JOURNAL OF COMPUTATIONAL BIOLOGY
Volume 12, Number 3, 2005

Pp. 370–390

Epigenetic Hypothesis Tests for Methylation and Acetylation in a Triple Microarray System

LANG LI,[1,2] HUIDONG SHI,[3] CONSTANTIN YIANNOUTSOS,[1,2]
TIM HUI-MING HUANG,[4] and KENNETH P. NEPHEW[2,5]

ABSTRACT

To fully elucidate the functional relationship between DNA methylation and histone hypoacetylation in gene silencing, we have developed an integrated "triple" microarray system that allows us to begin to decipher the influence of epigenetic hierarchies on the regulation of gene expression in cancer cells. Our hypothesis is that in the promoter region of a silenced gene, reversal of two epigenetic factors (i.e., DNA demethylation and/or histone hyperacetylation) is highly correlated with gene reexpression after treatment of the human epithelial ovarian cancer cell line CP70 with the drug combination 5-aza-2′-deoxycytidine (DAC), a demethylating agent, and trichostatin A (TSA), an inhibitor of histone deacetylases. To estimate the posterior probabilities for genes with altered expression, DNA methylation and histone acetylation status measured with a triple-microarray system, we have employed an established empirical Bayes model. Two methods have been proposed to test our hypothesis that DNA demethylation and histone hyperacetylation are highly correlated among those up-regulated genes. One method follows a weighted least squares regression, while the other is derived from a chi-square statistic. The data derived by these approaches, which have been further verified through bootstrap analyses, support the proposed epigenetic correlation (p-values are less than 0.001). Further simulations suggest that even if the constant variance and normality assumptions do not hold, the power of those two tests is robust.

Key words: acetylation, empirical Bayes, epigenetics, methylation, microarray.

[1]Division of Biostatistics, Department of Medicine, Indiana University, Indianapolis, IN 46022.
[2]Indiana University Cancer Center, Indianapolis, IN 46022.
[3]Department of Pathology and Anatomical Sciences, Ellis Fischel Cancer Center, University of Missouri, School of Medicine, Columbia, Missouri 65203.
[4]Division of Human Cancer Genetics, Department of Molecular Virology, Immunology, and Medical Genetics, Comprehensive Cancer Center, The Ohio State University, Columbus, OH 43210.
[5]Medical Sciences, Indiana University, Bloomington, IN 47405.

INTRODUCTION

Microarray approaches used to study functional DNA–protein interactions have revealed that many transcription regulators are linked to chromatin remodeling, placing this type of epigenetic change at the center of gene regulation (Ren *et al*., 2000; Weinmann *et al*., 2002; Suzuki *et al*., 2002; Cameron *et al*., 1999). Repressed chromatin and gene silencing are associated with changes in DNA methylation and histone hypoacetylation (Jones and Baylin, 2002). Even though these epigenetic modifications are widely recognized as contributing factors in human tumorigenesis, the molecular mechanism(s) is not well understood. To fully elucidate the functional relationship between DNA methylation and histone hypoacetylation in gene silencing, we have developed an integrated "triple" microarray system to decipher the influence of epigenetic hierarchies on regulation of gene expression in cancer cells (Shi *et al*., 2002, 2003). We used a panel of GC-rich promoter and first exon DNAs, called ECISTs (expressed CpG island sequence tags), to screen methylated sites within 1,507 CpG island fragments (Shi *et al*., 2003). In addition, we utilized the promoter sequences within the ECISTs to identify chromatins immunoprecipitated with antibodies against hyperacetylated histones and the exon-containing portions for measuring levels of the corresponding transcripts. This new microarray system provides an effective tool for segregating, at specific loci, changes in gene expression that are a consequence of promoter demethylation and histone hyperacetylation by epigenetic treatments (Shi *et al*., 2003).

To assess the status of gene expression, DNA methylation, and histone acetylation simultaneously, the human epithelial ovarian cancer cell line CP70 was treated with 5-aza-2′-deoxycytidine (DAC), a demethylating agent, and trichostatin A (TSA), an inhibitor of histone deacetylases (Shi *et al*., 2003). Two replicates were generated for each type of triple microarray experiment. To analyze our triple datasets, we chose the fold-change method described by Chen *et al*. (1997). The resulting analysis established an initial relationship between gene reexpression, DNA demethylation, and histone hyperacetylation. However, this preliminary analysis lacked both a formal hypothesis and the sound justification of the statistical tests used to support the fold-change criterion when using different types of microarray data.

Based on this prior analysis, here we have sought to develop rigorous statistical methods to test the hypothesis that the reversal of two epigenetic factors (i.e., DNA demethylation and/or histone hyperacetylation) in the promoter region of a silenced gene is highly correlated with gene reexpression after treatment with the combination of DAC+TSA. In order to test this hypothesis, it is crucial to know whether or not a gene is reexpressed, demethylated, or hyperacetylated. Because these indicator variables are not directly observable, it is desirable to estimate their change in status from the data. Statistically, for any gene, indicator variable estimates can be used to interpret the probability of a change in status (reversal of an epigenetic factor), given the data.

Several statistical methods have been used for microarray analysis of differentially expressed genes. These approaches are based primarily on multiple comparisons of different experimental conditions. For example, Kerr *et al*. (2000) and Wolfinger *et al*. (2001) employed analysis of variance (ANOVA) models to conduct a hypothetical test for different expression levels of individual genes in multiple microarray experiments. Dudoit *et al*. (2001) used a t statistic to address the problem of multiple comparisons through permutation analysis. These approaches yielded only p-values representing the probability of having an observed expression difference of a given gene, if the status is assumed to be the same before and after a treatment. However, for the triple-array data, these approaches cannot be used to estimate the status change. A Bayes or an empirical Bayes approach may be a more appropriate choice for this type of analysis. In this regard, Efron *et al*. (2001) proposed an empirical Bayes model for calculating the probability of a differentially expressed gene given the observed data. As the empirical distribution for their t-like statistic for each gene does not share variation information, it works well only in situations involving at least a few replicates. Newton *et al*. (2001) also proposed another empirical Bayes model for cDNA microarray experiments with only one replicate. According to their method, if a gene is differentially expressed under two conditions, its level of expression is independently generated from the same distribution (Newton *et al*., 2001); otherwise, the level of expression is the same between experimental and control samples. In their work, the observational component follows a gamma distribution with mean μ_g, and μ_g itself follows an inverse gamma distribution (prior component). This hierarchical model is often referred to as the gamma–gamma model. Specifically, all genes share the same distribution for the within-gene sampling errors. This is a crucial feature in their method, as no replicates were available in their data example.

Kendziorski *et al*. (2003) further extended the gamma–gamma model to situations where replicates were available. In addition, they developed a log-normal model for the observational component and a normal model for the prior component. They demonstrated a comparable performance for a log-normal–normal model and a gamma–gamma model.

There are two major advantages in the methods proposed by Newton *et al*. (2001) and Kendziorski *et al*. (2003). First, information sharing is a consequence of the empirical Bayes approach. The model pools the variation information across all genes, making it well suited for datasets containing only a few replicates (e.g., two replicates). Second, it is easy to introduce correlation structures in a log-normal model. Therefore, in this paper, we have adapted and improved the log-normal–normal model for the triple-microarray analysis. Instead of having one error term to describe variations as in their model, we have multiple error terms to represent variation between technical replicates (arrays) and measurement error within a single array, respectively (Churchill, 2002). In addition, it is important to develop a statistical approach to formally test our epigenetic hypothesis: demethylation and histone hyperacetylation status are strongly correlated among up-regulated genes after epigenetic treatment. Moreover, if the normality assumption or equal variance assumption do not hold, it is important to know the power of the statistical test and to evaluate its performance (Baggerly *et al*., 2001). This paper is organized as follows: 1) triple-microarray experiment, data, statistical models and hypothetical test methods (Materials and Methods); 2) data analysis, power analysis, and sensitivity analysis (Results); 3) Discussion.

MATERIALS AND METHODS

Cell culture

A human epithelial ovarian cancer cell line CP70 was cultured in the presence of vehicle (PBS) or DAC (0.5 μM; medium changed every 24 hours). After four days, cells were either harvested or treated with TSA (0.5 μM) for 12 hours. DNA and RNA were isolated using QIAarnp Tissue and RNeasy kits (Qiagen), respectively.

Methylation microarray analysis

Preparation of methylation amplicons was carried out as previously described (Yan *et al*., 2001). Briefly, CP70 DNA (1 μg) was digested with *MseI* and then ligated to a PCR-linker. The ligated DNA was digested with methylation-sensitive endonucleases *Bst*UI and *Hpa*II and amplified with a linker primer by PCR. Genomic fragments containing methylated sites were protected from enzymatic restrictions and could be amplified; however, fragments containing unmethylated sites were digested and not amplified. Both treated and control samples were cohybridized onto an ECIST microarray side with different labels (Cy5 or Cy3), as previously described (Yan *et al*., 2001; Shi *et al*., 2003).

Expression microarray analysis

In the expression microarray, total RNA (100 μg) was prepared from control CP70 cells or cell cultured with TSA/DAC. The RLCS method was used to generate full-length cDNAs. Both treated and control samples were cohybridized onto the ECIST microarray panel with different labels (Cy5 or Cy3).

Acetylation microarray analysis

The protocol used to identify immunoprecipitated E2F1 targets (Weinmann *et al*., 2002) was adapted for this study. To obtain a network of DNA–protein biopolymers, treated or untreated CP70 cells (2×10^7 cells/assay) were cross-linked using 1% formaldehyde. Cell nuclei were collected by microcentrifugation, and cross-linked chromatin fibers were isolated and fragmented to $\sim$600 bp by sonication. Immunoprecipitation was carried out with 5 μg of anti-acetylated histone H3 or H4 rabbit polyclonal antibody (Upstate) or no-antibody (negative control). Purified chromatin DNA was recovered from 10–15 preparations for fluorescent labeling. H3 or H4 treated samples were labeled with Cy5, while untreated samples (no-antibody) were labeled with Cy3. Microarray hybridization and posthybridization washings procedures have been previously described by Weinmann *et al*. (2002).

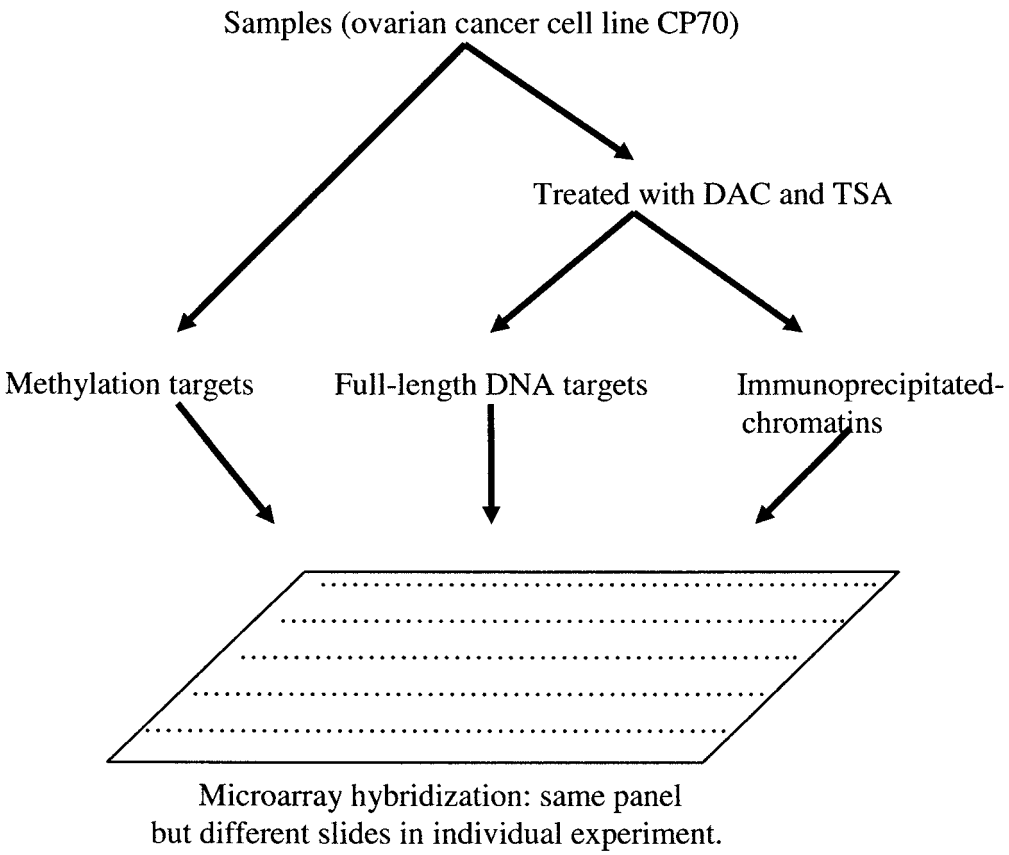

FIG. 1. Schematic flowchart for parallel assessment of gene expression, DNA methylation, and histone acetylation in ovarian cancer cell line CP70.

Data structure

The CP70 cell line, cultured with either vehicle or DAC+TSA, was divided into two aliquots for two technical replicates for the three microarray analyses. Each microarray used the same array panel but different slides in individual experiments. A schematic flowchart for parallel assessment of gene expression, DNA methylation, and histone acetylation is presented in Fig. 1. Among expression and methylation arrays, two replicates followed the same protocol in two parallel experiments, and Cy5 and Cy3 dyes were swapped between treated (DAC+TSA) and untreated (vehicle) samples. However, for the acetylation arrays, the no-antibody treated sample served as a negative control with the Cy3 label and antibody treated (H3 or H4) samples were labeled by Cy5. Among those four acetylation arrays, a balanced 2 by 2 factorial experiment was designed: (H3, H4) versus (vehicle, DAC+TSA). Although the no-antibody effect and the dye effect were totally confounded, the treatment effect (vehicle, DAC+TSA) was not.

Normalization

Our analysis started from background-justified intensities. All negative intensities were set at a positive threshold. Many statistical strategies have been proposed to normalize array-to-array variations. Kerr, Martin, and Churchill (2000) advocated a unified ANOVA model to perform normalization and estimate gene-treatment interaction effects on the cDNA microarray data simultaneously. However, with nonlinear array-to-array variation (discussed below), that unified model has been difficult to implement. On the other hand, Wolfinger *et al.* (2001) proposed a two-step analysis strategy. At step one, a linear model was employed to normalize the data; and at step two, another linear mixed model was adopted to estimate gene-treatment interaction effects. We would like to combine the method of Kerr, Martin, and Churchill (2000) with the strategy of Wolfinger *et al.* (2001). At step one, a nonlinear model was used to normalize

the array-to-array variation; at step two, we used a linear mixed model to normalize the dye bias and estimate gene-treatment interaction effects.

To address array-to-array variation, we have considered the following notation. Denote G_{ijkg} as the log-transformed expression level for gene g from replicate i ($i = 1, 2$) under treatment group j ($j = 1$ for control, 2 for DAC+TSA) with dye k ($k = 1$ for Cy5, 2 for Cy3). In particular, let $\{G_{1\bullet\bullet g} = G_{111g} + G_{122g}\}_{g=1,\ldots,n}$ and $\{G_{2\bullet\bullet g} = G_{212g} + G_{221g}\}_{g=1,\ldots,n}$. (Note: the dye assignment in replicates followed a Latin square design). Therefore, if two replicates are comparable, $\{G_{1\bullet\bullet g} - G_{2\bullet\bullet g}\}$ should vary around 0 and this difference should be independent of $\{G_{1\bullet\bullet g} + G_{2\bullet\bullet g}\}$. This is the A (average of log-expression values) versus M (difference in log-expression values) plot proposed by Dudoit *et al*. (2002), which is displayed in Fig. 2(a). We have applied the Loess nonlinear normalization technique discussed by Bolstad *et al*. (2003). In addition, we have added one extra step: denote $G'_{i\bullet\bullet g}$ as the normalized total intensity for gene g of replicate i. The treated or untreated gene g's intensity is $G'_{ijkg} = G'_{i\bullet\bullet g} \times G_{ijkg}/G_{i\bullet\bullet g}$. Figure 2(b) displays the expression MA plot after normalization. M is approximately around 0 and independent of A.

Similarly, the notation of methylation arrays has the same subscripts M_{ijkg} and the normalization follows the same procedures as those of expression arrays. Figure 2(c) and (d) are MA plots before and after normalization, respectively.

The notation for acetylation arrays is the same but its interpretation is different from the previous two. Denote A_{ijkg} as the log-transformed acetylation level for gene g from replicate i ($i = 1$ for H3, 2 for H4) under treatment group j ($j = 1$ for control, 2 for DAC+TSA) with antibody k ($k = 1$ for yes, 2 for no). In particular, let $\{A_{1j\bullet g} = A_{1j1g} + A_{1j2g}\}_{g=1,\ldots,n}$ and $\{A_{2j\bullet g} = A_{2j1g} + A_{2j2g}\}_{g=1,\ldots,n}$. Following the same procedures as with the other two arrays, Fig. 2(e) and (g) are MA plots before normalization for the

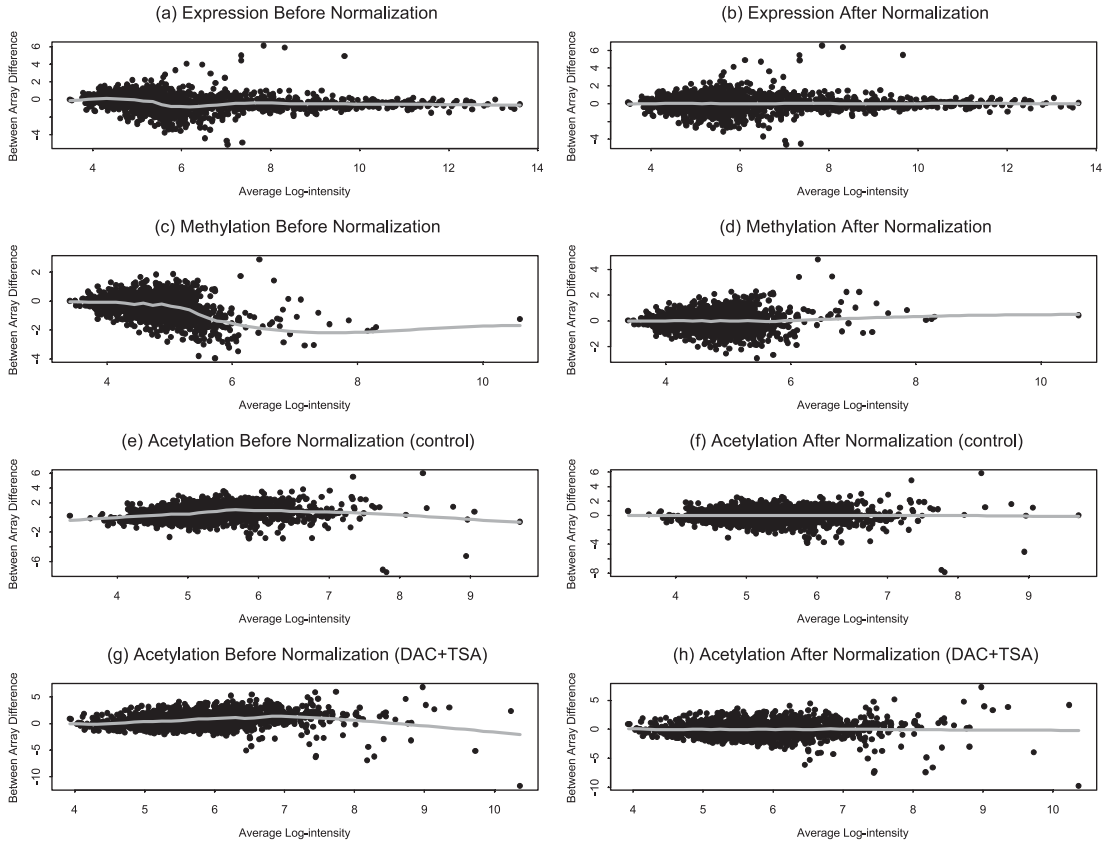

FIG. 2. Triple array normalization. (**a**) Gene expression before normalization. (**b**) Gene expression after normalization. (**c**) DNA methylation before normalization. (**d**) DNA methylation after normalization. (**e**) Histone acetylation before normalization (control). (**f**) Histone acetylation after normalization (control). (**g**) Histone acetylation before normalization (treated). (**h**) Histone acetylation after normalization (treated). The bold lines are based on the Loess fits.

control and DAC+TSA treated arrays respectively, and Fig. 2(f) and (h) are their corresponding MA plots after normalization. Without loss of generality, we still use $(G_{ijkg}, M_{ijkg}, A_{ijkg})$ to denote the normalized intensities.

There was no significant evidence of a weak hybridization region for any of our array samples, either before or after normalization. From Fig. 2, all Loess curves, before normalization, appear to be flat at the beginning, suggesting that low signals for low abundance or absent genes were comparable for both replicates from all three array types. In other words, weak signals were comparable between two replicates. However, some curvature for higher signals was evident, suggesting that biological and technical variations were higher for the higher abundance genes. The nonlinear normalization effectively detected this overall discrepancy pattern among all genes between two replicates and globally filtered out biological and technical variations. The impact of the normalization step to the second step analysis will be discussed in the Results section.

Model specification

Although array-to-array variations between replicates were approximately 0 on average after normalization, different genes showed differences between replicates, even after normalization (Fig. 2(b), (d), (f), (h)). In addition, dye bias for expression and methylation arrays was not yet corrected . Therefore, similar to the step-two of Wolfinger *et al*. (2001), our model for gene expression is

$$G_{ijkg} = 1_{\{k=1\}}\gamma_G + \mu_{G,jg} + b_{G,ig} + \varepsilon_{G,ijg}, \tag{1}$$

where the subscript G denotes the random variables and regression parameters for gene expression data. Subscripts M and A will be adopted for methylation and acetylation models, respectively, as the acetylation model has a different formulation from (1). In particular, G_{ijkg} is the normalized expression level after step one. The symbol $\varepsilon_{G,ijg} \sim$ i.i.d. $N(0, \sigma^2_{G,0})$ represents the measurement error within a single array (i.e., reading the fluorescent labels). It is referred to as the third component of variation in Churchill *et al*. (2002). The symbol $b_{G,ig} \sim$ i.i.d. $N(0, \sigma^2_{G,1})$ is the between-replicate sampling error (i.e., array-to-array variation). It is referred to as the second component of variation in Churchill *et al*. (2002). The symbol γ_G is the dye effect due to Cy5 comparing to Cy3, and $1_{\{k=1\}}$ is an indicator for Cy5. The vectors $(\mu_{G,1g}, \mu_{G,2g})^T$ are the expression levels for gene g in the control and treatment groups, respectively. It follows a mixture of two normal distributions (2),

$$\begin{aligned}\begin{pmatrix}\mu_{G,1g}\\ \mu_{G,2g}\end{pmatrix} \sim{}& (1 - z_{G,g})N\left\{\begin{pmatrix}\mu_G\\ \mu_G\end{pmatrix}, \begin{pmatrix}\sigma^2_{G,2} & \sigma^2_{G,2}\\ \sigma^2_{G,2} & \sigma^2_{G,2}\end{pmatrix}\right\}\\ &+ z_{G,g}N\left\{\begin{pmatrix}\mu_G\\ \mu_G\end{pmatrix}, \begin{pmatrix}\sigma^2_{G,2} & \rho_G\sigma^2_{G,2}\\ \rho_G\sigma^2_{G,2} & \sigma^2_{G,2}\end{pmatrix}\right\},\end{aligned} \tag{2}$$

where $z_{G,g} = 0$ if the expression levels of gene g are not different between two treatment conditions and $z_{G,g} = 1$ otherwise. If the expression levels are not different, model (2) implies that $\mu_{G,1g} = \mu_{G,2g} = \mu_G$ and they are sampled from distribution $N(\mu_G, \sigma^2_{G,2})$; otherwise, $(\mu_{G,1g}, \mu_{G,2g})^T$ are sampled from $N\left[\begin{pmatrix}\mu_{G,1}\\ \mu_{G,2}\end{pmatrix}, \begin{pmatrix}\sigma^2_{G,1} & \rho_G\sigma^2_{G,2}\\ \rho_G\sigma^2_{G,2} & \sigma^2_{G,2}\end{pmatrix}\right]$, i.e., a bivariate normal distribution with correlation ρ_G. In particular, $N(\mu_G, \sigma^2_{G,2})$ represents a latent distribution of expression levels among all genes.

Together with model (1), our mixture model (2) is an extension of the two-step linear mixed model of Wolfinger *et al*. (2001). Because of an incorporation of the indicator variable, $1_{\{G,g\}}$, we are able to estimate the probability of a gene being differentially expressed between two treatment conditions. In addition, the two error terms in model (1) are also an extension of the log-normal–normal model of Kendziorski *et al*. (2003) that had only one measurement error term to describe all technical and biological variation. Moreover, the mixture distribution (2) also has a term ρ_G to model the potential correlation of gene expression levels between treated and untreated groups, if they are differentially expressed, while Kendziorski and colleagues (2003) assumed the correlation to be 0. The gene–dye interaction effect has not been included in our current model (1). As this effect is orthogonal to the gene-treatment interaction

effect, $\mu_{G,2g} - \mu_{G,1g}$, it will not bias the estimation of the gene-treatment interaction effect. According to our initial analysis (not reported here), the gene-dye interaction effect was too small to have an impact on the gene-treatment interaction inference.

The complete data log-likelihood function is

$$
\begin{aligned}
l_G &= \sum_g z_{G,g}\{\log[f_1(\mathbf{G}_g)] + p_G\} + (1 - z_{G,g})\{\log[f_0(\mathbf{G}_g)] + 1 - p_G\} \\
\mathbf{G}_g &= (G_{111g}, G_{122g}, G_{212g}, G_{221g})^T, \\
f_1(\mathbf{G}_g) &= (2\pi|\mathbf{D}_{G,1}|)^{-1/2} \exp\{-(\mathbf{G}_g - \mathbf{x}_G\boldsymbol{\beta}_G)^T \mathbf{D}_{G,1}^{-1}(\mathbf{G}_g - \mathbf{x}_G\boldsymbol{\beta}_G)/2\}, \\
f_0(\mathbf{G}_g) &= (2\pi|\mathbf{D}_{G,0}|)^{-1/2} \exp\{-(\mathbf{G}_g - \mathbf{x}_G\boldsymbol{\beta}_G)^T \mathbf{D}_{G,0}^{-1}(\mathbf{G}_g - \mathbf{x}_G\boldsymbol{\beta}_G)/2\}, \\
\mathbf{x}_G &= (\mathbf{x}_1^T, \mathbf{x}_2^T)^T, \quad \mathbf{x}_1 = (1,1,1,1)^T, \quad \mathbf{x}_2 = (1,0,0,1)^T, \\
\boldsymbol{\beta}_G &= (\mu_G, \gamma_G), \\
\mathbf{D}_1 &= \mathrm{diag}_r(\sigma_{G,0}^2) + \sigma_{G,1}^2 \times \mathrm{diag}_2(1) \otimes I_2 + H_1 \otimes I_2 \\
\mathbf{D}_0 &= \mathrm{diag}_4(\sigma_{G,0}^2) + \sigma_{G,1}^2 \times \mathrm{diag}_2(1) \otimes I_2 + H_0 \otimes I_2 \\
H_{G,1} &= \begin{pmatrix} \sigma_{G,2}^2 & \rho_G\sigma_{G,2}^2 \\ \rho_G\sigma_{G,2}^2 & \sigma_{G,2}^2 \end{pmatrix}, \quad H_{G,0} = \begin{pmatrix} \sigma_{G,2}^2 & \sigma_{G,2}^2 \\ \sigma_{G,2}^2 & \sigma_{G,2}^2 \end{pmatrix}.
\end{aligned} \tag{3}
$$

I_n is an n by n matrix with all elements equal to 1. Denote p_G as the overall probability of changed gene expression. In particular, np_G represents the average number of genes with changed expression levels. Kendziorski *et al.* (2003) have assumed that the correlation $\rho_G = 0$. Under this assumption, $\mathbf{D}_1$ is

$$
\mathbf{D}_1 = \mathrm{diag}_4(\sigma_{G,0}^2) + \sigma_{G,1}^2 \times \mathrm{diag}_2(1) \otimes I_2 + \sigma_{G,2}^2 \times \mathrm{diag}_2(1) \otimes I_2. \tag{4}
$$

The log-likelihood function (3) is called a dependent model, while the log-likelihood function based on (4) is called an independent model. We have used the EM algorithm to perform the maximization. The detailed procedure is given in Appendix I. The estimation has been performed through function nlminb() in Splus.

The models for methylation microarrays are identical to those of expression arrays, except that subscript G is replaced by M. However, the acetylation array models are different from the other two models in the following way:

$$
A_{ijkg} = 1_{\{k=1\}} \times \mu_{A,jg} + \gamma_A + a_{A,g} + b_{A,ig} + \varepsilon_{A,ijkg}, \tag{5}
$$

where A_{ijkg} is the normalized acetylation level after step one; $1_{\{k=1\}} = 1$ if the antibody is applied and 0 otherwise; $\varepsilon_{A,ijkg} \sim$ i.i.d. $N(0, \sigma_{A,0}^2)$ represents the measurement error within a single array; $b_{A,ig} \sim$ i.i.d. $N(0, \sigma_{A,1}^2)$ is the between-replicate sampling error due to antibodies (H3 or H4); γ_A is the combination of dye and no-antibody effects averaged over all the genes; $a_{A,g} \sim N(0, \sigma_{A,a}^2)$ is the gene specific no-antibody and dye effect; and vectors $(\mu_{A,1g}, \mu_{A,2g})^T$ are the acetylation levels adjusted for both dye and no-antibody effects for gene g under control and treatment, respectively. It follows a mixture of two normal distributions,

$$
\begin{aligned}
\begin{pmatrix} \mu_{A,1g} \\ \mu_{A,2g} \end{pmatrix} \sim{}& (1 - z_{A,g}) N\left\{ \begin{pmatrix} \mu_A \\ \mu_A \end{pmatrix}, \begin{pmatrix} \sigma_{A,2}^2 & \sigma_{A,2}^2 \\ \sigma_{A,2}^2 & \sigma_{A,2}^2 \end{pmatrix} \right\} \\
&+ z_{A,g} N\left\{ \begin{pmatrix} \mu_A \\ \mu_A \end{pmatrix}, \begin{pmatrix} \sigma_{A,2}^2 & \rho_A\sigma_{A,2}^2 \\ \rho_A\sigma_{A,2}^2 & \sigma_{A,2}^2 \end{pmatrix} \right\}.
\end{aligned} \tag{6}
$$

The log-likelihood function is

$$
\begin{aligned}
l_A &= \sum_g z_{A,g}\{\log[f_1(A_g)] + p_A\} + (1 - z_{A,g})\{\log[f_0(A_g)] + 1 - p_A\} \\
A_g &= (A_{111g}, A_{112g}, A_{121g}, A_{122g}, A_{211g}, A_{212g}, A_{221g}, A_{222g})^T, \\
f_1(A_g) &= (2\pi|\mathbf{D}_{A,1}|)^{-1/2} \exp\{-(A_g - \mathbf{x}_A\boldsymbol{\beta}_A)^T \mathbf{D}_{A,1}^{-1}(A_g - \mathbf{x}_A\boldsymbol{\beta}_A)/2\}, \\
f_0(A_g) &= (2\pi|\mathbf{D}_{A,0}|)^{-1/2} \exp\{-(A_g - \mathbf{x}_A\boldsymbol{\beta}_A)^T \mathbf{D}_{A,0}^{-1}(A_g - \mathbf{x}_A\boldsymbol{\beta}_A)/2\}, \\
\mathbf{x}_A &= (\mathbf{x}_1^T, \mathbf{x}_2^T)^T, \quad \mathbf{x}_1 = (1,1,1,1,1,1,1,1)^T, \quad \mathbf{x}_2 = (1,0,1,0,1,0,1,0)^T, \\
\boldsymbol{\beta} &= (\gamma_A, \mu_A), \\
\mathbf{D}_1 &= \mathrm{diag}_8(\sigma_{A,0}^2) + \sigma_{A,1}^2 \times I_4 \otimes \mathrm{diag}_2(1) + \sigma_{A,a}^2 \times I_8 + H_1 \otimes I_2 \\
\mathbf{D}_0 &= \mathrm{diag}_8(\sigma_{A,0}^2) + \sigma_{A,1}^2 \times I_4 \otimes \mathrm{diag}_2(1) + \sigma_{A,a}^2 \times I_8 + H_0 \otimes I_2 \\
H_1 &= \begin{pmatrix} \sigma_{A,2}^2 & 0 & \rho_A\sigma_{A,2}^2 & 0 \\ 0 & 0 & 0 & 0 \\ \rho\sigma_{A,2}^2 & 0 & \sigma_{A,2}^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad H_0 = \begin{pmatrix} \sigma_{A,2}^2 & 0 & \sigma_{A,2}^2 & 0 \\ 0 & 0 & 0 & 0 \\ \sigma_{A,2}^2 & 0 & \sigma_{A,2}^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.
\end{aligned} \tag{7}
$$

where I_n is an n by n matrix with all elements equal to 1. Denote p_A as the overall probability of changed gene expression.

Epigenetic hypothesis test

Our primary epigenetic hypothesis is that among all up-regulated genes, DNA demethylation and histone hyperacetylation are highly correlated after treatment with DAC+TSA. Before testing, this hypothesis needs to be defined. Denote $(\Delta G_g, \Delta M_g, \Delta A_g)$ as three multinomial random variables that are equal to 1 if gene g is up-regulated, demethylated, and hyperacetylated, are equal to -1 if gene g is down-regulated, hypermethylated, and deacetylated, and are 0 if the status of all three is not changed. However, as $(\Delta G_g, \Delta M_g, \Delta A_g)$ are all unobserved from the triple-arrays, they need to be estimated from the data. Without loss of generality, we use ΔG_g as an illustration, as ΔM_g and ΔA_g follow in the same way:

$$
\begin{aligned}
\Delta\hat{G}_g^{(1)} &= \Pr\{\Delta G_g = 1 \mid \mathbf{G}_g\} \\
&= \Pr\{\mu_{G,g2} > \mu_{G,g1}) \cap (z_{G,g} = 1) \mid \mathbf{G}_g\} \\
&= \Pr\{\mu_{G,g2} > \mu_{G,g1} \mid z_{G,g} = 1, \mathbf{G}_g\} \Pr(z_{G,g} = 1 \mid \mathbf{G}_g) \\
&= \Phi(dG_g) \times \hat{z}_{G,g},
\end{aligned} \tag{8}
$$

$$
\begin{aligned}
\Delta\hat{G}_g^{(-1)} &= \Pr\{\Delta G_g = -1 \mid \mathbf{G}_g\} \\
&= \Pr\{(\mu_{G,g2} < \mu_{G,g1}) \cap (z_{G,g} = 1) \mid \mathbf{G}_g\} \\
&= \Pr\{\mu_{G,g2} < \mu_{G,g1} \mid z_{G,g} = 1, \mathbf{G}_g\} \Pr(z_{G,g} = 1 \mid \mathbf{G}_g) \\
&= \Phi(dG_g) \times \hat{z}_{G,g},
\end{aligned} \tag{9}
$$

$$
\begin{aligned}
\Delta\hat{G}_g^{(0)} &= \Pr\{\Delta G_g = 0 \mid \mathbf{G}_g\} \\
&= \Pr\{\mu_{G,g2} = \mu_{G,g1}) \cap (z_{G,g} = 0) \mid \mathbf{G}_g\} \\
&= \Pr(z_{G,g} = 0 \mid \mathbf{G}_g) = 1 - \hat{z}_{G,g},
\end{aligned} \tag{10}
$$

where $dG_g = (\frac{G_{111g} + G_{212g} - G_{122g} - G_{221g}}{\sigma_{G,0}^2 + \sigma_{G,1}^2}) / \sqrt{\frac{4\sigma_{G,0}^2}{(\sigma_{G,0}^2 + \sigma_{G,1}^2)^2} + \frac{2(1-\rho_G)}{\sigma_{G,2}^2}}$, Φ is the standard normal probability distribution (the detailed derivation for (8) is in Appendix II), and the three probabilities in (8), (9), and (10) add up to 1. In particular, $\hat{p}_G^{(1)} = \sum_g \Delta\hat{G}_g^{(1)}/n$, $\hat{p}_G^{(-1)} = \sum_g \Delta\hat{G}_g^{(-1)}/n$, and $\hat{p}_G^{(0)} = \sum_g \Delta\hat{G}_g^{(0)}/n$.

The same derivation can be applied to the methylation array indicator variable estimates, $(\Delta\hat{M}_g^{(1)}, \Delta\hat{M}_g^{(-1)}, \Delta\hat{M}_g^{(0)})$. However, the formulations of the $(\Delta\hat{A}_g^{(1)}, \Delta\hat{A}_g^{(-1)}, \Delta\hat{A}_g^{(0)})$ estimates are slightly different form (8), (9), and (10) in $\Phi(dA_g) = \Phi\left[\left(\frac{A_{111g}-A_{112g}+A_{211g}-A_{212g}-(A_{121g}-A_{122g}+A_{221g}-A_{222g})}{\sigma_{A,0}^2}\right) \Big/ \sqrt{\frac{4}{\sigma_{A,0}^2} + \frac{2(1-\rho_A)}{\sigma_{A,2}^2}}\right]$ (see Appendix III).

In order to test the association between estimated DNA demethylation and histone hyperacetylation status, our first method is simply a linear model,

$$\Delta\hat{A}_g^{(1)} = \beta_0 + \beta_1\Delta\hat{M}_g^{(1)} + e_g, \tag{11}$$

where e is an error term. The null hypothesis is that $\beta_1 = 0$, while the alternative is that $\beta_1 \neq 0$. Since our main goal is to focus on up-regulated genes, we have applied a weighted least squares estimate (Weinsberg, 1985), which maximizes the following objective function:

$$l_{WLS}(\beta_0, \beta_1) = -\sum_g \Delta\hat{G}_g^{(1)}(\Delta\hat{A}_g^{(1)} - \beta_0 - \beta_1\Delta\hat{M}_g^{(1)})^2. \tag{12}$$

If the up-regulation information ($\Delta\hat{G}_g^{(1)} = 0$ or 1) were known for all the genes, Equation (12) would be reduced to an ordinary least squares for all the up-regulated genes.

Our second method is based on a 2×2 contingency table with cell probability

$$p_{11|1} = \Pr(\Delta M_g \neq 1, \Delta A_g \neq 1 \mid \Delta G_g = 1), \quad p_{12|1} = \Pr(\Delta M_g \neq 1, \Delta A_1 = 1 \mid \Delta G_g = 1),$$
$$p_{21|1} = \Pr(\Delta M_g = 1, \Delta A_g \neq 1 \mid \Delta G_g = 1), \quad p_{22|1} = \Pr(\Delta M_g = 1, \Delta A_g = 1 \mid \Delta G_g = 1).$$

Under the null hypothesis, $\Delta M_g \perp \Delta A_g$ given that $\Delta G_g = 1$. Then the odds ratio

$$\text{OR} = (p_{11|1} \times p_{22|1})/(p_{11|1} \times p_{22|1}) = (p_{111} \times p_{221})/(p_{111} \times p_{221}) = 1. \tag{13}$$

The alternative hypothesis is OR $\neq 1$. To test this hypothesis, a natural choice is a chi-square statistic (Agresti, 1990). Denote the cell counts for the contingency table as

$$n_{11} = \sum_g 1\{\Delta G_g = 1, \Delta M_g \neq 1, \Delta A_g \neq 1\}, \quad n_{12} = \sum_g 1\{\Delta G_g = 1, \Delta M_g \neq 1, \Delta A_g = 1\},$$
$$n_{21} = \sum_g 1\{\Delta G_g = 1, \Delta M_g = 1, \Delta A_g \neq 1\}, \quad n_{22} = \sum_g 1\{\Delta G_g = 1, \Delta M_g = 1, \Delta A_g = 1\}, \tag{14}$$
$$n_{1\bullet} = n_{11} + n_{12}, \quad n_{2\bullet} = n_{21} + n_{22}, \quad n_{\bullet 1} = n_{11} + n_{21}, \quad n_{\bullet 2} = n_{12} + n_{22}, \quad n_{\bullet\bullet} = n_{\bullet 1} + n_{\bullet 2}.$$

The chi-square statistic is

$$X = \frac{(n_{11} - m_{11})^2}{m_{11}} + \frac{(n_{12} - m_{12})^2}{m_{12}} + \frac{(n_{21} - m_{21})^2}{m_{21}} + \frac{(n_{22} - m_{22})^2}{m_{22}}$$
$$\text{where} \quad m_{11} = \frac{n_{1\bullet} \times n_{\bullet 1}}{n_{\bullet\bullet}}, \quad m_{12} = \frac{n_{1\bullet} \times n_{\bullet 2}}{n_{\bullet\bullet}}, \quad m_{21} = \frac{n_{2\bullet} \times n_{\bullet 1}}{n_{\bullet\bullet}}, \quad m_{22} = \frac{n_{2\bullet} \times n_{\bullet 2}}{n_{\bullet\bullet}}. \tag{15}$$

As the joint Bernoulli random variable $1_{\{\Delta G_g=1, \Delta M_g\neq 1, \Delta A_g\neq 1\}}$ for gene g is unobserved, we have assumed that within each gene g, its up-regulation, demethylation, and hyperacetylation status are independent. Therefore, $1_{\{\Delta G_g=1, \Delta M_g\neq 1, \Delta A_g\neq 1\}} = 1_{\{\Delta G_g=1\}} \times 1_{\{\Delta M_g\neq 0\}} \times 1_{\{\Delta A_g\neq 0\}}$, and it can be estimated as

$$\begin{aligned}\hat{1}_{\{\Delta G_g=1, \Delta M_g\neq 1, \Delta A_g\neq 1\}} &= \hat{1}_{\{\Delta G_g=1\}} \times \hat{1}_{\{\Delta M_g\neq 1\}} \times \hat{1}_{\{\Delta A_g\neq 1\}} \\ &= \Pr\{\Delta M_g \neq 1 \mid \mathbf{M}_g\} \Pr\{\Delta A_g \neq 1 \mid \mathbf{A}_g\} \Pr\{\Delta G_g = 1 \mid \mathbf{G}_g\} \\ &= (1 - \Delta\hat{M}_g^{(1)})(1 - \Delta\hat{A}_g^{(1)})\Delta\hat{G}_g^{(1)}.\end{aligned} \tag{16}$$

This assumption will attenuate the power of our test statistics, as it forces the calculation of $\hat{1}_{\{\Delta G_g=1, \Delta M_g\neq 1, \Delta A_g\neq 1\}}$ toward the null hypothesis, and it will become clear later in our power analysis.

RESULTS

Data analysis

The results of the triple microarray data analysis are summarized in Table 1. The dependent model, our proposed method, allowed for a correlation between differentially expressed, methylated, and acetylated genes. The independent model assumed a zero correlation. For the gene expression data, ignoring the correlation resulted in a 30% underestimation of the overall probability of the genes with changed expression (from 0.434 when correlation was taken into account to 0.303 when it was not). On the other hand, both models produced similar results for the methylation and acetylation arrays, as the correlation parameters in both cases were close to zero. In summary, between control and the combined treatment of DAC and TSA, based on the dependent model, approximately 43.4% of genes were up-regulated or down-regulated, 72.3% of genes had altered methylation status, and 57.7% of genes had altered acetylation status. In addition, among genes with an altered expression level, 77% were up-regulated. Among genes with an altered methylation status, 60% were demethylated. And among genes with altered an acetylation status, 68% were hyperacetylated.

In order to assess the bias of the probability estimate for up-regulated or down-regulated genes when ignoring the correlation structure, we performed a simulation study. We chose the dependent model (1) as the true model and the parameter estimates from gene expression data as the true parameters. We generated 1,000 datasets, and each dataset contained 1,507 genes and two replicates per group. Its Monte Carlo estimate from the dependent model (averaged over 1,000 samples) was 0.441 ± 0.008, while its Monte Carlo estimate from the independent model was 0.313 ± 0.007. These results reinforce the importance of accounting for the correlation structure in the dependent model.

The between-gene variance is sometimes desirable, and in our situation, it includes the treatment effect. The more significantly a treatment alters gene expression, methylation, and acetylation, the larger the between-gene variance will be. On the other hand, it is important to control the within-replicates and between-replicates sampling errors, as lower sampling errors increase the chance of detecting differentially expressed, methylated and acetylated genes. In the gene-expression array, the variance among genes was $\hat{\sigma}^2_{G,2} = 1.343$, or about 6 to 10 times higher than the variance of the within and between-replicate sampling errors $\hat{\sigma}^2_{G,0} = 0.232$ and $\hat{\sigma}^2_{G,1} = 0.118$, respectively. In the methylation array, the variance among genes was $\hat{\sigma}^2_{M2} = 0.082$, about 1.5 to 2 times higher than the within and between-replicate sampling error $\hat{\sigma}^2_{M,0} = 0.057$ and $\hat{\sigma}^2_{M,1} = 0.044$, respectively. Similarly, in the acetylation array, the variance among

Table 1. Parameter Estimation from Mixture Models

	Gene expression model						
	p_G	γ_G	μ_G	$\sigma^2_{G,0}$	$\sigma^2_{G,1}$	$\sigma^2_{G,2}$	ρ_G
Independent model	0.303	0.657	3.060	0.255	0.114	1.292	0.000
Dependent model	0.434	0.655	3.046	0.232	0.118	1.343	0.494
	Gene methylation model						
	p_M	γ_M	μ_M	$\sigma^2_{M,0}$	$\sigma^2_{M,1}$	$\sigma^2_{M,2}$	ρ_M
Independent model	0.728	0.700	2.379	0.055	0.043	0.085	0.000
Dependent model	0.723	0.697	2.356	0.057	0.044	0.082	−0.153
	Gene acetylation model							
	p_A	γ_A	μ_A	$\sigma^2_{A,0}$	$\sigma^2_{A,1}$	$\sigma^2_{A,2}$	$\sigma^2_{A,a}$	ρ_A
Independent model	0.576	0.345	2.634	0.186	0.052	0.020	0.404	0.000
Dependent model	0.577	0.344	2.631	0.182	0.051	0.020	0.405	0.032

genes was $\hat{\sigma}^2_{A,2} = 0.405$, or about 2 to 8 times higher than the within and between-replicate sampling error $\hat{\sigma}^2_{A,0} = 0.182$ and $\hat{\sigma}^2_{A,1} = 0.051$, respectively. In addition, the variance of no-antibody and dye effects among all genes was $\hat{\sigma}^2_{A,a} = 0.02$ in the acetylation arrays.

To account for dye bias, log-Cy5 signals were 0.655 and 0.697 ($\hat{\gamma}_G, \hat{\gamma}_M$) higher than log-Cy3 in the expression and methylation arrays, respectively. On the other hand, $\hat{\gamma}_A = 0.344$ represented the average the no-antibody and dye effect among all genes in the acetylation arrays.

Figure 3 displays scatter plots for $(\Delta\hat{M}_g^{(1)}, \Delta\hat{A}_g^{(1)})$, after the genes were filtered by $\Delta\hat{G}_g^{(1)}$. Although the overall patterns were flat, the weighted least squares regressions showed an increasing trend. Particularly, the fitted line was $\Delta\hat{\hat{A}}_g^{(1)} = 0.37 + 0.164\Delta\hat{M}_g^{(1)}$ when all data were included in Fig. 3(a), where $\hat{\beta}_1 = 0.164 \pm 0.022$ (p-value < 0.00001 based on the standard normal distribution). Based on this, we reject the null hypothesis that DNA demethylation and histone hyperacetylation are independent among up-regulated genes.

In order to better understand the associations among gene expression, DNA demethylation, and histone hyperacetylation, contingency tables were constructed. Among genes with a large chance of up-regulation (Table 2a, b, c, and e), DNA demethylation and histone hyperacetylation status were positively correlated. Among genes with a small chance of up-regulation (Table 2d and f), there was no evidence of an association between DNA demethylation and histone hyperacetylation status. It is worthwhile to notice that in Table 2e, among 60 genes with a high chance of up-regulation ($\Delta\hat{G}^{(1)} \geq 0.8$), there was no strong evidence of both DNA demethylation and histone hyperacetylation change in 44 genes ($\Delta\hat{A}^{(1)} < 0.2$ and $\Delta\hat{M}^{(1)} < 0.2$). Biologically, while expression of these genes was not directly associated with epigenetic changes, these

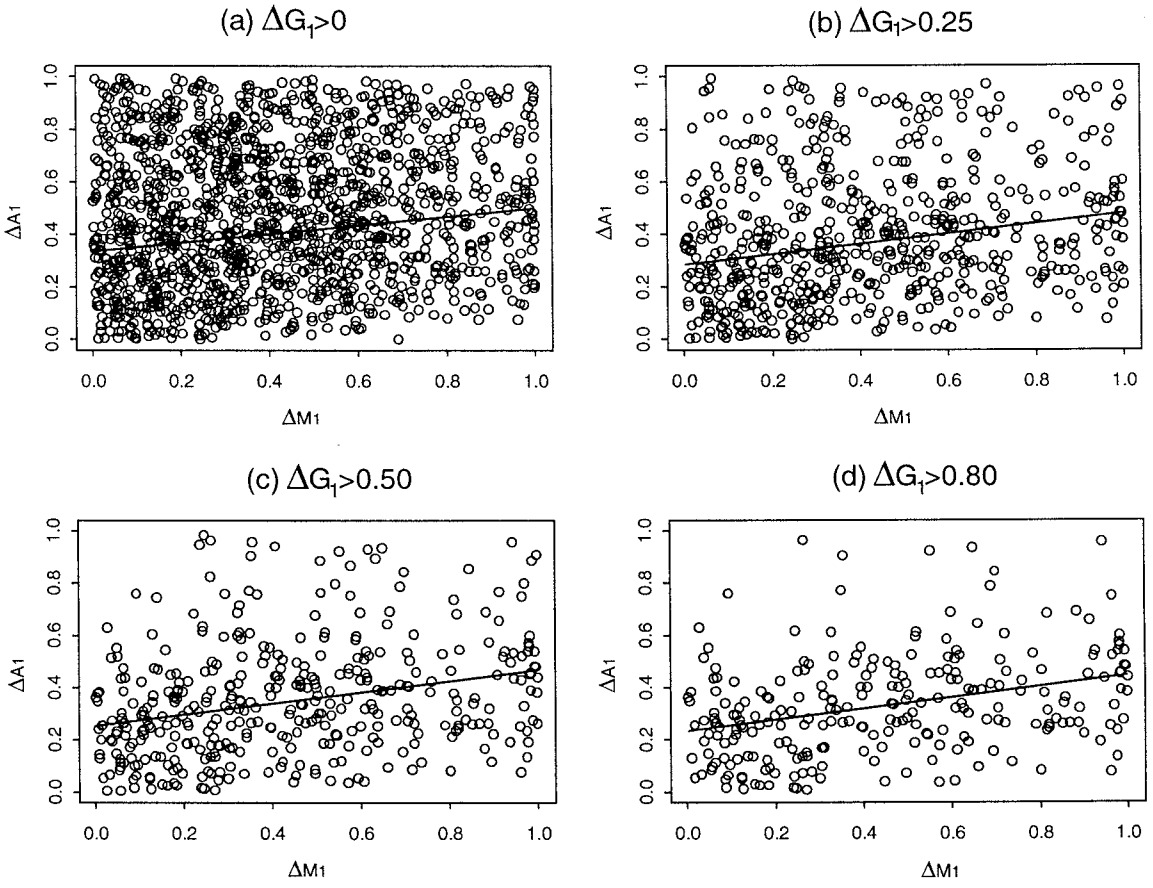

FIG. 3. Scatter plot of $(\Delta\hat{M}_g^{(1)}, \Delta\hat{A}_g^{(1)})$ (**a**) All genes; (b) genes with $\Delta G_g^{(1)} > 0.25$; (**c**) genes with $\Delta G_g^{(1)} > 0.50$; and (**d**) genes with $\Delta G_g^{(1)} > 0.80$.

Table 2. Contingency Tables of DNA-Methylation/Histone-Acetylation with Different Gene Expression Threshold

(a) $\Delta\hat{G}^{(1)} \geq 0$

	$\Delta\hat{A}^{(1)} \geq 0.5$	$\Delta\hat{A} < 0.5$
$\Delta\hat{M}^{(1)} \geq 0.5$	304	398
$\Delta\hat{M}^{(1)} < 0.5$	300	503

Chi-square p-value = 0.022

(b) $\Delta\hat{G}^{(1)} \geq 0.25$

	$\Delta\hat{A}^{(1)} \geq 0.5$	$\Delta\hat{A} < 0.5$
$\Delta\hat{M}^{(1)} \geq 0.5$	92	91
$\Delta\hat{M}^{(1)} < 0.5$	169	266

Chi-square p-value = 0.011

(c) $\Delta\hat{G}^{(1)} \geq 0.8$

	$\Delta\hat{A}^{(1)} \geq 0.5$	$\Delta\hat{A} < 0.5$
$\Delta\hat{M}^{(1)} \geq 0.5$	31	16
$\Delta\hat{M}^{(1)} < 0.5$	84	102

Chi-square p-value = 0.013

(d) $\Delta\hat{G}^{(1)} < 0.2$

	$\Delta\hat{A}^{(1)} \geq 0.5$	$\Delta\hat{A} < 0.5$
$\Delta\hat{M}^{(1)} \geq 0.5$	167	300
$\Delta\hat{M}^{(1)} < 0.5$	122	210

Chi-square p-value = 0.83

(e) $\Delta\hat{G}^{(1)} \geq 0.8$

	$\Delta\hat{A}^{(1)} \geq 0.8$	$\Delta\hat{A}^{(1)} < 0.2$
$\Delta\hat{M}^{(1)} \geq 0.8$	4	2
$\Delta\hat{M}^{(1)} < 0.2$	10	44

Chi-square p-value = 0.032

(f) $\Delta\hat{G}^{(1)} < 0.2$

	$\Delta\hat{A}^{(1)} \geq 0.8$	$\Delta\hat{A}^{(1)} < 0.2$
$\Delta\hat{M}^{(1)} \geq 0.8$	15	42
$\Delta\hat{M}^{(1)} < 0.2$	12	30

Chi-square p-value = 0.98

genes could be downstream targets of other genes controlled by DNA demethylation and histone hyperacetylation.

On the other hand, in Table 2f, among 99 genes with a small chance of up-regulation ($\Delta\hat{G}^{(1)} < 0.2$), there was some evidence of a change in either DNA demethylation or histone hyperacetylation for 69 genes ($\Delta\hat{A}^{(1)} \geq 0.8$ or $\Delta\hat{M}^{(1)} \geq 0.8$). Biologically, these 69 genes could be regulated by nonepigenetic mechanisms; therefore, a change in DNA methylation and histone acetylation status would not lead to up-regulation of these genes.

To consolidate our results, we performed three sets of bootstrap analyses. Our null hypothesis was that M_{ijkl} and A_{ijkl} were independent, regardless of G_{ijkl}. The first bootstrap analysis tried to address the effect of the normalization step on the follow-up data analysis. A bootstrap array contained 1,507 resampled genes (with replacement) from the raw data (before normalization). To preserve the treatment effect on gene expression, DNA demethylation, and histone hyperacetylation, the Cy3 and Cy5 signals for a given gene were drawn in pairs. A set of bootstrap samples (two replicates for each of three array types) had one bootstrap array generated from each array sample in our data, and these bootstrap samples were drawn independently. This independent resampling scheme ensured that gene expression, DNA demethylation, and histone hyperacetylation status were independent under the null hypothesis. In total, $k = 1, \ldots, 1{,}000$ sets of bootstrap samples, $\{\mathbf{G}_g^{(k)}, \mathbf{M}_g^{(k)}, \mathbf{A}_g^{(k)}\}_{g=1}^{1,507}$, were generated. Then, in each bootstrap sample, two replicates for each of three array types were normalized by the same Loess nonlinear method (Bolstad *et al.*, 2003). A mixture model (1) was fitted to the normalized expression and methylation arrays, and a mixture model (5) was fitted to the normalized acetylation array. The values $(\Delta\hat{G}_g^{(1)(k)}, \Delta\hat{M}_g^{(1)(k)}, \Delta\hat{A}_g^{(1)(k)})$ were estimated, a weighted least square regression was performed, and the ratio $\hat{\beta}_1^{(k)}/se(\hat{\beta}_1^{(k)})$ was calculated. Based on the empirical distribution of $\{\hat{\beta}_1^{(k)}/se(\hat{\beta}_1^{(k)})\}_{k=1,\ldots,1,000}$, the p-value was less than 0.001.

The main difference between the second and first bootstrap analysis was that the second analysis drew normalized instead of unnormalized signals. Similarly to the first analysis, $k = 1, \ldots, 1{,}000$ sets of normalized independent bootstrap samples, $\{\mathbf{G}_g^{(k)}, \mathbf{M}_g^{(k)}, \mathbf{A}_g^{(k)}\}_{g=1}^{1,507}$, were generated. This independent resampling scheme ensured that gene expression, DNA demethylation, and histone hyperacetylation status were independent under the null hypothesis. For each bootstrap sample, a mixture model (1) was fitted to expression and methylation arrays, and a mixture model (5) was fitted to the corresponding acetylation array. The

values $(\Delta\hat{G}_g^{(1)(k)}, \Delta\hat{M}_g^{(1)(k)}, \Delta\hat{A}_g^{(1)(k)})$ were estimated, a weighted least square regression was performed, and $\hat{\beta}_1^{(k)}/se(\hat{\beta}_1^{(k)})$ was calculated. Based on the empirical distribution of $\{\hat{\beta}_1^{(k)}/se(\hat{\beta}_1^{(k)})\}_{k=1,\ldots,1,000}$, the p-value was less than 0.001. By comparing this p-value to the one from the first analysis, we concluded that the normalization step had little if any effect on the follow-up data analysis.

Because the standard errors of the mixture model parameters in (3) and (7) were within 3% of their parameter estimates (not reported in Table 1), the parameter estimates were highly accurate with such a large sample size ($n = 1{,}507$). Therefore, in our third bootstrap analysis, we directly sampled ($n = 1{,}507$) $\Delta\hat{G}_g^{(1)(k)}$ from $\Delta\hat{G}_g^{(1)}$, $\Delta\hat{M}_g^{(1)(k)}$ from $\Delta\hat{M}_g^{(1)}$, and $\Delta\hat{A}_g^{(1)(k)}$ from $\Delta\hat{A}_g^{(1)}$ independently with replacement $k = 1, \ldots, 1{,}000$ times. This independent resampling scheme ensured that gene expression, DNA demethylation, and histone hyperacetylation status were independent under the null hypothesis. In each bootstrap sample, a weighted least squares regression was performed, and $\hat{\beta}_1^{(k)}/se(\hat{\beta}_1^{(k)})$ was calculated. The empirical distribution was formed $\{\hat{\beta}_1^{(k)}/se(\hat{\beta}_1^{(k)})\}$, and the p-value was the percent of $\{\hat{\beta}_1^{(k)}/se(\hat{\beta}_1^{(k)})\}$ replicates that were larger than the observed $\hat{\beta}_1/se(\hat{\beta}_1)$. The p-value was less than 0.001.

Therefore, all three bootstrap analyses confirmed the result using weighted least square analysis. By comparing bootstrap p-values between analysis one and two, we concluded that the normalization step had little effect on the follow-up data analysis. By comparing bootstrap p-values between analysis two and three, we concluded that the mixture model parameters in (3) and (7) were highly accurate such that their variances had little effect on the weighted least square inference.

The chi-square statistic $X = 0.357$ in the triple-array data (p-value $= 0.45$) was based on the chi-square distribution with one degree of freedom. However, as the approximation, $\hat{1}\{\Delta G_g = 1, \Delta M_g \neq 0, \Delta A_g \neq 0\} = \Pr\{\Delta M_g \neq 0 \mid \mathbf{M}_g\} \Pr\{\Delta A_g \neq 0 \mid \mathbf{A}_g\} \Pr\{\Delta G_g \neq 1 \mid \mathbf{G}_g\}$, may change the shape of the asymptotic distribution of X, we have performed three bootstrap analyses. These three bootstrap procedures were exactly the same as those in the weighted least squares method. In the first analysis, based on $k = 1{,}000$ sets of bootstrap samples from unnormalized data under the null hypothesis, $(\Delta\hat{G}_g^{(1)(k)}, \Delta\hat{M}_g^{(1)(k)}, \Delta\hat{A}_g^{(1)(k)})$ were estimated, and an empirical distribution of $\{X^{(k)}\}$ was generated. The p-value for the observed $X = 0.357$ was less than 0.001.

In the second bootstrap analysis, based on $k = 1{,}000$ sets of bootstrap samples from normalized data under the null hypothesis, $(\Delta\hat{G}_g^{(1)(k)}, \Delta\hat{M}_g^{(1)(k)}, \Delta\hat{A}_g^{(1)(k)})$ were estimated, and an empirical distribution of $\{X^{(k)}\}$ was generated. Again, the p-value for the observed $X = 0.357$ was less than 0.001.

In the third bootstrap analysis, 1,000 sets of $(\Delta\hat{G}_g^{(1)(k)}, \Delta\hat{M}_g^{(1)(k)}, \Delta\hat{A}_g^{(1)(k)})$ were directly sampled from $(\Delta\hat{G}_g^{(1)}, \Delta\hat{M}_g^{(1)}, \Delta\hat{A}_g^{(1)})$. An empirical distribution of $\{X^{(k)}\}$ was formed. The p-value for the observed $X = 0.357$ was less than 0.001.

The *conservative* nature of the chi-square statistic due to the independence assumption made in (16) was confirmed in three bootstrap analyses. The fact that the p-value from analysis one and two was the same suggested that the normalization step had little effect on the follow-up analysis. The equality of the p-values from analysis two and three suggested that the mixture model parameters in (3) and (7) were highly accurate such that their variances had little effect on the chi-square inference. Both bootstrap analysis one and two involved computationally intensive normalization or mixture model fitting steps. By contrast, the computational burden associated with the third bootstrap scheme was much lighter and is a viable replacement for the chi-square test that is of questionable validity in this context.

Power analysis

For an experimental design to test the epigenetic hypothesis, it is useful to have some statistical guidelines. The interesting issues are the relationships among the power, the number of genes, the number of replicates, and the size of the effect. In the power analysis, we adopted all the parameter estimates from the dependent model (Table 1). In addition, marginally, among genes with an altered expression level, 78% were set as up-regulated. Among genes with an altered methylation status, 60% were set as demethylated. And, among genes with altered acetylation status, 68% were set as hyperacetylated.

The odds of hyperacetylation were defined as the proportion of hyperacetylated genes over the proportion of nonhyperacetylated genes. In the experiment, when all of these genes were up-regulated, the odds of hyperacetylation in the demethylated genes were expected to be higher than for hyperacetylation in the methylated genes. If the genes were not up-regulated, they were assumed to have independent methylation

Step 1: $\Delta G_g \sim$ multi - nomial(0.339, 0.220, 0.441) for (1, 0, - 1), where $\Delta G_g = 1$ means gene up - regulation.

Step 2: If $\Delta G_g = 1$, then $\Delta M_g \sim$ multi - nomial(0.434, 0.400, 0.166) for (1,0,-1), where

$\Delta M_g = 1$ means DNA demethylation.

If $\Delta M_g = 1$, then $\Delta A_g \sim$ multi - nomial($p11$, $p12$, $p13$) for (1,0,-1),

where $\Delta A_g = 1$ means histone hyperactylation.

Else $\Delta A_g \sim$ multi - nomial($p21$, $p22$, $p23$) for (1,0,-1).

Else $\Delta M_g \sim$ multi - nomial(0.434, 0.400, 0.166) and

$\Delta A_g \sim$ multi - nomial(0.392, 0.320, 0.288).

In particular,

$$\begin{pmatrix} p11 & p12 & p13 \\ p21 & p22 & p23 \end{pmatrix} = \begin{pmatrix} 0.392, & 0.320, & 0.288 \\ 0.392, & 0.320, & 0.288 \end{pmatrix} \text{or} \begin{pmatrix} 0.605, & 0.208, & 0.187 \\ 0.230, & 0.405, & 0.365 \end{pmatrix} \text{or} \begin{pmatrix} 0.678, & 0.170, & 0.154 \\ 0.173, & 0.435, & 0.389 \end{pmatrix},$$

such that,

$$OR = \frac{P(\Delta M_g = 1, \Delta A_g = 1, \Delta G_g = 1) \times P(\Delta M_g \neq 1, \Delta A_g \neq 1, \Delta G_g = 1)}{P(\Delta M_g \neq 1, \Delta A_g = 1, \Delta G_g = 1) \times P(\Delta M_g = 1, \Delta A_g \neq 1, \Delta G_g = 1)} = 1, 5, 10 \text{ respectively.}$$

Step 3: Based on $(\Delta G_g, \Delta M_g, \Delta A_g)$, $\mathbf{G}_g$ and $\mathbf{M}_g$ are generated through model (1), and $\mathbf{A}_g$ is simulated through model (5) with parameters listed in Table 1.

FIG. 4. Simulation scheme.

and acetylation status. The odds ratio (OR) (Agresti, 1990) was employed to describe the association between methylation and acetylation status. The stronger the positive association, the larger the OR, while independence means OR = 1. In the power analysis, we used OR values of 1, 5, and 10 for the hyperacetylation and demethylation association among the up-regulated genes.

Gene expression, DNA methylation, and histone acetylation signals were simulated according to a scheme described in Fig. 4. In each simulated dataset, triple array signals for 500 or 1,500 genes were generated with either two or four replicates. In total, 1,000 simulated datasets were generated. A one-sided p-value under 0.05 was claimed to be statistically significant. The p-value for a bootstrap analysis was based on its empirical distribution. The bootstrap procedure was the same as the third bootstrap analysis outlined in the data analysis section. The results of the power analysis are shown in Table 3. First, the more genes or the more replicates, the higher the power. Second, the normal test in the weighted least squares regression had inflated type I errors ($\geq$0.05), which were calculated as the power when OR = 1. Hence, its higher power was biased compared to the weighted least squares bootstrap analysis. Third, while both methods maintained a 5% type I error rate, the weighted least squares bootstrap analysis was more powerful than the chi-square bootstrap analysis. For a dataset of 1,500 genes with two replicates for each array, the bootstrap WLS generated approximately 42% power to detect an OR of 5 for the correlation of the demethylation and hyperacetylation among up-regulated genes.

Sensitivity analysis

Model (1) assumes that $\mu_{G,jg}$, the log-transformed true expression for gene g, follows a normal distribution. However, according to Figs. 2(a) and (b), the distribution of gene expression is heavily skewed to the right, even after log-transformation. This is also true for both the methylation and acetylation arrays [Figs. 2(c), (d), (e), (f), (g), and (h)]. Newton *et al.* (2003) recognized this situation and applied a

TABLE 3. POWER ANALYSIS[a]

		WLS						*Chi-square*		
		Normal[b] *(OR)*			*Bootstrap*[c] *(OR)*			*Bootstrap*[c] *(OR)*		
# Gene	*# Replicate*	*1*	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*
1500	4	0.132	0.743	0.902	0.050	0.582	0.845	0.052	0.168	0.464
1500	2	0.114	0.612	0.865	0.053	0.423	0.694	0.050	0.123	0.227
500	4	0.135	0.573	0.797	0.051	0.349	0.575	0.049	0.101	0.203
500	2	0.123	0.432	0.590	0.050	0.153	0.332	0.051	0.068	0.112

[a]In the simulation, data are generated through models (1)–(3).
[b]Hypothesis tests and power calculations are based on a standard normal distribution from a weighted least-square regression.
[c]Hypothesis tests and power calculations are based on a bootstrap analysis.

nonparametric distribution instead of a normal or a gamma distribution. They demonstrated that a more flexible model leads to a better posterior probability calculation for gene selection. As our goal in this paper is an epigenetic hypothesis test, we would like to know whether the test is sensitive to the violation of the normality assumption.

Figures 5(a), (c), and (e) display the between-replicates variation for gene expression, DNA methylation, and histone acetylation signals, respectively. These features, as reported by many researchers (Baggerly *et al.*, 2001; Efron *et al.*, 2001; Roche and Durbin 2001; Tusher *et al.*, 2001), are not constant, suggesting that measurement errors for high-intensity gene expression follow an approximately constant coefficient

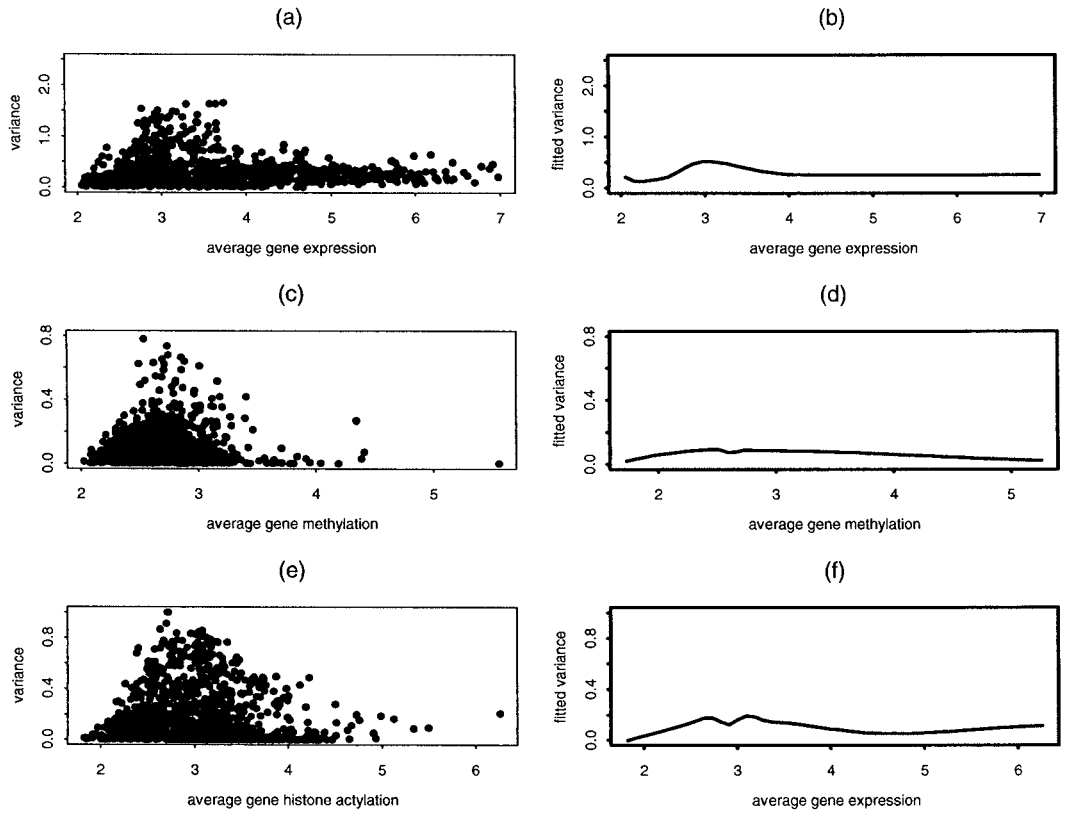

FIG. 5. Between replicate variances of three arrays: (**a**), (**c**), and (**e**) are between replicates variance for the gene expression, DNA methylation , and histone acetylation respectively, and (**b**), (**d**), and (**f**) are their loess fits respectively.

TABLE 4. SENSITIVITY ANALYSIS[a]

		WLS						*Chi-square*		
		Normal[b] *(OR)*			*Bootstrap*[c] *(OR)*			*Bootstrap*[c] *(OR)*		
# Gene	*# Replicate*	*1*	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*
1500	4	0.122	0.675	0.849	0.054	0.531	0.752	0.052	0.134	0.443
1500	2	0.115	0.560	0.715	0.055	0.374	0.603	0.047	0.076	0.204
500	4	0.127	0.456	0.634	0.049	0.274	0.473	0.046	0.079	0.189
500	2	0.133	0.325	0.543	0.052	0.103	0.305	0.050	0.066	0.094

[a]In the simulation, log-transformed signal is generated through a gamma distribution, and between replicate variance follows Equation (13).
[b]Hypothesis tests and power calculations are based on a standard normal distribution from a weighted least square regression.
[c]Hypothesis tests and power calculations are based on a bootstrap analysis.

of variation (CV), and the variation of log-transformed low signals is usually higher than that of log-transformed high signals. As illustrated in Figs. 5(b), (d), and (f), we observed a slightly different pattern compared to what has previously been reported. The variance of the lowest log-transformed signals was less than the median-level signals, though the variance of the highest signals remained low. This was mainly due to the initial normalization procedure, in which the negative signals were all set as a small positive number. This procedure inevitably forced some low-signal genes to have very small variance. The variance shapes of the three arrays were very similar. Furthermore, when the log-transformed signals were between 2 and 4 but otherwise approximately constant, the variance shapes followed a bell shape. Therefore, it was critical to know whether our tests were sensitive to the inequality in the variances.

To perform a sensitivity analysis, data were simulated to accommodate the above two features and the follow-up weighted least squares and chi-square tests were based on the mis-specified normality and equal variance assumptions. If our tests were robust, they would have comparable power regardless of the assumption violations. Therefore, a gamma distribution was employed to simulate the $(\mu_{G,jg}, \mu_{M,jg}, \mu_{A,jg})$ for each gene. This distribution had the same mean and variance as our initial normal random variable but was skewed unlike a normal distribution. In addition, the between-replicate variance was simulated as follows:

$$\sigma_1^2(\mu) = \begin{cases} -a + b & \text{otherwise,} \\ -a(\mu - 3)^2 + b & 2 \le \mu \le 4, \end{cases} \tag{17}$$

where μ represents the true log-transformed signal for log-transformed gene expression, methylation, or acetylation level. In particular, $(a, b) = (0.22, 0.31)$ for an expression array, $(0.07, 0.11)$ for a methylation array, and $(0.07, 0.10)$ for an acetylation array.

Gene expression, DNA methylation, and histone acetylation signals were simulated according to the same scheme described in Fig. 4, except that step 3 employed a gamma model and used the between-replicate variances described by (17). In each simulated dataset, 500 or 1,500 triple array gene signals were generated with either two or four replicates, and a total of 1,000 simulated datasets were generated.

The power of our proposed tests are listed in Table 4. The violations of those two assumptions only slightly reduced the power of the three tests. In particular, the type I error rate of the WLS bootstrap analysis was maintained at the 5% level.

DISCUSSION

Our empirical Bayes model (1) and (5) extend the log-normal–normal model of Kendziorski *et al.* (2003), the ANOVA model of Kerr *et al.* (2000), and the two-stage ANOVA model of Wolfinger *et al.* (2001). Furthermore, our model provides a flexible way to describe each individual microarray type of the

triple-microarray system. Based on this framework, by an EM algorithm between the group treated with TSA and DAC and the control, approximately 43.4% of genes were differentially expressed (up-regulated or down-regulated), a change in the methylation status was observed for 72.3% of the genes, and the acetylation status of 57.6% of the genes was also changed . In particular, we demonstrated the importance of accounting for the correlation structure in our proposed model (1). Furthermore, we show that ignoring the correlation leads to a 30% underestimation of the overall probability of differentially expressed genes (a reduction from 0.434 when correlation is taken into account to 0.303 when it is not). This has been verified through Monte Carlo simulations.

Two methods have been developed to test our null hypothesis, i.e., that promoter hypermethylation and histone hypoacetylation status are independent among up-regulated genes. One approach follows a weighted least squares regression and the other a chi-square statistic. Both methods provided strong evidence to reject the null hypothesis. Bootstrap analyses subsequently verified these results.

Based on the results of the power analysis, we showed that the WLS bootstrap approach is more powerful than the chi-square bootstrap methodology. The type-I error rate of a normal test in WLS is inflated and so is its power. Sensitivity analyses demonstrated that our proposed tests are highly robust to the violation of normality or equal variance assumptions.

APPENDIX A

Expression data was used to demonstrate our EM algorithm. The complete data log-likelihood function is given in Equation (A.1)

$$l_G(\boldsymbol{\theta}) = \sum_g z_{G,g}\{\log[f_1(\mathbf{G}_g)] + p_G\} + (1 - z_{G,g})\{\log[f_0(\mathbf{G}_g)] + 1 - p_G\} \tag{A.1}$$

where $\boldsymbol{\theta} = (\mu_G, \sigma^2_{G,0}, \sigma^2_{G,1}, \sigma^2_{G,2}, \rho_G)$.

In the M-step, for given $\hat{p}_G$ and $\{\hat{z}_{G,g}\}$, $l_G(\boldsymbol{\theta})$ is maximized with an Splus function **nlminb**. In E-step, based on current $\hat{\boldsymbol{\theta}}$,

$$\hat{z}_{G,g} = \frac{\hat{p}_G f_1(\mathbf{G}_g)}{\hat{p}_G f_1(\mathbf{G}_g) + (1 - \hat{p}_G) f_0(\mathbf{G}_g)} \quad \text{and} \quad \hat{p}_G = \frac{\sum_{g=1,\ldots,n} \hat{z}_{G,g}}{n}. \tag{A.2}$$

The EM algorithm for methylation and acetylation arrays follows similar procedures.

APPENDIX B

As ΔG_g was not observed from the expression arrays, it was necessary to estimate it from the data.

$$\begin{aligned} \Pr\{\Delta G_g = 1 \mid \mathbf{G}_g\} &= \Pr\{(z_{G,g} = 1) \cap (\mu_{G,1g} > \mu_{G,2g}) \mid \mathbf{G}_g\} \\ &= \Pr\{\mu_{G,1g} > \mu_{G,2g} \mid z_{G,g} = 1, \mathbf{G}_g) \Pr(z_{G,g} = 1 \mid \mathbf{G}_g) \end{aligned} \tag{B.1}$$

The second term of the right-hand side of Equation (B.1) is the probability of different expression levels for gene g between treatment and control given the observed data.

$$\Pr(z_{G,g} = 1 \mid \mathbf{G}_g) = \frac{p_G f_1(\mathbf{G}_g)}{p_G f_1(\mathbf{G}_g) + (1 - p_G) f_0(\mathbf{G}_g)} = \hat{z}_{G,g}, \tag{B.2}$$

where $f_1(\mathbf{G}_g)$ and $f_0(\mathbf{G}_g)$ are defined in Equation (3).

The first term of the right-hand side of Equation (B.1) is the posterior probability of up-regulation for gene g due to the treatment given that gene expression levels are different between treatment and control and observed data. The posterior mean and covariance of $(\mu_{G,1g}, \mu_{G,2g})$ are

$$E(\mu_{G,1g} \mid z_{G,g} = 1, \mathbf{G}_g) = \frac{\dfrac{G_{111g} + G_{212g}}{\sigma_{G,0}^2 + \sigma_{G,1}^2} + \dfrac{\mu_G}{\sigma_{G,2}^2}}{\dfrac{2}{\sigma_{G,0}^2 + \sigma_{G,1}^2} + \dfrac{1}{\sigma_{G,2}^2}},$$

$$E(\mu_{G,2g} \mid z_{G,g} = 1, \mathbf{G}_g) = \frac{\dfrac{G_{122g} + G_{221g}}{\sigma_{G,0}^2 + \sigma_{G,1}^2} + \dfrac{\mu_G}{\sigma_{G,2}^2}}{\dfrac{2}{\sigma_{G,0}^2 + \sigma_{G,1}^2} + \dfrac{1}{\sigma_{G,2}^2}},$$

$$\mathrm{var}(\mu_{G,1g} \mid z_{G,g} = 1, \mathbf{G}_g) = \mathrm{var}(\mu_{G,2g} \mid z_{G,g} = 1, \mathbf{G}_g) = \frac{1}{\dfrac{2}{\sigma_{G,0}^2 + \sigma_{G,1}^2} + \dfrac{1}{\sigma_{G,2}^2}},$$

$$\mathrm{cov}(\mu_{G,1g}, \mu_{G,2g} \mid z_{G,g} = 1, \mathbf{G}_g) = \frac{\dfrac{2\sigma_{G,1}^2}{(\sigma_{G,0}^2 + \sigma_{G,1}^2)^2} + \dfrac{\rho_G}{\sigma_{G,2}^2}}{\left(\dfrac{2}{\sigma_{G,0}^2 + \sigma_{G,1}^2} + \dfrac{1}{\sigma_{G,2}^2}\right)^2}.$$

The posterior mean and variance of $(\mu_{G,1g} - \mu_{G,2g})$ are

$$E(\mu_{G,1g} - \mu_{G,2g} \mid z_{G,g} = 1, \mathbf{G}_g) = \frac{\dfrac{G_{111g} + G_{212g}}{\sigma_{G,0}^2 + \sigma_{G,1}^2} - \dfrac{G_{122g} + G_{221g}}{\sigma_{G,0}^2 + \sigma_{G,1}^2}}{\dfrac{2}{\sigma_{G,0}^2 + \sigma_{G,1}^2} + \dfrac{1}{\sigma_{G,2}^2}},$$

$$\mathrm{var}(\mu_{G,1g} - \mu_{G,2g} \mid z_{G,g} = 1, \mathbf{G}_g) = \frac{\dfrac{4\sigma_{G,0}^2}{(\sigma_{G,0}^2 + \sigma_{G,1}^2)^2} + \dfrac{2(1-\rho_G)}{\sigma_{G,2}^2}}{\left(\dfrac{2}{\sigma_{G,0}^2 + \sigma_{G,1}^2} + \dfrac{1}{\sigma_{G,2}^2}\right)^2}.$$

Hence, the posterior probability can be written as

$$\Pr\{\mu_{G,1g} > \mu_{G,1g} \mid z_{G,g} = 1, \mathbf{G}_g\} = \Phi\left[\frac{\dfrac{G_{111g} + G_{212g}}{\sigma_{G,0}^2 + \sigma_{G,1}^2} - \dfrac{G_{122g} + G_{221g}}{\sigma_{G,0}^2 + \sigma_{G,1}^2}}{\sqrt{\dfrac{4\sigma_{G,0}^2}{(\sigma_{G,0}^2 + \sigma_{G,1}^2)^2} + \dfrac{2(1-\rho_G)}{\sigma_{G,2}^2}}}\right],$$

where $\Phi(\bullet)$ is the standard normal distribution function.

APPENDIX C

As ΔA_g is not observed from the expression arrays, it is necessary to be estimated from the data.

$$\begin{aligned}\hat{\Pr}\{\Delta A_g = 1 \mid \mathbf{A}_g\} &= \Pr\{(z_{A,g} = 1) \cap (\mu_{A,1g} > \mu_{A,2g}) \mid \mathbf{A}_g\} \\ &= \Pr\{\mu_{A,1g} > \mu_{A,2g} \mid z_{G,g} = 1, \mathbf{A}_g\} \Pr(z_{G,g} = 1 \mid \mathbf{A}_g)\end{aligned} \tag{C.1}$$

The second term of the right-hand side of Equation (C.1) is the probability of differential expression levels for gene g between treatment and control given the observed data.

$$\Pr(z_{A,g} = 1 \mid \mathbf{A}_g) = \frac{p_G f_1(\mathbf{A}_g)}{p_A f_1(\mathbf{A}_g) + (1 - p_A) f_0(\mathbf{A}_g)} = \hat{z}_{A,g}, \tag{C.2}$$

where $f_1(\mathbf{A}_g)$ and $f_0(\mathbf{A}_g)$ are defined in Equation (7).

The first term of the right-hand side of Equation (C.1) is the posterior probability of up-regulation for gene g due to the treatment, given that the expression levels of that gene are different between the treatment and control groups and the observed data. The posterior mean and covariance of $(\mu_{A,1g}, \mu_{A,2g})$ are

$$E(\mu_{A,1g} \mid z_{A,g} = 1, \mathbf{A}_g) = \frac{\dfrac{A_{111g} - A_{112g} + A_{211g} - A_{212g}}{\sigma^2_{A,0}} + \dfrac{\mu_A}{\sigma^2_{A,2}}}{\dfrac{2}{\sigma^2_{A,0}} + \dfrac{1}{\sigma^2_{A,2}}},$$

$$E(\mu_{A,2g} \mid z_{G,g} = 1, \mathbf{A}_g) = \frac{\dfrac{A_{121g} - A_{122g} + A_{221g} - A_{222g}}{\sigma^2_{A,0}} + \dfrac{\mu_A}{\sigma^2_{A,2}}}{\dfrac{2}{\sigma^2_{A,0}} + \dfrac{1}{\sigma^2_{A,2}}},$$

$$\operatorname{var}(\mu_{A,1g} \mid z_{A,g} = 1, \mathbf{A}_g) = \operatorname{var}(\mu_{A,2g} \mid z_{A,g} = 1, \mathbf{A}_g) = \frac{1}{\dfrac{2}{\sigma^2_{A,0}} + \dfrac{1}{\sigma^2_{A,2}}},$$

$$\operatorname{cov}(\mu_{A,1g}, \mu_{A,2g} \mid z_{A,g} = 1, \mathbf{A}_g) = \frac{\dfrac{\rho_A}{\sigma^2_{A,2}}}{\left(\dfrac{2}{\sigma^2_{A,0}} + \dfrac{1}{\sigma^2_{A,2}}\right)^2}.$$

The posterior mean and variance of $(\mu_{A,1g} - \mu_{A,2g})$ are

$$E(\mu_{A,1g} - \mu_{A,2g} \mid z_{A,g} = 1, \mathbf{A}_g) = \frac{\dfrac{A_{111g} - A_{112g} + A_{211g} - A_{212g} - (A_{121g} - A_{122g} + A_{221g} - A_{222g})}{\sigma^2_{A,0}}}{\dfrac{2}{\sigma^2_{A,0}} + \dfrac{1}{\sigma^2_{A,2}}},$$

$$\operatorname{var}(\mu_{A,1g} - \mu_{A,2g} \mid z_{A,g} = 1, \mathbf{A}_g) = \frac{\dfrac{4}{\sigma^2_{A,0}} + \dfrac{2(1 - \rho_A)}{\sigma^2_{A,2}}}{\left(\dfrac{2}{\sigma^2_{A,0}} + \dfrac{1}{\sigma^2_{A,2}}\right)^2}.$$

Hence, the posterior probability can be written as

$$\Pr(\mu_{A,1g} > \mu_{A,2g} \mid z_{A,g} = 1, \mathbf{A}_g) = \Phi\left[\frac{\dfrac{A_{111g} - A_{112g} + A_{211g} - A_{212g} - (A_{121g} - A_{122g} + A_{221g} - A_{222g})}{\sigma^2_{A,0}}}{\sqrt{\dfrac{4}{\sigma^2_{A,0}} + \dfrac{2(1-\rho_A)}{\sigma^2_{A,2}}}}\right]$$

where $\Phi(\bullet)$ is the standard normal distribution function.

ACKNOWLEDGMENTS

This work was supported in part by National Cancer Institute Grants CA-85289 (to K.P.N.), CA-82709 (to C.T.Y.), and CA-69065 (to T.H.-M.H.).

REFERENCES

Agresti A. 1990. *Categorical Data Analysis*, John Wiley, New York.

Baggerly, K.A., Coombes, K.R., Hess, K.R., Stivers, D.N., Abruzzo, L.V., and Zhang, W. 2001. Identifying differentially expressed genes in cDNA microarray experiments. *J. Comp. Biol.* 8, 639–659.

Bolstad, B.A., Irrizarry, R.A., Astrand, M., and Speed, T.P. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19, 185–193.

Cameron, E.E., Bachman, K.E., Myohanen, S., Herman, J.G., and Baylin, S.B. 1999. Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nature Genet.* 21, 103–107.

Chen, Y., Dougherty, E.R., and Bittner, M.L. 1997. Ratio-based decisions and the quantitative analysis of cDNA microarray images. *J. Biomed. Optics* 2, 364–374.

Churchill, G. A. 2002. Fundamentals of experimental design for cDNA microarrays. *Nature Genet. Suppl.* 32.

Dudoit, S, Yang, Y.H., Speed, T.P., and Callow, M.J. 2002. Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Stat. Sin.* 12, 111–139.

Efron, B., Tibshirani, R., Storey, J.D., and Tusher, V. 2001. Empirical Bayes analysis of a microarray experiment. *J. Am. Statist. Assoc.* 96, 1151–1160.

Jones, P.A., and Baylin, S.B. 2002. The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* 3, 415–428.

Kendziorski, C.M., Newton, M.A., Lan, H., and Gould, M.N. 2003. On parametric empirical Bayes methods for comparing multiple groups using replicated gene expression profiles. *Statist. Med.* In press.

Kerr, M.K., Martin M., and Churchill, G.A. 2000. Analysis of variance for gene expression microarray data. *J. Comp. Biol.* 7, 819–837.

Newton, M.A., Kendziorski, C.M., Richmond, C.S., Blattner, F.R., and Tsui, K.W. 2001. On differential variability of expression ratio: Improving statistical inference about gene expression changes from microarray data. *J. Comp. Biol.* 8, 37–52.

Newton, M.A., Noueiry, A., Sarkar, D., and Ahlquist, P. 2003. Detecting differential gene expression with a semiparametric hierarchical mixture method. Manuscript.

Ren, B., Robert, F., Wyrick, J.J., Aparicio, O., Jennings, E.G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin E., Volkert, T.L., Wilson, C.J., Bell, S.P., and Young, R.A. 2000. Genome-wide location and function of DNA binding proteins. *Science* 290, 2306–2309.

Roche, D.M., and Durbin, B. 2001. A model for measurement error for gene expression arrays. *J. Comp. Biol.* 8, 557–570.

Shi, H., Wei, S., Leu, Y., Rahmatpanah, F., Liu, J., Yan, P., Nephew, K.P., and Huang, T.H. 2003. Triple analysis of the cancer epigenome: An integrated microarray system for assessing gene expression, DNA methylation, and histone acetylation. *Cancer Res.* 63:2164–2171.

Shi, H., Yan, P.S., Chen, C.M., Rahmatpanah, F., Lofton-Day, C., Caldwell, C.W., and Huang, T.H-M. 2002. Expressed CpG island sequence tag microarray for dual screening of DNA hypermethylation and gene silencing in cancer cells. *Cancer Res.* 62, 3214–3220.

Suzuki, H., Gabrielson, E., Chen, W., Anbazhagan, R., van England, M., Weijenberg, M.P., Herman, J.G., and Baylin, S.B. 2002. A genomic screen for genes up-regulated by demethylation and histone deacetylase inhibition in human colorectal cancer. *Nature Genet*. 31, 141–149.

Tusher, V., Tibshirani, R., and Chu, G. 2001. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci*. 98, 5116–5121.

Weinberg, S. 1985. *Applied Linear Regression*, 2nd ed., John Wiley, New York.

Weinmann, A.S., Yan, P.S., Oberley, M.J., Huang, T.H-M., and Farnham, P.J. 2002. Isolating human transcription factor targets by coupling chromatin immunoprecipition and CpG island microarray analysis. *Genes Dev.* 16, 235–244.

Wolfinger, R.D., Gibson, G., Wolfinger, E.D., Bennett, L., Hamadeh, H., Bushel, P., Afshari, C., and Paules, R.S. 2001. Assessing gene significance from cDNA microarray expression data via mixed models. *J. Comp. Biol*. 8, 625–637.

Yan, P.S., Chen, C.-M., Shi, H., Rahmatpanah, F., Wei, S.H., Caldwell, C.W., and Huang, T.H-M. 2001. Dissecting complex epigenetic alternations in breast cancer using CpG island microarrays. *Cancer Res*. 61, 8375–8380.

Address correspondence to:
Lang Li
Department of Medicine
Division of Biostatistics
1050 Wishard Blvd., RG4101
Indianapolis, IN 46022-2872

E-mail: lali@iupui.edu