UHNMAC News

The University Health Network Microarray Centre Newsletter – Autumn 2008

UHN Microarray Centre

Welcome to the Autumn 2008 edition of the UHNMAC News!

This edition features a review of a recent publication that used UHNMAC Human CpG island (HCGI) 12K arrays to investigate the relationship between DNA hypermethylation and histone modifications, like H3K27triM, in gene silencing. A brief report about microarray software applications and the services offered by the UHNMAC Bioinformatics Team is also included.

Service Publications

Five articles have recently been published featuring data obtained using UHNMAC Services (Affymetrix, Agilent, and In-house array platforms). Click on the links to read more!

Affymetrix platform:

- Tone *et al.* Clin Cancer Res 2008 14(13):4067
- Jones *et al*. Mol Cancer Res 2008 6:819
- Lovegrove *et al*. Am J Pathol 2007 171(6):1

Agilent platform:

- Zippo *et al*. Nature Cell Biol 2007, 9(8):932

In-house cDNA arrays:

- Haulk *et al*. Small 2008 4(1):153

Feature Article & Review

Kondo Y, *et al.* Gene Silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. Nature Genetics, 2008, 40(6):741

Two epigenetic systems involved in the repression of gene activity are Polycomb group (PcG) proteins and DNA methylation[1]. A PcG protein, namely enhancer of zeste 2 (EZH2), is a member of the polycomb repressor complex 2 (PRC2) that methylates histone H3 lysine 27[2]. Polycomb-based histone H3 lysine 27 trimethylation (H3K27triM) is one of at least two histone modifications that mediate epigenetic silencing in cancer cells[3]. H3K27triM serves as a signal for the binding of

Please see H3L27triM, page 2

Data Analysis: Licensed software and freeware used at the UHNMAC

The purpose of this report is to familiarise researchers with the licensed software and freeware that is available for data analysis and to **inform researchers about the services offered** by the UHNMAC Bioinformatics Team.

The Bioinformatics Team is a valuable resource for both experimental design and data analysis. By consulting the group in the planning phase of the project, researchers can ensure they have the best experimental design in order to test

Please see Data Analysis, page 3

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.

We welcome any comments or suggestions about the newsletter, please contact general@microarrays.ca.

H3L27triM

from page 1

another polycomb repressor complex, PRC1, which blocks the recruitment of transcriptional activation factors and prevents the initiation of transcription[4]. The relationship between DNA hypermethylation and histone modifications, like H3K27triM, is not completely understood.

Prior to the publication of the study by Kondo and colleagues, it was thought that PcG-mediated methylation of H3K27 and *de novo* DNA methylation in cancers were linked[5-7]. Vire *et al*. suggest EZH2 controls CpG methylation through direct physical contact with DNA methyltransferases[6] and Schlesinger *et al*. used chromatin immunoprecipitation (ChIP) analysis to show that methylated genes in cancer cells are associated with nucleosomes containing H3K27triM[7].

... tumor-suppressor gene silencing in cancer by EZH2-mediated H3K27triM is mechanistically distinct from DNA methylation-associated silencing.

In contrast, this study by Kondo *et al*. suggests that tumor-suppressor gene silencing in cancer by EZH2-mediated H3K27triM is mechanistically distinct from DNA methylation-associated silencing. Using ChIP coupled with CpG promoter arrays from the UHN Microarray Centre, this study compared prostate cancer cells to normal prostate and found that up to 5% of the genes on the CpG promoter array were silenced in cancer cells by H3K27triM independent of DNA methylation. The global assessment of H3K27 modifications using CpG arrays was performed to find H3K27triM gene targets and 12 genes were found. To determine whether this enrichment at H3K27triM of these genes was a cancer-specific silencing event, Kondo *et al*. studied normal (non-cancer) cells and found that these 12 genes showed no or much less enrichment for H3K27triM. This data suggests that, for many genes, H3K27triM is used by cancer cells to silence genes that should be active in normal cells. The global assessment of H3K27 modifications by promoter array was also performed. Of the 200 genes most significantly bound to H3K27triM, 16% had CpG islands and 84% did not, suggesting that H3K27triM-mediated silencing mechanism tends to target non-CpG island promoters.

This study presents several lines of evidence that support the independence of H3K27triM and DNA methylation including: H3K27triM targets show no or low DNA methylation in their promoters (with a few exceptions); the inhibition of EZH2, the main methyltransferase responsible for H3K27triM, reduces H3K27triM, suppresses clonogenicity, and reactivates hundreds of genes including genes silenced by H3K27triM, but has no effect on silencing by DNA methylation; and transfection of one of the H3K27triM targets (RARβ2) into the cells led to rapid and stable silencing, accompanied by H3K27triM but no or minimal DNA methylation[8].

This research contradicts previous reports that PcG proteins and DNA methylation systems are mechanistically linked silencing pathways[6,7]. These contradictions may be explained by tissue- and cancer-specific differences related to the activation of specific silencing pathways. Future research in gene silencing mechanisms will need to investigate why DNA methylation in cancer affects some silenced PcG targets but not others and why there is variability between different cancers[8]. This study further adds to the complexity of epigenetic dysregulation in cancer and establishes H3K27triM-mediated silencing as a promising therapeutic target[8].

References

1. Jaenisch R and Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 2003, 33:245
2. Cao R and Zhang Y. The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Curr Opin Genet Dev 2004, 14(2):155
3. Jenuwein T and Allis CD. Translating the histone code. Science, 2001, 293:1074

4. Aoto T *et al*. Polycomb group protein-associated chromatin is reproduced in post-mitotic G1 phase and is required for S phase progression. J Biol Chem 2008, 283(27):18905
5. Widschwendter M *et al*. Epigentic stem cell signature in cancer. Nat Genet 2007, 39(2):157
6. Vire, E *et al*. The Polycomb group protein EZH2 directly controls DNA methylation. Nature 2006, 439:871
7. Schlesinger Y *et al*. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes to DNA hypermethylation and heritable silencing. Nat Genet 2007, 39:232
8. Kondo, Y *et al*. Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. Nature Genet 2008, 40(6):741

Data Analysis

from page 1

their hypothesis. There are also several advantages to having your data analysed by the Bioinformatics Team. First, experienced analysts have access to specialised licensed software for the analysis of gene expression, array Comparative Genomic Hybridisation (aCGH), hybridisation after differential methylation (DMH), chromatin immunoprecipitation (ChIP)-on-chip, and miRNA data, saving you thousands of dollars in licensing fees. Second, analysis protocols are frequently revised to ensure that end-users receive the most comprehensive analysis. In addition, the analysis protocol can be customised depending on the requirements of the customer. And lastly, years of experience working with microarray data and various analysis platforms will save you lots of time (and headaches). The UHNMAC data analysis service is available to customers and collaborators.

The licensed analysis software used at the UHNMAC includes GeneSpring, CGH Analytics, and ChIP Analytics, all from Agilent. GeneSpring, which is used for the analysis of Agilent, Affymetrix, and in-house arrays, provides statistical tools for the analysis of expression data and enables the understanding of microarray data in a biological context[1]. CGH Analytics and ChIP Analytics software is used for the analysis of Agilent aCGH and ChIP-on-chip data, respectively. CGH Analytics provides multiple views of the data including a data and experiment browser; three levels of zoom-in including Genome, Chromosome, and Gene views, and a tabbed view of the feature by feature probe data and annotations[2]. ChIP Analytics allows for the identification of probes, genes, and genomic loci that exhibit significant binding and provides reports that can be easily loaded into other analysis software[3]. In August 2008, Agilent released DNA Analytics 4.0, which allows users to visually explore, detect, and analyse aCGH, ChIP, and methylation data in context with the genome of interest[4]. The joint analysis tool will allow users to investigate the correlation between copy number measured by Agilent microarrays and expression levels measured in any platform using the same samples[4].

In addition to licensed software, the UHNMAC uses freeware that is available to the academic community. Freeware, such as Multiexperiment Viewer (MeV), Significance Analysis of Microarrays (SAM), Prediction Analysis of Microarray (PAM), Bioconductor, Onto-Express, Pathway-Express, and Cytoscape, have been used at the UHNMAC and are briefly described below. Other data analysis applications, such as Statistics for Microarray Analysis (SMA) and RMAExpress (Robust Multichip Average Expression summary), are also briefly described for those who want to learn more.

- The TM4 Suite consists of four applications; Microarray Data Manager (MADAM) facilitates the entry of data into a database and provides forms that simplify the tracking of experimental parameters and data needed for analysis; TIGR_Spotfinder is used for the rapid analysis of microarray images and quantification of gene expression; Microarray Data Analysis System

(MIDAS) reads files generated by Spotfinder and allows for the normalisation and filtering of data; Multiexperiment Viewer (MeV) reads several input file formats, including Affymetrix® and Genepix®, and incorporates algorithms for visualisation, classification, clustering, and statistical analysis[5]. These applications are available to the scientific research community at TIGR's Software Download Site.

- SAM, a statistical technique for finding significant genes in a set of microarray experiments, and PAM, a technique that performs class prediction and sample classification from expression data using the "nearest shrunken centroid method", were both created by a group at Stanford University[6-9].

- Bioconductor is primarily based on the R programming language and provides a number of packages (add-on modules for R) for the analysis and comprehension of genomic data[10]. Courses to learn more about Bioconductor and course material are available online.

- Onto-Tools, which include Onto-Express (OE) and Pathway-Express (PE), among many others, were developed at the Wayne State University Intelligent Systems and Bioinformatics Laboratory and can be downloaded from their website[11]. OE allows for the functional profiling of a set of genes based on Gene Ontology. OE consists of an annotations database and an associated web-accessible software tool. PE is a tool that translates a list of differentially expressed genes into a list of all associated pathways[11].

- Cytoscape is an open source bioinformatics platform for visualising biological pathways and molecular interaction networks, and integrating these networks with gene expression data[12]. The core platform allows for data integration and visualisation and additional features, such as network and molecular profiling analyses, are available as plugins[12]. Version 2.6.0 was released in April 2008, with the next major update (version 3.0) expected in 2009.

- Terry Speed's group at University of California, Berkeley has also developed several applications including an R package for the analysis of cDNA microarray data called SMA [13]. Group members also developed, a standalone program for Windows for Affymetrix Genechip® data using the Robust Multichip Average (RMA) expression (RMAExpress) summary[14]. RMAExpress does not require R, nor is it dependent on Bioconductor. Version 1.0 of RMAExpress was released in June 2008.

This report does not include a complete list of analysis software. Visit the bioinformatics links section of our website (www.microarrays.ca) for an extensive list of the data analysis software available for microarray analyses[15]. If you have any questions about data analysis or experimental design, please do not hesitate to contact us geneservice@microarrays.ca.

References

1. GeneSpring (Agilent) http://www.chem.agilent.com/Scripts/PDS.asp?lPage=27881
2. CGH Analytics (Agilent) http://www.chem.agilent.com/Scripts/PDS.asp?lPage=29457
3. ChIP Analytics (Agilent) http://www.chem.agilent.com/Scripts/PDS.asp?lPage=37324
4. DNA Analytics 4.0 (Agilent) http://www.chem.agilent.com/scripts/pds.asp?lpage=65223
5. TM4 Suite (TIGR) http://www.tm4.org/
6. Tusher, Tibshirani, Chu. Significance analysis of microarrays applied to the ionizing radiation response. PNAS 2001 98: 5116-5121
7. Significance Analysis of Microarrays (SAM) http://www-stat.stanford.edu/~tibs/SAM/
8. Tibshirani, Hastie, Narasimhan, Chu. Diagnosis of multiple cancer types by shrunken centroids of gene expression. PNAS 2002 99:6567-6572
9. Prediction Analysis of Microarrays (PAM) http://www-stat.stanford.edu/~tibs/PAM/
10. Bioconductor http://www.bioconductor.org/GettingStarted
11. Onto-Tools (Wayne State University) http://vortex.cs.wayne.edu/projects.htm
12. Cytoscape http://www.cytoscape.org/features2.php
13. Statistics for Microarray Analysis (SMA; Terry Speed's Group) http://www.stat.berkeley.edu/%7Eterry/Group/index.html
14. RMAExpress http://rmaexpress.bmbolstad.com/
15. UHNMAC Bioinformatics webpage http://www.microarrays.ca/links/revised_bioinf.html