Technical Note

UHN Microarray Centre

Comparison of Illumina, Affymetrix, and Agilent array-based platforms for miRNA expression

- miRNA expression profiles were generated for placenta and brain RNA samples on several array-based platforms (Affymetrix, Agilent, and Illumina)
- Illumina miRNA BeadChip data was highly reproducible among technical replicates (correlation >0.99)
- Study found that preliminary Illumina miRNA BeadChip data correlates well with the human miRNA data from Affymetrix GeneChips and Agilent miRNA arrays

Introduction

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.

miRBase is a public web-based resource for miRNA data established by the Sanger Institute. 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.

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.

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

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 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-array BeadChip format (12 arrays per slide)

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

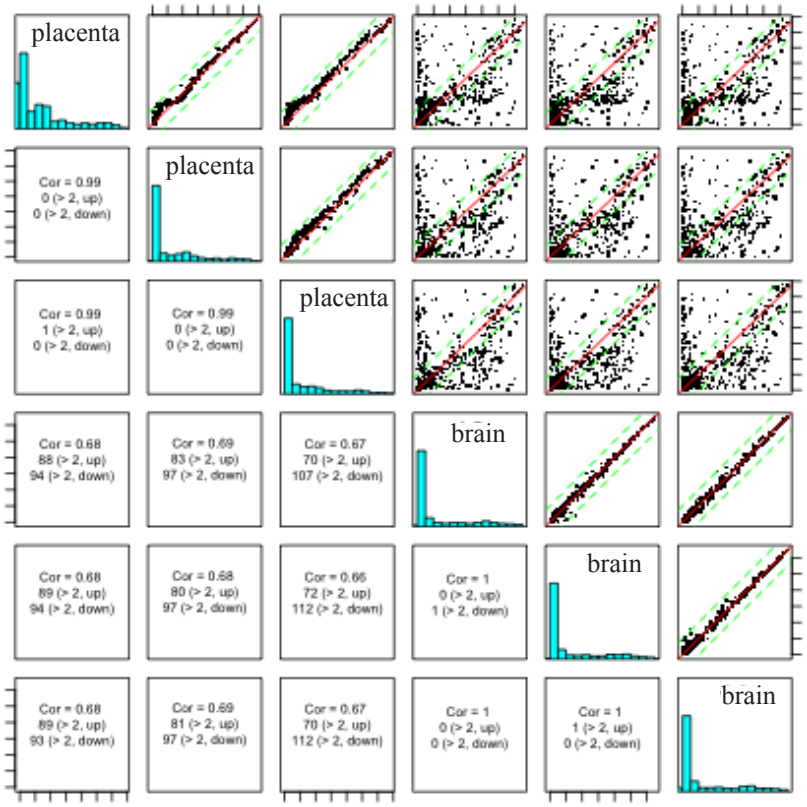

Figure 1. The normalised log2 signal intensity plots illustrate the pair-wise correlation of the expression profiles between the placenta (200 ng) and brain (200 ng) samples hybridised to the Illumina miRNA v1 BeadChips. The data from both samples (run in triplicate) have a very high correlation (>0.99).

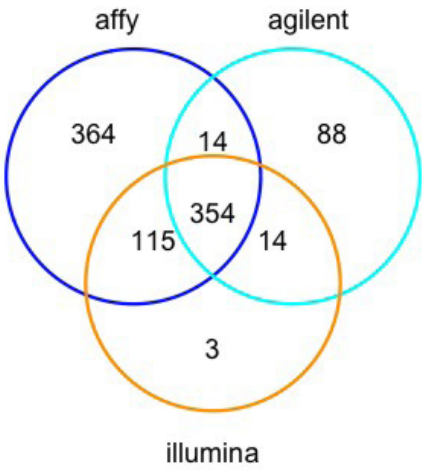

Figure 2. The overlap of human miRNA probes for the Affymetrix, Agilent, and Illumina platforms after controls and non-human miRNAs were filtered out.

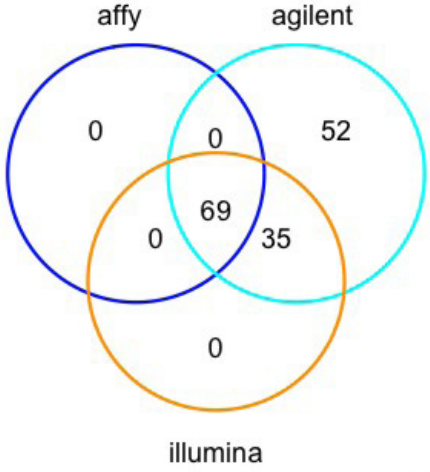

Figure 3. Following the t-test (p value < 0.05 and fold change >2), 156 of the 354 common human miRNAs were differentially expressed between placenta and brain samples, and 69 of the 156 probes (44%) were common among the three platforms. The data was not filtered on the basis of signal intensity.

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.

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