The Role of Non-Coding RNA in Plant Stress

Thesis by
Cameron Ross MacPherson

In Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy

King Abdullah University of Science and Technology, Thuwal, Kingdom of Saudi Arabia

December, 2012
EXAMINATION COMMITTEE APPROVAL FORM

The thesis of Cameron Ross MacPherson is approved by the examination committee.

Committee Chairperson: Prof. Vladimir Bajic
Committee Co-Chair: Prof. Nam-Hai Chua
Committee Member: Prof. Nina V. Fedoroff
Committee Member: Prof. Pierre Magistretti
COPYRIGHT

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License (CC BY-NC-SA 3.0).

You are free:
- to Share — to copy, distribute and transmit the work
- to Remix — to adapt the work

Under the following conditions:
- Attribution — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
- Noncommercial — You may not use this work for commercial purposes.
- Share Alike — If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar license to this one.

With the understanding that:
- Waiver — Any of the above conditions can be waived if you get permission from the copyright holder.
- Public Domain — Where the work or any of its elements is in the public domain under applicable law, that status is in no way affected by the license.
- Other Rights — In no way are any of the following rights affected by the license:
  - Your fair dealing or fair use rights, or other applicable copyright exceptions and limitations;
  - The author’s moral rights;
  - Rights other persons may have either in the work itself or in how the work is used, such as publicity or privacy rights.
- Notice — For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to this web page.

© 2012

Cameron Ross MacPherson

CC BY-NC-SA 3.0
Post-transcriptional gene silencing (PTGS) is a powerful mechanism that can be adapted to genetically modify crop plants. PTGS operates in many plant signaling pathways including those mediating stress responses. Given the small number of miRNAs known, research on the characterization of stress-related micro-RNA (miRNA) and their targets could provide the basis for engineering stress tolerant traits in crops. Indeed, several examples of miRNA mediated crop tolerance have been reported. In the research presented here, we aimed to analyze the role of small non-coding RNA (smRNA) pathways involved in plant stress. In particular, we focused on miRNA-mediated PTGS in phosphate (Pi) starvation. The analysis was split into two research projects. First, to identify potential miRNA targets we began by analyzing the response and recovery of coding and long non-coding RNAs (lncRNA) to Pi starvation in shoot and root. The results obtained were the first genome-wide description of the root’s Pi starvation response and recovery. We found that the root’s response involved a widely different set of genes than that of the shoot. In the second research project, the results of the first project were correlated with the responses of miRNA and trans-acting small-interfering RNA (tasiRNA) during Pi starvation. Many miRNA circuits have been predicted before, however, tasiRNA circuits are not as well defined. Therefore, we made use of the double-stranded RNA-binding protein 4 (DRB4) smRNA libraries to enhance our prediction of tasiRNAs. Altogether, we provided evidence to support the following
miRNA-mRNA pairs that may function in Pi starvation: IPS1:miR399:PHO2; miR399:RS4; miR399:NF-YA10; miR398:CSD1/2; miR2111:TPS11; miR164:NAC6; miR157:TMO7; miR157:PSB28; RPS2:miR169:IPS2; miR397:LAC2; TAS4:PAP1; NR1:PAP1; and Chr3_1967672:TMO7. In general, we found that non-miR399 related circuits were active only during the root’s recovery from Pi starvation. The functional roles of the genes targeted by these PTGS circuits suggested that the local root response to Pi starvation was influenced by the plant's systemic response pathways via PHR1-mediated PTGS. Finally, since many PTGS targets function to modulate concentrations of reactive oxygen species and sucrose, we hypothesized that the candidate PTGS circuits found in our research mediate a general stress recovery process by modulating metabolites involved in signaling pathways.
# TABLE OF CONTENTS

Examination committee approval form 2
Copyright 3
Abstract 4
Table of contents 6
List of abbreviations 8
List of figures 9
List of tables 10

**Chapter 1:** Introduction
  Justification and summary of research conducted 11
  Non-coding small-RNA 15
  Small-RNA biology 19
  The role of miRNA post-transcriptional gene silencing 22
  The role of tasiRNA post-transcriptional gene silencing 28
  Conclusion 30

**Chapter 2:** The molecular response to, and recovery from phosphate starvation 31
  Introduction 31
  Results 33
    General data observation 33
    Novel Pi-signaling patterns observed among genes involved in phosphate starvation 34
    The response and recovery of genes absent from the ATH1 microarray platform 38
    Analysis of functionally related genes 42
    Expression patterns of known Pi starvation responsive genes in roots and shoots 43
    Three gene expression phases uncover functional responses to Pi starvation 46
    Interaction between Pi starvation and various AtGenExpress treatments 52
    Genes affected only during Pi starvation 55
    Pi starvation responsive genes affected in multiple AtGenExpress treatments 58
  Discussion 61
Comparisons among different studies reveal the need to establish a reference dataset 61
Expression profiles of Pi starvation responsive genes absent from the ATH1 microarray 61
A novel root response suggests roles for energy metabolism 64
Transcription factors increase expression in root as the system tends toward recovery 65
Expression of IPT3, ARF9, and MYB85 may alter root-morphology during Pi starvation 66
Cross-talk among Pi starvation candidates uncovers commonality in response to plant hormone treatments 67
Conclusion 68

Chapter 3: PTGS mediates cross-talk between the systemic and local phosphate starvation response 70
Introduction 70
Results 73
Known miRNA circuits in stress 76
Novel miRNA:target and target-mimic circuits 83
Trans-acting small-interfering RNA in plant stress 100
DRB4-dependent tasiRNA-derived siRNAs target Pi starvation responsive loci 105
Discussion 107
Conclusion 113

Chapter 4: Final Remarks 115
References 123

Supplementary Chapter A.1: Materials and Methods for Chapter 2 134
Supplementary Chapter A.2: Materials and Methods for Chapter 3 135
Data-sources 135
Automatic hypothesis generation 136
Analysis of stress responsive circuits 138
Prediction of tasiRNA genes, siRNAs, and targets 139
DRB4 smRNA data processing 140
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEX</td>
<td>Differentially expressed</td>
</tr>
<tr>
<td>DRB</td>
<td>Double-stranded RNA binding</td>
</tr>
<tr>
<td>dsRNA</td>
<td>Double-stranded RNA</td>
</tr>
<tr>
<td>GEO</td>
<td>Gene Expression Omnibus</td>
</tr>
<tr>
<td>INR</td>
<td>Initial negative response</td>
</tr>
<tr>
<td>IPR</td>
<td>Initial positive response</td>
</tr>
<tr>
<td>IP-Seq</td>
<td>Immuno-precipitation sequencing</td>
</tr>
<tr>
<td>IncRNA</td>
<td>Long non-coding RNA</td>
</tr>
<tr>
<td>LNR</td>
<td>Latent negative response</td>
</tr>
<tr>
<td>LOF</td>
<td>Loss of function</td>
</tr>
<tr>
<td>LPR</td>
<td>Latent positive response</td>
</tr>
<tr>
<td>LR</td>
<td>Lateral root</td>
</tr>
<tr>
<td>mat-miRNA</td>
<td>Mature miRNA</td>
</tr>
<tr>
<td>mat-smRNA</td>
<td>Mature smRNA</td>
</tr>
<tr>
<td>MINP</td>
<td>Minimum p-value</td>
</tr>
<tr>
<td>miRNA</td>
<td>Micro-RNA</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger-RNA</td>
</tr>
<tr>
<td>ncRNA</td>
<td>Non-coding RNA</td>
</tr>
<tr>
<td>PC</td>
<td>Principal component</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal components analysis</td>
</tr>
<tr>
<td>Pi</td>
<td>Inorganic phosphate</td>
</tr>
<tr>
<td>piRNA</td>
<td>piwi-RNA</td>
</tr>
<tr>
<td>PNR</td>
<td>Persistent negative response</td>
</tr>
<tr>
<td>PPR</td>
<td>Persistent positive response</td>
</tr>
<tr>
<td>pre-smRNA</td>
<td>Premature smRNA</td>
</tr>
<tr>
<td>pri-smRNA</td>
<td>Primary smRNA</td>
</tr>
<tr>
<td>PTGS</td>
<td>Post-transcriptional gene silencing</td>
</tr>
<tr>
<td>Rec.</td>
<td>Recovery</td>
</tr>
<tr>
<td>Res.</td>
<td>Response</td>
</tr>
<tr>
<td>RISC</td>
<td>RNA-induced silencing complex</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>RREC</td>
<td>Root recovery</td>
</tr>
<tr>
<td>RRES</td>
<td>Root response</td>
</tr>
<tr>
<td>siRNA</td>
<td>small-interfering RNA</td>
</tr>
<tr>
<td>smRNA</td>
<td>small-RNA</td>
</tr>
<tr>
<td>SREC</td>
<td>Shoot recovery</td>
</tr>
<tr>
<td>SRES</td>
<td>Shoot response</td>
</tr>
<tr>
<td>tasiRNA</td>
<td>trans-acting small-interfering RNA</td>
</tr>
<tr>
<td>TF</td>
<td>Transcription factor</td>
</tr>
<tr>
<td>TGS</td>
<td>Transcriptional gene silencing</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

-- Chapter 1 --
Figure 1.1: Biogenesis of micro-RNAs in A. thaliana 18
Figure 1.2: Detailed biogenesis pathways of several small-RNA classes in A. thaliana 27

-- Chapter 2 --
Figure 2.1: Initial assessment of data and quality control 34
Figure 2.2: The distribution of probe-sets across tissues and response-and-recovery classes 42
Figure 2.3: Analysis of gene expression and function among known Pi responsive loci 45
Figure 2.4: Analysis of gene expression and function among novel Pi responsive loci 47
Figure 2.5: Comparative analysis between P_i starvation data and AtGenExpress 54
Figure 2.6: Cross-talk between P_i^{starv}, ABA treatment and cold, salt, and osmotic stress 59

-- Chapter 3 --
Figure 3.1: Experimental evidence for miR398:CSD1 and miR398:CSD2 circuits 82
Figure 3.2: Experimental evidence for the miR398:PSB28 circuit 83
Figure 3.3: Experimental evidence for the miR169:At1g67600 circuit 88
Figure 3.4: Experimental evidence for the miR2111:TPS11 circuit 89
Figure 3.5: Experimental evidence for the miR399:NF-YA10 circuit 91
Figure 3.6: Experimental evidence for the miR164:NAC6 circuit 93
Figure 3.7: Experimental evidence for the miR399:RS4 circuit 94
Figure 3.8: Experimental evidence for the miR157:TMO7 circuit 98
Figure 3.9: Experimental evidence for the miR397:LAC2 circuit 99
Figure 3.10: Tiling-array data from Pi starvation experiments for TAS1b and TAS2 104
Figure 3.11: Model of PTGS in Pi starvation stress (summary of findings) 112
# LIST OF TABLES

-- Chapter 2 --
Table 2.1: Response-and-recovery classification scheme for loci 36
Table 2.2: Responsive and recovering genes absent from the ATH1 micro-array 39

-- Chapter 3 --
Table 3.1: Micro-RNAs differentially expressed in the response to, or recovery from Pi starvation 76
Table 3.2: List of miRNA:target circuits and the stress conditions they are known to respond to 78
Table 3.3: List of hypotheses assigned to miRNA circuits based on their fold-change profiles 79
Table 3.4: Summary of results for known stress responsive miRNA circuits 80
Chapter 1

Introduction

Justification and summary of research conducted
The state of the global food supply is concerning. The area of arable land is decreasing (Chen et al., 2010); food demand is increasing (Tilman et al., 2011; Chen et al., 2010; Martinez-Porchas et al., 2012); and, billions of US dollars are lost in agricultural product every year due to plant stress (Shulaev et al., 2008). Genetic modification has the potential to alleviate stress-related crop loss. Post-transcriptional gene silencing (PTGS) is a powerful mechanism that can be adapted for use in genetic modification. Therefore, the overriding goal of the research presented in Chapters 2 and 3 was to investigate the role of PTGS in the plant's response to stress. The aim was to identify both novel small non-coding RNA (smRNA) pathways involved in plant stress, and to identify smRNA's functional role in plant stress.

As sessile organisms, plants have developed to survive ever-changing environments. Changes at the genetic and molecular levels reflect the adaptability required of plants to survive harsh conditions (Seki et al., 2002). This adaptability is exemplified by the behavior of smRNAs in response to plant stress. After the discovery of smRNA in 1991 (Wightman et al., 1991), it was found that smRNAs are involved in many plant stress responses (Sunkar et al., 2007; Moldovan et al., 2010a/b; Khraiwesh et al., 2012; Tang et
al., 2012). Thus, genetic modification of smRNA pathways (circuits) has the potential to alleviate crop loss. Of particular interest is the smRNA subclass termed micro-RNA (miRNA), for which several interesting candidates have been identified in stress tolerance research (Zhang et al., 2005; Katiyar-Agarwal et al., 2006; Sunkar et al., 2006; Katiyar-Agarwal et al., 2007). Additionally, PTGS regulates the abundance of messenger-RNA (mRNA) via many distinct miRNAs (Jones-Rhoades et al., 2006; Zhang et al., 2006b; Zhang et al., 2007) during plant development and stress (Jones-Rhoades et al., 2006; Sunkar et al., 2007; Shukla et al., 2008; Ruiz-Ferrer et al., 2009; Khraiwesh et al., 2012; Xia et al., 2012). Improving crop tolerance through genetic modification is an established method of improving crop yields (Uzogara, 2000; Liu and Chen, 2010). Importantly, the development of stress tolerant crops could lead to the agricultural development of previously unusable land. Given the pervasiveness of PTGS across many plant systems and the small number of miRNAs, research and modification of PTGS circuitry could offer a more efficient means of engineering stress tolerant crops than the traditional route of modifying protein coding genes. Therefore, the study of PTGS circuitry in relation to plant stress could help to address global concerns on food security.

In Chapter 3 our study of PTGS is discussed. In particular we analyzed candidate miRNAs for their putative role in phosphate (Pi) starvation stress. We then related Pi starvation PTGS to other stress conditions and found that PTGS could play a significant role in the general plant stress response. A large amount of preliminary work was required to
conduct the analyses presented in Chapter 3. This preliminary work is presented in Chapter 2, where both coding and non-coding genes were analyzed for their response to, and recovery from Pi starvation. The results in Chapter 2 provide the first genome-wide data available for use in studying the root response. Some of the key findings were that the shoot and root responses involved different gene sets; the root recovery process involved several fold more TFs than the initial response to Pi starvation; and, many genes differentially expressed in response to Pi starvation were similarly responsive to other stress conditions (as indicated by AtGenExpress data). The results presented in Chapter 3 used the gene expression data presented in Chapter 2 to analyze the response of miRNA targets. Chapter 2's tiling-array data was also used to infer the expression patterns of trans-acting small-interfering RNA (tasiRNA) during Pi starvation. In Chapter 3, the results show putative miRNA:target circuits active during Pi starvation stress. Several putative, well annotated miRNA:target circuits were discovered, whose target mRNAs showed marked differential expression in response to Pi starvation. These were IPS1:miR399:PHO2, miR399:RS4, miR399:NF-YA10, miR398:CSD1/2, miR2111:TPS11, miR164:NAC6, miR157:TMO7, miR157:PSB28, RPS2:miR169:IPS2, and miR397:LAC2. Moreover, we identified two novel tasiRNA genes as part of the following circuits: NR1:PAP1, and Chr3_1967672:TMO7. After analyzing the functional roles of the genes targeted by the above PTGS circuits, the results presented in Chapter 3 suggested that the local root response to Pi starvation was influenced by the plant's systemic response pathways via PHR1-mediated PTGS. Taken together, from the results presented in Chapter 3, we hypothesized that PHR1-mediated PTGS influences the local
root response by modulating concentrations of reactive oxygen species (ROS). We further hypothesized that miR2111 is capable of influencing shoot to root carbon signaling by silencing TPS11. Thus, the results presented in Chapter 3 provide testable hypotheses for how the two systemic and local responses to Pi starvation communicate.

In the results discussed in Chapter 2, cross talk was observed between Pi starvation and stress conditions known for eliciting oxidative stress. We found that several miRNA and tasiRNA circuits targeted genes modulating ROS concentrations. Recent reviews have highlighted that PTGS is observed across a broad spectrum of biotic and abiotic stress conditions (Khraiwesh et al., 2012). This is an important observation, as the analysis of cross talk and the general stress response is an essential step toward engineering stress tolerant crops. Understanding the network of metabolic changes of plants in response to multiple stress conditions is not a simple task; yet, it is necessary as plants are constantly exposed to multiple, ever-changing environmental conditions. Thus, if we could identify and understand the common genetic elements that govern these stress responses, we could engineer genetically enhanced crops with less trial and error. The results in this research ultimately provided a list of candidate genetic elements capable of mediating cross-talk between stress conditions and the recovery process by modulating metabolites, in particular ROS and sucrose concentrations, during Pi starvation. Therefore, future work should focus on verifying the listed candidates as actual regulators of ROS and sucrose concentrations during stress. Once verified, the
PTGS model presented in Chapter 3 could provide a good basis for studying the effects of metabolites on plant stress, in general.

**Non-coding small-RNA**

The broad aim of the research presented in Chapters 2 and 3 was to investigate the biological function of miRNAs and siRNAs with respect to plant stress. Both miRNAs and siRNAs are subclasses of smRNA, which is a subclass of non-coding RNA (ncRNA).

It is important to note the relative contribution of ncRNA to the functioning of the cell. From the perspective of rRNA, tRNA, and mRNA, only mRNA is diverse across species. This suggests that mRNA and its regulation are of critical importance to defining one organism from another. Transcription factors (TF) are proteins, derived from mRNA, that regulate the transcription of both mRNA and ncRNAs. TFs have therefore been a focus of many research activities as they are an important component of the genetic circuitry guiding the proper development of organisms. Similarly, some ncRNAs have been found to regulate the abundance of mRNA transcripts. However, little is known regarding the function of many ncRNAs. Some ncRNAs have little phenotypic effect, while others take affect only in an organism's off-spring (Miska et al., 2007). The discovery of ncRNAs changed the simple three-class RNA model (m/t/r-RNA) to a multi-class system. Scientists are still discovering and adding to the list of known RNA their classes and subclasses. The mechanisms underlying ncRNA-processing are still being elucidated, and the functions of ncRNAs in biological pathways are still being explored. As such, ncRNAs
have become an important research topic. The discovery of miRNA had a profound impact on molecular biology as it showed that not only was ncRNA a superclass of smRNA and long-ncRNA (lncRNA), but that smRNA had a functional role in different aspects of an organism’s biology.

One of the early discoveries that sparked research of smRNAs was performed on *Caenorhabditis elegans*, where the gene lin14 was found to be repressed by lin4 in lin14's 3'-UTR (Wightman et al., 1991). The *C. elegans* worm was popularized as a model organism largely by the contributions of Sidney Brenner (Brenner et al., 1974). Intimate knowledge of the worms anatomy and development was crucial for Wightman et al.'s 1991 success. It was Lee et al., (1993) that showed a non-protein coding version of lin4 could inhibit lin14 too; moreover, lin4 produced a 60-nucleotide (nt, 60nt) hairpin structure that contained a 25nt smRNA. Reciprocally, Wightman et al., (1993) showed that deletion of lin14's complementary sequence (same sequence as the lin4 25nt smRNA) in the 3'UTR was sufficient to remove lin4's control over lin14. The lin4 and lin14 results demonstrated the first RNA-to-RNA regulation process (Ambros, 2008). Lin4 was also the first smRNA known. However, at the time lin4 could have been a peculiarity of *C. elegans*. But, several studies followed in a remarkably short time span that served to highlight smRNA as a viable endogenous regulator of mRNA across phyla. First, Craig Mello and Andrew Fire (who won the 2006 Nobel Prize for their work) united to describe double-stranded-RNA (dsRNA) constructs as capable of specifically finding and down-regulating a pre-designed target (Fire et al., 1998). Second, plants were known to
acquire immunity to certain viral infections (Covey et al., 1997; Baulcombe et al., 2004), and also to eventually recognize artificially inserted DNA fragments and 'silence' their expression (Lindbo et al., 1993; Longstaff et al., 1993). David Baulcombe proposed that RNA was a key proponent of the observed silencing phenomenon and later went on to show that smRNAs approximately 25nt in length could be found for those mRNAs being silenced (Baulcombe and Mello share the 2006 Nobel Prize in Physiology or Medicine). Third, just over one decade ago another lin4-like smRNA was discovered, namely let7 (Reinhart et al., 2000). Hence, by the year 2000, smRNAs were known to exist and to function endogenously across species.
Lin4 and let7 were later classified as miRNAs. After their initial discovery, many more smRNAs with similar properties to miRNAs were discovered in several different genomes (Ambros, 2008). Some of these early miRNA properties were: they were found...
to be transcribed in clusters under the control of a single promoter; they were often found in the introns of protein coding genes, or scattered across the repetitive regions of centromeric, telomeric, and sub-telomeric regions; and finally, they were transcribed by different RNA polymerases. With respect to the latter property, miRNAs were initially observed to have different RNA-processing pathways as some were capped and polyadenylated, while others where neither (Ambros, 2008). This ambiguity in miRNA biogenesis was later accounted for by adding further smRNA sub-classifications. Today, miRNA biogenesis is well understood (Figure 1.1). Small-RNAs are now classified by considering both their biogenesis and biological function. The next section of this chapter will refer to smRNA biogenesis pathways using miRNAs as an example. The sections thereafter will highlight examples in the public literature of smRNAs and their roles in plant stress.

**Small-RNA biology**

In all smRNA processing, the initial transcript containing the functional smRNA is termed the primary-smRNA (pri-smRNA). The functional smRNA is always termed the mature-smRNA (mat-smRNA) and is processed from a penultimate smRNA termed the premature-smRNA (pre-smRNA). A mat-smRNA is between 18 and 30nt in length. Generally, subclasses of smRNA favor particular nucleotide lengths. In all sub-classes of smRNA the "sm-" prefix can simply be replaced by the specified sub-classes prefix, e.g. "mi-", for miRNA. Thus, the primary miRNA and mature miRNA are written pri-miRNA and mat-miRNA, respectively. The nomenclature of a specific miRNA such as miRNA-399
is written using "miR" to yield "pri-miR399", "pre-miR399", or simply "miR399" to denote the mature form. It should be noted that, as smRNA classes are constantly in flux, the nomenclature remains in development. For miRNA's, however, the nomenclature seems to have stabilized over the last several years. For the purpose of this thesis, the nomenclature described above will be maintained.

For many smRNA classes, the processing from pri-smRNA to mat-smRNA is class specific. In other words, a specific biogenesis pathway produces a specific class of smRNA. In the case of miRNA, it was initially found that the nuclear enzyme Drosher (in animals) and Drosher like protein 1 (DCL1, in plants) is responsible for the initial cleavage of the pri-miRNA -- this is the first RNA-processing step toward a mature miRNA (Lee et al., 2003). The DCL1-dependent processing of pri-miRNA results in a 60nt to 70nt long pre-miRNA (independent of the initial pri-miRNA's size). This intermediate pre-miRNA translocates to the cytoplasm as a tightly coiled dsRNA hairpin structure. In the cytoplasm the protein complex known as DICER, which contains both a helicase and catalytic subunit, is responsible for unwinding the hairpin structure, locating the ~22nt mat-miRNA sequence, and excising it as a short double-stranded moiety (dsmat-miRNA, or ds-miR) (Bernstein et al., 2001; Bartel, 2004). The ds-miR is then passed to RISC, another protein complex that chooses a single strand as a "guide" strand, and discards the other. Nothing is currently known about how the guide strand is selected other than it is a non-random process (Darnell et al., 2011). But, the discarded strand, the "passenger" strand (denoted as miR*), is mostly found to be non-functional. The guide-strand (mat-miRNA)
is, however, retained within the RISC complex and used to find, by complementarity, mRNA sequences targeted for site-directed cleavage (Liu et al., 2004). The miRNA pathway is therefore a post-transcriptional mechanism for mRNA-silencing (i.e. PTGS).

PTGS appears to mimic the function of TFs as a regulator of transcript abundance. Thus, one relevant question asks why both the PTGS and TF systems evolved. Clues to the answer to this question may lie in plant PTGS, which has both an endogenous and exogenous role. The exogenous role of PTGS is involved in viral defense where RNA viruses are directed through smRNA pathways to mediate the silencing of viral-born transcripts, thus acquiring immunity (Covey et al., 1997; Baulcombe, 2004). Hence, endogenous PTGS pathways may have evolved by co-opting the molecular machinery used for viral defense.

In plants, the production of miRNA remains largely the same as described above. But, peculiarities in smRNA biogenesis do exist across species. For instance, plants do not contain the piwi-RNAs (piRNA) found in Drosophila, but do produce trans-acting small-interfering RNAs (tasiRNA) (Vazquez, 2006). The following sections of this chapter will introduce several relevant plant smRNAs and their role in plant stress. While reading these sections please refer to Figure 1.2.
The role of miRNA post-transcriptional gene silencing

In animals, miRNAs are transcribed in clusters, whereas in plants miRNAs are transcribed separately (Bartel, 2004). Other than this difference, pri-miRNAs are still initially processed in the nucleus of plant cells before being exported to the cytoplasm by the HASTY (HST) transporter (Papp et al., 2003; Park et al., 2005). Micro-RNA biogenesis is a 2-stage process in Arabidopsis, which requires the consecutive action of two DICER-like 1 protein (DCL1) mediated cleavage events (Kurihara, 2004). In addition, the DCL1 partner protein, hyponastic leaves 1 (HYL1) -- a dsRNA-binding protein (DRB) -- is required for interaction with the dsRNA hairpin structure (Vazquez and Gasciolli et al., 2004; Kurihara et al., 2006). The exported pre-miRNA is then taken up by the RNA-induced silencing complex (RISC), which discards the passenger strand in favor of the guide strand that is then used to target mRNA for site-specific cleavage (Llave et al., 2002). Importantly, there is a large difference between plant and animal PTGS in terms of the degree of complementarity required between miRNA and target. Plant miRNAs have greater complementarity to their target mRNAs and mediate mRNA-cleavage and degradation (Bartel, 2004), whereas animal miRNAs are less complementary and tend to inhibit mRNA translation (translational repression) (Carrington and Ambrose, 2003; Pillai, 2005). Another important consideration is that miRNAs are sorted between RISC complexes depending on the particular ARGONAUTE1 protein (AGO1) it contains (Vaucheret et al., 2004; Baumberger and Baulcombe, 2005; Qi et al., 2005; Kim, 2008). Hence, there is a network of miRNA-processing that depends on the protein composition of the cell, which in turn depends on mRNA abundance. Therefore, it is
easy to imagine the possibility of feedback and feed-forward loops that involve miRNA intermediates. Indeed, miRNAs have been observed to promote and repress downstream targets respectively. An example of miRNAs promoting downstream targets can be found in the signal-boosting phenomenon that occurs during miRNA-mediated tasiRNA biogenesis (see the next section on tasiRNAs). In the following paragraphs, a description of miRNA-mediated repression is presented with particular focus on the role of miRNA in endogenous plant processes.

Over 100 different miRNAs in Arabidopsis regulate the translation of many mRNA targets during plant development (Jones-Rhoades et al., 2006; Zhang et al., 2006b; Zhang et al., 2007). Even environmental stimuli have been shown to result in the differential regulation of several miRNA species (Jones-Rhoades et al., 2006; Zhang et al., 2006b; Zhang et al., 2007). Systemic aberrations in the miRNA pathways are known to lead to developmental deformities, as observed in loss-of-function (LOF) dcl1 mutants (Liu et al., 2005; Kurihara et al., 2006). HST, the miRNA nuclear exporter, is also required for proper miRNA function (like DCL1). And, like DCL1, mutant hst leads to aberrant leaf formation, improper floral transition timing, and increased sterility (Bollman et al., 2003). These results all suggest the presence of individual miRNAs in pathways regulating each of the affected developmental programs. Over the years researchers have identified specific process-related miRNA species.
Micro-RNAs in somatic leaf development: The highly conserved (Zhang et al., 2006a) miRNAs (miR165 and miR166) target three homeodomain TFs important in leaf development (Juarez et al., 2004): PHABULOSA (PHB), PHAVOLUTA (PHV), and REVOLUTA (REV). Importantly, both LOF mutants of the TFs (McConnell et al., 2001; Emery et al., 2003) and aberrant miR165 and miR166 expression (Juarez et al., 2004) lead to similar leaf morphologies. These sets of evolutionary studies, coupled with LOF mutants and reverse genetics approaches, have tied miR165 and miR166 to the endogenous development of leaf tissues. In a similar study, miR159 was also linked to leaf development through the regulation of TFs (Palatnik et al., 2003). However, miR165 and miR166 were shown not to affect miR159 targets (Palatnik et al., 2003). Taken together, miRNAs may function in distinct molecular pathways that converge to form the same phenotype.

Micro-RNAs in other developmental patterns: Root, shoot, floral, and embryo development have all been shown to be regulated by TFs of the type NAM/ATAF/CUC(NAC) (Aida et al., 1997; Takada et al., 2001; Hibara et al., 2003) -- including auxin response factors (ARF2-4/10/16/17) involved in root patterning (Mallory et al., 2005; Sorin et al., 2005; Williams et al., 2005a; Yang et al., 2006a). These TFs are shown to be affected by aberrant miR164 expression (Guo et al., 2005), which is coupled to abnormalities in the developmental programs listed above (Laufs et al., 2004; Mallory, 2004a; Guo et al., 2005).
Micro-RNAs in transitioning from juvenile to adult: The juvenile to adult phase change in plants is marked by the development of sexual organs and mitosis; it is therefore worth considering separately from somatic developmental changes. In plants, flowering time may be altered to produce early-transitioning adults by down-regulating APETALA-2 protein (AP2), a regulator of floral-timing and floral-patterning (Lohmann and Weigal, 2002). Subsequently, it has been shown that miR172 targets AP2 and that aberrant miR172 alters floral-timing and -morphology in a predictable manner (Aukerman and Sakai, 2003; Chen, 2004; Lauter et al., 2005). The miRNA miR171 is almost exclusively transcribed in inflorescence and floral tissues, and is known to be involved in floral development along with miR159, and miR156 (Llave et al., 2002; Achard et al., 2004; Schwab et al., 2005). It is worth noting that miR159 has multiple roles and was previously reported to be involved in leaf development as well (Palatnik et al., 2003). In floral tissues, miR159 targets LEAFY (LFY3) to affect another development and floral-timing (Achard et al., 2004). Hence, two comments are worth making: Firstly, miR159 is not exclusively somatic nor is it specific to floral parts; and secondly, miRNAs are capable of having multiple functions in a tissue-dependent and/or time-dependent manner. Therefore, the complexities and rules governing mRNA regulation are likely transferable to miRNA regulation. Indeed, like mRNAs, miRNAs are transcribed by DNA-dependent RNA polymerase II (Pol-II) (Bartel, 2004).

Micro-RNAs in environmental signaling pathways: The economic impetus behind stress research in plants (agronomy) is massive -- reports have shown a 20 to 30 percent
decrease in crop yield, or even total loss, due to environmental factors (Zhang et al., 2000; Mahajan and Tuteja, 2005; Vinocur and Altman, 2005; Yamaguchi and Blumwald, 2005). Environmental factors include a wide range of phenomena that can mostly be classified into one of three categories: nutritive, abiotic, and biotic. Currently, several miRNAs are becoming interesting candidates for stress tolerance research in plants (Zhang et al., 2005). In particular, miR395 abundance has been shown to respond to sulfate starvation, a nutritive stress (Zhang, 2006b). Similarly, for abiotic stress, the abundance of miR319 and miR402 transcripts have both been shown to respond to cold, whereas miR402 has also been shown to respond to drought and salt stress (Sunkar et al., 2004). This demonstrates cross-talk between miRNA signaling pathways and suggests the possibility of miRNA-redundancy in stress responsive pathways. The majority of these environmentally associated miRNA's are based on reports inferring cause from expression patterns. Micro-RNA399 is one candidate whose interaction-circuit is well understood in nutrient stress. In phosphate starvation, miR399 is up-regulated alongside the phosphate transporter PHT1;8/9 (Fujii et al., 2005). Micro-RNA399 targets PHO2, which itself represses PHT1;8/9 via ubiquitination (Bari et al., 2006). Hence, miR399 silences PHO2 activity, thus preventing PHT1;8/9 ubiquitination and leading to root to shoot translocation. However, miR399 expression is slightly delayed in order to allow for root tissues to over-accumulate phosphate reserves (Raghothama, 1999; Abel et al., 2002; Bari et al., 2006).
Figure 1.2: Detailed biogenesis pathways of several small-RNA classes in *A. thaliana*

The above examples highlight the diverse set of roles to which miRNAs have evolved. This last point, taken together with the fact that there are far fewer miRNAs than actual protein coding genes suggests that miRNAs are much more versatile and adaptable than
their protein coding counterparts. Later, in Chapter 3, we investigate miRNA versatility in terms of its roles in plant stress. We also investigate tasiRNAs in the same regard, which will be introduced in the next and final section of this chapter.

**The role of tasiRNA post-transcriptional gene silencing**

Trans-acting small-interfering RNAs (tasiRNAs) follow a similar production pipeline to miRNAs: non-coding regions are transcribed by Pol-II, processed, and eventually used to guide the cleavage of target mRNAs (Vazquez et al., 2004; Peragine et al., 2004). However, multiple tasiRNAs are transcribed as a single sense-strand of a so-called "TAS" locus and processed into diverse sets of siRNAs (Vazquez, 2006). This sharply contrasts with plant miRNAs, which are usually each transcribed uniquely (unlike animal miRNAs, which are also transcribed in clusters on a single transcript). The exact role of tasiRNA is still unclear (Vazquez et al., 2004; Peragine et al., 2004; Allen et al., 2005; Xie et al., 2005; Yoshikawa et al., 2005), but this is hardly surprising given the current focus on miRNA biogenesis and function. The biogenesis of tasiRNA is also altered in miRNA-associated mutants. This is due to the requirement of miRNA-guided AGO1 cleavage of tasiRNA progenitors -- therefore tasiRNA biogenesis is downstream of miRNA biogenesis (Allen et al., 2005; Yoshikawa et al., 2005). AGO1 cleavage then initiates synthesis of the progenitors complementary strand by RNA-directed RNA polymerase 6 protein (RDR6); oddly enough, synthesis occurs in a 3'-direction for TAS genes TAS1 and TAS2, but in a 5'-direction for TAS3 (Vazquez et al., 2004; Peragine et al., 2004; Allen et al., 2005). This newly synthesized long dsRNA is then cleaved by a DCL1-family member, DCL4 (similarly to DCL1 for miRNA biogenesis). The unique, and interesting, part about DCL4-mediated
cleavage is that it excises many small interfering RNAs (siRNA) in a phased manner across the entire tasiRNA progenitor (Vazquez et al., 2004; Gasciolli et al., 2005). Some evidence for co-operation between DCL4 and DRB family members exists -- as manipulation of dsRNA is a requirement, this protein-protein interaction seems plausible (Hiraguri et al., 2005). The resulting double stranded tasiRNAs are then shunted to the canonical miRNA RISC-like pathway for mRNA targeting and degradation. In this way, it is conceivable that tasiRNAs can self-perpetuate beyond the initial miRNA guided cleavage of the tasiRNA progenitor -- this mode of genesis may have implications in viral defense, as noted below.

**Trans-acting siRNA in regulating auxin response factors and mediating viral defense:** A relevant example of tasiRNA biology is observed in the regulation of auxin response factor proteins ARF3 and ARF4. These two ARFs serve important roles during development (Allen et al., 2005; Yoshikawa et al., 2005). Actually, the miRNA pathway mediated by miR164 has been described (see above) to target these two ARFs during root development. The tasiRNA, tasiR-ARF, which targets ARF3 and ARF4, is well conserved throughout the plant kingdom suggesting that the tasiRNA pathway serves an important function across species (Allen et al., 2005; Yoshikawa et al., 2005). Several other lines of evidence support the notion of tasiRNA mediated immunity against RNA viruses, which must produce dsRNA during their life cycle. Coupled with a single miRNA-guided cleavage event, the tasiRNA system may offer persistent immunity against the
invading virus without the host needing to continuously spend resources on producing pri-miRNA.

**Conclusion**

Biologists are discovering PTGS pathways in many biotic and abiotic plant stress responses. PTGS is becoming an increasingly relevant mode of transcriptional regulation in plant stress. Moreover, PTGS is now a viable means for increasing stress tolerance in our staple crops. In Chapters 2 and 3, research focuses on Pi starvation stress. However, as PTGS appears to be involved in several stress responses, effort is given to relate the Pi starvation results to other stress conditions. As such, Chapter 2 describes the analysis of cross talk among protein coding genes in 20 AtGenExpress datasets; and, Chapter 3 presents an analysis that was guided by previously defined networks of stress responsive PTGS circuits. The combined results of Chapters 2 and 3 suggest that while PTGS is involved in exclusive Pi starvation pathways such as miR399:PHO2, PTGS is also involved in the general stress recovery process by mediating metabolites such as ROS and sucrose.
-- Chapter 2 --

The molecular response to, and recovery from phosphate starvation

Introduction
This chapter describes the response and recovery of genes in root and shoot of Arabidopsis to phosphate (Pi) starvation stress. The results are published (Woo and MacPherson et al., 2012) and also describe the first high-throughput experiments regarding the root's response to Pi starvation. Moreover, the tiling-array data is presented for the first time for plant Pi starvation and provides information on genomic regions previously only studied in low-throughput experiments.

This chapter generated results needed for Chapter 3, which deals with the role of non-coding RNA in plant stress. The Pi starvation response provides good perspective into the role of non-coding RNA in plant stress, as only one miRNA-circuit has been thoroughly researched (IPS1:miR399:PHO2) (Bari et al., 2006), yet smRNA deep-sequencing results suggest that many more miRNAs are involved (Chen et al., 2007). Small-RNA deep-sequencing data was also collected, but not presented here (see Chapter 3).
Even though the focus of this thesis is on non-coding RNA, the study of Pi starvation has important applications in agriculture and water conservation. A common practice in modern agriculture to overuse Pi fertilizers is contributing to the contamination of waterways and disruption of natural ecosystems (Moss, 2008; Chebud, Naja and Rivero, 2011). The study of plant Pi metabolism and its underlying genetic pathways is key to discovering methods of efficient fertilizer usage. The work presented here describes a genome-wide resource on the molecular dynamics underpinning the response and recovery in roots and shoots of Arabidopsis thaliana to Pi starvation.

Genome-wide profiling by microarray and tiling-array revealed minimal overlap between root and shoot transcriptomes suggesting two independent Pi starvation pathways. Significant gene expression patterns were detected for over 1000 candidates and displayed initial, persistent, and latent responses to Pi starvation. Comparisons to AtGenExpress identified cohorts of genes affected across multiple stress and hormonal stimuli. Indeed, abscisic acid (ABA) treatment displayed a dominant role in affecting many Pi starvation responsive candidates. Analysis of gene expression across 20 different treatments enabled determination of genes more specifically affected in the Pi starvation response as compared to their family members. Indeed, the analysis revealed that PHR1-regulated members of closely related Pi responsive families (PHT1;1, PHT1;7–9, SPX1-3, and PHO1;H1) were not differentially expressed (DEX) in other stress treatments; moreover, members of the same family not known to be PHR1-regulated did respond to other stress treatments.
As a genome-wide study, many genes were identified that were previously unknown to respond to and/or recover from Pi starvation. These results are likely to describe the highest resolution of genome-wide data related to Pi starvation in Arabidopsis available to the community to date (Woo and MacPherson et al., 2012). The accession code, GSE34004, may be used to access the microarray and tiling-array data from the Gene Expression Omnibus (GEO).

**Results**

**General data observation**

In microarray data, DEX genes were detected during the response (Pi\textsuperscript{starv}/Pi\textsuperscript{mock}) and recovery (Pi\textsuperscript{replete}/Pi\textsuperscript{starv}) in shoot and root. Principal component analysis (PCA) enabled observation along the 3rd principal component, variations between the sample groups that clearly distinguish Pi starvation treatment (Pi\textsuperscript{starv}) from mock treatment (Pi\textsuperscript{mock}) and Pi\textsuperscript{starv} from recovery (Pi\textsuperscript{replete}). Along the 2nd principal component root and shoot samples are clearly distinguished (Figure 2.1a,b). On the other hand, the Pi\textsuperscript{mock} and Pi\textsuperscript{replete} sample groups are not clearly separated in either shoot or root data, which demonstrates that the three-days recovery period is sufficient to reverse the effects of Pi\textsuperscript{starv} on the transcriptome to a great extent.
**Figure 2.1:** Initial assessment of data and quality control
Subfigure (a) and (b) summarizes all gene expression data from 18 micro-array’s by making use of a principal components analysis (PCA). (a) The amount of variation/information captured (blue) by each principle component (PC) is combined to obtain the total variation (black). (b) A plot of the two most biologically informative PCs, PC2 and PC3. Root (red) and shoot (green) organs are shown to separate along the x-axis (PC2), whereas samples for mock (diamond), treatment (square), and recovery (triangle) separate along the y-axis (PC3). (c) Dot plot of log₂ fold-change values for loci captured by micro-array (x-axis) and tiling-array (y-axis) technologies. The root response (top-left, blue) and recovery (top-right, red) each have $R^2$ values above 0.7, whereas the shoot response (bottom-left, green) and recovery (bottom-right, purple) each have $R^2$ values above 0.65.

**Novel Pi-signaling patterns observed among genes involved in phosphate starvation**

- Gene expression patterns in shoot and root

Previous studies have identified gene sets that initially respond to (Misson et al., 2005; Müller et al., 2007; Nilsson et al., 2010) or recover (Morcuende et al., 2007) from Pi starvation. However, they are mutually independent from one another and rarely agree in expression at the individual gene level. The advantage offered by this study is the combination of those previously used approaches (Misson et al., 2005; Müller et al.,...
2007; Nilsson et al., 2010; Morcuende et al., 2007) within one experimental design. This helps to more easily compare different studies. Moreover, for some genes, expression was observed to persist during the three-day recovery period, while other genes latently responded to Pi starvation during recovery. These persistent and latent responses have not been reported before and are therefore a source of interest.

We begin by considering the results obtained from gene loci present on the ATH1 platform as those have been most well characterized. Afterward, the expression of genes detected exclusively by tiling-array is analyzed. We note that given the exclusive use of ATH1 gene-chips in the previous Pi starvation research, results from the tiling-array are new, provide new insights into Pi starvation, and make a rich resource for future analyses.

A total of 1,257 genes were differentially expressed in roots during response or recovery. Table 2.1 highlights the criteria used to sub-classify these 1,257 genes. A large proportion (25%) of root-responsive genes increased expression during the response phase before returning to basal levels in the recovery phase (labeled "IPR", Figure 2.1e,f,g,h). Approximately 8% of the genes decreased expression in the response phase, but returned to the basal level in the recovery phase (labeled "INR"). A small portion of the genes continued to be DEX during recovery (labeled "PPR" or "PNR"). Over half the genes (58%) responded during the recovery phase only, suggesting that the conditions
for the latent response are set up during starvation. We grouped these genes into the "Latent Positive Response, LPR" (420 genes) or "Latent Negative Response, LNR" (320 genes) classes. The latent response totaled 740 genes and therefore seem of equal importance to the Pi starvation response as genes of the initial and persistent starvation classes (these together consists of combined 517 genes).

**Table 2.1: Response-and-recovery classification scheme for loci**

<table>
<thead>
<tr>
<th>Class</th>
<th>Code</th>
<th>Response</th>
<th>Recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basal Response</td>
<td>BAR</td>
<td>no change</td>
<td>no change</td>
</tr>
<tr>
<td>Continuous Positive Response</td>
<td>CPR</td>
<td>≤ 0.001</td>
<td>+</td>
</tr>
<tr>
<td>Initial Positive Response</td>
<td>IPR</td>
<td>≤ 0.001</td>
<td>+</td>
</tr>
<tr>
<td>Persistent Positive Response</td>
<td>PPR</td>
<td>≤ 0.001</td>
<td>+</td>
</tr>
<tr>
<td>Latent Positive Response</td>
<td>LPR</td>
<td>no change</td>
<td>≤ 0.001</td>
</tr>
<tr>
<td>Continuous Negative Response</td>
<td>CNR</td>
<td>≤ 0.001</td>
<td>-</td>
</tr>
<tr>
<td>Initial Negative Response</td>
<td>INR</td>
<td>≤ 0.001</td>
<td>-</td>
</tr>
<tr>
<td>Persistent Negative Response</td>
<td>PNR</td>
<td>≤ 0.001</td>
<td>-</td>
</tr>
<tr>
<td>Latent Negative Response</td>
<td>LNR</td>
<td>no change</td>
<td>≤ 0.001</td>
</tr>
</tbody>
</table>

By table headings: (Class) All loci are binned into 1 of 9 response-and-recovery classes; (Code) Classes are given a 3 letter code for convenience. (Response) The p-value and log₂ fold-change required by a locus’ response in order to qualify for the respective class. (Recovery) The p-value and log₂ fold-change required by a locus’ recovery in order to qualify for the respective class. A “no change” is indicated if the p-value is not less than or equal to 0.001. (Cartoon Representation) A cartoon diagram indicating the general expression pattern captured by the class.
In shoot ATH1 data, we classified 182 genes into response-and-recovery classes (Table 2.1, Figure 2.1a,b,c,d). Approximately 38% of these genes were classified as IPR. Briefly, we classified the remaining 112 genes as PPR (18 genes), PNR (2 genes), LPR (23 genes), and LNR (69 genes). We observed that shoot tissues initially up-regulated 48% of all responsive genes and the remainder were mostly latently down-regulated. Hence, shoot tissues show a distinct shift in gene expression patterns. It is interesting to observe that the number of latently responsive genes (92 genes) was similar to those of the initial and persistent responses combined (90 genes).

- **Systemic and organ-specific gene expression**
  The data above shows that the plant transcriptome systemically responds to Pi starvation as we observe significant gene expression changes in both root and shoot. Moreover, Pi starvation gene expression profiles change over time, but definite initial and persistent responses could be observed. Despite the systemic response of Arabidopsis to Pi starvation, only 89 genes were differentially expressed in both shoot and root data (Figure 2.2i-all rows and columns except the first). The 89 genes may be involved in the systemic Pi starvation response as 60% of these responders were differentially expressing in the same manner, in both organs (Figure 2.2i-diagonal).

  This small overlap between gene expression in two organs indicates the presence of two distinct organ-specific gene regulation pathways active during Pi starvation. However,
considering that the root response contributed 87% of the combined Pi starvation responsive genes in both organs, we focused our functional analyses on the root data.

In the following section, tiling-array specific results are presented. Then, our functional analysis on DEX genes in Pi starvation and recovery is presented.

The response and recovery of genes absent from the ATH1 microarray platform
The advantage of the 1.0R tiling-array over the ATH1 microarray is the possibility to observe expression of transcripts derived from genomic regions not covered by ATH1’s optimized probes. TAIR8 has annotations for 26,956 protein-coding genes of which, 21,912 (81%) genes are represented on ATH1.

All genes represented by less than three probes on the 1.0R tiling-array were excluded, leaving 4,730 genes for further analysis. The transcript abundance of genes in each treatment was calculated using the same median-polish methodology as used in the microarray analysis. The response of all genes (G) was assessed after normalizing data so that standard deviation (s) equals one. A gene \( k \) was declared as differentially expressed if \(|G_k/s| >= 3\). This ensured that only the most differentially expressed genes (top ~1.5% in the root response) were considered DEX.
Genes absent from ATH1 were classified into one of nine classes in the manner described by Table 2.1. This method classified 477 out of the 4,730 tiling-array specific genes as DEX (Table 2.2); 171 genes were root specific, 242 were shoot specific, and 64 were systemically (commonly) DEX. Details on the response and recovery for all 4,730 genes can be found in Supplemental data of Woo and MacPherson et al., (2012).

### Table 2.2: Responsive and recovering genes absent from the ATH1 micro-array

<table>
<thead>
<tr>
<th>Identifiers</th>
<th>Root</th>
<th>Shoot</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Systemically Regulated</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G01060</td>
<td>CPL3 / ETC3</td>
<td>4.291</td>
</tr>
<tr>
<td>AT5G43300</td>
<td>GDPD3</td>
<td>4.272</td>
</tr>
<tr>
<td>AT5G09922</td>
<td>IPS3</td>
<td>5.971</td>
</tr>
<tr>
<td>AT4G19038</td>
<td>LCR15</td>
<td>1.376</td>
</tr>
<tr>
<td>AT3G61172</td>
<td>LCR8</td>
<td>0.412</td>
</tr>
<tr>
<td>AT2G14240</td>
<td>BIHLH100</td>
<td>-2.631</td>
</tr>
<tr>
<td>AT2G40705</td>
<td>LCR43</td>
<td>1.420</td>
</tr>
<tr>
<td>AT1G75607</td>
<td>LCR65</td>
<td>-2.178</td>
</tr>
<tr>
<td>AT5G69790</td>
<td>ORG2</td>
<td>-1.625</td>
</tr>
<tr>
<td>AT3G61177</td>
<td>LCR58</td>
<td>1.963</td>
</tr>
</tbody>
</table>

| **Root Specific** | | |
| AT1G3165 | CLE4 | -1.977 | 1.924 | INR | -0.107 | -0.097 | BAR |
| AT2G31081 | CLE4 | -2.012 | 2.875 | INR | 1.168 | -0.746 | BAR |
| AT2G31082 | CLE7 | -1.952 | 2.312 | INR | 0.111 | 0.121 | BAR |
| AT5G24920 | GD13 | -2.877 | 3.590 | INR | 0.217 | 0.223 | BAR |
| AT3G69085 | LCR44 | -1.885 | 2.264 | INR | 1.002 | -0.650 | BAR |
| AT4G19970 | LSE3 | -2.308 | 1.684 | INR | 0.613 | 0.096 | BAR |
| AT4G18197 | PUP7 | -1.770 | 1.755 | INR | 0.544 | -0.300 | BAR |
| AT4G34760 | AGL58 | 1.941 | -2.146 | IPR | -0.136 | -0.280 | BAR |
| AT5G69095 | CYP712A2 | 2.089 | -2.979 | IPR | -0.110 | -0.599 | BAR |
| AT3G30723 | GD16 | 3.627 | -5.125 | IPR | -0.030 | -0.020 | BAR |
| AT2G29660 | PT14 | 1.810 | -1.706 | IPR | 1.185 | -1.223 | BAR |
| AT3G94000 | PLL3 | 2.523 | -2.929 | IPR | 0.978 | -0.054 | BAR |
| AT4G27920 | RCR4 | 2.259 | -2.132 | IPR | 0.734 | -0.431 | BAR |
| AT3G33130 | GRI | -0.373 | -1.756 | LNR | -0.056 | -0.157 | BAR |
| AT3G61182 | LCR54 | -0.226 | -1.759 | LNR | -0.763 | -0.202 | BAR |
| AT2G14365 | LCR84 | 0.519 | -1.936 | LNR | 0.756 | -0.101 | BAR |
| AT4G10115 | SCR1290 | 0.548 | -2.085 | LNR | -0.297 | -0.485 | BAR |
| AT4G29305 | LCR25 | -1.368 | 2.295 | LPR | 0.973 | -0.521 | BAR |
| AT4G67646 | RAP2.9 | -1.617 | 3.125 | LPR | -0.404 | 0.512 | BAR |
| AT4G23170 | EP1 | -2.487 | 0.662 | PNR | -0.033 | -0.134 | BAR |
| AT4G69795 | LCR13 | -2.395 | 1.119 | PNR | -0.939 | 0.163 | BAR |
| AT3G23167 | LCR9 | -1.694 | 0.502 | PNR | 0.556 | -1.297 | BAR |
| AT4G18195 | PUP8 | -1.820 | 1.335 | PNR | -0.136 | -0.001 | BAR |
| AT3G23715 | SCR113 | -2.088 | 0.971 | PNR | 1.289 | 0.501 | BAR |
| AT2G20825 | ULT2 | -1.640 | 0.337 | PNR | 0.339 | 0.139 | BAR |
| AT5G45105 | ZIP8 | -2.392 | 0.846 | PNR | 0.284 | -1.101 | BAR |
| AT4G22210 | LCR85 | 1.664 | 0.028 | PPR | 0.399 | -0.258 | BAR |
| AT4G68015 | R44F17 | 1.728 | -0.550 | PPR | 0.315 | 0.319 | BAR |

| **Shoot Specific** | | |
| AT4G74510 | AT5PTASE11 | 0.182 | -0.184 | BAR | -1.362 | 1.610 | INR |
| AT1G66145 | CLE38 | -0.570 | -0.135 | BAR | -1.422 | 1.680 | INR |
| AT4G11485 | LCR11 | -0.969 | -1.131 | BAR | -1.689 | 1.849 | INR |
| AT4G13890 | SHM5 | -0.177 | -0.165 | BAR | -1.534 | 1.421 | INR |
| AT4G10767 | SCR23 | -0.231 | -0.470 | BAR | 1.459 | -2.258 | IPR |
| AT2G30432 | TCL1 | 0.062 | -0.924 | BAR | 1.680 | -1.393 | IPR |
| AT4G98280 | LBD2 | -0.667 | -0.194 | BAR | 0.729 | -1.387 | LNR |
| AT4G29280 | LCR22 | -0.461 | -0.171 | BAR | 0.485 | -1.779 | LNR |
| AT3G64340 | NAC049 | 1.550 | -0.809 | BAR | 0.731 | 1.440 | LNR |
| AT5G03417 | RAZF13 | -0.093 | -0.201 | BAR | 1.362 | 2.259 | BAR |
| AT5G45873 | SCR27 | 0.864 | 0.235 | PPR | 0.754 | -1.712 | LNR |
| AT4G31380 | FLP1 | 0.549 | -0.073 | BAR | -0.234 | 1.555 | LPR |
This table presents 71 of the 477 tiling-array-specific genes identified as differentially expressed during the Pi\textsuperscript{starv} response and/or recovery. These 71 were chosen only for the reason that they have been annotated with a gene symbol. The remaining 406 genes do not have gene symbols and are poorly annotated. See Materials and Methods for a complete list. *: Abbreviations for response (Res.) and recovery (Rec.).
**Figure 2.2:** The distribution of probe-sets across tissues and response-and-recovery classes

Subfigures (a), (b), (e), and (f) display the relative distribution (no y-axis) of fold-change values (log₂) for significantly (black curve) and non-significantly (grey curve) regulated probe-sets in shoot (a, b) and root (e, f) for both response (a, e) and recovery (b, f). (c, g) Dot plot of fold-change values for response (x-axis) against recovery (y-axis) from shoot (c) and root (g) samples. Probe-sets tend to recovery as is evident by their trend (line of full recovery, dotted red line, y=-x). Significantly regulated probe-sets are colored according to their response-and-recovery classification (Table 2.1). (d, h) Histogram of probe-set counts (y-axis) according to response-and-recovery classes (Table 2.1) for shoot (d) and root (h). (i) Table showing the intersection between shoot (columns) and root (rows) response-and-recovery classes. The top leftmost cell being a count of probe-sets neither responsive in shoot nor root. The first column and row being counts of loci uniquely regulated by root and shoot, respectively. The diagonal being counts of ubiquitously regulated probe-sets. Positive and negative regulation is denoted by a blue and red background, respectively. Whereas, probe-sets differentially regulated between shoot and root are denoted using white backgrounds.

**Analysis of functionally related genes**

Gene annotations found in the GO-SLIM ontology (http://www.geneontology.org/) were used to calculate p-values for all DEX genes in both root and shoot data (1,350 genes).

We found that "cellular response to phosphate starvation" was the most significant (Bonferroni adjusted p-value <= 3.81x10⁻¹³); 13 of the 21 genes annotated by this term were identified in both organs and mostly exhibited an IPR pattern. Other Pi starvation related terms such as "galactolipid biosynthetic process" and "sugar:hydrogen symporter activity" were also found significant.

In this section we aimed to characterize DEX gene changes in root gene expression, using their annotations found in GO-SLIM. We first analyze genes previously described to be involved in the Pi starvation response, and then move on to characterize genes
that have not been known as Pi starvation responsive. In both cases, this study is the first to show both the response and recovery of expression of these genes.

**Expression patterns of known Pi starvation responsive genes in roots and shoots**

Through a literature survey we identified 84 genes to be involved in Pi starvation (Woo and MacPherson et al., 2012). We clustered these 84 genes in terms of their functional annotations (Figure 2.3-heatmap) and identified eight groups of interest. These clusters represented ‘transcription’, ‘plastid metabolism’, ‘response to wounding’, ‘Pi-transport’, ‘anthocyanin biosynthesis’, and ‘galacto-lipid biosynthesis’ and ‘glyco-lipid biosynthesis’. Figure 2.3a-h shows expression patterns for each 84 genes in each of the eight clusters. Approximately 30% of the 84 genes were IPR in roots, while 60% were non-responsive. This was not surprising given that these genes were identified from the results of many distinct studies. Similar results were obtained when we performed the same analysis on gene expression levels in shoots. Comparison of the two organs showed that 14 genes (ATPAP1, ATPAP17, PHT1;4, PHT1;7, PHO1;H1, PHF1, PLDP2, MGD2, MGD3, SQD1, SQD2, SPX1, SPX2, and SPX3) shared the IPR patterns in both organs. The majority of the remaining IPR genes were non-responsive in shoots and therefore represent part of the root specific response to Pi starvation (At3g03530, DGD2, PAP6, PHT1;3, PHT1;1, PHT1;8, PHT1;9, PHT1;5).

The genome-wide gene expression data shows 1,231 DEX genes, making them putative Pi starvation responsive genes. These 1,231 Pi starvation responsive genes did not
include those 84 genes identified by our literature survey. As given by Table 2.1, these putative Pi starvation responsive genes were distinguished from one another by initial, persistent, or latent expression patterns. These three classes of gene expression patterns describe Pi availability related to our experimental design and were therefore appropriate gene groups to use in our functional analysis. The following part will describe the results of the functional analysis on these three gene expression classes, bearing in mind that each gene considered is a putative Pi starvation responsive gene candidate.
Figure 2.3: Analysis of gene expression and function among known Pi responsive loci
(Top) Clustered heatmap of known Pi-responsive loci (rows) and GO-SLIM annotations (columns). Each color represents a different functional annotation and colors are only displayed, per row: if the rows gene is annotated by the GO-SLIM term. Selected clusters are marked from a-to-h. (Bottom, a-h) Transcription profiles of each functional cluster.
displaying the relative abundance of transcripts for particular known $P_i$-responsive genes in mock $(M)$, treatment $(T)$, and recovery $(R)$ samples.

Three gene expression phases uncover functional responses to Pi starvation

By definition, for a gene to be classified as either initially, persistently, or latently responsive, it had to be DEX in either the response or recovery phase. In this section we focus on root data and candidate genes involved in the Pi starvation response. To summarize the root data (1,231 candidate Pi starvation responsive genes): the initial response phase consisted of 292 positively (IPR) and 112 negatively (INR) responsive genes; 45 persistent positively (PPR) and 48 negatively (PNR) responsive genes; and, 423 latent positively (LPR) and 311 negatively (LNR) responsive genes.

Genes in the initial, persistent, and latent phases were annotated with GO-SLIM. tMeV software (Zhang and Forde, 1998; Folter, 2005) was used to cluster functionally similar genes in each phase and produce the heatmaps shown in Figure 2.4a,e,i. Expression data was overlapped with each cluster to determine which displayed the greatest differential response to Pi starvation. For each phase, the three most DEX clusters were selected for display: initial (Figure 2.4b-d); persistent (Figure 2.4f-h); and, latent (Figure 2.4j-l).

Initially responsive gene clusters were generally annotated by GO categories "transcription regulation", "ionic transport", and "transmembrane transport".
Figure 2.4: Analysis of gene expression and function among novel Pi responsive loci
(Top, a, e, i) Clustered heatmaps of novel Pi-responsive genes (rows) and GO-SLIM annotations (columns) separately displayed for the initial (a), persistent (e), and latent (i)
responses. (b-d, f-h, j-l) Summaries of selected clusters from clustered heatmaps showing annotations (blue: annotated; white: not annotated) and fold-change for response-and-recovery (red: positively regulated; green: negatively regulated) as measured by micro- and tiling-array platforms.

**Initial response, transcription regulation:** Members of the AGL, MYB, and WRKY TF families had decreased expression (Figure 2.4b-green arrow). The agamous-like (AGL) TFs are known to be involved with root morphological processes (Zhong et al., 2008), SOC1 is known to interact with both AGL44 and XAL1 (Zhou et al., 2009), and MYB85 to regulate lignin biosynthesis (Wilson et al., 1996; Franco-Zorrilla et al., 2005). On the other hand, ARF TFs, At1g71130 and At2g33710 showed increased expression (Figure 2.4b-red arrow). Both At1g71130 and At2g33710 are members of the ERF/AP2 family, At1g71130 is implicated in sugar:phosphate responses and At2g33710 in salt-stress.

**Initial response, ionic transport:** All genes with this annotation were with increased expression. Three of these genes encode proteins involved in calcium ion transport. ECA9 is an auto-regulated Ca2+ efflux pump and the ECA2 ATPase catalyzes Ca2+ efflux. Similarly, ATCHX17 is a sodium/proton anti-porter and was DEX 6.4 fold.

**Initial response, transporters:** Most notably, genes for sulfate and carnitine transporters, SULTR1;3 and OCT1, were DEX in Pi starvation by 7 and 71.5 fold, respectively. Indeed, OCT1 was the most DEX protein-coding gene in this study.
Persistently responsive genes were generally involved in "transcription regulation" and "ionic transport", similarly to the initially responsive genes. However, contrary to the initially responsive genes, several genes were also involved in "intra-cellular signaling". 

**Persistent response, transcription regulation:** Persistent MYB family members showed increased expression indicating that they may play a role in setting up conditions for the latent response. Similarly, the ethylene TF, TNY, showed increased expression and is involved in cytokinin biosynthesis (Qin et al., 2008), which is known to cross-talk with Pi starvation response (Albrecht et al., 2001). On the other hand, the DRIP1 TF showed reduced expression and encodes an E3 ligase known to mediate ubiquitination of TNY's family member DREB2A (Nelson et al., 2010). It may therefore be possible that DRIP1 also ubiquitinates TNY, which could explain why DRIP1 is repressed in light of TNY's increased expression.

**Persistent response, membrane-associated transport proteins:** Zinc-ion efflux and iron-reductase activity is repressed. Two peptide transporter genes, At1g62200 and At1g22550, have been implicated in zinc hyper-accumulation. Furthermore, the gene for another carnitine transporter, OCT4, increased expression by 2.5 fold. In contrast to OCT1's IPR pattern (fold-change of 71.5), OCT4 did not return to basal expression levels during the recovery suggesting that carnitine transport and energy-lipid metabolism is still a priority in the latent phase.
Persistent response, molecular signaling: All genes within this cluster showed reduced expression and did not recover after three-days in the replete medium. The most prominent members were WALK4 (membrane-bound receptor-like kinase), CIPK22 (associate with calcium-binding calcineurin B-like proteins) (Després et al., 2000), and At1g33590 (involved in signal transduction and the karrikin response) (Uno et al., 2000).

Latently responsive genes were annotated with a diverse set of functions. However, the most prominent expression patterns belong to clusters implicated in transcription regulation. A total of 48 TFs were latently responsive. This TF count exceeded the number of TFs identified in the initial and persistent response by 4 and 5.3 fold, respectively. Among proteins encoded by these 48 TF genes, several TF families were prominently represented: (1) the major leucine zipper family including 5 bZIP, 5 bHLH, 2 WRKY, and 2 HB TFs; (2) the ERF/AP2 family, with 5 TFs; (3) the zinc finger family, with 4 TFs; and (4), the NAC family, with 3 TFs. The remaining 22 TF genes were distributed among those encoding MYB, WUS, IAA, and other families.

The major leucine zipper family: Members of the bZIP sub-family are often involved in oxidative and pathogen defense responses and are commonly linked to ABA-related pathways (Zhang et al., 2009; Pandey and Somssich, 2009). In this group, genes for BZIP3, BZIP6, BZIP9, and BZIP24 were latently DEX. The bHLH sub-family is generally involved in plant development, circadian rhythm, and stress (Jung et al., 2010) included
genes for BHLH093, MYC3, At1g10610, At1g61660, and At2g42280. The WRKY sub-family mediate salicylic and jasmonic acid response and are involved in defense against pathogens and/or herbivores (Yang et al., 2011). Finally, we identified HD-ZIP18 and HD-ZIP52 as two latently responsive genes.

The ERF/AP2 family: Members of this family are involved in acclimation stress and are responsive to salicylic acid, jasmonic acid, cytokinin, and ethylene (Ascencio-Ibáñez et al., 2008; Wang et al., 2008). Two cytokinin and three ethylene response factors were latently DEX (CRF5/6, RAP2.1/2.11, HRE1).

As mentioned above, the latent phase is characterized by the activities of diverse TF families. To identify common processes where these TFs are involved, the TAIR database was used to provide a set of associations between peer-reviewed articles and the genes they describe. A simple text-mining approach was then used to group TF genes according to previous research topics (Woo and MacPherson et al., 2012). Using this technique we identified ten latent TFs previously reported as differentially expressed in studies on pathogen response and cell-cycle regulation during geminivirus infection (BZIP9, CRF6, MYBL2, NAC001, TGA1, ZF2, At1g61660, At2g03470, At4g32605, At4g37180) (Oono et al., 2006). In an analogous way, another set of seven latent TFs were previously implicated in pollen germination and tube growth (MDB2, TLP2, ZF2, At1g08170, At2g03470, At3g11100, At4g37180) (Hanada et al., 2011). Finally, we found
five latent TFs to be implicated by differential gene expression in a cold-acclimation study (BHLH093, BZIP9, HB52, RAP2.1, TLP2) (Lee et al., 2006). Additional smaller groups of TFs were identified in publications investigating topics such as "primary and secondary metabolites" (Osuna et al., 2007), "post-transcriptional regulation" (Ramel et al., 2007), "sucrose" (Journot-Catalino et al., 2006; Lelandais-Brière et al., 2007), and "basal resistance to pathogens" (Bari et al., 2006). Indeed, many TFs were found as differentially expressed across several studies.

Together, these results suggest novel roles for latent TFs in roots during recovery from Pi starvation. On the other hand, their presence across diverse studies highlights the importance to consider which genes are specific to the Pi starvation response and which are more general (non-specific) responders. In the following section we addressed these issues for all DEX genes by comparing our results with those published by the AtGenExpress initiative (Kilian et al., 2007).

**Interaction between Pi starvation and various AtGenExpress treatments**

To detect cross-talk (interaction) between different stress and hormone treatments we used the following procedure: (1) DEX genes from AtGenExpress treatments were detected using the same methods, software, and thresholds as for the analysis of our dataset; (2) DEX genes in our root experiments were associated to AtGenExpress treatments for which they were also DEX; (3) Each gene-node was weighted by the number of AtGenExpress treatments it was found to be associated with, and each
treatment-node by the number of Pi starvation responsive genes found to be DEX by that treatment (Figure 2.5a). With this procedure we were able to construct a network that links genes and the treatments. Moreover, groups of genes behaving in a similar manner could be identified. The results generally show that there is a subset of Pi starvation specific genes and another gene sub-set that interacts with as many as four other AtGenExpress treatments suggesting they may be more general stress responders (Figure 2.5b).
Figure 2.5: Comparative analysis between \( P_i \) starvation data and AtGenExpress

Subfigure (a) depicts a network of interactions between several types of treatment (blue squares) and genomic loci (other nodes) represented by probes on the ATH1 microarray. Edges/lines are drawn between treatments and loci if the locus is differentially regulated in the treatment as compared to its control (see Materials Methods). (b) A pie
chart summarizing the degree to which loci were found to interact with AtGenExpress treatments. Seventy percent of Pi-responsive loci are found to interact with nothing else, whereas 21% interact with 1 other AtGenExpress treatment and 1% interact with 4 AtGenExpress treatments. (c) Stacked histogram breaking down the degree of interactivity in terms of locus counts (y-axis) between various AtGenExpress treatments (x-axis). From bottom-to-top the stacked bars indicate the contributions made by the initial, persistent, and latent responses. (d) A plot describing the minimum threshold required (y-axis; p-value: blue; corrected: green) in order to detect a locus (x-axis; ranked by y-axis) as being differentially regulated in at least 1 AtGenExpress treatment. A dotted green line indicates the default threshold used by our analysis that results in 380 loci being detected in at least 1 AtGenExpress treatment and are therefore classified as non-specific. (c) A plot identical to (d), except only a subset of loci (x-axis) that are known to be Pi-responsive are displayed.

**Genes affected only during Pi starvation**

The 1,249 DEX genes in root were sorted into Pi starvation specific or non-specific bins according whether or not they were found to be DEX in at least one other non-Pi-starvation treatment (Figure 2.5b). ‘Specific’ should be interpreted as being affected by only Pi starvation. ‘Non-specific’ should be interpreted as being affected by at least one other treatment. Although there were actually 1,257 root Pi-responsive genes, eight genes were ambiguously mapped to AGI identifiers and were discarded from further analysis. At a Bonferroni corrected (for multiplicity testing) p-value threshold of 0.001, ~70% of the 1,249 genes were Pi starvation specific in roots.

The use of a p-value threshold binned genes, in a binary manner, into specific or non-specific categories. However, the use of a p-value threshold is essentially arbitrary, as can be seen by the shape of the blue curve in Figure 2.5d. Because this type of Boolean analysis did not distinguish between genes that are more affected by Pi starvation in root than in other treatments, we ascertained the minimum p-value threshold (MINP)
required to classify a gene as non-specific (Figure 2.5d). We looked at the p-value determined in other treatments for each of the 1,249 genes. In order to claim that a gene is more DEX, i.e. more affected, by Pi starvation, the p-value in Pi starvation should be smaller than in other treatments for the same gene. Thus, if MINP is greater than p-value for the gene in Pi starvation, then the higher MINP is, the greater the gene’s specificity to Pi starvation. Therefore, by ranking each gene by MINP, it is possible to determine the relative strength of influence of Pi starvation to the gene relative to other treatments. For instance, among the set of known Pi starvation responsive genes, we found PHO1, PHO1;H1, DGD1, and PHF1 to be specifically regulated in response to Pi starvation (Figure 2.5e). We also observed that PAP6, PHT1;3, and MGD2 ranked higher by MINP than any other known Pi starvation responsive gene and were therefore relatively more specific to Pi starvation than to any other treatments tested. Indeed, PAP6 was among the genes with the greatest specificity to Pi starvation, requiring a notably relaxed MINP (p-value = 0.2, ranked 7th) in order to be classified as Pi starvation non-specific.

Given the tendency of known Pi starvation -responsive genes to include multiple members of the same family (such as ion-transporters), we were able to ascertain the relative specificity of the family members (Figure 2.5e): (1) Among the PHT1 Pi-transporter family members, genes for PHT1;3 (42nd most Pi starvation specific), PHT1;8 (226th most Pi starvation specific), PHT1;1 (503rd), PHT1;9 (637th), and PHT1;7 (726th) were all Pi starvation specific; (2) Among the lipid processing families, both MGD2 and
MGD3 were highly ranked at 82nd and 189th, respectively, whereas, DGD2 and DGD1 were less specific with ranks of 760 and 811, respectively; (3) Three members of the SPX gene family were highly ranked at 330 (SPX1), 246 (SPX3), and 114 (SPX2); finally, (4) PHO1, a well-studied gene, was less specific (ranked 684th) than its close homolog, PHO1;H1 (ranked 394th).

After ranking all Pi starvation responsive genes in our study we found that of the top 100 most Pi starvation specific genes, there were only three members previously described as being Pi starvation responsive. These were PAP6 (ranked 7th), PHT1;3 (ranked 42nd), and MGD2 (ranked 82nd).

The most Pi starvation specific and initially responsive genes included At4g19770 (encoding a putative chitinase, responding at 3.5 fold and recovering at 4.3 fold) ranked 1st, At4g25160 (encoding a receptor-like cytoplasmic kinase with similarity to universal stress response protein of bacteria, responding at 7.5 fold and recovering at 16.5 fold) ranked 3rd, and At1g04700 (encoding a tyrosine kinase, responding at 1.8 fold and recovering at 2.2 fold) ranked 5th.

Genes that persistently responded to starvation during recovery include MTPA1 (encoding a zinc transporter, responding at 10.2 fold) ranked 2nd, At3g53770 (encoding a LEA3 family member, responding negatively at 3.8 fold) ranked 6th, and MYB72
(encoding a TF mediating systemic resistance, responding negatively at 7.3 fold) ranked 8th.

The top most specific latently responsive genes included At1g62090 (encoding a pseudogene-protein kinase, recovering negatively at 1.6 fold) ranked 10th, At5g66220 (encoding a chalcone-flavanone isomerase, recovering negatively at 1.6 fold) ranked 12th, and At3g26130 (encoding a cellulase, recovering at 1.8 fold) ranked 13th.

**Pi starvation responsive genes affected in multiple AtGenExpress treatments**

At p-value 0.001, 380 genes were non-specific as they were DEX in up to 4 additional AtGenExpress treatments. This behavior was exploited to elucidate: (1) treatments most similar to Pi starvation in terms of DEX genes; (2) gene interaction with hormonal treatments; and (3) cohorts of genes affected across different treatments.
Figure 2.6: Cross-talk between Pi^star, ABA treatment and cold, salt, and osmotic stress.

Subfigures (a) through (d) display highlighted portions of Figure 5a. Pi-responsive loci are highlighted (yellow) if they have additionally been found to be regulated by ABA.
treatment (a), cold stress (b) salt stress (c), and/or osmotic stress (d). The degree of overlap between highlighted regions is high indicating a large degree of cross-talk between treatments.

Botrytis cineria infection showed most similarity to Pi starvation data with 118 common DEX genes. This was closely followed by cold, salt, and osmotic stress with 106, 104, and 74 common DEX genes, respectively (histogram, Figure 2.5c). Most genes affected by these treatments were latently responsive in our experiments indicating that they may be involved in a general recovery from stress.

ABA treatment, with 76 common DEX genes, was the most similar to Pi starvation among the seven hormone treatments tested. This was notable as the next most similar response was elicited by methyl jasmonate (MJ) with approximately 3-fold fewer genes in common Pi starvation.

Overall, inspection of the network uncovered cohorts of common DEX genes between Pi starvation and other treatments. The most central cluster (see Figure 2.6a-d) of genes consisted of those associated with treatments most similar to Pi starvation (B. Cinerea, Cold, ABA, Salt, Osmotic). These genes included AZF2 (zinc finger protein, known to be regulated by ABA, salt, and desiccation), SPS2F (putative sucrose-phosphate synthase activity), TRI (tropinone reductase), At1g52560 (HSP20-like), At1g62570 (Flavin-Monooxygenase family implicated in multiple-stress studies), At3g01260 (Galactose-mutarotase family), and At5g35735 (Galactose-mutarotase family). These cohorts of
genes affected across several treatments suggest co-regulatory mechanisms and may lead to the elucidation of new Pi starvation response and recovery processes.

Discussion

Comparisons among different studies reveal the need to establish a reference dataset. By early 2012, 84 genes were well recognized to be a part of the stress response to Pi starvation (Woo and MacPherson et al., 2012). High throughput studies on Pi-starvation responsive genes identified as much as 33% of these 84 genes, and sometimes as few as 3% (Woo and MacPherson et al., 2012). In our study, 40% of the 84 genes were characterized by their response and recovery to Pi-starvation. These relatively low percentages highlight the need to standardize experimental designs, which was attempted by this study by analyzing different plant organs for their response and recovery to Pi starvation under highly controllable hydroponic conditions.

Expression profiles of Pi starvation responsive genes absent from the ATH1 microarray

The ATH1 microarray has been the default technology for transcriptomic studies thus far. Now, technologies such as the 1.0R tiling-array offer more gene coverage than ATH1. However, analyses and methods using the microarray architecture are dominantly used. Also, genes represented by ATH1 are better annotated. Therefore, the ATH1 platform was chosen as the primary means to analyze gene expression. Nevertheless, acquisition of tiling-array data in parallel served to both corroborate ATH1 results and to investigate those genetic elements not represented by ATH1 probes. As
such, the response and recovery profiles were determined for those 5,044 protein-coding genes absent from the ATH1 microarray (see Materials and Methods). Of these, 477 genes were DEX and may play a role during and post Pi starvation in roots and shoots (Table 2.2). Among those DEX genes, IPS1, GDPD3 and CPL3 are described below.

IPS1 is a non-coding RNA induced by Pi starvation and regulated by PHR1, a MYB TF (Bustos et al., 2010). IPS1 is well known for its ability to quench miR399 activity by acting as a target mimic. The quenching affect is due to IPS1’s partial complementarity to the mature miR399. This property allows binding of the RISC-loaded complex to IPS1 without cleavage occurring (Dello Ioiio et al, 2007). Our tiling-array data showed that IPS1’s transcript levels were inversely correlated to Pi availability in hydroponic media (Table 2.2). This result was confirmed by RT-PCR on the same samples. The IPS1 expression profile was the most differentially regulated profile among the 5,044 tiling-array-specific genes examined suggesting that the canonical IPS1:miR399:PHO2 circuit is active. Moreover, these IPS1-results are a good positive control for the remaining 476 DEX genes identified by the tiling-array analysis.

The glycerophosphodiester phosphodiesterase protein three (GDPD3) was implicated in Pi homeostasis in 2011 (Okushima et al., 2007). Our tiling-array data reflected this finding as well. Results show that GDPD3 exhibited a 19-fold increase in roots. The function of the GDPD family has been implicated in Pi-recycling through a membrane
remodeling process that is sensitive to Pi-flux (Cheng et al., 2011). Phospholipid bilayers are a major reservoir of intra-cellular Pi and are recycled during periods of limited Pi. Membrane Pi recycling replaces phospholipids in organelle membranes with alternate forms such as galacto- and sulfo-lipids (Thibaud et al., 2010). GDPD3 hydrolyzes glycerolphosphodiesters originating from phospholipids and produces a glycerol-3-phosphate (G-3-P) and a corresponding alcohol. The G-3-P byproduct is further degraded by acid phosphatases resulting in Pi accumulation (Okushima et al., 2007). GDPD3 is one of several genes DEX and involved in membrane remodeling, as described in the ATH1 results above.

The expression profiles of IPS1 and GDPD3 exemplify intra-cellular responses to Pi starvation. However, root architecture is also influenced by the availability of Pi. Under Pi-limiting conditions, root architecture changes by inhibiting primary root growth and promoting increased density of both lateral roots and root hairs (Rouached et al., 2010). CAPRICE LIKE MYB3 (CPL3) has been characterized as a positive regulator for root hair patterning (Rubio et al., 2001). It may therefore play a role in the Pi starvation response. However, since CPL3 is not represented by ATH1 it has not been described in previous Pi starvation studies. Thus, we aimed to gauge this gene's behavior from our tiling-array data and found that it was indeed over-expressed by 20-fold in roots before recovering to basal levels (classified as IPR). The shoot expression pattern was also IPR, although with less remarkable at 3-fold over-expression. It is therefore likely that the Pi starvation involves CPL3 to govern root hair patterning.
Altogether, IPS1 serves as a positive control for the tiling-array data. Also, IPS1 data will be used in Chapter 3 to model mimic:miRNA:target circuits. GDPD3 is proposed here as a novel gene involved in a process canonical to the Pi starvation response (membrane remodeling). CPL3 is a novel gene for which gene expression evidence suggests its role in remodeling root architecture. It is likely that among the 474 DEX genes identified in the tiling-array analysis, additional candidates with novel roles in Pi starvation are present. Our study provides the first tiling-array data available to the Pi starvation research community.

**A novel root response suggests roles for energy metabolism**

Previous genome-wide Pi starvation studies have used whole seedlings when sampling. Results from such research is likely biased toward shoot, as shoots contain 9-fold more RNA than root (Woo and MacPherson et al., 2012). The results reported here demonstrate the genome-wide response to Pi starvation in root for the first time. In fact, the number of responsive genes in root was 6.9-fold higher than those in shoots; root responsive genes were DEX to a much greater extent; and, 7% of the DEX genes in root were DEX in shoot. These observations emphasize the importance of analyzing distinct shoot and root samples, as the root response is overlooked in whole seedling analyses.

Lipid and energy metabolism seem to play a role in the initial and persistent responses as exemplified by DEX of OCT1 family members. Specifically, OCT1 and OCT4 encode
carnitine transporters involved in mitochondrial fatty acid metabolism. Arabidopsis mutants deficient in OCT1 and transgenic plants overexpressing 35S::OCT1 have been shown to promote and suppress lateral root hair development, respectively (Franco-Zorrilla et al., 2007). OCT1 is over-expressed 71-fold during Pi starvation, but lateral root hair development is promoted during Pi starvation. Therefore, OCT1 appears to be uncoupled from root hair development during Pi starvation and may hypothetically play a role in Pi starvation induced lipid/energy metabolism.

Transcription factors increase expression in root as the system tends toward recovery
Results show a general trend of AGL TFs to be DEX in response to Pi starvation and then return to basal levels after recovery. On the other hand, NF-YA TFs were also DEX but persisted in their response during the recovery period. The latent response showed a marked increase in the number of DEX TFs (48 genes): notably, 4-fold and 5.3-fold the numbers of TFs observed during the initial and persistent responses, respectively. Considering the number of DEX TFs, it is likely that the Pi starvation response is regulated by more than just the PHR1 and PHL1 TFs identified to date (Ji et al., 2008; Cheng et al., 2011); this hypothesis has been suggested before (Nilsson et al., 2010; Essigmann et al., 1998).

The over-expression of TFs was observed more frequently at the latent stages of Pi starvation. This suggests that significant regulatory changes occur during recovery. Text-
mining uncovered such DEX TFs to be involved in viral infection and cold acclimation. Therefore, the TFs identified may play a role in the general stress response.

**Expression of IPT3, ARF9, and MYB85 may alter root-morphology during Pi starvation**
During Pi starvation primary and lateral root (LR) development is significantly altered. Cytokinin (CK) represses root development at high concentrations (Cruz-Ramírez et al., 2006). Hence, we hypothesized that CK is reduced in Pi-limiting conditions. Without observing CK concentrations we used IPT3 as a marker gene and found it significantly repressed in roots suggesting that both IPT3 and CK may have a role in root remodeling. Similarly, ARF9 has increased gene expression in root tissues. Although ARF9 is not associated to root architecture it responds positively to auxin in the same manner as its family members ARF7/19 which are involved in root remodeling (Williamson, 2001). Therefore, ARF9 may have a similar role as IPT3 in this context as they both could facilitate root remodeling during Pi starvation.

During LR formation, removal of structural scaffolds within cell walls is required. This process represses MYB69 and MYB85. Our data shows that MYB69 and its upstream regulatory genes (SIZ1, SND1, NST1, and VND6) were basally expressed. However, expression of MYB85 (required for proper lignin deposition) was significantly reduced suggesting attenuation of lignin deposition resulting in a favorable environment for initiation of root hairs under Pi starvation.
Cross-talk among Pi starvation candidates uncovers commonality in response to plant hormone treatments

Cross-talk between stress responsive pathways was inferred by measuring similarities between the Pi starvation data presented here and 20 conditions in AtGenExpress. We showed that MYB72 is highly Pi-specific when compared to most other genes. Although both MYB72 and MYB74 are DEX, the former family member is ranked 8th for Pi-specificity whereas the latter, 1238th. In other words, MYB74 was found to be more DEX in conditions other than Pi starvation. Therefore, either MYB74 has evolved to become less Pi-specific, or MYB72 has evolved to be regulated by Pi-responsive pathways, exclusively. By contrast, the 3 cytochrome P450 genes (CYP94D1, ranked 13th; CYP735A1, ranked 24th; CYP76G1, ranked 27th) were highly specific to Pi starvation suggesting that they are functionally similar and may be able to complement one another in mutant studies.

Of note, the less specific a gene is to Pi starvation, the more often it is DEX in other stress responses. Therefore, groups of genes not Pi-specific but DEX in response to Pi starvation are more likely to be affected in other stress conditions as well. Such cases suggest co-regulation whereby upstream regulators respond to multiple stress conditions similarly. Investigating these cases would shed light on common regulatory responses. Indeed, some stress conditions have been better studied than others and the findings may be transferable if co-regulators are identified. For instance, we have determined groups of co-expressed Pi-responsive genes that are influenced by one or more of the following treatments: ABA, cold, salt, drought, and fungal infection. Indeed,
these five conditions yield the greatest number of commonly DEX genes. Furthermore, when compared to ABA treatment, other hormone treatments (ACC, IAA, MJ, BL, ZEA) show little cross-talk with Pi starvation, suggesting that ABA is the predominant hormone regulator for the general stress response.

**Conclusion**

The results presented describe Arabidopsis response and recovery to Pi starvation. For the first time we show detailed response in root, demonstrating that many more genes are differentially expressed than in shoot in response to Pi starvation. The analysis focused on genes represented on ATH1 for functional analyses and description of DEX genes. However, the 1.0R tiling-array implicated 477 genes affected by the Pi starvation that had not before been studied in genome-wide analyses. The efficacy of the tiling-array to identify novel genes in Pi starvation such as IPS1 highlights the utility of true genome-wide technologies in the detection of coding as well as non-coding transcript levels. ATH1 and 1.0R genechips describe a more varied response-and-recovery molecular phenotype than it was recognized in the literature. The data suggests the presence of an initial, persistent, and latent responses in both shoot and root organs. There is a progression of differentially regulated functional gene classes from initially responsive ion-transporters and persistent cellular signaling genes to latently responsive transcriptional regulators. Based on this progression we hypothesize that initially responsive genes are involved in immediate survival to Pi limiting stress as they mostly consist of Pi and metal ion transporters. We surmise that persistently responsive genes
participate in the transition between survival and recovery. The large number of TFs elicited during recovery suggests a return to non-stress conditions and normal state. However, the degree of TFs elicited during starvation conditions suggests the presence of pathways not controlled by the canonical TF, PHR1. Yet the list of genes regulated by PHR1 is large and prompted questions as to how specific its downstream genes were to the Pi starvation response. In answer, we found that for at least three of the well-known Pi responsive gene families (PHO1, PHT1;1, and SPX), the members that we observed to be most Pi starvation specific (PHO1;H1, PHT1;7–9, and SPX1-3) were regulated by PHR1. Finally, we demonstrated a large degree of cross-talk between clusters of co-expressed stress-responsive genes, which implicated ABA as the likely hormone mediator responsible for regulating common stress-responsive pathways.

As a genome-wide study of the Pi starvation response, the results, analyses, and interpretations presented here serve to enable future targeted research in the field of plant stress. In chapter three, the data described here will be used to interpret the role of non-coding RNA in plant stress, specifically the Pi starvation response.
Introduction
This chapter describes the analysis of small non-coding RNAs and their role in post
transcriptional gene silencing (PTGS) during Pi starvation stress. PTGS pathways are
discovered in many biotic and abiotic plant stress responses (Khraiwesh et al., 2012).
Moreover, PTGS is now a viable means for increasing stress tolerance in our staple crops
(Katiyar-Agarwal, 2006; Sunkar et al., 2006; Katiyar-Agarwal et al., 2007). Thus, PTGS is
becoming an increasingly relevant mode of transcriptional regulation in plant stress.
Among the most significant results presented in this chapter are those where we found
evidence for the activity of 11 putative miRNA and tasiRNA circuits during the response
and recovery of Arabidopsis to Pi starvation. Previously, only one miRNA circuit
(miRNA399:PHO2) was well characterized. The work in this chapter makes use of the
microarray and tiling-array data described in Chapter 2 (Woo and MacPherson et al.,
2012). Also, Pi starvation and mutant double-stranded RNA binding protein 4 (DRB4)
smRNA deep-sequencing libraries were donated by the Plant Molecular Biology
laboratory at Rockefeller University, NYC, NY, for use in this research.
As sessile organisms, plants have developed means to survive ever changing environments. Plant metabolism and changes at the genetic and molecular levels reflect the flexibility/adaptability required to survive conditions such as a drought, starvation, temperature, infection, or light stress (Seki et al., 2002). PTGS has recently been shown to occur across several biotic and abiotic stress conditions (Khraiwesh et al., 2012; Ruiz-Ferrer and Voinnet, 2009; Shukla et al., 2008; Sunkar et al., 2007; Jones-Rhoades et al., 2004; Xia et al., 2012). Given that there are far fewer pertinent PTGS miRNAs and siRNAs than there are protein coding genes in Arabidopsis (see http://mpss.udel.edu/at_sbs/), and that PTGS appears to pervade many stress responses, then by seeking to understand the role of PTGS in plant stress there is large scope to provide novel insights applicable to genetic improvement of crop tolerance. Several papers have already reported PTGS-based genetic modification (Katiyar-Agarwal et al., 2006; Sunkar et al., 2006; Katiyar-Agarwal et al., 2007). The motivation of improving crop tolerance is important as genetically modified crops could aid farmers in increasing their yields. Moreover, it could also allow agriculture on previously unusable land. Recent reports have shown the amount of arable land is decreasing every year, while food demand is on the rise (Chen et al., 2010). Thus, the study of PTGS in plant biology has the potential to address one the most pressing concerns of our time, food security.

The main aim of this chapter was to identify elements involved in PTGS that also function in the plant's response to stress. To do this we restricted our search to the analysis of miRNAs, and their closely related tasiRNAs. The data at our disposal
consisted of 36 microarray and tiling-array experiments measuring the transcriptomic response and recovery of Arabidopsis' root and shoot to Pi starvation (published by Woo and MacPherson et al., 2012). In addition, we used a related Pi starvation smRNA dataset and a list of previously predicted miRNA targets and target-mimics (e.g. IPS1 (Bari et al., 2006). This data was then used (see Materials and Methods section) to detect the presence of any miRNA:target circuits that could possibly be active during either the response or recovery from Pi starvation stress. From the list of candidates generated, we selected several pertinent examples based on their biology and the extent to which their putative targets were differentially expressed in our Pi starvation data. These examples included IPS1:miR399:PHO2, miR399:RS4, miR399:NF-YA10, miR398:CSD1/2 (known but not connected to Pi starvation before), miR2111:TPS11, miR164:NAC6, miR157:TMO7, miR157:PSB28, RPS2:miR169:IPS2, and miR397:LAC2. Interestingly, binding sites for the Pi starvation inducible MYB transcription factor (TF), PHR1, have been found in miR399, miR398, miR169, and TAS4 gene promoters suggesting that PHR1 may regulate the transcription of several PTGS elements.

For our analysis of tasiRNA activity in plant stress, we additionally analyzed double stranded RNA binding protein 4 (DRB4) smRNA data in order to determine putative tasiRNA genes in addition to the well-studied TAS1-4 genes. We then used DRB4 smRNA data to identify and construct a network of TAS genes and their respective target mRNAs. Targets in this network were then evaluated for putative roles in stress, based on their behavior in the Pi starvation. Using this methodology, we found evidence
supporting the activity of TAS4, TAS1b, and TAS2 in Pi starvation. Moreover, we identified two novel TAS genes, one located in nitrate reductase 1 (NR1), and the other located in an intergenic region on chromosome 3 at position 1967672 (Chr3_1967672). In summary, we identified the following TAS:target relationships, TAS4:PAP1, NR1:PAP1, and Chr3_1967672:TMO7.

Altogether, our results suggest that the local root response to Pi starvation is influenced by the plant's systemic response pathways via PHR1-mediated PTGS. We hypothesize that PHR1-mediated PTGS influences the local root response by modulating concentrations of ROS. Furthermore, we suggest that miR2111 is capable of influencing shoot to root carbon signaling pathways by silencing TPS11. Recent reviews in plant Pi metabolism have highlighted the need to further understand the link between the local and global Pi starvation response (Abel, 2011). In conclusion, our results provide testable hypotheses for how the two responses (systemic and local) interact during Pi starvation in order to adapt to changes in Pi-bioavailability through the acquisition, transport, and recycling of Pi.

**Results**
The role of non-coding RNA in plant stress is explored in this chapter with three questions in mind: 1/ What is the role of the protein machinery responsible for producing different classes of smRNA; 2/ What is known of miRNA stress responsive circuits; and 3/ Can putative miRNA circuits be identified? The Pi starvation data
described in the previous chapter is used to help answer these questions and will be referred to as the, "Pi-dataset". The sub-set of smRNA sequencing data used for the analysis presented here will be referred to as the, "smRNA-dataset". The smRNA-dataset was kindly contributed by the Plant Molecular Biology laboratory at Rockefeller University, NY, USA.

The Pi-dataset revealed few to no differentially expressed (DEX) genes encoding proteins responsible for smRNA biogenesis. This was the case in both the root and shoot responses and recovery to Pi starvation. These basal gene expression patterns demonstrate that among other forms of regulation, miRNA feedback loops such as the miR168:AGO1 circuit (Vaucheret et al., 2004) are not active during Pi starvation. However, miR168 feedback has been reported for several stress conditions such as drought, heat, salt, UV-radiation, and symbiotic nitrogen fixation (Liu et al., 2008). On the other hand, miR168 is only detected at one fourth the transcript levels of Pista-responsive miR399f, in the root response. Lack of fold change in miR168 response supports the hypothesis that feedback in miRNA biogenesis is not part of the Pi starvation response.

Despite the lack of response to Pi starvation by genes involved in smRNA biogenesis, many of their smRNA products are DEX during the Pi starvation response and recovery (Table 3.1). The smRNA-dataset contained expression data for 86 miRNAs and 3847
putative miRNA:target predictions. Among the 3847 predictions, 158 were identified as possible miRNA:target-mimic relationships due to the targets not containing open reading frames (for example see IPS1 (Bari et al., 2006)). From the set of miRNAs, target-mimics, and target mRNAs, we derived 5057 mimic:miRNA:target circuits.

In the results to follow, four miRNA:target circuits will be presented that are known to be involved in several biotic and abiotic stress responses. These relationships have been curated from the available literature and represent findings from several plant organisms, including A. thaliana. Of these known miRNA:target relationships, one was miR399:PHO2, currently the only known miRNA:target relationship conclusively shown to be involved in Pi starvation (Bari et al., 2006).
Table 3.1: Micro-RNAs differentially expressed in the response to, or recovery from Pi starvation

<table>
<thead>
<tr>
<th>mat-miRNA</th>
<th>Root</th>
<th>Shoot</th>
<th>mat-miRNA</th>
<th>Root</th>
<th>Shoot</th>
<th>mat-miRNA</th>
<th>Root</th>
<th>Shoot</th>
</tr>
</thead>
<tbody>
<tr>
<td>ath-miR158b</td>
<td>1.2</td>
<td>-0.5</td>
<td>0.2</td>
<td>30</td>
<td>ath-miR169m</td>
<td>1.1</td>
<td>0.2</td>
<td>-3.4</td>
</tr>
<tr>
<td>ath-miR395f</td>
<td>0.6</td>
<td>1.3</td>
<td>-2.8</td>
<td>0.4</td>
<td>31</td>
<td>ath-miR169m</td>
<td>1.1</td>
<td>0.2</td>
</tr>
<tr>
<td>ath-miR395d</td>
<td>0.9</td>
<td>2.1</td>
<td>-2.6</td>
<td>0.5</td>
<td>32</td>
<td>ath-miR169m</td>
<td>1.9</td>
<td>2.0</td>
</tr>
<tr>
<td>ath-miR395e</td>
<td>0.9</td>
<td>2.1</td>
<td>-2.6</td>
<td>0.5</td>
<td>33</td>
<td>ath-miR169m</td>
<td>1.9</td>
<td>2.0</td>
</tr>
<tr>
<td>ath-miR395b</td>
<td>0.6</td>
<td>1.3</td>
<td>-2.8</td>
<td>0.4</td>
<td>34</td>
<td>ath-miR169a</td>
<td>0.6</td>
<td>0.8</td>
</tr>
<tr>
<td>ath-miR395c</td>
<td>0.6</td>
<td>1.3</td>
<td>-2.8</td>
<td>0.4</td>
<td>35</td>
<td>ath-miR169a</td>
<td>0.9</td>
<td>0.7</td>
</tr>
<tr>
<td>ath-miR395a</td>
<td>0.9</td>
<td>2.1</td>
<td>-2.6</td>
<td>0.5</td>
<td>36</td>
<td>ath-miR169f</td>
<td>0.9</td>
<td>0.7</td>
</tr>
<tr>
<td>ath-miR167c</td>
<td>0.8</td>
<td>1.3</td>
<td>0.0</td>
<td>0.0</td>
<td>37</td>
<td>ath-miR169a</td>
<td>0.9</td>
<td>0.7</td>
</tr>
<tr>
<td>ath-miR167d</td>
<td>0.2</td>
<td>1.3</td>
<td>-2.8</td>
<td>0.4</td>
<td>38</td>
<td>ath-miR169a</td>
<td>0.9</td>
<td>0.7</td>
</tr>
<tr>
<td>ath-miR159d</td>
<td>2.2</td>
<td>-1.6</td>
<td>4.0</td>
<td>0.3</td>
<td>40</td>
<td>ath-miR159b</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>ath-miR159e</td>
<td>2.2</td>
<td>-1.6</td>
<td>4.0</td>
<td>0.3</td>
<td>41</td>
<td>ath-miR166f</td>
<td>2.5</td>
<td>0.6</td>
</tr>
<tr>
<td>ath-miR159f</td>
<td>2.2</td>
<td>-1.6</td>
<td>4.0</td>
<td>0.3</td>
<td>42</td>
<td>ath-miR156g</td>
<td>2.5</td>
<td>0.6</td>
</tr>
<tr>
<td>ath-miR159g</td>
<td>2.2</td>
<td>-1.6</td>
<td>4.0</td>
<td>0.3</td>
<td>43</td>
<td>ath-miR166d</td>
<td>2.5</td>
<td>0.6</td>
</tr>
<tr>
<td>ath-miR164c</td>
<td>2.1</td>
<td>-1.4</td>
<td>0.5</td>
<td>0.7</td>
<td>44</td>
<td>ath-miR166e</td>
<td>2.5</td>
<td>0.6</td>
</tr>
<tr>
<td>ath-miR164d</td>
<td>2.1</td>
<td>-1.4</td>
<td>0.5</td>
<td>0.7</td>
<td>45</td>
<td>ath-miR166b</td>
<td>2.5</td>
<td>0.6</td>
</tr>
<tr>
<td>ath-miR164a</td>
<td>2.1</td>
<td>-1.4</td>
<td>0.5</td>
<td>0.7</td>
<td>46</td>
<td>ath-miR166c</td>
<td>2.5</td>
<td>0.6</td>
</tr>
<tr>
<td>ath-miR164b</td>
<td>2.1</td>
<td>-1.4</td>
<td>0.5</td>
<td>0.7</td>
<td>47</td>
<td>ath-miR165a</td>
<td>2.5</td>
<td>0.6</td>
</tr>
<tr>
<td>ath-miR856</td>
<td>0.2</td>
<td>1.3</td>
<td>1.2</td>
<td>48</td>
<td>ath-miR165a</td>
<td>2.5</td>
<td>0.6</td>
<td>1.1</td>
</tr>
<tr>
<td>ath-miR172a</td>
<td>0.0</td>
<td>1.3</td>
<td>1.2</td>
<td>49</td>
<td>ath-miR165a</td>
<td>2.5</td>
<td>0.6</td>
<td>1.1</td>
</tr>
<tr>
<td>ath-miR172b</td>
<td>0.3</td>
<td>0.0</td>
<td>-1.4</td>
<td>1.2</td>
<td>50</td>
<td>ath-miR856</td>
<td>1.5</td>
<td>1.1</td>
</tr>
<tr>
<td>ath-miR856</td>
<td>0.3</td>
<td>0.0</td>
<td>-1.4</td>
<td>1.2</td>
<td>51</td>
<td>ath-miR159d</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>ath-miR182</td>
<td>2.0</td>
<td>-1.5</td>
<td>0.0</td>
<td>0.2</td>
<td>52</td>
<td>ath-miR159d</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>ath-miR183</td>
<td>2.0</td>
<td>-1.5</td>
<td>0.0</td>
<td>0.2</td>
<td>53</td>
<td>ath-miR182</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>ath-miR185</td>
<td>2.0</td>
<td>-1.5</td>
<td>0.0</td>
<td>0.2</td>
<td>54</td>
<td>ath-miR183</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>ath-miR186a</td>
<td>3.4</td>
<td>-1.5</td>
<td>0.3</td>
<td>0.5</td>
<td>55</td>
<td>ath-miR185</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>ath-miR186b</td>
<td>3.4</td>
<td>-1.5</td>
<td>0.3</td>
<td>0.5</td>
<td>56</td>
<td>ath-miR186a</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>ath-miR186c</td>
<td>3.4</td>
<td>-1.5</td>
<td>0.3</td>
<td>0.5</td>
<td>57</td>
<td>ath-miR186b</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>ath-miR186d</td>
<td>3.4</td>
<td>-1.5</td>
<td>0.3</td>
<td>0.5</td>
<td>58</td>
<td>ath-miR186c</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Values: \( \log_2(\text{Fold-change}) \), green shade: Increased expression, red shade: decreased expression, grey shade: basal expression (no change).

Known miRNA circuits in stress

Of the 78 miRNA:target relationships garnered from the public literature (see Materials and Methods section), this study had access to 34 in the form of 129 mimic:miRNA:target circuits (Table 3.2). Each mimic-RNA, miRNA, and target was associated to their fold-changes during the response and recovery of Arabidopsis shoot and root to Pi starvation (see Chapter 2). To determine if any of these known stress-
responsive miRNA:target circuits were a part of the Pi starvation response, we
developed several conditions required to be met by the fold-changes observed in mimic,
miRNA, and target (Table 3.3). Thus, Table 3.3 provides a list of hypotheses attributed to
a given mimic:miRNA:target circuit based on the fold-change of that circuit’s individual
RNA components. Using Table 3.3 to map hypotheses to all circuits, we discovered that
six of the 129 circuits displayed the necessary changes in expression expected for active
miRNA regulation during Pi starvation (highlighted in Table 3.4). The six circuits shared
three of the same miRNAs, and the abundance of their target-mimics did not suggest
quenching; therefore, after discarding the target-mimics the six circuits were collapsed
to three miRNA:target circuits. These three circuits were miR398:CSD1, mir398:CSD2,
and miR157:PSB28. These three circuits displayed similarities to the canonical Pi
starvation responsive circuit, IPS1:miR399:PHO2 (Table 3.4). Given prior association of
these three circuits to other stress responses, it is likely that they are generally involved
in plant stress. More specifically, since the three circuits appeared to be active only
during the root recovery then two conclusions were drawn: one, these circuits are
involved in the general recovery from stress; and two, they are regulated in a root
specific manner. Indeed, these circuits are distinct from the canonical miR399 circuit,
IPS1:miR399:PHO2, in that miR399 is Pi starvation specific, and the circuit is active in
both shoot and root (Table 3.4).
**Table 3.2**: List of miRNA:target circuits and the stress conditions they are known to respond to

Columns – (A)biotic: circuit is either involved in biotic or abiotic stress, Stress: name of stress circuit is known to be involved in, miRNA: the miRNA identifier, REG: the expected regulation of the miRNA in the stress (++) increased expression, (+): slightly increased expression, (--) decreased expression), targets: list of miRNA targets applicable to stress condition, Pi Stra (miRNA): the fold-change (log2) of the miRNA in our Pi starvation data (RES: response, REC: recovery).
Table 3.3: List of hypotheses assigned to miRNA circuits based on their fold-change profiles

<table>
<thead>
<tr>
<th>Observed</th>
<th>Hypothesis</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIMIC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UP</td>
<td>- Mimic is quenching miRNA, which inhibits target</td>
<td>H1</td>
</tr>
<tr>
<td>UP</td>
<td>- Synergistic effect of both mimic and miRNA, i.e. both are regulated to up-regulate target</td>
<td>H2.1</td>
</tr>
<tr>
<td>DOWN</td>
<td>- miRNA inhibits target (high support) but is down-regulated, mimic does not quench</td>
<td>H2.2</td>
</tr>
<tr>
<td>DOWN</td>
<td>- miRNA inhibits target (high support) but is down-regulated, no mimic activity</td>
<td>H2.3</td>
</tr>
<tr>
<td>-</td>
<td>- miRNA inhibits target (high support), mimic does not quench</td>
<td>H3.1</td>
</tr>
<tr>
<td>DOWN</td>
<td>- miRNA inhibits target (high support), miRNA is quenchable by mimic (low support compared to H1)</td>
<td>H3.2</td>
</tr>
<tr>
<td>-</td>
<td>- miRNA inhibits target, no mimic activity</td>
<td>H3.3</td>
</tr>
<tr>
<td>UP</td>
<td>- Mimic is masking the effect of the miRNA resulting in no change in target transcript levels</td>
<td>H4</td>
</tr>
<tr>
<td>DOWN</td>
<td>- No evidence of interaction</td>
<td>H5.1</td>
</tr>
<tr>
<td>DOWN</td>
<td>- No evidence of interaction</td>
<td>H5.10</td>
</tr>
<tr>
<td>-</td>
<td>- No evidence of interaction</td>
<td>H5.11</td>
</tr>
<tr>
<td>DOWN</td>
<td>- No evidence of interaction</td>
<td>H5.12</td>
</tr>
<tr>
<td>-</td>
<td>- No evidence of interaction</td>
<td>H5.13</td>
</tr>
<tr>
<td>DOWN</td>
<td>- No evidence of interaction</td>
<td>H5.14</td>
</tr>
<tr>
<td>-</td>
<td>- No evidence of interaction</td>
<td>H5.15</td>
</tr>
<tr>
<td>DOWN</td>
<td>- No evidence of interaction</td>
<td>H5.16</td>
</tr>
<tr>
<td>-</td>
<td>- No evidence of interaction</td>
<td>H5.17</td>
</tr>
<tr>
<td>DOWN</td>
<td>- No evidence of interaction</td>
<td>H5.18</td>
</tr>
<tr>
<td>-</td>
<td>- No evidence of interaction</td>
<td>H5.19</td>
</tr>
<tr>
<td>UP</td>
<td>- No evidence of interaction</td>
<td>H5.2</td>
</tr>
<tr>
<td>DOWN</td>
<td>- No evidence of interaction</td>
<td>H5.3</td>
</tr>
<tr>
<td>-</td>
<td>- No evidence of interaction</td>
<td>H5.4</td>
</tr>
<tr>
<td>UP</td>
<td>- No evidence of interaction</td>
<td>H5.5</td>
</tr>
<tr>
<td>DOWN</td>
<td>- No evidence of interaction</td>
<td>H5.6</td>
</tr>
<tr>
<td>-</td>
<td>- No evidence of interaction</td>
<td>H5.7</td>
</tr>
<tr>
<td>UP</td>
<td>- No evidence of interaction</td>
<td>H5.8</td>
</tr>
<tr>
<td>DOWN</td>
<td>- No evidence of interaction</td>
<td>H5.9</td>
</tr>
<tr>
<td>-</td>
<td>- No evidence of interaction</td>
<td></td>
</tr>
</tbody>
</table>

See Materials and Methods for detailed explanation of using this table. Columns – Observed MIMIC: the behaviour of the target-mimic based on fold-change (“UP” for increased expression, “DOWN” for decreased expression), Observed miRNA: the behaviour of the miRNA based on fold-change (“UP” for increased expression, “DOWN” for decreased expression), Observed TARGET: the behaviour of the target based on fold-change (“UP” for increased expression, “DOWN” for decreased expression), Hypothesis: A written description of why the observed behaviour supports the activity of the miRNA circuit, Code: A brief internal ID for summarising a circuits behaviour; Blue highlight: hypotheses that the target-mimic in their description.
Table 3.4: Summary of results for known stress responsive miRNA circuits

<table>
<thead>
<tr>
<th>PTGS Circuit</th>
<th>INTERACTIONS and HYPOTHESES</th>
<th>EXPRESSION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Score</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Root</td>
<td>Shoot</td>
</tr>
<tr>
<td>DARE</td>
<td>DARE</td>
<td>DARE</td>
</tr>
<tr>
<td>DARE</td>
<td>DARE</td>
<td>DARE</td>
</tr>
</tbody>
</table>

| Columns – PTGS Circuit: lists the target-mimic, miRNA, and target; INTERACTIONS and HYPOTHESES: lists the hypotheses detected for the fold-change exhibited by the circuit (see Table 3.3); Overall Score: the number of times a hypothesis supporting miRNA activity is made (ranged from 0 to 4); EXPRESSION: the relative* log2 fold-change of MIMIC, miRNA, and TARGET for the ROOT and SHOOT RESPONSE and RECOVERY; Response of miRNA in other stress: Indication if miRNA is known to be responsive in other stress conditions; *: all fold-change data is expressed relative to the canonical miRNA circuit IPS1:miR399f:PHO2, which is why that circuits fold-change is always equal to 1, or -1.

Two circuits involving miR398 and two targets of the copper/zinc superoxide dismutase protein (CSD) family were hypothesized to be active during the recovery from Pi starvation (Figure 3.1). These circuits are miR398:CSD1 and miR398:CSD2. MicroRNA398 has been shown to be repressed by Pi limitation and is known to target CSD family members (Hsieh et al., 2009). However, miR398 and CSD1 and CSD2 have not been observed in Pi starvation experiments until now. Interestingly, CSD1 and CSD2 are have increased expression in As(V) stress (Abercrombie et al., 2008). Pi is chemically similar to As(V), hence As(V) is highly toxic as it destabilizes Pi metabolism (mimicking Pi starvation) causing oxidative stress (Catarecha et al., 2007). Interestingly, As(V) represses Fe-SODs which are also repressed during the Pi starvation experiments presented in Chapter 2. Therefore, Cu/Zn and Fe-SODs act similarly in As(V) stress and Pi
starvation. Whether miR398 is involved in As(V) stress remains to be seen. However, from the number of similarities between the two stress conditions miR398 could well be differentially regulated under As(V) stress. In As(V) stress, the likely role of the CSD family is in reducing redox stress. Since the miR398:CSD1/2 circuit is detected in root during recovery from Pi starvation, redox stress is likely caused by the Pi starvation response, and As(V) stress likely causes oxidative damage by limiting Pi availability. Finally, miR398 has a P1BS binding element in its putative promoter and may therefore be affected by the canonical Pi starvation TF PHR1 (Hsieh et al., 2009).
Figure 3.1: Experimental evidence for miR398:CSD1 and miR398:CSD2 circuits

a – predicted miRNA circuit, Nodes – green flag: miRNA, blue oval: gene, white hexagon: metabolite; Edges – black: known, red: predicted by this study, solid arrow: induction, empty arrow: inhibition, open arrow: influences; b – expression of PTGS elements from subfigure (a) – x-axis: categorical: root response, root recovery, shoot response, and shoot recovery; y-axis: log₂ fold-change (relative to the IPS1:miR399:PHO2 circuit, i.e. all mimic RNAs are relative to IPS1 fold-change, all miRNAs are relative to miR399 fold-change, and all genes are relative to PHO2 fold-change).
Figure 3.2: Experimental evidence for the miR157:PSB28 circuit

a – predicted miRNA circuit, Nodes – green flag: miRNA, blue oval: gene, white box: system process; Edges – black: known, red: predicted by this study, solid arrow: induction, empty arrow: inhibition, open arrow: influences; b – expression of PTGS elements from subfigure (a) – x-axis: categorical: root response, root recovery, shoot response, and shoot recovery; y-axis: log2 fold-change (relative to the IPS1:miR399:PHO2 circuit, i.e. all mimic RNAs are relative to IPS1 fold-change, all miRNAs are relative to miR399 fold-change, and all genes are relative to PHO2 fold-change).

In contrast to miR398, miR157 has not previously been identified as responsive to Pi starvation. Yet, our method predicted that miR157 could putatively repress
photosystem II reaction center protein 28 (PSB28, see Figure 3.2). Given that two of three miRNAs predicted by the method are known to respond to Pi starvation, the miR157:PSB28 circuit is likely bona-fide and novel. PSB28 is thought to be involved in photosystem II's electron transport chain due to its sequence similarity to PSBW (Friso et al., 2004). However, PSB28 has been implicated in biotic stress (insect and viral) (Little et al., 2007; Ascencio-Ibáñez et al., 2008) and more significantly in severe nitrogen limitation known to induce similar root morphology as Pi starvation (Bi et al., 2007). Significantly, PSB28 has recently been annotated by TAIR and the Gene Ontology as being involved in the biosynthesis of phosphatidylglycerol, a precursor to the phospholipids making up organelle membranes. Implication of phosphatidylglycerol biosynthesis is intriguing evidence, since Pi starvation and many of the DEX genes presented in Chapter 2, are involved in phospholipid recycling. The miR157:PSB28 circuit predicted here indicates that miR157 down-regulates PSB28 during Pi starvation recovery but not during the initial response. Given that phospholipids are recycled during the response and not the recovery, PSB28 is likely a negative regulator of phosphatidylglycerol biosynthesis.

The above results have been presented for miRNA circuits known to be involved in plant stress. The following section will present similar results, but for predicted miRNA circuits not described above.
**Novel miRNA:target and target-mimic circuits**

In the previous results, we used a known set of stress responsive miRNA:target circuits to find out if they were active during Pi starvation. Here, we analyzed 5,057 predicted mimic:miRNA:target circuits for their potential role in Pi starvation. After employing the same methodology as used for the analysis of known stress-responsive circuits (above), we found that expression profiles of miRNA, mRNA, and target-mimics supported the activity of 167 circuits in Pi starvation (approximately 3% of the total predicted). These 167 circuits are composed of 48 mimic:miRNA:target circuits and 119 miRNA:target circuits. These miRNA circuits were, in turn, composed of 12 miRNAs, 63 target mRNAs (counted by AGI identifier), and three target-mimic RNAs.

The previously predicted miR398:CSD1/2, miR157:PSB28 circuits are among the 167 predicted Pi starvation responsive miRNA circuits. Furthermore, the canonical miR399f:PHO2 circuit was also predicted and included five of its family members (miR399a-d,f).

Given the large number of putatively active miRNA circuits in Pi starvation, we chose to analyze and present only a few of the circuits in detail. To choose the more relevant Pi starvation responsive miRNA circuits, we ranked the 167 circuits (by the degree that their miRNA target's gene expression changed) and selected the top 50. The canonical
IPS1:miR399f:PHO2 was ranked 14th. Moreover, miR399 was predicted to be involved in several previously unreported Pi starvation responsive circuits (see below).

Micro-RNA169 putatively silenced its target (At1g67600) during root recovery from Pi starvation (Figure 3.3). From the results presented in Chapter 2, At1g67600 was non-responsive to Pi starvation in shoot. However, At1g67600 was significantly DEX in both the root's response (promoted, 9th most DEX gene) and recovery (repressed, 21st most DEX gene). TAIR annotates At1g67600 as a putative acid phosphatase. We sought to find more information on At1g67600 from the literature but found little to no published results. Indeed, no mention of At1g67600 was found in the latest tables associating genes to the publications they are described in (tables curated by TAIR). Therefore, At1g67600 appears to be a novel Pi starvation responsive gene and, its role as an acid phosphatase (liberate Pi groups), may indeed play a part in the response to Pi starvation. Micro-RNAs miR169b/c were repressed during the Pi starvation response, and promoted during the recovery (Figure 3.3). This observation supports the hypothesis that miR169b/c acts to post-transcriptionally inhibit At1g67600 during the recovery from Pi starvation. Moreover, it suggests that miR169 itself is regulated by Pi availability, a hypothesis that is supported by prior observations of miR169 showing differential expression in Pi and nitrogen limitation experiments (Pant et al., 2009). Finally, this circuit is further supported by the response and recovery of miR169's putative target-mimic, ID_4715 (genomic co-ordinates, Chr2:15408383-15408660). ID_4715 displayed expression patterns similar to those of miR169's target, which is
necessary for target-mimicry to occur. It remains to be seen whether ID_4715 quenches miR169 activity as efficiently as IPS1 does miR399. Altogether, the expression patterns exhibited by the ID_4715:miR169b/c:At1g67600 circuit support the hypothesis that At1g67600 is down-regulated by the synergistic effect of increased miR169b/c and decreased ID_4715 expression. Given the novelty of At1g67600 as a putative Pi starvation responsive gene, this circuit is of interest. Therefore, we propose that At1g67600 be designated "induced by phosphate starvation 2 (IPS2)" and that ID_4715 (a long non-coding RNA) be designated "repressed by phosphate starvation 1 (RPS1)".

Our results predict that miR2111 targets trehalose phosphatase/synthase 11 (TPS11, see Figure 3.4). The TPS gene family is thought to be involved in carbon signaling (Eastmond et al., 2003), and is differentially expressed in response to changes in sucrose concentration (Usadel et al., 2008). Pi starvation data (Chapter 2) shows that TPS11 is repressed during the root response and promoted during its recovery. Micro-RNA2111 is inversely correlated with TPS11 supporting the hypothesis that it regulates TPS11's abundance post-transcriptionally. The data does not support target-mimicry and the circuit appears to only be between miRNA and target gene. Altogether, carbon signaling, already known to play a role in the Pi starvation response (Lei et al., 2011), may involve miR2111.
Figure 3.3: Experimental evidence for the miR169:At1g67600 circuit

a – predicted miRNA circuit, Nodes – red half-flag: target-mimic, green flag: miRNA, blue oval: gene, white pentagon: sensing; Edges – black: known, red: predicted by this study, solid arrow: induction, empty arrow: inhibition, open arrow: influences; b – expression of PTGS elements from subfigure (a) – x-axis: categorical: root response, root recovery, shoot response, and shoot recovery; y-axis: log₂ fold-change (relative to the IPS1:miR399:PHO2 circuit, i.e. all mimic RNAs are relative to IPS1 fold-change, all miRNAs are relative to miR399 fold-change, and all genes are relative to PHO2 fold-change).
**Figure 3.4:** Experimental evidence for the miR2111:TPS11 circuit

*a* – predicted miRNA circuit, *Nodes* – green flag: miRNA, blue oval: gene, white hexagon: metabolite; *Edges* – black: known, red: predicted by this study, solid arrow: induction, empty arrow: inhibition, open arrow: influences; *b* – expression of PTGS elements from subfigure (a) – *x*-axis: categorical: root response, root recovery, shoot response, and shoot recovery; *y*-axis: $\log_2$ fold-change (relative to the IPS1:miR399:PHO2 circuit, i.e. all mimic RNAs are relative to IPS1 fold-change, all miRNAs are relative to miR399 fold-change, and all genes are relative to PHO2 fold-change).

Micro-RNA399 was predicted by our method to target several genes other than the canonical PHO2. Here, we show that the target-mimic, ID_5602 (IPS1, Chr3:3050940-
3051670), is promoted by Pi starvation and correlated to miR399d/f's putative target, NF-YA10 (Figure 3.5). NF-YA10 is a TF binding to 5'-CCAAT-3' motifs and has been shown to be responsive to Pi starvation (Müller et al., 2007). Given that miR399d/f are also correlated to NF-YA10, we hypothesize that IPS1 quenches miR399d/f's silencing of NF-YA10. Of note, PHO2 gene expression does not support IPS1 quenching in the IPS1:miR399d/f:PHO2 circuit in root. However, the opposite is true in shoot data: IPS1 quenching is supported by PHO2, but not NF-YA10 gene expression in shoot. Therefore, either miR399d/f does not cleave NF-YA10 (a possibility easily tested in vitro) or, another unknown factor regulates IPS1 quenching in a tissue specific manner. The nature of this additional factor is likely that NF-YA10 is expressed in a separate tissue to that of PHO2. There is limited annotation of NF-YA10's function. The Gene Ontology has computationally annotated it as being involved in anthocyanin biosynthesis (a known response to Pi starvation), citing Heyndrickx and Vandepoele, (2012) (Heyndrickx and Vandepoele, 2012). Interestingly, NF-YA2 is another CCAAT binding TF that displays a strikingly similar gene expression profile to NF-YA10 during Pi starvation. Like NF-YA10, NF-YA2 has also been shown to be promoted by Pi starvation (Müller et al., 2007). Unfortunately, like NF-YA10, there has not been a lot of prior research on NF-YA2. However, genome wide studies have implicated NF-YA2 in nitrogen limitation stress, which bares similarity to Pi starvation in that it induces similar changes in root morphology (Krapp et al., 2011). Finally, evidence was presented at the 22nd International Conference on Arabidopsis Research by Siriwardana C., Kumimoto R., and Holt B., (2011) that NF-YA2 was involved in the ABA-mediated drought response. This
connection to ABA and drought stress warrants mentioning since ABA plays a dominant role in regulating many stress responsive genes elicited by both Pi starvation and drought stress (Woo and MacPherson et al., 2012).

**Figure 3.5: Experimental evidence for the miR399:NF-YA10 circuit**

a – predicted miRNA circuit, Nodes – green flag: miRNA, blue oval: gene, white hexagon: metabolite; Edges – black: known, red: predicted by this study, solid arrow: induction, empty arrow: inhibition, open arrow: influences; b – expression of PTGS elements from subfigure (a) – x-axis: categorical: root response, root recovery, shoot response, and shoot recovery; y-axis: log₂ fold-change (relative to the IPS1:miR399:PHO2 circuit, i.e. all mimic RNAs are relative to IPS1 fold-change, all miRNAs are relative to miR399 fold-change, and all genes are relative to PHO2 fold-change).
Our data supports the hypothesis that miR164 silences the NAC domain containing protein 6 (NAC6) gene transcript from being translated in the root's recovery from Pi starvation (Figure 3.6). We found no evidence to support the action of target-mimicry. NAC6 is a TF that has been found to be a positive regulator of lateral root development (He et al., 2005). NAC6 gene expression has been shown to be induced by salt stress and responsive to ABA, auxin, and ethylene treatment (He et al., 2005). Given our evidence of NAC6 and miR164's response to Pi starvation, that NAC6 is a positive regulator of lateral root development, and, that Pi starvation induces lateral root development, we suggest that NAC6 and miR164 may be involved in root development during abiotic stress.

In addition to targeting PHO2, miR399d/f were predicted to silence raffinose synthase 4 (RS4) during the root's response and recovery to Pi starvation (Figure 3.7). We found no evidence to support target-mimicry affecting this circuit. The RS family is traditionally associated with osmotic regulation in seed and leaf biology (Nishizawa et al., 2008). More recently, the RS family has been associated with plant stress such as cold, heat, drought, and saline conditions (Zuther et al., 2004; Nishizawa et al., 2008), which give rise to oxidative damage (Bowler et al., 1992; Foyer et al., 1994; Alscher et al., 1997; Shigeoka et al., 2002). Since RS4 synthesizes an antioxidant (Nishizawa et al., 2008), and RS4 is promoted during the Pi starvation recovery process in root (p-value of 1.5E-5, see data described in Chapter 2), we hypothesize that RS4 is involved in recovery from oxidative damage caused by Pi starvation. Furthermore, our results suggest that
miR399d/f has evolved to silence both PHO2 and RS4. Thus, miR399d/f, a Pi starvation responsive miRNA, putatively regulates both Pi transport and recovery from Pi starvation. Of note, miR399d/f:RS4 is not the first circuit identified by our study to be implicated in recovery from oxidative damage; in the previous section, we had already described similar results for the miR398:CSD1/2 circuit. Thus, ROS metabolism appears to be a theme in Pi starvation root recovery. Interestingly, Joo et al., (2001) (Joo et al., 2001) reported that ROS promoted gravitropic root remodeling and that antioxidants such as vitamin-C and Trolox inhibit it. Thus, we suggest that both miR398 and miR399 play a role in root remodeling during Pi starvation by adjusting ROS concentrations.

Figure 3.6: Experimental evidence for the miR164:NAC6 circuit

*a – predicted miRNA circuit, Nodes – green flag: miRNA, blue oval: gene, white box: system process; Edges – black: known, red: predicted by this study, solid arrow:
induction, empty arrow: inhibition, open arrow: influences; b – expression of PTGS elements from subfigure (a) – x-axis: categorical: root response, root recovery, shoot response, and shoot recovery; y-axis: log$_2$ fold-change (relative to the IPS1:miR399:PHO2 circuit, i.e. all mimic RNAs are relative to IPS1 fold-change, all miRNAs are relative to miR399 fold-change, and all genes are relative to PHO2 fold-change).

**Figure 3.7**: Experimental evidence for the miR399:RS4 circuit

*a* – predicted miRNA circuit, **Nodes** – green flag: miRNA, blue oval: gene, white hexagon: metabolite; **Edges** – black: known, red: predicted by this study, solid arrow: induction, empty arrow: inhibition, open arrow: influences; b – expression of PTGS elements from subfigure (a) – x-axis: categorical: root response, root recovery, shoot response, and shoot recovery; y-axis: log$_2$ fold-change (relative to the IPS1:miR399:PHO2 circuit, i.e. all mimic RNAs are relative to IPS1 fold-change, all miRNAs are relative to miR399 fold-change, and all genes are relative to PHO2 fold-change).
The results above have already described three putative root remodeling circuits: 1/miR398:CSD1/2; 2/miR164:NAC6; and 3/miR399d/f:RS4. Here, we present evidence for another putative root remodeling circuit (miR157d:TMO7) that is active during root recovery from Pi starvation (Figure 3.8). Target of monopteros 7 (TMO7) is a non-cell-autonomous basic helix-loop-helix TF involved in monopteros-dependent (MP) embryonic root initiation (Schlereth et al., 2010). TMO7 has been shown to be downstream of cytokinin (Hoth et al., 2003) and upstream of gibberellin (Lee, 2006b) and brassinosteroid (Wang, 2009a) dependent regulation, which are known for their role in growth and development. Zhang et al., (2009) proposed that the large TMO family evolved to regulate cell elongation in root under different environmental stimuli. Given that TMO7 is the only Pi starvation responsive gene of its six family members (At1g74500, At3g25710, At4g23750, At5g60200, At2g28510, At3g45610), the Zhang et al., (2009) hypothesis supports the notion that TMO7 has evolved to have a role in Pi starvation root remodeling. However, TMO7 is not specific to Pi starvation stress and has been shown to be differentially expressed in several other stress conditions, including osmotic stress (see Chapter 2) (Woo and MacPherson et al., 2012). Since MP is slightly promoted during the Pi starvation response and miR157d is repressed, we hypothesize that TMO7 is putatively modulated by the synergistic behavior of both TF and miRNA during Pi starvation.

Previously, we had reported the putative involvement of the miR398:CSD1/2 circuit in Pi starvation. We then showed that this was not the only miRNA circuit involved in Pi
starvation induced oxidative stress, and that miR399d/f:RS4 may also regulate ROS concentration. Here, we introduce laccase 2 (LAC2) and miR397 as a putative circuit involved in Pi starvation root recovery, miR397:LAC2 (Figure 3.9). Interestingly, both miR398:CSD1/2 and LAC2 circuits are sensitive to copper fluctuations (Abdel-Ghany and Pilon, 1997) suggesting that the two circuits share common upstream regulators. This hypothesis is supported by our results, in that both circuits are shown to be active during the root's recovery from Pi starvation. Moreover, the miR397:LAC2 circuit is a bona-fide circuit in copper limiting stress (Abdel-Ghany and Pilon, 1997), however our study is the first to report its putative function in Pi starvation. Together, these two circuits (miR398:CSD1/2, and miR397:LAC2) suggest that either copper availability is affected by Pi starvation, or that copper stress responsive pathways are co-opted during Pi starvation recovery in root.

In analyzing the putative miRNA circuits it became clear that there were several regulatory themes involved in Pi starvation that had not been explored before. For instance, miRNA circuits are generally involved in root remodeling and recovery from oxidative damage. Indeed, ROS has been linked to root remodeling, and so root remodeling is a dominant theme of miRNA-mediated regulation in Pi starvation. Interestingly, most miRNA circuits were predicted to be active during the root's recovery from Pi starvation. However, the canonical IPS1:miR399:PHO2 circuit was found active in all organs and all phases, except shoot recovery. Indeed, only circuits involving miR399 were predicted by us to be active outside of the root's recovery. This suggests
that miR399 is the primary Pi starvation responsive miRNA, and that the other miRNAs (miR157, miR164, miR169, miR397, miR398, and miR2111) act only to recover from Pi starvation stress in root. Other stress conditions can mimic Pi starvation symptoms such as increased ROS concentrations. Thus, miRNAs miR157, miR164, miR169, miR397, miR398, and miR2111 could function as parts of general stress recovery miRNA circuits. These results have not been presented before. This is likely due to our unique access to Pi starvation root recovery data and, in part, to our analysis taking target-mimicry into account. Our access to tiling-array, microarray, and smRNA deep-sequencing data was also a unique advantage.

Micro-RNAs are well known for their role in stress response pathways, which is why we focused on them above. However, another class of small non-coding RNA is also known for its role in several stress responses, the trans-acting small-interfering RNAs (tasiRNA). The following results will analyze tasiRNA circuits known to respond to stress and describe how they behave in our Pi starvation data.
Figure 3.8: Experimental evidence for the miR157:TMO7 circuit

a – predicted miRNA circuit, Nodes – green flag: miRNA, blue oval: gene, white box: system process; Edges – black: known, red: predicted by this study, solid arrow: induction, empty arrow: inhibition, open arrow: influences; b – expression of PTGS elements from subfigure (a) – x-axis: categorical: root response, root recovery, shoot response, and shoot recovery; y-axis: log₂ fold-change (relative to the IPS1:miR399:PHO2 circuit, i.e. all mimic RNAs are relative to IPS1 fold-change, all miRNAs are relative to miR399 fold-change, and all genes are relative to PHO2 fold-change).
Figure 3.9: Experimental evidence for the miR397:LAC2 circuit

a – predicted miRNA circuit, Nodes – green flag: miRNA, blue oval: gene, white hexagon: metabolite; Edges – black: known, red: predicted by this study, solid arrow: induction, empty arrow: inhibition, open arrow: influences; b – expression of PTGS elements from subfigure (a) – x-axis: categorical: root response, root recovery, shoot response, and shoot recovery; y-axis: log₂ fold-change (relative to the IPS1:miR399:PHO2 circuit, i.e. all mimic RNAs are relative to IPS1 fold-change, all miRNAs are relative to miR399 fold-change, and all genes are relative to PHO2 fold-change).
Trans-acting small-interfering RNA in plant stress

Trans-acting siRNA 1 (TAS1) and Trans-acting siRNA 2 (TAS2) are involved in hypoxia and cold stress (Moldovan et al., 2010a/b; Tang et al., 2012). A report made in 2009 found TAS4 to be differentially regulated during Pi starvation in the leafy parts of the plant (Hsieh et al., 2009). However, since then, no additional connection between tasiRNA and Pi starvation has been reported. In the previous section, we observed that several miRNAs responsive to Pi starvation were likely involved in altering root morphology. Of the four canonical tasiRNAs, only TAS3 has been linked to regulating plant morphology: TAS3 regulates leaf morphology and lateral root development through argonaute 7 (AGO7) (Adenot et al., 2006) and auxin response factors (ARF) (Marin et al., 2010), respectively. Therefore, tasiRNAs are involved in plant stress. However, less is known about tasiRNA pathways than what is known about miRNA pathways. Moreover, the four canonical tasiRNAs represent only a fraction of the total tasiRNAs thought to exist in the A. thaliana genome (Chen et al., 2007). There is, therefore, large scope for exploring potential roles of tasiRNAs in plant stress that has not been adequately addressed. Thus, the aim of the analyses presented below is a step in this direction.

TAS3 and TAS4: Given TAS3’s role in root morphology, it is a promising tasiRNA candidate for mediating Pi starvation induced lateral root development. TAS4’s role in Pi starvation has been shown to be involved in anthocyanin production by targeting purple acid phosphatase 1 (PAP1) (Hsieh et al., 2009). Our tiling-array data shows that TAS3 and TAS4 do not change their expression levels during or post Pi starvation. Hence, we
could not demonstrate a causal relationship between TAS3 expression and Pi starvation induced changes in root morphology. In the case of TAS4’s target, PAP1 transcripts are increased 2-fold in the shoot response to Pi starvation, as expected in lieu of TAS4 activity. Therefore, our results support Hsieh et al.’s (2009) finding that TAS4 has a role in Pi starvation biology. More importantly, the example set by TAS4 demonstrates that our tiling-array and microarray data can be used to infer tasiRNA circuitry.

TAS1b and TAS2: Both TAS1b and TAS2 are found to be differentially regulated in our tiling-array data. TAS1b displayed an 8-fold decrease in transcript levels during the root’s response to Pi starvation. Furthermore, TAS1b's response was reversed during the Pi starvation recovery phase and increased 8-fold (Figure 3.10a). In contrast to TAS1b, TAS2 tiling-array probes exhibited a lot of noise and were only slightly differentially expressed. However, TAS2 bares mentioning, as both the level of noise and degree of expression were markedly better than what was observed for TAS3 and TAS4 tiling-array data (Figure 3.10b). Interestingly, both TAS1b and TAS2 displayed greater differential expression toward their 5’ ends. The opposite phenomenon is observed in microarray data where 3' ends of RNA are more stable, which is why microarray probes are generally placed toward the 3’ end. However, in the case of tasiRNAs, this phenomenon likely reflects active miRNA cleavage, which is required to initiate tasiRNA processing (Allen et al., 2005). These expression profiles for TAS1b and TAS2 suggest that their 5' siRNAs are being differentially regulated by Pi starvation. Unfortunately, none of the known TAS1b and TAS2 targets are well annotated in terms of their biological function.
However, the hypoxia study carried out by Moldovan et al. (2009) suggests that the targets of TAS1 and TAS2 are related to mitochondrial ROS and Ca2+ signaling via members of the pentatricopeptide repeat protein (PPR) family. Moldovan et al. (2009) further hypothesize a model which suggests that TAS1 and TAS2 may influence root morphology. In the previous section, several miRNAs were listed that may also play a role in remodeling root morphology via ROS in a Pi starvation responsive manner. Finally, apetala 2 (AP2) has been previously predicted to be a target of TAS1b and, in our data, is inversely correlated with TAS1b's response to and recovery from Pi starvation. Of note, members of the AP2 family are differentially expressed in hypoxic environments (Yang et al., 2011), which supports a TAS1/2:AP2 circuit given that TAS1 and TAS2 are also differentially expressed in such conditions (Moldovan et al., 2010a/b). Taken together, these results suggest that TAS1b/2:AP2 is a Pi starvation and hypoxia responsive circuit in root.

In the following set of results we analyzed tasiRNA's role in Pi starvation from the perspective of the tasiRNA-derived siRNAs themselves. To do this, we made use of smRNA deep-sequencing libraries from double-stranded RNA (dsRNA) binding protein 4 (DRB4) wild-type, knockout, and over-expresser lines (data donated, see Materials and Methods). DRB4 was chosen as it is required for the biogenesis of tasiRNA-derived siRNAs (Adenot et al., 2006). We then mapped differentially regulated siRNAs to known and putative trans-acting siRNA (TAS) loci to determine the set of tasiRNA-derived
siRNAs. Finally, we performed a functional analysis on tasiRNA-derived siRNA targets differentially expressed in our Pi starvation data.
Figure 3.10: Tiling-array data from Pi starvation experiments for TAS1b and TAS2

a – probes are arrayed from 5’ to 3’ on the reference strand, however TAS1b is located on the opposite strand;
b - probes are arrayed from 5’ to 3’ on the reference strand, however TAS2 is
located on the opposite strand, moving average (window-size = 5 probes) is applied to smooth the data in order to increase the signal-to-noise ratio.

**DRB4-dependent tasiRNA-derived siRNAs target Pi starvation responsive loci**

In the previous section we analyzed known tasiRNA genes and their targets for possible roles in the Pi starvation response by using tiling-array and microarray data. Previous studies have predicted that there are many tasiRNA genes in Arabidopsis (Chen et al., 2007). In particular, Chen et al., (2007) used DICER-like protein 4 (DCL4) smRNA deep-sequencing libraries to determine 30 clusters of DCL4-dependent phased smRNAs, i.e. tasiRNA-like genes (Chen et al., 2007).

In our analysis we used DRB4 knockout and over-expresser smRNA deep-sequencing libraries to determine 88 clusters of DRB4-dependent phased smRNAs, i.e. tasiRNA-like genes (see Materials and Methods). Two of the clusters predicted by our method were known TAS genes, TAS1c and TAS2. We then combined Chen et al.'s (2007) predicted tasiRNA genes with our own (total of 114 unique clusters), mapped all DRB4-dependent smRNAs to each cluster, and annotated all smRNAs overlapping a cluster as putative tasiRNA-derived siRNAs (see Materials and Methods section). We then used the Plant smRNA Target server (psRNA-TS) to predict putative target transcripts from the Arabidopsis genome (Dai and Zhao, 2011). Clusters that did not overlap siRNAs with a predicted target were discarded. The remaining 94 clusters were considered putative tasiRNA genes as they contained phased smRNAs whose expression were dependent on DRB4 activity and were predicted to cleave target mRNA or inhibit translation.
Our analysis of DRB4 smRNA data resulted in a network of putative tasiRNA genes and their targets. This putative tasiRNA network can be useful for different tasiRNA related studies. However, we aimed to use the network to discover the role of tasiRNA in plant stress. Thus we analyzed the set of predicted tasiRNA targets for differential expression in our microarray data. Using a fold-change threshold of two we found that 16 targets were differentially expressed in both the root's response and recovery to Pi starvation. Nine targets were responsive to Pi starvation in root, and, 21 were differentially expressed during the root's recovery from Pi starvation. In shoot data, fewer targets were differentially expressed. We found three targets that responded and recovered from Pi starvation, three targets were only responsive, and two targets were differentially expressed in the shoot’s recovery from Pi starvation.

Among our predicted tasiRNA targets, purple acid phosphatase 1 (PAP1) is a known Pi starvation responsive gene (Woo and MacPherson et al., 2012). The DRB4 data suggests that nitrate reductase 1 (NR1) encodes a tasiRNA locus responsible for generating siRNAs targeting PAP1. In our microarray data, PAP1 responds to Pi starvation by increasing expression more than 4-fold before recovering to basal levels (see Chapter 2). Moreover, NR1 is differentially expressed in our Pi starvation data in a manner opposite to PAP1 which supports our hypothesis that NR1 regulates PAP1 via a tasiRNA biogenesis pathway. Interestingly, PAP1 has been recently suggested as a target of tasiRNA guided regulation in apple (Xia et al., 2012). Our results reflect this hypothesis
in Arabidopsis and suggest that NR1-mediated PAP1 tasiRNA regulation is a conserved pathway.

Surprisingly, we found that TMO7, a gene that we had previously implicated to be regulated by miR157d during root Pi starvation, was also predicted to be regulated by tasiRNA generated from an intergenic region near position 1967672 of chromosome three. We hypothesized that miR157d may be regulating TMO7 via tasiRNA. However, the predicted tasiRNA gene did not appear to harbor an miR157d cleavage site (as detected by psRNA-TS) (Dai and Zhao, 2011). Therefore, TMO7 may be post-transcriptionally regulated by two distinct siRNA pathways.

Discussion
The biological functions of small non-coding RNAs have been a topic of scientific interest since their discovery in nematodes, in 1991 (Wightman et al., 1991). It is now well established that smRNAs engage in transcriptional gene silencing (TGS) and post-TGS (PTGS). TGS and PTGS are general regulatory processes and take part in diverse biological pathways. It is generally thought that smRNAs have a role in the adaptive responses demonstrated by plants to ever changing conditions. In this chapter the role of PTGS in plant stress was investigated using data from a recently published Pi starvation study (Woo and MacPherson et al., 2012) and smRNA data donated by the Plant Molecular Biology laboratory at Rockefeller University, NYC, NY, USA. PTGS is caused by sequence-specific recognition of target mRNAs by several proteins and two
classes of smRNA, namely miRNA and siRNA. Upon recognition of target mRNAs, mRNAs are either degraded or their translation blocked depending on the level of complementarity between target and smRNA. Both miRNA and siRNA have conserved pathways in plants and both classes have been found to play a role in Pi starvation stress. Interestingly, long non-coding RNAs have also been found to function in PTGS-mediated plant stress responses; as exemplified by the Pi starvation inducible IPS1, which acts to mimic miR399’s target and thus competitively quench the miRNAs silencing activity. Given these examples of PTGS and the extent of the physiological and chemical changes that plants undergo when subjected to Pi starvation, Pi starvation offers a unique perspective into the role of non-coding RNAs in plant stress.

This chapter was separated into two parts. In the first part we mapped Pi starvation expression data from microarray, tiling-array, and smRNA deep-sequencing to 5057 previously predicted mimic:miRNA:target circuits. This mapping enabled us to select nine miRNA:target circuits for which the expression data supported PTGS activity. Similarly, two mimic:miRNA:target circuits were identified, including the IPS1:miR399f:PHO2 circuit. In the second part, we analyzed four TAS genes (TAS1-4) for potential PTGS activity and showed that TAS1b and TAS2 were influenced by Pi starvation stress. Since TAS genes actually represent clusters of phased siRNAs, with each siRNA potentially differentially regulated, their expression is difficult to ascertain using tiling-array data. Hence, we made use of DRB4 smRNA deep-sequencing libraries to predict novel TAS genes using similar methods as previously described for DCL4
smRNA deep-sequencing libraries (Chen et al., 2007). Pooling our predicted TAS genes with those previously predicted (Chen et al., 2007) yielded a set of 94 putative TAS genes. We analyzed our gene expression profiles for the targets of these 94 TAS genes using our microarray data and found support for several tasiRNA circuits being involved in Pi starvation. Two of our predicted tasiRNA circuits may play a role in root remodeling, one via TMO7, and the other through modulating ROS concentrations.

Taken together, this study, in addition to the well-studied IPS1:miR399:PHO2 circuit, identified 12 smRNA circuits likely to be involved in Pi starvation. These circuits are, miR2111:TPS11 (involved in carbon signaling), miR398:CSD1/2 (involved in ROS modulation), miR399:RS4 (involved in ROS modulation), miR399:NF-YA10 (involved in ROS modulation via anthocyanin biosynthesis), miR164:NAC6 (involved in root development), miR157:TMO7 (involved in root development), miR157:PSB28 (involved in Pi recycling), miR397:LAC2, TAS gene Chr3_1967672:TMO7 (involved in root development), NR1:PAP1 (involved in ROS modulation via anthocyanin biosynthesis), and TAS4:PAP1 (involved in ROS modulation via anthocyanin biosynthesis).

We observed several general trends in the behavior and function of the 12 predicted smRNA circuits: 1/ only circuits involving miR399 appeared to be systemically active, while all other circuits were predicted to be active in a root specific manner; 2/ all circuits with exception to a few miR399 circuits (including miR399:PHO2) were only
active during recovery from Pi starvation; 3/ several circuits were involved in remodeling the root; and 4/ several circuits were involved in modulating ROS concentrations.

In an attempted to place our findings in the context of the established model on the molecular response of Arabidopsis to Pi starvation, we compiled our findings with previously published work into one large model (Figure 3.11). The current established model on Pi starvation considers both a local (root) and systemic response. A recent, 2011, review article in Current Opinions in Plant Biology (Abel et al., 2011) highlighted the need to establish an understanding of the link between the global and local Pi starvation response. The article further proposed the Pi starvation responsive MYB-TF, PHR1, as the likely candidate for cross-talk between the local and systemic response. On the other hand, hormonal regulation via carbon signaling appears to be another link between the local root response and the systemic response (Hammond and White, 2011). As of yet, a direct genetic link between the two responses has not been found. Here, we propose that PTGS may provide at least one of the means by which the local and systemic Pi starvation responses communicate. Specifically, we find that PHR1-regulated miR399 putatively targets RS4, which is responsible for synthesis of raffinose (an antioxidant) and may lead to ROS-dependent root remodeling. Similarly, miR398 and TAS4 have a PHR1 binding element (P1BS) in their promoter regions (Hsieh et al., 2009). We predict miR398 and TAS4 to target two superoxide dismutases (CSD1/2) and PAP1, respectively. Micro-RNA169 also contains and P1BS element in its promoter and is known to be involved in oxidative stress (Hsieh et al., 2009). We found that miR169
targets a novel Pi starvation induced gene that we have renamed, induced by Pi starvation 2 (IPS2). Furthermore, we identified a likely candidate for miR169 target-mimicry. The target-mimic is a novel long non-coding RNA which we have found to be repressed by Pi starvation and have thus renamed it to, repressed by Pi starvation 1 (RPS1). Therefore, we have identified 4 putative siRNA circuits (two of which are regulated by target-mimicry) that link the systemic PHR1 TF to the local response via ROS-dependent signaling. On the other hand, our data suggests the activity of miR2111 targeting TPS11 in root. TPS11 depletes sucrose to form starch (Singh et al., 2011). Therefore, miR2111 activity could influence carbon signaling during Pi starvation. In summary, our analysis demonstrates plausible smRNA circuitry for PTGS-mediated crosstalk between the local and systemic response to Pi starvation. Therefore, PTGS appears to be a part of both of the previously suggested mechanisms of cross-talk between the local and systemic response to Pi starvation.
Figure 3.11: Model of PTGS in Pi starvation stress (summary of findings)

Nodes – green flag: miRNA, red half flag: targe-mimic, blue oval: protein coding gene, multi-blue ovals: gene family, multi-yellow flags: tasiRNAs gene, white box: system process, white hexagon: amount/concentration/metabolite, white pentagon: sensing,
**Conclusion**

PTGS relies on the production of short RNA sequences to guide its silencing machinery. Such short oligomers are easily produced and provide enough complexity for sequence specific complementarity to occur. PTGS is, therefore, a highly adaptable mechanism of genetic regulation. For this reason, PTGS has been proposed as a component of the adaptive responses exhibited by plants during periods of stress. Perhaps more importantly, artificial PTGS has already been demonstrated as a viable means of engineering agriculturally favorable plant phenotypes. Hence, the network of PTGS miRNA and tasiRNA circuitry proposed in this study will be a useful asset in determining the exact role of small non-coding RNAs in Pi homeostasis. Moreover, recent studies are finding connections between PTGS and the plants response to several stress conditions. Considering that there are far fewer pertinent siRNAs than protein coding genes, then unraveling the siRNA stress responsive network is of great importance to genetically engineering stress tolerant crops. However, even though sequence similarity and expression profiles were used to make our predictions, we must emphasize that actual PTGS has yet to be demonstrated. Therefore, these 11 putative circuits should be biochemically tested in Pi starvation stress, and indeed in other stress conditions where oxidative and carbon signaling play a role.
Final remarks

The study of non-coding RNA started in 1990, after the discovery of a new class of RNA in nematodes. The new RNA class turned out to be what we recognize today as miRNA, a subclass of smRNAs. The same smRNAs were discovered shortly thereafter in plants. Indeed, we now know that all higher life forms use smRNAs to regulate the expression of specific genes. Yet, plants were among the few model organisms that helped elucidate much on smRNA biology. One of the reasons for this being that plants require the exact, or near-exact complementarity between miRNA and target mRNA. This property has allowed for the accurate prediction of miRNA circuitry in silico before applying costly biochemical methods. Furthermore, our collective knowledge of plant biochemical pathways is impressive. Thus, since smRNA circuits target protein coding genes through PTGS, then prior knowledge of target gene biology can help to elucidate on the reasons why smRNAs silence particular targets, in particular tissues, and at particular times. The goal of our research has been to take advantage of accurate in silico predictions of PTGS circuitry, and prior knowledge of the mRNA targets in order to elucidate on PTGS's role in the context of plant stress. Many plant stress conditions have recently been observed to elicit several PTGS pathways. Likewise, several PTGS pathways are commonly observed across multiple stress responses. Given that there are
far fewer functionally recognized smRNAs than there are protein coding genes, the analysis of PTGS in plant stress could represent an efficient means of addressing crop tolerance through genetic modification. It was therefore the primary aim of this thesis to provide a network of smRNA circuitry that could (A) explain in part, the role of non-coding RNA in plant stress; and (B) predict a list of PTGS circuits for further research as candidates for improving crop tolerance. To do this, we focused on Pi starvation stress for its previously demonstrated PTGS response by miR399 and TAS4.

The economic impetus underlying plant stress research is exceptionally large -- reports have shown a 20 to 30 percent decrease in crop yield, or even total loss, due to environmental factors that could arise from nutritional, biotic, or abiotic stresses (Zhang et al., 2000; Mahajan and Tuteja, 2005; Vinocur and Altman, 2005; Yamaguchi and Blumwald, 2005). As mentioned above, genetic modification of smRNA circuits has the potential to alleviate crop loss caused by environmental factors. Currently, several miRNAs are becoming interesting candidates for stress tolerance research in plants (Zhang et al., 2005). Over 100 miRNAs in Arabidopsis are known to post transcriptionally regulate the abundance and/or translation of mRNA targets during plant development and stress (Jones-Rhoades et al., 2006; Zhang et al., 2006b; Zhang et al., 2007). In particular, miR395 abundance has been shown to respond to sulfate starvation (Zhang et al., 2006b). Similarly, for abiotic stress, the abundance of miR319 and miR402 transcripts responded to cold, and miR402 responded to drought and salt stress (Sunkar and Zhu, 2004). Micro-RNA402 demonstrates that miRNAs respond to multiple stress
conditions. Indeed, several review articles have compiled lists of miRNA stress responsive circuits. In our study such lists of miRNA circuitry were employed to analyze their behavior in the context of Pi starvation. Pi starvation was chosen for several reasons: 1/ Pi starvation responsive pathways are well studied; 2/ Pi starvation elicits large morphological as well as metabolic changes in plants, meaning that identifying PTGS-derived cross-talk with other stress responses is more likely; 3/ Pi is an economically important macro-nutrient commonly supplemented in fertilizers; and 4/ the Pi starvation response has been observed to elicit TAS4 and miR399 smRNA circuitry for which the later miRNA circuit is very well described.

To measure the transcriptomic response of Arabidopsis to Pi starvation, three datasets were collected: 1/ 18 microarrays; 2/ 18 tiling-arrays; and 3/ 18 smRNA libraries. The tiling-array and microarray datasets were described in detail in Chapter 2. The smRNA dataset (donated for use in our research) was derived from the same experimental protocols as described in Chapter 2. The smRNA dataset together with DRB4 WT and mutant smRNA libraries provided important assets for the analysis of smRNA PTGS circuits in Pi starvation. Altogether, the research described in this thesis integrates over 100 experiments from three different high throughput transcriptomic technologies in order to describe 11 novel PTGS circuits involved in Pi starvation, and potentially other stress conditions.
In Chapter 2, both coding and non-coding genes were analyzed for their response to, and recovery from Pi starvation. The two primary aims of Chapter 2's research were to (A) analyze the transcriptomic response and recovery of root tissues to Pi starvation, and (B) to provide high quality data we used in Chapter 3. The results presented in Chapter 2 were published in 2012 (Woo and MacPherson et al., 2012) and describe the first high-throughput experiments to analyze the root's response and recovery to Pi starvation. Given that Pi starvation elicits both a local root and a systemic response, the results in Chapter 2 provide the first genome-wide data available for use in studying the local root response. Moreover, tiling-array data is presented for the first time in Pi starvation research and provides data on genomic regions previously only studied in low-throughput experiments. For example, we detected the response of GDPD3 and CPL3 to Pi starvation, for which we hypothesize have roles in membrane remodeling and root architecture, respectively. In summary of Chapter 2, 89 genes were discovered to be differentially expressed in both shoot and root. Therefore, given that over 1000 genes were found to be differentially expressed, the shoot and root responses are largely distinct from one another. For example, it was found that DGD2, PAP6, and several members of the PHT1 family (PHT1;1/3/5/8/9) were specifically involved in the root's response and recovery to Pi starvation. We also found that the root recovery process differentially expresses several fold more TFs than the initial response to Pi starvation. After comparing our data with other stress data from AtGenExpress, we found that 70% of our root responsive genes did not respond to any treatment other than Pi starvation. Surprisingly, known Pi starvation responsive genes were responsive in
other AtGenExpress treatments, suggesting that they had not evolved to respond to Pi starvation stress exclusively. This non-exclusive use served as a basis to assay which stress conditions elicited similar responses to Pi starvation. We found that biotic infection by Botrytis cineria yielded the most similarity, followed by cold, salt, and osmotic stress. Finally, we found that of all the hormones, ABA treatment elicited the most similar response to Pi starvation. ABA treatment was also similar to many other stress conditions from AtGenExpress, which suggests that ABA is an important mediator in the general stress response. Altogether, Chapter 2 provides novel data in the form of genome-wide gene expression data for the root's response and recovery to Pi starvation. While shoot data was included, the fact that the genome-wide root response had not been reported before was important, especially considering the local root response to Pi starvation. One of the current questions regarding the Pi starvation response is on the nature of the interaction between the local and systemic response. Chapter 2 provides a description for a unique dataset that could be used to explore this and other root related questions. Chapter 3 described research that used the data presented in Chapter 2 to suggest how PTGS could mediate cross-talk between the local and systemic response.

The primary aim of Chapter 3 was to identify elements of PTGS that function in the plant's response to Pi starvation stress. Analysis was restricted to miRNAs and their closely related tasiRNAs. We identified the presence of miRNA:target circuits that were hypothetically active during Pi starvation stress. Several putative miRNA circuits were
discovered. However, we focused our analysis on those circuits whose targets were well annotated and showed marked differential expression in response to Pi starvation. These were: IPS1:miR399:PHO2, miR399:RS4, miR399:NF-YA10, miR398:CSD1/2 (known but not connected to Pi starvation before), miR2111:TPS11, miR164:NAC6, miR157:TMO7, miR157:PSB28, RPS2:miR169:IPS2, and miR397:LAC2. Interestingly, P1SB binding-sites for the Pi starvation inducible TF, PHR1, have been found in miR399, miR398, miR169, and TAS4 gene promoters suggesting that PHR1 regulates the transcription of several PTGS elements. Indeed, TAS4 has been shown to be Pi starvation responsive and our expression data supports the activity of TAS4, TAS1b, and TAS2 during Pi starvation. Moreover, we identified two novel TAS genes as part of the following circuits: NR1:PAP1, and Chr3_1967672:TMO7. Altogether, results of Chapter 3 suggested that the local root response to Pi starvation was influenced by the plant's systemic response pathways via PHR1-mediated PTGS. We hypothesized that PHR1-mediated PTGS influences the local root response by modulating concentrations of ROS. We further hypothesized that miR2111 is capable of influencing shoot to root carbon signaling by silencing TPS11. Thus, Chapter 3’s results provide testable hypotheses for how the two systemic and local responses to Pi starvation interact.

The combined results presented in Chapters 2 and 3 provide the first genome-wide evidence of the plant's root response and recovery to Pi starvation. Thus, predictions of novel roles for smRNAs during Pi starvation were possible. In Chapter 2, cross-talk was observed between Pi starvation and biotic, cold, salt, and drought stress. Interestingly,
these stress conditions are known for eliciting oxidative stress (Shulaev et al., 2008). This suggests that oxidative stress could be representative of an underlying general stress response. Indeed, Chapter 3 gives some discussion on the role of ROS in root remodeling, a process observed in plant responses to several stress conditions. Chapter 3 also suggests several hypotheses for how miRNAs and tasiRNAs could mediate cross-talk between Pi starvation of oxidative stress. Moreover, recent reviews have highlighted that PTGS is observed across a broad spectrum of biotic and abiotic stress conditions (Khraiwesh et al., 2012). Analysis of cross-talk and the general stress response is an essential step in engineering stress tolerant crops. Understanding the network of metabolic changes of plants in response to multiple stress conditions is not a simple task. Yet, it is a necessary task as plants are constantly exposed to ever changing and overlapping biotic and abiotic environments in nature. For example, drought is often coupled with heat stress, salt stress with osmotic stress and hypoxia, Pi starvation with metal stress. Thus, if we could identify and understand the common genetic elements that govern these stress responses, we could engineer genetically enhanced crops with less trial and error.

Plant stress elicits a change in the steady state of the metabolome, while the recovery from the stress conditions is the process of obtaining a new equilibrium (Shulaev et al., 2008). The results in this thesis ultimately provide a list of candidate genetic elements capable of mediating cross-talk between stress conditions and the recovery process by modulating metabolites, in particular ROS and sucrose concentrations during Pi
starvation. Therefore, future work should focus on verifying the listed candidates as actual regulators of ROS and sucrose concentrations during stress. Once verified, the model of Pi starvation PTGS presented in this thesis will provide a good basis for studying the effects of metabolites on plant stress, in general.


and carnitine


Kurihara, Y., Takashi, Y., and Watanabe, Y. (2006). The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. RNA 12, 206–212.


Materials and Methods for Chapter 2

Please refer to Woo and MacPherson et al., (2012) for a complete description of the Materials and Methods of the research presented in Chapter 2.
Materials and Methods for Chapter 3

Data-sources
The following data-sources were used during the research presented in Chapter 3:

1/ 78 miRNA:target relationships as well as miRNA:stress-condition relationships were extracted from Kraiwesh et al., (2012).

2/ 5,057 mimic:miRNA:target relationships were predicted and donated by Dr. Huan Wang, of PMB-RU*. The data is unpublished. For each mimic:miRNA:target relationship, the response and recovery of the miRNA to Pi starvation was reported for root and shoot tissues derived from Arabidopsis thaliana. This smRNA expression data was obtained during the same experiments as described in Chapter 2, using the same experimental design. Furthermore, this smRNA data is derived from the same biological samples as used for Chapter 2’s experiments. Therefore, the smRNA miRNA data is linked to the microarray and tiling-array data.

3/ DRB4 smRNA data was donated by Dr. Anna Jakubiec of PMB-RU* for use in this research project. The data is unpublished. The DRB4 smRNA data contains smRNA
sequencing data (IP-Seq) for wild-type DRB4 (WT-IP-Seq), drb4 mutant (drb4-IP-Seq), and DRB4 over-expresser (DRB4OE-IP-Seq) libraries. This data was received unprocessed. Processing steps are outlined at the end of this materials and methods section under, "DRB4 smRNA data processing".

4/ The local and systemic response to Pi starvation presented in Figure 3.11 was adapted from Abel S., (2011).

* - Plant Molecular Biology laboratory at Rockefeller University, NYC, NY, USA (Laboratory head: Professor Nam-Hai Chua).

**Automatic hypothesis generation**
Considering the mimic:miRNA:target relationship, we defined nine hypothesis to explain PTGS activity given only fold-change data. These hypotheses were (see Table 3.3, Chapter 3):

- **Hypothesis 1 (H1):** The miRNA's expression is increased, but the increased mimic expression is masking the miRNA's silencing activity and resulting in increased target abundance.
- **Hypothesis 2.1 (H2.1, synergism):** an increase of mimic and decrease of miRNA expression contribute to the increased abundance of the target.
• Hypothesis 2.2 (H2.2): A decreased expression of the miRNA results in increased abundance of the target. The mimic is differentially expressed, but has no observable effect.

• Hypothesis 2.3 (H2.3): A decreased expression of the miRNA results in increased expression of the target. The mimic is basally expressed.

• Hypothesis 3.1 (H3.1): The miRNA is increased and decreases target abundance. The mimic is differentially expressed, but has no observable effect.

• Hypothesis 3.2 (H3.2): The miRNA is increased and quenches both mimic and target, which are both decreased.

• Hypothesis 3.3 (H3.3): The miRNA is increased and silences the target, which is decreased. The mimic is basally regulated.

• Hypothesis 4 (H4): The mimic and miRNA are both increased, whereas the target remains basally regulated. The miRNA's silencing activity on the target is being masked by the mimic.

• Hypothesis 5 (H5): The fold-change data is too ambiguous to discern any relationship.

Mimic:miRNA:target circuits can be computationally evaluated to determine which hypothesis fits their fold-change data. Mimic, miRNA, and target fold-change were considered to be increased if fold-change was greater-than or equal-to 2-fold. Similarly, decreased expression was considered at less-than or equal-to 0.5-fold. And, basal
expression was considered to be between 0.5 and 2 fold. As annotated in table 3.3, increased expression is annotated as "UP", decreased expression as "DOWN", and basal expression as ".". Thus, the fold-change profile of any one mimic:miRNA:target circuit can be expressed in terms of "UP", "DOWN", and ",," and then compared to table 3.3 in order to assign the appropriate hypothesis.

**Analysis of stress responsive circuits**
The 78 miRNA:target circuits extracted from Kairi et al., (2012) were analyzed to determine whether they could possibly play a role in Pi starvation. These 78 circuits were represented by 129 mimic:miRNA:target circuits from our predicted dataset (see "Data-sources" above). To begin, we first collected fold-change data for the Pi starvation root response (RRES), root recovery (RREC), shoot response (SRES), and shoot recovery (SREC) for mimic (tiling-array data, see Chapter 2), miRNA (donated, see "Data-sources" above) and target genes (microarray data, see Chapter 2). Each of the 129 mimic:miRNA:target circuits were then evaluated for possible PTGS activity in Pi starvation by using the four fold-change datasets (RRES, RREC, SRES, SREC) and the automatic hypothesis generation described above.

The 129 circuits described above are part of a much larger set of 5,057 predicted circuits (see "Data-sources" above). Evaluation of the 5,057 circuits was performed in the same manner as described for the 129 circuits.
Prediction of tasiRNA genes, siRNAs, and targets
Initial prediction of tasiRNA genes: DRB4 smRNA data was derived from 3 different genetic backgrounds (see "Data-sources" above). To predict putative tasiRNA genes, we pooled all smRNA libraries together, and mapped all smRNAs to the genome using Bowtie software: http://bowtie-bio.sourceforge.net/index.shtml (with perfect matching). We then found clusters of phased 21-24nt smRNAs and annotated such clusters as putative tasiRNA genes. Within this predicted set we found TAS1c, and TAS2, which confirm that the method is identifying tasiRNA-like genes.

Initial determination of mat-tasiRNAs: Any smRNAs from the pooled DRB4 data mapping to the predicted tasiRNA genes were considered putative mat-tasiRNAs.

Determining mat-tasiRNAs targets: The Plant smRNA Target Server was used to determine the putative targets of the mat-tasiRNAs identified above (Dai and Zhao, 2011). Default parameters were used, except for the maximum expectation (ME) parameter, which controls the level of false-positives predicted. The ME parameter is ranged between 0 and 5, with conservative values (minimum false-positives) suggested between 0 and 2. We set the ME parameter to 1.

Filtering predicted mat-tasiRNAs: All mat-tasiRNAs without a target (other than themselves) were discarded and not considered bona-fide mat-tasiRNAs.
Filtering predicted tasiRNA genes: All tasiRNA genes with bona-fide mat-tasiRNAs were considered true tasiRNA genes.

**DRB4 smRNA data processing**
Raw sequencing data was returned from the Rockefeller University Sequencing-Platform with low-quality reads already filtered out. Data was split into three files consistent with their original sample: WT, drb4, DRB4-OE. The remainder of the data consisted of reads sized 35nt in length. All smRNA reads were scanned for adapter sequence, and trimmed of any such adapter sequence found. This resulted in a distribution of reads sized between 19 and 28nt.

Reads were mapped back to the Arabidopsis genome (version: TAIR9) using the algorithm implemented by Bowtie software: http://bowtie-bio.sourceforge.net. Bowtie was set to consider only mappings without any mismatches at all.