

Three Different Methods of Expression Profile Analysis in Plants

Mari L. Salmi, Thomas J. Bushart and Greg B. Clark

Department of Molecular, Cell and Developmental Biology, Scholl of Biological Sciences, The University of Texas at Austin, Austin TX 78712, USA

Abstract

In the last 15 years many advances have been made in techniques available to molecular biologists. In the field of gene expression in particular, several new methods have become common use for determining changes in gene expression over the course of plant development, changes induced by external stimulus, and differences in the various tissue types of a plant. Here we describe the theory, general methods, and some applications for three of these new techniques in the broad field of plant biology; DNA microarray, quantitative real-time reverse transcription PCR, and serial analysis of gene expression.

Introduction

Environmental stimuli often induce differential expression of genes that mediate the adaptive responses to those stimuli. For example, heat-shock proteins are up-regulated in response to increased temperature (Keeler et al. 2000), the expression of several calmodulin genes is altered in response to touch (Botella and Arteca 1994), and genes related to photosynthesis are up-regulated by light (Lopez-Juez et al. 1998). In these examples and others, genes whose expression is altered by a stimulus, play a key role in the signal transduction and physiological response to that stimulus. Currently, there are several methods available to researchers to monitor these changes in gene expression. Here we will review the applications of three of the most recently developed approaches for monitoring transcript abundance; DNA microarray, Quantitative Real-Time Reverse Transcription Polymerase Chain Reaction (Q RT-PCR), and Serial Analysis of Gene Expression (SAGE).

DNA Microarray: DNA microarray technology has provided a method of monitoring expression profiles in any experimental condition of thousands of genes simultaneously. In plants, microarray analysis has been used to study changes during floral development and development of reproductive structures (Hennig et al. 2004), growth under conditions of biotic and abiotic stress (Schenk et al. 2000; Marin et al. 2003; Reymond 2001), and how gene expression is altered in mutants (Mandaokar et al. 2003).

There are two basic types of DNA microarray, oligonucleotide arrays and cDNA arrays. Both can be used for similar types of analysis, including gene expression monitoring, but they differ in construction and design. Both types of microarray allow the user to identify quantitative differences in transcript abundance between two samples for all features, or genes, represented on the array.

DNA microarray technology is based on the principle that single strands of nucleic acid will anneal, or hybridize, with complementary strands. In all microarray production, some form of a gene, either amplified cDNA or an oligonucleotide, is laid down on a solid surface, these being the “features” on the array. The number of features in an array varies, depending on the method of construction. These features are the genes that are simultaneously monitored for changes in expression in a hybridization experiment.

Differential expression data are obtained in a cDNA microarray analysis. This is done by means of fluorescent probes, synthesized from a defined pool of expressed RNA. These probes are applied to the array allowing complementary sequences in the fluorescent probe to hybridize with the features on the array (Fig. 1). cDNA Microarray expression data is a comparison between two pools of RNA, for instance RNA isolated from leaves of *Arabidopsis* plants grown in normal light conditions and RNA from leaves of plants grown in high light conditions (Rossel et al. 2002). The two different pools are labeled with two different colors of fluorescent dyes, green and red, respectively. The abundance of each feature on the array can be measured by laser scanner for each of the treatment samples (represented by different colors of fluorescent probe) to determine the relative difference in abundance of each individual gene on the array.

Synthetic oligonucleotide arrays are produced by several private companies; e.g., Affymetrix, and can be purchased as a mass-produced catalog array, such as the Affymetrix GeneChip®, or as a custom designed array. These arrays may contain over one million “probes” which are generally short oligonucleotides, for instance 25-mers in the case of Affymetrix-produced chips. These short oligonucleotides are produced on a glass chip through a patented process of light-directed synthesis, or photolithography (Lipshutz et al. 1999). Many commercially produced oligonucleotide microarrays are constructed to be overlaid with one sample of chemically modified cDNA from one RNA sample, and control genes and spots within the array are used to eliminate background signals for the results.

The features of the oligonucleotide arrays are produced and catalogued, based on gene sequence information, and are completely dependent on the annotation of genes available for the organism. A consideration with

commercially produced oligonucleotide arrays is dependence on the company that produces them for array scanning hardware and scan analysis software. Although many array analysis programs are available for public use at no cost or can be purchased, the format of data obtained from oligonucleotide array companies is not always compatible with the publicly-available programs. A major advantage of the oligonucleotide system of DNA microarray over cDNA microarray is that the density of features possible with this type of array construction exceeds all other commonly used array types. Therefore, oligonucleotide arrays can be used to measure expression levels of more genes simultaneously, thereby increasing the amount of expression data obtained from each hybridization. Oligonucleotide arrays representative of the entire genome of several plant model organisms such as *Arabidopsis*, soybean and rice are available for purchase from several companies as catalog items at minimal cost (\$ 100 - \$ 500 a chip). However, producing custom arrays of other organisms through these companies may be cost prohibitive.

The alternative to synthetic oligonucleotide arrays is a printed cDNA microarray. The construction of this type of microarray requires PCR amplification of the feature genes, then depositing these genes on a solid surface, usually a microscope slide. Most commonly, cDNA libraries are randomly sampled and individual clones transferred to microtiter plates for partial sequencing to generate expressed sequence tags (ESTs). These clones of known sequence are then amplified by PCR in 96- or 384- well plates followed by cleanup and precipitation of the PCR products. The amplified, purified cDNA is then re-suspended in an appropriate salt buffer. A high speed arraying machine is used to deposit this cDNA on a glass microscope slide coated with a polymer that allows adherence of the DNA (Schena et al. 1995). Parts list and instructions for constructing this type of arraying machine are available publicly (<http://cmgm.stanford.edu/pbrown/mguide/index.html>), and there are numerous different types of scanning hardware and software available. One major advantage printed cDNA microarrays have over synthetic oligonucleotide arrays is the ability to make the arrays, as well as completing all steps in their production in-house. This means that, although there is substantial cost associated with initial production and set-up, generating additional arrays requires minimal time and cost.

The features present on a printed cDNA microarray are gene products. The method of obtaining these sequences, reveals important information about the temporal and physical locations of their expression. The method of random sampling of a cDNA library for partial sequencing and compiling an EST library means that the number of ESTs of the same gene obtained is related to the abundance of that gene in the cDNA library (Van der Hoeven et al. 2002). Often

the most abundant genes in the cDNA library are represented numerous times on a printed cDNA microarray serving as an internal control in the array.

The features printed on a cDNA microarray are typically PCR amplified products of 500 to 3,000 bases in length. The size of feature spots and the distance between spots varies greatly in cDNA microarrays, depending on the arraying machine and printing tips used. Up to 42,000 features can be printed on a cDNA microarray (<http://www.dkfz.de/mga/groups.asp?siteID=95>), less than half the number of features that oligo-arrays may contain. Printed cDNA microarrays from several different organisms as well as cDNA clone sets can be purchased from many companies and academic groups. There are also companies and groups that will produce custom-printed cDNA microarrays from any cDNA clone set. In addition, oligonucleotide clone sets representing the entire genome of a model organism such as *Arabidopsis* or mouse are available. These sets can be printed with a high speed arraying machine.

Quantitative Real-Time Reverse Transcription Polymerase Chain Reaction: Global surveys by microarray analyses can provide valuable information comparing transcript levels for large numbers of genes during different developmental stages or in response to a variety of environmental stimuli. However, quantitative real time RT-PCR analysis is a technique which offers a researcher the opportunity to focus on the expression changes of a particular gene or gene family. This technique allows for the quantification of the absolute or relative level of any plant gene transcript of interest and has the specificity and sensitivity to be used to successfully determine levels of expression for transcripts with very high or extremely low message abundance (reviewed by Gachon et al. 2004). It has been effectively used as a tool to accurately determine message levels for a variety of members of conserved plant gene families. This technique has enabled researchers to study and successfully analyze the expression of several enzyme gene families (Tan et al. 2003; Yokoyama and Nishitani 2001; Yokoyama et al. 2004). Quantitative real time RT-PCR is also well suited for expression profiling of signal transduction gene families and can be effectively used to determine expression changes in response to abiotic stresses (Charrier et al. 2002; Adams-Phillips et al. 2004; Cantero et al. 2006; Jang et al. 2004).

There are various types of quantitative real time RT-PCR but all are based on detection of PCR product as it is being amplified based on increased fluorescence (Fig. 2). Descriptions of the various types can be found at the web site: <http://www.gene-quantification.info>. The increase in signal follows a typical S-curve with the region of importance being the exponential phase. The number of cycles it takes to reach the point of exponential increase in signal is dependent upon the amount of starting template for the PCR reaction; lower template

concentrations will take longer (more cycles) to achieve the same signal level as compared to higher template concentrations. Since there are a variety of fluorescent tags available, it is possible to amplify more than one product in the same tube. Multiplexing requires that each product be labeled differently and monitored at the same time. Multiplex reactions have the advantage of each product experiencing identical amplification conditions, but this aspect is complicated by competitive amplification between products. It may be simpler to carefully control for variations between reactions than trying to optimize a multiplex reaction. Changes in the fluorescent signal are also useful for determining specificity of the reactions through analysis of a melting curve. As the temperature in an end-point PCR reaction increases, the fluorescent probes will undergo transition between high and low signals as their configuration changes. The specifics of this transition depend on the particular assay system used, as detailed below. Since annealing/dissociation temperatures are based on sequence, each detectable product should give only one signal peak.

The main difference between various methods of RT-PCR is in the detection of the product. Presence of PCR products can be detected through double stranded DNA (dsDNA) binding chemicals, hybridization probes, or hydrolysis probes. At the simplest, a dsDNA binding chemical indicates the addition of product since the starting cDNA template is single stranded (Fig. 2A). SYBR® Green (Molecular Probes, Invitrogen Corp. Carlsbad, California) is typically used in these instances. This method has the drawback of being very unspecific. Any extraneous PCR products produced in this case will increase the signal, leading to artificially high template concentration calculations. Methods employing dsDNA binding chemicals must be controlled very carefully to avoid this complication. For the same reason, multiplexing is not possible with dsDNA binders. The remaining methods are more specific, though they vary in terms of costs and ease of use.

Of the remaining two methods, hybridization probes are the more straightforward. Use of hybridizing probes increases the specificity of the measurements since increases in fluorescent signal will only occur in the presence of matching target sequences. These methods typically rely on the use of the Fluorescence Resonance Energy Transfer (FRET) principle to distinguish between free and hybridized probes. In the case of the LightCycler® (Roche, Basel Switzerland), two neighboring probes are labeled with fluorochromes that have FRET interactions. Successful hybridization by both probes allows for excitation of one fluorochrome and emission detection of the second (Fig. 2B). In a similar way, molecular beacons (The Public Health Research Institute, Newark, New Jersey) use a quenching moiety to decrease the signal from the unhybridized probe. Here FRET is used to lower the signal in an unbound probe, typically a hairpin structure brings the two elements in close proximity. Signal is

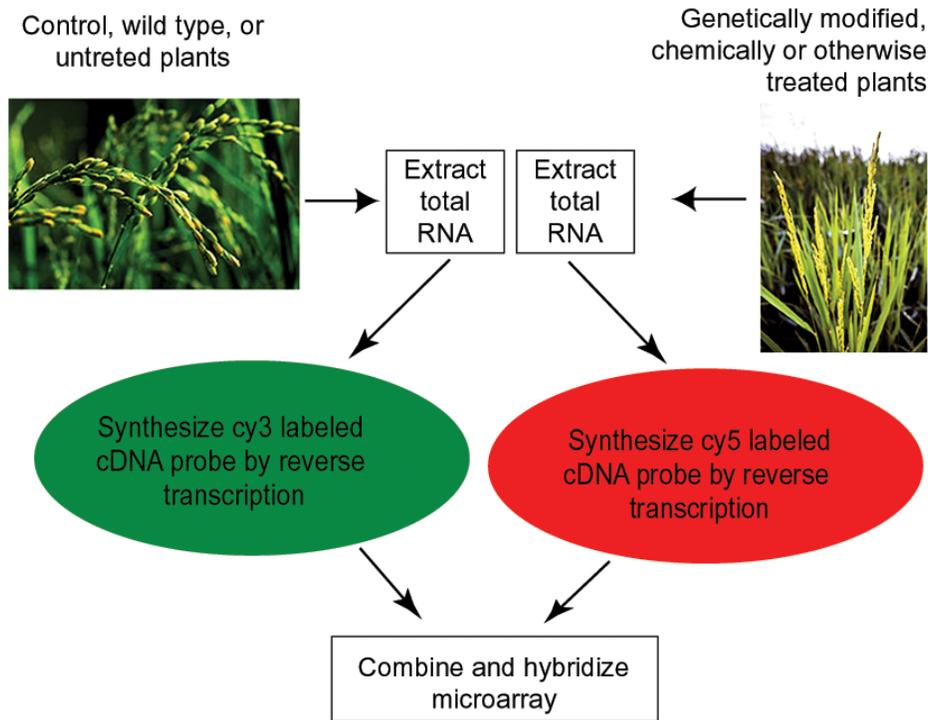
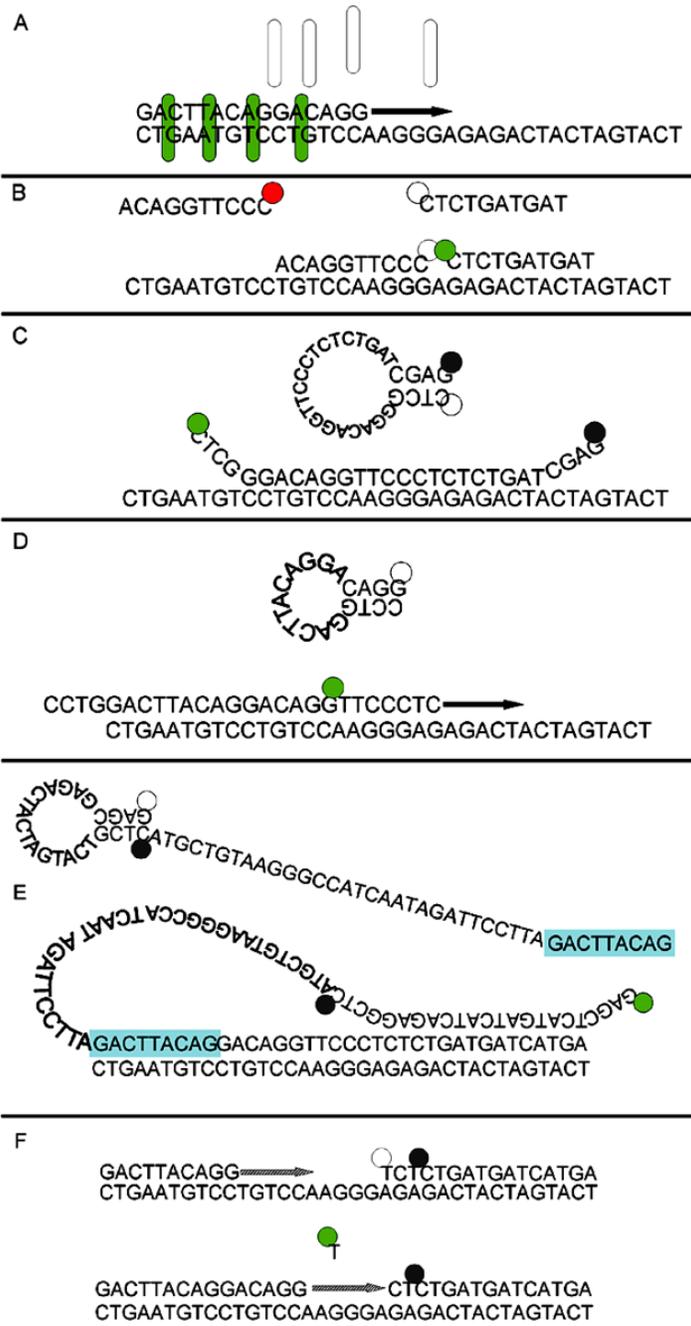


Fig. 1. Microarray Probe Synthesis Through Reverse Transcription Incorporation of Fluorescent Compounds.

Fig. 2. Various Methods of Quantitative Real Time RT-PCR.

Green circles represent the actively fluorescent compound being detected in each case. Empty circles represent quenched or non-fluorescent moieties. Grey circles represent quenching compounds. A. Double Stranded DNA binders such as SYBR® Green. Binding to double stranded product causes a compound to become fluorescent. The arrow represents the extension of a PCR primer by polymerase. B. LightCycler™ FRET activated fluorescent primers. Unbound LightCycler™ primers on top. The left primer is fluorescent when unbound, represented by red circle. When both primers are bound to the target (lower) FRET occurs and excitation of the first fluorochrome causes emission from the second which is detected by the PCR machine. C. Molecular Beacon. Unbound probe has hairpin structure (top) that keeps the quencher near the fluorochrome, preventing fluorescent emission. Binding to the target (lower) spatially removes the quencher from the fluorochrome thereby allowing fluorescence. D. LUX™ type primers. Hairpin structure of the primer quenches the fluorescent signal. Incorporation of the primer into the extending product increases fluorescent signal. The arrow represents the extension of a PCR primer by polymerase. E. Scorpions®. The blue box represents the primer portion of the Scorpion®. The hairpin at the end acts much like the Molecular Beacons, keeping the quencher in proximity to the fluorochrome. Scorpions® get incorporated into the product as primers and fluorescence occurs when Scorpion® tail binds to the same strand that the primer became incorporated into, thereby separating the quencher from the fluorochrome and allowing emission. F. TaqMan® type. Fluorochrome and quencher on the probe are in close proximity (top). As the polymerase exonuclease activity (represented by the striped arrow) approaches the probe its 5' exonuclease activity cleaves the probe releasing the fluorescent moiety from the quencher (bottom) and allowing fluorescence emission.



detected at each round of hybridization in the PCR cycle by the increase in fluorescence, when the probe is bound to the target product not allowing FRET to occur any longer (Fig. 2C).

Invitrogen Corp. (Carlsbad, California) uses a Light Upon eXtention (LUX™) method to detect successfully amplified products in a manner similar to molecular beacons. In lieu of FRET and the quencher however, one of the primers used for PCR is specially designed to have a terminal hairpin structure. The hairpin itself quenches the fluorescent signal such that detection of the product is measured by the increase of signal occurring from incorporation of the primer into a PCR product (Fig. 2D).

Scorpions® (Molecular Probes, Invitrogen Corp. Carlsbad, California) are another detection mechanism with elements similar to both molecular beacons and the LUX™ system. The structure of the Scorpion® probe has a primer region and a separate hairpin portion. The hairpin region is responsible for FRET quenching of the fluorochrome through proximity of a quencher. The literal twist to this process comes after incorporation of the primer into the product. The loop region of the quenching hairpin has matching sequence to the downstream amplified product. Through an intramolecular binding event, the loop causes the 5' fluorochrome to be moved away from the quencher, giving fluorescent signal. The binding of the 5' end of the Scorpion® primer to the same strand of the product appears much like a scorpion's arched tail, hence the choice of name (Fig. 2E).

The final variation of Real Time RT-PCR product detection uses cleavage of the probe as an indicator of successful amplification. In the case of TaqMan® probes (Applied Biosystems, Foster City, California), FRET quenches the signal from the fluorochrome at the 5' end of the probe. Signal is measured through the release of this fluorochrome via the 5' exonuclease activity of the polymerase as it approaches the hybridized probe (Fig. 2F). Signal in these instances is not dependent upon temperatures and annealing of the probe after its initial cleavage.

Choosing a method for detection depends largely on resources and time. SYBR Green® is certainly a cheaper option, but requires careful use and analysis due to the non-specific nature. LUX™ primers and Scorpions® have the benefit of requiring fewer components in the PCR reaction since the indicator becomes incorporated into the product. Specificity is excellent for molecular beacons and TaqMan® probes, but such experiments can be very costly and may also have difficult end-point analysis. TaqMan® probes, for example, require addition of SYBR Green® for melting curve analyses since fluorescence in this case becomes uncoupled from the actual products. When selecting a detection method,

researchers should take into account the required sensitivity, time available and monetary resources.

Regardless of the detection method used, analysis of the data involves conversion of signal into meaningful measurements of starting template. The two major methods for template calculations are the standard curve method or the comparative C_T method. For the first method, a series of measurements are made on known amounts of starting material in order to construct a standard curve reference for unknown measurement comparison. The standard curve method has the benefit of being useable for determining template copy numbers. Its main drawback is that the construction of the curve must be very accurate, which may be problematic when dealing with RNA samples. Standardization to an internal control, such as a housekeeping gene, may be required as well. The alternative, the comparative C_T method, does not give absolute copy numbers, but rather gives information regarding relative changes in gene expression. For this method, a control or reference gene is compared to the gene of interest between two conditions, typically being treated vs. untreated or different time points. The comparative C_T method requires that the gene of interest be amplified at a similar efficiency level to that of the control gene.

Care must be taken with selection of control genes in either case. It is necessary that the genes selected are not significantly altered in expression level in the situations being investigated. This may necessitate some additional work in novel experiments. It is also becoming more advisable to use several control genes for better confidence in outcomes. In a recent study a number of good candidates in *Arabidopsis* were determined (Czechowski et al. 2005). The increasing amounts of publicly available microarray data are invaluable for control gene selection for atypical treatments or conditions.

Serial Analysis of Gene Expression: Serial analysis of gene expression (SAGE) is another approach to high throughput gene expression profiling which has been used to study global expression changes in a variety of organisms (reviewed in Tuteja and Tuteja 2004). The use of this technique appears to be gaining popularity in plant studies and recent improvements may make it more cost effective and efficient (Robinson et al. 2004; Crawford et al. 2005; Diné et al. 2005). SAGE has some advantages over other high throughput systems, such as microarray and quantitative RT-PCR, especially when working on non-model plant species. The main advantage of the SAGE method is that when studying a plant for which there is no DNA database available, SAGE allows the researcher to obtain a short sequence of any cDNA of interest and use rapid amplification of cDNA ends (RACE) to obtain a full-length sequence of the cDNA of interest. SAGE has been used to analyze global gene expression in the model plant *Arabidopsis* (Lee and Lee 2003; Fizames et al. 2004; Robinson et al. 2004; Ekman et

al. 2005). One of the first SAGE studies in plant cells quantified the expression level of 5,291 genes in rice and identified genes differentially expressed after treatment with anaerobic conditions (Matsumura et al. 1999). Studies analyzing global expression in pine trees during wood formation, in maize root tips, and in mature potato tubers have also been done using the SAGE technique (Lorenz et al. 2002; Poroyko et al. 2005; Nielsen et al. 2005). SAGE has been particularly successfully in investigating expression changes during host-pathogen interactions. For example, SAGE was used to monitor expression changes in rice leaves during infection with blast fungus (Matsumura et al. 2003) and to examine expression differences between resistant and susceptible genotypes of cassava during viral infection with cassava mosaic disease (Fregene et al. 2004). SAGE has also been used to profile expression changes in certain plant fungal pathogens during key developmental stages (Thomas et al. 2002; Irie et al. 2003).

The SAGE method is very similar to EST sequencing and results in the construction of a cDNA library, but one in which short unique sequence tags (13-20 base pairs or more) are created for each message and then linked together in a library for sequencing. After cDNA synthesis, cDNAs are immobilized on streptavidin beads and a series of digests and ligations are performed. These steps create several strands of cDNA composed of the gene specific tags that are linked together called concatemers. This is followed by cloning, sequencing and bioinformatic analyses of that particular pool of cDNAs. This type of cDNA library is constructed and gene tag abundance is analyzed for different treatments, tissues, or experimental conditions of interest. Based on the abundance of gene specific tags in each library of clones, the relative amount of transcripts can be determined and compared for the large numbers of genes in each SAGE cDNA library.

One disadvantage of SAGE can occur with specificity of tags. This occurs if the individual tag for one gene overlaps with another gene. However, use of longer tags can partially overcome this problem (Ryo et al. 2000). Another problem with the SAGE method is the high number of biochemical steps required for constructing the SAGE cDNA library (Carulli et al. 1998). For example, incomplete enzyme digestions or mis-priming during amplification steps can lead to problems with the quality of the final product. It should also be noted that the short unique sequence tags obtained must be extracted from more complex sequences in the library and thus the analysis of the sequences requires special bioinformatics tools.

Conclusions

Recent advances in biotechnology have allowed plant biologists to get a more accurate and quantifiable estimate of transcription level of any gene than

previously available using techniques such as Northern Blot analysis, RNA protection assay and Reverse Transcription PCR. Microarray and SAGE analysis, described here, both provide a means of high-throughput analysis of gene expression level. These methods can be used to identify genes whose transcription level is altered in response to an abiotic or chemical stress from among several thousand genes. All three methods described here identify the relative expression level of a particular gene based on its abundance in various conditions and in relation to other genes. SAGE and quantitative RT-PCR can both be used to identify to copy number for a particular transcript (Velculescu et al. 1997; Wagatsuma et al. 2005).

As with all molecular biology methodologies, the data obtained from any of the techniques presented here should be independently verified by another technique. For example, results of differential expression from microarray analysis can be independently verified by quantitative RT-PCR (Salmi et al. 2005), or by Northern blot. The quantitative data obtained with these three techniques can be very useful for predicting proteins involved in the physiological response of a plant to stimulus or stress. However, it should be reiterated that all of these methodologies are used to monitor mRNA levels for specific genes, not protein levels directly. Not all mRNAs made by a cell get translated to proteins in a timely fashion. For example, mRNA can be stored in the cell, remaining inactive until some factor(s) triggers its activation and availability for translation (Spirin 1994). Moreover, mRNA can be targeted for degradation before it gets translated (Alizadeh et al. 2005). In all the techniques described here, the presence of the mRNA is monitored without regard to its translational activity. Suffice it to say that expression profiling using the three techniques reviewed here has successfully predicted potential biological roles for the protein products of many genes.

References

- Adams-Phillips LC, Barry CP, Kannan PJ, Leclercq LM, Bouzayen MJ and Giovannoni J** (2004) Evidence that CTR1-mediated ethylene signal transduction in tomato is encoded by a multigene family whose members display distinct regulatory features, *Plant Mol. Biol.* **54**: 387-404.
- Alizadeh Z, Kageyama S and Aoki F** (2005) Degradation of maternal mRNA in mouse embryos: selective degradation of specific mRNAs after fertilization. *Molec. Reprod. Devel.* **72**: 281-290.
- Botella JR and Arteca RN** (1994) Differential expression of 2 calmodulin genes in response to physical and chemical stimuli. *Plant Mol. Biol.* **24**: 757-766.
- Cantero A, Barthakur S, Bushart T, Morgan RO, Fernandez MP, Chou S, Clark G and Roux SJ** (2006) Expression profiling of the *Arabidopsis* annexin gene family during abiotic stress, germination and de-etiolation. *Plant Physiol. Biochem.* **44**: 13-24

- Carulli JP, Artinger M, Swain PM, Root CD, Chee L, Tulig C, Guerin J, Osborne M, Stein G, Lian J and Lomedico PT** (1998) High throughput analysis of differential gene expression. *J. Cell. Biochem. Suppl.* **30/31**: 286-296.
- Charrier B, Champion A, Henry Y and Kreis M** (2002) Expression profiling of the whole *Arabidopsis* Shaggy-like kinase multigene family by real-time reverse transcriptase-polymerase chain reaction. *Plant Physiol.* **130**: 577-590.
- Crawford AC, White J, Bundock P, Cordeiro G, McIntosh S, Pacey-Miller T, Rooke L and Henry RJ** (2005) Consistent production of cost-effective Long SAGE libraries. *Plant Mol. Biol. Rep.* **23**: 139-143.
- Czechowski T, Stitt M, Altmann T, Udvardi MK and Scheible WR** (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol.* **139**: 5-17
- Dinel S, Bolduc C, Belleau P, Borivin A, Yoshioka M, Calvo E, Piedboeuf B, Snyder EE, Labrie F and St-Amand J** (2005) Reproducibility, bioinformatic analysis and power of the SAGE method to evaluate changes in transcriptome. *Nucleic Acids Res.* **33**(3): Art. No. e26.
- Ekman DR, Wolfe NL and Dean JFD** (2005) Gene expression changes in *Arabidopsis thaliana* seedling roots exposed to the munition hexahydro-1,3,5-trinitro-1,3,5-triazine. *Environ. Sci. Tech.* **39**: 6313-6320.
- Fizames C, Munos S, Cazettes C, Nacry P, Boucherez J, Gaymard F, Piquemal D, Delorme V, Commes TS, Doumas P, Cooke R, Marti J, Sentenac H and Gojon A** (2004) The *Arabidopsis* root transcriptome by serial analysis of gene expression. Gene identification using the genome sequence. *Plant Physiol.* **134**: 67-80.
- Fregene M, Matsumura H, Akano A, Dixon A and Terauchi R** (2004) Serial analysis of gene expression (SAGE) of host-plant resistance to the cassava mosaic disease (CMD). *Plant Mol. Biol.* **56**: 563-571.
- Gachon C, Mingam A and Charrier B** (2004) Real-time PCR: what relevance to plant studies? *J. Exp. Bot.* **55**: 1445-1454
- Hennig L, Gruissem W, Grossniklaus U and Köhler C** (2004) Transcriptional programs of early reproductive stages in *Arabidopsis*. *Plant Physiol.* **135**: 1765-1775.
- Irie T, Matsumura H, Terauchi R and Saitoh H** (2003) Serial Analysis of Gene Expression (SAGE) of *Magnaporthe grisea*: genes involved in appressorium formation. *Molec. Gen. Genom.* **270**: 181-189.
- Jang JY, Kim DG, Kim YO, Kim JS and Kang HS** (2004) An expression analysis of a gene family encoding plasma membrane aquaporins in response to abiotic stresses in *Arabidopsis thaliana*. *Plant Mol. Biol.* **54**: 713-725.
- Keeler SJ, Boettger CM, Haynes JG, Kuches KA, Johnson MM, Thureen DL, Keeler CLb and Kitto SL** (2000) Acquired thermotolerance and expression of the HSP100/ClpB genes of lima bean. *Plant Physiol.* **123**: 1121-1132.
- Lee JY and Lee DH** (2003) Use of serial analysis of gene expression technology to reveal changes in gene expression in *Arabidopsis* pollen undergoing cold stress. *Plant Physiol.* **132**: 517-529.
- Lipshutz RJ, Fodor S, Gingeras TR and Lockhart DJ** (1999) High density synthetic oligonucleotide arrays. *Nature Genetics Suppl.* **21**: 20-24.

- Lopez-Juez E, Jarvis RP, Takeuchi A, Page AM and Chory J** (1998) New *Arabidopsis* cue mutants suggest a close connection between plastid- and phytochrome regulation of nuclear gene expression. *Plant Physiol.* **118**: 803-815.
- Lorenz WW and Dean JFD** (2002) SAGE Profiling and demonstration of differential gene expression along the axial developmental gradient of lignifying xylem in loblolly pine (*Pinus taeda*). *Tree Physiol.* **22**: 301-310.
- Mandaokar A, Kumar VD, Amway M and Browse J** (2003) Microarray and differential display identify genes involved in jasmonate-dependent anther development. *Plant Mol. Biol.* **52**: 775-786.
- Marin K, Suzuki L, Yamaguchi K, Ribbeck K, Yamamoto H, Kanesaki Y, Hagemann M and Murata N** (2003) Identification of histidine kinases that act as sensors in the perception of salt stress in *Synechocystis* sp. PCC 6803. *Proc. Natl. Acad. Sci.* **100**: 9061-9066.
- Matsumura H, Nirasawa S and Terauchi R** (1999) Transcript profiling in rice (*Oryza sativa* L.) seedlings using serial analysis of gene expression (SAGE). *Plant J.* **20**: 719-726.
- Matsumura H, Reich S, Ito A, Saitoh H, Kamoun S, Winter P, Kahl G, Reuter M, Kruger DH and Terauchi R** (2003) Gene expression analysis of plant host-pathogen interactions by SuperSAGE. *Proc. Natl. Acad. Sci. U.S.A* **100**: 15718-15723.
- Nielsen KL, Gronkjaer K, Welinder KG and Emmersen J** (2005) Global transcript profiling of potato tuber using LongSAGE. *Plant Biotech. J.* **3**: 175-185.
- Poroyko V, Hejlek LG, Spollen WG, Springer GK, Nguyen HT, Sharp RE and Bohnert HJ** (2005) The maize root transcriptome by serial analysis of gene expression. *Plant Physiol.* **138**: 1700-1710.
- Reymond, P** (2001) DNA microarray and plant defence. *Plant Physiol. and Biochem.* **39**: 313-321.
- Robinson SJ, Cram DJ, Lewis CT and Parkin IAP** (2004) Maximizing the efficacy of SAGE analysis identifies novel transcripts in *Arabidopsis*. *Plant Physiol.* **136**: 3223-3233.
- Rossel JB, Wilson IW and Pogson BJ** (2002) Global changes in gene expression in response to high light in *Arabidopsis*. *Plant Physiol.* **130**: 1109-1120.
- Ryo A, Kondoh N, Wakatsuki T, Hada A, Yamamoto N and Yamamoto M** (2000) A modified serial analysis of gene expression that generates longer sequence tags by nonpalindromic cohesive linker ligation. *Anal. Biochem.* **277**: 160-162.
- Salmi ML, Bushart TJ, Stout SC and Roux SJ** (2005) Profile and analysis of gene expression changes during early development in germinating spores of *Ceratopteris richardii*. *Plant Physiol.* **138**: 1734 -1745.
- Schena M, Shalon D, Davis RW and Brown PO** (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**: 467-470.
- Schenk PM, Kazan K, Wilson I, Anderson JP, Richmond T, Sommerville SC and Manners JM** (2000) Coordinated plant defense responses in *Arabidopsis* revealed by microarray analysis. *Proc. Natl. Acad. Sci.* **97**: 11655-11660.
- Spirin AS** (1994) Storage of messenger-RNA in eukaryotes - envelopment with protein, translational barrier at 5' side, or conformational masking by 3' side. *Mol. Reprod. Develop.* **38**: 107-117.

- Tan BC, Joseph LM, Deng WT, Liu LJ, Li QB, Cline K and McCarty DR** (2003) Molecular characterization of the *Arabidopsis* 9-cis epoxycarotenoid dioxygenase gene family. *Plant J.* **35**: 44-56.
- Thomas SW, Glaring MA, Rasmussen SW, Kinane JT and Oliver RP** (2002) Transcript profiling in the barley mildew pathogen *Blumeria graminis* by serial analysis of gene expression (SAGE). *MPMI* **15**: 847-856.
- Tuteja R and Tuteja N** (2004) Serial analysis of gene expression: Applications in malaria parasite, yeast, plant, and animal studies. *J. Biomed. Biotech.* **2**: 106-112.
- Van der Hoeven R, Ronning C, Giovannoni J, Martin G and Tanksley S** (2002) Deductions about the number, organization, and evolution of genes in the tomato genome based on analysis of a large expressed sequence tag collection and selective genomic sequencing. *Plant Cell* **14**: 1441-1456.
- Velculescu VE, Zhang L, Zhou W, Vogelstein J, Basrai MA, Bassett DE, Hieter P, Vogelstein B and Kinzler KW** (1997) Characterization of the yeast transcriptome. *Cell* **88**: 243-251.
- Wagatsuma A, Sadamoto H, Kitahashi T, Lukowiak K, Urano A and Ito E** (2005) Determination of the exact copy numbers of particular mRNAs in a single cell by quantitative real-time RT-PCR. *J. Exp. Biol.* **208**: 2389-2398.
- Yokoyama R and Nishitani K** (2001) A comprehensive expression analysis of all members of a gene family encoding cell-wall enzymes allowed us to predict *cis*-regulatory regions involved in cell-wall construction in specific organs of *Arabidopsis*. *Plant Cell Physiol.* **42**: 1025-1033.
- Yokoyama R, Rose JKC and Nishitani K** (2004) A surprising diversity and abundance of xyloglucan endotransglucosylase/hydrolases in rice. Classification and expression analysis. *Plant Physiol.* **134**: 1088-1099.