Curr Genet (2005) 47: 121–131
DOI 10.1007/s00294-004-0551-3

RESEARCH ARTICLE

Yan Lu · Jun Zhu · Pengyuan Liu

A two-step strategy for detecting differential gene expression in cDNA microarray data

Received: 4 September 2004 / Revised: 27 October 2004 / Accepted: 27 October 2004 / Published online: 10 December 2004

Abstract A mixed-model approach is proposed for identifying differential gene expression in cDNA microarray experiments. This approach is implemented by two interconnected steps. In the first step, we choose a subset of genes that are potentially expressed differentially among treatments with a loose criterion. In the second step, these potential genes are used for further analyses and data-mining with a stringent criterion, in which differentially expressed genes (DEGs) are confirmed and some quantities of interest (such as gene × treatment interaction) are estimated. By simulating datasets with DEGs, we compare our statistical method with a widely used method, the *t*-statistic, for single genes. Simulation results show that our approach produces a high power and a low false discovery rate for DEG identification. We also investigate the impacts of various source variations resulting from microarray experiments on the efficiency of DEG identification. Analysis of a published experiment studying unstable transcripts in *Arabidopsis* illustrates the utility of our method. Our method identifies more novel and biologically interesting unstable transcripts than those reported in the original literature.

Keywords cDNA microarray · Gene expression · Mixed-model approach · *Arabidopsis*

Introduction

Recent developments in microarray technology make it possible to rapidly capture all of the gene expression profiles in biological samples (Ross et al. 2000; Welsh et al. 2001; Bouton and Pevsner 2002; Guffanti et al. 2002). This technology results in large amounts of data, the interpretation of which is a major bottleneck in current studies. A natural step in extracting microarray data information is to examine the extremes, for example, genes with significant differential expression in two samples (case vs control) or in a time-series (such as cell cycles).

Microarray data are characterized by high dimensionality (thousands of genes) and small sample size (often <30). Systematic and stochastic fluctuations are usually involved in microarray experiments (Schuchhardt et al. 2000). Therefore, the raw dye intensity or ratio value has a high noise to signal ratio between probes. The *x*-fold change approach may induce high false positives and/or false negatives when used as a simple criterion to determine the genes differentially expressed between query and reference samples. Some biologically important genes with small *x*-fold changes are highly statistically significant when they are measured repetitively with high precision. Conversely, many genes with large *x*-fold changes in one array and high variability across multiple arrays have no statistical significance (Wolfinger et al. 2001). Various statistical methods have been proposed for identifying differentially expressed genes (DEGs; Chen et al. 1997, 2002; Ideker et al. 2000; Kerr et al. 2000; Newton et al. 2001; Thomas et al. 2001; Wolfinger et al. 2001; Efron et al. 2001; Churchill 2002; Ibrahim et al. 2002; West 2003; Smyth 2004), but none has yet gained widespread acceptance for the analysis of microarray data. The most basic statistical problem is that the measured differential expression cannot completely reflect a real biological shift in gene expression (Newton et al. 2001).

Discrimination and cluster analysis techniques have been very useful for searching patterns of gene expression that are highly correlated (Eisen et al. 1998; Spellman et al. 1998; Golub et al. 1999; Tamayo et al. 1999; Hastie et al. 2000). These methods are involved in using various types of clustering algorithms, such as self-organizing maps, *k*-means clustering and hierarchical clustering, to

Communicated by S. Hohmann

Y. Lu · J. Zhu (✉) · P. Liu
Institute of Bioinformatics, Zhejiang University,
Hangzhou, 310029, Peoples Republic of China
E-mail: jzhu@zju.edu.cn
Tel.: +86-571-86971444
Fax: +86-571-86049815

discriminate and characterize patterns of gene expression. However, such exploratory methods alone do not provide the opportunity to engage in statistical inference. Furthermore, the gene expression level or relative ratio level with sampling errors within experiments is performed directly in discrimination and cluster analyses; and thus the distance between data-points cannot reflect the true differential expression between genes.

Mixed-model approaches are widely used to partition various sources of variability. They have the flexibility to handle unbalanced data and can be easily extended to more complicated biological models which have been proven as powerful statistical tools in classic quantitative genetic analyses (Searle et al. 1992). The objectives of this paper are: (1) to propose a mixed-model approach to analyzing variance components for cDNA microarray data analysis, applying the method to selecting a target subset of DEGs that are of biological interest and (2) to assess the effectiveness of this method by extensive computer simulations, specifically compared with the widely used approach based on *t*-statistics for single genes (Dudoit et al. 2000). Analyzing data publicly available for the study of unstable transcripts in *Arabidopsis* demonstrates the utility of our method.

Materials ad methods

Each datum in a microarray experiment is associated with one particular combination of an array in the experiment: a fluorescence dye (red or green), a treatment and a gene. In our analysis, we used the logarithms of the original fluorescence measurements as phenotypic values, not the log ratio values, as used by some previous studies (Kerr et al. 2000; Wolfinger et al. 2001).

To alleviate the computation burden, we propose a two-step strategy for analyzing microarray data. In the first step, we choose a subset of genes that are potentially expressed differentially among treatments with a loose criterion. In the second step, these potential genes are combined for further analyses and data-mining with a stringent criterion, in which DEGs are confirmed and some quantities of interest (such as gene × treatment interaction) are estimated. Both types of the aforementioned analyses are performed using a mixed-model approach for a variance–component framework.

Choosing a subset of potential genes with differential expression

We first normalized the original fluorescence data before choosing a subset of genes. The purpose of normalization is to minimize systematic experimental biases so that the observed variation arises from biological differences. Let y_{ijkl} denote the logarithm of a measurement from the ith array, the jth treatment, the kth dye and the lth gene in a cDNA microarray experiment. The original fluorescence data are normalized as: $r_{ijkl} = y_{ijkl} - \left(\bar{y}_{i} \ldots + \bar{y}_{.j..} + \bar{y}_{..k.} - 2\bar{y}_{....}\right)$.

The normalized data, r_{ijkl}, can be viewed as a variation for each gene after removing systematic experimental errors and are the input data for the following single-gene model:

$$r_{ijkl} = \mu_l + A_{il} + T_{jl} + D_{kl} + \gamma_{ijkl} \quad (1)$$

Here, μ_l represents the overall average expression level of gene l (a fixed effect), A_{il} is the ith array effect of gene l (a random effect): $A_{il} \sim \left(0,\ \sigma^2_{A(l)}\right)$ T_{jl} is the jth treatment effect of gene l (a random effect): $T_{jl} \sim \left(0,\ \sigma^2_{T(l)}\right)$; D_{kl} is the kth dye effect of gene l (a random effect): $D_{kl} \sim \left(0,\ \sigma^2_{\gamma(l)}\right)$; γ_{ijkl} is the residual error of gene l: $\gamma_{ijkl} \sim \left(0,\ \sigma^2_{\gamma(l)}\right)$. The array effects account for differences among arrays. Differences among arrays may arise from differences in print quality or from differences in the ambient conditions when the plates were processed, which may increase or reduce the hybridization efficiencies of labeled cDNA. The treatment effects account for differences among treatments. Such differences can arise when some treatments (e.g., a specific cell line) have more transcription activity in general than others. The dye effects account for fluorescent signal differences. One dye may show consistently higher signal intensity than another. The single-gene model is fitted separately to the normalized data from each gene, allowing an elementary inference to be made, using a separate estimate of variability. The methods described here are for the prejudication of a subset of genes with differential expression. This procedure is similar to a variation filter that is commonly used to exclude genes with less than a certain x-fold variation among the collected samples (Golub et al. 1999). However, the x-fold variation filter is usually based on total gene expression variations. Instead, our procedure focuses on total treatment effects, which may increase the filter efficiency.

Combining analysis of multiple genes

A subset of genes potentially expressed differentially between one or more pairs of samples in the dataset can be used for further analysis as follows:

$$y_{ijkl} = \mu + G_l + A_i + T_j + D_k + GA_{li} + GT_{lj} + GD_{lk} + \varepsilon_{ijkl} \quad (2)$$

where μ is the average of overall expression levels (a fixed effect), G_l is the fixed effect of the lth gene, $A_i \sim \left(0,\ \sigma^2_A\right)$ is the random effect of the ith array, $T_j \sim \left(0,\ \sigma^2_T\right)$ is the random effect of the jth treatment and $D_k \sim \left(0,\ \sigma^2_D\right)$ is the random effect of the kth dye. $GA_{li} \sim \left(0, \sigma^2_{GA}\right)$ is the interaction between the lth gene and the ith array, $GT_{lj} \sim \left(0, \sigma^2_{GT}\right)$ is the interaction between the lth gene and the jth treatment and $GD_{lk} \sim \left(0, \sigma^2_{GD}\right)$ is the interaction between gene l and dye k. The random error

term ε_{ijkl} is the residual effect: $\varepsilon_{ijkl} \sim (0, \sigma^2_\varepsilon)$. Interpretations of A_i, T_j and D_k are similar to those in Eq. 1. The gene effects, G_l, account for differences in transcription level among the genes. Some genes may be inherently more active in mRNA transcription than others. The gene × array interactions, GA_{li}, account for the average effect of the spot on the ith array for the lth gene. It is a "spot" effect due to the potential incomplete control over the amount and concentration of cDNA immobilized from one array to the next. The gene × dye interactions, GD_{lk}, are gene-specific dye effects and account for the average effect of the kth fluorescence dye for the lth gene. This may contribute to the differential hybridization efficiencies of two chemically different fluorescence dyes for the same probe. The gene × treatment interactions, GT_{lj}, are of interest in microarray experiments. These effects capture the departure from the overall averages that are attributable to the specific combination of the jth treatment and the lth gene.

Similar interpretations of the aforementioned factors were also detailed by Kerr et al. (2000). Whether a specific factor is regarded as fixed or random depends not only on the levels of source variation but also on the investigator's particular interest in the study. A fixed effect is one that is repeatable. That is, if other researchers repeat a specific microarray experiment, they are estimating the same effects. A random effect is one that is not repeatable. That is, another researcher will not (probably cannot) estimate the same effects, but can estimate the variance of the effects from another sample. In our study, we treated gene effects as fixed, while others were treated as random. For example, the print quality of the arrays and the ambient conditions under which the arrays were probed varied from one microarray experiment to another. Such array effects may not be repeatable among different microarray experiments and thus are treated as random effects. The basic mRNA transcription level for a specific gene may remain inherently similar among different microarray experiments when there are no interference factors such as those from arrays and treatments. Such a basic transcription level is estimable with suitable experimental designs. Therefore, the gene effects are treated as fixed effects in our model.

Statistical assessment of gene significance

Both types of the above models can be analyzed by a mixed-model approach. The single-gene model (Eq. 1) can be rewritten in the following matrix form:

$$\begin{aligned}\mathbf{r}_{(l)} &= \mathbf{1}\mu_{(l)} + \mathbf{U}_{A(l)}\mathbf{e}_{A(l)} + \mathbf{U}_{T(l)}\mathbf{e}_{T(l)} + \mathbf{U}_{D(l)}\mathbf{e}_{D(l)} + \mathbf{e}_{\varepsilon(l)} \\ &= \mathbf{1}\mu_{(l)} + \sum_{u=1}^{4}\mathbf{U}_{u(l)}\mathbf{e}_{u(l)} \sim N\left(\mu_{(l)}, \mathbf{V}_{(l)}\right)\end{aligned} \tag{3}$$

with this variance–covariance matrix:

$$\begin{aligned}\mathrm{Var}\left(\mathbf{r}_{(l)}\right) &= \mathbf{V}_{(l)} \\ &= \sigma^2_{A(l)}\mathbf{U}_{A(l)}\mathbf{U}^{\mathrm{T}}_{A(l)} + \sigma^2_{T(l)}\mathbf{U}_{T(l)}\mathbf{U}^{\mathrm{T}}_{T(l)} \\ &\quad + \sigma^2_{D(l)}\mathbf{U}_{D(l)}\mathbf{U}^{\mathrm{T}}_{D(l)} + \sigma^2_{\varepsilon(l)}\mathbf{I}\end{aligned}$$

where $\mu_{(l)}$ is the population mean over all entries of gene l, $\mathbf{e}_{u(l)}$ is the vector of random effects: $\mathbf{e}_{u(l)} \sim (0, \sigma^2_{u(l)}\mathbf{I})$; $\mathbf{U}_{u(l)}$ is the known incidence matrix relating to the random vector $\mathbf{e}_{u(l)}$, $U^{\mathrm{T}}_{u(l)}$ is the transposition of $\mathbf{U}_{u(l)}$; $\mathbf{U}_{4(l)} = \mathbf{I}$ is an identity matrix. Similarly, the multi-gene model (Eq. 2) can also be expressed as the matrix form.

Variance components of the aforementioned models can be estimated using maximum likelihood estimation (ML), restricted maximum likelihood estimation (REML), and minimum norm quadratic unbiased estimation (MINQUE; Searle et al. 1992). Among these three methods, MINQUE possesses the advantages of unbiasedness, no assumption of normal distribution and less computation (Zhu and Weir 1994a). The prediction of random effects can be obtained using methods for best linear unbiased prediction (BLUP; Henderson 1963), linear unbiased prediction (LUP; Zhu and Weir 1994a) and adjusted linear unbiased prediction (AUP; Zhu 1993; Zhu and Weir 1996). The fixed effects can be obtained through the ordinary least square estimation (OLSE) method or the generalized least square estimation (GLSE) method. The Jackknife resampling procedure (Miller 1974; Searle et al. 1992) can be used for estimating the sampling variance of estimated variance components, predicted random effects and estimated fixed effects; and a t-test is then used for the significance test.

Microarray data are characterized by high dimensionality and small sample size, which may not warrant normal distribution of the data and usually requires intensive computation for ML or REML estimators. From this reason, MINQUE(1), an unbiased MINQUE method with all the prior values set at one (Zhu and Weir 1996), was used to estimate the variance components and the Jackknife resampling procedure was used for significance tests in our method. The AUP and OLSE methods were used for predicting random effects and estimating fixed effects, respectively.

In the single-gene model, a series of hypotheses can be made about the variance of treatment: H_0:$\sigma^2{}_{T(l)} = 0$ vs H_1:$\sigma^2{}_{T(l)} = 0$. If H_0 in the null hypothesis about gene l is rejected, the observation of this gene is retained for further analysis in the multi-gene model. In the subsequent multi-gene model, a t-test following the Jackknife resampling procedure is applied to test the null hypothesis of a specific gene without differential expression, that is, the gene × treatment interaction effect (i.e., e_{GT}) is not significantly different from zero. However, if at least one of the e_{GT} of gene l is not equal to zero, the gene l is considered a DEG. This resample-based t-test in the multi-gene model can capture the departure from the overall average that is attributable to the specific combination of the jth treatment and the lth gene.

Simulation design

A series of simulations for cDNA microarray experiments was conducted to evaluate the performance of the proposed approach. The loop design was adopted in our simulated experiments. The loop design involves constructing a cyclic sequence of n treatments on n arrays, with each treatment represented twice, each time labeled with a different fluorescence dye (Kerr and Churchill 2001). In all the simulations conducted, there were 4,000 genes and six treatments. The six treatments were divided into two groups of three each. Treatments T_1–T_3 were in one group and treatments T_4–T_6 were in another. For the first group, in the first array treatment T_1 was marked with Cy3 dye and treatment T_2 was marked with Cy5 dye, in the second array treatment T_2 was marked with Cy3 dye and treatment T_3 was marked with Cy5 dye and in the third array treatment T_3 was marked with Cy3 dye and treatment T_1 was marked with Cy5 dye. Note that in spotted cDNA microarrays the two treatments under comparison are labeled with two different dyes and co-hybridized to the same array. The design was similar for another group with treatments T_4–T_6 (Table 1). Each of them was replicated three times, giving 18 arrays in total.

Generating gene-expression data

To generate each dataset, we preset different magnitudes of source variations (i.e., variance components) in the simulated microarray experiments. The gene × treatment interaction variance was set as 50 and the ratio of the gene × treatment interaction variance (V_{GT}) to the total phenotypic variance (V_P), that is, V_{GT}/V_P, varied from 0.1 to 0.9 in all of the simulations. Four configurations of the remaining variance components (V_A, V_D, V_T, V_{GA}, V_{GD}, V_ε) were simulated for the remainder of the phenotypic variation (i.e., V_P–V_{GT}): (1) the effects of A, D, T, GA, GD and ε contribute equally to the remainder of phenotypic variation, that is, $V_A{:}V_D{:}V_T{:}V_{GA}{:}V_{GD}{:}V_\varepsilon = 1{:}1{:}1{:}1{:}1{:}1$ (denoting EQUAL), (2) the A and GA effects dominate in the remainder of phenotypic variation, that is, $(V_A + V_{GA})/(V_P - V_{GT}) = 0.9$ and $V_D{:}V_T{:}V_{GD}{:}V_\varepsilon = 1{:}1{:}1{:}1$ (denoting ARRAYDOM), (3) the D and GD effects dominate in the remainder of phenotypic variance, that is, $(V_D + V_{GD})/(V_P - V_{GT}) = 0.9$ and $V_A{:}V_T{:}V_{GA}{:}V_\varepsilon = 1{:}1{:}1{:}1$ (denoting DYEDOM) and (4) the T effects dominate the remainder of phenotypic variation, that is, $V_T/(V_P - V_{GT}) = 0.9$ and $V_A{:}V_D{:}V_{GA}{:}V_{GD}{:}V_\varepsilon = 1{:}1{:}1{:}1{:}1$ (denoting TREATDOM). Note that the efficiency of identifying DEGs is dependent on the relative proportions among different source variations rather than on the absolute magnitude of each of them. We assumed that there were only 40 DEGs among a total of 4,000 genes tested in the experiment (representing 1% of total genes), that is, 40 genes had gene × treatment interaction effects. The gene-expression value was obtained by the multi-gene model (Eq. 2) and the random effects in the model were drawn by generating a pseudo-random normal deviate with zero mean and different known variances.

Table 1 Experimental design of simulations

Array	Dye	
	D_1 (Cy3)	D_2 (Cy5)
Group 1		
A_1	T_1	T_2
A_2	T_2	T_3
A_3	T_3	T_1
Group 2		
A_4	T_4	T_5
A_5	T_5	T_6
A_6	T_6	T_4

Efficiency of identifying differentially expressed genes

We compared the proposed method with the conventional two-sample t-test method (Dudoit et al. 2000). For the t-test method, simulations were performed with and without x-fold filter. In the former case, we first excluded those genes with maximum x-fold changes of less than two among different treatments and then performed the t-test method on the remaining dataset. In the latter case, we performed the t-test method directly on the whole dataset. Power, false discovery rate and false number were used to evaluate the efficiency of these methods for identifying DEGs. Power refers to the probability of declaring a statistical significance when a true DEG exists. False discovery rate is the proportion that genes declared to be differentially expressed which are not differentially expressed in reality. False number is the total number of false positives (genes declared to be differentially expressed which in reality are not) and false negatives (genes truly differentially expressed but not declared as such). Global significant level was set at 0.05; and multiple testing was adjusted by Bonferroni's correction in both the mixed-model and the t-test methods.

Efficiency of predicting random effects and estimating fixed effects

We then evaluated the efficiency of predicting random effects and estimating fixed effects with our models, using the proportion of bias, $(\hat{\theta} - \theta)/|\theta|$, where θ is the true effect value and $\hat{\theta}$ is the mean of the predicted random effect or estimated fixed effect.

Results

Monte Carlo simulations were run 200 times for each case and the mean results of the 200 simulations are presented below.

Identifying DEGs

We first evaluated the performance of the mixed-model approach and *t*-test methods under different source variations resulting from microarray experiments. Powers and false discovery rates are summarized in Fig. 1 and false numbers are summarized in Fig. 2. There is a general tendency: the larger *GT* interactions account for the gene differential expression, higher power and lower false discovery rate; and fewer false numbers are achieved by each of these methods. Their efficiencies in identifying DEGs are apparently dependent on various source variations in the microarray experiments. In addition, the *t*-test method with the filtration procedure worked a little better than that without the filtration procedure in most cases, but the difference was quite small. For a simpler and clearer presentation of the results, in the following comparisons we applied the *t*-test methods to both of the above two methods, that is, *t*-test methods with and without the filtration procedure.

When the variances of *A*, *D*, *T*, *GA*, *GD* and ε are of similar magnitude (EQUAL), our method achieved consistently higher powers and lower false discovery rates than the *t*-test method. When the *A* and *GA* effects dominated in the remainder of the phenotypic variance (ARRAYDOM), our method produced dramatically higher powers and lower or similar false discovery rates than the *t*-test method. When the *D* and *GD* effects dominated in the remainder of phenotypic variance (DYEDOM), our method still gave dramatically higher powers than the *t*-test method. The false discovery rates of our method were slightly higher than the *t*-test method when the V_{GT}/V_P exceeded 0.3. When the *T* effects accounted for a majority of the remainder of the phenotypic variance (TREATDOM), the *t*-test method showed a higher power than our method but at the cost of extremely higher false discovery rates. In all of the four cases studied, our method always produced fewer false numbers than the *t*-test method. In particular, in the case of TREATMENT, about 2,500–3,000 genes of the total 4,000 genes were false positives or false negatives by the *t*-test method, while only 4–40 genes were false positives or false negatives by our method. These results indicate that, in most cases, our approach has a higher efficiency of identifying DEGs, while the odds of falsely declaring DEGs are lower.

We then classified differential expression into three categories with regard to individual *GT* variance of a specific gene: genes with a large *GT* variance, genes with a medium *GT* variance and genes with a small *GT* variance. Powers of the mix-model and the *t*-test method for identifying each of the three groups of genes are shown in Table 2. All methods showed higher powers of identifying DEGs having a large *GT* variation. Specifically, those genes with *GT* variation $>3\%$ of the total *GT* variation of all genes were more frequently declared to be differentially expressed in our simulated experiments. When $V_{GT}/V_P=0.8$, the powers for identifying DEGs with a large *GT* variation were similar in these methods. The differences in statistical powers between these methods were due to their ability to identify genes with medium or small *GT* variation. When $V_{GT}/V_P=0.4$, neither method could efficiently identify the DEGs with a medium or small *GT* variation, but there were differences in statistical powers for identifying genes with large *GT* variation. In the simulated experiments, our method generally had high efficiency in identifying genes with medium to large *GT* variation in most cases when $V_{GT}/V_P>0.6$.

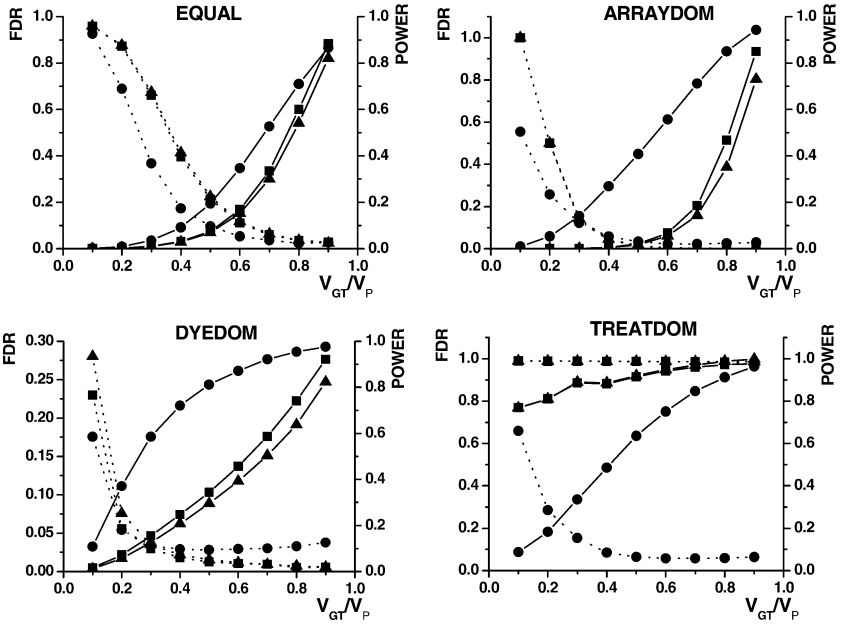

Fig. 1 False discovery rates (*FDR*) and powers of identifying DEGs using the mixed-model approach (*circles*) and the *t*-test method with (*squares*) and without (*triangles*) the filtration procedure. *Dotted lines* are false discovery rates and *solid lines* are powers

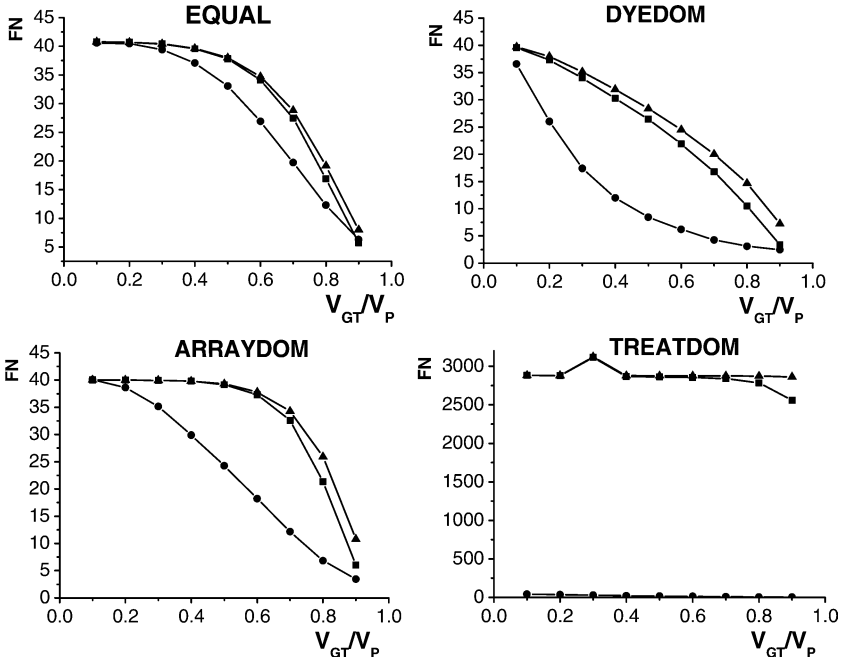

Fig. 2 False numbers (*FN*) when identifying DEGs by the mixed-model approach (*circles*) and the *t*-test method with (*squares*) and without (*triangles*) the filtration procedure

Table 2 Effects of individual *GT* variance on powers for identifying DEGs. *MM* Mixed-model approach, *t-testF* *t*-test method with filtration procedure, *t-test* *t*-test method without filtration procedure

V_{GT}/V_p	Variance component	>0.03[a]			0.01–0.03			<0.01		
		MM	*t*-testF	*t*-test	MM	*t*-testF	*t*-test	MM	*t*-testF	*t*-test
0.8	EQUAL	0.978	0.893	0.937	0.800	0.584	0.590	0.160	0.073	0.066
	ARRAYDOM	0.989	0.799	0.854	0.916	0.404	0.345	0.493	0.037	0.019
	DYEDOM	1.000	0.880	0.860	0.976	0.748	0.670	0.835	0.439	0.349
	TREATDOM	0.991	1.000	1.000	0.945	0.996	0.998	0.729	0.932	0.944
0.4	EQUAL	0.320	0.074	0.116	0.073	0.016	0.021	0.005	0.000	0.000
	ARRAYDOM	0.659	0.016	0.031	0.263	0.001	0.002	0.021	0.000	0.000
	DYEDOM	0.983	0.401	0.434	0.809	0.223	0.214	0.184	0.033	0.017
	TREATDOM	0.813	0.947	0.964	0.523	0.884	0.898	0.104	0.773	0.778

[a] The number indicates the proportion of individual variance of *GT* to total variance

Predicting random effects and estimating fixed effects

Table 3 shows the proportion of bias for *GT* effects predicted by the AUP method and for gene effects estimated by the OLSE method, respectively. For *GT* effects with large absolute sizes, the biases of their predictors were reasonably small (ca. 5%). However, for *GT* effects with small absolute sizes, the biases of their predictors were considerably larger. Similar results were also observed in the estimation of gene effects. These results suggest that our method can well predict *GT* effects with large absolute values, while prediction of *GT* effects with small absolute values should be treated with caution. This is also true for the estimation of gene effects.

Real example

We applied our method to analyze the publicly available datasets from the study of Gutiérrez et al. (2002), who examined mRNA degradation in intact *Arabidopsis thaliana* by cDNA microarrays containing 11,521 clones. In their study, three independent cordycepin treatments (biological replicas) were analyzed. Each pair of samples from 0 min and 120 min after cordycepin treatment was used in two microarray hybridizations, the second with reverse labeling relative to the first (technical replicas). Statistical analyses of the ratios were performed using the *t*-test. The data are available online at the Stanford microarray database (http://genome-www5.stanford.edu/; ExptID: 11374, 11333, 11339, 11323, 11375, 11342).

When using the *t*-test and the conservative Bonferroni method to adjust *P* values, 100 genes with unstable transcripts showed significantly different ratios from the mean of the population at $\alpha < 0.0001$ (see Gutierrez et al. 2002, supporting table 2). For a comparison of the results, the significance level of $\alpha = 0.0001$ was also adopted for single tests using the mixed-model approach. We found 90 genes with significant mRNA degradation from 0 min to 120 min, including 51 genes identified by both methods and 39 genes identified only by the mixed-model approach (Table 4).

Gutiérrez et al. (2002, Table 1) listed some *Arabidopsis* genes with unstable messages, including the

Table 3 Bias proportion of *GT* effects predicted by AUP and gene effects estimated by OLSE. *GT* effects and gene effects are divided into large, medium and small, according to their true absolute size

V_{GT}/V_p	Variance component	*GT* effect			Gene effect		
		Large	Medium	Small	Large	Medium	Small
0.8	EQUAL	0.058	0.140	0.650	−0.015	0.138	0.966
	ARRAYDOM	0.041	0.106	1.309	0.064	0.157	1.741
	DYEDOM	0.040	0.118	0.992	0.063	0.116	1.046
	TREATDOM	0.040	0.132	0.896	0.070	0.160	1.451
0.4	EQUAL	−0.058	0.182	0.687	0.026	0.183	−1.178
	ARRAYDOM	0.035	0.091	1.152	0.044	0.133	1.888
	DYEDOM	0.061	0.138	0.536	0.025	0.088	−0.645
	TREATDOM	0.065	0.166	−0.518	0.058	0.132	1.814

DNA-binding protein RAV1 gene at locus At1g13260 and the homeodomain transcription factor (ATHB-6) gene at locus At2g22430. AA395830 and N37328 are two expressed sequence tags (ESTs) from the gene at locus At1g13260; and H77088 and T04337 are two ESTs from the gene at locus At2g22430. They were all identified as unstable transcripts by our method, while only N37328 and T04337 were found by the *t*-test. AA720100, AA720105 and T76004 are all from the nucleotide sugar epimerase gene at locus At4g30440; and T20600, N65459 and T75944 are all from cytochrome P450 monooxygenase gene at locus At4g31500. The *t*-test only found that AA720100 and T20600 were unstable, whereas AA720105, T76004, N65459 and T75944 were identified as unstable genes by our method. T20543, AA720239 and AA720240 are three ESTs from the gene at locus At5g64260 which were identified as unstable genes by our method but not by the *t*-test. AA067525 and AA067498 are both from the calmodulin-related protein 2 gene at locus At5g37770, AA597715 and H36178 are both from the ethylene responsive element binding factor-like gene at locus At5g61590 and both AA597849 and T46143 are from the gene at locus At1g72450. Both of the methods identified one transcript from each of the three genes, respectively. However, the *t*-test did not find multiple transcripts from the same gene that were not found by the mixed-model approach. These EST identifications were searched in the *A. thaliana* annotation database and the *A. thaliana* gene index at the Institute for Genomic Research (http://www.tigr.org). Finding several unstable transcripts from the same gene is to be expected since the probes, coding for the same gene, should display very similar expression profiles (Liu et al. 2003). From this aspect, the mixed-model approach can identify more reasonable unstable transcripts.

In addition, polyA may play an important role in the translation of mRNA by increasing the stability of mRNA and allowing mRNA to function normally. Half-lives for histone mRNA that lacks a polyA tail were considerably lower than 30 min (Greenberg 1972). Two histone-related ESTs (H76940, AA720291) that were not identified as unstable genes by the *t*-test were found by our approach.

Discussion

Genome-wide identification of DEGs using conventional molecular techniques (e.g., Northern blot analysis) is expensive and time-consuming. Microarray technology represents one of the latest breakthroughs in experimental molecular biology which allows the monitoring of gene expression for tens of thousands of genes in parallel. It is already producing huge amounts of valuable data (Brazma and Vilo 2000). Many standard statistical methods have been used to mine such data. In the present study, we propose a method for microarray data analysis based on a mixed-model approach. As compared with the conventional *t*-test approach, our method tends to have a higher efficiency in identifying DEGs, while the odds of falsely declaring genes with differential expression are lower. Furthermore, some quantities of interests can be obtained by the AUP method for random effects or by the OLSE method for fixed effects. The method developed here has been implemented in the Windows-interface software QGA Station that is available at http://www.cab.zju.edu.cn/english/ics/faculty/zhujun.html.

Our method is an extension of recent groundwork by Kerr et al. (2000) and Wolfinger et al. (2001). The rationale underlying these methods is that total gene expression is partitioned into various source variations due to different factors, attempting to minimize and/or eliminate inherent "noise" in microarray experiments. However, the mixed linear models employed in our method are of a different form from previous studies. We implemented our method in two interconnected steps using a concise algorithm, MINQUE, with no requirement for assuming a normal distribution in the microarray data. In the first step, we choose a subset of potential DEGs, using the single-gene model. This procedure is similar to a *x*-fold variation filter. However, the *x*-fold variation filter is usually based on total gene-expression variations, while our procedure uses total treatment effects, which may increase the filter efficiency. In the second step, multiple gene-expression profiles are analyzed simultaneously and some interesting effects are estimated, using the multi-gene model. In our study, Bonferroni's method was used to set the cutoff for a

Table 4 *A. thaliana* genes with unstable transcripts identified by the mixed-model approach. Expressed sequence tags (*Locus*) were identified as differentially expressed genes by both the mixed-model approach and the *t*-test method

Accession	Locus	Gene information	*p*-value	Accession	Locus	Gene information	*p*-value	Accession	Locus	Gene information	*p*-value
AA042089	AT4g02380	Similar to several small proteins (ca. 100 aa) that are induced by heat, auxin, ethylene and wounding	3.58E−06	AA720291	At1g08880	Strong similarity to *Picea* histone H2A (gb|X67819). ESTs ATTS3874, T46627,T14194 come from this gene	6.77E−06	R30283	[a]AT4g31550	(AL080283) putative DNA-binding protein [*A. thaliana*]	2.35E−06
AA042408	[a]AT4g29950	Putative protein	5.07E−06	AA713153	At3g51360	Putative protein	7.88E−06	R29917	[a]At4g17230	Scarecrow-like 13 (SCL13)	2.55E−05
AA042412	[a]At1g32130	Unknown protein	6.79E−07	AA720105	At4g30440	Nucleotide sugar epimerase protein	1.30E−06	R30557	At5g42380	Putative protein	4.35E−05
AA042669	[a]At1g75900	Anter-specific proline-rich protein apg (protein cex)	4.80E−05	AA720239	At5g64260	(AB018441) phi-1 [*Nicotiana tabacum*]	7.10E−06	R64946	[a]At3g57930	Putative protein	3.27E−05
AA067498	At5g37770	Calmodulin-related protein 2, touch-induced	3.11E−07	AA720240	At5g64260	(AB018441) phi-1 [*N. tabacum*]	2.79E−05	R65120	At2g37940	(AC007661) unknown protein [*A. thaliana*]	1.96E−05
AA394319	[a]AT3g56880	Putative protein	1.03E−05	AA713007	[a] At2g41410	Calmodulin-like protein	3.26E−05	R87001	[a]AT5g61600	EREBP-4 [*N. tabacum*]	1.68E−05
AA394366	At3g15450	(AL078467) putative protein	3.43E−05	AI100427	[a] AT4g15760	(Z97339) hypothetical protein	7.79E−07	R89921	[a]AT5g19190	Putative protein	1.28E−06
AA394409	At4g17090	Putative beta-amylase	9.50E−05	AI100650	[a]At1g01550	Hypothetical protein	2.04E−07	R90579	At5g04610	Putative protein	5.94E−05
AA394587	[a]AT5g41080	Putative protein	2.92E−05	H36178	At5g61590	Ethylene responsive element binding factor-like	5.47E−06	T13984	At5g56980	Putative protein	1.38E−06
AA394829	At4g36500	(Z99708) putative protein [*Arabidopsis thaliana*]	3.84E−05	H36428	[a]AT3g11410	Protein phosphatase 2C (PP2C)	2.73E−05	T13839	[a]AT3g15450	Unknown [*Arabidopsis thaliana*]	1.07E−05
AA395006	[a]AT3g45970	(AL035539) putative pollen allergen [*A. thaliana*]	1.13E−06	H76905	At2g24790	(AC006585) CONSTANS-like B-box zinc finger protein	1.07E−05	T04337	[a]At2g22430	Homeobox-leucine zipper protein ATHB-6 (HD-ZIP protein ATHB-6)	2.86E−07
AA395343	[a]At2g23810	Similar to senescence-associated	6.88E−05	H36869	[a]AT3g56360	Putative protein	1.59E−05	T13991	[a]At2g29450	Glutathione S-transferase 103-1A	2.03E−06
AA395351	[a]AT5g67480	(AL035605) putative protein	5.03E−05	H37631	[a]AT5g63790	ATAF2 protein [-*A. thaliana*]	5.99E−05	T14209	[a]At1g19180	Unknown protein	6.05E−09
AA395830	At1g13260	(AB013886) RAV1	3.07E−07	H76698	[a]AT5g05440	Putative protein	3.98E−06	N96483	[a]At2g32150	(AC006223) putative hydrolase	1.28E−05
AA395910	[a]AT3g62550	Putative protein	1.57E−05	H36431	[a] AT4g32020	Putative protein	9.84E−06	T20543	At5g64260	(AB018441) Phi-1 [*Nicotiana tabacum*]	5.29E−08

AA585854	[a]AT3g54810	Similar to GATA transcription factor 3	4.73E−09	H76940	At1g06760	Histone H1.1	2.23E−06	T20842	At5g21940	Expressed protein	7.90E−05
AA585971	At2g32150	Putative hydrolase	5.58E−06	H77088	At2g22430	Homeobox-leucine zipper protein ATHB-6 (HD-ZIP protein ATHB-6)	2.74E-06	T21879	[a]AT5g04340	Putative c2h2 zinc finger transcription factor [*Arabidopsis thaliana*]	2.18E−06
AA597384	At1g19180	Unknown protein	1.22E−05	N37308	At1g63090	Unknown protein	2.86E−05	T21700	[a]At1g09070	(AJ007586) src2-like protein [*Arabidopsis thaliana*]	9.01E−07
AA597420	[a]At1g74950	Unknown protein	9.10E−06	N37328	[a] At1g13260	(AB013886) RAV1	7.93E−06	T22403	At1g66180	Unknown protein	1.89E−07
AA597982	[a]At5g03430	Putative protein	6.19E−06	N37850	[a] At1g69890	T7N9.12 [*Arabidopsis thaliana*]	3.85E−07	T22424	[a]At1g17420	Lipoxygenase [*Lycopersicon esculentum*]	2.73E−06
AA598115	At5g61900	Copine-like protein	3.50E−06	N37995	[a]AT5g56190	F22O2.16 [*Arabidopsis thaliana*]	9.84E−05	T22441	At5g63160	Putative protein	1.70E−05
AA598137	At1g07280	Unknown protein	1.66E−05	N38405	[a]At1g37130	Nitrate reductase2 (NR2)	1.55E−05	T45380	At1g62180	APS reductase [*A. thaliana*]	9.86E−06
AA605453	[a]At2g26190	Unknown protein	2.88E−05	N65459	At4g31500	(D78598) Cytochrome P450 monooxygenase	4.44E−05	T41662	At4g16920	(Z97342) Disease resistance RPP5- like protein [*Arabidopsis thaliana*]	1.43E−05
AA605476	At5g53050	Hydrolase alpha/ beta fold family protein	1.90E−05	N95945	At2g18440	Unknown protein	7.02E−07	T46143	At1g72450	Unknown protein	2.53E−05
AA650871	At1g19180	Unknown protein	1.85E−06	N95988	[a]At1g25550	Hypothetical protein	6.89E−07	T76004	At4g30440	Nucleotide sugar epimerase protein	1.58E−08
AA651102	[a]AT5g20880	Expressed protein	5.36E−07	N96265	[a] AT4g18010	Putative inositol polyphosphate 5-phosphatase At5P2	3.78E−09	T75944	At4g31500	(D78598) cytochrome P450 monooxygenase	4.69E−05
AA651342	[a]AT3g47960	(AL049658) putative peptide transporter [*Arabidopsis thaliana*]	8.67E−05	R29894	[a]At2g40000	(AF002109) putative nematode-resistance protein	2.71E−07	T76090	[a]AT3g15210	(AB008106) ethylene responsive element binding factor 4	2.94E−05
AA712424	[a]At1g49500	Unknown protein	1.16E−05	T20525	[a]AT3g52400	(AJ245407) putative syntaxin protein [*A. thaliana*]	3.08E−07	T76263	[a]AT4g24570	(AL035356) putative mitochondrial uncoupling protein	1.17E−05
AA712786	At5g61900	Copine-like protein	1.33E−05	N97061	At1g75860	Unknown protein	3.59E−06	T76510	[a]AT3g55240	Putative protein	3.59E−07
AA712865	[a]At1g32920	Unknown protein	3.34E−07	N96457	[a] AT5g64260	Phi-1-like protein	1.97E−06	W43654	At3g59350	Protein kinase-like protein	8.54E−05

significant P-value in both the single-gene and multi-gene models. The significance level, α, can be a little larger in the former than in the latter, which may reduce the risk of losing some interesting DEGs during the filtration procedure and thus increase statistical power. Other criteria such as Benjamini and Hochberg's procedure can also be used to adjust the P-value to control false discovery rates in these two models (Benjamini and Hochberg 1995). Our method can also handle designs with more than two dyes that can decrease the experimental costs (Forster et al. 2004). Another advantage of our method is its ability to handle missing data, a common problem in microarray experiments.

Replications of spot measurements either within or between arrays are essential in our method. Our method can be applied to the reference design and loop design and their modifications with replications. Replication is an important aspect of a good microarray design. There are basically two types of replication: (1) biological replication in which RNA samples from independent sources are used and (2) technical replication in which the same RNA sample is applied to different arrays. Whether biological or technical replication or both are used in microarray experiments depends on the relative magnitude of the biological and technical variability in the sample. Repeated spots on the same array are a kind of replication but apply the same RNA samples within the same array. This can reduce array effects due to the quality of robot-fabricated immobilized cDNA probes within the same array. Lee et al. (2000) recommended that at least three replicates be used in designing experiments using cDNA microarrays. In our simulated experiments with three replicates, although our method performed reasonably better than the t-test method, only those DEGs with large GT variation were consistently identified in most cases. Therefore, the number of genes identified in most microarray experiments likely represents an underestimate of DEGs when using a conservative significant level. If experimental outlay and sample are enough, six to eight replicates are likely the best (Pan 2002).

Various clustering methods are commonly used in microarray data analysis (Eisen et al. 1998; Spellman et al. 1998; Golub et al. 1999; Tamayo et al. 1999; Hastie et al. 2000; Pan 2002). In these methods, expression levels or ratios with sampling errors within experiments are usually analyzed directly, which may introduce noise and even bias in identifying groups of genes and thus result in the false interpretation of gene-expression patterns. Our method is complementary to the current clustering methods. In our method, interesting effects (such as the gene × treatment interactions here) can be predicted and/or estimated. Investigators can use these genetic effects in clustering to make sure the inputs are biologically meaningful. In our previous study, we also proposed a dissimilarity coefficient for clustering populations, using mixed linear models (such as the models proposed in our microarray study). The dissimilarity coefficient has two parameters, for squared difference of marginal mean and variance component of interaction, and has appropriate statistical properties (Zhu and Weir 1994b). Incorporation of such techniques in our method specifically for microarray data is straightforward and awaits further investigation.

In our simulations, we investigated the impact of various source variations on the efficiency of identifying genes expressed differentially among different treatments. We found that the same method resulted in dramatically different efficiencies (power, false discovery rate) under different configurations of the remaining source variations, given that the proportion of GT interactions accounting for the total gene-expression variations was fixed. For example, when $V_{GT}/V_P = 0.6$, the t-test method had 40% power in identifying DEGs when the dye effect and gene-specific dye effect accounted for a majority of the remainder variation, while this method had less than 10% power when the array effect and spot effect dominated the remainder variation (Fig. 1). A similar trend was observed in our method. This suggests that the efficiency of detecting DEGs is more affected by the systematic variation arising from arrays than that from dyes. If the experiment is finished for several batches within each array, the batch effects in the arrays may be considered to diminish the systematic errors. Modeling such effects or other appropriate effects in the single- and multi-gene models is straightforward in our method. Our studies have an important implication for the experimental design and execution of microarray studies. A desirable experimental design of a microarray should keep experiment-wise systematic errors as low as possible and, at the same cost, selectively diminish the systematic errors of some specific factors (such as the arrays here) that have more effect on the efficiency of detecting DEGs.

Treatments, genes, dyes, arrays and their interactions are well known as the source of effects contributing to variations in microarray data (Kerr et al. 2000; Churchill 2002). However, simulations of microarray data have not gained wide acceptance because, in the real world, a potential complexity may be involved in these source variations. This also makes difficulties for theoretical justifications of different statistical methods. In our study, in addition to simulated data, we compared experimentally the mixed-model approach with the t-test, using a real dataset for identifying unstable transcripts (Gutiérrez et al. 2002). The results showed that our method can identify more unstable transcripts than the t-test. We suggest researchers check their data distribution and pre-analyze various source variations in their experiments. Our method can be a competing candidate approach for those datasets which depart from normality and have moderate experimental errors.

Acknowledgements This research was supported in part by the National Natural Science Foundation of China. We greatly thank David Bartsch for his careful reading of this manuscript. We thank the Stanford microarray database for their opening data source.

References

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing, J R Stat Soc B 57:289–300

Bouton CM, Pevsner J (2002) DRAGON view: information visualization for annotated microarray data. Bioinformatics 18:323–324

Brazma A, Vilo J (2000) Gene expression data analysis. FEBS Lett 480:17–24

Chen Y, Dougherty ER, Bittner MI (1997) Ratio-based decisions and the quantitative analysis of cDNA microarray images. J Biomed Opt 2:364–374

Chen Y, Kamat V, Dougherty ER, Bittner ML, Meltzer PS, Trent JM (2002) Ratio statistics of gene expression levels and applications to microarray data analysis. Bioinformatics 18:1207–1215

Churchill GA (2002) Fundamentals of experimental design for cDNA microarrays. Nat Genet 32 [Suppl]:490–495

Dudoit S, Yang YH, Callow MJ, Speed TP (2000) Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. Stat Sin 12:111–139

Efron B, Tibshirani R, Storey JD, Tusher V (2001) Empirical Bayes analysis of a microarray experiment. J Am Stat Assoc 96:1151–1160

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95:14863–14868

Forster T, Costa Y, Roy D, Cooke HJ, Maratou K (2004) Triple-target microarray experiments: a novel experimental strategy. BMC Genomics 5:13

Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, et al (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286:531–537

Greenberg JR (1972) High stability of messenger RNA in growing cultured cells. Nature 240:102–104

Guffanti A, Reid JF, Alcalay M, Simon G (2002) The meaning of it all: web-based resources for large-scale functional annotation and visualization of DNA microarray data. Trends Genet 18:589–592

Gutierrez RA, Ewing RM, Cherry JM, Green PJ (2002) Identification of unstable transcripts in *Arabidopsis* by cDNA microarray analysis: rapid decay is associated with a group of touch- and specific clock-controlled genes. Proc Natl Acad Sci USA 99:11513–11518

Hastie T, Tibshirani R, Eisen MB, Alizadeh A, Levy R, Staudt L, Chan WC, Botstein D, Brown P (2000) 'Gene shaving' as a method for identifying distinct sets of genes with similar expression patterns. Genome Biol 1:3

Henderson CR (1963) Selection index and expected genetic advance. In: Hanson WD, Robinson HE (eds) Statistical genetics and plant breeding. National Academy of Sciences, Washington, DC

Ibrahim JG, Chen MH, Gray RJ (2002) Bayesian models for gene expression with DNA microarray data. J Am Stat Assoc 97:88–99

Ideker T, Thorsson V, Siegel AF, Hood LE (2000) Testing for differentially-expressed genes by maximum likelihood analysis of microarray data. J Comput Biol 7:805–817

Kerr MK, Churchill GA (2001) Experimental design for gene expression microarrays. Biostatistics 2:183–201

Kerr MK, Martin M, Churchill GA (2000) Analysis of variance for gene expression microarray data. J Comput Biol 7:819–837

Lee ML, Kuo FC, Whitmore GA, Sklar J (2000) Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. Proc Natl Acad Sci USA 97:9834–9839

Liu L, Hawkins DM, Ghosh S, Young SS (2003) Robust singular value decomposition analysis of microarray data. Proc Natl Acad Sci USA 100:13167–13172

Miller RG (1974) The Jackknife: a review. Biometrika 61:1–15

Newton MA, Kendziorski CM, Richmond CS, Blattner FR, Tsui KW (2001) On differential variability of expression ratios: improving statistical inference about gene expression changes from microarray data. J Comput Biol 8:37–52

Pan W (2002) A comparative review of statistical methods for discovering differentially expressed genes in replicated microarray experiments. Bioinformatics 18:546–554

Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, Iyer V, Jeffrey SS, Van de RM, Waltham M, et al (2000) Systematic variation in gene expression patterns in human cancer cell lines. Nat Genet 24:227–235

Schuchhardt J, Beule D, Malik A, Wolski E, Eickhoff H, Lehrach H, Herzel H (2000) Normalization strategies for cDNA microarrays. Nucleic Acids Res 28:E47

Searle SR, Casella G, McCulloch CE (1992) Variance components. Wiley, New York

Smyth GK (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 31

Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB, Brown PO, Botstein D, Futcher B (1998) Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. Mol Biol Cell 9:3273–3297

Tamayo P, Slonim D, Mesirov J, Zhu Q, Kitareewan S, Dmitrovsky E, Lander ES, Golub TR (1999) Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. Proc Natl Acad Sci USA 96:2907–2912

Thomas JG, Olson JM, Tapscott SJ, Zhao LP (2001) An efficient and robust statistical modeling approach to discover differentially expressed genes using genomic expression profiles. Genome Res 11:1227–1236

Welsh JB, Zarrinkar PP, Sapinoso LM, Kern SG, Behling CA, Monk BJ, Lockhart DJ, Burger RA, Hampton GM (2001) Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer. Proc Natl Acad Sci USA 98:1176–1181

West D (2003) Bayesian factor regression models in the "large ,*p* small *n*" paradigm. Bayesian Stat 7:723–732

Wolfinger RD, Gibson G, Wolfinger ED, Bennett L, Hamadeh H, Bushel P, Afshari C, Paules RS (2001) Assessing gene significance from cDNA microarray expression data via mixed models. J Comput Biol 8:625–637

Zhu J (1993) Methods of predicting genotype value and heterosis for offspring of hybrids. J Biomath 8:32–44

Zhu J, Weir BS (1994a) Analysis of cytoplasmic and maternal effects. I. a genetic model for diploid plant seeds and animals. Theor Appl Genet 89:625–637

Zhu J, Weir BS (1994b) Clustering populations by mixed linear models. J Biomath 9:1–14

Zhu J, Weir BS (1996) Diallel analysis for sex-linked and maternal effects. Theor Appl Genet 92:1–9