UHNMAC News

The University Health Network Microarray Centre Newsletter – Summer 2007

UHN Microarray Centre

Welcome to the Summer 2007 edition of the UHNMAC News!

This newsletter features a review of a recent publication by the MicroArray Quality Control (MAQC) Consortium. This study describes the experimental design and probe mapping efforts behind phase I of the MAQC project.
In addition, the feature report on Carbohydrate Microarrays highlights this emerging technique used in glycomic research.
We wish everyone a safe and happy summer!

ANNOUNCEMENTS

MicroRNA Array Service is coming soon
In the next few months, the UHNMAC will launch a new microRNA Array Service using the Exiqon and Agilent platforms. Please check our website, http://www.microarrays.ca/services/services.html for details.

1st Annual Toronto Functional Genomics Symposium postponed
The UHNMAC will re-schedule this event in the near future and details will be posted on our website. We thank all of our sponsors, speakers, and registered delegates for their support.

Prices for most Agilent Services have been lowered
In March 2007, the UHNMAC lowered most Agilent Service prices. Please visit our website, http://www.microarrays.ca/services/Agilentservice.html, for details.

FEATURE ARTICLE & REVIEW

MAQC Consortium. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nature Biotechnology, 2006, 24(9):1151

MAQC Project: Sparking scientific debate about the reliability of microarray technology

Although some publications have found excellent inter- and intraplatform reproducibility among different microarray platforms[1-3], other studies have found very little overlap in differentially expressed genes among various platforms[4,5]. Many researchers use DNA microarrays as a high-throughput screening tool for obtaining expression profiles and understand the limitations of microarray technology. It has even been suggested

Please see MAQC, page 2

FEATURE REPORT : Carbohydrate Microarrays

Glycomics, the functional study of carbohydrates in living organisms, has recently played a greater role in biological research and medical applications[1]. Carbohydrates are important components of glycoprotein and glycolipid cell-surface molecules, which are responsible for processes such as recognition, adhesion, and signalling[2]. These vital functions are known to play a role in growth and development, tissue repair[1], pathogen invasion[3], and tumour progression[4].

Please see Carbohydrate Microarrys, page 4

As always, general questions about microarrays can be addressed to help@microarrays.ca, orders for UHNMAC array products can be placed at orders@microarrays.ca, and questions about any of our services can be addressed to geneservice@microarrays.ca. If you have any suggestions for newsletter articles or questions you'd like addressed, please contact general@microarrays.ca.

MAQC

from page 1

that using a combination of microarray platforms may overcome the inherent biases of each method and that such an experimental approach will identify robust changes in gene expression[1]. However, members of the scientific community are concerned about the lack of concordance in results obtained from different microarray platforms. As microarrays evolve, and the idea of using them for diagnostic and regulated clinical experimentation becomes reality, such concerns must be addressed. The Microarray Quality Control (MAQC) project was initiated to provide quality control tools to the microarray community and to develop guidelines for microarray data analysis by providing large reference datasets along with accessible reference RNA samples[6].

International consortiums like the External RNA Control Consortium (ERCC) and Microarray Gene Expression Data (MGED) Society have provided the microarray community with RNA controls to enable consistency and reliability of gene expression platforms and standards for annotating microarray data, respectively. The MAQC project is also an international, community-wide effort with the goal of experimentally addressing the key issues surrounding the reliability of DNA microarray data and establishing operational metrics to assess the performance of seven microarray platforms[7].

In summary, Phase I of the MAQC project involved 137 participants from 51 organisations. Gene expression levels from two commercially available RNA samples (Universal Human Reference RNA (UHRR; Stratagene) and Human Brain Reference RNA (HBRR; Ambion)) in four titration pools (100% UHRR; 100% HBRR; 75% UHRR:25% HBRR; and 25% UHRR:75% HBRR) were evaluated on seven microarray platforms (Applied Biosystems, Affymetrix, Agilent Technologies, GE Healthcare, Illumina, Eppendorf, and National Cancer Institute). Each microarray platform was evaluated at three independent test sites and five replicates for each of the four sample types (6 platforms were one-colour) were assayed at each site (total of approximately 60 hybridisations per platform). The RNA samples were also tested on three alternative gene expression platforms (TaqMan Gene Expression Assays (Applied Biosystems), StaRT-PCR (Gene Express), and QuantiGene (Panomics)) to assess the relative accuracy of each microarray platform[8].

When analysing the data, the MAQC project based most of their results on a set of 12,091 common genes that are represented on all 6 of the high-density microarray platforms, albeit using different probe sequences[8]. Analysis of the MAQC data set found that one-colour microarray platforms had a median coefficient of variation (CV) of 5-15% for quantitative signal and a concordance rate of 80-95% for the qualitative detection call between sample replicates[8]. As expected, the variation increased when data from the different test sites using the same platform were included. The results indicate that, for the sample types chosen and these test sites, microarray results were repeatable within each test site, reproducible between test sites and comparable across platforms, even considering the difference in probe sequences across platforms as well as unique protocols for labelling and expression detection[8]. Other issues that may have affected the interplatform variability include possible annotation problems and the specificity of each probe on the array[8]. This study also highlighted differences in various performance metrics between microarray platforms. For instance, the data suggests that the Affymetrix platform had better intersite reproducibility, Illumina had better intrasite repeatability, and that some platforms (e.g., Agilent one-colour and Applied Biosystems) were more comparable to TaqMan assays[8].

...each microarray platform has made different trade-offs with respect to repeatability, sensitivity, specificity and ratio compression[8].

Much of the debate sparked by this study has to do with the way in which that was analysed. The MAQC Consortium suggests that relying on the

statistical significance (*P* value) instead of the actual measured quantity of differential expression (ratio or fold change) when identifying differentially expressed genes, may contribute to the lack of agreement between microarray platforms[9]. The MAQC group has been criticised for implying that prioritising genes by magnitude of effect is more productive than prioritising genes by the level of statistical significance[10], although the MAQC does recommend a nonstringent level of statistical significance ($P<0.05$ or 0.01) be used in conjunction with fold change[11]. MAQC emphasises that criterion such as sensitivity and specificity should also be considered when developing rules for determining which genes are differentially expressed, but that reproducibility is also critical as results which are not reproducible are of no use in scientific environments[11]. Another concern was that differences in the normalisation methods used by each platform might have led to discrepancies[10]. Another study has suggested that the impact of normalisation methods on the reproducibility of gene lists becomes minimal when the fold change, instead of the *P* value, is used as the ranking criterion for gene selection[9]. Also, the analysis of the MAQC data set does not include biology-based performance metrics such as Gene Ontology terms or pathways[8].

Critics of this study have suggested a flawed experimental design, specifically the small size (n=5)[10] and the choice of RNA samples[12]. The MAQC defends its sample selection; that such distinct reference RNA samples were deliberately chosen so that the technical accuracy of the different platforms could be determined. As part of the MAQC project, Shippy *et al*. have found that RNA titration samples are a valuable tool for assessing microarray platform performance and different analysis methods[13]. The MAQC Consortium points out that the comparability of microarray data in this study does not necessarily mean that the same level of consistency would be achieved in experiments where more biologically similar samples were compared[8].

Initial analysis of the MAQC data set indicates that each microarray platform has made different trade-offs with respect to repeatability, sensitivity, specificity and ratio compression[8]. The MAQC project provides a framework for assessing the potential of microarray technologies as a tool to provide reliable gene expression data for clinical and regulatory purposes[8]. This study concludes that the technical performance of microarrays as assessed in the MAQC project supports their continued use for gene expression profiling in basic and applied research and may lead to their use as a clinical diagnostic tool.

Further analyses of the MAQC reference data set (Phase I) has been published[14,15]. The results from Phase II of the MAQC project are scheduled for release in September 2008[6]. These results will most likely continue to fan the flames of debate over the reliability and reproducibility of microarray technology.

References:

1. Bosotti, R. et al. Cross platform microarray analysis for robust identification of differentially expressed genes. BMC Bioinformatics, 2007, 8 Suppl 1:S5
2. Wang, Y. et al. Large scale real-time PCR validation on gene expression measurements from two commercial long-oligonucleotide microarrays. BMC Genomics, 2006, 7:59
3. Kuo, W.P. et al. A sequence-oriented comparison of gene expression measurements across different hybridization-based technologies. Nature Biotechnology, 2006, 24(7):832
4. Tan, P.K. et al. Evaluation of gene expression measurements from commercial microarray platforms. Nucleic Acids Research, 2003, 31(19):5676
5. Miller, R.M. et al. Dysregulation of gene expression in the 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine-lesioned mouse substantia nigra. J Neurosci, 2004, 24(34):7445
6. http://www.fda.gov/nctr/science/centers/toxicoinformatics/maqc/
7. http://www.nature.com/nbt/focus/maqc/index.html
8. MAQC Consortium. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nature Biotechnology, 2006, 24(9):1151
9. Guo, L. et al. Rat toxicogenomic study reveals analytical consistency across microarray platforms. Nat Biotechnol, 2006, 24(9):1162
10. Klebanov, L. et al. Statistical methods and microarray data. Letter to the Editor, 2007, Nature Biotechnology, 25(1):25
11. Shi, L. et al. Reply to Letter to the Editor, Nature Biotechnology, 2007, 25(1):26\
12. Liang, P. MAQC papers over the cracks. Letter to the Editor, 2007, Nature Biotechnology, 25(1):27
13. Shippy, R. et al. Using RNA sample titrations to assess microarray platform performance and normalization techniques. Nature Biotechnology, 2007, 24(9):1123

14. Canales, R.D. et al. Evaluation of DNA microarray results with quantitative gene expression platforms. Nature Biotechnology, 2007, 24(9):1115
15. Patterson, T.A., et al. Performance comparison of one-color and two-color platforms within the MicroArray Quality Control (MAQC) Project. Nature Biotechnology, 2007, 24(9):1140

Carbohydrate Microarrays

from page 1

Applications of carbohydrate microarrays

Carbohydrate microarrays are emerging as a common technique used in glycomic research, as they can be used to characterise carbohydrate-cell interactions, determine the binding profile of carbohydrate-binding proteins, detect pathogens, and provide high-throughput screening of inhibitors of carbohydrate-protein interactions[1].

Insight into the specificity of cell-surface carbohydrate interaction with antibodies and receptors will advance the development of new therapeutics and diagnostic assays[4]. One study has shown that carbohydrate arrays can be used to monitor the immune response to carbohydrate epitopes at different stages during differentiation, metastasis, or treatment. Huang *et al*. synthesised the cancer antigen Globo H hexasaccharide, an epitope found on the cell surface of breast, prostate, and ovarian cancers, and its truncated sequences[4]. The arrayed saccharides were then used for the fluorescence-based binding analysis of two monoclonal anti-Globo H antibodies and the serum from breast cancer patients, to define the specificity of these antibodies[4]. This study found that the microarray platform was more effective and sensitive than the traditional ELISA method[4].

Another study has found that carbohydrate microarrays are ideal for whole-cell applications as the arrays present carbohydrate ligands in such a way that mimics cell-cell interactions[3]. Pathogen detection experiments can be completed in complex mixtures of cells or protein using the known carbohydrate binding epitopes of the pathogen of interest[3]. Since binding can be observed for low concentration bacterium in heterogeneous solutions, carbohydrate arrays could be used as a fast diagnostic tool[1].

Carbohydrate microarrays have also been used to screen for novel inhibitors of carbohydrate-protein interactions. A study by Bryan *et al*. used carbohydrate microarrays to screen for inhibitors of fucosyltransferases, enzymes critical to the synthesis of inflammation mediators[5]. Such studies could lead to novel therapeutics for inflammatory diseases such as arthritis and colitis.

Recently, a novel microarray technique called comprehensive microarray polymer profiling (CoMPP), has been used to provide insight into the structure and functions of plant cell walls[6]. The CoMPP technique combines the sequential extraction of glycans from various plant tissues and the generation of arrays which are then hybridised with monoclonal antibodies or carbohydrate-binding modules with specificities for cell-wall components[6].

Manufacture of carbohydrate microarrays

Carbohydrate microarrays can be made using standard robotic microarray printing technology. The most common method for preparing carbohydrate microarrays is by covalent attachment of chemically modified carbohydrates to derivatised (chemically modified) glass surfaces[7]. Other methods include non-covalent immobilisation of unmodified carbohydrates on underivatised surfaces, non-covalent immobilisation of chemically modified carbohydrates on underivatised surface, and the covalent immobilization of unmodified carbohydrate on a derivitised surface, a method that is still under investigation[7]. Since carbohydrate-protein interactions are relatively weak, the glycan immobilised on the array should be strongly recognised by the protein and properly oriented and spaced to allow multivalent interactions[7].

The Functional Glycomics Consortium has created a glycan microarray by coupling amine-functionalised glycans to N-hydroxysuccinimide (NHS)-activated glass slides[8]. The NHS-activated surface allows covalent attachment of glycans containing a terminal amine. This glycan array, which represents diverse and biologically relevant structures representing the terminal sequences of glycoprotein and glycolipid glycans, has been used to analyse most major classes of glycan binding proteins (GBP) including antibodies, intact viruses, and mammalian, plant, viral and bacterial lectins[8].

Glycan Binding Protein (GBP) assays

Following the manufacture of carbohydrate arrays, the slides can be used for GBP assays. The slides are incubated in either a one-step procedure with labelled proteins or a sandwich procedure in which the bound GBP is overlaid with a fluorescently labelled secondary antibody or GBPs pre-complexed with labelled antibodies[8]. Surface plasmon resonance (SPR), an optical technique for measuring the adsorption of material (in the case of carbohydrate arrays, the adsorption of GBP) on a metal surface, has also been used as an alternative to fluorescence-based detection methods[9]. Karamanska *et al*. found that SPR imaging of a glycoside array could be used to study plant lectin recognition[9]. The results found by SPR imaging were in agreement with those obtained by fluorescence-based carbohydrate arrays but with the added advantage of label-free analysis[9].

Lipopolysaccaride and glycoprotein microarrays

Lipopolysaccharide (LPS) arrays and glycoprotein arrays are also invaluable tools for glycomics research. Thirumalapura *et al*. published their investigation involving lipopolysaccharide (LPS) arrays for the detection of anti-LPS antibodies[10]. In this study, LPS, a major component of the outer membrane of Gram-negative bacteria, from several bacterial strains were immobilised on nitrocellulose-coated glass slides and hybridised with antibodies[10]. This study found that LPS arrays were about 100-fold more sensitive compared to conventional immunofluorescence assays. Zhao *et al*. used glycoprotein microarrays to screen a variety of lectins to identify glycosylation patterns in sera from normal, chronic pancreatitis, and pancreatic cancer patients[11]. The study suggests that altered glycan structures may have utility for the differential diagnosis of pancreatic cancer and chronic pancreatitis and identify critical differences between biological samples from patients with different clinical conditions[11].

References:

1. de Paz, J.L. and Seeberger, P.H. Recent Advances in Carbohydrate Microarrays. QSAR Comb Sci, 2006, 25(11):1027-1032
2. Dwek, R.A. Glycobiology: Toward Understanding the Function of Sugars. Chem Rev, 1996, 96:683-720.
3. Disney, M.D., and P.H. Seeberger. The Use of Carbohydrate Microarrays to Study Carbohydrate-Cell Interactions and to Detect Pathogens. Chem Biol, 2004, 11(12):1701-1707
4. Huang, C-Y *et al*. Carbohydrate microarray for profiling the antibodies interacting with Globo H tumor antigen. PNAS, 2006, 103(1):15-20
5. Bryan, M.C., *et al*. High-throughput identification of fucosyltransferase inhibitors using carbohydrate microarrays. Bioorg Med Chem Lett, 2004, 14:3185-3188
6. Moller, I. *et al*. High-throughput mapping of cell-wall polymers within and between plants using novel microarrays. Plant J, 2007, 50:1118-1128
7. Shin, I., *et al*. Carbohydrate Microarrays: An Advanced Technology for Functional Studies of Glycans. Chem Eur J, 2005, 11:2894-2901
8. Blixt, O. *et al*. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. PNAS, 2004, 101(49):17033-17038
9. Karamanska, R., *et al*. Surface plasmon resonance imaging for real-time, label-free analysis of protein interactions with carbohydrate microarrays. Glycoconj J, 2007, June 16; Epub ahead of print
10. Thirumalapura, N.R., *et al*. Lipopolysaccharide microarrays for the detection of antibodies. J Immunol Methods, 2005, 298(1-2):73-81
11. Zhao, J., *et al*. Glycoprotein Microarrays with Multi-Lectin Detection: Unique Lectin Binding Patterns as a Tool for Classifying Normal, Chronic Pancreatitis and Pancreatic Cancer Sera. J Proteome Res, 2007, 6(5):1864-1874