Plant Mol Biol
DOI 10.1007/s11103-006-9013-2

ORIGINAL PAPER

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

Ping Che · Tanzy M. Love · Bronwyn R. Frame · Kan Wang · Alicia L. Carriquiry · Stephen H. Howell

Received: 31 March 2005 / Accepted: 4 August 2005
© Springer Science+Business Media B.V. 2006

Abstract Gene expression patterns were profiled during somatic embryogenesis in a regeneration-proficient maize hybrid line, Hi II, in an effort to identify genes that might be used as developmental markers or targets to optimize regeneration steps for recovering maize plants from tissue culture. Gene expression profiles were generated from embryogenic calli induced to undergo embryo maturation and germination. Over 1,000 genes in the 12,060 element arrays showed significant time variation during somatic embryo development. A substantial number of genes were downregulated during embryo maturation, largely histone and ribosomal protein genes, which may result from a slowdown in cell proliferation and growth during embryo maturation. The expression of these genes dramatically recovered at germination. Other genes up-regulated during embryo maturation included genes encoding hydrolytic enzymes (nucleases, glucosidases and proteases) and a few storage genes (an α-zein and caleosin), which are good candidates for developmental marker genes. Germination is accompanied by the up-regulation of a number of stress response and membrane transporter genes, and, as expected, greening is associated with the up-regulation of many genes encoding photosynthetic and chloroplast components. Thus, some, but not all genes typically associated with zygotic embryogenesis are significantly up or down-regulated during somatic embryogenesis in Hi II maize line regeneration. Although many genes varied in expression throughout somatic embryo development in this study, no statistically significant gene expression changes were detected between total embryogenic callus and callus enriched for transition stage somatic embryos.

Keywords Somatic embryogenesis · Microarray · Transcriptome · Plant regeneration · Tissue culture · Corn

Supplementary material is available for this article at http://www.dx.doi.org/10.1007/s11103-006-9013-2

P. Che · S. H. Howell (✉)
Plant Sciences Institute, Iowa State University, Ames, IA 50011, USA
e-mail: SHH@iastate.edu

P. Che
e-mail: pche@iastate.edu

T. M. Love · A. L. Carriquiry
Department of Statistics, Iowa State University, Ames, IA 50011, USA

T. M. Love
e-mail: tanzy@iastate.edu

A. L. Carriquiry
e-mail: alicia@iastate.edu

B. R. Frame · K. Wang
Center for Plant Transformation and the Department of Agronomy, Iowa State University, Ames, IA 50011, USA

B. R. Frame
e-mail: bframe@iastate.edu

K. Wang
e-mail: kanwang@iastate.edu

Introduction

The regeneration of maize in tissue culture is important for the production of transgenic maize and for crop improvement using genetic engineering approaches. The first somatic embryos in maize tissue culture were produced by Green and Phillips (1975). 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 × B73 line (Gordon-Kamm et al. 1990).

Maize cell lines derived from transformation competent sources such as immature embryos are heterogeneous for cells with differing embryogenic potential. Friable (Type II) callus (Armstrong and Green 1985) was found to be highly embryogenic and readily produced plants. Induction of embryogenic callus is genotype-specific in many plant species, including maize. Most maize elite lines remain inaccessible to improvement using standard transformation techniques either because they fail to produce embryogenic callus from transformation competent tissues, or they fail to regenerate efficiently after embryogenic callus induction.

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 specifically expressed in embryogenic carrot cells. One such gene encoded a leucine-repeat receptor protein kinase and was dubbed as a somatic embryogenesis receptor kinase (SERK) (Schmidt et al. 1997). In Arabidopsis, five members of the SERK family have been identified (AtSERK1-5). AtSERK1 was expressed during somatic embryogenesis, and the embryogenic competence of callus derived from seedlings over-expressing AtSERK1 (driven by the CaMV35S promoter) was elevated 3 to 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 was preferentially expressed in reproductive tissues with the strongest expression in microspores, while ZmSERK2 expression was fairly uniform in all tissues investigated. Both genes were expressed in callus cultures whether they were embryogenic or not, which suggested that the genes might not be good markers for embryogenesis in maize (Baudino et al. 2001).

LEAFY COTYLEDON 1 (*LEC1*) in Arabidopsis is up-regulated during zygotic embryogenesis and promoted somatic embryogenesis when ectopically expressed in vegetative cells (Lotan et al. 1998). *LEC1* encodes a transcription factor, and *lec1* mutants prematurely germinate producing 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; Parcy et al. 1997; Vicient et al. 2000). Maize genes with sequences similar to *LEC1* have been identified, and the expression pattern of ZmLec1 has been profiled during somatic embryogenesis (Zhang et al. 2002). The expression of ZmLec1 during maize somatic embryogenesis was similar to *LEC1* during Arabidopsis zygotic embryogenesis with general expression throughout 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. The expression patterns of 900 other maize genes have also been profiled during zygotic embryogenesis in a study that focused primarily of the expression of metabolic genes (Lee et al. 2002).

In this study, we 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 somatic embryo maturation after removal from auxin-containing medium. However, no significant changes in gene expression were evident when comparing embryogenic callus enriched with transition stage somatic embryos and total callus on auxin-containing medium. The genes regulated during these later stages of somatic embryogenesis may serve as developmental markers for improving the regeneration of more recalcitrant lines.

Materials and methods

Materials and tissue culture methods

Somatic embryos were generated in embryogenic callus lines developed independently from immature Hi II zygotic 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, greenhouse-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 + 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 μM) 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 sub-culturing every 2 weeks on the same medium. Callus was then subjected to regeneration conditions by transferring about 15 small pieces (approximately 4 mm) of embryoid-enriched 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 (McCain and Hodges 1986). Petri-plates (100 × 25 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 $\mu E/m^2/s$) 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°C for 5 min. The mixture was centrifuged at 12,000*g* at 4°C for 10 min and to the supernatant was added 3 ml of chloroform. The mixture was vortexed for 15 s and allowed to sit at room temperature for 2–3 min. The mixture was centrifuged at 10,000*g* at 4°C for 15 min, and RNA was precipitated from the upper phase by adding ½ volume each of isopropanol and 0.8 M sodium citrate/1.2 M NaCl. The mixture was allowed to sit at room temperature for 10 min and centrifuged at 10,000*g* at 4° for 10 min. The pellet was washed with 70% EtOH, vortexed briefly and centrifuged again at 10,000*g* at 4°C for 10 min. The pellet was air dried for 5 min and dissolved in 250 μl 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 an 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 aminoallyl 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, 2 mM 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°C followed by base hydrolysis of RNA in 200 mM NaOH, 10 mM EDTA and incubation for 15 min at 65°C.

The aminoallyl-label cDNA was purified using a modified QIAquick (Qiagen) PCR purification procedure. The cDNA reaction was mixed with 5× 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 μl of 5 mM PB (pH 8.0) and 80% EtOH and centrifuged each time. Thirty microliter 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 μl 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 μl of freshly prepared 0.1 M sodium carbonate buffer (pH 9.0). Cy3- or Cy5-esters (AmershamPharmacia) were dissolved in 73 μl DMSO, and 4.5 μl of the appropriate NHS-Cy were added to the labeled cDNA. The mixture was incubated in the dark at room temperature for 1 h. Following the reaction uncoupled dye was removed using a QIAquick PCR purification kit (Qiagen). Thirty-five microliter of sodium acetate buffer (pH 5.2) and 250 μl 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 μl 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 (5× SSC, 0.1% SDS and 1% bovine serum albumin) and incubated at 42° for 45 min. The slides were washed 5× 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 μl of hybridization buffer (50% formamide, 5× SSC and 0.1% SDS) to which was added 1 μl of 20 μg/μl human COT1 DNA (LifeTechnologies) and 1 μl of 20 μg/μl poly A DNA (Invitrogen) to block non-specific hybridization. The sample was heated at 95°C 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 × 60 mm glass coverslip and placed in a sealed hybridization chamber with 20 μl of water added to the chamber at the end of the slide. The chamber was incubated overnight at 42°C. Following incubation the slide was carefully removed from the chamber and placed in a staining dish with wash buffer containing 1× SSC and 0.2% SDS at 42°C. The coverslip was gently removed, and the slide was agitated in the wash buffer for 4 min. The slide was further washed with 0.1× SSC and 0.2% SDS at room temperature for 4 min and then in 0.1× 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 a 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 65,535 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 65,535 (the maximum).

Data normalization

We assume that there is a systematic bias in the gene expression measurements between the two dyes. For gene j that 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 $A_j = \log(\mathrm{sqrt}(R_j G_j))$ 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.

We used print tip group-intensity dependent normalization to remove the systematic bias in our data related to dye and print tip group. We also fitted the loess curves of intensity for each print tip group and corrected each pair of expression values on a slide for the curve.

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 (12,060) 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. We do so using the P-values generated by the ANOVA test, P_j, for each gene.

We use an adjustment proposed by Benjamini and Hochberg (1995). This adjustment attempts to control the expected proportion of false positives out of genes concluded to be differentially expressed. This proportion is called the false discovery rate. The multiple comparisons adjustment used assumes that the test statistics generating the P-values are independent. The j-th gene is considered significantly differentially expressed over time if $P_j \le P_{(k)}$ where $P_{(k)}$ is the k-th ordered P-value for the genes, $k = \max\{j{:}P_{(j)} \le j^*\alpha/t\}$, t is the number of tests being performed, and $\alpha > 0$ is the predetermined target error rate. Using this rule, the expected false discovery rate will be less than α.

Results

Somatic embryogenesis and expression profiling

Callus derived from maize Hi II immature zygotic embryos can be propagated in vitro as Type II callus (Armstrong et al. 1991). During the growth of this callus on auxin (2,4-D)-containing medium, some of the callus cells form embryogenic cell clusters, which eventually differentiate into globular and transitional stage somatic embryos (Jiménez 2001), or so-called embryoids (globular-like embryos with conspicuous suspensor-like structures, Fig. 1A) (Armstrong and Green 1985). For routine maize regeneration, highly embryogenic callus, rich in its content of embryoids, is transferred onto Regeneration Medium I (no 2,4-D, 6% sucrose) to induce somatic embryo maturation (Fig. 1B). After 7 days on this medium, tissue destined to form mature somatic embryos appears milky or less translucent. Embryo development and maturation continues for 21 days, and when mature somatic embryos are transferred to light on Regeneration Medium II (no 2,4-D, 3% sucrose), the embryos germinate.

Two independent experiments were conducted to examine gene expression patterns during somatic embryogenesis in maize (Fig. 2A, B). The goal of the first experiment (Fig. 2A) was to profile 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. The second (Fig. 2B) was designed to determine whether gene expression differences could be detected between embryogenic callus enriched with embryoids and total embryogenic callus growing on 2,4-D-containing (N6E) medium.

Six independent, embryogenic 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. Thirty were used for the time course analysis following induction of somatic embryo maturation (Fig. 2A) and the remaining six arrays were used to compare embryoid-enriched and total callus prior to removal from auxin-containing medium (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

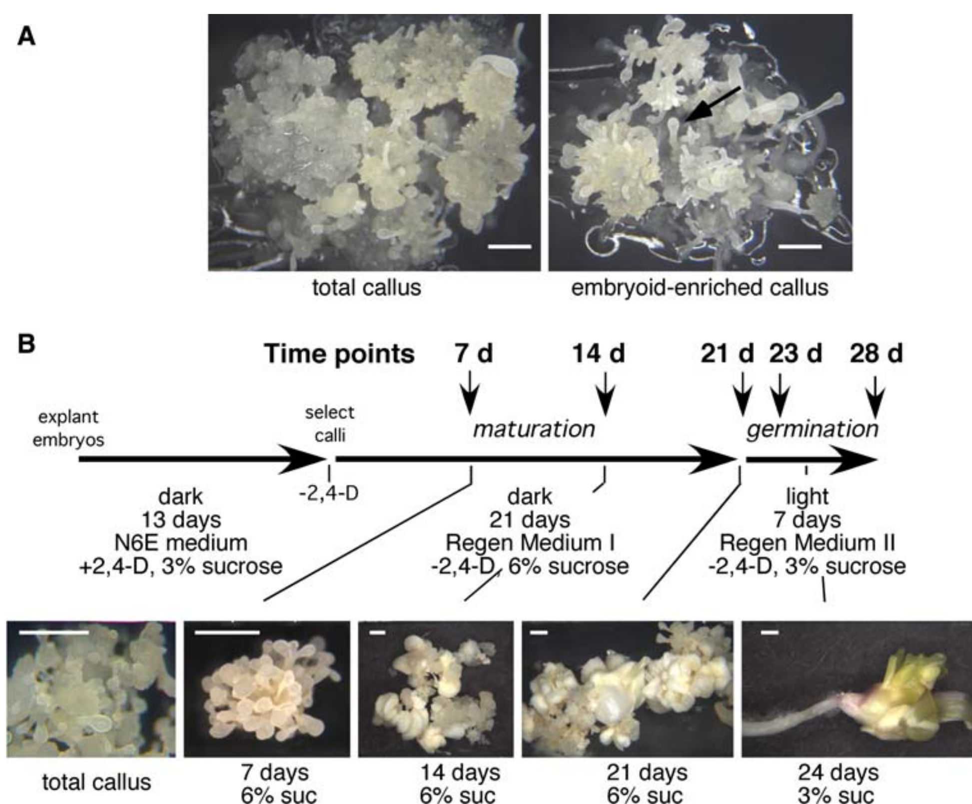


Fig. 1 Somatic embryo development in maize Hi II callus. (**A**) Examples of total and embryoid enriched callus growing on N6E medium. Arrow points out one of many transition stage somatic embryos in embryogenic callus. (**B**) Time course of somatic embryo development, maturation and germination. Somatic embryo maturation was initiated by transferring embryoid-enriched callus to 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

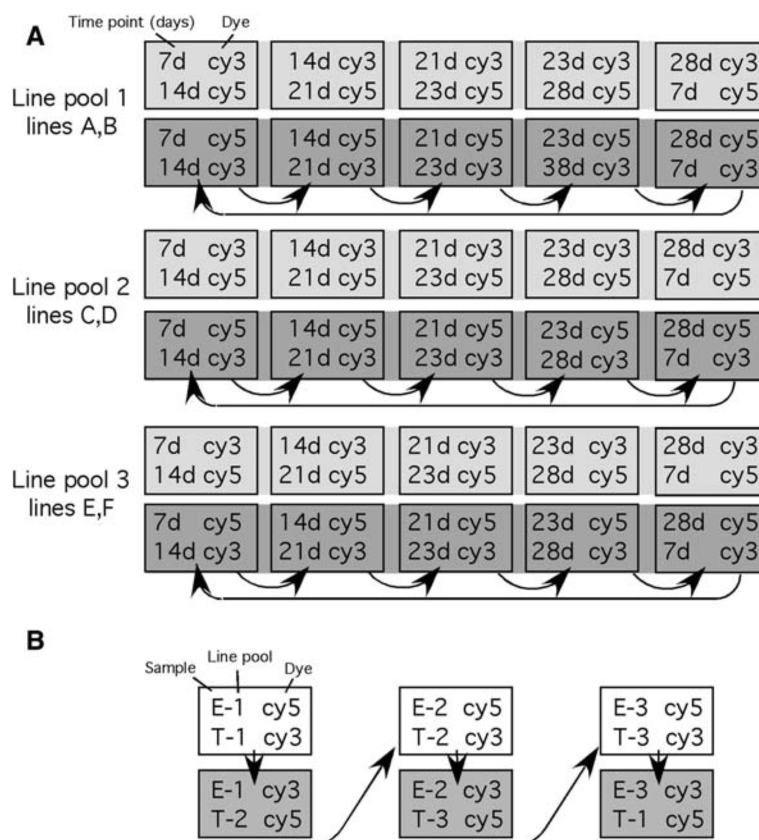


Fig. 2 Loop design for microarray hybridization experiments. Six independent callus lines (A–F) were initiated, and the lines were pooled into 2 lines per pool to obtain enough RNA in each pool for the microarray analysis. Each rectangle represents 1 chip. (**A**) Time course analysis. Time point and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

not to a single reference, such as a zero time sample. 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. However, it should be noted that of the 12 repeated measurements on each time point, only three are true biological replications (3 line pools). Thus, the power of our conclusions is lessened by the fact that the four replications at each time point within a line pool are technical replications using the same biological material.

Gene expression patterns following induction of somatic embryo maturation

Following induction of embryo maturation, somewhat more than a 1,000 genes out of 12,060 in the study showed significant time variation (at the $\alpha = 0.05$ level, considering multiple comparisons, see supplementary Table 1). During maturation and germination, increasing numbers of genes were up-regulated by 2-fold or more (Fig. 3). Likewise, an increasing number of genes were down regulated 2-fold or more during maturation, but that trend reversed itself during embryo germination.

The overall trends were made up of individual genes with varied expression patterns, and patterns for ~1,000 genes with significant time variation were organized into groups. The patterns were clustered into 12 groups on model-based clustering of the sequential expression ratios. 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 four groups of genes in which a category, other than unknown or primary metabolism, emerged or dominated the distributions are shown.

The first pattern group was characterized by genes down-regulated during embryo maturation, which then recovered during germination (Fig. 4A). Compared to other groups, this group had a larger number of genes encoding nuclear proteins, such as a gene encoding proliferating cell nuclear antigen, and histone genes. 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 accompanied 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. 4B).

Early in germination, stress-related genes, such as a gene encoding a heat shock protein, were transiently up-regulated (Fig 4C). Other genes that were up-regulated during germination (with expression levels generally higher at both time points) included a large group encoding channel proteins or membrane transporters, such as a gene encoding a water channel

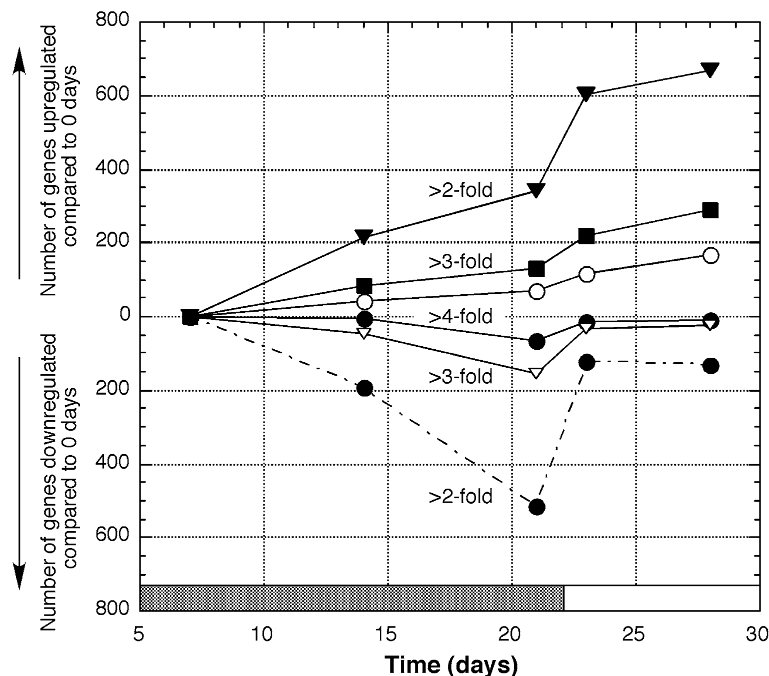


Fig. 3 Summary of gene expression during somatic embryo maturation and germination in maize. Number of genes out of 1,026 genes are shown that vary significantly with time and are either up or down-regulated more than 2-, 3- or 4-fold during embryo development. Shaded bar represents the period of somatic embryo development and maturation. Unshaded bar is the time of germination

(data not shown). Finally, genes encoding photosynthetic and other chloroplast components, such as a gene encoding a chlorophyll *a*/*b* binding protein, were up-regulated as shoots began to green (Fig. 4D). Thus, from a gene expression perspective, germination first involved the activation of expression of stress-related and transporter/channel-encoding genes followed by the up-regulation 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 α-zein, an embryo storage protein, was up-regulated 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, caleosin) 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 up-regulated, 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.

The expression patterns of a number of genes were stage-specific, i.e. peaking at a particular time point. Such genes might make good markers to track the progress of embryo development. The genes that make the clearest markers were those with the most significant time variation in expression (low P values for significance in variation with time). Genes expressed early in embryo maturation (7 days) and those expressed in germination clearly showed the most significant time variation of expression (Table 1 and supplementary Table 1). Markers for mid to late embryo maturation were not as clear (had higher P values). Only one gene in our study that showed significant time variation in expression peaked at 14 days and 28 days. This gene of unknown function (606040E09.x1, Genbank AI737443) could be used as a midmaturation marker if one ignores its other peak of expression during germination. Two other genes including a chitinase (MEST40-G02, GenBank BG842841) peak more broadly at 14–21 days and might be useful as mid to late maturation markers (Table 1).

The genes with the most significant time variation encode chloroplast proteins (supplementary Table 1). They rapidly rise in expression during germination and many peak at the last time point (28 days). A gene encoding a non-chloroplast protein that varies significantly with time and rises with the same time course as the chloroplast proteins is a gene that encodes a protein phosphatase 2C (707051D04.x1, GenBank AW399849) (supplementary Table 1). A gene encoding a transcription factor that varies most significantly with time encodes a GT-1 transcription factor (614025B02.y1, GenBankAW037041) (supplementary Table 1). The expression of this gene peaks at 23 days. A gene (603040B10.x1, Genbank AI964625) encoding a translation factor that varies most significantly with time peaks in expression at the end of the maturation period (21 days) (supplementary Table 1).

Line variation

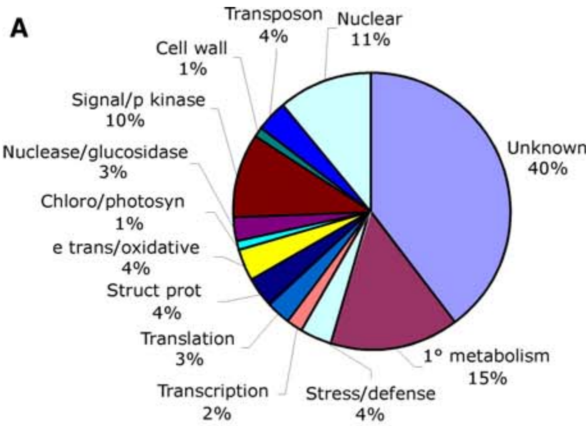
We also looked for genes with significant variation across time points and either with considerable or little variation across lines. Genes with expression patterns that vary considerably across lines might be useful to others if the variation correlates with regeneration competence or with other traits that vary from line to line. In our case, the Hi II maize lines were highly embryogenic, and we observed little variation in regeneration competency across the cell culture lines. Nonetheless, two genes, one encoding a lipid transfer protein and the other a bZIP family transcription factor, showed significant time variation in one line pool, but not in the other two line pools (Fig. 6, upper panel). Such genes might be useful developmental markers for distinguishing cultures that show line variation in later regeneration steps. Other genes, such as one encoding a putative disease response protein and a Photosystem I assembly protein, showed significant variation across time points but little variation between lines (Fig. 6, lower panel).

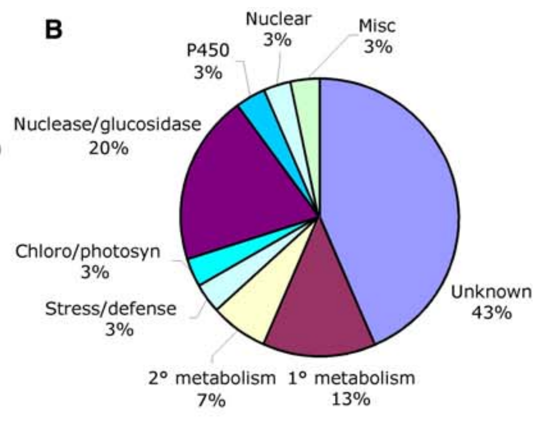
Comparison of embryoid enriched and total callus

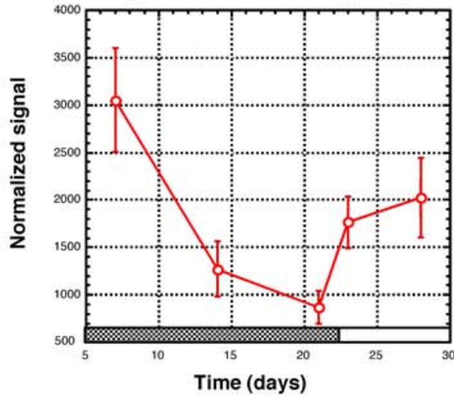
In our maize regeneration procedure (http://www.agron.iastate.edu/ptf/ Web/mainframe.htm), embryoid-enriched callus is selected (using a dissecting scope)

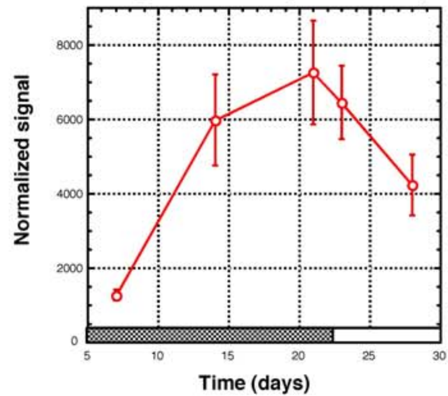
▶

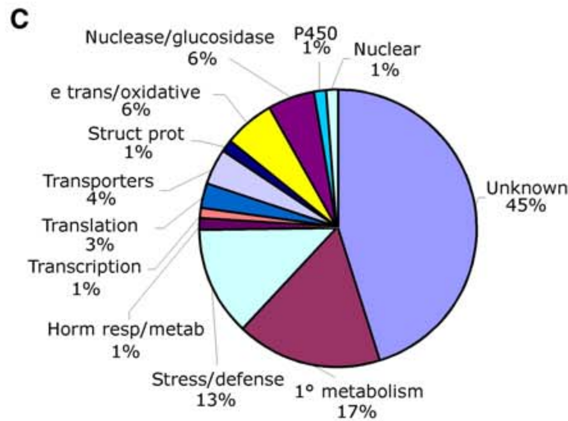
Fig. 4 Genes with different expression profiles. Expression profiles of the 1,026 genes with the greatest variation across time points cluster into 12 pattern groups. Examples of genes from four different pattern groups are shown here. The distributions of gene functions in the pattern group are shown in the pie charts. Genes were categorized into 19 functional groups. Means and standard errors (SEs) for 12 repeats at each time point are shown in the line graphs. Period of embryo maturation (stippled bar), embryo germination (unshaded bar)

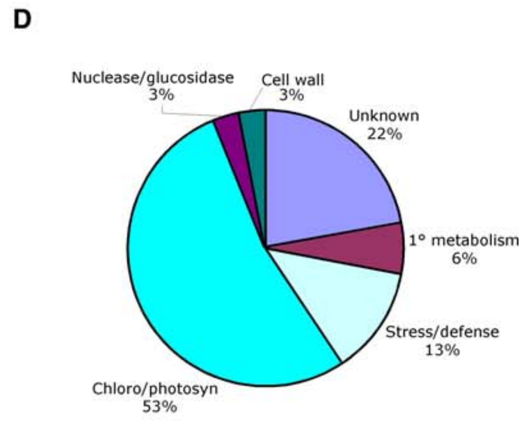

A
Transposon 4%
Nuclear 11%
Cell wall 1%
Signal/p kinase 10%
Nuclease/glucosidase 3%
Chloro/photosyn 1%
e trans/oxidative 4%
Struct prot 4%
Translation 3%
Transcription 2%
Stress/defense 4%
1° metabolism 15%
Unknown 40%

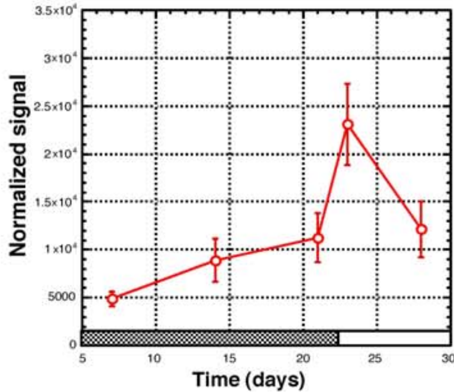

B
Nuclear 3%
Misc 3%
P450 3%
Nuclease/glucosidase 20%
Chloro/photosyn 3%
Stress/defense 3%
2° metabolism 7%
1° metabolism 13%
Unknown 43%

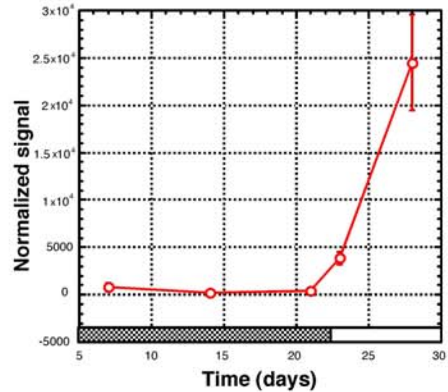

Proliferating cell nuclear antigen (AL734348)
Normalized signal
Time (days)


Beta-glucosidase (AW352489)
Normalized signal
Time (days)


C
Nuclease/glucosidase 6%
P450 1%
Nuclear 1%
e trans/oxidative 6%
Struct prot 1%
Transporters 4%
Translation 3%
Transcription 1%
Horm resp/metab 1%
Stress/defense 13%
1° metabolism 17%
Unknown 45%


D
Nuclease/glucosidase 3%
Cell wall 3%
Unknown 22%
1° metabolism 6%
Stress/defense 13%
Chloro/photosyn 53%


Heat shock protein (AI901570)
Normalized signal
Time (days)


Chlorophyll a/b binding prot (BG841274)
Normalized signal
Time (days)

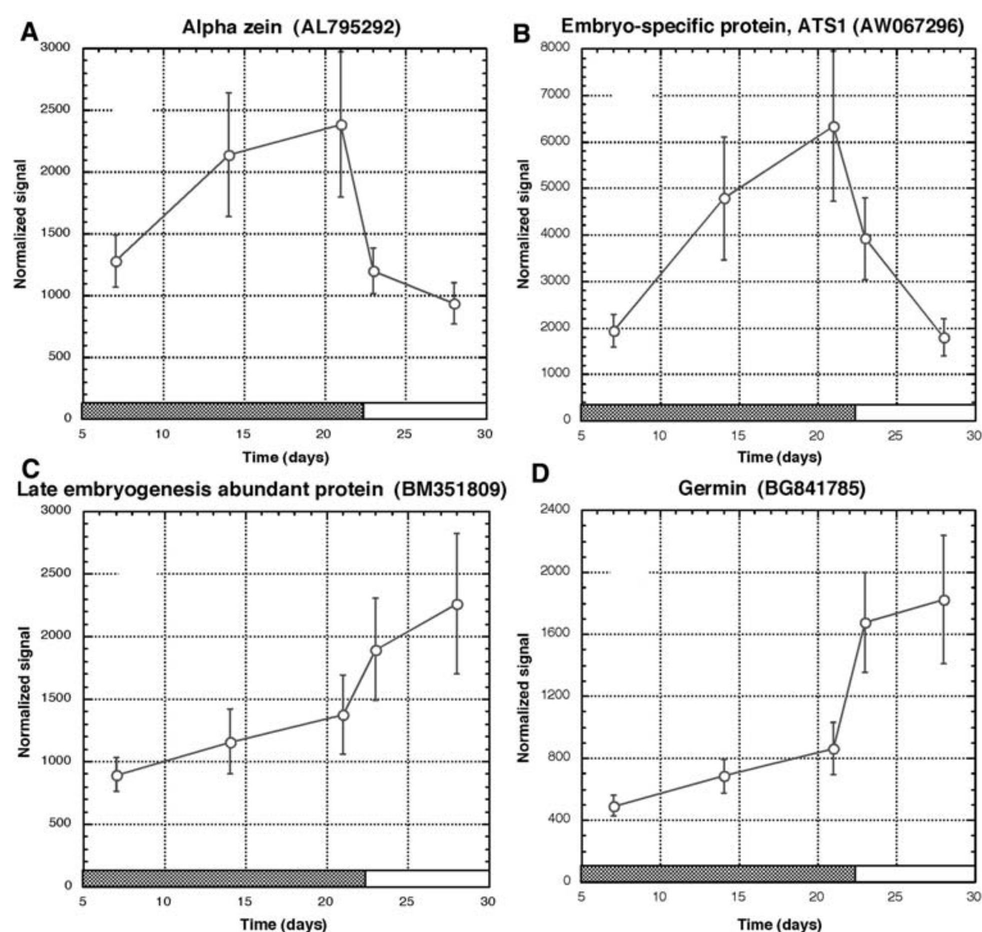


Fig. 5 Expression profiles for four genes belonging to classes of genes expressed during zygotic embryo development or germination

from friable embryogenic callus for transfer onto Regeneration Medium I. This regeneration approach was reported to produce over 30 times more plants per gram fresh weight of callus than did the indiscriminate transfer of total embryogenic callus to the same medium (McCain and Hodges 1986). However, when we compared gene expression levels in embryoid-enriched callus to those in total callus, we found that none of the

Table 1 Genes that vary significantly and peak in expression at different stages during embryo maturation and germination

Time Point	EST accession #	GenBank #	*P* value[a]	Gene function
Maturation (7 days)	606075A10.x1	AI855092	1.55e-09	phosphoenolpyruvate carboxylase
	683028H12.xl	AW091476	1.88e-09	unknown
Maturation (14 days)	606040E09.x1[b]	AI737443	4.33e-06	unknown
Maturation[c] (14–21 days)	MEST40-G02	BG842841	2.57e-04	chitinase
	603012D01.x1	AI861117	5.57e-05	unknown
Maturation (21 days)	MEST39-H04	BG842769	1.51e-08	unknown
	605018A02.x1	AI668188	5.86e-08	unknown
Germination (23 days)	614058D04.y1	AW065683	1.34e-12	unknown
	MEST22-F11	BG841531	1.51e-12	lipoxygenase
Germination (28 days)	496021F12.x1	AI967091	4.20e-14	chlorophyll *a*/*b* binding protein 1
	MEST140-B10	BM334653	7.90e-14	RuBPCase small subunit

[a] *P* values for significance of variation with time (from supplementary Table 1)

[b] The expression pattern for this gene peaks at two stages, once in maturation and once in germination. This gene can be included in the group that peaks at 14 days (maturation) if one ignores another peak of expression during germination

[c] This group represents genes that peak at 14–21 days. The group is included because the group that peaks at 14 days is underpopulated with genes

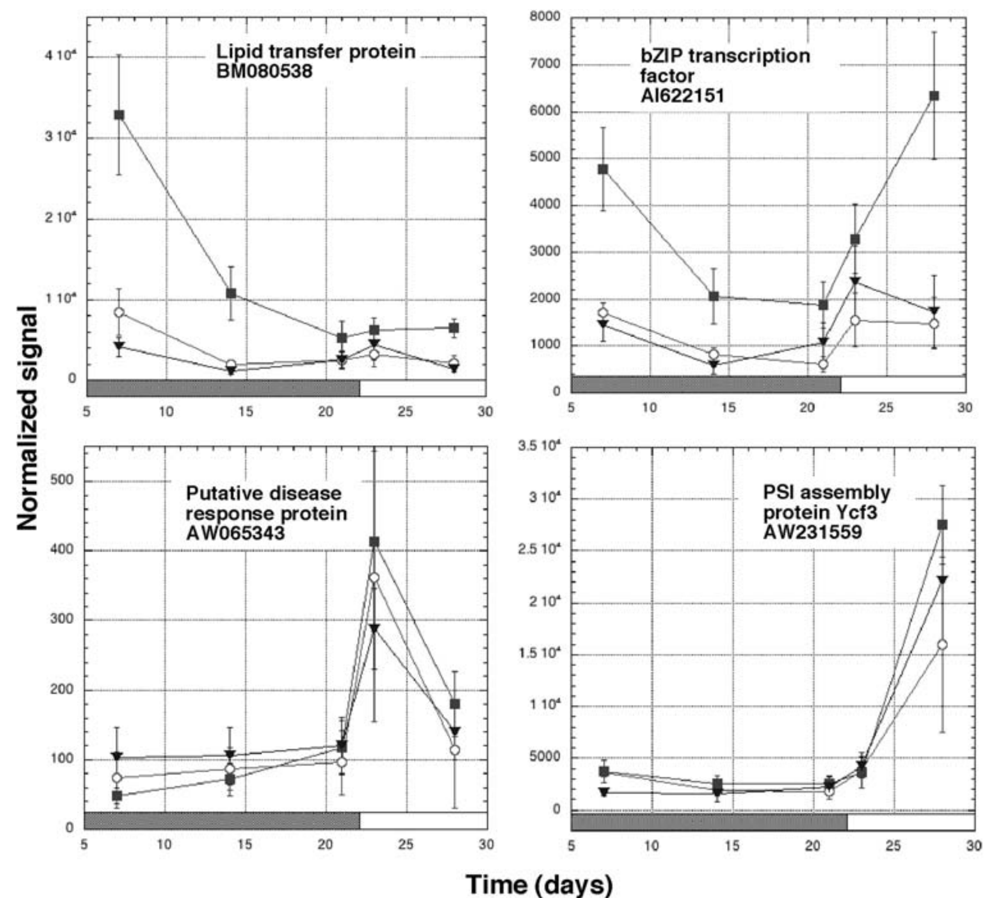


Fig. 6 Line variation in gene expression. Examples of genes with significant variation across time points and that (upper panel) show significant line pool variation or (lower panel) show little line pool variation. Period of embryo maturation (stippled bar), embryo germination (unshaded bar). Line pool 1 ■—■, line pool 2 ○—○, line pool 3 ▼—▼

12,060 genes in this study showed significant differences (at the $\alpha = 0.05$ level, considering multiple comparisons, see supplementary Table 2).

Discussion

Gene expression patterns change extensively during somatic embryo 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 (Fig. 7). 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 embryo maturation and 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 this stage of somatic embryo development. Unlike zygotic embryogenesis (Lending and Larkins 1989), we did not observe the large-scale up-regulation in expression of storage protein genes. Whether maize Hi II somatic embryos accumulate fewer storage proteins than their zygotic counterparts is a matter that deserves exploration. In any case, only a few storage protein genes for α-zein and a caleosin, a lipid body protein (Naested et al. 2000), appear to be good markers for somatic embryo maturation in these Hi II lines.

Some stress response genes, such as heat shock genes are up-regulated at the onset of germination (Fig. 7). Their up-regulation may be a normal developmental event or a response to the transfer of tissue to new culture medium. Of interest is the up-regulation of a group of genes that encodes various transporters and membrane channels. Finally, as expected, germination and shoot greening are accompanied by the activation in expression of a myriad of genes encoding photosynthetic and chloroplast components.

It was surprising to find little correlation between our study on somatic embryogenesis and those of Lee et al. (2002) who described gene expression patterns during zygotic embryogenesis in maize. We can rationalize this by the fact that Lee et al. focused on a

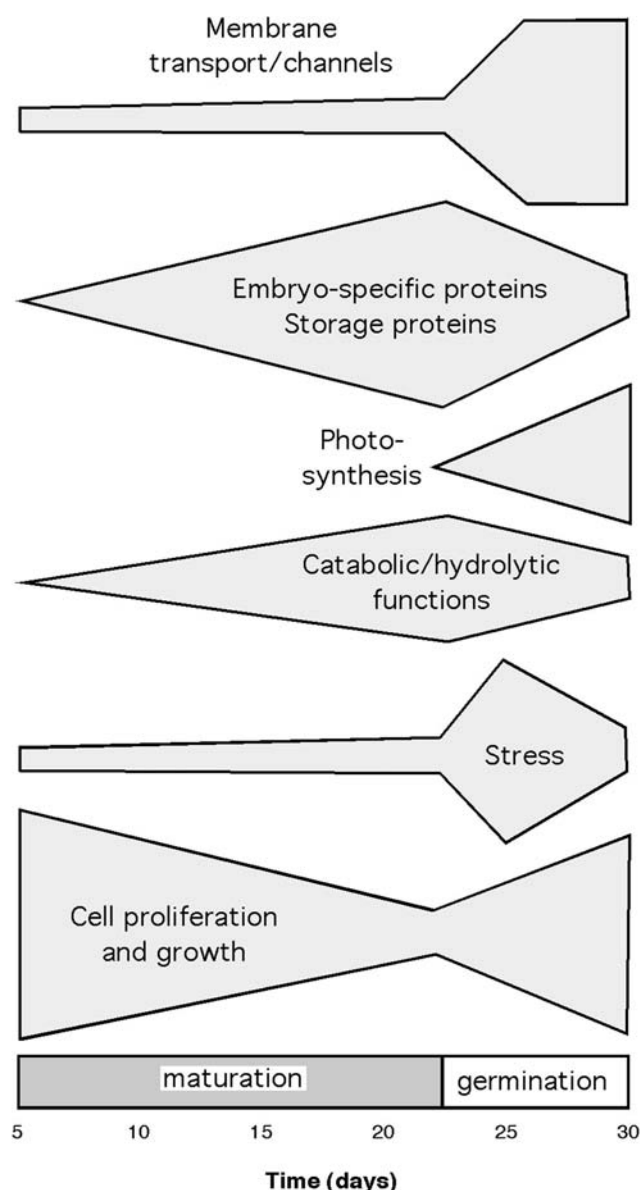


Fig. 7 Trends in expression of genes in various functional categories show significant time variation during embryo maturation and germination. Based on the expression profiles of genes that are typical of the pattern group, most of which are shown in Figs. 4 and 5. In descending order in the diagram: water channel, BM079333; alpha zein, AL795292; chlorophyll *a*/*b* binding protein, BG841274; beta-glucosidase, AW352489; heat shock protein, AI901570; proliferating cell nuclear antigen, AL734348

selected group of 900 genes mostly relating to metabolism. It is likely that the metabolism of embryos in planta is very different than in tissue culture (in nutrient media). What appears to be somewhat similar in the two studies are the expression profiles of genes encoding proteins involved in fatty acid metabolism, such as acyl carrier protein (MEST13-H03, GenBank BG840953) and acyl-CoA binding protein (MEST13-H03, GenBank BG840953) (supplementary table I). The expression patterns of these genes in our study track overall trends and those of genes encoding proteins involved in general growth and proliferation. The expression of these fatty acid metabolism genes declines during embryo maturation, but increases again upon germination. Other genes with expression patterns comparable between the study by Lee et al and ours encode late maturation proteins, such as RAB17 (603031G08.x1, GenBank AI947963) (supplementary table 1). These genes increase in expression during embryo maturation and decline during germination.

Some of the gene expression patterns we observed were significantly line-pool dependent and others were not. Highly regulated genes with expression patterns that are line independent may be good developmental markers across multiple lines. Line-dependent genes, on the other hand, may be useful if their expression patterns correlate with traits such as efficient regeneration of fertile plants.

We also looked for gene expression differences between embryoid-enriched and total callus and found none that were statistically significant. Two possible reasons for this may be: (1) total callus from the regeneration-proficient Hi II line is replete with viable embryoids but also contains embryogenic cells clusters and globular embryos (without suspensors), both of which also represent early stages of somatic embryogenesis. The developmental difference between embryoid enriched and total callus may therefore be modest. (2) Genes that differ in expression between embryoid enriched and total callus may not be present on the cDNA chips or may not be reflected by differences in the transcriptome. Gene expression differences between embryogenic and pre- or non-embryogenic callus might be more effectively detected in less regeneration-proficient lines.

In early attempts to identify markers for embryogenic competence, translation products of RNA from cultured carrot cells and somatic embryo were compared by 2D gel electrophoresis. 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 pro-embryonic 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.

Gene expression markers have been used more widely in recent years to characterize embryogenic lines and to describe embryo development. 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. A recent microarray study by Thibaud-Nissen et al. (2003) profiled gene expression patterns during somatic embryogenesis in soybean. 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 2 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).

Some of the general features of the gene expression program in soybean were also observed in the course of maize somatic embryogenesis such as the increase in expression of certain storage protein genes, the fall and subsequent rise in cell division gene expression and the mid-course expression of stress response genes. Because we measured gene expression patterns during the late stages of somatic embryo development in this study, we would not expect to observe expression of genes associated with oxidative burst, detoxification and cell wall modification that Thibaud-Nissen et al. (2003) attributed to the earlier, dedifferentiation stage of somatic embryogenesis from soybean cotyledon tissue.

Acknowledgements 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, and the Biotechnology Program all at Iowa State University.

References

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 Newslett 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 Roy 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. Curr Sci 83:715–730

Dobbin K, Simon R (2002) Comparison of microarray designs for class comparison and class discovery. Bioinformatics 18:1438–1445

Fransz 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

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 29:548–550 552–554, 556 passim

Jiménez VM, Bangerth F (2001) Hormonal status of maize initial explants and of the embryogenic and non-embryogenic callus cultures derived from them as related to morphogenesis in vitro. Plant Sci 160:247–257

Lee JM, Williams ME, Tingey SV, Rafalski JA (2002) DNA array profiling of gene expression changes during maize embryo development. Funct Integr Genom 2:13–27

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 RL, Goldberg RB, Harada JJ (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

McCain JW, Hodges TK (1986) Anatomy of somatic embryos from maize embryo cultures. Bot Gaz 147:453–460

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

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. Development 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 Biol-Plant 32:179–183

Sung ZR, Okimoto R (1981) Embryonic proteins in somatic embryos of carrot. Proc Natl Acad Sci USA 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

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, Booij 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