WHITEPAPER

Gene Expression Profiling on the Ziplex® System Demonstrates High Inter-platform Concordance

Michael Jones, Adam Dempsey and David Englert

Axela, Inc. Toronto, Ontario, CANADA

Abstract

The samples and titration mixtures used in the Microarray Quality Control (MAQC) project[1] were assayed on the Ziplex® Automated Workstation with a TipChip containing probes for 52 genes that were profiled using TaqMan® in the MAQC study. The multi-site Ziplex results were compared to the publicly available MAQC data generated from five global expression platforms and the TaqMan assay. Reproducibility of Ziplex, measured as the median coefficient of variation (CVs) of normalized Ziplex System probe data from all samples across four sites, was 13.5%, compared to a range of 17% to 34% for the same set of genes on the global expression array platforms. All expression platforms had higher median CVs relative to the TaqMan results. The number of genes determined to be differentially expressed was greatest on the TaqMan assay; however the concordance of differentially expressed genes was good between the Ziplex System and all the other platforms. The performance of the Ziplex System was consistent with other gene expression profiling platforms evaluated in the MAQC project.

Introduction

The MicroArray Quality Control (MAQC) study was undertaken by a large consortium of investigators to assess reproducibility within and between gene expression profiling platforms[1]. The two RNA samples that were profiled in the MAQC study are commercially available, and the data from all the platforms evaluated in the study is also publicly available. We have performed the MAQC experiment on a focused set of 52 genes. The study was performed at multiple sites using the Ziplex Automated Workstation and a single lot of reagents. The Ziplex System multi-site results were compared against the MAQC results for the same set of 52 genes from five global expression arrays and from real-time PCR (TaqMan®). We have examined reproducibility, analytical sensitivity, response to the titration mixtures, and sensitivity and concordance of differential expression measurements.

The Ziplex® Research System

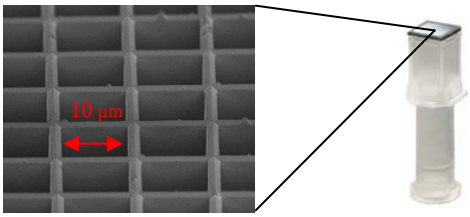
The Ziplex Automated Workstation (Figure 1) is designed to consolidate array molecular binding, imaging, data quantification, and quality control into a single, small bench-top unit. The System uses flow-thru chip technology in which probes are immobilized within a highly porous silicon matrix providing a very high surface area for binding probes. During hybridization/protein binding, washing and detection, samples and reagents are continually perfused through the chip within a microplate well to provide

Figure 1

(A) The Ziplex Research System: a fully automated, bench-top instrument oligonucleotide and protein microarray analysis. (B) Scanning EM of the porous silicon substrate and an assembled TipChip

A

B



axela

efficient mixing and transport of target molecules to the immobilized probes. Figure 1B shows an electron micrograph of the silicon chip which is fixed to a plastic tube to form a TipChip.

TipChip Arrays

The TipChip array is comprised of probes for 52 genes that were all represented in the MAQC study, including the TaqMan data. The array was designed with Axela Molecular's standard probe design and screening process in order to ensure optimal performance. One probe was designed for each gene with the exception of POLR2A gene for which two probes were designed. POLR2A was used to normalize the MAQC TaqMan data and hence was expected to be an invariant gene in the samples being profiled. All probes were spotted in triplicate on the array. The array also contained a number of quality control probes.

Samples

The same lot numbers of Universal Human Reference RNA (UHRR, Sample "A", Stratagene cat# 740000 lot #1130623) and Human Brain Reference RNA (HBRR, Sample "B", Ambion cat#AM6052 lot# 055201) as used in the MAQC project were acquired from commercial sources. The "C" and "D" samples ("titration samples") were prepared from the A and B samples in the 75:25 and 25:75 mass ratios as done in the MAQC project.[2] The relative proportions of A and B mRNA used in calculations to predict the expression values of the C and D samples were the consensus values (reflecting actual mRNA concentrations in the A and B samples) determined with the gene expression profiling platforms in the MAQC project. Messenger RNA was amplified using IVT from the total RNA and biotin-labeled with standard Axela procedures at each of four sites. A total of five technical replicates were produced from each of the four sample types. Each TipChip was hybridized with six µg of labeled aRNA using standard Ziplex System protocols. Hybridizations were performed in three runs across four sites using the Ziplex Automated Workstation using two lots of TipChips.

Quality Filtering and Data Normalization

To be included in the analysis a sample profile had to have at least 90% of the gene probes detected (i.e., signal greater than the 95th percentile of a set of eight unique negative control probes). This resulted in the exclusion of eight sample profiles (sample B n=2, C n=2 and D n=4). Probes that were not detected in at least 60% of the sample profiles for all samples across all four sites were also excluded from subsequent analysis.

The data from the six platforms in the MAQC project was normalized using methods specific to each platform.[2] The normalization methods used on global expression platforms are not applicable to the Ziplex System due to the relatively small number of probes on TipChip arrays. Normalization of the Ziplex System data to expression values of genes that are stably expressed across the samples of interest can eliminate any uncontrolled variability in sample preparation, aRNA quantification, pipetting, TipChips or processing on the Ziplex System. It is essential for effective normalization that the genes used for normalization are stably expressed across the samples in a study.

The probes selected to be used in the normalization method were chosen prior to the generation of data on Ziplex. POLR2A was chosen as the probe to be used for normalization to be consistent with the normalization method employed on the TaqMan platform. The geometric mean of two unique POLR2A probes on the TipChip array was used to normalize the Ziplex data in the MAQC study.

Analytical Sensitivity

Detection calls were generated for all probes on each of the MAQC platforms using the manufacturers' criteria. The proportion of genes scored as detectable (i.e., present) varied between platforms, reflecting in part to the different detection criteria that are used on each platform.

As with the platforms evaluated in the MAQC study, a present call was generated for each probe for each sample replicate. Overall, of the 52 probes analyzed in this study, 0, 1 or 2 were not detected on each of the platforms compared The number and type of gene probes not detectable was specific to each platform. However, some probes were found to be undetectable or have a relatively lower detection rate across a number of platforms, for example IGFBP1 and IL6. As each platform has specific criteria for determining whether a probe is considered present or not, the variance in the number of genes detected is as expected.

Reproducibility

The reproducibility, measured as coefficient of variation (CV), of each of the 52 probes was determined for each sample type and each platform separately. The CVs of the Ziplex System data were improved through normalization, as expected. The median of the probe CVs for all the platforms, including the non-normalized and normalized Ziplex System data, are presented in Table 1. Overall the CVs were smallest for the TaqMan assay. The CVs for the normalized Ziplex System data were slightly smaller than the Affymetrix platform data. The CVs of the other gene expression profiling platforms were similar to the Affymetrix platform data, with the exception of the Illumina platform data which showed CVs nearly double that of other expression platforms. All normalization strategies appeared to provide similar measures of reproducibility. Overall the reproducibility of the Ziplex System was similar to or better than all other platforms with the exception of the TaqMan data.

Table 1

Median %CVs of MAQC probes across multiple sites for each sample type for normalized (MPr) and unnormalized (Raw) Ziplex data and for the six addition gene expression profiling platforms assessed in the MAQC study.

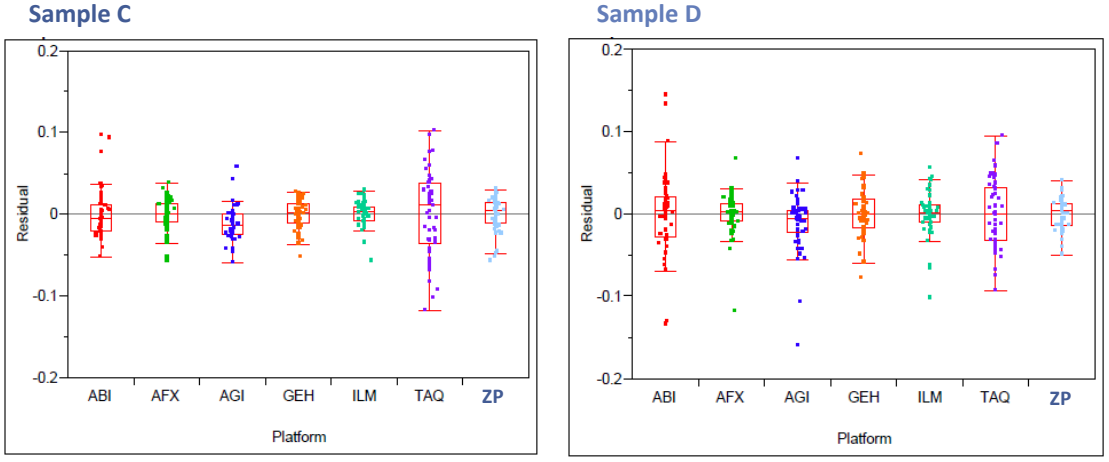
	%CV							
Sample	Ziplex		TAQ	AFX	ILM	ABI	AGI	GEH
	Raw	MPr						
A	27.1	13.5	2.8	16.3	33.4	17.3	17.2	19.7
B	25.4	12.7	3.1	18.7	34.1	20.3	17	19.1
C	30.6	14.5	2.7	17.3	34.4	14.9	15	18.8
D	24.2	13.0	3.01	16.5	33.7	18.2	21.4	19.5

Titration Response – Measured vs. Predicted Expression Values

The Ziplex System accurately predicted the expression values of the C and D titration mixtures from the measured expression values of the A and B samples. Linear regression was performed to assess the relationship between the measured signals from samples C and D and those signals predicted based on the percentage of samples A and B in these samples. This was done independently for each platform. Pearson's correlation between the mean signal intensities between samples C and D and the predicted C and D was >0.99 for all platforms. The exception was the Agilent platform which had an actual to predicted correlation of >0.98. Comparison of the residual error from the linear regression analysis between each global expression platform indicated that the titration response was consistent across platforms (Figure 2) with TaqMan showing greater variation overall.

Figure 2

Residual error from Linear Regression analysis comparing the log signals produced by the C and D samples and the values predicted for the C and D samples. The relative proportions of A and B mRNA in the C and D samples were assumed to be the consensus values determined with gene expression profiling platforms in the MAQC project (C=0.82A + 0.18B and D=0.33A + 0.67B) reflecting actual mRNA concentrations in the A and B samples). All platforms showed a similar titration response with TaqMan showing the largest variation.



Differential Expression

The mean log10 signal intensity was compared between samples A and B and between samples C and D for each probe on each platform using a two-tailed Student's t-test. Large expression differences are expected between the A and B samples whereas smaller differences are ex

pected between the C and D samples. With the exception of Illumina, all platforms detected a similar number of differentially expressed genes (Table 2). The Ziplex System detected fold change differences of 1.3 to 83.5 between samples A and B and differences of 1.2 to 2.4 between samples C and D ($p<0.0001$).

Table 2

The number of genes that were differentially expressed between the A and B samples and between the C and D samples on Ziplex (normalized data; MPr) and on each of the MAQC platforms at two levels of statistical significance based on a two-tailed Student's t-test.

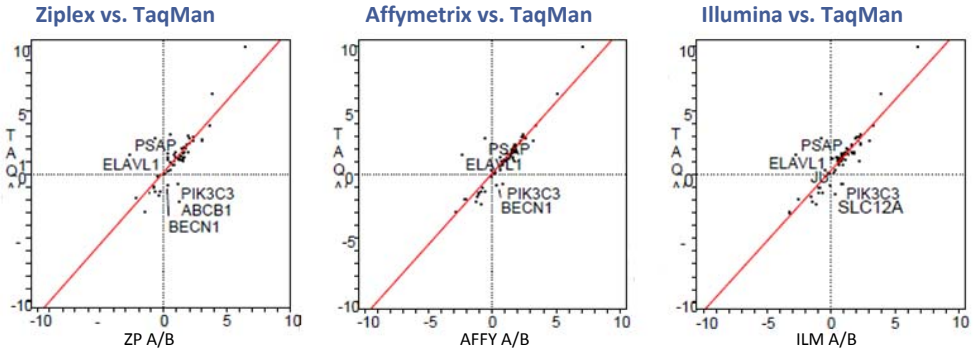
Platform	n	A vs B		C vs D	
		0.001	0.0001	0.001	0.0001
Ziplex MPr	52	46	44	40	38
AFX	52	49	48	42	33
ILM	52	40	34	18	9
ABI	52	44	40	35	32
AGI	52	47	47	41	35
GEH	52	48	45	31	26
TAQ	52	52	50	47	41

The fold change between samples A and B and between sample C and D was generally consistent across all expression platforms. The number of genes for which there were significantly different expression values between the A and B or the C and D samples was greatest with the TaqMan assay (Table 2). The number of significantly differentially expressed genes was similar on all the microarray platforms with the exception of Illumina, reflecting the higher CVs on that platform. The ability to detect small expression differences is directly related to a platform's reproducibility. The ability of the Ziplex System to detect differential expression for a relatively large number of genes shows that its reproducibility is at least equivalent to that of the other gene expression profiling platforms.

There were also a few genes for which the measured differential expression was discordant between the TaqMan assay and some of the gene expression profiling platforms (Figure 3), including Ziplex. There were five discordant probes between Ziplex and TaqMan for the comparison of samples A and B. Four of these probes were also discordant between TaqMan and the Affymetrix platform, with three discordant against the Illumina platform. There was good agreement between the Ziplex, Affymetrix and Illumina platforms, with only one discordant probe when Ziplex is compared to Affymetrix (ABCB1) and two compared against Illumina (ABCB1 and SLC12A2) (Figure 4).

Figure 3

Correlation of the log2 fold change between the A and B samples as measured by TaqMan (TAQ) vs. Ziplex (ZP) and two gene expression profiling platforms (Affymetrix (AFFY) and Illumina (ILM)). The gene probes for which the differential expression was discordant between TaqMan and the other gene expression profiling platforms are labeled.



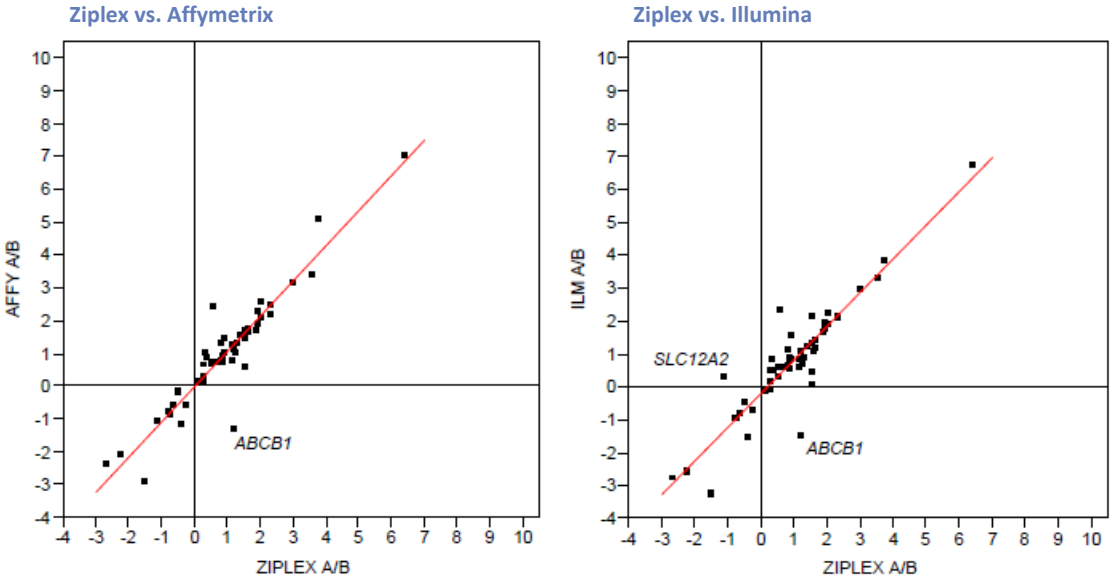
Conclusions

The analytical sensitivity, reproducibility and fold-change sensitivity of the Ziplex System were similar to or better than these metrics of the global expression platforms. The reproducibility of expression signals on Ziplex compared favorably to the global expression platforms, and the titration response was equivalent across all expression platforms. The Ziplex System ranked amongst the highest in its ability to detect differentially expressed genes and had excellent fold-change concordance with TaqMan and the other gene expression profiling platforms. Thus, the performance of the Ziplex automated workstation relative to other platforms indicates that

it is an equally viable tool for gene expression profiling. Furthermore, given it's focused and highly automated design, the Ziplex automated workstation is also strongly suited for verification of differential gene expression results obtained from whole genome expression profiling.

Figure 4

Correlation of the log2 fold change between the A and B samples as measured by Ziplex (XM) vs. Affymetrix (AFFY) and Illumina (ILM). The gene probes for which the differential expression was discordant between Ziplex and the other gene expression profiling platforms are labeled.



References

1. Nature Biotechnology, Volume 24, 2006.
2. Shippy, R., et al. (2006). "Using RNA sample titrations to assess microarray platform performance and normalization techniques." Nat Biotechnol 24(9): 1123-31.

About Axela, Inc.

Axela's platforms provide powerful new approaches to multiplexed protein and nucleic acid analysis designed to greatly simplify biomarker testing in clinical research and diagnostics. Axela's commercial research products significantly improve the amount and quality of information derived from traditional assays. This approach shortens time to result and provides access to unique categories of markers that form a pipeline of future diagnostic offerings.

For research use only.

axela

50 Ronson Dr, Suite 105
Toronto, Ontario
Canada
M9W 1B3

Ph: 1.866.94.AXELA (toll free)
or +1.416.798.1625
Fax: +1.416.798.8635

www.axela.com
info@axela.com