

Analysis of Unbalanced Microarray Data

Mei-Ling Ting Lee^{1,2,3,4}, G.A. Whitmore⁵, Rus Y. Yukhananov^{1,2}

¹*Brigham and Women's Hospital*, ²*Harvard Medical School*,

³*Harvard School of Public Health*

⁴*Massachusetts General Hospital*, and ⁵*McGill University*

Abstract: This paper investigates statistical procedures for analyzing microarray gene expression data obtained from studies with an unbalanced experimental design. We demonstrate the methods using microarray data from a study of opioid dependence in mice. The experiment was designed to investigate how morphine dependence alters gene expression in spinal cord mRNA. The aim was to identify genes that characterize the tolerance, withdrawal and two abstinence stages of dependence and to describe how gene expression is altered in moving from one stage to the next. The study design was unbalanced in several respects. First, for mice receiving morphine, arrays were made for four dependence stages, while for mice receiving placebo, arrays were made for only three stages. Second, administrative error led to an omitted replication for one treatment combination. Third, some expression readings were missing.

Extending the two-stage ANOVA model of Lee *et al* (2000, 2002a) this paper first uses a chi-square statistic to identify a small set of genes that exhibit differential expression over one or more treatment combinations. This gene set is then examined further using cluster analysis and novel inference methods to uncover specific genes and gene clusters that play a role at different stages of opioid dependence and, in particular, a role in the persistence of effect into the late abstinence stage. The latter effect implies that morphine dependence has a long-term genetic impact. The statistical power of the study to uncover differentially expressed genes is calculated as a prelude to further investigation. The analytical results proved useful to scientists in understanding the link between opioid dependence and gene function.

Key words: Differentially expressed genes, microarray data, power analysis, two-stage ANOVA, unbalanced experimental design.

1. Introduction

Linear statistical models, such as the methods of analysis of variance (ANOVA) as discussed by Kerr, Martin, et al (2000), Kerr and Churchill (2001), Lee et al (2000, 2002a), and mixed-effects statistical models discussed by Wolfinger, Gibson, et al (2001), have been widely used for analyzing microarray gene expression data. These investigations, however, have tended to focus on complete and balanced designs. A complete and balanced design is an ideal that is not always achieved in practice. The cost of conducting a microarray experiment, in terms of labor, time and money, sometimes makes it impractical to use a balanced design. Also, some studies set out with a balanced design but become unbalanced during implementation because of technical and administrative problems. Finally, it is common for scientists to embark on a study prior to consulting a statistician and the chosen design may be unbalanced. In recognition of this reality, this paper presents methods for analyzing microarray data from unbalanced designs. We illustrate the methods using data from a study of opioid dependence in mice.

This article extends the two-stage ANOVA model considered in Lee *et al* (2000, 2002a) to unbalanced designs. First, we use a chi-square statistic to identify a small set of genes that exhibit differential expression over one or more treatment combinations. This gene set is then examined further using cluster analysis and novel inference methods to uncover specific genes and gene clusters that play a role at different stages of opioid dependence and, in particular, a role in the persistence of effect into the late abstinence stage. The latter effect implies that morphine dependence has a long-term genetic impact. We calculate the statistical power of the study to uncover differentially expressed genes as an aid to planning the scale of future studies in this series of investigations. Finally, we show how the analytical results proved useful to scientists in understanding the link between opioid dependence and gene function.

2. Experimental method and design

The impact of opioid dependence on the genetic makeup of spinal cord and brain tissue is medically and scientifically important and has attracted the attention of a number of researchers – see Loguinov *et al.* (2001) and references therein. The experiment reported here was designed to investigate how morphine dependence in mice alters gene expression in spinal cord mRNA. The study design involved two treatments, *morphine* and *placebo*, and four time points corresponding to consecutive stages of opioid dependence, classified as *tolerance*, *withdrawal*, *early abstinence* and *late abstinence*. The aim of the study was to identify genes that characterize the tolerance, withdrawal and two abstinence stages and to describe how gene expression is altered in moving from one stage to the next.

Pellets containing 75 mg of morphine base (treatment) or placebo (control) were implanted subcutaneously in eight-week old male mice. Treatment mice were sacrificed sequentially at four time points corresponding to the four dependence stages. Control mice were sacrificed at the same time points, with the exception of the withdrawal stage which was omitted on the assumption that the tolerance and withdrawal stages are identical in placebo group. Thus, four arrays were made for mice given the morphine treatment, one for each dependence stage, while only three arrays were made for mice given the placebo treatment.

The microarray data we consider resulted from hybridization of the dorsal part of the mouse spinal cord samples to a custom-designed cDNA array. The cDNA array contains 1667 cDNA fragments and 61 control spots. At each time point (i.e., at each stage), in both the treatment and control groups, three mice were sacrificed, for a total of 21 mice. The paucity of spinal column mRNA in any single mouse required that the mRNA of the three mice sacrificed together be combined and blended into a single sample. The treatment and control samples were labeled with Cy-5 dye and read on the red channel. Other control samples, derived from mouse brain tissue, were labeled with the Cy-3 (green) dye. As the purpose of this investigation is to learn how morphine dependence alters gene expression in spinal cord mRNA, the Cy-3 readings from mouse brain tissue (a different tissue type) are not used in the analysis reported here. See Lee et al (2000) for

another illustration of the use of single-channel microarray readings. DNA microarray analysis of gene expression was done essentially as described by Eisen and Brown (1999) with some modifications (Loguinov *et al.* 2001).

For each *gene g*, fluorescent expressions were measured for *treatment group* (1= placebo, 2= morphine), and *time stage t* (1= tolerance, 2= withdrawal, 3= early abstinence, 4= late abstinence), except for the placebo-withdrawal combination where no array was created. The placebo-withdrawal combination was not implemented on the assumption that it would be identical to the placebo-tolerance combination. The original study plan called for each spinal column sample to be placed on two spots of the same slide, yielding a *replicated expression reading*. This plan was not carried through perfectly, however, as the spot replicate for the morphine-tolerance combination was omitted by mistake.

Finally, as often happens in microarray experiments, it was found that the array for the morphine-late abstinence combination had a large number of defective spots and several dozen spots on other arrays were faulty. The final microarray data set contains readings for only 1722 genes out of the original set of 1728 (=1667+61). Six genes were dropped because they had defective readings for the morphine-tolerance combination, the unreplicated treatment combination in the design. We did not choose to employ imputation to replace missing values. It was felt that any imputation assumption would be too speculative to be justified.

In summary, therefore, we find three elements of imbalance in this design: (1) an omitted treatment combination, (2) an omitted replicate and (3) several missing expression readings.

3. Analytical framework

As the design is not a complete factorial design (one cell is not complete), we shall initially view the $K = 7$ cells as seven individual experimental conditions, indexed by $k = k_1, \dots, k_7$. The reading Y_{gkr} represents the natural logarithm of the raw Cy-5 gene expression reading, without background correction, for gene g , treatment combination k and spot replicate r . Table 1 shows the correspondence of the treatment combinations to the index values.

Table 1. The experimental design for the study involves seven treatment combinations of time stage and morphine treatment. The placebo-withdrawal treatment combination was not implemented on the assumption it would be identical to the placebo-tolerance combination. All treatment combinations involved replicated spots except the morphine-tolerance combination which had no replicate because of administrative error.

Treatment	Dependence Time Stage			
	<i>t</i> = 1 <i>Tolerance</i>	<i>t</i> = 2 <i>Withdrawal</i>	<i>t</i> = 3 <i>Early Abstinence</i>	<i>t</i> = 4 <i>Late Abstinence</i>
<i>Placebo</i>	k_1	*	k_2	k_3
<i>Morphine</i>	k_4	k_5	k_6	k_7

As shown by Lee *et al* (2002a), the following two-stage ANOVA model is appropriate for this kind of data set.

$$Y_{gkr} = \mu + \tau_k + W_{gkr} \tag{3.1}$$

$$W_{gkr} = \gamma_g + (\gamma\tau)_{gk} + \epsilon_{gkr} \tag{3.2}$$

In equation (3.1), μ represents the overall population mean log-expression intensity, τ_k is the main effect for the array corresponding to experimental condition k , and W_{gkr} is a normalized (centered) quantity that captures all the gene-specific effects of the experimental condition. In (3.2), γ_g is the main effect for gene g , $(\gamma\tau)_{gk}$ is the interaction effect for experimental condition k and gene g , and ϵ_{gkr} is an error term. Interaction $(\gamma\tau)_{gk}$ reflects the differential expression of gene g in condition k . We assume that the error terms ϵ_{gkr} have mean 0, have a common variance and are mutually independent. We make no assumption about their distribution form.

Recall that the original study design called for two spot replicates per treatment combination. The replication in the design ensures ample degrees of freedom for estimating all parameters. We note for clarity that the notation in models (3.1) and (3.2) does not take explicit account of the

omitted spot replicate in treatment combination k_4 or the missing readings for defective spots.

The estimation of the parameters in equations (3.1) and (3.2) proceeds in two stages. The residuals from fitting the first model are estimates of the W_{gkr} . The estimates $\hat{\tau}_k$ are normalizing constants for the arrays. As there are $G = 1722$ genes in our microarray data set and the design is unbalanced, fitting the second model involves substantial calculation. In essence, a regression model is fitted for each gene g to provide estimates of γ_g and $(\gamma\tau)_{gk}$, for $k = k_1, \dots, k_7$, $g = 1, \dots, G$. We impose sum-to-zero estimability constraints on each of the sets of parameters for main effects and interactions.

We shall use the simpler notation \mathcal{I}_{gk} for the interaction term $(\gamma\tau)_{gk}$ and denote its estimate by $\hat{\mathcal{I}}_{gk}$. The $\hat{\mathcal{I}}_{gk}$ are the estimated differential effects for the seven treatment combinations.

4. Analysis and results

We now describe the results of several analyses of differential gene expression in this experiment.

4.1 Differentially Expressed Gene Pool

First, we investigate which genes, if any, are differentially expressed in any of the treatment combinations, reflecting effects arising from either the time stage or morphine treatment. This first analysis identifies a set of genes that are qualified for study in subsequent stages of analysis in the sense that they show evidence of some form of differential expression. The sum of the squared interaction effects, shown next, is a summary measure that captures the combined effects across all treatment combinations.

$$C_g = \sum_{k=k_1}^{k_7} \hat{\mathcal{I}}_{gk}^2 \quad (4.1)$$

We do not use measure C_g directly because of the imbalance in this experimental design. Instead, we use the *model sum of squares* from the regression analysis of (3.2) for gene g . The model sum of squares is a weighted-sum

version of C_g . We denote this quantity by SSM_g . With a balanced design having r readings for gene g on each of the treatment combinations, we note that C_g and SSM_g would be connected as follows: $SSM_g = rC_g$.

If the ANOVA error term in (3.2) were normally distributed, one might wish to use the regression F statistic $F_g^* = MSM_g/MSE_g$ to test the null hypothesis of no differential expression for gene g , i.e., the hypothesis $H_0 : \mathcal{I}_{gk} = 0$ for $k = k_1, \dots, k_7$. Here $MSM_g = SSM_g/6$ and MSE_g denote the regression mean squares for the model and error ANOVA components, respectively. The test would entail comparing F^* with an appropriate F percentile.

We wish to avoid an assumption about the error term distribution, however, and thus we use the following chi-square based rule. This rule requires that the estimated interaction effects $\hat{\mathcal{I}}_{gk}$ be approximately normally distributed. With replicated observations, the central limit theorem gives some assurance of approximate normality in this context.

Proceeding, we let \mathcal{M} denote the median value of the SSM_g values for all $G = 1722$ genes. We then calculate the statistic

$$S_g = \frac{\chi_{K-1}^2(0.5)}{\mathcal{M}} SSM_g, \quad (4.2)$$

where $K - 1 = 7 - 1 = 6$ in this application. Quantity $\chi_{K-1}^2(0.5)$ denotes the median of a χ_{K-1}^2 variable. The ratio of the respective medians in (4.2) scales the quantity SSM_g appropriately so that S_g will have an approximate χ_{K-1}^2 distribution under the null hypothesis of no differential expression, i.e., under $H_0 : \mathcal{I}_{gk} = 0$ for $k = k_1, \dots, k_7$. The distributional approximation holds if at least 50 percent of the genes are not differentially expressed, which we believe to be a reasonable assumption in this study.

Now, we choose the following chi-square percentile as a cutoff for S_g that separates differentially expressed genes from those that are not

$$\chi_{K-1}^2 \left(\frac{G}{G+1} \right) \quad (4.3)$$

In this application, this cutoff value equals $\chi_6^2(1722/1723) = 23.75$. Criterion (4.3) has the property that it tends to produce one false positive in each application. Another criterion could be used to control the family

Table 2: Differentially Expressed Gene Pool: 43 genes exhibiting significant differential expression for one or more of the seven treatment combinations according to the criterion that model sum of squares SSM_g is significantly large, ordered by chi-square statistic S_g .

Gene Name	Chi-square S_g	Gene Name	Chi-square S_g
g1145	23.80181	g1087	29.99177
g339	24.03087	g234	30.62762
g1201	24.14126	g870	30.99461
g33	24.16163	g1698	31.61664
g898	24.23594	g914	32.54018
g1519	24.48179	g1105	32.92247
g830	24.98025	g900	32.98985
g1526	25.36965	g278	34.11127
g946	25.45033	g1169	35.229
g922	25.75307	g1465	35.61224
g708	25.76863	g1649	36.63601
g453	25.83311	g1073	38.83312
g1175	25.9591	g1199	41.85474
g1669	26.39848	g1063	42.49206
g1623	26.62695	g1185	42.66629
g939	29.0621	g438	43.74333
g899	29.1816	g1257	48.36805
g633	29.36458	g1134	48.95993
g17	29.50138	g1191	52.02124
g311	29.63199	g1024	55.24098
g371	29.81269	g1202	71.21092
g1704	29.87632		

type I error rate but the study team found this one to be acceptable in this application. Based on the cutoff value of 23.75, we find that the 43 genes listed in Table 2 are differentially expressed. We refer to this set as the *differentially expressed gene pool*. The fact that only one false positive is

Table 3: Gene clusters sharing similar differential expression profiles across the *seven treatment combinations, irrespective of morphine or placebo treatments*. The 43 genes are listed here.

Cluster	Gene Name
A1 :	{g278, g311, g339, g371, g453}
A2 :	{g438, g870, g898, g900, g914, g946}
A3 :	{g526, g1704}
A4 :	{g17, g33, g1169, g1201, g1519}
A5 :	{g708, g1063, g1257, g1649, g1669}
A6 :	{g1191, g1465}
A7 :	{g939, g1134, g1175}
Singleton	{g234}, {g633}, {g830}, {g899}, {g922}, {g1024}, {g1073}, {g1087}
Clusters:	{g1105}, {g1145}, {g1185}, {g1199}, {g1202}, {g1623}, {g1698}

expected in this list implies that there is a substantial number of genes in the list that are truly differentially expressed.

4.2 Cluster analysis

Next, we employ cluster analysis to identify subsets of the 43 genes in the differentially expressed gene pool that have similar expression profiles across the seven treatment combinations. For this purpose, we employ the routine for hierarchical cluster analysis from the JMP statistical software system of the SAS Institute. The clustering routine uses the standardized values of the \hat{I}_{gk} and the centroid linkage criterion. The results are shown by a gene-wise dendrogram in Figure 1. We see a number of distinct clusters. We cut the tree at a distance of 1.423, creating 22 clusters. Table 3 shows the membership of these clusters. Seven clusters contain multiple genes and 15 are singleton clusters.

4.3 Comparing genes at the late abstinence stage

Continuing, we focus our investigation on differential expression for the late abstinence stage, i.e. time stage 4, to see if morphine administration has a lasting impact on gene expression. For this purpose, we are interested

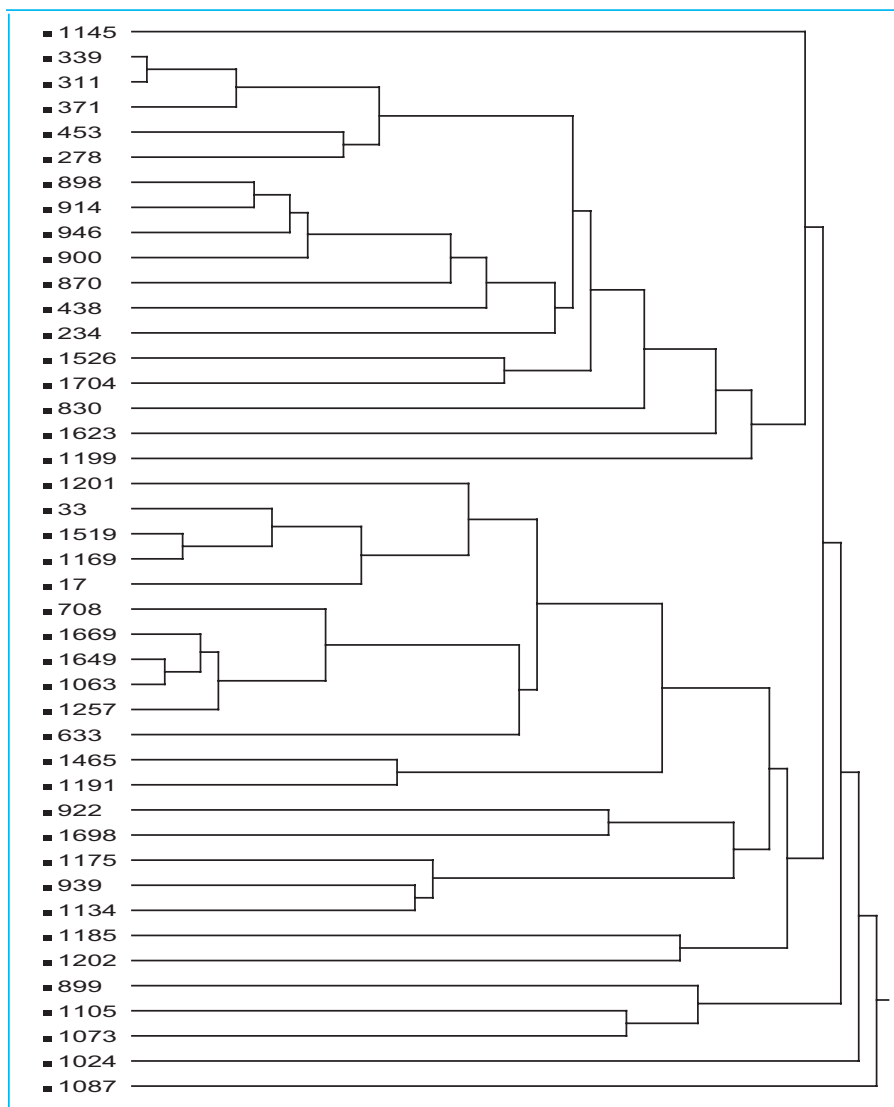


Figure 1. Dendrogram corresponding to Table 3. Vertical line shows the selected cluster solution.

in the following difference between condition k_3 , placebo at time 4, and condition k_7 , morphine at time 4.

$$d_g = \mathcal{I}_{gk_7} - \mathcal{I}_{gk_3} \tag{4.4}$$

The estimate of this difference is denoted by \hat{d}_g , where $\hat{d}_g = \hat{\mathcal{I}}_{gk_7} - \hat{\mathcal{I}}_{gk_3}$.

Again, based on the approximate normality of the estimated interaction effects, we have that \hat{d}_g follows an approximate $N(0, \sigma_0^2)$ distribution under the null hypothesis $H_0 : d_g = 0$, where σ_0^2 denotes the null variance parameter. This hypothesis states that there is no differential expression between the morphine and placebo treatments in the late abstinence stage.

We look at both the sample standard deviation of the \hat{d}_g and the standard deviation estimated from the interquartile range of the \hat{d}_g for the 1722 genes as alternative estimates of the null standard deviation σ_0 . The respective values are 0.5564 and 0.5942. They differ little. To be conservative, we shall use the smaller value as our estimate of σ_0 , denoting the estimate by $\hat{\sigma}_0$. We then compute the standardized values $z_g = \hat{d}_g / \hat{\sigma}_0$ and compare $|z_g|$ for each gene g with the standard normal percentile $z[G/(G + 1)] = z(1722/1723) = 3.2483$. Values of $|z_g|$ that exceed this percentile are judged to be differentially expressed. This cutoff criterion will tend to give one false positive in each tail (two false positives in total) with each application. Table 4 lists the four genes that are up-regulated and the one gene that is down-regulated in the morphine treatment, relative to the placebo, in the late abstinence stage. All of these genes are in the differentially expressed gene pool identified in the first analysis. As Table 4 lists only five genes in total and two are expected to be false positives, three genes are expected to be true positives.

4.4 Summarizing differences across all time stages

As a final analysis, differences between morphine and placebo treatment combinations of the type shown in formula (4.4) were calculated for all four time stages for the whole gene pool. For the withdrawal stage, the interaction estimates $\hat{\mathcal{I}}_{gk_1}$ from the placebo-tolerance combination were used as estimates for the missing placebo-withdrawal combination in the experimental design.

Table 4: Five genes exhibiting a significant difference in differential expression *between the morphine and placebo treatments* in the *late abstinence stage* as assessed by the standard normal statistic z_g .

Gene Name	Standard Normal Statistic z_g
Down-regulated in <i>morphine</i>	
g1087	-3.608409
Up-regulated in <i>morphine</i>	
g278	3.632209
g438	3.312954
g1199	3.915316
g1704	3.323965

We wish to test the null hypothesis $H_0 : d_{gt} = 0$ for all t . Here d_{gt} is the difference for gene g in time stage t and $t = 1, \dots, T$, with $T = 4$. The standardized difference corresponding to estimate \hat{d}_{gt} is denoted by z_{gt} and is calculated here using the standard deviation of all differences \hat{d}_{gt} at time stage t . As an approximate test procedure, we employ the following statistic, which is an analog of the earlier chi-square statistic (4.1).

$$D_g = \sum_{t=1}^T z_{gt}^2 \quad (4.5)$$

A gene is judged to be differentially expressed between the morphine and placebo treatments at one or more dependence time stages if its statistic D_g exceeds $\chi_T^2[G/(G+1)] = \chi_4^2(1722/1723) = 19.67$. In this application of the test procedure, D_g will be only approximately distributed as χ_4^2 under H_0 . With the approximation, however, we know the procedure is expected to yield about one false positive in testing the whole set of $G = 1722$ genes.

Table 5 shows the nineteen genes identified as differentially expressed by this criterion. The nineteen genes have been grouped into nine clusters in Table 5 according to the pattern of expression across the four time stages.

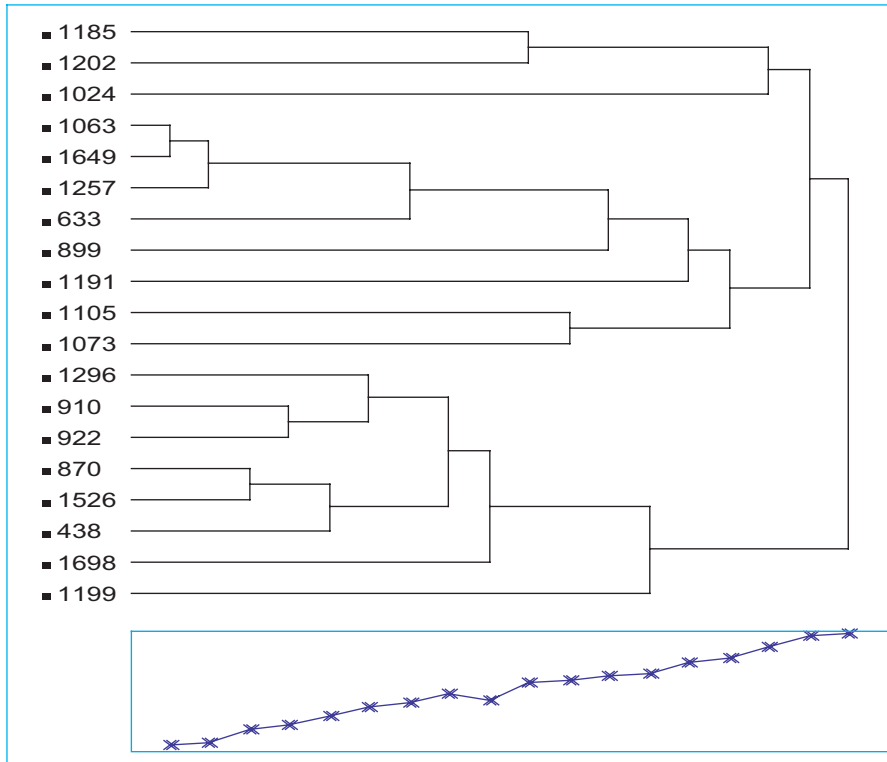


Figure 2. Dendrogram corresponding to Table 5. Vertical line shows the selected cluster solution.

The nine clusters were cut from the dendrogram in Figure 2 that was generated using the same hierarchical clustering routine described earlier. The cluster patterns are not internally uniform but some consistency is present. For example, the cluster with the largest membership, cluster B8 as listed in Table 5, shows genes that are fairly consistently up-regulated at all time stages under morphine relative to placebo. There is no guarantee that all genes that appear as differentially expressed by this criterion will be from the differentially expressed gene pool, although there should be a high degree of correspondence. We note that two of the 19 genes (g910, g1296) in cluster B8 are not in the pool of 43 genes listed in Table 2.

Table 5: Differences in differential expression *between the morphine and placebo treatments* in each of *four dependence time stages* as assessed by the standard normal statistic z_g for each stage. The clusters group genes with similar profiles across the *four time stages*.

Cluster	Gene Name	Standardized Normal Statistic z_g			
		<i>Tolerance</i>	<i>Withdrawal</i>	<i>Early Abstinence</i>	<i>Late Abstinence</i>
B1 :	g1185	-2.555528	.4801729	2.578545	-2.515898
	g1202	-1.891121	3.915715	.2488594	-1.984508
B2 :	g1024	.5734071	6.774648	-1.205199	-1.01564
B3 :	g633	.0371928	-.9443634	-4.435981	.3645121
	g1063	-1.920154	-.7747319	-4.065078	-1.267348
	g1257	-1.863966	-1.286642	-4.496387	-.9962125
	g1649	-1.747722	-.8297673	-4.239462	-.9612752
B4 :	g899	-3.404126	-4.072057	-1.705509	-1.007635
B5 :	g1191	2.517021	.840284	-4.309214	-.4327433
B6 :	g1105	-2.514672	-3.631862	-3.366916	2.004774
B7 :	g1073	-2.711815	-2.119922	.4428932	2.917885
B8 :	g438	1.58086	3.956466	.5766376	3.312954
	g870	2.501067	3.291142	.5538081	2.004692
	g910	3.12235	2.857303	1.773949	.2036051
	g922	3.532286	1.859219	2.587013	.9108489
	g1296	1.939579	3.970597	.8866081	-1.1409477
	g1526	2.472733	2.719066	1.742795	2.413913
	g1698	1.73313	1.428013	3.779864	2.034925
B9 :	g1199	3.352447	-.2531143	.9331462	3.915316

5. Estimated power

Using methods described in Lee *et al* (2002b), we have estimated the statistical power of the experimental design used in this study as a guide to designing further studies in this series of genetic investigations. To make this estimate, we use S_g in (4.2) as the summary measure of differential expression. Under H_0 , this measure is approximately a χ_6^2 variate.

As the specified alternative hypothesis H_a , we postulate a single treatment condition (say, the morphine-late abstinence condition k_7) that is differentially expressed relative to all other treatment conditions which are assumed to have uniform expression levels. Specifically, we assume that the

differential expression in this pattern is a four-fold difference. Under this specification of H_a , S_g follows a non-central chi-square distribution with six degrees of freedom and a non-centrality parameter which is estimated to be 25.0.

We have controlled the family type I error probability so that only one false positive is expected. We also have assumed that the estimation errors in the interaction estimates \hat{T}_{gk} are independent for different genes. (We stress that the independence here relates to statistical estimation error and not gene co-regulation. Co-regulation is almost certainly present.)

Finally, calculation of the relevant area under the non-central chi-square distribution gives the estimated power of this study under the specified alternative hypothesis as 0.739. This power is moderate to large, suggesting that differential expression levels consistent with a non-centrality parameter of 25.0 had more than seven chances in ten of being uncovered. The sum of the degrees of freedom and non-centrality value assumed under H_a equals $6 + 25.0 = 31$. Referring to the values of S_g reported in Table 2 for differentially expressed genes, we see that a value of 31 is quite typical. If H_a were changed to an hypothesis with an eight-fold differential expression then the noncentrality parameter increases to 56.4 and the power jumps to 0.999.

6. Discussion of results

Genes implicated in long-term genetic changes following morphine administration are expected to have a consistent and major differential impact across all time stages and, in particular, to be clearly significant in the early and late stages of abstinence. With this thought in mind, the results of the preceding statistical analysis were subjected to review by the principal scientist, giving the following findings and those summarized in Table 6.

6.1 Sorting out image artifacts

Upon review of the 43 genes in the differentially expressed gene pool in Table 2, 14 were observed to have image artifacts that may have caused them to appear as differentially expressed. Specifically, the artifact was identified as a six-fold or greater difference between the two duplicate spots

Table 6: Functional classification of genes in the differentially expressed gene pool.

Category	Gene Name
Energy metabolism:	g234, g278, g438, g633, g830, g900, g910, g914 g946, g1199, g1526, g1623
Regulatory function:	g453, g898, g1519, g1704
Structural function:	g870, g922, g1145, g1185, g1296
Image artifacts:	g17, g33, g939, g1063, g1073, g1087, g1134, g1169 g1201, g1257, g1465, g1649, g1669, g1698
Unknown function:	g311, g339, g371, g708, g899, g1024, g1105 g1175, g1191, g1202

for a gene (in those treatment combinations that were replicated). No attempt was made to repeat the analysis excluding these genes. Considerable subjectivity would be introduced by judging what kind of remedy was appropriate for these spots.

As noted in Section 2, 61 of the 1728 spots were used as negative controls. Three of these negative-control genes (g1063, g1087 and g1134) appear in the list of 43 differentially expressed genes and are clearly false positives. As these three negative controls are among the 14 spots showing image artifacts, the finding suggests that it was faulty replication of the spot that allowed these negative controls to appear in the differentially expressed gene pool.

6.2 Functional classification

Table 6 classifies the 43 genes in the differentially expressed gene pool into functional groups. The 14 genes with image artifacts are segregated in one class of the table. The remaining 29 genes are classified into functional groups related to energy metabolism, regulatory function or structural function or as having unknown function. Recall from section 4.4 that two additional genes (g910, g1296) appeared in Table 5 as having differences in expression between morphine and placebo treatments but are not among the

43 differentially expressed genes in Table 2. These two genes have also been classified in Table 6 into the energy metabolism and structural functions groups, respectively.

The energy metabolism group in Table 6 is characterized by increased expression during the tolerance stage and then a slight decrease during the early abstinence stage. Expression for the regulatory group also increases during the tolerance stage, decreases during withdrawal and then increases during late abstinence. Finally, expression for structural genes increases during tolerance and decreases during late abstinence.

Matching these categories with the findings in Tables 3, 4 and 5, reveals some interesting features. A small subset of genes associated with energy metabolism appears in Cluster A2 of Table 3 (g438, g900, g914, g946). The only gene spot down-regulated in morphine in Table 4 turns out to be a negative control from the image artifact category (g1087). Of the four genes up-regulated in morphine in Table 4, three are associated with energy metabolism (g278, g438, g1199) and the last is found in the regulatory group (g1704). In Table 5, singleton cluster B6={g1105} (a gene in the unknown function category) is noteworthy for its pattern of strong down-regulation for the first three time stages and then moderate up-regulation in the late abstinence stage. Also, oddly, singleton cluster B2={g1024} (having unknown function) is strongly up-regulated in morphine in the withdrawal stage. Whether this occurrence reflects a data anomaly or a real expression response requires further investigation. We observe that cluster B3 in Table 5 consists almost completely of genes with image artifacts (including one negative control gene, g1063). Gene g633, a member of the energy-metabolism group, is the only exception. Singleton cluster B7={g1073} is also in the image artifact category. Cluster B8, a large cluster, is a mixed lot consisting of genes associated with energy metabolism and structural function and one spot on the image artifact list (g1698).

Tables 3 and 5 pose a contrast in clustering of differential expression profiles. Table 3 clusters genes that have similar gene expression profiles *across the seven treatment combinations* as listed in Table 1, irrespective of how these profiles relate to time stage or morphine-placebo treatment. On the other hand, Table 5 clusters genes on the basis of similar differential expression patterns between morphine and placebo *across the four time*

stages. Matching the cluster memberships in Tables 3 and 5, we see that cluster A5 in Table 3 and cluster B3 in Table 5 share three gene spots having image artifacts (g1063, g1257, g1649), the first being a negative control. Several of the singleton clusters also match up. Cluster A2 in Table 3 and cluster B8 in Table 5 share two genes (g438, g870).

Acknowledgements

We acknowledge with thanks the financial support provided for this research by the National Institutes of Health grants CA89756-01, HL66795-02 (Lee), and the Natural Sciences and Engineering Research Council of Canada (Whitmore).

References

- Eisen, M. B. and Brown, P. O. (1999). DNA arrays for analysis of gene expression, *Methods in Enzymology*, **303**:179-205.
- JMP Statistical Discovery Software, SAS Institute, (2000).
- Kerr, M.K., Martin, M. and Churchill, G.A. (2000). Analysis of variances for gene expression microarray data, *Journal of Computational Biology*, **7**: 819-837.
- Kerr, M.K. and Churchill, G.A. (2001). Experimental design issues for gene expression microarrays. *Biostatistics*, **2**: 183-201.
- Lee M.-L. T., Kuo, F. C., Whitmore, G. A. and Sklar, J. L. (2000). Importance of replication in microarray gene expression studies: Statistical methods and evidence from repetitive cDNA hybridizations, *Proceedings of the National Academy of Sciences*, **97**, 9834-9839.
- Lee, M.-L. T., Lu, W, Whitmore, G. A. and Beier, D. (2002a). Models for microarray gene expression data, *Journal of Biopharmaceutical Statistics*, **12**, 1-19.
- Lee M.-L. T. and Whitmore G. A. (2002b). Power and sample size for DNA microarray studies, *Statistics in Medicine*, (In Press).

- Loguinov, A, Anderson, L, Crosby, G. J, Yukhananov, R. Y. (2001). Gene expression profile following acute administration of morphine, *Physiological Genomics*, **6**, 169-181.
- Wolfinger, R. D., Gibson, G., Wolfinger, E. D., Bennett, L., Hamadeh, H., Bushel, P., Afshari, C., Paules, R. (2001). Assessing gene significance from cDNA microarray expression data via mixed models. *Journal of Computational Biology*, **8**, 625-637.
- Yukhananov, R. Y., Klodt, P. M, Shemanov, A. Y, Maisky, A and Nyberg, F. (1994). Decreased concentration of dynorphin A in striatum and increased activity of dynorphin converting enzyme in the spinal cord following repeated morphine treatment, *Regulatory Peptide*, **50**:235-236.

Received July 29, 2002; accepted August 21, 2002

Mei-Ling Ting Lee, Ph.D.
Channing Laboratory
Brigham and Women's Hospital
Harvard Medical School
181 Longwood Avenue
Boston, MA, 02115 USA
Email: meiling@channing.harvard.edu

G. A. Whitmore, Ph.D.
Faculty of Management
McGill University,
1001 Sherbrooke St. West
Montreal, Quebec H3A 1G5, CANADA
Email: george.whitmore@mcgill.ca

Rus Yukhananov, MD, PhD.
Neurogenomic Laboratory
Anesthesia Research, MRB 611
Brigham and Women's Hospital
Harvard Medical School
75 Francis St., Boston MA 02115 USA
Email: rryukhan@zeus.bwh.harvard.edu