



A Comprehensive Large-Scale Evaluation of Gene Expression Measurements Across Different Hybridization-Based Technologies



Winston Patrick Kuo^{1,2,3}, Fang Liu⁴, Tor Kristian-Jenssen⁴, Steven Perrin⁵, Sangdun Choi⁶, Roger Bumgarner⁷, Juhan Kim⁸, Glenn F. Short III³, Mason W. Freeman³, Brian Seed³, Roderick Jensen⁹, George M. Church³, Eivind Hovig⁴, Connie L. Cepko³, Lucila Ohno-Machado²

¹Dept. Developmental Biology, Harvard School of Dental Medicine, ²Decision Systems Group, ³Dept. of Genetics, Harvard Medical School, Boston, MA, ⁴The Norwegian Radium Hospital, Dept. of Tumor Biology, Oslo, Norway, ⁵Research Molecular Discovery, Biogen Idec, Cambridge, MA, ⁶Division of Biology, California Institute of Technology, Pasadena, CA, ⁷Dept. of Microbiology, University of Washington, Seattle, WA, ⁸Seoul National University College of Medicine, Seoul, Korea, ⁹Dept. of Physics, University of Massachusetts Boston, Boston, MA

Introduction

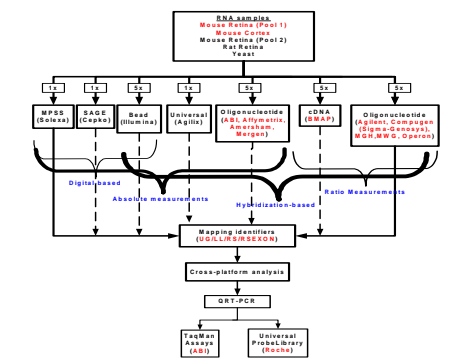
Gene expression microarrays have made a significant impact in many areas of research. The diversity of platforms and analytical methods has made comparison of data from multiple platforms very challenging. In this large-scale study, we developed a comprehensive framework for cross-platform comparison that included data from ten different platforms. As part of this framework, probe sequences were obtained from each platform.

The different platforms include single- and dual- dye platforms, and cDNA and oligonucleotide microarrays that were either fabricated "in-house" or were commercially available. Each laboratory received aliquots from two different RNA samples, mouse retina (MR) and mouse cortex (MC), that were prepared in the Cepko Laboratory at Harvard Medical School. The biological variation within tissue RNA preparations was minimized by pooling tissue from many animals prior to extracting the RNA. Hybridization was conducted in replicates of five to enhance statistical reliability.

QRT-PCR was used as an independent method to quantify the results of the microarray data. Two approaches were taken, one using ABI TaqMan assays and the other, Roche's Universal ProbeLibrary (UPL). TaqMan assays were conducted at ABI, whereas the UPL was conducted in-house. UPL was chosen over SYBR Green due to its flexibility, ease of use, cost, and result time from assay design to results.

Ideally, it would have been preferred if all experiments were conducted at one site, but since this was not practical given resource constraints, experiments were conducted at multiple sites. This ensured that the experiments were carried out by experienced technicians for each platform. Additionally it permitted us to evaluate the level of the variation of different platforms among laboratories and concurrently, the variation of similar platforms used at different sites. Using this approach, if the variation exists between platforms, we can get a general assessment whether it was attributable to a particular platform.

Figure 1: Experimental Design



Microarray Experiments

Sample preparation and hybridization steps were conducted following the protocols provided for each platform. Eight of the ten microarray platforms are currently commercially available: Affymetrix, Agilent, Applied Biosystems (ABI), Amersham (now GE Healthcare), CompuGen (now Sigma-Genosys), Mergen, MWG BioTech (now Ocum Biosolutions), and Operon. Probes from CompuGen and Operon were printed together on the same slide. The remaining two platforms are from academic laboratories: cDNA and long oligonucleotide (MGH) arrays. Six of the ten microarray platforms (Agilent, cDNA, CompuGen, MGH, MWG, and Operon arrays) are considered as two-dye platforms as they require the hybridization of two samples, whereas the others are single-dye platforms.

A subset of the experiments was conducted independently at a second laboratory using identical samples. This portion of the study is still ongoing, but results from Affymetrix, Amersham, and Mergen platforms have been completed.

All labeling and hybridization methods were completed as specified by each manufacturer's hybridization protocol. Image processing of the scanned images were conducted using the recommended scanners and settings.

Mapping of Probes Across Platforms

Probe mapping was conducted using two approaches: annotation-based and sequence-based. The probe sequences from each microarray platform were mapped to the mouse genome using the BLAT stand-alone program. The sequence alignment results were parsed so that only probe-to-exon matched pairs were extracted for the cross-platform comparisons. Probe-to-exon means that only the aligned sequences positioned completely within an exon were considered as a match.

Figure 2: Cross-platform matched probes within an exon

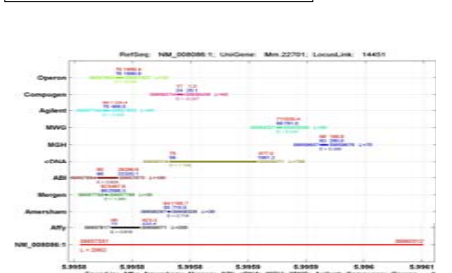


Table 1: Number of overlaps across platforms

| Unique # | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
|------------------|-------|-------|------|------|------|------|------|------|------|------|-----|
| UniGene Clusters | 30120 | 12749 | 5060 | 2157 | 2089 | 1058 | 1823 | 1972 | 1551 | 686 | 175 |
| LocaLink | 21844 | 8777 | 2834 | 2149 | 2069 | 1982 | 1812 | 1409 | 1544 | 1248 | 175 |
| RefSeq | 15510 | 2729 | 2835 | 2215 | 2180 | 1588 | 1814 | 1282 | 985 | 129 | 13 |
| RefSeq exon | 28874 | 14932 | 4966 | 2670 | 1776 | 1385 | 981 | 534 | 227 | 39 | 4 |

Table 2: UniGene overlaps

| | Affy | GE | Mergen | ABI | cDNA | MGH | MWG | Agilent | Comp | Operon |
|-----------|------|------|--------|-------|------|-------|------|---------|------|--------|
| Affy | 3541 | 6119 | 5892 | 4055 | 1882 | 6486 | 5080 | 4972 | 1593 | 8964 |
| GE Health | 0.47 | 9750 | 6575 | 4438 | 1665 | 7796 | 6764 | 4988 | 1776 | 7494 |
| Mergen | 0.49 | 0.56 | 8305 | 4118 | 1241 | 6715 | 7079 | 4160 | 1591 | 6000 |
| ABI | 0.21 | 0.23 | 0.22 | 14210 | 4674 | 5060 | 4322 | 3400 | 1108 | 4804 |
| cDNA | 0.11 | 0.11 | 0.09 | 0.27 | 7519 | 2060 | 1486 | 1765 | 365 | 1824 |
| MGH | 0.42 | 0.54 | 0.47 | 0.26 | 0.12 | 12513 | 7460 | 5951 | 1678 | 5546 |
| MWG | 0.49 | 0.58 | 0.69 | 0.23 | 0.1 | 0.54 | 8768 | 4534 | 1822 | 5647 |
| Agilent | 0.39 | 0.38 | 0.33 | 0.18 | 0.12 | 0.4 | 0.36 | 8509 | 971 | 4807 |
| Comp | 0.16 | 0.18 | 0.23 | 0.07 | 0.04 | 0.13 | 0.2 | 0.1 | 2022 | 1877 |
| Operon | 0.41 | 0.54 | 0.49 | 0.22 | 0.09 | 0.55 | 0.48 | 0.31 | 0.16 | 17111 |

Table 3: RefSeq exon overlaps

| | Affy | GE | Mergen | ABI | cDNA | MGH | MWG | Agilent | Comp | Operon |
|-----------|------|------|--------|------|------|------|------|---------|------|--------|
| Affy | 4859 | 2993 | 2712 | 2868 | 416 | 1441 | 1441 | 2488 | 682 | 3374 |
| GE Health | 0.19 | 7996 | 2771 | 2788 | 338 | 1305 | 1832 | 1955 | 711 | 3516 |
| Mergen | 0.29 | 0.32 | 7216 | 3178 | 368 | 1916 | 2079 | 2295 | 903 | 4090 |
| ABI | 0.24 | 0.19 | 0.23 | 976 | 544 | 1879 | 1903 | 3165 | 721 | 4177 |
| cDNA | 0.08 | 0.04 | 0.05 | 0.05 | 927 | 131 | 182 | 484 | 83 | 449 |
| MGH | 0.06 | 0.08 | 0.07 | 0.12 | 0.22 | 7601 | 394 | 446 | 297 | 1418 |
| MWG | 0.18 | 0.15 | 0.21 | 0.14 | 0.03 | 0.08 | 4656 | 1184 | 583 | 2483 |
| Agilent | 0.28 | 0.16 | 0.2 | 0.24 | 0.07 | 0.05 | 0.12 | 6529 | 444 | 2988 |
| Comp | 0.12 | 0.08 | 0.11 | 0.07 | 0.03 | 0.03 | 0.1 | 0.06 | 1712 | 987 |
| Operon | 0.34 | 0.27 | 0.35 | 0.3 | 0.05 | 0.1 | 0.23 | 0.25 | 0.11 | 8532 |

Data Analysis

The normalization, transformation, and filtering steps constituted the pre-processing procedure. The methods were performed independently for each platform and in several ways for different purposes. All raw data values were normalized and the method of normalization was selected based on the dye format. Normalization:

- Single-dye platforms - quantile normalization
- Two-dye platforms - loess normalization.

Transformation:

- Linear scaling
- Percentile scaling
- Log₂ ratios from the two samples.

Filtering:

- Filtering at the spot (image) level, taking into account quality flags and signal-to-noise ratio (SNR) thresholds

Pearson and Spearman correlation coefficients were calculated for intra- and inter- platform comparisons.

Results

| | Pearson | Spearman | | Pearson | Spearman |
|------------|---------|----------|------|---------|----------|
| | mean | sd | mean | sd | sd |
| Affymetrix | 0.78 | 0.12 | 0.71 | 0.15 | |
| GE | 0.89 | 0.06 | 0.84 | 0.08 | |
| Mergen | 0.68 | 0.17 | 0.71 | 0.16 | |
| ABI | 0.81 | 0.10 | 0.70 | 0.15 | |
| cDNA | 0.71 | 0.22 | 0.72 | 0.21 | |
| MGH | 0.86 | 0.08 | 0.84 | 0.09 | |
| MWG | 0.87 | 0.07 | 0.80 | 0.11 | |
| Agilent | 0.95 | 0.03 | 0.88 | 0.07 | |
| CompuGen | 0.83 | 0.09 | 0.87 | 0.07 | |
| Operon | 0.87 | 0.07 | 0.85 | 0.06 | |

Table 4: Intra-platform Correlations for log₂ ratios - (a) non-filtered (b) filtered

| | Affy | GE | Mergen | ABI | cDNA | MGH | MWG | Agilent | Comp | Operon |
|-----------|------|------|--------|------|------|------|------|---------|------|--------|
| Affy | 1 | 0.78 | 0.71 | 0.79 | 0.31 | 0.65 | 0.6 | 0.87 | 0.39 | 0.64 |
| GE Health | 0.78 | 1 | 0.73 | 0.79 | 0.28 | 0.64 | 0.59 | 0.67 | 0.38 | 0.65 |
| Mergen | 0.74 | 0.75 | 1 | 0.73 | 0.32 | 0.58 | 0.54 | 0.65 | 0.35 | 0.58 |
| ABI | 0.84 | 0.8 | 0.79 | 1 | 0.29 | 0.66 | 0.62 | 0.7 | 0.38 | 0.63 |
| cDNA | 0.35 | 0.31 | 0.37 | 0.29 | 1 | 0.3 | 0.35 | 0.25 | 0.05 | 0.26 |
| MGH | 0.67 | 0.66 | 0.59 | 0.68 | 0.3 | 1 | 0.53 | 0.63 | 0.34 | 0.51 |
| MWG | 0.59 | 0.6 | 0.61 | 0.64 | 0.36 | 0.55 | 1 | 0.67 | 0.34 | 0.49 |
| Agilent | 0.72 | 0.69 | 0.73 | 0.73 | 0.25 | 0.64 | 0.66 | 1 | 0.35 | 0.58 |
| Comp | 0.36 | 0.37 | 0.37 | 0.39 | 0.07 | 0.4 | 0.33 | 0.33 | 1 | 0.35 |
| Operon | 0.66 | 0.65 | 0.61 | 0.66 | 0.3 | 0.55 | 0.5 | 0.61 | 0.33 | 1 |

Table 5: Correlation coefficient: UniGene clusters

| | Affy | GE | Mergen | ABI | cDNA | MGH | MWG | Agilent | Comp | Operon |
|-----------|------|------|--------|------|------|------|------|---------|------|--------|
| Affy | 1 | 0.84 | 0.87 | 0.85 | 0.29 | 0.8 | 0.7 | 0.79 | 0.57 | 0.81 |
| GE Health | 0.89 | 1 | 0.84 | 0.84 | 0.3 | 0.79 | 0.68 | 0.79 | 0.61 | 0.8 |
| Mergen | 0.31 | 0.37 | 1 | 0.88 | 0.31 | 0.78 | 0.68 | 0.74 | 0.65 | 0.81 |
| ABI | 0.52 | 0.58 | 0.3 | 1 | 0.28 | 0.77 | 0.68 | 0.74 | 0.65 | 0.8 |
| cDNA | 0.82 | 0.38 | 0.36 | 0.28 | 1 | 0.31 | 0.36 | 0.26 | 0.22 | 0.3 |
| MGH | 0.82 | 0.79 | 0.77 | 0.79 | 0.38 | 1 | 0.61 | 0.68 | 0.6 | 0.69 |
| MWG | 0.71 | 0.66 | 0.71 | 0.68 | 0.14 | 0.63 | 1 | 0.68 | 0.56 | 0.64 |
| Agilent | 0.82 | 0.77 | 0.8 | 0.79 | 0.24 | 0.72 | 0.65 | 1 | 0.54 | 0.69 |
| Comp | 0.65 | 0.62 | 0.67 | 0.55 | 0.21 | 0.68 | 0.53 | 0.61 | 1 | 0.69 |
| Operon | 0.83 | 0.82 | 0.82 | 0.84 | 0.34 | 0.74 | 0.63 | 0.73 | 0.64 | 1 |

Table 6: Correlation coefficient: RefSeq exons

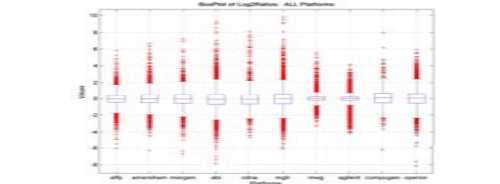


Figure 3: Box-plots of the log₂ ratios (filtered) for all platforms

| | Affy | GE | Mergen | ABI | cDNA | MGH | MWG | Agilent | Comp | Operon | Roche UPL |
|-----------|------|------|--------|------|------|------|------|---------|------|--------|-----------|
| Affy | 1 | 0.10 | 0.53 | 0.93 | 0.89 | 0.56 | 0.58 | 0.33 | - | 0.70 | 0.81 |
| GE Health | 0.64 | 1 | 0.69 | 0.88 | 0.43 | 0.96 | 1.00 | 0.97 | - | 0.28 | 0.46 |
| Mergen | 0.93 | 0.54 | 1 | 0.86 | 0.77 | 0.99 | 0.48 | 0.49 | - | 0.83 | 0.85 |
| ABI | 0.70 | 0.48 | 0.84 | 1 | 0.67 | 0.96 | 0.49 | 0.18 | - | 0.79 | 0.76 |
| cDNA | 0.89 | 0.69 | 0.94 | 0.90 | 1 | 0.53 | 0.44 | 0.19 | - | 0.52 | 0.72 |
| MGH | 0.92 | 1.00 | 0.99 | 0.98 | 0.59 | 1 | 0.94 | 0.94 | - | 0.81 | 0.88 |
| MWG | 0.26 | 1.00 | 0.25 | 0.48 | 0.40 | 0.38 | 1 | 0.88 | - | 0.44 | 0.11 |
| Agilent | 0.53 | 0.40 | 0.48 | 0.60 | 0.38 | 0.99 | 0.99 | 1 | - | 0.50 | 0.84 |
| Comp | - | - | - | - | - | - | - | - | 1 | - | - |
| Operon | 0.71 | 0.38 | 0.46 | 0.91 | 0.84 | 0.99 | 0.25 | 0.70 | - | 1 | 0.55 |
| Roche UPL | 0.75 | 0.88 | 0.75 | 0.73 | 0.67 | 0.61 | 0.80 | 0.20 | - | 0.81 | 1 |

Table 7: Validated Retina-Related Genes using Roche UPL

| | Affy | GE | Mergen | ABI | cDNA | MGH | MWG | Agilent | Comp | Operon | Taqman |
|-----------|------|------|--------|------|------|------|------|---------|------|--------|--------|
| Affy | 1 | 0.89 | 0.92 | 0.94 | 0.41 | 0.85 | 0.74 | 0.85 | 0.71 | 0.92 | 0.93 |
| GE Health | 0.88 | 1 | 0.90 | 0.90 | 0.46 | 0.84 | 0.59 | 0.80 | 0.66 | 0.93 | 0.92 |
| Mergen | 0.90 | 0.86 | 1 | 0.92 | 0.44 | 0.86 | 0.56 | 0.79 | 0.65 | 0.91 | 0.94 |
| ABI | 0.94 | 0.90 | 0.89 | 1 | 0.44 | 0.91 | 0.62 | 0.82 | 0.78 | 0.92 | 0.94 |
| cDNA | 0.48 | 0.45 | 0.40 | 0.48 | 1 | 0.35 | 0.14 | 0.48 | 0.40 | 0.47 | 0.40 |
| MGH | 0.91 | 0.87 | 0.79 | 0.89 | 0.46 | 1 | 0.58 | 0.71 | 0.71 | 0.88 | 0.85 |
| MWG | 0.60 | 0.59 | 0.52 | 0.58 | 0.20 | 0.36 | 1 | 0.68 | 0.12 | 0.65 | 0.63 |
| Agilent | 0.84 | 0.79 | 0.82 | 0.85 | 0.42 | 0.86 | 0.90 | 1 | 0.62 | 0.82 | 0.84 |
| Comp | 0.74 | 0.69 | 0.74 | 0.82 | 0.34 | 0.76 | 0.81 | 0.68 | 1 | 0.60 | 0.75 |
| Operon | 0.90 | 0.84 | 0.87 | 0.88 | 0.92 | 0.85 | 0.98 | 0.91 | 0.69 | 1 | 0.92 |
| Taqman | 0.92 | 0.89 | 0.92 | 0.93 | 0.44 | 0.88 | 0.95 | 0.87 | 0.79 | 0.88 | 1 |

Table 8: Validated Microarray results using ABI TaqMan® Assays

| | Affy 1 | Affy 2 | GE 1 | GE 2 | Mer 1 | Mer 2 |
|--------|--------|--------|------|------|-------|-------|
| Affy 1 | 1 | 0.89 | 0.81 | 0.81 | 0.75 | 0.76 |
| Affy 2 | 0.89 | 1 | 0.78 | 0.78 | 0.74 | 0.74 |
| GE 1 | 0.76 | 0.73 | 1 | 0.93 | 0.78 | 0.72 |
| GE 2 | 0.75 | 0.71 | 0.87 | 1 | 0.77 | 0.72 |
| Mer 1 | 0.74 | 0.72 | 0.73 | 0.70 | 1 | 0.79 |
| Mer 2 | 0.69 | 0.68 | 0.65 | 0.66 | 0.74 | 1 |

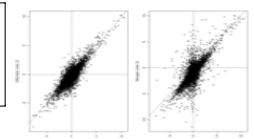


Figure 4: Cross-Laboratory scatter-plots

Table 9: Cross-Laboratory comparison

Conclusion

Stringent pre-processing techniques (normalization, transformation, filtering) are necessary any cross-platform comparisons.

- Consistency of measurements mapped at the exon level demonstrates that the data from different platforms can be compared.
- Dynamic ranges for all platforms were similar, whereas Agilent and MWG arrays displayed the highest ratio compression.
- Ability to reliably detect low expressed genes is still a limitation of microarrays in general, in contrast to the accuracy of measuring highly expressed genes.
- Cross-laboratory variations are significantly smaller than the cross-platform variations
- QRT-PCR by ABI TaqMan and Roche Universal ProbeLibrary confirmed our results. We found that UPL has really enabled high-throughput quantitative validations for our cross-platform study.
- Extensions to our analytical approach are continuing with the goal of improving data quality and exchangeability.

Acknowledgements

We like to thank Applied Biosystems, GE Healthcare, Illumina, Mergen, MWG-Biotech, and Roche Diagnostics for supplying materials for this study. WPK was supported by NIH-EY014466 grant. CLC was supported by the Howard Hughes Medical Institute. FL and EH were supported by the functional genomics program (FUGE) in the Research council of Norway. GMC was supported by NIH-NHGRI-CEES. MF, BS, GS were supported by PGA grants HL66678 and HL72358. RB was supported by NIH grants HL072370 and ESO11387.

References

- (1) Kuo, W.P., Jenssen, T.K., Butte, A.J., Ohno-Machado, L., Kohane, I.S. (2002) Bioinformatics, 18(3):405-12.
- (2) Shi, L., Tong, W., Goodrich, F., Frueth, F.W., Fang, H., Han, T., Fusco, J.C., Casciano, D.A. (2004) Expert Rev Mol Diagnostics, 4(4):761-77

2006 ABRF Poster Award Finalist; We like to thank the ABRF Evaluation Committee and GE Healthcare for sponsoring this event.