

DNA MICROARRAY WORKSHOP

**VIDO, University of Saskatchewan
SASKATOON,
FEBRUARY 26-27, 2004**

FEB 26 PM:

DNA MICROARRAY DESIGN AND DATA ANALYSIS

Neil Winegarden, University Health Network, Toronto

Sorin Draghici, Wayne State University, Detroit

Nicholas Tinker, Agriculture and Agri-Food Canada, Ottawa

Christopher Bowman, NRC-Institute for Biondiagnostics, Winnipeg

FEB 27 AM:

DNA MICROARRAY APPLICATIONS

Branimir Gjetvaj, Agriculture and Agri-Food Canada, Saskatoon

Palok Aich, VIDO

Brad Cooney, Amersham

Matt Links, University of Saskatchewan

Wei Dong, University of Saskatchewan

Janet Hill, NRC-Plant Biotechnology Institute

NRC · CNRC

Plant Biotechnology Institute



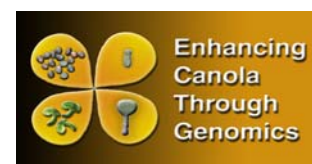
Agriculture and
Agri-Food Canada

Agriculture et

Agroalimentaire Canada



Ag-West Biotech Inc.



Thursday Feb. 26, 12:30pm-5:00pm
DNA Microarray experimental design and data analysis

12:30-1:00 Registration

1:00-1:05 Andy Potter, VIDO
Welcome remarks

1:05-1:10 Gijs van Rooijen, Genome Prairie
Introduction to the workshop

1:10-2:00 Neil Winegarden, University Health Network, Toronto
Microarrays: Data By Design

Microarrays, when used properly, can be an extraordinarily powerful tool for genomics research. However, the complexity of the technology provides numerous potential pitfalls to the uninitiated. Proper experiment design is critical to ensuring that there is a possibility of answering the question being asked. Despite the many wonderful and powerful data analysis tools now available - it is impossible to pull relevant data out of a poorly designed experiment. This talk will discuss some of the more common issues experienced by researchers wanting to design a microarray experiment. The concepts are generic in that they can be applied to nearly any type of study - be it an oncology or agricultural experiment being conducted. By fully examining the question at hand - and realising the power and limitation of each of the downstream tools being used, an appropriate experiment can be designed to allow the researcher to harness the power of microarray technology.

2:00-2:50 Sorin Draghici, Wayne State University, Detroit
Data Analysis Tools for DNA microarrays

We have developed a number of tools for data mining and functional analysis of high throughput gene expression experiments such as microarray and SAGE. The typical result of such an experiment is a list of tens or hundreds of genes found to be differentially regulated in the condition under study. Independently of the methods used to select these genes, the common task faced by any researcher is to translate these lists of genes into a better understanding of the biological phenomena involved. Currently, this is done through a tedious combination of searches through the literature and a number of public databases. We developed Onto-Express (OE) as a novel tool able to automatically translate such lists of differentially regulated genes into functional profiles characterizing the impact of the condition studied. OE constructs functional profiles (using Gene Ontology terms) for the following categories: biochemical function, biological process, cellular role, cellular component, molecular function and chromosome location. Statistical significance values are calculated for each category. Onto-Express is complemented by another set of tools including Onto-Design, Onto-Compare and Onto-Convert. Onto-Design is aimed at the researcher who are working in a hypothesis driven setup and wish to use in-house custom arrays. Onto-Design allows such researchers to select the best set of genes representing given selected pathways and biological processes. Onto-Compare allows the user to perform a functional comparison of any number of commercially available microarrays. This is useful if the research is focused on a small number of biological processes (e.g. 10-15) but the user wishes to use commercial arrays. Finally, Onto-Convert allows a quick translation between UniGene cluster IDs, NCBI accession IDs and custom, array specific IDs such as Affymetrix IDs. Onto-Convert is useful in various stages of common data analysis procedures or for instance, to those researchers who wish to use Affymetrix's custom printing service.

2:50-3:20 *Break, Networking*

3:20-4:10

**Nicholas Tinker, Ghislaine Allard, Jas Singh, AAFC, Ottawa
Progress and Pitfalls in the Analysis of a Brassica Transgene
Expression Study**

Microarrays have given scientists an unprecedented opportunity for discovery, but they have also delivered new challenges and false expectations. Scientists understandably ask for "instructions" on microarray analysis. Unfortunately, instructions are as varied as the problems being investigated, and are always contingent on factors such as error tolerance, resource availability, expertise, and what's already been done before the question is asked. A generalized set of instructions might go like this: (1) Plan your experiment, (2) do the wet stuff, (3) normalize and inspect data, (4) filter and average, (5) test hypotheses and/or identify new hypotheses. Rigorous hypothesis tests with meaningful error control are seldom done properly. A number of subtle statistical pitfalls can lead to false confidence and misleading conclusions. Moreover, many of the questions asked of microarrays are "fuzzy", so the answers are just as fuzzy, and many microarray studies should be considered primarily as exploratory investigations. The purpose of this presentation is to provide an honest examination of a single microarray experiment to identify potential pitfalls, realistic conclusions, and further recommendations.

4:10-5:00

**Christopher Bowman, Institute for Biodiagnostics NRC, Winnipeg
Multivariate Methods for Microarrays**

Microarrays can provide information on the simultaneous expression of thousands of RNAs in parallel. Standard univariate statistical tests (SAM, t test, fold change) can be used to sort through these expression patterns to extract a list of significantly expressed genes for further study. Univariate techniques have their shortcomings, however, in that they do not take gene interactions into account. If a single gene's expression pattern does not discriminate between the classes of interest, univariate tests will not be sufficiently powerful to analyze the data. Multivariate methods, both supervised (classification, neural nets, SVM), and unsupervised (clustering, PCA, SOM) allow one to detect the coordinated action of several genes simultaneously, and offer greater discriminatory power. This power comes at a price, and fully multivariate methods developed originally for data where the number of samples run into the tens of thousands are not always appropriate when analyzing microarray datasets with hundreds, dozens, or only a handful of samples. This talk will discuss multivariate methods that have been applied to microarray data analysis, illustrate their improved power over univariate methods, and present some of the cautions and caveats that anyone who intends to use such methods should be aware of.

5:00-7:00

Wine and Cheese Reception, Networking

**Friday Feb. 27, 8:30 am-12:30 pm
DNA Microarray applications**

- 8:30-9:00** **Branimir Gjetvaj and Andrew Sharpe, AAFC, Saskatoon**
Oligo design for microarray transcript profiling in polyploid Brassica
- 9:00-9:30** **Palok Aich, VIDO**
Microarray detection systems: Resonance Light Scattering (RLS) vs. fluorescence system
- 9:30-10:00** **Brad Cooney, Amersham**
Various labeling techniques of microarray probes for hybridization
- 10:00-10:30* *Break, Networking*
- 10:30-11:00** **Matt Links, University of Saskatchewan**
The application of microarrays to the study of abiotic stress in Triticum
- 11:00-11:30** **Wei Dong, University of Saskatchewan**
The role of RIZ1, a Rb interacting/ histone methyltransferase in human acute leukemia and identification of its molecular targets by cDNA microarray
- 11:30-12:00** **Janet Hill, NRC-Plant Biotechnology Institute**
DNA arrays application in diagnostics
- 12:00-12:10* *Closing remarks*

Thank you to the workshop sponsors: NRC-PBI, AAFC-Saskatoon, Genome Prairie, Ag-West Biotech Inc., FGAS project at Genome Prairie, VIDO and Amersham.