Technical Note

UHN Microarray Centre

Validation of the Illumina iScan System for gene expression

- Gene expression profiles were generated for Universal Human Reference RNA (UHRR; Stratagene), Human Brain Total RNA (HBTR; Ambion) and 25:75 mixtures of UHRR and HBTR, and the data was compared to published MAQC data
- Data demonstrated that gene ratios are consistent with MAQC data and that overall reproducibility is exceptionally high
- RNA amplification from as little as 25 ng quality total RNA provided data comparable to that obtained from 500 ng total RNA

Introduction

Prior to the launch of the Illumina iScan System at the University Health Network Microarray Centre (UHNMAC), an internal validation of the platform was completed. The purpose of this report is to briefly provide the results of our validation of Illumina's gene expression platform. By repeating an experiment previously performed by the MicroArray Quality Control (MAQC) Consortium, we were able to compare our results in order to determine the robustness of the system and to determine the reproducibility of the technology. A growing trend toward the analysis of smaller amounts of RNA has been witnessed (and promoted) by the UHNMAC. As such, this study also looked at the amplification of smaller amounts of total RNA to approximate the lowest starting amount of total RNA required for quality data.

In 2006, the MAQC Consortium published the results of a comprehensive comparison of multiple array platforms (1). The goal of this community-wide study, which involved 137 participants from 51 organisations, was to establish QC metrics and thresholds for objectively assessing the performance of various microarray platforms and to evaluate the advantages and disadvantages of various data analysis methods. Differential gene expression levels between various RNA samples (including Stratagene Universal Human Reference RNA (UHRR), Ambion Human Brain Total RNA (HBTR), and two mixtures: 75:25 mixture of UHRR and HBTR, and 25:75 mixture of UHRR and HBTR) were assessed using seven microarray platforms. The resulting datasets were used to assess the precision and cross-platform comparability of microarrays (1).

Method

For the samples listed in Table 1, total RNA was labelled using the Illumina® TotalPrep™ RNA Amplification Kit. At least 10 µg cRNA was obtained following amplification (including from 25 ng total RNA input). 1.5 µg cRNA was hybridised to a Human Whole-Genome BeadChip 48K, version 3.0. In some cases, samples were hybridised in duplicate or quadruplicate. Data was normalised (rsn method), quantified in BeadStudio™ and analysed using BeadStudio™ and R-package software applications.

Table 1. List of samples used for the internal validation of the Illumina iScan System for gene expression studies.

Total RNA Sample	Starting amount of total RNA used for amplification
HBTR	500 ng
75% UHRR/25% HBTR	500 ng
25% UHRR/75% HBTR	500 ng
Illumina sample #1	500 ng
Illumina sample #2	500 ng
HBTR	50 ng
HBTR	25 ng

Results

- All 12 hybridisations passed Illumina's internal quality control metrics (including labelled hybridisation controls, low and high stringency controls, signal generation (biotin) controls, and negative controls) and all 12 arrays met the Illumina acceptance criteria (greater than 11,000 genes were detected at p-values <0.1); data not shown
- The identification of differentially expressed genes (two-fold up and down) from this validation study was highly consistent with MAQC data, but due to differences in the BeadChip versions, a direct comparison of the data sets was beyond the scope of this preliminary report
- Hierarchical clustering of all samples indicates that the variability between technical replicates was very low and that technical replicates cluster closely together; Figure 1
- Data from brain samples (amplified from 25 ng, 50 ng, and 500 ng total RNA) were highly reproducible as correlation values were >0.99 (excluding one replicate of the HBTR sample amplified from 50 ng, which had a correlation value of 0.96); Figure 2
- Reproducibility and amplification fidelity deteriorate as the amount of starting RNA decreases, however in cases where sample is limiting, relatively robust data can still be obtained from as little as 25 ng of total RNA; Figure 3

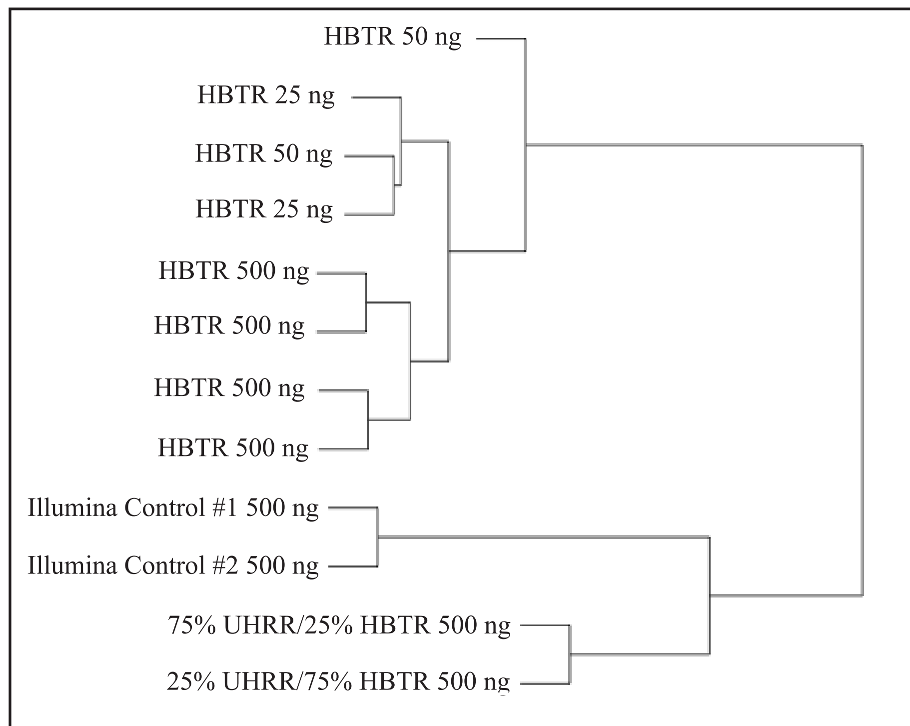

Figure 1. Hierarchical cluster of the 12 samples based on 11068 genes with a sd/mean > 0.1. Technical replicates of the HBTR samples (amplified from 500 ng) clustered together, as did the HBTR samples amplified from 25 ng total RNA. When box-plots for each sample were examined, one of the replicates of HBTR amplified from 50 ng appeared more different than the other HBTR samples and thus did cluster as tightly as others. However, all HBTR samples clustered together.

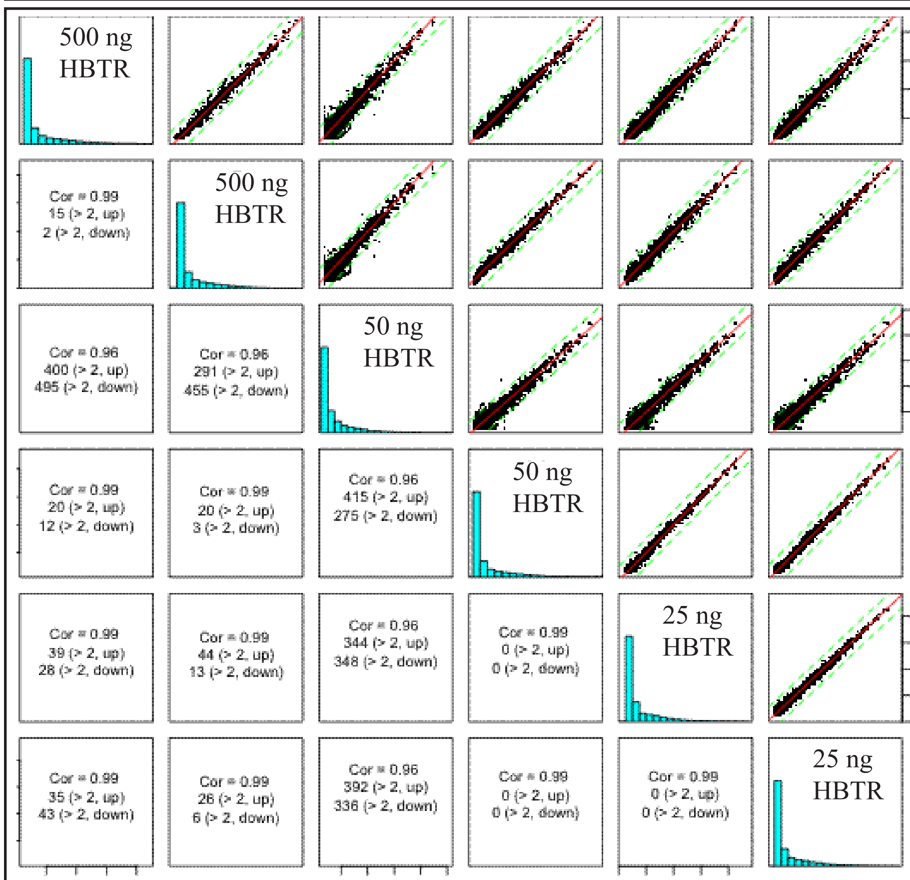

Figure 2. The normalised $\log_2$ signal intensity plots (generated from a random sample of 5000 data points) illustrate the pair-wise correlation of the expression profiles between the following data sets: HBTR amplified from 500 ng, HBTR amplified from 50 ng, and HBTR amplified from 25 ng. The correlation values of the HBTR samples (amplified from 25 ng, 50 ng, and 500 ng total RNA) are between 0.96 and 1. When one of the HBTR replicates is excluded, the correlation value for all sample pairs is 0.99.

Figure 3. Venn diagram showing the overlap of differentially expressed genes between the 75% UHRR/25% HBTR sample and HBTR samples amplified from 500 ng, 50 ng, and 25 ng total RNA. The number of common differentially expressed genes among brain samples amplified from 25 ng, 50 ng, and 500 ng indicates that robust data can be obtained from as little as 25 ng of total RNA.

Conclusions

- UHNMAC validation data was highly consistent with MAQC data
- The validation data was highly reproducible among technical replicates (correlation >0.96)
- Although relatively robust data can be obtained from as little as 25 ng of total RNA, the UHNMAC requests a minimum of 50 ng of total RNA for Illumina gene expression studies
- This study was one of the preliminary validation tests of the iScan System prior to launching the Illumina service; further testing is ongoing

References

1. MAQC Consortium. Nature Biotechnology 2006, 24(9):1151