Funct Integr Genomics (2003) 3:135–143
DOI 10.1007/s10142-003-0093-7

ORIGINAL PAPER

John Pufky · Yang Qiu · Mulpuri V. Rao · Patrick Hurban · Alan M. Jones

The auxin-induced transcriptome for etiolated *Arabidopsis* seedlings using a structure/function approach

/ Received: 23 June 2003 / Revised: 15 September 2003 / Accepted: 7 October 2003 / Published online: 26 November 2003

Abstract To increase our understanding of the mode of action of auxin, we analyzed auxin-induced changes in the *Arabidopsis* transcriptome with microarrays representing 20,426 *Arabidopsis* genes. Treatment of etiolated seedlings with low concentrations of the auxin, indole-3-acetic acid (IAA), decreased the expression levels of 23 genes, whereas it increased the expression levels of 47 genes within 20 min. After 40 min, the directional trend in genomic change was predominantly an increase in gene expression. Among the most rapidly induced changes are those in genes encoding transcription factors. Promoter regions of transiently induced genes contained DNA motifs that bind auxin response (ARFAT) and silence element binding factors whereas genes induced by IAA during the entire experimental period contained MYC and ARFAT DNA motifs at higher frequencies. Six structurally diverse auxins were analyzed to determine genes that are unique to a specific auxin, as well as a common set of genes that are rapidly regulated by all tested auxins, thus enabling the identification of shared DNA motifs. In addition to ARFAT, analysis of promoter regions of genes induced by all six auxins revealed the presence of an abscisic-acid-responsive DC3 promoter-binding factor and low temperature responsive elements suggesting a possible role for abscisic acid in modulating auxin-induced responses.

J. Pufky and Y. Qiu contributed equally to this work

J. Pufky · Y. Qiu · M. V. Rao · P. Hurban (✉)
Paradigm Genetics,
108 Alexander Drive, Research Triangle Park, NC 27709-4528, USA
e-mail: phurban@paragen.com
Tel.: +1-919-4253000
Fax: +1-919-5448094

A. M. Jones
Department of Biology and Curriculum in Genetics and Molecular Biology,
University of North Carolina,
Chapel Hill, NC 27599-3280, USA

Keywords Abscisic acid · *Arabidopsis* · Auxin · Gene expression · Auxenic herbicides

Electronic Supplementary Material Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s10142-003-0093-7

Introduction

Because auxins are hormones that are central to plant growth and development and because they represent an important class of growth regulators for agriculture, the mode of auxin action is of major interest. However, little is known about auxin receptors and elements that reside apically in this signal transduction network. It is clear that auxins evoke rapid changes in the cell (Jones 1994; Napier 1995). The plasma membrane becomes hyperpolarized, the pH of the cell wall space and cytoplasm, respectively, decreases and transiently increases, cell wall viscoelastic properties change, and the movement of cell wall precursors and cell wall enzymes occurs. Nuclear changes are also equally rapid (Abel and Theologis 1996; Hagen and Guilfoyle 2002). Using standard differential screening approaches several genes whose expression are altered by auxin application have been identified, most notably a large gene family called *Aux/IAA*, a small family defined by the prototype gene called *GH3*, and a large family of small auxin-up-regulated genes (*SAUR*s). Some members of these gene families show increased expression within 5 min of auxin application, others are not detectably increased until 30 or 60 min suggesting a cascade of gene expression control. Presumably, those genes that are expressed the earliest are candidates for the primary response genes that in turn control expression of downstream gene expression. Consistent with this notion is that most of the *Aux/IAA* genes, encoding components of positive and negative transcriptional complexes, are rapidly induced by auxin even in the presence of translational inhibitors. Recently, Che et al. (2002)

analyzed gene expression in *Arabidopsis* explants on auxin- and cytokinin-containing media over a time course of days using a microarray containing approximately 8,000 gene probes. Numerous auxin response genes were found to be expressed in tissues exposed to auxin. Expression of these genes is associated with global developmental changes that occur as cells change competency for organogenesis.

We took a genomic approach to address the mode of auxin action. Specifically, we asked what are the genes whose expression is changed by applied auxin within the rapid time period known to alter growth of etiolated *Arabidopsis* seedlings? When do these changes occur? What set of genes are specific to auxin action? By asking this last question, we anticipated that auxins may have indirect effects on the cell that are attributable to chemical structures present in specific auxins that are irrelevant to their auxenic mode of action. Therefore, we included six structurally diverse compounds in order to extract the "auxin signature" among the gene profiles. These compounds are indole-3-acetic acid (IAA), naphthalene-1-acetic acid (NAA), 1,4 dichloro, 5-methoxy benzoic acid (dicamba), 3,7 dichloroquinoline-8-carboxylic acid (quinclorac), 1,2,4 trichloro, 3-aminopicoline-5-carboxylic acid (pichloram), and 2,4-dichlorophenoxy-1-acetic acid (2,4-D). At low concentrations, all of these compounds exhibit an auxenic mode of action in that they stimulate cell elongation as defined for auxin classification (Weed Society of America 1989; van Overbeek 1964).

By focusing on the early changes induced by auxin, as revealed by a microarray that presently covers approximately 75% of the *Arabidopsis* genome, and by replicating the experimental design of earlier studies using *Arabidopsis*, we hoped to obtain a nearly complete set of genes that are rapidly induced by auxin. Completion of this set frames studies on the auxin-regulated DNA motifs (ARFAT). In addition to the known ARFAT, we found that DNA motifs previously noted for regulation by the plant hormone abscisic acid (ABA) were frequently associated with these gene promoters.

Materials and methods

Plant growth and auxin application

Approximately 500 surface-sterilized seeds of *Arabidopsis* (accession Col-0) were sown in a sterile environment on filter disks overlaying solid 1/2 MS media 1% sucrose, stratified at 4°C for 48 h and grown in darkness at 25°C for 4 days. Seedlings were gently submerged by application to the plates of the different auxin solutions (made in ethanol carrier; 0.1% final concentration) and, at the indicated times, the seedlings were lifted from the plate en masse and flash frozen in liquid nitrogen.

RNA extraction and microarray manufacture

Frozen seedlings were lyophilized, and then pulverized with a bead beater. Ten milligrams lyophilized and pulverized tissue was rehydrated using 100 μl RNALater (Ambion), followed by extraction of total RNA with Trizol reagent (GibcoBRL, Rockville, Md.) following the manufacturer's protocol. RNA quality was determined using the Bioanalyzer 2100 and the RNA 6000 assay (Agilent Technologies, Palo Alto, Calif.) according to the manufacturer's instructions. RNA concentrations were determined by measuring the absorption at 260 nm using an Ultrospec 2000 (Pharmacia Biotech, Piscataway, N.J.). To minimize biological variation and increase the accuracy of our results, total RNA was isolated from tissue pooled from approximately 1,500 seedlings receiving the same treatment in three independent experiments. Microarrays containing 20,426 *Arabidopsis thaliana* gene probes as specified by Paradigm Genetics (representing 75% of the *Arabidopsis* genome assuming a total of 27,117 genes) were manufactured by Agilent Technologies by in situ synthesis of 60-mer oligonucleotides using an ink-jet printing method as described by Hughes et al. (2001). The microarray design file was deposited at http://www.ncbi.nlm.nih.gov/geo/ with platform accession number GPL483.

Microarray analyses

cRNA labeled with either cyanine 3-CTP or cyanine 5-CTP was generated from each RNA sample using the Fluorescent Linear Amplification Kit (Agilent) according to the manufacturer's protocol. Hybridization solutions that contained 2,000 ng each of labeled cRNA from vehicle-treated control plants and auxin-treated plants were prepared using the In situ Hybridization Reagent Kit (Agilent). Each pair of samples to be hybridized was independently labeled and hybridized utilizing fluor-reversal for a total of two hybridizations per sample pair. Hybridized microarrays were scanned using the Agilent LP2 DNA Microarray Scanner at 10 μm resolution.

Microarray images were visually inspected for defects, and were analyzed using Feature Extraction Software (Version 6.1.1, Agilent), using the average foreground signal intensity from a group of negative control features to determine background. The resulting dataset was inspected for unusual trends in the foreground and background signal values. Individual features with undesirable pixel intensity statistics (e.g., non-uniformity or saturation in either channel) were excluded from the data set and from all downstream analyses. In the event that greater than 2.5% of gene features on an array were excluded for any reason, then the entire microarray result was rejected, and the hybridization was repeated.

Data from individual arrays was loaded into the Resolver Gene Expression Data Analysis System (Version 3.0, Rosetta Biosoftware, Kirkland, Wash.), wherein data from fluor-reversal replicate hybridizations were combined. Dual criteria consisting of a p-value threshold of $\leq$0.005 and a $\geq$2-fold change for the time course experiment (Figs. 1, 2) and $\geq$ 1.5-fold change for the six-auxin experiment (Fig. 4) were applied to define signature data points (both more abundant and less abundant). For the six-auxin treatment experiment, we also applied the criterion that a gene must be represented in at least one of the six treatments. These loosened criteria enable us to create the extended gene set needed for the subsequent agglomerative cluster analysis performed in Resolver using Wards Minimum Variance as the heuristic criteria and Euclidean Distance as the distance metric. Experimental data were deposited at http://www.ncbi.nlm.nih.gov/geo/ with series number GSE664.

Functional annotation

To assess global changes in expression patterns of genes related by function, both auxin-up-regulated and auxin-down-regulated genes were sorted into categories based on their assigned function by Gene Ontology Consortium (available at http://www.arabidopsis.org/info/ ontologies/go/). Where possible and necessary, additional functional annotation was provided for genes listed as null or unknown by iterative BLAST searches for suggested function. In approximately one third of the cases a clear function was suggested by this analysis. When the result remained ambiguous, the gene was

left unannotated (null). This conservative approach yielded annotations for between 40% and 65% for the genes in our sets.

Promoter analysis

Arabidopsis genomic sequence was obtained from TIGR (February 2002 release, http://www.tigr.org/tdb/e2k1/ath1/) and custom Perl scripts were written to retrieve sequences 1.1 kb upstream from known or predicted start codons. For gene clusters (see Results), the 1.1-kb presumptive promoter sequence of each gene in the cluster was queried for the presence of discrete DNA motifs against the PLACE database (Higo et al. 1999) using the Signal Scan program provided at http://www.dna.affrc.go.jp/htdocs/PLACE/signalscan.html. The frequency of occurrence (*f*) for each motif within a gene cluster was calculated by dividing the number of occurrences of a particular motif within the gene cluster by the total number of genes in the cluster. To estimate the significance of *f* for a gene cluster, 20 randomly generated gene clusters containing the same number of genes were produced, *f* was determined for each random gene cluster, and the average *f* and standard deviation were derived for the 20 random gene clusters. A significant difference in the *f* between test gene clusters and random gene clusters was calculated by pair wise, Student's *t*-tests. The *t*-test results represented in units of standard error (SE) indicate the normalized differences between the *f* of the test gene cluster and of the random gene clusters. A DNA motif was considered significantly overrepresented in a given cluster if its random gene cluster *f* is <1.0, and it is present in at least 50% of test genes within the cluster with 2.2 or more SE ($p \leq 0.05$) above the *f* of that element in a random gene set. To ensure that these results were not skewed by the inclusion of both evidence-based and ab initio predicted genes in our randomly generated gene clusters, a subset of tests was conducted using clusters containing only evidence-based genes. The results were not significantly different.

Results

In an attempt to understand the early mode of action for auxin, we combined genomic and structure/function approaches. Our study incorporated two compatible variables, time after auxin application and six structurally different auxins, to reveal differences in the expression profiles of 20,426 *Arabidopsis* genes. To identify genes responsive to auxin, 4-day-old Col-0 seedlings were treated with 10 μM IAA and samples collected at 0, 20, 40 and 60 min after the initiation of auxin treatments. Care was taken to replicate as closely as reasonable the conditions of the original experiments that identified the first auxin-regulated genes from two decades earlier (Hagen et al. 1984; Theologis et al. 1985). These studies utilized etiolated seedlings which, in several species, have been shown to respond to exogenous auxin (Melhado et al. 1981). Etiolated seedlings have also been used in genetic screens for auxin-sensitive mutants (Boerjan et al. 1995; Reed et al. 1998; Watahiki and Yamamoto 1997).

To minimize process-related variation, 1,500 seedlings receiving the same treatment from three independent experiments were pooled twice, first by combining the tissue and then by preparing three separate RNA samples and combining these after quality control to generate a single sample for each treatment and time point. A preliminary study with lower density (8K genes) Agilent oligonucleotide microarrays indicated that the correlation of log ratio calls between experiments reproducibly exceeds 0.95. Total RNA was isolated, labeled with cyanine-3 and cyanine-5 and hybridized to microarrays consisting of 60-mer oligonucleotide probes representing 20,426 genes (an estimated 75% of the annotated genome). The microarrays used in these studies include nearly all known auxin-responsive genes. A comparison of the content of our microarray against a compilation of auxin-responsive genes (Abel and Theologis 1996) revealed that the compiled list was well represented on our array. Our array included 27 sequences in the Aux/IAA family, 15 sequences in the SAUR family, 14 sequences in the GH3 gene family, 4 sequences coding for ACC synthase, ARG1, pCNT115, 16 ARFs, and 34 miscellaneous auxin-induced factors.

We applied stringent control measures during all steps of data analysis. These include fluor-reversal replicate hybridizations and establishment of dual criteria as described in Materials and methods. Data from replicate fluor-reversal hybridizations were combined and all primary analysis was performed with Resolver software using an error model to determine boundaries for significance of the results. The *p*-value generated by the error model for each data point estimates the confidence with which a data point can be considered to reflect differential expression of the corresponding gene. By using $p \leq 0.005$ as a cutoff value, we found that treatment of seedlings with auxin for 20, 40 and 60 min altered the expression patterns of 477, 933 and 311 genes, respectively (Fig. 1; for entire probe set see Electronic Supplementary Material 1). However, to increase the relevance of the auxin-responsive transcriptome, we imposed a second criterion of twofold change on these 1,421 auxin-responsive genes. Using dual criteria, we found that the expression of 70 genes was altered by at least a factor of 2 within 20 min of the application of 10 μM IAA in 0.1% ethanol to 4-day-old etiolated seedlings. The ratios of gene expression between IAA-treated and control ethanol-treated samples for each time point were generated and genes with at least a twofold change in expression relative to the controls are summarized in Fig. 1A. Note that the unclustered distribution of genes with at least a twofold differential expression across the range of hybridization intensities, which is roughly proportional to expression levels, indicates an unbiased representation. Furthermore, we have effectively reduced the number of false positives, since many genes that have a greater than twofold change in expression were not selected (arrows, Fig. 1A) because the *p*-values did not meet the criteria described above.

After 20 min of auxin application, 47 genes had increased while 23 genes decreased expression (Fig. 1B). Twenty minutes later, the character of the gene expression profiles dramatically shifted so that the preponderant direction was positive, due primarily to a set of genes that were transiently expressed at 40 min (discussed later). The expression of 106 genes increased twofold or more while only 23 genes had expressions that were at least twofold less than the control. This trend remained at

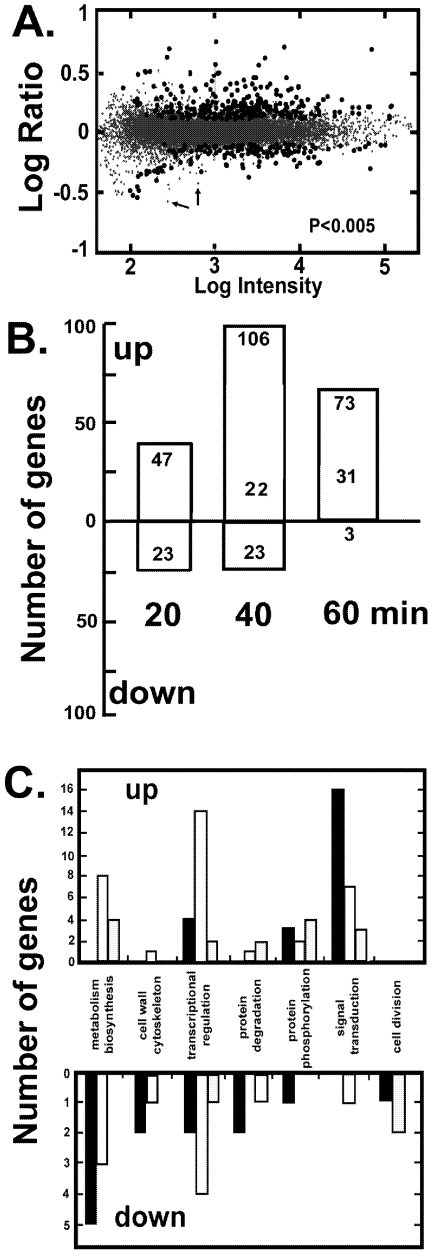

Fig. 1A–C RNA transcript profiling of *Arabidopsis* etiolated seedlings treated with indole-3-acetic acid (IAA) over time using microarrays containing 20,426 open reading frames. **A** The ratios of expression in auxin treatment vs control (0.1% ethanol) treatment were determined for 20,426 genes and plotted as a function of their normalized intensity (relative expression) for the indicated times after auxin application. Genes that met the criteria of having a fold-change ≥ 2 and a p value ≤ 0.005 are *emboldened* (see Materials and methods for criteria details), against the background of points representing ratios at intensities that do not meet these strict criteria. Points indicated by an *arrow* are two examples of values that appear to meet the fold-change requisite but fail the p value criterion. **B** The number of genes with a twofold or more change in gene expression compared to seedlings treated with 0.1% ethanol (control) is shown for three time points after auxin treatments. Down- and up-regulated genes were binned as gene sets. *Numbers within the bars* represent the gene set size. *Shaded areas* of bars indicates genes (*number indicated*) that were already up-regulated at the earlier time point. *Bold numbers* below the pairs of bars represent time after auxin application in minutes. **C** The genes at each time point represented in **B**, were grouped into functional categories according to their annotations as described in Materials and methods and plotted to illustrate the number of genes. *Black bars* 20 min, *open bars* 40 min, *grey bars* 60 min after auxin application. The entire gene sets with annotations are provided in the Electronic Supplementary Material

60 min and genes in this final time point had relatively large fold-changes (Fig. 1B). Over the entire 60-min time course, 183 genes had altered expression after auxin application as determined by both p-value and fold-change.

As expected, the 183 genes identified in the time course experiment (see Electronic Supplementary Material 2) included known auxin-regulated genes such as *Aux/IAA*, *GH3*, and *SAUR*s. However, approximately 90% of the genes in this set have not been previously recognized as auxin-inducible. For example, among the genes with the greatest differential expression at 20 min were genes encoding a receptor kinase (At3g45420), homeobox factors (At4g17460, At5g47370), a zinc-finger transcription factor (At1g68360), an F-box containing protein (At1g78100) and an NBR LRR disease-resistance protein (AT4g27190). Expression of genes encoding pectate lyase (At1g14420), calcium-dependent kinase (At5g19360), heme oxygenase (At1g69720), an additional NBR LRR disease-resistance protein (At1g63350), and ubiquitin (At5g09340) decreased the most. Expression of a gene encoding an FH2-containing protein (At5g58160) increased nearly tenfold after 40 min of auxin application. FH2-containing proteins are involved in actin re-arrangement through interaction with Rho-GTPases and profilin (Imamura et al. 1997). Three replicated genes (At1g29440, At1g29450, At1g29460), annotated as putative auxin regulated genes, show similar expression patterns supplanting ab initio evidence for gene function and also suggesting that this gene replication was relatively recent in *Arabidopsis* genome evolution.

To help elucidate the biological roles of the genes determined to be auxin responsive, we assembled groups based on gene ontology classifications. Figure 1C illustrates the changes in assigned cell functions over time following auxin application. At 20 min, the cellular priority is to increase expression of the *Aux/IAA*, *SAUR*, and *GH3* genes (categorized as "signal transduction"), while decreasing expression of genes encoding metabolic enzymes. A large percentage of these genes encoded elements of transcriptional control. By 40 min after auxin application, the priority switches to transcription. Genes encoding these factors are potential targets of the Aux/IAA proteins in the 20-min gene set. Many of these encoded functions relate to wall growth (e.g. At1g0565, At1g1442, At1g3322, At1g5149, At1g6506) and cellular metabolism (e.g. At1g4380, At3g6151, At4g2542, At5g4888).

Since each gene identified as auxin responsive can have a different pattern of expression through the three time points examined, we visualized these trends to illustrate any emergent dynamic patterns in gene expres-

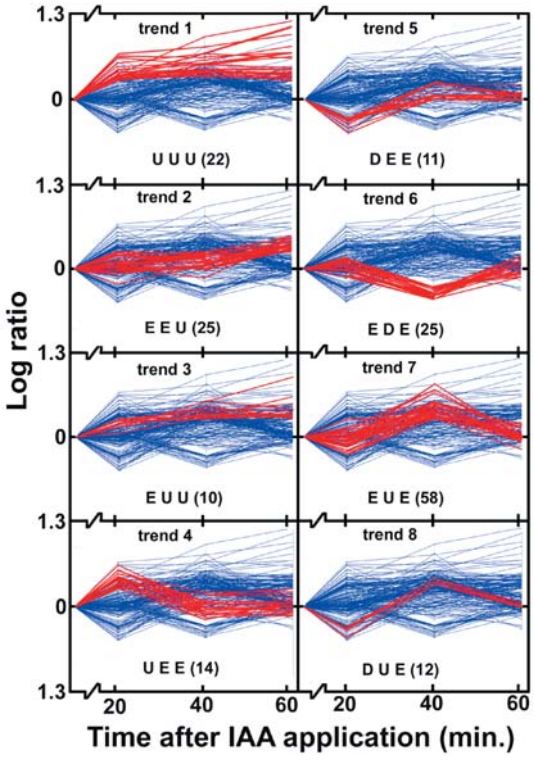

Fig. 2 Analysis of coordinated gene regulation. Eight expression pattern trends over time for those genes meeting the auxin-induced criteria were identified and plotted as log ratios (from Fig. 1) vs time after auxin application. The trend is represented by the pattern for individual genes highlighted in *red* against the entire auxin-regulated gene set shown in *blue*. Each trend is designated a number representing the gene subset subsequently used in the promoter analysis described in Fig. 3. The number of genes in this subset is indicated in *parentheses*. *U* Up, *D* down, *E* even (statistically unchanged). For example, among genes in the auxin-regulated gene set, "UUU" designates a subset of genes whose expression is significantly greater than control at each time point over the previous point. The eight trends shown here cover 177 of the genes in the 183 gene set described in Fig. 1; one trend of 6 genes was too small for consideration

sion (Fig. 2; gene IDs for each trend are shown in the Electronic Supplementary Material 3). A group of 22 genes was up-regulated within 20 min of auxin exposure, and their expression levels remained persistently higher than controls (trend 1). However, the most prominent pattern was transient up-regulation, with 54 genes comprising this set. For example, a group of 14 genes (trend 4) is up-regulated at 20 min after auxin exposure, with expression levels equivalent to the control at the remaining times. In contrast, a group of 58 genes (trend 7) is comparable to control expression levels at 20 min, up-regulated in response to auxin by 40 min, declining by 60 min. Altogether, 47, 64, and 25 genes are increased at 20, 40 and 60 min, respectively. Taken together, this trend analysis suggests a potential cascade of regulatory factors that are directly responsive to auxin, which then propagate the physiological response to auxin by, in turn, regulating a discrete set of secondary response genes.

Regulation of plant gene expression is generally assumed to result from a highly coordinated binding of different types of transcription factors in specific combinations. These regulatory pathways require the combinatorial coordination of highly specific DNA-protein and protein-protein interactions. (Singh et al. 2002) Therefore, to understand the molecular mechanisms that regulate gene expression in response to auxin treatments, we analyzed DNA motifs in the 1.1-kb 5′-upstream regions of all auxin-responsive genes grouped into different trends based on the kinetics of their expression. Promoter analysis of transiently induced genes (trend 4; UEE) revealed that binding sites for silencing element binding factor (SEBF) and auxin response factor (ARF; 64% of gene set) occur at higher frequencies compared to those of randomly generated gene sets of similar size. Binding sites for MYC/MYB and auxin response factors (ARFAT) were enriched in genes in trend 1, representing persistently induced genes (U,U,U; Fig. 3, Electronic Supplementary Material 3). Approximately 68% of the genes in trend 1 contained at least one or more copies of MYC-binding motifs, whereas 54% of the genes contained one or more copies of ARFAT-binding DNA motifs (Fig. 3B). Analysis of promoter regions of transiently induced genes (trend 4; U,E,E) revealed that binding sites for SEBF and ARF (64% of gene set) occur at higher frequencies compared to the randomly generated gene set (Fig. 3A). In plants, auxin plays a key role in a wide variety of growth and developmental processes. Although the regulatory mechanisms by which exogenously applied auxenic compounds modulate gene expression are not clearly understood, several studies have shown that treatment of plants with auxin modulates gene expression within 4 min of treatment. (Abel et al. 1995) Studies investigating *cis*-acting elements discovered that many of the auxin-responsive genes contain either a TGTCTC or TGTCCCAT DNA motif. Site-directed mutations within the TGTCTC motif have revealed that TGTC bases are critical for ARF binding (Guilfoyle et al. 1998; Hagen and Guilfoyle 2002; Ulmasov et al. 1997).

In order to distinguish between changes induced specifically via the auxin mode of action and non-auxin and to apply a structure function approach to gene expression profiling, we compared the expression profiles for six structurally different auxins. The six auxins were chosen based on their distinct ring-backbones and their similar growth-promoting potential (Fig. 4A). The primary structural differences reside in the type of ring-backbone upon which a critical carboxylic acid group is placed (Edgerton et al. 1994). The endogenous auxin is indole-3-acetic acid (IAA). Besides indoles, auxin ring-backbones can be naphthalenes, quinolines, picolines, and benzenes. The critical structural features at the atomic level for auxin binding have only been determined for one candidate auxin receptor, namely, auxin-binding protein 1 (Woo et al. 2002). For experiments comparing the six auxins, replicate samples collected at the 20-min time point were processed and the expression patterns of combined gene sets for all six compounds is shown in

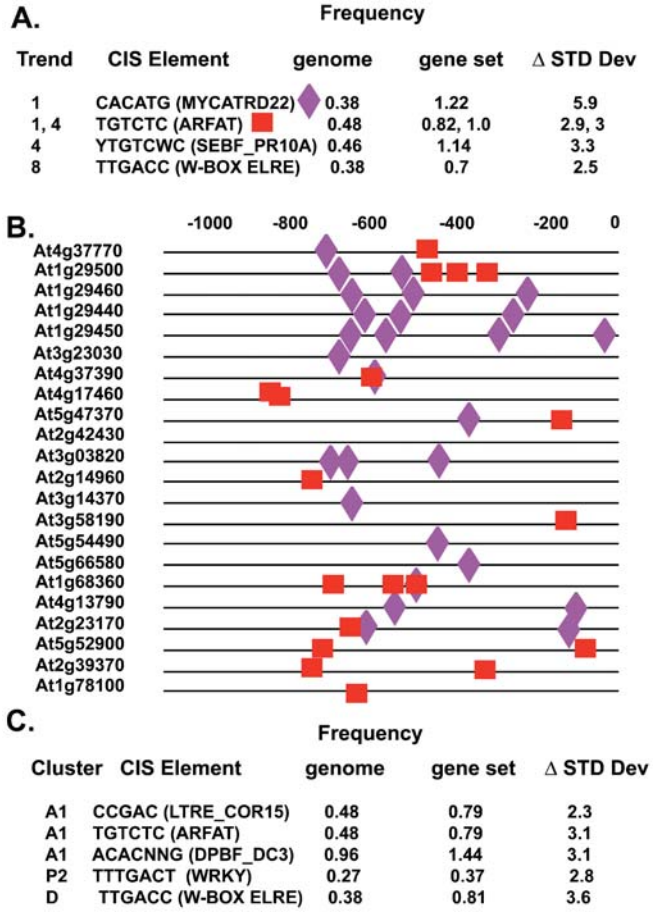

Fig. 3A-C Analysis of common and/or shared DNA motifs. **A** Promoter regions (1.1-kb 5′ UTR) for each gene in the trend subset were analyzed for DNA motifs as described in the Materials and methods. To be listed, a DNA motif had to meet three criteria: (1) it had to be present within 1.1 kb 5′ to the start codon in at least half of the genes in the trend subset, (2) the frequency of this motif within the subset had to be at least 2.2 standard deviations over the frequency of this element in the genome, and (3) the frequency of the element could not occur at a frequency greater than 0.96 in the genome. The DNA motif designation is as described in the PLACE database (http://www.dna.affrc.go.jp/htdocs/PLACE/). **B** The location of two DNA motifs meeting these criteria for the 22 genes in trend 1 are plotted as a function of distance in base pairs 5′ to the start codon. The two motifs represented by *symbols* are indicated with the corresponding symbol in **A**. **C** DNA motifs meeting the criteria described in **A** for the trend sets shown in Fig. 4 are indicated. *A1* Set of genes with expression induced by all six auxins (see Fig. 4A); *D* set of genes with expression increased by dicamba; *P2* set of genes with expression reduced by pichloram

Fig. 4B. The criteria for inclusion in this gene set are: (1) at least a 1.5-fold change in gene expression over the controls, (2) a p value ≤ 0.005, and (3) the first two criteria met in at least one of the six auxin treatments. Four hundred and eighteen genes comprise this auxin gene set (Electronic Supplementary Material 4). Not surprisingly, there is a great degree of similarity in gene expression patterns between the six auxins. However, the larger gene set we employed enabled cluster analysis to reveal that the six compounds can be weakly segregated

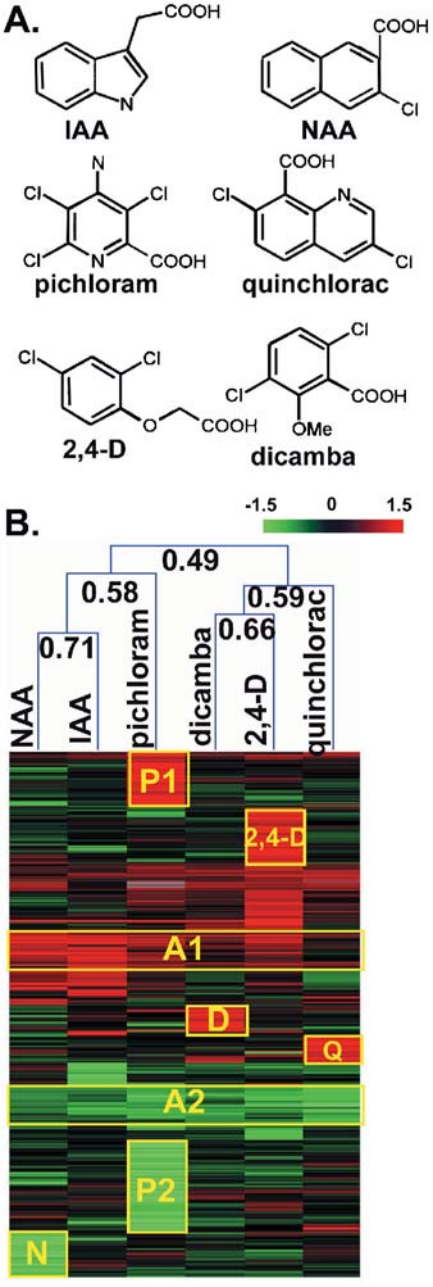

Fig. 4A, B Cluster analysis of genes altered in expression by six structurally different auxins. **A** Chemical structures and names of the six auxins used in this comparative study. **B** Cluster analyses (agglomerative algorithm) of gene expression profiles induced in etiolated *Arabidopsis* seedlings after 20 min application of the six auxins illustrated in **A**. The relative change in expression of 418 genes meeting the criteria described in the Materials and methods are displayed *top to bottom*. The dynamic range in change is shown at the bottom in log scale with −1.5 (*green*) to 1.5 (*red*). Nearest neighbor joining of the genes is omitted. The relation of the six auxins is indicated *along the top* with the correlation coefficients noted for the degree of relatedness between auxins shown. Eight clusters of genes were isolated and boxed: *A1* induced by all six auxins; *A2* repressed by all six auxins; *N* repressed by 1-NAA; *P1* induced by pichloram; *P2* repressed by pichloram; *D* induced by dicamba; *2,4-D* induced by 2,4-D; *Q* induced by quinchlorac

into two groups. Group 1 contains IAA, NAA and pichloram while group 2 contains 2,4-D, quinchlorac, and dicamba. The largest correlation coefficients tightly group IAA and NAA genes, indicating that these two auxins induce the most similar expression profiles. Smaller clusters were useful in distinguishing genes specific to a particular auxin (Electronic Supplementary Material 5) and are delineated in Fig. 4 (boxed).

To increase our understanding of the molecular mechanisms that regulate gene expression in response to various auxenic compounds, the promoter regions of genes induced by six different auxins were analyzed as described for common and shared DNA motifs. Analysis of genes induced by all six auxins (denoted as A1 in Fig. 4b) indicated the predominant occurrence of DNA motifs such as ARFAT, *Dc3* promoter-binding factor (DPBF) and low temperature responsive elements (LTRE) in a majority of genes.

Discussion

We sought a nearly complete auxin-induced transcriptome in etiolated seedlings. Our analysis identified all genes previously known to be auxin-induced, validating our methodology through replication of published work. Nonetheless, approximately 90% of the genes identified here were not previously known to be auxin-regulated. (Abel and Theologis 1996; Hagen and Guilfoyle 2002) Thus, our analysis has significantly expanded our knowledge of the auxin-regulated transcriptome, and given the comprehensiveness of our microarray, we may be approaching completion of the auxin-induced gene set for etiolated *Arabidopsis* seedlings. Several lines of evidence support the idea that these genes are indeed auxin-regulated. First, our criteria for differential expression included both an estimate of confidence based on the quality of the underlying result beginning at the pixel level, and a fold-change cutoff. We set our confidence level to $\geq$ 99% ($p \leq 0.005$) and our fold-change level to ≥ 2, thresholds that are probably overly conservative. In preliminary experiments, we labeled the same RNA sample with both cyanine-3 and cyanine-5 and hybridized both to the same array, with the expectation that no genes would show a significant difference in expression level—a self vs self experiment. In individual hybridizations, no genes were differentially expressed by our *p*-value criteria, fewer than 0.1% of genes had fold-changes greater than or equal to 2, and less than 1.4% of genes had fold-changes greater than or equal to 1.6 (data not shown). We addressed spurious gene expression, biological variation, and process variation by combining (1) biological samples to average, in effect, the response, (2) discrete RNA extractions to smooth out process variation, and (3) data from dye-flip replicate hybridizations. Vehicle controls and rapid freezing were used to eliminate genes that have altered expression from physical and carrier manipulation. A two-dimensional approach, expression over time and expression as a function of auxin structure, was employed wherein six structurally-different auxins altered a common set of genes. The observation that many genes were found to have altered expression sustained over more than one time point after auxin application further reinforces our confidence in these results. The multidimensionality of experimental design and multiple criteria to be met for selection increase confidence that genes identified in this study are genuinely regulated by auxin.

Given the profound physiological changes induced by auxin, it was surprising to find that the expression of only a relatively few genes was altered by auxin in these tissues. Based on their Gene Ontology categorizations, we conclude that the presumptive role of genes that change in expression over time as a result of auxin presentation suggest a set of cellular events set in motion. A rapid increase in auxin over basal levels in etiolated seedlings first alters the profile of transcription factors. Most of the genes with altered expression at 20 min encoded transcription factors or components of a nuclear complex such as members of the Aux/IAA family, categorized as "signal transduction" in Fig. 1C. The trend was an increase in expression although there are many genes with decreased transcriptional activity. The peak of change occurred at 40 min, suggesting a rapid response with associated desensitization for the preponderance of genes.

Coordinated expression of genes implies that these gene promoters share common DNA motifs. Not surprisingly, the well-defined ARFAT motif (Tiwari et al. 2003; Ulmasov et al. 1997, 1999) was found at least once in most of the genes that were rapidly up-regulated by auxin at the three time points assayed (trend 1; Fig. 2B). Several of these genes were previously known to be auxin regulated, but this analysis revealed two new transcription factors (At4g17460, At1g68360), an F-box containing protein (At1g78100), a LOB-domain protein (At3g58190; Shuai et al. 2002) and three unknown proteins (At2g23170, At5g52900, At2g39370) that contain at least one ARFAT motif. In addition, 9 of 14 genes in trend 4 (UEE) contained ARF motifs illustrating a correlation between genes that are up-regulated within 20 min and the presence of a DNA motif that binds ARF.

In addition to the ARFAT motif, analysis of transiently regulated genes revealed the presence of an SEBF motif whose sequence (YTGTCWC) is similar to the ARF DNA motif. Similar to the known function of auxin as a negative regulator of the expression of several defense genes, SEBF has been shown to act as a transcriptional repressor of plant defense genes (Boyle and Brisson 2001). Furthermore, analysis of genes responsive to pathogen, wounding and different stress factors by Boyle and Brisson (2001) revealed the presence of the SEBF motif which lead them to hypothesize that the coordinated binding at ARFAT and SEBF DNA motifs regulate auxin-induced gene expression. Consistent with the above hypothesis, a majority of genes transiently induced by IAA also contained DNA motifs that bind SEB factors.

Analysis of genes with increased expression during the entire experimental period indicated the occurrence of

another DNA motif that was previously not reported to play a role in auxin-induced responses at higher frequency. A large portion of genes induced by auxin on a long-term basis (represented by trends U,U,U and E,U,U) contained MYC factors previously characterized for their responsiveness to abscisic acid (ABA), hydrogen peroxide (H_2O_2) and other stress factors such as drought and salinity (Abe et al. 1997; Busk and Pages 1998; Seki et al. 2002). Furthermore, genes induced by all auxenic herbicides also contained, in addition to ARFAT, DNA motifs such as LTRE and DPBF previously characterized for ABA responsiveness. Several studies have demonstrated that the core LTRE DNA motif, CCGAC, confers gene inducibility to low temperature, drought and ABA (Jiang et al. 1996; Stockinger et al. 1997; Busk and Pages 1998). Similarly, the DPBF DNA motif, which is initially found in the carrot *DC3* gene and belongs to a novel class of bZIP transcription factors, is also characterized as responsive to ABA (Kim et al. 1997; Finkelstein and Lynch 2000).

Eight distinct sets of co-regulated genes illustrate the complexity involved in auxin-induced gene expression and shared DNA motifs in these sets reveals a possible junction for crosstalk between hormones. For example, these results suggest a possible interaction between auxin and ABA signaling pathways at the level of transcriptional control. Previous studies have shown that auxin treatment triggers the production of H_2O_2 and ethylene, which in turn triggers the biosynthesis of ABA (Hansen and Grossmann 2000). Furthermore, analysis of several of the early or primary auxin-responsive genes revealed the presence of a DNA-binding domain with similarity to maize transcriptional activator VP1 and its *Arabidopsis* ABI3 relative (Giraudat et al. 1992). The predominant enrichment of ABA-responsive DNA motifs such as LTRE, MYC and DPBF in the majority of genes induced by all auxenic herbicides clearly suggests a role for ABA in modulating auxin-induced gene expression.

Our analyses also reveal a possible feedback mechanism in auxin-induced gene expression through control of hormone pool sizes by auxin-regulated genes encoding hormone modifying enzymes. Recently, it was shown that JAR1 plus 18 related proteins are similar to adenylate-forming enzymes and some have adenylate transfer activity to jasmonic (JA), salicylic (SA) and indole-3-acetic acid substrates (Staswick et al. 2002). The JAR1 family falls into three protein clades based on sequence homologies and clade II includes eight proteins, which have been shown to use IAA as substrate; one of these also utilizes SA as a substrate. We show here that genes for six of these family members (At5g54510, At4g27260, At4g37390, At1g59500, At2g23170, At2g14960) are induced by auxin and three of these increase expression over the entire time course (trend 1). While it is not clear yet what the effect of adenylation has on IAA, increased expression of this gene by auxin suggests a feedback mechanism for both the auxin and JA pathways.

There was no a priori means to predict if the six structurally different auxins would induce similar or different expression patterns since, to our knowledge, this is the first study to compare these auxins at a molecular level. The observation here that the gene expression patterns induced by these six auxins were similar indicates for the first time that each of these compounds shares a similar mode of action.

Two decades ago the search for genes regulated by auxin in etiolated seedlings began (Hagen et al. 1984; Theologis et al. 1985) and here we report its near complete transcriptome in this tissue. The resulting 16 gene sets, distinguished by their timing of expression and by specificity to six different auxins, now provides the material for extensive analyses of the promoters for novel controlling elements. This awaits refinements in current sampling techniques (Thijs et al. 2001, 2002) to robustly predict these elements.

Acknowledgements This work was supported by the Kenan Foundation and the North Carolina Biotechnology Center by a grant awarded to A.M.J. and Paradigm Genetics and by grants from NSF and NIGMS to A.M.J. We are grateful to the North Carolina Collaborative Funding Program for the major support. We thank Mike Vernon and Ani Chatterjee for technical assistance in a preliminary study utilizing the 8 K Affymetrix arrays, Doug Boyes for the suggestion of a structure/function approach in the early phase of this project and Neil Hoffmann for his discussion and guidance.

References

Abe H, Yamaguchi-Shinozaki K, Urao T, Iwasaki T, Hosokawa D, Shinozaki K (1997) Role of *Arabidopsis* MYC and MYB homologs in drought- and abscisic acid-regulated gene expression. Plant Cell 9:1859–1868

Abel S, Theologis A (1996) Early genes and auxin action. Plant Physiol 111:9–17

Abel S, Nguyen MD, Theologis A (1995) The PS-IAA4/5-like family of early auxin-inducible mRNAs in *Arabidopsis thaliana*. J Mol Biol 251:533–549

Boerjan W, Cervera MT, Delarue M, Beeckman T, Dewitte W, Bellini C, Caboche M, Van Oncklen H, Inze D (1995) *superroot*, a recessive mutation in *Arabidopsis*, confers auxin overproduction. Plant Cell 7:1405–1419

Boyle B, Brisson N (2001) Repression of the defense gene PR-10a by the single-stranded DNA binding protein SEBF. Plant Cell 13:2525–2537

Busk PK, Pages M (1998) Regulation of abscisic acid-induced transcription. Plant Mol Biol 37:425–443

Che P, Gingerich DJ, Lall S, Howell SH (2002) Global and hormone-induced gene expression changes during shoot development in *Arabidopsis*. Plant Cell 14:2771–2785

Edgerton MD, Tropsha A, Jones AM (1994) The auxin-binding site of auxin-binding protein 1. Phytochemistry 35:1111–1123

Finkelstein RR, Lynch TJ (2000) The *Arabidopsis* abscisic acid response gene ABI5 encodes a basic leucine zipper transcription factor. Plant Cell 12:599–609

Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM (1992) Isolation of the *Arabidopsis ABI3* gene by positional cloning. Plant Cell 4:1251–1261

Guilfoyle TJ, Ulmasov T, Hagen G (1998) The ARF family of transcription factors and their role in plant hormone responsive transcription. Cell Mol Life Sci 54:619–627

Hagen G, Guilfoyle T (2002) Auxin-responsive gene expression: genes, promoters and regulatory factors. Plant Mol Biol 49:373–385

Hagen G, Kleinschmidt A, Guilfoyle T (1984) Auxin regulated gene expression in intact soybean hypocotyl and excised hypocotyl sections. Planta 162:147–153

Hansen H, Grossmann K (2000) Auxin-induced ethylene triggers abscisic acid biosynthesis and growth inhibition. Plant Physiol 124:1437–1448

Higo K, Ugawa Y, Iwamoto M, Korennaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database. Nucleic Acids Res 27:297–300

Hughes TR, Mao M, Jones AR, Burchard J, Marton MJ, Shannon KW, Lefkowitz SM, Ziman M, Schelter JM, Meyer MR, Kobayashi S, Davis C, Dai H, He YD, Stephaniants SB, Cavet G, Walker WL, West A, Coffey E, Shoemaker DD, Stoughton R, Blanchard AP, Friend SH, Linsley PS (2001) Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. Nat Biotechnol 19:342–347

Imamura H, Tanaka K, Hihara T, Umikawa M, Kamei T, Takahashi K, Sasaki T, Takai Y (1997) Bni1p and Bnr1p: downstream targets of the Rho family small G-proteins which interact with profilin and regulate actin cytoskeleton in *Saccharomyces cerevisiae*. EMBO J 16:2745–2755

Jiang C, Iu B, Singh J (1996) Requirement of a CCGAC *cis*-acting element for cold induction of the *BN115* gene from winter *Brassica napus.* Plant Mol Biol 30:679–684

Jones AM (1994) Auxin-binding proteins. Annu Rev Plant Physiol Plant Mol Biol 45:393–420

Kim SY, Chung HJ, Thomas TL (1997) Isolation of a novel class of bZIP transcription factors that interact with ABA-responsive and embryo-specification elements in the Dc3 promoter using a modified yeast one-hybrid system. Plant J 11:1237–1251

Melhado LL, Jones AM, Ho T-HD, Leonard NJ (1981) Azido auxins: synthesis and biological activity of fluorescent labeling agents. Plant Physiol 68:469–475

Napier R (1995) Towards an understanding of ABP1. J. Exp Bot 46:1787–1795

Overbeek J van (1964) Survey of mechanism of herbicide action. Academic Press, London

Reed JW, Elumalai RP, Chory J (1998) Suppressors of an *Arabidopsis thaliana phyB* mutation identify genes that control light signaling and hypocotyl elongation. Genetics 148:1295–1310

Seki M, Ishida J, Narusaka M, Fujita M, Nanjo T, Umezawa T, Kamiya A, Nakajima M, Enju A, Sakurai T, Satou M, Akiyama K, Yamaguchi-Shinozaki K, Carninci P, Kawai J, Hayashizaki Y, Shinozaki K (2002) Monitoring the expression pattern of around 7,000 *Arabidopsis* genes under ABA treatments using a full-length cDNA microarray. Funct Integr Genom 2(6):282–291

Shuai B, Reynaga-Pena CG, Springer PS (2002). The lateral organ boundaries gene defines a novel, plant-specific gene family. Plant Physiol. 129:747–761

Singh KB, Foley RC, Onate-Sanchez L (2002) Transcription factors in plant defense and stress responses. Curr Opin Plant Biol 5:430–436

Staswick P, Tiryaki I, Rowe M (2002) Jasmonate response locus, JAR1 and several related *Arabidopsis* genes encode enzymes of the firefly luciferase superfamily that show activity on jasmonic, salicylic, and indole-3-acetic acids in an assay for adenylation. Plant Cell 14:1405–1415

Stockinger EJ, Gilmour SJ, Thomashow MF (1997) *Arabidopsis thaliana CBF1* encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a *cis*-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. Proc Natl Acad Sci USA 94:1035–1040

Theologis A, Huynh TV, Davis RW (1985) Rapid induction of specific mRNAs by auxin in pea epicotyl tissue. J Mol Biol 183:53–68

Thijs G, Lescot M, Marchal K, Rombauts S, De Moor B, Rouze P, Moreau Y (2001) A higher order background model improves the detection of regulatory elements by Gibbs sampling. Bioinformatics 17:1113–1122

Thijs G, Marchal K, Lescot M, Rombauts S, De Moor B, Rouze P, Moreau Y (2002) A Gibbs sampling method to detect over-represented motifs in upstream regions of coexpressed genes. J Comp Biol 9:447–464

Tiwari SB, Hagen G, Guilfoyle T (2003) The roles of auxin response factor domains in auxin-responsive transcription. Plant Cell 15:533–543

Ulmasov T, Hagan G, Guilfoyle T (1997) ARF1, a transcription factor that binds to auxin response elements. Science 276:1865–1868

Ulmasov T, Hagen G, Guilfoyle TJ (1999) Activation and repression of transcription by auxin-response factors. Proc Natl Acad Sci USA 96:5844–5849

Watahiki MK, Yamamoto KT (1997) The *massugu1* mutation of *Arabidopsis* identified with failure of auxin-induced growth curvature of hypocotyl confers auxin insensitivity to hypocotyl and leaf. Plant Physiol 115:419–426

Weed Society of America (eds) (1989) Herbicide handbook. Weed Society of America, Champaign, Ill.

Woo EJ, Marshall J, Bauly JM, Chen J-G, Venis MA, Napier RM, Pickersgill RW (2002) Crystal structure of auxin-binding protein 1 in complex with auxin. EMBO J 21:2877–2885