Histone deacetylase inhibitors (HDACI) are anticancer agents that inhibit the activity of enzymes that acetylate histone tails and other proteins. Acetylation is thought to play a role in cancer development and progression (1). Several HDACIs are in clinical development for solid malignancies. In preclinical studies, these agents have shown antitumor activity in vitro and in vivo in prostate cancer models. Phase I and II clinical trials with HDACI resulted in stable disease and partial responses in a wide range of solid malignancies, including prostate cancer (2, 3). Recently, the Food and Drug Administration approved HDACI vorinostat for the treatment of advanced cutaneous T-cell lymphoma and several clinical trials are ongoing in solid malignancies. Since its initial success as a tool for quantitative transcription level analysis in 1995 (4), microarray technology has become instrumental for global and parallel analysis of cellular biology (5). Although microarray technology has been rapidly adopted by the scientific community, experimental design and interpretation of the resulting data remain challenging. The choice of experimental design can be crucial in determining the extent to which biological and technical variation affect the results of a microarray experiment. Commonly used experimental designs in dual-color arrays include the "loop design" and "reference design." In the loop design, all samples are hybridized in such a way that the connected arrays form a loop pattern (1-2, 2-3, 3-4, and 4-1). In a reference design, all samples of interest are cohybridized with the same reference sample (1-R, 2-R, and 3-R; ref. 6).

In this article, we report results of a microarray experiment conducted using a more complex design as a case study to increase understanding of the technical aspects in experimental design and analysis of high-throughput gene expression studies. Before our microarray experiment, we tested the sensitivity of prostate cancer cells to HDACIs. PC3 cells were more resistant than DU145 cells to HDACI treatment. To improve understanding of the mechanism underlying this difference in resistance, we designed an experiment to simultaneously compare global gene expression in 16 different biological samples with 22 arrays. To determine the relative contribution of three experimental...
variables of interest on gene expression, all samples were hybridized according to a tailor-made “multidimensional loop” design. Independent comparison of all 16 samples was made possible with an adjusted ANOVA model for each feature on the array. This article discusses the pros and cons of our microarray design and its feasibility for other researchers.

Materials and Methods

Cell culture and chemicals

The following human prostate cancer cell lines (e.g., LNCaP, DU145, PC3, and CWR22R) were obtained from the American Type Culture Collection and cultured in RPMI 1640 (Life Technologies) supplemented with 10% fetal bovine serum. Vorinostat [also known as suberoylanilide hydroxamic acid (SAHA)] was obtained from Aton Pharma/Merck and valproic acid (VPA) and trichostatin A (TSA) were obtained from Sigma-Aldrich.

Proliferation assays

DU145 and PC3 cells were plated in 96-well plates at 4,000 cells per well and CWR22R and LNCaP cells were plated at 10,000 cells per well. Per treatment arm, 8 wells were used. All cells were allowed to adhere for 24 h before treatment, except LNCaP, which was allowed to adhere for 48 h. At the start of the experiment, freshly prepared drug was added to each well and incubated for 48 h. The proliferative activity was quantified using the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) kit (Promega; Cell Titer 96 Non-Radioactive Cell Proliferation Assay) which is based on reduction of tetrazolium salt MTT to colored formazan by mitochondrial enzymes present only in living, metabolically active cells (7). In congruence with the manufacturers’ description, cells were incubated for 4 h with the dye solution after treatment with HDACi followed by overnight incubation with the stop solution. To acquire a uniform color, a shaker was used. The absorbance at 570 nm was read and background was subtracted using the 650 nm reading and a blank reading. P values were calculated using the two-sample t test assuming equal variances (Student’s t test). All MTT assays were conducted in triplicate.

Microarray experiment

Experimental design. To increase our insight in the mechanism of action of HDACis on prostate cancer cells and to gain insight in possible resistance mechanisms against these agents, we compared the transcription levels of PC3 and DU145 cell lines after treatment with vorinostat and VPA. For the microarray experiment, 1 μmol/L vorinostat and 1 mmol/L VPA were chosen as the treatment doses. It has been shown that these doses are achievable in humans. Forty-eight or 96 h treatment with either 1 μmol/L vorinostat or 1 mmol/L VPA in both DU145 and PC3 cell lines were compared.

In total, we compared 16 conditions with 22 arrays. We used dual-color microarrays to take advantage of the ability to hybridize two samples per array. We preferred direct comparison on the more important questions while also allowing answering additional questions via indirect comparison. Some samples were hybridized three or four times to facilitate the comparisons of interest. In a loop design, a larger number of edges between samples is associated with higher variance in the corresponding comparison (6). No assumptions were made concerning (in)dependence of our experimental conditions. Taking into account dye swaps and technical replicates, the microarray setup resulted in a double cube. The directed graph (8) in Fig. 2B gives a schematic overview of the experimental set up. Each arrow represents one microarray slide comparing expression between two RNAs and is considered a block in the experimental layout (6). The sample on the tail side of the arrow is labeled with Cy5 (red). The sample on the head side of the arrow is labeled with Cy3 (green). The variables of interest included cell line (DU145 and PC3), HDACi (vorinostat and VPA), and duration of treatment (48 and 96 h). Control samples within each cube use the solvent specific for the HDACi. All samples are labeled and hybridized at least once with Cy3 and once with Cy5 to correct for dye effects.

Cell culture. On day 0, DU145 and PC3 cells (low passage) were trypsinized at ~85% confluency and replated at 50,000 cells per 100 mm dish. Each treatment was applied in triplicate. Aliquots for each drug concentration and time point were prepared in RPMI 1640 on day 0 and stored in -20°C for use in the following 4 days. Fetal bovine serum was aliquoted separately and stored in -20°C. The concentration of DMSO for the vorinostat aliquots was 0.05%.

On day 1, the medium was removed and replaced with the above-described (drug-containing) medium. The following 4 days, the medium was carefully aspirated each day and replaced with fresh (drug-containing) medium. After 48 and 96 h, dishes were harvested on ice using scrapers and 1 mL TRIzol reagent (Invitrogen) per dish and the pooled samples were stored in -80°C until RNA isolation at the end of the experiment.

RNA isolation. Total RNA was isolated from DU145 and PC3 cells using a combination of TRIzol and RNeasy Kit (Qiagen) with slight modifications to manufacturer’s protocol. Briefly, cells were harvested in TRIzol. After addition of chloroform and centrifugation, the aqueous phase was separated. This was ~50% of the initial TRIzol volume. Next, an equal volume of 70% ethanol was added to the supernatant. Because the final alcohol concentration now was only 30%, the RNA did not precipitate. The mixture was then passed through RNeasy column and further purification carried out according to the manufacturer’s instructions. A DNase step was also included to ensure removal of any contaminating DNA.

RNA quality assessment. Total RNA of all samples was quantified by the Johns Hopkins Sidney Kimmel Comprehensive Cancer Center Microarray Core Facility with NanoDrop ND-100 followed by quality assessment with 2100 Bioanalyzer (Agilent Technologies) according to manufacturer’s protocol. rRNA ratio (28S/18S) was 1.7 to 2.2 and RNA integrity number value was 9.4 to 9.9.

Sample amplification and labeling. Sample amplification and labeling procedures were carried out by the Sidney Kimmel Comprehensive Cancer Center Microarray Core facility using Low RNA
Input Fluorescent Linear Amplification Kit (Agilent Technologies). Total RNA (1 μg) was reverse transcribed into first- and second-strand cDNA by MMLV-RT using an oligo(dT) primer that incorporates a T7 promoter sequence. The cDNA was then used as a template for in vitro transcription in the presence of T7 RNA polymerase and cyanine-labeled CTPs. The labeled cRNA was purified using RNeasy mini kit (Qiagen) and followed by quantification on both concentrations of cRNA and dye labeled. RNA spike-in controls (Agilent Technologies) were added to the RNA samples before amplification and labeled according to manufacturer’s protocol. In total, 0.75 μg of each sample labeled with Cy3 or Cy5 was mixed with control targets (Agilent Technologies).

Target fragmentation and microarray hybridization. Because the size of our microarray experiment was too large to process on one day, it was decided to split the experiment into manageable portions. For target fragmentation, hybridization, and scanning, the experiment was split evenly. Fragmentation was carried out by the Sidney Kimmel Comprehensive Cancer Center Microarray Core Facility by incubating at 60°C for 30 min and stopped by adding equal volume of 2× hybridization buffer (Agilent Technologies). Fragmented targets were added on 22 human 1Av2 Microarrays (Agilent Technologies) all from the same batch. Agilent’s Human 1Av2 microarray consists of 21,073 60-mer probes printed at random on a glass slide. In total, there are 1,000 replicates (10 × 100 gene targets) randomly printed on the array. In addition, 1,080 control spots are printed on the array. The microarrays were assembled into a hybridization chamber (Agilent Technologies) and hybridized at 60°C for 17 h in a hybridization oven with rotation. Hybridized microarrays were washed and dried according to the Agilent microarray processing protocol.

Microarray scanning and data acquisition. Microarrays were scanned by the Sidney Kimmel Comprehensive Cancer Center Microarray Core Facility with the Agilent G2565BA microarray scanner under default settings recommended by Agilent Technologies for expression microarrays with 100% PMT and 10 μm resolutions. Data were extracted using Feature Extraction Software version 8.1 (Agilent Technologies).

Data analysis. Normalized values of the median red and green foreground were calculated from M and A: $R_f = A + (M / 2)$ and $G_f = A - (M / 2)$. Partitioning of the total variation in the foreground signal into different sources enables comparison of the variance between levels of a factor to the variance within each level due to noise. Our ANOVA model for each probe represents how different factors contribute to the observed signal. All two-way interactions of the experimental factors, treatment, cell line, and length of treatment were included in the model, because we could not assume the effects of each factor to be independent.

5. www.bioconductor.org
6. www.r-project.org
factor were consistent across the levels of other factors. In summary, we assumed:

\[ y_{\text{fg}} = \mu + I_i + D_j + T_k + C_l + L_m + (TC)_{ijkl} + (TL)_{klm} + (CL)_{imn} + e_{ijklm} \]

where \( y_{\text{fg}} \) is the log 2 of the foreground fluorescent intensity from array (I) and dye (D). \( I_i \) and \( D_j \) representing the fixed effects of treatment \( T \), cell line \( C \), and length of treatment \( L \).

Because between-array normalization was not done during preprocessing, factors for the effects of dye \( D \) and array \( I \) were included in the model.

Quantities of interest were contrast coefficients reflecting comparisons of treatment and controls within cell line and duration of treatment. For significance analysis, we used the moderated \( t \) statistic. This has the same interpretation as an ordinary \( t \) statistic, except that it is adjusted for multiple comparisons with the Benjamini-Hochberg method.

**Western blots**

Cells were plated and treated as described in the cell culture section of the microarray experiment. After 24, 48, 72, and 96 h, cells were washed once with ice-cold PBS and harvested on ice using scrapers and the lysis M-PER solution of Pierce containing a cocktail of protease inhibitors (Roche). After 10 min incubation on ice, the lysates were centrifuged at 6,000 rpm at 4°C for 10 min. The protein-containing supernatant was stored in -80°C until further analysis with the BCA assay (Pierce). Laemmli sample buffer at a ratio of 1:2 was mixed to the sample and heated at 95°C for 5 min. The denatured proteins were stored in -80°C until use. Using a 4% to 15% Tris-SDS-PAGE gel (Bio-Rad), 10 to 20 μg proteins were separated at 100 V for 1 h. The separated proteins in the gel were electrophoretically transferred onto polyvinylidene difluoride membranes (Bio-Rad) at 380 mA for 1 h. Blocking was carried out with TBS plus 5% nonfat milk [20 mmol/L Tris-HCl (pH 7.6), 137 mmol/L NaCl]. After washing, the membrane was incubated overnight with primary antibodies for p21WAF1 (Santa Cruz Biotechnology) and Ac-H3K9 (Cell Signaling) diluted in TBS. Appropriate homodisulfide-disulfide-conjugated secondary antibodies (Amersham) were incubated for 1 h. For detection, autoradiography using enhanced chemiluminescence was done according to the manufacturer’s instructions (Amersham).

**Results**

**Cell-specific sensitivity for SAHA and TSA using MTT assays.** To test the sensitivity of prostate cancer cells to HDACIs, a standard MTT proliferation assay was used. Four epithelial prostate cancer cell lines (LNCaP, PC3, CWR22R, and DU145) were treated with 3 μmol/L vorinostat or 300 nmol/L TSA (Fig. 1). The assay revealed that, compared with LNCaP, CWR22R, and DU145, the proliferation rate of PC3 cells was significantly less affected after 48 h treatment with either vorinostat or TSA.

For all other experiments, we decided to compare PC3 cells with DU145 cells because the doubling time, morphology, androgen responsiveness, and p53 status of PC3 and DU145 is relatively comparable. Figure 2A shows that, when DU145 and PC3 cells are treated for 96 h, growth of DU145 cells is completely inhibited at the highest concentration of HDACI treatment, whereas proliferation of PC3 cells is only inhibited by 50%.

**Microarray experiment.** Three independent variables were studied in our experiment: two prostate cancer cell lines (DU145 and PC3) with varying sensitivity to HDACI, two structurally different HDACIs [vorinostat (also known as SAHA) and small-chain fatty acid VPA], and two treatment intervals (48 and 96 h). To determine the relative contribution of these three independent variables of interest on gene expression, all samples were hybridized according to a novel “multidimensional loop” (Fig. 2B). This setup allowed for independent analysis of each variable without assuming dependency between each variable and with minimal use of resources.

**Preprocessing summary.** Image plots of raw expression data for each individual microarray (Fig. 3) revealed some spatial artifacts in the green background signal. Plots of the foreground log2(red/green) versus background log2(red/green) revealed little correlation (Fig. 4A and B), indicating that the spatial artifacts seen in the green background did not significantly affect foreground signal (11). Background subtraction was therefore omitted.

In MA plots of foreground signals, we noted for all arrays a deviation from zero in the low expression signals (Fig. 4C). We used Loess, as implemented in the limma package (12), to adjust for this systematic technical variation associated with intensity (Fig. 4D).

Box plots of the Loess normalized red and green foreground signals from all arrays visualized a batch effect (Fig. 5A and B), which was most likely caused by our splitting target fragmentation, hybridization, and scanning of the arrays into three
different days. This batch effect was not visible in box plots of the Loess normalized fold changes ($M$; Fig. 5C), indicating that the batch effect did not favor one channel. Box plots of the Loess normalized fold changes ($M$) also indicated that most arrays had similar variability, except for all four arrays with both cell lines on the same array. Although across-array normalization could correct for the batch effect, it might also eliminate true cell line differences. Therefore, an array factor was added to the statistical model to adjust for this effect without performing a separate across-array normalization (see Materials and Methods).

**Gene expression data.** Using a custom-made ANOVA model (methods), genes with significantly different expression were identified. The comparison of interest in this study was differential expression after treatment. Table 1 displays the differential expression caused by these variables.

**Validation: Western blots of differentially expressed genes.** Western blot analysis was used to validate the up-regulation of the classic HDACi activity marker p21 and the acetylation status of histone 3 after treatment with VPA or vorinostat. Figure 6 shows that the acetylation status of H3 and the protein expression of p21 are increased throughout the course of the experiment. Treatment of DU145 cells compared with PC3 cells resulted in a much higher protein expression of acetylated H3 and p21.

**Discussion**

Although clinical trials with HDACIs are ongoing, little is known about the mechanisms of action and resistance to these compounds. To improve our understanding of underlying mechanisms of resistance, we studied the effect of HDACIs on gene expression profiles of prostate cancer cell lines with different drug sensitivity. We decided to prepare 16 RNA samples of interest that represented three variables: cell line, duration of treatment, and type of HDACI. To extract data from these samples to answer our complex questions, we needed bioinformaticists.

In this article, we presented our microarray experiment as a case study to increase understanding of experimental design and analysis in microarray experiments. Each step during data generation and statistical analysis of a microarray experiment may influence the expression data that are generated. Below, we will address four issues of importance in the context of microarray experiments: design, biological and technical variation, batch effect, and pooling.

One of the first steps in a microarray experiment that may influence the eventual results is its design. Several groups have highlighted the importance of a solid design, but due to the large variety of research questions, general rules are difficult to define. In this study, we highlight the importance of a solid design using our own experience. As mentioned above, the two most common set ups in dual-color arrays are the loop design and the reference design.

One of the first steps in a microarray experiment that may influence the final results is its design. Several groups have highlighted the importance of a solid design, but due to the large variety of research questions, general rules are difficult to define. In this study, we highlight the importance of a solid design using our own experience. As mentioned above, the two most common approaches in dual-color arrays are the loop design and the reference design. The advantages of loop designs include repetitive analysis of samples of interest and the ability to assess dye effect without the use of additional resources. Disadvantage of a loop design include that they provide less information on comparisons between samples that have indirect contact and that the efficiency of these comparisons is distance dependent (6). In addition, they are often complicated, not robust to chip loss due to technical misadventures, and limited in the hypotheses that can be

![Image plots of unprocessed red and green signals](#)

**Fig. 3.** Image plots of unprocessed red and green signals: log$_2$ $R_b$ (A), log$_2$ $G_b$ (B), log$_2$ $R_f$ (C), and log$_2$ $G_f$ (D) of a representative microarray visualize spatial artifacts in log$_2$ $G_b$. Image plots visualize the intensity of each expression value on the glass slide. Their purpose is to identify spatial defects in the data sets such as scratches. $R_b$, red background; $G_b$, green background; $R_f$, red foreground; $G_f$, green foreground.
tested. In contrast, reference designs are robust to chip loss and are able to address a much wider range of hypotheses. Furthermore, reference designs are much easier to explain to investigators and are harder to misexecute either in the planning, execution, or analysis stage. There are two main disadvantages of a classic reference design: (a) all comparisons between the conditions of interest are indirect, as those are never hybridized on the same array, and (b) unless the reference itself is of interest, half the power of the experimental technique is invested in measurements that are not of direct scientific interest. These two factors generally result in a lower power to identify changes of interest and in a larger statistical variance on the estimated effects. Recently, it has been proposed that reference designs show primarily advantage in experiments that cannot be completed at once (6, 14). In clinical trials for instance, tumor samples might come available throughout the course of the trial. With the use of a (commercially available) reference sample that is cohybridized with each tumor sample, the measurement can be done, whereas technical variation and quality can be assessed. For each experiment, a careful deliberation on the pros and cons of each setup in the light of the experimental question will aid the eventual data that is generated. In our experiment, we concluded that the use of a reference design would have resulted in the use of ~50% more microarray slides. Instead, we chose for a modified loop design (Fig. 2B). Because 16 samples needed to be analyzed, a classic loop design would have resulted in too great a distance between some of the samples (6). To achieve a lower residual variance, we decided to hybridize more samples directly.

The choice of experimental design can be crucial in determining the extent to which biological and technical variation effect results of a microarray experiment (7). Depending on the experimental setup, some technical variation can be adjusted by normalization steps. Within-array normalization may correct for variation that is array-specific, such as scratches, printing effects, and local background. Between-array normalization aims at correcting for systematic variation that occurs across arrays, such as time effect and technical variation in mRNA pools (15). In general, one can identify four factors that contribute to variation of two-channel microarray data (6), varieties V (factors of interest, such as time points of a biological process or treatment), genes G (sequences spotted on the arrays), dyes D (Cy3 and Cy5), and arrays A (overall variation in fluorescent signal from array to array). Depending on the interaction between these factors, up to $2^4 = 16$ causes of variation are possible. The interaction between the factors of interest (V) with genes (G) is of primary interest to study differential gene expressions (6). All other interactions, such as the interaction of dye (D) with genes (G; ref. 16), determine the sources of experimental variability and should be corrected before quantification of differential expression. In our experiment, the use of our multiple-loop, double-cube design for the analysis of 16 samples enabled profiling of transcription in prostate cancer cell lines DU145 and PC3 during treatment with two HDACIs vorinostat and VPA. An ANOVA model per
gene that included additional factors for dye (D) and array (A) retained cell line-specific changes while correcting for between-array variability such as our batch effect (Fig. 5A and B). Larger microarray experiments that aim to answer complex questions can cause additional challenges that need to be adequately addressed. An example of this is the batch effect that is visible in the Gf and Rf box plots of our data (Fig. 5A and B). Batch or time effects can occur when not all arrays are hybridized at the same time and/or by the same person. In our case, the experiment included 22 arrays; therefore, a batch

![Gene expression box plots](image)

Fig. 5. A and B, box plots of green (Gf) and red (Rf) foreground signals after Loess normalization visualizing a batch effect in each channel. C, box plots of Loess normalized fold changes (M) for all 22 arrays using foreground data. Increased variability in four arrays having both cell lines cohybridized is visible: array 5+6 both have vorinostat-treated cell lines cohybridized; array 17+18 both have VPA-treated cell lines cohybridized.

<table>
<thead>
<tr>
<th>Table 1.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(A) Differential expression after treatment with VPA and vorinostat</strong></td>
</tr>
<tr>
<td><strong>Cell line</strong></td>
</tr>
<tr>
<td>DU145</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>PC3</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

| **(B) Expression pattern of commonly changed genes (refs. 1, 13)** |
| **Gene name** | **Systematic name** |
| tVPAda2DU | tVPAda4DU | tVPAda2PC3 | tVPAda4PC3 | tSAHAda2DU | tSAHAda4DU | tSAHAda2PC3 | tSAHAda4PC3 |
| 2.00 | 2.56 | 1.70 | 2.18 | 2.10 | 3.48 | 1.25 | 2.07 | CDKN1A | NM_000389 |
| 2.04 | 2.07 | 1.82 | 1.85 | 1.73 | 2.02 | 1.56 | 1.83 | GLRX | NM_002064 |
| 1.55 | 1.59 | 1.16 | 1.18 | 1.59 | 1.61 | 1.18 | 1.19 | FUC1A | NM_000147 |
| 1.32 | 1.42 | 1.21 | 1.30 | 1.39 | 1.83 | 1.29 | 1.69 | VEGFC | NM_005429 |
| 0.53 | 0.53 | 0.82 | 0.81 | 0.45 | 0.43 | 0.78 | 0.75 | TYMS | NM_001071 |
| 0.66 | 0.68 | 0.92 | 0.95 | 0.71 | 0.69 | 0.90 | 0.90 | CTPS | NM_001905 |
| 1.01 | 0.94 | 1.09 | 1.02 | 1.02 | 0.99 | 1.02 | 0.99 | RARB | NM_000965 |
| 1.17 | 1.22 | 1.20 | 1.25 | 1.03 | 1.01 | 1.23 | 1.20 | SEMA3A | NM_006080 |
| 0.96 | 0.92 | 1.01 | 0.97 | 0.99 | 0.96 | 0.97 | 0.94 | NR1 | NM_003873 |
| 1.01 | 0.97 | 1.00 | 0.96 | 1.02 | 1.07 | 0.99 | 1.04 | TF | NM_005994 |
| 1.00 | 1.02 | 1.03 | 1.05 | 1.01 | 1.01 | 1.00 | 1.00 | IL2 | NM_00586 |
| 1.10 | 1.41 | 1.92 | 2.48 | 1.17 | 1.69 | 1.33 | 1.93 | IL6 | NM_000600 |
| 1.00 | 0.96 | 1.04 | 1.01 | 1.01 | 1.02 | 1.00 | 1.00 | IFNG | NM_00619 |
| 0.95 | 0.97 | 1.02 | 1.04 | 1.02 | 1.00 | 1.04 | 1.02 | AR | NM_000444 |
| 0.99 | 1.03 | 0.97 | 1.01 | 1.08 | 1.09 | 0.96 | 0.96 | KLK3 | NM_001648 |
| 0.49 | 0.59 | 0.83 | 0.99 | 0.52 | 0.59 | 0.80 | 0.92 | HIST2H2AA | NM_003516 |

NOTE: Differential expression (P < 0.001) after 48 or 96 h treatment with 1 μmol/L vorinostat or 1 mmol/L VPA in PC3 or DU145 cell lines.
effect was inevitable. Time-specific changes may affect both signals equally, and if a ratio (fold change) is taken during analyses, the time effect may be missed (Fig. 5C). For a loop design, the separate expression values are used instead of fold change during ANOVA analysis. A correction step for time effects is very important in this setup, because true differential expression will be confounded otherwise. The “normalizeBetweenArrays” function in Limma includes a normalization step to correct for scale differences across arrays. The basic assumption of this procedure is that the distribution of log2 intensities between arrays is similar. Major distribution and scale differences in the intensities are therefore assumed to be due to technical variation. In our experimental setup, this basic assumption could have eliminated cell line-specific differences of interest (Fig. 5C). Therefore, we corrected for the batch effect that was visible in the separate channels (Fig. 5A and B) by adding a factor for array (A) to the ANOVA model.

While setting up designs for microarray experiments, most investigators face the question of whether to pool replicate samples or to measure each biological replicate separately. Biological replicates are independent preparations of samples taken from the same biosource that have been treated identically. Biological replicates assess the natural variability in a system. The separate measurement of biological replicates versus pooling of replicate samples is under continuing debate (17–19). The main argument in favor of measuring biological replicates separately is that biological variation may occur from sample to sample. This variation is especially prominent when tissue samples from different patients are being studied. The basic assumption when one decides to pool samples is that the expression data of pooled samples is equal to the average of the same measure taken from each individual sample that contributed to the pool. We assumed that the biological variation within our cell lines is limited. Variation from experiment to experiment will primarily come from technical variation. Based on this assumption, we decided to pool three samples from the same group. To assess the technical variation arising from the microarray technology, each pooled sample was split and multiple technical replicate measurements were carried out. As a result of our experimental setup, there were no degrees of freedom available to estimate the biological variance. We did, however, look at expression between the DMSO control and the water control and did not find any differentially expressed genes (data not shown).

When data generation, normalization, and statistical analysis have taken place, the data need to be analyzed and interpreted. To determine which genes are differentially expressed and which gene expressions remained unchanged, arbitrary cutoffs for fold change and/or P value (or B statistic) are often used. However, Hughes et al. showed that the basa expression of some genes fluctuates much more than others (20, 21). In addition, Pritchard et al. found a functional relationship between genes whose expression is most variable (22).

During the process of data interpretation, it is important to realize that the strength of microarray technology is to generate comprehensive data that precede the formulation of hypotheses. In contrast to the data generation, data interpretation is often guided by preexisting ideas in microarray experiments (5). In this study, we found several differentially expressed genes that have been linked to HDACI treatment in previous studies (refs. 1, 13), although some previously published genes showed not to be differentially expressed in our data set (Table 1B). Likely reason for these differences in expression profiles include differences in type of HDACI, cell line used, duration of treatment, and dose. Further studies are necessary to determine the importance of these differences.

Not to lose sight of the purpose of our multiple-loop, double-cube microarray experiment, earlier in vitro and in vivo experiments with VPA by our group and others (23) have shown that PC3 cells are relatively resistant to VPA and other HDACIs. In this article, we describe that treatment of prostate cancer cell lines PC3 and DU145 with HDACIs causes distinct changes in expression profiles. The changes in expression levels arising from the microarray technology, each pooled sample was split and multiple technical replicate measurements were carried out. As a result of our experimental setup, there were no degrees of freedom available to estimate the biological variance. We did, however, look at expression between the DMSO control and the water control and did not find any differentially expressed genes (data not shown).

When data generation, normalization, and statistical analysis have taken place, the data need to be analyzed and interpreted. To determine which genes are differentially expressed and which gene expressions remained unchanged, arbitrary cutoffs for fold change and/or P value (or B statistic) are often used. However, Hughes et al. showed that the basal expression of some genes fluctuates much more than others (20, 21). In addition, Pritchard et al. found a functional relationship between genes whose expression is most variable (22).

During the process of data interpretation, it is important to realize that the strength of microarray technology is to generate comprehensive data that precede the formulation of hypotheses. In contrast to the data generation, data interpretation is often guided by preexisting ideas in microarray experiments (5). In this study, we found several differentially expressed genes that have been linked to HDACI treatment in previous studies (refs. 1, 13), although some previously published genes showed not to be differentially expressed in our data set (Table 1B). Likely reason for these differences in expression profiles include differences in type of HDACI, cell line used, duration of treatment, and dose. Further studies are necessary to determine the importance of these differences.
in the relatively resistant PC3 were significantly lower than that of DU145. Our microarray design will help us discern the contribution of duration of treatment and type of HDACI to the resistance pattern seen in PC3. Next steps in our analysis will include bioinformatics tools that aid interpretation of microarray data such as Gene Set Enrichment Analysis and clustering algorithms.

**Conclusion**

In this study, one experiment was used to compare differential expression profiles of 16 different biological samples that represented three factors of interest. We conclude that our multiple-loop, double-cube microarray design can be used to identify molecular pathways that may be relevant in mechanisms of resistance to HDACIs.

**Disclosure of Potential Conflicts of Interest**

No potential conflicts of interest were disclosed.

**Acknowledgments**

We thank Morgan A. Marks for helpful discussions and assistance with R and Bioconductor, Conover Talbot, Wayne Yu, Catharine van Tussenbroek Fonds, Jo Keurfonds, Leiden International Study Fund, and Aegon.

**References**

A Multiple-Loop, Double-Cube Microarray Design Applied to Prostate Cancer Cell Lines with Variable Sensitivity to Histone Deacetylase Inhibitors

Madeleine S.Q. Kortenhorst, Marianna Zahurak, Shabana Shabbeer, et al.