

## Representation of Individual Gene Expression in Completely Pooled mRNA Samples

Änne GLASS,<sup>1,†</sup> Jeannette HENNING,<sup>2</sup> Thomas KAROPKA,<sup>1</sup> Thomas SCHEEL,<sup>1</sup>  
Sven BANSEMER,<sup>1</sup> Dirk KOCZAN,<sup>3</sup> Lothar GIERL,<sup>1,\*</sup> Arndt ROLFS,<sup>2</sup> and Ulrike GIMSA<sup>2</sup>

<sup>1</sup>Institute for Medical Informatics and Biometry, University of Rostock, Medical Faculty,  
Rembrandtstr. 16/17, 18055 Rostock, Germany

<sup>2</sup>Department of Neurology, Neurobiological Laboratory, University of Rostock, Medical Faculty,  
Gehlsheimer Str. 20, 18147 Rostock, Germany

<sup>3</sup>Institute for Immunology, University of Rostock, Medical Faculty, Schillingallee 70, 18055 Rostock, Germany

Received November 8, 2004; Accepted February 28, 2005

**Designing microarray experiments, scientists are often confronted with the question of pooling due to financial constraints, but discussion of the validity of pooling tends toward a sub-pooling recommendation. Since complete pooling protocols can be considered part of sub-pooling designs, gene expression data from three complete pooling experiments were analyzed. Data from complete pooled versus individual mRNA samples of rat brain tissue were compared to answer the question whether the pooled sample represents individual samples in small-sized experiments. Our analytic approach provided clear results concerning the Affymetrix<sup>®</sup> MAS 5.0 signal and detection call parameters. Despite a strong similarity of arrays within experimental groups, the individual signals were evidently not appropriately represented in the pooled sample, with slightly more than half of all the genes considered. Our analysis reveals problems in cases of small complete pooling designs with less than six subjects pooled.**

**Key words:** complete pooling; oligonucleotide microarrays; gene expression

To pool or not to pool RNA samples from individuals for microarray hybridization is a question that has to be dealt with in experimental design if there are either financial or technological constraints,<sup>1)</sup> or if the experimental aim (e.g., screening through tissue libraries) demands this strategy. Experimental pooling has a variety of meanings in biology and microarray studies. First, it can be focused on adding a number of libraries from different tissues onto the same microarray slide to search for differentially expressed genes for all the libraries together—instead of on a library-by-library basis. Second, pooling can be used for bulk segregant

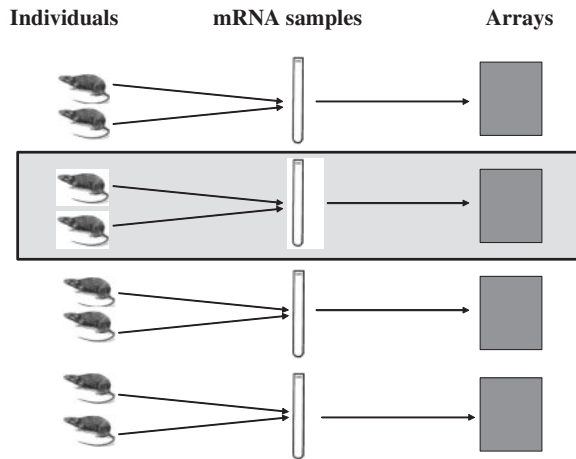
analysis for identifying and mapping genes that control simple inherited traits. Third, its aim may be to find group characteristics in the pooled sample (i.e., mean or median expression). Regardless of the reasons for mRNA pooling, it affects the gene expression data analysis and the conclusions to be drawn. The exact effects have not been systematically studied in the context of microarray experiments.<sup>2)</sup> The resulting lively and controversial discussion of the validity of pooling should therefore consider the special focus of pooling and the experimental design until this open problem has been solved. Pooling to add libraries as described above has been characterized as a non-convincing strategy.<sup>3)</sup> By contrast, pooling to find group characteristics has been characterized as a powerful, cost-effective, and rapid means of identifying the most common changes in the gene expression profiles of human colon tumors.<sup>4)</sup> In the latter case, pooling has been focused on obtaining one representative of the individual samples. This is our focus, too. Naturally, one cannot expect that each individual is equally well represented, since variations among them might not be uniformly distributed. Consequently, their representation in the pool will be biased toward outliers.

It is noteworthy that in the case of complete pooling, there is no replication of pooled samples for each treatment group. In contrast, in a sub-pooling approach there are a few subsets of samples randomly selected and pooled onto one microarray. This approach provides multiple pooled samples within each group and allows an estimation of variability among them and hence a statistical analysis. A complete pool can thus be considered a basic module of an experimental sub-pooling design (Fig. 1). Therefore, it appeared logical to analyze the complete pooling approach as described

\* This manuscript is dedicated to Lothar Gierl, deceased December, 2004.

† To whom correspondence should be addressed. Fax: +49-381-494-7203; E-mail: aenne.glass@medizin.uni-rostock.de

Abbreviations: A, absent call; C, control group; L, laboratory effect group; M, marginal call; MAS, Microarray Suite software; P, present call; r<sup>2</sup>, coefficient of determination; S, stimulation effect group; STN, subthalamic nucleus



**Fig. 1.** Example of an Experimental Sub-Pooling Design. A complete pool (light gray) can be considered a basic module of an experimental sub-pooling design.

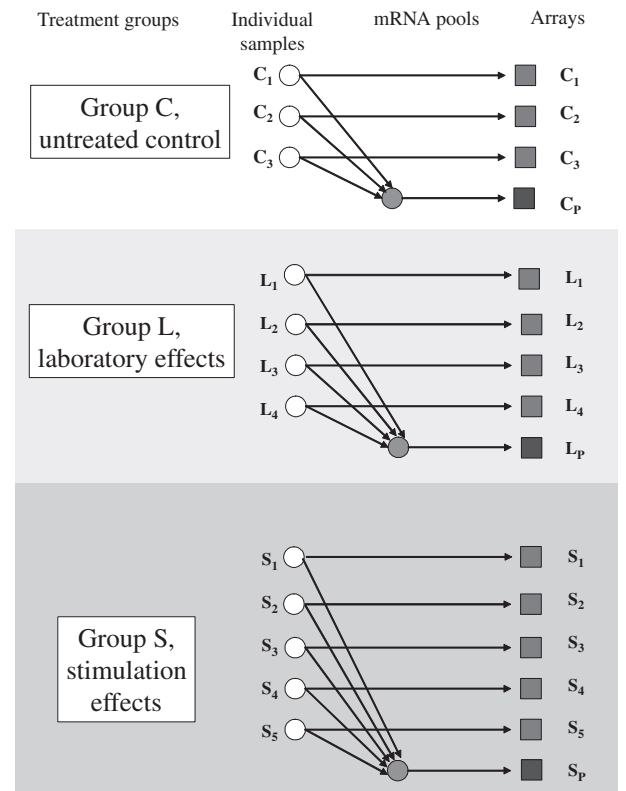
here for quality control.<sup>5)</sup> The crucial question is whether the pooled sample can be considered a representative of a set of biological replicates (individual samples which are combined in one experimental group).

Our analysis was based on three small-sized microarray experiments (Affymetrix® technology) with an untreated control group and two different effect groups. For each group, the individual mRNA samples and a sample containing pooled mRNA from all the individual samples were hybridized. The microarray data of the pooled sample were compared to those of the individual samples within the groups. The quality characteristics of these comparisons were then used to compare the three different experimental groups to ensure reproducibility of the results.

**Materials and Methods**

All experimental procedures were conducted in accordance with the regulations and licensing of local authorities. Adult male Wistar rats (Harlan-Winkelmann, Borchon, Germany) with a body weight of 280–330 g were used, and were housed under a 12 h light/dark cycle with free access to water and standard diet. Three of 12 rats were randomized for an untreated control group C (no treatment effects on gene expression). Using a stereotaxic frame (Stoelting, Wood Dale, IL), a concentric bipolar microelectrode (CB ASC; FHC, Bowdoinham, ME) was placed into the right STN of nine anesthetized rats. Four rats were randomized for an effect group L with laboratory effects caused by electrode implantation, the electrode itself, and psychological as well as mechanical stress. The remaining five anesthetized rats were stimulated for 3 h with biphasic constant-current pulses with a repetition frequency of 130 Hz and a pulse duration of 60 μs at 500 μA with a stimulus generator (Multichannel Systems, Reutlingen,

Germany). They were grouped into a stimulation effect group S. For these rats, the effects as described for group L plus effects of stimulation were to be expected. After decapitation the brains were removed. Slices of 1 mm thickness containing the STN and parts of the striatum were cut out using the guidance of a brain matrix (WPI, Berlin, Germany), and parts of the frontal cortex and the cerebellum were removed. The tissue of individual rats was incubated in RNA-later (Ambion/Europe, Huntingdon, Cambridgeshire, UK) overnight at 4 °C and stored at –80 °C until mRNA preparation. Whole RNA was extracted by tissue disruption (FastPrep™ FP120, Savant, Qbiogene, Heidelberg, Germany), and phenol/chloroform extraction followed by RNeasy™ cleanup (Qiagen, Hilden, Germany). The cRNA preparation and hybridization on Rat Expression 230 A Arrays (Affymetrix®, St. Clara, CA) was conducted according to Affymetrix® standard protocols. For each treatment group, all the individual samples as well as a complete pool of them were hybridized. See Fig. 2 for details of the experimental design. The complete pool was obtained by pooling mRNA from all individual samples of the respective group stoichiometrically prior to



**Fig. 2.** Experimental Design. Each individual mRNA and a pooled mRNA sample were hybridized onto microarrays for each treatment group. Group C: three completely untreated individuals made up the control group. Group L: electrodes were implanted into four rats without stimulation treatment. For these rats, laboratory effects caused by the implantation and stress were to be expected. Group S: five rats were stimulated for 3 h, which resulted in stimulation effects in addition to laboratory effects.

hybridization onto one microarray.

The analysis was performed within groups C (rats C<sub>1,2,3</sub> and the pooled sample C<sub>P</sub>), L (rats L<sub>1,2,3,4</sub> and the pooled sample L<sub>P</sub>), and S (rats S<sub>1,2,3,4,5</sub> and the pooled sample S<sub>P</sub>), and additionally among the groups to compare the representation results for increasing effects on gene expression and an increasing number of subjects that make up an mRNA pool. We used MAS 5.0 to process raw microarray probe set data and to generate gene expression values (“signals”) and “detection calls” with their associated p-values for every transcript presented on the arrays. Transcript detection calls were generated with Wilcoxon’s signed rank test. Qualitative calls (*i.e.*, whether a gene expression is considered present, P; absent, A; or marginal, M) are derived from these p-values. Data were scaled based on total intensity (for details see [www.affymetrix.com](http://www.affymetrix.com), Statistical Algorithms Reference Guide, Part Number 701137 Rev 3). The data from the resulting pivot tables for each group were analyzed for the signal and detection call parameters according to the question of pool representation. The statistical package SPSS Win 12.0 was used for calculation of confidence intervals for signal means.

**Signal analysis.** The coefficient of determination  $r^2$  was calculated between pairs of arrays to determine array similarity in relation to the signal parameter within groups as a prerequisite for further array analysis. We analyzed all array pairs (*e.g.*, six array pairs for the control group: C<sub>1</sub>–C<sub>2</sub>, C<sub>1</sub>–C<sub>3</sub>, C<sub>2</sub>–C<sub>3</sub>, C<sub>1</sub>–C<sub>P</sub>, C<sub>2</sub>–C<sub>P</sub>, and C<sub>3</sub>–C<sub>P</sub>).

After verifying array similarity within groups, the representation of the individuals in the pool array concerning signal values was analyzed. A gene was defined to be represented by the pooled sample regarding the parameter signal, if the pool signal value met the range [ $x_{\min}$ ,  $x_{\max}$ ] of the individual signals. For the purpose of representation, the pool signal should certainly fulfill this requirement. Otherwise, it was defined to be not represented. Each group was examined

to determine how many of the individual signals were (“insight genes”) and were not (“outside genes”) represented by the pooled sample, of the 15,923 genes.

**Detection call analysis.** The similarity of arrays for the parameter detection call was analyzed by determining the distribution of the parameter values P, A, and M for each array and comparing the results within and between groups. For representation analysis within groups, the pivot table of all group data was sorted by the three call values P, A, and M of the pooled sample.

It was examined for how many genes the individual calls matched the pool call. It was defined that the detection call of the pooled sample represented the individual samples exactly if for a given gene each of the individual detection calls had the same value as the pooled sample (*e.g.*, for a P constellation: P–PPPPP). Conversely, the representation was defined as false if all individual detection calls differed from the pooled sample call (*e.g.*, P–AAAAA). The pool representation was defined to be acceptable if most of the individual calls reflected the pool call (*e.g.*, P–PPPAA), and not acceptable for the remaining constellations like this P constellation: P–PPAAA, or this A constellation: A–PPPAA.

## Results

### *All arrays showed similar detection call patterns*

First the similarity of arrays was confirmed. For all of the 15 individual or pool arrays alike, the detection call statistics showed a call pattern of strong homogeneity. Nearly 60% of the genes had a P call, 38% had an A call, and 2% had an M call.

### *Arrays showed similar signals within each treatment group*

The arrays showed a high degree of similarity within each group in relation to the signal parameter (Table 1). The high similarity of arrays was reflected in the linear

**Table 1.** Array Similarity Regarding the Parameter Signal

$r^2$	Rat C <sub>1</sub>	Rat C <sub>2</sub>	Rat C <sub>3</sub>	Pool C <sub>P</sub>	$r^2$	Rat L <sub>1</sub>	Rat L <sub>2</sub>	Rat L <sub>3</sub>	Rat L <sub>4</sub>	Pool L <sub>P</sub>	$r^2$	Rat S <sub>1</sub>	Rat S <sub>2</sub>	Rat S <sub>3</sub>	Rat S <sub>4</sub>	Rat S <sub>5</sub>	Pool S <sub>P</sub>
Rat C <sub>1</sub>		0.97	0.96	0.97	Rat L <sub>1</sub>		0.98	0.97	0.97	0.98	Rat S <sub>1</sub>		0.98	0.98	0.96	0.97	0.97
Rat C <sub>2</sub>			0.98	0.97	Rat L <sub>2</sub>			0.98	0.97	0.99	Rat S <sub>2</sub>			0.98	0.97	0.98	0.98
Rat C <sub>3</sub>				0.97	Rat L <sub>3</sub>				0.98	0.99	Rat S <sub>3</sub>				0.97	0.97	0.98
					Rat L <sub>4</sub>					0.98	Rat S <sub>4</sub>					0.97	0.98
											Rat S <sub>5</sub>						0.98

Pairwise comparison of array similarity (coefficient of determination  $r^2$ ) for all arrays within the treatment groups C (control), L (laboratory effects), and S (stimulation effects). For details regarding the groups, see “Materials and Methods”.

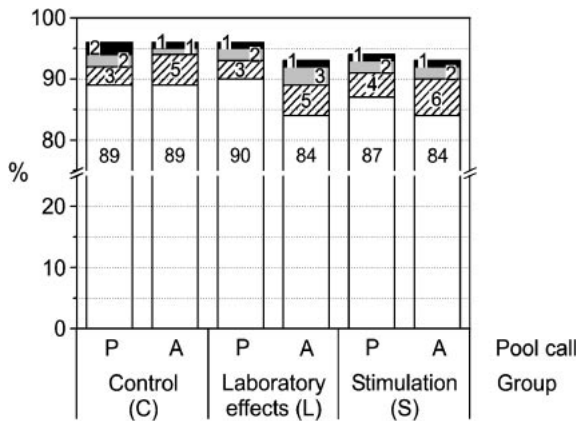
scatter plots with the indication of  $r^2$  for C: [0.96, 0.98], L: [0.97, 0.99], and S: [0.96, 0.98]. For  $r^2$ , see Table 1 (plots not shown). For comparison, the values of  $r^2$  for array pairs between the groups C and S ranged between [0.93, 0.96].

*Detection calls but not signals of individuals were acceptably represented by pool arrays in treatment groups*

Based on the array similarity, the pool representation for the individuals was analyzed for signal and detection call in groups. The representation analysis of the

parameter detection call was performed for A and P constellations of the pooled samples in groups. Constellations including M at any position were negligibly few (4%). An exact matching of the pool with all the individual arrays was detected for 84–90% of the genes (Fig. 3). A total of 3–6% of all individual detection call combinations were acceptably reflected in the pool. For 2% of the genes, the pool could not be accepted as representative, and 1% of the genes were falsely represented.

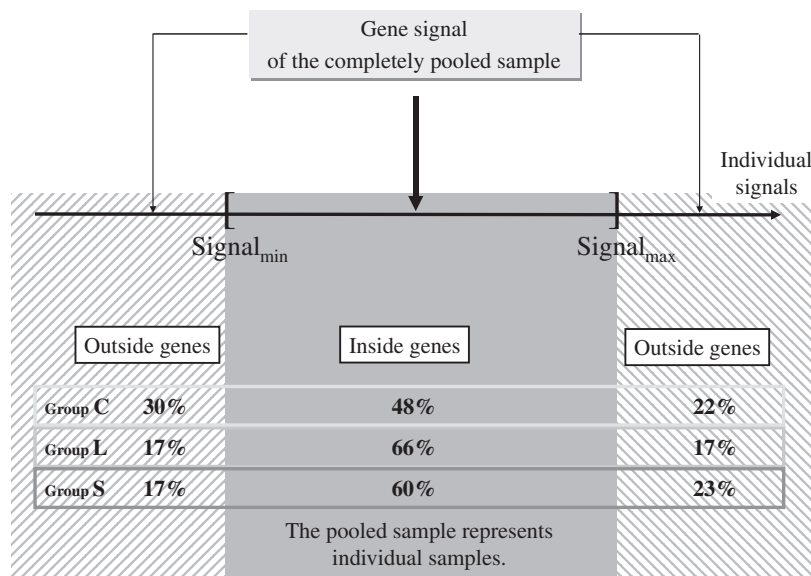
According to the definition of the representation, the pooled sample signals were detected inside the range of individual signal values for only about half to two-thirds of all genes for all groups alike (Fig. 4), which does not argue for an acceptable representation. Since weak signals are not of biological interest, we filtered only the P signals, and got the same representation. The remaining pool signals, which lay outside the ranges, were symmetrically distributed on both sides of the individual signal range. Even if the distributions of gene signals were asymmetrical,<sup>6</sup> the pool signal should clearly range between the minimum and maximum of individual signals in order to represent characteristics of individuals. Both groups of genes, whether represented or not by the signal, had the same detection call pattern as in Fig. 3.



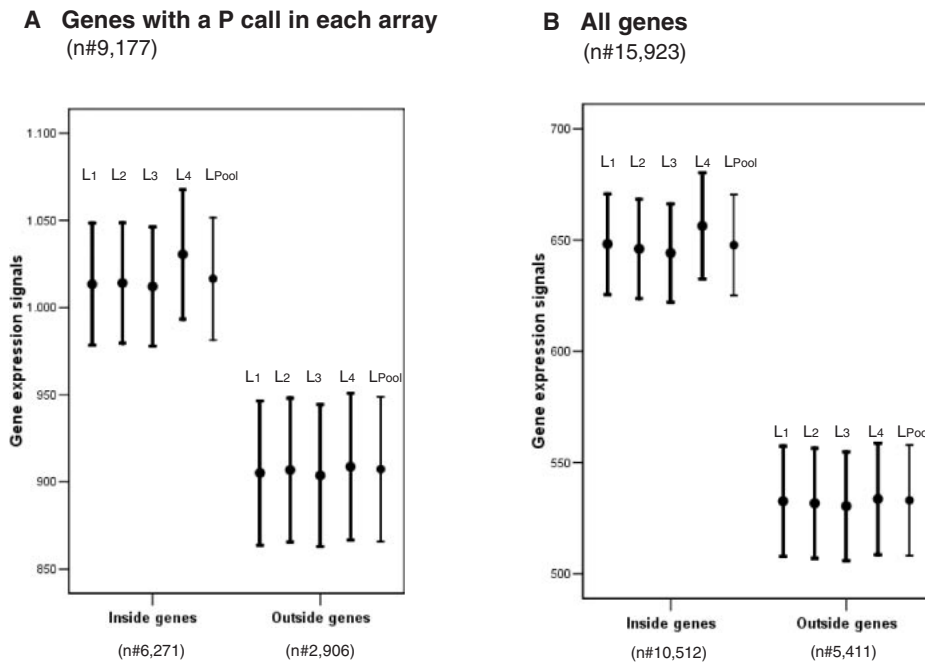
**Fig. 3.** Pool Representation of the Individual Detection Calls. Abundance of call pattern, *i.e.*, the percentage of genes for which the call of the pooled sample present (P) and absent (A) was reflected/contrasted by individual calls. Bottom-to-top, exactly (white), acceptably (crosshatched), not acceptably (gray), or falsely (black). Note that the remaining genes had a marginal call (see “Materials and Methods”, detection call analysis).

**Discussion**

This manuscript has addressed the question of the validity of complete pooling in order to obtain group characteristics in gene expression in the pool in small-sized Affymetrix experiments. Behind it is the effort to reduce the experimental costs of hybridization. Our



**Fig. 4.** Pool Representation of the Individual Signal Values. The gene expression signal of the pooled sample meets the range of the respective individual signals for about half to two-thirds of all genes (see “Materials and Methods”, signal analysis).



**Fig. 5.** Genes of the Pooled Sample Which Meet the Range of the Respective Individual Signals (inside genes) Show Higher Signals Than Those Which Do Not (outside genes).

The 95% confidence intervals of signal means for treatment group L show the same pattern for inside and outside genes for valid genes with a P call (A) and for all genes (B) alike.

microarray study with three groups of increasing interference with gene expression in genetically homogeneous rats and an accordingly increasing number of individuals making up the mRNA pool showed an acceptable representation (for about 97% of the genes) regarding detection call values and a questionable pool representation of individual samples concerning signal values (for slightly more than half of all genes) in every group. The largely constant detection call pattern (Fig. 3) throughout the experiments suggests a high quality of sample handling (RNA preparation, sample mixing, and cRNA synthesis). Therefore, we assume that a general technology problem can be excluded.

The apparent contradiction between the high quality of detection call representation and low signal representation for individual genes in the pooled arrays is due to the different ways of obtaining these two parameters. Detection call values are categorical data. They are defined by a whole range of fluorescence intensities and corrected by several control data on the microarray by MAS 5.0 software. Thus it is valid for a wider range of varying intensities, and hence is less sensitive to deviations in detected values than the signals themselves. Therefore, it is possible that the detection call shows a good representation while the signal is more prone to errors, due to deviations in scanning, for example.

It is most likely that the problems in signal representation arise from the expression value calculation by the MAS 5.0 software. Especially the background correction bears the risk of miscalculation if expression is low. This software uses a “mismatch” oligonucleotide probe whose fluorescence intensity value is subtracted from

the value of the “perfect match” probe. If the mismatch shows a higher value than the perfect match, an estimated value is used for further calculations. Thus the software itself introduces errors that might be involved in the problems observed by us. Hence we checked the characteristics of genes outside compared to genes inside. Indeed, they showed a clear difference concerning their expression levels. The outside genes have markedly lower signal levels than the inside genes (Fig. 5). We did not find any preferred incidence of certain genes in the outside areas among the C, L, and S groups.

Since the analyzed arrays were highly similar and the results obtained were comparable for each of the different treatment groups, we conclude that experimental designs with a completely pooled sample obtained from the mRNA of three to five individuals cannot be expected to represent group characteristics. The gene expression information obtained by the pooled sample is not representative of the individuals. This can be attributed to the small number of individual samples that are pooled. Indeed, in the case of unaffected gene expression in our control group C, the pooled sample from three individuals represented the typical gene expression patterns with only 48%. Interestingly, an increase in gene expression by a treatment such as in group S compared to group L does not necessarily improve the quality of representation of individuals by the pool. Perhaps the number of individuals is more important for the quality of representation than the treatment effects. While this manuscript describes an analytic approach to the still open question of the

validity of pooling in order to obtain representative data for the individual samples, Kendziorski *et al.*<sup>2)</sup> discussed the pooling of mRNA across individuals statistically. They provided formulas and calculations for the total number of arrays and a lower boundary for the total number of individuals required in a sub-pooling design to achieve a confidence interval comparable to those obtained from the no-pooling case. According to their data, the smallest sub-pooling design requires the pooling of eight individuals in a balanced manner (Fig. 1) onto four arrays instead of five individuals onto five arrays in the corresponding no-pooling case. Understandably, from the statistical point of view, an analysis of less than 5 subjects is not considered. Our analytical approach to pools obtained from three, four, or five individual samples fills the non-addressed gap of small designs. Our experimental results confirm the statistical point of view.

Nevertheless, since our analyzed design of one array of pooled mRNA per group of three to five subjects provided an almost 60% pool representation, the question arises as to whether representation would improve in the case of comparable small balanced sub-pooling designs, as mentioned by Kendziorski *et al.*<sup>2)</sup> Generally, it would be interesting to study analytically in a coordinated, well-defined experimental approach, whether sub-pooling as opposed to complete pooling<sup>7)</sup> provides an improved representation of individuals, particularly, since the complete pooling protocol used here can be part of a sub-pooling experimental design.

This manuscript refers to Affymetrix microarray technology. Whether the problems described here apply to other technologies such as Amersham or Agilent is unresolved, since results from the different technologies are not comparable.<sup>8)</sup>

## Acknowledgments

This work was supported by the German Federal Ministry of Education and Research (FKZ 01ZZ0108).

## References

- 1) Kerr, M. K., Design considerations for efficient and effective microarray studies. *Biometrics*, **59**, 822–828 (2003).
- 2) Kendziorski, C. M., Zhang, Y., Lan, H., and Attie, A. D., The efficiency of pooling mRNA in microarray experiments. *Biostatistics*, **4**, 465–477 (2003).
- 3) Anderssen, R. S., Wu, Y., Dolferus, R., and Saunders, I., An a posteriori strategy for enhancing gene discovery in anonymous cDNA microarray experiments. *Bioinformatics*, **20**, 1721–1727 (2004).
- 4) Agrawal, D., Chen, T., Irby, R., Quackenbush, J., Chambers, A. F., Szabo, M., Cantor, A., Coppola, D., and Yeatman, T. J., Osteopontin identified as lead marker of colon cancer progression, using pooled sample expression profiling. *J. Natl. Cancer Inst.*, **94**, 513–521 (2002).
- 5) Lee, J. K., Analysis issues for gene expression array data. *Clin. Chem.*, **47**, 1350–1352 (2001).
- 6) Nunez-Garcia, J., Mersinias, V., Cho, K. H., Smith, C. P., and Wolkenhauer, O., The statistical distribution of the intensity of pixels within spots of DNA microarrays: What is the appropriate single-value representative? *Appl. Bioinformatics*, **2**, 229–239 (2003).
- 7) Peng, X., Wood, C. L., Blalock, E. M., Chen, K. C., Landfield, P. W., and Stromberg, A. J., Statistical implications of pooling RNA samples for microarray experiments. *BMC Bioinformatics*, **4**, 26 (2003).
- 8) Marshall, E., Getting the noise out of gene arrays. *Science*, **306**, 630–631 (2004).