FEBS Letters 560 (2004) 120–124 FEBS 28097

Gene-specific dye bias in microarray reference designs

Alan A. Dombkowski*, Bryan J. Thibodeau, Susan L. Starcevic, Raymond F. Novak

Institute of Environmental Health Sciences, Wayne State University, 2727 Second Ave, Detroit, MI 48201, USA

Received 30 October 2003; revised 19 December 2003; accepted 7 January 2004

First published online 3 February 2004

Edited by Lukas Huber

Abstract The most widely used microarray experiment design includes the use of a reference standard. Comparisons of gene expression between samples are facilitated because each sample is directly measured against the reference standard, using two fluorescent dyes. Numerous reports indicate that some genes incorporate the two commonly used dyes with different efficiencies, contributing to inaccurate data. However, it is widely assumed that these effects will not corrupt results if the reference standard is labeled with the same dye on each microarray. We demonstrate that this assumption is not reliable and that dye orientation can significantly influence measured changes in gene expression.

Key words: Microarray; Dye bias; Reference design; Dye swap

1. Introduction

The most widely used microarray experiment design is a dual-color experiment using a common reference sample [1]. In a dual-color microarray experiment each mRNA sample is reverse-transcribed into cDNA and labeled with a fluorescent dye. The two dyes most commonly used are Cy3 (green) and Cy5 (red). Two samples, each with a different fluorescent label, are then co-hybridized on a single microarray. For each gene on a microarray the relative abundance of mRNA in the two co-hybridized samples is expressed as a ratio of the fluorescent intensities. There are two general experiment designs that enable the comparison of gene expression between biological samples: direct and indirect [2]. In the former, the two biological samples to be compared are co-hybridized on the same microarray and the ratio obtained provides a direct measure of relative expression. In an indirect comparison, each sample of interest is co-hybridized with a common reference sample used for each microarray in the experiment, thus the term reference design. The ratios obtained from each microarray in the reference design experiment can then be compared to identify genes with differential expression. One advantage of the reference design is that it allows an investigator to easily expand a study to include any number of biological samples, as long as the same reference sample is used on each microarray [1]. However, reference designs are known to increase the variance of microarray data, as compared with direct designs [2].

Systematic and gene-specific dye bias effects have been observed in dual-color experiments [2–8]. Expression intensities can be biased by differences related to incorporation and detection efficiencies of the two dyes. Systematic effects can be controlled with appropriate data normalization techniques [6,8–12]. Dye swaps, where opposite dye orientations are used on replicate microarrays, are commonly employed to control gene-specific dye bias in direct comparison designs. However, it is widely assumed that reference design experiments do not require dye swaps. It is presumed that if biological samples to be compared are labeled with the same dye, then gene-specific dye bias is eliminated in reference designs. This is based on the assumption that, if a given gene incorporates one of the dyes preferentially, the same effect will be observed in each microarray, as long as the dye orientation is consistent. Thus, when comparing expression ratios obtained from different biological samples, the gene-specific dye bias is expected to be eliminated when the expression difference is calculated according to Eq. 1, where A and B represent two biological samples of interest, and R represents the common reference sample [2].

$$\left[\mathrm{Log}\left(\frac{A}{R}\right)+\text{dye bias}\right]-\left[\mathrm{Log}\left(\frac{B}{R}\right)+\text{dye bias}\right]=\mathrm{Log}\left(\frac{A}{B}\right) \quad (1)$$

Most reference designs have relied on this assumption. It has been noted that with a conventional reference design using a single dye orientation the effects of dye and treatment on measured expression are confounded [1]. Dye swaps have been suggested as a means to avoid the confounding effect, yet most reference designs have used a single dye orientation. A small number of published reference design experiments have incorporated dye swaps to preclude effects from gene-specific dye bias [13,14].

Here, we investigate the influence of dye orientation on measured differential expression in a microarray reference design. Data were used from a microarray experiment where replicate arrays were performed in each dye orientation for each of two cell lines. The cell lines were taken from the MCF-10AT model of breast cancer progression and represent different stages of human proliferative breast disease [15,16]. Fluorescent-labeled cDNAs were prepared from total RNA samples via the widely used direct labeling protocol. Analysis of variance (ANOVA) was used to identify genes where the measured change in expression between cell lines is dependent on the dye orientation.

2. Materials and methods

2.1. Experiment design

Replicate microarrays were completed for each of two cell lines,

*Corresponding author. Fax: (1)-313-577 0082.
E-mail address: domski@wayne.edu (A.A. Dombkowski).

0014-5793/04/$30.00
doi:10.1016/S0014-5793(04)00083-3

MCF-10AT (AT) and MCF-10ATG3B (3B). A total of six microarrays were performed per cell line, with three in each dye orientation. Fig. 1A represents the experiment design. Following the convention of Churchill, each arrow represents a microarray [1]. The tail of an arrow indicates a sample labeled with Cy5, and the head of the arrow points to the sample labeled with Cy3. Each cell line was cultured on three individual plates. The RNA isolated from each plate was separated into two aliquots, one labeled with Cy5 and the other with Cy3. Each of these samples was co-hybridized on a microarray with the reference sample labeled by the alternative dye. The Agilent cDNA microarrays used in this study contain two independent microarrays per glass slide, each with a complete probe set. The two arrays can be hybridized independently with different samples. Fig. 1B shows the microarray slide layout. For the pair of oppositely labeled samples taken from a single cell culture plate, one sample was co-hybridized against the reference on the left side of the glass slide, and the other hybridized against the reference on the right side. Thus, variation between dye swap microarrays due to differences in hybridization conditions was minimized.

2.2. Cell culture

AT and 3B cell lines were used. These cell lines are derived from MCF-10A cells which are preneoplastic human breast epithelial cells [15,16]. The AT cell line was created with the insertion of mutated T24 Ha-ras gene into MCF-10A cells. AT cells sporadically progress to carcinomas when implanted subcutaneously in nude/beige mice. Cells obtained from derived carcinomas were re-established in culture and subsequently implanted back into mice three times to generate the 3B cell line. When implanted in nude/beige mice 3B cells ultimately progress to invasive carcinoma. The AT and 3B cells were cultured in Dulbecco's modified Eagle's medium/F-12 medium (Invitrogen, Carlsbad, CA, USA) supplemented with 10 μg/ml human insulin, 20 ng/ml of epidermal growth factor, 0.5 μg/ml of hydrocortisone, 5% horse serum, 100 U/ml of penicillin, and 100 μg/ml of streptomycin. The cells were maintained for 24 h in a humidified environment of 5% CO_2/95% air at 37°C. Each cell line was cultured on three separate plates and the cultures were 90% confluent at the time of RNA isolation.

2.3. RNA isolation

Total RNA was purified with the RNeasy Mini kit (Qiagen, Valencia, CA, USA), according to the manufacturer's protocol, after isolation with Trizol reagent (Invitrogen). The protocol included DNase treatment. RNA quality was assessed using an Agilent Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA) prior to use in microarray assays.

2.4. Microarrays

Agilent Human 1 cDNA microarrays were used (product number G4100A, Agilent Technologies). The microarrays contain more than 12 000 sequence-verified cDNAs from the Incyte Genomics Human UniGene 1 and Human Drug Target clone sets. For each sample, Cy3- and Cy5-labeled cDNAs were generated from 20 μg of total RNA using the Agilent Direct Labeling Kit, product number G2557A. The labeling protocol uses Cy3-dCTP, Cy5-dCTP, MMLV-RT, and oligo-dT priming. Universal human reference RNA (Stratagene, La Jolla, CA, USA), consisting of RNA isolated from 10 cell lines, was used as a common reference. On each microarray, labeled cDNA derived from the reference RNA was mixed with labeled cDNA from cell culture cells and co-hybridized. Hybridization, blocking and washing of the microarrays was accomplished using Agilent's cDNA Microarray Kit Protocol. Microarrays were scanned using an Agilent dual laser DNA microarray scanner, model G2565AA, with 10 μm resolution.

2.5. Data analysis

Microarray image analysis was performed with Agilent Feature Extraction software, version A.5.1.1. Fluorescent intensity values were adjusted using local background subtraction. Lowess intensity-dependent normalization was used to correct for systematic dye bias. Log ratios representing the relative abundance of cell culture RNA compared with reference sample RNA were derived from the normalized intensities. Log ratios were obtained for a total of 12 654 unique probes on the microarray. For probes spotted in more than one location on the array the intensity values were averaged. The data were analyzed using GeneSpring software (Silicon Genetics). Standard parametric one-way and two-way ANOVAs were performed on log-transformed data. ANOVA analysis was used to identify genes differentially expressed between the AT and 3B cell lines, with $P \leq 0.005$.

2.6. Reverse transcription polymerase chain reaction (RT-PCR)

Comparative RT-PCR was performed using the TaqMan 5′ nuclease fluorogenic assay, according to the manufacturer's instructions (Applied Biosystems, Foster City, CA, USA). Reverse transcription was carried out using 2 μg of total RNA following the protocol for the Taqman Reverse Transcription Master Mix. TaqMan primer and probe sets were obtained from Applied Biosystems using the Assay-on-Demand service. Taqman Universal PCR master mix, Taqman primer and probes, and 10 ng of cDNA were used in each reaction. Each sample was assayed in triplicate for each gene. Assays were completed using the ABI Prism 7000 Sequence Detection System (Applied Biosystems). Initial setup conditions for the thermal cycler were 2 min at 50°C followed by 10 min at 95°C. After initial setup 40 cycles were run with 15 s at 95°C followed by 1 min at 60°C. Data were analyzed using the comparative cycle threshold (C_t) method. Relative quantitation was performed comparing gene expression between the two cell lines and using GAPDH for normalization. Relative expression was calculated as $2^{-\Delta\Delta C_t}$, where ΔC_t was calculated by subtracting the average normalization C_t from the average target C_t value. The $\Delta\Delta C_t$ was then obtained by the difference in ΔC_t values obtained for each of the two cell lines.

3. Results and discussion

ANOVA has been used extensively for the analysis of replicate microarray data [13,17,18]. Here, we used two-way ANOVA to identify genes where the calculated *change* in expression between cell lines is dependent on the dye orientation. In this experiment the two ANOVA factors were cell line and dye orientation. A significant cell line × dye orientation interaction indicates a gene where the observed *change* in expression between cell lines is dependent on the dye orientation of the microarrays. Using the data from all 12 microarrays, we performed a two-way ANOVA, selecting genes with an interaction P value ≤ 0.005. Of the 12 654 unique probes on the microarrays, 698 probes had a dye orientation × cell line interaction effect with the specified significance value. Given the number of genes on the array and the significance value, approximately 63 false positives were expected by chance. Therefore, the false discovery rate of the 698 genes is 9%. We refer to this set of genes as the interaction subset.

Using the 698 genes with a significant cell line × dye orientation interaction, the fold changes between the AT and 3B cell lines were calculated in each dye orientation. Genes having ≥ 1.5-fold change in either dye orientation were selected, resulting in a total of 237 genes. For each gene, the $\log_2$(3B/AT) value was calculated according to Eq. 1. Fig. 2 presents a scatter plot comparison of the $\log_2$(3B/AT) values in each dye orientation for the 237 genes. Gene-specific dye bias is evident by the poor correlation. In many cases the direction of expression change is opposite in the two dye orientations. Other genes have the same direction of change; however, the magnitude of change is significantly different in the two orientations.

We chose several genes as examples to highlight the gene-specific dye bias effect. Fig. 3 presents the microarray data in each dye orientation for eight example genes. These genes were selected to demonstrate the effect is not unique to one dye orientation, or to expression level or fold change. Coloring is consistent with the experiment design convention used in Fig. 1. Common gene names are listed in each example.

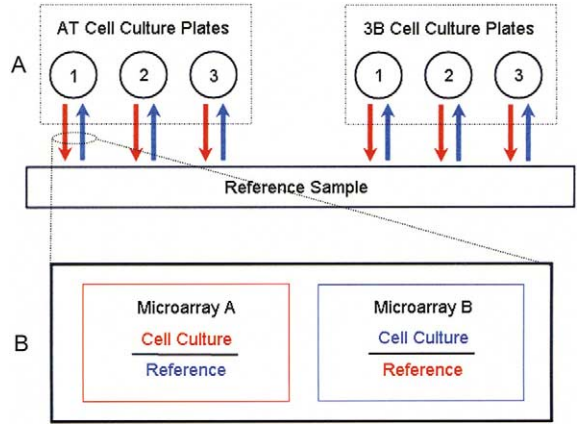

Fig. 1. Experiment design. A: Red arrows represent microarrays with cell culture sample labeled with Cy5 and reference sample with Cy3. Blue arrows indicate the opposite dye orientation. B: Layout of dye swap microarrays. Each slide contains two microarrays, each with a complete probe set. Microarray A was co-hybridized with the samples labeled per the red arrows. Microarray B was co-hybridized with the samples labeled per the blue arrows.

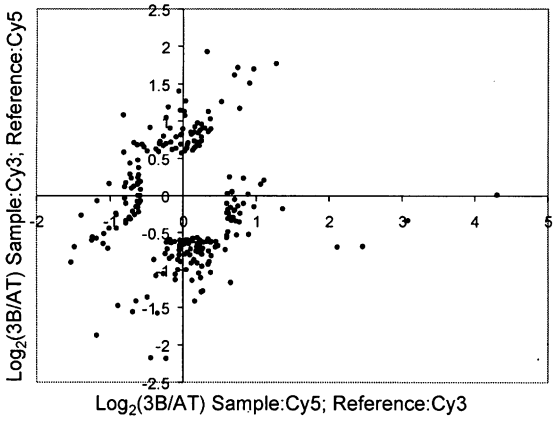

Fig. 2. Comparison of measured change in expression for 237 genes selected from the dye bias set. Each point represents a gene from the interaction set having a minimum 1.5-fold change in either dye orientation. Data values represent the average $\log_2$ ratio of gene expression in the 3B cell line with respect to the AT cell line. The *x*-axis indicates values where the cell culture samples were labeled with Cy5 and the reference sample with Cy3. The *y*-axis represents the values for the opposite dye orientation.

Data for the eight example genes are presented in Table 1, where the *P* value indicates the significance of the two-way ANOVA interaction. A fold change of 1 indicates no change in expression between the two cell lines. We note the dramatic differences in measured change of expression when comparing the two dye orientations. For example, gene AF1Q had a nearly 20-fold change in expression in one dye orientation, but in the opposite dye orientation no change was observed. The error bars in Fig. 3 indicate good data reproducibility; therefore, it is unlikely the observed effect is due to experiment noise. Interestingly, some genes have very similar ratios in both dye orientations for one cell line but distinctly different ratios for the other cell line, as observed with gene PHAP1.

Validation of fold changes was performed for several genes using TaqMan RT-PCR. Table 1 includes the fold changes measured with RT-PCR for four genes. Each dye orientation provided a more reliable measure of differential expression in half the cases, as compared with the opposite orientation. For gene ESDN, the RT-PCR data indicate a fold change of +1.5, consistent with the +1.7 observed in the orientation where the cell line samples were labeled with Cy5 and the reference with Cy3. The same orientation was also the most reliable for gene IGF1, where the microarray data indicated a fold change of +1.1 reflecting virtually no change in expression. The RT-PCR data for this gene reveal that the transcript was not abundant enough in either cell line to provide a significant measurement. The alternative dye orientation showed a −1.9-fold change in expression for IGF1. Examination of the microarray data revealed low intensity measurements for this gene. The dye orientation using cell line samples labeled with Cy3 and the reference with Cy5 was the most reliable for genes NK4 and AF1Q. This orientation provided a fold change of −2.4 for gene NK4, while the opposite orientation indicated +1.2. The RT-PCR data reveal significant repression of transcription for NK4. For AF1Q the sample-Cy3 reference-Cy5 orientation indicated no change in expression, roughly comparable with the −1.3-fold change from RT-PCR. However, the opposite orientation indicated a +19.7-fold change in gene expression.

To ascertain whether the interaction subset genes were biased by expression level, we compared intensity scatter plots of the 698 genes to scatter plots of the bulk microarray data. Fig. 4 illustrates the results for the replicate microarrays of the 3B cell line where the cell culture samples were labeled with Cy3 and the reference was labeled with Cy5. Fig. 4A shows one data point for each of the 12 654 genes on the microarray. For each gene, the intensity values were averaged using the data from the three biological replicates. The horizontal axis indicates the average intensity for the Cy5-labeled reference sample, and the vertical axis represents the average intensity for Cy3-labeled cell culture sample. Fig. 4B provides the equivalent data using only the 698 genes in the interaction

Table 1
Examples of gene-specific dye bias

Gene Name	Accession	*P*-value	Probe ID	Cell line: Cy5, Ref: Cy3 (fold change)	Cell line: Cy3, Ref: Cy5 (fold change)	RT-PCR (fold change)
PHAP1	AW468096	5.34E-07	1994721	1	+1.5	
ZG16	AAC08708	4.54E-05	2851127	+1.3	+3.8	
PFDN4	NM_002623	0.0001	43871	−2.2	−1.1	
IGF1	M37484	0.0002	1499549	+1.1	−1.9	1[a]
ESDN	D29810	4.17E-06	1402715	+1.7	−1.2	+1.5
NK4	M59807	0.0001	2418490	+1.2	−2.4	−6.5
KIAA1131	AB032957	3.64E-06	4936477	+1.2	−1.6	
AF1Q	AL038143	0.0002	1403041	+19.7	1.0	−1.3

[a]No change based on lack of detectable transcript in both cell lines.

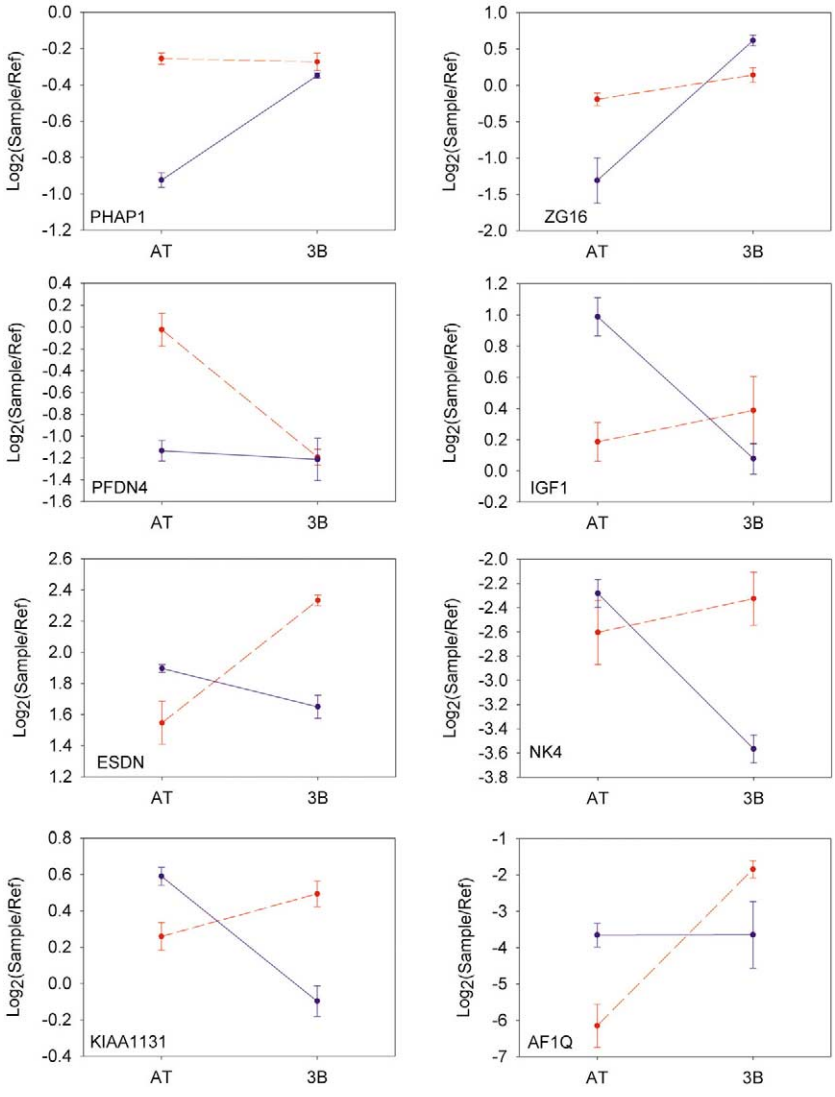

Fig. 3. Examples demonstrating gene-specific dye bias. Each plot represents one gene, with the common gene name shown. The *y*-axis reflects the average $\log_2$ ratio of cell culture sample with respect to reference sample. The *x*-axis indicates the cell culture sample. Red, dashed lines present data where the cell culture samples were labeled with Cy5 and the reference sample with Cy3. Blue lines reflect the opposite dye orientation. Error bars show standard deviation.

subset. The subset data follow the same general distribution of the bulk microarray data indicating that the interaction subset is not biased by intensity. Data from the other three cell line/dye orientation combinations were also analyzed and no observable expression-dependent bias was found in this set of genes (data not shown).

To simulate the conventional reference design approach, we segregated the data by dye orientation. For each dye orientation we performed one-way ANOVA on the six microarrays to identify genes that are differentially expressed between the two cell lines. The significance level was set at $P \leq 0.005$. With cell culture samples labeled with Cy3 and the reference with Cy5, 845 genes were identified as differentially expressed. Of these genes, 24% were identified in the interaction set, indicating that the observed change in expression for these genes may not be reliable. In the analysis where cell culture samples were labeled with Cy5 and the reference with Cy3, 889 genes were identified as having a statistically significant change in expression between cell lines. Of these, 20% were identified in the interaction set. Therefore, with the conventional single-dye orientation approach we find that over 20% of the genes identified as differentially expressed are cases where the measured change in expression would be different if observed in the other dye orientation. We note that these results were obtained using statistical methodology commonly employed to analyze microarray data, and in all cases our false discovery rate was under 9%.

4. Conclusions

The most widely used microarray experiment design is the dual-color reference design. With this approach it is usually assumed that gene-specific dye bias will be 'canceled out' when the expression ratios from two or more microarrays are compared, as long as the samples to be compared are labeled with the same fluorescent dye. Thus, most investigators have completed these experiments using a single dye orientation. We demonstrate that dye orientation can have a significant influence on the measurement and inference of differential gene expression. While we have not identified the

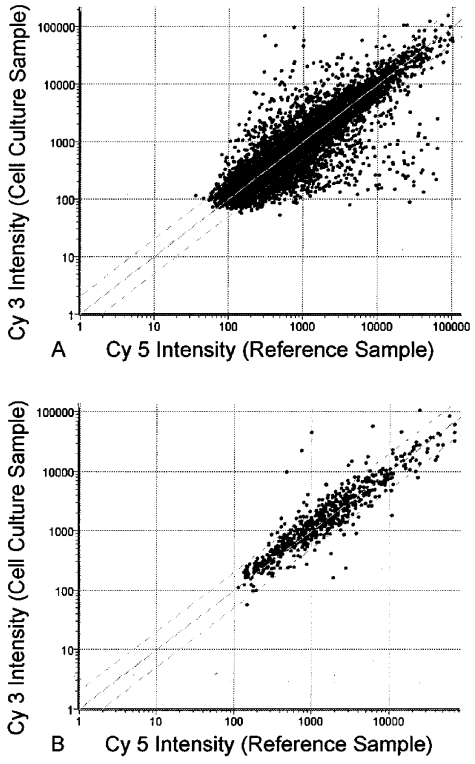

Fig. 4. Distribution of intensity values. A: Averaged data for the three microarrays where 3B cell culture samples were labeled with Cy3 and reference sample with Cy5. Each data point represents one of 12 654 genes. Each axis reflects the normalized intensity of the respective dye. B: Same data as A but limited to the 698 genes with a dependence on dye orientation.

cause of the bias, the labeling method likely contributes. In this work we utilized direct labeling which is commonly employed in microarray protocols. This method incorporates the fluorescent dyes during a reverse transcription reaction. Some investigators are now using methods that rely on chemical labeling after the required reverse transcription step, with the goal of avoiding dye bias effects [19,20]. Until dye bias effects are better characterized, our findings suggest that dye swaps are warranted in microarray reference designs. In this study, over 20% of conclusions regarding differential expression may be inaccurate using a conventional approach with a single dye orientation. This has significant implications for the interpretation of data from the growing repositories of microarray experiments.

Acknowledgements: We thank Dr. Fred Miller for the MCF-10AT and MCF-10ATG3B cell lines. This work was supported by NIH Grant ES10595 and NIEHS Center Grant P30 ES06639.

References

[1] Churchill, G.A. (2002) Nat. Genet. 32, 490–495.
[2] Yang, Y. and Speed, T. (2002) Nat. Rev. Genet. 3, 579–588.
[3] Bartosiewicz, M., Trounstine, M., Barker, D., Johnston, R. and Buckpitt, A. (2000) Arch. Biochem. Biophys. 376, 66–73.
[4] Taniguchi, M., Miura, K., Iwao, H. and Yamanaka, S. (2001) Genomics 71, 34–39.
[5] Kerr, M.K. and Churchill, G.A. (2001) Biostatistics 2, 183–201.
[6] Tseng, G.C., Oh, M.K., Rohlin, L., Liao, J.C. and Wong, W.H. (2001) Nucleic Acids Res. 29, 2549–2557.
[7] Simon, R., Radmacher, M.D. and Dobbin, K. (2002) Genet. Epidemiol. 23, 21–36.
[8] Nadon, R. and Shoemaker, J. (2002) Trends Genet. 18, 265–271.
[9] Dudoit, Y., Yang, Y.H., Callow, M.J. and Speed, T.P. (2000) Technical Report 578, Department of Statistics, UC Berkeley, CA.
[10] Yang, Y.H., Dudoit, Y., Luu, P., Lin, D.M., Peng, V., Ngai, J. and Speed, T.P. (2002) Nucleic Acids Res. 30, e15.
[11] Fang, Y., Brass, A., Hoyle, D.C., Hayes, A., Bashein, A., Oliver, S.G., Waddington, D. and Rattray, M. (2003) Nucleic Acids Res. 31, e96.
[12] Benes, V. and Muckenthaler, M. (2003) Trends Biochem. Sci. 28, 244–249.
[13] Pritchard, C.C., Hsu, L., Delrow, J. and Nelson, P.S. (2001) Proc. Natl. Acad. Sci. USA 98, 13266–13271.
[14] Brem, R.B., Yvert, G., Clinton, R. and Kruglyak, L. (2002) Science 296, 752–755.
[15] Starcevic, S.L., Elfrink, C. and Novak, R.F. (2001) J. Natl. Cancer Inst. 93, 776–782.
[16] Dawson, P.J., Wolman, S.R., Tait, L., Heppner, G.H. and Miller, F.R. (1996) Am. J. Pathol. 148, 313–319.
[17] Kerr, M.K., Martin, M. and Churchill, G.A. (2000) J. Comput. Biol. 7, 819–837.
[18] Draghici, S., Kulaeva, O., Hoff, B., Petrov, A., Shams, S. and Tainsky, M.A. (2003) Bioinformatics 19, 1348–1359.
[19] Yang, I.V., Chen, E., Hasseman, J.P., Liang, W., Frank, B.C., Wang, S., Sharov, V., Saeed, A.I., White, J., Li, J., Lee, N.H., Yeatman, T.J. and Quackenbush, J. (2002) Genome Biol. 3, research0062.1–0062.12.
[20] Richter, A., Schwager, C., Hentze, S., Ansorge, W., Hentze, M.W. and Muckenthaler, M. (2002) BioTechniques 33, 620–628.