



Sixth Framework Programme for Quality of Life and Management of
Living Resources

Project no. LSHG-CT-2006-037686

EMERALD

Empowering the Microarray-Based European
Research Area to Take a Lead in Development
and Exploitation

EU Deliverable: D2.1

Due Date: Sept 2007

Delivery Date: Dec 2007

Version 1

Partner responsible: LGC

Sixth Framework Programme for Quality of Life and Management of
Living Resources

Development of microarray reference standards

Table of Contents

Executive Summary	3
Introduction.....	3
Reference materials available for array standardisation	4
Universal reference samples.....	4
Spike-in reference controls.....	6
Array Standardisation Initiatives.....	7
External RNA Controls Consortium (ERCC)	7
Microarray Quality Control (MAQC) Initiative	8
LGC microarray standardisation programme	9
Conclusion.....	11

Executive Summary

The objective of this workpackage is to plan and advocate the use of standards by the microarray community. This will involve the identification of suitable reference materials (external spike-ins, universal reference samples etc.), the assessment of analytical “best practice” guidelines and standardised approaches to experimental design and execution. This report summarises some of the ongoing global efforts to develop array reference standards and describes some of the reference materials currently available.

Introduction

Despite the widespread adoption of DNA microarray technology, there remains considerable uncertainty regarding the quality of data obtained using these technologies. The data originating from these assays frequently exhibits large and often unexplained variations between experiments, platforms and laboratories. The multitude of methods available to analyse the data can further result in drastically different interpretations of the same dataset.

The complex nature of a microarray experiment introduces many potential sources of variability. These include:

- Experimental design (including level of replication)
- Sample extraction
- Sample quality
- Array design
 - Probe design, position within the transcript, length and homology to other sequences
 - Replicates and controls
- Labelling protocol
- Hybridisation conditions
- Wash conditions
- Scanning instrument
- Image processing
- Data normalisation and analysis
- Data quality assessment
- Interpretation of results

All of these variables may contribute to the overall uncertainty of the conclusions drawn from an array experiment. Comparing results from seemingly identical experiments between different laboratories or even days can prove challenging, these challenges increase further when data from different array platforms needs to be compared. For all of these stages, particularly when related to spotted arrays, there is currently a lack of consensus and standardisation amongst the scientific community regarding best practice. Consequently, comparability and reproducibility issues often arise, not only from data produced in different laboratories or by different users, but also from data produced by the same operator and using the same protocol.

Adoption of DNA microarrays by an increasing number of disciplines outside of the research setting, coupled with growing confusion regarding the value of the large datasets generated by such experiments has now attracted the attention of standards and regulatory bodies who have started to issue guidance on data generation and handling. The following table lists several of the guidance documents currently available.

Organisation	Title	Reference
Food And Drug Administration (FDA)	Guidance for Industry Pharmacogenomic Data Submissions —Companion Guidance	http://www.fda.gov/cder/guidance/7735dft.pdf
European Medicines Agency (EMA)	Reflection paper on pharmacogenomic samples, testing and data handling	EMA/CHMP/PGxWP/201914/2006
Organisation for Economic Co-operation and Development (OECD)	Report of the OECD/IPCS workshop on toxicogenomics	http://www.oilis.oecd.org/oilis/2005doc.nsf/LinkTo/NT00001002/\$FILE/JT00183336.PDF
Clinical and Laboratory Standards Institute (CLSI)	Diagnostic Microarrays	http://www.clsi.org/source/orders/free/mm12AF.pdf
	Use of External RNA Controls in Gene Expression Assays	http://www.clsi.org/source/orders/free/mm16a.pdf

Table 1 Guidance documents for array based assays

To gain regulatory approval and become fully integrated in to drug development and medical practice array platforms need to demonstrate sufficient sensitivity, specificity, reproducibility, robustness, reliability, accuracy, precision and clinical relevance. To demonstrate this there is a need for:

- Calibrated reference materials
- Analytical “best practice” guidelines
- Standardised approaches to experimental design and execution
- Benchmark datasets
- Metrics/thresholds for assessing experimental performance
- Independent validation
- Guidelines for QC and data analysis

Reference materials available for array standardisation

Current approaches for incorporating reference materials in to microarray experiments include the use of universal reference samples and spike-in controls. The following sections outline some of the approaches and initiatives underway to develop and incorporate reference standards and materials in to the array process.

Universal reference samples

A universal reference sample works by generating uniform fluorescent signals across the vast majority of probes on an array, thus providing a base level against which the relative

abundance of transcripts from test samples can be measured. The ideal standard should be readily available, inexpensive, invariant over time and from laboratory to laboratory, and should represent all genes with a uniform signal.

Pooled RNA (either from multiple tissues or cell lines) is typically used as universal reference samples in gene expression microarray experiments. Cell lines in particular provide an abundant source of RNA and may be more stable over time and from batch to batch than tissue samples. However, tissue samples may be more biologically representative than cell lines which have undergone several passages. Concerns over this approach, however, have been raised as the pooling of RNA from several cell lines or tissues may result in some low abundance transcripts being diluted beyond the limit of detection.

From a standardisation point-of-view, the use of in vivo RNA is sub-optimal for several reasons:

- 1) The relative expression of transcripts within and between cell lines is not necessarily stable over time and may result in batch-to-batch variation;
- 2) the application of a primary RNA standard against which rigorous quality control can be performed is challenging due to the unstable nature of the raw material;
- 3) the difficulty of reliably measuring absolute quantities of each individual transcript within the reference sample makes sample characterisation a challenge.

Commercially available sources of universal reference RNA include:

- **Stratagene** - Universal Reference RNA
(<http://www.stratagene.com/manuals/740000.pdf>)
 - Prepared from 10 cancer cell lines:
 - Adenocarcinoma (mammary gland), hepatoblastoma (liver), adenocarcinoma (cervix), embryonal carcinoma (testis), glioblastoma (brain), melanoma, liposarcoma, histiocytic lymphoma (macrophase; histocyte), lymphoblastic leukaemia, plasmacytoma (myeloma; B lymphocyte)
- **SuperArray** - XpressRef Universal Total RNA
(<http://www.superarray.com/ReferenceRNA.php>)
 - Prepared from 20 different human and fetal normal major organs
- **Biochain** – Universal RNA
(http://www.biochain.com/biochain/shop2.aspx?sub_category=Universal%20RNA)
 - Produced from major human organs (available as male and female)
 - Placenta, testis, ovary, spleen, thymus, peripheral blood lymphocytes, brain (Frontal lobe), brain (Occipital lobe), brain (Parietal lobe), brain (Temporal lobe), brain (Cerebellum), spinal cord, heart, skeletal muscle, uterus, liver, pancreas, prostate, stomach, small intestine, colon, kidney, lung, skin, fetal brain (Frontal lobe), fetal brain (Occipital lobe), fetal brain (Parietal lobe), fetal brain (Temporal lobe), fetal brain (Cerebellum), fetal liver

The use of universal reference samples derived from genomic DNA is an alternative option to RNA. The advantages of this approach in terms of standardisation are that genomic

DNA is readily available, relatively stable and provides a reasonably uniform signal across different sequences.

Spike-in reference controls

Spike-in controls typically consist of exogenous RNA transcripts or oligonucleotides, which are added to the test samples either prior to, and/or post-reverse transcription and labelling. Spike-in controls can be used to evaluate the quality of both the mRNA reverse transcription and the microarray process including sensitivity, specificity, signal linearity, and consistency of the assay. In addition, the expected dye ratios can be determined and the differences in signal intensities due to the differences in dye incorporation and quantum yield normalized.

In vitro transcription of cRNA provides an approach for producing reference RNA samples. This procedure can be performed under controlled circumstances and allows for a thorough quality control of the material.

Labelled reverse complementary oligonucleotides can also be used for reference sample purposes. Oligonucleotides are stable, easy to quality control and quantify, and can provide consistent signals across arrays.

The following table lists several of the spike-in control materials currently available, or due to be available soon. These consist of materials that are independent of platform and several that are platform specific.

Control material	Features	Uses
External RNA Controls Consortium (ERCC) materials http://www.cstl.nist.gov/biotech/Cell&TissueMeasurements/GeneExpression/ERCC.htm	96 synthetic transcripts Spike-in external, exogenous sequences that do not naturally appear in eukaryotic samples 700 – 2000nt long, poly-adenylated	Performance characterization Track “whole-system” microarray technical performance
LGC microarray performance indicators http://www.mfbprog.org.uk/themes/theme_projects.asp?intThemeID=22&intProjectCatID=81	Panel of 12 synthetic exogenous 40mer sequences for spotting, and 6 fluorescently labelled oligonucleotides for spiking in to hybridisation mix Varying GC content and sequence homology to each other.	Hybridisation behaviour of the standards can be predicted over a range of conditions and deviations from the expected profile used to identify problems arising from spatial and pin effects, spot deposition, hybridisation/washing protocols, and scanner performance.
Stratagene Spot Report http://www.stratagene.com/lit_items/SpotReport_TB106_Q105_(LR).pdf	Exogenous sequences and corresponding RNA spikes derived from <i>A. Thaliana</i>	Allows data comparisons within and between microarray experiments (e.g. sensitivity, specificity, signal linearity, and consistency of the assay).

		Differences in signal intensities due to differences in dye incorporation and quantum yield can be normalized.
Platform specific controls (Affymetrix, Agilent, Codelink etc.)	<p>Agilent RNA Spike-In Kit - two sets of ten <i>in-vitro</i> synthesized RNA mixtures derived from Adenovirus E1A transcriptome with different concentrations in each set.</p> <p>Affymetrix GeneChip® Hybridization Control Kit- four pre-mixed biotin-labeled transcripts from <i>E.coli</i> and P1 bacteriophage (BioB, BioC, BioD, Cre) in staggered concentrations, for addition to hybridization cocktail.</p> <p>CodeLink Whole Genome control probes. Panel of 6 control probes designed against <i>E.coli</i> genes with corresponding mRNA spikes.</p>	<p>expected versus observed log ratios used to monitor a microarray workflow for linearity, sensitivity, and accuracy.</p> <p>Spike-in controls used to monitor hybridisation. BioB is only present half of the time, but BioC, BioD, & Cre should always have a present (P) call.</p> <p>Controls for sensitivity and dynamic range</p>

Table 2. Spike-in control reference materials

Array Standardisation Initiatives

Several of the initiatives currently underway to standardise and incorporate reference controls in to the array process are described in more detail in the following sections.

External RNA Controls Consortium (ERCC)

The External RNA Controls Consortium (ERCC) is composed of approximately 70 members originating from the public, private, and academic sectors. The group was initiated in 2003 to develop a set of external RNA control transcripts that can be used to assess technical performance in gene expression assays.

The remit of the ERCC is to:

- Develop and establish a certified reference material (CRM) consisting of a reference set of approximately 100 well-characterised clones comprising RNA transcripts from random unique sequences.
- To make them readily available to the scientific community
- To provide all data concerning sequence composition and any corresponding test data
- To provide standard operating procedures for their production and use.

- To provide suitable algorithms and bioinformatics tools for their assessment and evaluation.

The external RNA controls were designed to be added after RNA isolation, but prior to cDNA synthesis in order to allow monitoring of all aspects of technical performance of the assay, and to evaluate whether the results for a given experiment are consistent with defined performance criteria.

All ERCC work is intended to apply to all techniques for gene expression profiling including: quantitative, real-time reverse transcriptase polymerase chain reaction (QRT-PCR) assays as well as one-colour and two-colour microarray experiments. The ultimate goal of the study is to provide the expression profiling research community with a means of obtaining a reliable quantitative method for establishing performance and confidence in data produced, regardless of the platform or assay employed.

Microarray Quality Control (MAQC) Initiative

The MicroArray Quality Control (MAQC) project involves six FDA Centres, major providers of microarray platforms and RNA samples, EPA, NIST, academic laboratories, and other key microarray stakeholders. The purpose of the first phase of the MAQC (MAQC I) project was to provide quality control tools to the microarray community in order to identify procedural failures and to develop guidelines for microarray data analysis by providing the public with large reference datasets along with readily accessible reference RNA samples.

The MAQC project generated two sources of standard reference material representing commercially available human total RNA samples. Differential gene expression levels between the two samples were measured using microarrays and other technologies (e.g., QRT-PCR); with large amounts of array datasets generated.

These two samples consisted of RNA extracted from brain (commercially available from Ambion <http://www.ambion.com/catalog/CatNum.php?6050>) and a universal RNA reference sample (commercially available from Stratagene <http://www.stratagene.com/manuals/740000.pdf>) made from a collection of RNA pooled from ten cell lines selected to provide broad gene coverage.

The human reference RNA from the MAQC project can be used to calibrate arrays in ongoing quality control and performance validation efforts and provide a universal set of reference materials that can be used by individual laboratories to standardise and validate their operating procedures. In addition, the data set can be used for testing new data analysis and normalization algorithms to improve the microarray platforms further. The standard reference material can be used to test laboratory proficiency and will be valuable for troubleshooting microarray experiments in the field.

The major conclusions which have been drawn from this work are that, provided careful experimental design is implemented, and appropriate data analysis is performed, hybridisation array data is both reproducible and comparable between different platforms and laboratories. In addition, the study demonstrated that fold change results from microarray experiments closely correlated with those from alternatives assay formats such as qRT-PCR. Other factors of particular note include findings that indicate that inter run

variability between MAQC experiments could be principally attributed to cross platform differences, including probe cross hybridisation issues and the binding of splice variants. Interestingly, although the presence of inter day and operator factors could be identified from the data set, the magnitude and variance of this factor is within that which can be adequately coped with statistically.

A second key finding has been confirmation of the view long held in the literature that simple statistical testing as a means for ranking differentially expressed genes is no longer appropriate. Instead, the findings of the MAQC report recommend a two stage approach where appropriate filtering of the data set is initially carried out in order to reduce the transcript pool, which can then be ranked on fold-change using a non-stringent cut off. However, this is not without its caveats since genes expressed at the lower end or near to the limit of detection will tend to be missed. The take home message from the report is that there is no one unique statistical solution for dealing with microarray gene expression array data.

Further details and the results of the MAQC project have been described in detail in a series of manuscripts published in *Nature Biotechnology*, September 8, 2006. PDF files of the manuscripts are freely available at <http://www.nature.com/nbt/focus/maqc/>

The second phase of the MAQC project (MAQC II) is concerned with evaluating the predictive classifiers used to analyse data.

LGC microarray standardisation programme

A series of initiatives to address microarray standardisation issues are currently being undertaken by LGC (the UK's National Measurement Institute for chemical and biochemical measurements). One initiative is concerned with the development of "best practice" protocols, reference standards and toolkits to increase confidence in array technologies. A set of spike-in oligonucleotide performance indicators has been developed by LGC to monitor hybridisation efficiency and specificity and assist in the development, optimisation, validation and comparability of array-based assays.

The basic principle of this material is that a panel of 12 spotting standards of various sequence contexts are spotted on to a microarray in replicate. The 12 spotting standards consist of three sets of four oligonucleotides, with the four oligonucleotides within each set differing from each other in sequence by a varying number of bases at specified positions. The three sets of oligonucleotides contain 40, 50 or 60 % GC content respectively.

For each of the three sets of spotting standards a fluorescently labelled reverse complement oligonucleotide has been designed which is perfectly matched to one of the 4 standards within each set and differs in sequence by one mismatch, three mismatches or five mismatches to the remaining three standards within each set respectively. Cy3 labelled and Cy5 labelled reverse complements for each of the three sets of standards are hybridised to the microarray (6 targets in total). The relative amounts of Cy3 and Cy5 labelled oligonucleotides hybridised to the array can be varied between the sets of standards such that dose response measurements can be made.

The hybridisation behaviour of the standards can be predicted over a range of conditions and deviations from the expected profile used to identify problems arising from spatial and

pin effects, spot deposition, hybridisation/washing protocols, and scanner performance. The benefit of this approach is to allow the objective identification of factors affecting the performance of an assay and aids in the development, optimisation, and validation of array based assays and facilitates the ability to compare data between experiments, platforms and laboratories.

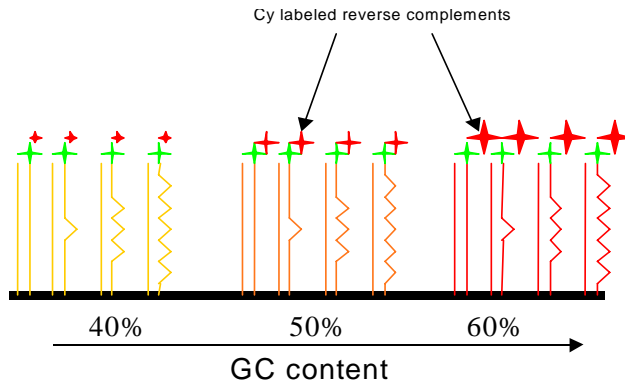


Figure 1 Schematic representation of the array performance indicators (kinks represent number of base mismatches with the target probe)

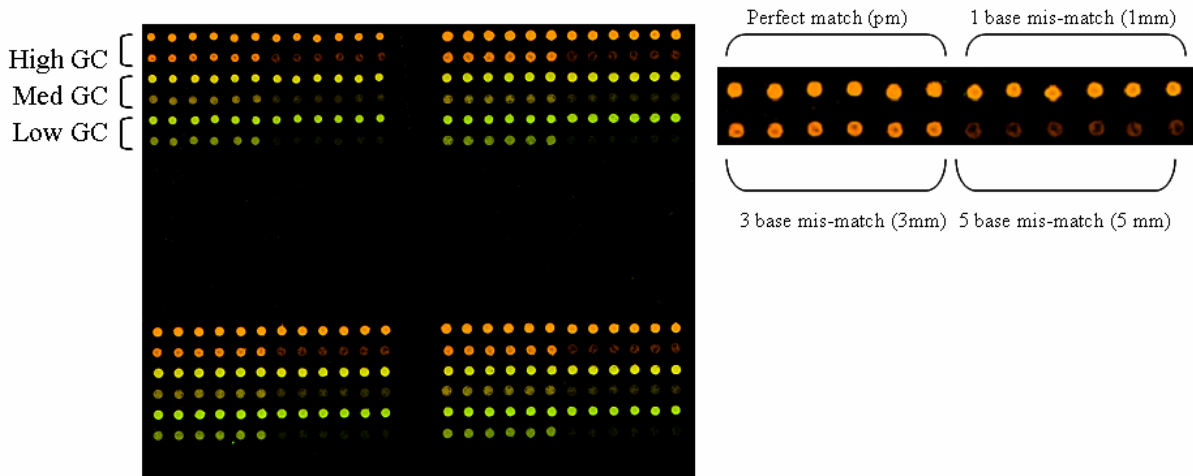


Figure 2 Typical microarray image of array performance indicators

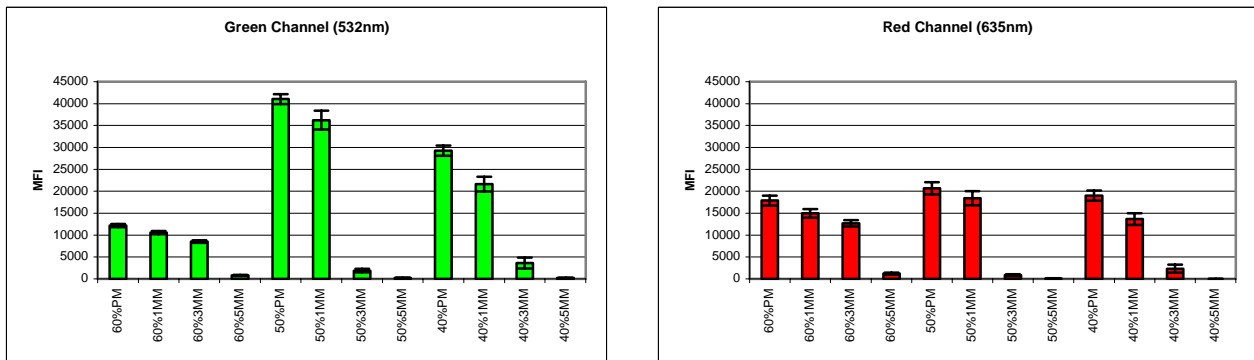


Figure 3 Example of array performance profile using Median Fluorescence Intensity (MFI) of probes

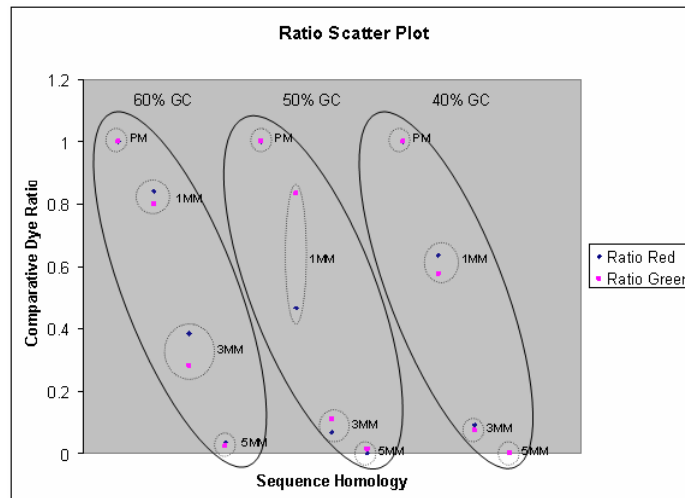


Figure 4 Performance indicators used to demonstrate discriminatory power of array assay. Ratio of signal from each spotting standard compared to "perfect match" spike in control is plotted for each set of standards and each dye layer

Conclusion

DNA microarrays facilitate the simultaneous determination of gene expression levels for large numbers of genes, however, considerable concerns have been raised as to the reliability and reproducibility of the data produced.

Inherent technical differences exist between different microarray platforms and experiments, which may have an impact on the reproducibility and reliability of the data produced. Direct comparison of different microarray platforms is not a recent consideration in terms of high-through put biological monitoring technologies. Numerous studies have highlighted problems associated with achieving a significant degree of concordance of differentially expressed genes using different platforms. Possible causes include: probe sequence differences, variation in labelling and hybridisation conditions, but more from an overall lack of industrial standards across the current multitude of array platforms employed. Currently there are very few appropriate tools available to the array community for the objective assessment of microarray data. The lack of calibrated reference materials and units of measurements have made standardisation between platforms, laboratories, and experiments a particular issue.

The report of the recently completed MAQC project supports the continued use of microarray gene expression profiling for basic and applied research but stops short of acknowledging their application in clinical diagnosis. Initiatives such as those orchestrated by organisations such as ERCC, MGED, LGC, and MAQC have provided the community with standards for data reporting, analysis tools and some useful controls.

Clearly, microarrays still require additional refinement before they can be accepted for use in the regulatory decision making process, or in areas such as clinical diagnosis. The findings and data reported by the MAQC project will in part allow microarrays to take the

next step along this path by providing a comprehensive and well characterised data set for use by the community. What is now needed alongside guidance on technical protocol standardisation is the development of appropriate quality metrics and thresholds to objectively assess the quality of microarray data along with well characterised reference control materials. Consensus agreement on the metrics to be used, followed by universal uptake and implementation by those performing array-based experiments and analysing the data will be crucial for the wider acceptance of microarrays as a valid measurement tool outside of the research setting. Initiatives are now underway through collaborative efforts by the grass-roots array community, metrology institutes, regulatory agencies and technology providers to try to achieve this goal.