

Author to whom all correspondence should be sent:

Stephen H. Howell

1073 Roy J. Carver Co-Laboratory

Iowa State University

Ames IA 50011

Telephone: 515 294-5267

FAX 515 294-5256

Email: SHH@iastate.edu

Gene expression patterns during somatic embryo development and germination in maize Hi II cultures

Ping Che¹, Tanzy M. Love², Bronwyn R. Frame³, Kan Wang³, Alicia L. Carriquiry² and Stephen H. Howell¹

¹Plant Sciences Institute, ²Department of Statistics, ³Center for Plant Transformation and the Department of Agronomy, Iowa State University, Ames IA 50011, USA

Financial Source: PC is supported by the Plant Sciences Institute, Iowa State University; TML is supported by NSF DMS-0091953. BRF is supported by the Department of Agronomy, the Plant Sciences Institute, the Biotechnology Program all at Iowa State University.

Corresponding author:

Stephen H. Howell
FAX 515 294-5256
Email: SHH@iastate.edu

ABSTRACT

Gene expression changes associated with embryogenic callus formation and with somatic embryo maturation and germination were examined in a regeneration-proficient hybrid line of *Zea mays*, Hi II. 12,060 element maize cDNA microarrays were used to generate gene expression profiles from embryogenic calli induced to undergo embryo maturation and germination. No statistically significant gene expression changes were detected in comparing embryogenic with total callus. On the other hand, over 1,000 genes showed significant time variation during somatic embryo development. In general, a substantial number of genes were downregulated during embryo maturation, largely histone and ribosomal protein genes, which may result from a slow down in cell proliferation and growth during embryo maturation. The expression of these genes dramatically recovered at germination. Other genes upregulated during embryo maturation included genes encoding hydrolytic enzymes (nucleases, glucosidases and proteases) and a few storage genes (zein and caleosin), which are good candidates for developmental marker genes. Germination is accompanied by the upregulation of a number of stress response and membrane transporter genes, and, as expected, greening is associated with the upregulation of many genes encoding photosynthetic and chloroplast components. Thus, some, but not all genes, typically associated with zygotic embryogenesis are significantly up or downregulated during somatic embryogenesis in Hi II maize line regeneration.

INTRODUCTION

Somatic embryos were first produced in tissue culture in the classic experiments by Steward et al. (1958). Carrot cells grown in suspension cultures in the presence of 2,4-dichlorophenoxyacetic acid (2,4-D) differentiate into embryogenic “proembryos” or “proembryonic masses” (PEMs) (Halperin, 1964). Upon removal of 2,4-D, PEMs differentiate into somatic embryos (de Vries et al., 1988).

The first somatic embryos in maize tissue culture were produced by Green and Phillips (1975). Maize cell lines derived from sources such as immature embryos are heterogeneous for cells with different embryogenic competence. Friable callus (type II) in these lines was found to be more embryogenic (Armstrong and Green, 1985). The generation of embryogenic callus is genotype-specific in many plant species, including maize, and quite often the most desirable and/or economically important are recalcitrant to regeneration. Because of that it has been important to recognize embryogenic cells or tissues and to identify markers for them.

Some of the early attempts to find indicators for embryogenic competence relied on biochemical markers. Isozyme differences between embryogenic and non-embryogenic cultures were demonstrated for glutamate dehydrogenase isoperoxidase, esterase and malate dehydrogenase isozymes (Fransz et al., 1989; Rao et al., 1990). Schmidt et al. (1997) employed differential display to identify genes differentially expressed in embryogenic carrot cells. One such gene encoded a leucine-repeat receptor (LRR) protein and was dubbed as somatic embryogenesis receptor kinase (SERK).

SERK is a leucine-rich repeat (LRRs) protein kinase (Schmidt et al., 1997). In Arabidopsis, five members of SERK family have been identified (AtSERK1-5). AtSERK1 is expressed during somatic embryogenesis, and the embryogenic competence of callus derived from seedlings overexpressing AtSERK1 (driven by the CaMV35S promoter) is elevated 3-4 fold when compared with the wild-type callus (Hecht et al., 2001). At least two related genes have been identified in maize, ZmSERK1 and 2 (Baudino et al., 2001). ZmSERK1 is preferentially expressed in reproductive tissues with the strongest expression in microspores, while ZmSERK2 expression is fairly uniform in all tissues investigated. Both genes are expressed in callus cultures whether they are embryogenic or not, which suggests that the genes might not be good markers for embryogenesis in maize (Baudino et al., 2001).

LEAFY COTYLEDON 1 (LEC1) is an example of an Arabidopsis transcription factor that is upregulated during embryogenesis and that promotes somatic embryogenesis when ectopically expressed in vegetative cells (Lotan et al., 1998). *LEC1* is expressed throughout early stage zygotic embryos, and in later stages its expression is confined to the embryo periphery, outside the procambrial zone (Lotan et al., 1998). Loss-of-function mutations in *LEC1* affect many aspects of embryogenesis. *lec1* mutants prematurely germinate and are desiccation intolerant. Mutant embryos produce cotyledons with characteristics of later postgerminative development (Meinke, 1992; Meinke et al., 1994; West et al., 1994). *lec1* affects the expression of certain maturation phase genes including those encoding storage proteins (Meinke et al., 1994; West et al., 1994; Percy et al., 1997; Vicent et al., 2000).

Maize genes with sequences similar to *LEC1* have been identified, and the expression pattern of ZmLec1 has been profiled during somatic embryo genes (Zhang et al., 2002). The expression of ZmLec1 during maize somatic embryogenesis is similar to *LEC1* during Arabidopsis zygotic embryogenesis with general expression through the embryo up to the globular stage of development (Zhang et al., 2002). Lowe et al. (2000) reported that ectopic expression of the ZmLec1 greatly improved the recovery of transformants in maize tissue culture.

With the advent of microarray technologies, one major study has been conducted profiling gene expression patterns during somatic embryogenesis in soybean (Thibaud-Nissen et al., 2003). Soybean somatic embryos are formed on the adaxial surface of immature cotyledons placed on high levels of 2,4-D. Thibaud-Nissen et al. (2003) compared gene expression during embryo development on the adaxial side of cotyledons to callus formation on the abaxial side. Their results suggest that cotyledons dedifferentiate for two weeks prior to the development of somatic embryos. Genes involved in oxidative stress responses and cell division change in expression on the adaxial side of the cotyledons indicating that events involving cell proliferation and cell death are played out during somatic embryo development (Thibaud-Nissen et al., 2003).

In this study, we compared gene expression patterns between embryogenic and total callus and profiled gene expression patterns during somatic embryo maturation and germination in a regeneration-proficient maize line, Hi II. We found significant gene expression changes during embryo development, but none in the comparison between embryogenic and total callus. The genes regulated during somatic embryogenesis may have use as markers for genetically improving the regeneration of more recalcitrant lines.

RESULTS

Somatic embryogenesis and expression profiling

During the growth of embryo-derived Hi II maize callus on 2,4-D-containing medium, some of the callus differentiates into a form that is, itself, highly embryogenic. Embryogenic clumps of Hi II maize callus are characterized by cells that are spindle-shaped or polar, while the cells in non-embryogenic material are diametric or amorphous (Fig. 1A). For routine maize regeneration, embryogenic callus composed almost entirely of polar cells is selected and transferred onto Regeneration Medium I (-2,4-D, 6% sucrose) that promotes the development of somatic embryos (Fig. 1B). After 7 days on this medium, tissue destined to become somatic embryos appears milky or less translucent (Fig. 1B). Embryo development and maturation continues for 21 days, and when mature somatic embryos are transferred to light on Regeneration Medium II (-2,4-D, 3% sucrose), the embryos germinate (Fig. 1B).

Two independent experiments were conducted to examine gene expression patterns during somatic embryogenesis in maize. The first was designed to determine whether gene expression differences could be detected between selected embryogenic callus and total, unselected callus growing on 2,4-D-containing (N6E) medium. The second experiment was aimed at profiling gene expression patterns during somatic embryo maturation and germination – with the aim of understanding the gene expression events underlying somatic embryogenesis and possibly identifying developmental markers.

Six independent callus lines (A-F) were sampled and two lines were pooled (creating 3 line pools) to obtain sufficient amounts of RNA for microarray analysis (without amplification). Gene expression patterns were profiled using maize cDNA microarrays. Thirty-six microarray chips were each spotted with 12,060 maize cDNAs. Six arrays were used to compare embryogenic and total callus (Fig. 2A), and the remaining 30 were used for the time course analysis (Fig. 2B). The chips were hybridized with cy3 and cy5 cDNAs using a loop design strategy (Dobbin and Simon, 2002) in which samples were compared to each other and not to a single reference, such as a zero time sample (Fig. 2A and B). In the time course experiment, the strategy allows for more repetition of time points with the same number of chips. In each line pool, each time point is sampled 4 times – twice with a cy3 labeled probe and twice with a cy5 labeled probe. Thus, across all three line pools, a time point sample is repeated 12 times. Such a scheme permits analysis of both time and line pool variation.

Gene expression difference

Hi II callus was used for these studies because it is highly embryogenic. We estimated that about half of the callus is embryogenic with cellular morphologies described above. We compared the gene expression levels in selected embryogenic callus to total callus, and found that none of the 12060 genes in this study showed significant differences (at the $\alpha=0.05$ level, considering multiple comparisons). Given that about half of the callus appeared embryogenic, one might expect, at best, a two-fold difference between total and embryogenic callus for a gene that was upregulated in

embryogenic callus. However, if a gene is downregulated in embryogenic callus one might expect to see larger differences because embryonic callus appears largely free of nonembryogenic material. However, no significant differences were observed whether one considered genes up or down regulated.

The situation was very different during somatic embryo development. Somewhat more than a 1000 genes out of 12,060 in the study showed significant time variation (at the $\alpha=0.05$ level, considering multiple comparisons) during somatic embryo maturation and germination. During that time increasing numbers of genes were upregulated by 2-fold or more during maturation and germination (Fig. 3). Likewise, an increasing number of genes are down regulated 2-fold or more during maturation, but that trend reversed itself during embryo germination.

The overall trends are made up of individual genes with varied expression patterns, and patterns for ~1000 genes with significant time variation were organized into groups. The patterns clustered into 12 groups based on model-based clustering of the sequential expression ratios. The identities and membership for each group were iteratively estimated until the most likely arrangement of groups was found. This was repeated for a set number of groups from 5 to 20, and it was estimated that 12 was the most likely number of groups. Each gene was assigned to one of 19 functional categories. The largest category was for genes with unknown function and usually the second largest was for genes involved in primary metabolism. The functional distribution for six groups of genes in which a category, other than unknown or primary metabolism, emerged or dominated the distributions are shown.

The first two pattern groups were characteristic of genes downregulated during embryo maturation, which then recover during germination (Fig 4A and B). The first group had a larger number than other groups of genes encoding nuclear proteins, such a gene encoding proliferating cell nuclear antigen, and histone genes (Fig. 4A). The second pattern group included a large category of genes encoding proteins involved in translation, such as 40S ribosomal protein S15 (Fig. 4B). The decline in expression of these genes during maturation likely indicates a reduction in cell proliferation and growth during maturation. The recovery in expression of these genes later on accompanies the growth spurt during germination. In another group of genes, dominated by a category of glucosidases, nucleases and proteases, expression rose during maturation but then dropped off during germination (Fig. 4C).

Early in germination, a group of stress-related genes, such as a gene encoding a heat shock protein, are transiently upregulated (Fig 4D). Other genes that are upregulated during germination (with expression levels generally higher at both time points) include a large group encoding channel proteins or membrane transporters, such as a gene encoding a water channel (Fig. 4E). Finally, genes encoding photosynthetic and other chloroplast components, such as a gene encoding a chlorophyll a/b binding protein, are upregulated as the shoot begins to green (Fig. 4F). Thus, from a gene expression perspective, germination first involves the activation of expression of stress-related and transporter/channel-encoding genes followed by the upregulation of photosynthetic/chloroplast genes.

We also examined the expression patterns of genes typically associated with zygotic embryogenesis and germination to determine if they would be good markers for somatic embryogenesis. Some showed expected expression patterns – others did not. Of

those that did, a gene encoding an alpha zein, an embryo storage protein, was upregulated as expected during embryo maturation, and then its expression levels fell during germination (Fig. 5A). Another gene encoding an embryo-specific Ca⁺⁺-binding protein (ATS1) showed a somewhat similar pattern of expression (Fig. 5B). A gene encoding a late embryogenesis abundant protein was expressed at increasing levels during maturation, but transcript levels continued to rise during germination (Fig. 5C). Finally, a gene encoding a protein related to germins was upregulated, as expected during germination (Fig. 5D). Surprisingly, most other genes encoding zeins and late embryogenesis abundant proteins did not show significant time variations in expression during embryo maturation and germination.

Line variation

We also looked for genes with significant variation across time points and either with considerable or little variation across lines. Genes with expression patterns that vary considerably across lines might be useful if the variation correlates with embryonic competence or with other traits that vary from line to line. For example, two genes, one encoding a lipid transfer protein and the other a bZIP family transcription factor, show significant time variation in one line pool, but not in the other two line pools (Fig 6, upper panel). Other genes, such as one encoding a putative disease response protein and photosystem I assembly protein, show significant variation across time points but little variation between lines (Fig. 6, lower panel). Such genes might be useful developmental markers for cultures that show line variation.

DISCUSSION

The regeneration of maize in tissue culture is important for the production of transgenic maize and for crop improvement using genetic engineering approaches. Reports of fertile maize plants regenerated from protoplasts (Prioli and Sondahl, 1989; Shillito et al., 1989) were closely followed by the production of transgenic, fertile maize from transformed suspension cell cultures of the hybrid A188 x B73 line (Gordon-Kamm et al., 1990). While other lines have since been transformed, most elite lines remain inaccessible to improvement using standard transformation techniques because they are recalcitrant or very difficult to regenerate.

Hi-II is frequently used in transformation and regeneration studies because it is a regeneration-proficient line. We looked for gene expression differences between embryogenic and total callus in this line and found none that were statistically significant. There may be several reasons for this. 1) Hi II is a highly regeneration-proficient line, and there may be little difference between embryogenic and non-embryogenic Hi II callus, 2) the genes that significantly differ in expression between embryogenic and total callus may not present on the cDNA chips, or 3) the difference between embryogenic and total callus may not be reflected by differences in the transcriptome. It might be better to search for gene expression differences between embryogenic and nonembryogenic callus by analyzing less regeneration-proficient lines, i.e., where smaller amounts of callus are embryogenic.

In contrast, gene expression patterns change extensively during somatic embryo development, maturation and germination. Following transfer to medium lacking 2,4-D and throughout embryo maturation, there is a progressive decline in the expression of genes involved in cell proliferation and growth, such as genes encoding histones and ribosomal proteins. Strikingly, the expression levels of these genes recover at the onset of germination. The changes in expression may reflect a slowdown in cell proliferation and growth during somatic embryogenesis and a resurgence in expression of these genes at germination. During maturation, expression rises for a group of genes encoding hydrolytic enzymes, such as nucleases, glucosidases and proteases, suggesting a breakdown, and perhaps a retooling, of cell components during somatic embryo development. Unlike, zygotic embryogenesis (Lending and Larkins, 1989), we did not observe the large-scale upregulation in expression of storage protein genes, although one of the γ -zein genes is upregulated. Whether maize Hi II somatic embryos accumulate less storage proteins than their zygotic counterparts is a matter that deserves exploration. In any case, only a few storage protein genes for γ -zein and a caleosin, a lipid body protein (Naested et al., 2000), appear as good markers of somatic embryo maturation in the Hi II line.

Some stress response genes, such heat shock genes are upregulated at the onset of germination. Their upregulation may be a normal developmental event or may simply be a response to the transfer of tissue to new culture medium. Of interest is a group of genes upregulated at this time that encode various transporters and membrane channels. Finally, as expected, germination and shoot greening is accompanied by the activation in expression of a myriad of genes encoding photosynthetic and chloroplast components.

Earlier attempts to identify markers for embryogenic competence involved translating RNA from cultured carrot cells and somatic embryos and comparing the translation products by electrophoresis on 2D gels. With the exception of two polypeptides, called E1 and E2, Sung and Okimoto (1981) found few differences, which led Choi et al. (1987) to suggest that the similarities in gene expression patterns may reflect the fact that proembryonic masses (PEMs) in cultured cells may already be “committed to the embryogenic program.” Wilde et al. (1988) also used 2D gel electrophoresis of translation products to arrive at similar conclusions that “many of the molecular processes of embryogenesis are already established in PEMs in the presence of auxin.”

Monoclonal antibodies directed at cell-wall arabinogalactan-proteins (AGPs) have also been used to identify markers associated with somatic embryogenesis. Extracts containing seed AGPs enhance the formation of protoplast-derived somatic embryos (Kreuger and Van, 1993; Kreuger and van Holst, 1996; van Hengel et al., 2001) and certain AGPs, such as those bearing the epitope recognized by the monoclonal antibody Jim8, are produced in embryogenic carrot cultures (Pennell et al., 1992). Both AGPs and EP3 endochitinases are released from carrot cells into the culture medium, and conditioned medium from carrot cultures enhances somatic embryo formation (van Hengel et al., 2001). These and other findings support the concept that activation of AGPs by hydrolytic enzymes such as endochitinase promotes somatic embryogenesis (Van Hengel et al., 2002). Gene expression markers have been used more widely in recent years to characterize embryogenic lines and to describe the embryo development process. Chugh and Khurana (2002) reviewed the state of knowledge on gene expression

in somatic embryogenesis in higher plants prior to the extensive use of global gene profiling technologies.

METHODS

Materials and tissue culture methods

Somatic embryos were generated in embryogenic callus lines developed independently from immature Hi II embryo explants using protocols described at the Plant Transformation Facility website for the production of transgenic corn (<http://www.agron.iastate.edu/ptf/web/system.htm>).

Briefly, green house grown ears from the Hi-II hybrid line (Armstrong et al., 1991) were dehusked and surface sterilized for 20 min (50% commercial bleach in water plus 1 drop/L of Tween 20) then rinsed three times with sterilized water. Immature zygotic embryos were excised and cultured embryo-axis side down (scutellum side up) on N6E media (N6 salts and vitamins (Chu, 1975), 2 mg/L 2,4-D, 100 mg/L myo-inositol, 2.76 g/L proline, 30 g/L sucrose, 100 mg/L casein hydrolysate, 2.5 g/L gelrite, pH 5.8 after Songstad et al. (1996)). Silver nitrate (25 μ M) was added after autoclaving. The plates were wrapped with vent tape and incubated at 28°C in the dark for 2 weeks.

Friable type II callus was bulked up from 6 separate embryo explants over 8 weeks by subculturing every two weeks on the same medium. Callus was then subjected to regeneration conditions by transferring about 15 small pieces (approximately 4 mm) of embryogenic callus to Regeneration Medium I (MS salts and vitamins (Murashige and Skoog, 1962), 100 mg/L myo-inositol, 60 g/L sucrose, 3 g/L gelrite, pH 5.8) and incubating for 3 weeks at 25°C in the dark. Petri-plates (100x25 mm) were wrapped with vent tape. After 3 weeks, matured somatic embryos were identified using a light microscope, transferred to Regeneration Medium II (as for Regeneration Medium I but with 3% sucrose), and placed in the light (~80 μ E/m²/s) for germination. Plantlets sprouted leaves and roots on this medium.

RNA extraction and microarray analysis

RNA was extracted using a TRIzol method modified from Chomzynski and Sacchi (1987) and described in TAIR protocols (<http://www.arabidopsis.org/servlets/TairObject?type=protocol&id=501683718>). In this procedure 1 g of maize callus tissue was ground with liquid nitrogen in a mortar and pestle. The ground powder was mixed with 15 ml TRIzol reagent (Life Technologies) and incubated at 60° for 5 min. The mixture was centrifuged at 12,000 g at 4° for 10 min and to the supernatant was added 3 ml of chloroform. The mixture was vortexed for 15 sec and allowed to sit at room temperature for 2-3 min. The mixture was centrifuged at 10,000 g at 4° for 15 min, and RNA was precipitated from the upper phase by adding 1/2 volume each of isopropanol and 0.8M sodium citrate/1.2M NaCl. The mixture was allowed to sit at room temperature for 10 min and centrifuged at 10,000g at 4° for 10 min. The pellet was washed with 70% EtOH, vortexed briefly and centrifuged again at 10,000g at 4° for 10 min. The pellet was air dried for 5 min and dissolved in 250 μ l of

DEPC-treated water. The RNA sample was centrifuged in a microcentrifuge for 5 min at room temperature and the insoluble pellet discarded. The RNA sample was cleaned up by passing through a RNeasy column (Qiagen) according to manufacturer's instructions.

cDNA was synthesized and labeled according to procedures described by Hegde et al. (2000). The procedure is an indirect labeling method in which first-strand cDNA is synthesized in the presence of amino-allyl labeled dUTP, and then NHS-esters of the appropriate cyanine fluor are covalently coupled to the substituted cDNA strand. The reaction mix for first strand synthesis consisted of Superscript II buffer (Life Technologies), 10 mM DTT, 5 mM dATP, dCTP and dGTP, 3 mM dTTP, 2mM aminoallyl-dUTP, 0.3 mg/ml oligo dT (Invitrogen) and 400 units of Superscript II reverse transcriptase (Invitrogen). The reaction was incubated overnight at 42° followed by base hydrolysis of RNA in 200 mM NaOH, 10 mM EDTA and incubation for 15 min at 65°.

The aminoallyl-label cDNA was purified using a modified QIAquick (Qiagen) PCR purification procedure. The cDNA reaction was mixed with 5X volume of 5 mM potassium phosphate (PB, pH 8.0) and transferred to a QIAquick column. The column was centrifuged for 1 min in a collection tube at 14,000 rpm in a microcentrifuge, washed twice with 750 μ l of 5 mM PB (pH 8.0) and 80% EtOH and centrifuged each time. 30 μ l of 4 mM potassium phosphate (pH 8.5) were added to the column, incubated for 1 min, and RNA was eluted by centrifugation at 14,000 rpm for 1 min. The elution step was repeated once more with another 30 μ l of 4 mM PB (pH 8.5). The sample was dried in a SpeedVac.

The aminoallyl-label cDNA was coupled to the Cy dyes by dissolving the dried cDNA in 4.5 μ l of freshly prepared 0.1 M sodium carbonate buffer (pH 9.0). Cy3- or Cy5-esters (AmershamPharmacia) were dissolved in 73 μ l DMSO, and 4.5 μ l of the appropriate NHS-Cy were added to the labeled cDNA. The mixture was incubated in the dark at room temperature for 1 hr. Following the reaction, uncoupled dye was removed using a QIAquick PCR purification kit (Qiagen). 35 μ l of sodium acetate buffer (pH 5.2) and 250 μ l 5 mM PB (pH 8.0) were added to the reaction and transferred to a QIAquick column. The dye-coupled cDNA was eluted with 2 aliquots of 30 μ l of elution buffer (Qiagen) and dried in a SpeedVac.

Maize cDNA chips were prepared in the Iowa State University microarray facility by spotting aminosilane coated slides with a Cartesian PixSys 5500 Arrayer. The maize chips contained over 12,000 spotted cDNA inserts obtained from the NSF Plant Genome EST projects led by Virginia Walbot (Stanford) and Patrick Schnable (Iowa State). The cDNAs included in the chip (Gen II, Version B) are listed at <http://www.plantgenomics.iastate.edu/maizechip/>. The slides to be hybridized were placed in Coplin jar with prehybridization buffer (5XSSC, 0.1% SDS and 1% bovine serum albumin) and incubated at 42° for 45 min. The slides were washed 5X by dipping in MilliQ water (Millipore) at room temperature, followed by dipping in isopropanol and air drying.

For hybridization, each labeled probe was resuspended in 19 μ l of hybridization buffer (50% formamide, 5X SSC and 0.1% SDS) to which was added 1 μ l of 20 μ g / μ l human COT1 DNA (LifeTechnologies) and 1 μ l of 20 μ g / μ l poly A DNA (Invitrogen) to block non-specific hybridization. The sample was heated at 95° for 3 min to denature the probe and centrifuged at 13,000 rpm for 1 min in a microcentrifuge at room temperature. The probe was applied to a microarray slide, covered with a 22 X 60 mm glass coverslip

and placed in a sealed hybridization chamber with 20 μ l of water added to the chamber at the end of the slide. The chamber was incubated overnight at 42°. Following incubation the slide was carefully removed from the chamber and placed in a staining dish with wash buffer containing 1X SSC and 0.2% SDS at 42°. The coverslip was gently removed, and the slide was agitated in the wash buffer for 4 min. The slide was further washed with 0.1X SSC and 0.2% SDS at room temperature for 4 min and then in 0.1X SSC for another 4 min. The slides were allowed to air dry.

Microarray data analysis

Imagene software (Biodiscovery) was used to read image files from the General Scanning ScanArray 5000 scanner. Imagene employs a fixed circle method to segment spots by positioning a circle of fixed diameter for the greatest difference between pixels inside and outside the spot. The mean signal pixel intensity computed from approximately 120 pixel intensity measurements was obtained for each spot. Background was selected using a concentric-circle-band method in which a second circle is placed around the first and pixels within the halo are designated as background. The intensity of each background pixel was recorded, and the median background pixel intensity was used to estimate the background effect. (The median was used rather than the mean because some pixels designated as background may actually have fluorescent probe in them. These pixels, therefore, have much higher intensity values than the pixels from empty regions of the slide.)

In the time course study, all of the slides from each line pool were prepared in order and read in the same batch. This was done for job scheduling reasons and is not recommended for an experimental setup. It would have been better to randomize the slides with respect to experimental order, because time effects (learning, machine calibration) may be present and confound with treatment effects. Substantially more effort to randomize preparation between line pools would allow more precise estimates of the line variation.

Different laser and sensor settings were used to scan each slide to adjust the dynamic range of the scanner to the overall fluorescence intensity of the slide. Higher laser settings create more fluorescence and higher photomultiplier settings amplify the light signal. However, low range settings miss spots with low signals and in high range settings, high intensity signals are saturated. (There is an upper limit of 65535 to the measurement of fluorescence so that signal from spots that are brighter will be truncated.) Preliminary work indicates that a significant reduction in the variability of expression estimates can be obtained when analyzing the data from multiple readings with the appropriate statistical model. However, for speed and simplicity, we included only one reading for each slide by choosing the one with the highest median intensity among readings with the fewest intensities reported as 65535 (the maximum).

Data normalization

We assume that there is a systematic bias in the gene expression measurements between the two dyes. For gene j which is not differentially expressed, we do not expect $R_j = G_j$ on average. Instead, we expect $R_j = k_j G_j$ for some k_j . The total signal intensity for

each gene on a single slide is the sum of the fluorescent intensity in both R and G channels. An alternative measure of intensity defined as $A_j = \log(\sqrt{R_j G_j})$ can be plotted against the log ratio, $M_j = \log(R_j/G_j)$, because both measures are defined on the log scale. The dye bias has been shown to be dependent on the intensity level (Yang et al., 2002). Additionally, each print tip has characteristics, which can result in spots printed by the same print tip to be correlated. As a consequence, spots in the same print tip group (metarow and metacolumn combination) appear in spatially similar groups within the slide. Thus print tip groups may account for bias due to print tips and act as a surrogate for spatial effects (Yang et al., 2002). The effects of intensity dependent, print tip group related dye bias should be removed in normalization.

Print tip group-intensity dependent normalization assumes that the normalizing constant is a function of intensity for each print tip group i , $k_j = f_i(R_j + G_j)$. Because it is assumed that only a small proportion of genes in our experiment are differentially expressed, a robust estimator of $\log(k_j)$ is the loess curve of M against A using only the middle range of the data in each print tip group (Yang et al., 2002).

Print tip group-intensity dependent normalization has the following characteristics:

?? I functions of intensity per slide where I is the number of print tip groups; each gene takes its own value within a group.

?? The factor k_j is interpreted as the dye bias against Cy5 at intensity $R_j + G_j$.

?? Accounts for intensity dependent effects.

?? Includes some spatial effects.

?? Does not rescale the data to have a similar variability.

Estimation of treatment effects

Analysis of variance (ANOVA) was used to determine whether several groups (or treatments) have equal mean expression. Under the null hypothesis, all groups have a common mean and standard deviation and ANOVA is used to test whether any of the groups violate that assumption. In the design of the time course experiment, each group corresponds to a different time point in embryonic development, and there were 12 observations for each gene at each time point. Therefore, we can, in principle, conduct a test of the null hypothesis for each of the genes, to investigate whether mean expression varies across treatments (or time points).

Because there are a large number of elements (12060) in the arrays, conducting so many hypothesis tests would likely result in a large proportion of false positive conclusions. A false positive occurs when we erroneously conclude that a gene exhibits different expression levels at different time points. In experiments such as this, it is very important to control the experiment-wise error rate at a predetermined level by carrying out an adjustment that accounts for the erosion in confidence levels in multiple comparisons. Most of the approaches that can be implemented to account for multiple comparisons rely on the p-values generated by the ANOVA procedure, P_j , for each gene. The most conservative approach is the Bonferroni adjustment. The Bonferroni adjustment controls the experiment-wise error rate by assuming that expression of all genes remains constant over time (no treatment effect). The j th gene is considered significantly differentially expressed over time if $P_j < \alpha/t$ where t is the number of tests being performed and $\alpha > 0$ is the predetermined target error rate (Ishwaran and Rao 2003).

A less conservative adjustment was proposed by Benjamini and Hochberg (1995). This adjustment only assumes that the genes which have large p-values have no differential expression over time. The j th gene is considered significantly differentially expressed over time if $P_j < P_{(k)}$ where $P_{(k)}$ is the k^{th} ordered p-value for the genes, $k = \max\{j: P_{(j)} \leq j \cdot \alpha / t\}$, t is the number of tests being performed, and $\alpha > 0$ is the predetermined target error rate (Ishwaran and Rao, 2003).

LITERATURE CITED

- Armstrong CL, Green CE** (1985) Establishment and maintenance of friable embryogenic maize callus and the involvement of L-proline. *Planta* **164**: 207-214
- Armstrong CL, Green CE, Phillips RL** (1991) Development and availability of germplasm with high Type II culture formation response. *Maize Gen Coop Newsletter* **65**: 92-93
- Baudino S, Hansen S, Brettschneider R, Hecht VF, Dresselhaus T, Lorz H, Dumas C, Rogowsky PM** (2001) Molecular characterisation of two novel maize LRR receptor-like kinases, which belong to the SERK gene family. *Planta* **213**: 1-10
- Benjamini Y, Hochberg Y** (1995) Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc Ser B* **57**: 289-300
- Choi JH, Liu LS, Borkird C, Sung ZR** (1987) Cloning of genes developmentally regulated during plant embryogenesis. *Proc. Natl. Acad. Sci. USA* **84**: 1906-1910
- Chomczynski P, Sacchi N** (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* **162**: 156-159
- Chu CC** (1975) Establishment of an efficient medium for anther culture of rice through comparative experiments on the nitrogen sources. *Sci. Sinica* **18**: 659-668
- Chugh A, Khurana P** (2002) Gene expression during somatic embryogenesis -- recent advances. *Current Science* **83**: 715-730
- de Vries SC, Booij H, Meyerink P, Huisman G, Wilde HD, Thomas TL, van Kammen A** (1988) Acquisition of embryonic potential in carrot cell-suspension cultures. *Planta* **176**: 196-204
- Dobbin K, Simon R** (2002) Comparison of microarray designs for class comparison and class discovery. *Bioinformatics* **18**: 1438-1445
- Franz PF, de Ruijter NCA, Schel JHN** (1989) Isozymes as biochemical and cytogenetic markers in embryogenic callus cultures of maize (*Zea mays* L.). *Plant Cell Rep.* **8**: 67-70
- Gordon-Kamm W, Spencer T, Mangano ML, Adams T, Daines R, Start W, O'Brien J, Chambers S, Adams Jr. W, Willetts N, Rice T, Mackey C, Krueger R, Kausch A, Lemaux PG** (1990) Transformation of maize cells and regeneration of fertile transgenic plants. *Plant Cell* **2**: 603-618
- Green CE, Phillips RL** (1975) Plant regeneration from tissue cultures of maize. *Crop Sci.* **15**: 417-420
- Halperin W** (1964) Morphogenetic studies with partially synchronized cultures of carrot embryos. *Science* **146**: 408-410

- Hecht V, Vielle-Calzada JP, Hartog MV, Schmidt ED, Boutilier K, Grossniklaus U, de Vries SC** (2001) The Arabidopsis *SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1* gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture. *Plant Physiol* **127**: 803-816
- Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R, Hughes JE, Snesrud E, Lee N, Quackenbush J** (2000) A concise guide to cDNA microarray analysis. *Biotechniques* **29**: 548-550, 552-544, 556 passim
- Ishwaran H, Rao JS** (2003) Detecting differentially expressed genes in microarrays using Bayesian model selection. *J Amer Stat Assoc* **98**: 438-455
- Kreuger M, Van HGJ** (1993) Arabinogalactan proteins are essential in somatic embryogenesis of *Daucus carota* l. *Planta* **189**: 243-248
- Kreuger M, van Holst GJ** (1996) Arabinogalactan proteins and plant differentiation. *Plant Mol Biol* **30**: 1077-1086
- Lending CR, Larkins BA** (1989) Changes in the zein composition of protein bodies during maize endosperm development. *Plant Cell* **1**: 1011-1023
- Lotan T, Ohto MA, Yee KM, West MAL, Lo R, Kwong RW, Yamagishi K, Fischer R, L., Goldberg R, B., Harada J, J.** (1998) Arabidopsis *LEAFY COTYLEDON1* is sufficient to induce embryo development in vegetative cells. *Cell* **93**: 1195-1205
- Lowe K, Abbit S, Glassman K, Gregory C, Hoerster G, Rasco-Gaunt S, Sun X, Lazzeri P, Gordon-Kamm W** (2000) Use of maize *Lec1* to improve transformation. *In Vitro Cell. Dev. Biol* **36**: W15
- Meinke DW** (1992) A homeotic mutant of *Arabidopsis thaliana* with leafy cotyledons. *Science* **258**: 1647-1650
- Meinke DW, Franzmann LH, Nickle TC, Yeung EC** (1994) Leafy cotyledon mutants of Arabidopsis. *Plant Cell* **6**: 1049-1064
- Murashige T, Skoog F** (1962) A revised medium for rapid growth and bioassay with tobacco tissue cultures. *Physiol. Plant.* **15**: 437-497
- Naested H, Frandsen GI, Jauh GY, Hernandez-Pinzon I, Nielsen HB, Murphy DJ, Rogers JC, Mundy J** (2000) Caleosins: Ca²⁺-binding proteins associated with lipid bodies. *Plant Mol Biol* **44**: 463-476
- Parcy F, Valon C, Kohara A, Misera S, Giraudat J** (1997) The *ABSCISIC ACID-INSENSITIVE3*, *FUSCA3*, and *LEAFY COTYLEDON1* loci act in concert to control multiple aspects of Arabidopsis seed development. *Plant Cell* **9**: 1265-1277
- Pennell RI, Janniche L, Scofield GN, Booij H, de Vries SC, Roberts K** (1992) Identification of a transitional cell state in the developmental pathway to carrot somatic embryogenesis. *J Cell Biol* **119**: 1371-1380
- Prioli LM, Sondahl MR** (1989) Plant regeneration and recovery of fertile plants from protoplasts of maize (*Zea mays* L.). *Bio/Technology* **7**: 589-594
- Rao KV, Suprasanna P, Reddy GM** (1990) Biochemical changes in embryogenic and non-embryogenic calli of *Zea mays* L. *Plant Sci.* **66**: 127-130
- Schmidt EDL, Guzzo F, Toonen MAJ, de Vries SC** (1997) A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. *Devel.* **124**: 2049-2062

- Shillito RD, Carswell GK, Johnson CM, Di Maio JJ, Harms CT** (1989) Regeneration of fertile plants from protoplasts of elite inbred maize. *Bio/Technology* **7**: 581-587
- Songstad D, Armstrong CL, Petersen WL, Hairston B, Hinchee MAW** (1996) Production of transgenic maize plants and progeny by bombardment of Hi II immature embryos. *In Vitro Cell Dev Bio-Plant* **32**: 179-183
- Steward D, Mapes MO, Mears JS** (1958) Growth and organized development of cultured cells. II. Organization in cultures grown from freely suspended cells. *Am. J. Bot.* **45**: 705-708
- Sung ZR, Okimoto R** (1981) Embryonic proteins in somatic embryos of carrot. *Proc. Natl. Acad. Sci. U.S.A.* **78**: 3683-3687
- Thibaud-Nissen F, Shealy RT, Khanna A, Vodkin LO** (2003) Clustering of microarray data reveals transcript patterns associated with somatic embryogenesis in soybean. *Plant Physiol* **132**: 118-136
- van Hengel AJ, Tadesse Z, Immerzeel P, Schols H, van Kammen A, de Vries SC** (2001) N-acetylglucosamine and glucosamine-containing arabinogalactan proteins control somatic embryogenesis. *Plant Physiol* **125**: 1880-1890
- Van Hengel AJ, Van Kammen A, De Vries SC** (2002) A relationship between seed development, Arabinogalactan-proteins (AGPs) and the AGP mediated promotion of somatic embryogenesis. *Physiol Plant* **114**: 637-644
- Vicient CM, Bies-Etheve N, Delseny M** (2000) Changes in gene expression in the *leafy cotyledon1* (*lec1*) and *fusca3* (*fus3*) mutants of *Arabidopsis thaliana* L. *J Exp Bot* **51**: 995-1003
- West MALW, Yee KM, Danao J, Zimmerman JL, Fischer RL, Goldberg RB, Harada JJ** (1994) *LEAFY COTYLEDON1* is an essential regulator of late embryogenesis and cotyledon identity in *Arabidopsis*. *Plant Cell* **6**: 1731-1745
- Wilde HD, Nelson WS, Booiij H, de Vries SC, Thomas TL** (1988) Gene expression programs in embryonic and non embryonic carrot cultures. *Planta* **176**: 205-211
- Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, Speed TP** (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucl Acids Res* **30**: e15
- Zhang S, Wong L, Meng L, Lemaux PG** (2002) Similarity of expression patterns of knotted1 and ZmLEC1 during somatic and zygotic embryogenesis in maize (*Zea mays* L.). *Planta* **215**: 191-194

FIGURE CAPTIONS

Figure 1. Somatic embryo development in maize Hi II tissue culture. (A) Examples of total and embryogenic callus growing on N6E medium. Arrow points out polarized cell characteristic of cells in embryogenic callus. (B) Time course of somatic embryo development, maturation and germination. Somatic embryo development and maturation was initiated by transferring embryogenic callus onto Regeneration Medium I (-2,4-D, 6% sucrose). Embryos were germinated by transfer to the light on Regeneration Medium

II (-2,4-D, 3% sucrose). Samples were taken at time points as indicated during embryo maturation and germination for profiling gene expression patterns. Bars = 1 mm.

Figure 2. Loop design for microarray hybridization experiments. Six independent culture lines (A-F) were initiated and the lines were pooled into 2 lines per pool to obtain enough RNA in each pool for the microarray analysis. Each rectangle represents 1 chip. (A) Comparison between embryogenic and total callus. Sample source (embryogenic, E, or total, T) #, line pool # (1, 2 or 3) and the probe dye type (either cy3 or cy5) are indicated for each chip. (B) Time course analysis. Time point and probe dye type are indicated for each chip.

Figure 3. Trends in gene expression during somatic embryo maturation and germination in maize. Number of genes out 1026 genes which show significant time variation, that are either up or downregulated more than 2-, 3- or 4-fold during embryo development. Stipled bar represents the period of somatic embryo development and maturation. Unshaded bar is the time of germination.

Figure 4. Genes with different expression profiles. Expression profiles of the 1026 genes with the greatest variation across time points cluster into 12 pattern groups. Examples of genes from six different pattern groups are shown here. The distribution of function of genes in the pattern group are shown. Genes were categorized into 19 functional groups. Means and standard errors (SEs) for 12 repeats at each time point. Period of embryo maturation (stippled bar), embryo germination (unshaded bar).

Figure 6. Expression profiles for four genes belonging to classes of genes expressed during zygotic embryo development or germination.

Figure 5. Line variation in gene expression. Examples of genes with significant variation across time points and that (upper panel) show significant line pool variation or (lower panel) show little line pool variation. Period of embryo maturation (stippled bar), embryo germination (unshaded bar).

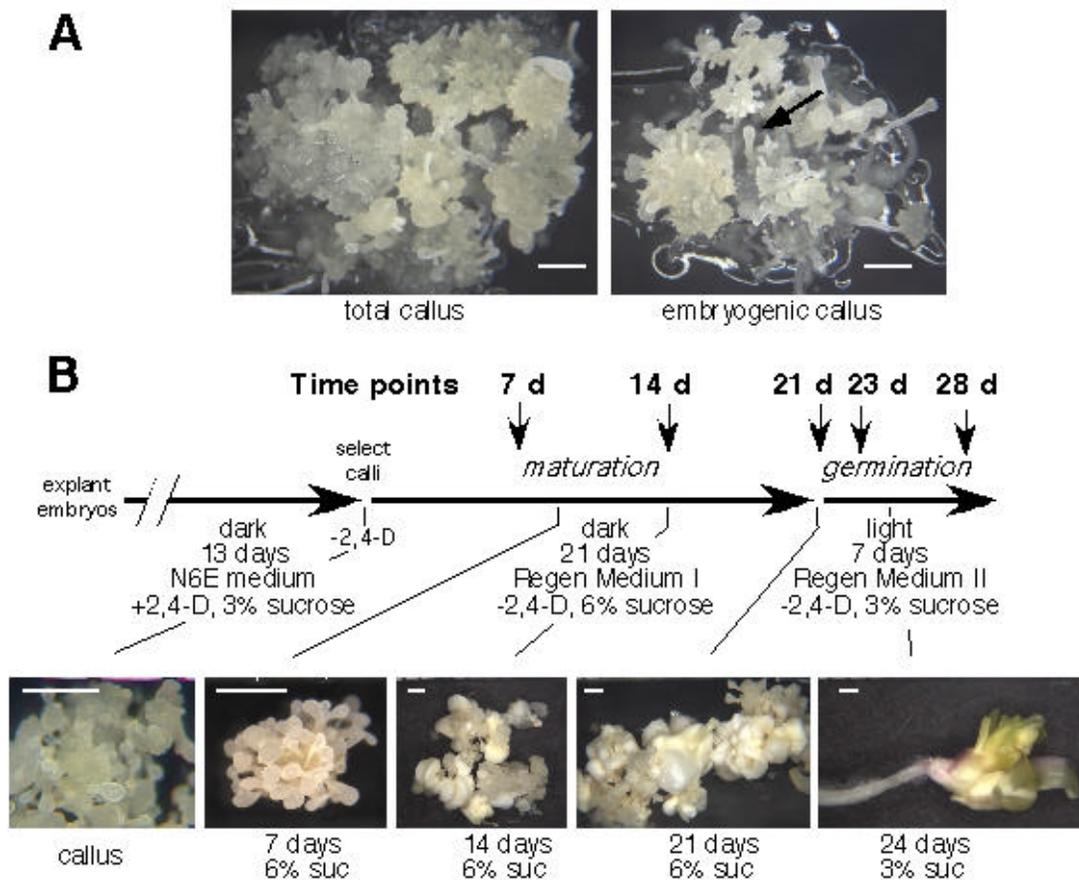


Figure 1

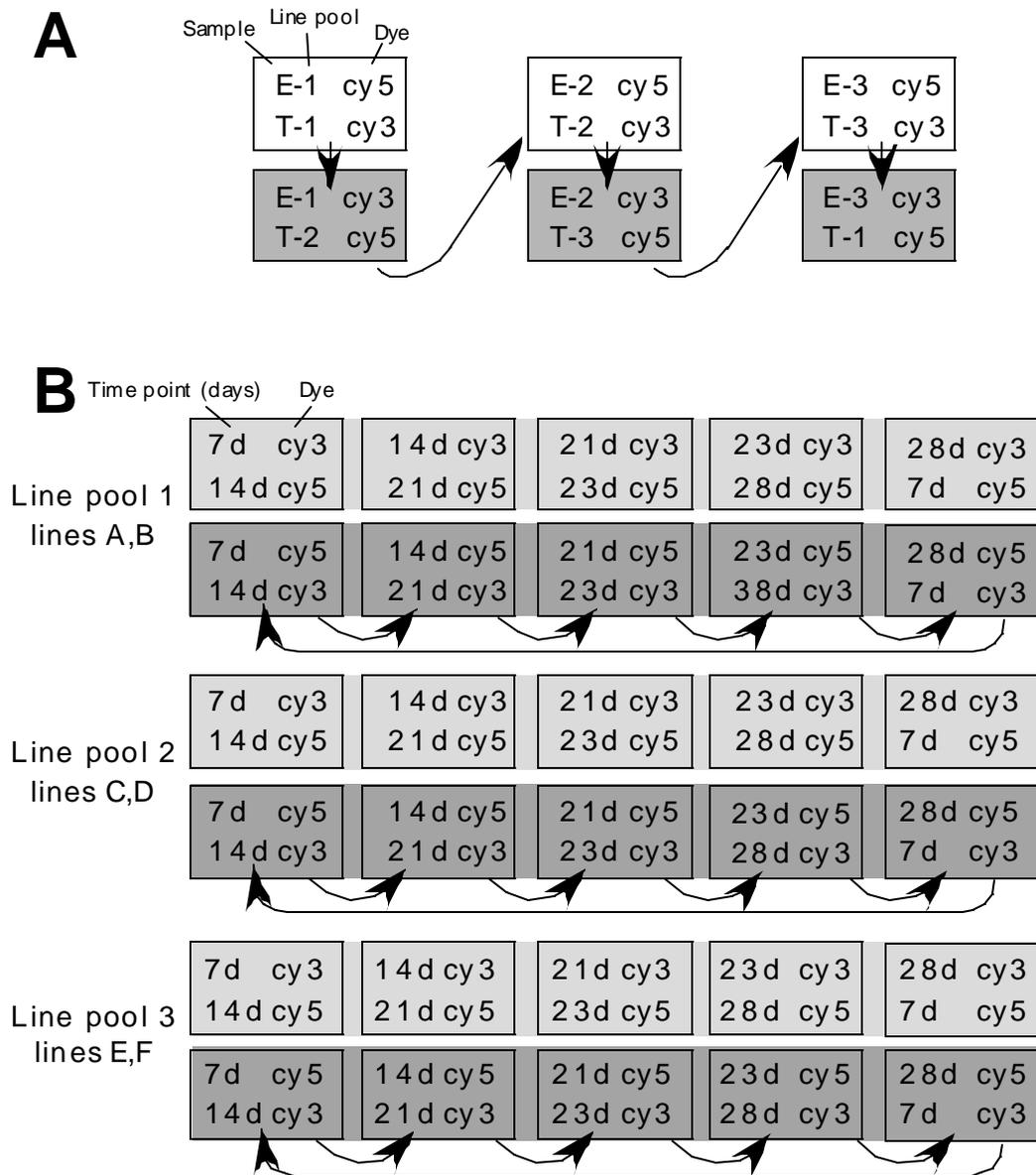


Figure 2

Figure 3

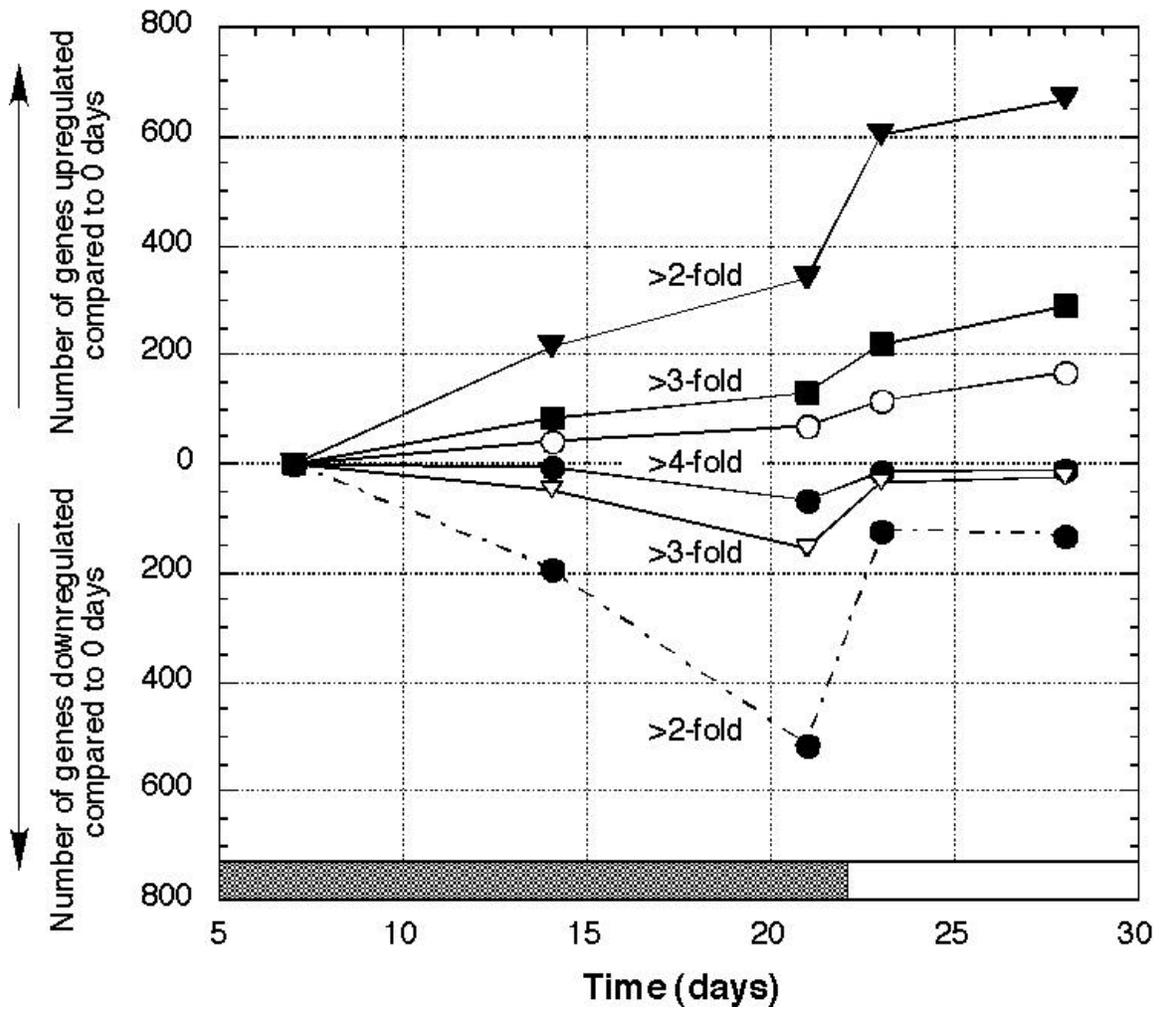


Figure 5



Figure 6

