

FOCUS PAPER

Insights into the genomic nitrate response using genetics and the Sungear Software System

Rodrigo A. Gutiérrez^{1,3}, Miriam L. Gifford¹, Chris Poultney², Rongchen Wang⁴, Dennis E. Shasha², Gloria M. Coruzzi¹ and Nigel M. Crawford^{4,*}

¹ Department of Biology, New York University, New York, NY 10003, USA

² Courant Institute of Mathematical Sciences, New York University, New York, NY 10003, USA

³ Departamento de Genética Molecular y Microbiología, Pontificia Universidad Católica de Chile, Santiago, Chile

⁴ Section of Cell and Developmental Biology, Division of Biological Sciences, University of California at San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0116, USA

Received 3 January 2007; Revised 16 March 2007; Accepted 20 March 2007

Abstract

Nitrate is both a nutrient and a potent signal that stimulates plant growth. Initial experiments in the late 1950s showing that nitrate enhances nitrate reductase (NR) activity after several hours of treatment have now progressed to transcriptome studies identifying over 1000 genes that respond to μM levels of nitrate within minutes. The use of an *Arabidopsis* NR-null mutant allowed the identification of genes that respond to nitrate when the production of downstream metabolites of nitrate is blocked. Further dissection of the nitrate response is now possible using new bioinformatic tools such as Sungear to perform comparative studies of multiple transcriptome responses across different laboratories and environmental conditions. These analyses have identified genes and pathways (e.g. nitrate assimilation, pentose phosphate pathway, and glycolysis) that respond to nitrate under a variety of conditions (context-independent). Most of these genes and pathways are ones that were identified using the NR-null mutant as responding directly to nitrate. By contrast, other processes such as protein synthesis respond only under a subset of conditions (context-dependent). Data from the NR-null mutant suggest these latter processes may be regulated by downstream nitrogen metabolites.

Key words: Genomics, microarray, nitrate, Sungear.

Introduction

The molecular analysis of the nitrate response began in the late 1950s with measurements of activity of a key enzyme in the nitrate assimilation pathway, nitrate reductase (NR). NR was found to be an adaptive enzyme as its activity was induced after exposure to nitrate for several hours to several days (Tang and Wu, 1957; Hewitt and Afridi, 1959). It was later shown that nitrate induced the *de novo* synthesis of NR (Zielke and Filner, 1971; Somers *et al.*, 1983; Remmler and Campbell, 1986). Once the gene coding for NR was cloned, it was shown that NR mRNA is rapidly induced by nitrate even in the presence of protein synthesis inhibitors (reviewed by Redinbaugh and Campbell, 1991; Wang *et al.*, 2000). Subsequently, nitrate responses were demonstrated for other genes whose products were involved in nitrate and general nitrogen (N) metabolism including nitrate transporters, nitrite reductase (NiR), glutamine synthetase, and glutamate synthase (reviewed by Redinbaugh and Campbell, 1991; Wang *et al.*, 2000). Genes involved in synthesizing organic acids for ammonium assimilation and genes involved in energy metabolism, especially in reduced ferredoxin production, and in NAD(P)H synthesis by the pentose phosphate pathway were also shown to be nitrate-induced in early work (reviewed by Redinbaugh and Campbell, 1991; Stitt, 1999; Wang *et al.*, 2000).

These foundational studies established the outlines of the nitrate response: induction of the nitrate assimilation pathway along with supporting systems to produce reductant and organic acids. To build on these early studies in

* To whom correspondence should be addressed. E-mail: ncrawford@ucsd.edu

the new genomic era, microarray and quantitative real time PCR (Q-PCR) analyses of *Arabidopsis* mRNA were employed. These studies were performed using a variety of environmental and treatment conditions. These growth and treatment conditions are of key importance as they dictate the biological context of the response.

Biological context for nitrate treatments

In general, hydroponics is the system of choice for studies involving rapid changes of specific nutrients, although other approaches have been employed (e.g. agarose plates). Plants are initially grown in liquid media under conditions where they have sufficient N, which can be provided as nitrate, ammonium, urea, or amino acids, or which can come from internal seed stores. For plants such as barley or corn, the seed reserves provide sufficient reduced N so that seedlings can be grown initially on N-free media. For plants with minimal seed reserves such as *Arabidopsis*, an external supply of N is required. Once plants have reached the desired size (1–6 weeks are typical), they are treated with nitrate. For *Arabidopsis*, two different approaches have been used for nitrate treatments. In the first approach, plants are deprived of N for a limited time (typically 1–2 d) to deplete any nitrate before the treatment. Alternatively, seedlings can be grown continuously on media with a nitrate-free N source such as ammonium then treated with nitrate. Both approaches have advantages and disadvantages. In the first approach, which employs N deprivation, plants are shifting their metabolism to salvaging internal N supplies and, if the N deprivation is prolonged, become stressed. The addition of nitrate not only provides nitrate but also an N source that relieves the N deprivation. In the second approach, the plants always have an external N supply, but the ammonium causes acidification and can repress some nitrate responses such as induction of nitrate uptake.

Another component of the media that has to be considered is sucrose. To achieve the most rapid growth and best germination for *Arabidopsis* in hydroponics, sucrose is included in the medium. Inclusion of sucrose requires that plants be grown under axenic, closed conditions, and is most useful for studying young seedlings that are 10–15-d-old. Alternatively, plants can be grown for 5–6 weeks in open trays without sucrose. Growth without sucrose better mimics soil-grown plants, but has the disadvantage that growth and nitrate response of young seedlings are less robust. For this review, we focus exclusively on experiments where sucrose is present in the media and nitrate is the only variable. However, it is clear that the nitrate response is greatly affected by the carbon (C) status of the plants (Palenchar *et al.*, 2004; Price *et al.*, 2004; Gutiérrez *et al.*, 2007). Responses to C treatments are much stronger than to N-only treatments in plants starved of carbon

(Palenchar *et al.*, 2004; Price *et al.*, 2004), and most genes that respond to N do so in a carbon-dependent manner (Gutiérrez *et al.*, 2007).

Biological context for microarray experiments

The first genome-scale experiments to examine the nitrate response were carried out in submerged cultures with high ammonium concentrations (20 mM) (Wang *et al.*, 2000). To protect plants from acidification induced by ammonium, a pH buffer and succinate were included. After 10 d of growth, plants were treated with nitrate using two different protocols to identify two classes of genes. The first condition used low nitrate (250 μ M) for a short period of time (20 min) to identify genes that are early responders, which would include primary response genes that are induced or depressed directly by nitrate. The second treatment used high nitrate (5–10 mM) for a longer time (2 h) to capture as many nitrate-responsive genes as possible. Because the plants were submerged in liquid, it was not possible to separate roots from shoots.

In the second set of microarray experiments, plants were grown on a platform so that only roots would be submerged (Wang *et al.*, 2003). This improved system allowed separation of roots from shoots and analysis of an organ-specific response. In addition, the media had lower levels of ammonium (5 mM), which was completely depleted by day 10 when the nitrate was supplied. This experiment focused on genes that responded rapidly to low levels of nitrate (250 μ M for 20 min).

In a third set of experiments, which employed N deprivation, plants were grown on 3 mM nitrate, 1 mM ammonium, and 1 mM Gln for 7 d, then transferred to very low N (0.15 mM nitrate and 0.05 mM ammonium) for 2 d (Scheible *et al.*, 2004). After 2 d of N deprivation, several physiological markers were checked, and it was found that the plants showed signs of stress (e.g. reduced chlorophyll content, accumulation of anthocyanins, 50-fold lower Gln). Plants were then treated with 3 mM nitrate for 30 min and 3 h. Because plants were grown in submerged culture, roots could not be separated from shoots, and thus only whole plants were analysed. Among other things, this report provided important genomic and metabolite data on how plants respond to N deprivation.

Lastly, an NR-null mutant was used to identify genes that respond specifically to nitrate (Wang *et al.*, 2004). If grown on ammonium succinate as the sole N source, this mutant showed equivalent growth to wild-type plants. This mutant is valuable because nitrate is not metabolized. When wild-type plants are treated with nitrate, nitrate is converted to ammonium and incorporated into amino acids, thus producing downstream metabolites. Thus, one cannot be sure that a response to nitrate treatment is due to nitrate directly or to the synthesis of a downstream metabolite of

nitrate. The mutant allows one to distinguish between these possibilities as no downstream metabolites are produced from nitrate. These experiments examined the response in roots and shoots to 5 mM nitrate after 2 h. Those genes that respond equivalently in the wild-type and mutant plants are considered direct responders to nitrate.

Microarray analyses of the nitrate response

By 2000 it was known that nitrate induced about a dozen genes involved in nitrate assimilation, ammonium assimilation, energy metabolism, the switch from starch to organic acid production, and making non-symbiotic haemoglobin (reviewed by Stitt, 1999; Wang *et al.*, 2000). The first microarray analysis of the nitrate response used a cDNA array for 5.5K genes, which encompasses about 20% of the *Arabidopsis* genome, on whole plants in submerged cultures (Wang *et al.*, 2000). Reassuringly, it identified known nitrate-induced genes including the NR genes *NIA1* and *NIA2*, NiR, and the nitrate transporter gene *NRT1.1*. A minimum of 2-fold response averaged over two biological replicates was used to designate genes as nitrate-responsive based on statistical analyses of arrays using uniprobe controls (i.e. analysis of variation found using the same probe for both experiments showed that a ratio of 1.74 was minimal for detection of differential expression) (Wang *et al.*, 2000).

The results from the cDNA array more than doubled the number of known nitrate-responsive genes. The newly identified nitrate-responsive genes could be grouped into metabolic genes and potential regulatory genes. Among the metabolic genes, the two that most clearly affected a specific pathway were transketolase and transaldolase, which are part of the non-oxidative branch of the pentose phosphate pathway. The oxidative branch converts glucose-6-phosphate to ribulose-5-phosphate and produces NAD(P)H. The non-oxidative branch converts ribulose-5-phosphate to glyceraldehyde-3-phosphate, which is a glycolytic intermediate. All of the remaining nitrate-responsive metabolism genes were scattered over the metabolic map and included asparagine synthetase (*ASN2*), a malate dehydrogenase, and sulphite reductase. The ammonium transporter gene *AMT1.1* was strongly repressed by high nitrate at 2 h.

Besides metabolic genes, a small group of potential regulatory genes was identified. These included a potential MYB transcription factor similar to the two circadian regulators *LHY* and *CCA1*, two potential protein kinases, and the response regulator *ARR6*.

The main observations from this first set of experiments were that there were many more nitrate-responsive genes than previously identified and that some encoded potential regulatory proteins. The most responsive genes were ones involved in nitrite reduction (e.g. a gene encoding NiR

and pentose phosphate pathway genes), presumably to prevent the accumulation of toxic nitrite. About twice as many genes were found to respond to high nitrate at 2 h than low nitrate after 20 min in whole plants. There was evidence that the high ammonium in the medium was repressing some of the nitrate response as *NRT2.1*, a highly nitrate-induced gene, was not responding to nitrate except under the most vigorous treatments.

The next round of experiments improved both the culture conditions and the microarray analysis, resulting in a vastly expanded data set (Wang *et al.*, 2003). Plants were grown in hydroponic conditions so that roots could be separated from shoots, and the Affymetrix ATH1 chip encoding probes for almost 23K genes (about 85% of the genome) was used. The Affymetrix system is much more sensitive than the cDNA array because statistically significant response ratios at about 1.3 or greater could be detected. For the culture conditions the initial ammonium concentration was decreased to 5 mM, and the final concentration was below detection when nitrate was added at 10 d. Treatment with low nitrate (250 μ M) for 20 min was used to identify rapidly responding genes. Two biological replicates were performed.

These improvements led to a major increase in the number of nitrate-responsive genes identified and to detection of organ specificity in the response. Almost 1200 genes showed a significant increase or decrease in roots (555 genes induced and 621 depressed using the Affymetrix software for making I and D calls). Only 183 responsive genes were found in the shoots. Genes involved in glycolysis, trehalose-6-phosphate metabolism, iron transport and metabolism, and sulphate uptake and reduction were identified as nitrate-responsive. The responses in iron and sulphate metabolic genes were not unexpected as nitrate influences sulphate assimilation (Reuveny *et al.*, 1980), and iron is needed by many of the enzymes required for nitrate assimilation including NR, NiR, and ferredoxin. The effects on glycolytic and trehalose metabolic genes were unexpected. It is possible that the activation of the glycolytic genes may support the pentose phosphate pathway and energy production. Another finding was that not all genes within a gene family were nitrate-responsive under these conditions. For example, among the *NRT2* nitrate transporters, *NRT2.1*, *2.2*, and *2.4* were induced but *NRT2.3*, *2.6*, and *2.7* showed no response. Among the ferredoxin:NADP(H) oxidoreductase (*FNR*) genes, two root *FNR* genes were induced, but the two leaf *FNR* genes were not. In addition, there were almost 60 putative regulatory genes that responded in roots.

The major findings from this study were that almost 10% of the detectable transcriptome responds rapidly to low nitrate and that roots were much more responsive than shoots under these conditions. It was reasoned that it was the slower delivery of nitrate via the xylem to the shoots

compared with the immediate exposure of the roots that was responsible for the lower response in shoots; however, it was possible that shoots have a weaker response.

Microarray analysis using an NR-null mutant

To develop these studies further, an NR-null mutant was constructed (Wang *et al.*, 2004). An NR-null mutant is unable to metabolize nitrate, and thus no downstream metabolites are made from it. *Arabidopsis* has two NR genes, *NIA1* and *NIA2*. Previously, an NR double mutant (*G'4-3*) had been made that contained a deletion of the major gene *NIA2* and a point mutation in the minor gene *NIA1* (Wilkinson and Crawford, 1993). However, this mutant was not a true null as it showed some growth on nitrate as the sole N source and retained 1% of wild-type activity in shoots and almost 10% of wild-type activity in roots (Wilkinson and Crawford, 1993; Lejay *et al.*, 1999). Thus, a true null mutant was needed; however, it was not known if an NR-null mutant would be viable. By crossing the *nia2* deletion mutant with an *nia1-Ds* insertion mutant and by testing a variety of media to rescue the double mutant, it was found that a viable NR-null mutant could be obtained if it was germinated on plates before transfer to autoclaved vermiculite soil and if it was provided with ammonium succinate throughout its life cycle (Wang *et al.*, 2004). This mutant had no detectable NR activity and did not grow on nitrate as the sole N source.

The nitrate response in the NR-null mutant was then tested. One typical response was absent in the mutant: breakdown of starch upon relief of N deprivation. When wild-type plants are deprived of N, they accumulate starch, which is broken down when N in the form of nitrate or ammonium is reapplied. In the mutant, nitrate failed to induce starch depletion but ammonium still did, indicating that downstream metabolites of nitrate were signalling starch depletion.

A microarray analysis was then performed. Both wild-type and NR-mutant plants were grown with ammonium succinate, which results in equivalent growth. Plants were treated with high nitrate (5 mM) for 2 h, then mRNA in roots and shoots were examined using the ATH1 chip. A total of 1596 genes were induced or repressed in wild-type shoots and roots in both biological replicates (Table 1). Out of these 1596 genes, 595 were equivalently induced or repressed in the NR-null mutant (Class I, Table 1). This established a set of genes that could be identified as direct

nitrate responders as the response does not require nitrate reduction. Upon grouping genes into functional categories as specified by the Munich Information Center for Protein Sequences (MIPS), it was found that these genes were most over-represented in the categories of energy, general metabolism, amino acid metabolism, glycolysis, and N and S metabolism and transport, suggesting that these processes were most impacted directly by nitrate.

The gene set described above (responding in both wild-type and mutant plants) is only one of three classes of genes. Class II genes respond in the wild type but not in the mutant (Table 1). This class may respond to downstream metabolites of nitrate as they require nitrate reduction to be induced or depressed. However, because the NR-null mutant was a hybrid between two different ecotypes (Columbia and Landsberg), any difference in response could be due to polymorphisms between these ecotypes and not to the NR mutation. An NR-null mutant with a pure ecotype background is needed to examine this group of genes. The third class responded only in the NR-null mutant (Class III, requires no nitrate reduction to respond). As yet the significance of this class is not known.

When organ-specific responses were examined, several interesting findings were obtained. First, the total number of responding genes in roots with high nitrate after 2 h (979) was no more than for low nitrate after 20 min (1176) found in the previous study. Thus, the quantitative root response appears to saturate by 20 min because longer times and higher nitrate do not increase the number of responsive genes. Shoots, however, show a large

Table 1. *Classes of genes that respond to nitrate*

| Gene class | No. of genes | Genotype |
|------------|--------------|----------------------|
| Total | 1596 | Wild type |
| Class I | 595 | Wild type and mutant |
| Class II | 492 | Wild type only |
| Class III | 392 | Mutant only |

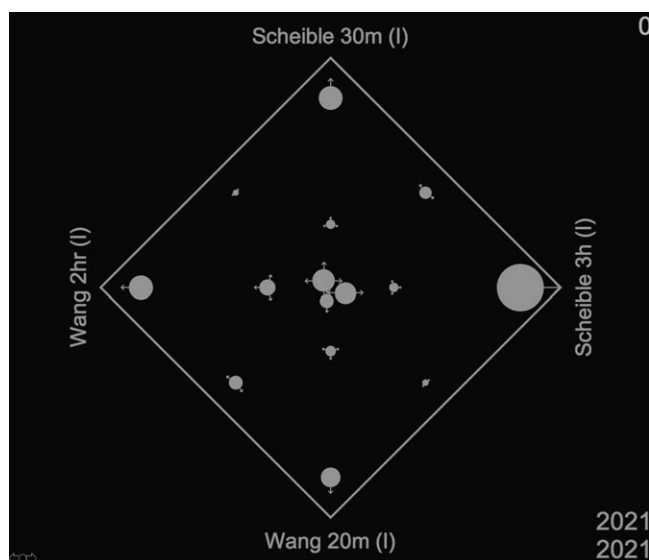


Fig. 1. Sungear representation of four data sets. The Sungear polygon is shown with names of the four data sets at the vertices (anchors). The circles inside the polygon (vessels) represent genes that are shared by the anchors indicated by the arrows around the vessels. The area of each vessel is proportional the number of genes associated with that vessel.

increase from 183 at low nitrate for 20 min to 897 at high nitrate for 2 h. Thus, shoots can respond to the same approximate extent as roots; it is concluded that shoots are not weaker but simply slower responders. One caveat for such comparisons of total numbers of responding genes is that the number is heavily skewed by the most weakly responding genes. However, the same pattern is found when a cut-off of 2-fold is used.

A surprising finding from this analysis was that the majority of induced or depressed genes did not overlap in roots and shoots (Wang *et al.*, 2004). For example, 353 genes were uniquely induced in shoots versus roots while only 174 genes overlapped. Thus, there is substantial

organ-specificity to the nitrate response under these conditions.

One of the most interesting genes responding directly to nitrate is isopentenyl transferase 3 (*IPT3*). This protein catalyses the first and rate-limiting step in cytokinin biosynthesis, and the promoter of *IPT3* is expressed in phloem and is induced by nitrate (Miyawaki *et al.*, 2004). In our microarray data, *IPT3* is much more strongly induced in roots (23- and 7-fold in wild type and mutant, respectively) than in shoots (about 2-fold). It has been proposed that this gene is responsible for the nitrate-enhanced synthesis of cytokinin in *Arabidopsis* (Miyawaki *et al.*, 2004; Wang *et al.*, 2004).

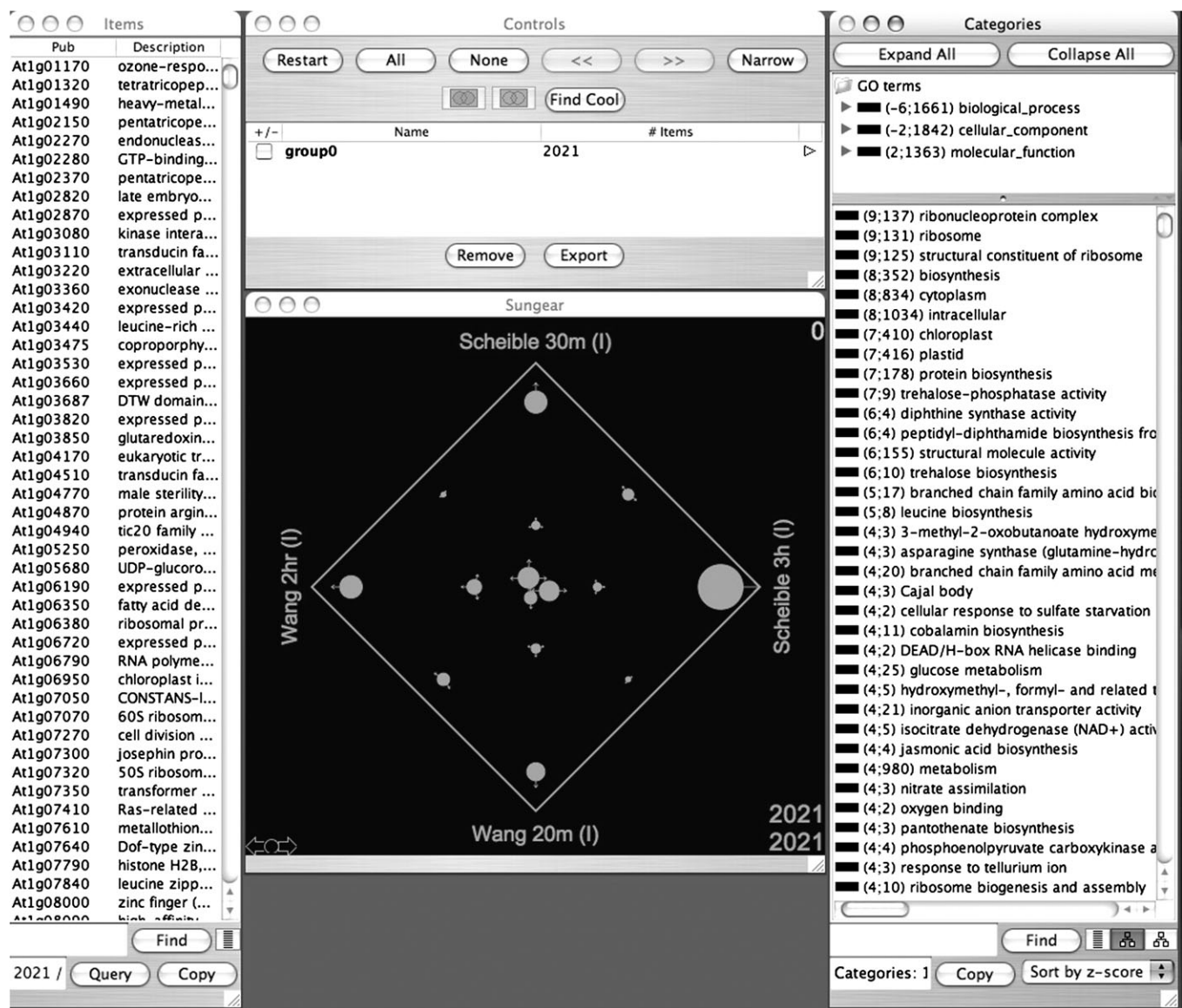


Fig. 2. Full Sungear window for the analysis of four data sets. Shown is the complete Sungear page for the polygon in Fig. 1. The left panel provides a list of genes corresponding to the selected vessels (in this case all four data sets). The panel on the right shows the GO terms (functional categories) to which these genes belong. The GO terms are ranked by the *z*-score value, which represents the extent of over-representation of genes in that category. The numbers in parentheses (#,#) next to each GO term refer to the *z*-score and the number of genes in that category, respectively.

Nitrate response following N deprivation

In the last set of microarray experiments, plants were N-deprived for 2 d before nitrate treatment (Scheible *et al.*, 2004). Submerged cultures were used so that only whole plants were analysed. These experiments provided several key data sets: the transcriptome response during N deprivation and N resupply and concomitant changes in key metabolite levels. It was found that N deprivation resulted in a decline in mRNA for genes involved in photosynthesis, chlorophyll synthesis, plastid protein synthesis, and in an increase in mRNA for many genes involved in secondary metabolism. Gln and Glu levels were substantially lower (50-fold and 6-fold, respectively), and 2-oxoglutarate levels were 2- to 3-fold higher in the deprived plants. Re-addition of nitrate led to a rapid increase (30 min) in genes involved in nitrate uptake and metabolism and in the production of reductant and organic acids, even though actual levels of primary metabolites did not change significantly. Longer nitrate exposure (3 h) resulted in induction of genes involved in amino acid and nucleotide synthesis and scavenging, RNA synthesis and processing, and protein synthesis. Genes involved in cell expansion and growth were also induced while genes involved in amino acid and nucleotide breakdown were depressed. Many potential regulatory genes were also uncovered. An extensive RNA analysis of transcription factor genes was performed using Q-PCR, and genes that were below the detection of the ATH1 chip were found to be nitrate-responsive. The 3 h treatment resulted in a 2–2.5-fold increase in the number of responding genes (e.g. 1292 induced genes) compared with the 30 min treatment (e.g. 645 induced genes).

Now that multiple large data sets of nitrate-responsive genes are available, there is a need to perform comparative analyses of such data sets to identify genes and pathways shared in common and those that are unique among these data sets. Simple Venn diagrams or spreadsheets are not sufficient for this type of analysis. A new tool that can take advantage of existing bioinformatic resources is needed.

Analysis of multiple transcriptome data sets using Sungear

Currently, several thousand microarray data sets are publicly available for several model organisms (Parkinson *et al.*, 2005). The challenge ahead is to derive robust biological insights and biological hypotheses from this vast amount of data. A new bioinformatics tool called Sungear has been developed to do just this (Poultney *et al.*, 2007). Sungear generalizes Venn diagrams to view multiple collections of genes, relates those collections to functional categories, and permits visual, real-time, statistically based, data exploration (Poultney *et al.*, 2007).

Whereas Venn diagrams are, in practice, limited to comparisons of two or three data sets, Sungear can represent a larger number and is limited only by the researcher's willingness to understand a visual display having many components (anchors and associated vessels). The Sungear software is freely available over the web (<http://virtualplant-prod.bio.nyu.edu/cgi-bin/sungear/index.cgi>) and is one of the analysis tools on the VirtualPlant web site (<http://www.virtualplant.org>). Users can also obtain an individual copy of the software upon request. Documentation for Sungear is available online (<http://virtualplant-prod.bio.nyu.edu/public/docs/>).

To use Sungear, lists of genes are uploaded as TAIR locus identifiers (At#g#####) into the VirtualPlant web site. Gene lists of interest (called gene sets) are selected and then imported into Sungear, which shows in graphic form how many genes are shared by two or more gene sets. Sungear creates a polygon and then assigns each vertex of the polygon to a gene set. These vertices are called 'anchors'. Inside the polygon there are circular forms called 'vessels', each of which corresponds to a group of genes present in all the gene sets corresponding to a particular set of anchors. The position of the vessel is dictated by the positions of the anchors, and the area of the vessel is proportional to the number of genes associated with that vessel.

Sungear was used to analyse the nitrate response in *Arabidopsis* plants from four of the data sets described above (rapidly induced genes after 20 min and 30 min; Wang *et al.*, 2003; Scheible *et al.*, 2004) and genes induced after 2 h and 3 h in wild-type plants only;

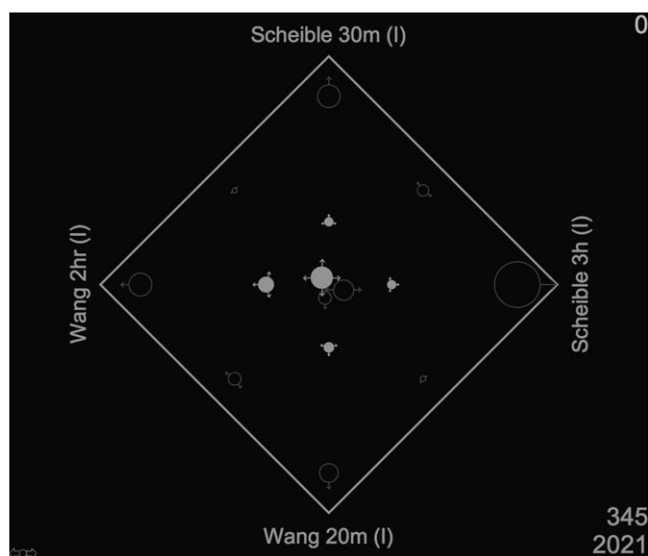


Fig. 3. Sungear polygon for the intersection among four data sets. The same polygon in Fig. 1 is shown after narrowing it down to the genes that appear in three or more sets (have three or more anchors in common). Vessels containing genes from at least three anchors are highlighted.

Scheible *et al.*, 2004; Wang *et al.*, 2004). These data sets were selected as they all share the features that *Arabidopsis* seedlings were transiently treated with nitrate and that gene responses were analysed using the ATH1 Affymetrix chips. However, the data came from experiments conducted by two different research groups, and thus experimental conditions such as growth, pretreatment, and treatment conditions vary as described above. The root and shoot data sets of Wang *et al.* were combined for each condition in order to compare directly with the whole plant data sets of Scheible *et al.*

Sungear analysis of these data sets generated a four-sided polygon corresponding to 2021 total genes (Fig. 1). The anchors are labelled with the names of each data set at each vertex. Sungear shows the complete list of genes from all four data sets (union of the data sets) to the left of the polygon window (Fig. 2). To the right of the polygon are shown the functional categories to which these genes belong. The categories are ranked by their statistical over-representation as indicated by a *z*-score (number of standard deviations from the mean). The higher the *z*-score, the more significant is the result. The most significant

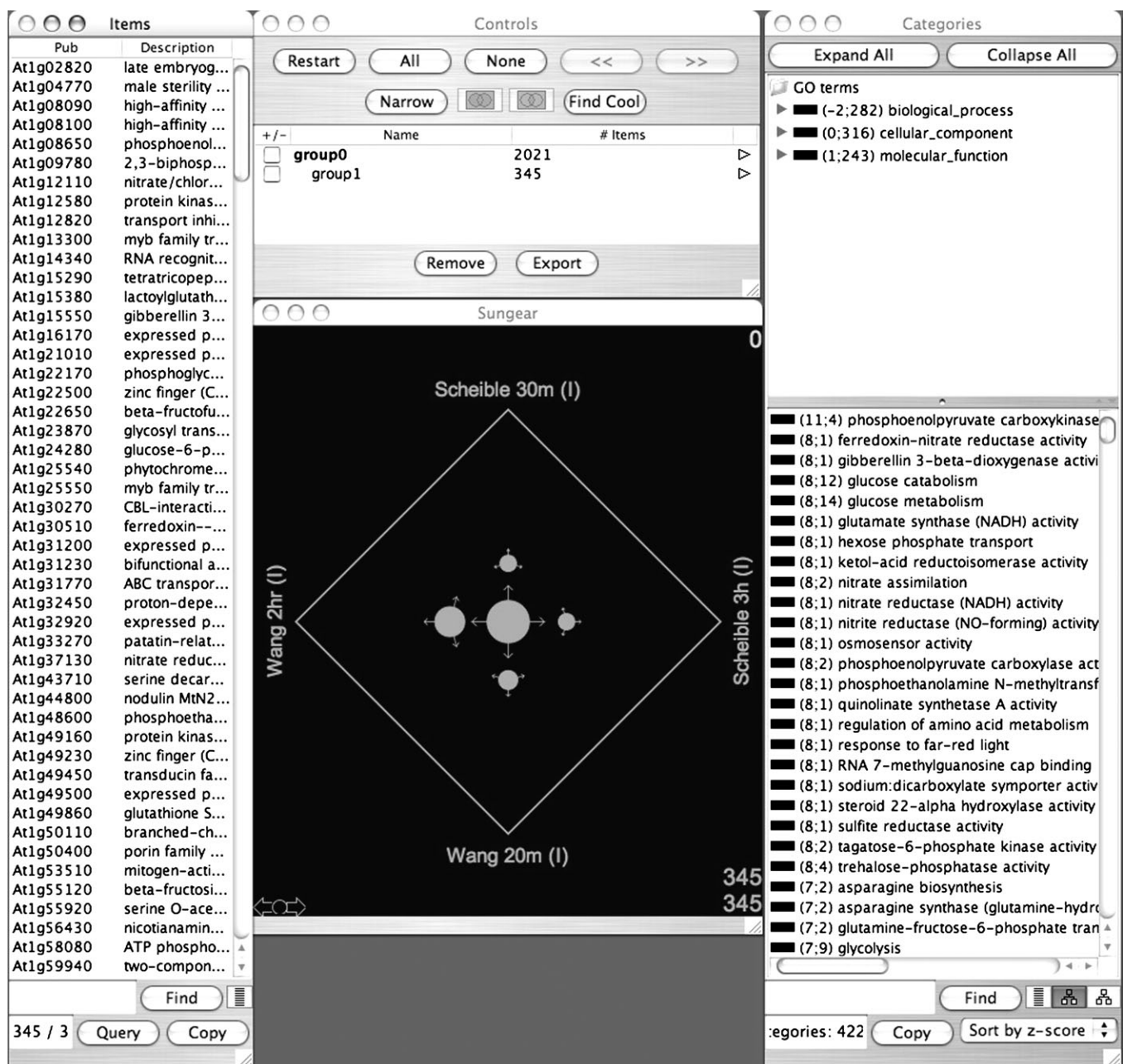


Fig. 4. Full Sungear analysis of intersection sets. Shown is the complete Sungear window for the polygon in Fig. 3.

categories were related to ribosomes, protein synthesis, and chloroplasts/plastids. Trehalose metabolism was also included in the top groups. These results, using the union of all the data sets, would suggest that nitrate most strongly affects genes involved in protein synthesis and ribosome structure.

Context-dependent and independent nitrate-induced genes

Instead of analysing all the genes in the union of these data sets, one can select those genes that are shared by three or more sets (intersection). This will identify genes that are commonly regulated across different experiments (context-independent). Using Sungear, one can easily select vessels corresponding to three or more anchors, which creates a group of 345 genes (see highlighted vessels in Figs 3, 4). The analysis of these genes gives a different conclusion to that obtained from the 2021 genes comprising the union of the data sets. The Sungear analysis shows that 345 genes are most over-represented in the categories that involve nitrate assimilation, glucose metabolism/glycolysis, and organic acid production (PEP carboxykinase and carboxylase) (Fig. 4). These 345 genes were exported to BioMaps, another analysis tool at the Virtual-Plant site. BioMaps allows a rigorous non-parametric statistical analysis of over-representation using hypergeometric distribution and generates a *P*-value. The lower the *P*-value is, the higher the significance of the over-representation. A BioMaps analysis of the 345 gene set showed that the highest ranking functional categories were energy, glycolysis and gluconeogenesis, general and C metabolism, pentose-phosphate pathway, N and S metabolism, amino acid metabolism, and assimilation of ammonia (Table 2). These categories are almost identical to those found for genes that respond directly to nitrate (induced in shoots and roots of both wild-type and NR-null mutant plants, 368 total) (Table 3). In fact, comparison of the specific genes shows that 75% (260 out of 345) of the genes that respond in at least three of the gene sets were also induced in the NR-null mutant.

This analysis identifies genes and associated pathways that are nitrate-induced in different laboratories using different treatment conditions. These genes are likely to be those that are universally induced by nitrate and include genes involved in nitrate assimilation, glycolysis, amino acid and organic acid metabolism, and energy production. This finding is confirmed by mutant analysis showing that most of these genes respond directly to nitrate and do not require NR to be induced.

Those genes not shared by the majority of data sets would be more experiment-specific (context-dependent). For example, genes that were over-represented in the categories of ribosomes and protein synthesis (from the

Sungear analysis of the 2021 genes obtained by the union of all data sets) were much more prevalent in the 2–3 h treatments. Out of the 131 genes in the ribosome category, 117 were in the Wang 2 h and Scheible 3 h groups. Out of the 178 genes in the protein synthesis category, 152 were in the Wang 2 h and Scheible 3 h groups. Interestingly, these categories also contain over-represented genes that are induced in wild-type but not the NR-null mutant, suggesting that they might respond to a downstream metabolite, which would explain why they require some time after nitrate treatment to respond.

Conclusion

The application of transcriptome technologies and bioinformatics has resulted in a comprehensive picture of the nitrate response. We know almost all the *Arabidopsis*

Table 2. *BioMaps analysis of context-independent genes*

| Category | Observed frequency | Expected frequency | <i>P</i> -value |
|---------------------------------|-----------------------------------|------------------------------|-----------------|
| Energy | 31 genes, 9.0% | 1.50% | 1.1E-13 |
| Glycolysis and gluconeogenesis | 16 genes, 4.6% | 0.60% | 1.1E-08 |
| Metabolism | 55 genes, 15.9% | 6.50% | 3.9E-08 |
| C and carbohydrate metabolism | 29 genes, 8.4% | 2.80% | 1.0E-05 |
| Pentose-phosphate pathway | 6 genes, 1.7% | 0.10% | 2.6E-05 |
| Nitrogen and sulphur metabolism | 9 genes, 2.6% | 0.30% | 3.0E-05 |
| Amino acid metabolism | 15 genes, 4.3% | 0.90% | 3.9E-05 |
| Assimilation of ammonia | 7 genes, 2.0% out of 345 genes | 0.20% out of 26 444 genes | 7.3E-05 |

Table 3. *BioMaps analysis of direct nitrate-responding genes*

| Category | Observed frequency | Expected frequency | <i>P</i> -value |
|---------------------------------|-----------------------------------|------------------------------|-----------------|
| Energy | 37 genes, 10.1% | 1.50% | 4.4E-18 |
| Metabolism | 69 genes, 18.8% | 6.50% | 7.8E-14 |
| Glycolysis and gluconeogenesis | 21 genes, 5.7% | 0.60% | 2.3E-13 |
| Transported compounds | 24 genes, 6.5% | 1.10% | 2.6E-10 |
| C and carbohydrate metabolism | 35 genes, 9.5% | 2.80% | 2.1E-08 |
| Amino acid metabolism | 19 genes, 5.2% | 0.90% | 7.6E-08 |
| Ion transport | 15 genes, 4.1% | 0.50% | 8.4E-08 |
| Nitrogen and sulphur metabolism | 11 genes, 3.0% | 0.30% | 3.7E-07 |
| Cellular transport | 31 genes, 8.4% | 2.70% | 2.4E-06 |
| Assimilation of ammonia | 8 genes, 2.2% | 0.20% | 7.0E-06 |
| Transport facilitation | 21 genes, 5.7% | 1.60% | 3.2E-05 |
| Anion transport | 8 genes, 2.2% | 0.20% | 3.4E-05 |
| Pentose-phosphate pathway | 6 genes, 1.6% out of 368 genes | 0.10% out of 26 444 genes | 4.0E-05 |

genes detectable by microarray technologies that are responsive to nitrate treatment in both roots and shoots under a variety of conditions. We know which genes are direct nitrate responders based on experiments using rapid treatments with low nitrate and using a NR-null mutant. We know which processes and pathways are universal responders (responding to nitrate regardless of the N growth and treatment conditions) and which are context-dependent (responding to a subset of conditions). We also have clues about which processes and pathways respond to downstream N metabolites of nitrate. Other work has provided insights into how nitrate responses are influenced by C status. Future work will identify regulatory genes that mediate the nitrate response and further integrate the nitrate response into the overall signalling network of the plant.

Acknowledgements

This work was funded by grants from the National Science Foundation (IIS-9988345, IIS-0414763, DBI-0445666) and 0115586 to DES; grants from the National Science Foundation – N2010 (IBN0115586) and (DBI-0445666) and the National Institutes of Health (GM32877) to GMC; a grant from the National Science Foundation (DBI-0445666) to RAG; an EMBO postdoctoral fellowship ALTF107-2005 to MLG; and grants from the National Institutes of Health (GM40672) and the National Science Foundation (DBI-0115586 and IOB-0519985) to NMC.

References

- Gutiérrez RA, Lejay L, Dean D, Chiaromonte F, Shasha DE, Coruzzi GM.** 2007. Qualitative network models and genome-wide expression data define carbon/nitrogen-responsive molecular machines in Arabidopsis. *General Biology* (in press).
- Hewitt EJ, Afridi MMRK.** 1959. Adaptive synthesis of nitrate reductase in higher plants. *Nature* **183**, 57–58.
- Lejay L, Tillard P, Lepetit M, Olive FD, Filleur S, Daniel-Vedele F, Gojon A.** 1999. Molecular and functional regulation of two NO₃-uptake systems by N- and C-status of Arabidopsis plants. *The Plant Journal* **18**, 509–519.
- Miyawaki K, Matsumoto-Kitano M, Kakimoto T.** 2004. Expression of cytokinin biosynthetic isopentenyltransferase genes in Arabidopsis: tissue specificity and regulation by auxin, cytokinin, and nitrate. *The Plant Journal* **37**, 128–138.
- Palenchar PM, Kouranov A, Lejay LV, Coruzzi GM.** 2004. Genome-wide patterns of carbon and nitrogen regulation of gene expression validate the combined carbon and nitrogen (CN)-signaling hypothesis in plants. *Genome Biology* **5**, R91.
- Parkinson H, Sarkans U, Shojatalab M, et al.** 2005. ArrayExpress—a public repository for microarray gene expression data at the EBI. *Nucleic Acids Research* **33**, D553–555.
- Poultney CS, Gutiérrez RA, Katari MS, Gifford ML, Paley WB, Coruzzi GM, Shasha DE.** 2007. Sungear: interactive visualization and functional analysis of genomic datasets. *Bioinformatics* **23**, 259–261.
- Price J, Laxmi A, St Martin SK, Jang JC.** 2004. Global transcription profiling reveals multiple sugar signal transduction mechanisms in Arabidopsis. *The Plant Cell* **16**, 2128–2150.
- Redinbaugh MG, Campbell WH.** 1991. Higher plant responses to environmental nitrate. *Physiologia Plantarum* **82**, 640–650.
- Remmler JL, Campbell WH.** 1986. Regulation of corn leaf nitrate reductase. II. Synthesis and turnover of the enzyme's activity and protein. *Plant Physiology* **80**, 442–447.
- Reuveny Z, Dougall DK, Trinity P.** 1980. Regulatory coupling of nitrate and sulfate assimilation pathways in cultured tobacco cells. *Proceedings of the National Academy of Sciences, USA* **77**, 6670–6672.
- Scheible WR, Morcuende R, Czechowski T, Fritz C, Osuna D, Palacios-Rojas N, Schindelasch D, Thimm O, Udvardi MK, Stitt M.** 2004. Genome-wide reprogramming of primary and secondary metabolism, protein synthesis, cellular growth processes, and the regulatory infrastructure of Arabidopsis in response to nitrogen. *Plant Physiology* **136**, 2483–2499.
- Somers DA, Kuo T-M, Kleinhofs A, Warner BL, Oaks A.** 1983. Synthesis and degradation of barley nitrate reductase. *Plant Physiology* **72**, 949–952.
- Stitt M.** 1999. Nitrate regulation of metabolism and growth. *Current Opinion in Plant Biology* **2**, 178–186.
- Tang PS, Wu HY.** 1957. Adaptive formation of nitrate reductase in rice seedlings. *Nature* **179**, 1355–1356.
- Wang R, Guegler K, LaBrie ST, Crawford NM.** 2000. Genomic analysis of a nutrient response in Arabidopsis reveals diverse expression patterns and novel metabolic and potential regulatory genes induced by nitrate. *The Plant Cell* **12**, 1491–1509.
- Wang R, Okamoto M, Xing X, Crawford NM.** 2003. Microarray analysis of the nitrate response in Arabidopsis roots and shoots reveals over 1000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. *Plant Physiology* **132**, 556–567.
- Wang R, Tischner R, Gutiérrez RA, Hoffman M, Xing X, Chen M, Coruzzi G, Crawford NM.** 2004. Genomic analysis of the nitrate response using a nitrate reductase-null mutant of Arabidopsis. *Plant Physiology* **136**, 2512–2522.
- Wilkinson JQ, Crawford NM.** 1993. Identification and characterization of a chlorate-resistant mutant of *Arabidopsis* with mutations in both *NIA1* and *NIA2* nitrate reductase structural genes. *Molecular and General Genetics* **239**, 289–297.
- Zielke HR, Filner P.** 1971. Synthesis and turnover of nitrate reductase induced by nitrate in cultured tobacco cells. *Journal of Biological Chemistry* **246**, 1772–1779.