Genome-Wide Analysis of Gene Expression Profiles Associated with Cell Cycle Transitions in Growing Organs of Arabidopsis[1][w]

Gerrit T.S. Beemster, Lieven De Veylder, Steven Vercruysse, Gerrit West, Debbie Rombaut, Paul Van Hummelen, Arnaud Galichet, Wilhelm Gruissem, Dirk Inzé*, and Marnik Vuylsteke

Department of Plant Systems Biology, Flanders Interuniversity Institute for Biotechnology, Ghent University, B-9052 Ghent, Belgium (G.T.S.B., L.D.V., S.V., G.W., D.R., D.I., M.V.); Microarray Facility, Flanders Interuniversity Institute for Biotechnology, 3000 Louvain, Belgium (P.V.H.); and Institute of Plant Sciences, Swiss Federal Institute of Plant Sciences, Swiss Federal Institute of Technology, CH–8092 Zurich, Switzerland (A.G., W.G.)

Organ growth results from the progression of component cells through subsequent phases of proliferation and expansion before reaching maturity. We combined kinematic analysis, flowcytometry, and microarray analysis to characterize cell cycle regulation during the growth process of leaves 1 and 2 of Arabidopsis (*Arabidopsis thaliana*). Kinematic analysis showed that the epidermis proliferates until day 12; thereafter, cells expand until day 19 when leaves reach maturity. Flowcytometry revealed that endoreduplication occurs from the time cell division rates decline until the end of cell expansion. Analysis of 10 time points with a 6k-cDNA microarray showed that transitions between the growth stages were closely reflected in the mRNA expression data. Subsequent genome-wide microarray analysis on the three main stages allowed us to categorize known cell cycle genes into three major classes: constitutively expressed, proliferative, and inhibitory. Comparison with published expression data obtained from root zones corresponding to similar developmental stages and from synchronized cell cultures supported this categorization and enabled us to identify a high confidence set of 131 proliferation genes. Most of those had an M phase-dependent expression pattern and, in addition to many known cell cycle-related genes, there were at least 90 that were unknown or previously not associated with proliferation.

Growth is one of the most widely studied processes of plant biology, depending both on genetic predisposition and a wide range of environmental and physiological parameters. A multitude of regulatory pathways converge and interact to control growth of individual parts of the plant through their effects on just two cellular processes: cell division and expansion. The precise association between these processes and growth of the organ as a whole is far from resolved (Tsukaya, 2002; Beemster et al., 2003). In general terms, growth results from the progression of component cells through a succession of developmental phases: proliferation (cells divide at a rate matching their expansion and maintain cell size homeostasis), expansion (cells stop dividing but continue to expand, typically to a size much larger than that of meristematic cells), and maturity (cells no longer expand). In indeterminate organs like roots, initial cells continuously produce new cells, which enter this developmental sequence and are subsequently displaced by their successors. This results in a spatial developmental gradient. In determinate organs, particularly dicotyledonous leaves, founder cells proliferate only for a limited time, and the cells they produce progress through development more or less synchronously. For both types of organ, the rate and duration of cell division and postmeristematic expansion determine growth through their effect on final cell number and mature cell size, respectively. In Arabidopsis (*Arabidopsis thaliana*), the size of mature leaf pavement cells is correlated to their ploidy levels, which vary from 2C to 32C, due to differences in the number of endoreduplication cycles they have undergone (Melaragno et al., 1993). It is currently not clear when cells endoreduplicate and if endoreduplication drives the expansion process or vise versa. Hypothetically, this possibility provides, next to cell production, a second link between cell cycle regulation and organ growth. Thus, to understand the genetic network that controls growth and the role of cell cycle-regulating genes therein, we

[1] This work was supported by the Interuniversity Poles of Attraction Programme-Belgian Science Policy (grant no. P5/13) and by the European Commission Quality of Life and Management of Living Resources program (grant no. QLK5–CT–2001–01871).

* Corresponding author; e-mail dirk.inze@psb.ugent.be; fax 32–(0)9–33–13–809.

[w] The online version of this article contains Web-only data.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.104.053884.

need to investigate the changes that occur as cells progress through these growth phases.

Based on sequence homology with known cell cycle genes from yeast and animal systems, we identified 61 core cell cycle genes encoded by the genome of Arabidopsis (Vandepoele et al., 2002). Later, 19 additional cyclins were added to this list (Wang et al., 2004). For some of those (like P- and T-type cyclins), no function in the cell cycle has been identified to date. There is a high degree of duplication, particularly of cyclin-dependent kinases (CDKs) and cyclins. Although the general function of each cell cycle gene is known from homologous systems, it is unclear how the individual genes interact to regulate cell cycle progression in the context of growing plant organs. Do the duplicates within each gene family fulfill the same role in different organs, or do they function in parallel in all organs? How are these genes connected to the up- and downstream regulatory circuitry? Recently, it was revealed that besides these core cell cycle genes, over 1,000 other genes show differential expression patterns in synchronized cell cultures (Menges et al., 2003). For most of these genes, it is unknown what their relationship is with the core cell cycle machinery, let alone what their function is in the context of an intact plant.

To start addressing such questions, we need to study the cell cycle in the context of growing organs. Unfortunately, the study of cell cycle activity in multicellular organs is not straightforward and there are many pitfalls, mainly related to heterogeneity between tissues and the sometimes considerable velocity of cells at the base of meristems (Webster and Macleod, 1980). Kinematic analysis, pioneered half a century ago (Goodwin and Stepka, 1945; Erickson and Sax, 1956) and since theoretically refined (Gandar, 1980; Silk, 1984), is based on the dynamic behavior of cells and therefore facilitates accurate calculation of cell division and expansion rates in individual tissues. In maize (*Zea mays*) leaves, kinematically determined cell division parameters correlate closely with molecular data obtained from the same material (Granier et al., 2000). We recently adapted the kinematic approach to the small size of the Arabidopsis leaf (De Veylder et al., 2001). In dicotyledonous leaves, the growth processes are primarily separated in time (Donnelly et al., 1999; De Veylder et al., 2001). Here, we explore this feature and show that it is possible, by harvesting leaf material throughout development, to characterize changes in DNA content and transcriptional activity accompanying the transition between subsequent growth phases. Based on this, we categorized core cell cycle genes into three distinct functional classes that help to understand their function in plant growth and development. Combining our own data with earlier published data from root tips and synchronized cell cultures, using the core cell cycle genes as a guide, we identified 131 high-confidence proliferation genes, for many of which the relationship with cell proliferation has, to our knowledge, not been established yet.

RESULTS

Kinematic Analysis

We set out to understand the relationship between gene expression and cell cycle regulation during organ growth. Therefore, we first characterized cell cycle and expansion parameters using the kinematic approach developed earlier (De Veylder et al., 2001). The area of leaves 1 and 2, which develops synchronously and are practically indistinguishable, was determined by quantitative image analysis. From day 5, when they could first be visualized in whole-mount specimen, until around day 10 after sowing, the area increased exponentially (Fig. 1A). This was due to a nearly constant rate of cell expansion during this period (Fig. 1B). After day 10, expansion rates gradually declined until a mature area of 14 ± 2 mm^2 was reached at about 19 d after sowing (Fig. 1, A and B). We determined cell size in the abaxial epidermis because it lacks trichomes. Until day 9, average cell area remained approximately constant, around 100 μm^2 (Fig. 1C), indicating that during this period leaf growth was proportional to the increase in epidermal cell number. After that, cell size increased to 10-fold that of dividing cells. From leaf size and average cell area, the total number of abaxial epidermal cells was estimated. It increased exponentially until day 9; the rate of increase then dropped steeply until the final number ($14{,}800 \pm 800$) was reached at about day 16 (Fig. 1D). Consequently, cell division rates, calculated from the change in cell number over time, were approximately constant at about 0.055 cell cell^{-1} h^{-1} until day 9. This value corresponds with an average cell cycle duration of about 18 h, similar to that observed earlier under comparable conditions in the leaves (De Veylder et al., 2001) and roots (Beemster and Baskin, 1998) of the same ecotype. Cell division rates rapidly declined between days 9 and 12. After that residual values around 0.01 cell cell^{-1} h^{-1} were observed until day 16 (Fig. 1E). As cell division rates began to decline, the first stomata appeared (Fig. 1F). The stomatal index, the fraction of guard cells, increased steadily to a final value of just over 20% around day 19, i.e. at the same time the leaf reaches its mature size. Thus, the growth process in the abaxial epidermis can be summarized as follows: proliferation until day 12 (active cell division accompanied by cell expansion); expansion from day 12 to 19 (ongoing cell expansion and differentiation, as evidenced by the appearance of stomata), and maturity from day 19 onward.

Flow Cytometry

From day 8 (when they could first be dissected) until day 25, we harvested whole-leaf samples for flow cytometry and evaluated the distribution of nuclear DNA content as a function of time. The correspondence of whole-leaf flow cytometry data with the kinematic data obtained from the abaxial epidermal layer only was remarkably close (Fig. 2): On days 8 and 10, when epidermal cells were proliferating, 80% of the

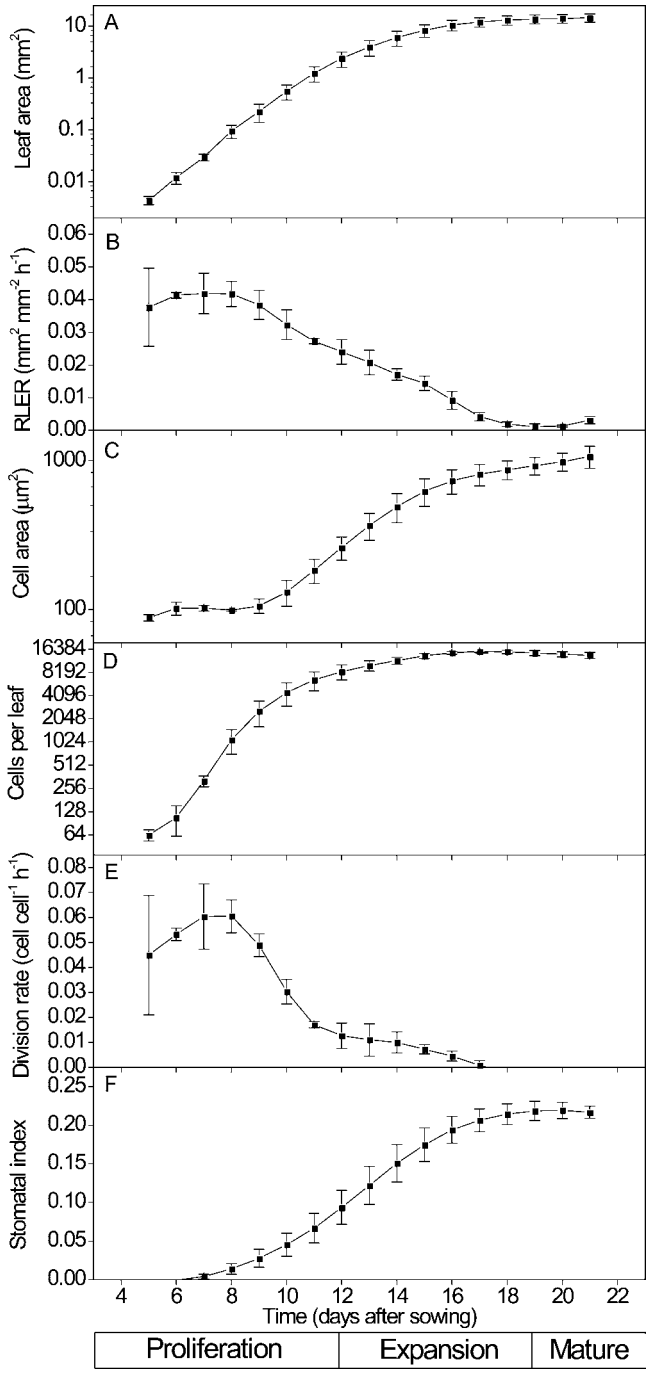

Figure 1. Kinematic analysis of leaf growth. A, Leaf area. B, Relative leaf expansion rate (RLER). C, Cell area. D, Number of cells per leaf. E, Cell division rate. F, Stomatal index. Symbols denote averages ± SE of three replicate experiments. At the bottom the growth phases are indicated proliferation (cells divide and expand simultaneously), expansion (expansion in absence of division), and mature (no more cell growth).

cells were of 2C DNA content; the remainder was of 4C content (Fig. 2). The amount of 8C cells observed was negligible, indicating nearly all cells of the leaf were engaged in mitotic cell cycle activity. From day 10 onwards, when average cell division rates were rapidly declining, the 2C fraction decreased to around 25%, typical for mature Arabidopsis leaves (Galbraith et al., 1991). At the same time, the 4C fraction increased to reach a maximum of 60% on day 14, thereafter declining to just below 50% and remaining constant after day 16. From day 14 onwards, 8C cells were observed, unequivocally demonstrating endoreduplication had started. The 8C fraction increased to about 30% on day 17 and, after day 22 an additional fraction (1%) of 16C cells was found. These data indicate that as average division rates declined, endoreduplication commenced and continued approximately until leaf growth was completed.

Microarray Analysis

For mRNA transcript analysis, we first analyzed an extensive time series with limited coverage (6,008 genes) to obtain high temporal resolution, allowing the selection of appropriate samples for subsequent genome-wide analysis. Based on the above characterization of the growth process, we selected 10 time points, covering the entire growth process and well into maturity. Day 9 was the earliest stage at which we could harvest sufficient material to extract RNA. From then until day 31, we harvested whole-leaf blades every second or third d from 30 to 250 plants grown on replicate plates. The majority (5,473 or approximately 91%) of genes on the array gave a signal above background levels for at least one sample. For a small number of genes, there were duplicate probes of different length and sequence. The expression pattern for these probes was nearly identical in each case (data not shown), demonstrating the reproducibility of the hybridization. Of the genes that gave a positive signal, 2,061 (34% of all genes on the array) were significantly modulated ($P \leq 0.001$). The largest difference in

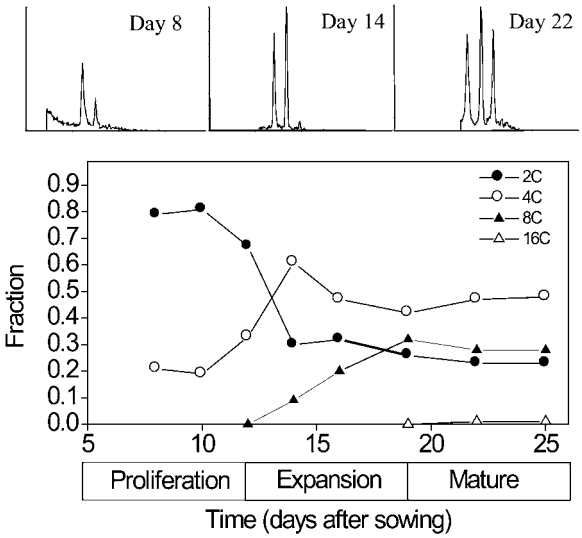

Figure 2. Flow cytometry analysis of nuclear DNA content. The results of a representative experiment are plotted. The bar at the bottom denotes the growth phases based on the kinematic analysis (Fig. 1).

expression level was 83-fold. For 114 genes, the difference was more than 10-fold, for 376 genes more than 5-fold, and for 1,642 genes more than 2-fold.

To examine the nature of transcriptional changes associated with progression through leaf development, we performed a support tree analysis on the significantly modulated genes ($P \leq 0.001$), essentially using the expression data as developmental fingerprints. The topology of the resulting tree exactly matches the growth phases identified by the kinematic analysis; the main branch subdivides growing and mature samples, and the second-order branching in the samples of growing leaves separates dividing from expanding leaves (Fig. 3). The higher support within the branch of the growing leaves suggests a higher degree of variation between these than between the mature samples.

Having demonstrated the clear correlation between gene expression and leaf development at the global level, we focused on the expression profiles of the genes. Quality threshold (QT) clustering divided the significantly modulated genes into 16 clusters of 20 or more genes that shared a similar pattern and one (cluster 17) containing the remaining genes (Fig. 4; Supplemental Table I). The two largest clusters (1 and 2), each containing approximately 20% of the differentially expressed genes, were specifically expressed in mature and growing tissue, respectively. Most of the remaining clusters also contained genes whose expression was closely related to the developmental stages: proliferation (clusters 9, 12, and 15), expansion (clusters 4, 6, 11, 13, and 14), growing (proliferation + expansions; cluster 7), and mature tissues (clusters 3 and 8). Clusters 5 and 10 are specifically up- or down-regulated during the first stages of maturity, explaining the branching in the mature half of the support tree (Fig. 3). Overall, these expression patterns closely matched the kinematically determined growth parameters and suggest that the majority of differences occur between the three main stages: proliferation, expansion, and mature.

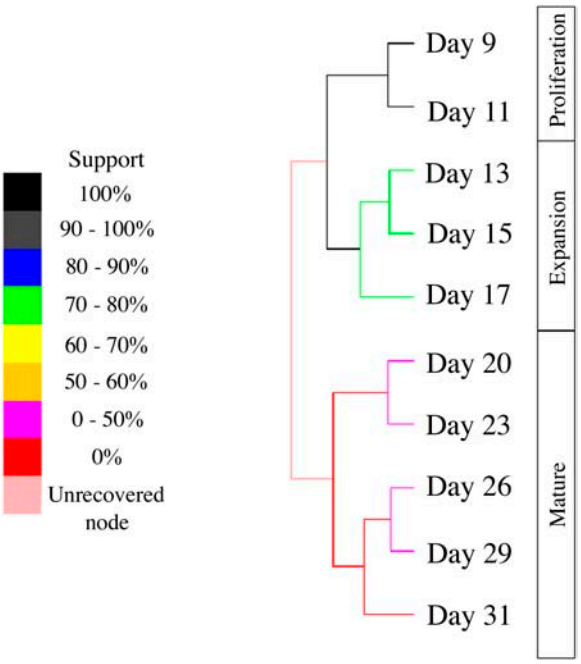

Figure 3. Clustering of developmental time series support tree analysis (Graur and Li, 2000) of the expression data of 2,061 significantly modulated genes ($P \leq 0.001$) comparing time points. The level of support for each branch of the tree is color coded as indicated in the legend; the level of support indicates the confidence for each branch. The bar on the right denotes the growth phases based on the kinematic analysis (Fig. 1).

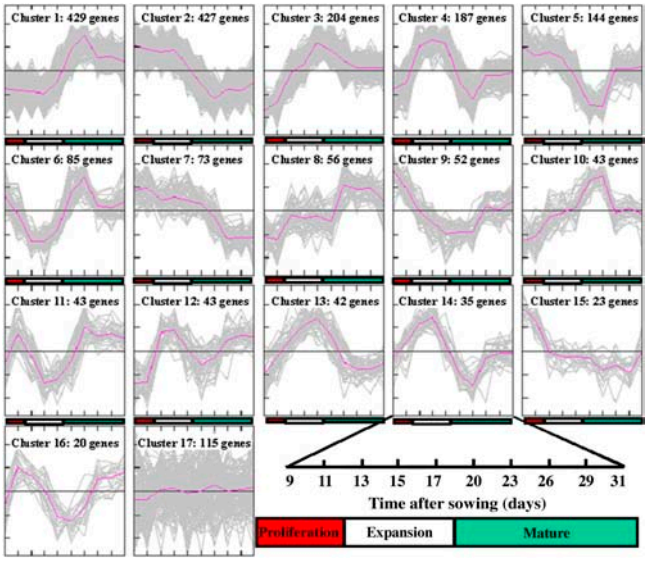

Figure 4. Clustering of gene expression profiles by QT-Clust analysis (Heyer et al., 1999) of the expression profiles of 2,061 significantly modulated genes ($P \leq 0.001$). Cluster number and size are indicated. The abscissa denotes the time after sowing, which is enlarged for Cluster 14, and the ordinate indicates normalized and median-centered expression levels. The colored bar shows the corresponding growth phases based on the kinematic analysis (Fig. 1).

Core Cell Cycle Genes

The above clearly shows that global gene expression patterns during leaf development reflected the transitions in cell cycle mode (proliferation, endoreduplication, and off). Therefore, we focused our attention on cell cycle genes. To establish the role of these genes in mitotic and endoreduplication cycli, we performed a second set of analyses using the Affymetrix ATH1 Genechip on leaf blades from of 9-, 15-, and 22-d-old seedlings, representing the proliferating, endoreduplicating, and mature stages, respectively. Seventeen cell cycle genes were present on both the cDNA and Affymetrix arrays, and the obtained expression patterns closely matched between both platforms (data not shown). Unfortunately, 8 cell cycle genes (CDKB1;1, CYCB1;2, CYCB2;3, CYCD7;1, CYCT1;1, CYCT1;2, DEL1, and KRP6) were not present on the Affymetrix arrays.

Of the genes on the array, 10 (CDKD;1, CYCA2;1, CYCA2;4, CYCA3;3, CYCD4;2, CYCP3;1, CYCP3;2, CYCP4;2, CYCP4;3, and SDS) were not detected above background levels in any of the samples (based on the Affymetrix present calls; $P < 0.04$). For the remaining 62 genes, 2 main functional classes can be distinguished (Table I):

(1) Constitutive expression. No significant variation ($P > 0.05$) or less than 1.5-fold difference between minimum and maximum expression values. This class encompassed all A-, C-, D-, E-, and F-type CDKs (except for CDKD;2); CKS1; all E2Fs; DPs and RB; most KRPs; all expressed H, J18, L, P, and T-type cyclins; and a relatively small number of A, C, and D-type cyclins.
(2) Expression during proliferation. Significant ($P \leq 0.05$) variation and over 1.5-fold differences between minimum and maximum expression levels.
(3) The highest expression occurred at day 9 and the expression level on day 15 was closer to that of day 22 than of day 9. In this class, we found the B-type CDKs; CKS2; the majority of A, B, and D-type cyclins; DEL2 and 3; and WEE1.

The expression of the cell cycle inhibitor KRP1 and CDKD;2 increased with time, whereas CYCP4;1 was the only cell cycle gene that is expressed highly in both proliferating and endoreduplicating leaves.

Recently, Birnbaum et al. (2003) published a gene expression map for the Arabidopsis root tip. They obtained root samples that roughly correspond to division zone, elongation zone, and mature region. To test the generality of the expression patterns obtained in the leaf, we analyzed the expression of the cell cycle genes in their data using the same criteria as for the leaf samples.

Of the 72 cell cycle genes on the array, the expression patterns of 37 were identical (Table II). Specifically, A-, B-, C-, E-, and F-type CDKs; CKS1 and 2; B- and T-type cyclins behaved similar in both systems. In contrast the D-type CDKs that had diverse expression patterns in the leaf were all classified proliferative in the root tip. Within the A- and D-type cyclin families, specific members shifted from constitutive to proliferative expression patterns and vice versa. Interestingly, both E2Fa and E2Fb were constitutively expressed during leaf development but were specifically expressed in the meristem of the root tip. Finally, in addition to KRP1, 2 other inhibitors, KRP2 and 5, also increased in expression levels as cell cycle activity decreases in the root tip, indicating that these inhibitory proteins form a third functional class of cell cycle genes next to proliferation and constitutively expressed genes.

Besides organ-specific features, there is a high degree of similarity between the two datasets. Most CDKs and E2F pathway genes are constitutively expressed, whereas the activating regulators, cyclins A, B, and D, are mainly expressed in proliferating tissue. Inhibitory KRPs are either constitutively expressed or at increasing levels as cells progress from mitosis to endoreduplication and shutdown of the cell cycle.

Proliferation Genes

These observations of known cell cycle regulatory genes demonstrate that expression data from developing leaves and the root tips from Birnbaum et al. (2003) form a good basis to identify proliferation genes in growing organs. Another approach for finding such genes was used by Menges et al. (2003). Also using Affymetrix microarrays, they identified a set of 1,081 transcripts whose expression was cell cycle phase dependent in synchronized Arabidopsis cell cultures.

Table I. *Expression patterns of core cell cycle genes during leaf development*

Definition of the categories was as follows. Not detected are those genes for which no P calls were obtained in any of the samples. (P calls are based on significance of differences between match and mismatch probes $P \leq 0.04$.) Constitutive are those genes for which no significant difference between three developmental stages (proliferation, endoreduplication, and mature) could be detected ($P \leq 0.04$) or where the fold difference between highest and lowest average value was smaller than 1.5. Proliferation pattern are genes that had significant differences between the 3 stages, where the highest values were in the proliferation and the difference between expression in proliferation and endoreduplication was larger than between endoreduplication and mature. The category ''Proliferation and endocycle'' encompasses genes significantly different between samples with similar levels in proliferating and endoreduplicating tissues. Increasing are those genes for which the expression levels significantly increase in subsequent developmental stages.

Expression Pattern	Genes
Constitutive	CDKA;1, CDKC;1, CDKC;2, CDKD;3, CDKE;1, CDKF;1, CKS1, CYCA2;2, CYCA3;4, CYCC1;1, CYCD2;1, CYCD5;1, CYCH;1, CYCJ18, CYCL1;1, CYCP1;1, CYCP2;1, CYCT1;3, CYCT1;4, CYCT1;5, E2Fa, E2Fb, E2Fc, DPa, DPb, RBR, KRP2, KRP3, KRP4, KRP5, KRP7
Proliferation	CDKB1;2, CDKB2;1, CDKB2;2, CKS2, CYCA1;1, CYCA1;2, CYCA2;3, CYCA3;1, CYCA3;2, CYCB1;1, CYCB1;3, CYCB1;4, CYCB1;5, CYCB2;1, CYCB2;2, CYCB2;4, CYCB2;5, CYCB3;1, CYCC1;2, CYCD1;1, CYCD3;1, CYCD3;2, CYCD3;3, CYCD4;1, CYCD6;1, DEL2, DEL3, WEE1
Proliferation and endocycle	CYCP4;1
Increasing	CDKD;2, KRP1
Not detected	CDKD;1, CYCA2;1, CYCA2;4, CYCA3;3, CYCD4;2, CYCP3;1, CYCP3;2, CYCP4;2, CYCP4;3, SDS
Not on array	CDKB1;1, CYCB1;2, CYCB2;3, CYCD7;1, CYCT1;1, CYCT1;2, DEL1, KRP6

Table II. *Expression patterns of core cell cycle genes in different root zones*

Expression categories were determined using the same criteria as for Table I. Indicated in bold are those genes that had a similar pattern in both experiments.

Expression Pattern	Genes
Constitutive	**CDKA;1**, **CDKC;1**, **CDKC;2**, **CDKE;1**, **CDKF;1**, **CKS1**, CYCA1;2, CYCA2;1, CYCA2;3, CYCA3;1, CYCA3;2, CYCB1;4, **CYCC1;1**, CYCC1;2, CYCD1;1, **CYCD2;1**, CYCD3;1, **CYCD5;1**, **CYCL1;1**, **CYCP1;1**, CYCP2;1, CYCP3;1, CYCP4;1, **CYCT1;3**, **CYCT1;4**, **CYCT1;5**, **E2Fc, DPb**, **RBR**, DEL2, SDS, WEE1, **KRP3**, **KRP7**
Proliferation	**CDKB1;2**, **CDKB2;1**, **CDKB2;2**, CDKD;1, CDKD;2, CDKD;3, **CKS2**, **CYCA1;1**, CYCA2;2, CYCA2;4, CYCA3;3, **CYCB1;1**, **CYCB1;3**, **CYCB1;5**, **CYCB2;1**, **CYCB2;2**, **CYCB2;4**, **CYCB2;5**, CYCB3;1, **CYCD3;2**, **CYCD3;3**, CYCH;1, CYCP3;2, E2Fa, E2Fb
Proliferation and endocycle	CYCD4;1, DEL3, DPa
Increasing	CYCJ18, CYCP4;2, CYCP4;3, **KRP1**, KRP2, KRP5
Not detected	**CYCA3;3**, **CYCD4;2**, CYCD6;1, CYCP2;1, KRP4
Not on array	**CDKB1;1**, **CYCB1;2**, **CYCB2;3**, **CYCD7;1**, **CYCT1;1**, **CYCT1;2**, **DEL1**, **KRP6**

Individually, each of these three datasets presumably also contains a significant number that do not play a role in cell cycle-related events, but for which the expression level covaries. To filter out such genes, we decided to combine the information contained in these three datasets to identify a high-confidence set of proliferation-specific genes. To have an idea about the enrichment in potential cell cycle genes, we used the core cell cycle genes as a reference. Using the same criteria as for the core cell cycle genes, 2,066 genes were specifically expressed in proliferating leaves (Supplemental Table II), among which were 28 (1.4%) core cell cycle genes (Table I; Fig. 5). In the root system, a similar number of transcripts (2,148) were specifically expressed in the apical region (Supplemental Table II; Fig. 5), among which were 25 (1.2%) core cell cycle genes (Table 1). Finally, there were 19 core cell cycle genes among the 1,081 genes with cell cycle-dependent expression patterns (1.8%; Fig. 5). While, overall, most genes are present in only one of the datasets, most core cell cycle genes are shared by at least two sets (Fig. 5). Consequently, the percentage of core cell cycle genes is similar to that of the array as a whole when the genes present in only one of these three sets are considered. Thus, although the individual datasets are clearly enriched in cell cycle genes, this is primarily due to the genes that are in the overlap between two or more experiments. Similarly, the fraction of core cell cycle genes shared between all three data sets is again significantly enriched compared to that in the overlaps between only two datasets. The highest percentage of core cell cycle genes (8%) is found in the overlap between all three sets. Therefore, we consider the 131 genes shared by all 3 datasets as high-confidence proliferation genes. Interestingly, 74% of these genes are expressed in M phase, while this is only 33% in the whole set of cell cycle-regulated genes. Consistently, all known cell cycle genes present in this set are M phase specific (CYCA1;1, 7 B-type cyclins, two B-type CDKs, 2 mitotic checkpoint genes, and 4 anaphase-promoting complex [APC] components). The presence of 18 cytoskeleton genes in this set nearly all with M phase-specific expression is clearly associated with mitotic spindle and phragmoplast formation. Notable exceptions that are not M phase specific: 4 histones (S phase) and 5 cell wall synthesis-related genes (G1 phase), and an minichromosome maintenance (MCM) gene (G1 phase). The remaining genes encompass the PINHEAD/ZWILLE gene; 3 high mobility group and one SCARECROW family transcription factor; 2 chromatin remodeling genes (topoisomerase2 [TOP2] and methyltransferase1 [MET1]); 3 nodulation-related genes; the KNOLLE syntaxin gene; an unnamed gene with homology to the kinetochore assembly protein CENPCA from maize (Dawe et al., 1999); 5 protein kinases, among which is BRI1, a brassinosteroid-response regulator (Clouse et al., 1996); 2 thaumatin-like genes; a Rac GTP-binding protein; 2 electron transport genes; a ribosomal protein; 7 genes with a transport function; and a total of 51 unknown genes.

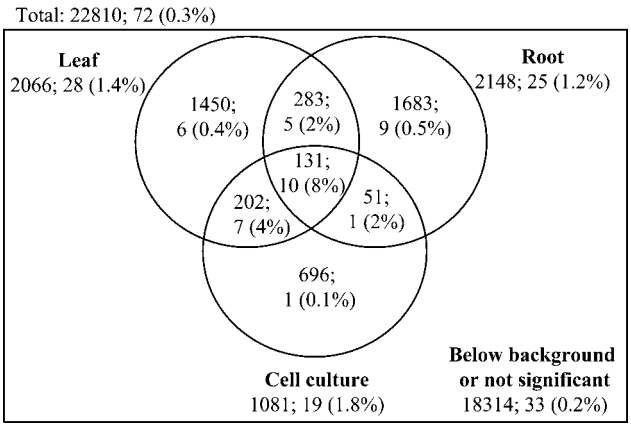

Figure 5. Enrichment of cell cycle genes among those that show modulation during leaf development, in different developmental zones of the root tip and in synchronized cell cultures. The distribution of all genes on the array is compared to that of core cell cycle genes defined by Vandepoele et al. (2002) and Wang et al. (2004). Numbers subsequently indicate total number of genes in a population; the number of core cell cycle genes and the percentage of genes that are core cell cycle genes (in parentheses).

DISCUSSION

Experimental System

To study in detail the molecular basis of cell cycle regulation during organ development, we chose the first leaf pair of Arabidopsis because its growth phases are separated in time rather than spatially, as with root and shoot apices. Kinematic and flow cytometry analyses were used to determine cell cycle, cell expansion, and differentiation parameters during the progression from fully meristematic to mature. Implicitly, we treated the leaf blade as uniform cell material for the flow cytometry and transcriptome analyses. However, earlier research revealed significant differences between leaf tissues. It was shown that division stops first in the epidermis and last in the vascular tissue (Donnelly et al., 1999) and in a tip-to-base gradient (Pyke et al., 1991; Donnelly et al., 1999). Presumably, the downward part of the division rate curve (Fig. 1) is therefore primarily caused by a gradual decrease in the fraction of dividing cells rather than an increase in cell cycle duration.

Earlier, we showed that stomatal development in the base of the leaf was only 1 to 2 d behind on the tip (De Veylder et al., 2001). Differences in cell expansion have also been documented: Dorsoventral cell expansion generally continues longer than area expansion (which we measured), and the mesophyll expands more in the dorso-ventral direction than the epidermis (Pyke et al., 1991). Finally, it has been shown that gene expression is often cell type-specific (e.g. CYCA2;1 expression is restricted to the vascular tissue; Burssens et al., 2000). Nevertheless, we report that, despite these well-documented sources of heterogeneity between tissues, whole-leaf measurements with flow cytometry and microarray analysis correspond surprisingly well with the parameters determined kinematically. This implies that, relative to the time steps that we used, the development of the majority of cells in the leaf parallels that of the epidermis. This is an important result and implies that whole-leaf samples that are easily obtained can be used for investigating the molecular basis of the growth processes quantified in the epidermis.

Transcript Profiling

We first used a cDNA microarray spotted with expressed sequence tags of 20% to 25% of the genes encoded in the Arabidopsis genome. Nearly 90% of these genes were detected and more than one-third were significantly modulated over time. When we used the genome-wide Affymetrix arrays on a limited number of time points, 71% of the genes on the array were expressed at significant levels in at least one of the time points. Thus, a large portion of genes is expressed in the course of the growth process, underscoring the value of the model system. Moreover, the overall pattern of gene expression reflects the presence of the three growth phases, which correspond closely to proliferation, expansion/endoreduplication, and maturity (compare with Figs. 1 and 3). The functions of many of the genes encoded by the Arabidopsis genome are still largely hypothetical, and some 30% do not even show homology that enables classification. Therefore, the data we obtained, along with other available datasets, can serve as a catalog of gene expression from which the putative function of a particular gene can be inferred and inspire further research. For this reason, all data have been submitted to Array-Express (http://www.ebi.ac.uk/arrayexpress/; E-MEXP-144).

Cell Cycle Transitions

Here we used the transcriptome data in concert with the kinematic and flow cytometry data to unravel the cell cycle regulation in the context of a growing multicellular organ. The obtained data show that cell cycle activity reflects the three different developmental stages. During the proliferation phase, cells execute a mitotic cell cycle where cells go through successive cycles of G1-S-G2-M phases. This is illustrated by the presence of cells with a 2C (G1) and 4C (G2) DNA content during this phase (Fig. 2). When proliferation stops, endoreduplication cycles, comprising of successive phases of G and S phase in absence of karyokinesis, commenced immediately as can be seen by the appearance of 8C and 16C (Fig. 2). The last transition occurred when all cell cycle activity finished, as evidenced by a stable DNA distribution and cell number (Figs. 1 and 2). This coincided at the whole-organ level with the blade reaching its final size and ending its growth. The process of endoreduplication occurs in many plants (D'Amato, 1952) including Arabidopsis, where it happens in most organs (Galbraith et al., 1991). In the epidermis of Arabidopsis leaves, a linear correlation between the level of endoreduplication and cell size was observed (Melaragno et al., 1993), suggesting that endoreduplication drives expansion. In maize endosperm and tomato fruit, endoreduplication has also been shown to occur when mitosis is completed (Grafi and Larkins, 1995; Joubès et al., 1999), suggesting that endoreduplication results from an inhibition of the G2 to M transition of the cell cycle while S phase entry continues. Our observations that the onset of endoreduplication in Arabidopsis leaves occurs when cell division stops and that the majority of proliferation specific genes are M phase-specific strongly supports this model. Roots also endoreduplicate (Beemster et al., 2002), and the similarity of cell cycle gene expression in both organs demonstrated here suggests that it also takes place in the expansion phase.

Cell Cycle Gene Expression

It is well established that transcriptional regulation plays an important role in cell cycle control. Therefore, it was not surprising that the transitions between cell

cycle modes were associated with distinct patterns of cell cycle gene expression. Of the 80 designated core cell cycle genes, 72 were present on the Affymetrix ATH1 array and of those, 63 were detected above background levels in at least 1 of the 3 developmental stages. This nearly 90% detection is well above average, underlining that cell cycle genes can be effectively studied in this experimental system by means of microarray analysis.

To enable discrimination between organ-specific and more general expression patterns, we compared our data with those recently obtained with the same platform from corresponding developmental stages in the root tip by Birnbaum et al. (2003).

The data indicate that, at the expression level, members of each gene family behave very similarly during vegetative organ development and three categories of expression profiles can be distinguished:

(1) Constitutive expression: CDKs (with the exception of the B-type) and the downstream transcriptional control machinery formed by E2Fs, DPs, and RB, determine the potential to divide (Hemerly et al., 1993). Moreover, many D-type cyclins are also constitutively expressed ensuring a constitutive drive into S phase.
(2) Proliferation specific: most of the activating A- and B-type cyclin partners that determine the activity of the CDKs and downstream transcription complexes.
(3) Increasing during development: inhibitory proteins that will inhibit the activity of CDK/cyclin complexes, thereby inhibiting cell cycle progression.

Based on these three categories, a model can be proposed that explains the transitions between successive cell cycle modes. First, the repression of M phase activators in concert with the continuing activation of S phase activity by the CDKA/CYCD complexes determines the transition from proliferation to endoreduplication. Second, the increasing levels of KRP proteins reach a threshold that terminates endoreduplication when cells become mature.

This model is also supported by our recent findings that the balance between proliferation and endoreduplication can be effectively controlled by varying the relative S and M phase activity by altering the activity of E2Fa/DPa and CDKB1;1, respectively (Boudolf et al., 2004).

A little bit less straightforward is the role of the inhibitory proteins. Plants overproducing KRP2 indeed endoreduplicate less (De Veylder et al., 2001) as would be predicted from the above model. However, division rates are also reduced, implying that these inhibitory proteins not only affect exit from endoreduplication but also the rates at which the cycle progresses as such.

We found little evidence for cell cycle genes that were specifically expressed during endoreduplication. Based on recent work on *Medicago truncatula* and maize, the cell cycle switch CCS52a and WEE1 proteins, respectively, were proposed as positive regulators of the endocycle and such an expression pattern (Cebolla et al., 1999; Sun et al., 1999; Vinardell et al., 2003). Both types of proteins inhibit M phase-specific CDK activity. For the model presented here, it would implicate a double-control mechanism to control down-regulation of M phase activity both at the transcriptional and protein levels.

However, none of the 3 Arabidopsis, CCS52 genes (CCS52a1, CCS52a2, and CCS52b), or WEE1 is expressed specifically in endoreduplicating tissue (Supplemental Table II). This indicates that species-specific differences may exist in the function of these genes and that in Arabidopsis there is no evidence that inhibition of M phase CDK activity at the protein level is necessary for endoreduplication.

Functional analysis will be necessary to resolve the function of these genes in Arabidopsis.

Proliferation Genes

An important aspect of this paper is the genome-wide search for proliferation genes. To enrich for genes that are involved in the regulation of proliferation, rather than coincidentally expressed in proliferating leaf tissue, we combined the expression data from three datasets performed on the same Affymetrix platform. As a reference, we used a set of known core cell cycle genes. Of the 22,810 genes on the chips, 3,800 genes were differentially expressed in at least 1 of the experiments (i.e. 16.7%). However, the majority of these genes are present in only one of the datasets (Fig. 5). A total of 667 genes are present in at least 2 and only 131 in all 3 datasets.

When the same analysis is performed for core cell cycle genes, 39 of the 72 genes on the array were significant in at least 1 set (i.e. 54%), while the majority of significant genes are present in more than 1 dataset, the highest number of these genes being present in all 3 sets (Fig. 5). These data illustrate that a large part of cell cycle-related genes on the array could be identified this way. However, at least the same number of core cell cycle genes were not picked up this way, showing this dataset is by no means a complete inventory of cell cycle related genes.

However, it is clear that a high portion of core cell cycle genes is present in the intersection of at least two of the three datasets, and therefore it seems likely that among the genes in these categories a high percentage is somehow associated with the regulation of the cell cycle.

Here we focused on the 131 genes that are present in all 3 datasets and dubbed them high-confidence proliferation genes. The strong enrichment in M phase-specific genes is consistent with the presented model based on core cell cycle genes, suggesting that genes to be switched off at the transition between proliferation and endoreduplication would be M phase activators.

Based on these experiments, it proved impossible to effectively select for new cell cycle genes that are constitutively expressed or induced as cells progress through successive developmental stages. Nevertheless, the identification of well over 50 unknown genes or unknown relationships to cell cycle regulation will be a fruitful basis for expanding our knowledge on cell cycle regulation, particularly in the context of growing multicellular organs.

MATERIALS AND METHODS

Culture Conditions

Seeds of Arabidopsis (*Arabidopsis thaliana*) ecotype Columbia were plated on agar-solidified culture medium (1× Murashige and Skoog [Duchefa, Haarlem, The Netherlands], 0.5 g/L MES, pH 6.0, 1 g/L Suc, and 0.6% plant tissue culture agar [LabM, Bury, UK]) in 150- × 25-mm round petri dishes (type Integrid, Falcon, Franklin Lakes, NJ). These plates were placed horizontally in a growth chamber kept at 22°C with a 16-h photoperiod of 65 μE m^{-2} s^{-1} photosynthetically active radiation supplied by white fluorescent tubes.

Kinematic Analysis

Kinematic analysis was performed as described earlier (De Veylder et al., 2001) on the abaxial epidermis of leaf 1 and 2 blades harvested daily from days 5 to 21 in 3 replicate experiments.

Flow Cytometry Analysis

At 8 and 10 d after sowing, primordia of leaves 1 and 2 were dissected from the shoots of 250 and 125 plants under a binocular. On days 12, 14, 16, 19, 22, and 25, the blades of approximately 30 plants were dissected by eye. The tissue was chopped with a razorblade in 200 to 400 μL of buffer (45 mM $MgCl_2$, 30 mM sodium citrate, 20 mM MOPS, pH 7, and 1% Triton X-100; Galbraith et al., 1991), filtered over a 30-μm mesh, and 1 μL of 1 mg/mL of 4,6-diamidino-2-phenylindole was added. The nuclear DNA content distribution was analyzed with a BRYTE HS flow-cytometer (Bio-Rad, Hercules, CA).

Microarray Analysis

Primordia of leaves 1 and 2 were dissected from the shoots of 250 plants under a binocular at 9 d after sowing. On days 11, 13, 15, 17, 20, 23, 26, 29, and 31, the blades of approximately 125 (day 11), 60 (day 13), and 30 (day 15 onwards) plants were dissected by eye. Total RNA was isolated using Trizol Reagent (Life Technologies, Gaithersburg, MD).

We used microarrays spotted in duplicate with 6,008 Arabidopsis genes derived from the unigen clone collection from Incyte (Arabidopsis Gem I, distributed originally by Incyte, now available through Open Biosystems, Huntsville, AL) and 520 positive and negative controls (for details, see www.microarrays.be/service/currently available arrays). For each sample, a minimum of 5 μg total RNA was amplified as described previously (Puskás et al., 2002).

Hybridization and posthybridization washing were performed at 45°C using an automated slide processor, the program for which can be downloaded from www.microarrays.be/service.htm. Arrays were scanned at 532 and 635 nm using a Generation III scanner (Amersham BioSciences, Little Chalfont, UK). Image analysis was performed with ArrayVision (Imaging Research, St. Catherines, Ontario, Canada).

As a common reference used for reciprocal labeling, we used a mixture of equal amounts of RNA derived from day 9 to day 23 samples. Spot intensities were measured as artifact removed total intensities without correction for background. For 24 negative control spots containing a *Bacillus subtilis*-specific cDNA and 6,008 Arabidopsis spots, we first performed within-slide normalization by plotting for each single slide an MA plot (Yang et al., 2002). The Lowess normalization (f = 0.2) was applied to correct for dye intensity differences. Based on the 96 adjusted $\log_2 R$ and $\log_2 G$ signal intensities of the negative control spots, the background median and the 95th percentile were calculated. The 95th percentile was defined as signal threshold.

The adjusted signal intensities were compared to the signal threshold; 535 genes were uniformly below the signal threshold and removed. The values of the remaining 5,473 genes were set to the background median whenever they were below this signal threshold.

To normalize between slides and to identify differentially expressed genes between the different time points, we performed 2 sequential ANOVAs, proposed by Wolfinger et al. (2001) as follows: (1) let y_{iklm} be log2 of the Lowess-transformed spot measurement from gene i (i = 1–5,473); we first applied a linear normalization ANOVA model of the form $y_{iklm} = \mu + A_k + (ADR)_{klm} + \varepsilon_{iklm}$ to estimate global variation of the collection of i selected cDNA fragments, where μ is the sample mean, A_k the effect of the *k*th Array (k = 1–20), $(ADR)_{klm}$ the channel effect (Array*Dye; $AD)_{kl}$ for the mth replication of the total collection of cDNA fragments on a slide (m = 1–2), and ε_{iklm} the stochastic error; (2) we then averaged the residuals from this model coming from duplicate spots on each slide and subjected them to 5,473 gene-specific models of the form $r_{ijklmn} = \mu + (GD)_{il} + (GS)_{ij} + (GST)_{ijn} + (GA)_{ik} + \gamma_{ijklmn}$, partitioning gene-specific variation into Gene*Dye effects ($[GD]_{il}$, l = 1–2), Gene*Sample effects ($[GS]_{ij}$, j = 1–2; reference and test sample), Gene*Sample*Time effects ($[GST]_{ijn}$; n either the number of harvested samples [n = 1–10] or to the number of paired observations [we took day 9 and 11, 13 and 15, 17 and 20, 23 and 26, and 29 and 31 as independent replications; n = 1–5]), Gene*Array spot effects ($[GA]_{ik}$, k = 1–20) and error γ_{ijklmn}. We made standard assumptions about the preceding linear models. In particular, A_k, $(ADR)_{klm}$, $(GA)_{ik}$, ε_{iklm}, and γ_{ijklmn} were considered as random effects, while $(GD)_{il}$, $(GS)_{ij}$, and $(GST)_{ijn}$ were assumed to be fixed.

As a measure of variability in expression levels between time points, we calculated for each gene the Wald statistic for the parameter $[GST]_{ijn}$, where n was set equal to five [the number of grouped observations]). The Wald statistic was tested against the χ^2-distribution. The P-value cutoff was set at 0.001 and no adjustments for multiple testing were made. We used the restricted maximum likelihood (REML) procedure of Genstat (Release 6.1 for Windows, VSN International, Hemel Hempstead, UK) for normalization, gene model fits, and to test the Wald statistic.

For all significantly modulated expression profiles, the values of $(GST)_{ijn}$, where n equals the 10 samples harvested, were used to calculate the expression ratios relative to the reference sample. These ratios were mean centered, normalized, and clustered. We clustered the data for the different samples using support tree (Graur and Li, 2000) and the genes in function of time with QT-Clust (Heyer et al., 1999) implemented in the MeV software (Saeed et al., 2003). The support tree was based on 1,000 jackknife permutations using Euclidian distance as a criterion. For the QT-Clust we used Pearson correlation as a distance measure, a d-value of 0.4, and a minimal cluster size of 20 genes.

Genome-Wide Analysis on the Affymetrix ATH1 Array

In 3 independent experiments, plants were harvested on day 9, 15, and 22 and RNA was isolated as outlined before. The RNA was labeled and hybridized, whereafter the slides were scanned according to standard Affymetrix procedures. The resulting CEL files were imported into the statistical package R (www.r-project.org), in which the Bio-conductor (www.bioconductor.org) libraries Affy, Stats, and FactDesign were loaded. Using this the quality of the slides was checked, the data were normalized, and expression values were calculated using the robust multichip average procedure, and present calls were calculated and ANOVA analysis of variance was performed on the log2 expression values to determine the significance of difference in expression between sampling times.

The same procedure was followed with the two independent sets of CEL files from three zones Birnbaum et al. (2003), where the difference in expression between these zones was evaluated.

Received September 22, 2004; revised December 23, 2004; accepted December 23, 2004; published April 29, 2005.

LITERATURE CITED

Beemster GTS, Baskin TI (1998) Analysis of cell division and elongation underlying the developmental acceleration of root growth in *Arabidopsis thaliana.* Plant Physiol **116:** 515–526

Beemster GTS, De Vusser K, De Tavernier E, De Bock K, Inzé D (2002) Variation in growth rate between *Arabidopsis thaliana* ecotypes is correlated with cell division and A-type cyclin dependent kinase activity. Plant Physiol **129:** 854–864

Beemster GTS, Fiorani F, Inzé D (2003) Cell cycle: the key to plant growth control? Trends Plant Sci **8:** 154–158

Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, Benfey PN (2003) A gene expression map of the Arabidopsis root. Science **302:** 1956–1960

Boudolf V, Vlieghe K, Beemster GTS, Magyar Z, Torres-Acosta JA, Maes S, Van Der Schueren E, Inze D, De Veylder L (2004) The plant-specific cyclin-dependent kinase CDKB1;1 and transcription factor E2Fa-DPa control the balance of mitotically dividing and endoreduplicating cells in Arabidopsis. Plant Cell **16:** 2683–2692

Burssens S, de Almeida Engler J, Beeckman T, Richard C, Shaul O, Ferreira P, Van Montagu M, Inzé D (2000) Developmental expression of the *Arabidopsis thaliana* CycA2;1 gene. Planta **211:** 623–631

Cebolla A, Vinardell JM, Kiss E, Oláh B, Roudier F, Kondorosi A, Kondorosi E (1999) The mitotic inhibitor *ccs52* is required for endoreduplication and ploidy-dependent cell enlargement in plants. EMBO J **18:** 4476–4484

Clouse SD, Langford M, McMorris TC (1996) A brassinosteroid-insensitive mutant in *Arabidopsis thaliana* exhibits multiple defects in growth and development. Plant Physiol **111:** 671–678

D'Amato F (1952) Polyploidy in the differentiation and function of tissues and cells in plants. Caryologia **4:** 311–358

Dawe RK, Reed LM, Yu HG, Muszynski MG, Hiatt EN (1999) A maize homolog of mammalian CENPC is a constitutive component of the inner kinetochore. Plant Cell **11:** 1227–1238

De Veylder L, Beeckman T, Beemster GTS, Krols L, Terras F, Landrieu I, Van der Schueren E, Maes S, Naudts M, Inzé D (2001) Functional analysis of cyclin-dependent kinase inhibitors of Arabidopsis. Plant Cell **13:** 1653–1668

Donnelly PM, Bonetta D, Tsukaya H, Dengler RE, Dengler NG (1999) Cell cycling and cell enlargement in developing leaves of Arabidopsis. Dev Biol **215:** 407–419

Erickson RO, Sax KB (1956) Rates of cell division and cell elongation in the growth of the primary root of *Zea mays*. Proc Am Philos Soc **100:** 499–514

Galbraith DW, Harkins KR, Knapp S (1991) Systemic endopolyploidy in *Arabidopsis thaliana*. Plant Physiol **96:** 985–989

Gandar PW (1980) The analysis of growth and cell production in root apices. Bot Gaz **141:** 131–138

Goodwin RH, Stepka W (1945) Growth and differentiation in the root tip of *Phleum pratense*. Am J Bot **32:** 36–46

Grafi G, Larkins BA (1995) Endoreduplication in maize endosperm: involvement of M phase-promoting factor inhibition and induction of S phase-related kinases. Science **269:** 1262–1264

Granier C, Inzé D, Tardieu F (2000) Spatial distribution of cell division rate can be deduced from that of $p34^{cdc2}$ kinase activity in maize leaves grown at contrasting temperatures and soil water conditions. Plant Physiol **124:** 1393–1402

Graur D, Li W-H (2000) Fundamentals of Molecular Evolution, Ed 2. Sinauer Associates, Sunderland, MA

Hemerly AS, Ferreira P, de Almeida-Engler J, Van Montagu M, Engler G, Inzé D (1993) *cdc2a* expression in Arabidopsis is linked with competence for cell division. Plant Cell **5:** 1711–1723

Heyer LJ, Kruglyak S, Yooseph S (1999) Exploring expression data: identification and analysis of coexpressed genes. Genome Res **9:** 1106–1115

Joubès J, Phan T-H, Just D, Rothan C, Bergounioux C, Raymond P, Chevallier C (1999) Molecular and biochemical characterisation of the involvement of cyclin-dependent kinase A during the earlydevelopment of tomato fruit. Plant Physiol **121:** 857–869

Melaragno JE, Mehrotra B, Coleman AW (1993) Relationship between endopolyploidy and cell size in epidermal tissue of Arabidopsis. Plant Cell **5:** 1661–1668

Menges M, Hennig L, Gruissem W, Murray JAH (2003) Genome-wide gene expression in an Arabidopsis cell suspension. Plant Mol Biol **53:** 423–442

Puskás LG, Zvara Á, Hackler JL, Van Hummelen P (2002) RNA amplification results in reproducible microarray data with slight ratio bias. Biotechniques **32:** 1330–1341

Pyke KA, Marrison JL, Leech RM (1991) Temporal and spatial development of the cells of the expanding first leaf of *Arabidopsis thaliana* (L.) Heynh. J Exp Bot **42:** 1407–1416

Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, et al (2003) TM4: a free, open-source system for microarray data management and analysis. Biotechniques **34:** 374–378

Silk WK (1984) Quantitative descriptions of development. Annu Rev Plant Physiol **35:** 479–518

Sun Y, Dilkes BP, Zhang H, Dante RA, Newton P, Lowe KS, Jung R, Gordon-Kamm WJ, Larkins BA (1999) Characterization of maize (*Zea mays* L.) Wee1 and its activity in developing endosperm. Proc Am Philos Soc **96:** 4180–4185

Tsukaya H (2002) Interpretation of mutants in leaf morphology: genetic evidence for a compensatory system in leaf morphogenesis that provides a new link between cell and organismal theories. Int Rev Cytol **217:** 1–39

Vandepoele K, Raes J, De Veylder L, Rouzé P, Rombouts S, Inzé D (2002) Genome-wide analysis of core cell cycle genes in Arabidopsis. Plant Cell **14:** 903–916

Vinardell JM, Fedorova E, Cebolla A, Kevei Z, Horvath G, Kelemen Z, Tarayre S, Roudier F, Mergaert P, Kondorosi A, et al (2003) Endoreduplication mediated by the anaphase-promoting complex activator CCS52A is required for symbiotic cell differentiation in *Medicago truncatula* nodules. Plant Cell **15:** 2093–2105

Wang G, Kong H, Sun Y, Zhang X, Zhang W, Altman N, dePamphilis CW, Ma H (2004) Genome-wide analysis of the cyclin family in Arabidopsis and comparative phylogenetic analysis of plant cyclin-like proteins. Plant Physiol **135:** 1084–1099

Webster PL, Macleod RD (1980) Characteristics of root apical meristem cell population kinetics: a review of analyses and concepts. Environ Exp Bot **20:** 335–358

Wolfinger RD, Gibson G, Wolfinger ED, Bennett L, Hamadeh H, Bushel P, Afshari C, Paules RS (2001) Assessing gene significance from cDNA microarray expression data via mixed models. J Comput Biol **8:** 625–637

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. Nucleic Acids Res **30:** e15