IJMPR 2018, 7(1), 01-08

International Journal of Modern Pharmaceutical Research

www.ijmpronline.com

ISSN: 2319-5878

IJMPR

Review Article

REVIEW ON GENE EXPRESSION STUDIES USING MICROARRAY

Yidnekachew Wolde*[1] and Abirham Bekele[2]

[1]University of Gondar, Collage of Natural and Computational Science, Department of Biotechnology, Gondar, Ethiopia.

[2]Welkite University, Collage of Natural and Computational Science, Department of Biotechnology, Welkite, Ethiopia.

Received on: 15/12/2017

Revised on: 01/01/2018

Accepted on: 22/01/2018

***Corresponding Author**

Yidnekachew Wolde

University of Gondar, Collage of Natural and Computational Science, Department of Biotechnology, Gondar, Ethiopia, B.O.B. 196.

yidibiot19@gmail.com,

ABSTRACT

The aim of this review such technologies allow the analysis of different constituents namely the transcripts, proteins and metabolites of the cell that help to deduce gene function. Microarray is one of the technologies permitting simultaneous interrogation of complex nucleic acid mixtures. The genome-wide examination of gene expression with this technology has provided transcript profiles of numerous biological samples for studying population variations, genotype-phenotype relationships, knock-out pools, whole genome genetic foot printing, and genome mismatch scanning. The information obtained by coupling of a microarray expression database and data analysis and viewing tools can serve as a foundation for proposing hypotheses and performing dedicated research. DNA microarray will substantially increase the speed at which differential gene expression can be analyzed and gene functions are elucidated. DNA microarray technology, however, limits expression studies to the mRNA level. Biological processes regulated at RNA degradation, protein synthesis, protein degradation or protein modification will be inert to the microarray approach. Ideally, it should be accompanied by analyses at the protein level.

KEYWORD: Microarray, cDNA, mRNA, genotype-phenotype relationships.

1. INTRODUCTION

Complete genome sequence information for dozens of organisms has opened a new avenue for biologists to explore these organisms at the genome level. This genome information has yielded the ability to perform high throughput, genome-wide screens of the gene function and has boosted application of a range of new technologies to functional plant gene analysis.

Such technologies allow the analysis of different constituents namely the transcripts, proteins and metabolites of the cell that help to deduce gene function. This is where genomics turns out to be functional genomics. Functional genomics, which includes transcriptomics, proteomics and metabolomics, aims to determine the biological function of genes and their products.

Different screening approaches and methodologies have been developed forming separate fields within functional genomics. Such screening approaches are transcriptomics, proteomics and metabolomics. Other methodologies of functional plant genomics like forward and reverse genetic approaches, RNA interference and transgenic are inefficient when it comes to performing high throughput functional genomics analysis, despite the existence of dense genetic and physical maps in model organisms (Allison *et al.*, 2006).

Therefore this seminar paper is initiated with the following objectives:

- ✓ To recognize efficient use of microarray technology will ultimate rely on the interconnection each expression methods.
- ✓ To analysis of diverse and accurate transcriptional profiles to data produced by other biological functional genomics tools.

2. Gene expression Analysis

Gene expression analysis is considered to be the most prominent and powerful tool for functional genomics. Gene expression might be divided into two main categories according to the number of genes investigated in a single experiment. First one of these categories is directed to the analysis of a single gene at a time, whereas the second one provides a global view of genome and its transcript profile. Whole genome transcript profiling, having advanced from one gene at a time methods, supplies information on both physical and functional annotation of genome and its components like regulatory regions (Donson *et al*., 2002).

Until the development of high throughput technologies, molecular analysis and functional annotation of genes generally focused on a single gene level. Measurement of mRNA abundance and changes in gene expression at the

level of single mRNA are performed by various techniques (Lockhart and Winzeler., 2000). These techniques are time consuming, labor intensive and less informative compared to high throughput screening methods. Moreover interactions between different genes and regulatory sequences or relations of pathways working together or in an opposite manner cannot be resolved by analyses at a single gene level. Most widely employed expression analyses at single gene level include Northern blotting and quantitative real time PCR (qRT-PCR).

3. Most commonly used techniques for gene expression analysis in plants cDNA-AFLP

RNA is converted into double stranded cDNA and then digested with two restriction enzymes: a frequent-cutter and a rare-cutter. Synthetic adapters are ligated to the cDNA ends and primers complementary to the adapter sequences (plus small extensions of 1, 2, or 3 nucleotides) are used to amplify fragments with asymmetric ends. These fragments are displayed on sequencing gels and compared. Specific fragments can be eluted from gels and sequenced to identify genes with differential expression (Bachem *et al.*, 1996).

3.1 SAGE

RNA is converted to double-stranded cDNA with a biotin attached to the oligo(dT) first strand synthesis primer and cleaved with a restriction enzyme, leaving 3'-most fragments immobilized onto streptavidin beads. After ligation with linkers onto the non-biotinylated end cDNAs are released, ligated together and amplified by PCR. Primer regions are removed from PCR products and the resulting fragments are ligated together into concatemers, cloned and sequenced. Finally, a software package identifies and counts the relative frequency of the sequences in the samples (Yamamoto *et al.*, 2001).

3.2 MPSS

Individual 3´ restriction fragments from a cDNA library are coupled to one of a million beads, amplified, arrayed and sequenced simultaneously for 20 residues to provide a million signature sequences. Transcripts can then be identified and the corresponding transcriptome quantitatively characterized (Brenner *et al*., 2000).

3.3 Real-time PCR

A variant from conventional PCR based on the detection and quantification of the fluorescence emitted by PCR products accumulated through the amplification process (Higuchi *et al*., 1993).

3.4 Microarrays

Marked samples are tested against sequences from thousands of different genes fixed on small solid supports (usually glass microscope slides). Depending on their sequence, the samples will hybridize with different spots in the array, which isanalyzed by specialized image software (Schena *et al*., 1995).

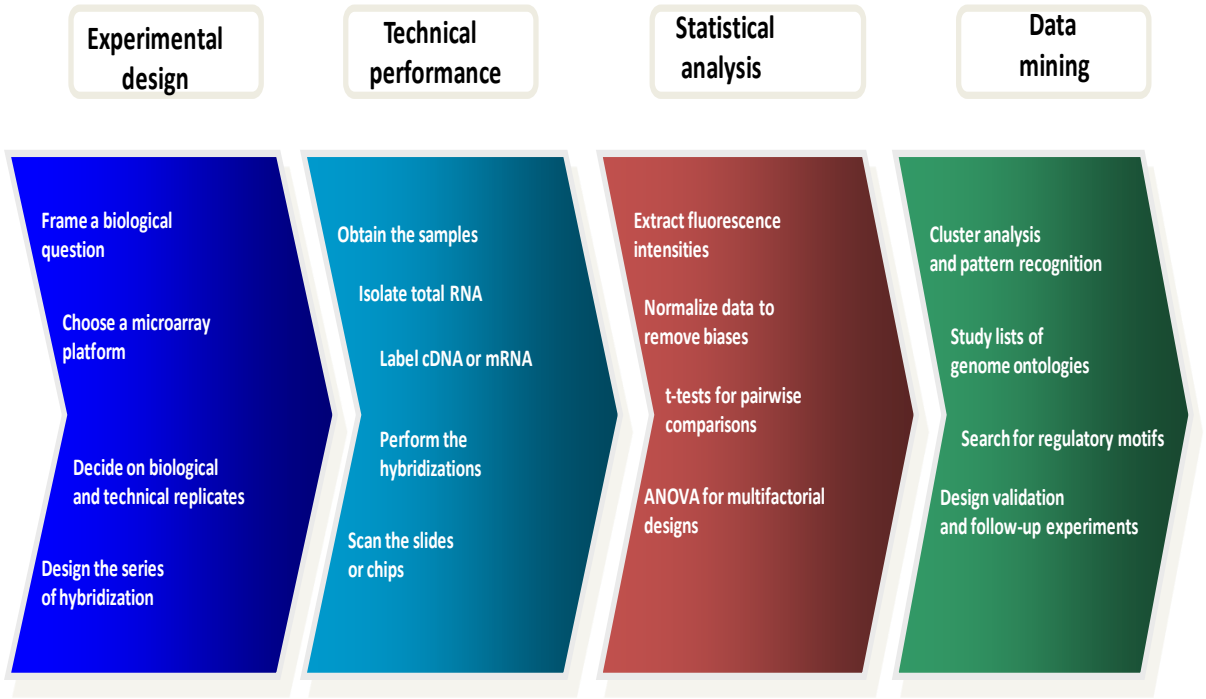


Figure 1: General scheme of microarray.

Microarray for Gene Expression Analysis

The microarray is one of the revolutionary technologies enabling comprehensive and high throughput surveys of

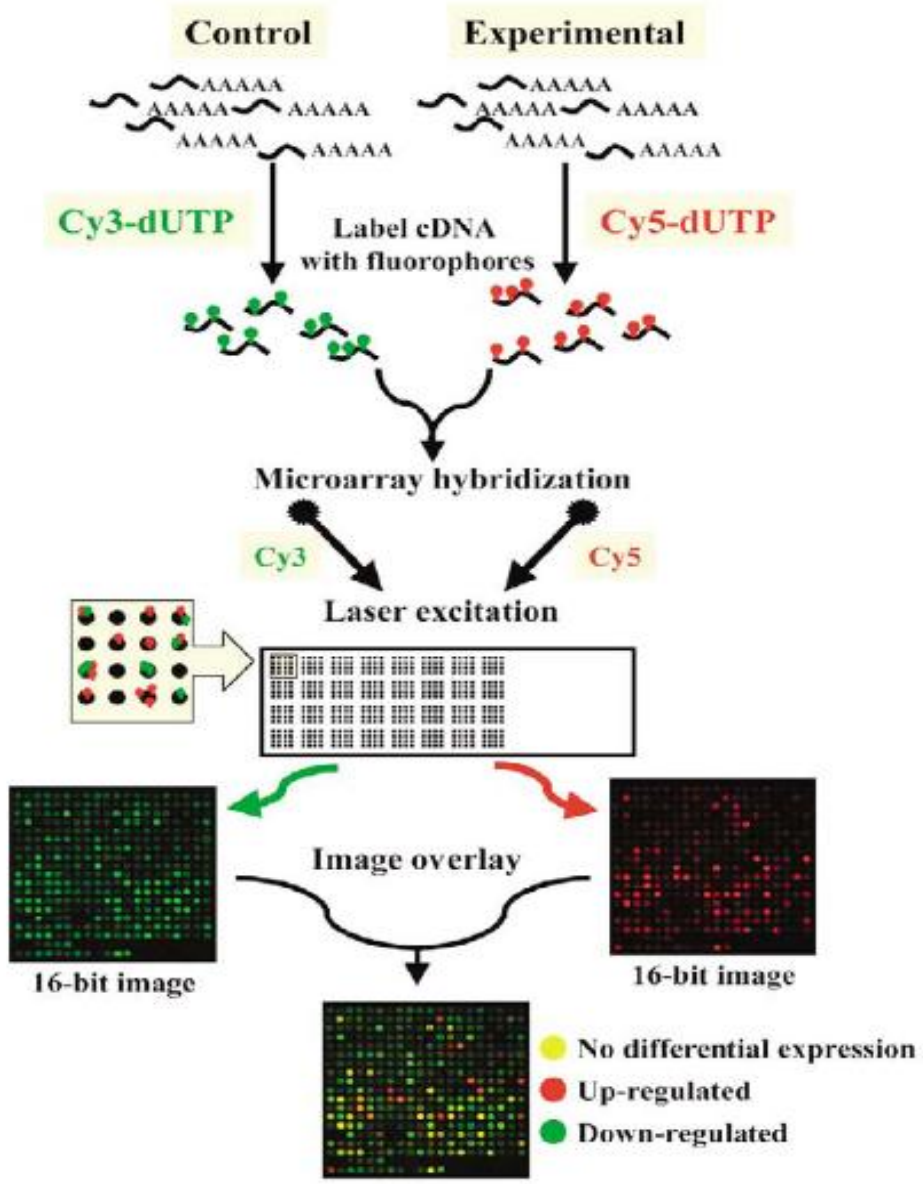
DNA or RNA molecules on a genome-wide scale (Lockhart and Winzeler., 2000). DNA microarray uses between hundreds and thousands of DNA probes arrayed on a solid surface to interrogate the abundance and/or binding ability of DNA or RNA target molecules in a single experiment. Coupled with bioinformatics tools and internet based infrastructure, this technology greatly extends the scope of biological research in the post-genomics era. The DNA probes that are used in a DNA The samples to be interrogated are labeled fluorescently or radioactively before the detection. When samples are applied to the DNA microarray, the DNA probes will capture the nucleic acid target molecules through sequence complementation. The strength of the fluorescent or radioactive signal from the captured targets reflects the abundance of the target molecules and/or the binding compatibility between the probe and target molecules. A charge-coupled device (CCD) or laser scanner can then be used to record this fluorescence or radioactive signals quantitatively. Hundreds and possibly thousands of DNA probes are arrayed in a single microarray to analyze many different targets in parallel, significantly improving the efficiency of this detection strategy. DNA microarrays are being increasingly used in plant biology research for the parallel analysis of gene expression.

cDNA microarrays have been the predominant method for the parallel analysis of gene expression in various biological processes, such as the maintenance of circadian rhythms, plant disease resistance, environmental stress responses, fruit and seed development, signaling in photomorphogenesis, and nitrate assimilation (Donson *et al*., 2002). GeneChip is a specialized microarray that uses *in-silico* synthesized DNA oligonucleotides as probes to detect the sequence similarity and abundance of target-DNA or -RNA molecules through complementary-sequence binding (Lipshultz *et al*., 1999). The standardized fabrication process and data-processing techniques used in this technology not only ensure data quality but also make data mining across a normalized database feasible (Zhu and wang, 2000).

Technical improvements in microarray production and data collection instruments are introduced frequently making this area very dynamic. The development of mask-free fabrication promises to reduce the cost and increase the efficiency of making oligonucleotide microarrays relative to current methods based on photolithography. A CCD camera-based slide scanner using white light illumination is available, providing greater versatility in fluorophore choice and minimizing the photobleaching often associated with laser-based scanners. Similarly, introduction of fluorophores with reduced susceptibility to photo-bleaching or other enhanced properties will also be beneficial. A DNA microarray robotic gridding instrument equipped with microarray could be amplified cDNA fragments or synthesized DNA oligonucleotides that have sequences that complement the target sequences. Thus, DNA microarrays are categorized into cDNA microarrays and DNA oligonucleotide probe microarrays. These DNA oligomers are either directly synthesized on the surface of the microarray or deposited onto the surface mechanically and covalently immobilized on the surface (Schena *et al*., 1995).
non-contact, piezoelectric dispensing system permits flexibility in printing surface choice.

4. Steps involved in Construction of DNA Microarray

According to (Alba *et al*.,2004). The basic steps involved in the preparation of DNA microarrays are: (1) the cloned DNA template collection, which is maintained in a 96-well format, is used for PCR amplification of the clone inserts with universal primers complementary to the clone vector, (2) Quality control of the purified PCR product includes DNA quantitation and DNA gel electrophoresis where the gels are manually scored as to whether PCR amplification led to the production of a single band as expected, (3) PCR products that fail to meet the quality control criteria stored in a database for future reference at the time of data analysis, (4) Four 96-well plates of PCR products are usually re-racked into one single 384-plate for high through-put robotic deposition onto glass slides.



Images by Dr. John Bennett, IRRI
Figure 2: DNA microarray hybridization and data acquisition.

5.1 Probe Design and array manufacturing

Probe design is an important issue for oligonucleotide microarrays. The sequence information is gathered from genomic EST sequence databases and handeled in silico for determination of probe complementary regions. The probes are designed to complement to the 3' end of expressed sequences. GeneChips involve probe pairs that each consists of 25 bp oligonucleotides. On the other hand, computer-based sequence design is not employed for cDNA microarray; rather the probes are PCR amplified from an already available cDNA library. Therefore the sequences of the probes are not known exactly.

Construction of arrays and placement of probes onto specific locations on the solid surface are done via separate techniques for oligonucleotides and cDNA array. GeneChips use light directed synthesis for the construction of high density DNA probe arrays using two techniques: photolithography and solid-phase DNA synthesis. Synthetic linkers, modified with photochemically removable protecting groups, are attached to a glass substrate and light through a photographic mask is directed to specific areas on the surface to produce localized photodeprotection (Lockhart *et al.*, 2000).

The first of the series of chemical building blocks, hydroxyl protected deoxynucleosides, is incubated with the surface, and chemical coupling occurs at those sites that have been illuminated in the precedingstep. Next, light is directed to the different regions of the substrate by a new mask, and the chemical cycle is repeated (Lockhart *et al.*, 2000).

cDNA microarrays use spotters or arrayers to transfer and place probe solution onto specific regions of the substrate. Principal components of an arrayer are computer controlled three axis robot and a unique pen tip assembly (Cheung *et al*., 1999).

5.2 Target preparation, Hybridization and scanning

Two types of microarrays also differ with respect to sample preparation and labeling. geneChip microarrays require isolation of RNA, ds CDNA synthesis, in vitro transcription of cRNA and labeling of cRNA with biotin during transcription. Similarly, cDNA microarray sample preparation starts with isolation of RNA and conversion to ss cDNA, however during cDNA synthesis, the target is labeled with fluorescent dyes, Cy3 or Cy5. Labeled cDNA targets are mixed and directly hybridized to arrays (Schulze and Downward., 2001).

The most common way to generate the fluorescent targets is direct incorporation of nucleotide-fluorophore conjugates (e.g. Cy3/Cy5-dUTP) into first strand cDNA with reverse transcriptase (Paolocci *et al*., 2005). Although this is the most time efficient procedure, the high Km of reverse transcriptase for dNTPs makes this a costly procedure.

5.3 Acquisition, transformation, and processing of microarray data

According to (Alba *et al*., 2004). The schematic steps involved for the analysis of cDNA microarrays is shown on Figure 6. Array scan was done immediately after they are washed/dried using a two- channel confocal microarray scanner and the associated ScanArray software. Scans were conducted at a resolution of 10µm with the laser power typically set between 70 and 85% of maximum and the photomultiplier tube typically set at 80% of maximum.

Excitation/emission settings were 543/570 µm and 633/670 µm for the Cy3 and Cy5 fluors, respectively. Raw fluorescence image data is typically saved as .tif files, which are subsequently converted to numerical signal data (.txt files and/or .xml files) using ImaGene software. To facilitate data transfer between investigators raw microarray data should be deposited in public repositories designed for expression data (preferably in the form of .tif files, .txt files, and/or .xml files); the importance of timely public release of microarray data cannot be overemphasized.

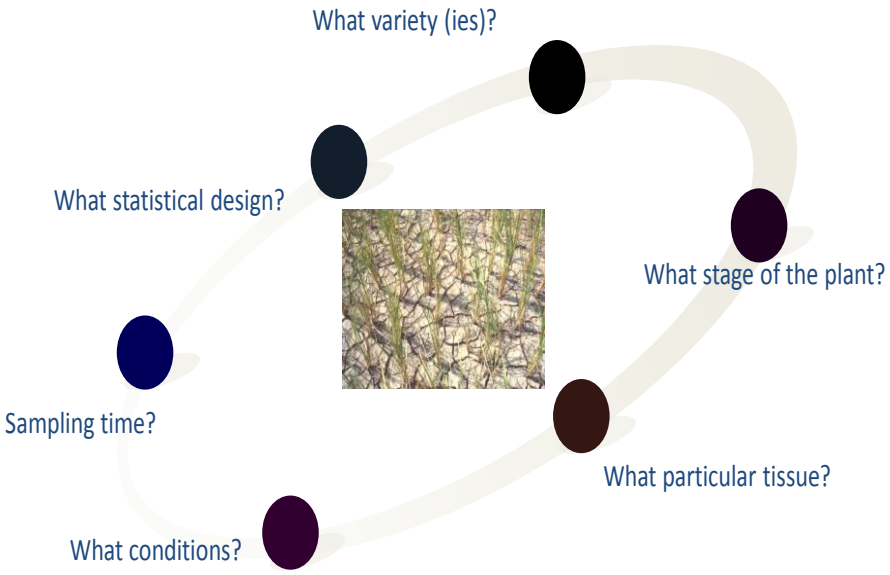


Figure 3: Sample question during gene expression.

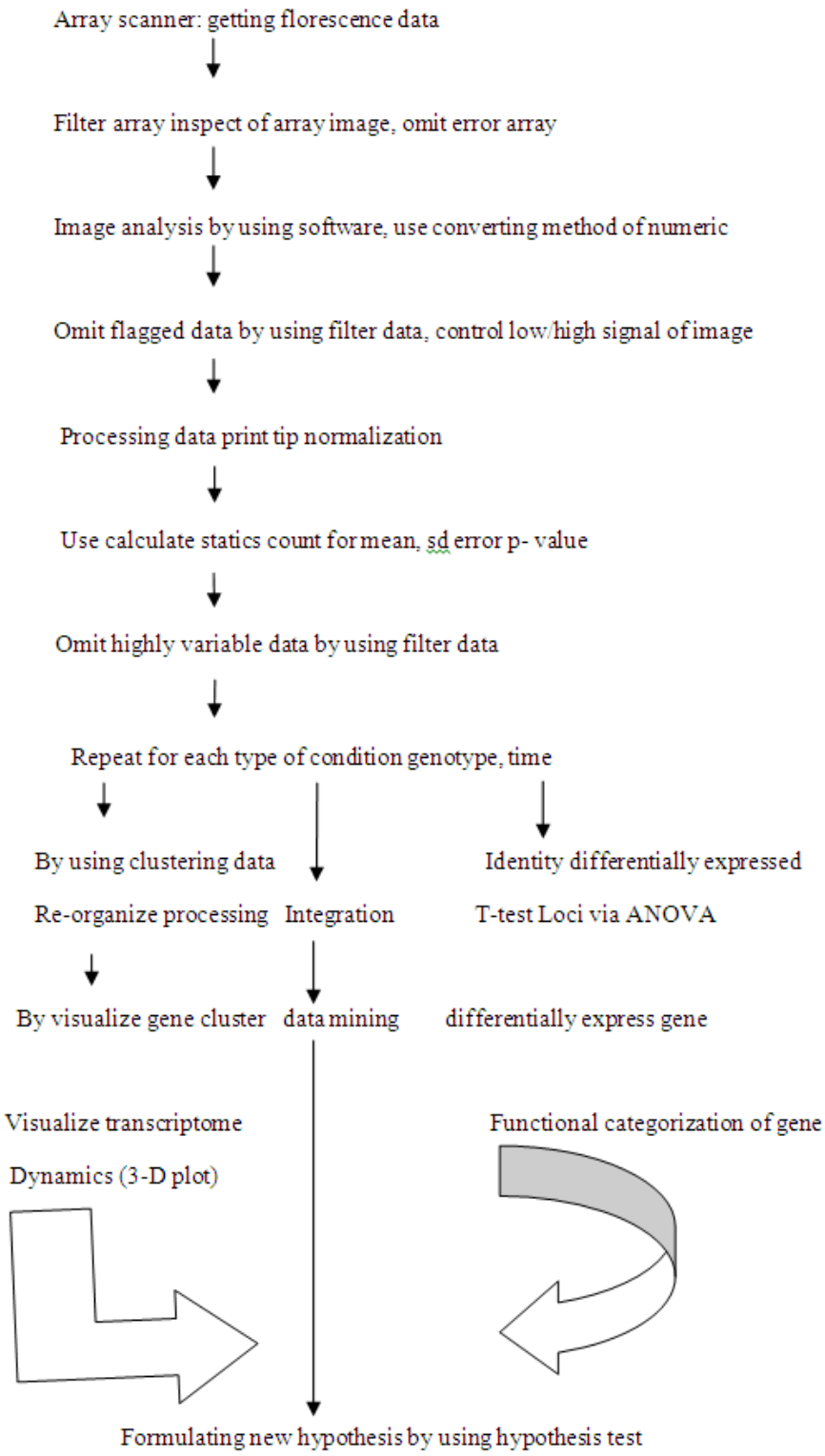
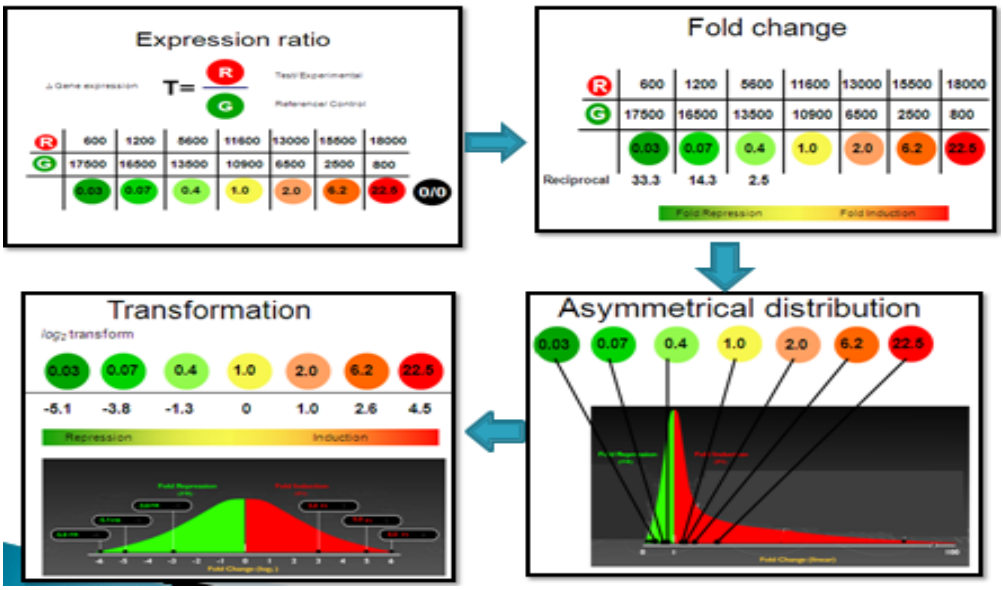


Figure 4: Analysis pipeline for microarray data.

6 Data Pre-Processing

Resolves systematic errors and bias introduced by experimental platform

- Data cleaning and transformation: removing flagged spots, background subtraction
- Within-array normalization: removes dye and spatial bias, brings cy3 and cy5 channels on equal footing
- Between-array normalization: enables comparison of multiple arrays, brings samples hybridized to different arrays on equal footing.



Images by Dr. John Bennett, IRRI

Figure 5: Data transformation of Microarray data.

7 Normalization

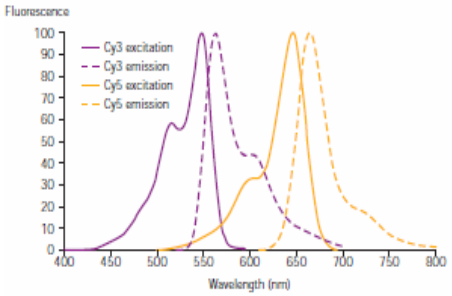
Compare ratio data from one microarray slide to another microarray slide, the ratio data needs to be normalized to correct for experimental variation.

Differences between the relative Cy3 and Cy5 signals in microarray slides may be due to one or more of the following:

- ✓ The amounts of mRNA used in the Cy3 and Cy5 labeling reaction.
- ✓ Efficiency of detection of the Cy3 and Cy5 by the detection system within the scanner.
- ✓ Relative incorporation differences of the Cy3 and Cy5 reverse transcriptase.

8 Visualization and Clustering

- ➢ Normalized microarray data is subsequently exported to any number of data visualization software for further analysis.
- GeneMaths (Applied Maths) software.
- Spotfire DecisionSite (Amersham Biosciences).
- ➢ Several standard statistical techniques are used to help interpret microarray data.



Excitation/emission settings are 543/570 nm and 633/670 nm for the Cy3 and Cy5 fluors, respectively.

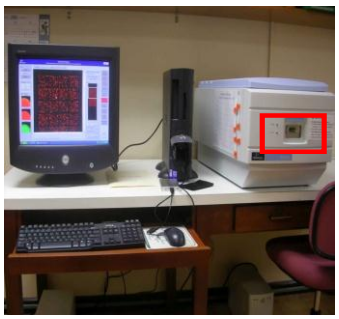
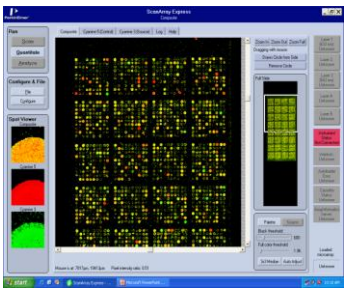
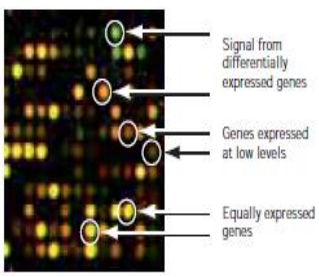


Image processing by Laser scanning





Images captured by scanner

Images by Dr. John Bennett, IRRI

Figure 6: Visualization of expression of gene.

Two main public repositories for expression data

- The Gene Expression Omnibus (GEO) at the National Center of Biotechnology Information (NCBI). GEO allows for the retrieval of gene profiles for several experiments these databases provide a repository and archive for expression data rather than a primary resource for expression data analysis.
- Array Express at the European Bioinformatics Institute (EBI)

9 REFERENCE

1. Alba, R., Fei, Z., Payton1, P., Liu, Y., Moore, S.L., Debbie, P., Cohn, J., D'Ascenzo, M., Gordon, S., Rose, J. K. C., Martin, G., Tanksley, S. D., Bouzayen, M., Jahn, M.M., and Giovannoni, J., ESTs, cDNA microarrays, and gene expression profiling: tools for dissecting plant physiology and development. *Plant J*., 2004; 39: 697-714.
2. Allison, D. B., Cui, X., Page, G. P., and Sabripour, M., Microarray data analysis: from disarray to consolidation and consensus. *Nature Rev. Genet.*, 2006; 7: 55-65.
3. Bachem, C. W. B., Van Der Hoeven, R. S., De Bruijn, S. M., Vreugdenhil, D., Zabeau, M., and Visser, R. G. F., Visualization of differential gene expression using a novel method of RNA finger printing based on AFLP: Analysis of gene expression during potato tuber development. *Plant J*., 1996; 9: 745-753.
4. Brenner, S., M. Bridgham, J. J., Golda, G., Lloyd, D. H., Johnson, D., Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.*, 2000; 18: 630-634.
5. Cheung, V. G., Morley, M., Aguilar, F., Massimi, A., Kucherlapati, R., Childs, G., Making and reading microarrays. *Nat. Genet*., 1999; 21: 15-19.
6. Donson, J., Fang, Y., Espiritu-Santo, G., *et al.*, Comprehensive gene expression analysis by transcript profiling. *Plant Mol. Biol.*, 2002; 48: 75-97.
7. Higuchi, R., C., Dollinger, F. G., and Watson, R., Kinetic PCR analysis: Real-time monitoring of DNA amplifcation reactions. *Nat. Biotechnol.*, 1993; 11: 1026-1030.
8. Lipshultz, R. J., Fodor, S. P. A., Gingeras, T. R., Lockhart, D. J, High density synthetic oligonucleotide arrays. *Nat Genet.*, 1999; 21: 20-24.
9. Lockhart, D. J., Winzeler, E. A., Genomics, gene expression and DNA arrays. *Nature*, 2000; 405: 827-836.
10. Paolocci, F., Bovone, T., Tosti, N., Arcioni, S., Damiani, F., Light and an exogenous transcription factor qualitatively and quantitatively affect the biosynthetic pathway of condensed tannins in Lotus corniculatus leaves. *J Exp. Bot.*, 2005; 56: 1093-1103.
11. Schulze, A., and Downward, J., Navigating gene expression using microarrays: a technology review. *Nature Cell Biol*., 2001; 3: 190-195.
12. Schena, M., Shalon, D., Davis, R. W., Brown, P. O., Quantitative monitoring of gene expression patterns with complementary DNA microarray. *Science*, 1995; 270: 467-470.
13. Yamamoto, M., Wakatsuki, T., Hada, A., and Ryo, A., Use of serial analysis of gene expression (SAGE) technology. *J. Immunol. Methods*, 2001; 250: 45-66.
14. Zhu, T., Wang, X., Large-scale profiling of the *Arabidopsis* transcriptome. *Plant Physiol.*, 2000; 124: 1472-1476.