UHNMAC News

The University Health Network Microarray Centre Newsletter – Summer 2009

UHN Microarray Centre

Affymetrix Authorized Service Provider

Agilent Certified Services Provider Microarray-Based Genomic Analysis

This edition of the newsletter summarises the internal validation of the Illumina iScan System for miRNA expression profiling. Part of this validation involved a comparison with Affymetrix and Agilent miRNA platforms. All three miRNA platforms are now available at the UHNMAC. If you missed the report on the validation of the Illumina gene expression platform, it can be found in the Spring 2009 newsletter. Also included is the summary of a paper that discusses miRNA array data normalisation and the impact of various methods on sensitivity, specificity, and fold-change measurements relative to qPCR.

Validation of the Illumina miRNA platform: comparison with Affymetrix and Agilent

As part of the internal validation of the Illumina iScan System, the goal of this study was to compare the Illumina Human microRNA (miRNA) array with two other array-based platforms, Affymetrix and Agilent, which are also available at the UHNMAC. This report briefly describes the results obtained from the miRNA platform comparison.

Please see miRNA platforms, page 2

Comparison of Normalisation Methods for miRNA Microarray Data

Summary of: Pradervand S, et al. Impact of normalisation on miRNA microarray expression profiling. RNA 2009, 15(3):493-501

Micro-RNA (miRNA) are regulators of mRNA translation and stability that play key roles in many biological processes. There are three forms of miRNA: long pri-miRNA, hairpin pre-miRNA, and the 19- to 25-nucleotides-long mature miRNA. miRBase (http://microrna.sanger.ac.uk/) is the central online repository for miRNA nomenclature, sequence data, annotation and target

Please see Normalisation, page 5

Announcements

Thanks to everyone who attended the Open House in June for your support. We had over 250 visitors!

Nanostring Service

The Nanostring Service will be launched in August. To learn more about this technology, join us on Sept 24th for a technical seminar or visit our Nanostring webpage.

Upcoming Events

Fall 2009 User Group Schedule:
Thursday, September 24th
Thursday, October 22nd
Thursday, November 26th

Meetings are held in TMDT 4-204 from 3-4pm. Please check our website for topics/guest speakers.

miRNA platforms

from page 1

miRBase is a public web-based resource for miRNA data established by the Sanger Institute(1). miRBase includes a searchable database of published miRNA sequences and annotation, a database that predicts miRNA targets in animals, and provides gene hunters with unique names for novel miRNA genes (1). As new miRNA genes are identified, updated versions of the miRBase database are periodically released (Table 1). Manufacturers of miRNA arrays use the miRBase database content to design the array probes.

Table 1. Total number of miRNAs in the Sanger miRBase database (1).

miRBase version	Release date	Number of entries (all species)
9.0	October 2006	4361
9.1	February 2007	4449
10.0	August 2007	5071
11.0	April 2008	6396
12.0	September 2008	8619
13.0	March 2009	9539

Materials and Methods

For the Illumina and Affymetrix platforms, 200 ng of placenta (Ambion) and brain (Ambion) total RNA were labelled and hybridised according to the manufacturer's protocol, briefly described below (Table 2). Experiments were performed in triplicate. The Agilent protocol recommended the labelling of 100 ng placenta and brain total RNA, and these experiments were performed in duplicate.

Platform comparison:

- Signal from each miRNA array/platform was normalised using the quantile method
- Data was then filtered to include only probes that were present at least once across all samples
- Control probes, as well as non-human probes from the Affymetrix GeneChip, were removed; the set of human ("hsa") miRNA probes common among all three platforms remained
- T-tests were then performed on the common human miRNAs to identify differentially expressed miRNAs with a p value < 0.05 and fold change >2

Results

- For the Illumina platform, all hybridisations passed internal QC metrics (including array hybridisation controls, negative controls, polyadenylation controls, oligo annealing controls, mismatch controls, extension controls, and contamination detection controls) and all arrays met the Illumina acceptance criteria
- Illumina data was highly reproducible among technical replicates (correlation >0.99); Figure 1
- The three platforms had a set of 354 common human miRNA probes (after filtering as described above); Figure 2
- Following t-tests ($p<0.05$ and fold change >2; Benjamini & Hochberg method used for multiple testing correction) on the 354 common human miRNA probes, 156 were found to be differentially expressed between the placenta and brain samples. Of the 156 differentially expressed miRNAs, 69 (44%) were common among all three platforms; Figure 3

Discussion

In this study, each platform utilised a different labelling and hybridisation protocol and each manufacturer used a different version of the Sanger miRBase database for probe design. As a result, this platform comparison was limited to the human miRNAs common among all three platforms. It is also important to remember that although the probes representing the same miRNA gene were compared, the probes could have different sequence lengths and/or configurations, depending on the platform. Of the 354 common human miRNA probes, 156 were differentially expressed (t-test with p value < 0.05 and fold change >2) and 69 (44%) of the 156 probes were common among the three platforms.

Table 2. A brief description of the arrays and labelling methods used for each of the platforms compared in this study.

Manufacturer & Array Type	miRBase version	Number of human miRNAs	Labelling method	Unique array features and advantages of the platform
Affymetrix miRNA GeneChip v1	11.0	846 + 21 viral miRNA	FlashTag™ Biotin RNA Labelling kit (Genisphere), which uses Genisphere's proprietary 3DNA signal amplification technology	GeneChip also includes > 900 probes representing human small nucleolar RNAs and other sequences; array covers 71 organisms including human, mouse, rat, monkey and canine; over 46K probes per array; 4 copies of each miRNA probe per array
Agilent Human miRNA array v1	9.1	470 + 64 viral miRNA	Agilent Labelling (v1) uses CIP (GE) & T4 RNA ligase, Bio-Rad Micro Bio-Spin 6 columns; this protocol does not involve fractionation or amplification	All probes on the Agilent miRNA array contain a 5` hairpin, abutting the probe-target region, to increase target and size specificity; probes represent mature miRNA species; total of 15K probes per array arranged in an "8-up" (8 arrays per slide) format; custom arrays are also available
Illumina Human miRNA BeadChip v1	9.0	735	Adaptation of the DASL (cDNA-mediated Annealing, Selection, Extension, and Ligation) assay; two-step discrimination based on hybridisation followed by enzymatic primer extension; suitable for clinical FFPE samples	High feature redundancy; in addition to miRNA probes based on miRBase sequences, probes also designed using Illumina sequencing data; probes arranged in a "12-up" (12 arrays per slide) BeadChip format

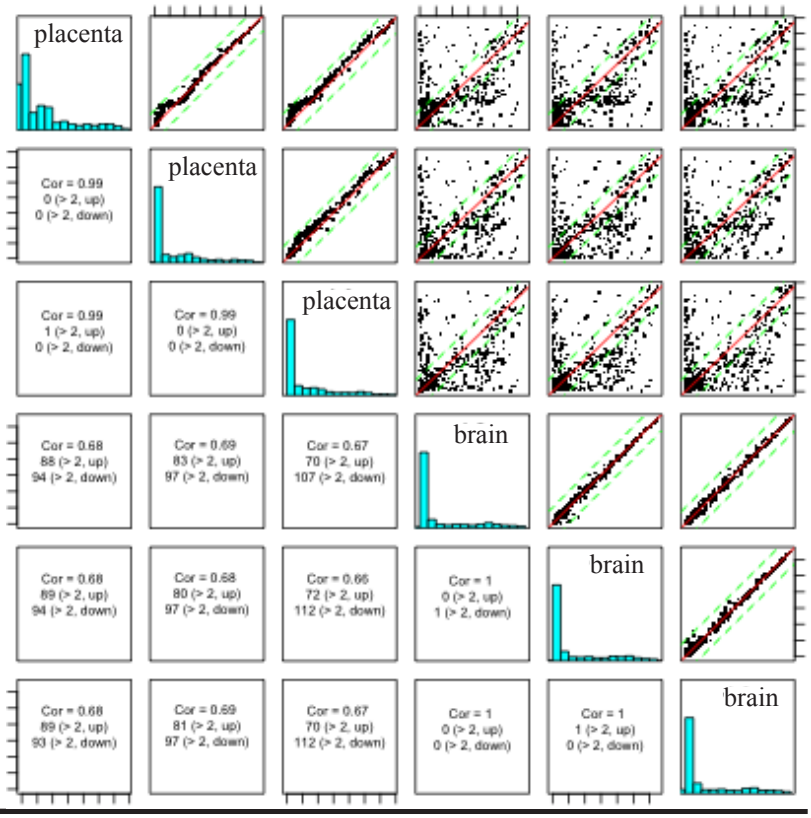

Figure 1. The normalised log2 signal intensity plots illustrate the pair-wise correlation of the expression profiles between the placenta (200 ng) and brain (200 ng) samples hybridised to the Illumina miRNA v1 BeadChips. The data from both samples (run in triplicate) have a very high correlation (>0.99).

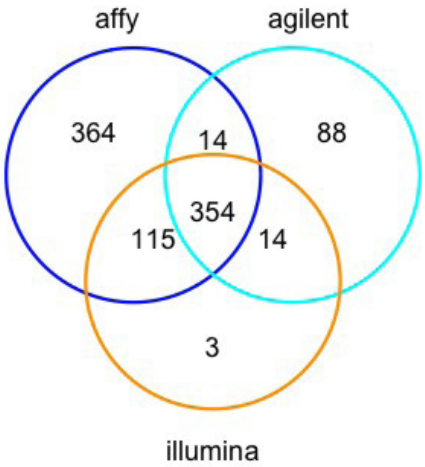

Figure 2. The overlap of human miRNA probes for the Affymetrix, Agilent, and Illumina platforms after controls and non-human miRNAs were filtered out.

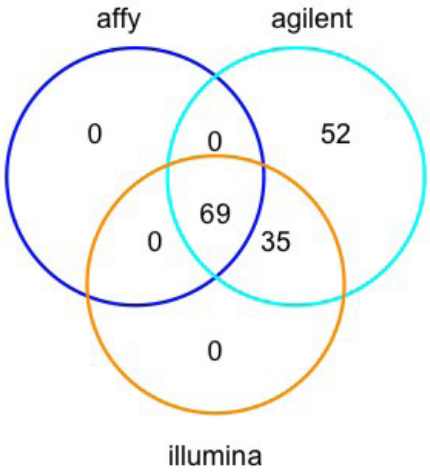

Figure 3. Following the t-test (p value < 0.05 and fold change >2), 156 of the 354 common human miRNAs were differentially expressed between placenta and brain samples, and 69 of the 156 probes (44%) were common among the three platforms. The data was not filtered on the basis of signal intensity.

Given that the probe content of the Affymetrix miRNA GeneChip was based on a more recent version (v11) of miRBase, compared to the Agilent and Illumina arrays (versions 9.1 and 9.0, respectively), it was expected that there would be more overlap between the miRNAs identified using the Agilent and Illumina platforms. It is possible that 52 of the differentially expressed miRNAs were identified only by the Agilent platform because the Agilent experiments were performed in duplicate; this number may have been lower if the Agilent experiments were performed in triplicate. This result may also be due to the fact that the Agilent probes have a unique hairpin design and represent only mature miRNA species.

The results of this study led to the conclusion that the Illumina Human miRNA platform provides data that correlates well to the Affymetrix and Agilent platforms when considering miRNA that are common among the three platforms.

Conclusion

- The Illumina miRNA data was highly reproducible among technical replicates (correlation >0.99)
- Preliminary Illumina miRNA BeadChip data correlated well with the human miRNA data from Affymetrix GeneChips and Agilent miRNA arrays
- This experiment will be repeated using the most recent array versions of the three platforms and the data will be validated

References

1. miRBase webpage. July 20, 2009. ftp://ftp.sanger.ac.uk/pub/mirbase/sequences/CURRENT/README

Normalisation

from page 1

prediction (1). Version 13 (released in March 2009) contains 9539 entries representing hairpin precursor miRNAs, expressing 9169 mature miRNA products, in 103 species (2).

Array-based platforms are often used to profile miRNA expression; however, the normalisation of miRNA profiling data has not yet been investigated in detail. Early miRNA array normalisation methods involved centring to median values (3,4) and scaling based on total array intensities (5), and more recently, quantile normalisation (6,7) and variance stabilising normalisation (VSN) (8,9) have been applied to miRNA data. Pradervand and colleagues have published a study that applied different normalisation methods to miRNA array data, generated using the Agilent platform, and then assessed the impact of each method on sensitivity, specificity, and fold-change measurement relative to qPCR (10).

This study also describes the VSN-Inv normalisation method, which involves selecting invariants (non changing miRNAs) and using them to compute linear regression normalisation coefficients or VSN parameters. In contrast to VSN with default parameters, which assumes that the expression of most genes do not change, the method where VSN parameters are calculated from invariant probes is based on the assumption that there exists a set of miRNAs whose expression is constant across all arrays in the experiment (probes with high mean expression across arrays and low standard deviation) (10). Data normalised using the invariant-based method was compared to data normalised by scaling, quantile, and VSN with default parameters as well as to data that was not normalised.

For the comparison of the normalisation methods, this study used two distinct data sets: a comparison of two tissues (heart and brain) where a large percentage of the miRNAs are differentially expressed and a data set from a squamous carcinoma cell line transfected with two different constructs where a smaller percentage of the miRNAs are differentially expressed. To evaluate sensitivity and specificity following data normalisation, a set of 59 miRNAs differentially expressed between the brain and heart samples were monitored in mixed heart-brain samples. qPCR was performed for 17 of these miRNAs in order to compare fold change measurement using the different normalisation methods. A similar evaluation was conducted for the transfected cell lines.

The study concludes that normalisation based on the set of invariants and quantile were the most robust over the experimental conditions tested, and suggests the method of invariant selection and normalisation can also be applied to data sets from one-colour miRNA arrays, focused gene expression arrays, and expression profiling using qPCR.

References

1. Griffiths-Jones S, et al. miRBase: tools for microRNA genomics. NAR 2008, 36:D154-D158
2. miRBase Sequence Database, Release 13.0. July 11, 2009 ftp://ftp.sanger.ac.uk/pub/mirbase/sequences/CURRENT/README
3. Sun Y, et al. Development of a microarray to detect human and mouse microRNAs and characterization of expression in human organs. Nucleic Acids Res 2004, 32:e188
4. Castoldi M, et al. A sensitive array for microRNA expression profiling (miChip) based on locked nucleic acids (LNA). RNA 2006, 12:913-920
5. Miska EA, et al. Microarray analysis of microRNA expression in the developing mammalian brain. Genome Biol 2004, 5:R68
6. Laurent LC, et al. Comprehensive microRNA profiling reveals a unique human embryonic stem cell signature dominated by a single seed sequence. Stem Cells 2008, 26:1506-1516
7. Sengupta S, et al. MicroRNA 29c is down-regulated in nasopharyngeal carcinomas, up-regulating mRNAs encoding extracellular matrix proteins. PNAS 2008, 105:5874-5878
8. Davison TS, et al. Analyzing Micro-RNA Expression Using Microarrays. Methods Enzymol. 2006, 411:14-34

9. Huber W, et al. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics 2002, 18(Suppl 1):S96-104
10. Pradervand S, et al. Impact of normalisation on miRNA microarray expression profiling. RNA 2009, 15(3):493-501

The UHNMAC will soon be launching several new services, including:

- Nanostring Technology
- Bio-Rad Bio-Plex
- Caliper LabChip GXII

Please visit our website to learn more!

Please visit our new bioinformatics website,

http://data.microarrays.ca