Explaining the Sources of Discrepancies in Gene Expression Profiles Generated on three Whole-Genome Gene Expression Microarray Platforms

Maouche S. [1], Poirier O. [1], Godefroy T. [1], Olaso R. [2], Gut I. [2], Montalescot G. [3], and Cambien F. [1]

[1] INSERM UMR S525, Faculté de Médecine Pitié-Salpêtrière, 91 Boulevard de l'Hôpital, Paris, 75013, France.
[2] Centre National de Génotypage, Evry, France.
[3] Institut de Cardiologie. Groupe Hospitalier Pitié-Salpêtrière, Paris, France

ABSTRACT

Previous small-scale cross-platforms comparative studies have discussed several issues of microarray-based gene expression data, including comparability between platforms, repeatability between labs, performance, and concordance to non-array based gene expression. Recently, results of the MicroArray Quality Control (MAQC) project (**1**), the first large-scale cross-platforms study conducted with the goal of establishing quality control metrics for microarray data and of assessing the reliability of gene expression profiles generated on different platforms, showed that using standardized procedures, microarray results from different platforms are reproducible. We conducted a study to compare gene expression data generated on three platforms: Illumina Bead Chip Human-6 V1, Affymetrix HGU133plus 2.0, and the academic RNG/MRC two-color chip. 10 RNA samples from human monocyte and monocyte-derived macrophage were hybridized in parallel to the 3 platforms. In addition, a list of differentially expressed genes generated using a larger number of hybridizations to the RNG/MRC platform was included in the cross-platforms comparisons and used as a reference to assess the 3 platforms.

INTRODUCTION

Several academic and commercial microarray platforms are available; this situation has raised the need for cross-platforms and inter-laboratories reproducibility assessment. In addition to the studies discussing quality of microarray issues which continue to appear regularly, several standardization efforts that aim to improve the quality of microarray-based gene expression and the comparability of microarray data among different platforms and laboratories have been initiated. In this study, we assessed the internal consistency of each platform, the overlap in lists of genes identified as differentially expressed generated on each array type, and biological relevance of significant genes identified on each platform (Gene Ontology (GO) classes enrichments). Impact of genes lists selection criteria (statistical significance threshold) on the degree of concordance between the 3 platforms was investigated.

MATERIALS & METHODS

RNA samples. Samples were obtained from patients with symptoms of acute coronary syndrome who had undergone coronary angiography at the department of cardiology of the Pitié-Salpêtrière Hospital, Paris and who had one stenosis >50 % diagnosed in at least one major coronary artery. This study was approved by the ethic committee of Pitié-Salpêtrière Hospital and informed consent was obtained from all participants.

Experimental design and hybridization. 176 RNA samples were extracted from monocyte and monocyte-derived macrophage cells. Five biological replicates for each type of sample were hybridized in parallel to the three platforms (Affymetrix, Illumina and RNG-10). The remaining 166 samples (83 monocyte samples and 83 macrophages) were hybridized only to the RNG/MRC platform (RNG-98) and expression profiles generated from this experiment were used as reference to assess the performance of the 3 platforms based on 10 samples.

Data pre-processing and analysis. For each platform, raw data was background corrected and normalized using appropriate methods. Statistical analysis was carried out in R & Bioconductor environments. Cross-platforms comparisons were performed both on the whole-content of each array and on a subset of well matched transcripts. Gene ontology

(GO) enrichment was also performed to assess the 3 platforms in terms of biological relevance of genes identified as differentially expressed.

Filtering, annotation, and between platform probe mapping. Qualitative detection calls for the two one-color arrays were calculated, compared and used to remove absent probes. For the two-color array, we used a filtering based on spots quality indicators to remove bad spots from the analysis. In addition, control probes present on all array types were removed prior to the statistical analysis.

RESULTS

The analysis revealed a large number of genes differentially expressed between monocytes and macrophages. Gene lists were generated at different levels of statistical significance to examine the impact of gene list selection criteria on the level of agreement between the 3 platforms. At an adjusted p-value < 0.001, the three platforms identified in common 92 genes, representing 53.49%, 7.36%, 6.45% of all differentially expressed genes identified on the RNG-10, Affymetrix and Illumina platforms, respectively. The list of genes generated by the RNG-98 experiments confirmed 91% of the 1139 genes present on the Illumina list and 88 % of the 1001 genes present on the Affymetrix list (adj. p-value < 0.001) (**Figure 1 (B)),** whereas the overlap among the 3 platform was of 576 genes. These results suggest that the Illumina and Affymetrix technologies are complementary.

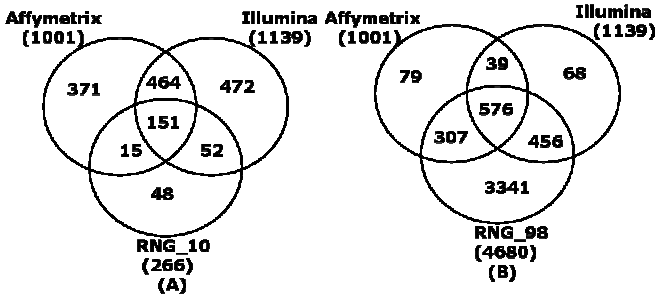

Figure 1. Overlap in lists of differentially expressed genes. Results are showed both for RNG-10 and RNG-98 and only a subset of well matched transcripts represented on the 3 platforms are included in this comparaison.

Results from the RNG platform using 10 samples were less convincing, but the analyses based on GO categories enrichment demonstrated that the gene lists delivered by the 3 platforms were highly correlated with the biological question examined in this study (**Table 1**).

Table1. GO categories (* number of genes present and $ adjusted P-value) enriched in the lists of differentially expressed genes selected using 2 thresholds: A. Adj P < 0,001 , B. Adj P < 0,05 and 20 % largest fold-change. Results are showed for only 2 GO categories.

Microarray platform		Number of gene in the list	Cell adhesion (622 genes)	Macrophage-mediated immunity (140 genes)
Affymetrix	A	1203	54 * ($<10^{-03}$) $	22 ($<10^{-03}$)
	B	751	45 ($<10^{-05}$)	19 ($<10^{-04}$)
Illumina	A	1420	55 (<0.05)	21 (<0.01)
	B	637	39 ($<10^{-05}$)	23 ($<10^{-09}$)
RNG_10	A	160	18 ($<10^{-05}$)	6 (0.05)
	B	275	24 ($<10^{-05}$)	12 ($<10^{-05}$)
RNG_98	A	6961	211 (0.05)	66 (<0.01)
	B	1770	90 ($<10^{-08}$)	37 ($<10^{-08}$)

The lack in agreement of gene lists was strongly influenced by the criteria (statistical significance threshold and fold-change) used to select the gene lists. The best criterion was found to be a combination of non stringent P-value (Adj P < 0.05) and largest platform-specific fold-change (20% largest values) (**table 1**). On the other hand the criterion recommended by MAQC's authors, combining a non-stringent P-value and a fold-change of 2 (**1**) was inappropriate for Illumina data.

REFERENCES

1. Shi,L., Reid,L.H., Jones,W.D., Shippy,R., Warrington,J.A., Baker,S.C., Collins, P.J., de Longueville,F., Kawasaki,E.S., Lee,K.Y. et al. (2006) MAQC Consortium. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nature Biotechnol*., **24**(9), 1151-6.