

Global Relief One Chip at a Time

Lauren Stevenson

“We can banish extreme poverty in our generation — yet 8 million people die each year because they are too poor to survive. The tragedy is that, with a little help, they could even thrive.”

- Jeffrey D. Sachs

Poverty, hunger, and disease plague millions of people worldwide. Consequently, many organizations such as the United Nations have generated plan upon plan suggesting possible solutions for the alleviation of these international problems: one of which promises to reduce the extremes of these problems by the year 2015. Yet, the somewhat ambitious aims of the UN's Millennium Project may prove more successful than once thought. With its incorporation of technological advances such as microarrays, it may be possible to, one day, distribute phenotypically improved crops, characterize unknown diseases, and administer tailored medicines. How could this be possible? Much of the answer lies in the information revealed by the fully-sequenced genomes of model organisms and, as of late, our own.

With the unveiling of the genomes of model organism after model organism came a rapid development of techniques and equipment that could adequately support and analyze the masses of genetic information uncovered. On the forefront of these technologies was the microarray. This tool provides a way of looking at the expression levels of all the genes in an organism at one time when the organism is subjected to a particular condition. Before such technology existed, it was only possible to examine one gene, or at most, a few genes at a time. Thus, this technology has proven invaluable: comparing levels of gene expression in a healthy person versus those of a diseased person, for example, could provide insight into the gene(s) responsible for that particular disease, and consequently, enhance our understanding of the underlying mechanisms. Ultimately, this technology has the potential to revolutionize the way the world values and utilizes genetic information.

Microarrays are spots of DNA fixed onto a glass microscope slide, silicon chip, or nylon membrane in an orderly "array" such that each spot

corresponds to precisely one gene in the organism under study. There are many different types of microarrays available including cDNA, oligonucleotide, and Affymetrix arrays, yet all are essentially variations of a common theme. Perhaps the most commonly used arrays are the cDNA (complementary DNA) arrays simply because they are significantly less expensive. These cDNA microarrays are produced by amplifying an organism's genomic DNA by PCR. These PCR products are purified using purification columns to remove unwanted salts, detergents, PCR primers, and unwanted proteins present in the PCR mixture. These purified PCR products are then printed onto a glass microscope slide in approximately 5 nl aliquots using a computer-controlled robot (Duggan). In the end, the microarray is simply a glass slide with an organism's entire genome, or at least most of it, printed onto it. Arrays are available for a variety of model organisms such as fruit flies, rats, roundworms, and yeast.

After the array is successfully made, the sample to be tested must be generated. Two samples are needed for a microarray experiment: one sample that has been subjected to a particular condition and another sample that has remained under normal conditions, or rather, a control. If the organism under study were the model plant organism *Arabidopsis thaliana* and this plant was treated with chitin, a pathogen that elicits a defense response in most plants, one sample would be treated with chitin while the other would remain untreated.

The preparation of these samples is no easy task. Roughly outlined, the procedure begins with the mRNA of the treated and untreated samples. In separate experiments, the mRNAs of both samples are reversely transcribed with reverse transcriptase. Reverse transcription creates double-stranded cDNAs of the mRNAs, also known as first-strand DNA. This first-strand DNA is then purified and treated with a large subunit of DNA polymerase I

called Klenow and some special nucleotides that are tagged with fluorescent markers. This creates second-strand cDNAs that now have fluorescent tags attached to certain nucleotides in their sequences. Each sample is tagged with a different color: the treated sample is typically tagged with red and the untreated sample is typically tagged with green (Duggan).

Once the two fluorescently labeled samples are prepared, they are mixed together resulting in what is known as a complex probe. This complex probe is then incubated with the microarray containing the immobilized genes of the organism under study. The microarray and samples then undergo a process called hybridization probing, or rather, a technique that utilizes fluorescently labeled nucleic acid molecules as mobile probes to identify complementary molecules. During this hybridization probing, two complementary sequences, such as the immobilized target DNA on the glass slide and the fluorescently labeled mobile probes, find each other and hybridize, or lock together. The fluorescently labeled samples bind to the sites on the array corresponding to the genes that are expressed in the organism during that particular condition.

After hybridization, a type of scanner consisting of lasers, a special microscope, and a camera “excite” the fluorescent labels and capture the digital image of the array. A special computer program is then used to analyze the gene expression levels in both samples. In the microarray below (Fig. 1), the red spots represent genes that became more active, or up-regulated, when subjected to a particular condition. The green spots represent genes that lost activity, or became downregulated, and the yellow spots represent genes that did not undergo significant change in activity.

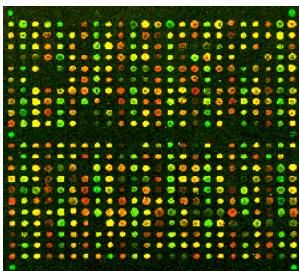


Figure 1: Example of a microarray from *Arabidopsis thaliana*.

Typically, if the treated sample expresses a particular gene more intensely than does the untreated sample, the gene in the treated sample is then investigated to determine why it becomes so active when subjected to that particular condition.

DNA microarrays provide a useful method for exploring the genomes of multiple organisms at the molecular level. Our present knowledge of how genes are regulated stems from the analysis of a limited number of genes. Thus, microarrays will herald a new area of investigation of gene regulation that promises to provide a much deeper understanding of how cells coordinate expression of thousands of genes. With microarrays, we can now identify all the genes of an organism that change expression when subjected to a particular condition, and hope to make sense of the cell's response to that condition. Such information provides key clues to the function of these genes, and ultimately their exact roles in certain defense, disease, and drug-response mechanisms. Once these mechanisms are known, the problems of hunger, disease, and ineffective medications should be easier to address.

One of the most promising solutions to the devastating hunger problem facing millions throughout the world has been the creation of genetically engineered crops. With the invention of technologies such as microarrays, it has become possible to pinpoint key components involved in plant-pathogen interactions. Knowledge of these particular components could allow for the creation of crops with improved phenotypes, or rather, crops with phenotypes that increase rates of survival when subjected to particular pathogens. These phenotypically improved crops, if distributed internationally could play a major role in satiating world hunger.

Much of the research in the field of plant genomics is conducted on the model organism *Arabidopsis thaliana*, the genetic analog of many crops. Presently, comparative plant genomics is contributing substantially to the analysis of genetically complex crops such as wheat. A comparative analysis between *Arabidopsis* and wheat shows that the exons in industrially and commercially important components such as starch synthase III are largely conserved in both plants (Franki). This degree of conservation indicates that,

in some *instances*, genes present in *Arabidopsis* can be used to identify corresponding genes in wheat.

Comparative plant genomics, coupled with microarray technology, seems to be the path towards the reality of these phenotypically improved crops: microarray technology has the capacity to show which genes are activated and to what degree in each species during a given condition via comparative experiments. One comparative experiment may show that the genes in *Arabidopsis* defense mechanisms are present, but not utilized in wheat. Therefore, the consequent activation or eventual incorporation of these defense genes into wheat could increase its pathogen resistance and, ultimately, increase its yield. These possibilities make it necessary to rapidly identify and characterize the genes and cellular components involved in *Arabidopsis* defense. Indeed, 800 million people worldwide are food insecure with 180 million being preschool children (Liang).

At present, many in the scientific community are utilizing microarray technology, not only in relation to model organisms such as *A. thaliana*, but as their primary means of analyzing our own genetic code. Comparing the levels of gene expression in a healthy person versus those in a diseased person can provide insight into the gene(s) responsible for that particular disease. With this insight, scientists can make way towards designing "tailored medications," or rather, medications specifically formulated according to a person's genetic composition (Lagay).

Responsible for the recent interest in tailored medications was the discovery of Single Nucleotide Polymorphisms (SNPs) in the late 1990s (Lagay). These SNPs are nothing more than mistaken pairings of nucleotides within a DNA sequence: a guanine paired with a thymine instead of with a cytosine. What makes SNPs important to medicine is the fact that they provide clues about the components involved in disease mechanisms. They also connect drug reactions to genes.

Each person's genetic material contains a unique SNP pattern that is made up of many different genetic variations. Most SNPs are not responsible for a disease state; rather, they act as tools for locating elusive disease genes. The fact that a SNP co-segregates in a pedigree with a disease phenotype implies that the SNP is not only

on the same chromosome as the disease gene, but that it is physically close (Lagay). Therefore, SNPs have successfully provided clues as to the whereabouts of other "candidate genes" in the vicinity that may also be contributing to the disease mechanism. Occasionally, a SNP may actually cause a disease and, therefore, can be used to search for and isolate the actual disease-causing gene.

Currently, researchers are utilizing SNP microarray analysis or mutation microarray analysis in efforts to establish SNP patterns for a variety of diseases. Once these SNP patterns are defined, SNP microarray technology can be used to test an individual for that SNP expression pattern and, consequently, determine whether he/she is susceptible to that disease.

Aside from disease-gene location, SNPs are also key components in drug response. A person's response to a particular drug depends heavily upon whether the drug's target cells have the proper receptors and also on how the individual metabolizes the drug. If a person rapidly metabolizes a drug, it can be rendered ineffective. Conversely, a person with slow metabolism can accumulate the drug in toxic amounts. Genes, and in some cases, SNPs control both of these factors—receptor binding sites and the enzymes involved in metabolism.

In a study of a set of enzymes called CYP3A, the enzymes that metabolize about fifty percent of all drug compounds, two SNPs were found that "quash" the production of these metabolizing enzymes (Lagay). People who carry either of the two SNPs were found to metabolize drugs more sluggishly. As noted before, sluggish metabolism can lead to the accumulation of certain drugs in highly toxic amounts, therefore, if a person underwent a SNP microarray analysis that would screen for the presence of these culprit SNPs, the adverse drug reactions caused by the sluggish metabolism could be eliminated simply by that person ceasing to take that particular drug. Better still; drugs could be designed to accommodate patients with certain SNP patterns who experience adverse drug reactions to the one-formula-fits-all drugs of today.

Since a single microarray can be used to screen 100,000 SNPs found in a patient's genome in a matter of hours, it is no surprise that right now

there is a race to catalogue as many of these genetic variations found within the human genome as possible (NCBI). Yet, not surprisingly, this “cataloguing race” has generated many obstacles for scientists. Microarray technology, as previously mentioned, is one of the most important experimental breakthroughs in molecular biology, and consequently, the generated data is fast overcoming the capacity for storage and analysis (NCBI). If researchers plan to combine the profiles of dozens or even hundreds of patients, then a database in which to store and catalogue this plethora of information is absolutely necessary.

Presently, a multitude of organizations have already recognized this problem and have taken steps towards establishing solutions. The National Center for Biotechnology Information (NCBI) launched the Gene Expression Omnibus (GEO) project in 1999 in response to the growing demand for a public repository for data generated from microarray experiments. The primary purpose of GEO is to provide a place where information from various microarray experiments could be shared by the scientific community at large, and ultimately, GEO is to become a resource that aids in the understanding of fundamental molecular and genetic processes that control health and disease (NCBI).

New technology has led directly to improved standards of living. With the aid of microarray technology, these standards can do nothing but increase. Microarray technology will hopefully, one day, touch the lives of millions.

References

Duggan, David J. *et al.* Expression Profiling using cDNA Microarrays. *Nature Genetics* vol. 21. 10-14. (1999).

Franki, Michael and Appels, Rudi. *Genome Biology* 2002. 3: reviews 1013.11013.5. <<http://genomebiology.com/2002/3/5/reviews/1013.1>>.

Lagay, Faith. “Pharmacogenomics: Revolution in a Bottle?” 3-10. February 5, 2003. <<http://www.ama-assn.org/ama/pub/category/7459.html>>.

Liang, G. H. and D. Z. Skinner. *Genetically Modified Crops: Their Development, Uses, and Risks.* 2-3. Food Products Press. New York. 2004.

National Center for Biotechnology (NCBI). Microarrays: “Chipping Away at the Mysteries of Science.” March 30, 2004. <<http://www.ncbi.nlm.nih.gov/About/primer/microarrays.html>>.

Sachs, Jeffrey D. "The End of Poverty." Mar. 14, 2005. <<http://www.time.com/time/archive/preview/0,10987,1034738,00.html>>.

Lauren Stevenson, a senior from Daphne, AL, is a Howard Hughes Medical Institute Undergraduate Research Intern. As a biology major and Blount Undergraduate Initiative minor, Lauren is active in the University Honors Program and Tri-Beta.