2009

Extracting biologically relevant information from microarray data as related to nitrate response in maize roots and node development in rice

Michael Eugene Miller
Iowa State University

Extracting biologically relevant information from microarray data as related to nitrate response in maize roots and node development in rice

by

Michael Eugene Miller

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Molecular, Cellular, and Developmental Biology

Program of Study Committee:
Patrick S Schnable, Major Professor
Yanhai Yin
James T Colbert

Iowa State University

Ames, Iowa

2009

Dedication

I would like to dedicate this work to my family for their endless love and support, to my lab mates, without whom none of this would have been possible, and to my future students, may the experience gained through this process make me a better guide, mentor, and instructor.

Table of Contents

List of Figures

List of Tables

Chapter 1: General Introduction

Microarray technology and maize genomics

Genomic sequencing and analytical tools have opened new and exciting avenues through which to study the response to environmental stimuli and the process of development. DNA microarrays are one of the most popular tools for performing such studies. These arrays allow for differential analysis of transcription between two or more similar samples exposed to various environmental or developmental stimuli. Key to determining the genomic response to such stimuli is placing the transcriptional changes in some meaningful biological context. The following thesis details two studies seeking to accomplish exactly that: determining the biological response to nitrate exposure in the roots of maize seedlings and sketching the transcriptional changes associated with development in rice.

Genomics has become a major field of scientific interest and has led to the mapping and sequencing of numerous genomes including the relatively recent publications of the maize [1] and rice [2, 3] genomes. As the sequencing of these genomes has progressed, so has the understanding of the composition and interaction of the genes (i.e. functional genomics) held in each. It has been estimated that the genomes of both rice and maize encode roughly 50,000 different expressed genes [4, 5]. Ultimately, the goals are to understand where, when, and how each of these genes is activated and interact with each other, and what functions each performs. Global expression profiling, made possible by the use of technologies such as DNA microarrays, has become one of the first steps for these goals.

There are a number of methods by which researchers can investigate expression values of various genes, including: northern blots, S1 nuclease protection assays, differential display, serial analysis of gene expression (SAGE) [6], and RNA-seq. DNA microarrays and SAGE have been utilized for analyzing transcription levels for thousands of genes in parallel. Microarray technology works by exploiting the ability of an mRNA or cDNA to hybridize specifically to its corresponding DNA template. Generally speaking, there are two types of DNA microarray designs: cDNA and oligonucleotide [7]. In both types of arrays thousands of DNA probes of known sequence are printed on a solid substrate, commonly glass, in a pre-defined order, such that each spot on the array consists of a known sequence at a known position. Each spot, or probe, contains hundreds to thousands of copies of the DNA sequence that can specifically bind to an mRNA or cDNA of complementary sequence. The binding mRNA or cDNA are tagged by fluorescent or radioactive markers. When comparing transcription between two biological samples, separate labels are incorporated into each sample and both are hybridized to the same array. The difference in label intensity at any given spot corresponds to the difference in transcript abundance between the two samples [7]. Thus, transcript levels for the thousands of genes printed on the array can be compared between two biological

samples simultaneously. Such microarrays allow for the exploration of gene expression in at least ways: static and dynamic [6]. Static experiments provide information regarding in which tissue or at what developmental point a gene is expressed, while dynamic studies indicate the interactions between various genes.

While microarrays have the potential to greatly enhance the understanding of genetic interactions, they also pose their own unique sets of challenges and limitations. Chief among the limitations of performing such studies are the need for careful and detailed experimental design and the statistical handling of the results [8]. The design of the microarray experiment must be robust enough to minimize any undesired variance between samples. Variance can be minimized by the specificity with which tissues are harvested and collected, as well as by the inclusion of an adequate number of biological and technical replicates. However, the larger limitation to producing interpretable and/or meaningful results is the statistical handling of the results [8]. Various statistical tests have been performed on microarray results, including two-sample t tests, Welsh t-tests, non-parametric tests, and ANOVA-type methods for determining significant differences in transcript abundance [9]. It is important to exclude as many false-positives as possible from the final results. When considering p-values derived from t-tests it is common practice to include the proportion of false positives using established methods [10]. It has also become increasingly common to build consideration of false discovery rate (FDR) into the statistical analysis by displaying the results as q-values [11]. Both FDR and q-values attempt to control for multiple testing. That is, in a microarray experiment thousands of null hypotheses are being tested simultaneously, potentially resulting in a large number of genes incorrectly called significant. If a microarray experiment tested differential expression among 800 genes using a p-value cutoff of .05, one would expect 40 of these genes to be incorrectly called significant (800 x .05 = 40). This is known as a Type I error. FDRs and q-values attempt to control for this type of error.

After the experiment is properly designed and implemented and appropriate statistics are applied to the results the major challenge becomes determining what the results indicate about the biological underpinnings of the response. A number of pre-existing tools allow the lists of genes to be organized and overlaid on pre-existing biological pathways. However, most of these programs were developed for microbial or animal systems, limiting their usefulness in interpreting microarray data from plant systems as irrelevant pathways are imported and plant-specific pathways and processes are absent [12]. One of the first plant-specific programs for microarray interpretation was created by combining a database of microbial and animal pathways (www.metacyc.org) with the annotated Arabidopsis genome, allowing the exploration of plant biochemical pathways [12, 13]. The concept was later extended by Thimm et al, (2004) in the creation of the MapMan tool that was designed specifically for use with the Affymetrix 22K Arabidopsis array, though this tool can be expanded to other species as genomic annotations become available. This tool allows for the visualization of differentially expressed genes on biochemical pathways,

including those specific to plants, and has increased power over AraCyc in that MapMan can resolve specific members of an enzyme family instead of listing all family members as a single entity.

To date, MapMan has not been adapted to work with the annotations of either the maize or rice genome, and no known systems exists for accomplishing the same tasks in either maize or rice. Interpretation is further complicated by a number of other factors, including; incomplete annotation of genomes and the difficulty by which annotations from one species can be overlain against annotations from another. This means the biological interpretation of microarray results in maize and rice currently require large amounts of manual manipulation.

Introduction to studies

The Center for Plant Genomics at Iowa State University has designed, printed, and implemented a series of cDNA arrays for maize, as well as an oligonucleotide array for rice. In particular, the enclosed reports utilize a maize chip (GEO No. GPL1984) containing 7,888 informative cDNA probes including 10 spots intentionally developed to aid in studying response to nitrate, and a combination of rice oligonucleotide chips containing 23,040 informative probes (GEO Nos. GPL6939 and GPL6940).

As mentioned above, there are two general classes of studies that can be performed using microarray technology: static and dynamic. The two enclosed studies provide examples of both. The first, nitrate response in maize roots, is a more or less static approach to identifying gene activity following exposure to nitrate. This study is static in that it demonstrates if genes become more active in roots or not following exposure to nitrate. Though it should be noted the study does include two time points as well as the two environmental settings, so it also takes into account a certain degree of dynamic change. The focus of the second study is to investigate the dynamic interactions of gene activities over spatiotemporal development in rice. Specifically, the complex series of comparisons made allowed for the interrogation of transcriptional change across spatially separated anatomical positions (in this case, the nodes) as well as across four time points. While this study does give an indication of which genes are active in these particular tissues at a given time (static), the analytical emphasis was placed on evaluating how gene activity changed and elucidating any genetic interactions that may be controlling the process (dynamic).

Thanks to sequence similarity among homologous and orthologous genes, it is possible to assign putative functions to a large set of DNA sequences based on the amount of overlapping identity they share with genes of known function [14]. This ability to assign functions to entire gene sets allows for cross-referencing a list of significant genes with a list of putative functions assigned to those genes. This, in turn, allows for the exploration of genetic activity at the functional level of such gene lists. Such explorations can focus on individual genes in isolation or can focus on genetic groups with similar functions. Here, both methods are utilized to examine the biological implications in the respective studies.

References

1. Pennisi E: **Plant sciences. Corn genomics pops wide open**. *Science* 2008, **319**(5868):1333.
2. Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y, Zheng L *et al*: **The TIGR Rice Genome Annotation Resource: improvements and new features**. *Nucleic Acids Res* 2007, **35**(Database issue):D883-887.
3. Yuan Q, Ouyang S, Wang A, Zhu W, Maiti R, Lin H, Hamilton J, Haas B, Sultana R, Cheung F *et al*: **The institute for genomic research Osa1 rice genome annotation database**. *Plant Physiol* 2005, **138**(1):18-26.
4. Fu Y, Emrich SJ, Guo L, Wen TJ, Ashlock DA, Aluru S, Schnable PS: **Quality assessment of maize assembled genomic islands (MAGIs) and large-scale experimental verification of predicted genes**. *Proc Natl Acad Sci U S A* 2005, **102**(34):12282-12287.
5. Livingstone K, Rieseberg LH: **Rice genomes: a grainy view of future evolutionary research**. *Curr Biol* 2002, **12**(13):R470-471.
6. Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM: **Expression profiling using cDNA microarrays**. *Nat Genet* 1999, **21**(1 Suppl):10-14.
7. Stears RL, Martinsky T, Schena M: **Trends in microarray analysis**. *Nat Med* 2003, **9**(1):140-145.
8. Pan JZ, Eckenhoff RG, Eckenhoff MF: **Limitations of microarray studies**. *Anesth Analg* 2007, **104**(5):1300-1301; author reply 1301-1302.
9. Forster T, Roy D, Ghazal P: **Experiments using microarray technology: limitations and standard operating procedures**. *J Endocrinol* 2003, **178**(2):195-204.
10. Fernando RL, Nettleton D, Southey BR, Dekkers JC, Rothschild MF, Soller M: **Controlling the proportion of false positives in multiple dependent tests**. *Genetics* 2004, **166**(1):611-619.
11. Cheng C, Pounds SB, Boyett JM, Pei D, Kuo ML, Roussel MF: **Statistical significance threshold criteria for analysis of microarray gene expression data**. *Stat Appl Genet Mol Biol* 2004, **3**:Article36.
12. Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Kruger P, Selbig J, Muller LA, Rhee SY, Stitt M: **MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes**. *Plant J* 2004, **37**(6):914-939.
13. Mueller LA, Zhang P, Rhee SY: **AraCyc: a biochemical pathway database for Arabidopsis**. *Plant Physiol* 2003, **132**(2):453-460.
14. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool**. *J Mol Biol* 1990, **215**(3):403-410.

Chapter 2: Global Transcriptomic Responses of Maize Seedling Roots to Nitrate

Michael E Miller[a,b] Sudhansu Dash[c,1], Yan Fu[c,d], and Dan Nettleton[e, f], Patrick S. Schnable[a, b, c, d , f, 2]

[a] Molecular, Cellular and Developmental Biology Graduate Program, Iowa State University, Ames, IA 50011
[b] Department of Genetics, Development and Cell Biology, Iowa State University, Ames, IA 50011
[c] Department of Agronomy, Iowa State University, Ames, Iowa 50011
[d] Center for Carbon Capturing Crops, Iowa State University, Ames, Iowa 50011-3650
e Department of Statistics, , Iowa State University, Ames, Iowa 50011
[f] Center for Plant Genomics, Iowa State University, Ames, Iowa 50011-3650
1. Current address: Virtual Reality Applications Center, Iowa State University, Ames, Iowa 50011-2274
2. Corresponding author: Roy J. Carver Co-Laboratory Iowa State University Ames, IA 50011-3650.
E-mail: schnable@iastate.edu

Abstract

For both economic and environmental reasons it is desirable to reduce levels of nitrogen application to maize, while maintaining yields. Here, a cDNA microarray experiment was conducted to identify maize genes that exhibit differential expression at two time points (early: 0.5 hour and late: 24 hours) following exposure to nitrate. A total of 465 genes exhibited differential expression, 44 at the early time point ($p \leq .05$, $q \leq .27$) and 436 at the late time point ($p \leq .001$; $q \leq ..023$), with 15 exhibiting differential expression at both time points. The findings from this first study of global transcript responses to nitrate in a C4 plant, overall exhibit a high degree of similarity to previous studies performed in the C3 plants Arabidopsis, tomato, and rice, and in addition contribute novel observations into the transcriptomic responses of plants to nitrate. Comparative analyses across both C3 and C4 species highlight the importance of the pentose phosphate, glycolysis, and Calvin cycle pathways in the nitrate response. This study identified novel genes in these pathways as being differentially expressed following nitrate exposure. The current study also suggests that malate dehydrogenase may play a different role in N metabolism in C4 as compared to C3 plants, consistent with differences in metabolism in these species. This study also identified differentially expressed genes in various other pathways with less immediately obvious

connections to nitrogen utilization.

Background

Nitrogen (N), an integral component of every nucleic acid and protein, is one of the most important macronutrients for plant growth. By boosting yields, N-containing fertilizer helps to supply food for at least 40% of the global population [1]. However, N that is not utilized by the crops is vulnerable to loss via volatization, denitrification, and leaching [2]. It has been estimated that roughly 37% of the N applied to maize is recovered in above-ground biomass [2, 3], leaving substantial amounts of N vulnerable to loss. N lost from agricultural systems represents both an economic loss as well as having negative environmental impacts, such as contamination of groundwater [4], eutrophication of surface waters [5], and emission of N-containing greenhouse gases [6].

Worldwide, maize is one of the most important contributors to human caloric intake, along with rice and wheat [2]. In the US, maize receives the largest amount of N-fertilizer among all crops. An average of 135 lbs of N/acre were applied to ~79 million acres of maize in the US between 2000 and 2006 [7]. Driven by demand for ethanol, US maize acreage increased in 2007 by nearly 14 million acres (18%) from the 2000 to 2006 average [7]. Assuming consistent fertilization rates, this corresponds to an increase of over 940,000 tons of applied N. As a well developed genetic model with a draft genome sequence, maize is an appropriate system for developing approaches to reduce N application to cereals. Its large acreage and substantial N-fertilization requirements will help ensure that methods to increase the efficiency of N utilization will have profound economic and ecological impacts.

In soils subjected to annual cropping systems, nitrate (NO_3-) is the largest source of N reaching crops [8]. Several reviews detail the pathways important for nitrate uptake and assimilation [8-11]. Previous studies in Arabidopsis, tomato, and rice have demonstrated that nitrate exposure induces numerous genes involved in the uptake and assimilation of nitrate, as well as more far-ranging effects on transcription and metabolism. These studies also indicate nitrate may have pronounced effects on carbon metabolism as evidenced by changes in expression patterns of genes involved in the pentose phosphate pathway, glycolysis, and the Calvin cycle [12-16]. The existence of interactions between nitrogen and carbon

metabolism is further supported by the association of nitrate applications with a reduction in starch synthesis and increased synthesis of organic acids [17], the acceleration of nitrate uptake by sucrose-feeding [18], and the association of low concentrations of N with decreases in enzymes required for photosynthesis [19].

In this study, microarray technology was used to define the effects of nitrate application on global transcript levels in the roots of maize seedlings. In all, 464 genes represented on the microarray exhibited significantly different transcript levels following exposure to nitrate as compared to control seedlings. The physiological functions of these differentially regulated genes indicate the importance of metabolic processes such as glycolysis and the pentose phosphate pathway, reduction in phenylpropanoid synthesis, and non-symbiotic hemoglobin in the process of nitrate assimilation. Further, the results presented here indicate malate dehydrogenase may play a distinctly different role in N metabolism in C4 compared to C3 plants, and that internal protein turnover may be important to cell response under N-limited conditions.

Results and Discussion

Transcriptomic responses of maize seedling roots to nitrate

To identify genes exhibiting differential regulation in the presence of nitrate, maize seedlings were placed into solutions containing either calcium sulfate or calcium nitrate, and root tissue was harvested at two time points: 30 minutes (early) and 24 hours (late) following treatment (see Methods). Prior to microarray analysis, plant root responses to this N treatment were verified via two independent methods: a nitrate reductase (NR) enzyme assay and qRT-PCR for the uroporphyrinogen III methyltransferase (UPM1) gene (Fig. 2.1). UPM1, which is responsible for synthesizing a subunit of nitrite reductase (NiR), is nitrate-inducible in Arabidopsis [14, 15]. Increases in both NR activity and UPM1 transcript accumulation were observed following nitrate treatment, demonstrating that the induction conditions used in this experiment were suitable for studying the effects of nitrate on gene expression in maize.

Following induction a microarray experiment was conducted (see Methods). 485 cDNA spots exhibited significant differential expression. Forty-four and 456 spots were

differentially expressed in the early and late time points, respectively, and 15 probes exhibited significant expression differences at both time points (Suppl. Table 2.1). The ESTs associated with these significant spots were then aligned against the longest available sequences and duplications were removed to determine the number of putatively unique genes exhibiting differential regulation. After removal of duplications, 44 and 436 genes were determined to be differentially expressed in the early and late time points respectively, with 15 genes overlapping between the two for a total of 465 differentially expressed genes (Suppl. Table 2.2). Interestingly, included among the 15 overlapping genes are two annotated as potential transcription factors (Table 2.1). Additionally, multiple modes of differential regulation were exhibited, with up-regulation (increased transcript accumulation) being the most prominent (Fig. 2.2).

Validation of Microarray Results

Initially, eight genes exhibiting high levels of differential expression (≥ 2 fold change) at the late time point were validated via qRT-PCR using three biological replications. All eight genes assayed exhibited significant up-regulation following induction with either calcium nitrate or potassium nitrate (Suppl. Table 2.3), demonstrating that a high percentage of the genes that exhibit high fold changes in the microarray experiment can be validated Further validation was performed using SEQUENOM's MassARRAY platform (www.sequenome.com). A total of 11 genes selected from the two time points and exhibiting a range of fold changes were analyzed (Suppl. Table 2.4). The expression levels of a high percentage of genes assayed via this method concurred with the results of the microarray. Viewed together, these results suggest that a high percentage of differentially regulated spots from the microarray can be validated.

Non-symbiotic hemoglobin

At both time points, the gene with the highest fold change was a non-symbiotic hemoglobin. Although the function of this gene is not well understood, it has been found to be up-regulated following nitrate exposure in other studies [14, 16, 20]. Class-1 non-symbitoic hemoglobin expression has been linked to the expression of NR and has been hypothesized to aid in the detoxification of nitric oxide and/or nitrite [20]. Alternatively,

because dissolved molecular oxygen impairs the enzymatic efficiency of NR, non-symbiotic hemoglobin may serve to bind O_2 in the vicinity of NR, thus increasing the efficiency of this enzyme [21]. It should be noted, however, that nitrate induces non-symbiotic hemoglobin even under hypoxic conditions [21], making this hypothesized function less convincing.

Early Response Genes

After 30 minutes of exposure to nitrate, forty-four genes exhibited significant differential expression compared to the control using a p-value cutoff of .05 and a q-value <.27, meaning approximately a quarter of these genes would be expected to be false positives. Many of the most highly differentially regulated transcripts appear to be involved in nitrite (as opposed to nitrate) reduction (Table 2.2), similar to the findings of Wang et al. (2000); transcripts encoding NiR, ferredoxin, ferredoxin NADP reductase, and UPM1 were all up-regulated nearly or more than 2-fold. High fold changes were also observed for two genes potentially involved in DNA replication or repair: minichromosome maintenance factor 5 (+4.99 fold) and Noc3p (+2.97 fold). Two genes (DV490164, DV490781) involved in phenylpropanoid synthesis/metabolism were up-regulated, as were two genes (BG842802, DV492115) involved in amino acid biosynthesis. A number of genes involved in signal transduction exhibited up-regulation including a protein phosphatase 2C-like gene (+3.69 fold), while one Zn-RING finger protein was down-regulated. More interestingly, two potential transcription factors, one containing a MYB-domain and the other containing a basic leucine zipper (bZIP)-like domain were up-regulated within 30 minute of inductions. These genes do not share significant sequence similarities to the early response MYB and bZIP genes identified following nitrate induction of Arabidopsis and tomato [15, 16]. The four early response genes that were down-regulated act in diverse physiological pathways.

Late Response Genes

After twenty-four hours exposure to nitrate, 436 genes exhibited significantly different expression compared to the control ($p<0.001$; q,0.03). As was true for the early response genes, many of the late response genes that exhibited the highest fold changes correspond to genes involved in nitrite reduction (Table 2.3). In addition, at this time point, genes active in nitrate reduction and ammonia bioassimilation also exhibited significant

increases, incluidng: nitrate reductase (NR), cytoplasmic and plastidic forms of both glutamine synthetase (GS1 and GS2, respectively) and glutamine:2-oxoglutarate aminotransferase (NADH-GOGAT and Ferredoxin-Dependent GOGAT, respectively), as well as two aspartate aminotransferases. Induction was also observed for one isoform of glutamate dehydrogenase (GDH1); in contrast Scheible et al. (2004) reported that GDH3 was repressed by nitrate. It is also interesting to note that while isoforms of both GS1 and GS2 were up-regulated, one isoform of GS1 was repressed following nitrate exposure. These findings suggest these isozymes have distinct roles in nitrogen metabolism.

Pentose phosphate pathway genes

The bio-assimilation of nitrate requires the production of reducing equivalents and carbon skeletons, largely in the form of organic acids. In non-green tissues reducing equivalents are mainly produced by the pentose phosphate pathway. Genes encoding the enzymes of both the oxidative and non-oxidative arms of this pathway are known to respond to nitrate [22]. Here, we demonstrate that an additional gene of the oxidative arm (6-phosphogluconolactonase 1) is up-regulated (1.52 fold) following exposure to nitrate. Previous studies reported that few genes involved in the early steps of glycolysis were induced by nitrate. Further, these genes were often involved in both glycolysis and the pentose phosphate pathway, leading to the hypothesis that these genes were induced as a means to recycle carbon back into or metabolize byproducts of the pentose phosphate pathway [15]. Our findings indicate this may not be the case. From hexokinase to pyruvate kinase, genes controlling nearly every step of glycolysis were induced by nitrate in the current study (Fig. 2.3), including phosphofructokinase, which catalyzes the committed step of glycolysis. In addition, the few glycolytic genes which were not identified as being differentially regulated in this study were all previously shown to be induced by nitrate, including enolase and phosphoglycerate mutase [22].

TCA cycle genes and differences between C3 and C4 plants

Multiple genes involved in the TCA cycle were induced following nitrate exposure, as had been observed previously [13]. However, in contrast to earlier studies, in the current

study a putative malate dehydrogenase gene was down-regulated following nitrate induction. This is unexpected, because this enzyme not only catalyzes a reaction that produces the reducing equivalent NADH but is also crucial to the cyclical nature of the TCA cycle. This finding may be due to the fact that all previous reports of expression changes in response to nitrate have involved C3 plants. C3 and C4 plants differ in their metabolism of malate and oxaloacetate. Hence, in C4 plants malate dehydrogenase may play an important role during N starvation in coordinating C and N metabolism.

Other pathways and genes

Many of the strongly up-regulated genes have little or no annotation (Table 2.4). Others are involved in diverse pathways and functional annotations, including genes annotated as being hormone responsive (e.g. DV550670: Auxin-repressed protein-like, transcription factors, and having roles in RNA processing). The bulk of differentially expressed DNA replication/repair genes were up-regulated, but notable exceptions were observed. Interestingly, four transposons were differentially expressed (two up- and two down-regulated). In agreement with the findings of Wang et al. (2001), a number of genes encoding ribosomal proteins were induced, as were genes encoding translation initiation and elongation factors. As suggested by Wang et al. (2001), this may reflect a general increase in translation. Not noted in previous studies, a number of genes encoding histones were significantly up-regulated. Seven probes with sequence similarity to histones were up-regulated, including H1, H2B, and H4.

Consistent with the observation that the application of nitrate inhibits lateral root elongation [23], two cellulose synthases exhibited down-regulation, as did genes for actin and a tubulin alpha chain. In contrast, cell wall modifying genes such as pectinesterases, expansins, and glucanases exhibited both up and down regulation.

Genes involved in amino acid, protein, and lipid metabolism as well as signal transduction exhibited both up and down-regulation. A number of genes involved in aromatic amino acid metabolism appear to be down-regulated, though two genes potentially involved in tryptophan synthesis were up-regulated. A cysteine protease was down-regulated and two cysteine protease inhibitors were up-regulated.

Consistent with the findings of Scheible et al. (2004), most of the differentially expressed phenylpropanoid metabolism genes were down-regulated, many strongly. One of the most strongly down-regulated of these genes encodes chalcone synthase. Because this enzyme acts at a fork in the pathways leading to synthesis of anthocyanins or monolignols and N-starved plants accumulate anthocyanins [24], this gene may represent an important control point relative to N metabolism

Conclusions

The current work adds to our expanding knowledge of the transcriptomic response to nitrate. General trends in nitrate response observed across species include: 1) differential regulation of glycolysis, the pentose phosphate pathway and the TCA cycle, which produce reducing equivalents and carbon skeletons for nitrate assimilation, 2) substantial up-regulation of non-symbiotic hemoglobin, 3) Repression of genes involved in phenylpropanoid synthesis. The current study also suggests that GS and GDH isoforms may play specific roles during N assimilation and utilization, phenylpropanoids may be reduced, and malate dehydrogenase may play a fundamentally different role in N metabolism in C3 and C4 plants.

Materials and Methods

Plant Material and Growth Conditions

The maize B73 inbred was used for all experiments. Kernels were disinfected by soaking for 30 minutes in 10% commercial bleach solution (containing 5% sodium hypochlorite), then imbibed for 24 hours in aerated distilled water. The imbibed kernels were wrapped in germination paper [25], soaked in distilled water and placed in a beaker containing distilled water for germination in the dark at 28°C. After 72 hours, the solution was changed to 5mM calcium sulfate with the beakers transferred to a growth chamber (16 hours light at 25°C and 8 hours dark at 20°C).

Nitrate Exposure

Two time points of nitrate exposure were considered in these experiments: 30 minutes

(early) and 24 hours (late). Six biological replications were included for each time point and each replication consisted of a control sample and a nitrate-treated sample. After 10 days of growth on germination paper in 5mM calcium sulfate, six biological replications were chosen at random and placed in either fresh 5mM calcium sulfate or 5mM calcium nitrate [22], these replications constituted the late treatment groups. On day 11 following imbibition the remaining six biological replications were randomly assigned to either 5mM calcium sulfate or 5mM calcium nitrate for 30 minutes, constituting the early treatment groups. The root tissue was collected and frozen with liquid N_2, and stored at -80°C. The efficacy of the nitrate treatment was monitored according to a simple nitrate reductase assay protocol [26]. Total RNA was extracted by homogenizing roughly 3g of root tissue in 30ml Trizol Reagent (Invitrogen, Catalog No. 15596-026) and RNA extracted subsequently following the manufacturer's protocol. mRNA was isolated using Qiagen Oligotex kits (Catalog No. 70042), following the manufacturer's protocol. Total RNA was reverse transcribed using SuperscriptII (Invitrogen, Catalog No. 18064-022) with poly d(T) primers.

Microarray

Hybridizations were performed using Generation II version C microarray chips (GEO No. GPL1984) generated at the Center for Plant Genomics at Iowa State University (www.plantgenomics.iastate.edu/maizechip), This chip was designed specifically for this project by the inclusion of 10 genes that were known to exhibit a nitrogen-response. Portions of these genes were PCR amplified from B73 cDNA and spotted on the array. Hybridizations were conducted using a the protocol of [27].

Microarray Data Analysis

Six biological replicates were analyzed for each time point. Each array was scanned three times at 10um resolution for both the Cy3 and Cy5 channels using increasing intensity [28] using a ScanArray 5000 (Packard BioScience, now PerkinElmer,http://www.perkinelmer.com/). Prior to statistical analysis non-informative spots were removed from the data set. Signal intensities were normalized and mean centered [29]. On a spot-by-spot basis significance was determined by the criteria Early = $p \leq .05$, $q \leq .27$ and Late = $p \leq .001$, $q \leq ..023$, using the smallest value from the three scans as the determinant

of significance.

Quantitative Real-Time PCR

Total RNA was reversed transcribed with SuperScriptII using gene-specific primers. PCR primers were tested for specificity using genomic DNA as template prior to quantitative RT-PCR reactions. A human gene (Ac# AA418251) was spiked into the RNA samples in equal amounts prior to cDNA synthesis for normalization.

MassARRAY QGE Assay

MassARRAY QGE (Sequenom) was performed for selected genes that exhibited statistically significant change in gene expression in response to nitrate (viz., BM073725, DV490607, BM333948, DQ011869, DV491035, BG841282, BG841893, BM078981, DV491210, BG874123, BM072886). Primers were designed using the SEQUENOM QGE software using gene sequences downloaded from the MAGI database (http://magi.plantgenomics.iastate.edu/).

5' RACE

Full length transcripts of some of the probes identified as having high fold-changes by the microarray experiment were cloned using the GeneRacer kit (Invitrogen, catalog no. L1502-01), following the manufacturer's protocol. Gene-specific primers were designed to target the 3' ends of the genes using Primer3.

Sequence analysis

To determine a putative unique gene set, the cDNA sequences corresponding to the spots on the microarray were aligned against available EST and genomic sequences for extension. Sequences from the chip were initially aligned against ESTs using the MEC-P95-Mar06 data set (http://magi.plantgenomics.iastate.edu/), then subsequently aligned with maize assembled genomic islands (MAGIs) using the GenSeqer program [29]. A link was determined if at least 1 exon of the EST hit the MAGI with >=95% identity and >=80% EST coverage. The MAGI's were then BLASTed against sequenced BACs and links were formed

if the MAGI shared >=400bp of alignment length and exhibited >=99% identity with aligned region and had <=20bp of non-aligned regions. For those instances where no EST to MAGI relationship was observed, ESTs were aligned directly to the BACs and links were formed if the EST has all exons at >=95% identity, >=90% EST coverage with tails <= 60bp. MEC and MAGI data sets are publicly available at http://magi.plantgenomics.iastate.edu/downloadall.html. These ESTs were then aligned against the longest available sequences and duplications were removed to reveal the number of putatively unique genes exhibiting differential regulation.

Following microarray analysis, interesting patterns in genetic and biochemical responses were explored by assigning putative functions to the genes sequences with significant expression difference. The probes were initially aligned against sequences maintained by the Center for Plant Genomics (magi.plantgenomics.iastate.edu) to identify the longest available sequences. The identification of the longest sequence was then used to trim the number of differentially regulated probes down to the likely unique genes by removal of duplicated identities. The longest sequences were then subjected to BLASTx and BLASTn alignments against available public databases to assign putative functions to the differentially regulated elements (Supplemental Table 2.5). The putative functions were subsequently explored to discern interesting patterns in genetic and biochemical responses to nitrate exposure.

Acknowledgments

We would like to thank Mr. Eric Weatherford for technical support with qRT-PCR and QGE validation experiments, Dr. A. Assibi Mahama and Ms. Chunfang Wang for technical support with the 5' RACE experiments, Mr. Pengcheng Lu and Ms. Lisa Borsuk for performing statistical analyses under the guidance of Dr. Dan Nettleton, Cheng-Ting Yeh and Karthik Viswanathan for bioinformatics support, and Ruth Swanson-Wagner for constructive discussions. We thank the maize genome sequence project (DBI-0527192) for sharing the sequences of BACs prior to publication. This study was supported by a grant to P.S. from the Iowa Corn Promotion Board (ICPB) and funds from the State of Iowa.

References

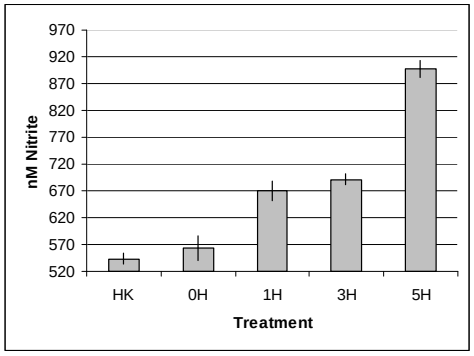
1. Smil V: **Enriching the earth : Fritz Haber, Carl Bosch, and the transformation of world food production**. Cambridge, Mass.: MIT Press; 2001.
2. Cassman KG, Dobermann A, Walters DT: **Agroecosystems, nitrogen-use efficiency, and nitrogen management**. *Ambio* 2002, **31**(2):132-140.
3. Olk DC, Cassman KG, Simbahan GC, Sta. Cruz PC, Abdulrachman S, Nagarajan R, Tan PS, Satawathananont S: **Interpreting fertilizer-use efficiency in relation to soil nutrient-supplying capacity, factor productivity, and agronomic efficiency** *Nutrient Cycling in Agroecosystems* 1999, **53**(1):35-41.
4. Burkart MR, Stoner JD: **Nitrate in aquifers beneath agricultural systems**. *Water Sci Technol* 2002, **45**(9):19-28.
5. Rabalais NN, Turner RE, Wiseman WJ, Jr.: **Hypoxia in the Gulf of Mexico**. *J Environ Qual* 2001, **30**(2):320-329.
6. Hirel B, Le Gouis J, Ney B, Gallais A: **The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches**. *J Exp Bot* 2007, **58**(9):2369-2387.
7. **U.S. Fertilizer Use and Price** [www.ers.usda.gov/Data/Fertilizer/Use/]
8. Crawford N, Glass A: **Molecular and physiological aspects of nitrate uptake in plants**. *Trends in Plant Science* 1998, **3**(10):389-395.
9. Crawford NM: **Nitrate: nutrient and signal for plant growth**. *Plant Cell* 1995, **7**(7):859-868.
10. Lam HM, Coschigano KT, Oliveira IC, Melo-Oliveira R, Coruzzi GM: **The Molecular-Genetics of Nitrogen Assimilation into Amino Acids in Higher Plants**. *Annu Rev Plant Physiol Plant Mol Biol* 1996, **47**:569-593.
11. Stitt M: **Nitrate regulation of metabolism and growth**. *Curr Opin Plant Biol* 1999, **2**(3):178-186.
12. Lian X, Wang S, Zhang J, Feng Q, Zhang L, Fan D, Li X, Yuan D, Han B, Zhang Q: **Expression profiles of 10,422 genes at early stage of low nitrogen stress in rice assayed using a cDNA microarray**. *Plant Mol Biol* 2006, **60**(5):617-631.
13. Scheible WR, Morcuende R, Czechowski T, Fritz C, Osuna D, Palacios-Rojas N, Schindelasch D, Thimm O, Udvardi MK, Stitt M: **Genome-wide reprogramming of primary and secondary metabolism, protein synthesis, cellular growth processes, and the regulatory infrastructure of Arabidopsis in response to nitrogen**. *Plant Physiol* 2004, **136**(1):2483-2499.
14. Wang R, Guegler K, LaBrie ST, Crawford NM: **Genomic analysis of a nutrient response in Arabidopsis reveals diverse expression patterns and novel metabolic and potential regulatory genes induced by nitrate**. *Plant Cell* 2000, **12**(8):1491-1509.
15. Wang R, Okamoto M, Xing X, Crawford NM: **Microarray analysis of the nitrate response in Arabidopsis roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism**. *Plant Physiol* 2003, **132**(2):556-567.
16. Wang YH, Garvin DF, Kochian LV: **Nitrate-induced genes in tomato roots. Array analysis reveals novel genes that may play a role in nitrogen nutrition**. *Plant Physiol* 2001, **127**(1):345-359.
17. Scheible WR, Gonzalez-Fontes A, Lauerer M, Muller-Rober B, Caboche M, Stitt M: **Nitrate Acts as a Signal to Induce Organic Acid Metabolism and Repress Starch Metabolism in Tobacco**. *Plant Cell* 1997, **9**(5):783-798.
18. Morcuende R, Krapp A, Hurry V, Stitt M: **Sucrose-feeding leads to increased rates of nitrate assimilation, increased rates of α-oxoglutarate synthesis, and increased synthesis of a wide spectrum of amino acids in tobacco leaves** *Planta* 1998, **206**(3):394-409.
19. Delgado E, Mitchell RAC, Parry MA, Driscoll SP, Mitchell VJ, Lawlor DW: **Interacting effects of CO2 concentration, temperature and nitrogen supply on the photosynthesis and**

composition of winter wheat leaves. *Plant, Cell & Environment* 1994, **17**(11):1205-1213.

20. Ohwaki Y, Kawagishi-Kobayashi M, Wakasa K, Fujihara S, Yoneyama T: **Induction of class-1 non-symbiotic hemoglobin genes by nitrate, nitrite and nitric oxide in cultured rice cells**. *Plant Cell Physiol* 2005, **46**(2):324-331.
21. Kenis JD, Campbell WH: **Oxygen Inhibition of Nitrate Reductase Biosynthesis in Detached Corn Leaves via Inhibition of Total Soluble Protein Synthesis**. *Plant Physiol* 1989, **91**(3):883-888.
22. Wang R, Tischner R, Gutierrez RA, Hoffman M, Xing X, Chen M, Coruzzi G, Crawford NM: **Genomic analysis of the nitrate response using a nitrate reductase-null mutant of Arabidopsis**. *Plant Physiol* 2004, **136**(1):2512-2522.
23. Zhang H, Jennings A, Barlow PW, Forde BG: **Dual pathways for regulation of root branching by nitrate**. *Proc Natl Acad Sci U S A* 1999, **96**(11):6529-6534.
24. Martin T, Oswald O, Graham IA: **Arabidopsis seedling growth, storage lipid mobilization, and photosynthetic gene expression are regulated by carbon:nitrogen availability**. *Plant Physiol* 2002, **128**(2):472-481.
25. Wen TJ, Schnable PS: **Analyses of Mutants of 3 Genes That Influence Root Hair Development in Zea-Mays (Gramineae) Suggest That Root Hairs Are Dispensable**. *American Journal of Botany* 1994, **81**(7):833-842.
26. Harley SM: **Use of a Simple, Colorimetric Assay to Demonstrate Conditions for Induction of Nitrate Reductase in Plants**. *The American Biology Teacher* 1993, **55**(3):161-164.
27. Nakazono M, Qiu F, Borsuk LA, Schnable PS: **Laser-capture microdissection, a tool for the global analysis of gene expression in specific plant cell types: identification of genes expressed differentially in epidermal cells or vascular tissues of maize**. *Plant Cell* 2003, **15**(3):583-596.
28. Skibbe DS, Wang X, Zhao X, Borsuk LA, Nettleton D, Schnable PS: **Scanning microarrays at multiple intensities enhances discovery of differentially expressed genes**. *Bioinformatics* 2006, **22**(15):1863-1870.
29. Swanson-Wagner RA, Jia Y, DeCook R, Borsuk LA, Nettleton D, Schnable PS: **All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents**. *Proc Natl Acad Sci U S A* 2006, **103**(18):6805-6810.

Figures

A



B

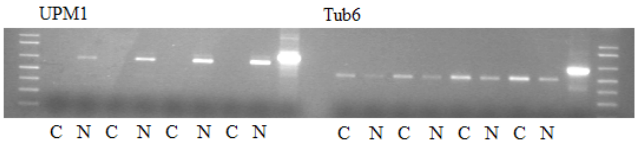
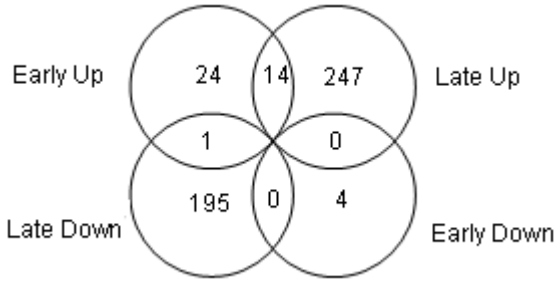


Figure 2.1: (A) Results of NR assay presented as nano-moles of nitrate. HK = heat killed; 0H, 1H, 3H, 5H = duration of nitrate exposure in hours. (B) RT-PCR of UPM1 following 30min exposure to nitrate. Tub6 = tubulin.

A



B

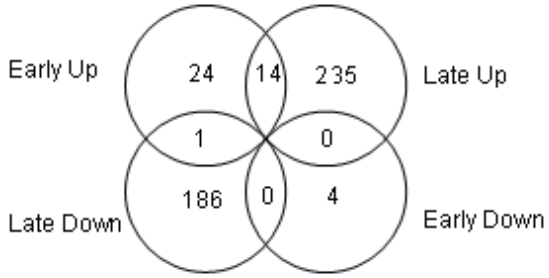


Figure 2.2: (A) Total number of differentially regulated probes and the direction of observed fold change following exposure to nitrate. (B) Number of putatively unique genes and direction of observed fold change following exposure to nitrate.

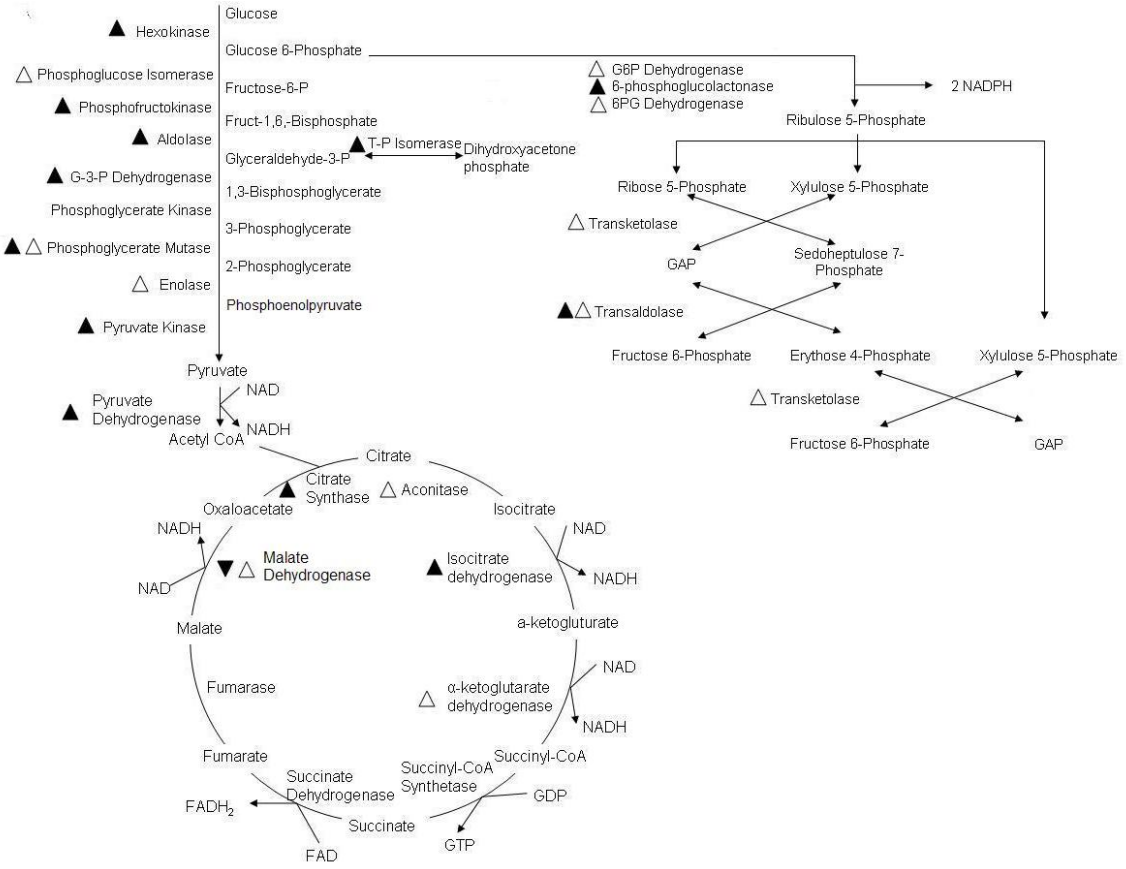


Figure 2.3: Model of biochemical pathways depicting observed directional changes in transcript abundance and putative function of various microarray spots following nitrate exposure. ▲= increase in transcript abundance observed in the current study; ▼= decrease in transcript abundance observed in the current study; Δ= increase in transcript abundance observed in previous published studies.

Tables

Table 2.1: Genes exhibiting significant differential expression at both time points and their putative functions. FC = fold change

	FC		
GeneID	**30 Min**	**24 Hrs**	**Putative Function**
BM333948	11.7	84	Non-symbiotic hemoglobin
DQ011869	8.5	14.4	Uroporphyrinogen III methyltransferase
AI622662	7.6	8	Uroporphyrinogen III methyltransferase
BM334491	5.0	12.7	Ferredoxin-6
DV492546	5.0	4.4	Minichromosome maintenance deficient protein 5
M23456	4.8	6.5	Nitrite reductase
DV491035	4.6	14.4	Homeodomain leucine zipper-like
BM336095	3.2	3.6	Ferredoxin--NADP reductase
BG841986	2.5	2.5	Myb, DNA-binding domain containing protein
BM074654	2.0	0.7	Indole-3-acetic acid-regulated protein
BG840791	1.8	1.5	Auxin-regulated protein-like
DV551235	1.7	1.6	Unnamed protein product
DV492115	1.7	1.5	Anthranilate phosphoribosyltransferase
DV491056	1.5	4.5	Aspartate aminotransferase
DV490747	1.5	3.9	Aspartate aminotransferase

Table 2.2: Genes showing at least a two-fold increase in transcript abundance after 30 minutes exposure to nitrate. FC = fold change; * = not previously reported

GeneID	P-value	FC	Putative Function
BM333948	0.0001	11.7	Non-symbiotic hemoglobin
DQ011869	0.0000	8.5	Uroporphyrinogen III methyltransferase
AI622662	0.0005	7.6	Uroporphyrinogen III methyltransferase
BM334491	0.0000	5.0	Ferredoxin-6
DV492546*	0.0003	5.0	Minichromosome maintenance deficient protein 5
M23456	0.0001	4.8	Nitrite reductase
DV491035*	0.0001	4.6	Homeodomain leucine zipper-like
DV492356*	0.0006	3.7	Protein phosphatase type 2-C
BM336095	0.0001	3.2	Ferredoxin--NADP reductase
BM350754*	0.0005	3.1	Hypothetical
DV494689*	0.0001	3.0	Noc3p
DV491802*	0.0002	2.8	Acid phosphatase-like
DV490425*	0.0002	2.7	Serine carboxypeptidase II
DV490164*	0.0005	2.5	4-coumarate coenzyme A ligase
BG841986*	0.0001	2.5	Myb, DNA-binding domain containing protein
DV942423*	0.0004	2.3	Hypothetical

Table 2.3: Genes showing at least a five-fold increase in transcript abundance after 24 hours exposure to nitrate. FC = fold change; * = not previously reported

GeneID	P-value	FC	Putative Function
BM333948	0.0003	84.0	Non-symbiotic hemoglobin
BG874013*	0.0001	18.1	Unknown protein
DV491035*	0.0001	14.4	Homeodomain leucine zipper-like
DQ011869	0.0000	14.4	Uroporphyrinogen III methyltransferase
DV489508*	0.0002	14.1	B12D family protein
BM334491	0.0000	12.7	Ferredoxin-6
BM348849	0.0001	10.9	Early nodulin
AI622662	0.0002	8.0	Uroporphyrinogen III methyltransferase
M23456	0.0000	6.5	Nitrite reductase
BM350368*	0.0001	6.4	Unknown protein
DV494692*	0.0001	6.3	Cp-thionin
AW066985*	0.0002	6.1	Cis,cis-muconate cycloisomerase -like
BM338819*	0.0009	5.7	AP2 domain containing protein RAP2.1
BM073657*	0.0009	5.6	Serine/threonine kinase
BM074220*	0.0001	5.1	Alanine aminotransferase 2

Table 2.4: Partial list of genes with poorly defined functions exhibiting high fold change increases. FC = fold change

Gene ID	P-Value	FC	Time Point	Putative Function
DV491035	0.0001	4.6	30 Min	Leucine zipper-like
DV492356	0.0006	3.7	30 Min	Phosphatase type 2-C
DV491802	0.0002	2.8	30 Min	Acid phosphatase-like
BG841986	0.0001	2.5	30 Min	Myb domain
BG874013	0.0001	18.1	24 Hrs	Hypothetical Protein
DV491035	0.0001	14.4	24 Hrs	Leucine zipper-like
DV489508	0.0002	14.1	24 Hrs	B12D family protein
BM348849	0.0001	10.9	24 Hrs	Early nodulin
BM350368	0.0001	6.4	24 Hrs	Hypothetical Protein
DV494692	0.0001	6.4	24 Hrs	Thionin
BM338819	0.0009	5.7	24 Hrs	AP2 domain
BM073657	0.0009	5.6	24 Hrs	Serine/threonine kinase

Chapter 3: Global Transcript Analysis of Node Development in Rice

Michael E Miller[a,b], An-Ping Hsia[c,d], Hui-Hsien Chou[b,d], Dan Nettleton[e], Yan Fu[c,f], Patrick S. Schnable[a, b, d, f , 1]

[a] Molecular, Cellular and Developmental Biology Graduate Program, Iowa State University, Ames, IA 50011
[b] Department of Genetics, Development and Cell Biology, Iowa State University, Ames, IA 50011
[c] Department of Agronomy, Iowa State University, Ames, Iowa 50011
[d] Center for Plant Genomics, Iowa State University, Ames, Iowa 50011-3650
[e] Department of Statistics, Iowa State University, Ames, Iowa 50011
[f] Center for Carbon Capturing Crops, Iowa State University, Ames, Iowa 50011-3650

1. Corresponding author: Roy J. Carver Co-Laboratory Iowa State University Ames, IA 50011-3650. E-mail: schnable@iastate.edu

Abstract

Plant development and architecture, which are controlled by a complex web of underlying genetic regulation, are important determinants of crop yield. To elucidate the molecular genetics of nodal development in rice, global profiling was used to compare transcription levels between all combinations of four discrete anatomical positions (nodes 1, 2, 3, and 4) and four time points (46, 53, 60, and 67 days post-planting). A total of 1,945 genes were found to be significantly differentially regulated in at least one of the 38 possible comparisons ($q<0.05$). Further, these genes were found to cluster into 10 groups of co-regulated expression. Exploration of these 10 clusters as well as consistency of differential expression between comparisons indicate that transcription is relatively stable over time for a given node, but varies to a much wider extent among nodes on a given day. The difference in expression between spatially divided nodes is especially pronounced when comparing the basal node to higher nodes. In addition, the current study has identified five putative transcription factors (Os01g13740, Os01g68900, Os02g39140, Os03g54160, and Os11g47460) that may play important roles in regulating differential expression between node 1 and higher nodes.

Introduction

The major cereal crops rice, maize, and wheat provide the bulk of the world's nutrition, with rice alone feeding over half of the world's population [1]. As that population increases, so will the demand for food, making it imperative to identify the underlying genetic mechanisms controlling yield. Plant architecture, the three-dimensional organization of the plant's aerial tissue, is one of the most important factors affecting crop yield [2, 3]. This architecture is defined by the plant's degree of branching, intermodal elongation, shoot determinacy [1, 4], and branching angles [5]. While external

environmental factors such as light, temperature, humidity, nutrition, and plant density influence these characteristics, the overall architecture of the plant is determined by genetics. Tiller buds are formed at the axil of each leaf on its mother stem of the rice plant, but only those formed at the unelongated basal internodes have the capacity to develop into tillers; a process controlled by the axial meristems [1]. Thus, the genetic interactions of the basal-most portions of the plant would be expected to play important roles in determining plant architecture by controlling tiller and/or leaf outgrowth.

Recent progress has been made in uncovering some of the molecular mechanisms controlling rice development [3, 5]. Mutants have been identified that affect various developmental processes, including axillary initiation [6], axillary outgrowth [7], branching [8], and leaf initiation [9-11]. Here, global transcript profiling was conducted to explore the complex genetic interactions regulating the development of rice. An rice oligo-array consisting of 23,040 probes (www.ricearray.org) was utilized to compare transcript abundance between four time points (46, 53, 60, and 67 days post-planting) and four anatomical positions (nodes 1, 2, 3, and 4) to give an indication of the spatiotemporal changes in transcription that occur during rice development.

Results

Global Transcription Profiles in Nodes of Developing Rice

To assess transcriptomic variations between nodes of developing rice plants, a series of microarray studies were performed to profile and compare genome-wide transcript abundance across developmental space and time. In all, thirteen samples were available when considering the development of four nodes across four time points (Table 3.1). These thirteen treatments allowed thirty-eight comparisons (a combination of pair-wise and serial statistical analyses, according to the experimental design; Suppl Fig. 3.1). A total of 1,945 probes (out of 23.040) exhibited significant differential expression in at least one of these thirty-eight comparisons (Suppl Table 3.1) using a criteria of qvalue $\leq .05$.

Differential Accumulation of Transcripts During Development

The numbers of probes exhibiting significantly different transcript accumulation for all thirty-eight comparisons are given in Table 3.2. There is a striking difference between the number of differentially regulated probes between temporal comparisons and spatial comparisons. In fact, the largest number of differentially regulated probes in a comparison between time points for a given node is 68, while several comparisons between nodes at a given day exhibit hundreds of differentially regulated probes.

Variation at Node 1

A closer look at Table 2 reveals that all of the comparisons resulting in large numbers of differentially regulated probes involve node 1. In fact, when comparing higher nodes to each other there is little evidence for differential transcription. Nodes 2 and 3 do not exhibit differential transcription at any time point. When comparing nodes 3 and 4 there was no significant difference in transcript abundance at day 60 and only four probes were significant at day 67. Even when comparing the spatially separated nodes 2 and 4 few probes showed significant differences (13 at day 60, 45 at day 67).

Coordinated Transcription Across Treatments

To identify those genes exhibiting coordinated regulation, standardized transcript levels for the 1,945 differentially regulated spots were clustered across all thirteen treatments. Ten clusters of co-regulated transcripts were thus identified (Suppl Table 3.2). The accumulation of transcripts for these genes was plotted across both spatial (Fig. 3.1) and temporal (Fig. 3.2) development. When the clusters are organized spatially, large discrepancies can be observed between expression levels at node 1 compared to higher nodes, whereas there is commonly a more gradual change in expression patterns between the higher nodes.

Consistency of Differential Regulation

Venn diagrams were generated for overlaying the lists of differentially regulated genes which identified 37 genes that were always significant across different nodes at the same developmental stage (Suppl Fig. 3.3). However, a closer inspection of the pair-wise comparisons for these genes indicated that most of them were significantly differentially expressed probably because the serial comparisons included node 1. To further explore this, all pair-wise and serial comparisons involving node 1 at day 60 were diagrammed (Suppl Fig. 3.4) to identify 32 genes which were consistently significant at this time point. Of these 32 genes, five had annotations indicating potential roles in transcriptional regulation (Table 3.3), indicating these genes may play pivotal roles in differentiating transcription between node 1 and higher nodes.

When these results are expanded to other days the observations are less consistent. None of the transcription factors identified above appear in the list of significant genes when comparing nodes 1 and 2 at days 46, 53, and 67; though the low numbers of significant genes observed at days 53 and 67 may be masking these transcription factors. However, most of the transcription factors do appear to be significant when comparing node 1 to nodes 3 and 4 at days 60 and 67, further indicating they may play integral roles in differentiating gene activity between these nodes.

Discussion

As can be observed in Table 2, a larger number of differentially regulated probes were evident in spatial rather than temporal comparisons. That is, when comparing transcript levels between nodes at a given day there is often more significant differentially regulated probes than when comparing the same node across different time points. This may indicate that transcription at a given node is relatively stable over time, whereas spatially divided nodes show more unique transcription patterns. Though it is interesting that the temporal comparisons pitting the first and last dates a node is available against each other often result in an increase in significant probes. This may indicate that transcription level in the nodes are gradually changing across time but the change is not large enough to be considered significant until the time points considered are far enough apart.

This difference in spatial transcription is most pronounced when comparing the basal-most node (node 1) to higher nodes, while higher nodes exhibited little or no significant differences in transcription. This may indicate pronounced unique morphological and/or physiological aspects of the basal node. Previous studies on the morphological and anatomical architecture of rice have shown that leaf length and width vary depending on position on the main culm, with leaf length increasing at higher positions on the culm, and that tiller initiation is usually restricted to three nodes below the top emerging leaf [12]. Thus, it would be expected to observe large differences between nodes, especially when comparing basal nodes which can develop not only leaves but tillers and/or roots [12].

All pair-wise comparisons resulting in high numbers of differentially transcribed probes involved node 1, indicating transcriptional regulation at this node is vastly different from higher nodes. This difference in transcript abundance is readily evident when comparing expression patterns in clusters of co-regulated genes. In the 10 clusters, expression levels at node 1 appear to show little variation. However, in most cases there is a large discrepancy between expression levels at node 1 and node 2. Expression levels between node 2 and higher nodes appear to be relatively stable, though some variations can be observed. For instance, Cluster 6 and Cluster 10 show a pronounced peak in gene expression at node 2 and node 3, respectively, compared to other nodes. A jagged peak in expression for node 2 at day 46 was found in cluster 9 (Fig. 3.1). There is no evidence for enrichment of a particular class of genes in these clusters, making it difficult to postulate why these genes exhibit such patterns. In fact, there is no evidence for enrichment of genetic classes in any of the given clusters. It could be interesting to see how a mutation in any gene in a cluster affects expression of the other genes. This may be especially true of those genes annotated as transcription factors.

Certain transcription factors have previously been shown to affect development of plant architecture. MOC1 encodes a gene showing sequence similarity to a class of putative transcription factors known as GRAS [13], and mutations in this gene result in impaired axillary initiation [6]. In addition, OsTB1 encodes a putative transcription factor of the TCP protein family, and over-

expression of this gene has been shown to limit tiller outgrowth [7]. Here, the consistency of differential expression patterns of at least five genes annotated as transcription factors indicate these genes may play important roles in the transcriptional segregation between node 1 and higher nodes. At 60 days post-planting these five genes are among 32 the show consistent differential regulation. When expanding these observations to other days, they become less consistent largely due to the fluctuations seen in significant transcripts between node 1 and 2. However, when comparing node 1 to nodes 3 or 4 these transcription factors are regularly observed on the list of differentially regulated probes. Further, these five transcription factors are found in four different co-regulation clusters, indicating they may play pivotal roles in coordinating regulation for these particular sets of genes. Lending credence to this hypothesis is the fact that one of these transcription factors (TR030972) maps to the rice locus Os01g13740 which encodes a protein showing sequence similarity to the OsGLK2 gene, which has previously been shown to regulate cell-type differentiation processes [14]. The other four transcription factors of interest here do not show sequence similarity to mutants that have been previously described. It would therefore be of interest to explore the phenotypes related to mutations in these genes.

Conclusion

When comparing differences in transcript accumulation across the thirteen samples considered in this study, a total of 1,945 genes were found to accumulate to significantly different levels in at least one of the 38 possible comparisons. These significant differences indicate that transcription varies to a greater degree across anatomical space than over time, especially when comparing the basal-most node to higher nodes. There is also evidence for coordinated transcript accumulation among clusters of genes across spatiotemporal development. It could be interesting to explore the affects mutations to genes in these cluster have on expression patterns, particularly the five genes (Os01g13740, Os01g68900, Os02g39140, Os03g54160, and Os11g47460) annotated as putative transcription factors and showing consistent differential expression between the basal-most node and higher nodes. It should also be noted that while few genes exhibit differential expression at a node across time points, there is evidence for a gradual change in expression levels.

Materials & Methods

Growth Conditions

Seeds of *Oryza sativa* L cv. Nipponbare were incubated in water at 42C for 24 hours prior to planting. Initial watering contained a mixture of 1 to 1 Peters Excel 21-5-20 and Peters Excel 15-5-15 water soluble fertilizers (Grace Sierra, Milpitas, CA). Subsequent watering contained no fertilizer mix

and was applied to maintain water levels high enough to submerse the seed. Each planting date consisted of 64 plants planted as 16 blocks of 4 plants to each block. Each planting date was considered one replication, from which all time points were harvested and compared. Each time point consisted of three blocks chosen at random (for a total of 12 plants/time point/replicate). In all, 6 replications were considered per time point (Suppl Fig. 3.1). Plants were grown in a Percival PGC-105 growth chamber set for 12 hours light at 28C with 85% humidity, and 12 hours dark at 22C with 60% humidity.

Harvesting

Plants were harvested at four time points: 46, 53, 60, and 67 days post planting. Three blocks (12 plants) were chosen at random for harvest at each time point. Seven to nine of the harvested plants were used for RNA isolation, with the remainder being used for imaging or stored for biochemical analysis. Following removal of the three randomly chosen blocks, the blocks remaining in the growth chamber were rearranged such that the growing space was maintained (Suppl Fig. 3.2). Plants were cut at the roots, below node one. Then the leaves were removed, leaving a naked stalk. Nodes and internodes were sectioned using razor blades, and the sections were placed in foil and frozen in liquid nitrogen. As nodes appear over developmental time, not all nodes were available for every given harvest date. In all, thirteen different combinations of harvest date and node placement were compared (Table 1).

Sample Preparation

Frozen nodes were ground using mortars and pestles. RNA was isolated using RNeasy Plant Mini kit by Qiagen following the manufacturer's suggested protocol. RNA were reverse transcribed using poly-T as well as random primers and SuperScript II (Invitrogen, Catalog No. 18064-022), then labeled and pooled according to established protocols [15]. Each sample was labeled with Cy3 or Cy5 dye and hybridized to the GPL6939 and GPL6940 platforms, containing a combined total of 23,040 spots. The arrays for all replications were scanned using a ScanArray 5000 (Packard, Meriden, CT) at 10um resolution for both Cy3 and Cy5 channels. Scans were performed at three intensities [16]. Prior to statistical analysis non-informative spots were removed from the data set. Signal intensities were normalized and mean centered according to previously established protocols [17].

Determining q-values

Transcription profiles for the thirteen developmental points considered were compared to determine differential expression. Each transcription profile was compared individually and as a set across time and across developmental space. In all, 38 different comparisons were made (Suppl Fig 3.1). q-values were then determined according to the established methods [18] (Suppl Table 3.2).

Clustering of Transcript Expressions

Transcript expression values from the thirteen developmental points considered were then clustered to determine possible patterns of coordinated expression. Expression values form the thirteen points were standardized, and the full set of standardized values was clustered using K-medoids clustering [19]. A variation of the gap statistic [20] was used to estimate the appropriate number of clusters. The R program codes used to perform this clustering can be accessed at http://www.public.iastate.edu/~dnett/microarray/rfunctions.txt. The patterns of expression values were plotted using the JMP7 software package.

Venn Diagrams

Venn diagrams overlaying lists of differentially regulated transcripts between sets of comparisons were generated using a publicly available program at http://www.pangloss.com/seidel/Protocols/venn.cgi.

Acknowledgements

We would like to thank Mr. Joshua Shendelman and Ms. Marianne Smith for technical support in growing and harvesting materials, Mr. Pengchung Lu for performing statistical analyses, Ms. Hailing Jin and Ms. Mitzi Wilkening for technical support in microarray experiments, Mr. Cheng-Ting Yeh for bioinformatic support, and Ms. Ruth Swanson-Wagner for constructive discussions. This research was funded by a grant from the NSF (DBI-0313887) to Drs. P Ronald and P.S. Schnable.

References

1. Wang Y, Li J: **The plant architecture of rice (Oryza sativa)**. *Plant Mol Biol* 2005, **59**(1):75-84.
2. Komorisono M, Ueguchi-Tanaka M, Aichi I, Hasegawa Y, Ashikari M, Kitano H, Matsuoka M, Sazuka T: **Analysis of the rice mutant dwarf and gladius leaf 1. Aberrant katanin-mediated microtubule organization causes up-regulation of gibberellin biosynthetic genes independently of gibberellin signaling**. *Plant Physiol* 2005, **138**(4):1982-1993.
3. Wang Y, Li J: **Molecular basis of plant architecture**. *Annu Rev Plant Biol* 2008, **59**:253-279.
4. Ward SP, Leyser O: **Shoot branching**. *Curr Opin Plant Biol* 2004, **7**(1):73-78.
5. Yang XC, Hwa CM: **Genetic modification of plant architecture and variety improvement in rice**. *Heredity* 2008, **101**(5):396-404.
6. Li X, Qian Q, Fu Z, Wang Y, Xiong G, Zeng D, Wang X, Liu X, Teng S, Hiroshi F *et al*: **Control of tillering in rice**. *Nature* 2003, **422**(6932):618-621.
7. Takeda T, Suwa Y, Suzuki M, Kitano H, Ueguchi-Tanaka M, Ashikari M, Matsuoka M, Ueguchi C: **The OsTB1 gene negatively regulates lateral branching in rice**. *Plant J* 2003, **33**(3):513-520.
8. Mao C, Ding W, Wu Y, Yu J, He X, Shou H, Wu P: **Overexpression of a NAC-domain**

protein promotes shoot branching in rice. *New Phytol* 2007, **176**(2):288-298.

9. Kawakatsu T, Itoh J, Miyoshi K, Kurata N, Alvarez N, Veit B, Nagato Y: **PLASTOCHRON2 regulates leaf initiation and maturation in rice**. *Plant Cell* 2006, **18**(3):612-625.
10. Miyoshi K, Ahn BO, Kawakatsu T, Ito Y, Itoh J, Nagato Y, Kurata N: **PLASTOCHRON1, a timekeeper of leaf initiation in rice, encodes cytochrome P450**. *Proc Natl Acad Sci U S A* 2004, **101**(3):875-880.
11. Zhu QH, Dennis ES, Upadhyaya NM: **Compact shoot and leafy head 1, a mutation affects leaf initiation and developmental transition in rice (Oryza sativa L)**. *Plant Cell Rep* 2007, **26**(4):421-427.
12. Smith CW, Dilday RH: **Rice: origin, history, technology, and production**: John Wiley and Sons; 2002.
13. Bolle C: **The role of GRAS proteins in plant signal transduction and development**. *Planta* 2004, **218**(5):683-692.
14. Rossini L, Cribb L, Martin DJ, Langdale JA: **The maize golden2 gene defines a novel class of transcriptional regulators in plants**. *Plant Cell* 2001, **13**(5):1231-1244.
15. Nakazono M, Qiu F, Borsuk LA, Schnable PS: **Laser-capture microdissection, a tool for the global analysis of gene expression in specific plant cell types: identification of genes expressed differentially in epidermal cells or vascular tissues of maize**. *Plant Cell* 2003, **15**(3):583-596.
16. Skibbe DS, Wang X, Zhao X, Borsuk LA, Nettleton D, Schnable PS: **Scanning microarrays at multiple intensities enhances discovery of differentially expressed genes**. *Bioinformatics* 2006, **22**(15):1863-1870.
17. Swanson-Wagner RA, Jia Y, DeCook R, Borsuk LA, Nettleton D, Schnable PS: **All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents**. *Proc Natl Acad Sci U S A* 2006, **103**(18):6805-6810.
18. Storey JD, Tibshirani R: **Statistical significance for genomewide studies**. *Proc Natl Acad Sci U S A* 2003, **100**(16):9440-9445.
19. Kaufman L, Rousseeuw P: **Finding Groups on Data: an Introduction to Cluster Analysis**: John Wiley & Sons; 1990.
20. Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, Botstein D, Altman RB: **Missing value estimation methods for DNA microarrays**. *Bioinformatics* 2001, **17**(6):520-525.

Figures

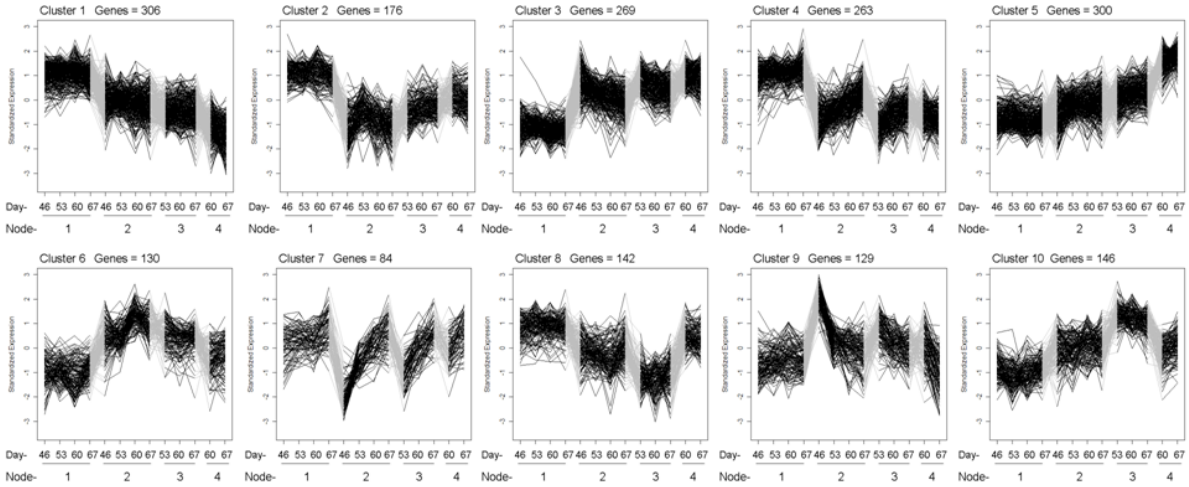


Figure 3.4: Standardized expression values organized spatially. N = node number; D = Days post-planting; Genes = Number of significant genes in that cluster.

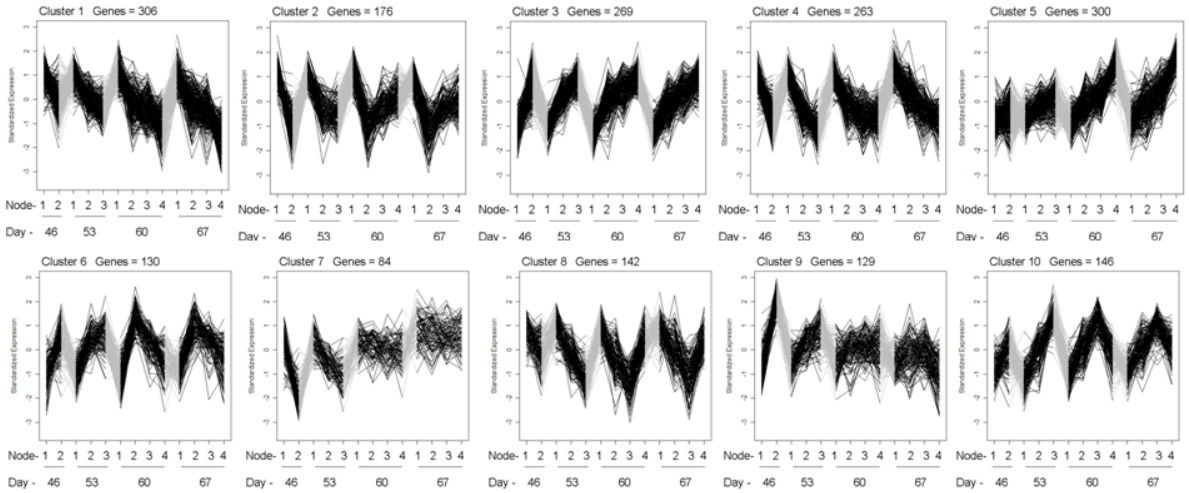


Figure 3.5: Standardized expression values organized temporally. N = node number; D = Days post-planting; Genes = Number of significant genes in that cluster.

Tables

Table 3.5: Treatments analyzed.

	46 Days	53 Days	60 Days	67 Days
Node 4	N/A	N/A	12	13
Node 3	N/A	9	10	11
Node 2	5	6	7	8
Node 1	1	2	3	4

Table 3.6: Numbers of significant genes found in each of the 38 possible comparisons. # = Number of significant genes in that comparison.

Across Time (Temporal)	**#**	**Across Nodes (Spatial)**	**#**
46Day vs 53Days in Node1	0	Node1 vs Node2 in 46Days	439
46Day vs 60Days in Node1	0	Node1 vs Node2 in 53Days	30
46Day vs 67Days in Node1	23	Node1 vs Node3 in 53Days	671
53Day vs 60Days in Node1	0	Node2 vs Node3 in 53Days	0
53Day vs 67Days in Node1	0	Node1 vs Node2 in 60Days	181
60Day vs 67Days in Node1	0	Node1 vs Node3 in 60Days	435
46Day vs 53Days in Node2	7	Node1 vs Node4 in 60Days	423
46Day vs 60Days in Node2	27	Node2 vs Node3 in 60Days	0
46Day vs 67Days in Node2	68	Node2 vs Node4 in 60Days	13
53Day vs 60Days in Node2	0	Node3 vs Node4 in 60Days	0
53Day vs 67Days in Node2	0	Node1 vs Node2 in 67Days	7
60Day vs 67Days in Node2	0	Node1 vs Node3 in 67Days	295
53Day vs 60Days in Node3	0	Node1 vs Node4 in 67Days	550
53Day vs 67Days in Node3	38	Node2 vs Node3 in 67Days	0
60Day vs 67Days in Node3	0	Node2 vs Node4 in 67Days	45
60Day vs 67Days in Node4	0	Node3 vs Node4 in 67Days	4
Across Day(46,53,60,67) in Node1	4	Across Node(1,2,3) at 53Days	370
Across Day(46,53,60,67) in Node2	20	Across Node(1,2,3,4) at 60Days	539
Across Day(53,60,67) in Node3	16	Across Node(1,2,3,4) at 67Days	460

Table 3.7: 32 probes showing consistent differential regulation in comparisons involving node 1 at day 60. LOC_Os refers to genomic location. Functions are as defined by ricearray.org. Cluster number refers to the co-regulation cluster in which each gene is found.

Probe ID	LOC_Os	Function	Cluster	GO Annotation
TR030195	LOC_Os01g03680	Bowman-Birk type bran trypsin inhibitor precursor	1	N/A
TR030972*	LOC_Os01g13740	Myb-like DNA-binding domain, SHAQKYF class	3	GO:0003677
TR031467	LOC_Os01g22336	Peroxidase 1 precursor	1	GO:0003674
TR031595	LOC_Os01g24710	Salt stress-induced protein	4	N/A
TR034674*	LOC_Os01g68900	Zinc finger, C3HC4 type family protein	1	GO:0005488
TR035996	LOC_Os02g13290	Phosphoethanolamine/phosphocholine phosphatase	5	GO:0003824
TR037629*	LOC_Os02g39140	DNA binding protein	4	GO:0003677
TR039388	LOC_Os03g03790	Acyl-activating enzyme 11	3	GO:0000166
TR039389	LOC_Os03g03810	Flower-specific gamma-thionin precursor	3	N/A
TR040668	LOC_Os03g19600	PE_PGRS family protein	1	GO:0000003
TR041667	LOC_Os03g36560	Peroxidase 27 precursor	1	GO:0003674
TR042885	LOC_Os03g52860	Lipoxygenase 2	4	GO:0003824
TR043000*	LOC_Os03g54160	MADS-box transcription factor 14	5	GO:0003677
TR044247	LOC_Os04g09390	Lectin precursor	4	N/A
TR044354	chr04:5985588	N/A	4	N/A
TR045376	LOC_Os04g31460	Expressed protein	3	N/A
TR045750	LOC_Os04g36720	Ferric reductase-like transmembrane component	3	GO:0003824
TR052687	LOC_Os08g34700	GDU1	5	N/A
TR054086	LOC_Os10g11500	Pathogenesis-related protein PRB1-3	4	GO:0005576
TR057782	LOC_Os06g30370	Homogous to Arabidopsis Mother of FT and TFL	3	GO:0008289
TR057856	LOC_Os06g32370	Leaf-specific thionin precursor	4	N/A
TR068989	LOC_Os11g32650	Chalcone synthase	3	GO:0006519
TR069030	LOC_Os11g33270	Xyloglucan endotransglucosylase/hydrolase	5	GO:0005618
TR070058	AK058861	PsbM	3	N/A
TR070066	AF143941	Cytochrome b559 beta subunit	3	N/A
TR071321	LOC_Os04g02980	Cucumisin precursor	3	GO:0005618
TR072044	LOC_Os10g05490	Cytochrome P450 76C4	5	GO:0003824
TR072121	LOC_Os10g31330	Glycine-rich cell wall structural protein 2	5	GO:0000003
TR072557	LOC_Os09g31080	Induced stolen tip protein TUB8	3	GO:0005215
TR072733	LOC_Os05g34240	Expressed protein	3	N/A
TR073006	LOC_Os12g19470	Ribulose bisphosphate carboxylase small chain C	3	GO:0003824
TR073286*	LOC_Os11g47460	MYB59	5	GO:0000003

Chapter 4: General Conclusions

Summary of Experimental Findings

Both studies described here add to our understanding of fundamental biology in various ways. The exposure of maize roots to nitrate aids the steady advance in our understanding of nitrogen utilization in plants, while the spatiotemporal analysis of rice node development using microarrays has not previously been reported. Each experiment presented its own unique set of challenges and achievements.

The findings presented here advance our understanding of nitrate response by indicating differential regulation for novel genes of the pentose phosphate, glycolytic, and Calvin cycle pathways; pathways previously known to respond to nitrate. These findings also suggest a fundamentally different role for malate dehydrogenase in regulating nitrogen metabolism between C3 and C4 plants. Additionally, novel genes in various pathways with less obvious connections to nitrogen metabolism were revealed. Hopefully, this advancement in our knowledge of nitrogen response and metabolism may one day be used to engineer plants with higher nitrogen utilization efficiencies, thus reducing the environmental impact of annual cropping systems.

Rice development was explored in a novel manner using microarrays to assess changes in transcript abundance across time and developmental space. By doing so, it was possible to identify sets of genes showing coordinated control, as well as trends in spatiotemporal changes in transcript abundance. Chief among these observations is the fact that transcript accumulation appears to be much more variable across anatomical positions than across time at a given position, especially when comparing the basal-most node to higher nodes. In fact, transcript abundance between higher nodes appears to be relatively stable across time. Also, although transcript abundance appears to be relatively stable over time, there is evidence that changes in transcript abundance do occur over time, because the comparisons between the first and last date a node was available generally gave the largest number of significantly different genes. In addition, transcription factors were identified that may play important roles in controlling these processes.

Limitations of Microarray-Based Studies

Microarrays have been in use for roughly 15 years [1], and in that time have revolutionized biological research [2]. Microarrays have enabled a number of studies to be performed on a nearly genome-wide scale, including the monitoring of steady-state gene expression [3], comparative studies such as those presented here, locating regions of copy number changes in cancers [4], mapping the binding sites of transcriptional regulators (ChiP-on-Chip) [5], and surveying long-range DNA

interactions [6]. This myriad of studies has vastly improved our understanding of genetic interactions and control. However, microarray-based techniques are not without limitations.

Microarray studies are limited by reliance on existing knowledge of genome sequence, high background levels caused by cross-hybridizations, and limited dynamic range [7]. Microarrays are designed *a priori* based on known genomic or mRNA sequences, meaning the investigator determines which genes to include on the array, thus limiting their ability to investigate transcriptional activities of other genes in the same experiment [6]. Cross hybridization makes it difficult to compare findings across experiments, or even across technical replications, without complicated mechanisms of normalization [7]. Microarrays also have limited dynamic range, meaning it is difficult to detect transcription at very low or at very high levels [2], though this limitation is addressed to a certain extent by performing scans at multiple power settings [8]. These limitations along with the fact that microarrays do not return absolute expression levels [2] restrict their utility in certain studies. Emerging technologies have been devised that overcome these limitations and can move the field of transcriptomics to a new plane.

Further, genomic-scale studies, including microarrays, are limited by a lack of functional annotation which restricts the interpretation of biological underpinnings. While the goal of comparative transcriptional studies may be to uncover physiological or biochemical functions of genes, that goal would be more easily achieved if the functions of more genes were known in greater detail. Perhaps it would be optimal to use a combination of classical genetics and transcriptomics to better elucidate genetic functions. This could be accomplished by a step-wise regimen of identifying interesting target genes through comparative microarray analyses then altering these genes by classical techniques, such as reverse genetics or over-expression. In certain cases microarray studies should be able to build off each other, For instance, the binding sites of the transcription factors identified in the rice development study could be explored through ChiP-on-Chip analysis to further determine what genes, if any, they regulate directly.

The Future of Transcriptomics: RNA-Seq

The emerging next-generation sequencing technologies have been formidable challengers to microarrays as the technology of choice for comparative transcriptomic studies. In essence, massively parallel sequencing technologies are used to determine the sequence of all transcripts from a cell, these sequences are then aligned against a reference genome or a comprehensive set of genes to determine what genes are transcribed and in what amount, a process termed RNA-Seq [7]. This open-platform technology is not limited in the same ways as microarrays such as a pre-determined set of probes. In addition, dye labeling and hybridizations are not needed, which significantly reduce the experimental artifacts. Since sequencing is not based on a pre-determined chip format, this method

can survey all possible transcripts from a sample, though a reference genome or EST library is usually required to do so. It also does not have an upper or lower bounds for the number of transcripts that can be identified, thus offering a far greater dynamic range than is possible with microarrays. The ability to detect low or high transcript levels by this method is determined by the "depth" of sequencing performed, that is, by the number of runs the sample is put through. This ability is to a certain extent offset by the cost of performing multiple runs on a sample. In addition, samples processed by RNA-Seq show little to no background noise and so do not require intensive normalization to allow for comparisons to be made across samples or between replicates. Further, it has been determined that RNA-Seq has a greater capacity to determine absolute expression levels than does microarray-based approaches [2], making this technique more valuable in terms of modeling and evaluating regulatory processes.

However, no technology is without limits. While RNA-Seq offers several biologically relevant advantages when compared to microarrays, it also has obstacles to overcome. First, is deciding the depth at which to run the sequencing experiment. While the cost of sequencing has rapidly been declining [6], each run adds substantial cost to the overall experiment. In addition, massively parallel sequencing generates enormous amounts of data, so much so that it can be difficult to process, transmit, or even store. For instance, the raw image files from a single run of next generation sequencers can require terabytes of storage space [6], and the bioinformatic challenge of processing this much data is not trivial. Finally, RNA-Seq shares at least one limitation with microarray-based studies, the lack of functional genetic annotation.

Conclusion

The work described herein adds to our understanding of fundamental biological processes. Through our analyses, we have identified various genes that may play important roles in controlling and/or coordinating the genetic interactions underlying nitrogen response in maize roots and development in rice using traditional microarray techniques. There is reason to believe that future studies using RNA-Seq will further improve our understanding of the genetic networks controlling these processes, but that transcriptomics will eventually need to be paired with classical genetics to truly elucidate the functions of the genes involved.

References

1. Schena M, Shalon D, Davis RW, Brown PO: **Quantitative monitoring of gene expression patterns with a complementary DNA microarray**. *Science* 1995, **270**(5235):467-470.
2. Fu X, Fu N, Guo S, Yan Z, Xu Y, Hu H, Menzel C, Chen W, Li Y, Zeng R *et al*: **Estimating accuracy of RNA-Seq and microarrays with proteomics**. *BMC Genomics* 2009, **10**:161.
3. Cheadle C, Fan J, Cho-Chung YS, Werner T, Ray J, Do L, Gorospe M, Becker KG: **Stability regulation of mRNA and the control of gene expression**. *Ann N Y Acad Sci* 2005,

1058:196-204.

4. Nannini M, Pantaleo MA, Maleddu A, Astolfi A, Formica S, Biasco G: **Gene expression profiling in colorectal cancer using microarray technologies: results and perspectives**. *Cancer Treat Rev* 2009, **35**(3):201-209.
5. Wang JC: **Finding primary targets of transcriptional regulators**. *Cell Cycle* 2005, **4**(3):356-358.
6. Wilhelm BT, Landry JR: **RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing**. *Methods* 2009.
7. Wang Z, Gerstein M, Snyder M: **RNA-Seq: a revolutionary tool for transcriptomics**. *Nat Rev Genet* 2009, **10**(1):57-63.
8. Skibbe DS, Wang X, Zhao X, Borsuk LA, Nettleton D, Schnable PS: **Scanning microarrays at multiple intensities enhances discovery of differentially expressed genes**. *Bioinformatics* 2006, **22**(15):1863-1870.