Statistical Applications in Genetics and Molecular Biology

Volume 6, *Issue* 1 2007 *Article* 20

Experimental Design for Two-Color Microarrays Applied in a Pre-Existing Split-Plot Experiment

G. A. Milliken* K. A. Garrett†

S. E. Travers‡

*Kansas State University, milliken@ksu.edu

†Kansas State University, kgarrett@ksu.edu

‡Kansas State University, travers@ksu.edu

Experimental Design for Two-Color Microarrays Applied in a Pre-Existing Split-Plot Experiment*

G. A. Milliken, K. A. Garrett, and S. E. Travers

Abstract

Microarray applications for the study of gene expression are becoming accessible for researchers in more and more systems. Applications from field or laboratory experiments are often complicated by the need to superimpose sample pairing for two-color arrays on experimental designs that may already be complex. For example, split-plot designs are commonly used in biological systems where experiments involve two types of treatments that are not readily applied at the same scale. We demonstrate how effects that are confounded with arrays can still be estimated when there is sufficient replication. To illustrate, we evaluate three methods of sample pairing superimposed on a split-plot design with two treatments, deriving the variance associated with parameter estimates for each. Design A has levels of the whole plot treatment paired on the same microarray within a level of the subplot treatment. Design B has crossed levels paired on the same microarray. Design C has levels of the treatment applied to subplots paired on the same microarray within a whole plot. Designs A and B have lower variance than design C for comparing the levels of the whole plot treatment. Designs B and C have lower variance for comparing the levels of the subplot treatment and design C has lower variance for comparing the levels of the subplot treatment within each level of the whole plot treatment. We provide SAS code for the analyses of variance discussed.

KEYWORDS: ecological genomics, experimental design, gene expression, microarray analysis, split-plot design

---

*We thank SAGMB reviewers and Z. Tang for comments that led to improvements in this work and members of the KSU Ecological Genomics community for inspiring discussions. It's also a pleasure to acknowledge support by the Ecological Genomics Initiative of Kansas through U.S. National Science Foundation Grant No. EPS-0236913 with matching funds from the Kansas Technology Enterprise Corporation, by the Office of Science (Program in Ecosystem Research), U.S. Department of Energy, Grant No. DE-FG02-04ER63892, by the NSF under Grants DBI-0421427, DEB-0130692, DEB-0516046, and EF-0525712, and by the NSF Long Term Ecological Research Program at Konza Prairie. This is contribution 07-93-J of the Kansas State Experiment Station.

1. Introduction

Evaluating gene expression using microarrays generally involves a two step process. The first step is to use a designed experiment or ecological experiment to obtain samples, such as plant tissues. The second step is to superimpose a microarray pairing design on the ecological design in order to extract gene expression data. The use of two color microarrays on samples from an ecological design involving blocking factors poses problems because important effects may be confounded with blocks and/or microarrays. The statistical consequences of placing samples from an ecological design on two-color microarrays can be evaluated in some situations. This paper evaluates the statistical properties of three two-color microarray pairing strategies superimposed onto samples obtained when the ecological experiment consists of a split-plot design. The variances of the treatment effects from a split-plot ecological experiment are evaluated to provide information as to which strategy for using the two-color microarray provides the smallest variance for the most important comparison.

Recent advances in genomics have led to the application of new technologies and data measurement into a wide range of conceptual fields. For example, high throughput estimates of gene expression and descriptions of gene sequences for natural populations of organisms has led to the development of fields such as ecological and evolutionary genomics (Feder and Mitchell-Olds 2003; Purugganan and Gibson 2003). Microarrays have been particularly effective for ecologists and evolutionary biologists interested in mining the transcriptome of organisms for candidate genes involved in ecologically important processes and in understanding genetic constraints (Thomas and Klaper 2004). As the use of microarrays in new biological disciplines becomes easier and less expensive, there will be an increasing trend for application of microarrays to study gene expression in pre-existing experimental designs. Transcriptional profiling is a useful tool for measuring the responses of organisms to stress, environmental variation and conflicting demands. Many long-term ecological and field based studies would benefit from increased understanding of these responses at the genomic level (Garrett et al. 2006; Travers et al. 2007). Ecotoxicological studies measuring responses of organisms to environmental pollutants and stressors have already benefited from transcriptional profiling through the use of microarrays (e.g., Greer et al. 2001).

Much attention has been given to the design and analysis of microarray applications for the study of gene expression (e.g., Chu et al. 2002; Jin et al. 2001; Kerr and Churchill 2001; Wolfinger et al. 2001), but the context has generally been new experiments designed specifically for microarray analyses (e.g., Blum et al. 2004) using a completely randomized design. But as microarrays are applied more widely to evaluate gene expression differences between treatments

where the treatments are applied in a designed experiment, new designs that merge the microarray pairing design with the design producing the treatments will be needed. For example, ecological field experiments are often implemented in a split-plot design since it may be difficult to realistically apply treatments at the same scales (e.g., Fay et al. 2002). The design and treatment structures of the pre-existing experiment must be used to construct an appropriate model that also includes the microarray pairing and, for two-color microarrays, the two dye colors. There are many ways to superimpose the microarray pairing with the designed experiment. Moreover, each interface influences the model in its own way.

Research in microarray experimental design has begun to address more complex design structures (e.g., Bueno Filho et al. 2006; Glonek and Solomon 2004; Kerr 2006; Rosa et al. 2005; Tempelman 2005; Wit et al. 2005), but most researchers have not dealt with the more complex features of an ecological design used to collect the samples when evaluating gene expressions. Some types of confounding discussed in microarray design literature can generate a split-plot design. In other cases an incomplete block design analysis may combine information for an effect that is confounded with some microarrays and not confounded with other microarrays. When an incomplete block design is used and blocks are included as a random effect, the treatment effects are estimated by combining intra-block and inter-block information. The split-plot design is a special case of incomplete block designs where one chooses to confound either a main effect or an interaction with blocks and then uses inter-block information to obtain information about the confounded effect. When the microarray design is superimposed on the ecological design, the confounding becomes more complicated. Here, we illustrate how to estimate effects and their standard errors in a split-plot design when there is sufficient replication, even when an effect is confounded with microarrays. We use the fact that split-plot designs are incomplete block designs with some effect(s) confounded with blocks. When the microarray design is superimposed onto the split-plot design, other effects may be confounded with microarrays. The between-array information can be used to obtain estimates of those effects that are confounded with microarrays. The variances of the estimated effects that are confounded with blocks and microarrays are larger than those that are not confounded with blocks and microarrays, as is illustrated in the discussion of the various designs below.

2. 1. Evaluation of variance in three designs

To demonstrate the basic analysis of the split-plot design, consider an experiment that has two types of treatments, each with two levels, where the design structure is a split-plot design with six blocks with the levels of Treatment 1 as the whole

plot factor and the levels of Treatment 2 as the subplot factor. A model that can be used to describe data from this split-plot design is

$y_{ijkl} = \mu_{ik} + b_j + w_{ij} + \varepsilon_{ijkl}$

where $i = 1, 2,\ j = 1, 2, ..., 6,\ k = 1, 2,\ l = 1, 2$

μ_{ik} is the mean response from level i of Treatment 1 and level k of treatment 2

b_j is the random block effect with distribution $N(0, \sigma^2_{blk})$

w_{ij} is the wholeplot effect assigned to the ith level of Treatment 1 within the jth block with distribution $N(0, \sigma^2_{wp})$

ε_{ijkl} is the random effect of the residual with distribution $N(0, \sigma^2_{\varepsilon})$

The analysis of variance table for this split-plot design is in Table 1, which includes the degrees of freedom, expected mean squares with noncentrality parameters for each of the fixed effects denoted by ϕ^2_{trt1}, say, for treatment 1 effects and the SAS Proc Mixed code needed to fit the model. The whole plot error term is computed as the block by Treatment 1 interaction.

Table 1. Analysis of Variance Table and SAS Proc Mixed code for Split-plot Design

Source	df	EMS
Blocks	5	$\sigma^2_{\varepsilon} + 2\sigma^2_{w} + 4\sigma^2_{blk}$
trt1	1	$\sigma^2_{\varepsilon} + 2\sigma^2_{w} + \phi^2_{trt1}$
trt2	1	$\sigma^2_{\varepsilon} + \phi^2_{trt2}$
trt1*trt2	1	$\sigma^2_{\varepsilon} + \phi^2_{trt1*trt2}$
blk*trt1	5	$\sigma^2_{\varepsilon} + 2\sigma^2_{w}$
Residual	10	σ^2_{ε}

```
proc mixed;
class blk trt1 trt2;
model y=trt1|trt2/ddfm=kr;
random blk trt1*blk;
```

The next step in this process is to use microarrays to evaluate gene expression. There are several methods one could use to apply the microarrays to the treatments within each block. Figure 1 displays the split-plot design with the levels of treatments 1 and 2 as well as candidate assignment of treatments to microarrays and colors. The arrows indicate the two samples that are paired on an individual microarray, with the head of the arrow indicating the sample labeled with the red dye (Cy5) and the tail of the arrow indicating the sample labeled with

the green dye (Cy3). Two microarrays are used per block for a total of 12 microarrays in the whole experiment (at one sampling date). These diagrams illustrate the assignment of treatment levels as the same within each whole plot or block to make the assignment of microarrays clearer, though in a real experiment the placement of treatment levels would be randomized within whole plots and blocks. Without the microarrays the design is a split-plot with two sizes of experimental units, the whole plot (the entity to which the levels of treatment 1 were applied) and the subplot (the entity to which the levels of the treatment 2 were applied). By including the microarrays, additional experimental units are generated and added to the random effects of the model. The two dye colors add another factor to the fixed effects, so the treatment structure is a three way factorial arrangement with the levels of color crossed with the levels of treatment 1 crossed with the levels of treatment 2. The addition of dye color to the treatment structure changes the overall design as there are not eight treatment combinations and there are blocks of size four, thus the resulting design is an incomplete block design where four of the treatment combinations are included in a block. In particular, four of the treatment combinations are included in one set of blocks and the other four treatment combinations are included in another set of blocks. The structure of the treatment assignment to blocks depends on the strategy of assigning the microarrays as indicated in Figure 1.

In general, a model that could be used to describe data from one of these designs is

$$y_{ijklm} = \mu_{ikm} + b_j + w_{ij} + a_{l(j)} + \varepsilon_{ijklm}$$

where $i = 1,2,\ j = 1,2,...,6,\ k = 1,2,\ l = 1,2,\ m = R,G$

μ_{ikm} is the mean response from level i of Treatment 1, level k of Treatment 2 and color m

b_j is the random block effect with distribution $N(0, \sigma^2_{blk})$

w_{ij} is the wholeplot effect assigned to the *ith* level of Treatment 1 within the *jth* block with distribution $N(0, \sigma^2_{wp})$

$a_{l(j)}$ is the random effect of the *lth* array within the *jth* block with distribution $N(0, \sigma^2_{array})$

ε_{ijklm} is the random effect of the residual with distribution $N(0, \sigma^2_{\varepsilon})$

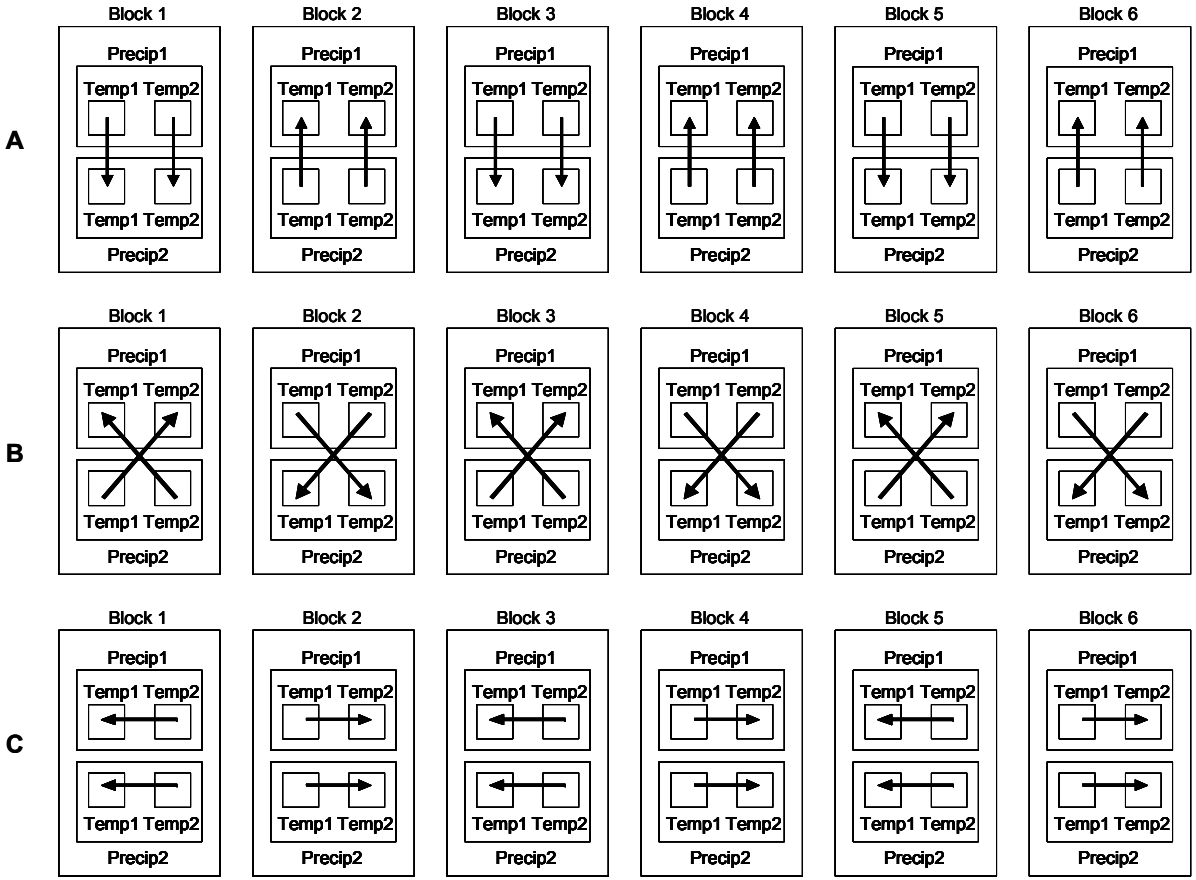

Figure 1. Schematic of pairing of microarray samples from a split-plot design with two levels of each of two treatments, where the treatments in this example are precipitation ('precip') applied at the whole plot level and temperature ('temp') applied at the subplot level. Arrows indicate two samples that are paired for hybridization on the same microarray, with the head of the arrow indicating one dye and the tail of the arrow indicating the other dye.

Design A. Pairs are assigned within a level of Treatment 2 (temp) at the subplot level and across levels of Treatment 1 (precip) at the whole plot level. **Design B**. Pairs are assigned across levels of Treatment 1 and across levels of Treatment 2. **Design C**. Pairs are assigned within a level of Treatment 1 (precip) at the whole plot level and across levels of Treatment 2 (temp) at the subplot level.

The subscripts corresponding to the random effects can change depending on the microarray assignment pattern. Not all subscripts are used since only four of the eight treatment combinations occur within a given block. Further, the set of blocks can be divided into two types of blocks where the treatment combinations are common to the blocks within a block type. For example, the two types of blocks within each of the designs in Figure 1 are blocks 1, 3, and 5 of block type I and blocks 2, 4, and 6 of block type II. The block types are generated by the assignment of the colors of the microarrays and each assignment method has its own effect on the resulting model and resulting analysis. There are several mean comparisons of interest, including the main effect of the levels of Treatment 1 ($\mu_{1\bullet\bullet} - \mu_{2\bullet\bullet}$), the main effect of the levels of Treatment 2 ($\mu_{\bullet 1\bullet} - \mu_{\bullet 2\bullet}$), the main effect of color ($\mu_{\bullet\bullet R} - \mu_{\bullet\bullet G}$), the interaction between the levels of Treatment 1 and the levels of Treatment 2 ($\mu_{11\bullet} - \mu_{21\bullet} - \mu_{12\bullet} + \mu_{22\bullet}$), comparisons of the levels of Treatment 1 at each level of Treatment 2 ($\mu_{11\bullet} - \mu_{21\bullet}$, $\mu_{12\bullet} - \mu_{22\bullet}$), comparisons of the levels of Treatment 2 at each level of Treatment 1 ($\mu_{11\bullet} - \mu_{12\bullet}$, $\mu_{21\bullet} - \mu_{22\bullet}$), etc. Because of the incomplete block design aspect of this study, information about the comparisons of interest can come from within block comparisons, between block comparisons, between whole plot comparisons, and between microarray comparisons. It is imperative that the analysis exhibits the variability associated with each estimated comparison of interest.

2.2. Analysis of Design A

Four of those combinations are in block type I (blocks 1, 3, and 5) and four are in block type II (blocks 2, 4, and 6). Thus the resulting design is an incomplete block design with treatment combinations (1,1,G), (1,2,G), (2,1,R) and (2,2,R) in block type I and treatment combinations (1,1,R), (1,2,R), (2,1,G), and (2,2,G) in block type II. A model that can be used to describe the data from these treatment combinations in the two types of blocks is

$$y_{ijklm} = \mu_{ilm} + b_{jk} + w_{ijk} + a_{jkl} + \varepsilon_{ijklm}$$

where $i = 1, 2$ (levels of Treatment 1)

$j = I, II$ (types of blocks)

$k = 1, 2, 3$ (block within type of block)

$l = 1, 2$ (levels of Treatment 2),

$m = R, G$ (Color)

There are only four treatment combinations per block type, so not all combinations of *i, l,* and *m* occur within each block. The block effect, the whole

plot effect and the array effect have subscripts indicating which block type and block within a block type each belongs.

Table 2 contains the models for the treatment combinations in the *kth* block of Type I blocks.

Table 2. Models for observations from the *k*th block of a Type I block of design A	
	Treatment 2 Level 1
Treatment 1 Level 1	$y_{1Ik1G} = \mu_{11G} + b_{Ik} + w_{1Ik} + a_{Ik1} + \varepsilon_{1Ik1G}$
Treatment 1 Level 2	$y_{2Ik1R} = \mu_{11R} + b_{Ik} + w_{2Ik} + a_{Ik1} + \varepsilon_{2Ik1R}$
	Treatment 2 Level 2
Treatment 1 Level 1	$y_{1Ik2G} = \mu_{12G} + b_{Ik} + w_{1Ik} + a_{Ik2} + \varepsilon_{1Ik2G}$
Treatment 1 Level 2	$y_{2Ik2R} = \mu_{22R} + b_{Ik} + w_{2Ik} + a_{Ik2} + \varepsilon_{2Ik2R}$

This is a two-way treatment structure, levels of Treatment 1 by levels of Treatment 2, when the levels of color are ignored. The effect of Treatment 1 in the *k*th block of block type I is

$$T_{1Ik} = \tfrac{1}{2}\left[y_{1Ik1G} + y_{1Ik2G} - y_{2Ik1R} - y_{2Ik2R}\right]$$

and the variance of T_{1Ik} is

$$Var(T_{1Ik}) = \tfrac{1}{4}Var(2w_{1Ik} - 2w_{2Ik} + \varepsilon_{1Ik1G} + \varepsilon_{1Ik2G} - \varepsilon_{2Ik1R} - \varepsilon_{2Ik2R})$$
$$= \sigma_\varepsilon^2 + 2\sigma_{wp}^2$$

There are three type I blocks, so there are three values T_{1Ik}, k=1,2,3, and the sample variance of these three values provides two degrees of freedom for estimating $\sigma_\varepsilon^2 + 2\sigma_{wp}^2$.

The effect of treatment 2 in the *kth* block of block type I is

$$T_{2Ik} = \tfrac{1}{2}\left[y_{1Ik1G} + y_{2Ik1R} - y_{1Ik2G} - y_{2Ik2R}\right]$$

and the variance of T_{2Ik} is

$$Var(T_{2Ik}) = \tfrac{1}{4}Var(2a_{Ik1} - 2a_{Ik2} + \varepsilon_{1Ik1G} - \varepsilon_{1Ik2G} + \varepsilon_{2Ik1R} - \varepsilon_{2Ik2R})$$
$$= \sigma_\varepsilon^2 + 2\sigma_{array}^2$$

There are three type I blocks, so there are three values T_{2Ik}, k=1,2,3, and the sample variance of these three values provides two degrees of freedom for estimating $\sigma_\varepsilon^2 + 2\sigma_{array}^2$

The interaction of Treatment 1 by Treatment 2 in the *k*th block of block type I is

$T_1 \times T_{2Ik} = \frac{1}{2}\left[y_{1Ik1G} + y_{2Ik2R} - y_{1Ik2G} - y_{2Ik1R} \right]$

and the variance of $T_1 x T_{2Ik}$ is

$$Var(T_1 \times T_{2Ik}) = \tfrac{1}{4} Var(\varepsilon_{1Ik1G} - \varepsilon_{1Ik2G} + \varepsilon_{2Ik2R} - \varepsilon_{2Ik1R})$$
$$= \sigma_\varepsilon^2$$

There are three type I blocks, so there are three values $T_1 x T_{2Ik}$, k=1,2,3 and the sample variance of these three values provides two degrees of freedom for estimating σ_ε^2

The block I effect is $Blk_{Ik} = \frac{1}{2}\left[y_{1Ik1G} + y_{2Ik2R} + y_{1Ik2G} + y_{2Ik1R} \right]$ and the variance is

$$Var(Blk_{Ik}) = \tfrac{1}{4} Var(4b_{Ik} + 2w_{1Ik} + 2w_{2Ik} + 2a_{Ik1} + 2a_{Ik2} + \varepsilon_{1Ik1G} + \varepsilon_{1Ik2G} + \varepsilon_{2Ik1R} + \varepsilon_{2Ik2R})$$
$$= \tfrac{1}{4}\left[16\sigma_b^2 + 8\sigma_{wp}^2 + 8\sigma_{array}^2 + 4\sigma_\varepsilon^2 \right]$$
$$= \sigma_\varepsilon^2 + 2\sigma_{wp}^2 + 2\sigma_{array}^2 + 4\sigma_b^2$$

There are three type I blocks, so there are three values Blk_{Ik}, k=1,2,3, and the sample variance of these three values provides two degrees of freedom for estimating $\sigma_\varepsilon^2 + 2\sigma_{wp}^2 + 2\sigma_{array}^2 + 4\sigma_b^2$.

Table 3 contains the models for the treatment combinations in the *kth* block of Type II blocks.

Table 3. Models for observations from the *kth* block of a Type II block of design A

	Treatment 2 Level 1
Treatment 1 Level 1	$y_{1IIk1R} = \mu_{11R} + b_{IIk} + w_{1IIk} + a_{IIk1} + \varepsilon_{1IIk1R}$
Treatment 1 Level 2	$y_{2IIk1G} = \mu_{11G} + b_{IIk} + w_{2IIk} + a_{IIk1} + \varepsilon_{2IIk1G}$
	Treatment 2 Level 2
Treatment 1 Level 1	$y_{1IIk2R} = \mu_{12R} + b_{IIk} + w_{1IIk} + a_{IIk2} + \varepsilon_{1IIk2R}$
Treatment 1 Level 2	$y_{2IIk2G} = \mu_{22G} + b_{IIk} + w_{2IIk} + a_{IIk2} + \varepsilon_{2IIk2G}$

The treatment effects and variances from the Type II blocks can be determined similarly as was done for the Type I blocks. The Type II effects and variances are included in Table 4.

Table 4. Effects and variances for Type II blocks	
Effect	Variance
$T_{1IIk} = \frac{1}{2}\left[y_{1IIk1R} + y_{1IIk2R} - y_{2IIk1R} - y_{2IIk2G}\right]$	$\sigma^2_\varepsilon + 2\sigma^2_{wp}$
$T_{2IIk} = \frac{1}{2}\left[y_{1IIk1R} + y_{2IIk1G} - y_{1IIk2R} - y_{2IIk2G}\right]$	$\sigma^2_\varepsilon + 2\sigma^2_{array}$
$T_1 \times T_{2IIk} = \frac{1}{2}\left[y_{1IIk1R} + y_{2IIk2G} - y_{1IIk2R} - y_{2IIk1G}\right]$	σ^2_ε
$Blk_{IIk} = \frac{1}{2}\left[y_{1IIk1R} + y_{2IIk2G} + y_{1IIk2R} + y_{2IIk1G}\right]$	$\sigma^2_\varepsilon + 2\sigma^2_{wp} + 2\sigma^2_{array} + 4\sigma^2_b$

The sum of squares due to blocks within block type is computed as the sample variance of the Blk_{Ik}, k=1,2,3, pooled with the sample variance of the Blk_{IIk}, k=1,2,3, which is based on 4 degrees of freedom.

The sum of squares for the whole plot error is computed as the sample variance of the T_{1Ik}, k=1,2,3, pooled with the sample variance of the T_{1IIk}, k=1,2,3, which is based on 4 degrees of freedom.

The sum of squares for the microarray error is computed as the sample variance of the T_{2Ik}, k=1,2,3, pooled with the sample variance of the T_{2IIk}, k=1,2,3, which is based on 4 degrees of freedom.

The residual sum of squares is computed as the sample variance of the T_1xT_{2Ik}, k=1,2,3, pooled with the sample variance of the T_1xT_{2IIk}, k=1,2,3, which is based on 4 degrees of freedom.

The effect of Treatment 1 from block type I is $E(T_{1Ik}) = \frac{1}{2}[\mu_{11G} + \mu_{12G} - \mu_{21R} - \mu_{22R}]$, which is aliased with the G-R Color effect.

The effect of Treatment 1 from block type II is $E(T_{1IIk}) = \frac{1}{2}[\mu_{11R} + \mu_{12R} - \mu_{21G} - \mu_{22G}]$, which is aliased with the R-G Color effect.

The mean of these two effects provides the main effect for the levels of treatment 1 as $E[\frac{1}{2}(T_{1Ik} + T_{1IIk})] = \bar{\mu}_{1\bullet\bullet} - \bar{\mu}_{2\bullet\bullet}$.

The effect of Treatment 2 from block type I is $E[T_{2Ik}] = \frac{1}{2}[\mu_{11G} + \mu_{21R} - \mu_{12G} - \mu_{22R}]$, which is aliased with part of the Treatment 1 by Color interaction. The effect of Treatment 2 from block type II is $E[T_{2IIk}] = \frac{1}{2}[\mu_{11R} + \mu_{21G} - \mu_{12R} - \mu_{22G}]$ and is aliased with another part of the Treatment 1 by Color interaction.

The sum of these two effects provides the main effect for the levels of Treatment 2 as, $E[\frac{1}{2}(Tr_{2a} + Tr_{2b})] = \bar{\mu}_{\bullet 1\bullet} - \bar{\mu}_{\bullet 2\bullet}$

The difference between T_{1Ik} and T_{1IIk} is the main effect of Color as $E[\frac{1}{2}(Tr_{1a} - Tr_{1b})] = \bar{\mu}_{\bullet\bullet G} - \bar{\mu}_{\bullet\bullet R}$.

The difference between T_{2Ik} and T_{2IIk} is an estimate of the Treatment 1 by Treatment 2 by Color interaction as $E[T_{2Ik} - T_{2IIk}] = [\mu_{11G} + \mu_{22G} - \mu_{12G} - \mu_{21G}] - [\mu_{11R} + \mu_{22R} - \mu_{12R} - \mu_{21R}]$.
The effects of Treatment 2 within a block type are confounded with microarray. The interaction effect of Treatment 1 with Treatment 2 from block type I is $E[T_1 x T_{2Ik}] = \mu_{11G} + \mu_{22R} - \mu_{21R} - \mu_{12G}$ and from block type II $E[T_1 x T_{2IIk}] = \mu_{11R} + \mu_{22G} - \mu_{21G} - \mu_{12R}$
The mean of these two interaction effects provides the Treatment 1 by Treatment 2 interaction as $E[\frac{1}{2}(T_1 x T_{2Ik} + T_1 x T_{2IIk})] = \bar{\mu}_{11\bullet} + \bar{\mu}_{22\bullet} - \bar{\mu}_{21\bullet} - \bar{\mu}_{12\bullet}$
The mean difference of these two interaction effects provides the Treatment 2 by Color interaction as $E[\frac{1}{2}(T_1 x T_{2Ik} - T_1 x T_{2IIk})] = \bar{\mu}_{\bullet 1G} + \bar{\mu}_{\bullet 2R} - \bar{\mu}_{\bullet 1R} - \bar{\mu}_{\bullet 2G}$
Finally, the mean difference of the block I effect and the block II effect provides the Treatment 1 by Color interaction as

$$E[\tfrac{1}{2}(Blk_{Ik} - Blk_{IIk})] = \frac{\mu_{11G} + \mu_{12G} + \mu_{21R} + \mu_{22R}}{4} - \frac{\mu_{11R} + \mu_{12R} + \mu_{21G} + \mu_{22G}}{4}$$
$$= \bar{\mu}_{1\bullet G} + \bar{\mu}_{2\bullet R} - \bar{\mu}_{1\bullet R} - \bar{\mu}_{2\bullet G}$$

The whole plot error for the split-plot design was computed as the block*treatment interaction. Now the blocks are split into two block types, thus the appropriate code for the whole plot error in SAS Proc Mixed is `trt1*blk(blktype)`. The treatment 2 effect within each block type is confounded with microarray and the microarray is another type of split-plot in each block. If the levels of Treatment 2 were identical (no Treatment 2 effect) then the microarray error term would be computed as array*blk(blktype), thus, the appropriate code for the microarray error term in SAS Proc Mixed is `array*blk(blktype)`.
The SAS Proc Mixed code (Gibson and Wolfinger 2004; Jin et al. 2001; Littell et al. 2006; Wolfinger et al. 2001) needed to fit this model is

```
proc mixed;
class blk trt1 trt2 array color blktype;
model y=trt1|trt2 blktype color trt2*color trt2*blktype;
random blk(blktype) trt1*blk(blktype) array*blk(blktype);
```

where blktype is the trt1*color interaction, trt2*blktype is the trt1*trt2*color interaction, blk(blktype) is the block within block type random effect, trt1*blk(blktype) is the wholeplot error and array*blk(blktype) is the microarray random effect.

The above information is summarized in Table 5

Table 5. Analysis of Variance Table for Design A		
Source	df	EMS
trt1	1	$\sigma^2_\varepsilon + 2\sigma^2_w + \phi^2_{trt1}$
trt2	1	$\sigma^2_\varepsilon + 2\sigma^2_{array} + \phi^2_{trt2}$
trt1*trt2	1	$\sigma^2_\varepsilon + \phi^2_{trt1*trt2}$
blktype	1	$\sigma^2_\varepsilon + 2\sigma^2_w + 2\sigma^2_{array} + 4\sigma^2_{blk} + \phi^2_{trt1*color}$
color	1	$\sigma^2_\varepsilon + 2\sigma^2_w + \phi^2_{color}$
trt2*color	1	$\sigma^2_\varepsilon + \phi^2_{trt2*color}$
trt2*blktype	1	$\sigma^2_\varepsilon + 2\sigma^2_{array} + \phi^2_{trt1*trt2*color}$
blk(blktype)	4	$\sigma^2_\varepsilon + 2\sigma^2_w + 2\sigma^2_{array} + 4\sigma^2_{blk}$
blk*trt1*color(blktype)	4	$\sigma^2_\varepsilon + 2\sigma^2_w$
trt2*array*blk(blktype)	4	$\sigma^2_\varepsilon + 2\sigma^2_{array}$
Residual	4	σ^2_ε

The sources trt1*trt2 and trt2*color are tested by the residual, the sources trt1 and color are between whole plot effects and are tested by blk*trt1(blktype), the sources trt2 and trt1*trt2*color are between microarray effects and are tested by array*blk(blktype) and trt1*color which is designated as blktype is a between block effect and is tested by blk(blktype).

Table 6 contains the parameters of interest and the variances of their estimates. The main effects of Treatment 1 and Color are between whole plot effects. The main effect of Treatment 2 is a between microarray effect. The comparisons of the levels of Treatment 1 at each level of Treatment 2 have a variance that involves the residual and whole plot variance components. The variances of the comparisons of the levels of Treatment 2 at each level of Treatment 1 involve the residual and microarray variance components.

Table 6. Parameters of Interest and Variances of Estimates of Parameters of Interest	
Parameter	Variance of Estimate of the Parameter
$\bar{\mu}_{1\bullet\bullet} - \bar{\mu}_{2\bullet\bullet}$	$\left[\sigma^2_\varepsilon + 2\sigma^2_{wp}\right]/6$
$\bar{\mu}_{\bullet 1\bullet} - \bar{\mu}_{\bullet 2\bullet}$	$\left[\sigma^2_\varepsilon + 2\sigma^2_{array}\right]/6$
$\bar{\mu}_{\bullet\bullet R} - \bar{\mu}_{\bullet\bullet G}$	$\left[\sigma^2_\varepsilon + 2\sigma^2_{wp}\right]/6$
$\bar{\mu}_{11\bullet} - \bar{\mu}_{21\bullet}$	$\left[\sigma^2_\varepsilon + \sigma^2_{wp}\right]/3$

$\bar{\mu}_{12\bullet} - \bar{\mu}_{22\bullet}$	$\left[\sigma_{\varepsilon}^2 + \sigma_{wp}^2\right]/3$
$\bar{\mu}_{11\bullet} - \bar{\mu}_{12\bullet}$	$\left[\sigma_{\varepsilon}^2 + \sigma_{array}^2\right]/3$
$\bar{\mu}_{21\bullet} - \bar{\mu}_{22\bullet}$	$\left[\sigma_{\varepsilon}^2 + \sigma_{array}^2\right]/3$

2.3. Analysis of Design B

Design B is similar to design A in that there are two block types and each block type contains four treatment combinations. The four treatment combinations in block type I (blocks 1, 3, and 5) are [(1,1,R),(1,2,R),(2,1,G), (2,2,G)] and the four treatment combinations in block type II (blocks 2, 4, and 6) are [(1,1,G),(1,2,G),(2,1,R),(2,2,R)]. A model that can be used to describe the data from these treatment combinations in the two types of blocks is

$y_{ijklmn} = \mu_{ilm} + b_{jk} + w_{ijk} + a_{jkn} + \varepsilon_{ijklmn}$

where $i = 1, 2$ (levels of Treatment 1)

$j = I, II$ (types of blocks)

$k = 1, 2, 3$ (block within type of block)

$l = 1, 2$ (levels of Treatment 2),

$m = R, G$ (Color)

$n = 1, 2$ (microarray)

An important difference between the model for design B and the model for design A is that arrays and levels of Treatment 2 were confounded and a single subscript was needed. But for design B, a subscript is needed to denote the different arrays in addition to the one needed to denote the levels of Treatment 2. There are only four treatment combinations per block type, so not all combinations of *i, l,* and *m* occur within each block. The block effect, the whole plot effect and the array effect have subscripts indicating to which block type and block within a block type each belongs.

Tables 7 and 8 contain the models for the treatment combinations in the *k*th block of Type I and Type II blocks, respectively, for the design B.

Table 7. Models for observations from the *k*th block of a Type I block of design B	
	Treatment 2 Level 1
Treatment 1 Level 1	$y_{1Ik1R2} = \mu_{11R} + b_{Ik} + w_{1Ik} + a_{Ik2} + \varepsilon_{1Ik1R2}$
Treatment 1 Level 2	$y_{2Ik1G1} = \mu_{11G} + b_{Ik} + w_{2Ik} + a_{Ik1} + \varepsilon_{2Ik1G1}$

	Treatment 2 Level 2
Treatment 1 Level 1	$y_{1Ik2R2} = \mu_{12R} + b_{Ik} + w_{1Ik} + a_{Ik1} + \varepsilon_{1Ik2R1}$
Treatment 1 Level 2	$y_{2Ik2G2} = \mu_{22G} + b_{Ik} + w_{2Ik} + a_{Ik2} + \varepsilon_{2Ik2G2}$

Table 8. Models for observations from the *k*th block of a Type II block of design B

	Treatment 2 Level 1
Treatment 1 Level 1	$y_{1IIk1G1} = \mu_{11G} + b_{IIk} + w_{1IIk} + a_{IIk1} + \varepsilon_{1IIk1G1}$
Treatment 1 Level 2	$y_{2IIk1R2} = \mu_{21R} + b_{IIk} + w_{2IIk} + a_{IIk2} + \varepsilon_{2IIk1R2}$
	Treatment 2 Level 2
Treatment 1 Level 1	$y_{1IIk2G2} = \mu_{12G} + b_{IIk} + w_{1IIk} + a_{IIk2} + \varepsilon_{1IIk2G2}$
Treatment 1 Level 2	$y_{2IIk2R1} = \mu_{22R} + b_{IIk} + w_{2IIk} + a_{IIk1} + \varepsilon_{2IIk2R1}$

The technique utilized for dissecting the models for design A can be employed with the models of design B. That analysis provides the following results. The Blktype effect corresponds to the Treatment 1 by color interaction. The comparison of the levels of Treatment 1 and the levels of color are between whole plot comparisons. Comparisons of the levels of Treatment 2 and the Blktype by Treatment 2 interaction (which is the Treatment 1 by Treatment 2 by Color interaction) are within whole plot and array comparison or residual comparisons. The Treatment 1 by Treatment 2 interaction and the Treatment 2 by Color interaction are between array comparisons. The Blktype comparison is a between block comparison. These comparisons are summarized in the analysis of variance table in Table 9.

Table 9. Analysis of Variance Table for Design B using Type I Sums of Squares to Evaluate Expected Mean Squares

Source	df	EMS
trt1	1	$\sigma^2_\varepsilon + 2\sigma^2_{wp} + \phi^2_{trt1}$
trt2	1	$\sigma^2_\varepsilon + \phi^2_{trt2}$
trt1*trt2	1	$\sigma^2_\varepsilon + 2\sigma^2_{array} + \phi^2_{trt1*trt2}$
blktype (trt1xcolor)	1	$\sigma^2_\varepsilon + 2\sigma^2_{wp} + 2\sigma^2_{array} + 4\sigma^2_{blk} + \phi^2_{trt1*color}$

color	1	$\sigma_\varepsilon^2 + 2\sigma_{wp}^2 + \phi_{color}^2$
trt2*color	1	$\sigma_\varepsilon^2 + 2\sigma_{array}^2 + \phi_{trt2*color}^2$
trt2*blktype	1	$\sigma_\varepsilon^2 + \phi_{trt1*trt2*color}^2$
blk(blktype)	4	$\sigma_\varepsilon^2 + 2\sigma_w^2 + 2\sigma_{array}^2 + 4\sigma_{blk}^2$
blk*trt1*(blktype)	4	$\sigma_\varepsilon^2 + 2\sigma_w^2$
array*blk(blktype)	4	$\sigma_\varepsilon^2 + 2\sigma_{array}^2$
Residual	4	σ_ε^2

The SAS Proc Mixed code that can be used to fit the above model is

```
proc mixed;
class blk trt1 trt2 array color blktype;
model y=trt1|trt2 blktype color trt2*color trt2*blktype;
random blk(blktype) trt1*blk(blktype) array*blk(blktype);
```

Table 10 contains the parameters of interest and the variances of estimates of those parameters. The variances of estimates of the parameters, $\bar{\mu}_{1\bullet\bullet} - \bar{\mu}_{2\bullet\bullet}$, $\bar{\mu}_{\bullet\bullet R} - \bar{\mu}_{\bullet\bullet G}$, $\bar{\mu}_{11\bullet} - \bar{\mu}_{12\bullet}$, and $\bar{\mu}_{21\bullet} - \bar{\mu}_{22\bullet}$ are identical to those in Table 6 for design A. The variances of the estimates of the other parameters are different.

Table 10. Parameters of Interest and Variances of Estimates of Parameters of Interest based on Design B

Parameter	Variance of Estimate of the Parameter
$\bar{\mu}_{1\bullet\bullet} - \bar{\mu}_{2\bullet\bullet}$	$\left[\sigma_\varepsilon^2 + 2\sigma_{wp}^2\right]/6$
$\bar{\mu}_{\bullet 1\bullet} - \bar{\mu}_{\bullet 2\bullet}$	$\left[\sigma_\varepsilon^2\right]/6$
$\bar{\mu}_{\bullet\bullet R} - \bar{\mu}_{\bullet\bullet G}$	$\left[\sigma_\varepsilon^2 + 2\sigma_{wp}^2\right]/6$
$\bar{\mu}_{11\bullet} - \bar{\mu}_{21\bullet}$	$\left[\sigma_\varepsilon^2 + \sigma_{wp}^2 + \sigma_{array}^2\right]/3$
$\bar{\mu}_{12\bullet} - \bar{\mu}_{22\bullet}$	$\left[\sigma_\varepsilon^2 + \sigma_{wp}^2 + \sigma_{array}^2\right]/3$
$\bar{\mu}_{11\bullet} - \bar{\mu}_{12\bullet}$	$\left[\sigma_\varepsilon^2 + \sigma_{array}^2\right]/3$
$\bar{\mu}_{21\bullet} - \bar{\mu}_{22\bullet}$	$\left[\sigma_\varepsilon^2 + \sigma_{array}^2\right]/3$

For $\bar{\mu}_{\bullet 1\bullet} - \bar{\mu}_{\bullet 2\bullet}$ the variance of the estimate is $\left[\sigma_\varepsilon^2\right]/6$ for design B while it is $\left[\sigma_\varepsilon^2 + 2\sigma_{array}^2\right]/6$ for design A. For $\bar{\mu}_{11\bullet} - \bar{\mu}_{21\bullet}$, and $\bar{\mu}_{12\bullet} - \bar{\mu}_{22\bullet}$ the variance of the estimates is $\left[\sigma_\varepsilon^2 + \sigma_{wp}^2 + \sigma_{array}^2\right]/3$ for design B while it is $\left[\sigma_\varepsilon^2 + \sigma_{wp}^2\right]/3$ for design A. If one uses design B instead of design A the variance of $\bar{\mu}_{\bullet 1\bullet} - \bar{\mu}_{\bullet 2\bullet}$ is smaller while the variances of $\bar{\mu}_{11\bullet} - \bar{\mu}_{21\bullet}$ and $\bar{\mu}_{12\bullet} - \bar{\mu}_{22\bullet}$ are larger.

2.4. Analysis of Design C

Design C is similar to designs A and B in that there are two block types and each block type contains four treatment combinations. The four treatment combinations in block type I (blocks 1, 3, and 5) are [(1,1,R),(1,2,G),(2,1,R), (2,2,G)] and four treatment combinations in block type II (blocks 2, 4, and 6) are [(1,1,G),(1,2,R),(2,1,G),(2,2,R)]. A model that can be used to describe the data from these treatment combinations in the two types of blocks is

$$y_{ijklm} = \mu_{ilm} + b_{jk} + w_{ijk} + a_{ijk} + \varepsilon_{ijklm}$$

where $i = 1, 2$ (levels of Treatment 1)

$j = I, II$ (Types of Blocks)

$k = 1, 2, 3$ (block within type of block)

$l = 1, 2$ (levels of Treatment 2),

$m = R, G$ (Color)

An important difference between the model for design C and the model for design B is that arrays and levels of Treatment 1 were confounded which means the whole plot effect and the array effect are confounded or inseparable. Thus the model can be simplified by combining the whole plot term and the array term into a single term in the model. The simplified model is

$$y_{ijklm} = \mu_{ilm} + b_{jk} + (w+a)_{ijk} + \varepsilon_{ijklm}$$

where $i = 1, 2$ (levels of Treatment 1)

$j = I, II$ (Types of Blocks)

$k = 1, 2, 3$ (block within type of block)

$l = 1, 2$ (levels of Treatment 2),

$m = R, G$ (Color)

where $(w+a)_{ijk}$ denotes the combined whole plot and array effects with variance component $\sigma_{wp+array}^2$. In a lot of situations it is likely that this combined variance component is equal to the sum of the two variance components

or $\sigma^2_{wp+array} = \sigma^2_{wp} + \sigma^2_{array}$. Having the whole plot and array inseparable means this design is a split-plot design within each of the block types. There are only four treatment combinations per block type, so not all combinations of *i, l,* and *m* occur within each block. The block effect, and combined whole plot - microarray effect have subscripts indicating to which block type and block within a block type each belongs.

Tables 11 and 12 contain the models for the treatment combinations in the *k*th block of Type I and Type II blocks, respectively, for design C. The technique utilized for dissecting the models for design A can be employed with the models of design C. That analysis provides the following results. The Blktype effect corresponds to the Treatment 2 by Color interaction. The comparison of the levels of Treatment 1 is a between whole plot comparison. Comparisons of the levels of Treatment 2, Treatment 1 by Treatment 2 interaction, comparison of the levels of Color, and Treatment 1 by Color interaction are within whole plot comparisons (as well as within array) The Blktype or Treatment 2 by Color comparison is a between block comparison. These comparisons are summarized in the analysis of variance table in Table 9.

Table 11. Models for observations from the *k*th block of a Type I block of design C	
	Treatment 2 Level 1
Treatment 1 Level 1	$y_{1Ik1R1} = \mu_{11R} + b_{Ik} + (w+a)_{Ik1} + \varepsilon_{1Ik1R1}$
Treatment 1 Level 2	$y_{2Ik1R2} = \mu_{21R} + b_{Ik} + (w+a)_{Ik2} + \varepsilon_{2Ik1R2}$
	Treatment 2 Level 2
Treatment 1 Level 1	$y_{1Ik2G1} = \mu_{12G} + b_{Ik} + (w+a)_{Ik1} + \varepsilon_{1Ik2G1}$
Treatment 1 Level 2	$y_{2Ik2G2} = \mu_{22G} + b_{Ik} + (w+a)_{Ik2} + \varepsilon_{2Ik2G2}$

Table 12. Models for observations from the *k*th block of a Type II block of design C	
	Treatment 2 Level 1
Treatment 1 Level 1	$y_{1IIk1G1} = \mu_{11G} + b_{IIk} + (w+a)_{IIk1} + \varepsilon_{1IIk1G1}$
Treatment 1 Level 2	$y_{2IIk1G2} = \mu_{21G} + b_{IIk} + (w+a)_{IIk2} + \varepsilon_{2IIk1G2}$

	Treatment 2 Level 2
Treatment 1 Level 1	$y_{1IIk2R1} = \mu_{12R} + b_{IIk} + (w+a)_{IIk1} + \varepsilon_{1IIk2R1}$
Treatment 1 Level 2	$y_{2IIk2R2} = \mu_{22R} + b_{IIk} + (w+a)_{IIk2} + \varepsilon_{2IIk2R2}$

The SAS Proc Mixed code required to fit this model is as follows where blktype denotes the treatment 2 by color interaction and trt1*blktype denotes the three way interaction.

```
proc mixed;
class blk trt1 trt2 array color blktype;
model y=trt1|trt2 blktype color trt1*color trt1*blktype;
random blk(blktype) trt1*blk(blktype) ;
```

The whole-plot (or whole plot and array combined) error term is computed as trt1*blk(blktype).

Table 13. Analysis of Variance Table for Design C using Type I Sums of Squares to Evaluate Expected Mean Squares

Source	df	EMS
trt1	1	$\sigma^2_\varepsilon + 2\sigma^2_{wp+array} + \phi^2_{trt1}$
trt2	1	$\sigma^2_\varepsilon + \phi^2_{trt2}$
trt1*trt2	1	$\sigma^2_\varepsilon + \phi^2_{trt1*trt2}$
blktype (T2xColor)	1	$\sigma^2_\varepsilon + 2\sigma^2_{wp+array} + 4\sigma^2_{blk} + \phi^2_{trt1*color}$
color	1	$\sigma^2_\varepsilon + \phi^2_{color}$
trt1*color	1	$\sigma^2_\varepsilon + \phi^2_{trt1*color}$
trt1*blktype (t1xt2xcolor)	1	$\sigma^2_\varepsilon + 2\sigma^2_{wp+array} + \phi^2_{trt1*trt2*color}$
blk(blktype)	4	$\sigma^2_\varepsilon + 2\sigma^2_{wp+array} + 4\sigma^2_{blk}$
blk*trt1*(blktype)	4	$\sigma^2_\varepsilon + 2\sigma^2_{wp+array}$
Residual	8	σ^2_ε

Table 14 contains the parameters of interest and the variances of estimates of those parameters. The variances of estimates of the parameters, $\bar{\mu}_{1\bullet\bullet} - \bar{\mu}_{2\bullet\bullet}$, $\bar{\mu}_{\bullet\bullet R} - \bar{\mu}_{\bullet\bullet G}$, $\bar{\mu}_{11\bullet} - \bar{\mu}_{12\bullet}$, and $\bar{\mu}_{21\bullet} - \bar{\mu}_{22\bullet}$ are identical to those in Table 6 for design A. The variance of the estimates of the other parameters are different.

Table 14. Parameters of Interest and Variances of Estimates of Parameters of Interest based on Design C	
Parameter	Variance of Estimate of the Parameter
$\bar{\mu}_{1\bullet\bullet} - \bar{\mu}_{2\bullet\bullet}$	$\left[\sigma_\varepsilon^2 + 2\sigma_{wp+array}^2\right]/6$
$\bar{\mu}_{\bullet 1\bullet} - \bar{\mu}_{\bullet 2\bullet}$	$\left[\sigma_\varepsilon^2\right]/6$
$\bar{\mu}_{\bullet\bullet R} - \bar{\mu}_{\bullet\bullet G}$	$\left[\sigma_\varepsilon^2\right]/6$
$\bar{\mu}_{11\bullet} - \bar{\mu}_{21\bullet}$	$\left[\sigma_\varepsilon^2 + \sigma_{wp+array}^2\right]/3$
$\bar{\mu}_{12\bullet} - \bar{\mu}_{22\bullet}$	$\left[\sigma_\varepsilon^2 + \sigma_{wp+array}^2\right]/3$
$\bar{\mu}_{11\bullet} - \bar{\mu}_{12\bullet}$	$\left[\sigma_\varepsilon^2\right]/3$
$\bar{\mu}_{21\bullet} - \bar{\mu}_{22\bullet}$	$\left[\sigma_\varepsilon^2\right]/3$

3. Concluding Remarks

The variances of estimates of parameters of interest for each of the three designs are presented in Table 15. Designs B and C have the smallest variance for comparing the levels of Treatment 2. Design C has the smallest variances for comparing the levels of Color, and for comparing the levels of Treatment 2 within each level of Treatment 1. Designs A and B have a smaller variance than design C for comparing the levels of Treatment 1. Thus, the choice of which way to superimpose the microarrays with the split-plot design depends on selection of effects for which one wishes to have more powerful comparisons.

Table 15. Parameters of interest and variances of the estimated parameters based on designs A, B, and C			
	Variance of Estimate of the Parameter		
Parameter	Design A	Design B	Design C
$\bar{\mu}_{1\bullet\bullet} - \bar{\mu}_{2\bullet\bullet}$	$\left[\sigma_\varepsilon^2 + 2\sigma_{wp}^2\right]/6$	$\left[\sigma_\varepsilon^2 + 2\sigma_{wp}^2\right]/6$	$\left[\sigma_\varepsilon^2 + 2\sigma_{wp+array}^2\right]/6$
$\bar{\mu}_{\bullet 1\bullet} - \bar{\mu}_{\bullet 2\bullet}$	$\left[\sigma_\varepsilon^2 + 2\sigma_{array}^2\right]/6$	$\left[\sigma_\varepsilon^2\right]/6$	$\left[\sigma_\varepsilon^2\right]/6$
$\bar{\mu}_{\bullet\bullet R} - \bar{\mu}_{\bullet\bullet G}$	$\left[\sigma_\varepsilon^2 + 2\sigma_{wp}^2\right]/6$	$\left[\sigma_\varepsilon^2 + 2\sigma_{wp}^2\right]/6$	$\left[\sigma_\varepsilon^2\right]/6$
$\bar{\mu}_{11\bullet} - \bar{\mu}_{21\bullet}$	$\left[\sigma_\varepsilon^2 + \sigma_{wp}^2\right]/3$	$\left[\sigma_\varepsilon^2 + \sigma_{wp}^2 + \sigma_{array}^2\right]/3$	$\left[\sigma_\varepsilon^2 + \sigma_{wp+array}^2\right]/3$
$\bar{\mu}_{12\bullet} - \bar{\mu}_{22\bullet}$	$\left[\sigma_\varepsilon^2 + \sigma_{wp}^2\right]/3$	$\left[\sigma_\varepsilon^2 + \sigma_{wp}^2 + \sigma_{array}^2\right]/3$	$\left[\sigma_\varepsilon^2 + \sigma_{wp+array}^2\right]/3$

$\bar{\mu}_{11\bullet} - \bar{\mu}_{12\bullet}$	$\left[\sigma_\varepsilon^2 + \sigma_{array}^2\right]/3$	$\left[\sigma_\varepsilon^2 + \sigma_{array}^2\right]/3$	$\left[\sigma_\varepsilon^2\right]/3$
$\bar{\mu}_{21\bullet} - \bar{\mu}_{22\bullet}$	$\left[\sigma_\varepsilon^2 + \sigma_{array}^2\right]/3$	$\left[\sigma_\varepsilon^2 + \sigma_{array}^2\right]/3$	$\left[\sigma_\varepsilon^2\right]/3$

In the designs evaluated here, the assignment of dyes to treatments followed the same pattern within a block in the sense that, for designs A and B, the same dye was assigned to a particular precipitation treatment for both microarrays evaluated within a block and, for design C, the same dye was assigned to a particular temperature treatment within a block. This is illustrated by the directionality of arrows in Figure 1. Other variations on dye assignment within a block would be possible and, in some cases would have the effect of equalizing variances. For example, one could superimpose a combination of all three of the microarray designs, perhaps with the first pair of blocks having pairing as in design A, the second with pairing as in design B, and the third with pairing as in design C. Such a composite design should provide estimates of effects that have similar variances. The three microarray designs considered here enable one to explicitly evaluate the variances of each of the effects. Superimposing other microarray designs on this split-plot design, such as combinations of the above, does not enable the variances of the comparisons to be readily evaluated explicitly and would generally make explicit comparisons impossible. In that case one would have to resort to simulation to get a sense of the magnitudes of the variances of the interesting comparisons.

These analyses are specific to a split-plot design with two treatments, each having two levels, but they illustrate some general processes for developing designs and analyses for application of microarrays in pre-existing experiments. First, when variance among arrays is substantial, comparisons of treatment levels will be more powerful if treatment levels can be included on the same array. Of course this is the basis for loop designs which include pairs of each interesting comparison on arrays. If more than two levels are present, other strategies must be devised to include representative levels on the same arrays, unless more than two colors are used (Woo et al, 2005). But, second, it is also clearly feasible to compare levels of treatments that do not appear on the same microarray when sufficient replication is included. In cases where a large percentage of genes respond to treatment levels, care may need to be taken in the normalization process to avoid obscuring differences between treatment levels when comparisons of levels are made across microarrays rather within microarrays. But now as microarrays become more readily available so that sufficient replication is feasible, designs in which arrays are confounded with an effect can be added to the microarray analyst's repertoire.

References

Blum J. E., P. Casati, V. Walbot, and A. E. Stapleton (2004) Split-plot microarray design allows sensitive detection of expression differences after ultraviolet radiation in the inbred parental lines of a key maize mapping population. Plant Cell and Environment 27:1374-1386.

Bueno Filho, J. S. S., S. G. Gilmour, and G. J. M. Rosa (2006) Design of microarray experiments for genetical genomics studies. Genetics 174: 945-957.

Chu, T. M., B. Weir, and R. Wolfinger (2002) A systematic statistical linear modeling approach to oligonucleotide array experiments. Mathematical Biosciences 176:35-51 Special Issue.

Fay, P. A., J. D. Carlisle, B. T. Danner, M. S. Lett, J. K. McCarron, C. Stewart, A. K. Knapp, J. M. Blair, and S. L. Collins (2002) Altered rainfall patterns, gas exchange and growth in grasses and forbs. International Journal of Plant Sciences 163:549-557.

Feder, M. E. and T. Mitchell-Olds (2003) Evolutionary and ecological functional genomics. Nature Reviews Genetics 4:649-655.

Garrett, K. A., S. H. Hulbert, J. E. Leach, and S. E. Travers (2006) Ecological genomics and epidemiology. European Journal of Plant Pathology 115:35-51.

Gibson G., and R. Wolfinger (2004) Gene expression profiling using mixed models. Ch. 11, pp 251-278, A. M. Saxton, ed. in Genetic analysis of complex traits using SAS, SAS Users Press, Cary NC.

Glonek, G. F. V., and P. J. Solomon (2004) Factorial and time course designs for cDNA microarray experiments. Biostatistics 5:89-111.

Greer, C. W., L. G. Whyte, J. R. Lawrence, L. Masson and R. Brousseau (2001) Genomics technologies for environmental science. Environmental Science and Technology 35:360A-366A.

Jin, W., R. M. Riley, R. D. Wolfinger, K. P. White, G. Passador-Gurgel, and G. Gibson (2001) The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. Nature Genetics 29:389-395.

Kerr, K. F. (2006) 2k factorials in blocks of size 2, with application to two-color microarray experiments. UW Biostatistics Working Paper Series, University of Washington, Paper 227.

Kerr, M. K., and G. A. Churchill (2001) Statistical design and the analysis of gene expression microarray data. Genetical Research 77:123-128.

Littell, R. C., G. A. Milliken, W. W. Stroup, R. D. Wolfinger, and O. Schabenberger (2006) SAS System Mixed Models, 2nd Edition. SAS Institute, Cary, North Carolina.

Purugganan, M., and G. Gibson (2003) Merging ecology, molecular evolution and functional genetics. Molecular Ecology 12:1109-1112.

Rosa, G. J. M., J. P. Steibel, and R. J. Tempelman (2005) Reassessing design and analysis of two-colour microarray experiments using mixed effects models. Comparative and Functional Genomics 6:123-131.

Tempelman, R. J. (2005) Assessing statistical precision, power, and robustness of alternative experimental designs for two color microarray platforms based on mixed effects models. Veterinary Immunology and Immunopathology 105:175-186.

Thomas, M. A., and R. Klaper (2004) Genomics for the ecological toolbox. Trends in Ecology and Evolution 19:439-445.

Travers, S. E., M. D. Smith, J. Bai, S. H. Hulbert, J. E. Leach, P. S. Schnable, A. K. Knapp, G. A. Milliken, P. A. Fay, A. Saleh, and K. A. Garrett (2007) Ecological genomics: making the leap from model systems in the lab to native populations in the field. Frontiers in Ecology and the Environment 5:19-24.

Wit, E., A. Nobile, and R. Khanin (2005) Near-optimal dual channel microarray studies. Applied Statistics 54:817-830.

Wolfinger, R. D., G. Gibson, E. D. Wolfinger, L. Bennett, H. Hamadeh, P. Bushel, C. Afshari, and R. S. Paules (2001) Assessing gene significance from cDNA microarray expression data via mixed models. Journal of Computational Biology 8:625-637.

Woo, W., W. Krueger, A. Kaur, and G. Churchill (2005) Experimental design for three-color and four-color gene expression microarrays. Bioinformatics, 21:i459 - i467.

Yang, Y. H, and T. Speed (2002) Design issues for cDNA microarray experiments. Nature Reviews Genetics 3:579-588.