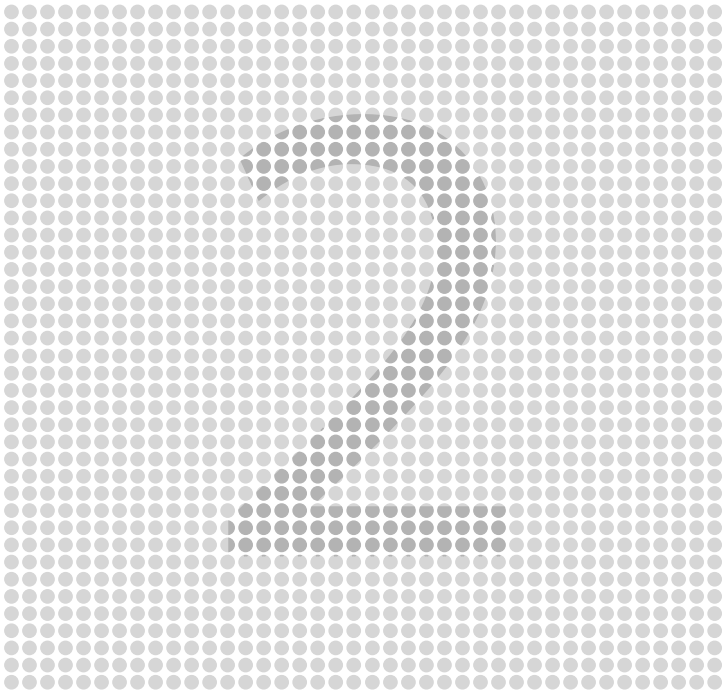

Monitoring global messenger RNA changes in externally controlled microarray experiments

Jeroen van de Peppel[1,3], Patrick Kemmeren[1,3],Harm van Bakel[2],Marijana Radonjic[1],Dik van Leenen[1] & Frank C.P.Holstege[1]

[1]Genomics Laboratory, Department of Physiological Chemistry, University Medical Centre Utrecht, Utrecht, the Netherlands.
[2]DMG Section Research, Department of Biomedical Genetics, University Medical Centre Utrecht, Utrecht, the Netherlands.
[3]These authors contributed equally to this work.

EMBO Rep. 2003 Apr;4(4):387-93

Monitoring global messenger RNA changes in externally controlled microarray experiments

Jeroen van de Peppel[1,3], Patrick Kemmeren[1,3], Harm van Bakel[2], Marijana Radonjic[1], Dik van Leenen[1] & Frank C.P.Holstege[1]

[1]Genomics Laboratory, Department of Physiological Chemistry, University Medical Centre Utrecht, Utrecht, the Netherlands.
[2]DMG Section Research, Department of Biomedical Genetics, University Medical Centre Utrecht, Utrecht, the Netherlands.
[3]These authors contributed equally to this work.

ABSTRACT

Expression profiling is a universal tool, with a range of applications that benefit from the accurate determination of differential gene expression. To allow normalization using endogenous transcript levels, current microarray analyses assume that relatively few transcripts vary, or that any changes that occur are balanced. When normalization using endogenous genes is carried out, changes in expression levels are calculated relative to the behavior of most of the transcripts. This does not reflect absolute changes if global shifts in messenger RNA populations occur. Using external RNA controls, we have set up microarray experiments to monitor global changes.

The levels of most mRNAs were found to change during yeast stationary phase and human heat shock when external controls were included. Even small global changes had a significant effect on the number of genes reported as being differentially expressed. This suggests that global mRNA changes occur more frequently than is assumed at present, and shows that monitoring such effects may be important for the accurate determination of changes in gene expression.

INTRODUCTION

Microarray expression profiling is a universal tool, with a range of applications that benefit from the accurate determination of differential gene expression (Brown and Botstein, 1999; Young, 2000). The detection of changes in messenger RNA expression requires normalization between samples. This counters non-biological variation, such as differences in labeled material, local array differences, dye-specific biases and so on (Quackenbush, 2001; Tseng et al., 2001; Yang et al., 2002).

Normalization methods consist of an algorithm and the features on the array to which the algorithm is applied. Ideally, such features should have identical signals in all the samples under investigation. Most researchers carrying out microarray analyses have dismissed the idea of using invariant house-keeping genes that are stably expressed across a wide range of experimental conditions (Lee et al., 2002). With a few exceptions (Talaat et al., 2002; Yang et al., 2002), most microarray experiments make use of the expression levels of all genes as normalization features. The assumption underlying this 'all-genes' approach is that relatively few transcript levels vary between samples, or that any changes that occur are balanced. Normalization using the expression levels of endogenous genes means that changes are calculated relative to the majority of transcripts. These relative changes do not reflect the absolute changes at the cellular level that occur if global shifts in mRNA populations take place.

The aim of this study was to set up microarray experiments, incorporating external controls, to monitor the effects of inactivating components of the generally required transcription machinery, such as RNA polymerase II (Holstege et al., 1998; Wang et al., 2002). The purpose of external normalization controls is to derive a set of signals for which the final outcome is known to be

equal among samples. This can be achieved by the addition of equal amounts of control RNA molecules to samples before processing. Here, we present results from experiments in which we observed global changes occurring under conditions that, in the past, have been studied without the use of external controls. As well as showing how global effects can be monitored, the results suggest that global mRNA changes occur more frequently than is presently assumed by researchers carrying out microarray analyses. This has important implications for the interpretation of such experiments.

RESULTS

External normalization controls

A generalized scheme for carrying out microarray analyses using external controls for normalization is shown in **Figure 2.1a**. In this study, total RNA was spiked with a mixture of nine control RNAs. The concentrations of the control RNAs were varied over three orders of magnitude to cover a range of mRNA expression levels (**Figure 2.1b**). Oligonucleotide probes (of 70 nucleotides in length) representing each control were spotted at least twice onto the microarray subgrids to generate sufficient data points (960 in total; 20 per subgrid). This allows local, expression-dependent or intensity-dependent normalization (Yang et al., 2002). The microarrays also incorporated other controls in addition to the gene probes (6,371 for *Saccharomyces cerevisiae*, each spotted twice, and 16,735 for the human arrays).

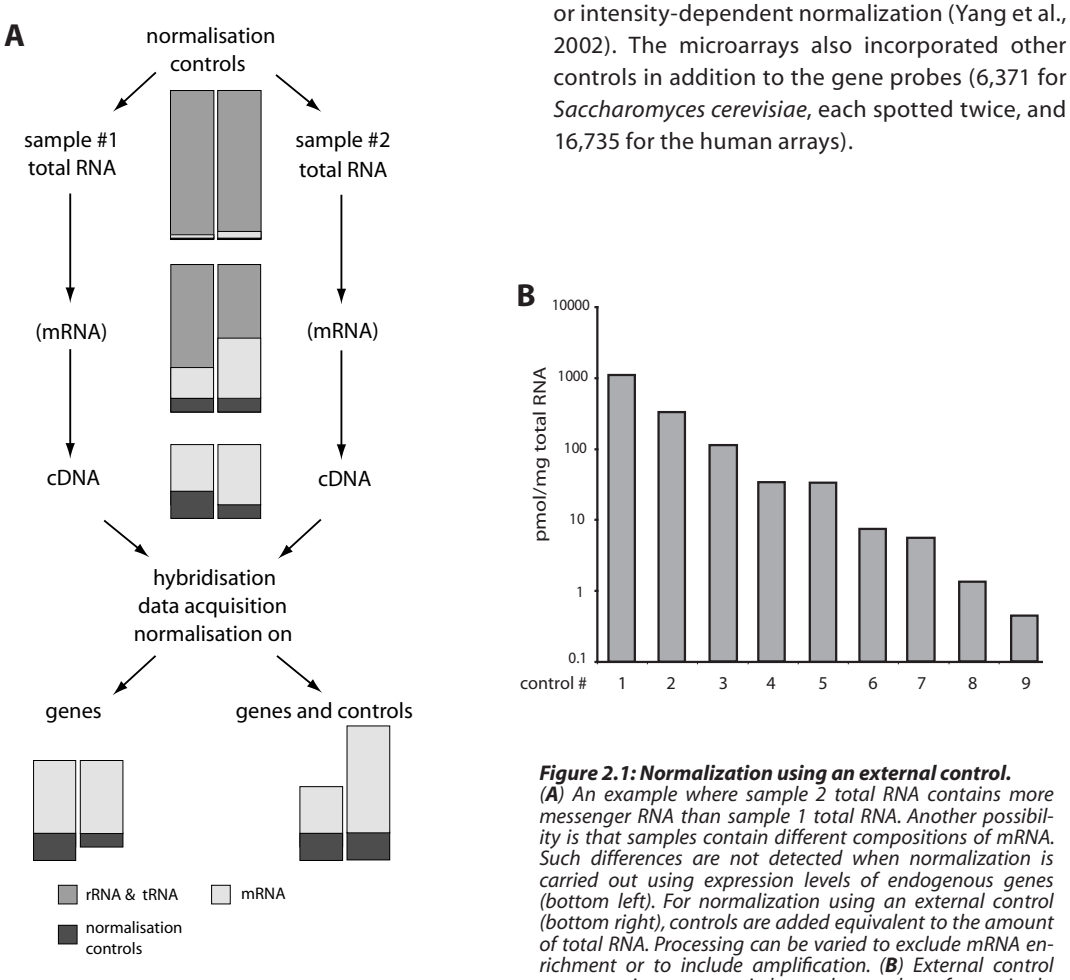

Figure 2.1: Normalization using an external control. *(**A**) An example where sample 2 total RNA contains more messenger RNA than sample 1 total RNA. Another possibility is that samples contain different compositions of mRNA. Such differences are not detected when normalization is carried out using expression levels of endogenous genes (bottom left). For normalization using an external control (bottom right), controls are added equivalent to the amount of total RNA. Processing can be varied to exclude mRNA enrichment or to include amplification. (**B**) External control concentrations were varied over three orders of magnitude. The amounts of each in vitro transcribed control RNA added as a single mix to total RNA are shown. rRNA, ribosomal RNA; tRNA, transfer RNA.*

Global changes during mammalian heat shock

One study that has shown global changes in the mRNA population is that of the heat-shock response of primary human umbilical vein endothelial cells (HUVECs). This is similar to many previous microarray studies that have examined cellular responses, and was investigated as part of our program to understand transcriptional regulation. When external controls were used, a change in global mRNA levels during heat shock was shown, as indicated by the gradual separation of the green gene spots away from the blue control spots (**Figure 2.2a-d**). This culminated in an almost twofold median drop in mRNA levels on average in the dye-swap experiments (**Figure 2.2d**). Typically, these responses have previously been normalized using expression levels of all genes. Approximately equal numbers of genes were interpreted as being up- or downregulated after normalization of the experiment in this way (**Figure 2.2e**, left column), which is a consequence of making the assumption that there is no overall change. Such changes in individual transcript levels are relative to the behavior of most of the transcripts. If a global change occurs, as is the case here (**Figure 2.2d**), normalization using genes does not correlate with changes at the cellular level. The interpretation of these experiments differs markedly when external controls are used (**Figure 2.2e**, middle column). The number of genes reported as being downregulated increased from 506 to 6,872 when an arbitrary twofold threshold was used. These changes are relative to equivalent amounts of total RNA because spiking of the controls was carried out at the total RNA stage. Counting cells before harvesting allows differences in RNA content in the two

states to be taken into account (**Figure 2.2e**, right column).

Results of small global changes during serum starvation

A second example shows the results of subjecting human mammary gland adenocarcinoma (MCF7) cells to serum deprivation for 30 h (**Figure 2.3**). Although this resulted in only a small change in global mRNA levels (1.3-fold median drop; **Figure 2.3a**), it had a significant effect on the outcome. More than twice as many genes were found to be differentially expressed as when normalization using the all-genes approach was carried out (**Figure 2.3b**).

Yeast stationary phase

To rule out the possibility that global changes are restricted to mammalian cell cultures, we also studied stationary phase in *S. cerevisiae*. This showed a 1.8-fold median drop in mRNA levels when normalized using external controls (**Figure 2.4a**). Significantly, there is a group of more than 1,000 genes that appear to be upregulated when normalized using the all-genes method, although their mRNA levels had actually decreased. This was only revealed after normalization using external controls was carried out (**Figure 2.4c**, left and middle columns). The apparent upregulation when the experiment was normalized using the all-genes approach was seen because the amount of downregulation for these genes is less than that of most transcripts. When cellular RNA content was also taken into account, the changes were found to be more extreme (**Figure 2.4c**, right column).

Changes dependent on expression levels

An important issue highlighted by the stationary phase experiment is that of an inconsistent distribution of changes across the range of expression levels: that is, along the x axis in **Figure 2.4a,b**. This results in the apparently aberrant behavior of the controls seen when normalization using the all-genes approach was carried out (**Figure 2.4b**). Highly expressed genes (on the right halves of the graphs in **Figure 2.4**) show far more downregulation. One group of highly expressed mRNAs are the ribosomal protein genes (Holstege et al., 1998).The high degree of downregulation of these genes is consistent with the significant decrease in translation that occurs during stationary phase (Dickson and Brown, 1998). Unlike approaches using endogenous transcripts, normalization using external controls is not confounded by an uneven distribution of changes, and also contributes to a more accurate determination of mRNA changes in this respect.

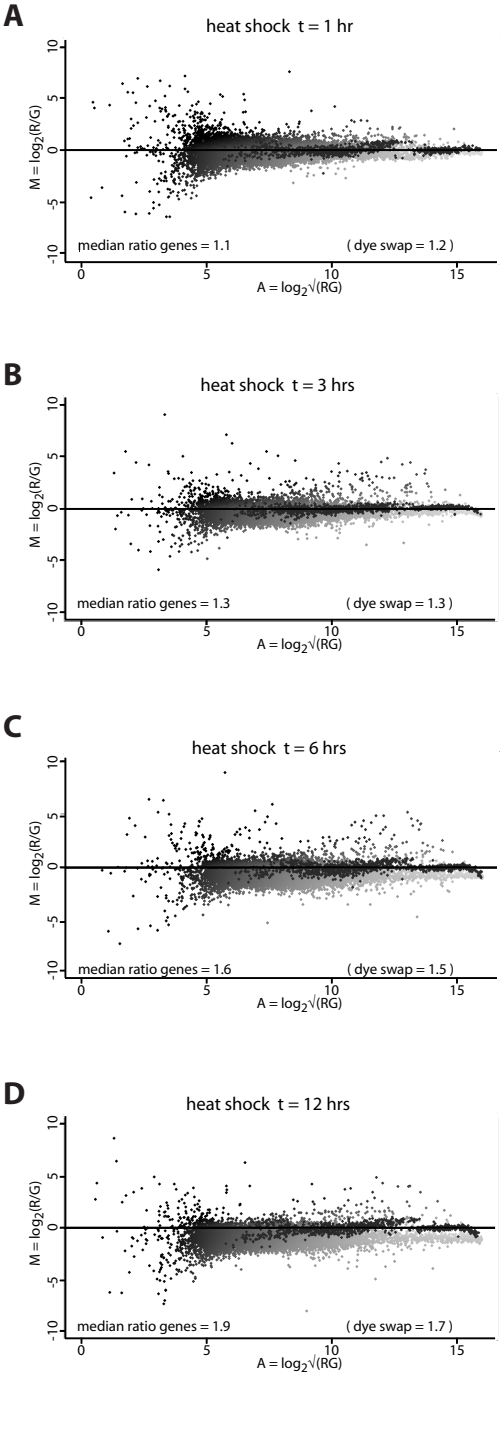

Figure 2.2: Global messenger RNA changes during mammalian heat shock.
*Human umbilical vein endothelial cells were heat-shocked and RNA was isolated at the timepoints shown (**A–D**). The result of hybridization of each sample (R) against that of the non-heatshocked reference cells (G) is shown. Each graph is an MA scatterplot (where M= log2(R/G) and A = log2√(RG); Yang et al., 2002). The y axis shows the log2 ratio (mean spot intensity minus mean local area background) after normalization using genes and controls (see Methods). The values plotted on the x axes are derived from the intensities of both channels. The median change for all genes is indicated, as well as the result of a corresponding dye-swap experiment. (**E**) Comparison of different normalization strategies. From left to right are the results of three strategies applied to the 12-h timepoint. Each column is made up of 16,735 coloured lines, stacked vertically. Each line represents the change for a single gene. Upregulated genes are shown in red and downregulated genes are shown in green. Numbers below the bars indicate the median change of all genes after normalization (as a ratio) and how many genes are reported as being upregulated or downregulated if an arbitrary twofold cut-off is applied. The first column shows the result of Lowess normalization per subgrid using endogenous genes (see Methods). The second column shows the result of incorporating external controls into the normalization strategy by carrying out normalization to equivalent amounts of total RNA (normalization using genes and controls; see Methods). This leads to a markedly altered perception of the changes that have taken place, with many more transcripts reported as being downregulated. The third column shows the result when the slight drop in the total RNA content of the cells is taken into account.*

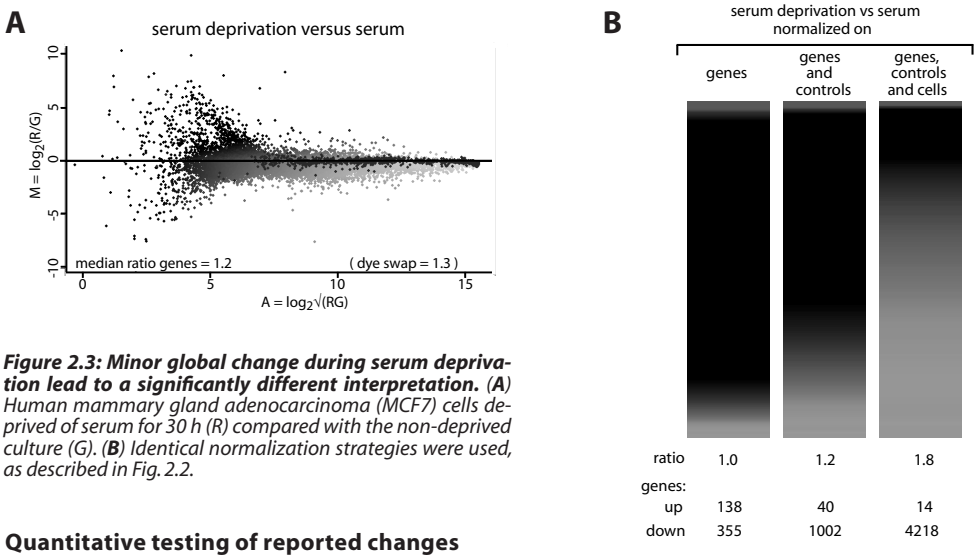

Figure 2.3: Minor global change during serum deprivation lead to a significantly different interpretation. *(**A**) Human mammary gland adenocarcinoma (MCF7) cells deprived of serum for 30 h (R) compared with the non-deprived culture (G). (**B**) Identical normalization strategies were used, as described in Fig. 2.2.*

Quantitative testing of reported changes

How accurate are the changes observed when normalization is carried out using external controls? Carrying out RT–PCR (reverse transcription followed by PCR) on a selected set of mRNAs leads to the question of what reference transcript should be chosen for the normalization of such experiments. The accuracy of the reported global mRNA changes was therefore tested by spiking the control mixtures in different ratios in identical pairs of RNA samples. The ratios chosen were 1:1, 1:2 and 1:10, thereby simulating 1-, 2- and 10-fold changes in the mRNA population relative to the controls. Normalization using external controls allowed the accurate determination of changes in mRNA levels relative to the controls. The gene spots drifted away from the control spots as the spiking ratio was increased to 1:10 (**Figure 2.5a-c**). The median amount of change observed in the expression of the genes after normalization was almost identical to the spiking ratios, with values of 1.0-, 1.9- and 9.3-fold, respectively, for the average of each dye swap. This shows that the global changes reported here (up to twofold median changes; **Figure 2.2-4**) are likely to be correct.

DISCUSSION

This study focuses on the features used for single-slide normalization and shows how the incorporation of external controls can markedly alter the interpretation of microarray experiments.

External controls have been used previously for normalization in studies examining the artificial inactivation of RNA polymerase II (Holstege et al., 1998; Wang et al., 2002). In these cases, global changes were expected. The experiments presented here show that global changes can also occur under more conventional experimental conditions. We have also shown how these controls can be applied to cope with local, intensity-dependent systematic variation (Yang et al., 2002) by representation in sufficient numbers on each microarray subgrid, and by spiking over a range of levels. As well as being useful for normalization, controls such as these are useful for monitoring sample labeling, optimization of all microarray protocols, and as external controls for the reported ratios.

Recent papers have discussed the importance of experimental design and normalization algorithm choice (Kerr and Churchill, 2001; Kroll and Wolfl, 2002; Quackenbush, 2001; Tseng et al., 2001; Yang et al., 2002; Yang and Speed, 2002). An important improvement has been the adoption of normalization algorithms that take into account local, intensity-dependent systematic variation (Yang et al., 2002). Regardless of the algorithm

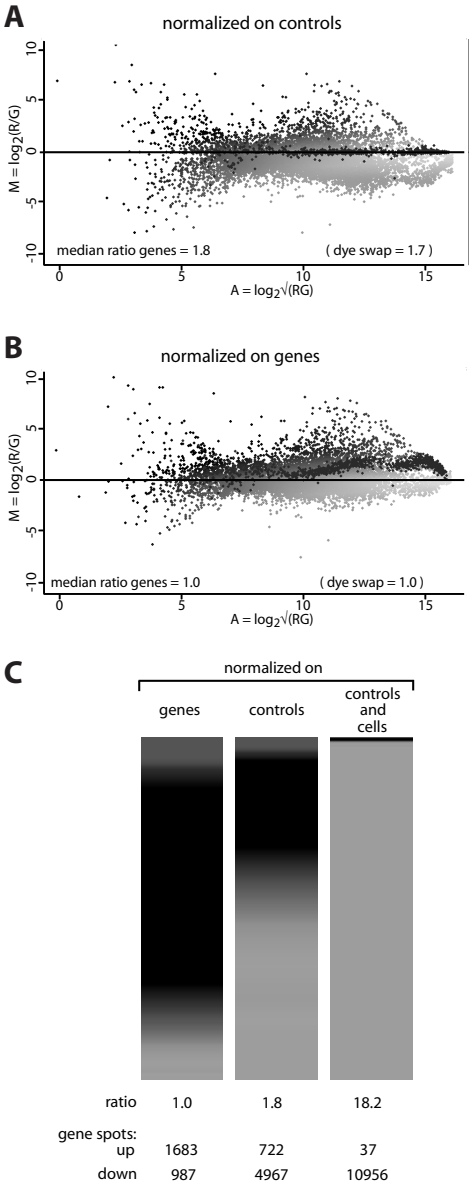

Figure 2.4: Yeast stationary phase culture compared with mid-log phase culture.
*(**A**) MA scatterplot (where M = log2(R/G) and A = log2√(R/G)) after Lowess normalization for each subgrid using controls (see Methods). (**B**) Lowess normalization for each subgrid using genes (see Methods). The aberrant pattern of external control spots (blue) occurred because messenger RNA levels had not changed uniformly across the entire range of expression levels (along the x axis). RNA levels for genes with higher expression levels had dropped to a greater degree than those of genes expressed at lower levels. Due to saturation of the signals in the scanned images, this effect seemingly decreases for a small group of genes expressed at the highest levels, resulting in the curve of the control spots to the far right of the graph. (**C**) Comparison of different normalization strategies. The left and middle columns correspond to the graphs shown in (**B**) and (**A**), respectively. The drop in total RNA yields per cell (see Methods) can also be taken into account (right column). Because each probe was spotted twice on the yeast arrays, the numbers reflect how many spots have changed. R, stationary phase culture; G, mid-log phase culture.*

applied, most current analyses rely on assumptions of evenly distributed changes and/or on the absence of global shifts. As well as the wide use of the level of expression of all genes as an invariant feature, alternative normalization features that have been proposed include housekeeping genes, spotted microarray sample pools or spots containing genomic DNA (for overviews, see (Kroll and Wolfl, 2002; Yang et al., 2002)). Normalization using such features does not reveal global, unbalanced changes. The use of a common reference sample, made up of a collection of all the DNAs represented on arrays, has also been proposed (Dudley et al., 2002; Sterrenburg et al., 2002). This allows better comparisons between slides and also overcomes the problem of obtaining negligible signals for underrepresented mRNAs. However, such approaches do not address the possibility of global, unbalanced changes. It is likely that the use of a combination of external controls and a common pooled, spotted reference sample may be the best way of overcoming several problems simultaneously.

Monitoring global effects is not required, or feasible, for every microarray experiment. The advantage of the use of external controls is that it allows the possibility of detecting such changes in an experimental set-up that is otherwise unchanged. The disadvantage is the requirement for the robust preparation of RNA samples and the reliable quantitation of yields. When RNA yields are too low to be monitored reliably, or when RNA preparations vary qualitatively, the use of external controls for normalization will not be reliable.

The importance of monitoring global and/or unbalanced mRNA changes is dictated by the goals of the experiment. For example, in disease classification studies, determination of the most extreme (relative) markers of a particular state without consideration of the actual changes does not require external control normalization.

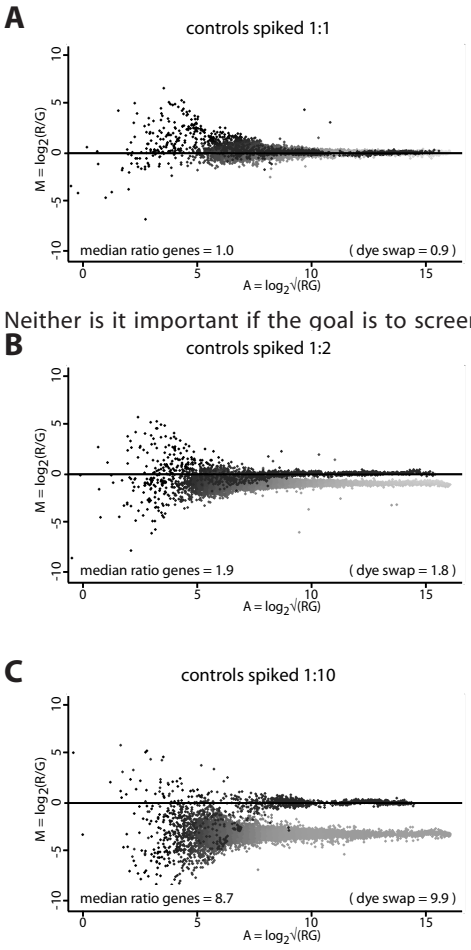

Figure 2.5: External controls can accurately detect global transcript changes.
*The external control mix was added at a ratio of (**A**) 1:1, (**B**) 1:2 and (**C**) 1:10 to paired aliquots of a single yeast total RNA preparation.*

only the most responsive genes in particular experimental conditions. However, as experimental biology becomes more comprehensive and quantitative due to the availability of whole-genome sequences and to the increased focus on systems biology (Ideker et al., 2001), it will become more of a necessity to monitor with greater precision the actual, rather than the relative, changes in different cellular states. Methods that take into account the possibility of global changes will contribute towards such goals. Examples of studies that will benefit from the use of external controls include comprehensive studies of gene regulation and analyses of drug side-effects, as well as microarray studies incorporating only limited sets of genes.

METHODS

Accession numbers and protocols

For MIAME (minimum information about a microarray experiment)-compliant (Brazma et al., 2001) protocols, data sets in Microarray Gene Expression Markup Language (MAGEML) (Spellman et al., 2002) and normalization scripts, see our website (http://www.genomics.med.uu.nl/pub/jvp/ext_controls) or the public microarray database ArrayExpress (http://www.ebi.ac.uk/microarray/ArrayExpress/arrayexpress.html) (**Table 2.1**).

Table 2.1 Arrayexpress accession numbers

Accession number	Description
A-UMCU-1	UMC Utrecht *Saccharomyces cerevisiae* 16-k array, version 1.1
A-UMCU-2	UMC Utrecht *Homo sapiens* 19-k array, version 1.0
P-UMCU- 1	*S. cerevisiae* culture
P-UMCU-2	Human umbilical vein endothelial cell culture
P-UMCU-3	MCF7 cell culture
P-UMCU-4	Total RNA isolation (*S. cerevisiae)*
P-UMCU-5	Total RNA isolation (mammalian cell culture)
P-UMCU-6	Messenger RNA enrichment (*S. cerevisiae*)
P-UMCU-7	Amino-allyl labelling
P-UMCU-8	Microarray production
P-UMCU-9	Hybridization
P-UMCU-10	Scanning protocol
P-UMCU-11	Image analysis
E-UMCU-1	Yeast spiked controls
E-UMCU-2	Human umbilical vein endothelial cell culture heat shock
E-UMCU-3	MCF7 serum deprivation
E-UMCU-4	Yeast stationary phase

MCF7, human mammary gland adenocarcinoma cell line; UMCU, University Medical Centre Utrecht.

External controls

Constructs containing *Bacillus subtilis* genes (*ycxA*, *yceG*, *ybdO*, *ybbR*, *ybaS*, *ybaF*, *ybaC*, *yacK* and *yabQ*) cloned between the *Xba*I and *Bam*HI sites in pT7T3 (Amersham Pharmacia Biotech) were made, with an additional 30-nucleotide poly(A) sequence between the gene and the *Xho*I site. For making RNA, plasmids were digested with *Xho*I for use in *in vitro* transcription reactions using MEGAscript-T7 (Ambion).

Cell culture

HUVECs were isolated as described in Jaffe et al. (Jaffe et al., 1973). Cells were cultured in endothelial

growth medium (EGM-2) at 37°C in the presence of 5% CO2. Before heat shock, the medium was removed, preheated EGM-2 was added and cells were incubated at 42.5°C in EGM-2.

MCF7 cells were cultured in DMEM/Ham's F12 medium (1:1) containing 5% fetal calf serum, glutamine (300 mg/ml), penicillin (100 inhibitory units/ml) and streptomycin (100 mg/ml). Cells at 70% confluence were grown for 30 h in phenol-red-free, serumfree medium with 0.2% BSA, transferrin (10 mg/ml) and 30 nM sodium selenite.

S. cerevisiae S288c (*MATa*; *met15*; *ura3*; *his3*Δ1; *leu2*) (Research Genetics) was grown in YEP medium (containing yeast extract and peptone) supplemented with 2% glucose. Cultures for the spiking experiments (Figure 5) were grown to mid-log phase (OD_{600} = 0.5), and for the stationary phase experiment were grown to mid-log phase or to stationary phase (OD_{600} = 10.0; 10-day culture).

RNA isolation and labeling

For mammalian cells, total RNA was prepared using Trizol (Gibco BRL) in accordance with the manufacturer's instructions. External controls were added in an appropriately diluted 5-µl mixture to 10 µg of total RNA. Yeast total RNA was prepared using hot phenol. External controls were added, as an appropriately diluted 5-µl mixture, to 500 µg of total RNA, and mRNA was isolated using Oligotex (Qiagen). Complementary DNA synthesis was carried out using 10 µg of mammalian total RNA, or 3 µg of yeast mRNA, in the presence of 2-aminoallyl-dUTP. Samples were purified using Microcon-30 (Millipore) columns and were coupled to Cy3 and Cy5 fluorophores. Before hybridization, free dyes were removed using Chromaspin-30 (Clontech) columns, and the efficiency of cDNA synthesis and dye incorporation was measured using a spectrophotometer (UV1240mini, Shimadzu).

Microarray hybridization

From each sample, 300 ng cDNA (with a specific activity of 2–4% dye-labeled nucleosides) was hybridized for 16–20 h at 42°C. Slides were scanned in a Scanarray 4000 XL (Perkin Elmer Biosystems). Image analysis was carried out using Imagene 4.0 (Biodiscovery).

Microarray production

C6-amino-linked oligonucleotides (70 nucleotides in length), the Yeast Genome ArrayReady and the Human Genome ArrayReady Oligo set (version 1.1) were purchased from Qiagen, and were printed on Corning UltraGAPS slides with a MicroGrid II (Apogent Discoveries) using 48-quill pins

(Microspot2500; Apogent Discoveries) in 3 x SSC at 50% humidity and at 18°C, and were processed by ultraviolet crosslinking (2,400 millijoules, 10 min) with a Stratalinker2400 (Stratagene).

Normalization

Algorithms were based on Lowess print-tip normalization (Yang et al., 2002), applied in the statistical package R (Ihaka and Gentleman, 1996), using the existing packages SMA (http://www.stat.berkeley.edu/users/terry/zarray/Software/smacode.html) and com.braju.sma (http://www.maths.lth.se/help/R/com.braju.sma). Alterations were made for the import of Imagene 4.0 files, flagging of control spots, Lowess line calculation on subsets of spots (either controls or genes) and extrapolation to all spots in the subgrid. This has been incorporated in an R package called genomics.sma.

The first method, normalization using the expression levels of endogenous genes, uses gene spots to calculate the Lowess line for each subgrid and then applies these lines to all spots. The second method, normalization using the expression levels of all genes, and external controls, also shifts all spots linearly according to the median ratio of the control spots. The third method, normalization using external controls, uses the controls to calculate the Lowess line for each subgrid and then applies these lines to all spots for normalization.

The second and third methods gave almost identical results for the experiments shown in **Figures 2.2, 2.3 and 2.5**. Due to the non-uniform distribution of changes in mRNA levels in the stationary phase experiment, only the third method (normalization using external controls) gave accurate results overall. This would be the method of choice for all experiments. However, there was a significant variation in the amount of the individual oligonucleotides supplied in the collections compared with the control oligonucleotides. This resulted in suboptimal extrapolation of the Lowess line to those gene spots that have extremely low intensity values, as a result of both low amounts of oligonucleotides and low or non-existent levels of mRNA. The variation in the amount of oligonucleotides supplied has been rectified in later versions of the collections.

To take into account differences in RNA yield per cell, cells were counted, and a linear shift was applied to the normalized data using the ratio derived from the total RNA yield per cell from each pair of samples.

ACKNOWLEDGEMENTS

We thank A. Dijk, R. de Roos, and I. Hamelers for cell cultures, K. Duran for technical assistance, E. Holloway, H. Parkinson and U. Sarkans for assistance with MAGE-ML generation and submission to ArrayExpress, C. Wilson and R.A. Young for external control constructs and P.C. van der Vliet, H.Th.M. Timmers, W. van Driel, R. Kerkhoven and C. Wijmenga for advice and discussions. This work was supported by the following grants from the Netherlands Organization for Scientific Research (NWO): 05050205, 90101238, 90101226, 016026009 and 90104219.

REFERENCES

Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C. A., Causton, H. C., *et al.* (2001). Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet *29*, 365-371.

Brown, P. O., and Botstein, D. (1999). Exploring the new world of the genome with DNA microarrays. Nat Genet *21*, 33-37.

Dickson, L. M., and Brown, A. J. (1998). mRNA translation in yeast during entry into stationary phase. Mol Gen Genet *259*, 282-293.

Dudley, A. M., Aach, J., Steffen, M. A., and Church, G. M. (2002). Measuring absolute expression with microarrays with a calibrated reference sample and an extended signal intensity range. Proc Natl Acad Sci U S A *99*, 7554-7559.

Holstege, F. C., Jennings, E. G., Wyrick, J. J., Lee, T. I., Hengartner, C. J., Green, M. R., Golub, T. R., Lander, E. S., and Young, R. A. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. Cell *95*, 717-728.

Ideker, T., Galitski, T., and Hood, L. (2001). A new approach to decoding life: systems biology. Annu Rev Genomics Hum Genet *2*, 343-372.

Ihaka, R., and Gentleman, R. (1996). R: a language for data analysis and graphics. J Comp Graph Statist *5*, 299-314.

Jaffe, E. A., Nachman, R. L., Becker, C. G., and Minick, C. R. (1973). Culture of human endothelial cells derived from umbilical veins. Identification by morphologic and immunologic criteria. J Clin Invest *52*, 2745-2756.

Kerr, M. K., and Churchill, G. A. (2001). Statistical design and the analysis of gene expression microarray data. Genet Res *77*, 123-128.

Kroll, T. C., and Wolfl, S. (2002). Ranking: a closer look on globalisation methods for normalisation of gene expression arrays. Nucleic Acids Res *30*, e50.

Lee, P. D., Sladek, R., Greenwood, C. M., and Hudson, T. J. (2002). Control genes and variability: absence of ubiquitous reference transcripts in diverse mammalian expression studies. Genome Res *12*, 292-297.

Quackenbush, J. (2001). Computational analysis of microarray data. Nat Rev Genet *2*, 418-427.

Spellman, P. T., Miller, M., Stewart, J., Troup, C., Sarkans, U., Chervitz, S., Bernhart, D., Sherlock, G., Ball, C., Lepage, M., *et al.* (2002). Design and implementation of microarray gene expression markup language (MAGE-ML). Genome Biol *3*, RESEARCH0046.

Sterrenburg, E., Turk, R., Boer, J. M., van Ommen, G. B., and den Dunnen, J. T. (2002). A common reference for cDNA microarray hybridizations. Nucleic Acids Res *30*, e116.

Talaat, A. M., Howard, S. T., Hale, W. t., Lyons, R., Garner, H., and Johnston, S. A. (2002). Genomic DNA standards for gene expression profiling in Mycobacterium tuberculosis. Nucleic Acids Res *30*, e104.

Tseng, G. C., Oh, M. K., Rohlin, L., Liao, J. C., and Wong, W. H. (2001). Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. Nucleic Acids Res *29*, 2549-2557.

Wang, Y., Liu, C. L., Storey, J. D., Tibshirani, R. J., Herschlag, D., and Brown, P. O. (2002). Precision and functional specificity in mRNA decay. Proc Natl Acad Sci U S A *99*, 5860-5865.

Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J., and Speed, T. P. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res *30*, e15.

Yang, Y. H., and Speed, T. (2002). Design issues for cDNA microarray experiments. Nat Rev Genet *3*, 579-588.

Young, R. A. (2000). Biomedical discovery with DNA arrays. Cell *102*, 9-15.