *TAS* transcripts (Allen and Howell 2010). Their synthesis is dependent on a miRNA or siRNA trigger (Allen et al. 2005) and requires the involvement of

DCL4, RDP6 (RNA-dependent RNA polymerase6) and SGS3 (suppressor of gene silencing3) (Allen and Howell 2010).

NAT-siRNAs are produced from complementary mRNA transcripts (or natural antisense transcripts, NATs), and their synthesis is dependent on RDP6, and DCL1, DCL2 and/or DCL3 (Borsani et al. 2005; Zhang et al. 2012). There are two types of NAT-siRNAs, *cis*-NAT-siRNAs and *trans*-NAT-siRNAs (Borsani et al. 2005; Zhang et al. 2012). *cis*-NAT-siRNAs are produced from NATs transcribed from the same genomic loci, while *trans*-NAT-siRNAs arise from NATs that are transcribed from different genomic loci (Axtell 2013).

Based the characteristics of biogenesis, hpRNAs can be divided into two subgroups, micro RNAs (miRNAs) and other hpRNAs (ohpRNAs) (Axtell 2013). Canonical miRNAs are ~21nt small RNAs that are precisely cleaved by DCL1, and posttranscriptionally regulate the gene expression (Meyers et al. 2008; Axtell et al. 2011). Whereas so-called long miRNAs (lmiRNAs), ~24nt miRNAs, are cleaved by DCL3 and mediate DNA methylation (Wu et al. 2010). ohpRNAs arise from imprecisely/precise cleavage of one or two different DCL proteins from long inverted repeat sequences (up to several kb in length), leading to one or more sizes of small RNAs (Henderson et al. 2006; Dunoyer et al. 2010; Yelina et al. 2010; Axtell 2013). ohpRNAs mediate the expression of genes either in a transcriptional or in a posttranscriptional way (Henderson et al. 2006; Wang et al. 2008; Dunoyer et al. 2010). To date, the biogenesis of ohpRNAs remains largely unknown, but it appears that their processing pathway is determined by their promoter components (Wang et al. 2008).

miRNAs are a well-studied subclass of hpRNAs that are produced by precise cleavage of DCLs in plants and that function to regulate the expression of genes at the posttranscriptional

level (Huang et al. 2013; Rogers and Chen 2013). miRNAs are transcribed by RNA polymerase II from miRNA genes (Meyers et al. 2008), and the primary transcripts (pri-miRNAs) experience modification by adding 3'-ployadenosine and a 5'-cap (Chen 2009). Pri-miRNAs are processed by DCLs into precursor miRNAs, and then the miRNA-miRNA* (miRNA star, the pattern strand of miRNA) duplexes (Naqvi et al. 2012). The miRNA-miRNA* duplexes are stabilized through methylation at 3'-ends by HEN1 (Yang et al. 2006) and then exported into cytoplasm by the Exportin 5 homologue, HASTY (Park et al. 2005). In cytoplasm, mature miRNAs are recruited into an Argonaute protein complex (RISC) in which they bind to target mRNAs by Watson-Crick base pairs (Rogers and Chen 2013). This consequently results in their degradation or in translation inhibition of target mRNAs (Rogers and Chen 2013).

Based on the information reviewed above, it is possible to annotate the complete set of sRNAs including hairpin RNAs for a species including each class of sRNAs. An efficiently way to accomplish this task is to produce a so called pipeline of scripts to accomplish this task. Such a pipeline has been created for the purpose of discovering and annotating the small regulatory RNAs within the genus *Gossypium*.

As reviewed above heat stress caused by elevated temperature adversely affects an array of physiological processes from anatomical, biochemical, to gene levels, and often leads to a drastic reduction in economic yield (Wahid et al. 2007). The mechanism of plant heat response has been well-documented at the biochemical and gene levels and has been recently reviewed in (Hasanuzzaman et al. 2013; Qu et al. 2013). The expression of genes regulated by small RNAs, such as miRNAs, at transcriptional and post transcriptional levels has been shown to be highly conserved during plant responses to a variety of stresses (Khraiwesh et al. 2012; Guan et al. 2013; Lu et al. 2013).

Studies have shown that miRNAs are specifically involved in plant responses to thermal stress (Sunkar et al. 2006; Xin et al. 2010; Khraiwesh et al. 2012; Yu et al. 2012). Of stress-responsive miRNAs, miR398 is a well-studied stress-responsive miRNAs and regulates plant tolerance to various oxidative stresses by controlling the expression of Cu/Zn superoxide dismutase (CSD) (Sunkar et al. 2006; Zhu et al. 2011). Overexpressing a miR398-resistant form of *CSD* genes induces stress tolerance due to the accumThulation of CSD mRNAs (Sunkar et al. 2006). In addition to miR398, Yu et al. show that miR156h/g of *Brassica rapa* are induced by heat stress and their target is *BracSPL2* (Squamosa Promoter Binding Protein-Like). This transcription factor is down-regulated by heat stress (Yu et al. 2012). Recently, some ohpRNAs were also shown to be responsive to heat stress and pathogen infection (Xin et al. 2011).

The second objective of the work summarized here will be to identify changes in each class of small RNAs of *Gossypium* as discovered and annotated above. Characterizing these changes on a genome-wide basis to generate genotype specific patterns of gene regulation by small regulatory RNAs that can make such results available for plant improvement is the ultimate goal of the work to be subsequently described.

Reference

Abdel-Ghany SE, Pilon M. 2008. MicroRNA-mediated systemic down-regulation of copper protein expression in response to low copper availability in Arabidopsis. *J Biol chem.* **283**(23): 15932-15945.

Adeli K. 2011. Translational Control Mechanisms in Metabolic Regulation: Critical Role of RNA binding proteins, miRNAs, and Cytoplasmic RNA granules. *Am J Physiol Endocrinol Metab*. **301**(6):E1051-64

Al-Whaibi MH, Siddiqui MH, Basalah MO. 2012. Salicylic acid and calcium-induced protection of wheat against salinity. *Protoplasma.* **249**(3): 769-778.

Aloni B, Peet M, Pharr M, Karni L. 2001. The effect of high temperature and high atmospheric CO2 on carbohydrate changes in bell pepper (Capsicum annuum) pollen in relation to its germination. *Physiol Plant.* **112**(4): 505-512.

Alscher RG, Erturk N, Heath LS. 2002. Role of superoxide dismutases (SODs) in controlling oxidative stress in plants. *J Exp Bot.* **53**(372): 1331-1341.

Anderson P, Kedersha N. 2006. RNA granules. *J Cell Biol.* **172**(6): 803-808.

Ansorge WJ. 2009. Next-generation DNA sequencing techniques. *New Biotech.* **25**(4): 195-203.

Apel K, Hirt H. 2004. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Ann Rev Plant Biol.* **55**: 373-399.

Asada K. 2006. Production and scavenging of reactive oxygen species in chloroplasts and their functions. *Plant Physiol.* **141**(2): 391-396.

Aukerman MJ, Sakai H. 2003. Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *Plant Cell.* **15**(11): 2730-2741.

Aung K, Lin SI, Wu CC, Huang YT, Su CL, Chiou TJ. 2006. pho2, a phosphate overaccumulator, is caused by a nonsense mutation in a microRNA399 target gene. *Plant Physiol.* **141**(3): 1000-1011.

Axtell MJ. 2013. Classification and Comparison of Small RNAs from Plants. *Ann Rev Plant Biol.* **64**: 137-159.

Axtell MJ, Westholm JO, Lai EC. 2011. Vive la difference: biogenesis and evolution of microRNAs in plants and animals. *Genome Giol.* **12**(4): 221.

Bari R, Datt Pant B, Stitt M, Scheible WR. 2006. PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants. *Plant Physiol.* **141**(3): 988-999.

Barnabas B, Jager K, Feher A. 2008. The effect of drought and heat stress on reproductive processes in cereals. *Plant Cell Environ.* **31**(1): 11-38.

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

Bassel GW, Lan H, Glaab E, Gibbs DJ, Gerjets T, Krasnogor N, Bonner AJ, Holdsworth MJ, Provart NJ. 2011. Genome-wide network model capturing seed germination reveals coordinated regulation of plant cellular phase transitions. *Proc Natl Acad Sci U S A.* **108**(23): 9709-9714.

Baumberger N, Baulcombe DC. 2005. Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proc Natl Acad Sci U S A.* **102**(33): 11928-11933.

Bilski P, Li MY, Ehrenshaft M, Daub ME, Chignell CF. 2000. Vitamin B6 (pyridoxine) and its derivatives are efficient singlet oxygen quenchers and potential fungal antioxidants. *Photochem Photobiol.* **71**(2): 129-134.

Blevins T, Pontes O, Pikaard CS, Meins FJ. 2009. Heterochromatic siRNAs and DDM1 independently silence aberrant 5S rDNA transcripts in Arabidopsis. *PloS One* **4**(6): e5932.

Bondino HG, Valle EM, Ten Have A. 2012. Evolution and functional diversification of the small heat shock protein/alpha-crystallin family in higher plants. *Planta.* **235**(6): 1299-1313.

Boonburapong B, Buaboocha T. 2007. Genome-wide identification and analyses of the rice calmodulin and related potential calcium sensor proteins. *BMC Plant Biol.* **7**: 4.

Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, Madden TL, Matten WT, McGinnis SD, Merezhuk Y et al. 2013. BLAST: a more efficient report with usability improvements. *Nucleic Acids Res.* **41**(Web Server issue):W29-33

Borchert GM, Lanier W, Davidson BL. 2006. RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol.* **13**(12): 1097-1101.

Borsani O, Valpuesta V, Botella MA. 2001. Evidence for a role of salicylic acid in the oxidative damage generated by NaCl and osmotic stress in Arabidopsis seedlings. *Plant Physiol.* **126**(3): 1024-1030.

Borsani O, Zhu J, Verslues PE, Sunkar R, Zhu JK. 2005. Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell.* **123**(7): 1279-1291.

Bouche N, Fait A, Bouchez D, Moller SG, Fromm H. 2003. Mitochondrial succinic-semialdehyde dehydrogenase of the gamma-aminobutyrate shunt is required to restrict levels of reactive oxygen intermediates in plants. *Proc Natl Acad Sci U S A.* **100**(11): 6843-6848.

Bouche N, Fromm H. 2004. GABA in plants: just a metabolite? *Trends Plant Sci.* **9**(3): 110-115.

Brodersen P, Voinnet O. 2009. Revisiting the principles of microRNA target recognition and mode of action. *Nat Rev Mol Cell Biol.* **10**(2): 141-148.

Buchan JR, Muhlrad D, Parker R. 2008. P bodies promote stress granule assembly in Saccharomyces cerevisiae. *J Cell Biol.***183**(3): 441-455.

Buchan JR, Parker R. 2009. Eukaryotic stress granules: the ins and outs of translation. *Mol Cell.* **36**(6): 932-941.

Buchan JR, Yoon JH, Parker R. 2011. Stress-specific composition, assembly and kinetics of stress granules in Saccharomyces cerevisiae. *J Cell Sci.* **124**(Pt 2): 228-239.

Cabiscol E, Piulats E, Echave P, Herrero E, Ros J. 2000. Oxidative stress promotes specific protein damage in Saccharomyces cerevisiae. *J Biol Chem.* **275**(35): 27393-27398.

Cai P, Hou N, Piao X, Liu S, Liu H, Yang F, Wang J, Jin Q, Wang H, Chen Q. 2011. Profiles of small non-coding RNAs in Schistosoma japonicum during development. *PLoS Negl Trop Dis.* **5**(8): e1256.

Camp VW, Capiau K, Montagu VM, Inze D, Slooten L. 1996. Enhancement of oxidative stress tolerance in transgenic tobacco plants overproducing Fe-superoxide dismutase in chloroplasts. *Plant Physiol.***112**(4): 1703-1714.

Cao J, Barbosa JM, Singh NK, Locy RD. 2013. GABA shunt mediates thermotolerance in Saccharomyces cerevisiae by reducing reactive oxygen production. *Yeast* **30**(4): 129-144.

Cao WH, Liu J, He XJ, Mu RL, Zhou HL, Chen SY, Zhang JS. 2007. Modulation of ethylene responses affects plant salt-stress responses. *Plant Physiol.***143**(2): 707-719.

Cao WH, Liu J, Zhou QY, Cao YR, Zheng SF, Du BX, Zhang JS, Chen SY. 2006. Expression of tobacco ethylene receptor NTHK1 alters plant responses to salt stress. *Plant Cell Environ.***29**(7): 1210-1219.

Cao YR, Chen SY, Zhang JS. 2008. Ethylene signaling regulates salt stress response: An overview. *Plant Signal Beh.* **3**(10): 761-763.

Cheikh N, Jones RJ. 1994. Disruption of Maize Kernel Growth and Development by Heat Stress (Role of Cytokinin/Abscisic Acid Balance). *Plant Physiol.***106**(1): 45-51.

Chen C, Wan H, Zhou Q. 2010. The next generation sequencing technology and its application in cancer research. *Ch J Lung Cancer.* **13**(2): 154-159.

Chen L, Ren Y, Zhang Y, Xu J, Sun F, Zhang Z, Wang Y. 2012. Genome-wide identification and expression analysis of heat-responsive and novel microRNAs in Populus tomentosa. *Gene.* **504**(2): 160-165.

Chen PY, Lee KT, Chi WC, Hirt H, Chang CC, Huang HJ. 2008. Possible involvement of MAP kinase pathways in acquired metal-tolerance induced by heat in plants. *Planta.* **228**(3): 499-509.

Chen YF, Randlett MD, Findell JL, Schaller GE. 2002. Localization of the ethylene receptor ETR1 to the endoplasmic reticulum of Arabidopsis. *J Plant Chem.***277**(22): 19861-19866.

Chen X. 2009. Small RNAs and their roles in plant development. *Ann Revi Cell Dev Biol.* **25**: 21-44.

Cheval C, Aldon D, Galaud JP, Ranty B. 2013. Calcium/calmodulin-mediated regulation of plant immunity. *Biochim Biophysi.* **1833**(7):1766-71.

Chiou TJ, Aung K, Lin SI, Wu CC, Chiang SF, Su CL. 2006. Regulation of phosphate homeostasis by MicroRNA in Arabidopsis. *Plant Cell.* **18**(2): 412-421.

Chiou TJ, Lin SI. 2011. Signaling network in sensing phosphate availability in plants. *Ann Rev Plant Biol.* **62**: 185-206.

Choi MS, Kim MC, Yoo JH, Moon BC, Koo SC, Park BO, Lee JH, Koo YD, Han HJ, Lee SY et al. 2005. Isolation of a calmodulin-binding transcription factor from rice (*Oryza sativa* L.). *J Plant Chem.* **280**(49): 40820-40831.

Cholewa E, Cholewinski AJ, Shelp BJ, Snedden WA, Bown AW. 1997. Cold shock-stimulated gamma-aminobutyrate synthesis is mediated by an increase in cytosolic Ca^{2+}, not by an increase cytosolic H^{+}. *Can J Bot.* **75**: 375-382.

Chorostecki U, Crosa VA, Lodeyro AF, Bologna NG, Martin AP, Carrillo N, Schommer C, Palatnik JF. 2012. Identification of new microRNA-regulated genes by conserved targeting in plant species. *Nucleic Acids Res.* **40**(18): 8893-8904.

Chuck GS, Tobias C, Sun L, Kraemer F, Li C, Dibble D, Arora R, Bragg JN, Vogel JP, Singh S et al. 2011. Overexpression of the maize Corngrass1 microRNA prevents flowering, improves digestibility, and increases starch content of switchgrass. *Proc Natl Acad Sci U S A.* **108**(42): 17550-17555.

Colaiacovo M, Bernardo L, Centomani I, Crosatti C, Giusti L, Orru L, Tacconi G, Lamontanara A, Cattivelli L, Faccioli P. 2012. A Survey of MicroRNA Length Variants Contributing to miRNome Complexity in Peach (Prunus Persica L.). *Front Plant Cci.* **3**: 165.

Coller J, Parker R. 2004. Eukaryotic mRNA decapping. *Annu Rev Biochem.* **73**: 861-890.

Crafts-Brandner SJ, Law RD. 2000. Effect of heat stress on the inhibition and recovery of the ribulose-1,5-bisphosphate carboxylase/oxygenase activation state. *Planta.* **212**(1): 67-74.

Cutler SR, Rodriguez PL, Finkelstein RR, Abrams SR. 2010. Abscisic acid: emergence of a core signaling network. *Ann Rev Plant Biol.***61**: 651-679.

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

Dai X, Zhuang Z, Zhao PX. 2011. Computational analysis of miRNA targets in plants: current status and challenges. *Brief Bioinformatics.* **12**(2): 115-121.

Damgaard CK, Lykke-Andersen J. 2011. Translational coregulation of 5'TOP mRNAs by TIA-1 and TIAR. *Genes & development* **25**(19): 2057-2068.

Day IS, Reddy VS, Shad Ali G, Reddy AS. 2002. Analysis of EF-hand-containing proteins in Arabidopsis. *Genome Biol* **3**(10): RESEARCH0056.

De Las Rivas J, Barber J. 1997. Structure and thermal stability of photosystem II reaction centers studied by infrared spectroscopy. *Biochemistry* **36**(29): 8897-8903.

De Ronde JA, Cress WA, Kruger GH, Strasser RJ, Van Staden J. 2004. Photosynthetic response of transgenic soybean plants, containing an Arabidopsis P5CR gene, during heat and drought stress. *J Plant Physiol.***161**(11): 1211-1224.

DeFalco TA, Bender KW, Snedden WA. 2010. Breaking the code: Ca^{2+} sensors in plant signalling. *Biochem J.* **425**(1): 27-40.

Delaunay A, Isnard AD, Toledano MB. 2000. H_2O_2 sensing through oxidation of the Yap1 transcription factor. *The EMBO journal* **19**(19): 5157-5166.

Desikan R, Cheung MK, Bright J, Henson D, Hancock JT, Neill SJ. 2004. ABA, hydrogen peroxide and nitric oxide signalling in stomatal guard cells. *J Exp Bot.* **55**(395): 205-212.

Desikan R, S AH-M, Hancock JT, Neill SJ. 2001. Regulation of the Arabidopsis transcriptome by oxidative stress. *Plant Physiol.***127**(1): 159-172.

Deupi X, Standfuss J, Schertler G. 2012. Conserved activation pathways in G-protein-coupled receptors. *Biochem Soc Trans.* **40**(2): 383-388.

DeWald DB, Torabinejad J, Jones CA, Shope JC, Cangelosi AR, Thompson JE, Prestwich GD, Hama H. 2001. Rapid accumulation of phosphatidylinositol 4,5-bisphosphate and inositol 1,4,5-trisphosphate correlates with calcium mobilization in salt-stressed arabidopsis. *Plant Physiol.***126**(2): 759-769.

Di Paolo G, De Camilli P. 2006. Phosphoinositides in cell regulation and membrane dynamics. *Nature.* **443**(7112): 651-657.

Diaz J, Alvarez-Buylla ER. 2009. Information flow during gene activation by signaling molecules: ethylene transduction in Arabidopsis cells as a study system. *BMC Syst Biol.* **3**: 48.

Drinnenberg IA, Weinberg DE, Xie KT, Mower JP, Wolfe KH, Fink GR, Bartel DP. 2009. RNAi in budding yeast. *Science.* **326**(5952): 544-55

Du L, Ali GS, Simons KA, Hou J, Yang T, Reddy AS, Poovaiah BW. 2009. Ca^{2+}/calmodulin regulates salicylic-acid-mediated plant immunity. *Nature.* **457**(7233): 1154-1158.

Dubiella U, Seybold H, Durian G, Komander E, Lassig R, Witte CP, Schulze WX, Romeis T. 2013. Calcium-dependent protein kinase/NADPH oxidase activation circuit is required for rapid defense signal propagation. *Proc Natl Acad Sci U S A.* **110**(21): 8744-8749.

Dugas DV, Bartel B. 2008. Sucrose induction of Arabidopsis miR398 represses two Cu/Zn superoxide dismutases. *Plant Mo Biol.* **67**(4): 403-417.

Duysens LN, Amesz J, Kamp BM. 1961. Two photochemical systems in photosynthesis. *Nature.* **190**: 510-511.

Evrard A, Kumar M, Lecourieux D, Lucks J, von Koskull-Doring P, Hirt H. 2013. Regulation of the heat stress response in Arabidopsis by MPK6-targeted phosphorylation of the heat stress factor HsfA2. *PeerJ.* **1**: e59.

Falcone DL, Ogas JP, Somerville CR. 2004. Regulation of membrane fatty acid composition by temperature in mutants of Arabidopsis with alterations in membrane lipid composition. *BMC Plant Biol.* **4**: 17.

Fang Y, Spector DL. 2007. Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living Arabidopsis plants. *Curr Biol.* **17**(9): 818-823.

Fehniger TA, Wylie T, Germino E, Leong JW, Magrini VJ, Koul S, Keppel CR, Schneider SE, Koboldt DC, Sullivan RP et al. 2010. Next-generation sequencing identifies the natural killer cell microRNA transcriptome. *Genome Res.* **20**(11): 1590-1604.

Finka A, Cuendet AF, Maathuis FJ, Saidi Y, Goloubinoff P. 2012. Plasma membrane cyclic nucleotide gated calcium channels control land plant thermal sensing and acquired thermotolerance. *Plant Cell.* **24**(8): 3333-3348.

Firon N, Pressman E, Meir S, Khoury R, Altahan L. 2012. Ethylene is involved in maintaining tomato (Solanum lycopersicum) pollen quality under heat-stress conditions. *AoB plants* **2012**: pls024.

Floris M, Mahgoub H, Lanet E, Robaglia C, Menand B. 2009. Post-transcriptional regulation of gene expression in plants during abiotic stress. *Int J Mol Sci.* **10**(7): 3168-3185.

Foyer CH, Noctor G. 2005. Redox homeostasis and antioxidant signaling: a metabolic interface between stress perception and physiological responses. *Plant Cell.* **17**(7): 1866-1875.

Foyer CH, Noctor G. 2011. Ascorbate and glutathione: the heart of the redox hub. *Plant Physiol.***155**(1): 2-18.

Foyer CH, Shigeoka S. 2011. Understanding oxidative stress and antioxidant functions to enhance photosynthesis. *Plant Physiol.***155**(1): 93-100.

Friedlander MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, Rajewsky N. 2008. Discovering microRNAs from deep sequencing data using miRDeep. *Nat Biotech.* **26**(4): 407-415.

Galon Y, Finkler A, Fromm H. 2010. Calcium-regulated transcription in plants. *Mol Plant.* **3**(4): 653-669.

Gandikota M, Birkenbihl RP, Hohmann S, Cardon GH, Saedler H, Huijser P. 2007. The miRNA156/157 recognition element in the 3' UTR of the Arabidopsis SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. *Plant J.* **49**(4): 683-693.

Gao P, Bai X, Yang L, Lv D, Li Y, Cai H, Ji W, Guo D, Zhu Y. 2010. Over-expression of osa-MIR396c decreases salt and alkali stress tolerance. *Planta* **231**(5): 991-1001.

Giacomelli JI, Weigel D, Chan RL, Manavella PA. 2012. Role of recently evolved miRNA regulation of sunflower HaWRKY6 in response to temperature damage. *New phytol* **195**(4): 766-773.

Gill SS, Tuteja N. 2010. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol.Biochem.* **48**(12): 909-930.

Ginzberg I, Barel G, Ophir R, Tzin E, Tanami Z, Muddarangappa T, de Jong W, Fogelman E. 2009. Transcriptomic profiling of heat-stress response in potato periderm. *J Exp Bot.* **60**(15): 4411-4421.

Gong M, Yong-Jun L, Shun-Zhong C. 1998. Abscisic acid-induced thermotolerance in maize seedlings is mediated by calcium and associated with antioxidant systems. *J Plant Physiol.***153**(3-4): 488-496.

Gorantla M, Babu PR, Lachagari VB, Reddy AM, Wusirika R, Bennetzen JL, Reddy AR. 2007. Identification of stress-responsive genes in an indica rice (Oryza sativa L.) using ESTs generated from drought-stressed seedlings. *J Exp Bot.* **58**(2): 253-265.

Goswami A, Banerjee R, Raha S. 2010. Mechanisms of plant adaptation/memory in rice seedlings under arsenic and heat stress: expression of heat-shock protein gene HSP70. *AoB Plant.* **2010**: plq023.

Grefen C, Stadele K, Ruzicka K, Obrdlik P, Harter K, Horak J. 2008. Subcellular localization and in vivo interactions of the Arabidopsis thaliana ethylene receptor family members. *Mol Plant.* **1**(2): 308-320.

Guan Q, Lu X, Zeng H, Zhang Y, Zhu J. 2013. Heat stress induction of miR398 triggers a regulatory loop that is critical for thermotolerance in Arabidopsis. *Plant J.* **74**(5):840-51

Guo H, Ecker JR. 2004. The ethylene signaling pathway: new insights. *Curr Opin Biol.* **7**(1): 40-49.

Guo J, Yang X, Weston DJ, Chen JG. 2011. Abscisic acid receptors: past, present and future. *J Integr Plant Biol.* **53**(6): 469-479.

Haag JR, Ream TS, Marasco M, Nicora CD, Norbeck AD, Pasa-Tolic L, Pikaard CS. 2012. In vitro transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. *Mol cell.* **48**(5): 811-818.

Han MH, Goud S, Song L, Fedoroff N. 2004. The Arabidopsis double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. *Proc Natl Acad Sci U S A.* **101**(4): 1093-1098.

Harper JF, Breton G, Harmon A. 2004. Decoding Ca^{2+} signals through plant protein kinases. *Ann Rev Plant Biol.***55**: 263-288.

Harper JF, Harmon A. 2005. Plants, symbiosis and parasites: a calcium signalling connection. *Nature reviews Mol Cell Biol.* **6**(7): 555-566.

Hays DB, Do JH, Mason RE, Morgan GM, Finlayson SA. 2007. Heat stress induced ethylene production in developing wheat grains induces kernel abortion and increased maturation in a susceptible cultivar. *Plant Sci.* **172**(6): 1113-1123.

Heerklotz D, Doring P, Bonzelius F, Winkelhaus S, Nover L. 2001. The balance of nuclear import and export determines the intracellular distribution and function of tomato heat stress transcription factor HsfA2. *Mol Cell Biol.* **21**(5): 1759-1768.

Henderson IR, Zhang X, Lu C, Johnson L, Meyers BC, Green PJ, Jacobsen SE. 2006. Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nat Genet.* **38**(6): 721-725.

Henz SR, Cumbie JS, Kasschau KD, Lohmann JU, Carrington JC, Weigel D, Schmid M. 2007. Distinct expression patterns of natural antisense transcripts in Arabidopsis. *Plant Physiol.* **144**(3): 1247-1255.

Hilton GR, Lioe H, Stengel F, Baldwin AJ, Benesch JL. 2013. Small heat-shock proteins: paramedics of the cell. *Top Curr Chem.* **328**: 69-98.

Hofmann NR. 2009. The plasma membrane as first responder to heat stress. *Plant Cell.* **21**(9): 2544.

Holmes LE, Campbell SG, De Long SK, Sachs AB, Ashe MP. 2004. Loss of translational control in yeast compromised for the major mRNA decay pathway. *Mol Cell Biol.* **24**(7): 2998-3010.

Hong SW, Vierling E. 2001. Hsp101 is necessary for heat tolerance but dispensable for development and germination in the absence of stress. *Plant J.* **27**(1): 25-35.

Hooijmaijers C, Rhee JY, Kwak KJ, Chung GC, Horie T, Katsuhara M, Kang H. 2012. Hydrogen peroxide permeability of plasma membrane aquaporins of Arabidopsis thaliana. *J Plant Res.* **125**(1): 147-153.

Horvath I, Glatz A, Nakamoto H, Mishkind ML, Munnik T, Saidi Y, Goloubinoff P, Harwood JL, Vigh L. 2012. Heat shock response in photosynthetic organisms: membrane and lipid connections. *Prog lipid Res* **51**(3): 208-220.

Howell MD, Fahlgren N, Chapman EJ, Cumbie JS, Sullivan CM, Givan SA, Kasschau KD, Carrington JC. 2007. Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in Arabidopsis reveals dependency on miRNA- and tasiRNA-directed targeting. *Plant Cell.* **19**(3): 926-942.

Hu S, Claud EC, Musch MW, Chang EB. 2010. Stress granule formation mediates the inhibition of colonic Hsp70 translation by interferon-gamma and tumor necrosis factor-alpha. *Am J Physiol GastroLliver Physiol.* **298**(4): G481-492.

Hu X, Jiang M, Zhang A, Lu J. 2005. Abscisic acid-induced apoplastic H2O2 accumulation up-regulates the activities of chloroplastic and cytosolic antioxidant enzymes in maize leaves. *Planta.* **223**(1): 57-68.

Huang XS, Liu JH, Chen XJ. 2010. Overexpression of PtrABF gene, a bZIP transcription factor isolated from *Poncirus trifoliata*, enhances dehydration and drought tolerance in tobacco via scavenging ROS and modulating expression of stress-responsive genes. *BMC Plant Biol.* **10**: 230.

Huang Y, Li H, Hutchison CE, Laskey J, Kieber JJ. 2003. Biochemical and functional analysis of CTR1, a protein kinase that negatively regulates ethylene signaling in Arabidopsis. *Plant J.* **33**(2): 221-233.

Iyer NJ, Jia X, Sunkar R, Tang G, Mahalingam R. 2012. microRNAs responsive to ozone-induced oxidative stress in Arabidopsis thaliana. *Plant Signal Beh.* **7**(4).

Jacquier-Sarlin MR, Polla BS. 1996. Dual regulation of heat-shock transcription factor (HSF) activation and DNA-binding activity by H2O2: role of thioredoxin. *Biochem J.* **318 (Pt 1)**: 187-193.

Jang JY, Rhee JY, Chung GC, Kang H. 2012. Aquaporin as a membrane transporter of hydrogen peroxide in plant response to stresses. *Plant Signal Beh.***7**(9): 1180-1181.

Jeong DH, Park S, Zhai J, Gurazada SG, De Paoli E, Meyers BC, Green PJ. 2011. Massive analysis of rice small RNAs: mechanistic implications of regulated microRNAs and variants for differential target RNA cleavage. *Plant Cell.* **23**(12): 4185-4207.

Jiang M, Zhang J. 2003. Cross-talk between calcium and reactive oxygen species originated from NADPH oxidase in abscisic acid-induced antioxidant defence in leaves of maize seedlings. *Plant Cell Environ.***26**(6): 929-939.

Jin H. 2008. Endogenous small RNAs and antibacterial immunity in plants. *FEBS L.* **582**(18): 2679-2684.

Jin H, Vacic V, Girke T, Lonardi S, Zhu JK. 2008. Small RNAs and the regulation of cis-natural antisense transcripts in Arabidopsis. *BMC Mol Biol.* **9**: 6.

Jones JC, Duffy JW, Machius M, Temple BR, Dohlman HG, Jones AM. 2011. The crystal structure of a self-activating G protein alpha subunit reveals its distinct mechanism of signal initiation. *Sci. Signal.***4**(159): ra8.

Jones-Rhoades MW, Bartel DP. 2004. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol Cell.* **14**(6): 787-799.

Ju C, Yoon GM, Shemansky JM, Lin DY, Ying ZI, Chang J, Garrett WM, Kessenbrock M, Groth G, Tucker ML et al. 2012. CTR1 phosphorylates the central regulator EIN2 to control ethylene hormone signaling from the ER membrane to the nucleus in Arabidopsis. *Proc Natl Acad Sci U S A.* **109**(47): 19486-19491.

Jung JH, Seo PJ, Park CM. 2009. MicroRNA biogenesis and function in higher plants. *Plant Biotech Rep.* **3**: 111-126.

Kaiser WM. 1979. Reversible inhibition of the Calvin cycle and activation of oxidative pentose phosphate cycle in isolated intact chloroplasts by hydrogen peroxide. *Planta.* **145**: 377-382.

Kalachova T, Iakovenko O, Kretinin S, Kravets V. 2013. Involvement of phospholipase D and NADPH-oxidase in salicylic acid signaling cascade. *Plant Physiol Biochem.* **66**: 127-133.

Kang G, Li G, Zheng B, Han Q, Wang C, Zhu Y, Guo T. 2012. Proteomic analysis on salicylic acid-induced salt tolerance in common wheat seedlings (*Triticum aestivum* L.). *Biochim Biophysi.* **1824**(12): 1324-1333.

Kasschau KD, Fahlgren N, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Carrington JC. 2007. Genome-wide profiling and analysis of Arabidopsis siRNAs. *PLoS Biol.* **5**(3): e57.

Katiyar-Agarwal S, Gao S, Vivian-Smith A, Jin H. 2007. A novel class of bacteria-induced small RNAs in Arabidopsis. *Genes Dev.* **21**(23): 3123-3134.

Katiyar-Agarwal S, Morgan R, Dahlbeck D, Borsani O, Villegas A, Jr., Zhu JK, Staskawicz BJ, Jin H. 2006. A pathogen-inducible endogenous siRNA in plant immunity. *Proc Natl Acad Sci U S A.* **103**(47): 18002-18007.

Kedersha N, Anderson P. 2002. Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. *Biochem Soc Trans.* **30**(Pt 6): 963-969.

Kedersha N, Chen S, Gilks N, Li W, Miller IJ, Stahl J, Anderson P. 2002. Evidence that ternary complex (eIF2-GTP-tRNA(i)(Met))-deficient preinitiation complexes are core constituents of mammalian stress granules. *Mol Biol. Cell.***13**(1): 195-210.

Khraiwesh B, Zhu JK, Zhu J. 2012. Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochim Biophys Acta.* **1819**(2): 137-148.

Kiang JG, McClain DE. 1993. Effect of heat shock, [Ca2+]i, and cAMP on inositol trisphosphate in human epidermoid A-431 cells. *Am J Physiol.***264**(6 Pt 1): C1561-1569.

Kieber JJ, Rothenberg M, Roman G, Feldmann KA, Ecker JR. 1993. CTR1, a negative regulator of the ethylene response pathway in Arabidopsis, encodes a member of the raf family of protein kinases. *Cell.* **72**(3): 427-441.

Kim H, Helmbrecht EE, Stalans MB, Schmitt C, Patel N, Wen CK, Wang W, Binder BM. 2011. Ethylene receptor ETHYLENE RECEPTOR1 domain requirements for ethylene responses in Arabidopsis seedlings. *Plant Physiol.***156**(1): 417-429.

Kim KH, Alam I, Lee KW, Sharmin SA, Kwak SS, Lee SY, Lee BH. 2010. Enhanced tolerance of transgenic tall fescue plants overexpressing 2-Cys peroxiredoxin against methyl viologen and heat stresses. *Biot Letters.* **32**(4): 571-576.

Kim KN. 2013. Stress responses mediated by the CBL calcium sensors in plants. *Plant Biot Rep.* **7**(1): 1-8.

Kim MC, Chung WS, Yun DJ, Cho MJ. 2009. Calcium and calmodulin-mediated regulation of gene expression in plants. *Mol Plant.* **2**(1): 13-21.

Kim VN, Han J, Siomi MC. 2009. Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol.* **10**(2): 126-139.

Kimura S, Kawarazaki T, Nibori H, Michikawa M, Imai A, Kaya H, Kuchitsu K. 2013. The CBL-interacting protein kinase CIPK26 is a novel interactor of Arabidopsis NADPH oxidase AtRbohF that negatively modulates its ROS-producing activity in a heterologous expression system. *J Biochem.* **153**(2): 191-195.

Kirkinezos IG, Moraes CT. 2001. Reactive oxygen species and mitochondrial diseases. *Cell Dev Biol.* **12**(6): 449-457.

Kliebenstein DJ, Monde RA, Last RL. 1998. Superoxide dismutase in Arabidopsis: an eclectic enzyme family with disparate regulation and protein localization. *Plant Physiol.* **118**(2): 637-650.

Kleinhenz MD, Palta JP. 2002. Root zone calcium modulates the response of potato plants to heat stress. *Physiol Plant.***115**(1): 111-118.

Kliebenstein DJ, Monde RA, Last RL. 1998. Superoxide dismutase in Arabidopsis: an eclectic enzyme family with disparate regulation and protein localization. *Plant Physiol.***118**(2): 637-650.

Ko CB, Woo YM, Lee DJ, Lee MC, Kim CS. 2007. Enhanced tolerance to heat stress in transgenic plants expressing the GASA4 gene. *Plant Physiol Biochem.* **45**(9): 722-728.

Kotak S, Larkindale J, Lee U, von Koskull-Doring P, Vierling E, Scharf KD. 2007a. Complexity of the heat stress response in plants. *Curr Opin Biol.* **10**(3): 310-316.

Kotak S, Vierling E, Baumlein H, von Koskull-Doring P. 2007b. A novel transcriptional cascade regulating expression of heat stress proteins during seed development of Arabidopsis. *Plant Cell.* **19**(1): 182-195.

Kozomara A, Griffiths-Jones S. 2011. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* **39**(Database issue): D152-157.

Krinke O, Novotna Z, Valentova O, Martinec J. 2007. Inositol trisphosphate receptor in higher plants: is it real? *J Exp Bot.* **58**(3): 361-376.

Kudla J, Batistic O, Hashimoto K. 2010. Calcium signals: the lead currency of plant information processing. *Plant Cell.* **22**(3): 541-563.

Kuner R, Kohr G, Grunewald S, Eisenhardt G, Bach A, Kornau HC. 1999. Role of heteromer formation in GABAB receptor function. *Science* **283**(5398): 74-77.

Kuo HF, Chiou TJ. 2011. The role of microRNAs in phosphorus deficiency signaling. *Plant Physiol.* **156**(3): 1016-1024.

Kushwaha R, Singh A, Chattopadhyay S. 2008. Calmodulin7 plays an important role as transcriptional regulator in Arabidopsis seedling development. *Plant Cell.* **20**(7): 1747-1759.

Lan A, Huang J, Zhao W, Peng Y, Chen Z, Kang D. 2013. A salicylic acid-induced rice (Oryza sativa L.) transcription factor OsWRKY77 is involved in disease resistance of Arabidopsis thaliana. *Plant Biol.* **15**(3): 452-461.

Langmead B. 2010. Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinformatics.* **Chapter 11**: Unit 11.17.

Larkindale J, Hall JD, Knight MR, Vierling E. 2005. Heat stress phenotypes of Arabidopsis mutants implicate multiple signaling pathways in the acquisition of thermotolerance. *Plant Physiol.***138**(2): 882-897.

Larkindale J, Knight MR. 2002. Protection against heat stress-induced oxidative damage in Arabidopsis involves calcium, abscisic acid, ethylene, and salicylic acid. *Plant Physiol.***128**(2): 682-695.

Law RD, Crafts-Brandner SJ. 1999. Inhibition and acclimation of photosynthesis to heat stress is closely correlated with activation of ribulose-1,5-bisphosphate Carboxylase/Oxygenase. *Plant Physiol.***120**(1): 173-182.

Lee JH, Schoffl F. 1996. An Hsp70 antisense gene affects the expression of HSP70/HSC70, the regulation of HSF, and the acquisition of thermotolerance in transgenic Arabidopsis thaliana. *Mol General Genet.* **252**(1-2): 11-19.

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

Lemtiri-Chlieh F, MacRobbie EA, Webb AA, Manison NF, Brownlee C, Skepper JN, Chen J, Prestwich GD, Brearley CA. 2003. Inositol hexakisphosphate mobilizes an endomembrane store of calcium in guard cells. *Proc Natl Acad Sci U S A.* **100**(17): 10091-10095.

Li D, Wang L, Liu X, Cui D, Chen T, Zhang H, Jiang C, Xu C, Li P, Li S et al. 2013. Deep sequencing of maize small RNAs reveals a diverse set of microRNA in dry and imbibed seeds. *PloS One.* **8**(1): e55107.

Li T, Li H, Zhang YX, Liu JY. 2011. Identification and analysis of seven H(2)O(2)-responsive miRNAs and 32 new miRNAs in the seedlings of rice (Oryza sativa L. ssp. indica). *Nucleic Acids Res.* **39**(7): 2821-2833.

Li T, Li H, Zhang YX, Liu JY. 2011. Identification and analysis of seven H(2)O(2)-responsive miRNAs and 32 new miRNAs in the seedlings of rice (Oryza sativa L. ssp. indica). *Nucleic Acids Res.* **39**(7): 2821-2833.

Li Y, Zhang Z, Liu F, Vongsangnak W, Jing Q, Shen B. 2012. Performance comparison and evaluation of software tools for microRNA deep-sequencing data analysis. *Nucleic Acids Res.* **40**(10): 4298-4305.

Li Z, Zhang L, Wang A, Xu X, Li J. 2013. Ectopic overexpression of SlHsfA3, a heat stress transcription factor from tomato, confers increased thermotolerance and salt hypersensitivity in germination in transgenic Arabidopsis. *PloS one* **8**(1): e54880.

Liang G, Yang F, Yu D. 2010. MicroRNA395 mediates regulation of sulfate accumulation and allocation in Arabidopsis thaliana. *Plant J.***62**(6): 1046-1057.

Lin TY, Markhart AH. 1990. Temperature effects on mitochondrial respiration in *Phaseolus acutffolius* A. Gray and *Phaseolus vulgaris* L. *Plant Physiol.* (94): 54-58.

Lindquist S, Craig EA. 1988. The heat-shock proteins. *Annual Rev Genet.* **22**: 631-677.

Liu HT, Gao F, Cui SJ, Han JL, Sun DY, Zhou RG. 2006. Primary evidence for involvement of IP3 in heat-shock signal transduction in Arabidopsis. *Cell Res.* **16**(4): 394-400.

Liu HT, Li B, Shang ZL, Li XZ, Mu RL, Sun DY, Zhou RG. 2003. Calmodulin is involved in heat shock signal transduction in wheat. *Plant Physiol.***132**(3): 1186-1195.

Los DA, Murata N. 2004. Membrane fluidity and its roles in the perception of environmental signals. *Biochim Biophysi.* **1666**(1-2): 142-157.

Lu S, Yang C, Chiang VL. 2011. Conservation and diversity of microRNA-associated copper-regulatory networks in Populus trichocarpa. *J Iintegr Plant Biol.* **53**(11): 879-891.

Luan S. 2009. The CBL-CIPK network in plant calcium signaling. *Trends Plant Sci.***14**(1): 37-42.

Luan S, Kudla J, Rodriguez-Concepcion M, Yalovsky S, Gruissem W. 2002. Calmodulins and calcineurin B-like proteins: calcium sensors for specific signal response coupling in plants. *Plant Cell.* **14 Suppl**: S389-400.

Ma W, Smigel A, Tsai YC, Braam J, Berkowitz GA. 2008. Innate immunity signaling: cytosolic Ca2+ elevation is linked to downstream nitric oxide generation through the action of calmodulin or a calmodulin-like protein. *Plant Physiol.***148**(2): 818-828.

Ma Z, Coruh C, Axtell MJ. 2010. Arabidopsis lyrata small RNAs: transient MIRNA and small interfering RNA loci within the Arabidopsis genus. *Plant Cell.* **22**(4): 1090-1103.

Mackova H, Hronkova M, Dobra J, Tureckova V, Novak O, Lubovska Z, Motyka V, Haisel D, Hajek T, Prasil IT et al. 2013. Enhanced drought and heat stress tolerance of tobacco plants with ectopically enhanced cytokinin oxidase/dehydrogenase gene expression. *J Exp Bot.* **64**(10):2805-15

Mackowiak SD. 2011. Identification of novel and known miRNAs in deep-sequencing data with miRDeep2. *Curr Protoc Bioinformatics.* **Chapter 12**: Unit 12.10.

Maestri E, Klueva N, Perrotta C, Gulli M, Nguyen HT, Marmiroli N. 2002. Molecular genetics of heat tolerance and heat shock proteins in cereals. *Plant Mol Biol.* **48**(5-6): 667-681.

Mahajan S, Tuteja N. 2005. Cold, salinity and drought stresses: an overview. *Arch Biochem and Biophysi.* **444**(2): 139-158.

Mallory AC, Bartel DP, Bartel B. 2005. MicroRNA-directed regulation of Arabidopsis AUXIN RESPONSE FACTOR17 is essential for proper development and modulates expression of early auxin response genes. *Plant Cell.* **17**(5): 1360-1375.

Mallory AC, Reinhart BJ, Jones-Rhoades MW, Tang G, Zamore PD, Barton MK, Bartel DP. 2004. MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. *EMBO J.* **23**(16): 3356-3364.

Malone JH, Oliver B. 2011. Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol.* **9**: 34.

Manavella PA, Weigel D, Wu L. 2011. Argonaute10 as a miRNA locker. *Cell* **145**(2): 173-174.

Marino D, Dunand C, Puppo A, Pauly N. 2012. A burst of plant NADPH oxidases. *Trends Plant Sci.***17**(1): 9-15.

Martiniere A, Shvedunova M, Thomson AJ, Evans NH, Penfield S, Runions J, McWatters HG. 2011. Homeostasis of plasma membrane viscosity in fluctuating temperatures. *New Phyt.***192**(2): 328-337.

Mathelier A, Carbone A. 2010. MIReNA: finding microRNAs with high accuracy and no learning at genome scale and from deep sequencing data. *Bioinformatics.* **26**(18): 2226-2234.

Matthewman CA, Kawashima CG, Huska D, Csorba T, Dalmay T, Kopriva S. 2012. miR395 is a general component of the sulfate assimilation regulatory network in Arabidopsis. *FEBS L.* **586**(19): 3242-3248.

Matzke M, Kanno T, Daxinger L, Huettel B, Matzke AJ. 2009. RNA-mediated chromatin-based silencing in plants. *Curr Opin Cell Biol.* **21**(3): 367-376.

Matzke MA, Birchler JA. 2005. RNAi-mediated pathways in the nucleus. *Nat Rev Genet.* **6**(1): 24-35.

Mayerhofer H, Panneerselvam S, Mueller-Dieckmann J. 2012. Protein kinase domain of CTR1 from Arabidopsis thaliana promotes ethylene receptor cross talk. *J Mol Biol.***415**(4): 768-779.

Mazars C, Thuleau P, Lamotte O, Bourque S. 2010. Cross-talk between ROS and calcium in regulation of nuclear activities. *Mol Plant.* **3**(4): 706-718.

Mazroui R, Sukarieh R, Bordeleau ME, Kaufman RJ, Northcote P, Tanaka J, Gallouzi I, Pelletier J. 2006. Inhibition of ribosome recruitment induces stress granule formation independently of eukaryotic initiation factor 2alpha phosphorylation. *Mol Biol. Cell.***17**(10): 4212-4219.

Mejia-Teniente L, de Dalia Duran-Flores F, Chapa-Oliver AM, Torres-Pacheco I, Cruz-Hernandez A, Gonzalez-Chavira MM, Ocampo-Velazquez RV, Guevara-Gonzalez RG. 2013. Oxidative and Molecular Responses in Capsicum annuum L. after Hydrogen Peroxide, Salicylic Acid and Chitosan Foliar Applications. *Int J Mol Sci.* **14**(5): 10178-10196.

Meng Y, Chen D, Ma X, Mao C, Cao J, Wu P, Chen M. 2010. Mechanisms of microRNA-mediated auxin signaling inferred from the rice mutant osaxr. *Plant signal Beh.* **5**(3): 252-254.

Meng Y, Shao C, Wang H, Chen M. 2012. Are all the miRBase-registered microRNAs true? A structure- and expression-based re-examination in plants. *RNA Biol.* **9**(3): 249-253.

Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, Bowman JL, Cao X, Carrington JC, Chen X, Green PJ et al. 2008. Criteria for annotation of plant MicroRNAs. *Plant Cell.* **20**(12): 3186-3190.

Mikoshiba K. 2007. IP3 receptor/Ca2+ channel: from discovery to new signaling concepts. *Journal of neurochemistry* **102**(5): 1426-1446.

Miller EW, Dickinson BC, Chang CJ. 2010. Aquaporin-3 mediates hydrogen peroxide uptake to regulate downstream intracellular signaling. *Proc Natl Acad Sci U S A.* **107**(36): 15681-15686.

Miller G, Suzuki N, Rizhsky L, Hegie A, Koussevitzky S, Mittler R. 2007. Double mutants deficient in cytosolic and thylakoid ascorbate peroxidase reveal a complex mode of interaction between reactive oxygen species, plant development, and response to abiotic stresses. *Plant Physiol.***144**(4): 1777-1785.

Mishkind M, Vermeer JE, Darwish E, Munnik T. 2009. Heat stress activates phospholipase D and triggers PIP_2 accumulation at the plasma membrane and nucleus. *Plant J.* **60**(1): 10-21.

Misra S, Wu Y, Venkataraman G, Sopory SK, Tuteja N. 2007. Heterotrimeric G-protein complex and G-protein-coupled receptor from a legume (*Pisum sativum*): role in salinity and heat stress and cross-talk with phospholipase C. *Plant J.* **51**(4): 656-669.

Mittler R, Vanderauwera S, Suzuki N, Miller G, Tognetti VB, Vandepoele K, Gollery M, Shulaev V, Van Breusegem F. 2011. ROS signaling: the new wave? *Trends Plant Sci.***16**(6): 300-309.

Mohsenzadeh S, Saupe-Thies W, Steier G, Schroeder T, Fracella F, Ruoff P, Rensing L. 1998. Temperature adaptation of house keeping and heat shock gene expression in Neurospora crassa. *Fungal Genet Biol.* **25**(1): 31-43.

Mosher RA, Melnyk CW, Kelly KA, Dunn RM, Studholme DJ, Baulcombe DC. 2009. Uniparental expression of PolIV-dependent siRNAs in developing endosperm of Arabidopsis. *Nature.* **460**(7252): 283-286.

Muhlrad D, Parker R. 1994. Premature translational termination triggers mRNA decapping. *Nature.* **370**(6490): 578-581.

Nag A, King S, Jack T. 2009. miR319a targeting of TCP4 is critical for petal growth and development in Arabidopsis. *Proc Natl Acad Sci U S A.* **106**(52): 22534-22539.

Nakashima K, Ito Y, Yamaguchi-Shinozaki K. 2009. Transcriptional regulatory networks in response to abiotic stresses in Arabidopsis and grasses. *Plant Physiol.***149**(1): 88-95.

Napoli E, Wong S, Giulivi C. 2013. Evidence of reactive oxygen species-mediated damage to mitochondrial DNA in children with typical autism. *Mol. Autism.* **4**(1): 2.

Nash D, Miyao M, Norio M. 1985. Heat inactivation of oxygen evolution in Photosystem II particles and its acceleration by chloride depletion and exogenous manganese. *Biochim Biophysi.* **807**(2): 127-133.

Nie H, Zhao C, Wu G, Wu Y, Chen Y, Tang D. 2012. SR1, a calmodulin-binding transcription factor, modulates plant defense and ethylene-induced senescence by directly regulating NDR1 and EIN3. *Plant Physiol.***158**(4): 1847-1859.

Nishiyama Y, Yamamoto H, Allakhverdiev SI, Inaba M, Yokota A, Murata N. 2001. Oxidative stress inhibits the repair of photodamage to the photosynthetic machinery. *EMBO J.* **20**(20): 5587-5594.

Noriega G, Caggiano E, Lecube ML, Cruz DS, Batlle A, Tomaro M, Balestrasse KB. 2012. The role of salicylic acid in the prevention of oxidative stress elicited by cadmium in soybean plants. *Biometal.* **25**(6): 1155-1165.

North HM, De Almeida A, Boutin JP, Frey A, To A, Botran L, Sotta B, Marion-Poll A. 2007. The Arabidopsis ABA-deficient mutant aba4 demonstrates that the major route for stress-induced ABA accumulation is via neoxanthin isomers. *Plant J.* **50**(5): 810-824.

Nover L, Scharf KD, Neumann D. 1983. Formation of cytoplasmic heat shock granules in tomato cell cultures and leaves. *Mol Cell Biol.* **3**(9): 1648-1655.

Nover L. 1989. Cytoplasmic heat shock granules are formed from precursor particles and are associated with a specific set of mRNAs. *Mol Cell Biol.* **9**(3): 1298-1308.

Okamura K, Ladewig E, Zhou L, Lai EC. 2013. Functional small RNAs are generated from select miRNA hairpin loops in flies and mammals. *Genes Dev.* **27**(7): 778-792.

Onomoto K, Jogi M, Yoo JS, Narita R, Morimoto S, Takemura A, Sambhara S, Kawaguchi A, Osari S, Nagata K et al. 2012. Critical role of an antiviral stress granule containing RIG-I and PKR in viral detection and innate immunity. *PloS one.* **7**(8): e43031.

Oukarroum A, Strasser RJ, Schansker G. 2012. Heat stress and the photosynthetic electron transport chain of the lichen Parmelina tiliacea (Hoffm.) Ach. in the dry and the wet state: differences and similarities with the heat stress response of higher plants. *Photosynth Res.* **111**(3): 303-314.

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

Panchuk, II, Volkov RA, Schoffl F. 2002. Heat stress- and heat shock transcription factor-dependent expression and activity of ascorbate peroxidase in Arabidopsis. *Plant Physiol.***129**(2): 838-853.

Pandey D, Gratton JP, Rafikov R, Black SM, Fulton DJ. 2011. Calcium/calmodulin-dependent kinase II mediates the phosphorylation and activation of NADPH oxidase 5. *Mol Pharm.* **80**(3): 407-415.

Paradies G, Petrosillo G, Pistolese M, Ruggiero FM. 2002. Reactive oxygen species affect mitochondrial electron transport complex I activity through oxidative cardiolipin damage. *Gene.* **286**(1): 135-141.

Park MY, Wu G, Gonzalez-Sulser A, Vaucheret H, Poethig RS. 2005. Nuclear processing and export of microRNAs in Arabidopsis. *Proc Natl Acad Sci U S A.* **102**(10): 3691-3696.

Parsell DA, Lindquist S. 1993. The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. *Annual Rev Genet.* **27**: 437-496.

Perez-Quintero AL, Quintero A, Urrego O, Vanegas P, Lopez C. 2012. Bioinformatic identification of cassava miRNAs differentially expressed in response to infection by Xanthomonas axonopodis pv. manihotis. *BMC Plant Biol.* **12**: 29.

Perrot-Rechenmann C. 2010. Cellular responses to auxin: division versus expansion. *Cold Spring Harb Pers Biol.* **2**(5): a001446.

Pikaard CS, Haag JR, Ream T, Wierzbicki AT. 2008. Roles of RNA polymerase IV in gene silencing. *Trends Plant Sci.* **13**(7): 390-397.

Pitzschke A, Hirt H. 2006. Mitogen-activated protein kinases and reactive oxygen species signaling in plants. *Plant Physiol.***141**(2): 351-356.

Potlakayala SD, Reed DW, Covello PS, Fobert PR. 2007. Systemic acquired resistance in canola is linked with pathogenesis-related gene expression and requires salicylic Acid. *Phytopath.* **97**(7): 794-802.

Puig S, Andres-Colas N, Garcia-Molina A, Penarrubia L. 2007. Copper and iron homeostasis in Arabidopsis: responses to metal deficiencies, interactions and biotechnological applications. *Plant Cell Environ.* **30**(3): 271-290.

Qin D, Wu H, Peng H, Yao Y, Ni Z, Li Z, Zhou C, Sun Q. 2008. Heat stress-responsive transcriptome analysis in heat susceptible and tolerant wheat (Triticum aestivum L.) by using Wheat Genome Array. *BMC Genomics.*. **9**: 432.

Qu AL, Ding YF, Jiang Q, Zhu C. 2013. Molecular mechanisms of the plant heat stress response. *Biochem Biophysi Res Commun.* **432**(2): 203-207.

Queitsch C, Hong SW, Vierling E, Lindquist S. 2000. Heat shock protein 101 plays a crucial role in thermotolerance in Arabidopsis. *Plant Cell.* **12**(4): 479-492.

Ranty B, Aldon D, Galaud JP. 2006. Plant calmodulins and calmodulin-related proteins: multifaceted relays to decode calcium signals. *Plant Signal Beh.***1**(3): 96-104.

Reddy AS, Ali GS, Celesnik H, Day IS. 2011. Coping with stresses: roles of calcium- and calcium/calmodulin-regulated gene expression. *Plant Cell.* **23**(6): 2010-2032.

Rhoads DM, Umbach AL, Subbaiah CC, Siedow JN. 2006. Mitochondrial reactive oxygen species. Contribution to oxidative stress and interorganellar signaling. *Plant Physiol.***141**(2): 357-366.

Richardson A, Taylor CW. 1993. Effects of Ca2+ chelators on purified inositol 1,4,5-trisphosphate (InsP3) receptors and InsP3-stimulated Ca2+ mobilization. *J Plant Chem.* **268**(16): 11528-11533.

Rochon A, Boyle P, Wignes T, Fobert PR, Despres C. 2006. The coactivator function of Arabidopsis NPR1 requires the core of its BTB/POZ domain and the oxidation of C-terminal cysteines. *Plant Cell.* **18**(12): 3670-3685.

Rodriguez FI, Esch JJ, Hall AE, Binder BM, Schaller GE, Bleecker AB. 1999. A copper cofactor for the ethylene receptor ETR1 from Arabidopsis. *Science.* **283**(5404): 996-998.

Romero P, Rodrigo MJ, Alferez F, Ballester AR, Gonzalez-Candelas L, Zacarias L, Lafuente MT. 2012. Unravelling molecular responses to moderate dehydration in harvested fruit of sweet orange (Citrus sinensis L. Osbeck) using a fruit-specific ABA-deficient mutant. *J Exp Bot.* **63**(7): 2753-2767.

Ron M, Alandete Saez M, Eshed Williams L, Fletcher JC, McCormick S. 2010. Proper regulation of a sperm-specific cis-nat-siRNA is essential for double fertilization in Arabidopsis. *Genes Dev.* **24**(10): 1010-1021.

Sagi M, Fluhr R. 2006. Production of reactive oxygen species by plant NADPH oxidases. *Plant Physiol.***141**(2): 336-340.

Saidi Y, Finka A, Goloubinoff P. 2011. Heat perception and signalling in plants: a tortuous path to thermotolerance. *New Phyt.***190**(3): 556-565.

Saidi Y, Finka A, Muriset M, Bromberg Z, Weiss YG, Maathuis FJ, Goloubinoff P. 2009. The heat shock response in moss plants is regulated by specific calcium-permeable channels in the plasma membrane. *Plant Cell.* **21**(9): 2829-2843.

Saidi Y, Peter M, Finka A, Cicekli C, Vigh L, Goloubinoff P. 2010. Membrane lipid composition affects plant heat sensing and modulates Ca(2+)-dependent heat shock response. *Plant signaling and behavior* **5**(12): 1530-1533.

Salvucci ME, Crafts-Brandner SJ. 2004. Inhibition of photosynthesis by heat stress: the activation state of Rubisco as a limiting factor in photosynthesis. *Physiol Plant.***120**(2): 179-186.

Sarot E, Payen-Groschene G, Bucheton A, Pelisson A. 2004. Evidence for a piwi-dependent RNA silencing of the gypsy endogenous retrovirus by the Drosophila melanogaster flamenco gene. *Genetics.* **166**(3): 1313-1321.

Sato K, Hamada M, Asai K, Mituyama T. 2009. CENTROIDFOLD: a web server for RNA secondary structure prediction. *Nucleic Acids Res.* **37**(Web Server issue): W277-280.

Savchenko GE, Klyuchareva EA, Abramchik LM, Serdyuchenko EV. 2002. Effect of Periodic Heat Shock on the Inner Membrane System. *Ru J Plant Physiol.***49**(3): 349-359.

Scharf KD, Heider H, Hohfeld I, Lyck R, Schmidt E, Nover L. 1998. The tomato Hsf system: HsfA2 needs interaction with HsfA1 for efficient nuclear import and may be localized in cytoplasmic heat stress granules. *Mol Cell Biol.* **18**(4): 2240-2251.

Schramm F, Ganguli A, Kiehlmann E, Englich G, Walch D, Koskull-Döring Pv. 2006. The heat stress transcription factor HsfA2 serves as a regulatory amplifier of a subset of genes in the heat stress response in Arabidopsis. *Plant Mol Biol.* **60**(5): 759-772.

Schramm F, Larkindale J, Kiehlmann E, Ganguli A, Englich G, Vierling E, von Koskull-Doring P. 2008. A cascade of transcription factor DREB2A and heat stress transcription factor HsfA3 regulates the heat stress response of Arabidopsis. *Plant J.* **53**(2): 264-274.

Schreiber AW, Shi BJ, Huang CY, Langridge P, Baumann U. 2011. Discovery of barley miRNAs through deep sequencing of short reads. *BMC Genomics.* **12**: 129.

Shakirova FM, Bezrukova MV, Yuldashev RA, Fatkhutdinova RA, Murzabaev AR. 2013. Involvement of lectin in the salicylic acid-induced wheat tolerance to cadmium and the role of endogenous ABA in the regulation of its level. *Doklady Biol.* **448**: 49-51.

Sharabi-Schwager M, Porat R, Samach A. 2011. Overexpression of the *CBF2* transcriptional Activator Enhances Oxidative Stress Tolerance in Arabidopsis Plants. *Int J Bio.* **3**(2): 94-105.

Sharkey TD. 2005. Effects of moderate heat stress on photosynthesis: importance of thylakoid reactions, rubisco deactivation, reactive oxygen species, and thermotolerance provided by isoprene. *Plant Cell Environ.* **28**: 269-277.

Sharma P, Jha AB, Dubey RS, Pessarakli M. 2012. Reactive Oxygen Species, Oxidative Damage, and Antioxidative Defense Mechanism in Plants under Stressful Conditions. *J Bot.* **2012**(2012): 1-26.

Shelp BJ, Bown AW, McLean MD. 1999. Metabolism and functions of gamma-aminobutyric acid. *Trends Plant Sci.***4**(11): 446-452.

Shen J, Xie K, Xiong L. 2010. Global expression profiling of rice microRNAs by one-tube stem-loop reverse transcription quantitative PCR revealed important roles of microRNAs in abiotic stress responses. *Mol Genet Genomics.* **284**(6): 477-488.

Shi SQ, Shi Z, Jiang ZP, Qi LW, Sun XM, Li CX, Liu JF, Xiao WF, Zhang SG. 2010. Effects of exogenous GABA on gene expression of Caragana intermedia roots under NaCl stress: regulatory roles for H2O2 and ethylene production. *Plant Cell Environ.***33**(2): 149-162.

Shigeoka S, Ishikawa T, Tamoi M, Miyagawa Y, Takeda T, Yabuta Y, Yoshimura K. 2002. Regulation and function of ascorbate peroxidase isoenzymes. *J Exp Bot.* **53**(372): 1305-1319.

Si-Ammour A, Windels D, Arn-Bouldoires E, Kutter C, Ailhas J, Meins F, Jr., Vazquez F. 2011. miR393 and Secondary siRNAs Regulate Expression of the TIR1/AFB2 Auxin Receptor Clade and Auxin-Related Development of Arabidopsis Leaves. *Plant Physiol.* **157**(2): 683-691.

Simões-Araújo JL, Rumjanek NG, Margis-Pinheiro M. 2003. Small heat shock proteins genes are differentially expressed in distinct varieties of common bean. *Brazi J Plant Physiol.***15**(1): 33-41.

Sinha AK, Jaggi M, Raghuram B, Tuteja N. 2011. Mitogen-activated protein kinase signaling in plants under abiotic stress. *Plant Signal Beh.* **6**(2): 196-203.

Snider JL, Oosterhuis DM, Skulman BW, Kawakami EM. 2009. Heat stress-induced limitations to reproductive success in Gossypium hirsutum. *Physiol Plant.***137**(2): 125-138.

Stevenson JM, Perera IY, Heilmann II, Persson S, Boss WF. 2000. Inositol signaling and plant growth. *Trends Plant Sci.***5**(8): 357.

Sung DY, Kaplan F, Lee KJ, Guy CL. 2003. Acquired tolerance to temperature extremes. *Trends Plant Sci.* **8**(4): 179-187.

Sunkar R, Jagadeeswaran G. 2008. In silico identification of conserved microRNAs in large number of diverse plant species. *BMC Plant Biol.* **8**: 37.

Sunkar R, Kapoor A, Zhu JK. 2006. Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in Arabidopsis is mediated by downregulation of miR398 and important for oxidative stress tolerance. *Plant Cell.* **18**(8): 2051-2065.

Suzuki N, Koussevitzky S, Mittler R, Miller G. 2012. ROS and redox signalling in the response of plants to abiotic stress. *Plant Cell Environ.* **35**(2): 259-270.

Suzuki N, Miter R. 2006. Reactive oxygen species and temperature stresses: A delicate balance between signaling and destruction. *Physiol Plant.***126**(1): 45-51.

Swindell WR, Huebner M, Weber AP. 2007. Transcriptional profiling of Arabidopsis heat shock proteins and transcription factors reveals extensive overlap between heat and non-heat stress response pathways. *BMC Genomics.* **8**: 125.

Szymanski DB, Liao B, Zielinski RE. 1996. Calmodulin isoforms differentially enhance the binding of cauliflower nuclear proteins and recombinant TGA3 to a region derived from the Arabidopsis Cam-3 promoter. *Plant Cell.* **8**(6): 1069-1077.

Takahashi F, Mizoguchi T, Yoshida R, Ichimura K, Shinozaki K. 2011. Calmodulin-dependent activation of MAP kinase for ROS homeostasis in Arabidopsis. *Mol cell.* **41**(6): 649-660.

Takahashi S, Tamashiro A, Sakihama Y, Yamamoto Y, Kawamitsu Y, Yamasaki H. 2002. High-susceptibility of photosynthesis to photoinhibition in the tropical plant Ficus microcarpa L. f. cv. Golden Leaves. *BMC Plant Biol.* **2**: 2.

Tan W, Meng Q, Brestic M, Olsovska K, Yang X. 2011. Photosynthesis is improved by exogenous calcium in heat-stressed tobacco plants. *J Plant Physiol.***168**(17): 2063-2071.

Tang W, Charles TM, Newton RJ. 2005. Overexpression of the pepper transcription factor CaPF1 in transgenic Virginia pine (Pinus Virginiana Mill.) confers multiple stress tolerance and enhances organ growth. *Plant Mol Biol.* **59**(4): 603-617.

Temple BR, Jones AM. 2007. The plant heterotrimeric G-protein complex. *Ann Rev Plant Biol.***58**: 249-266.

Thakur V, Wanchana S, Xu M, Bruskiewich R, Quick WP, Mosig A, Zhu XG. 2011. Characterization of statistical features for plant microRNA prediction. *BMC Genomics.* **12**: 108.

Tian Y, Zhang H, Pan X, Chen X, Zhang Z, Lu X, Huang R. 2011. Overexpression of ethylene response factor TERF2 confers cold tolerance in rice seedlings. *Transgenic Res.* **20**(4): 857-866.

Trindade I, Capitao C, Dalmay T, Fevereiro MP, Santos DM. 2010. miR398 and miR408 are up-regulated in response to water deficit in Medicago truncatula. *Planta.* **231**(3): 705-716.

Triantaphylides C, Krischke M, Hoeberichts FA, Ksas B, Gresser G, Havaux M, Van Breusegem F, Mueller MJ. 2008. Singlet oxygen is the major reactive oxygen species involved in photooxidative damage to plants. *Plant Physiol.***148**(2): 960-968.

Turrens JF. 2003. Mitochondrial formation of reactive oxygen species. *J Physiol.* **552**(Pt 2): 335-344.

Tuteja N. 2007. Abscisic Acid and abiotic stress signaling. *Plant Signal Beh.***2**(3): 135-138.

Tuteja N. 2009. Signaling through G protein coupled receptors. *Plant Signal Beh.***4**(10): 942-947.

Tuteja N, Sopory SK. 2008a. Chemical signaling under abiotic stress environment in plants. *Plant Signal Beh.***3**(8): 525-536.

-. 2008b. Plant signaling in stress: G-protein coupled receptors, heterotrimeric G-proteins and signal coupling via phospholipases. *Plant Signal Beh.***3**(2): 79-86.

Uniacke J, Zerges W. 2008. Stress induces the assembly of RNA granules in the chloroplast of Chlamydomonas reinhardtii. *J Cell Biol.* **182**(4): 641-646.

Uno Y, Furihata T, Abe H, Yoshida R, Shinozaki K, Yamaguchi-Shinozaki K. 2000. Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. *Proc Natl Acad Sci.* **97**(21): 11632-11637.

Urano D, Chen JG, Botella JR, Jones AM. 2013. Heterotrimeric G protein signalling in the plant kingdom. *Open Biol.* **3**(3): 120186.

Urano D, Jones AM. 2013. "Round up the usual suspects": a comment on nonexistent plant g protein-coupled receptors. *Plant Physiol.***161**(3): 1097-1102.

Urano D, Jones JC, Wang H, Matthews M, Bradford W, Bennetzen JL, Jones AM. 2012a. G protein activation without a GEF in the plant kingdom. *PLoS Genet.* **8**(6): e1002756.

Urano D, Phan N, Jones JC, Yang J, Huang J, Grigston J, Taylor JP, Jones AM. 2012b. Endocytosis of the seven-transmembrane RGS1 protein activates G-protein-coupled signalling in Arabidopsis. *Nat Cell Biol.* **14**(10): 1079-1088.

Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. 2006. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science.* **313**(5785): 320-324.

Valdes-Lopez O, Arenas-Huertero C, Ramirez M, Girard L, Sanchez F, Vance CP, Luis Reyes J, Hernandez G. 2008. Essential role of MYB transcription factor: PvPHR1 and microRNA: PvmiR399 in phosphorus-deficiency signalling in common bean roots. *Plant Cell Environ.* **31**(12): 1834-1843.

Vanderauwera S, Suzuki N, Miller G, van de Cotte B, Morsa S, Ravanat JL, Hegie A, Triantaphylides C, Shulaev V, Van Montagu MC et al. 2011. Extranuclear protection of chromosomal DNA from oxidative stress. *Proc Natl Acad Sci U S A.* **108**(4): 1711-1716.

Verslues PE, Zhu JK. 2005. Before and beyond ABA: upstream sensing and internal signals that determine ABA accumulation and response under abiotic stress. *Biochem Soc Trans.* **33**(Pt2): 375-379.

Vigh L, Horvath I, Maresca B, Harwood JL. 2007. Can the stress protein response be controlled by 'membrane-lipid therapy'? *Trends Biochem Sci.* **32**(8): 357-363.

Vigh L, Los DA, Horvath I, Murata N. 1993. The primary signal in the biological perception of temperature: Pd-catalyzed hydrogenation of membrane lipids stimulated the expression of the desA gene in Synechocystis PCC6803. *Proc Natl Acad Sci U S A.* **90**(19): 9090-9094.

Volkov RA, Panchuk, II, Mullineaux PM, Schoffl F. 2006. Heat stress-induced H_2O_2 is required for effective expression of heat shock genes in Arabidopsis. *Plant Mol Biol.* **61**(4-5): 733-746.

Volkov RA, Panchuk, II, Schoffl F. 2003. Heat-stress-dependency and developmental modulation of gene expression: the potential of house-keeping genes as internal standards in mRNA expression profiling using real-time RT-PCR. *J Exp Bot.* **54**(391): 2343-2349.

Wahid A, Gelani S, Ashraf M, Foolad MR. 2007. Heat tolerance in plants: An overview. *Environ Expl Bot.* **61**(3): 199-223.

Walker DS, Gower NJ, Ly S, Bradley GL, Baylis HA. 2002. Regulated disruption of inositol 1,4,5-trisphosphate signaling in Caenorhabditis elegans reveals new functions in feeding and embryogenesis. *Mol Biol. Cell.***13**(4): 1329-1337.

Wan LC, Wang F, Guo X, Lu S, Qiu Z, Zhao Y, Zhang H, Lin J. 2012. Identification and characterization of small non-coding RNAs from Chinese fir by high throughput sequencing. *BMC Plant Biol.* **12**: 146.

Wang C, Han J, Liu C, Kibet KN, Kayesh E, Shangguan L, Li X, Fang J. 2012. Identification of microRNAs from Amur grape (vitis amurensis Rupr.) by deep sequencing and analysis of microRNA variations with bioinformatics. *BMC Genomics.* **13**: 122.

Wang LJ, Fan L, Loescher W, Duan W, Liu GJ, Cheng JS, Luo HB, Li SH. 2010a. Salicylic acid alleviates decreases in photosynthesis under heat stress and accelerates recovery in grapevine leaves. *BMC Plant Biol.* **10**: 34.

Wang L, Liu H, Li D, Chen H. 2011. Identification and characterization of maize microRNAs involved in the very early stage of seed germination. *BMC Genomics.* **12**: 154.

Wang MB, Helliwell CA, Wu LM, Waterhouse PM, Peacock WJ, Dennis ES. 2008. Hairpin RNAs derived from RNA polymerase II and polymerase III promoter-directed transgenes are processed differently in plants. *RNA.* **14**(5): 903-913.

Wang P, Du Y, Li Y, Ren D, Song CP. 2010b. Hydrogen peroxide-mediated activation of MAP kinase 6 modulates nitric oxide biosynthesis and signal transduction in Arabidopsis. *Plant Cell.* **22**(9): 2981-2998.

Wang W, Vinocur B, Shoseyov O, Altman A. 2004. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci.***9**(5): 244-252.

Wang Y, Li Y, Ma Z, Yang W, Ai C. 2010. Mechanism of microRNA-target interaction: molecular dynamics simulations and thermodynamics analysis. *PLoS Comp Biol.* **6**(7): e1000866.

Weber C, Nover L, Fauth M. 2008. Plant stress granules and mRNA processing bodies are distinct from heat stress granules. *Plant J.* **56**(4): 517-530.

West AH, Stock AM. 2001. Histidine kinases and response regulator proteins in two-component signaling systems. *Trends Biochem Sci.* **26**(6): 369-376.

Wheeler GL, Brownlee C. 2008. Ca2+ signalling in plants and green algae--changing channels. *Trends Plant Sci.***13**(9): 506-514.

Wierzbicki AT, Ream TS, Haag JR, Pikaard CS. 2009. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nat Genet.* **41**(5): 630-634.

Williams Z, Ben-Dov IZ, Elias R, Mihailovic A, Brown M, Rosenwaks Z, Tuschl T. 2013. Comprehensive profiling of circulating microRNA via small RNA sequencing of cDNA libraries reveals biomarker potential and limitations. *Proc Natl Acad Sci U S A.* **110**(11): 4255-4260.

Williamson V, Kim A, Xie B, McMichael GO, Gao Y, Vladimirov V. 2013. Detecting miRNAs in deep-sequencing data: a software performance comparison and evaluation. *Brief Bioinformatics.* **14**(1): 36-45.

Windels D, Vazquez F. 2011. miR393: integrator of environmental cues in auxin signaling? *Plant Signal Beh.* **6**(11): 1672-1675.

Wollmann H, Mica E, Todesco M, Long JA, Weigel D. 2010. On reconciling the interactions between APETALA2, miR172 and AGAMOUS with the ABC model of flower development. *Development* **137**(21): 3633-3642.

Wolozin B. 2012. Regulated protein aggregation: stress granules and neurodegeneration. *Mol Neuro.* **7**: 56.

Wu C. 1995. Heat shock transcription factors: structure and regulation. *Ann Rev Cell Dev Biol.* **11**: 441-469.

Wu HC, Jinn TL. 2010. Heat shock-triggered Ca2+ mobilization accompanied by pectin methylesterase activity and cytosolic Ca2+ oscillation are crucial for plant thermotolerance. *Plant Signal Beh.* **5**(10): 1252-1256.

Wu HC, Jinn TL. 2012. Oscillation regulation of Ca2+ /calmodulin and heat-stress related genes in response to heat stress in rice (*Oryza sativa* L.). *Plant Signal Beh.***7**(9): 1056-1057.

Wu L, Zhou H, Zhang Q, Zhang J, Ni F, Liu C, Qi Y. 2010. DNA methylation mediated by a microRNA pathway. *Mol Cell.* **38**(3): 465-475.

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

Wu P, Wang X. 2008. Role of OsPHR2 on phosphorus homeostasis and root hairs development in rice (Oryza sativa L.). *Plant Signal Beh.* **3**(9): 674-675.

Wu YQ, Chen DJ, He HB, Chen DS, Chen LL, Chen HC, Liu ZF. 2012. Pseudorabies Virus Infected Porcine Epithelial Cell Line Generates a Diverse Set of Host MicroRNAs and a Special Cluster of Viral MicroRNAs. *PloS One* **7**(1): e30988.

Wu X, Shiroto Y, Kishitani S, Ito Y, Toriyama K. 2009. Enhanced heat and drought tolerance in transgenic rice seedlings overexpressing OsWRKY11 under the control of HSP101 promoter. *Plant Cell Rep.* **28**(1): 21-30.

Wu Y, Zhang D, Chu JY, Boyle P, Wang Y, Brindle ID, De Luca V, Despres C. 2012. The Arabidopsis NPR1 protein is a receptor for the plant defense hormone salicylic acid. *Cell Rep.* **1**(6): 639-647.

Xia R, Meyers BC, Liu Z, Beers EP, Ye S. 2013. MicroRNA Superfamilies Descended from miR390 and Their Roles in Secondary Small Interfering RNA Biogenesis in Eudicots. *Plant Cell.* **25**(5):1555-72

Xie Z, Allen E, Fahlgren N, Calamar A, Givan SA, Carrington JC. 2005. Expression of Arabidopsis MIRNA genes. *Plant Physiol.* **138**(4): 2145-2154.

Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC. 2004. Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol.* **2**(5): E104.

Xin M, Wang Y, Yao Y, Xie C, Peng H, Ni Z, Sun Q. 2010. Diverse set of microRNAs are responsive to powdery mildew infection and heat stress in wheat (Triticum aestivum L.). *BMC Plant Biol.* **10**: 123.

Xu ZS, Chen M, Li LC, Ma YZ. 2011. Functions and application of the AP2/ERF transcription factor family in crop improvement. *J Integr Plant Biol.* **53**(7): 570-585.

Xuan P, Guo M, Liu X, Huang Y, Li W. 2011. PlantMiRNAPred: efficient classification of real and pseudo plant pre-miRNAs. *Bioinformatics.* **27**(10): 1368-1376.

Yamane Y, Kashino Y, Koike H, Satoh K. 1998. Effects of high temperatures on the photosynthetic systems in spinach: oxygen-evolving activities, fluorescence characteristics and the denaturation process. *Photosynth Res.* **57**: 51-59.

Yamasaki H, Abdel-Ghany SE, Cohu CM, Kobayashi Y, Shikanai T, Pilon M. 2007. Regulation of copper homeostasis by micro-RNA in Arabidopsis. *J Biol chem.* **282**(22): 16369-16378.

Yamasaki H, Pilon M, Shikanai T. 2008. How do plants respond to copper deficiency? *Plant Signal Beh.* **3**(4): 231-232.

Yan J, Tsuichihara N, Etoh T, Iwai S. 2007. Reactive oxygen species and nitric oxide are involved in ABA inhibition of stomatal opening. *Plant Cell Environ.***30**(10): 1320-1325.

Yang H, Zhang H, Zhu L, Zhang C, Li D. 2012. Identification and Characterization of MicroRNAs in Macaca fascicularis by EST Analysis. *Comp Funct Genomics.* **2012**: 957607.

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

Yang T, Poovaiah BW. 2000. An early ethylene up-regulated gene encoding a calmodulin-binding protein involved in plant senescence and death. *J Plant Chem.***275**(49): 38467-38473.

Yang T, Xue L, An L. 2007. Functional diversity of miRNA in plants. *Plant Sci.* **172**(3): 423-432.

Yang X, Li L. 2011. miRDeep-P: a computational tool for analyzing the microRNA transcriptome in plants. *Bioinformatics.* **27**(18): 2614-2615.

Yang X, Zhang H, Li L. 2011. Global analysis of gene-level microRNA expression in Arabidopsis using deep sequencing data. *Genomics.* **98**(1): 40-46.

Yang XJ, Finnegan PM. 2010. Regulation of phosphate starvation responses in higher plants. *Ann Bot.* **105**(4): 513-526.

Yin Z, Li Y, Han X, Shen F. 2012. Genome-wide profiling of miRNAs and other small non-coding RNAs in the Verticillium dahliae-inoculated cotton roots. *PloS One* **7**(4): e35765.

Yoo SD, Sheen J. 2008. MAPK signaling in plant hormone ethylene signal transduction. *Plant Signal Beh.***3**(10): 848-849.

Yoshimura K, Yabuta Y, Ishikawa T, Shigeoka S. 2000. Expression of spinach ascorbate peroxidase isoenzymes in response to oxidative stresses. *Plant Physiol.***123**(1): 223-234.

Yoshikawa M, Peragine A, Park MY, Poethig RS. 2005. A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. *Genes Dev.* **19**(18): 2164-2175.

Yost HJ, Lindquist S. 1988. Translation of unspliced transcripts after heat shock. *Science* **242**(4885): 1544-1548.

Youm JW, Jeon JH, Choi D, Yi SY, Joung H, Kim HS. 2008. Ectopic expression of pepper CaPF1 in potato enhances multiple stresses tolerance and delays initiation of in vitro tuberization. *Planta* **228**(4): 701-708.

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

Yu X, Wang H, Lu Y, de Ruiter M, Cariaso M, Prins M, van Tunen A, He Y. 2012. Identification of conserved and novel microRNAs that are responsive to heat stress in Brassica rapa. *J Exp Bot.* **63**(2): 1025-1038.

Zaratiegui M, Irvine DV, Martienssen RA. 2007. Noncoding RNAs and gene silencing. *Cell.* **128**(4): 763-776.

Zeng Y, Yi R, Cullen BR. 2003. MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proc Natl Acad Sci U S A.* **100**(17): 9779-9784.

Zhai J, Jeong DH, De Paoli E, Park S, Rosen BD, Li Y, Gonzalez AJ, Yan Z, Kitto SL, Grusak MA et al. 2011. MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes Dev.* **25**(23): 2540-2553.

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

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

Zhang BH, Pan XP, Wang QL, Cobb GP, Anderson TA. 2005. Identification and characterization of new plant microRNAs using EST analysis. *Cell Res.* **15**(5): 336-360.

Zhang H, Gao Z, Zheng X, Zhang Z. 2012. The role of G-proteins in plant immunity. *Plant Signal Beh.***7**(10): 1284-1288.

Zhang J, Du X, Wang Q, Chen X, Lv D, Xu K, Qu S, Zhang Z. 2010. Expression of pathogenesis related genes in response to salicylic acid, methyl jasmonate and 1-aminocyclopropane-1-carboxylic acid in Malus hupehensis (Pamp.) Rehd. *BMC Res Not.* **3**: 208.

Zhang L, Zheng Y, Jagadeeswaran G, Li Y, Gowdu K, Sunkar R. 2011a. Identification and temporal expression analysis of conserved and novel microRNAs in Sorghum. *Genomics.* **98**(6): 460-468.

Zhang PJ, Broekgaarden C, Zheng SJ, Snoeren TA, van Loon JJ, Gols R, Dicke M. 2013. Jasmonate and ethylene signaling mediate whitefly-induced interference with indirect plant defense in Arabidopsis thaliana. *New Phyt.***197**(4): 1291-1299.

Zhang R, Marshall D, Bryan GJ, Hornyik C. 2013. Identification and characterization of miRNA transcriptome in potato by high-throughput sequencing. *PloS One.* **8**(2): e57233.

Zhang W, Gao S, Zhou X, Chellappan P, Chen Z, Zhang X, Fromuth N, Coutino G, Coffey M, Jin H. 2011b. Bacteria-responsive microRNAs regulate plant innate immunity by modulating plant hormone networks. *Plant Mol Biol.* **75**(1-2): 93-105.

Zhang X, Xia J, Lii YE, Barrera-Figueroa BE, Zhou X, Gao S, Lu L, Niu D, Chen Z, Leung C et al. 2012. Genome-wide analysis of plant nat-siRNAs reveals insights into their distribution, biogenesis and function. *Genome Biol.* **13**(3): R20.

Zhao DL, Reddy KR, Kakani VG, Koti S, Gao W. 2005. Physiological causes of cotton fruit abscission under conditions of high temperature and enhanced ultraviolet-B radiation. *Physiol Plant.***124**: 189-199.

Zhao Z, Chen G, Zhang C. 2001. Interaction between reactive oxygen species and nitric oxide in drought-induced abscisic acid synthesis in root tips of wheat seedlings. *Au J Plant Physiol.***28**(10): 1055-1061

Zheng D, Ezzeddine N, Chen CY, Zhu W, He X, Shyu AB. 2008. Deadenylation is prerequisite for P-body formation and mRNA decay in mammalian cells. *J Cell Biol.* **182**(1): 89-101.

Zhong S, Lin Z, Grierson D. 2008. Tomato ethylene receptor-CTR interactions: visualization of NEVER-RIPE interactions with multiple CTRs at the endoplasmic reticulum. *J Exp Bot.* **59**(4): 965-972.

Zhou M, Li D, Li Z, Hu Q, Yang C, Zhu L, Luo H. 2013. Constitutive expression of a miR319 gene alters plant development and enhances salt and drought tolerance in transgenic creeping bentgrass. *Plant Physiol.* **161**(3): 1375-1391.

Zhou X, Ren L, Meng Q, Li Y, Yu Y, Yu J. 2010. The next-generation sequencing technology and application. *Protein Cel.l* **1**(6): 520-536.

Zhu C, Ding Y, Liu H. 2011. MiR398 and plant stress responses. *Physiol Plant.* **143**(1): 1-9.

Zhu QH, Helliwell CA. 2011. Regulation of flowering time and floral patterning by miR172. *J Exp Bot.* **62**(2): 487-495.

Zou J, Liu A, Chen X, Zhou X, Gao G, Wang W, Zhang X. 2009. Expression analysis of nine rice heat shock protein genes under abiotic stresses and ABA treatment. *J Plant Physiol.***166**(8): 851-861.

Zubko E, Meyer P. 2007. A natural antisense transcript of the Petunia hybrida Sho gene suggests a role for an antisense mechanism in cytokinin regulation. *Plant J. New phytol.* **52**(6): 1131-1139

CHAPTER II

IDENTIFICATION OF SMALL REGULATORY RNAS IN *GOSSYPIUM RAIMONDII* L. AND DETERMINATION OF THEIR ROLE IN THE LEAF HEAT STRESS RESPONSE

Abstract

Six small RNA libraries from heat-stressed and non-heat-stressed leaves of *G. raimondii* seedlings were sequenced using Illumina deep sequencing. This generated approximately 90 million raw reads, and about 43% of them exactly matched the *G. raimondii* genome. To analyze the *G. raimondii* small RNA transcriptome, a bioinformatic pipeline was built and used to analyze the *G. raimondii* small RNA data. This allowed us to profile and to annotate the five major classes of small RNAs in the *G. raimondii* genome yielding 146 miRNAs, 2,236 ohpRNAs, 268,246 ra-siRNAs, 791 pha-siRNAs, and 147,058 *cis*-NAT-siRNAs. Statistical data analysis (Baggerley's test) revealed that 45 miRNAs, 51 ohpRNAs, 4,278 ra-siRNAs, 28 pha-siRNAs and 1,415 *cis*-nat siRNAs were significantly differentially expressed in response to either 4 h or 24 h of heat stress treatment. This is consistent with a role for small RNAs in the *G. raimondii* response to heat stress deriving from involvement of ra-siRNA in changing DNA methylation and hpRNA and cis-NAT-siRNA mediated changes at the transcriptional and/or posttranscriptional levels.

Introduction

RNA silencing (RNAi) is a specific mechanism that regulates the developmental, stress-responses, and defense functions in eukaryotes by guiding mRNA cleavage/translational inhibition, or chromatin modification (Vazquez et al. 2010; Khraiwesh et al. 2012) (Saze et al. 2012). RNAi is mediated by a class of 20 to 24 nt small RNAs (or small regulatory RNAs) that are processed from long precursors by DICER, an RNaseIII type ribonuclease, in animals or DICER-like proteins (DCLs) in plants (Axtell et al. 2011) . These small RNAs are loaded into

Argonaute (AGO) proteins to form RNA induced silencing complex (RISC) in which they complimentarily bind to target RNAs or DNAs (Rogers and Chen 2013).

Small RNAs can be classified into two distinct families, short interfering RNAs (siRNAs) and hairpin RNAs (hpRNAs), on the basis of their origin and function (Axtell 2013). siRNAs arise from linear double stranded precursors, while hpRNAs are produced from single-stranded, hairpin-shaped precursor (Axtell 2013).

Based on their biogenesis and function, siRNAs can be divided into three major subgroups, repeat associated siRNAs (ra-siRNA) (also referred to as heterochromatic siRNAs or hc-siRNAs), phased secondary siRNAs (pha-siRNAs) (including trans-acting siRNAs that have known targets in gene silencing also referred to as ta-siRNAs), and natural antisense siRNAs (NAT-siRNAs) (Jamalkandi and Masoudi-Nejad 2009; Allen and Howell 2010). ra/hc-siRNAs typically are 24 nt small RNAs that are produced from repetitive sequences by the RNA polymerase IV/RDP2/DCL3 pathway (Chapman and Carrington 2007; Xie and Qi 2008). pha/ta-siRNAs are a class of secondary siRNAs that are produced from so-called phased secondary RNA gene transcripts (PSRG or *TAS* transcripts)(Allen and Howell 2010). Their synthesis is dependent on a miRNA or siRNA trigger (Allen et al. 2005) and requires the involvement of DCL4, RDP6 (RNA-dependent RNA polymerase6) and SGS3 (suppressor of gene silencing3) (Allen and Howell 2010). NAT-siRNAs are produced from complementary mRNA transcripts (or natural antisense transcripts, NATs), and their synthesis is dependent on RDP6, and DCL1, DCL2 and/or DCL3 (Borsani et al. 2005; Zhang et al. 2012b). There are two types of NAT-siRNAs, *cis*-NAT-siRNAs and *trans*-NAT-siRNAs (Borsani et al. 2005; Zhang et al. 2012b). *cis*-NAT-siRNAs are produced from NATs transcribed from the same genomic loci, while *trans*-NAT-siRNAs arise from NATs that are transcribed from different genomic loci (Axtell 2013).

Based the characteristics of biogenesis, hpRNAs can be divided into two subgroups, micro RNAs (miRNAs) and other hpRNAs (ohpRNAs) (Axtell 2013). Canonical miRNAs are ~21nt small RNAs that are precisely cleaved by DCL1 and posttranscriptionally regulate the gene expression (Meyers et al. 2008; Axtell et al. 2011), Whereas so-called long miRNAs (lmiRNAs), ~24nt miRNAs, are cleaved by DCL3 and mediate DNA methylation(Wu et al. 2010). ohpRNAs arise from cleavage by possibly one or two different DCL proteins from long inverted repeat sequences up to several kb in length, leading to one or more sizes of small RNAs (Henderson et al. 2006; Dunoyer et al. 2010; Yelina et al. 2010). ohpRNAs mediate the expression of genes either transcriptionally or posttranscriptionally (Henderson et al. 2006; Dunoyer et al. 2010). To date, the control of ohpRNAs biogenesis remains largely unknown, but it appears that this may be determined by promoter components (Wang et al. 2008).

miRNAs are a well-studied subclass of hpRNAs that are produced by precise cleavage of DCLs in plants and that function to regulate the expression of genes at the posttranscriptional level (Huang et al. 2013; Rogers and Chen 2013). miRNAs are transcribed by RNA polymerase II from miRNA genes (Meyers et al. 2008), and the primary transcripts (pri-miRNAs) experience modification by adding 3'-ploydenosine and a 5'-cap (Chen 2009). Pri-miRNAs are processed by DCLs into precursor miRNAs, and then the miRNA-miRNA* (miRNA star, the pattern strand of miRNA) duplexes (Naqvi et al. 2012). The miRNA-miRNA* duplexes are stabilized through methylation at 3'-ends by HEN1 (Yang et al. 2006) and then exported into cytoplasm by the Exportin 5 homologue, HASTY (Park et al. 2005). In the cytoplasm, mature miRNAs are recruited into an Argonaute protein complex (RISC) in which they bind to target mRNAs by Watson-Crick base pairs (Rogers and Chen 2013). This subsequently results in the degradation of the target mRNA or in mRNA translation inhibition (Rogers and Chen 2013).

Heat stress caused by elevated temperature adversely affects an array of physiological processes from anatomical, biochemical, to gene levels, and often leads to a drastic reduction in economic yield (Wahid et al. 2007). The mechanism of plant heat responses have been well-documented at the biochemical and gene levels and have recently been reviewed in (Hasanuzzaman et al. 2013; Qu et al. 2013). The expression of genes regulated by small RNAs during plant responses to a variety of stresses is conserved among the various stresses and common to many if not all plants (Khraiwesh et al. 2012; Guan et al. 2013; Lu et al. 2013).

Studies have shown that miRNAs are specifically involved in plant responses to thermal stress (Sunkar et al. 2006; Xin et al. 2010; Khraiwesh et al. 2012; Yu et al. 2012). Of stress-responsive miRNAs, miR398 is a well-studied stress-responsive miRNAs and regulates plant tolerance to various oxidative stresses by controlling the expression of Cu/Zn superoxide dismutase (CSD) (Sunkar et al. 2006; Zhu et al. 2011). Over-expressing a miR398-resistant form of *CSD* genes induces stress tolerance due to the accumulation of CSD mRNAs (Sunkar et al. 2006). In addition to miR398, Yu et al. have shown that miR156h/g of *Brassica rapa* are induced by heat stress and one of their targets is *BracSPL2* (Squamosa Promoter Binding Protein-Like). This transcription factor is down-regulated by heat stress (Yu et al. 2012). Recently, some ohpRNAs were also shown to be responsive to heat stress and pathogen infection (Xin et al. 2011).

Gossypium raimondii is a diploid species that is considered one of the progenitor species of modern fiber-bearing, commercial cottons, including upland (*G. hirsutum*) and Pima (*G. barbadense*) cottons. A number of studies have been conducted that have identified several hundred miRNAs in *G. hirsutum,* and a subset of cotton miRNAs were shown to be involved in cotton fiber development (Pang et al. 2009; Wang et al. 2012). However, little is known about miRNAs of *G. raimondii* and their roles in the heat response.

In this study, we applied high-throughput, deep sequencing to profile *G. raimondii* small RNAs and subsequently used a bioinformatic pipeline to annotate the small RNA population expressed in the stressed and unstressed leaves of *G. raimondii*. This identified a large number of hpRNAs and siRNAs, and their expression in response to heat stress was subsequently analyzed. Our findings provide an insight into small RNAs that are associated with regulatory roles in the *G. raimondii* heat stress response.

Materials and Methods

Growth of plant material, total RNA preparation, and sRNA library sequencing

The youngest fully expanded leaves of four-week-old *G. raimondii* seedlings were used to obtain total RNA. Samples were harvested from at least 3 seedlings treated at 46^{o}C. for 0 h (no heat stress treatment), and after 4 or 24 h of heat stress (heat stress treatment). The total RNA was separately isolated from each sample using hot borate method (Wan and Wilkins 1994). The constructing and sequencing of small RNAs was performed at the HudsonAlpha Institute Genome Sequencing Laboratory in Huntsville, AL, USA using an Illumina Small RNA Preparation Kit following manufacturer's protocol. Each small RNA sample was sequenced twice generating 2 sequencing replicates of each sample.

G. raimondii transcriptome sequences

For subsequent bioinformatic work it was necessary to produce a set of transcriptome sequences derived from the *G. raimondii* genome since existing publically available unigene and EST datasets were limited in scope. A GraTS dataset was built using the GORAGI.release_1 *G. raimondii* unigene sets (ftp://occams.dfci.harvard.edu/pub/bio/tgi/

data/Gossypium_raimondii/GORAGI.release_1.zip), the G. raimondii unigene set at NCBI (ftp://ftp.ncbi.nih.gov/repository/UniGene/), and an ESTs dataset constructed from Trinity assembled Illumina sequencing reads from *G. raimondii* leaf (unpublished data).

Processing of sequencing data

The raw sequencing data was processed using a pipeline consisting of a combination of custom-designed scripts, open source scripts, and CLC WorkBench 5.8 (http://www.clcbio.com/). Raw sequence data was filtered to remove sequencer reads lacking a 3' sequencing adaptor sequence, and the adaptor sequences and low quality bases were removed using CLC WorkBench 5.8. After filtering sequences 17 to 28 nt were retained for further analysis and designated as the preliminary sRNA libraries. Sequences from each of the 6 libraries and from a combined total library were separately collapsed to create read counts for each unique sequence in each of the single libraries and for the combined library.

The unique libraries were mapped to the genome of *G. raimondii*, version 2 (*G.raimondii*_v2) (Paterson et al. 2012), using the Bowtie mapping utility, version 1.0.0 (Langmead 2010), and only sequences exactly matching the *G. raimondii* genome were retained for further analysis. Unique sequences that exactly matched the non-protein coding sequences (rRNAs, tRNAs, snRNAs, and snoRNAs) were removed from the genome matching sequences above using the Bowtie mapping utility leaving total sRNA libraries for each treatment and replication and a combined total sRNA library that were used for subsequent analysis.

Bioinformatics determination of hairpin RNAs

The total sRNA library above was used to extract hairpin precursors using the open source miREAP script (http://sourceforge.net/projects/mireap/). To obtain the maximal number of hpRNAs, miREAP was employed by searching the *G.raimondii*_v2 genome sequence (Paterson et al. 2012) with the collapsed sRNA library using three combinations of two parameters: 1) the maximum distance between the putative sRNA and the corresponding complementary putative sRNA* sequences, and 2) the maximal number of copies of small RNAs found in the genome. The three combinations used were: 1a) 350nt and 2a) 12 loci, 1b) 350nt and 2b) 232 loci, and 1c) 500 nt and 2c) 500 loci. The derived putative hairpin precursors were examined using CentroidFold0.9 (Sato et al. 2009), considering two aspects: 1) unpaired nucleotides in the stem of the hairpin were less than 50% of the total; and 2) no more than four mismatches between hpRNA and hpRNA* were allowed. The putative hpRNAs passing the above filter were collapsed by removing duplicated sequences the summed dataset from the 3 runs using a custom Perl script. This generated a list of all predicted hpRNAs in the total sRNA library, and a library of the remaining non-hpRNAs that was used for siRNA determination as indicated below.

Identification of miRNAs

MiRNAs were identified using a pipeline of custom scripts and the miREAP open source script (http://sourceforge.net/projects/mireap) as follows. The hpRNA dataset (see above) was submitted to a custom Perl script that retained unique sequences with a read count of at least eight and that mapped to the *G.raimondii*_v2 genome at no more than 12 loci. These parameters were empirically determined to be the minimal values that permitted determination of known *Gossypium* sequences in the total sRNA library without inclusion of low abundance sequences that had questionable significance as legitmate sRNAs.

The read count ≥ 8, genome loci ≤ 12 sequences were submitted to the miREAP script and putative miRNA precursors were identified based on four characteristics: 1) minimum free energy of the predicted precursor ≤ -30 kcal/mol; 2) ≤ 4 four mismatches between the putative miRNA and putative miRNA*; 3) ≤ 2 "bulges" between the putative miRNA and putative miRNA*; and 4) the maximal distance between miRNAs and miRNA* ≤ 350nt. These parameters were chosen based on previously published values (Thakur et al. 2011) and empirical optimization with the datasets used in this study.

The miREAP output was further filtered using a custom Perl script that selected unique precursor sequences with a stand bias ≥ 0.9, calculated by dividing the count of sRNA reads mapping to a putative miRNA precursor in the forward manner by the total sRNA reads mapping to the putative precursor. Additionally, this script selected precursors based on an abundance bias ≥ 0.6, calculated by dividing the count of the most abundant and the second most abundant sRNA reads for a precursor by the total count of sRNA reads mapping that precursor according to Jeong et al. (Jeong et al. 2011; Zhai et al. 2011). The most abundant sRNAs mapping to a given precursor were condsidered to be miRNA candidates, and the second most abundant were considered either the miRNA* or in the miRNA region but truncated [16, with empirical optimization for these datasets].

The miRNA candidates passing the strand-bias and abundance bias filters were subsequently filtered based on secondary structure using CentroidFold0.9 (Sato et al. 2009), and an additional Perl script considering six criteria: 1) the maximal unpaired nucleotide bulges located in the miRNA:miRNA* regions are ≤ 2, and there are ≤ 3 nucleotides per bulge; 2) there are ≤ 4 mismatches allowed between miRNA and miRNA*; 3) MFEI (minimum free energy index) (Zhang et al. 2006) is between 0.38 and 1.6; 4) the unpaired nucleotides in the secondary

structure are ≥ 50% of total length; 5) CG content is ≥ 25%; 6) AMFE (adjusted minimum free energy) (Zhang et al. 2006) is ≤ -22 kcal/mol/100nt.

Candidate hpRNAs passing the secondary structure filter above were annotated as putative miRNAs, and those hpRNAs not annotated as miRNAs were annotated as other hpRNAs (ohp-RNAs). Subsequently, those miRNAs with lengths of 23-24 nt were annotated as long-miRNAs, while those 20 through 22 nt in length were considered as canonical miRNAs, and those canonical miRNAs with two or fewer mismatches to mature miRNAs registered as Viridiplantae miRNAs in miRBase version 20 (http://www.mirbase.org) were annotated as conserved miRNAs (con-miRNAs). The remainder of the canonical miRNAs were annotated as novel *G. raimondii* lineage-specific miRNAs (ls-miRNA). Note that only 891 of the 7385 miRNAs found in miRBase version 20 (12.1%) were ≥ 23 nt in length so a separate conserved long-miRNA class is not necessary at this time.

Bioinformatics analysis of repeat-associated siRNAs siRNAs.

A first step in identifying repeat-associated siRNAs was to generate a plant repeat dataset for use. Plant repeat sequences derived from both the Repbase Update database housed at Genetic Information Research Institute (GIRI, http://www.girinst.org) (Jurka et al. 2005) and the Plant Repeat Database (http://plantrepeats.plantbiology.msu.edu) (Ouyang and Buell 2004) were combined to generate a Viridiplantae repeat database (VRD) consisting of 21,141 putative repeat sequences found in plant genomes.

To identify repeat-associated siRNAs (ra-siRNAs), the *G. raimondii* non hpRNA sequences (see above) with read counts ≥ 2 were compared to the above VRD dataset, allowing for a maximal substitution rate of 20% (default) using ReaptMasker3.2.7 (http://www.repeatmasker.org)

(Tempel 2012). The sequences resulting from this were annotated as ra-siRNAs, while the remaining sequences in the non hpRNA library were left as the non ra-siRNA library.

Bioinformatics analysis of phased siRNAs (ta-siRNAs)

Phased siRNAs (phasiRNAs) and phasiRNA-producing transcripts were identified by submitting the non-ra-siRNA library from the previous step to the ta-siRNA prediction tool in the plant version of the UEA sRNA Workbench version 3.0 (Chen et al. 2007; Stocks et al. 2012). Sequences found in the *G.raimondii*_v2 genome (Paterson et al. 2012) and the GraTS dataset (see above) using a phasing register of 21, 22, or 24 nt and a p-value $\leq$ 0.01 were obtained along with their putative secondary siRNA producing transcrips. A custom Perl script was subsequently used to identify those transcripts where the abundance of sRNAs in the phased loci divided by the total sRNAs matching the transcript is $\geq$ 0.6. Such transcripts were annotated as phased-siRNA-producing transcripts (sometimes historically referred to as TAS genes), and sRNAs in phased loci were considered as phased-siRNAs. Phase-initiating miRNAs were predicted using a combination of the psRNATarget (Dai and Zhao 2011) utility and a custom Perl script that determines potential cleavage site of the phased-siRNA-producing transcripts based on previously published methods (Dai and Zhao 2008; Zhang et al. 2012a). All non ra-siRANs not annotated as phased-siRNAs were considered non-phased-siRNAs.

Bioinformatics analysis of *cis*-natural antisense siRNAs

Putative natural antisense transcripts *cis* natural antisense transcripts (*cis*-NATs) were identified as overlapping, opposite strand transcript pairs using a Perl script that was developed to qualify the putative *cis*-NATs by mapping the putative pairs to the *G.raimondii*_v2 genome by using the SSAHA2 utility, version2.5.5 (Ning et al. 2001). *cis*-NAT sequences matching the genome with greater than 98% identity having an overlap of $>$ 50 bp located on opposite strands at the same

non-repetitive locus in the genome and that formed a stable RNA-RNA heteroduplex using the RNA cofolder utility in the Vienna RNA Package (http://www.tbi.univie.ac.at/RNA/) (Lorenz et al. 2011) were annotated as *cis*-NAT-pairs. Sequencces found in the non-phased-siRNA dataset from the previous step exactly matching the *cis*-NATs were annotated as *cis*-NAT-siRNAs, while all remaining sequences in the non-phased-siRNA dataset, not annotated as *cis*-NAT-siRNAs were considered as uncharacterized siRNAs.

Statistical analysis

Baggerley's test was conducted using the statistical package built into CLC Genomic WorkBench 5.5.8. (Baggerly et al. 2003).

Target prediction and GO analysis of target genes

The prediction of miRNA and siRNA targets was performed with psRNATarget (Dai and Zhao 2011) using the GraTS dataset (see above) with a maximum expectation of 2.5 and other parameters set at default.

The predicted target genes were subjected to Gene Ontology analysis performed using Blast2GO (Conesa and Gotz 2008). A Blastx search of the NCBI protein database was conducted, and target genes hitting matching the database were annotated according to the GO database.

Results

Deep sequencing of small RNA libraries from heat stress-treated *G. raimondii* seedlings

Illumina deep sequencing was employed to generate six *G. raimondii* small RNA libraries. The six small RNA libraries were made from the youngest, fully expanded leaves from *G. raimondii* seedlings at 0h (before), 4h, or 24h of heat-treatment at 45 ^{o}C. Each small RNA sample was sequenced twice creating sequencing replicates. This generated a total sequence library for all treatments of just over 90 million total reads that were used for sRNA prediction, and the raw sequencing reads of individually sequenced libraries varied from 12,494,803 to 16,275,490 reads (**Table 1** and **Supplemental Table S1**).

The first step was to filter out reads missing sequencing adaptors, remove the adaptors from the remaining reads, and subsequently to filter out reads that contain low-quality nucleotide predictions. Reads shorter than 17 nt, and sequences longer than 28 nt were excluded from the analysis at this point as well. As shown in **Table 2** and **Supplementary Table S1**, this left 46,738,621 total reads which could be assembled into 8,057,647 unique putative sRNA sequences. Of those, 39,534,260 reads (84.6%), represented by 5,609,335 unique sequences, exactly matched the *G. raimondii* genome.

The total small regulatory RNA library was established by removing the sequences that corresponded to the non-protein coding sequences from the sequences exactly matching the *G. raimondii* genome (EMG sequences). This left a total of 25,885,344 reads (65.5% of EMG reads) that corresponded to 5,490,184 unique sequences. These sequences will subsequently be referred to as the total sRNA library.

Diversity of small regulatory RNAs in *G. raimondii*

As shown in **Table 2**, 475,241 reads (1.20% of EGM), which can be collapsed into 2,375 unique sequences, arose from hairpin-shaped precursors and were annotated into hairpin RNAs

(hpRNAs). Of those, 448,133 reads, corresponding to 146 unique sequences, met the criteria of plant miRNAs (Zhang et al. 2006; Meyers et al. 2008; Jeong et al. 2011) and were annotated as miRNAs, whereas 27,699 reads represented by 2,236 unique sequences failed to meet the criteria for consideration as plant miRNAs (Zhang et al. 2006; Meyers et al. 2008; Jeong et al. 2011), and these were annotated as other hairpin RNAs (ohpRNAs). Notably, seven unique sequences were annotated into both miRNAs and ohpRNAs, because they were produced from multiple genomic loci, and a subset of these loci failed to meet the criteria for annotation as plant miRNA precursors.

To identify the ra-siRNAs, the repeat sequence loci in the *G. raimondii* genome were first identified by sequence comparison with plant repeat sequences deposited in the Plant Repeats Database (http://plantrepeats. plantbiology.msu.edu and http://www.girinst.org) using Repeat Masker 3.2.7 (http://www.repeatmasker.org/) (Tempel 2012) allowing up to 20% base substitution in the homologous repeat regions. Consequently, 12,645,751 repeat-associated reads were identified corresponding to 159,701 unique repeat-associated sRNA (ra-siRNA) sequences. Repeat-associated siRNAs were the most abundant class of the *G. raimondii* sRNAs, accounting for 58.9% of the total sRNA reads and 13.2% of the unique sequences.

The non-ra-siRNA library was subjected to phasing analysis, and as a result, we obtained 52,060 putative secondary phased siRNA (pha-siRNA) reads, which can be collapsed into 791 unique sequences derived from 296 phased siRNA-producing transcripts. These 791 putative pha-siRNA sequences were annotated as phased siRNAs, and the remaining sequences in the non-ra-siRNA library were considered the non-pha-siRNAs.

A total of 5,247,152 reads (1.0% of the total sRNAs), corresponding to 147,058 unique sequences were identified by the sRNA pipeline as *cis*-nat-siRNAs, making the *cis*-nat-siRNAs the second most abundant group of sRNAs in the *G. raimondii* sRNA library. These *cis*-nat-siRNAs were derived from a 1,001 expressed *cis*-NAT transcript pairs from the 12,114 putative total *cis*-NAT pairs found in the G. raimondii genome.

Unexpectedly, 33.1% of the total reads or 92.4% unique sequences were not categorized in this study. Most of the uncharacterized reads were 24-nt in length, which suggests that these sequences may be siRNAs, because 24-nt siRNAs are the most abundant sRNA species found in plants (Kasschau et al. 2007). However, the average read number of the uncharacterized sequences was 1.4 indicating that many of these sequences have low expression in only one library. Thus, it is impossible to annotate these low abundance reads at this time based on the available data.

Characteristics of *G. raimondii* miRNAs

As shown in **Supplemental Table S2**, 146 mature miRNAs derived from 213 distinct loci were identified. These included 46 mature miRNA sequences that were conserved in one or more species of Viridiplantae according to miRBase version 20 that were derived from 96 distinct *G. raimondii* lineage-specific loci, and 100 novel, *G. raimondii* lineage-specific mature miRNAs (ls-miRNA) produced from 107 *G. raimondii* lineage-specific loci. The *G. raimondii* miRNAs ranged from 19 nt to 24 nt in length, but no miRNAs 23 nt in length were observed. miRNA < 23 nt in length were considered canonical miRNAs, whereas miRNAs (> 23 nt) were annotated as long miRNAs (Axtell et al. 2011).

As shown in **Figure 2A**, conserved mature miRNAs of *G. raimondii* ranged from 19 nt to 22 nt in length, and most of the conserved miRNAs (about 71%) were 21 nt long, whereas only 46% of ls-miRNAs were 21 nt long; approximately 40% of ls-miRNAs were 24 nt in length. Uridine was the most common 5'-terminal nucleotide for miRNAs less than 23 nt in length (**Figure 2B**) and a low occurrence of 5'-adenine (less than 22%) appeared in the canonical miRNAs (Figure 1B). In contrast to canonical miRNAs, 5'-adenine was predominant among long miRNAs and counted for approximately 80%. Interestingly, a low occurrence of 5'-cystine (less than 19%) and no 5'-guanine was detected among *G. raimondii* miRNAs. Previous studies show that 5'-nucleotide is essential for sorting miRNAs and loading onto AGO proteins (Mi et al. 2008). Thus, this characteristic may be functionally important for *G. raimondii* miRNAs.

G. raimondii is a putative progenitor of economically important, allotetraploid cottons, such as upland cotton (*G. hirsutum* L.) and Pima (*G. barbadense* L.). Currently, 80 miRNAs of upland cotton derived from 78 precursors are registered in miRBase (Release 20. To investigate sequence variations of miRNA genes between *G. raimondii* and *G. hirsutum*, we performed an analysis of sequence alignment using BLASTN (Boratyn et al. 2013). As a result, 14 *G. raimondii* miRNA precursors were exactly mapped to 15 *G. hirsutum* precursors, whereas 22 *G. raimondii* miRNA precursors can be aligned to 17 *G. hirsutum* miRNA precursors with one or more mismatches. As shown in **Figure 3A**, the sequence variations appeared at different regions of miRNA precursors, including miRNA, miRNA* (the pattern strands of miRNAs), loop or/and stem regions. Phylogenetic analysis showed that gra-MIR7459#a (the precursor for gra-miR7459#a of *G. raimondii*) and ghr-MIR7459a/b (the precursors for ghr-miR7459a/b) were in different clades (**Figure 3B**). It should be noted that variations in miRNA regions influenced the regulatory function of miR7495. As shown in **Figure 3C,** among the 38 putative targets of

miR7495, only 5 (approximately 13%) were shared between gra-miR7495#a and ghr-miR7495a/b. These findings revealed that the divergence between miRNAs of *G. raimondii* and *G. hirsutum* is present and the variation in sequences may be functionally important for gene regulation in cotton.

miRNA-directed phased secondary siRNA networks

In this study, a total of 296 pha-siRNA gene (PSRG) candidates were identified, of which 177 PSRGs (60%) produced 21 nt pha-siRNAs, 48 PSRGs (16%) generated 22 nt pha-siRNAs, and 71 others (24%) yielded 24 nt pha-siRNAs (**Figure 4** and **Supplemental Table S3-S5**). Among these PSRGs, 255 (~87%) were detected in the GraTS dataset, and were therefore considered transcriptionally active. Of the 177 PSRGs that produced pha-siRNAs 21 nt in length, 89 (50.3%) were found to code for known protein coding genes, while 53 (29.9%) corresponded to non-protein-coding sequences, and 35 (19.8%) corresponded to uncharacterized protein coding genes. Of the 89 protein coding genes, the largest group (32) was putative NBS-LRR disease resistance proteins. Other PSRGs included 22 mutator proteins, 12 auxin response factor/signaling F-box proteins, 7 MYB or bHLH transcription factors, 6 polyproteins, 4 NAC domain protein, and 7 others (**Figure 4A**). By contrast, 24/48 and 60/71 PSRGs produced pha-siRNAs 22nt and 24nt in length respectively that corresponded to non-protein-coding genes, and 12/48 and 2/71 PSRGs produced pha-siRNAs 22nt and 24nt in length respectively that corresponded to uncharacterized/hypothetical protein genes, while 11/48 and 9/71 PSRGs produced pha-siRNAs 22nt and 24nt in length respectively that corresponded to known protein coding genes (**Figure 4B** and **4C**).

A total of 123 PSRGs (~39% of the total) were predicted to be triggered by 55 miRNAs (**Supplemental Figure 1A**, **1B**, and **1C**). The miRNA triggers included 19 conserved and 36 lineage-specific miRNAs ranging from 19 nt to 22 nt in length, and most of them (~79%) were 21 nt in length. Interestingly, a subset of nine *G. raimondii* PSRGs were targeted by at least two or more, distinct miRNAs.

A PSRG, TAS3, found other plants (Montgomery et al. 2008; Marin et al. 2010), which encodes auxin response factor (ARF), is an apparently conserved plant PSRG (TAS gene) triggered by miR390. As shown in **Figure 5A and 5B**, the cleavage site of miR390 is highly conserved across at least 5 plant species. Three *G. raimondii* mR390-targeted TAS3-like sequences were character-ized, and each had the ability to produced 7 to 9 pha-siRNAs. In order to better understand the roles of these putative pha-siRNAs, we performed a target analysis using pssRNA Target (Dai and Zhao 2011). These pha-siRNAs were predicted to regulate diverse protein-coding genes, of which 22 to 41% were proteins associated with transcription, such as auxin response factors. Other pha-siRNA targets are enzymes, such as protein kinases and other signaling enzymes (**Figure 5C** and **Supplemental Table S9**).

In addition to known, conserved miRNA-TAS pathways, such as miR390-TAS3 (ARF), miR482-NBS-LRR, miR393-MYB, a number of putative novel pathways were found, and a subset of pha-siRNAs yielded by these pathways were abundantly represented in the total small RNA library (**Supplemental Table S3 to S5**). As shown in **Figure 6A**, miR2218/miRC136 that targeted TMV resistance protein N-LIKE (TMV RPN) defined a novel pha-siRNA-producing PSRG pathway found only in *G. raimondii*. In this pathway, six pha-siRNAs were highly expressed (1.5- 69 RPM) (**Figure 6A**). This putative PSRG is a conserved plant protein, and the

cleavage site of miRNAs is also conserved in plants (**Figure 6B**). This is consistent with this being a potential conserved PSRG/TAS pathway, though this is not reported in other plants.

In the absence of DCL4, DCL2 and DCL3 may act to produce 22 nt or 24 nt pha-siRNAs respectively (Axtell et al. 2006; Song et al. 2012; Zhang et al. 2012a). Forty-eight PSRGs can produce pha-siRNAs 22 nt in length and 71 PSRGs can produce pha-siRNAs 24 nt in length (**Supplemental Table 5**). A number of pha-siRNAs in these groups accumulated to a high levels, for example, the D1(+) siRNA produced from a MYB-like gene on chromosome 9 produced 4,728 reads (**Supplemental Table S4**). Combined with previous studies, this provides additional evidence that 22 nt and 24 nt pha-siRNAs have functional significance in plants.

Natural antisense transcripts and *cis*-NAT derived siRNAs

A total of 12,114 *cis*-NAT gene pairs were identified in the *G. raimondii* genome (see Methods). According to the overlapping patterns, these NATs can be classified into three major types, tail-to-tail (**Figure 7A),** head-to-head **(Figure 7B**) and enclosed (**Figure 7C)**. As shown in **Figure 7**, sRNAs were not randomly distributed, but displayed obvious expression peaks at distinct locations in both the overlapping and non-overlapping regions of the *cis*-NAT pair. Additionally, 1,001 (8%) of the total *cis*-NAT pairs produced siRNAs found in our leaf databases. The average total *cis*-nat-siRNA read count produced from each expressed *cis*-NAT pair was 616, while only 755 (6%) of the *cis*-NAT gene pairs produced total *cis*-nat-siRNA read counts greater than 500 reads per NAT as shown in **Figure 8A**, and only 230 *cis*-NAT gene pairs (1.9%) accumulated *cis*-nat-siRNAs reflecting greater than 50,0000 counts per sequence.

The 1,001 cis-NAT gene pairs yielded 147,058 unique *cis*-nat-siRNA sequences, corresponding to 5,247,152 total reads (**Table 2**). According to the expression value, a subset of *cis*-nat siRNAs

accumulated to a very high level, of which the top 10 expressed *cis*-nat-siRNAs are shown in **Figure 8B**. The most abundantly expressed cis-nat-siRNAs, *cis*-nat-siR49762 and *cis*-nat-siR145324 were among the most abundantly expressed sequences (307,626 and 210,953 reads per sequence) in the entire library, and the ten most abundantly expressed *cis*-nat-siRNAs showed significant changes in expression during heat stress treatment (**Figure 8C**). Specifically, five of the abundantly expressed *cis*-nat-siRNAs were down-regulated by as much as 2.5 fold and five were up-regulated by as much as 3 fold after 4h heat stress treatment, but after 24 h heat stress treatment all ten of these *cis*-nat-siRNAs returned to within 20% of their unstressed levels.

Global expression of sRNAs during heat stress

As shown in **Figure 9A**, sRNAs were not evenly distributed along the chromosome, but displayed site-specific expression patterns (so-called "hot-spots", defined as the loci in 1,000 nt on the chromosomes where the sRNAs accumulated to at least 5,000 RPM). Among the 13 *G. raimondii* chromosomes a total of 111 hotspots (**Supplemental Figure S2 to S14**) were identified similar to those shown in Figure 8A for chromosome 1. In a 770 nt region between position 55,849,200 and 55,849,970 nt) on chromosome 1, there were 12 loci where individual or multiple sRNAs accumulated to levels of more than 1,000 RPMs. Among these "hot-spots", shown sRNAs that were in the regions around 55,849,237 on chromosome 1 accumulated to the highest level (~23,000 RPM) although a second "hot spot around position $21{,}000\times10^3$ nt also showed high levels of expression of multiple sRNAs. It can also be seen that there are other regions on chromosome 1 (see **Figure 9A**) where expression of fewer individual sRNAs is shown at nearly or even higher levels. Consistent with sRNAs playing an important regulatory role in the heat stress response, the expression patterns of *G. raimondii* sRNAs in this region of

chromosome 1 displayed significantly altered expression patterns during heat stress (**Figure 9B**). The sRNAs in the hotspots showed, in various degrees, expression changes (**Figure 9C**).

Differentially expressed small regulatory RNAs during heat stress

As shown in **Figure 10** and **Supplemental Table S10-S13**, 45 miRNAs, 51 ohpRNAs, 4,278 ra-siRNAs, 28 pha-siRNAs, and 1,415 *cis*-nat siRNAs were significantly regulated during heat treatments. Of these, 30 miRNAs, 35 ohpRNAs, 3,498 ra-siRNAs, 24 pha-siRNAs, 1,274 *cis*-nat siRNAs showed more than 2 fold changes (up or down) in expression level in either 4 hrs or 24 hrs of heat stress compared to unstressed control.

miRNAs showing differential expression patterns were classified into five clades (**Figure 11**). Clade 1 included 7 miRNAs that were up-regulated in the 4h heat stress library and down-regulated in the 24 h heat stress library. Clade 2 consisted of 10 miRNAs that were down-regulated in both 4h and 24h library. Clade 3 was composed of 5 miRNAs that were up-regulated in both 4h and 24h libraries. Clade 4 included 14 miRNAs that were down-regulated in the 4h library, whereas returned to higher levels in the 24h libraries. Clade 5 included 9 miRNAs that were down-regulated in 4h and up-regulated in the 24h heat treatment.

Among differentially expressed miRNAs, 29 were conserved miRNAs, including known stress-responsive miRNAs reported in other plants, such as miR156, miR172, miR396 and miR398 (Zhou et al. 2008; Zhou et al. 2010; Eldem et al. 2012; Guan et al. 2013). miR398 is a well-characterized miRNA that plays a crucial role in plant adaption to thermal stress by controlling the expression of Copper/Zinc Superoxide dismutases (Cu/Zn-SODs) (Sunkar et al. 2006). Notably, two isoforms of miR398 were 5-fold up-regulated in the 4h heat stress treatment, and then 5-fold down-regulated in the 24 h treatment.

Only 51 ohpRNAs were differentially expressed during heat stress, including four miRNA-like ohpRNAs (**Figure 12**). Of these, the miR167-like ohpRNAs were most highly expressed with about 219 to 380 RPM in different libraries, and others were mostly expressed less than 50 rpm and had less than 2 fold changes.

In contrast to miRNAs or ohpRNAs, a larger number of ra-siRNAs and *cis*-nat siRNAs were differentially expressed during heat stress. The expression level of these siRNAs with more than 50 RPM in at least a library was shown in **Figure 13 and 14**. A subset of these siRNAs showed a differential expression during heat stress. For example, ra-siR98018 (7,369 RPM in the 0h heat stress library) was the most abundant among all repeat-associated siRNAs, and this siRNA dramatically decreased 2-fold in the 4h library and then increased 2 folds in the 24 h library compared to the 4h library. Most of the ra-siRNAs tended to show similar expression patterns, i.e. a decreased during heat stress. Similar to ra-siRNAs, Most of the highly expressed *cis*-nat siRNAs were down-regulated in both 4 h and 24 h heat stressed libraries.

As shown in **Figure 15**, 28 pha-siRNAs (3.5% of the total) were differentially expressed during the heat stress treatments. Among these, 18 were 21 nt long pha-siRNAs, and others were 22 nt or 24 nt long. These pha-siRNAs can be categorized into four clades. Pha-siRNAs in Clade 1 were up-regulated in 4h but not in 24 h. Clade 2 contained pha-siRNAs were up-regulated in both 4 h and 24 h libraries. Pha-siRNAs in Clade 3 were down-regulated in the 4 h library and returned to the expression levels similar to the 0h (before heat), while clade 4 was up-regulated in both the 24 h libraries (Clade 4). A few pha-siRNAs showed a great variation in expression in response to heat stress. For example, pha-siRNA-21-290 was one of the most abundantly expressed ta-siRNAs and was upgraded approximately 12-fold in the 4 h library and down-regulated about 6-fold in the 24 h library.

Discussion

Discussion

In this study, a bioinformatic pipeline was developed and implemented to identify the various classes of sRNAs found in the *G. raimondii* genome. This pipeline was used to demonstrate the great diversity of sRNAs that were found in this genome (**Table 2**), and quantitatively estimate the number of each species of sRNA found. Many of these sRNAs were expressed at significant levels in one or more of our libraries, although a large number of sRNAs demonstrate very low expression levels in only one or a few libraries making it impossible to demonstrate that these can be meaningfully classified into specific functional sRNA types at this time. It is more likely that many of these sRNAs with low abundance result from random or nonrandom RNA degradation processes occurring inside cells or during the preparation of total RNA from tissues. However, low abundance reads were retained in our original total sRNA library since many sRNAs are produced from dsRNA precursors, but only 1 of the 2 strands has functional significance and is consequently more stable (e.g. is loaded into a RISC complex). However, the existence of the complementary strand in the sRNA dataset can assist in determination of the functional sRNA (e.g. the existence of miRNA* improves the likelihood of authentic miRNA prediction). This led to a significant number of specific sequences in the total sRNA libraries that remain unclassified at the end of the pipeline (**Table 2**) as evidenced by the relatively low abundance of these reads (average abundance below 1.6 reads per sequence). Nevertheless, a significant number of *G. raimondii* sRNA sequences have been discovered in these leaf libraries, many of which have abundances well above minimal, random levels. However, since we only used leaf tissues for our library preparation, the list of discovered sRNAs cannot be considered exhaustive of all sRNAs found in the *G. raimondii* genome. By analyzing six leaf small RNA

libraries from *G. raimondii* with/without heat stress treatments, we sought to identify the genome-wide spectrum of sRNAs present in *G. raimondii* and determine the role of various sRNA species in the regulation of gene expression during heat stress. The major species of sRNAs identified, included miRNAs, ohpRNAs, ra-siRNAs, and cis-nat siRNAs. Conserved and novel putative miRNA-directed ta-siRNAs regulatory network were identified, and expression analysis of these sRNAs showed that a subset of sRNAs was significantly regulated during heat stress.

Hairpins RNAs in *G. raimondii*

It is apparent that based on numbers miRNAs are a minor class of the entire hpRNA family. Unlike ohpRNAs, which are expressed at only a few reads per sequence, many miRNAs are highly expressed in the leaf small RNA libraries of *G. raimondii*. Thus, even though the number of miRNAs is smaller, their expression levels suggest that they may be more functionally important than ohpRNAs.

Among *G. raimondii* miRNAs identified in this study, only 32% were conserved across plants. Similarly, a high proportion of lineage-specific, non-conserved miRNAs are also observed in rice [58], *Medicago* [59], and *Glycine* [60]. In Arabidopsis, approximately 30% miRNA families are not shared between *A. thaliana* and *A. lyrate* [61]. Therefore, miRNAs are evolved and lost at a high frequency [61]. miR7495 is a newly identified miRNA family that is found only in *G. hirsutum*, a tetraploid hybrid derived from a diploid *G. raimondii*-like and a diploid G. *arboreum*-like progenitor [29]. A miR7495 locus was identified in the genome of *G. raimondii*, whereas nucleotide variations were found between this miRNA locus in *G. raimondii* and *G. hirsutum* consistent with evolution of *G. raimondii* miRNAs following the formation of the *G.*

hirsutum tetraploid,, and this change potentially influences the recognition of target mRNAs by their miRNAs (**Figure 3C**).

miRNA-directed pha-siRNA regulatory networks

Pha-siRNAs, traditionally referred to as phased siRNAs, have been identified in a wide range of plants, from mosses to eudicots [63-65], including *G. hirsutum* [66,67], but pha-siRNAs have not previously been reported in *G. raimondii*. In this study, we made a global investigation of pha-siRNA loci in the genome of *G. raimondii*, which enabled us to identified 294 pha-siRNA genes (PSRGs) that generated 21 nt, 22 nt or 24 nt ta-siRNAs.

Among the 21 nt generating PSRGs, 28% corresponded to previously reported PSRG/ TAS genes, such as Arabidopsis TAS3 (Auxin Response Factor, ARF), NB-LRR disease resistance genes and MYB transcription factors . These PSRG/TAS pathways are conserved across plants and play crucial roles in plant development and stress responses [34,63,68].

Approximately 80% of the 24 nt-generating pha-siRNA genes reported here are non-protein coding, which is consistent with previous studies [39,69]. Previous studies have shown that 24 nt pha-siRNAs are produced from the DCL3/5-assocatied pathway (previously DCL-3 now considered as DCL5) [48], and these pha-siRNAs specifically accumulate in reproductive tissues (Johnson et al. 2009; Song et al. 2012) and are associated with transcriptional regulation of repeated sequence elements (transposons) [69]. Our results demonstrate that 24 nt pha-siRNAs are also abundantly made in leaf tissues, and likely play a role in the regulation of transposable elements in leaf tissues as they do in reproductive tissues in other systems.

The *G. raimondii* genome encodes six DCL proteins (two more than Arabidopsis) (Henderson et al. 2006) and appears to be a eudicots with two DCL3 homologs [29]. Although the functions of *G. raimondii* DCL proteins are not fully characterized, they generate a complicated population of

G. raimondii pha-siRNAs that likely have play important roles in sRNA regulation of posttranscriptional silencing in this species.

Only a small fraction of PSRGs appeared to be cleaved by predicted *G. raimondii* miRNAs. This may be due to two major reasons: 1) the miRNAs are not fully annotated in the genome of *G. raimondii* genome, and 2) the mechanism of biogenesis of pha-si/RNAs may be not fully understood. In the other words, other types of sRNAs might be capable of functionally acting to cleave the PSRG transcripts [70].

Ra-siRNAs and cis-nat siRNAs are major species of *G. raimondii* sRNA population

ra-siRNAs and cis-nat-siRNAs, the least represent two least understood classes of sRNA classes, whereas theyse siRNAs are shownappear to be associated with a variety of physiological processes, including abiotic and biotic stress responses [11,65,71]. In the our *G. raimondii* leaf sRNA libraries, ra-siRNAs and cis-nat-siRNAs accounted for approximately 51% and 16% of the total sRNA reads, respectively (**Table 1**). This is consistent with the previous study in Arabidopsis in which most of sRNAs are derived from repeat-associated loci [50] and also might reflect their significance for *G. raimondii*.

Repeat-derived associated ra-siRNAs (or also known as heterochromaticc-siRNAs) largely act to maintain epigenetic situationhave been shown to regulate by inducing or maintaining DNA methylation, heterochromatin formation, and chromatin remodeling associated with epigenetic inheritance controling many aspects of plant growth, development, and stress responses [72,73]. Although ra-siRNA loci in the genome of *G. raimondii* are poorly characterized, the recent studies show that DNA methylation is associated with cotton fiber development [72].

Cis-nat siRNAs are associated with plant organ development and stress [11,74] and have been identified in several plant species [10,65], though their existence has been questioned [48].

Using previously described methods [10,49], we identified a large number of cis-NAT gene pairs. However, 94% of these *cis*-NAT pairs did not produce cis-nat-siRNAs at a significant level in our leaf libraries, but a small subset of the cis-nat-siRNAs were highly expressed in the *G. raimondii* sRNA libraries. This was previously observed in Arabidopsis [75] and thus we can conclude that *cis*-NAT expression may be tissue-specific and only a fraction of the possible *cis*-nat-siRNAs is produced in leaf.

In other plant species *cis*-NAT pairs can be produced from protein-coding genes and environmental cues may be crucial for the formation of NAT pairs [75]. Surprisingly, sRNAs sequenced from natural antisense small RNA biogenesis mutants including *dcl1*, *dcl2*, *dcl3*, and *rdr6* map to *cis*-NATs as frequently as small RNAs sequenced from wild-type plants [75]. Nonetheless, DCL1 and/or DCL3 are required for the biogenesis of cis-nat siRNAs [10]. This might reflect that different types of NATs may be processed through different DCL-associated pathways.

Regulatory roles of small RNAs in *G. raimondii* heat stress response

Plant heat stress responses haves been well-documented in various levels, from biochemical, protein to gene levels, whereas recent studies show the importance of sRNAs in the regulation of plant thermal tolerance [24,25]. Although damaging at high concentrations, ROS (reactive oxygen species) appears to be a crucial signaling molecules that transduce stress signal to downstream effectors [76] and rapidly increase in the early stages of heat stress [77,78]. ,In *G. raimondii* miR398 increased 5-fold, the highest level of all heat-responsive miRNAs, and then was down-regulated by 24 h of heat stress. This result confirms results with Arabidopsis that implicate miR398 the regulation of plant oxidative stress by targeting Cu/Zn superoxide dismutase [24,25]. In addition to miR398, other conserved miRNAs, including miR166, miR396,

miR482, and novel miRNAs, such as miRC114 and miRC136 were differentially expressed during heat stress. This result is consistent with a role of conserved miRNAs, such as miR166, miR396 and miR482 in various stresses [57,79,80] and further shows that other lineage-specific miRNAs are coordinately regulated by heat stress in *G. raimondii*, consistent with a diversity of miRNAs being involved in the *G. raimondii* heat stress response.

In addition to miRNAs, several ohpRNAs were shown to be regulated during heat stress. Previous studies have shown that DCL2, DCl3 and DCL4 are required for the biogenesis ohpRNAs and produced small RNA populations with sizes from 21 nt to 24 small RNAs [13,81]. These ohpRNAs can be functionally loaded into AGO proteins and perform post-transcriptional gene silencing, but these genes show more species- or tissue-specific expression [13].

A subset of pha-siRNAs was also responsive to heat stress, including pha-siRNAs that were processed from MYB transcription factors. Although the biological function of many pha-siRNAs remains largely unknown, previous studies and our data showed that PSR/TAS genes, such as auxin response factor, NB-LRR and MYB TF, are functionally important for plant growth, development, and stress responses (Zhai et al. 2011; Cho et al. 2012).

Environmental stress influences genetic and epigenetic regulation in eukaryotes [82,83], and repeat-derived siRNAs (or heterochromatic siRNAs) play crucial roles in this epigenetic regulation [84,85]. In insects heat stress induces the rapid increase of ra-siRNAs as well as epigenetic modification of histones accompanied by changes in heterochromatin during the heat stress response [71]. Our data showed that repeat-derived siRNAs displayed different expression patterns. For example, ra-siRNAs-86281 was up-regulated 16-fold in 4 h library and then down-regulated 2-fold (compared to unstressed) in 24 h. However, most of ra-siRNAs tended to be down-regulated during heat stress. It might suggest the following: 1) different ra-siRNAs might

be associated with different heterochromatins that regulates different genes, and thus displayed different expression patterns; 2) *G. raimondii* heat response might be divided into different phases in which different genes are involved to perform different functions. In the future work, we will continue to work on elucidating the potential mechanism of ra-siRNA-directed heterochromatins and their roles in *G. raimondii* response during heat stress.

In Arabidopsis and rice, *cis*-nat siRNAs are differentially expressed in various tissues and organs and during stresses, such as salt and pathogen infection [9,74,86], and appear to play critical roles in histone acetylation and methylation [75]. In this study, we developed a bioinformatic pipeline that permits us to find potential *cis*-NAT gene pairs derived the same genomic loci. Our data showed that cis-nat siRNAs are the second largest class of *G. raimondii* sRNA population, and many were highly expressed in the leaf and differentially expressed during heat stress. This suggested that the expression of cis-nat siRNAs is associated with heat stress. However, future works will be needed to discover the mechanism of cis-nat siRNAs in response to heat stress.

When the sRNAs found in our libraries are mapped to specific chromosomal locations, many map to distinct proximal locations on the 13 *G. raimondii* chromosomes, creating genomic "hot spots" from which possibly functionally related sRNAs are produced (Qi et al. 2009). This phenomenon, and possible roles of chromosomal clustering of sRNAs in stress and developmental sRNA regulation are currently under further investigation since sRNAs often have multiple targets, or define regions of interest relative to heterochromatin formation (Smith et al. 2008) suggesting that specific chromosomal locations where "sRNA hot spots" exist produce signals that have the capacity to control numerous other genes throughout the genome.

This preliminary bioinformatic in this study, an investigation of the nature of and possible roles for small regulatory RNAs of *G. raimondii* was performed by analyzing six small RNA libraries

constructed from *G. raimondii* seedlings treated with heat stress. Our findings showed that a diversity of sRNAs is encoded in the genome of *G. raimondii*, and we also found that ta-siRNAs are organically connected to miRNAs to form sRNA regulatory networks, and that these networks may be further chromosomally related. A subset of these sRNAs were differentially expressed during heat stress, suggesting that sRNAs are playing a role in mediating the *G. raimondii* leaf heat stress response consistent with the emerging role of sRNAs in mediating stress responses in other plant species.

Reference

Allen E, Howell MD. 2010. miRNAs in the biogenesis of trans-acting siRNAs in higher plants. *S. Cell Dev. Bio.* **21**(8): 798-804.

Allen E, Xie Z, Gustafson AM, Carrington JC. 2005. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell.* **121**(2): 207-221.

Axtell MJ, Jan C, Rajagopalan R, Bartel DP. 2006. A two-hit trigger for siRNA biogenesis in plants. *Cell.* **127**(3): 565-577.

Axtell MJ, Westholm JO, Lai EC. 2011. Vive la difference: biogenesis and evolution of microRNAs in plants and animals. *Genome Bio.* **12**(4): 221.

Axtell MJ. 2013. Classification and Comparison of Small RNAs from Plants. *Ann Rew Plant Biol.* **64**: 137-159.

Baggerly KA, Deng L, Morris JS, Aldaz CM. 2003. Differential expression in SAGE: accounting for normal between-library variation. *Bioinformatics.* **19**(12): 1477-1483.

Banti V, Mafessoni F, Loreti E, Alpi A, Perata P. 2010. The heat-inducible transcription factor HsfA2 enhances anoxia tolerance in Arabidopsis. *Plant Physiol.* **152**(3): 1471-1483.

Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, Madden TL, Matten WT, McGinnis SD, Merezhuk Y et al. 2013. BLAST: a more efficient report with usability improvements. *Nucleic Acids Res*. 1-5.

Borsani O, Zhu J, Verslues PE, Sunkar R, Zhu JK. 2005. Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell.* **123**(7): 1279-1291.

Chapman EJ, Carrington JC. 2007. Specialization and evolution of endogenous small RNA pathways. *Nature Rev Genetics.* **8**(11): 884-896.

Chen HM, Li YH, Wu SH. 2007. Bioinformatic prediction and experimental validation of a microRNA-directed tandem trans-acting siRNA cascade in Arabidopsis. *PNAS.* **104**(9): 3318-3323.

Chen L, Ren Y, Zhang Y, Xu J, Sun F, Zhang Z, Wang Y. 2012. Genome-wide identification and expression analysis of heat-responsive and novel microRNAs in *Populus tomentosa*. *Gene.* **504**(2): 160-165.

Chen X. 2009. Small RNAs and their roles in plant development. *Ann Rev Cell Dev. Bio.* **25**: 21-44.

Cho SH, Coruh C, Axtell MJ. 2012. miR156 and miR390 regulate tasiRNA accumulation and developmental timing in Physcomitrella patens. *Plant Cell.* **24**(12): 4837-4849.

Conesa A, Gotz S. 2008. Blast2GO: A comprehensive suite for functional analysis in plant genomics. *J Plant Genomics.* **2008**: 619832.

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

Dai X, Zhao PX. 2008. pssRNAMiner: a plant short small RNA regulatory cascade analysis server. *Nucleic Acids Res.* **36**(Web Server issue): W114-118.

Dai X, Zhao PX. 2011. psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res.* **39**(Web Server issue): W155-159.

Dunoyer P, Brosnan CA, Schott G, Wang Y, Jay F, Alioua A, Himber C, Voinnet O. 2010. An endogenous, systemic RNAi pathway in plants. *EMBO J.* **29**(10): 1699-1712.

Eldem V, Celikkol Akcay U, Ozhuner E, Bakir Y, Uranbey S, Unver T. 2012. Genome-wide identification of miRNAs responsive to drought in peach (*Prunus persica*) by high-throughput deep sequencing. *PloS ONE.* **7**(12): e50298.

Fagegaltier D, Bouge AL, Berry B, Poisot E, Sismeiro O, Coppee JY, Theodore L, Voinnet O, Antoniewski C. 2009. The endogenous siRNA pathway is involved in heterochromatin formation in Drosophila. *PNAS.* **106**(50): 21258-21263.

Fei Q, Xia R, Meyers BC. 2013. Phased, Secondary, Small Interfering RNAs in Posttranscriptional Regulatory Networks. *Plant Cell*. 25(7):2400-15

Feliciello I, Parazajder J, Akrap I, Ugarkovic D. 2013. First evidence of DNA methylation in insect Tribolium castaneum: environmental regulation of DNA methylation within heterochromatin. *Epigenetics.* **8**(5): 534-541.

Fusaro AF, Matthew L, Smith NA, Curtin SJ, Dedic-Hagan J, Ellacott GA, Watson JM, Wang MB, Brosnan C, Carroll BJ et al. 2006. RNA interference-inducing hairpin RNAs in plants act through the viral defence pathway. *EMBO Rep.* **7**(11): 1168-1175.

Guan Q, Lu X, Zeng H, Zhang Y, Zhu J. 2013. Heat stress induction of miR398 triggers a regulatory loop that is critical for thermotolerance in Arabidopsis. *Plant J*. 74(5):840-5

Hasanuzzaman M, Nahar K, Alam MM, Roychowdhury R, Fujita M. 2013. Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. *J Mol Sci.* **14**(5): 9643-9684.

Henderson IR, Zhang X, Lu C, Johnson L, Meyers BC, Green PJ, Jacobsen SE. 2006. Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nat Genetics.* **38**(6): 721-725.

Huang Y, Zhang JL, Yu XL, Xu TS, Wang ZB, Cheng XC. 2013. Molecular functions of small regulatory noncoding RNA. *Biochem Bio.* **78**(3): 221-230.

Jamalkandi SA, Masoudi-Nejad A. 2009. Reconstruction of Arabidopsis thaliana fully integrated small RNA pathway. *Funct Int Genomics.* **9**(4): 419-432.

Jeong DH, Park S, Zhai J, Gurazada SG, De Paoli E, Meyers BC, Green PJ. 2011. Massive analysis of rice small RNAs: mechanistic implications of regulated microRNAs and variants for differential target RNA cleavage. *Plant cell.* **23**(12): 4185-4207.

Jin X, Pang Y, Jia F, Xiao G, Li Q, Zhu Y. 2013. A potential role for CHH DNA methylation in cotton fiber growth patterns. *PloS ONE.* **8**(4): e60547.

Johnson C, Kasprzewska A, Tennessen K, Fernandes J, Nan GL, Walbot V, Sundaresan V, Vance V, Bowman LH. 2009. Clusters and superclusters of phased small RNAs in the developing inflorescence of rice. *Genome Res.* **19**(8): 1429-1440.

Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. 2005. Repbase Update, a database of eukaryotic repetitive elements. *Cyto Genome Res.* **110**(1-4): 462-467.

Kasschau KD, Fahlgren N, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Carrington JC. 2007. Genome-wide profiling and analysis of Arabidopsis siRNAs. *PLoS Biol.* **5**(3): e57.

Katiyar-Agarwal S, Morgan R, Dahlbeck D, Borsani O, Villegas A, Jr., Zhu JK, Staskawicz BJ, Jin H. 2006. A pathogen-inducible endogenous siRNA in plant immunity. *PNAS.* **103**(47): 18002-18007.

Khraiwesh B, Zhu JK, Zhu J. 2012. Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biol Biophysic Acta.* **1819**(2): 137-148.

Konigshofer H, Tromballa HW, Loppert HG. 2008. Early events in signalling high-temperature stress in tobacco BY2 cells involve alterations in membrane fluidity and enhanced hydrogen peroxide production. *Plant Cell Envir.* **31**(12): 1771-1780.

Langmead B. 2010. Aligning short sequencing reads with Bowtie. *Cur Prot Bioinformatics.* **Chapter 11**: Unit 11 17.

Lee TF, Gurazada SG, Zhai J, Li S, Simon SA, Matzke MA, Chen X, Meyers BC. 2012. RNA polymerase V-dependent small RNAs in Arabidopsis originate from small, intergenic loci including most SINE repeats. *Epigenetics.* **7**(7): 781-795.

Lelandais-Briere C, Naya L, Sallet E, Calenge F, Frugier F, Hartmann C, Gouzy J, Crespi M. 2009. Genome-wide Medicago truncatula small RNA analysis revealed novel microRNAs and isoforms differentially regulated in roots and nodules. *Plant Cell.* 21(9): 2780-2796.

Lorenz R, Bernhart SH, Honer Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA Package 2.0. *AMB.* **6**: 26.

Lu X, Guan Q, Zhu J. 2013. Downregulation of CSD2 by a heat-inducible miR398 is required for thermotolerance in Arabidopsis. *Plant Sig Beh.* **8**(8).

Marin E, Jouannet V, Herz A, Lokerse AS, Weijers D, Vaucheret H, Nussaume L, Crespi MD, Maizel A. 2010. miR390, Arabidopsis TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. *Plant Cell.* **22**(4): 1104-1117.

Matzke M, Kanno T, Daxinger L, Huettel B, Matzke AJ. 2009. RNA-mediated chromatin-based silencing in plants. *Cur Opin Ccell Bio.* **21**(3): 367-376.

Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, Bowman JL, Cao X, Carrington JC, Chen X, Green PJ et al. 2008. Criteria for annotation of plant MicroRNAs. *Plant Cell* .**20**(12): 3186-3190.

Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C et al. 2008. Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. *Cell.* **133**(1): 116-127.

Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, Carrington JC. 2008. Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. *Cell.* **133**(1): 128-141.

Moxon S, Schwach F, MacLean D, Dalmay T, Studholme DJ, Moulton V (2008) A toolkit for analysing large-scale plant small RNA datasets. *Bioinformatic*s. 24:2252-2253.

Naqvi AR, Sarwat M, Hasan S, Roychodhury N. 2012. Biogenesis, functions and fate of plant microRNAs. *J Cell Physiol.* **227**(9): 3163-3168.

Neill SJ, Desikan R, Clarke A, Hurst RD, Hancock JT. 2002. Hydrogen peroxide and nitric oxide as signalling molecules in plants. *J Eexp Bot.* **53**(372): 1237-1247.

Ning Z, Cox AJ, Mullikin JC. 2001. SSAHA: a fast search method for large DNA databases. *Genome Res.* **11**(10): 1725-1729.

Ouyang S, Buell CR. 2004. The TIGR Plant Repeat Databases: a collective resource for the identification of repetitive sequences in plants. *Nucleic Acids Res.* **32**(Database issue): D360-363.

Pang M, Woodward AW, Agarwal V, Guan X, Ha M, Ramachandran V, Chen X, Triplett BA, Stelly DM, Chen ZJ. 2009. Genome-wide analysis reveals rapid and dynamic changes in miRNA and siRNA sequence and expression during ovule and fiber development in allotetraploid cotton (Gossypium hirsutum L.). *Genome Biol.* **10**(11): R122.

Park MY, Wu G, Gonzalez-Sulser A, Vaucheret H, Poethig RS. 2005. Nuclear processing and export of microRNAs in Arabidopsis. *PNAS.* **102**(10): 3691-3696.

Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, Llewellyn D, Showmaker KC, Shu S, Udall J et al. 2012. Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres. *Nature.* **492**(7429): 423-427.

Pezer Z, Ugarkovic D. 2012. Satellite DNA-associated siRNAs as mediators of heat shock response in insects. *RNA Biol.* **9**(5): 587-595.

Qi X, Bao FS, Xie Z. 2009. Small RNA deep sequencing reveals role for Arabidopsis thaliana RNA-dependent RNA polymerases in viral siRNA biogenesis. *PloS ONE.* 4(3): e4971.

Qu AL, Ding YF, Jiang Q, Zhu C. 2013. Molecular mechanisms of the plant heat stress response. *Biochem Biophysic Res Comm.* **432**(2): 203-207.

Quintero A, Perez-Quintero AL, Lopez C. 2013. Identification of ta-siRNAs and Cis-nat-siRNAs in Cassava and Their Roles in Response to Cassava Bacterial Blight. *Genomics Prot Bioinformatics.* **11**(3): 172-181.

Rogers K, Chen X. 2013. microRNA Biogenesis and Turnover in Plants. *Cold Spring Harbor Symp Quant Bio*. 77:183-94.

Ron M, Alandete Saez M, Eshed Williams L, Fletcher JC, McCormick S. 2010. Proper regulation of a sperm-specific cis-nat-siRNA is essential for double fertilization in Arabidopsis. *Gene Dev.* **24**(10): 1010-1021.

Sato K, Hamada M, Asai K, Mituyama T. 2009. CENTROIDFOLD: a web server for RNA secondary structure prediction. *Nucleic Acids Res.* **37**(Web Server issue): W277-280.

Saze H, Tsugane K, Kanno T, Nishimura T. 2012. DNA methylation in plants: relationship to small RNAs and histone modifications, and functions in transposon inactivation. *Plant Cell Physiol.* **53**(5): 766-784.

Smith DR, Quinlan AR, Peckham HE, Makowsky K, Tao W, Woolf B, Shen L, Donahue WF, Tusneem N, Stromberg MP et al. 2008. Rapid whole-genome mutational profiling using next-generation sequencing technologies. *Genome Res.* 18(10): 1638-1642.

Song X, Li P, Zhai J, Zhou M, Ma L, Liu B, Jeong DH, Nakano M, Cao S, Liu C et al. 2012. Roles of DCL4 and DCL3b in rice phased small RNA biogenesis. *Plant J.* **69**(3): 462-474.

Stocks MB, Moxon S, Mapleson D, Woolfenden HC, Mohorianu I, Folkes L, Schwach F, Dalmay T, Moulton V. 2012. The UEA sRNA workbench: a suite of tools for analysing and visualizing next generation sequencing microRNA and small RNA datasets. *Bioinformatics.* **28**(15): 2059-2061.

Subramanian S, Fu Y, Sunkar R, Barbazuk WB, Zhu JK, Yu O. 2008. Novel and nodulation-regulated microRNAs in soybean roots. *BMC Genomics.* **9**: 160.

Sunkar R, Kapoor A, Zhu JK. 2006. Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in Arabidopsis is mediated by downregulation of miR398 and important for oxidative stress tolerance. *Plant Cel.l* **18**(8): 2051-2065.

Tempel S. 2012. Using and understanding RepeatMasker. *Methods Mol Biol* **859**: 29-51.

Thakur V, Wanchana S, Xu M, Bruskiewich R, Quick WP, Mosig A, Zhu XG. 2011. Characterization of statistical features for plant microRNA prediction. *BMC Genomics.* **12**: 108.

Tittel-Elmer M, Bucher E, Broger L, Mathieu O, Paszkowski J, Vaillant I. 2010. Stress-induced activation of heterochromatic transcription. *PLoS Genet.* **6**(10): e1001175.

Vazquez F, Legrand S, Windels D. 2010. The biosynthetic pathways and biological scopes of plant small RNAs. *Trends Plant Sci.* **15**(6): 337-345.

Wahid A, Gelani S, Ashraf M, Foolad MR. 2007. Heat tolerance in plants: An overview. *Envir Exp Bot.* **61**(3): 199-223.

Wan CY, Wilkins TA. 1994. A modified hot borate method significantly enhances the yield of high-quality RNA from cotton (*Gossypium hirsutum* L.). *Ana Biochem.* **223**(1): 7-12.

Wang MB, Helliwell CA, Wu LM, Waterhouse PM, Peacock WJ, Dennis ES. 2008. Hairpin RNAs derived from RNA polymerase II and polymerase III promoter-directed transgenes are processed differently in plants. *RNA.* **14**(5): 903-913.

Wang ZM, Xue W, Dong CJ, Jin LG, Bian SM, Wang C, Wu XY, Liu JY. 2012. A comparative miRNAome analysis reveals seven fiber initiation-related and 36 novel miRNAs in developing cotton ovules. *Mol Plant.* **5**(4): 889-900.

Wu L, Zhou H, Zhang Q, Zhang J, Ni F, Liu C, Qi Y. 2010. DNA methylation mediated by a microRNA pathway. *Mol Cell.* **38**(3): 465-475.

Xia R, Zhu H, An YQ, Beers EP, Liu Z. 2012. Apple miRNAs and tasiRNAs with novel regulatory networks. *Genome Biol.* **13**(6): R47.

Xie Z, Qi X. 2008. Diverse small RNA-directed silencing pathways in plants. *Bioch Biophysic Acta.* **1779**(11): 720-724.

Xin M, Wang Y, Yao Y, Song N, Hu Z, Qin D, Xie C, Peng H, Ni Z, Sun Q. 2011. Identification and characterization of wheat long non-protein coding RNAs responsive to powdery mildew infection and heat stress by using microarray analysis and SBS sequencing. *BMC Plant Biol.* **11**: 61.

Xin M, Wang Y, Yao Y, Xie C, Peng H, Ni Z, Sun Q. 2010. Diverse set of microRNAs are responsive to powdery mildew infection and heat stress in wheat (*Triticum aestivum* L.). *BMC Plant Bio.* **10**: 123.

Xue W, Wang Z, Du M, Liu Y, Liu JY. 2013. Genome-wide analysis of small RNAs reveals eight fiber elongation-related and 257 novel microRNAs in elongating cotton fiber cells. *BMC Genomics.* **14**(1): 629.

Yang X, Wang L, Yuan D, Lindsey K, Zhang X. 2013. Small RNA and degradome sequencing reveal complex miRNA regulation during cotton somatic embryogenesis. *J Exp Bot.* **64**(6): 1521-1536.

Yang Z, Ebright YW, Yu B, Chen X. 2006. HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. *Nucleic Acids Res.* **34**(2): 667-675.

Yelina NE, Smith LM, Jones AM, Patel K, Kelly KA, Baulcombe DC. 2010. Putative Arabidopsis THO/TREX mRNA export complex is involved in transgene and endogenous siRNA biosynthesis. *PNAS.* **107**(31): 13948-13953.

Yu X, Wang H, Lu Y, de Ruiter M, Cariaso M, Prins M, van Tunen A, He Y. 2012. Identification of conserved and novel microRNAs that are responsive to heat stress in Brassica rapa. *J Exp Bot.* **63**(2): 1025-1038.

Zhai J, Jeong DH, De Paoli E, Park S, Rosen BD, Li Y, Gonzalez AJ, Yan Z, Kitto SL, Grusak MA et al. 2011. MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes Dev.* **25**(23): 2540-2553.

Zhan S, Lukens L. 2013. Protein-coding cis-natural antisense transcripts have high and broad expression in Arabidopsis. *Plant Physiol.* **161**(4): 2171-2180.

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

Zhang C, Li G, Wang J, Fang J. 2012a. Identification of trans-acting siRNAs and their regulatory cascades in grapevine. *Bioinformatics.* **28**(20): 2561-2568.

Zhang C, Li G, Wang J, Zhu S, Li H. 2013a. Cascading cis-cleavage on transcript from trans-acting siRNA-producing locus 3. *J Mol Sci.* **14**(7): 14689-14699.

Zhang J, Wu T, Li L, Han S, Li X, Zhang S, Qi L. 2013b. Dynamic expression of small RNA populations in larch (Larix leptolepis). *Planta.* **237**(1): 89-101.

Zhang X, Lii Y, Wu Z, Polishko A, Zhang H, Chinnusamy V, Lonardi S, Zhu JK, Liu R, Jin H. 2013c. Mechanisms of Small RNA Generation from Cis-NATs in Response to Environmental and Developmental Cues. *Mol Plant.* **6**(3): 704-715.

Zhang X, Xia J, Lii YE, Barrera-Figueroa BE, Zhou X, Gao S, Lu L, Niu D, Chen Z, Leung C et al. 2012b. Genome-wide analysis of plant nat-siRNAs reveals insights into their distribution, biogenesis and function. *Genome Biol.* **13**(3): R20.

Zhou L, Liu Y, Liu Z, Kong D, Duan M, Luo L. 2010. Genome-wide identification and analysis of drought-responsive microRNAs in Oryza sativa. *J Exp Bot.* **61**(15): 4157-4168.

Zhou X, Sunkar R, Jin H, Zhu JK, Zhang W. 2009. Genome-wide identification and analysis of small RNAs originated from natural antisense transcripts in Oryza sativa. *Genome Res.* **19**(1): 70-78.

Zhou X, Wang G, Sutoh K, Zhu JK, Zhang W. 2008. Identification of cold-inducible microRNAs in plants by transcriptome analysis. *Biochim Biophysic Acta.* **1779**(11): 780-788.

Zhu C, Ding Y, Liu H. 2011. MiR398 and plant stress responses. *Physiol Plant.* **143**(1): 1-9.

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

Table 1. Summary of small RNA library of *G. raimondii*

	Total reads	Unique reads
Raw sequencing	90,795,602	
Sequences after adaptor, low quality, & length trimming	46,738,621	8,057,647
Exactly matching the genome	39,534,260	5,609,535
ncRNA-matched	13,648,916	119,351
sRNA	25,885,344	5,490,184
	100.0%	100.0%
hpRNA	475,241	2,375
	1.8%	0.2%
miRNAs	448,133	146
	1.7%	0.0%
Other hpRNA	27,669	2,236
	0.1%	0.2%
siRNA	17,944,963	307,550
	69.3%	5.6%
Repeat-associated	12,645,751	159,701
	48.9%	13.2%
Phased siRNAs	52,060	791
	0.3%	0.2%
21 nt phasiRNA	36,488	441
	0.1%	0.0%
22 nt ta-siRNA	10,395	139
	0.0%	0.0%
24 nt ta-siRNA	5,177	211
	0.0%	0.0%
cis-nat-20-to-24nt	5,247,152	147,058
	20.3%	2.7%
Uncharacterized	7,465,140	5,180,259
	33.1%	92.4%

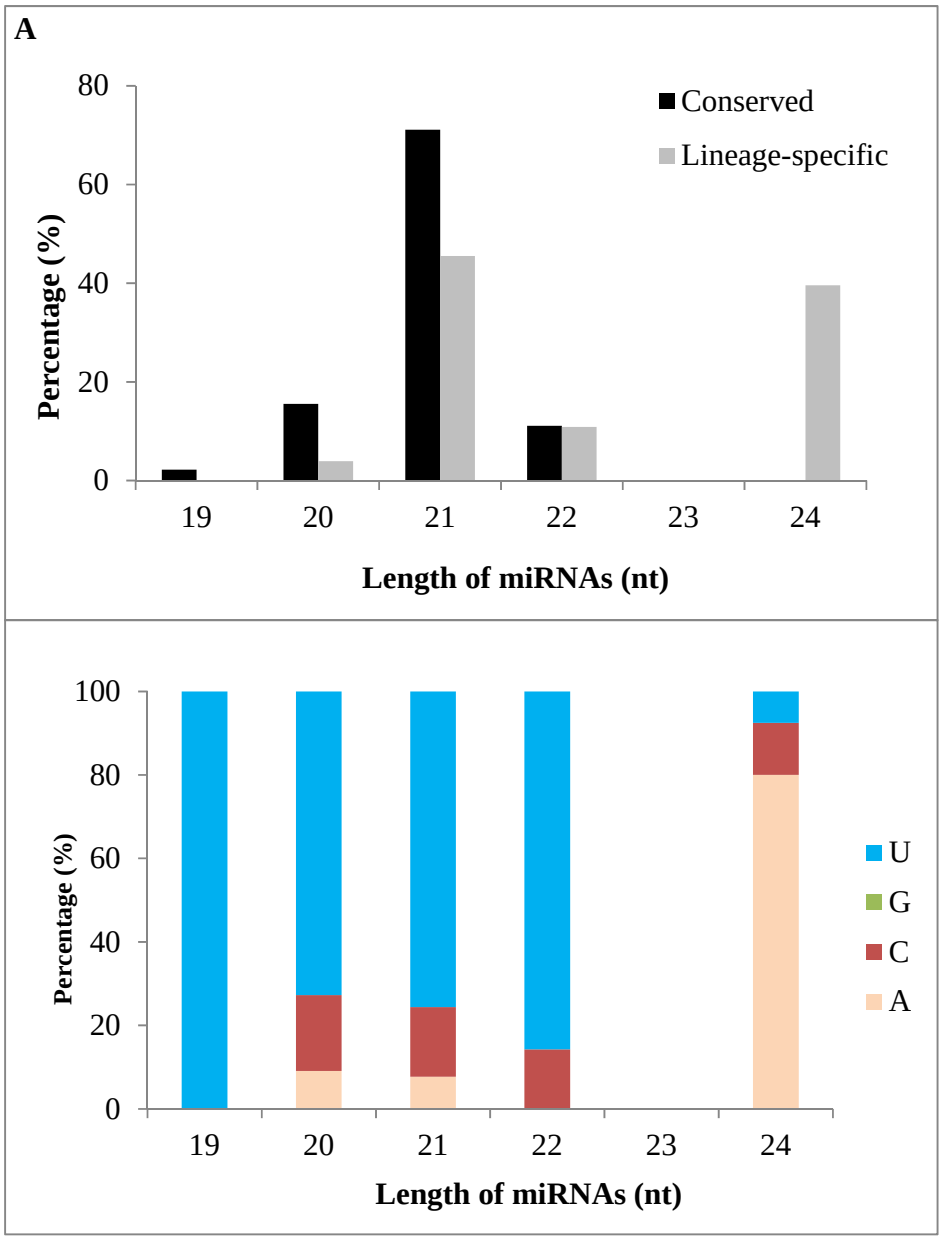


Figure 2. Length distribution and 5' end nucleotide of *G. raimondii* miRNAs

The length distribution of *G. raimondii* miRNAs is shown in A; *G. raimondii* miRNAs rage from 19 nt to 24 nt in length, but no 23 nt miRNA. The 5' nucleotide of miRNAs is shown in B; The 5' nucleotide of miRNAs are represented by different colors, U (blue), A (light orange), C (light red), and G (green)

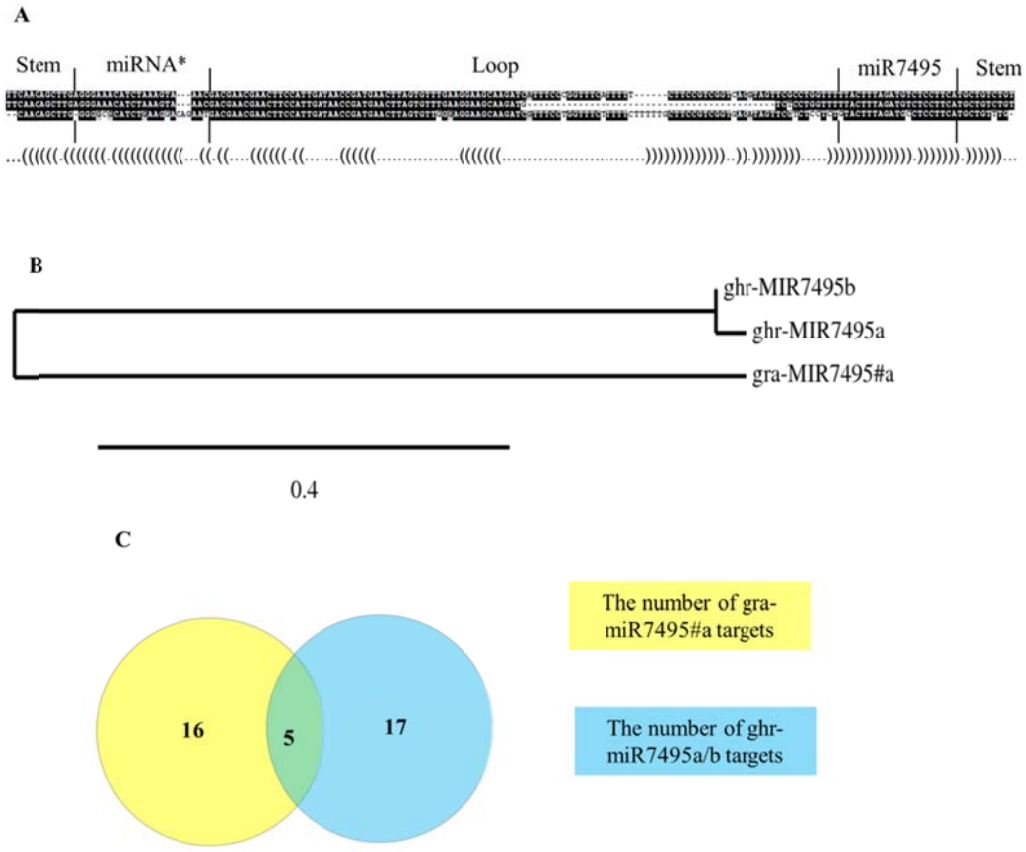


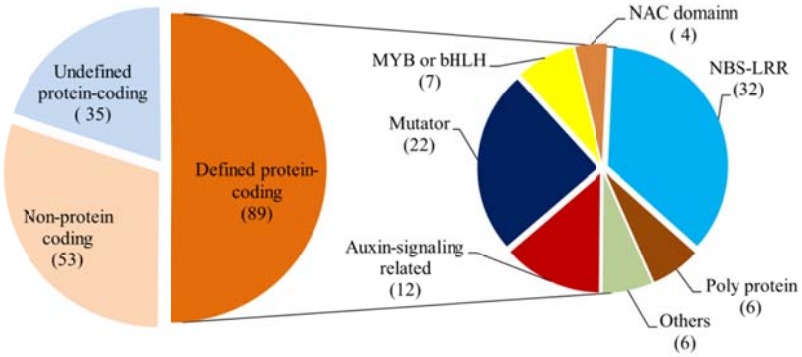
Figure 3. Sequence comparison of miRNA precursors between *G. raimondii* and *hirsutum*

The sequence alignment for two *G. hirsutum* miRNAs, ghr-miR7495a and ghr-miR7495b, a the *G. raimondii* gra-miR7495, is shown in A; the top and middle strand are ghr-MIR7495a a ghr-MIR7495b of *G. hirsutum*, and the bottom strand is gra-MIR7495a.

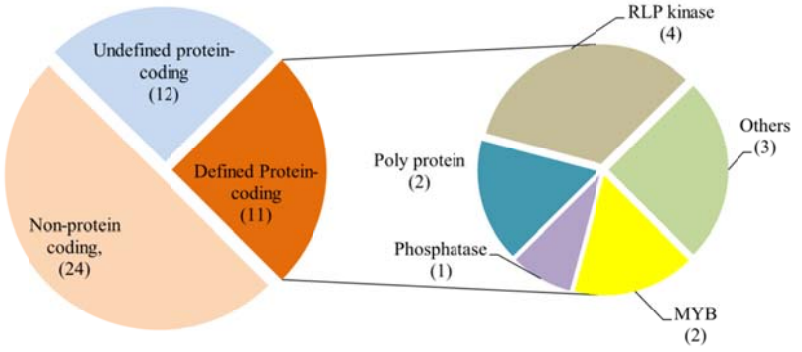
The phylogenetic analysis for these miRNA precursors is shown in B.

Venn diagram (C) shows that the numbers of common and different target genes that a regulated by ghr-mi749a/b and gra-miR7495

A


Undefined protein-coding (35)
Non-protein coding (53)
Defined protein-coding (89)
NAC domainn (4)
MYB or bHLH (7)
NBS-LRR (32)
Mutator (22)
Auxin-signaling related (12)
Poly protein (6)
Others (6)

B


Undefined protein-coding (12)
Non-protein coding, (24)
Defined Protein-coding (11)
RLP kinase (4)
Others (3)
Poly protein (2)
Phosphatase (1)
MYB (2)

C

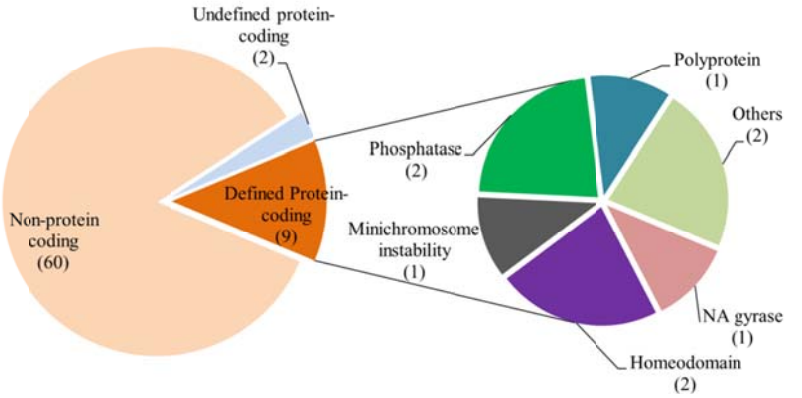

Undefined protein-coding (2)
Non-protein coding (60)
Defined Protein-coding (9)
Polyprotein (1)
Others (2)
Phosphatase (2)
Minichromosome instability (1)
NA gyrase (1)
Homeodomain (2)

Figure 4. Summary for phasiRNA-generating genes found in *G. raimondii*

The numbers of phasiRNA-generating genes for 21 nt, 22 nt, and 24 nt are shown in A, B, C, respectively. These genes are divided into three major groups, defined protein-coding genes (encoding known protein products), undefined protein-coding genes (encoding proteins that are not well characterized or hypothetical), and non-protein coding genes whose translated products have no hit in NCBI protein database.

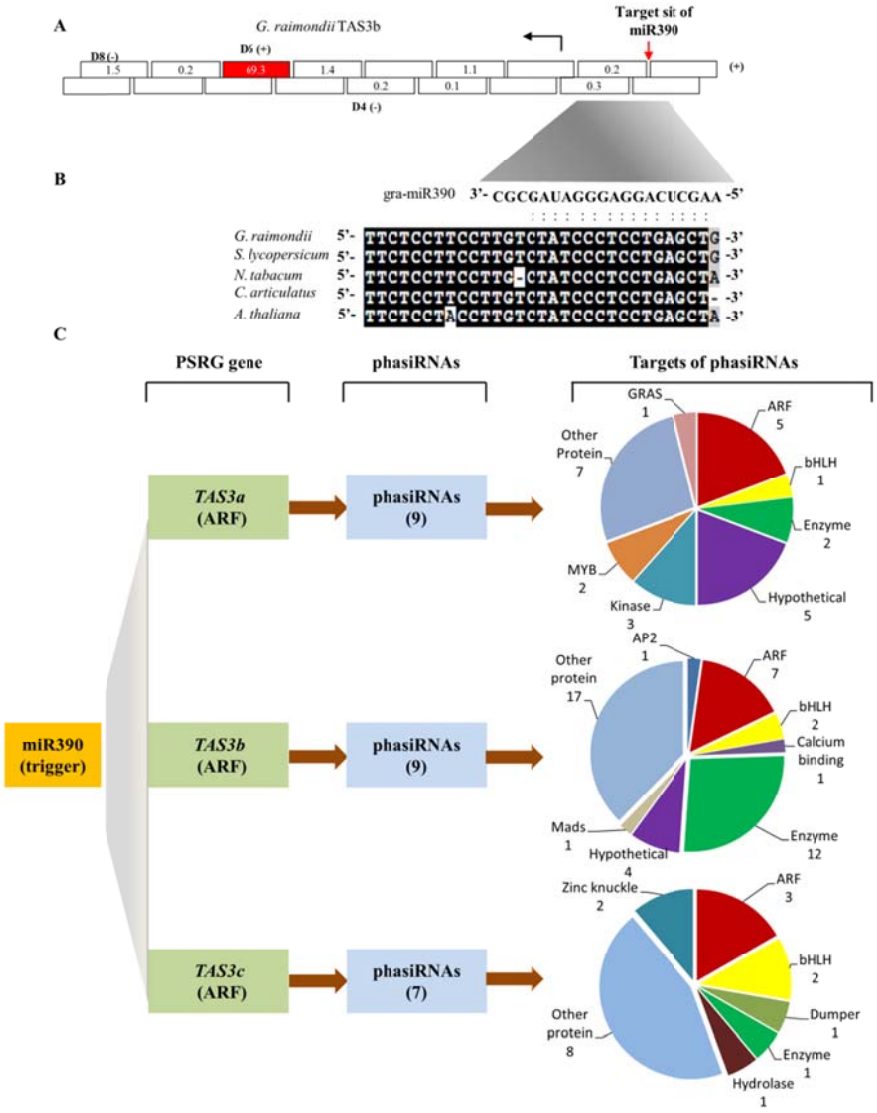


Figure 5. The proposed biogenesis of *G. raimondii TAS3*-like gene (PSRG) and potential regulatory pathway of *TAS*3-like-derived phasiRNAs

The proposed biogenesis of *G. raimondii TAS*3-like gene is shown in A; the target site of miR390 located at the 3' end, the red arrow shows the initial cleavage site of miR390. The black arrow shows the direction of cleavage direction. The D6(+), represented by the red box, is the most abundant expressed phasiRNAs, while others are detected 0.1 to 69 RPM.

Panel B shows the target site of miR390 of *G. raimondii TAS3*-like gene is conserved across five plant species.

The regulatory pathway of miR390-directed *TAS3*-like phasiRNA genes is shown in C; The pie plots on the right show the number of target genes of phasiRNAs derived *TAS3*-like transcripts.

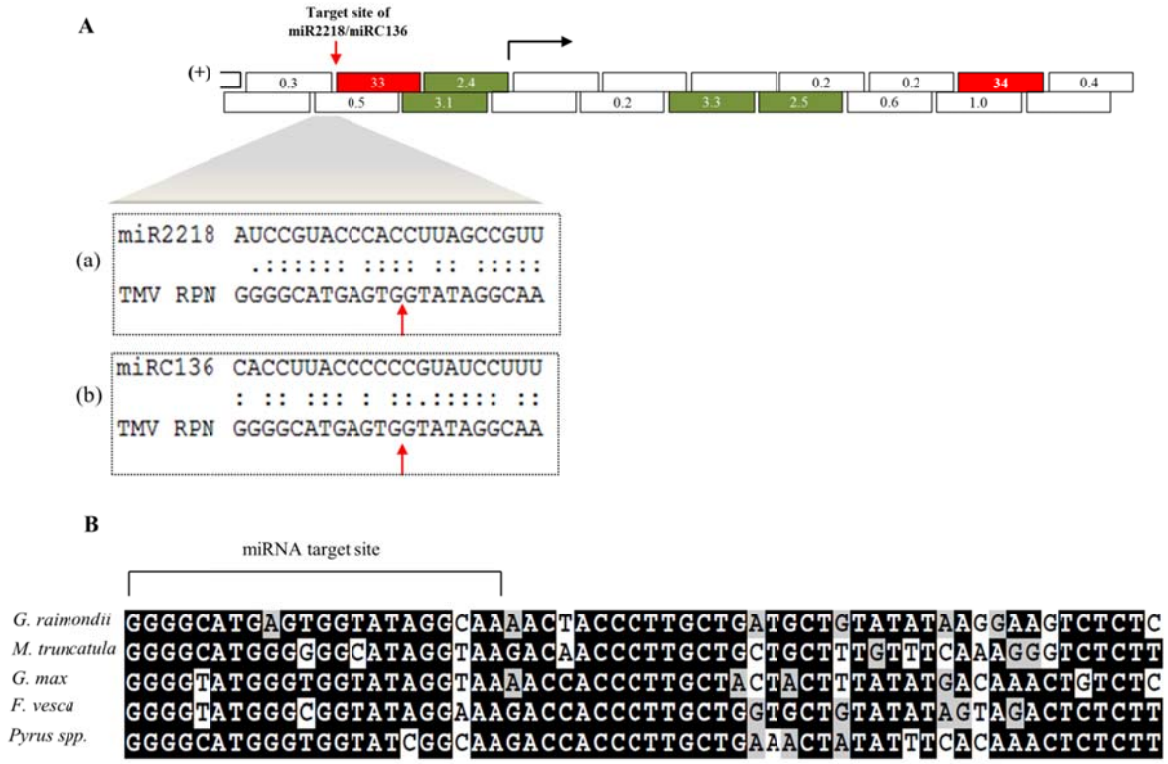


Figure 6. miR2218/miRC136 triggers the biogenesis of a novel phasiRNA-generating gene (TMV resistance protein N-LIKE)

The diagram for miR2218/miRC136-directed TMV RPN pathway is shown in A; the red arrow points the initial cleavage site of miRNAs and black arrow shows the direction of cleavage; the red box and dark green boxes represent highly and medially expressed ta-siRNAs, respectively; The numbers in the boxes represent the normalize expression value (RPM); the alignment of miR2218 and miRC316 are shown in a and b, respectively.

The alignment of TMV RPN genes in five distinct plants is shown in B; and the cleavage site of miRNAs is labeled by the black bar.

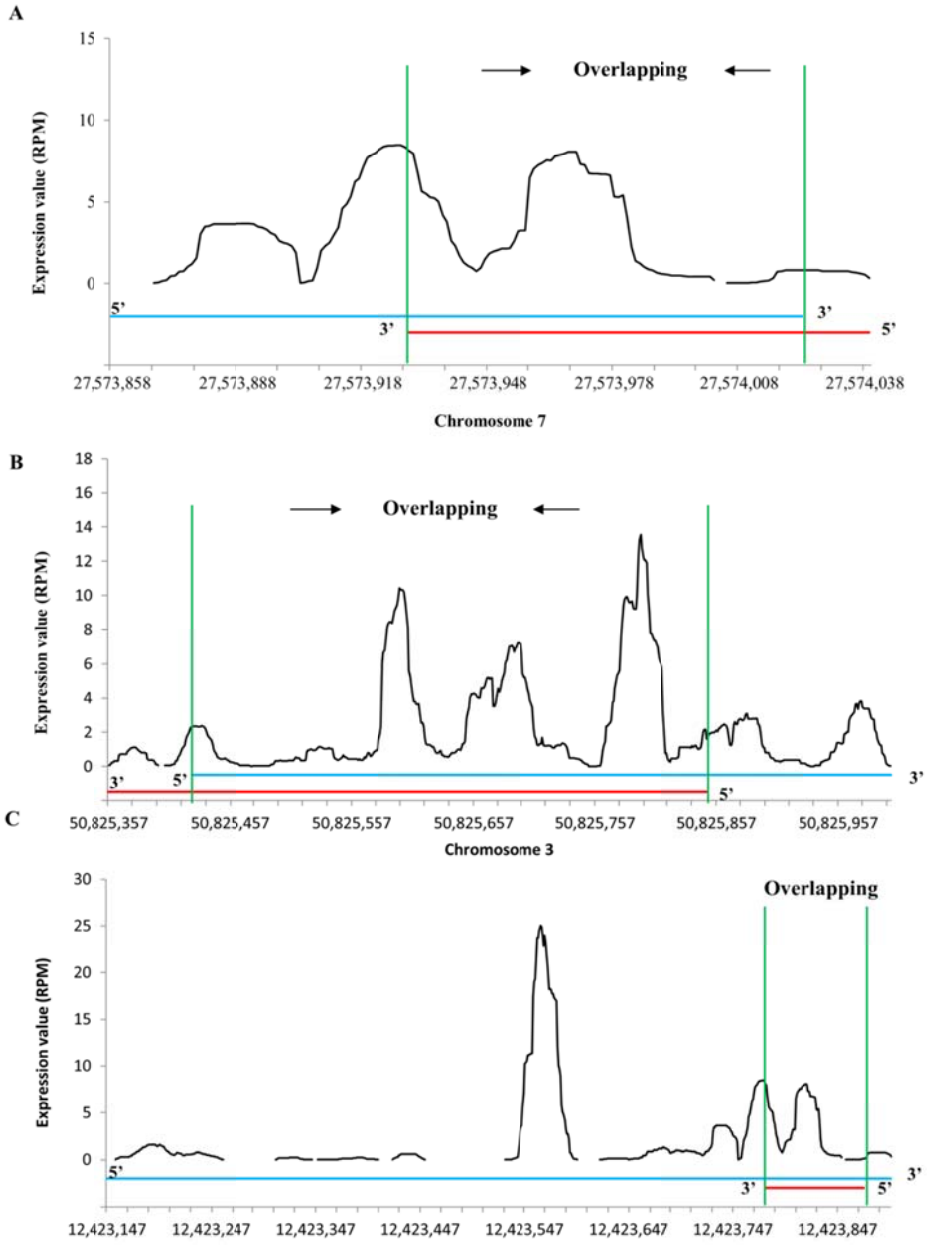


Figure 7. Three major types of cis natural antisense pairs in *G. raimondii*

Three major types of overlapping of cis-NATs are shown, in A (tail to tail), B (head to head), and C (enclosed). The blue and red line represent a pair of transcripts that transcribed from sense and antisense chromosomes, respectively. The regions between two green bars represent the overlapping between two transcripts. The normalize expression values (RPM) are shown in y-axis.

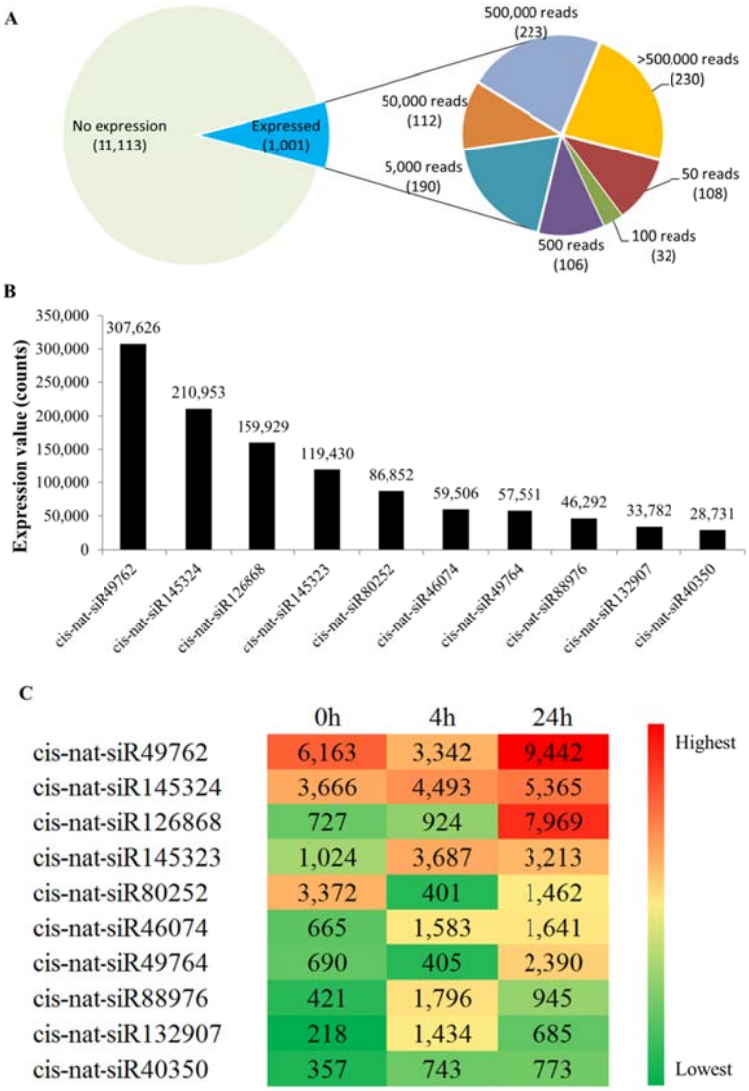


Figure 8. Expression of cis-nat siRNAs in the *G. raimondii* leaf

The number of cis-NATs with different expression value are shown in A; the number for each category represents the expression value (read counts) for the NATs . The top 10 of the most abundantly expressed cis-nat siRNAs are shown in B; the value of y-axis represents the detected counts of cis-nat siRNAs in the total sRNA library.

The expression change of the top 10 of the most abundantly expressed cis-nat siRNAs is shown in C; the colors represent the expression level of cis-nat siRNAs: red (higher expression) and green (lower expression).

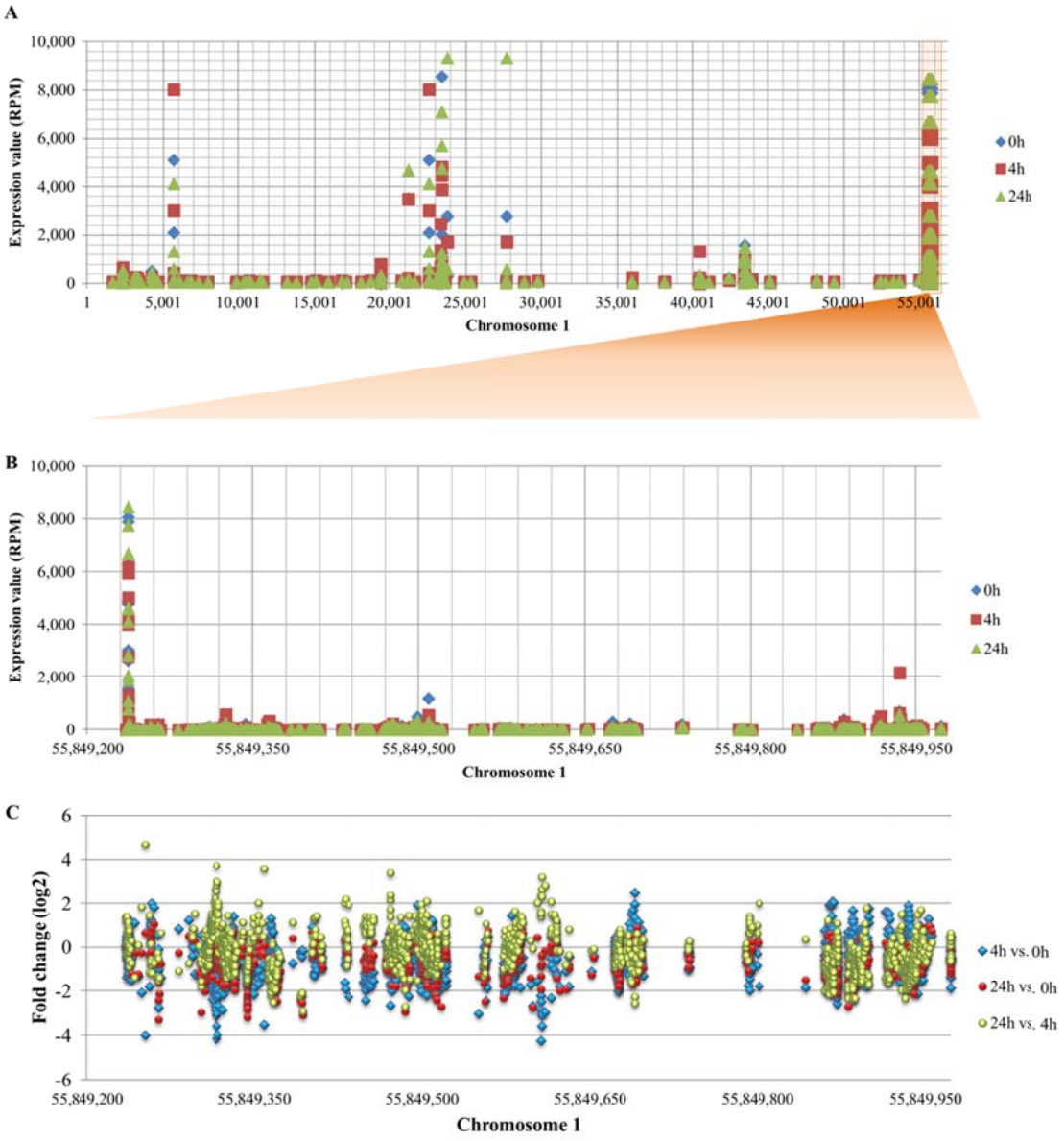


Figure 9. Distribution of sRNAs along the chromosome 1

The distribution of sRNAs (expressed at least 10 RPM) along chromosome 1 is shown in A, and the sRNAs derived from different libraries (treatments) are shown in blue (0h before heat), red (4h after heat) and green (24h after heat), repetitively.

One of hotspots located between 55,849,200 and 55,849,950 is shown in B.

The expression change of the sRNA in the hotspot is shown in C.

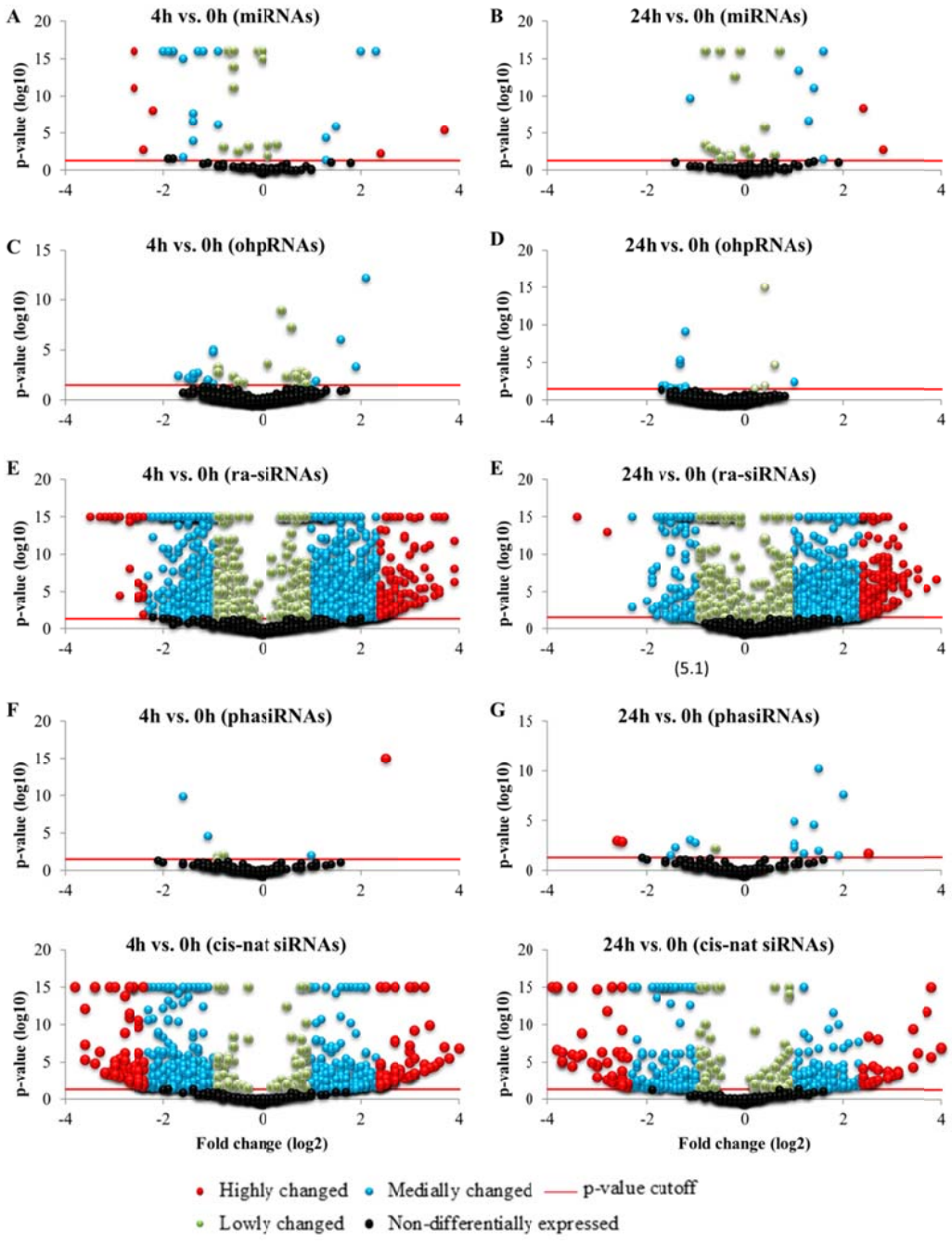


Figure 10. Expression of sRNAs during heat stress

Volcano plots show that the expression change of small RNAs during heat stress; A and B (miRNAs), C and D (ohpRNAs), E and F (ra-siRNAs), F and G (phasiRNAs), and H and I (cis-nat-siRNAs). The p-value (log10) and fold change (log2) are shown in x- and y-axis, respectively.

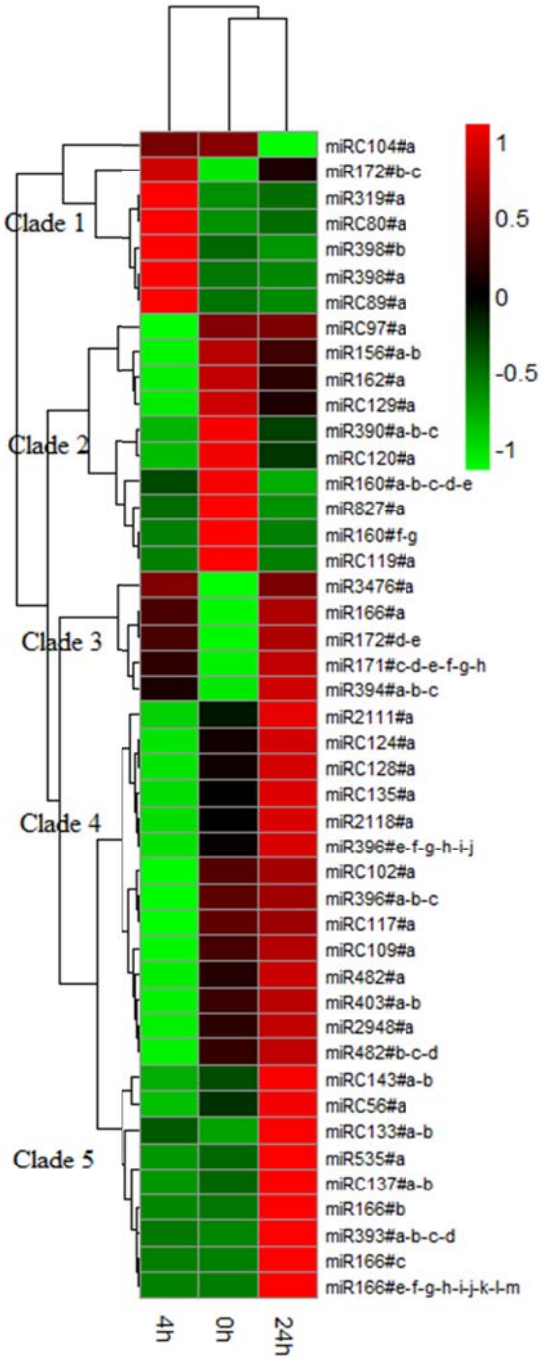


Figure 11. Expression of miRNAs during heat stress

Heatmap shows the expression change of miRNAs during heat stress. The miRNAs that were significantly regulated during heat stress are selected: 0h (0h before heat), 4h (4h after heat) and 24h (24h after heat). The colors represent the Z-score: higher expression (more red) and

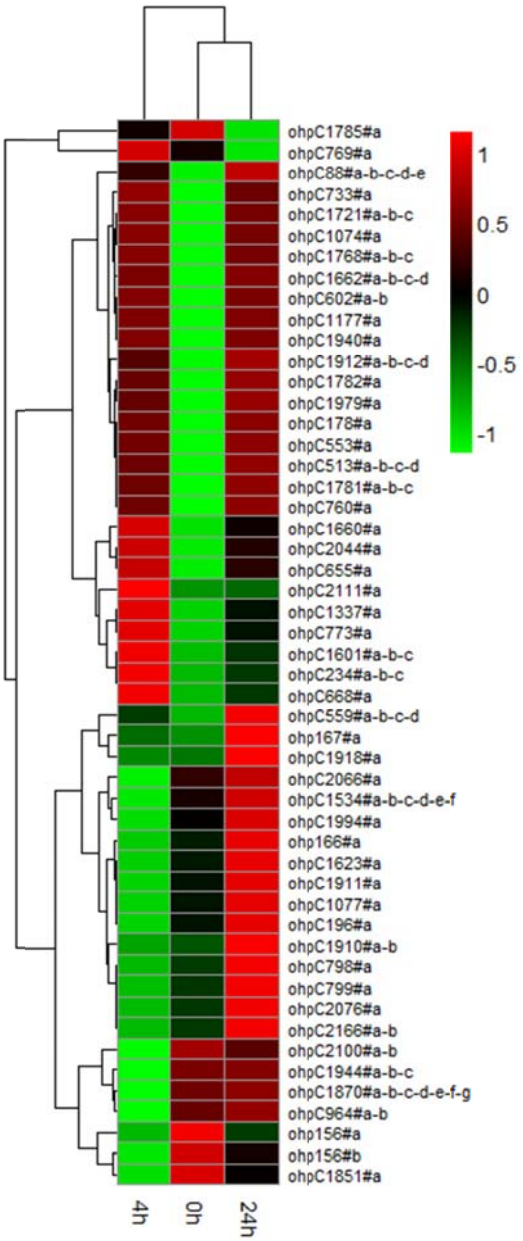


Figure 12. Expression of ohRNAs during heat stress

Heatmap shows the expression change of ohpRNAs during heat stress. The ohpRNAs that were significantly regulated during heat stress are selected: 0h (0h before heat), 4h (4h after heat) and 24h (24h after heat). the colors represent the Z-score: higher expression (more red) and lower expression (more green).

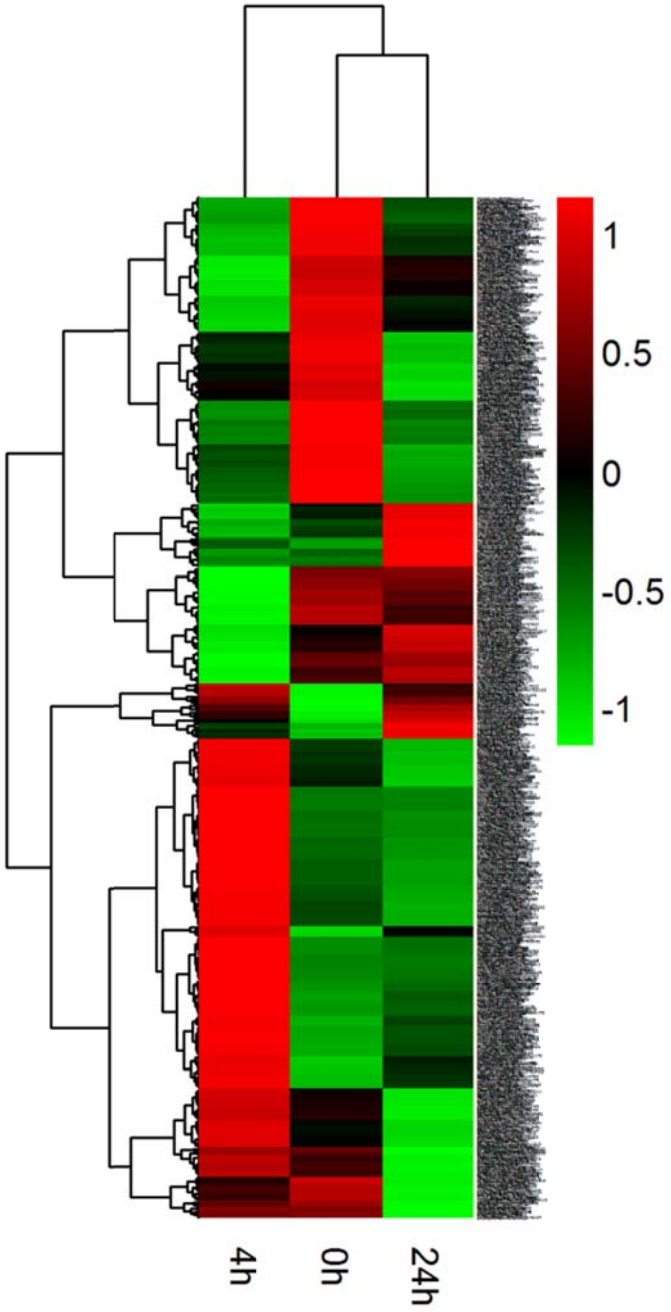


Figure 13. Expression of ra-siRNAs during heat stress

Heatmap shows the expression change of ra-siRNAs during heat stress. The ra-siRNAs that were significantly regulated during heat stress are selected: 0h (0h before heat), 4h (4h after heat) and 24h (24h after heat). the colors represent the Z-score: higher expression (more red) and lower expression (more green).

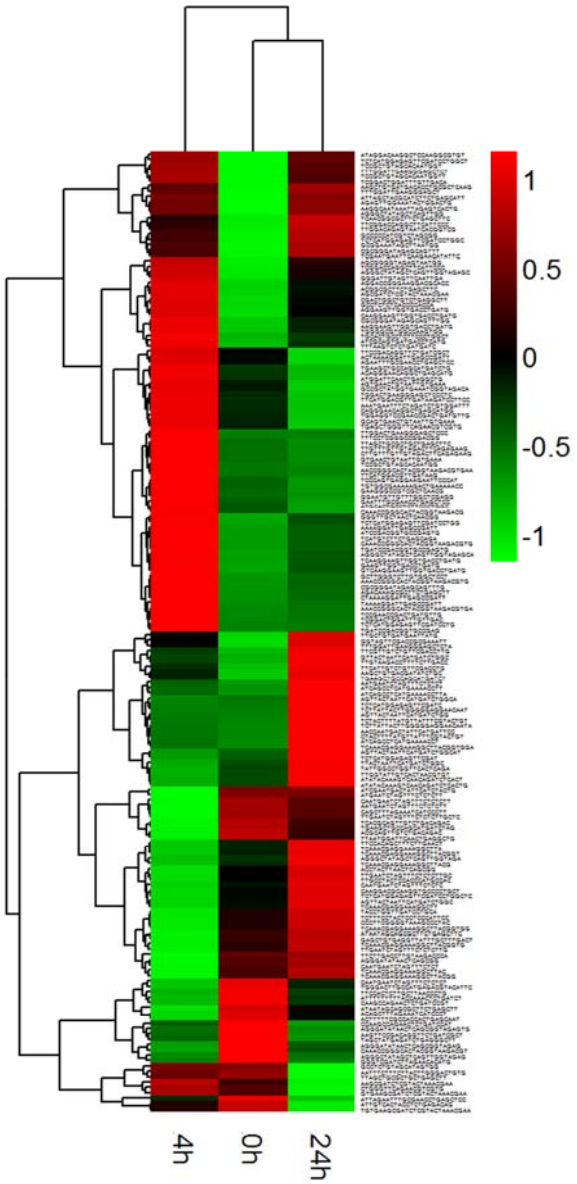


Figure 14. Expression of cis-nat siRNAs during heat stress

Heatmap shows the expression change of cis-nat siRNAs during heat stress. The cis-nat siRNAs that were significantly regulated during heat stress are selected: 0h (0h before heat), 4h (4h after heat) and 24h (24h after heat). The colors represent the Z-score: higher expression (more red) and lower expression (more green).

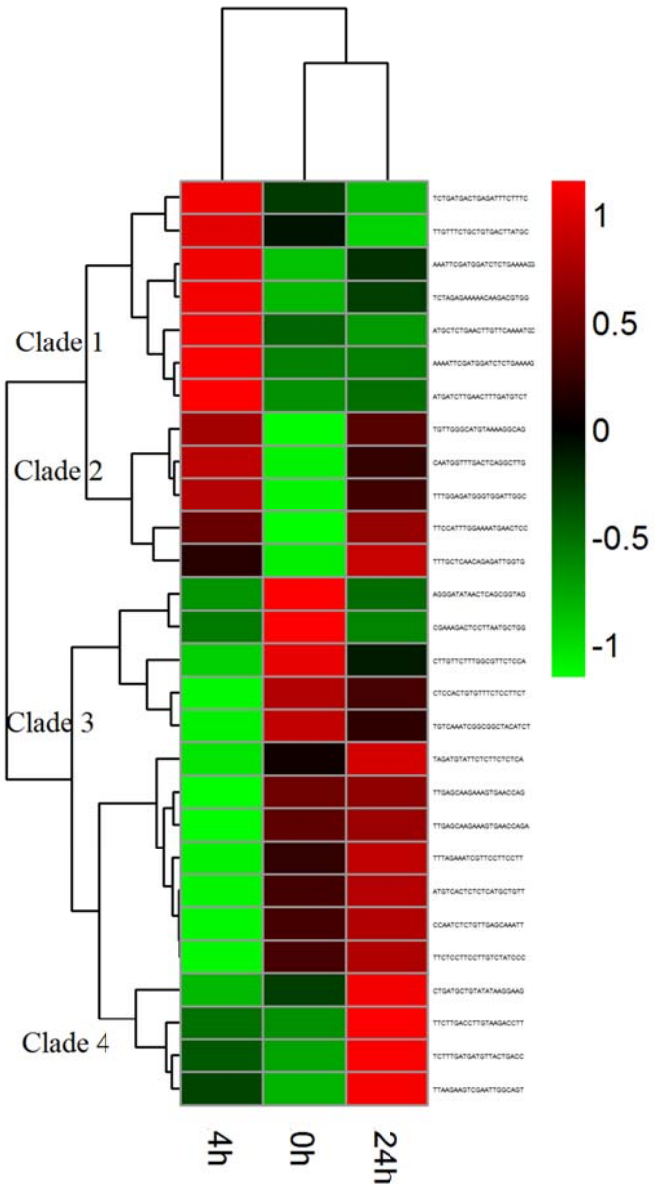


Figure 15. Expression of pha-siRNAs during heat stress

Heatmap shows the expression change of pha-siRNAs during heat stress. The pha-siRNAs that were significantly regulated during heat stress are selected: 0h (0h before heat), 4h (4h after heat) and 24h (24h after heat). The colors represent the Z-score: higher expression (more red) and lower expression (more green).

CHAPTER III

SMALL REGULATORY RNAS IN *GOSSYPIUM HIRSUTUM*

Abstract

In the present study, we constructed and sequenced forty four *G. hirsutum* small RNA libraries from eight types of tissues, using ABI Solid deep sequencing. These samples were prepared from heat-stressed/non-heat-stressed seedlings grown in green house and growth chamber, or nematode-infected/non-infected roots from the seedlings grown in green house, or reproductive tissues at six different developmental stages from the plants grown in EV Smith Experiment Station in summer. This generated a total of 708 million reads, corresponding to 85 million unique sequences. The analysis of these small RNAs identified a large number of small RNAs, which included 303 miRNAs, 6,114 ohpRNAs, 390,064 ra-siRNAs, 3,122 phased siRNAs, 238,813 *cis*-nat siRNAs, and 150,426 trans-nat siRNAs. Analysis of expression of small RNAs showed a subset of these small RNAs were differentially expressed in different tissues, either highly expressed in the root or leaf, or highly expressed in the reproductive tissues.

This work systematically investigated and profiled the major species of currently-documented small RNAs and analyzed their expression patterns in different tissues from the root to the boll. The identification of differentially expressed small RNAs would facilitate the further functional studies of small RNAs in the organ development and stress response. This provides an insight into the molecular mechanisms of nematode/heat tolerance and genetic breeding in cotton.

Introduction

Deep sequencing technology has become a powerful method for global analysis of transcriptomes of small regulatory RNAs (srRNAs) and for the study in their roles in plant growth, development and stress response (Kasschau et al. 2007; Zhang et al. 2013). Plant small RNAs are generated from diverse, complicated pathways that require Dicer-like

proteins (DCLs), Argonaute proteins (AGO), and/or RNA dependent polymerases (RDPs) (Axtell 2013). Based on the origin, biogenesis and function, plant small RNAs can be divided into two major families, hairpin RNAs (hpRNAs) and short interfering RNAs (siRNAs), and each family can be subdivided into different subgroups (Axtell 2013).

hpRNAs are generated from single-stranded precursors that are capable of forming hairpin-shape secondary structures, while siRNAs are processed from double-stranded precursors and require RNA dependent RNA polymerases (Axtell 2013). On the basis of biogenesis mechanism, hpRNAs and siRNAs can be subdivided into two and three groups, respectively (Axtell 2013).

hpRNAs consist of two major groups, microRNAs (miRNAs) and other hairpin RNAs (ohpRNAs). miRNAs are considered as the products that are precisely processed by DCLs, while ohpRNAs are produced from the imprecise cleavage by DCLs (Axtell 2013). miRNAs are well-documented in a variety of plants. The majority of plant miRNAs is ~21nt in length and regulates gene expression at post-transcription level. However, the biogenesis and function of ohpRNAs remain largely unknown.

In plants, siRNAs are classified into three major groups, heterochromatic siRNAs (or repeat-associated siRNAs), secondary siRNAs (sec-siRNAs), and natural antisense siRNAs (nat-siRNAs). Heterochromatic siRNAs (hc-siRNAs) are generated from heterochromatic or repetitive regions and associated with DNA methylation. sec-siRNAs are a specific type of siRNAs whose synthesis are dependent on the trigger or initiator of small RNAs, such as miRNAs. sec-siRNAs are demonstrated to be involved in the plant development and resistance to pathogen infection (Zhang C et al. 2012; Quintero et al. 2013). Natural antisense siRNAs (nat-siRNAs) are produced from a pair of transcripts, which are overlapping from opposite polarity. Nat-siRNAs are divided into *cis*-nat siRNAs and trans-nat siRNAs. *cis*-nat

siRNAs are processed from a pair of NAT that are transcribed from the same loci, while the precursors of *trans*-nat siRNAs are transcribed from different loci.

Growing evidences have shown the importance of siRNAs in plant growth, development, and stress response (Ron et al. 2010; Ito et al. 2011; Zhang X et al. 2012). Although miRNAs have been relatively-well profiled in *G. hirsutum* (Li et al. 2012; Yang et al. 2013), the species, expression, and function of siRNAs in *G. hirsutum* remain largely unknown. Especially, the molecular mechanism of fiber development still remains elusive, though some genes have been shown to be highly regulated in the fiber development (Taliercio and Boykin 2007; Rapp et al. 2010).

To gain insight into the roles of small RNAs, a large scale of deep sequencing was performed to sequence and profile *G. hirsutum* small RNAs from either stressed/unstressed seedlings, or reproductive tissues of heat tolerance/susceptible genotypes. The analysis of these data led to the identification of a large number of small RNAs, and a subset was highly regulated in different tissues, or between the vegetable and reproductive tissues. The data presented here will serve as foundation for future studies addressing fundamental molecular and developmental mechanisms that govern the organ development and regulatory mechanisms of plant abiotic/biotic tolerance.

Results

Deep sequencing of small RNA libraries from heat stress-treated *G. raimondii* seedlings

As shown in **Table 3** and **Supplementary Table S15**, ABA Solid Deep sequencing was employed to produce 708 million raw sequencing reads. After the preliminary treatment, 84,964,962 unique sequences, representing 230,202,658 total reads, were retained in the library. Of those, 34 million unique sequences (corresponding to 131 million reads) exactly

matched the genome of either *G. raimondii* (classified as D5 genome), *G. arboreum* (classified as A2 genome), or *G. hirsutum*-derived ESTs. Note that an almost equal percentage of unique sequences (or total reads) exactly matched the *G. raimondii* or *G. arboreum* genome. A total small RNA (sRNA) library was established by removing the sequences that corresponded to the non-protein coding sequences from the exactly genome-matching sequences (EMG). This left 34,363,412 unique sequences corresponding to 120,943,269 reads, and all these sequences will subsequently be referred to as the total sRNA library.

From the total sRNA library, 6,402 hpRNAs were identified, including 303 miRNAs (representing 8,376,873 reads) and 6,114 ohpRNAs (corresponding to 336,116 reads) (**Table 3 and Supplemental Table S15 and S16**). Notably, fourteen unique sequences were annotated into both miRNAs and ohpRNAs, because they were produced from multiple precursors, and a subset of the precursors failed to meet the criteria for annotation as plant miRNA precursors. Of the annotated miRNAs, 254 were lineage-specific miRNAs (ls-miRNAs), and the other 49 were long miRNAs, and among the ls-miRNAs, 140 putative miRNAs were homologous to known Viridiplantae miRNAs reported in miRBase version 20 (http://www.mirbase.org), and the other 163 were annotated as novel miRNAs currently found in only *G. raimondii* or *G. arboreum*.

After removal of the hpRNAs, the remaining sRNA sequences (34,357,009 unique sequences) were mapped to the repeat sequence loci in the *G. raimondii* genome. Consequently, 390,064 repeat-associated siRNAs (ra-siRNAs), corresponding to 9,299,246 reads, were identified, which accounted for 8% of the total sRNA reads or 1.25% of the unique sRNA sequences subsequently discovered in the total sRNA library.

The sRNAs after excluding hpRNAs and ra-siRNAs, the remaining sequences were utilized for identification of phasiRNAs. Consequently, 3,122 phased siRNAs (phasiRNAs), which represents 479,676 reads, were identified (**Table 3**), and these included 700, 146, and 2,276 siRNAs that were 21 nt, 22 nt, and 24 nt long, respectively. These phasiRNAs were produced from 510 phasiRNA-generating transcripts (also known as trans-acting siRNA-generating transcripts, TASs) derived from either protein-coding or non-protein-coding genes (**Supplemental Table S17-S19**). Most of 21nt phasiRNA-generating transcripts (184 of 232 21nt-phasiRNA-generating transcripts) were derived from various protein-coding genes, *TAS3* homologs (Auxin Response factors), NBS-LRR disease proteins, and MYB transcription factors, et al. (Supplemental Table S4). Of these, NBS-LRR genes (79) were the biggest source for 21nt-phasiRNA-generating transcripts. By contrast, most of 22nt or 24nt phasiRNAs were derived from non-protein coding genes (Supplemental Table S5 and S6).

To identify cis-nat or trans-nat siRNAs, sRNAs that have been categorized were excluded, and the retained sRNAs were used for further analysis. This resulted in identification of 389,239 siRNAs derived from a pair of transcripts that were transcribed from either the same or different genomic loci. 238,813 siRNAs that were derived from pairs of transcripts transcribed from the same loci were assigned cis-nat siRNAs, and others (150,426) derived from transcripts transcribed from different loci were trans-nat siRNAs (**Supplemental Table S25 and S26**).

The sequences in the total sRNA not annotated as hpRNAs or siRNA but annotated as uncharacterized sRNAs, represent what is most likely to be novel uncharacterized sRNAs found in the *G. hirsutum* genome. As the sequencing and assembly of the *G. hirsutum* genome is completed, it can be expected that more sRNAs would be characterized and annotated.

Expression of small RNAs of *G. hirsutum* among different tissues

To analyze the extent of the expression patterns of small RNAs, the sequencing data for the different classes of small RNAs were subjected to an analysis of hierarchical clustering. The abundance of each class of sRNAs within the different tissues was compared by averaging normalized read counts from libraries of the same treatments. Because the lowly expressed small RNAs represent the limit of detection rather than differential abundance analysis (Jeong et al. 2011), the small RNAs with at least one RPM at least one type of tissue were selected for the hierarchical analysis. To understand the expression change of sRNAs in different tissues, a detail analysis of each group of sRNAs is given below.

a. Expression of miRNAs in the different types of tissues

Based the hierarchical analysis, the miRNAs that were detected to be expressed in at least one TPM at least one type of tissue. As shown in **Figure 16**, these miRNAs were relatively well clustered by tissues according to tissue developmental stages and were grouped by the vegetable or reproductive tissues.

Almost all miRNAs, in a somewhat degree, displayed differential changes among different tissues (**Supplemental Table S21**). 197 miRNAs expressed with at least one RPMA in at least a tissue were selected to show in the Figure 1. According to the expression patterns, these miRNAs were divided into four major clades, clade 1 to 4 (**Figure 15A**). miRNAs in **Clade 1,** including 14 miRNAs, were relatively highly expressed in the reproductive tissues, with the highest expressed TPM value (257 TPM) being detected for miRC166 in the large square. **Clade 2** contained 43 miRNAs that were relatively higher expressed in the root, and the highest expression level (18,606 TPM) was detected for miR399 in the root and was more than twenty times than the leaf. but a subset also were detected to be relatively high

expressed in the square, boll or flower. The **Clade 3** included 18 miRNAs that were relatively higher expressed in the root and leaf. Interestingly, conserved miRNAs in this clade, such as miR171, miR156, and miR482, and these miRNAs were detected to be expressed more than 1,000 TPM in the root. Surprisingly, miRC1448 were overexpressed with more than 12,000 TPM in both root and leaf. The clade 4 contained 50 miRNAs that were higher expressed in the leaf, and 31 miRNAs (62%) were conserved miRNAs. Of these, miR2949 and miRC158 were intensively expressed in the leaf with more than 12,000 TPM, which was more than fifteen times than reproductive tissues. The clade 5 was the biggest clade and included 69 miRNAs, and 49 miRNAs (71%) were conserved miRNAs. The highest expression value (17,678 TPM) was detected for miR482b/c in the white flower. Notably, a subset of miRNAs such as miR396a and miR167a, was higher expressed in the root as well.

On the basis of the total abundance of miRNAs (**Supplemental Table S21**), miRNAs in the root and leaf accumulated to the highest level (186,4235 TPM in the root and 165,987 TPM in the leaf), while those were the lowest accumulation in the square or pink flower. Surprisingly, the abundance of miRNAs in the white flower and large boll were approximately more than two folds than other reproductive tissues. Although the total abundant dramatically varied between distinct tissues, 88% miRNAs (267/303) were detected at all types of tissues (**Figure 16B**). In addition, a few miRNAs were specifically detected in the specific tissues. For example, miR395a and miR395p were only detected in reproductive tissues, whereas miRC297 was detected in only flower and boll.

Based on the above analysis, it appeared that miRNAs were highly regulated between different tissues, and a subset of miRNAs was specifically expressed in the specific tissues, suggesting miRNAs were associated with *G. hirsutum* organ development.

b. Expression of ohpRNAs among different types of tissues

In present work, 6,114 ohpRNAs were identified through bioinformatical method, and 5,174 of those were expressed with at least one TPM in at least one type tissue. To better show the results, ohpRNAs that were detected at least 5 TPM in at least one library and were differentially expressed between any two different tissues ($p<0.05$) were selected as the representatives to demonstrate the expression change of ohpRNAs in different tissues. This remained 861 ohpRNAs for the hierarchical analysis, and the details were described in Supplemental Table S8.

As shown in **Figure 17A**, these ohpRNAs were clustered into three groups, the reproductive-tissue-specific (**Clade 1**), leaf-specific ohpRNAs (**Clade 2**) and leaf- and root-specific ohpRNAs (**Clade 3**), root-specific ohpRNAs (**clade 4**), and root-specific ohpRNAs. It should be noted that this classification only represented the trend of expression of ohpRNAs and a subset of ohpRNAs was higher expression on both vegetable and reproductive tissues.

Based on the total abundance, the ohpRNAs were higher expressed in the reproductive tissues than the vegetable tissues, root or leaf (**Supplemental Table S22**). Similar to miRNAs, 75% ohpRNAs were detected to be expressed among all types of tissues, while a subset of ohpRNAs appeared to be accumulated in only one or more specific tissues (**Figure 17B**).

The biogenesis and function of ohpRNAs are not well documented, but our analysis showed that ohpRNAs were differentially regulated between the vegetable tissues and reproductive tissues and that ohpRNAs were preferentially expressed in the reproductive tissues rather than the vegetable tissues. Although there is no evidence that ohpRNAs are associated with the tissue development in plants, our data implied that ohpRNAs might be vital for the development of *G. hirsutum* reproductive tissues.

c. Differentially expressed repeat-associated siRNAs

In this work, we profiled 390,064 repeat-associated siRNAs. But most of them were detected only one or two counts in only one type of tissue. To better analyze these siRNAs, ra-siRNAs that were with at least 8 TPM in at least one library as the representative to analyze their expression in different tissues. This remained 2,846 ra-siRNAs for the further analysis (**Supplemental Table S23**).

Based on the hierarchical analysis (**Figure 18A**), these ra-siRNAs can be divided into four clades, **Clade 1** to **4**. ra-siRNAs in **Clade 1** were relatively highly expressed in the root, while ra-siRNAs in **Clade 2** were higher expressed in both root and leaf. The **Clade 3** contained the ra-siRNAs that were relatively higher expressed in the leaf. The clade 4 contained ra-siRNAs that were relatively higher expressed in reproductive tissues.

d. Expression of phasiRNAs in different tissues

In this study, 3,122 phasiRNAs were identified from the *G. hirsutum* small libraries, whereas most of them were lowly expressed with one or two count. To find phasiRNAs with biological significance, the phasiRNAs with at least five TPM were selected for the analysis of hierarchical clustering. This remained 170 phased-siRNAs for the further analysis (**Supplemental Table S24**).

These siRNAs were classified into six clades, **Clade 1** to **4**, on the basis of the expression patterns (**Figure 19A**). **Clade 1** contained phasiRNAs that were relatively higher expressed in the root. The clade 2 included phasiRNAs that were higher expressed in the leaf. The clade 3 included phasiRNAs that were higher expressed in both leaf and small square. The clade 4was the biggest subgroup that contained phasiRNAs that were higher expressed in reproductive tissues. Of these phasiRNAs, 149 (87%) were detected among all eight tissues, and only 20 (12%) were detected in only the reproductive tissues (**Figure 19B**).

e. Expression of cis-nat siRNAs in different tissues

In this study, 238,813 cis-nat siRNAs were identified from the *G. hirsutum* small libraries. To find cis-nat siRNAs with biological significance, the cis-nat siRNAs with at least five TPM were selected for the analysis of hierarchical clustering. This remained 806 cis-nat siRNAs for the further analysis (**Supplemental Table S25**).

These cis-nat siRNAs were classified into three groups/clades, **Clade 1** to **3**, on the basis of the expression patterns (**Figure 20A**). **Clade 1** included cis-nat siRNAs that were preferentially expressed in the root, while **Clade 2** and **3** contained cis-nat siRNAs that were higher expressed in the reproductive tissues and leaf, respectively. Although the abundance of each cis-nat siRNAs appeared to vary in a large degree from tissues to tissues, most of cis-nat siRNAs (678 out of 806) were detected in all tissue types, and very few were detected in only one specific tissue (**Figure 20B**).

f. Expression of trans-nat siRNAs in different tissues

In this study, 150,426 trans-nat siRNAs were identified from the *G. hirsutum* small libraries. To find trans-nat siRNAs with biological significance, the trans-nat siRNAs with at least five TPM were selected for the analysis of hierarchical clustering. This remained 830 cis-nat siRNAs for the further analysis (**Supplemental Table S26**).

These cis-nat siRNAs were classified into three groups/clades, **Clade 1** to **3**, on the basis of the expression patterns (**Figure 21A**). **Clade 1** included cis-nat siRNAs that were preferentially expressed in the root, while **Clade 2** and **3** contained cis-nat siRNAs that were higher expressed in the reproductive tissues and leaf, respectively. Although the abundance of each cis-nat siRNAs appeared to vary in a large degree from tissues to tissues, most of cis-nat

siRNAs (678 out of 806) were detected in all tissue types, and very few were detected in only one specific tissue (**Figure 21B**).

Discussion and conclusion

Small RNAs in *Gossypium hirsutum*

Upland cotton (*G. hirsutum*) is a worldwide, economically importance crop, because it produces the largest source of the irreplaceable nature fiber for textile industry. The mechanism of development of cotton fiber has been intensively studied in anatomical, biochemical, and gene levels (Wang et al. 2004; Taliercio and Boykin 2007). Recently, the importance of gene expression regulated by small RNAs is shown to be crucial for the fiber development (Pang et al. 2009; Wang et al. 2012). Due to a lack of the completely sequenced and fully assembled genome, the study of the small RNAs in *G. hirsutum* is impeded and lagging behind the model species, such as Arabidopsis and rice. In this work, we used a newest, fully sequenced and the currently best assembled genome of *G. raimondii*, the model of D_T subgenome in *G. hirsutum* ($_T$ represents tetraploid) and a partially sequenced and assembled genome of *G. arboreum* (the model of A_T subgenome in *G. hirsutum*), combined a large scale of deep sequencing of small RNAs, to profile the small RNAs of *G. hirsutum*.

The work in this study constructed and sequenced the largest small RNA profiles of *G. hirsutum*, which comprised forty four small RNA libraries prepared from eight distinct of tissues or developmental stages, including root, leaf, square, flower and boll. This generated 708 M raw reads and left 230 M small RNAs. We noticed that the total reads of sRNAs matched the A2 or D5 genome were almost equal, suggesting the expression of sRNAs derived from the subgenome of A_T or D_T were equally expressed in the *G. hirsutum* genome. However, only 55.5% total reads or 40.1% unique reads exactly matched the A2 or D5 genome, while 85% total reads of *G. raimondii* small RNAs exactly matched the D5 genome.

This reflected that the genome of A2 (*G. arboreum*) or D5 (*G. raimondii*) might not be fully sequenced and assembled or the A2/D5 genome is substantially different with the subgenomes of A_T or D_T of polyploidy *G. hirsutum* genome.

It is known that 60% genes remained unchanged between *G. hirsutum* and diploid progenitors after polyploidization (Paterson et al. 2012). We noticed that ~63% conserved miRNAs or siRNAs are found in both A2 and D5 genome, suggesting at least 60% small RNA genes remained unchanged after divergence of the genome. However, only ~23% novel miRNAs or uncharacterized siRNAs were found in both A2 and D5 genome. This indicated a large number of *G. hirsutum* small RNAs that are evolutionarily non-conserved between different genomes. Among the uncharacterized sRNAs, 54% are 24 in length and characterized to be more like siRNAs. Due to importance of siRNAs in heterochromatic modification, it could be reasonably speculated that these non-conserved small RNAs might be especially biologically significant or important in the *G. hirsutum* genome for maintaining the epigenetic states.

Differentially expressed small RNAs in *G. hirsutum*

In this study, a large number of small RNAs were identified, and a subset of small RNAs was highly regulated in different tissues or developmental stage. This suggested these small RNAs are highly associated with the tissues or developmental stage.

Among the miRNAs highly expressed in reproductive tissues, miR156a-b and miR172 were highly upgraded at least 5-folds in square, flower and boll. These miRNAs have been shown to regulate the plant phase transition and flower timing (Wang et al. 2009; Zhu and Helliwell 2011). These miRNAs were very lowly expressed in the root or leaf samples, suggesting their crucial roles in the development of reproductive tissues in *G. hirsutum*. However, some miRNAs, such as miR164 and miR827, were highly expressed in the root or leaf, while lowly expressed in the reproductive tissues. This suggested that miRNAs were differentially

expressed in different tissues and might be associated with the development of vegetative and reproductive tissues.

Comparing to miRNAs, ohpRNAs were relatively lower expressed, whereas most of ohpRNAs were relatively higher expressed in reproductive tissues, rather than vegetable tissues (root and leaf). The function of ohpRNAs were not well-documented currently, but it is reported that these sRNAs are produced from complex, DCL-involved pathways in which multiple functional sRNAs ranging from 20 nt to 24 nt are produced and regulate gene expression by either DNA methylation and/or cleavage of target mRNAs (Dunoyer et al 2010; Henderson et al. 2006). Thus, we can infer that these differentially expressed ohpRNAs might play tissue-specific roles in the development of G. hirsutum reproductive tissues.

Unexpectedly, most of ra-siRNAs appeared to be highly expressed in the root or leaf, rather than in other reproductive tissues. A subset of ra-siRNAs was expressed with more than 1000 TPM in the leaf or roots, while down-graded two to five folds in the reproductive tissues. Previous studies have shown that ra-siRNAs regulate the heterochromatin formation (Fagegaltier et al. 2009) and varied greatly between parents and hybrids (Barber et al. 2012). However, our data showed that ra-siRNAs were associated with the organ development of *G. hirsutum*. Although the regulatory mechanism and roles of ra-siRNAs remained elusive, the data in this study provided an insight into the regulatory roles of ra-siRNAs in plant development.

Sec-siRNAs and cis-nat siRNAs have been recently found in various plant species (Zhai et al. 2011; Zhang et al. 2012). But none has been found in *Gossypium* plants to date. In this study, we bioinformatically identified a large number of putative sec-siRNA and cis-nat siRNAs. Unexpectedly, most of them were lowly expressed in the sequenced libraries. However, it appeared that the sec-siRNAs or cis-nat siRNAs were preferentially expressed in the leaf,

instead of other tissues, root or reproductive tissues. It indicated that the siRNAs were primarily produced from the leaf, rather than other tissues. This might be caused by the bias of EST or Unigene libraries of *G. hirsutum*, because most of *G. hirsutum* ESTs or Unigenes are produced from the leaf, while the identification of cis-nat siRNAs or sec-siRNAs are intensively dependent on the ESTs or Unigenes. Although a bias is present, our data still showed that a subset of sec-siRNAs or cis-nat was differentially expressed between vegetative and reproductive tissue. This suggested that sec-siRNAs or cis-nat siRNAs might be involved in the organ development of *G. hirsutum*. However, the further experimental evidences would be needed to confirm their roles.

In conclusion, we constructed and sequenced the forty four small RNA libraries of *G. hirsutum* that included eight different tissues or developmental stages. This generated the largest profiles of *G. hirsutum* small RNAs. The work in this study bioinformatically identified the five major classes of small RNAs currently documented in plants and profiled their expression in different tissues or developmental stages. This revealed that a large number of miRNAs or other siRNAs were differentially regulated between vegetative and reproductive tissues, suggesting small RNAs are involved in the organ development of *G. hirsutum*. This work would facilitate the further studies in the regulatory mechanism of organ development and molecular breeding of better-quality fiber or more stress-tolerant cotton.

Materials and Methods

Growth of plant material, total RNA preparation, and sRNA library sequencing

The youngest fully expanded leaves of four-week-old *G. raimondii* seedlings were used to obtain total RNA. Samples were harvested from at least 3 seedlings treated at 46^{o}C. for 0 h (no heat stress treatment), and after 4 or 24 h of heat stress (heat stress treatment). The total RNA was separately isolated from each sample using hot Borate method (Wan and Wilkins

1994). The constructing and sequencing of small RNAs was performed at the HudsonAlpha Institute Genome Sequencing Laboratory in Huntsville, AL, USA using an Illumina Small RNA Preparation Kit following manufacturer's protocol. Each small RNA sample was sequenced twice generating 2 sequencing replicates of each sample.

Processing of sequencing data

The raw sequencing data was processed using a pipeline consisting of a combination of custom-designed scripts, open source scripts, and CLC WorkBench 5.8 (http://www.clcbio.com/). Raw sequence data was filtered to remove sequencer reads lacking a 3' sequencing adaptor sequence, and the adaptor sequences and low quality bases were removed using CLC WorkBench 5.8. After filtering sequences 17 to 28 nt were retained for further analysis and designated as the preliminary sRNA libraries. Sequences from each of the 6 libraries and from a combined total library were separately collapsed to create read counts for each unique sequence in each of the single libraries and for the combined library.

Due to the unavailability of the genome of *G. hirsutum*, three reference database were used for further analysis, including the genome of *G. raimondii*, version 2 (*G.raimondii*_v2) (Paterson et al. 2012), and the preliminarily assembled genome of *G. arboreum* (the raw sequencing reads were downloaded from Comparative Evolutionary Genomics of Cotton, http://128.192.141.98/CottonFiber/) and the *G. hirsutum*-derived ESTs/unigenes downloaded from NCBI (ftp://ftp.ncbi.nlm.nih.gov/repository/UniGene/), DFCI (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=cotton), Comparative Evolutionary Genomics of Cotton, http://128.192.141.98/CottonFiber/) and PlantGDB (http://www.plantgdb.org/download/download.php?dir=/Sequence/ESTcontig/Gossypium_hirsutum/current_version). Bowtie mapping utility, version 1.0.0 (Berrier et al. 2010) was emplyoed to map sRNAs to the reference database, and only sequences exactly matching the

reference database were retained for further analysis. Unique sequences that exactly matched the non-protein coding sequences (rRNA, tRNA, snRNA, and snoRNA) were removed from the genome matching sequences above using the Bowtie mapping utility leaving total sRNA libraries for each treatment and replication and a combined total sRNA library that were used for subsequent analysis.

Bioinformatics determination of sRNAs

The identification of sRNAs, including hpRNAs, ra-siRNAs, phasiRNAs and cis-nat siRNAs, was performed using the methods and pipelines described in Chapter II, but the reference database above was used for this study.

The trans-nat siRNAs were identified by referring to the method previously described (Zhou et al. 2009). The pairs of transcripts that were not derived from the same loci from either *G. raimondii* or *G. arboreum* genome were used for trans-NAT analysis. Those putative NAT pairs that were able to form double stranded RNAs in which at least 100 nt long, complementary regions were found were annotated to trans-NATs, using a combination of in-house Perl script, BLASTN 2.2.28+ (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/) and RNAcofold version 2.0.7 that was built in Vienna RNA Package 2.0.7 (Lorenz et al. 2011).

Uncharacterized sRNAs

Those non hpRNAs that could not be annotated as ra-siRNAs, sec-siRNAs, or cis-NAT siRNAs as described above were grouped as currently uncharacterized sRNAs. Note that this remains the bulk of sRNAs in our library, and it is expected that many of these will ultimately be annotated to the current groups of siRNAs or to other new categories of siRNAs through time, but at the present time the bioinformatic tools available are not particularly robust, and

this forces the use of a conservative approach that may fail in full recognition of significant siRNAs.

Statistical analysis

kal's test is conducted by using statistical package built in commercial software CLC Genomic WorkBench 5.8.

miRNA and siRNA target prediction and GO analysis of target genes

The prediction of miRNAs and siRNAs targets was performed using psRNATarget (Dai and Zhao 2011) to search in the *G. raimondii* transcript database collected by JGI genomic Project, Phytozomev8.0 (http://www.jgi.doe.gov/), with maximum expectation of 2.0 and default for other parameters. The predicted target genes were subjected to Gene Ontology analysis performed using Blast2GO (Conesa and Gotz 2008) with default setts.

Reference

Axtell MJ, Westholm JO, Lai EC. 2011. Vive la difference: biogenesis and evolution of microRNAs in plants and animals. *Genome Biol.* **12**(4): 221.

Axtell MJ. 2013. Classification and Comparison of Small RNAs from Plants. *Annu Rev Plant Biol.* **64**: 137-159.

Barber WT, Zhang W, Win H, Varala KK, Dorweiler JE, Hudson ME, Moose SP. 2012. Repeat associated small RNAs vary among parents and following hybridization in maize. *PNAS.* **109**(26): 10444-10449.

Berrier A, Ulbricht R, Bonn M, Rivas JG. 2010. Ultrafast active control of localized surface plasmon resonances in silicon bowtie antennas. *Opt Express.* **18**(22): 23226-23235.

Conesa A, Gotz S. 2008. Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int J Plant Genomics.* **2008**: 619832.

Dai X, Zhao PX. 2011. psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res.* **39**(Web Server issue): W155-159.

Dunoyer P, Brosnan CA, Schott G, Wang Y, Jay F, Alioua A, Himber C, Voinnet O. 2010. An endogenous, systemic RNAi pathway in plants. *EMBO J.* **29**(10): 1699-1712.

Fagegaltier D, Bouge AL, Berry B, Poisot E, Sismeiro O, Coppee JY, Theodore L, Voinnet O, Antoniewski C. 2009. The endogenous siRNA pathway is involved in heterochromatin formation in Drosophila. *PNAS.* **106**(50): 21258-21263.

Henderson IR, Zhang X, Lu C, Johnson L, Meyers BC, Green PJ, Jacobsen SE. 2006. Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nat Gene.* **38**(6): 721-725.

Ito H, Gaubert H, Bucher E, Mirouze M, Vaillant I, Paszkowski J. 2011. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature.* **472**(7341): 115-119.

Jeong DH, Park S, Zhai J, Gurazada SG, De Paoli E, Meyers BC, Green PJ. 2011. Massive analysis of rice small RNAs: mechanistic implications of regulated microRNAs and variants for differential target RNA cleavage. *Plant Cell.* **23**(12): 4185-4207.

Kasschau KD, Fahlgren N, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Carrington JC. 2007. Genome-wide profiling and analysis of Arabidopsis siRNAs. *PLoS Biol.* **5**(3): e57.

Li Q, Jin X, Zhu YX. 2012. Identification and analyses of miRNA genes in allotetraploid Gossypium hirsutum fiber cells based on the sequenced diploid G. raimondii genome. *J Genet Genomics.* **39**(7): 351-360.

Lorenz R, Bernhart SH, Honer Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA Package 2.0. *AMB.* **6**: 26.

Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, Bowman JL, Cao X, Carrington JC, Chen X, Green PJ et al. 2008. Criteria for annotation of plant MicroRNAs. *Plant Cell* **20**(12): 3186-3190.

Pang M, Woodward AW, Agarwal V, Guan X, Ha M, Ramachandran V, Chen X, Triplett BA, Stelly DM, Chen ZJ. 2009. Genome-wide analysis reveals rapid and dynamic changes in miRNA and siRNA sequence and expression during ovule and fiber development in allotetraploid cotton (*Gossypium hirsutum* L.). *Genome Biol.* **10**(11): R122.

Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, Llewellyn D, Showmaker KC, Shu S, Udall J et al. 2012. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. *Nature.* **492**(7429): 423-427.

Quintero A, Perez-Quintero AL, Lopez C. 2013. Identification of ta-siRNAs and Cis-nat-siRNAs in Cassava and Their Rolein Response to Cassava Bacterial Blight. *Genomics Prot Bioinformatics*. **11**(3):172-81.

Rapp RA, Haigler CH, Flagel L, Hovav RH, Udall JA, Wendel JF. 2010. Gene expression in developing fibres of Upland cotton (Gossypium hirsutum L.) was massively altered by domestication. *BMC Biol.* **8**: 139.

Rogers K, Chen X. 2013. microRNA Biogenesis and Turnover in Plants. *Cold Spring Harb Sym Quant Biol.* **77**:183-94.

Ron M, Alandete Saez M, Eshed Williams L, Fletcher JC, McCormick S. 2010. Proper regulation of a sperm-specific cis-nat-siRNA is essential for double fertilization in Arabidopsis. *Genes Deve.* **24**(10): 1010-1021.

Taliercio EW, Boykin D. 2007. Analysis of gene expression in cotton fiber initials. *BMC Plant Biol.* **7**: 22.

Wan CY, Wilkins TA. 1994. A modified hot borate method significantly enhances the yield of high-quality RNA from cotton (*Gossypium hirsutum* L.). *Anal Biochem.* **223**(1): 7-12.

Wang H, Zhang X, Liu J, Kiba T, Woo J, Ojo T, Hafner M, Tuschl T, Chua NH, Wang XJ. 2011a. Deep sequencing of small RNAs specifically associated with Arabidopsis AGO1 and AGO4 uncovers new AGO functions. *Plant J.* **67**(2): 292-304.

Wang JW, Czech B, Weigel D. 2009. miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. *Cell.* **138**(4): 738-749.

Wang S, Wang JW, Yu N, Li CH, Luo B, Gou JY, Wang LJ, Chen XY. 2004. Control of plant trichome development by a cotton fiber MYB gene. *Plant Cell.* **16**(9): 2323-2334.

Yang X, Wang L, Yuan D, Lindsey K, Zhang X. 2013. Small RNA and degradome sequencing reveal complex miRNA regulation during cotton somatic embryogenesis. *J Exp Bot.* **64**(6): 1521-1536.

Zhai J, Jeong DH, De Paoli E, Park S, Rosen BD, Li Y, Gonzalez AJ, Yan Z, Kitto SL, Grusak MA et al. 2011. MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes Dev.* **25**(23): 2540-2553.

Zhang C, Li G, Wang J, Fang J. 2012. Identification of trans-acting siRNAs and their regulatory cascades in grapevine. *Bioinformatics* **28**(20): 2561-2568.

Zhang X, Li Y, Wu Z, Polishko A, Zhang H, Chinnusamy V, Lonardi S, Zhu JK, Liu R, Jin H. 2013. Mechanisms of Small RNA Generation from Cis-NATs in Response to Environmental and Developmental Cues. *Mol Plant.* **6**(3): 704-715.

Zhang X, Xia J, Lii YE, Barrera-Figueroa BE, Zhou X, Gao S, Lu L, Niu D, Chen Z, Leung C et al. 2012. Genome-wide analysis of plant nat-siRNAs reveals insights into their distribution, biogenesis and function. *Genome Biol.* **13**(3): R20.

Zhou X, Sunkar R, Jin H, Zhu JK, Zhang W. 2009. Genome-wide identification and analysis of small RNAs originated from natural antisense transcripts in Oryza sativa. *Genome Res.* **19**(1): 70-78.

Zhu QH, Helliwell CA. 2011. Regulation of flowering time and floral patterning by miR172. *J Exp Bot.* **62**(2): 487-495.

Table 3. Categories of G. hirsutum small RNAs

	Total	Unique
Raw reads	707,671,830	
Reads after trmimming adaptor, low qualtity nucleotides and reads shorter than 17nt or longer than 28nt	230,202,658	84,964,962
Genome/uniqge matched	130,922,432	34,478,006
G. raimondii	92,826,037	21,354,725
	71%	62%
G. arboreum	93,741,833	21,964,906
	72%	64%
Unigenes of *G. hirsutum*	46,768,762	4,082,028
	36%	12%
Small RNAs (excluding non-protein RNAs)	120,943,269	34,363,412
	100%	100%
Hairpin RNAs	8,708,893	6,403
	7%	0%
MicroRNAs (miRNAs)	8,376,873	303
	7%	0%
Other hairpin RNAs (ohpRNAs)	336,116	6,114
	0%	0%
Short interfering RNAs (siRNAs)	13,060,052	631,999
	9%	2%
Repeat-associated RNAs	9,299,246	390,064
	8%	1%
Phased siRNAs	479,676	3,122
	0%	0.00
Phased siRNAs (21nt)	263,732	700
	0%	0%
Phased siRNAs (22nt)	39,292	146
	0%	0%
Phased siRNAs (24nt)	176,652	2,276
	0%	0%
cis natural antisense siRNAs	3,281,130	238,813
(cis-nat siRNAs)	3%	1%
Uncharacterized	101,408,187	33,784,795
	84%	98%

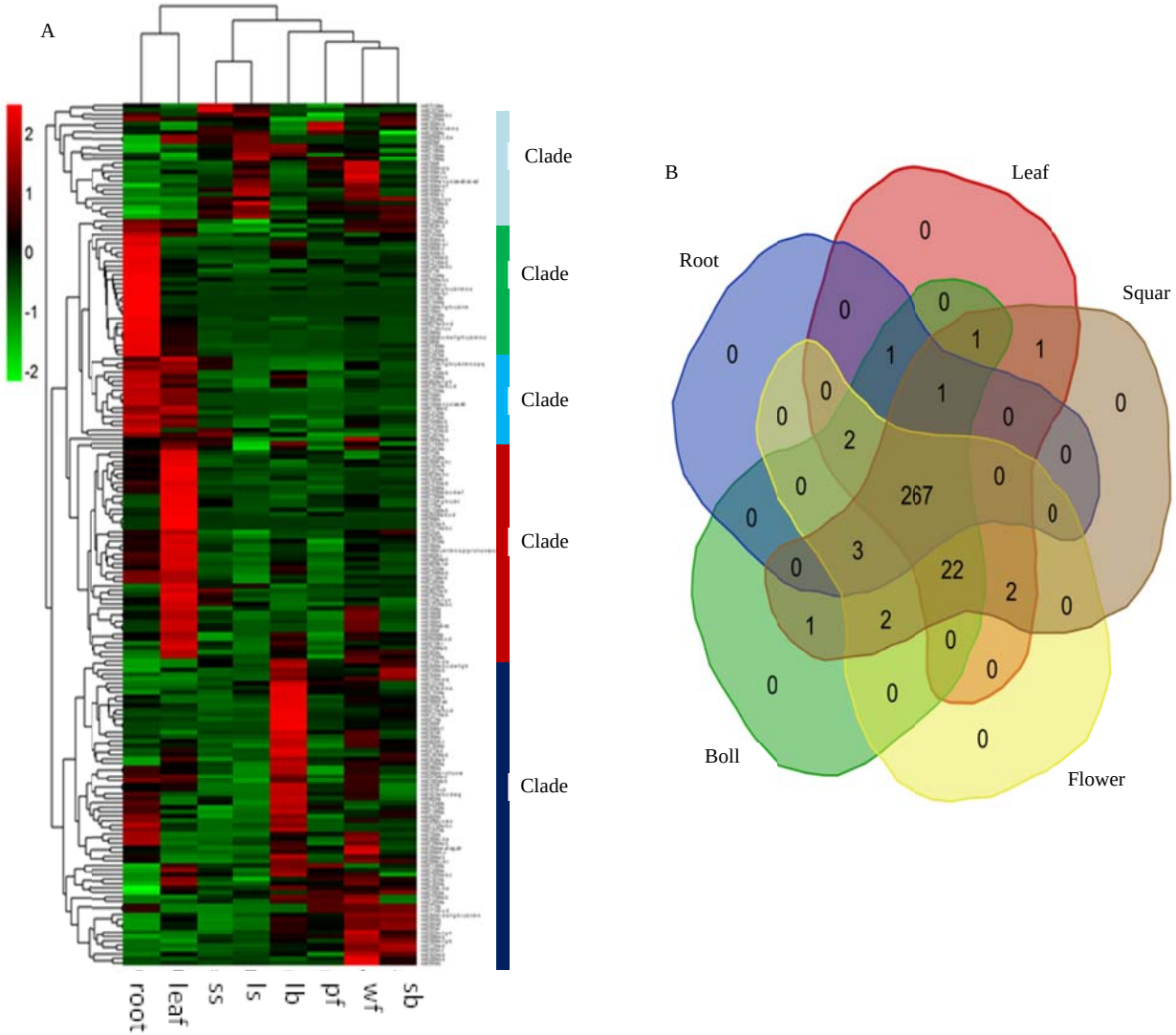


Figure 16. Expression of miRNAs among different tissues

The Hierarchical analysis of expression of miRNAs among different tissues is shown in A; the Z-score of the expression value for each miRNA is shown in different colors; The tissues were denoted as root, leaf, ss (small square), ls (large square), wf (white flower), pf (pink flower), sb (small boll), and lb (large boll).

The Venn diagram (Panel B) shows the number of miRNAs that were detected among different tissues.

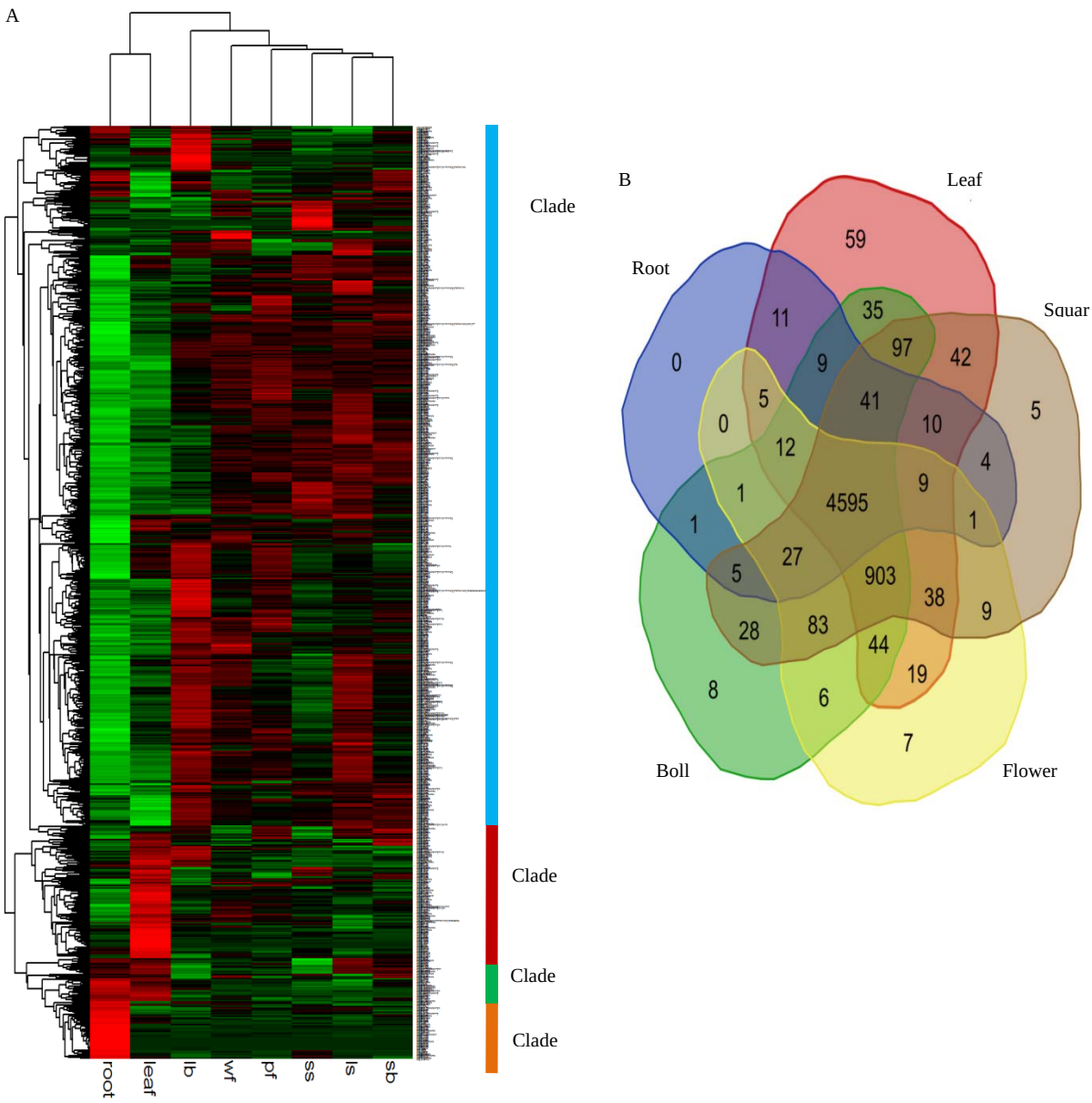


Figure 17. Expression of ohpRNAs among different tissues

The Hierarchical analysis of expression of miRNAs among different tissues is shown in A; the Z-score of the expression value for each miRNA is shown in different colors; The tissues were denoted as root, leaf, ss (small square), ls (large square), wf (white flower), pf (pink flower), sb (small boll), and lb (large boll).

The Venn diagram (Panel B) shows the number of miRNAs that were detected among different tissues.

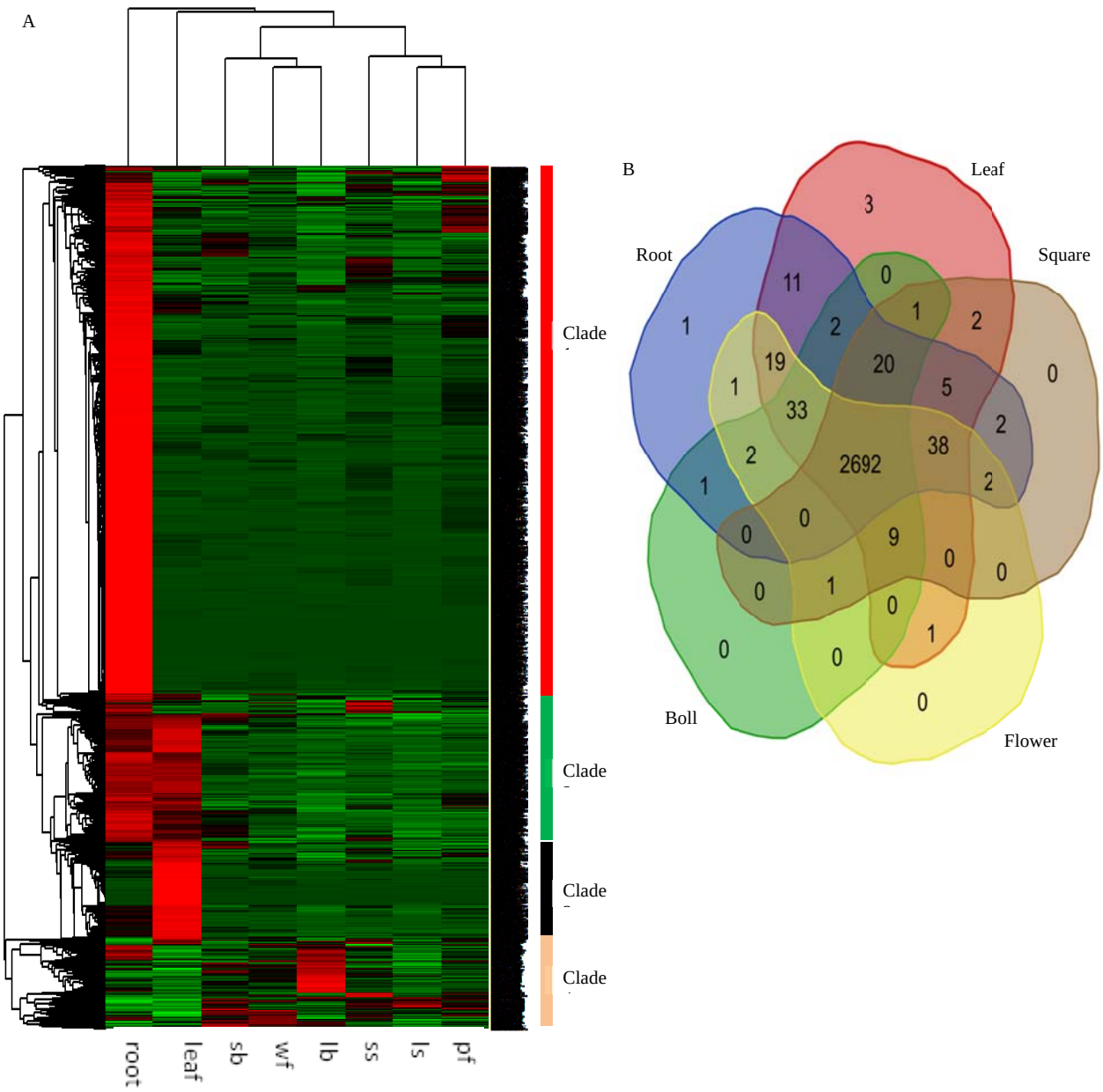


Figure 18. Expression of ra-siRNAs among different tissues

The Hierarchical analysis of expression of miRNAs among different tissues is shown in A; the Z-score of the expression value for each miRNA is shown in different colors; The tissues were denoted as root, leaf, ss (small square), ls (large square), wf (white flower), pf (pink flower), sb (small boll), and lb (large boll).

The Venn diagram (Panel B) shows the number of miRNAs that were detected among different tissues.

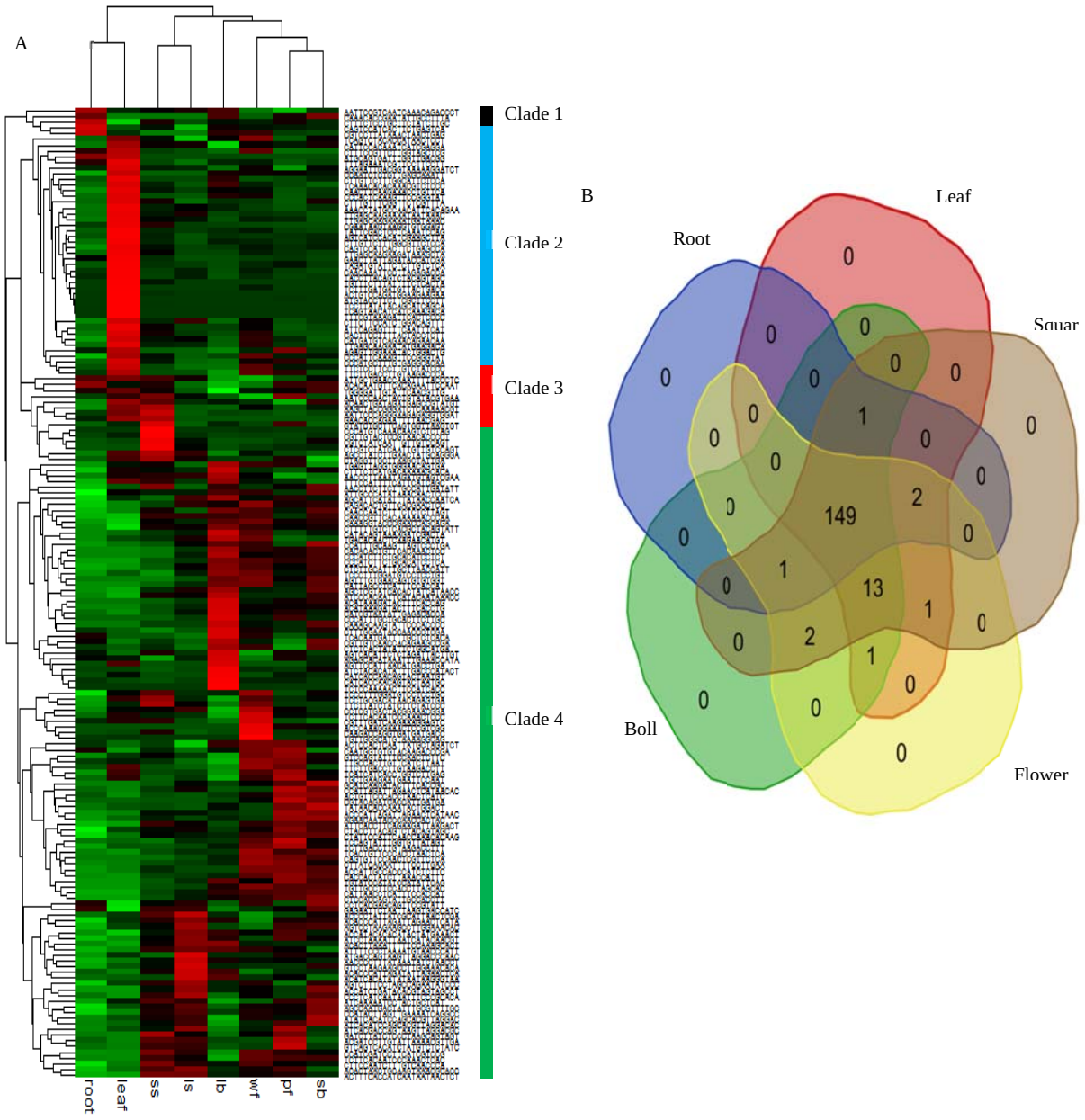


Figure 19. Expression of phasiRNAs among different tissues

The Hierarchical analysis of expression of miRNAs among different tissues is shown in A; the Z-score of the expression value for each miRNA is shown in different colors; The tissues were denoted as root, leaf, ss (small square), ls (large square), wf (white flower), pf (pink flower), sb (small boll), and lb (large boll).

The Venn diagram (Panel B) shows the number of miRNAs that were detected among different tissues.

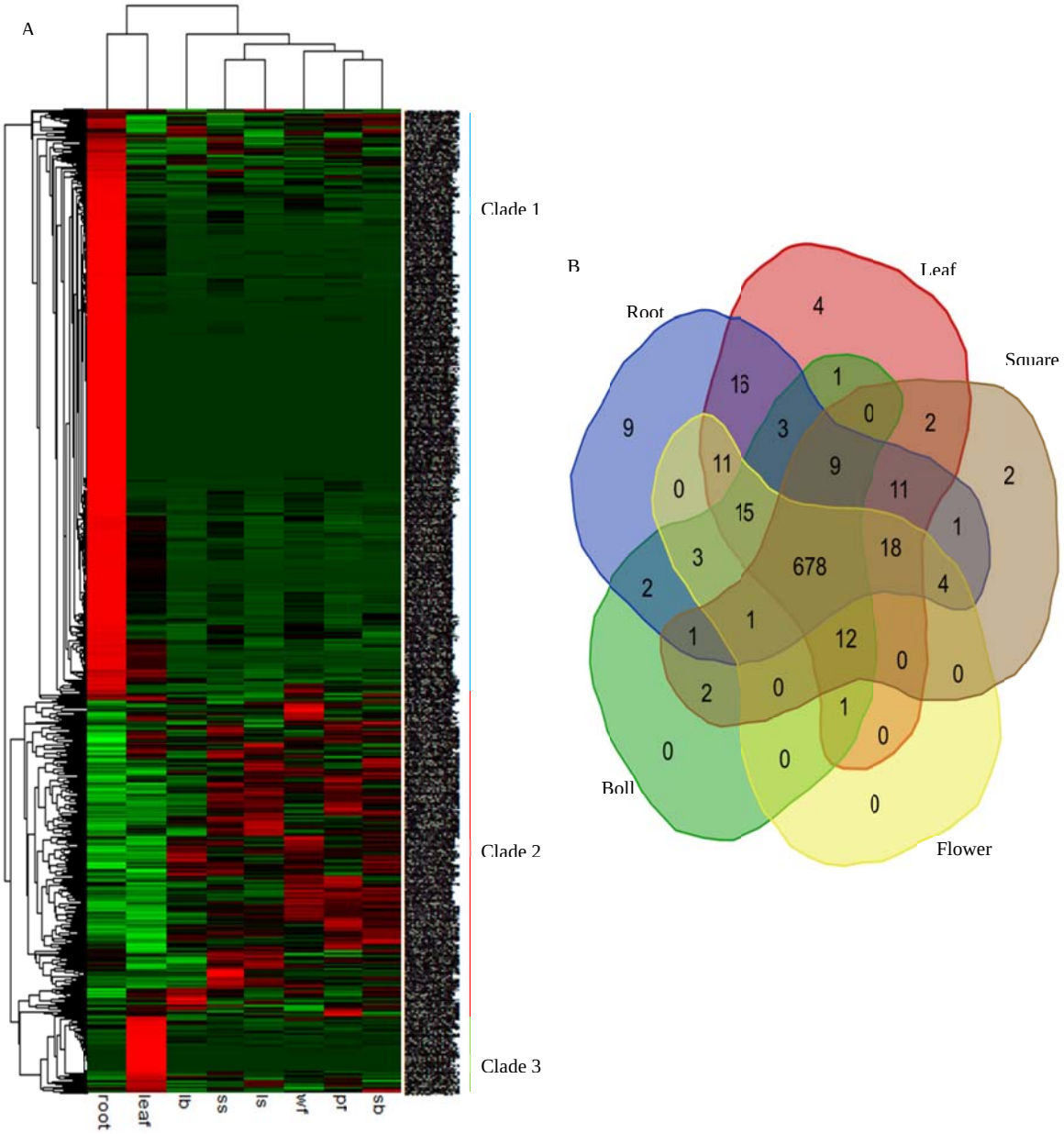


Figure 20. Expression of cis-nat siRNAs among different tissues

The Hierarchical analysis of expression of miRNAs among different tissues is shown in A; the Z-score of the expression value for each miRNA is shown in different colors; The tissues were denoted as root, leaf, ss (small square), ls (large square), wf (white flower), pf (pink flower), sb (small boll), and lb (large boll).

The Venn diagram (Panel B) shows the number of miRNAs that were detected among different tissues.

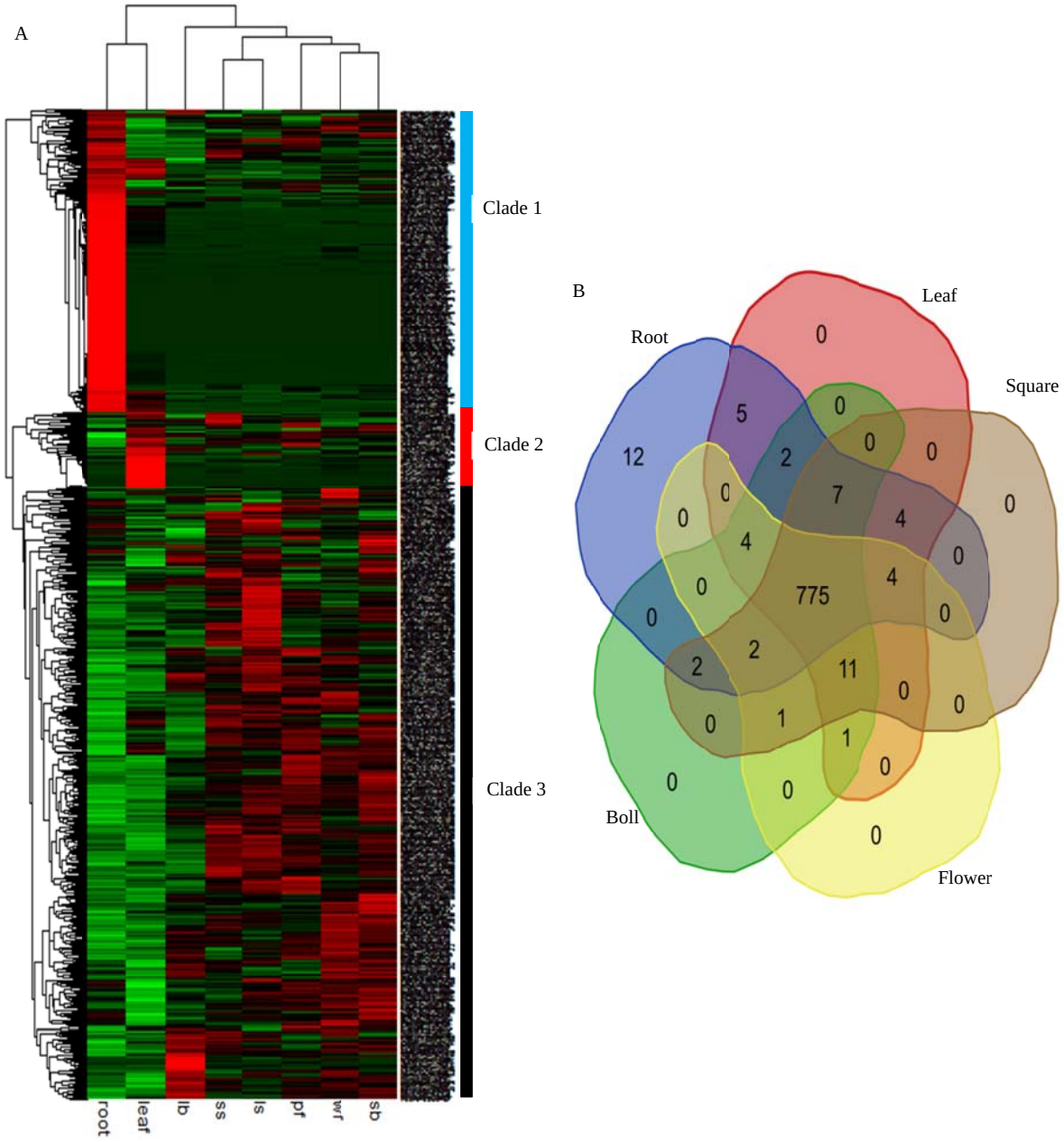


Figure 21. Expression of trans-nat siRNAs among different tissues

The Hierarchical analysis of expression of miRNAs among different tissues is shown in A; the Z-score of the expression value for each miRNA is shown in different colors; The tissues were denoted as root, leaf, ss (small square), ls (large square), wf (white flower), pf (pink flower), sb (small boll), and lb (large boll).

The Venn diagram (Panel B) shows the number of miRNAs that were detected among different tissues

Legends for Supplemental Figures

Supplemental Figure 1. Diagram for *G. raimondii* pha-siRNAs biogenesis

The diagrams for the biogenesis of *G. raimondii* pha-siRNAs are shown in A (21nt), B (22nt), and C (24). The line beginning with the sign ">" represent the genomic loci of pha-siRNA transcripts, the "Initiator" represent miRNA initiator of pha-siRNA transcripts from 3'-> 5. The plus (+) or negative signs (-) indicates the derivation of pha-siRNAs from plus strand or complementary strand, respectively, and the number represents the expression value of each pha-siRNA.

Supplemental Figure S2-14. Global changes of sRNAs along the 13 chromosomes during heat stress

The expression changes of sRNAs along the 13 chromosomes are shown in Supplemental Figure S2 to S14. The x-axis represents the chromosome and the number on the x-axis represents the base position. The expression value of sRNAs (rpm) at each locus is shown in y-axis. The black, red and blue cycles represent the different treatments, 0h (before heat treatment), 4h (after heat treatment), and 24h (after heat treatment).