Plant Cell Monogr (2)
A. Mujib · J. Šamaj: Somatic Embryogenesis
DOI 10.1007/7089_039/Published online: 20 October 2005

Genome-Wide Expression Analysis of Genes Involved in Somatic Embryogenesis

W. Tang (✉) · R. J. Newton

Department of Biology, East Carolina University, Howell Science Complex, Greenville, NC 27858-4353, USA
tangw@mail.ecu.edu

Abstract Genome-wide expression analysis is an important tool for identifying and analysing genes involved in various biological processes, including cell division, growth and development, signal transduction, transcript regulation, and responses to environmental cues. In this review, we discuss and compare the merits and limitations of the different genome-wide expression analysis technologies, including (1) complementary DNA (cDNA) microarrays, (2) oligonucleotide microarrays, (3) serial analysis of gene expression, (4) massively parallel signature sequencing, and (5) cDNA-amplified fragment-length polymorphism. Particular attention will be given to the genome-wide expression analysis of genes involved in somatic embryogenesis.

1 Introduction

Genome-wide expression analysis is an important tool for analysing genes involved in cellular, molecular, and developmental biological processes in microorganisms, plants, and animals (Hegde et al. 2000; Schena et al. 1995). Somatic embryogenesis is an asexual form of plant propagation in nature that mimics many of the events of sexual reproduction. The control of somatic embryo development involves the temporal expression of different sets of genes that allow the dividing cell to progress through the different stages of somatic embryogenesis. DNA microarrays provide a convenient tool for genome-wide expression analysis (Hegde et al. 2000; Schena et al. 1995). Studies using DNA microarrays to follow the patterns of genes allowed the identification of thousands or hundreds of genes that are involved in specific developmental processes. Although DNA microarrays are rapidly becoming the standard tool for genome-wide expression analysis, their application is still limited to a restricted number of experimental systems where the complete genome sequence or a large complementary DNA (cDNA) collection is available (Breyne and Zabeau 2001; Hegde et al. 2000; Schena et al. 1995). Several alternative technologies for expression profiling based on DNA sequencing or cDNA fragment analysis have been developed and successfully used in other biological systems, including plant species. DNA fragment analysis based methods, such as cDNA-amplified fragment-length polymorphism (AFLP), provide

Table 1 Comparison of methods used for genome-wide gene expression analysis

	cDNA microarray	Oligonucleotide microarray	SAGE	MPSS	cDNA-AFLP
Sensitivity	Moderate	Moderate	Moderate/ high	Moderate/ high	High
Specificity	Low	Low	High	High	High
Expression-level measurement	Relative	Relative	Absolute	Absolute	Relative
Possibility to integrate data	Yes	Yes	Yes	Yes	No
Necessity of molecular resources	Yes	Yes	Yes	Yes	No
Labour intensity	Low	Low	High	High	High
Cost	High	High	High	High	Low

a more appropriate tool for genome-wide expression analysis. Moreover, cDNA-AFLP exhibits properties that complement DNA microarrays and can be a useful tool for gene discovery (Breyne and Zabeau 2001). In this study, we overview the different genome-wide expression analysis technologies, including (1) cDNA microarrays, (2) oligonucleotide microarrays, (3) serial analysis of gene expression, (4) massively parallel signature sequencing (MPSS), and (5) cDNA-AFLP (Table 1). Particular attention will be given to the genome-wide expression analysis of genes involved in somatic embryogenesis.

2 Somatic Embryogenesis

Somatic embryogenesis is an important prerequisite for the use of many biotechnological tools for genetic improvement, as well as for clonal propagation (Schenk and Hildebrandt 1972; Yeung and Meinke 1993). Somatic embryogenesis may be induced by the manipulation of tissues and cells in vitro. Some of the most important factors for a successful plant regeneration are the culture medium and environmental incubation conditions. In angiosperms, the zygote divides transversally into two cells. The apical cell is small and dense with an intense activity of DNA synthesis (Yeung and Meinke 1993). This cell gives rise to the embryo head that will be the new plant. The basal cell is a large and highly vacuolated one that will form the suspensor complex, which plays an important role during the early stages of the young embryo (Yeung and Sussex 1979). Somatic embryos generally follow the same pattern

and are initiated from a somatic cell. Somatic embryos are formed from single cells cultivated in liquid or solid medium. Embryos can be distinguished from adventitious shoots, because they are bipolar, having both a shoot and root pole, and they do not have any vascular connections with the underlying parental tissue (Haccius 1978). Somatic embryo production is steadily being increased as essential factors become better understood (Williams and Maheswaran 1986). The ability to recover plants from single cells has made possible the genetic improvement. The most important advantages of somatic embryogenesis used in plant biology, including the ability to handle large numbers of individual cells in very small spaces and genetic variability, can be created deliberately in cultured cells by using genetic-engineering techniques (Yeung and Meinke 1993).

3
Late Embryogenesis Abundant Proteins

Late embryogenesis abundant (LEA) proteins are developmentally induced during the different stages of embryogenesis and are environmentally induced in embryos by desiccation or culture with abscisic acid (ABA) or high osmoticum (Hughes and Galau 1991). LEA proteins comprise a large group of probable desiccation protectants that are induced by similar stresses in vegetative tissues of different plant species (Skriver and Mundy 1991). In cotton (*Gossypium hirsutum*), 18 *Lea* and *LeaA* messenger RNAs (mRNAs) were cloned and identified to be environmentally induced by water stress; two of them, *Lea5* (cDNA D73) and *Leal4* (cDNA D95) are highly induced in mature leaves of water-stressed plants (Galau et al. 1986). In *Craterostigma plantagineum*, the desiccation-induced cDNA pcC27-45 were identified to encode proteins that are very hydrophilic (Baker et al. 1988; Piatkowski et al. 1990). *Lea* genes encode proteins with significant hydropathic character. Their hydropathic profiles are unremarkable; the amino-terminal half is somewhat hydropathic, possibly with a membrane-spanning region, and the carboxy-terminal half is somewhat hydrophilic (Galau et al. 1993). The proteins encoded by cotton *Leal4* and *Craterostigma* pcC27-45 thus define an additional family of water-stress-related proteins (Baker et al. 1988), the group 4 LEA proteins. An ACGT-containing element has been shown to be involved in the ABA induction of a wheat *Lea* gene (Guiltinan et al. 1990). *Leal4-A* contains sequences at nucleotides – 58 and – 14 from the transcription start that are similar to this element and similar sequences that are in many cotton Leu genes (Galau et al. 1992). *LeaZ4-A* encodes a 16.4-kD protein that is exactly collinear, with 66% identity, with that encoded by the *Craterostigma* cDNA pcC27-45, which is induced in leaves and roots during desiccation and in ABA-treated and NaCl-treated callus (Piatkowski et al. 1990). These proteins are slightly hydropathic throughout.

4
cDNA Microarray

Microarray expression analysis has become one of the most widely used functional genomics tools (Schaffer et al. 2000). Efficient application of this technique requires the development of robust and reproducible protocols, including PCR amplification of target cDNA clones, microarray printing, probe labelling, and hybridization cDNA microarrays (Hegde et al. 2000; Schena et al. 1995). cDNA microarrays have been developed that allow mRNA expression to be assessed on a global scale, allowing the parallel assessment of gene expression for hundreds or thousands of genes in a single experiment (Baldwin et al. 1999). The commonest use of these is for the determination of patterns of differential gene expression, comparing differences in mRNA expression levels between identical cells subjected to different stimuli or between different cellular phenotypes or developmental stages (Laub et al. 2000).

Microarray expression analysis is the most widely used method for profiling mRNA expression (Laub et al. 2000). cDNA segments representing the collection of genes are amplified by PCR and mechanically spotted at high density on glass microscope slides using robotic systems, creating a microarray containing thousands of elements (Hegde et al. 2000). Microarrays containing tens of thousands of cDNA clones can be easily constructed. The kinetics of hybridization allows relative expression levels to be determined based on the ratio with which each probe hybridizes to an individual array element. Hybridization is assayed using a confocal laser scanner to measure fluorescence intensities, allowing simultaneous determination of the relative expression levels of all the genes represented in the array (Hegde et al. 2000; Schena et al. 1995). The process of expression analysis can be broadly divided into three stages: (1) array fabrication; (2) probe preparation and hybridization; (3) data collection, normalization, and analysis (Hegde et al. 2000; Schena et al. 1995).

The cDNA microarrays (Schena et al. 1995) have proven powerful and are now widely used for genome-wide expression analysis in a wide range of organisms, including plants (Baldwin et al. 1999; Richmond and Somerville 2000; Schaffer et al. 2000). cDNA microarrays allow up to tens of thousands of genes to be analysed simultaneously. Microarrays comprising complete gene sets are available for a number of organisms, such as yeast (Wodicka et al. 1997), a number of bacteria (Laub et al. 2000; Selinger et al. 2000), and *Caenorhabditis elegans* (Jiang et al. 2001), for which the entire genome sequence has been determined. For example, it was reported that gene expression during the cell cycle in bacteria is strictly regulated at the level of transcription and that the expression profiles of cell cycle modulated genes are coincident with the functional activity of the genes (Laub et al. 2000). For a few other well-studied animal and plant species, the current gener-

ation of microarrays is limited to a subset of the genes, namely those for which a cDNA clone or an expressed sequence tag (EST) sequence is available. Hegde et al. (2000) developed protocols that had been standardized and that had been used regularly in many laboratories for microarray analysis. The procedures described have been tested and refined over the past year and have been optimized using hybridization of RNA derived from cell lines to give reproducible and consistent results. It should be noted that a number of alternative protocols have been published (Eisen and Brown 1999), but the system developed by Hegde et al. (2000) has a number of advantages over these. In particular, the combination of printing, labelling, and hybridization conditions that have allowed a significant reduction in the quantity of starting total RNA required for analysis.

5 Oligonucleotide Microarrays

Oligonucleotide microarray based hybridization analysis is a promising new technology which potentially allows rapid and cost-effective screens for all possible mutations and sequence variations in genomic DNA (Roberts et al. 2000; Saiki et al. 1989). Identifying and cataloguing these variations is a critical part of approaches that seek to identify the genetic basis for resistance to disease. These sequence variations will serve as genetic markers in studies of diseases and traits with complex inheritance patterns (Golub et al. 1999; Roberts et al. 2000). Large-scale sequence analysis is needed for population-based genetic risk assessment and diagnostic tests once mutations have been identified, because traditional technologies cannot easily meet the demands for rapid and cost-effective large-scale comparative sequence and mutational analysis (Hacia 1999). To perform thousands of separate hybridization reactions to evaluate each sample makes an oligonucleotide microarray more amenable to a large-scale clinical diagnostic laboratory than a common research laboratory setting (Lockhart et al. 1996). The current scientific literature largely centres on arrays manufactured using photolithographic-based methodologies developed by Affymetrix (Fodor et al. 1991; Hacia 1999). However, technologies such as mass spectroscopy based hybridization detection, could have an important role in coming years.

Oligonucleotide array based detection of known genomic DNA sequence variations was first reported in 1989 (Saiki et al. 1989). Probes complementary to six *HLA-DQA* alleles as well as nine mutations in *HBB* (encoding β-globin) were spotted onto nylon filters and incubated with biotin-labelled PCR products (Yershov et al. 1996). Advanced oligonucleotide array manufacturing processes have opened the way to evaluating more complex systems (Yershov et al. 1996). Arrays of 1480 oligonucleotide probes synthesized in situ by photolithographic-based processes were designed to detect 37 known muta-

tions in the coding region of *CFTR*, as well as all possible single-nucleotide substitutions (Yershov et al. 1996). In a blinded study, ten genomic DNA samples were successfully genotyped by characterizing fluorescent hybridization signals from test and wild-type reference samples at mutation-specific probes relative to those from wild-type samples. In a separate study, arrays of six oligonucleotide probes, generated by spotting oligonucleotides onto activated surfaces, were used to detect three different mutations in *HBB* (Yershov et al. 1996).

In *Arabidopsis*, defence and wounding responses have been analysed using cDNA microarrays (Schenk et al. 2000), whereas oligonucleotide arrays were used to study circadian-rhythm-modulated gene expression (Harmer et al. 2000). The analysis of the processes underlying fruit ripening in strawberries (Aharoni et al. 2000) was the first application of microarrays in a non-model plant species. The most important advantage of microarray-based technology is that gene expression profiles from either different samples or samples obtained using different treatments can be compared with each other and analysed together (Golub et al. 1999). Another striking example is presented in the landmark paper that describes the construction of a compendium of yeast expression profiles, combining data from both a number of mutant strains and treatments with different chemical compounds (Hughes et al. 2000). The power of microarrays was clearly illustrated by the characterization of a number of novel yeast genes solely on the basis of the gene expression profiles of the mutant strains. Similarly, the crosstalk and interaction among multiple mitogen-activated protein kinase pathways could be revealed by integrating gene expression profiles obtained under different experimental conditions (Roberts et al. 2000).

6
Serial Analysis of Gene Expression

Serial analysis of gene expression (SAGE) is a technique designed to take advantage of high-throughput sequencing technology to obtain a quantitative profile of cellular gene expression (Fig. 1). The SAGE technique measures not the expression level of a gene, but quantifies a tag that is a nucleotide sequence of a defined length adjacent to the 3′-most restriction site for a particular restriction enzyme and represents the transcription product of a gene (Velculescu et al. 1995). The SAGE technique is based on counting sequence tags of 14–15 bases from cDNA libraries (Velculescu et al. 1995; Zhang et al. 1997). This technology has been widely used to monitor gene expression in human cell cultures and tissue samples (Lash et al. 2000; Velculescu et al. 2000), but not in other organisms. In plants, this method has been applied only sporadically (Matsumura et al. 1999). The principle advantage of SAGE is that it gives an absolute measure of gene expression instead of measuring

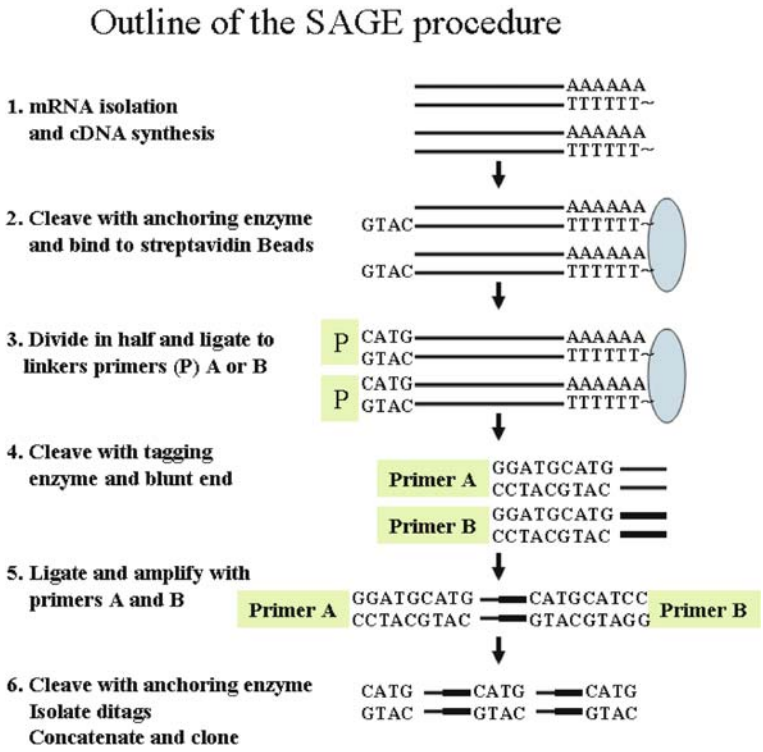

Fig. 1 Schematic of the serial analysis of gene expression (*SAGE*)

relative expression levels. Indeed, by counting the number of tags from each cDNA, one obtains an accurate measure of the number of transcripts present in the mRNA sample. As in the case of microarrays, independent data sets can be compiled in a single database, allowing the comparative analysis of data from different experiments (Lash et al. 2000; Velculescu et al. 2000). The public database SAGEmap already contains a comprehensive quantity of SAGE data from different cDNA libraries (Lash et al. 2000). Newly obtained data can be merged with the records already present in the database, enabling a more significant analysis of gene expression profiles.

SAGE required high amounts of input RNA, restricting its utility to large tissue samples. Recent improvements, however, now allow the use of 500–5000-fold less starting material and permit work with minute quantities of tissue containing only a few hundred or thousand cells (Datson et al. 1999; Matsumura et al. 1999). Although *Nla*III remains the most widely used restriction enzyme, enzyme substitutions are possible. The data product of the SAGE technique is a list of tags, with their corresponding count values, and thus is a digital representation of cellular gene expression. The principal limitation of SAGE is the need to sequence large numbers of tags in order to

monitor the scarcely expressed genes. Another drawback of SAGE is that the tags obtained are very short and hence not always unambiguous. Gene identification on the basis of short sequence tags relies on the availability of large databases of well-characterized ESTs. So there are two problems to be tackled when dealing with SAGE data in the form of tags and counts. The first deals with ensuring that the tags and their counts are a valid representation of transcripts and their levels of expression, and the second with making valid tag-to-gene assignments.

7
Massively Parallel Signature Sequencing

The recently developed MPSS technology holds the promise of a major improvement over SAGE (Brenner et al. 2000). MPSS is a parallel sequencing method that can generate hundreds of thousands of short sequence signatures in a single analysis, thus overcoming the principal shortcoming of SAGE (Brenner et al. 2000). Because the method generates longer, 16–20-base signatures, it should also be more accurate. Technically, however, the method is rather complex and not yet readily available to the broad scientific community (Brenner et al. 2000). The genomic sequence of *A. thaliana* has been completed in recent years (*Arabidopsis* Genome Initiative 2000). Experimental analyses and comprehensive descriptions of plant transcriptomes continue in parallel (Haas et al. 2003; Yamada et al. 2003). No plant transcriptome has been extensively characterized experimentally with both quantitative and qualitative expression data. Computational approaches to genome annotation can miss or incorrectly predict many genes, and validation of genome annotations with experimental data is essential (Andrews et al. 2000; Guigo et al. 2000).

As genomic sequencing becomes faster and more economical, it is critically important that methods are developed to detect and quantify every gene and alternatively spliced transcript within a genome (Adams et al. 1995). Large-scale sequencing of short mRNA-derived tags can establish the qualitative and quantitative characteristics of a complex transcriptome (Meyers et al. 2004). Meyers et al. (2004) sequenced 12 304 362 tags from five diverse libraries of *A. thaliana* using MPSS. A total of 48 572 distinct signatures, each representing a different transcript, were expressed at significant levels (Meyers et al. 2004). These signatures were compared with the annotation of the *A. thaliana* genomic sequence; in the five libraries, this comparison yielded between 17 353 and 18 361 genes with sense expression, and between 5487 and 8729 genes with antisense expression (Meyers et al. 2004). An additional 6691 MPSS signatures mapped to unannotated regions of the genome. Expression was demonstrated for 1168 genes for which expression data were previously unknown (Meyers et al. 2004). Alternative polyadenylation was observed for

more than 25% of *A. thaliana* genes transcribed in these libraries. The MPSS expression data suggest that the *A. thaliana* transcriptome is complex and contains many as-yet uncharacterized variants of normal coding transcripts (Meyers et al. 2004).

8 cDNA-Amplified Fragment-Length Polymorphism

The differential display technique developed by Liang and Pardee (1992) has been widely used to screen for genes that are differentially expressed. After the first publication of the differential display technique (Liang and Pardee 1992), several improved PCR-based methods, using restriction enzymes to generate cDNA specific tags, were described (Bachem et al. 1996; Kawamoto et al. 1999; Shimkets et al. 1999; Sutcliffe et al. 2000). The most widely used method, cDNA-AFLP, has been applied with success to the systematic analysis of genes involved in particular biological processes (Breyne and Zabeau 2001; Durrant et al. 2000). The cDNA-AFLP is based on the principle that a complex starting mixture of cDNAs is fractionated into smaller subsets, after which cDNA tags are PCR-amplified and separated on high-resolution gels (Breyne and Zabeau 2001; Durrant et al. 2000). The observed differences in the intensity of the bands provide a good measure of the relative differences in the levels of gene expression (Breyne and Zabeau 2001; Durrant et al. 2000). In a study of fungal pathogen response in tobacco cells, the screening of approximately 30 000 transcript tags identified a total of 273 modulated gene tags (Durrant et al. 2000). These differential display methods have proven useful for discovering differentially expressed genes, but not for quantitative genome-wide transcription analysis (Breyne and Zabeau 2001).

cDNA-AFLP analysis has been used to reveal early gene expression associated with the commitment and differentiation of a plant tracheary element by Milioni et al. (2002). The exogenous growth factors, auxin and cytokinin, are not required in the first 48 h after isolation of *Zinnia* mesophyll cells; furthermore, as little as 10 min of exposure to the growth factors at 48 h is both necessary and sufficient to commit cells to the tracheary element's differentiation pathway (Milioni et al. 2001). These findings suggest that the first 48 h of culture represents a time in which the cells adapt to liquid culture and acquire the competence to respond to the inductive signals (McCann 1997; Milioni et al. 2001). The precise transdifferentiation process provides a new and improved context in which to discover the earliest genes involved in switching on the developmental programme. In this project, a total of 652 differentially accumulated transcript-derived fragments (TDFs), ranging in length from 50 to 450 bp, were recovered from gels and reamplified, subcloned, and sequenced (Milioni et al. 2002). A total of 349 fragments (53.5%) of the differentially expressed genes showed close matches to database entries with assigned

identities. Thirteen groups were classified from these sequences based on functional categories established for *Arabidopsis* (*Arabidopsis* Genome Initiative 2000). The major group is involved in primary and secondary metabolism and energy generation (19.2%), whereas a slightly higher proportion (8%) is cell-wall-related. An additional 9.7% of the TDFs are involved in information processing and constitute genes involved in transcriptional control and signal transduction. In addition, 12.4% of the sequences share significant similarity to unknown or hypothetical genes with no assigned function from various genome projects, which represent new candidate proteins involved in cell fate determination, differentiation, cell wall remodelling, and cell death.

To understand how embryonic cells differentiate into the 40 or so cell types that constitute plants (Hulskamp and Kirik 2000), one approach is to study mutants in which meristematic function has been compromised (Haecker and Laux 2001). Another approach is to study mutants in which a clear developmental phenotype for a particular cell type can be identified, for example, root hairs (Parker et al. 2000), trichomes (Hulskamp and Kirik 2000), or xylem (McCann and Roberts 2000), based on identification of genes that are differentially expressed. Global gene expression technologies may permit the dissection of downstream events through comparisons of mutants in these pathways; however, to date, only a few genes have been identified that are specific to particular cell types (Milioni et al. 2001). Genes involved in vascular cell fates have been identified in cDNA-sequencing projects using material derived from young xylem tissue of loblolly pine (Allona et al. 1998) and poplar (Sterky et al. 1998). Tissue-specific transcript profiles have been obtained using DNA microarray analysis of 3000 ESTs of poplar (Hertzberg et al. 2001). To elucidate genetic programmes that control embryogenesis and regeneration of rice, Ito et al. (2002) conducted genome-wide expression analysis of genes involved in somatic embryogenesis. Functional analyses of genes demonstrated that five *KNOX* family class 1 homeobox genes were involved in somatic embryogenesis (Ito et al. 2002). The *KNOX* family class 1 homeobox genes encode transcription factors and protein kinases. Expression patterns of these genes during early embryogenesis and regeneration were analysed by reverse transcription PCR and in situ hybridization (Ito et al. 2002). It was found that constitutive expression of these genes is sufficient to maintain cells in a meristematic undifferentiated state (Ito et al. 2002).

9 Conclusion

Genome-wide expression analysis allows scientists to identify genes that are involved in somatic embryogenesis in plants. The control of somatic embryogenesis involves the temporal expression of different sets of genes through the different phases of the embryo development. A landmark study using

genome-wide expression analysis to follow the patterns of gene expression in rice has allowed the identification of hundreds of genes that are involved in somatic embryogenesis (Ito et al. 2002). Different genome-wide expression analysis technologies, including (1) cDNA microarray, (2) oligonucleotide microarrays, (3) serial analysis of gene expression, (4) MPSS, and (5) cDNA-AFLP, provide opportunities to explore the mechanism of somatic embryogenesis. DNA microarrays provide a convenient tool for genome-wide expression analysis; however, their use is limited to organisms for which the complete genome sequence or a large cDNA collection is available. Alternative technologies for expression profiling based on DNA sequencing or cDNA fragment analysis have been developed and successfully used in different biological systems. For example, cDNA-AFLP exhibits properties that complement DNA microarrays and may provide a more appropriate tool for genome-wide expression analysis, gene discovery, and transcript profiling. Somatic embryogenesis has been induced in some pine species (Tang 2000; Tang et al. 2001). We are using different genome-wide expression analysis technologies to identify genes involved in somatic embryogenesis.

References

Adams MD, Kerlavage AR, Fleischmann RD, Fuldner RA, Bult CJ, Lee NH, Kirkness EF, Weinstock KG, Gocayne JD, White O, Sutton G, Blake JA, Brandon RG, Chiu MW, Clayton RA, Cline RT, Cotton MD, Earlehughes J, Fine LD, Fitzgerald LM, Fitzhugh WM, Fritchman JL, Geoghagen NSM, Glodek A, Gnehm CL, Hanna MC, Hedblom E, Hinkle PS, Kelley JM, Klimek KM, Kelley JC, Liu LI, Marmaros SM, Merrick JM, Morenopalanques RF, Mcdonald LA, Nguyen DT, Pellegrino SM, Phillips CA, Ryder SE, Scott JL, Saudek DM, Shirley R, Small KV, Spriggs TA, Utterback TR, Weldman JF, Li Y, Barthlow R, Bednarik DP, Cao LA, Cepeda MA, Coleman TA, Collins EJ, Dimke D, Feng P, Ferrie A, Fischer C, Hastings GA, He WW, Hu JS, Huddleston KA, Greene JM, Gruber J, Hudson P, Kim A, Kozak DL, Kunsch C, Ji HJ, Li HD, Meissner PS, Olsen H, Raymond L, Wei YF, Wing J, Xu C, Yu GL, Ruben SM, Dillon PJ, Fannon MR, Rosen CA, Haseltine WA, Fields C, Fraser CM, Venter JC (1995) Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. Nature 377:3–174

Aharoni A, Keizer LCP, Bouwmeester HJ, Sun ZK, Alvarez-Huerta M, Verhoeven HA, Blaas J, van Houwelingen AMML, De Vos RCH, van der Voet H, Jansen RC, Guis M, Mol J, Davis RW, Schena M, van Tunen AJ, O'Connell AP (2000) Identification of the SAAT gene involved in strawberry flavor biogenesis by use of DNA microarrays. Plant Cell 12:647–662

Allona I, Quinn M, Shoop E, Swope K, St Cyr S, Carlis J (1998) Analysis of xylem formation in pine by cDNA sequencing. Proc Natl Acad Sci USA 95:9693–9698

Andrews J, Bouffard GG, Cheadle C, Lu JN, Becker KG, Oliver B (2000) Gene discovery using computational and microarray analysis of transcription in the Drosophila melanogaster testis. Genome Res 10:2030–2043

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408:796–815

Bachem CW, van der Hoeven RS, de Bruijn SM, Vreugdenhil D, Zabeau M, Visser RG (1996) Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. Plant J 9:745–753

Baker J, Steele C, Dure L (1988) Sequence and characterization of 6 Leu proteins and their genes from cotton. Plant Mol Biol 11:277–291

Baldwin D, Crane V, Rice D (1999) A comparison of gel-based, nylon filter and microarray techniques to detect differential RNA expression in plants. Curr Opin Plant Biol 2:96–103

Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo SJ, McCurdy S, Foy M, Ewan M, Roth R, George D, Eletr S, Albrecht G, Vermaas E, Williams SR, Moon K, Burcham T, Pallas M, DuBridge RB, Kirchner J, Fearon K, Mao J, Corcoran K (2000) Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nat Biotechnol 18:630–634

Breyne P, Zabeau M (2001) Genome-wide expression analysis of plant cell cycle modulated genes. Curr Opin Plant Biol 4:136–142

Datson NA, van der Perk-de Jong J, van den Berg MP, de Kloet ER, Vreugdenhil E (1999) MicroSAGE: a modified procedure for serial analysis of gene expression in limited amounts of tissue. Nucleic Acids Res 27:1300–1307

Durrant WE, Rowland O, Piedras P, Hammond-Kosack KE, Jones JD (2000) cDNA-AFLP reveals a striking overlap in race-specific resistance and wound response gene expression profiles. Plant Cell 12:963–977

Eisen MB, Brown PO (1999) DNA arrays for analysis of gene expression. Methods Enzymol 303:179–205

Fodor SPA, Read JL, Pirrung MC, Stryer L, Lu AT, Solas D (1991) Light-directed spatially addressable parallel chemical synthesis. Science 251:767–773

Galau GA, Hughes DW, Dure L (1986) Abscisic acid induction of cloned cotton late embryogenesis-abundant (Leu) mRNAs. Plant Mol Biol 7:155–170

Galau GA, Wang HYC, Hughes DW (1992) Cotton Leu4 (D19) and LeaA2 (D132) Group 1 Leu genes encoding water stress-related proteins containing a 20-amino acid motif. Plant Physiol 99:783–788

Galau GA, Wang HYC, Hughes DW (1993) Cotton fea5 and fea74 encode typical late embryogenesis-abundant proteins. Plant Physiol 101:695–696

Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286:531–537

Guigo R, Agarwal P, Abril JF, Burset M, Fickett JW (2000) An assessment of gene prediction accuracy in large DNA sequences. Genome Res 10:1631–1642

Guiltinan MJ, Marcotte WR Jr, Quatrano RS (1990) A plant leucine zipper protein that recognizes an abscisic acid response element. Science 250:267–271

Haas BJ, Delcher AL, Mount SM, Wortman JR, Smith RK, Hannick LI, Maiti R, Ronning CM, Rusch DB, Town CD, Salzberg SL, White O (2003) Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. Nucleic Acids Res 31:5654–5666

Haccius B (1978) Question of unicellular origin of non-zygotic embryos in callus cultures. Phytomorphology 28:74–81

Hacia JG (1999) Resequencing and mutational analysis using oligonucleotide microarrays. Nat Genet 21(S):42–47

Haecker A, Laux T (2001) Cell–cell signaling in the shoot meristem. Curr Opin Plant Biol 4:441–446

Harmer SL, Hogenesch JB, Straume M, Chang HS, Han B, Zhu T, Wang X, Kreps JA, Kay SA (2000) Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. Science 290:2110–2113

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 29:548–562

Hertzberg M, Aspeborg H, Schrader J, Andersson A, Erlandsson R, Blomqvist K, Bhalerao R, Uhlen M, Teeri TT, Lundeberg J, Sundberg B, Nilsson P, Sandberg G (2001) A transcriptional roadmap to wood formation. Proc Natl Acad Sci USA 98:14732–14737

Hughes DW, Galau GA (1991) Developmental and environmental induction of Leu and LeuA mRNAs and the postabscission program during embryo culture. Plant Cell 3:605–618

Hughes TR, Marton MJ, Jones AR, Roberts CJ, Stoughton R, Armour CD, Bennett HA, Coffey E, Dai HY, He YDD, Kidd MJ, King AM, Meyer MR, Slade D, Lum PY, Stepaniants SB, Shoemaker DD, Gachotte D, Chakraburtty K, Simon J, Bard M, Friend SH (2000) Functional discovery via a compendium of expression profiles. Cell 102:109–126

Ito Y, Hirochika H, Kurata N (2002) Organ specific alternative transcripts of KNOX family class 2 homeobox genes of rice. Gene 288:41–47

Hulskamp M, Kirik V (2000) Trichome differentiation and morphogenesis in Arabidopsis. Adv Bot Res 31:237–260

Jiang M, Ryu J, Kiraly M, Duke K, Reinke V, Kim SK (2001) Genome-wide analysis of developmental and sex-regulated gene expression profiles in Caenorhabditis elegans. Proc Natl Acad Sci USA 98:218–223

Kawamoto S, Ohnishi T, Kita H, Chisaka O, Okubo K (1999) Expression profiling by iAFLP: a PCR-based method for genome-wide gene expression profiling. Genome Res 9:1305–1312

Lash AE, Tolstoshev CM, Wagner L, Schuler GD, Strausberg RL, Riggins GJ, Altschul SF (2000) SAGEmap: a public gene expression resource. Genome Res 10:1051–1060

Laub MT, McAdams HH, Feldblyum T, Fraser CM, Shapiro L (2000) Global analysis of the genetic network controlling a bacterial cell cycle. Science 290:2144–2148

Liang P, Pardee AB (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257:967–971

Lockhart DJ, Dong HL, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang CW, Kobayashi M, Horton H, Brown EL (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. Nat Biotechnol 14:1675–1680

Matsumura H, Nirasawa S, Terauchi R (1999) Technical advance: transcript profiling in rice (Oryza sativa L.) seedlings using serial analysis of gene expression (SAGE). Plant J 20:719–726

McCann MC (1997) Tracheary element formation: Building up to a dead end. Trends Plant Sci 2:333–338

McCann MC, Roberts K (2000) Xylogenesis: the birth of a corpse. Curr Opin Plant Biol 3:517–522

Meyers BC, Vu TH, Tej SS, Ghazal H, Matvienko M, Agrawal V, Ning J, Haudenschild CD (2004) Analysis of the transcriptional complexity of Arabidopsis thaliana by massively parallel signature sequencing. Nat Biotechnol 22:1006–1011

Milioni D, Sado PE, Stacey NJ, Domingo C, Roberts K, McCann MC (2001) Differential expression of cell-wall-related genes during the formation of tracheary elements in the Zinnia mesophyll cell system. Plant Mol Biol 47:221–238

Milioni D, Sado PE, Stacey NJ, Roberts K, McCann MC (2002) Early gene expression associated with the commitment and differentiation of a plant tracheary element is revealed by cDNA-Amplified Fragment Length Polymorphism analysis. Plant Cell 14:2813–2824

Parker JS, Cavell AC, Dolan L, Roberts K, Grierson CS (2000) Genetic interactions during root hair morphogenesis in Arabidopsis. Plant Cell 12:1961–1974

Piatkowski D, Schneider K, Salamini F, Bartels D (1990) Characterization of five abscisic acid-responsive cDNA clones isolated from the desiccation-tolerant plant Craterostigma plantagineum and their relationship to other water-stress genes. Plant Physiol 94:1682–1688

Richmond T, Somerville S (2000) Chasing the dream: plant EST microarrays. Curr Opin Plant Biol 3:108–116

Roberts CJ, Nelson B, Marton MJ, Stoughton R, Meyer MR, Bennett HA, He YDD, Dai HY, Walker WL, Hughes TR, Tyers M, Boone C, Friend SH (2000) Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. Science 287:873–880

Saiki RK, Walsh PS, Levenson CH, Erlich HA (1989) Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. Proc Natl Acad Sci USA 86:6230–6234

Schaffer R, Landgraf J, Perez-Amador M, Wisman E (2000) Monitoring genome-wide expression in plants. Curr Opin Biotechnol 11:162–167

Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467–470

Schenk RV, Hildebrandt HC (1972) Medium and techniques for induction and growth of monocotyledoneous and dicotyledoneous plant cell cultures. Can J Bot 50:199–204

Schenk PM, Kazan K, Wilson I, Anderson JP, Richmond T, Somerville SC, Manners JM (2000) Coordinated plant defense responses in Arabidopsis revealed by microarray analysis. Proc Natl Acad Sci USA 97:11655–11660

Selinger DW, Cheung KJ, Mei R, Johansson EM, Richmond CS, Blattner FR, Lockhart DJ, Church GM (2000) RNA expression analysis using a 30 base pair resolution Escherichia coli genome array. Nat Biotechnol 18:1262–1268

Shimkets RA, Lowe DG, Tai JTN, Sehl P, Jin HK, Yang RH, Predki PF, Rothberg BEG, Murtha MT, Roth ME, Shenoy SG, Windemuth A, Simpson JW, Simons JF, Daley MP, Gold SA, McKenna MP, Hillan K, Went GT, Rothberg JM (1999) Gene expression analysis by transcript profiling coupled to a gene database query. Nat Biotechnol 17:798–803

Skriver K, Mundy J (1991) Gene expression in response to abscisic acid and osmotic stress. Plant Cell 2:503–512

Sterky F, Regan S, Karlsson J, Hertzberg M, Rohde A, Holmberg A, Amini B, Bhalerao R, Larsson M, Villarroel R, Van Montagu M, Sandberg G, Olsson O, Teeri TT, Boerjan W, Gustafsson P, Uhlen M, Sundberg B, Lundeberg J (1998) Gene discovery in the wood-forming tissues of poplar: Analysis of 5692 expressed sequence tags. Proc Natl Acad Sci USA 95:13330–13335

Sutcliffe JG, Foye PE, Erlander MG, Hilbush BS, Bodzin LJ, Durham JT, Hasel KW (2000) TOGA: an automated parsing technology for analyzing expression of nearly all genes. Proc Natl Acad Sci USA 97:1976–1981

Tang W (2000) Peroxidase activities of desiccation tolerant loblolly pine somatic embryos. In Vitro Cell Dev Biol Plant 36:488–491

Tang W, Guo Z, Ouyang F (2001) Plant regeneration from embryogenic cultures initiated from mature loblolly pine zygotic embryos. In Vitro Cell Dev Biol Plant 37:558–563

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW (1995) Serial analysis of gene expression. Science 270:484–487

Velculescu VE, Vogelstein B, Kinzler KW (2000) Analysing uncharted transcriptomes with SAGE. Trends Genet 16:423–425

Williams EG, Maheswaran G (1986) Somatic embryogenesis: factors influencing coordinated behaviour of cells as an embryogenic group. Ann Bot 57:443–462

Wodicka L, Dong H, Mittmann M, Ho MH, Lockhart DJ (1997) Genomewide expression monitoring in Saccharomyces cerevisiae. Nat Biotechnol 15:1359–1367

Yamada K, Lim J, Dale JM, Chen HM, Shinn P, Palm CJ, Southwick AM, Wu HC, Kim C, Nguyen M, Pham P, Cheuk R, Karlin-Newmann G, Liu SX, Lam B, Sakano H, Wu T, Yu GX, Miranda M, Quach HL, Tripp M, Chang CH, Lee JM, Toriumi M, Chan MMH, Tang CC, Onodera CS, Deng JM, Akiyama K, Ansari Y, Arakawa T, Banh J, Banno F, Bowser L, Brooks S, Carninci P, Chao QM, Choy N, Enju A, Goldsmith AD, Gurjal M, Hansen NF, Hayashizaki Y, Johnson-Hopson C, Hsuan VW, Iida K, Karnes M, Khan S, Koesema E, Ishida J, Jiang PX, Jones T, Kawai J, Kamiya A, Meyers C, Nakajima M, Narusaka M, Seki M, Sakurai T, Satou M, Tamse R, Vaysberg M, Wallender EK, Wong C, Yamamura Y, Yuan SL, Shinozaki K, Davis RW, Theologis A, Ecker JR (2003) Empirical analysis of transcriptional activity in the Arabidopsis genome. Science 302:842–846

Yershov G, Barsky V, Belgovskiy A, Kirillov E, Kreindlin E, Ivanov I, Parinov S, Guschin D, Drobishev A, Dubiley S, Mirzabekov A (1996) DNA analysis and diagnostics on oligonucleotide microchips. Proc Natl Acad Sci USA 93:4913–4918

Yeung EC, Meinke DW (1993) Embryogenesis in angiosperms: development of the suspensor. Plant Cell 5:1371–1381

Yeung EC, Sussex IM (1979) Embryogeny of Phaseolus coccineus: the suspensor and the growth of the embryo-proper in vitro. Z Pflanzenphysiol 91:423–433

Zhang L, Zhou W, Velculescu VE, Kern SE, Hruban RH, Hamilton SR, Vogelstein B, Kinzler KW (1997) Gene expression proles in normal and cancer cells. Science 23:1268–1272