Test of Significant Differences with *a priori* Probability in Microarray Experiments

Kyoko TODA,[†] Seiichi ISHIDA,[†] Kotoko NAKATA, Rieko MATSUDA, Yukari SHIGEMOTO-MOGAMI, Kayoko FUJISHITA, Shogo OZAWA, Jun-ichi SAWADA, Kazuhide INOUE, Koichi SHUDO, and Yuzuru HAYASHI^{††}

National Institute of Health Sciences, 1-18-1 Kami-Yoga, Setagaya, Tokyo 158-8501, Japan

A traditional method for comparing two expression levels of genes in microarray experiments is the two-sample *t*-test. Because of the difficulty in using a large number of microarrays, an alternative method is required which can provide a reliable judgment of the comparison from a small number of replicates, even from a single pair of control and treatment. We present a method for detecting the changes in the gene expression levels under two different conditions in microarray experiments. Our method targets a single experiment for each condition, while retaining the statistical advantages of the *t*-test. The new proposals are: 1) standard deviation (SD) estimates of the expression levels which are an indicator for significant differences are given *a priori* as a function of the expression levels; 2) the limit of detection (LOD) for the expression levels is used to eliminate the majority of genes expressed at extremely low levels. The *a priori* SD estimates are obtained from six replicates under a fixed condition and are shown to be the approximate, but proper description of the expression uncertainty covering diverse conditions (*e.g.*, different samples (human and rat) and different DNA chips). The LOD is defined as three times blank SD according to the IUPAC recommendation. A cell line (HL60) which will undergo macrophage differentiation on treatment with 12-*O*-tetradecanoylphorbol 13-acetate (TPA) is taken as an example. Our method is compared with the *t*-test for the data on duplicate TPA experiments and the former alone is evaluated with the data on a single TPA experiment. The errors from sample preparation and instrumental analysis are discussed.

(Received June 6, 2003; Accepted September 16, 2003)

Introduction

The advent of high-throughput array technology has now made it possible to collect data on thousands to tens of thousands of genes simultaneously. However, methods for detecting the genuine changes in the gene expression levels in cells or tissues are still evolving.¹⁻¹¹

A straightforward method for comparing two expression levels of genes is the traditional two-sample *t*-test. The basic problem with the *t*-test in microarray experiments, however, is that the repetition is restricted within a small number in most cases, because experiments are costly or tedious to repeat. Although the importance in replication has been illustrated,^{1,3,6,9} situations often arise where only single or duplicate experiments for each condition are allowed.

The purpose of this paper is to put forward a method for testing the significant differences of the gene expression levels under a single pair of experiments (a control experiment and treatment experiment). In order to take into account the stochastic aspects of gene expressions, we model our algorithm on the *t*-test. In our approach and the *t*-test, the SD estimates of the gene expression levels are a criterion for the statistical judgment, but one of the key differences is how to estimate the SD for each gene.

In the *t*-test, the SD estimates are derived from the same data set as those to be judged by the *t*-test, itself. This fact can be closely connected with the above-mentioned problem of replication. In our approach, the SD estimates, referred to here as *a priori* SD, are obtained from experimental results which are different from the target data set of the judgment.

Statistics tells that the variability in the estimates of SD obeys the chi-squares distribution and is much larger than the variability in the estimates of averages, as long as the estimates are obtained by repetition. That is, the estimates of averages are more reliable. The *t*-test uses the SD estimates directly, but in this paper, the *a priori* SD is given as a function of the average of the gene expression levels. Then, we can easily expect that our approach can provide more stable judgment, but needs a sound model for the *a priori* SD.

The idea of the *a priori* SD is not novel in the area of analytical chemistry. Since more than three decades ago, there have been published many theories and methods for estimating SD with no recourse to repetition, especially in instrumental analyses.¹²⁻²⁷ In spite of varied symbols and terminology in the literature, the largest part of uncertainty equations proposed can take a universal form:²⁴

$$\text{RSD}^2 = \frac{s_B^2}{A^2} + I^2, \quad (1)$$

where RSD denotes the relative standard deviation of measurements, s_B denotes blank SD, A measurements (*e.g.*, area), and I independent error. To our knowledge, Huber *et al.*

[†] Co-first authors.

^{††} To whom correspondence should be addressed.

K. S. present address: Japan Pharmaceutical Information Center.

first used Eq. (1) in 1971. The identical error models were adopted in microarray experiments.^{4,6} This paper also follows suit.

The mathematical formalism of uncertainty like Eq. (1) has wide applicability. Examples are LOD,^{21,24,26,28} confidence intervals of linear calibration²⁸ and tests of significant differences.^{3,4,6} Theoretical SD descriptions elaborated so far include Winefordner's theory,¹⁸⁻²⁰ Ingle's theory,¹⁵⁻¹⁷ Bouman's theory^{21,22} and FUMI theory (FUNction of Mutual Information).²³⁻²⁵

Proceeding along the lines suggested by the FUMI theory, our approach is named after it. The other salient feature of the FUMI theory in this paper is the introduction of LOD which is helpful to remove the vast majority of genes expressed at exceedingly low levels. This point is also a problem with which simple fold-change methods are accompanied.^{3,8} The FUMI theory is applied to a cell line (HL60) which will undergo macrophage differentiation on exposure to a tumor promotor, 12-*O*-tetradecanoylphorbol 13-acetate (TPA).^{29,30}

The errors due to sample preparation, before the instrumental measurement, are often a critical problem in practice. Typical experiments are planned to discern the contributions of the preparation and measurement processes to the total analytical error. This paper demonstrates that the error magnitude of preparation is even smaller than that of the measurement in our analytical system.

Materials and Methods

There are about ten thousand genes on a DNA chip used (GeneChip, Affymetrix). A probe set for a given gene on the DNA chip usually contains sixteen probe pairs, each of which is made up with perfect match and mismatch probe cells. The total RNAs prepared from a sample are enzymatically converted into fragmented, biotin-labeled cRNAs and hybridized to the probe sets. After washing and staining with phycoerythrin conjugated streptavidin, the amount of hybridized cRNAs is quantified by scanning the DNA chip with the argon-ion laser scanner. The resulting fluorescence image data are processed and given as "Signal" by a software (Microarray Suite 5.0, Affymetrix). The values of "Signal" can directly be related with the expression levels of the genes and are used as measurements of samples by the *t*-test and FUMI theory.

All the experiments including RNA isolation, hybridization, *etc.* were carried out according to the manufacturer's protocol. The cell lines used were human hepatocellular carcinoma cell line (HepG2), human promyelocytic leukemia cell line (HL60) and rat microglia. The rat cell line was obtained from primary cell cultures of neonatal Wister rat brains as described previously.³¹ The combinations with DNA chips (GeneChip, Affymetrix) were: HepG2 (U95A); HepG2 (U95B); HL60 (U95A); rat microglia (U74A).

The biotin-labeled cRNA for each cell line was stocked and later used for the repeated experiments ($n = 6$) which began with the hybridization (six arrays for each cell line).

The HL60 cells were exposed to 20 nM TPA for 1 h and the biotin-labeled cRNA was prepared and stored as a stock solution. A total of four U95A arrays were used (two with the TPA-exposed stock solution and two with the control stock solution).

The model experiments for the entire microarray analyses were carried out as follows: the total RNA was prepared from eleven culture dishes of HL60 by the RNeasy Mini total RNA preparation kit (Qiagen, Germany); cRNA was synthesized

from 10 μ g of total RNA on each dish according to the Affymetrix protocol; the cRNA was determined by ultra-violet absorption spectrometry; the RSD was calculated from the measurements ($n = 11$).

Theory

A brief review of the *t*-test and an in-depth explanation of our test are given below.

t-test and Cochran-Cox method

It is assumed that the number of replicates is two for exposure and control experiments, respectively. Let $\bar{X}_E$ be the mean of the expression levels (measurements), X_E , of a gene for exposed samples and $\bar{X}_C$ be the mean of measurements, X_C , of the gene for control samples. In the *t*-test, the expression levels of a gene are judged to be significantly different, if the absolute difference between $\bar{X}_E$ and $\bar{X}_C$, meets the condition:³²

$$\frac{|\bar{X}_E - \bar{X}_C|}{s} > 9.925, \quad (2)$$

where 9.925 is the critical value of $|t|$ at a significant level of 1%. Here, the SD estimate, s , takes the form:³²

$$s = \sqrt{(s_E^2 + s_C^2)/2}, \quad (3)$$

where s_E and s_C are the SD estimates of individual measurements, X_E and X_C , of exposure and control, respectively.

Before the *t*-test, the *F*-test is carried out for the SD estimates, s_E and s_C . If the homoscedasticity assumption ($s_E = s_C$) is rejected by the *F*-test, another critical value, 63.657, is used in Eq. (2) instead of 9.925 (Cochran-Cox method).³³ If the homoscedasticity is accepted, the *t*-test (Eqs. (2) and (3)) follows.

FUMI theory

The FUMI theory has an equivalent formalism of judgment:

$$\frac{|\bar{X}_E - \bar{X}_C|}{s} > 2.58 \quad (4)$$

where s means the SD estimate of numerator, $\bar{X}_E - \bar{X}_C$, and 2.58 is the critical value for a significant level of 1% under the assumption that the distribution of $\bar{X}_E - \bar{X}_C$ is normal.

We define σ (*a priori* SD) as the SD of the individual measurements, X_E and X_C (see Eq. (6)). If the SD of $\bar{X}_E$ is equal to the SD of $\bar{X}_C$, then the SD, s , of $\bar{X}_E - \bar{X}_C$ can be given:

$$s = \sigma \quad (= \sqrt{2} (\sigma \div \sqrt{2})); \text{ note that } n = 2). \quad (5)$$

Figure 1 illustrates the algorithm of the FUMI theory which consists of three types of judgment based on the *a priori* SD, σ (gray rhombi).

Step 1 (Difference $> 3s$; $s = \sqrt{2}\sigma$): If the duplicate expression levels for a gene under the same conditions (*e.g.*, X_C for Control 1 and Control 2) are even more different than those expected by the *a priori* SD, the gene is eliminated from the analysis.

Step 2 (Difference $> 2.58s$; $s = \sigma$; see Eqs. (4) and (5)): If the difference in the mean expression levels ($= \bar{X}_E - \bar{X}_C$) is regarded as being significant at 1% level, the gene goes to the next step. If not, it is discarded.

Step 3 (Larger data $> 2 \times \text{LOD}$; $\text{LOD} = 3(\sigma \div \sqrt{2})$): If the largest level of the control and exposure means is less than

Fig. 1 Flow chart of probabilistic significance test (FUMI theory). s , SD; LOD, limit of detection.

twice the limit of detection (LOD), the gene is removed from the analysis.

According to the IUPAC recommendation,³⁴ LOD is defined as three times blank SD, s_0 ($s_0 = \sigma$ where $X = 0$ in Eq. (6)). The blank measurements correspond to the expression levels of house-keeping genes and are assumed to vary due to a random fluctuation of detector noise or other error sources. The above LOD definition implies that the probability for a noise-created false signal being above the LOD is at most 0.13%.³⁴ If the smallest level of the control and exposure means is just the LOD, the minimum level which can be distinguished from the LOD level at a significant level of 0.13% is $2 \times \text{LOD}$.

To a single pair of experiments, the FUMI theory can also be applied. However, Step 1 should be skipped and the test begins at Step 2. Moreover, the critical values should be changed: $s = \sqrt{2} \sigma$ in Step 2; $\text{LOD} = 3 \sigma$ in Step 3.

Results and Discussion

Precision of microarray measurement

Figure 2A shows the precision plot for human HepG2 using U95A chips. The X axis is the average of 6 expression levels (measurements) for each gene (total 12559 genes). The Y axis denotes the SD values estimated statistically from the 6 measurements each. The SD estimates (●) are not randomly scattered, but seem to increase with increasing expression level. This trend of the precision plot is quite common to many instrumental analyses such as ultraviolet-visible absorption spectrometry, atomic absorption spectrometry and high performance liquid chromatography.^{15,23,25} The similar precision plots for microarrays were observed.^{3,4,6}

Our microarray experiments were repeated over a part of the entire process, ranging from the hybridization on the chips to the data processing which gives the measurements, X . The least squares fitting to the observed SD values in Fig. 2A can lead to the SD dependence on X (for details, see the legend of Fig. 2):

$$\sigma = \sqrt{0.009639 X^2 + 91897.8} \quad (6)$$

This is the *a priori* SD defined in the preceding section and is shown in Fig. 2A (—). As for Fig. 2A, exceptionally large SD estimates are spotted frequently at high expression levels and the region of the least squares fitting is limited as described in the figure legend to guarantee the goodness of fit. Li and Wong revealed outliers due to various reasons including image artifacts in oligonucleotide microarrays.¹⁰

The dots (●) of the precision plots in Figs. 2B–D are the SD estimates observed under the conditions different from Fig. 2A (i.e., different samples, chips; see the legend). However, the lines (—) in Figs. 2B–D are just the *a priori* SD, σ , drawn in Fig. 2A (Eq. (6)). From this fact, we can see that although the *a priori* SD is phenomenological without knowledge about the causality of errors appearing on X , the *a priori* SD can provide a general aspect in the microarray experiments conducted here. The results of the TPA experiments are analyzed below with Eq. (6).

The distribution of microarray measurements is shown in Fig. 3A. The data are collected from the genes which give almost the same averages, $X_D (= 1100)$, of measurements (for details, see the legend). Figure 3B illustrates the normal distribution with the SD obtained by substituting X_D for X in Eq. (6). Although there are slight differences between the observed and normal distributions, especially around the center and on the edges, the measurements can be considered normally distributed. This normality makes the SD scattering pattern in Fig. 2 interpretable in terms of the chi-square distribution. A log-normal distribution was observed for the bulk of Affymetrix microarray spot intensities.³⁵

Underlying the *t*-test is the gene-specific SD, or rather changeable SD from gene to gene. However, the intensity-dependent SD of the FUMI theory (Eq. (6)) is not surprising in the field of instrumental analysis. A photomultiplier can know the intensity of light, but can never know anything else, e.g., the origin of light is a human gene or rat gene.

Error sources and evaluation of *a priori* SD

The error sources of the microarray measurements have yet to be identified. The constant term of Eq. (6) ($= 91897.8$)

Fig. 2 Precision plots for microarray measurements in different conditions. Six DNA chips are used for each condition. The average and SD estimate of 6 measurements for a gene leads to the values of X and Y axes, respectively. ●, the SD estimates from six replicates; —, the fitted line (*a priori* SD, Eq. (6)). Conditions (samples, chips): A, human HepG2, U95A; B, human HL60, U95A; C, human HepG2, U95B; D, rat microglia, U74A. The intercept (= 91897.1) of Eq. (6) is the average of the variance estimates over X from 1000 to 2000. The coefficient (= 0.009639) is obtained from the least squares fitting of a straight line passing the origin to the average-subtracted variance estimates over X from 1000 to 100000.

Fig. 3 Distribution of expression levels (top) and normal distribution (bottom). The top figure uses 498 genes which are located in the X region from 1000 to 1200. The total number of measurements used is 2988 (= 6 × 498).

dominates at low fluorescence intensities and will correspond to the background noise which comes mainly from the photomultiplier of the detection unit. The coefficient of X^2 (= 0.009639) plays an important role at high intensities where the RSD of measurements is almost invariant (~10% here). Promising candidates of error sources are some procedures before the light detection such as the incorporation of fluorescent tags and hybridization.

The model experiments, using no microarrays (see “Materials and Methods”), include the former part of typical entire analysis (preparation of total RNA and synthesis of cRNA), but the procedure corresponding to detection is quite simple (ultraviolet-visible absorption). The RSD for the model experiments was observed to be about 5%. This result implies that the experimental error originates mainly from the total RNA preparation and cRNA synthesis, since the precision of the UV detection is usually high (RSD < 1%).

Our microarray experiments lack the former part of the entire analysis (total RNA preparation and cRNA synthesis). However, even if they included it, the error of the entire analysis ($RSD = (10^2 + 5^2)^{1/2} = 11.2$) would be almost equal to the error of the experiments without the former part (RSD ~10%, see above).

From the above discussion, it follows that in our microarray experiments, the most important error sources are the background noise at low fluorescence intensities and the incorporation of fluorescent tags and hybridization at high intensities. Since the former part of analysis does not affect the precision substantially, the *a priori* SD (Eq. (6)) can be considered to be applicable to the usual microarray experiments including the sample preparation.

t-test and FUMI theory for duplicate pairs of TPA experiments

If the sample size is large, the *t*-test and FUMI theory would

Fig. 4 Correlation between FUMI theory and t -test in TPA experiments. X axis, t value of FUMI theory (Eq. (4)); Y axis, t value of t -test (Eq. (2)). The original data of the test results ($\square$) marked with letters A – D are shown in Fig. 5.

yield almost indistinguishable results. With the duplicate pairs of experiments, however, the power of the methods is limited and some differences are found between them. The FUMI theory may occasionally implicate genes as being differentially expressed, even if the t -test rejects the significant difference, and will not always implicate genes having large t values.

The F -test at 1% probability level, carried out before the t -test, denies the homoscedasticity for 309 genes out of 12559 in the TPA experiments. The Cochran-Cox method suggests that either of 309 genes shows no significant differences in the mean expression levels between the TPA exposure, $\bar{X}_E$, and control, $\bar{X}_C$. Therefore, the 309 genes are eliminated from the subsequent analysis.

The t -test at 1% level is conducted for the remaining 12250 genes. Consequently, the expression levels between the two conditions are significantly different for 205 genes. Without the F -test described above, the t -test alone selected additional 7 genes and the total sum was 212 genes.

The FUMI theory, starting with the original data (12559 genes), demonstrates that 391 genes show significant differences at 1% probability level. The genes selected by the FUMI theory are almost twice in number those by the t -test.

Figure 4 shows the correlation between the t values, $|\bar{X}_E - \bar{X}_C|/s$, of the t -test and FUMI theory (see Eqs. (2) and (4)). In the bottom figure, the horizontal line ($Y = 9.925$) and vertical line ($X = 2.58$) represent the critical values at 1% level of the t -test and FUMI theory, respectively. The number of genes

Table 1 Significance tests for control and TPA exposure ($n = 2$)

t -test	FUMI theory		
	99% significant	Not significant	Total
99% significant	56	149	205
Not significant	335	12019	12354
Total	391	12168	12559

Table 2 FUMI theory for control and TPA exposure ($n = 1$)

	Ctrl-1, TPA-1	Ctrl-1, TPA-2	Ctrl-2, TPA-1	Ctrl-1, TPA-2
99%	204	294	306	309

included in the four regions separated by the two lines is listed in Table 1. The top figure shows that the t values spread more widely for the t -test than for the FUMI theory: 50 times the critical value for the t -test and 7 times that for the FUMI theory.

Figure 5 stresses the similarities and differences between the t -test and FUMI theory at 1% level. The bar charts (A – D) show the original data of the genes indicated by the squares (A – D) in Fig. 4. The error bars on the Ctrl-1 and Ctrl-2 bars denote the 99% confidence intervals of the t -test ($= 9.925s$) and FUMI theory ($= 2.58s$), respectively.

In Fig. 5B, the expression levels are decided to be significantly different by both the methods. This judgment can also be made by the visual inspection of the error bars: if the duplicate measurements are both higher than the error bars on the control measurements, the difference can be significant.

Figure 5A presents the situation where the significant difference is indicated by the t -test, but not by the FUMI theory. The accidental coincidence of the measurements causes the t -test to estimate such a small SD, s , that the 99% confidence interval ($= 9.925s$) becomes much narrower than the observed difference, $|\bar{X}_E - \bar{X}_C|$. On the other hand, the FUMI theory refers to the *a priori* SD (Eq. (6)) through the average of all the four measurements and the predicted 99% confidence interval ($= 2.58s$) is wider than the real difference, $|\bar{X}_E - \bar{X}_C|$.

The above decision is reversed in Fig. 5D. For this example, the discrepancy between the methods springs from the large scattering of the duplicate measurements (see the bars of TPA-1 and TPA-2). In examples A and D, the judgment inconsistency is attributable to accidental events.

Even if the *a priori* SD is small enough to affirm a significant difference, the FUMI theory can reject it. This situation arises at which the scattering of the duplicate measurements is even larger than that expected (see Fig. 5C). The decision by the t -test is also reasonable due to the large scattering.

FUMI theory for a single pair of TPA experiments

We examine the validity of the FUMI theory with each possible combination of the results of Table 1. Table 2 lists the number of expressed genes chosen by the FUMI theory. About 200 – 310 expressed genes are selected, but they are less than those selected from the duplicate pairs of experiments (391 genes in Table 1). The similar tendency was observed by Ideker *et al.*⁶

The expressed genes commonly included in all the combinations in Table 2 are 60 in number. About 20 – 30% of the expressed genes out of 200 – 310 genes are always selected irrespective of experiments. This fact implies the high

Fig. 5 Individual expression levels judged by *t*-test and FUMI theory differently (A and D) and equally (B and C) in TPA experiments. Figures A–D correspond to the data ($\square$) with the same letters in Fig. 4. Bars (left to right): measurement of control 1; measurement of control 2; measurement of TPA 1; measurement of TPA 2. Error bars: 9.925s on Ctrl-1 bar (see Eq. (2)); 2.58s on Ctrl-2 bar (see Eq. (4)). Judgment (*t*-test, FUMI theory; s means significant and ns not significant): A: s, ns; B: s, s; C: ns, ns; D: ns, s.

reproducibility of not only the microarray measurement carried out here but also the FUMI theory.

The product set of the 60 commonly expressed genes from the single pair of experiments and 391 expressed genes from the duplicate experiments contains 58 genes. Almost all the genes common to the single experiment are also selected by the more reliable test for the duplicate experiments.

Table 3 lists the number of false positive expressions judged by the FUMI theory. As is expected from the critical level (1%), almost 1% of the entire genes (= 12559) is falsely selected as positive. This result corroborates the normality assumption of the distribution shown in Fig. 3.

Tables 2 and 3 lead to the conclusion, though slightly rough, that among the expressed genes judged by the FUMI theory, a quarter will be selected again from another experimental set, a quarter will be variant and a half will be false positive.

In an analysis proposed by Lee *et al.*,¹ false positives exceedingly dominated over false negatives in experiments using cDNA microarrays and 9% of 288 genes were classified as false positives. The performance of their model is dissimilar from that of the FUMI theory, but the straightforward comparison will make no sense because of different settings of experiments and the idiosyncracies of the classification rules.

Conclusion

The FUMI theory for microarray experiments can handle the data on a single pair of treatment and control experiments, while retaining the statistical characteristics of the *t*-test which can only cope with the data on plural pairs of experiments. Being straightforward, the FUMI theory can be adapted to a broad range of experimental settings. The inevitable requisite is the uncertainty model (*i.e.*, the *a priori* SD) which is intrinsic to an instrument employed in a laboratory.

Table 3 False positive expressions judged by FUMI theory ($n = 1$)

	Ctrl-1, Ctrl-2	TPA-1, TPA-2
99%	123	134

The correctness of the FUMI theory has not been verified completely in this paper. However, as far as a single pair of our TPA experiments is concerned, a quarter of expressed genes selected by the FUMI theory is selected again for another experiments under the same conditions. Furthermore, almost all of these common genes are also chosen by the more reliable test for duplicate experiments. These experimental facts are strong, though circumstantial, pieces of evidence for the FUMI theory.

Some biologically relevant genes are known to be regulated on treatment with TPA in the extensive literature.^{29,30} Among them, Erg-1, Erg-2, JUNB, TNFAIP1, c-myc, c-myc, *etc.* have been found by the FUMI theory. The biological evaluation of the FUMI theory will be performed in a subsequent paper.

Acknowledgements

This work was supported in part by the Program for Promotion of Fundamental Studies in Health Sciences (MPJ-6 and MF-16) of the Organization for Pharmaceutical Safety and Research.

References

1. M.-L. T. Lee, F. C. Kuo, G. A. Whitmore, and J. Sklar, *Proc. Natl. Acad. Sci. USA*, **2000**, *97*, 9834.
2. W. Pan, *Bioinformatics*, **2000**, *18*, 546.
3. W. Pan, J. Lin, and C. T. Le, *Genome Biology*, **2002**, *3*.
4. D. M. Rocke and B. Durbin, *J. Comput. Biol.*, **2001**, *8*, 557.
5. M. A. Newton, C. M. Kendzierski, C. S. Richmond, F. R. Blatiner, and K. W. Tsui, *J. Comput. Biol.*, **2001**, *8*, 37.
6. T. Ideker, V. Thorsson, S. F. Siegel, and L. E. Hood, *J. Comput. Biol.*, **2000**, *7*, 805.
7. P. Baldi and A. D. Long, *Bioinformatics*, **2001**, *17*, 509.
8. V. G. Tusher, R. Tibshirani, and G. Chu, *Proc. Natl. Acad. Sci. USA*, **2001**, *98*, 5116.
9. R. D. Wolfinger, G. Gibson, E. D. Wolfinger, L. Bennett, H. Hamadeh, P. Bushel, C. Afshari, and R. S. Paules, *J. Comput. Biol.*, **2001**, *8*, 625.
10. C. Li and W. H. Wong, *Proc. Natl. Acad. Sci. USA*, **2001**, *98*, 31.
11. L. Wernisch, S. L. Kendall, S. Soneji, A. Wietzorrek, T. Parish, J. Hinds, P. D. Butcher, and N. G. Stoker, *Bioinformatics*, **2003**, *19*, 53.
12. J. F. K. Huber, J. A. R. J. Hulsman, and C. A. M. Meijers, *J. Chromatogr.*, **1971**, *62*, 79.
13. H. Barth, E. Dallmeier, G. Courtois, H. E. Keller, and B. L. Karger, *J. Chromatogr.*, **1973**, *83*, 289.
14. E. Grushka and I. Zamir, *Chem. Anal.*, **1989**, *98*, 529.
15. J. D. Ingle, Jr. and S. R. Crouch, "Spectrochemical analysis", **1988**, Prentice-Hall, Inc., New Jersey.
16. N. W. Bower and J. D. Ingle, Jr., *Anal. Chem.*, **1976**, *48*, 686.
17. N. W. Bower and J. D. Ingle, Jr., *Anal. Chem.*, **1977**, *49*, 574.
18. C. Th. J. Alkemade, W. Snelleman, G. D. Boutilier, B. D. Pollard, J. D. Winefordner, T. L. Chester, and N. Omenetto, *Spectrochim. Acta*, **1978**, *33B*, 383.
19. G. D. Boutilier, B. D. Pollard, J. D. Winefordner, T. L.

- Chester, and N. Omenetto, *Spectrochim. Acta*, **1978**, 33B, 401.
20. C. Th. J. Alkemade, W. Snelleman, G. D. Boutilier, and J. D. Winefordner, *Spectrochim. Acta*, **1980**, 35B, 261.
21. P. W. J. M. Boumans, *Anal. Chem.*, **1994**, 66, 459A.
22. P. W. J. M. Boumans, *Spectrochim. Acta*, **1981**, 36B, 1031.
23. Y. Hayashi and R. Matsuda, *Anal. Chem.*, **1994**, 66, 2874.
24. Y. Hayashi, R. Matsuda, and R. B. Poe, *Chromatographia*, **1995**, 41, 66.
25. R. Matsuda, Y. Hayashi, S. Sasaki, K. Saito, K. Iwaki, H. Harakawa, M. Satoh, Y. Ishizuki, and T. Kato, *Anal. Chem.*, **1998**, 70, 319.
26. E. D. Prudnikov, J. W. Elgersma, and H. C. Smit, *J. Anal. At. Spectrom.*, **1994**, 9, 619.
27. H. C. Smit and H. L. Walg, *Chromatographia*, **1976**, 9, 483.
28. Y. Hayashi, R. Matsuda, and R. B. Poe, *Analyst*, **1995**, 121, 591.
29. H. Q. Nguyen, B. Hoffman-Liebermann, and D. A. Liebermann, *Cell*, **1993**, 72, 197.
30. P. Tamayo, D. Slonim, J. Mesirov, Q. Zhu, S. Kitareewan, E. Dmitrovsky, E. Lander, and T. R. Golub, *Proc. Natl. Acad. Sci. USA*, **1999**, 96, 2907.
31. K. Nakajima, M. Hamanoue, M. Shimojo, N. Takei, and S. Kohsaka, *Biomed. Res.*, **1989**, 10, 411.
32. J. C. Miller and J. N. Miller, "Statistics for analytical chemistry", **1988**, Ellis Horwood, West Sussex.
33. A. Sakuma, "Bioassay-Design and analysis", **1964**, University of Tokyo.
34. G. L. Long and J. D. Winefordner, *Anal. Chem.*, **1983**, 55, 712A.
35. D. C. Hoyle, M. Rattray, R. Jupp, and A. Brass, *Bioinformatics*, **2002**, 18, 576.
-