# DATA ANALYSIS FOR QUANTITATIVE DETERMINATIONS OF POLAR LIPID MOLECULAR SPECIES

by

TINGTING SONG

B.S., SHANDONG NORMAL UNIVERSITY, 1996

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

DEPARTMENT OF STATISTICS
COLLEGE OF ARTS AND SCIENCES

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2010

Approved by:

Major Professor
GARY GADBURY

# Copyright

# Abstract

This report presents an analysis of data resulting from a lipidomics experiment. The experiment sought to determine the changes in the lipidome of big bluestem prairie grass when exposed to stressors. The two stressors were drought (versus a watered condition) and a rust infection (versus no infection), and were whole plot treatments arranged in a 2 by 2 factorial. A split plot treatment factor was the position on a sampled leaf (top half versus bottom half). In addition, samples were analyzed at different times, representing a blocking factor. A total of 110 samples were used and, for each sample, concentrations of 137 lipids were obtained. Many lipids were not detected for certain samples and, in some cases, a lipid was not detected in most samples. Thus, each lipid was analyzed separately using a modeling strategy that involved a combination of mixed effect linear models and a categorical analysis technique, with the latter used for certain lipids to determine if a pattern of observed zeros was associated with the treatment condition(s). In addition, p-values from tests of fixed effects in a mixed effects model were computed three different ways and compared. Results in general show that the drought condition has the greatest effect on the concentrations of certain lipids, followed by the effect of position on the leaf. Of least effect on lipid concentrations was the rust condition.

# Table of Contents

# List of Figures

# List of Tables

# Acknowledgements

Thank for the help and support from Dr. Gary Gadbury who is my major advisor, Dr. Ruth Welti who is my supervisor at the Kansas Lipidomics Research Center and my committee member, Dr. Leigh Murray and Dr. Karen Garrett whose student, Erin Frank, generated the data.

I am very grateful to work with Dr. Welti's research group in Kansas Lipidomics Research Center in the Division of Biology, especially thanks for Dr. Richard Jeannotte, Mary Roth and Pam Tamura.

I appreciate the department of Statistics and the Division of Biology.

My parents and my husband are always giving me support and encouragement. And I thank all my teachers and friends.

# Chapter 1 - Experiment design

In the field or in a natural environment, plants are exposed to multiple stresses. Drought stress is a common environmental factor limiting plant growth and yield. Biotic stresses also produce effects to plants in the field, such as the presence of a pathogen resulting in disease. Research on the effect on plants of the interactions between drought and rust infection is limited. The experiment that produced the data analyzed herein was designed to test drought and pathogen effects on the lipidome of the dominant tall grass prairie plant species. The experiment is briefly described here, with further details available in Frank (2007). The chosen plant is big bluestem (*Andropogon gerardii*). Big bluestem seed was supplied by the USDA Plant Materials Center. Seeds were sown in flats of vermiculite and transferred to yellow cones (12 cm) with soil from Konza Praire Nature Research Area. Plants in cones were tested for rust susceptibility when approximately 10 cm tall. Planting and sampling was done within blocks, representing time periods staggered by approximately a week. Efforts were made to sample plants at the same time of day for all blocks. Pots were whole plot experimental units within six blocks and the position of pots on a bench was re-randomized within a block each week. Drought and rust treatments were whole plot treatments and were applied in a factorial arrangement, with presence or absence of two intervals of drought stress as one factor and presence or absence of two rust inoculations as the other. The leaf below the youngest fully expanded leaf was treated for lipidomic or phytohormone analysis. To account for heavier infection closer to the tip of the leaf, each sampled leaf was gently folded in half and cut about one third of the way from the ends, alternating which part was used for polar lipid analysis and phytohormone analysis. The middle section was then cut in half, tip half and base half and each representing 1/6 of the leaf from the center one-third. The split plot treatment was the position on a leaf. The experiment design is thus a split-plot in a randomized complete block design with six blocks. Approximately 2-3 samples for each treatment combination within block were obtained. More details on the treatment and blocking variables are included in Chapter 3 that describes the data analyses that were conducted. First, Chapter 2 describes the measurements for lipid concentrations and preprocessing steps performed on the data.

# Chapter 2 - Clean the dataset

The dataset generated contains concentrations of lipids that are zero or near zero. The values near zero could be actually zero or less than or equal to the detection limit of the instrument for a corresponding lipid. Each lipid measured by the mass spectrometer will have a different detection limit. However, since it was not possible to have a standard for each lipid, the limits of detection were estimated from the analysis of samples that do not contain the lipid (blank samples). Each lipid targeted in the analysis will have a background value in analysis of the blank samples. The detection limit for each lipid was computed as follow: Limit of detection (LOD) = mean of the signals of a given lipid in blank runs + 3 x standard deviation for the values. A limit of detection was calculated for each lipid using their corresponding background signals in the 13 blank runs that were used in the analysis of the samples. The mean of all the limits of detection was computed and the resulting value was 0.002 nmol. Then, in the dataset, each value that was below 0.002 was replaced by a 0. This represents a choice; the alternative would have been to label as "non-detectable" the lipids with signals that were below the detection limit.

In order to deal with the numerous zero values present in the dataset, a series of criteria were set to clean the dataset. These criteria were established arbitrarily in order to minimize the impact of the zeros on the subsequent statistical analysis:

-A treatment with only two samples in which one or both samples were zeros, nothing was changed to the situation;

-A treatment with three samples in which two of the samples were zeros, the third nonzero value was replaced with a zero. In this situation, it could imply that the given lipid level was below the detection limit;

-A treatment with three samples and only one is a zero, and then it was replaced by the average value of the two nonzero values;

-If the data for a given lipid contain only zeros in one treatment combination, then the data for this given lipid stayed in the data set. If more than one treatment combination contained all zeros, the data for the given lipid were deleted from the dataset;

Finally, there still may be lipids with a concentration of zero in many samples. Different methods of analysis for lipids were conducted based on a proportion of zeros across samples. This is described in Chapter 3.

# Chapter 3 - Data analysis

## §3.1 overview

A total of 110 samples was obtained and analyzed according to a Randomized Complete Block with split-plot with six blocks. A whole plot experimental unit (WPEU) was a pot containing approximately 5 big bluestem plans. Samples were obtained from leaves selected from the plants. Blocks represented sampling times that were staggered by at least one week. The leaf below the youngest fully expanded leaf was prepared for lipidomic or phytohormone analysis. To evaluate effects of stressors closer to the tip of the leaf, each sampled leaf was cut in half, alternating which half was used for polar lipid analysis and phytohormone analysis. The leaf was gently folded in half and cut about one third of the way from each end. The middle section was then cut in half giving a tip half and a base half. The split plot treatment was the position of the sample on a leaf where 1 = top of leaf and 0 = base of leaf. The top/base samples were not paired on a leaf. That is, a top of a leaf was randomly sampled and a base of a different leaf randomly sampled because the remainder of each leaf was used for a different analysis. The two whole plot treatment factors were first drought/water where a drought condition = 0 and watered condition = 1, and second rust where rust = 1 indicates a rust infection applied to the unit and rust = 0 indicates no infection applied. The two whole plot treatment factors were arranged in a 2 way factorial design. Approximately 2-3 samples were sampled at the split plot level, but the design is not balanced as the number of samples varied due to technical difficulties with sample integrity. The response variable for each sample was the concentration of a lipid. The concentrations of 137 lipids were obtained. Each lipid was analyzed separately to assess the effect of treatments on the concentration of that particular lipid.

A complicating feature of the data set was the number of zero levels remaining for the lipid concentration in several samples for particular lipids, even after the preprocessing described in Chapter 2. These zeros could be present for several reasons: 1. The particular lipid was simply not present in the samples; 2. The concentration of a lipid was below the detection limit for the instrument; or 3. The treatment(s) blocked the formation of the lipid or caused its degradation. The planned primary analysis method was to use a mixed factor effects modeling approach with the blocking factor (sampling time) as a random effect and the treatment factors as fixed. There was concern about the convergence of likelihood estimates and/or the validity of inferences

based on such estimates when many response values were tied at zero. Thus, some initial screening was done to group lipids into categories based on the presence of zero levels across the 110 samples. Lipids with 100 or more zeros were not considered further. There were 44 of these (out of 137). Lipids with 11 to 99 zero concentrations across the 110 samples were analyzed using a categorical method to be discussed later. There were 42 of these. Finally, the remaining lipids were analyzed using a mixed effects modeling approach. However, these lipids were separated into two groups: lipids with 1 to 10 zeros and lipids with no zeroes. Though the same modeling strategy was used for lipids in both of these groups, it does allow for any follow-up study if results suggest that the presence of any zeros (versus no zeros) warrants more detailed investigation. There were 24 lipids containing 1 to 10 zeros among the 110 samples and 27 containing no zeros (i.e., the lipid was detected in all 110 samples). Figure 1 summarizes the categorization of the lipids into the four groups, and Table 1 shows the specific lipids appearing in categories for further analysis.

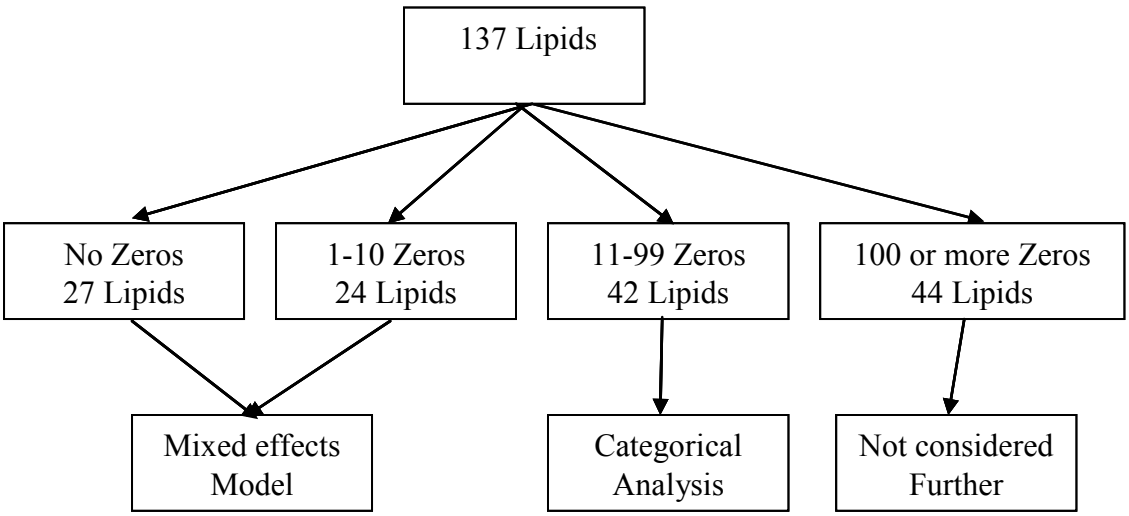


**Figure 1. Partitioning of lipids into modeling strategies based on the number of zero concentrations out of 110 samples.**

**Table 1. List of the lipids in the no zeros group, the 1-10 zeros group, and the 11-99 zeros group.**

| No Zeros (27 Lipids) | 1-10 Zeros (24 Lipids) | 11-99 Zeros (42 Lipids) |
|---|---|---|
| DGDG 34_3 | DGDG 34_4 | DGDG 36_1 |
| DGDG 34_2 | DGDG 36_2 | DGDG 38_4 |
| DGDG 34_1 | DGDG 38_6 | MGDG 38_4 |
| DGDG 36_6 | DGDG 38_5 | PG 34_1 |
| DGDG 36_5 | DGDG 38_3 | PG 34_0 |
| DGDG 36_4 | MGDG 34_1 | LysoPG 16_0 |
| DGDG 36_3 | MGDG 36_5 | LysoPG 18_2 |
| MGDG 34_4 | MGDG 38_6 | LysoPC 16_1 |
| MGDG 34_3 | MGDG 38_5 | LysoPE 16_0 |
| MGDG 34_2 | PG 32_1 | PC 36_1 |
| MGDG 36_6 | PG 32_0 | PC 38_6 |
| MGDG 36_4 | PG 34_3 | PC 38_5 |
| MGDG 36_3 | PG 34_2 | PE 34_4 |
| PG 34_4 | PC 34_4 | PE 38_6 |
| PC 34_3 | PC 34_1 | PE 40_3 |
| PC 34_2 | PC 36_2 | PE 40_2 |
| PC 36_6 | PC 38_4 | PE 42_3 |
| PC 36_5 | PC 38_2 | PI 34_1 |
| PC 36_4 | PE 36_3 | PI 36_6 |
| PC 36_3 | PE 36_2 | PI 36_5 |
| PC 38_3 | PE 38_3 | PI 36_4 |
| PE 34_3 | PE 38_2 | PI 36_3 |
| PE 34_2 | PE 42_2 | PI 36_2 |
| PE 36_6 | PI 34_2 | PI 36_1 |
| PE 36_5 | | PS 34_3 |
| PE 36_4 | | PS 34_2 |
| PI 34_3 | | PS 38_2 |

| | |
|---|---|
| Table 1 continued | PS 40_3 |
| | PS 42_3 |
| | PS 42_2 |
| | PA 34_3 |
| | PA 34_2 |
| | PA 36_5 |
| | LysoPG 16_1 |
| | LysoPC 16_0 |
| | LysoPC 18_3 |
| | LysoPC 18_2 |
| | PC 32_0 |
| | PC 40_2 |
| | PE 38_5 |
| | PS 40_2 |
| | PA 36_6 |

## §3.2 The mixed effects model analysis

This section describes the mixed effects model that was used for lipids in the 1-10 zeros group and in the no zeros group. The goal was to analyze the effect of treatments on the concentration of each of the 51 lipids analyzed by this technique (24 lipids in the 1-10 zeros group and 27 in the no-zeros group). A sequence of models was used to first investigate significant three-way interactions between the three fixed effect treatment factors. Next, two-way interactions were investigated followed by main effects. The first model considered is given by (1) below.

$$Y_{ijklm} = \mu + B_l + \alpha_i + \beta_j + (\alpha\beta)_{ij} + (B\alpha\beta)_{lij} + \gamma_k + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (B\alpha\gamma)_{lik} + (B\beta\gamma)_{ljk} + (\alpha\beta\gamma)_{ijk} + (B\alpha\beta\gamma)_{lijk} + \varepsilon_{ijklm}$$

……..(1)

A model is being fit to each lipid and there are 51 lipids resulting in 51 separate models. Blocks are random effects while drought, rust and position are fixed effects.

$\alpha_i$ , i=0, 1 denotes the drought main effect, with i=1 denoting the water condition and i=0 the drought condition,

$\beta_j$ , j=0, 1 denotes the rust main effect, with j=0 denoting the no rust condition and j=1 the rust condition,

$\gamma_k$ , k=0, 1 denotes the position main effect, with k=0 denoting the bottom position and k=1 the top condition.

$B_l$ , $(B\alpha\beta)_{lij}$ , $(B\alpha\gamma)_{lik}$ , $(B\beta\gamma)_{ljk}$ and $(B\alpha\beta\gamma)_{lijk}$ l=1, 2, 3, 4, 5, 6 denote random effects associated with the blocking where $B_l \sim N(0, \sigma_B{}^2)$ and other random effects are assumed to be independent normal random variables with mean zero and their own variance component. The terms $(\alpha\beta)_{ij}$ , $(\alpha\gamma)_{ik}$ and $(\beta\gamma)_{jk}$ denote the fixed two-way interaction

effects and $(\alpha\beta\gamma)_{ijk}$ denotes the fixed three way interaction effect. Finally, $\varepsilon_{ijklm}$ represents the residual error and is assumed to be normally distributed with mean zero and the variance $\sigma_{\varepsilon}^{2}$, and to be independent of other random effect terms.

The model in (1) was first tested against the model below in (2) that does not contain the three-way interaction term. Note that the variance component structure is the same in both models (1) and (2). Throughout all of the following analyses, p-values were computed three different ways: 1. within SAS proc mixed using the Kenward and Roger (1997) method for degrees of freedom; 2. using the asymptotic distribution for the likelihood ratio test statistic which, for this test, is chi-square with one degree of freedom; and 3. using a parametric bootstrap approach (500 bootstrap samples) for the likelihood ratio test statistic that has been reported in Faraway (2006). Some implications of also dropping the variance component associated with the three-way interaction term in model (2) are discussed in a later section. Any lipid with a statistically significant three-way interaction was set aside for further analyses involving determining which contrasts among treatments are contributing to the significance. A liberal significance level (i.e., unadjusted for multiple testing) of 0.05 was used for all tests to favor detection of any interactions requiring further investigation. Later in the summary section, p-values were adjusted using false discovery rate control reported in Storey (2003).

$$
\begin{aligned}
Y_{ijklm} &= \mu + B_l + \alpha_i + \beta_j + (\alpha\beta)_{ij} + (B\alpha\beta)_{lij} \\
&+\gamma_k + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (B\alpha\gamma)_{lik} + (B\beta\gamma)_{ljk} \\
&+(B\alpha\beta\gamma)_{lijk} + \varepsilon_{ijklm}
\end{aligned}
$$

…… …… (2)

All lipids with no significant three-way interaction were then analyzed using model (2) against a null model that drops one of the two-way interactions. Thus, for each lipid analyzed in this way, three hypotheses were tested, one for each interaction, comparing model (2) to a model without that fixed effect interaction. Again, the variance components were kept the same in both models and p-values are computed as discussed above for the tests of three-way interactions. The same significance level of 0.05 was used for all tests.

Any lipid having a statistically significant two-way interaction was set aside for further analyses of contrasts that contribute to the interaction. Standard errors were computed for each contrast using the estimated variance-covariance matrix for estimated fixed effects. Approximate 95% normal distribution based confidence intervals were computed. Any lipids that did not have any significant two-way interactions were then analyzed using the main-effects model (3) below.

$$Y_{ijklm} = \mu + B_l + \alpha_i + \beta_j + (B\alpha\beta)_{lij} + \gamma_k + (B\alpha\gamma)_{lik} + (B\beta\gamma)_{ljk} + (B\alpha\beta\gamma)_{lijk} + \varepsilon_{ijklm} \quad \text{.....(3)}$$

Once again, for each lipid analyzed with this model, the main effect terms were tested against a null model that does not contain a main effect. Thus, three hypotheses were tested for each lipid, one for each treatment factor (similar to the testing for two-way interactions). If a main effect was significant for a lipid, an approximate 95% normal distribution based confidence interval for the mean effect was constructed. Calculations were done using R (www.r-project.org) and the SAS MIXED procedure. Code is given in Appendix A and B.

## §3.3 Categorical analyses for lipids with more zero concentrations

There were 42 lipids with 11 to 99 zero concentrations across the 110 samples. Too many zeros could affect the convergence of algorithms for fitting a mixed effects model and make the assumptions for the model questionable. So instead a simple categorical analysis was used for these lipids to determine if a pattern of zeros depended on the fixed effects treatment structure. For each lipid with 11 to 99 zeros, the numbers of zeros were tallied and the numbers of positive concentrations were tallied (note the two tallies sum to 110). Contingency tables were then set up to determine if the zero (versus nonzero) lipid concentrations depended on each of the three treatment factors. So there were 126 contingency tables to test for a drought effect, a rust effect, and a position effect for the 42 lipids analyzed in this way. Table 2 is the contingency table for lipid DGDG 36_1 to test for a drought effect. Table 3 is the contingency table for lipid PG 34_1 to test for a drought effect. If the occurrence of a zero concentration does not depend on the treatment (for example the Table 2 for DGDG 36_1), there is no significant drought effect for DGDG 36_1. If the occurrence of a zero value does depend on a treatment (for example the Table 3 for PG 34_1), there may be a drought effect. P-values were computed two ways using R software (www.r-project.org) and the functions, *prop.test* (a chi-square test for independence) and *fisher.test* (Fisher's exact test, cf. Conover, 1999). A liberal significance level of 0.05 is used for all tests to favor detection of any significant effects.

**Table 2. Example of the distributions of zeros for a drought effect for lipid DGDG 36_1. The two columns are yes for a lipid being present and no for a zero concentration.**

| Drought | Lipid Present | | Total |
|---|---|---|---|
| | Yes | No | |
| Drought | $n_{00}$=44 | $n_{01}$=12 | $n_{0+}$=56 |
| Water | $n_{10}$=36 | $n_{11}$=18 | $n_{1+}$=54 |
| Total | $n_{+0}$=80 | $n_{+1}$=30 | n=110 |

**Table 3. Example of the distributions of zeros for a drought effect for lipid PG 34_1. The two columns are yes for a lipid being present and no for a zero concentration.**

| Drought | Lipid Present | | Total |
|---|---|---|---|
| | Yes | No | |
| Drought | $n_{00}$=56 | $n_{01}$=0 | $n_{0+}$=56 |
| Water | $n_{10}$=40 | $n_{11}$=14 | $n_{1+}$=54 |
| Total | $n_{+0}$=96 | $n_{+1}$=14 | n=110 |

# Chapter 4 - Results

## §4.1 Mixed effects result

Figure 2 shows a diagram summarizing the results of the mixed-effects model analyses.

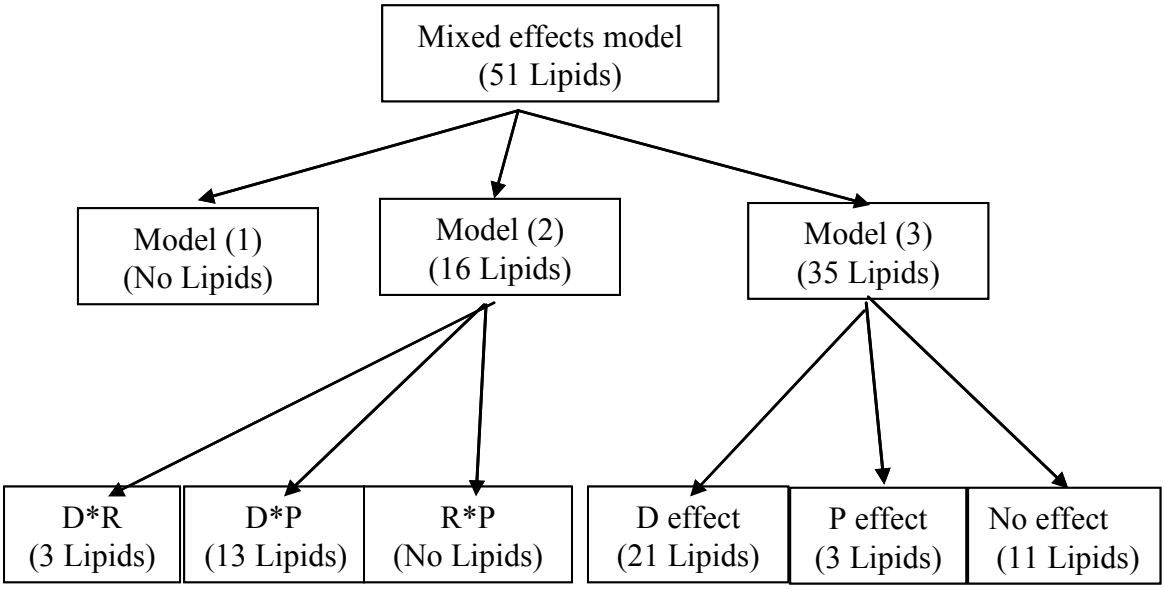


**Figure 2. Summary of results of the mixed effects modeling procedure, where D=Drought, R=Rust, and P=Position. Model 1 was a test for three way interactions, model 2, two way interactions, and model 3 tested for main effects. The diagram shows effects that were significant at a 0.05 significance level. For example, 3 lipids had a significant drought-rust interaction (D*R) in the mixed effects model.**

There are 51 lipids that were analyzed in 51 separate mixed effects models. First, three-way interaction of fixed effects, $(\alpha\beta\gamma)_{ijk}$, was tested using model (1) against the model (2). If the three-way interaction was significant for a lipid, that lipid was put into a category for model (1), meaning that it was set aside for follow-up analyses using estimates of contrasts. Otherwise, the lipid was tested for two-way interactions using model (2). As mentioned earlier, three different techniques were used to compute p-values: (i) within SAS proc mixed using the Kenward and

Roger (1997) method for degrees of freedom, (ii) using the asymptotic distribution for the likelihood ratio test statistic which, for this test, is chi-square with one degree of freedom, and (iii) using a parametric bootstrap approach with 500 bootstrap samples for the likelihood ratio test statistic that has been reported in Faraway (2006). Some comparison of these three p-values is done in results that follow and discussed in more detail in a later section.

### 4.1.1Test of three-way interactions

Of the 51 lipids tested for a three-way interaction, none were significant at a level 0.05. Figure 3 shows p-values for the 10 lipids with the smallest p-values having no zeroes and Figure 4 shows the 10 smallest p-values for lipids with 1-10 zeros in their concentration levels. It can be seen that in the tests shown below, the bootstrap technique is more conservative than the other two for computing p-values, and the use of asymptotic theory for the likelihood ratio test statistic is more liberal for testing fixed effects. This result has been seen by others as well (cf. Faraway, 2006).

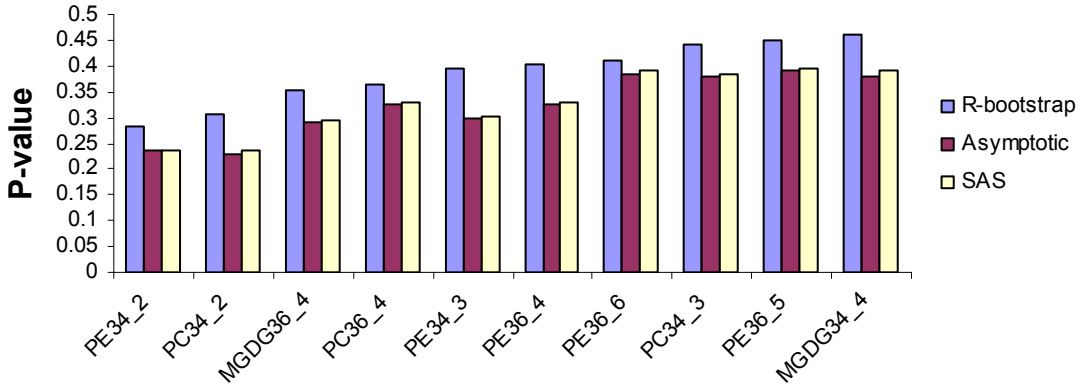


**Figure 3. Comparison of p-values computed three ways for the 10 lipids with smallest p-values in the no zeros group, testing for a significant three-way interaction in model 1.**

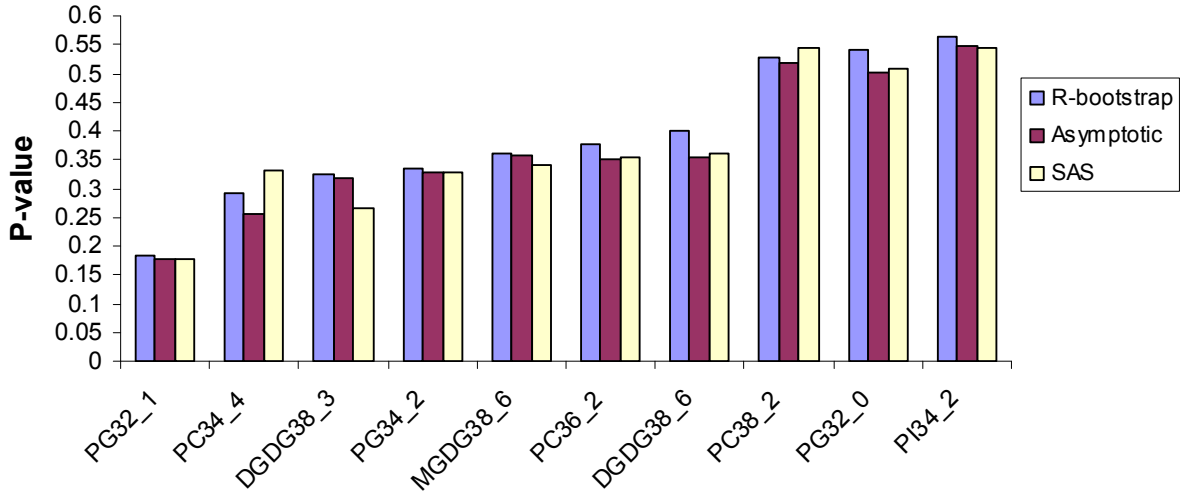


**Figure 4. Comparison of p-values computed three ways for the 10 smallest p-values in 1-10 zeros group testing for a significant three-way interaction in model 1.**

Since no lipids were significant for three-way interactions, all 51 were then tested for two-way interactions. Thus, for each lipid analyzed in this way, three hypotheses are tested, one for each interaction comparing model (2) to a model without that fixed effect interaction. Again, the variance components are kept the same in both models and p-values are computed as discussed above for the tests of three-way interactions. The same liberal significance level of 0.05 was used for all tests so as not to miss any interesting interactions.

### 4.1.2 Tests of two-way interactions

We did not observe anything unusual or different regarding model fits between the 1-10 zeros group and the no-zeros group. For these two groups there were only 3 lipids having a significant drought-rust interaction and 13 lipids having a drought-position interaction in both groups. So the results are reported together from the analyses of these two groups. The 3 lipids having a significant drought-rust interaction, i.e., an $(\alpha\beta)_{ij}$ effect are DGDG 34_1, DGDG 36_2 and MGDG 36_5. The 13 lipids having a significant drought-position interaction, i.e., an $(\alpha\gamma)_{ik}$ effect are PG 34_4, PG 32_1, PC 36_2, PC 36_4, PC 34_2, PC 36_3, MGDG 36_6, PI 34_2, PG 34_3, PE 42_2, PC 34_4, PE 36_3 and PI 34_3. So these 16 lipids are categorized as

model (2) lipids to be evaluated further by contrasts to determine why those particular interactions were significant. There were no lipids having a significant rust-position interaction. As a side note, there were 153 hypotheses tested for a two-way interaction. If a Bonferroni correction was used to adjust p-values, then none of the sixteen significant results mentioned above would remain significant. A less conservative adjustment is discussed in Chapter 6.

#### 4.1.2.1 Analysis of the drought-rust interaction

Figure 5 shows p-values for the 3 lipids having a significant drought-rust interaction in their concentration level. If a p-value was less than 0.05 by two of the methods, then it was declared a significant result.

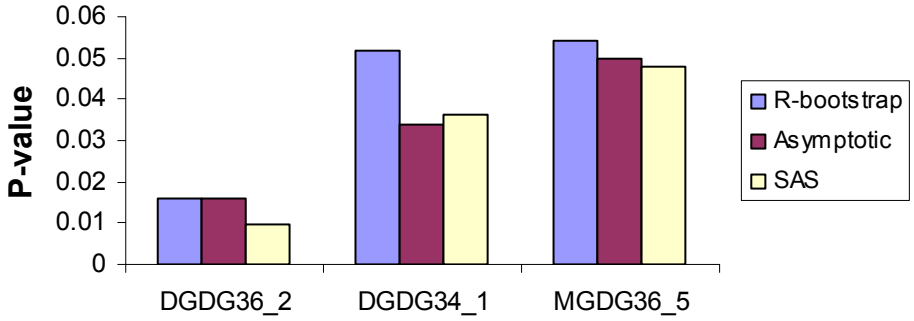


**Figure 5. P-values computed for a drought-rust interaction for lipids DGDG 34_1, DGDG 36_2 and MGDG 36_5.**

Contrasts of interest for the lipids shown in Figure 5 were a test for mean drought effect within rust = 0 (no rust), drought effect within rust=1 (rust present), rust effect within drought=0 (drought) and rust effect within drought=1 (water). Table 4 gives the confidence intervals for these three lipids. Lipids with intervals not covering zero are highlighted in bold. $\mu_{DRP}$ is the notation for the mean lipid concentration for drought condition D, rust condition R and position P. The upper left box is an approximate 95% confidence interval for the contrast, $\mu_{10.}-\mu_{00.}$, which is an interval estimate for a mean drought effect within rust condition equal to zero. Recall that D=1 indicates the watered condition, so the contrast is the change in concentration for the watered condition minus the drought condition. Results are similarly shown for the upper right and lower contrasts shown in Table 4.

**Table 4. 95% confidence intervals for contrasts for which the drought-rust interaction term was significant at the 0.05 level. The contrasts for the drought effect are for the watered condition minus the drought condition. For the rust effect, it is the rust present condition minus the absent condition.**

| $\mu_{10.}-\mu_{00.}$ | Drought effect within rust=0 | | $\mu_{11.}-\mu_{01.}$ | Drought effect within rust=1 | |
|---|---|---|---|---|---|
| | Lower CL | Upper CL | | Lower CL | Upper CL |
| **DGDG 34_1** | **-0.038** | **-0.006** | **DGDG 34_1** | **-0.061** | **-0.029** |
| DGDG 36_2 | -0.006 | 0.000 | **DGDG 36_2** | **-0.011** | **-0.005** |
| MGDG 36_5 | -0.221 | 0.343 | **MGDG 36_5** | **-0.602** | **-0.038** |
| | | | | | |
| $\mu_{01.}-\mu_{00.}$ | Rust effect within drought=0 | | $\mu_{11.}-\mu_{10.}$ | Rust effect within drought=1. | |
| | Lower CL | Upper CL | | Lower CL | Upper CL |
| **DGDG 34_1** | **0.002** | **0.033** | DGDG 34_1 | -0.021 | 0.011 |
| **DGDG 36_2** | **0.001** | **0.007** | DGDG 36_2 | -0.005 | 0.001 |
| MGDG 36_5 | -0.127 | 0.432 | MGDG 36_5 | -0.512 | 0.055 |

From the confidence intervals, we can see there are no drought effects within rust=0 for DGDG 36_2 and MGDG 36_5, but the drought factor shows a negative effect (watered condition minus drought condition) for DGDG 34_1 when rust=0. There are similar negative drought effects within rust=1 for DGDG34_1, DGDG 36_2 and MGDG 36_5. There is a positive rust effect (rust condition minus no rust condition) within drought=0 for DGDG 34_1 and DGDG 36_2, and there are no rust effects within drought=1 for DGDG 34_1, DGDG 36_2 and MGDG 36_5.

For the lipids having a positive or negative mean treatment effect as shown by above interval estimates that do not cover zero, point estimates of a mean ratio were also computed to show the magnitude of change in the mean concentration levels due to the stress condition. This estimate may be more interpretable because many lipids are present in very small concentrations and a mean difference may be less interpretable than a mean ratio.

The estimated mean ratio for the negative drought effect when rust=0 is 0.67 (drought=1/drought=0). So under the no-rust condition, lipid DGDG 34_1 is estimated to be present in the watered condition at a proportion of 0.67 with respect to the drought condition.

The estimated mean ratio for the negative drought effect within rust=1 for DGDG 34_1, DGDG 36_2 and MGDG 36_5 is 0.47, 0.46 and 0.72 (drought=1/drought=0), respectively. So under the rust condition being present (rust =1), lipid DGDG 34_1 has estimated mean concentration in the watered condition at 0.47 with respect to the drought condition. Lipid DGDG 36_2 under the watered condition has estimated mean concentration at 0.46 the estimated level of the drought condition. Lipid MGDG 36_5 under water present has an estimated mean concentration at 0.72 with respect to the drought condition. For a quick visual comparison, Figure 6 summarizes these results with a bar chart.

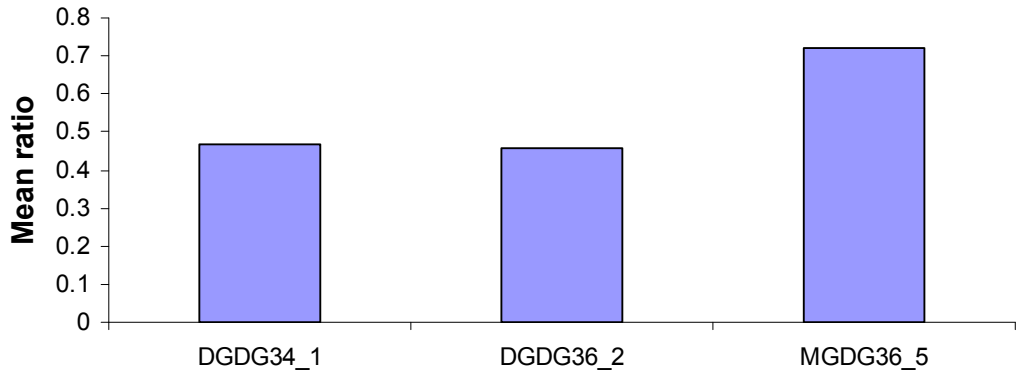


**Figure 6. Estimated mean ratio test computed for drought=1/drought=0 within rust=1 condition for lipids DGDG 34_1, DGDG 36_2 and MGDG 36_5.**

The mean ratio for the positive rust effect within drought=0 for DGDG 34_1 and DGDG 36_2 is 1.266 and 1.331 (rust=1/rust=0). So under the drought condition (drought=0),

lipid DGDG 34_1 will have estimated mean concentration with rust at 1.266 with respect to the no rust condition, lipid DGDG 36_2 under the rust condition has as estimated mean concentration at l.331 with respect to the level without rust. Figure 7 is the bar chart comparing these two results.

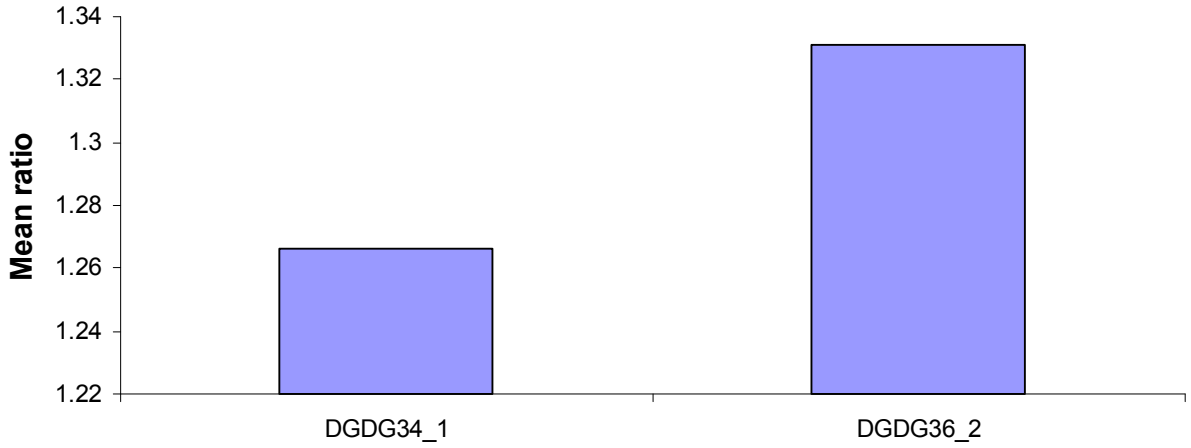


**Figure 7. Mean ratio test computed rust1/rust0 under drought condition for lipid DGDG 34_1 and DGDG 36_2.**

#### 4.1.2.2 Analysis of the drought-position interaction

There are 13 lipids having a significant drought-position interaction, i.e., the $(\alpha\gamma)_{ik}$ effect. These lipids are PG 34_4, PG 32_1, PC 36_2, PC 36_4, PC 34_2, PC 36_3, MGDG 36_6, PI 34_2, PG 34_3, PE 42_2, PC 34_4, PE 36_3 and PI 34_3. Figure 8 shows p-values for the 13 lipids having a significant drought-position interaction in their concentration level.

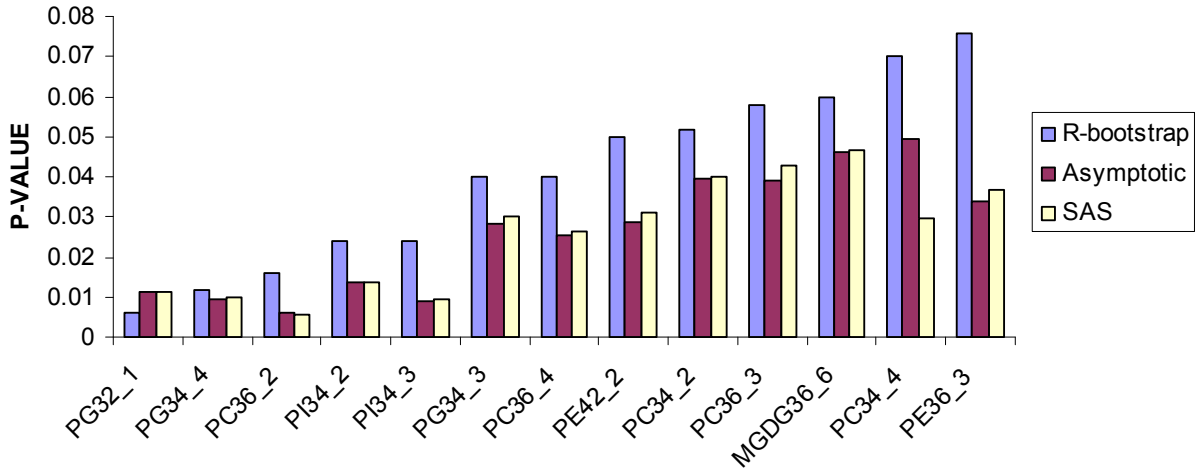


**Figure 8. P-values computed for a drought-position interaction for lipids PG 34_4, PG 32_1, PC 36_2, PC 36_4, PC 34_2, PC 36_3, MGDG 36_6, PI 34_2, PG 34_3, PE 42_2, PC 34_4, PE 36_3 and PI 34_3.**

Of interest for lipids PG 34_4, PG 32_1, PC 36_2, PC 36_4, PC 34_2, PC 36_3, MGDG 36_6, PI 34_2, PG 34_3, PE 42_2, PC 34_4, PE 36_3 and PI 34_3 is the drought effect within position =1 (top), drought effect within position=0 (bottom), position effect within drought=0 (drought) and position effect within drought=1 (water). Table 5 shows the confidence intervals for contrasts for these thirteen lipids. Lipids with intervals that do not cover zero are highlighted in bold. Note that, again, $\mu_{DRP}$ denotes the mean lipid concentration for drought condition D, rust condition R and position P. The upper left box is an approximate 95% confidence interval for the contrast, $\mu_{1.1}-\mu_{0.1}$ which means testing for a mean drought effect within position condition equal to 1, that is, the watered condition minus the drought condition at the top of the leaf. Contrasts are similarly denoted for the other three boxes.

**Table 5. 95% confidence intervals for contrasts for which the drought-position interaction term was significant at the 0.05 level. Again, drought effects are the watered condition minus the drought condition. A position effect is the top of the leaf minus the bottom.**

| $\mu_{1.1}-\mu_{0.1}$ | Drought effect within position=1 | | $\mu_{1.0}-\mu_{0.0}$ | Drought effect within position=0 | |
|---|---|---|---|---|---|
| | Lower CL | Upper CL | | Lower CL | Upper CL |
| PG 34_4 | -0.089 | 0.131 | **PG 34_4** | **-0.268** | **-0.050** |
| PG 32_1 | -0.051 | 0.061 | **PG 32_1** | **-0.123** | **-0.013** |
| PC 36_2 | -0.006 | 0.004 | **PC 36_2** | **-0.016** | **-0.006** |
| PC 36_4 | -0.029 | 0.015 | **PC 36_4** | **-0.064** | **-0.020** |
| PI 34_3 | -0.010 | 0.090 | PI 34_3 | -0.097 | 0.001 |
| PC 34_2 | -0.024 | 0.090 | PC 34_2 | -0.106 | 0.007 |
| PC 36_3 | -0.018 | 0.004 | **PC 36_3** | **-0.033** | **-0.012** |
| MGDG 36_6 | -6.543 | 0.554 | **MGDG 36_6** | **-10.538** | **-3.490** |
| PI 34_2 | -0.011 | 0.017 | **PI 34_2** | **-0.035** | **-0.007** |
| PG 34_3 | -0.082 | 0.063 | **PG 34_3** | **-0.187** | **-0.044** |
| PE 42_2 | -0.003 | 0.001 | **PE 42_2** | **-0.005** | **-0.002** |
| **PC 34_4** | **0.000** | **0.003** | PC 34_4 | -0.002 | 0.001 |
| PE 36_3 | -0.004 | 0.002 | **PE 36_3** | **-0.008** | **-0.002** |

Table 5 continued.

| | | | | | |
|---|---|---|---|---|---|
| $\mu_{1,1}-\mu_{1,0}$ | Position effect within drought=0 | | $\mu_{0,1}-\mu_{0,0}$ | Position effect within drought=1 | |
| | Lower CL | Upper CL | | Lower CL | Upper CL |
| PG 34_4 | -0.119 | 0.068 | **PG 34_4** | **0.060** | **0.249** |
| PG 32_1 | -0.072 | 0.010 | **PG 32_1** | **0.000** | **0.084** |
| PC 36_2 | -0.006 | 0.004 | **PC 36_2** | **0.004** | **0.014** |
| PC 36_4 | -0.026 | 0.018 | **PC 36_4** | **0.010** | **0.054** |
| PI 34_3 | -0.085 | 0.005 | **PI 34_3** | **0.002** | **0.094** |
| PC 34_2 | -0.055 | 0.057 | **PC 34_2** | **0.026** | **0.140** |
| PC 36_3 | -0.013 | 0.008 | **PC 36_3** | **0.002** | **0.024** |
| MGDG 36_6 | -4.155 | 1.439 | MGDG 36_6 | -0.163 | 5.487 |
| PI 34_2 | -0.017 | 0.010 | **PI 34_2** | **0.007** | **0.034** |
| PG 34_3 | -0.056 | 0.077 | **PG 34_3** | **0.049** | **0.184** |
| PE 42_2 | -0.002 | 0.002 | **PE 42_2** | **0.001** | **0.004** |
| PC 34_4 | -0.002 | 0.000 | PC 34_4 | -0.000 | 0.002 |
| PE 36_3 | -0.004 | 0.002 | **PE 36_3** | **0.001** | **0.007** |

From the confidence intervals, there is a positive drought effect (watered condition minus drought condition) for PC 34_4 when position=1. There are negative drought effects within position=0 (bottom) for PG 34_3, PG 32_1, PC 36_2, PC 36_4, PE 36_3, PE 42_2, MGDG 36_6, PG 34_4, PC 36_3 and PI 34_2. There are positive position effects within drought=1 (watered condition) for PC 36_2, PG 34_4, PG 32_1, PI 34_3, PC 34_2, PC 36_3, PI 34_2, PG 34_3, PE 42_2, PE 36_3 and PC 36_4 and there is no position effect within drought=0 (drought

condition) for PG 34_4, PG 32_1, PC 36_2, PC 36_4, PC 34_2, PC 36_3, MGDG 36_6, PI 34_2, PG 34_3, PE 42_2, PC 34_4, PE 36_3 and PI 34_3.

As before, for the lipids having a positive or negative effect as shown by interval estimates that do not cover zero, point estimates of a mean ratio were also computed to show the magnitude of change in the mean concentration levels due to the stress condition. This estimate may be more interpretable because many lipids are present in very small concentrations and a mean difference may be less interpretable than a mean ratio. The estimated mean ratio for the positive drought effect when position=1 for PC 34_4 is 1.374 (drought=1/drought=0). So on the leaf top, lipid PC 34_4 is estimated to be present in the watered condition at a proportion of 1.374 with respect to the drought condition. There is only one lipid with this particular effect.

The estimated mean ratio for the negative drought effect within position=0 (bottom) for PG 34_3, PG 32_1, PC 36_2, PC 36_4, PE 36_3, PE 42_2, MGDG 36_6, PG 34_4, PC 36_3 and PI 34_2 is 0.662, 0.605, 0.588, 0.660, 0.640, 0.496, 0.674, 0.631, 0.647 and 0.700 (drought=1/drough=0) respectively, are similarly computed. So at the leaf's bottom (position=0), lipid PG 34_3 has estimated mean concentration under the water condition (drought=1) at 0.662 with respect to the estimated mean concentration under the drought condition (drought=0). And similarly for lipids PG 32_1, PC 36_2, PC 36_4, PE 36_3, PE 42_2, MGDG 36_6, PG 34_4, PC 36_3 and PI 34_2. Figure 9 summarizes these results with a bar chart.

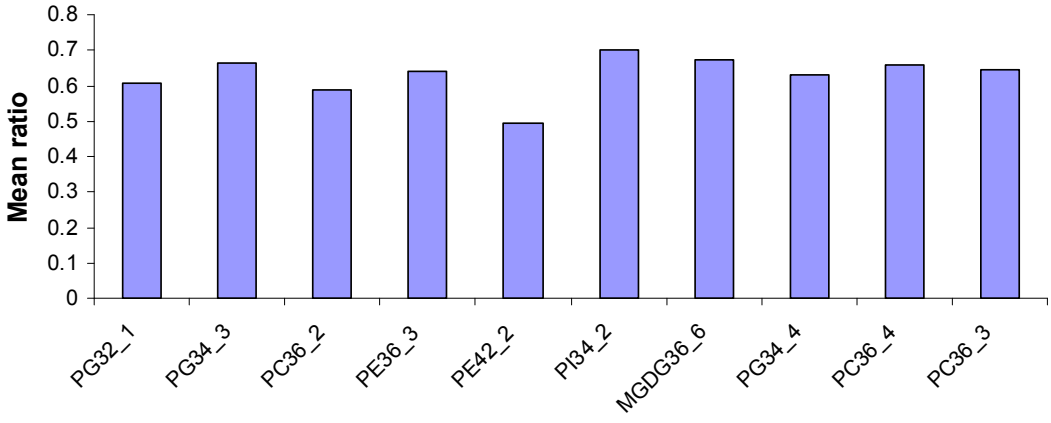


**Figure 9. Estimated mean ratio test computed for drought=1/drought=0 within position=0 condition for lipids PG 34_3, PG 32_1, PC 36_2, PC 36_4, PE 36_3, PE 42_2, MGDG 36_6, PG 34_4, PC 36_3 and PI 34_2.**

The estimated mean ratio for the positive position effect within drought=1 (water condition) for PC 36_2, PG 34_4, PG 32_1, PI 34_3, PC 34_2, PC 36_3, PI 34_2, PG 34_3, PE 42_2, PE 36_3 and PC 36_4 is 1.545, 1.567, 1.402, 1.271, 1.326, 1.318, 1.417, 1.514, 1.727, 1.370 and 1.385 (position=1/position=0) respectively, are similarly computed. So when the water condition is present (drought=1), lipid PC 36_2 has estimated mean concentration on the top (position=1) at 1.545 with respect to the estimated mean concentration at the bottom (position=0). And similarly for lipids PG 34_4, PG 32_1, PI 34_3, PC 34_2, PC 36_3, PI 34_2, PG 34_3, PE 42_2, PE 36_3 and PC 36_4. Figure 10 summarizes these results with a bar chart.

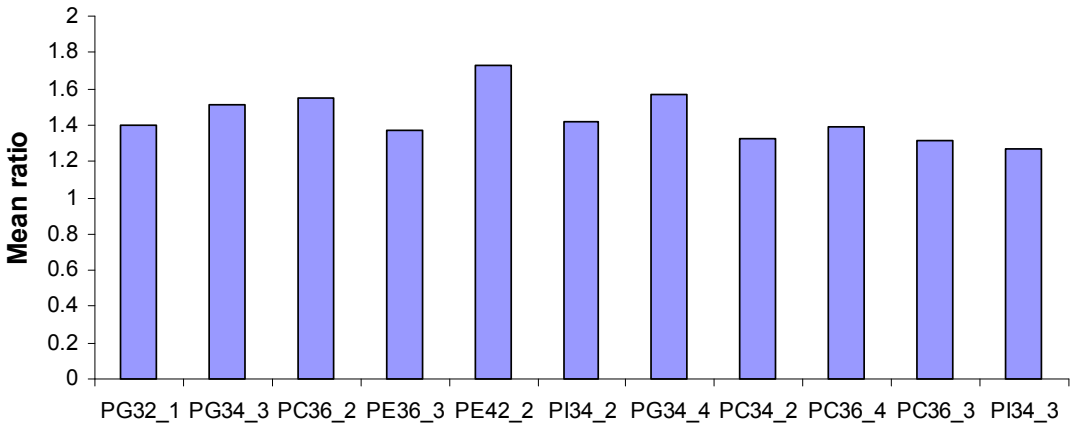


**Figure 10. Estimated mean ratio test computed position1/position0 within drought=1 condition for lipid PC 36_2, PG 34_4, PG 32_1, PI 34_3, PC 34_2, PC 36_3, PI 34_2, PG 34_3, PE 42_2, PE 36_3 and PC 36_4.**

There were no lipids with a significant rust*position interaction in the model (2). So no further estimation of contrasts associated with this interaction was pursued.

### 4.1.3 Test of main effects

Finally, the 35 lipids with no significant three-way interaction $(\alpha\beta\gamma)_{ijk}$ and no significant two way interactions $(\alpha\beta)_{ij}$ , $(\alpha\gamma)_{ik}$ or $(\beta\gamma)_{jk}$ were then analyzed using model (3) against a null model that drops one of the main effects. Again, the variance components were kept the same in both models and p-values were computed as discussed above for the test of three-way and two-way interactions. The same significance level of 0.05 was used for all tests. To show results for the 35 analyses more clearly, results for the no zeros group and 1-10 zeros group are separately reported, even though we did not observe anything unusually different in the model fits between these two groups.

#### 4.1.3.1 Analyses of main effects for the no zeros group of lipids

Figure 11, 12 and 13 are the bar charts to summarize the lowest 10 p-values computed for a drought effect, position effect and rust effect, respectively, in the no-zeros group. Recall that the parametric bootstrap procedure only used 500 bootstrap samples, so when p-values are very small, p-values computed this way may be equal to zero.

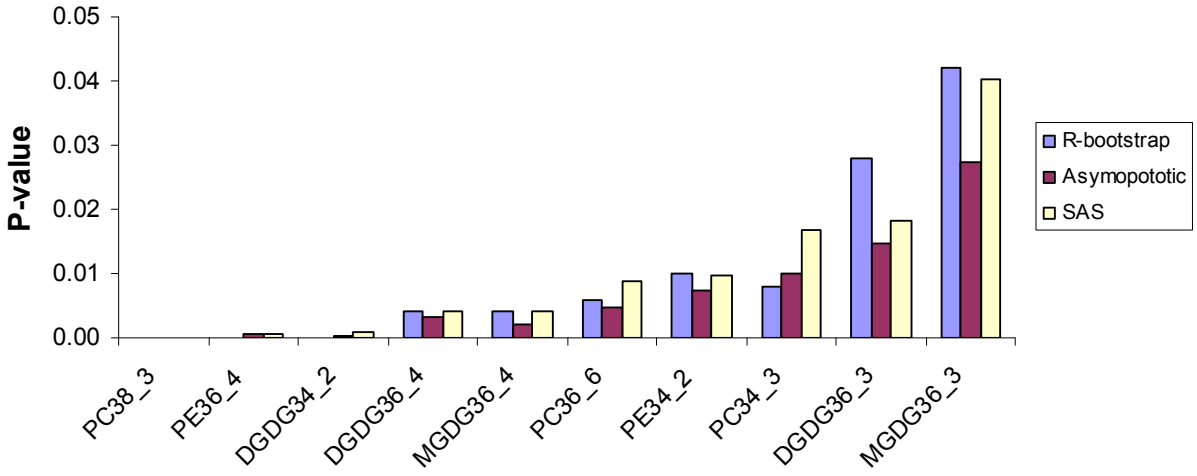


**Figure 11** Comparison of p-values computed for a drought effect for the lipids with the 10 smallest p-values in the no-zeros group using model 3.

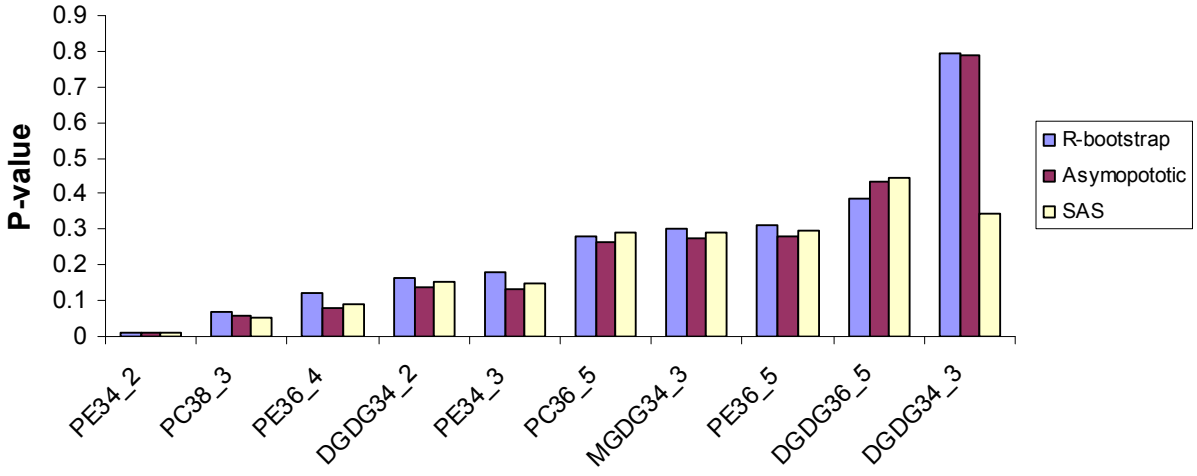


**Figure 12. Comparison of p-values computed for a position effect for the lipids with 10 smallest p-values in the no zeros group using model 3.**

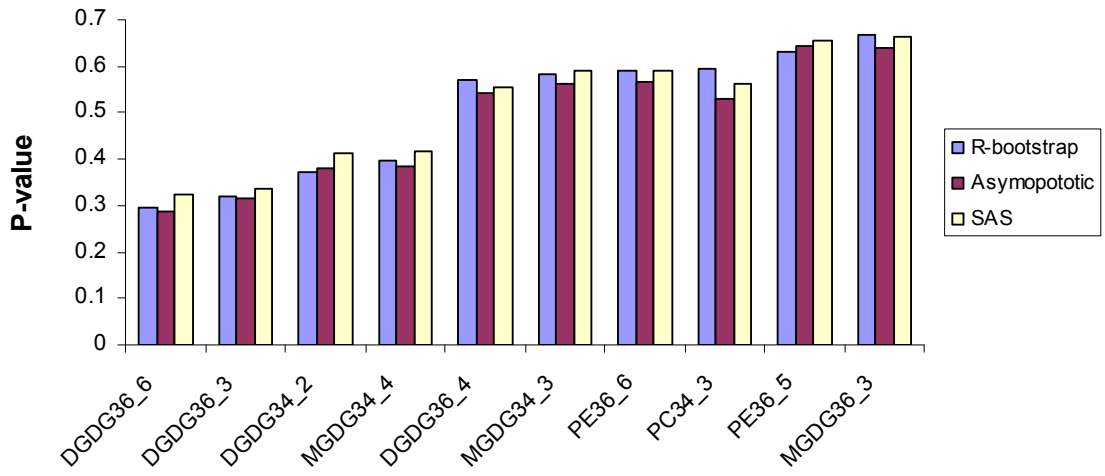


**Figure 13. Comparison of p-values computed for a rust effect for the lipids with the 10 smallest p-values in the no zeros group using model 3.**

Based on analyses of fixed effects, there are 12 lipids in the no zeros group having a significant drought effect, 1 lipid has a significant position effect, and no lipids have a significant rust effect. Table 6 shows the confidence intervals for associated contrasts for these effects.

**Table 6. 95% confidence intervals for the lipids not covering zero in the main effects model of no zero group lipids. $\mu_{DRP}$ represents the mean lipid concentration for drought condition D, rust condition R and position P. The upper left box is an approximate 95% confidence interval for the contrast, $\mu_{1..}-\mu_{0..}$ which means an interval estimate for a mean drought main effect (watered condition minus drought condition). $\mu_{..1}-\mu_{..0}$ denotes an estimate for a mean position effect (top minus bottom of leaf).**

| $\mu_{1..}-\mu_{0..}$ | | | | | |
|---|---|---|---|---|---|
| | Lower CL | Upper CL | | Lower CL | Upper CL |
| **DGDG34_2** | **-0.085** | **-0.029** | **PC34_3** | **0.028** | **0.196** |
| **DGDG36_4** | **-0.066** | **-0.013** | **PC36_6** | **0.018** | **0.091** |
| **DGDG36_3** | **-0.158** | **-0.014** | **MGDG34_4** | **-0.023** | **-0.000** |
| **MGDG34_2** | **-0.024** | **-0.000** | **PC38_3** | **0.013** | **0.028** |
| **MGDG36_4** | **-0.149** | **-0.036** | **PE34_2** | **-0.085** | **-0.012** |
| **MGDG36_3** | **-0.033** | **-0.002** | **PE36_4** | **-0.044** | **-0.013** |

| $\mu_{..1}-\mu_{..0}$ | | |
|---|---|---|
| | Lower CL | Upper CL |
| **PE34_2** | **0.012** | **0.083** |

From the confidence intervals, we can see that there are 9 lipids with an estimated negative mean drought effect, DGDG 34_2, DGDG 36_4, DGDG 36_3, MGDG 34_2, MGDG 36_4, MGDG 36_3, MGDG 34_4, PE 34_2 and PE 36_4, (watered condition minus drought condition). There are 3 lipids with a positive drought effect, PC 34_3, PC 36_6 and MGDG 34_4. For the lipids having a positive or negative drought effect as shown by interval estimates that do not cover zero, point estimates were also computed to show the magnitude of change in

the mean concentration levels. Figure 14 is the chart to show the magnitude of change for a drought effect (drought=1 (water)/drought=0 (drought)) for these 12 lipids.

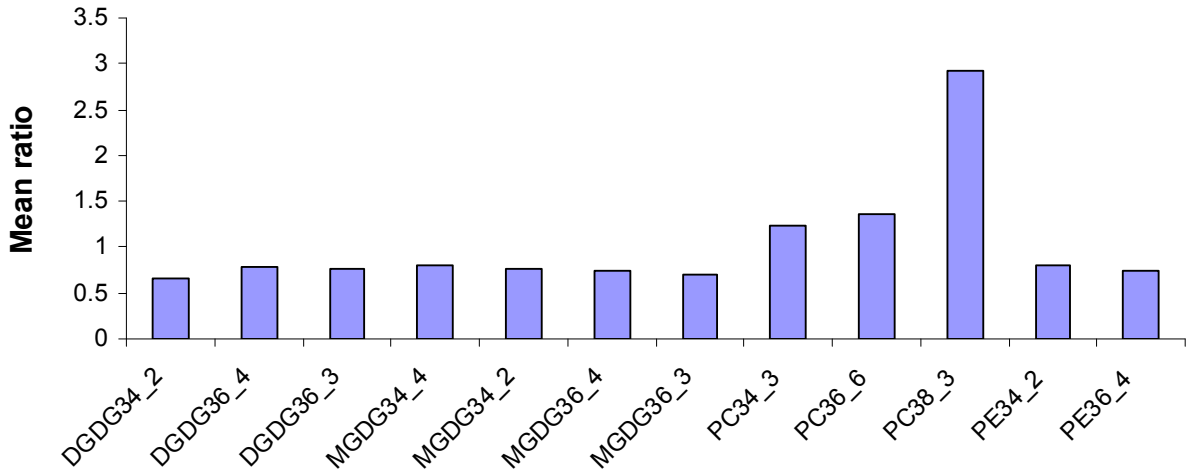


**Figure 14. Mean ratio estimate computed for drought=1 (water)/drought=0 (drought) for 12 lipids.**

There was only 1 lipid with a significant positive mean position effect in the no zeros group. The estimated magnitude of change for the position effect (position=1 (top)/position=0 (bottom)) for this particular lipid is 1.2376, meaning that it was present at the top of the leaf at 1.2 times that at the bottom.

There were no lipids having a significant rust effect in no zeros group.

#### 4.1.3.2 Analysis of main effects in the 1-10 zeros group

There are 15 lipids in 1-10 zeros group that are analyzed for significant main effects. Figure 15, 16 and 17 are the bar charts to summarize the lowest 10 p-values computed for a drought effect, position effect and rust effect, respectively, in 1-10 zeros group. Again, very small p-values are sometimes rounded or computed (i.e., in the bootstrap procedure) as zero.

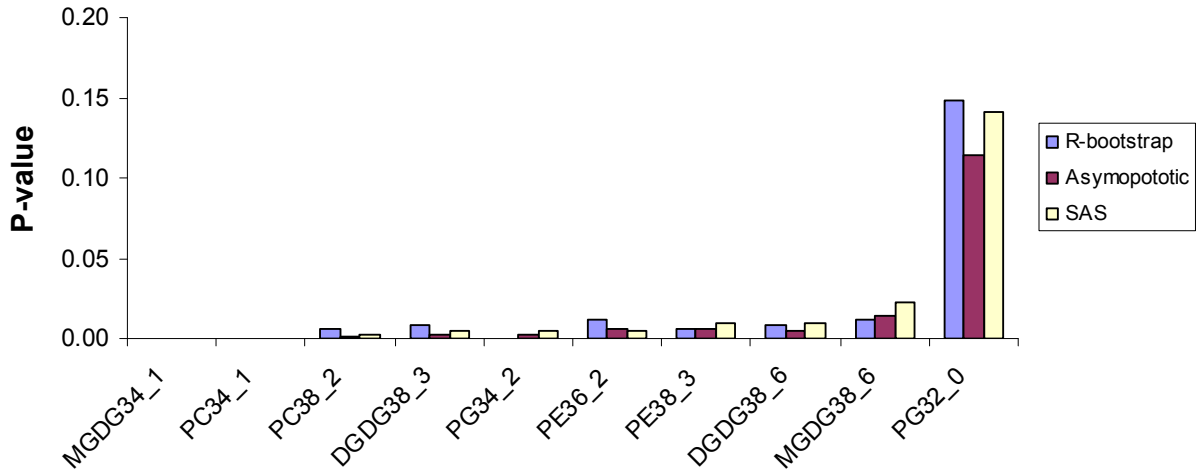


**Figure 15. Comparison of p-values computed for a mean drought effect main for the 10 smallest p-values in 1-10 zeros group.**

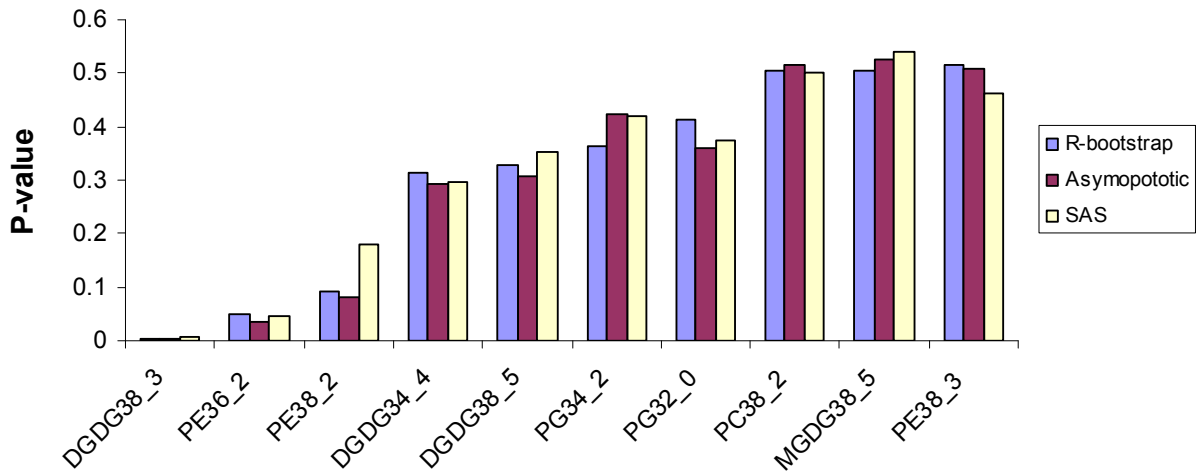


**Figure 16. Comparison of p-values computed for a mean position main effect for the 10 smallest p-values in 1-10 zeros group.**

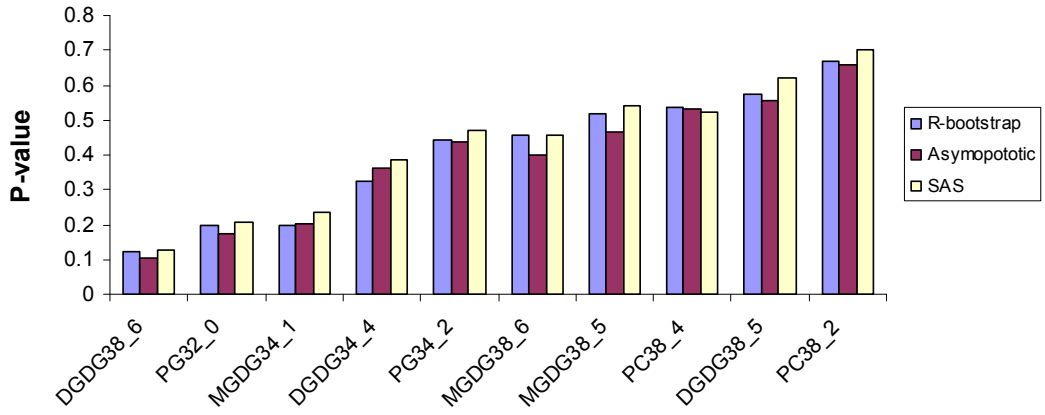


**Figure 17 Comparison of p-values computed for a mean rust main effect for the 10 smallest p-values in 1-10 zeros group.**

There are 9 lipids in the 1-10 zeros group having a significant mean drought effect, 2 lipids have a significant position effect, and no lipids have a significant rust effect. Table 7 shows the confidence intervals for the lipids with a significant main effect.

**Table 7. 95% confidence intervals for the lipids with a significant main effect in the 1-10 zeros group.**

| $\mu_{1..}-\mu_{0..}$ | | | | | |
|---|---|---|---|---|---|
| | Lower CL | Upper CL | | Lower CL | Upper CL |
| **DGDG 38_6** | **-0.018** | **-0.003** | **PC 34_1** | **-0.023** | **-0.010** |
| **DGDG 38_3** | **0.001** | **0.004** | **PC 38_2** | **0.003** | **0.010** |
| **MGDG 34_1** | **-0.022** | **-0.010** | **PG 34_2** | **-0.076** | **-0.018** |
| **MGDG 38_6** | **-0.007** | **-0.001** | **PE 36_2** | **-0.003** | **-0.001** |
| **PE 38_3** | **0.000** | **0.002** | | | |
| $\mu_{..1}-\mu_{..0}$ | | | | | |
| | Lower CL | Upper CL | | Lower CL | Upper CL |
| **DGDG 38_3** | **-0.003** | **-0.001** | **PE 36_2** | **0.000** | **0.003** |

There were 6 lipids with a negative mean drought effect (watered condition minus drought condition), DGDG 38_6, MGDG 34_1, MGDG 38_6, PG 34_2, PC 34_1 and PE 36_2 . There were 3 lipids with a positive drought effect, DGDG 38_3, PC 38_2 and PE 38_3. Figure 18 shows the bar charts representing the magnitude of change for a drought effect (drought=1 (water)/drought=0 (drought)) for these 9 lipids.

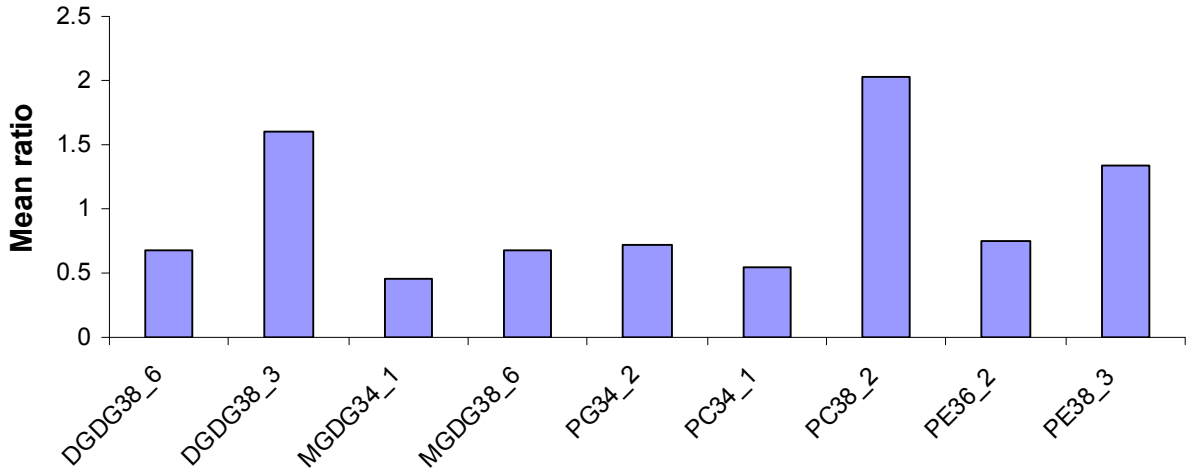


**Figure 18. Mean ratio estimate computed for a drought=1 (water)/drought=0 (drought) for 9 lipids with a significant drought main effect.**

There were 2 lipids having a significant position effect in the 1-10 zeros group, 1 with a negative position effect (top minus bottom position, DGDG 38_3, and 1 lipids with a positive position effect, PE 36_2. Figure 19 is the bar chart of the magnitude of change for the position effect (position=1 (top)/position=0 (bottom)) for these 2 lipids.

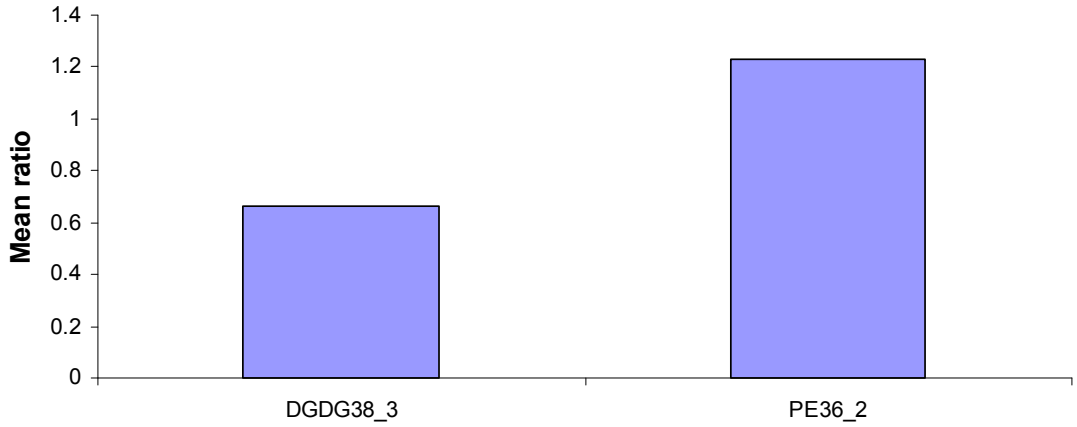


**Figure 19 Mean ratio estimate computed for a position=1 (top)/position=0 (bottom) for lipid with a significant position main effect.**

There were no lipids with significant rust main effect in 1-10 zeros group.

## §4.2 Categorical analyses result

For the 42 lipids with 11 to 99 zero concentrations across the 110 samples, a categorical analysis method was used to test for a drought effect, a rust effect, and a position effect. The objective is to determine if the pattern of zero levels in concentration was associated with the main effect treatments. P-values were computed two ways using the R functions, *prop.test* (a chi-square test for independence) and *fisher.test* (Fisher's exact test, cf. Conover, 1999). A liberal significance level of 0.05 was used for all tests to favor detection of any drought, rust, and position significant effects. There were 7 lipids with a significant drought effect. They are PG 34_1, LysoPC 16_0, LysoPC 18_3, PC 32_0, PS 42_2, PS 40_2 and PC 40_2. When determining significance, p-values from Fisher's test were used, though most of these agreed closely with those from the chi-square test.

**Table 8. Distribution of lipids present/not present for a drought effect for lipids PG 34_1, LysoPC 16_0, LysoPC 18_3, PC 32_0, PS 42_2, PS 40_2 and PC 40_2. Drought = 0 means drought condition and drought = 1 means watered. Present = 1 means the lipid was detected as present in the sample and 0 means the lipid was not detected**

| Drought | Lipid Present /not Present | PG34_1 | LysoPC 16_0 | LysoPC 18_3 | PC32_0 | PS42_2 | PS40_2 | PC40_2 |
|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 56 | 28 | 12 | 17 | 47 | 25 | 11 |
| 0 | 0 | 0 | 28 | 44 | 39 | 9 | 31 | 45 |
| 1 | 1 | 40 | 47 | 36 | 35 | 36 | 14 | 29 |
| 1 | 0 | 14 | 7 | 18 | 19 | 18 | 40 | 25 |

Table 8 shows the seven lipids for which the distribution of zeros is detected as dependent on the drought condition. Figure 20 is the bar chart for these effects.

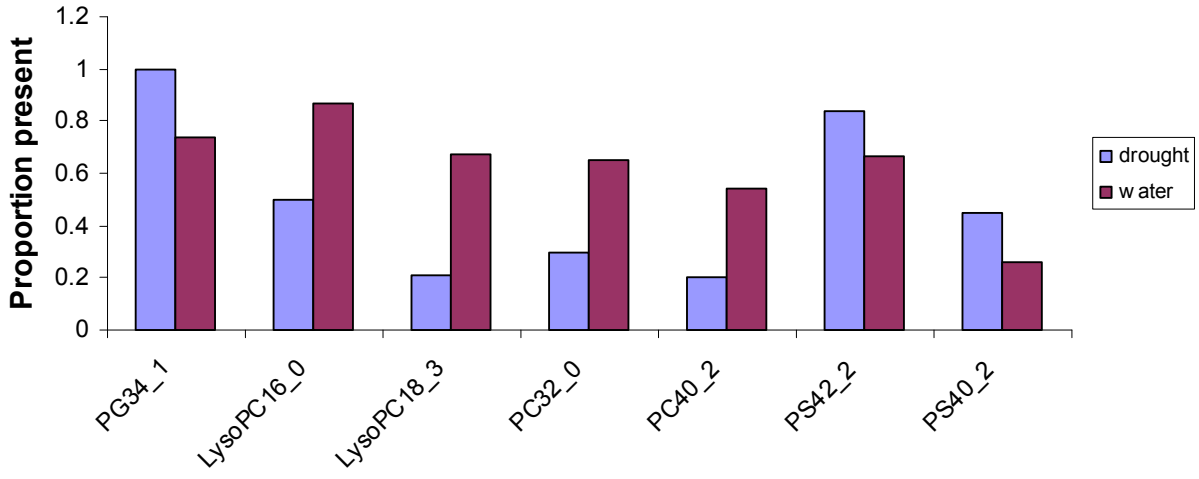


**Figure 20. Categorical data analysis of drought effect for lipid PG 34_1, LysoPC 16_0, LysoPC 18_3, PC 32_0, PS 42_2, PS 40_2 and PC 40_2. The y-axis is the proportion of samples for which the lipid was detected.**

Lipid PG 34_1, PS 42_2 and PS 40_2 are present at a proportion under the drought condition that is significantly higher than the corresponding proportion under the watered condition. Lipids LysoPC 16_0, LysoPC 18_3, PC 32_0, and PC 40_2 are present at a proportion under the drought condition that is significantly lower than that of the watered condition. Table 9 reports the average concentration for the drought and watered condition for these 7 lipids, drought=0 is drought condition, drought=1 is watered condition.

**Table 9. The average concentration for the drought and watered condition for lipid PG 34_1, LysoPC 16_0, LysoPC 18_3, PC 32_0, PS42_2, PS40_2 and PC 40_2.**

| Drought | PG34_1 | LysoPC 16_0 | LysoPC 18_3 | PC32_0 | PS42_2 | PS40_2 | PC40_2 |
|---|---|---|---|---|---|---|---|
| 0 | 0.074 | 0.001 | 0.001 | 0.001 | 0.021 | 0.002 | 0.000 |
| 1 | 0.032 | 0.003 | 0.002 | 0.003 | 0.012 | 0.002 | 0.001 |

There were 5 lipids with a significant position effect. They were LysoPE 16_0, PI 36_2, PA 34_2, LysoPG 18_2 and PA 36_5.

**Table 10. Similar to table 8 but for a position effect for lipids LysoPE 16_0, PI 36_2, PA 34_2, LysoPG 18_2 and PA 36_5.**

| Position | Lipid Present /not Present | LysoPE16_0 | PI36_2 | PA34_2 | LysoPG18_2 | PA36_5 |
|---|---|---|---|---|---|---|
| 1 | 1 | 11 | 17 | 12 | 14 | 11 |
| 1 | 0 | 45 | 39 | 44 | 42 | 44 |
| 0 | 1 | 23 | 7 | 32 | 24 | 23 |
| 0 | 0 | 31 | 47 | 22 | 30 | 31 |

Figure 21 is the bar chart to show the pattern in table 10.

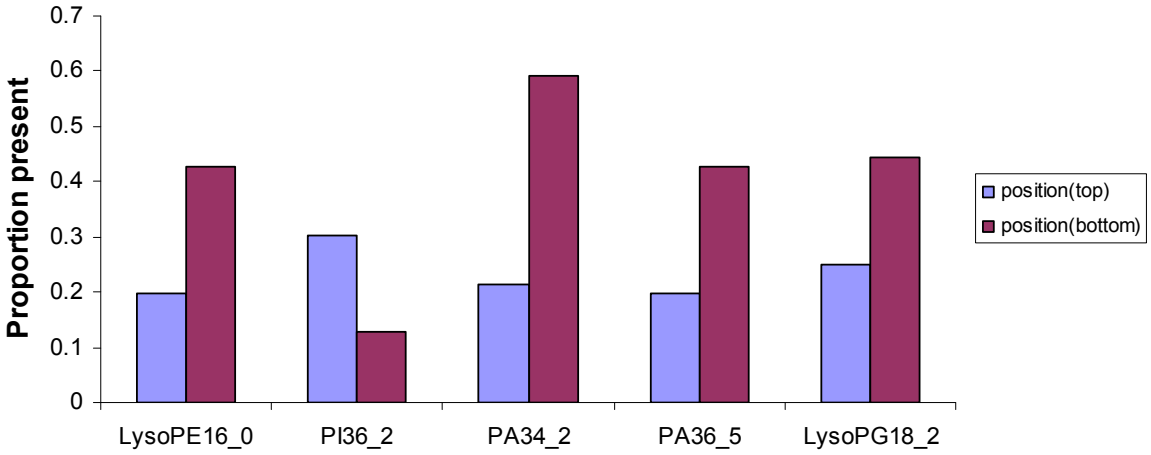


**Figure 21 Categorical data analysis of a position effect for lipids LysoPE 16_0, PI 36_2, PA 34_2, LysoPG 18_2 and PA 36_5. The y-axis is the proportion of lipids that were detected as present in each of the two positions.**

Figure 21 displays lipid PI 36_2 as present at a proportion that is higher at the top position versus the lower position. Lipids LysoPE 16_0, PA 34_2, LysoPG 18_2 and PA 36_5 are present in samples at a proportion that is lower at the top of the leaf versus the bottom. Table 11 reports the average concentration for the top of the leaf and the bottom of the leaf for these 5 lipids, position=1 is the top of the leaf, position=0 is the bottom of the leaf.

**Table 11. The average concentration for the top of the leaf and the bottom of the leaf for lipid LysoPE 16_0, PI 36_2, PA 34_2, LysoPG 18_2 and PA 36_5.**

| Position | LysoPE 16_0 | PI 36_2 | PA 34_2 | LysoPG 18_2 | PA 36_5 |
|---|---|---|---|---|---|
| 1 | 0.000 | 0.001 | 0.001 | 0.002 | 0.001 |
| 0 | 0.001 | 0.000 | 0.007 | 0.003 | 0.002 |

There were no lipids with a significant rust effect in these categorical analyses.

# Chapter 5 - Additional comparison of P-values

This section presents some additional comparison of p-values. In particular, some unusual behavior of the bootstrapped likelihood ratio test for a fixed effect when random effects differ between a null and alternative model is shown for a test of the three-way interaction term. Model (2) is shown below, which includes all random effects and all fixed effects except for the three way interaction, $(\alpha\beta\gamma)_{ijk}$ . In the earlier section, this null model was compared with a full model that included the three-way interaction fixed effect. The asymptotic distribution of the likelihood ratio test statistic is chi-square with 1 degree of freedom. This distribution is known to produce liberal p-values (i.e., smaller than nominal) when used as a reference distribution for a test of fixed effects in a mixed effects model on data from an unbalanced design (see Faraway, 2006).

$$Y_{ijklm} = \mu + B_l + \alpha_i + \beta_j + (\alpha\beta)_{ij} + (B\alpha\beta)_{lij} + \gamma_k + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (B\alpha\gamma)_{lik} + (B\beta\gamma)_{ljk} + (B\alpha\beta\gamma)_{lijk} + \varepsilon_{ijklm}$$

……..(2)

A question that sometimes arises is whether the variance component for the three-way interaction, $(B\alpha\beta\gamma)_{lijk}$ , should be dropped as well from the full model, when testing for a significant fixed effects three-way interaction. The asymptotic chi-square test would either have one degree of freedom if the variance components did not change between the two models (as was done earlier), or a two degree of freedom test if the variance component was also not included in the null model. This was evaluated with a parametric bootstrap routine that was used earlier, and reported in Faraway (2006).

Q-Q plots are used to compare the bootstrapped distribution of the likelihood ratio test statistic to the asymptotic chi-square distribution. A few cases are shown below.

Figure 22 is the QQ plot from an analysis of data for lipid DGDG 34_3 with the variance component, $(B\alpha\beta\gamma)_{ijkl}$ , included in the null model (i.e., model (2)). Figure 23 is QQ plot of lipid DGDG 34_3 that drops the variance component in the null model.

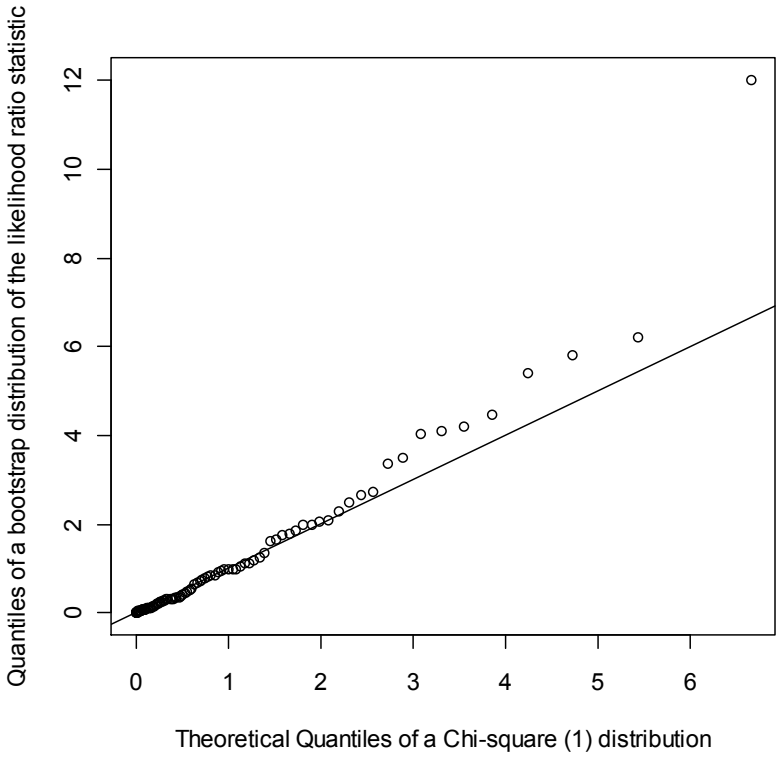


**Figure 22. Bootstrapped LRT approximations to the $\chi^2$ distribution with 1 degree of freedom. QQ plot of lipid DGDG34_3 with $(B\alpha\beta\gamma)_{ijkl}$ is in the model (2).**

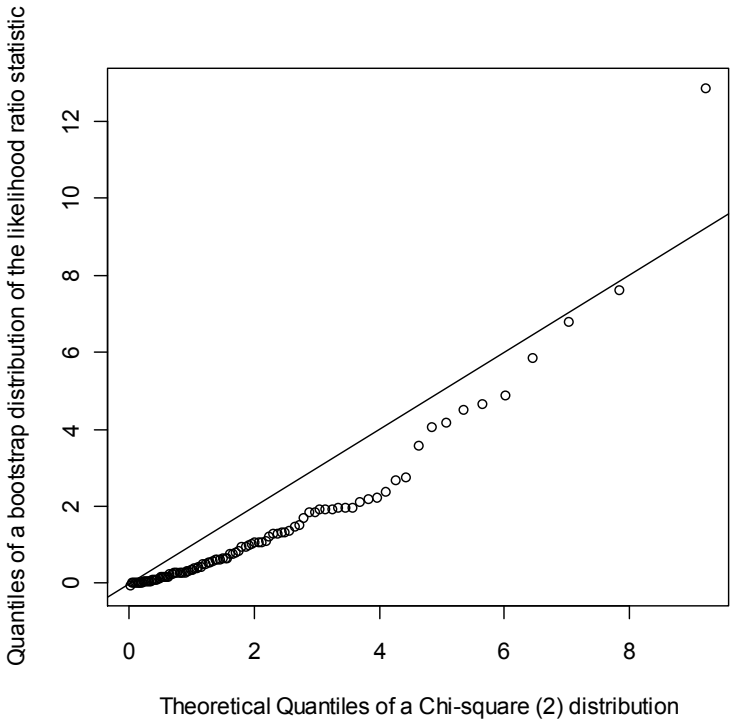


**Figure 23. Bootstrapped LRT approximations to the $\chi^2$ distribution with degree freedom of two. QQ plot of lipid DGDG34_3 with $(B\alpha\beta\gamma)_{ijkl}$ is not in the model (2).**

If the test statistic's theoretical reference distribution is appropriate, the points in the Q-Q plot will approximately lie on a line. Figure 25 shows that the use of the theoretical distribution will produce lower p-values than nominal for testing the significance of the fixed effect three way interactions. This is because the denominator of an appropriate "F-type" of statistic is assumed to be infinite.

Figure 23 exhibits behavior that is not expected. It is known that when testing for only variance components, p-values can be conservative but it is surprising how much conservative behavior is seen when testing a fixed and random effect simultaneously. An even more extreme case is shown below in Figures 24 and 25 where there is a significant departure from the theoretical distribution when the variance components differ between the null and alternative models.

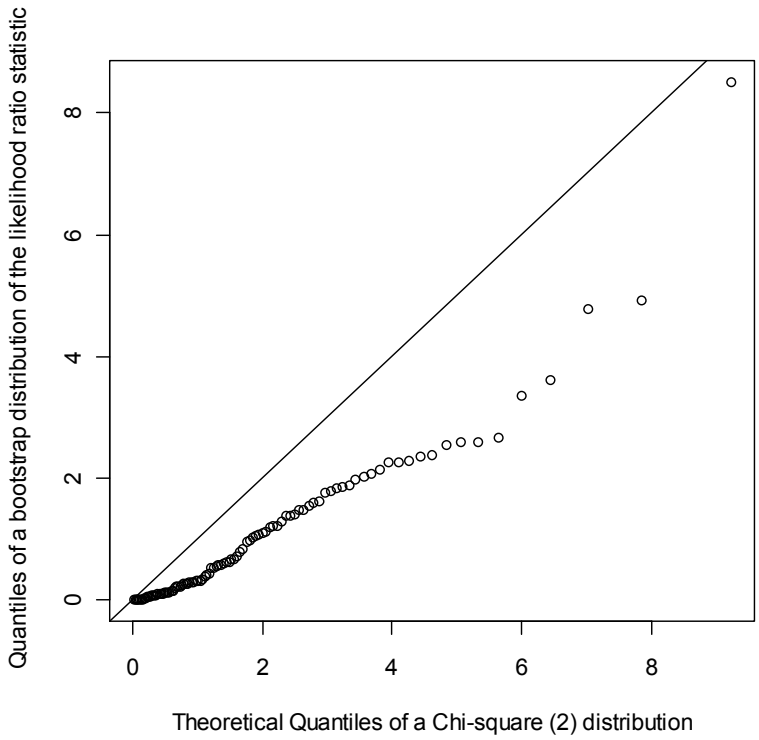


**Figure 24 Bootstrapped LRT approximations to the $\chi^2$ distribution with 2 degrees of freedom. QQ plot of lipid PC 36_6 with $(B\alpha\beta\gamma)_{ijkl}$ not in the model (2).**

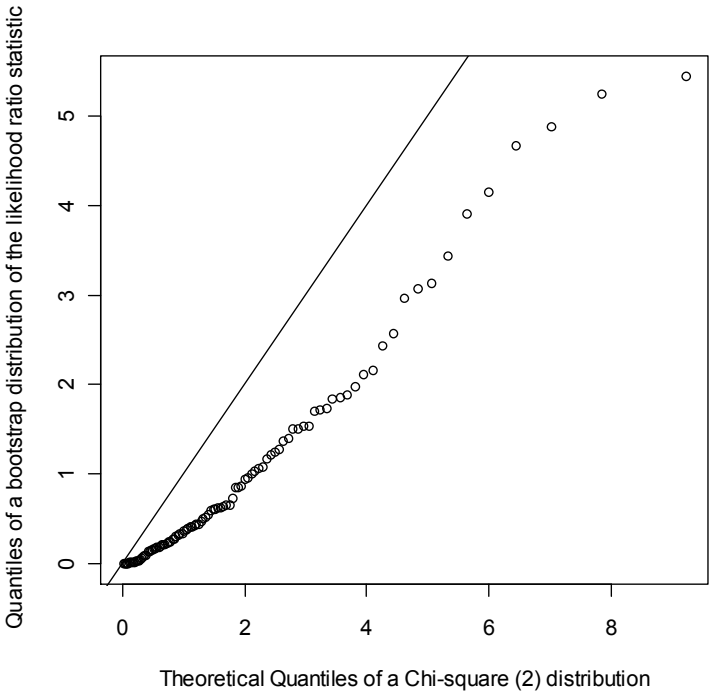


**Figure 25 Bootstrapped LRT approximations to the $\chi^2$ distribution with 2 degrees of freedom. QQ plot of lipid PC 36_5 with $(B\alpha\beta\gamma)_{ijkl}$ not in the model (2).**

For the parametric bootstrap routine, the R function *simulate* is used to parametrically simulate responses under the null model. What is not precisely known at this time is how this is done. More investigation is needed to determine the algorithms used to simulate responses and more simulations to determine if this is a common pattern, or if it is due to small numbers of replicates and unreliable estimates of variance components. For all earlier analyses being reported here, variance components were kept the same in all models under both null and alternative hypotheses.

# Chapter 6 - Summary

For the total 51 lipids in the mixed model, there was no lipid with a significant three way interactions effect, so all 51 lipids were tested for two way interactions. For the 51 lipids being tested, there were three lipids with a significant drought-rust interaction effect. These were DGDG 34_1, DGDG 36_2 and MGDG 36_5. Four 95% confidence intervals were computed for contrasts for each lipid where the drought-rust interaction term was significant at the 0.05 level. For these twelve confidence intervals, there were seven confidence intervals that were significant at the 0.05 level. For the lipids: DGDG 34_1, DGDG 36_2 and MGDG 36_5, the mean lipid concentration was higher in the drought condition versus the watered condition, when rust was present. The mean lipid concentration was higher in the drought condition versus watered condition for lipids DGDG 34_1 and DGDG 36_2, when rust was absent. For lipids DGDG 34_1 and DGDG 36_2, the mean lipid concentration was higher in the rust condition versus no rust condition, when it was in the drought condition. There were thirteen lipids with a significant drought*position interaction effect, PG 34_4, PG 32_1, PC 36_2, PC 36_4, PI 34_3, PC 34_2, PC 36_3, MGDG 36_6, PI 34_2, PG 34_3, PE 42_2, PC 34_4 and PE 36_3. Four 95% confidence intervals were also computed for contrasts for each lipid where the drought-position interaction term was significant at the 0.05 level. For these 52 confidence intervals, there were 22 confidence intervals that were significant at the 0.05 level. For the lipid PC 34_4, the mean lipid concentration was higher in the watered condition versus the drought condition at top of the leaf. The mean lipid concentration was lower in the watered condition versus the drought condition at the bottom of the leaf for lipids PG 34_4, PG 32_1, PC 36_2, PC 36_4, PC 36_3, MGDG 36_6, PI 34_2, PG 34_3, PE 42_2 and PE 36_3. For lipids PG 34_4, PG 32_1, PC 36_2, PC 36_4, PI 34_3, PC 34_2, PC 36_3, PI 34_2, PG 34_3, PE 42_2 and PE 36_3, the mean lipid concentration was higher at the top of the leaf versus at the bottom of the leaf, at the watered condition.

There were a total of 153 hypotheses that were tested for a two-way interaction. A false discovery rate (FDR) (Storey, 2003) adjustment was used to adjust p-values for multiple testing controls. The 153 two-way interaction p-values were combined and the R function, *qvalue* (www.r-project.org) used to obtain q-values. The lowest q-value was 0.233. So there

were no significant two-way interactions at a 0.1 level of FDR control. Figure 26 shows the plot of the p-values (x-axis) versus the qvalues (y-axis) to show how high the FDR values stay even at lower p-values.

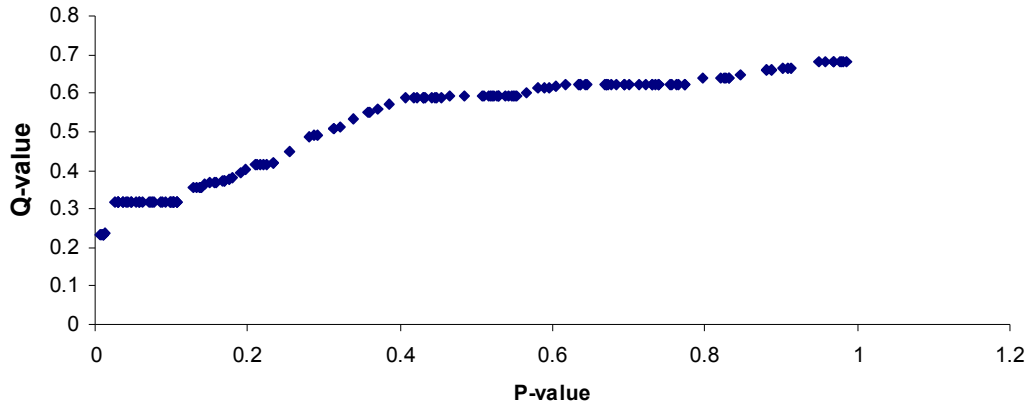


**Figure 26. Plot of the 153 p-values (x-axis) for two way interactions versus the corresponding 153 q-values (y-axis)**

There were 105 hypotheses tested for a main effect (drought, position and rust). The FDR method was used to adjust these p-values. The 105 main effect p-values were combined and the q-values computed as above. There were 20 significant q-values at the 0.1 level of FDR control. Table 12 shows these 20 p-values and the corresponding q-values.

**Table 12. Lipids from model (1) with a significant main effect at an FDR level of 0.1.**

| Variable | P-value | Main effect | Q-value |
|---|---|---|---|
| MGDG 38_6 | 0.0222 | drought | 0.091639 |
| DGDG 36_3 | 0.0183 | drought | 0.079516 |
| PC 34_3 | 0.0167 | drought | 0.076595 |
| PE 34_2 | 0.0113 | position | 0.054877 |
| DGDG 38_6 | 0.0099 | drought | 0.051083 |
| PE 34_2 | 0.0097 | drought | 0.051083 |
| PE 38_3 | 0.0092 | drought | 0.051083 |
| PC 36_6 | 0.0088 | drought | 0.051083 |
| DGDG 38_3 | 0.0059 | position | 0.040591 |
| PE 36_2 | 0.0051 | drought | 0.038277 |
| DGDG 38_3 | 0.005 | drought | 0.038277 |
| PG 34_2 | 0.005 | drought | 0.038277 |
| DGDG 36_4 | 0.0042 | drought | 0.038277 |
| MGDG 36_4 | 0.0042 | drought | 0.038277 |
| PC 38_2 | 0.0018 | drought | 0.024767 |
| DGDG 34_2 | 9.00E-04 | drought | 0.01486 |
| PE 36_4 | 6.00E-04 | drought | 0.012384 |
| PC 38_3 | 1.00E-04 | drought | 0.002752 |
| MGDG 34_1 | 1.00E-04 | drought | 0.002752 |
| PC 34_1 | 1.00E-04 | drought | 0.002752 |

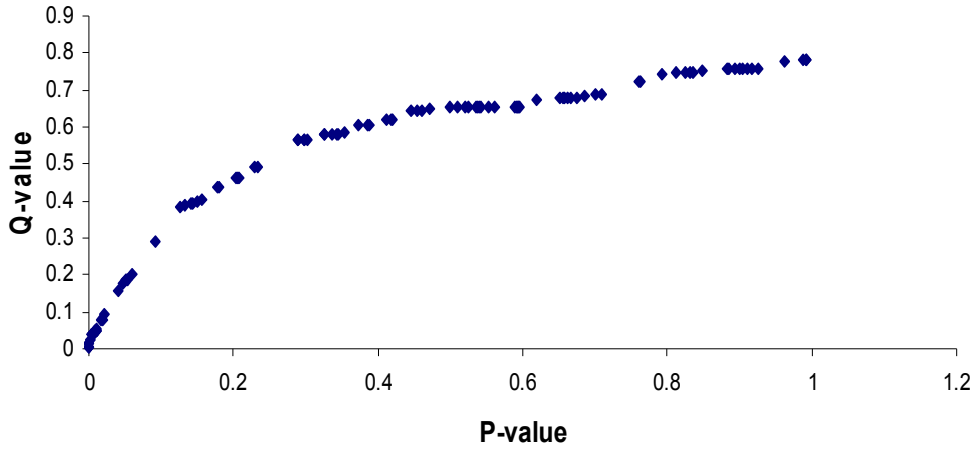


**Figure 27. Plot of the 105 p-values (x-axis) for a main effect versus the corresponding 105 q-values (y-axis)**

There were 126 hypotheses tested for a main effect (drought, position and rust) using the categorical analysis method.. The FDR method was also used to adjust these p-values. The 126 p-values were combined and the q-values computed as above. There were 6 significant q-values at the 0.1 level of FDR control. Table 13 shows the 6 p-values and the corresponding q-values.

**Table 13. Lipids from categorical data model with a significant main effect at an FDR level of 0.1.**

| Variable | P-value | Main effect | Q-value |
|---|---|---|---|
| PC32_0 | 0.000525 | drought | 0.011031 |
| PC40_2 | 0.00031 | drought | 0.007814 |
| PA34_2 | 8.10E-05 | position | 0.002552 |
| LysoPC16_0 | 3.33E-05 | drought | 0.001399 |
| PG34_1 | 1.77E-05 | drought | 0.001115 |
| LysoPC18_3 | 2.66E-06 | drought | 0.000335 |

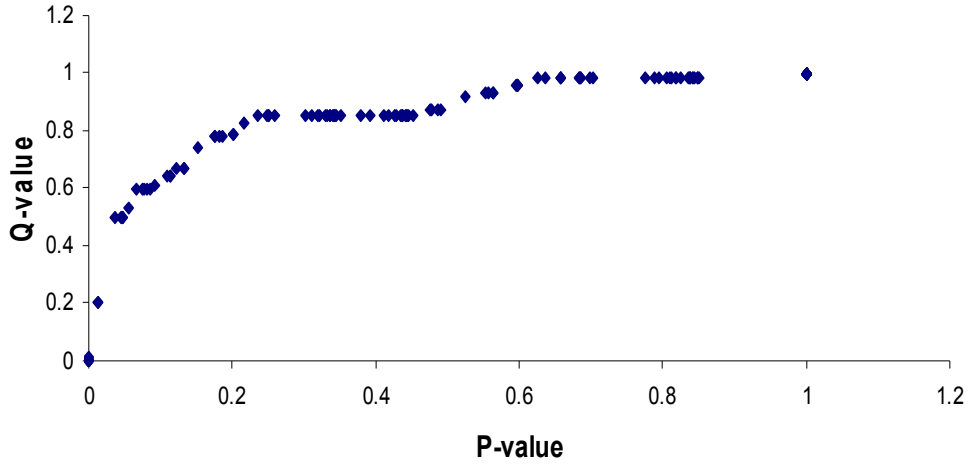


**Figure 28. Plot of the 126 p-values (x-axis) for the categorical analysis versus the corresponding 126 q-values (y-axis).**

Thus, in summary, if controlling for false discovery rate using the method reported in Storey. (2003), no two-way interactions are significant at an FDR level of 0.1, but 20 main effects remain significant at this level. An FDR level of control at 0.1, loosely stated, means that there is a 10% chance that a finding (i.e., a statistically significant result) will be a false discovery. Results from the analyses herein show that drought has the most significant effect on the lipidome of big bluestem, followed by the position on the leaf. The rust infection had little effect. More analyses could be conducted under more elaborate methods of FDR control, and to include a main effect analyses of all lipids grouped under model (2), i.e., those that showed a two way interaction at a significance level of 0.05.

# References

Benjamini, Yoav and Hochberg, Yosef (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B (Methodological) 57 (1): 289–300

Benjamini, Yoav and Yekutieli, Daniel (2001) The control of the false discovery rate in multiple testing under dependency. Annals of Statistics 29 (4): 1165–1188

Conver, W.J. (1999) Practical Nonparametric Statistics. John Wiley & Sons, Inc: New York.

Frank, Erin (2007) Rust and drought effects on gene expression and phytohormone concentration in big bluestem. Masters Thesis. Kansas State University.

Faraway, J. (2006) Extending the Linear Model with R .Chapman & Hall/CRC: New York.

Kenward and Roger (1997) Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood. Biometrics 53, 983-997

Storey, John D. (2003) The positive false discovery rate: A Bayesian interpretation and the q-value. Annals of Statistics 31 (6): 2013–2035

Yates, F. (1984). Tests of Significance for 2 x 2 Contingency Tables (with discussion). *Journal* of the Royal Statistical Society, Ser. *A* 147 (3): 426–463

# Computer Codes of SAS

## Mixed model effect

### *Test three way interaction*

```
proc import out= work.test
datafile = "C:\Documents and Settings\Lab\Desktop\nmol-ave-stdv.xls"
dbms = excel replace;
SHEET="no zero";
GETNAMES=YES; MIXED=NO; SCANTEXT=YES; USEDATE=YES; SCANTIME=YES;
run;
data one;
set work.test;
run;
proc print data=one;
run;
%let measurelist=DGDG34_3 DGDG34_2 DGDG34_1 DGDG36_6 DGDG36_5
DGDG36_4 DGDG36_3 MGDG34_4 MGDG34_3 MGDG34_2 MGDG36_6
MGDG36_4 MGDG36_3 PG34_4 PC34_3 PC34_2 PC36_6 PC36_5
PC36_4 PC36_3 PC38_3 PE34_3 PE34_2 PE36_6
PE36_5 PE36_4 PI34_3;
;
%let measurenum=27;
%macro lipidanalysis;
%do t=1 %to %eval(&measurenum) ;
 %let measure1 = %scan(&measurelist, %eval(&t)) ;
data tmp1; set one;
%fitmodel;
%end;
%mend ;
%macro fitmodel;
```

```
proc mixed  method=ml  cl covtest  data=tmp1;
class blk drought rust position;
model &measure1= drought|rust|position /ddfm=kr;
random blk blk*drought*rust position*blk*drought*rust;
ods listing  select  tests3;
ods output tests3=est;
run;
quit;
data est;
set est;
variable="&measure1";
sign=" ";
if probf<.01 then sign="*";run;
data estall;
set  estall est;
%mend fitmodel;
%lipidanalysis;
proc sort data=estall;
by effect;
run;
ods csv file='c:\Documents and Settings\Lab\Desktop\copy1.csv';
proc print data=estall;
run;
ods csv close;
```

***Test two way interaction***

```
proc import out= work.test
datafile = "C:\Documents and Settings\Lab\Desktop\nmol-ave-stdv.xls"
 dbms = excel replace;
SHEET="sheet1";
GETNAMES=YES; MIXED=NO; SCANTEXT=YES; USEDATE=YES; SCANTIME=YES;
```

```
run;
data one;
set work.test;
run;
proc print data=one;
run;
%let measurelist= DGDG34_3 DGDG34_2 DGDG34_1 DGDG36_6 DGDG36_5
DGDG36_4 DGDG36_3 MGDG34_4 MGDG34_3 MGDG34_2 MGDG36_6 MGDG36_4
MGDG36_3 PG34_4 PC34_3 PC34_2 PC36_6 PC36_5 PC36_4 PC36_3
PC38_3 PE34_3 PE34_2 PE36_6
PE36_5 PE36_4 PI34_3;
;
%let measurenum=27;
%macro lipidanalysis;
 %do t=1 %to %eval(&measurenum) ;
%let measure1 = %scan(&measurelist, %eval(&t)) ;
data tmp1; set one;
%fitmodel;
%end;
%mend ;
%macro fitmodel;
proc mixed  method=ml  cl covtest  data=tmp1;
class blk drought rust position;
model &measure1= drought rust position drought*rust drought*position rust*position /ddfm=kr;
random blk blk*drought*rust position*blk*drought*rust;
ods listing  select  tests3;
ods output tests3=est;
run;
quit;
data est;
set est;
```

```
variable="&measure1";
sign=" ";
if probf<.05 then sign="*";run;
data estall;
set  estall est;
%mend fitmodel;
%lipidanalysis;
proc sort data=estall;
by effect;
run;
ods csv file='c:\Documents and Settings\Lab\Desktop\copy1.csv';
proc print data=estall;
run;
ods csv close;
```

***Test main effect***

```
PROC IMPORT OUT= WORK.one
DATAFILE= "\\statsrvr\home\ttsong\Desktop\9-8-09-2.xls"
DBMS=EXCEL REPLACE;
SHEET="main effect-no-zero";
GETNAMES=YES;
 MIXED=NO;
 SCANTEXT=YES;
 USEDATE=YES;
SCANTIME=YES;
RUN;
proc print data=one;run;
%let measurelist=DGDG34_3          DGDG34_2    DGDG36_6    DGDG36_5    DGDG36_4
DGDG36_3    MGDG34_4    MGDG34_3    MGDG34_2    MGDG36_4    MGDG36_3
PC34_3          PC36_6          PC36_5          PC38_3  PE34_3  PE34_2  PE36_6 PE36_5
PE36_4;
```

```
;
%let measurenum=20;
%macro lipidanalysis;
 %do t=1 %to %eval(&measurenum) ;
 %let measure1 = %scan(&measurelist, %eval(&t)) ;
data tmp1; set one;
%fitmodel;
%end;
%mend ;
%macro fitmodel;
proc mixed  method=ml  cl covtest  data=tmp1;
class blk drought rust position;
model &measure1= drought rust position /ddfm=kr;
random blk blk*drought*rust position*blk*drought*rust;
ods listing  select  tests3;
ods output tests3=est;
run;
quit;
data est;
set est;
variable="&measure1";
sign=" ";
if probf<.01 then sign="*";run;
data estall;
set  estall est;
%mend fitmodel;
%lipidanalysis;
proc sort data=estall;
by effect;
run;
ods csv file='\\statsrvr\home\ttsong\Desktop\maineffectnozero.csv';
```

```
proc print data=estall;
run;
ods csv close;
```

# Computer Codes of R

## Mixed model effect

### *Test three way interaction*

```
/* ML method */
library(lme4)
setwd("C:\\Documents and Settings\\Lab\\My Documents\\ttsong\\lab-work\\data_2\\lipid\\9-8-09\\bootstrap")
data=read.table("9-8-09-NOZERO.txt",header=T)
for(i in 1:4){
data[,i]=as.factor(data[,i])}
# here data is the data matrix with column names
# as given in the model statement below
# cut is a cutoff for significance
PVALUE=0
tt=0
for(i in 5:dim(data)[2]){
y=data[,i]
lipid1=lmer(y~drought*rust*position+(1|blk)+(1|blk:drought:rust)+
(1| blk:position:drought:rust),data=data,REML=FALSE)
lipid2=lmer(y~drought+rust+position+drought:rust+drought:position+
rust:position +(1|blk)+(1|blk:drought:rust)+
(1|blk:drought:rust:position),data=data,REML=FALSE)
tt[i-4]=2*(logLik(lipid1)-logLik(lipid2))
PVALUE[i-4]=1-pchisq(tt[i-4],1)
}
write.csv(PVALUE, file = "R-result-for-three-way-nmol-nozero.csv")
/*bootstrap method*/
library(lme4)
```

```
setwd("C:\\Documents and Settings\\Lab\\My Documents\\ttsong\\lab-work\\data_2\\lipid\\9-8-09")
data=read.table("9-8-09-NOZERO.txt",header=T)
for(i in 1:4){
data[,i]=as.factor(data[,i])}
# here data is the data matrix with column names
# as given in the model statement below
# cut is a cutoff for significance
PVALUE=0
tvalue=0
for(j in 5:dim(data)[2]){
y=data[,j]
lipid1=lmer(y~drought*rust*position+(1|blk)+(1|blk:drought:rust)+
(1| blk:position:drought:rust),data=data,REML=FALSE)
lipid2=lmer(y~drought+rust+position+drought:rust+drought:position+
rust:position +(1|blk)+(1|blk:drought:rust)+
(1|blk:drought:rust:position),data=data,REML=FALSE)
anova(lipid1,lipid2)
stat=numeric(1000)
for(i in 1:1000){
ry=unlist(simulate(lipid2))
lipid1r=lmer(ry~drought*rust*position+(1|blk)+(1|blk:drought:rust)+
(1| blk:position:drought:rust),data=data,REML=FALSE)
lipid2r=lmer(ry~drought+rust+position+drought:rust+drought:position+
rust:position +(1|blk)+(1|blk:drought:rust)+
(1|blk:drought:rust:position),data=data,REML=FALSE)
stat[i]=2*(logLik(lipid1r)-logLik(lipid2r))
}
PVALUE[j-4]=mean(stat>anova(lipid1,lipid2)["lipid1","Chisq"])
tvalue[j-4]=(fixef(lipid1)/sqrt(diag(vcov(lipid1))))[8]
}
```

```
final=cbind(PVALUE,tvalue)
write.csv(final,file="final.csv")
```

### *Test two way interaction*

```
/* ML method test drought-rust*/
library(lme4)
setwd("C:\\Documents and Settings\\Lab\\My Documents\\ttsong\\lab-work\\data_2\\lipid\\9-8-09\\bootstrap")
data=read.table("9-8-09-NOZERO.txt",header=T)
for(i in 1:4){
data[,i]=as.factor(data[,i])}
# here data is the data matrix with column names
# as given in the model statement below
# cut is a cutoff for significance
PVALUE=0
tt=0
for(i in 5:dim(data)[2]){
y=data[,i]
lipid2=lmer(y~drought+rust+position+drought:position+rust:position+(1|blk)+(1|blk:drought:rust)+
(1| blk:position:drought:rust),data=data,REML=FALSE)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+
rust:position +(1|blk)+(1|blk:drought:rust)+
(1|blk:drought:rust:position),data=data,REML=FALSE)
tt[i-4]=2*(logLik(lipid1)-logLik(lipid2))
PVALUE[i-4]=1-pchisq(tt[i-4],1)
}
final2=cbind(PVALUE,tt)
write.csv(final2, file = "R-result-for-two-way-nmol-nozero-drought rust.csv")
/* ML method test rust-position*/
library(lme4)
```

```
setwd("C:\\Documents and Settings\\Lab\\My Documents\\ttsong\\lab-work\\data_2\\lipid\\9-8-09\\bootstrap")
data=read.table("9-8-09-NOZERO.txt",header=T)
for(i in 1:4){
data[,i]=as.factor(data[,i])}
# here data is the data matrix with column names
# as given in the model statement below
# cut is a cutoff for significance
PVALUE=0
tt=0
for(i in 5:dim(data)[2]){
y=data[,i]
lipid2=lmer(y~drought+rust+position+drought:rust+drought:position+
(1|blk)+(1|blk:drought:rust)+
(1| blk:position:drought:rust),data=data,REML=FALSE)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+
rust:position +(1|blk)+(1|blk:drought:rust)+
(1|blk:drought:rust:position),data=data,REML=FALSE)
tt[i-4]=2*(logLik(lipid1)-logLik(lipid2))
PVALUE[i-4]=1-pchisq(tt[i-4],1)
}
final2=cbind(PVALUE,tt)
write.csv(final2, file = "R-result-for-TWO-way-nmol-nozero-RUST-POSITION.csv")
```

**/* ML method test drought-position*/**

```
library(lme4)
setwd("C:\\Documents and Settings\\Lab\\My Documents\\ttsong\\lab-work\\data_2\\lipid\\9-8-09\\bootstrap")
data=read.table("9-8-09-NOZERO.txt",header=T)
for(i in 1:4){
data[,i]=as.factor(data[,i])}
# here data is the data matrix with column names
```

```
# as given in the model statement below
# cut is a cutoff for significance
PVALUE=0
tt=0
for(i in 5:dim(data)[2]){
y=data[,i]
lipid2=lmer(y~drought+rust+position+drought:rust+rust:position+(1|blk)+(1|blk:drought:rust)+
(1| blk:position:drought:rust),data=data,REML=FALSE)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+
rust:position +(1|blk)+(1|blk:drought:rust)+
(1|blk:drought:rust:position),data=data,REML=FALSE)
tt[i-4]=2*(logLik(lipid1)-logLik(lipid2))
PVALUE[i-4]=1-pchisq(tt[i-4],1)
}
final2=cbind(PVALUE,tt)
write.csv(final2, file = "R-result-for-two-way-nmol-nozero-drought position.csv")
/* bootstrap method test drought-rust*/
library(lme4)
setwd("C:\\Documents and Settings\\Lab\\My Documents\\ttsong\\lab-work\\data_2\\lipid\\9-8-
09\\bootstrap")
data=read.table("9-8-09-NOZERO.txt",header=T)
for(i in 1:4){
data[,i]=as.factor(data[,i])}
# here data is the data matrix with column names
# as given in the model statement below
# cut is a cutoff for significance
PVALUE=0
tvalue=0
for(j in 5:dim(data)[2]){
y=data[,j]
```

```
lipid2=lmer(y~drought+rust+position++drought:position+rust:position+(1|blk)+(1|blk:drought:ru
st)+(1| blk:position:drought:rust),data=data,REML=FALSE)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=data,REML=FALSE)
anova(lipid1,lipid2)
stat=numeric(500)
for(i in 1:500){
ry=unlist(simulate(lipid2))
lipid2r=lmer(ry~drought+rust+position+drought:position+rust:position+(1|blk)+(1|blk:drought:r
ust)+(1| blk:position:drought:rust),data=data,REML=FALSE)
lipid1r=lmer(ry~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=data,REML=FALSE)
stat[i]=2*(logLik(lipid1r)-logLik(lipid2r))
}
PVALUE[j-4]=mean(stat>anova(lipid1,lipid2)["lipid1","Chisq"])
tvalue[j-4]=(fixef(lipid1)/sqrt(diag(vcov(lipid1))))[5]
}
final=cbind(PVALUE,tvalue)
write.csv(final,file="final-TWO WAY-drought rust.csv")
/* bootstrap method test rust-position*/
library(lme4)
setwd("C:\\Documents and Settings\\Lab\\My Documents\\ttsong\\lab-
work\\data_2\\lipid\\9-8-09\\bootstrap")
data=read.table("9-8-09-NOZERO.txt",header=T)
for(i in 1:4){
data[,i]=as.factor(data[,i])}
# here data is the data matrix with column names
# as given in the model statement below
# cut is a cutoff for significance
PVALUE=0
tvalue=0
```

```
for(j in 5:dim(data)[2]){
y=data[,j]
lipid2=lmer(y~drought+rust+position++drought:rust+drought:position+(1|blk)+(1|blk:droug
ht:rust)+(1| blk:position:drought:rust),data=data,REML=FALSE)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=data,REML=FALSE)
anova(lipid1,lipid2)
stat=numeric(500)
for(i in 1:500){
ry=unlist(simulate(lipid2))
lipid2r=lmer(ry~drought+rust+position+drought:rust+drought:position+(1|blk)+(1|blk:drou
ght:rust)+(1| blk:position:drought:rust),data=data,REML=FALSE)
lipid1r=lmer(ry~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=data,REML=FALSE)
stat[i]=2*(logLik(lipid1r)-logLik(lipid2r))
}
PVALUE[j-4]=mean(stat>anova(lipid1,lipid2)["lipid1","Chisq"])
tvalue[j-4]=(fixef(lipid1)/sqrt(diag(vcov(lipid1))))[7]
}
final=cbind(PVALUE,tvalue)
write.csv(final,file="final-TWO WAY-1.csv")
```

**/* bootstrap method test drought-position*/**

```
library(lme4)
setwd("C:\\Documents and Settings\\Lab\\My Documents\\ttsong\\lab-
work\\data_2\\lipid\\9-8-09\\bootstrap")
data=read.table("9-8-09-NOZERO.txt",header=T)
for(i in 1:4){
data[,i]=as.factor(data[,i])}
# here data is the data matrix with column names
# as given in the model statement below
# cut is a cutoff for significance
```

```
PVALUE=0
tvalue=0
for(j in 5:dim(data)[2]){
y=data[,j]
lipid2=lmer(y~drought+rust+position++drought:rust+rust:position+(1|blk)+(1|blk:drought:r
ust)+(1| blk:position:drought:rust),data=data,REML=FALSE)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=data,REML=FALSE)
anova(lipid1,lipid2)
stat=numeric(500)
for(i in 1:500){
ry=unlist(simulate(lipid2))
lipid2r=lmer(ry~drought+rust+position+drought:rust+rust:position+(1|blk)+(1|blk:drought:r
ust)+(1| blk:position:drought:rust),data=data,REML=FALSE)
lipid1r=lmer(ry~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=data,REML=FALSE)
stat[i]=2*(logLik(lipid1r)-logLik(lipid2r))
}
PVALUE[j-4]=mean(stat>anova(lipid1,lipid2)["lipid1","Chisq"])
tvalue[j-4]=(fixef(lipid1)/sqrt(diag(vcov(lipid1))))[6]
}
final=cbind(PVALUE,tvalue)
write.csv(final,file="final-TWO WAY-drought position.csv")
```

**/* confidence interval and mean ratio test*/**

```
setwd("G:\\tingting song from computer\\ttsong\\lab-work\\data_2\\lipid\\9-8-09\\bootstrap\\two
way test")
x=read.table("droughtrust.txt",header=T)
library(lme4)
for(i in 1:4){
x[,i]=as.factor(x[,i])}
#Drought=0 rust = 0
```

```
low=0
up=0
for (i in 5:7){
y=x[,i]
d1=c(1,0,0,.5,0,0,0)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d1)%*%lipid1@fixef
error=sqrt(t(d1)%*%vcov(lipid1)%*%d1)[,1]
low[i]=mean-qt(0.975, 15) *error
up[i]=mean+qt(0.975, 15) *error
}
cbind(low,up)[5:7,]
#Drought=1 rust = 0
low=0
up=0
for (i in 5:7){
y=x[,i]
d2=c(1,1,0,.5,0,.5,0)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d2)%*%lipid1@fixef
error=sqrt(t(d2)%*%vcov(lipid1)%*%d2)[,1]
low[i]=mean-qt(0.975, 15) *error
up[i]=mean+qt(0.975, 15) *error
}
cbind(low,up)[5:7,]
#Drought= effect within rust = 0
low=0
up=0
for (i in 5:7){
```

```
y=x[,i]
d3=c(0,1,0,0,0,.5,0)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d3)%*%lipid1@fixef
error=sqrt(t(d3)%*%vcov(lipid1)%*%d3)[,1]
low[i]=mean-qt(0.975, 15) *error
up[i]=mean+qt(0.975, 15) *error
}
cbind(low,up)[5:7,]
#Drought=1 rust = 1
low=0
up=0
for (i in 5:7){
y=x[,i]
d4=c(1,1,1,.5,1,.5,.5)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d4)%*%lipid1@fixef
error=sqrt(t(d4)%*%vcov(lipid1)%*%d4)[,1]
low[i]=mean-qt(0.975,15) *error
up[i]=mean+qt(0.975,15) *error
}
cbind(low,up)[5:7,]
#Drought=0 rust = 1
low=0
up=0
for (i in 5:7){
y=x[,i]
d5=c(1,0,1,.5,0,0,.5)
```

```
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d5)%*%lipid1@fixef
error=sqrt(t(d5)%*%vcov(lipid1)%*%d5)[,1]
low[i]=mean-qt(0.975,15) *error
up[i]=mean+qt(0.975,15) *error
}
cbind(low,up)[5:7,]
#Drought effect within rust = 1
low=0
up=0
for (i in 5:7){
y=x[,i]
d6=c(0,1,0,0,1,.5,0)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d6)%*%lipid1@fixef
error=sqrt(t(d6)%*%vcov(lipid1)%*%d6)[,1]
low[i]=mean-qt(0.975,15) *error
up[i]=mean+qt(0.975,15) *error
}
cbind(low,up)[5:7,]
#Rust effect within drought = 0
low=0
up=0
for (i in 5:7){
y=x[,i]
d31=c(0,0,1,0,0,0,.5)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d31)%*%lipid1@fixef
```

```
error=sqrt(t(d31)%*%vcov(lipid1)%*%d31)[,1]
low[i]=mean-qt(0.975,15) *error
up[i]=mean+qt(0.975,15) *error
}
cbind(low,up)[5:7,]
#Rust effect within drought = 1
low=0
up=0
for (i in 5:7){
y=x[,i]
d61=c(0,0,1,0,1,0,.5)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d61)%*%lipid1@fixef
error=sqrt(t(d61)%*%vcov(lipid1)%*%d61)[,1]
low[i]=mean-qt(0.975,15) *error
up[i]=mean+qt(0.975,15) *error
}
cbind(low,up)[5:7,]
data=read.table("droughtposition.txt",header=T)
library(lme4)
for(i in 1:4){
data[,i]=as.factor(data[,i])}
#Drought=0  position= 1
low=0
up=0
for (i in 5:17){
y=data[,i]
d7=c(1,0,.5,1,0,0,.5)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
```

```
mean=t(d7)%*%lipid1@fixef
error=sqrt(t(d7)%*%vcov(lipid1)%*%d7)[,1]
low[i]=mean-qt(0.975, 21) *error
up[i]=mean+qt(0.975, 21) *error
}
cbind(low,up)[5:17,]
#Drought=1  position= 1
low=0
up=0
for (i in 5:17){
y=data[,i]
d8=c(1,1,.5,1,.5,1,.5)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d8)%*%lipid1@fixef
error=sqrt(t(d8)%*%vcov(lipid1)%*%d8)[,1]
low[i]=mean-qt(0.975, 21) *error
up[i]=mean+qt(0.975, 21) *error
}
cbind(low,up)[5:17,]
#Drought effect within position= 1
low=0
up=0
for (i in 5:17){
y=data[,i]
d9=c(0,1,0,0,.5,1,0)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d9)%*%lipid1@fixef
error=sqrt(t(d9)%*%vcov(lipid1)%*%d9)[,1]
low[i]=mean-qt(0.975, 21) *error
```

```
up[i]=mean+qt(0.975, 21) *error
}
cbind(low,up)[5:17,]
#Drought=0  position=0
low=0
up=0
for (i in 5:17){
y=data[,i]
d10=c(1,0,.5,0,0,0,0)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d10)%*%lipid1@fixef
error=sqrt(t(d10)%*%vcov(lipid1)%*%d10)[,1]
low[i]=mean-qt(0.975, 21) *error
up[i]=mean+qt(0.975, 21) *error
}
cbind(low,up)[5:17,]
#Drought=1  position=0
low=0
up=0
for (i in 5:17){
y=data[,i]
d11=c(1,1,.5,0,.5,0,0)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d11)%*%lipid1@fixef
error=sqrt(t(d11)%*%vcov(lipid1)%*%d11)[,1]
low[i]=mean-qt(0.975, 21) *error
up[i]=mean+qt(0.975, 21) *error
}
cbind(low,up)[5:17,]
```

```
#Drought effect within position=0
low=0
up=0
for (i in 5:17){
y=data[,i]
d12=c(0,1,0,0,.5,0,0)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d12)%*%lipid1@fixef
error=sqrt(t(d12)%*%vcov(lipid1)%*%d12)[,1]
low[i]=mean-qt(0.975, 21) *error
up[i]=mean+qt(0.975, 21) *error
}
cbind(low,up)[5:17,]
#position effect within drought=0
low=0
up=0
for (i in 5:17){
y=data[,i]
d91=c(0,0,0,1,0,0,.5)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d91)%*%lipid1@fixef
error=sqrt(t(d91)%*%vcov(lipid1)%*%d91)[,1]
low[i]=mean-qt(0.975, 21) *error
up[i]=mean+qt(0.975, 21) *error
}
cbind(low,up)[5:17,]
#position effect within drought=1
low=0
up=0
```

```
for (i in 5:17){
y=data[,i]
d121=c(0,0,0,1,0,1,.5)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean=t(d121)%*%lipid1@fixef
error=sqrt(t(d121)%*%vcov(lipid1)%*%d121)[,1]
low[i]=mean-qt(0.975, 21) *error
up[i]=mean+qt(0.975, 21) *error
}
cbind(low,up)[5:17,]
#mean ratio test
#Drought=0 rust = 0
mean00=0
for (i in 5:7){
y=x[,i]
d1=c(1,0,0,.5,0,0,0)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean00[i]=t(d1)%*%lipid1@fixef
}
mean00[5:7]
#Drought=1 rust = 0
mean10=0
for (i in 5:7){
y=x[,i]
d2=c(1,1,0,.5,0,.5,0)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean10[i]=t(d2)%*%lipid1@fixef
}
```

```
mean10[5:7]
#Drought=1 rust = 1
mean11=0
for (i in 5:7){
y=x[,i]
d4=c(1,1,1,.5,1,.5,.5)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean11[i]=t(d4)%*%lipid1@fixef
}
mean11[5:7]
#Drought=0 rust = 1
mean01=0
for (i in 5:7){
y=x[,i]
d5=c(1,0,1,.5,0,0,.5)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean01[i]=t(d5)%*%lipid1@fixef
}
mean01[5:7]
#####mean ratio for drought effect within rust=0
mean10[5:7]/mean00[5:7]
##########mean ratio drought effect within rust=1
mean11[5:7]/mean01[5:7]
#########mean ratio rust effect within drought=0
mean01[5:7]/mean00[5:7]
##########mean ratio rust effect within drought=1
mean11[5:7]/mean10[5:7]
data=read.table("droughtposition.txt",header=T)
library(lme4)
```

```
for(i in 1:4){
data[,i]=as.factor(data[,i])}
#Drought=0  position= 1
mean001=0
for (i in 5:17){
y=data[,i]
d7=c(1,0,.5,1,0,0,.5)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean001[i]=t(d7)%*%lipid1@fixef
}
mean001[5:17]
#Drought=1  position= 1
mean101=0
for (i in 5:17){
y=data[,i]
d8=c(1,1,.5,1,.5,1,.5)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean101[i]=t(d8)%*%lipid1@fixef
}
mean101[5:17]
#Drought=0  position=0
mean000=0
for (i in 5:17){
y=data[,i]
d10=c(1,0,.5,0,0,0,0)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean000[i]=t(d10)%*%lipid1@fixef
}
```

```
mean000[5:17]
#Drought=1  position=0
mean100=0
for (i in 5:17){
y=data[,i]
d11=c(1,1,.5,0,.5,0,0)
lipid1=lmer(y~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=x,REML=FALSE)
mean100[i]=t(d11)%*%lipid1@fixef
}
mean100[5:17]
#####mean ratio for drought effect within position=0
mean100[5:17]/mean000[5:17]
##########mean ratio drought effect within position=1
mean101[5:17]/mean001[5:17]
#########mean ratio position effect within drought=0
mean001[5:17]/mean000[5:17]
##########mean ratio position effect within drought=1
mean101[5:17]/mean100[5:17]
```

***Test main effect***

```
/*ML method */
library(lme4)
setwd("H:\\tingting song from computer\\ttsong\\lab-work\\data_2\\lipid\\9-8-09\\main
effect\\R")
data=read.table("main effect no zero-2.txt",header=T)
for(i in 1:4){
data[,i]=as.factor(data[,i])}
# here data is the data matrix with column names
# as given in the model statement below
# cut is a cutoff for significance
```

```
####test position effect
PVALUE=0
tt=0
for(i in 5:dim(data)[2]){
y=data[,i]
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+
(1| blk:position:drought:rust),data=data,REML=FALSE)
lipid2=lmer(y~drought+rust+(1|blk)+(1|blk:drought:rust)+(1|
blk:position:drought:rust),data=data,REML=FALSE)
tt[i-4]=2*(logLik(lipid1)-logLik(lipid2))
PVALUE[i-4]=1-pchisq(tt[i-4],1)
}
final5=cbind(PVALUE,tt)
write.csv(final5, file = "R-result-for-maineffectposition-nmol-nozero-2.csv")
#######test drought effect
PVALUE=0
tt=0
for(i in 5:dim(data)[2]){
y=data[,i]
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|
blk:position:drought:rust),data=data,REML=FALSE)
lipid2=lmer(y~rust+position+(1|blk)+(1|blk:drought:rust)+(1|
blk:position:drought:rust),data=data,REML=FALSE)
tt[i-4]=2*(logLik(lipid1)-logLik(lipid2))
PVALUE[i-4]=1-pchisq(tt[i-4],1)
}
final6=cbind(PVALUE,tt)
write.csv(final6, file = "R-result-for-maineffectdrought-nmol-nozero-2.csv")
#######test rust effect
PVALUE=0
tt=0
```

```
for(i in 5:dim(data)[2]){
y=data[,i]
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|
blk:position:drought:rust),data=data,REML=FALSE)
lipid2=lmer(y~drought+position+(1|blk)+(1|blk:drought:rust)+(1|
blk:position:drought:rust),data=data,REML=FALSE)
tt[i-4]=2*(logLik(lipid1)-logLik(lipid2))
PVALUE[i-4]=1-pchisq(tt[i-4],1)
}
final7=cbind(PVALUE,tt)
write.csv(final7, file = "R-result-for-maineffectrust-nmol-nozero-2.csv")
/* bootstrap method*/
library(lme4)
setwd("H:\\tingting song from computer\\ttsong\\lab-work\\data_2\\lipid\\9-8-09\\main
effect\\R")
data=read.table("main effect no zero-2.txt",header=T)
for(i in 1:4){
data[,i]=as.factor(data[,i])}
# here data is the data matrix with column names
# as given in the model statement below
# cut is a cutoff for significance
####test position effect
PVALUE=0
tvalue=0
for(j in 5:dim(data)[2]){
y=data[,j]
lipid2=lmer(y~drought+rust+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=data
,REML=FALSE)
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),
data=data,REML=FALSE)
anova(lipid1,lipid2)
```

```
stat=numeric(500)
for(i in 1:500){
ry=unlist(simulate(lipid2))
lipid2r=lmer(ry~drought+rust+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=da
ta,REML=FALSE)
lipid1r=lmer(ry~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position
),data=data,REML=FALSE)
stat[i]=2*(logLik(lipid1r)-logLik(lipid2r))
}
PVALUE[j-4]=mean(stat>anova(lipid1,lipid2)["lipid1","Chisq"])
tvalue[j-4]=(fixef(lipid1)/sqrt(diag(vcov(lipid1))))[4]
}
final=cbind(PVALUE,tvalue)
write.csv(final,file="final-maineffect-nozerobootstrap-position-2.csv")
#####test drought effect
PVALUE=0
tvalue=0
for(j in 5:dim(data)[2]){
y=data[,j]
lipid2=lmer(y~position+rust+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=data
,REML=FALSE)
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),
data=data,REML=FALSE)
anova(lipid1,lipid2)
stat=numeric(500)
for(i in 1:500){
ry=unlist(simulate(lipid2))
lipid2r=lmer(ry~position+rust+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=da
ta,REML=FALSE)
lipid1r=lmer(ry~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position
),data=data,REML=FALSE)
```

```
stat[i]=2*(logLik(lipid1r)-logLik(lipid2r))
}
PVALUE[j-4]=mean(stat>anova(lipid1,lipid2)["lipid1","Chisq"])
tvalue[j-4]=(fixef(lipid1)/sqrt(diag(vcov(lipid1))))[2]
}
final=cbind(PVALUE,tvalue)
write.csv(final,file="final-maineffect-nozerobootstrap-drought-2.csv")
#####test rust effect
PVALUE=0
tvalue=0
for(j in 5:dim(data)[2]){
y=data[,j]
lipid2=lmer(y~drought+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=
data,REML=FALSE)
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),
data=data,REML=FALSE)
anova(lipid1,lipid2)
stat=numeric(500)
for(i in 1:500){
ry=unlist(simulate(lipid2))
lipid2r=lmer(ry~drought+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),dat
a=data,REML=FALSE)
lipid1r=lmer(ry~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position
),data=data,REML=FALSE)
stat[i]=2*(logLik(lipid1r)-logLik(lipid2r))
}
PVALUE[j-4]=mean(stat>anova(lipid1,lipid2)["lipid1","Chisq"])
tvalue[j-4]=(fixef(lipid1)/sqrt(diag(vcov(lipid1))))[3]
}
final=cbind(PVALUE,tvalue)
```

```
write.csv(final,file="final-maineffect-nozerobootstrap-rust-2.csv")
/* Confidence interval*/
setwd("H:\\tingting song from computer\\ttsong\\lab-work\\data_2\\lipid\\9-8-09\\main
effect\\R")
x=read.table("main effect no zero-2.txt",header=T)
library(lme4)
for(i in 1:4){
x[,i]=as.factor(x[,i])}
#Drought effect
low=0
up=0
for (i in 5:24){
y=x[,i]
d1=c(0,1,0,0)
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),
data=x,REML=FALSE)
mean=t(d1)%*%lipid1@fixef
error=sqrt(t(d1)%*%vcov(lipid1)%*%d1)[,1]
low[i]=mean-qt(0.975, 16) *error
up[i]=mean+qt(0.975, 16) *error
}
cbind(low,up)[5:24,]
#rust effct
low=0
up=0
for (i in 5:24){
y=x[,i]
d2=c(0,0,1,0)
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),
data=x,REML=FALSE)
mean=t(d2)%*%lipid1@fixef
```

```
error=sqrt(t(d2)%*%vcov(lipid1)%*%d2)[,1]
low[i]=mean-qt(0.975, 16) *error
up[i]=mean+qt(0.975, 16) *error
}
cbind(low,up)[5:24,]
#position effct
 low=0
up=0
for (i in 5:24){
 y=x[,i]
d2=c(0,0,0,1)
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),
data=x,REML=FALSE)
mean=t(d2)%*%lipid1@fixef
error=sqrt(t(d2)%*%vcov(lipid1)%*%d2)[,1]
low[i]=mean-qt(0.975, 23) *error
up[i]=mean+qt(0.975, 23) *error
}
cbind(low,up)[5:24,]
data=read.table("maineffect1-10zero.txt",header=T)
library(lme4)
for(i in 1:4){
 data[,i]=as.factor(data[,i])}
#Drought effect
 low=0
 up=0
 for (i in 5:24){
 y=data[,i]
d1=c(0,1,0,0)
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),
data=data,REML=FALSE)
```

```
mean=t(d1)%*%lipid1@fixef
error=sqrt(t(d1)%*%vcov(lipid1)%*%d1)[,1]
low[i]=mean-qt(0.975, 16) *error
up[i]=mean+qt(0.975, 16) *error
}
cbind(low,up)[5:24,]
#rust effct
 low=0
up=0
for (i in 5:24){
y=data[,i]
d2=c(0,0,1,0)
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position)
data=data,REML=FALSE)
mean=t(d2)%*%lipid1@fixef
error=sqrt(t(d2)%*%vcov(lipid1)%*%d2)[,1]
low[i]=mean-qt(0.975, 16) *error
up[i]=mean+qt(0.975, 16) *error
}
cbind(low,up)[5:24,]
#position effct
low=0
up=0
for (i in 5:24){
 y=data[,i]
 d2=c(0,0,0,1)
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position)
data=data,REML=FALSE)
mean=t(d2)%*%lipid1@fixef
error=sqrt(t(d2)%*%vcov(lipid1)%*%d2)[,1]
low[i]=mean-qt(0.975, 23) *error
```

```
up[i]=mean+qt(0.975, 23) *error
}
cbind(low,up)[5:24,]
#mean ratio test for no zero
 #Drought effect
  mean100=0
 for (i in 5:24){
 y=x[,i]
d1=c(1,1,0.5,0.5)
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),
data=x,REML=FALSE)
mean100[i]=t(d1)%*%lipid1@fixef
}
mean100[5:24]
mean0001=0
for (i in 5:24){
y=x[,i]
d1=c(1,0,0.5,0.5)
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),
data=x,REML=FALSE)
mean0001[i]=t(d1)%*%lipid1@fixef
}
mean0001[5:24]
#####mean ratio for drought effect
mean100[5:24]/mean0001[5:24]
#rust effect
mean010=0
for (i in 5:24){
y=x[,i]
 d1=c(1,0.5,1,0.5)
```

```
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),
data=x,REML=FALSE)
mean010[i]=t(d1)%*%lipid1@fixef
}
mean010[5:24]
mean0002=0
for (i in 5:24){
y=x[,i]
d1=c(1,0.5,0,0.5)
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),
data=x,REML=FALSE)
mean0002[i]=t(d1)%*%lipid1@fixef
}
mean0002[5:24]
#####mean ratio for rust effect
 mean010[5:24]/mean0002[5:24]
#position effect
mean001=0
for (i in 5:24){
 y=x[,i]
 d1=c(1,0.5,0.5,1)
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),
data=x,REML=FALSE)
mean001[i]=t(d1)%*%lipid1@fixef
}
mean001[5:24]
mean0003=0
for (i in 5:24){
y=x[,i]
d1=c(1,0.5,0.5,0)
```

```
lipid1=lmer(y~drought+rust+position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),
data=x,REML=FALSE)
mean0003[i]=t(d1)%*%lipid1@fixef
}
mean0003[5:24]
#####mean ratio for position effect
mean001[5:24]/mean0003[5:24]
```

## *Categorical Data Analysis*

```
setwd("C:\\Documents and Settings\\Lab\\My Documents\\ttsong\\lab-work\\data_2\\lipid\\9-8-
09")
data=read.table("drought.txt",header=T)
final1=0
final2=0
for(i in 1:2){
data[,i]=as.factor(data[,i])}
for(j in 3:44){
y=data[,j]
x=cbind(c(y[1],y[3]),c(y[2],y[4]))
final1[j-2]=prop.test(x)$p.value
final2[j-2]=fisher.test(x)$p.value
}
final.drought=cbind(final1,final2)
write.csv(final.drought,file="final-drought.csv")
final3=0
final4=0
data2=read.table("rust.txt",header=T)
for(i in 1:2){
data2[,i]=as.factor(data[,i])}
for(j in 3:44){
y=data2[,j]
x=cbind(c(y[1],y[3]),c(y[2],y[4]))
```

```
final3[j-2]=prop.test(x)$p.value
final4[j-2]=fisher.test(x)$p.value
}
final.rust=cbind(final3,final4)
write.csv(final.rust,file="final-rust.csv")
final5=0
final6=0
data2=read.table("position.txt",header=T)
for(i in 1:2){
data2[,i]=as.factor(data[,i])}
for(j in 3:44){
y=data2[,j]
x=cbind(c(y[1],y[3]),c(y[2],y[4]))
final5[j-2]=prop.test(x)$p.value
final6[j-2]=fisher.test(x)$p.value
}
final.rust=cbind(final3,final4)
write.csv(final.position,file="final-position.csv")
```

## *QQ plot for DGDG34_3*

```
/* QQ plot with one degree of freedom*/
library(lme4)
setwd("G:\\tingting song from computer\\ttsong\\lab-work\\data_2\\lipid\\9-8-09")
data=read.table("DGDG34_3.txt",header=T)
for(i in 1:4){
data[,i]=as.factor(data[,i])}
# here data is the data matrix with column names
# as given in the model statement below
# cut is a cutoff for significance
lipid1=lmer(DGDG34_3~drought*rust*position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust
:position),data=data,REML=FALSE)
```

```
lipid2=lmer(DGDG34_3~drought+rust+position+drought:rust+drought:position+rust:position+(
1|blk)+(1|blk:drought:rust)+
(1|blk:drought:rust:position),data=data,REML=FALSE)
anova(lipid1,lipid2)
stat=numeric(100)
for(i in 1:100){
ry=unlist(simulate(lipid2))
lipid1r=lmer(ry~drought*rust*position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position
),data=data,REML=FALSE)
lipid2r=lmer(ry~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position),data=data,REML=FALSE)
stat[i]=2*(logLik(lipid1r)-logLik(lipid2r))
}
plot(qchisq((1:100)/101,2),sort(stat), main = "lipidDGDG34_3", xlab="Theoretical Quantiles of
a Chi-square (1) distribution", ylab=" Quantiles of a bootstrap distribution of the likelihood ratio
statistic"))
abline(0,1)
/* QQ plot with two degree of freedom*/
library(lme4)
setwd("G:\\tingting song from computer\\ttsong\\lab-work\\data_2\\lipid\\9-8-09")
data=read.table("DGDG34_3.txt",header=T)
for(i in 1:4){
data[,i]=as.factor(data[,i])}
# here data is the data matrix with column names
# as given in the model statement below
# cut is a cutoff for significance
lipid1=lmer(DGDG34_3~drought*rust*position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust
:position),data=data,REML=FALSE)
lipid2=lmer(DGDG34_3~drought+rust+position+drought:rust+drought:position+rust:position
+(1|blk)+(1|blk:drought:rust),data=data,REML=FALSE)
anova(lipid1,lipid2)
```

```
stat=numeric(100)
for(i in 1:100){
ry=unlist(simulate(lipid2))
lipid1r=lmer(ry~drought*rust*position+(1|blk)+(1|blk:drought:rust)+(1|blk:drought:rust:position
),data=data,REML=FALSE)
lipid2r=lmer(ry~drought+rust+position+drought:rust+drought:position+rust:position+(1|blk)+(1|
blk:drought:rust),data=data,REML=FALSE)
stat[i]=2*(logLik(lipid1r)-logLik(lipid2r))
}
plot(qchisq((1:100)/101,2),sort(stat), main = "lipidDGDG34_3", xlab="Theoretical Quantiles of
a Chi-square (2) distribution", ylab=" Quantiles of a bootstrap distribution of the likelihood ratio
statistic")
abline(0,1)
```

***FDR test for two way interaction***

```
setwd("H:\\tingting song from computer\\ttsong\\lab-work\\data_2\\lipid\\9-8-09\\bootstrap\\two
way test")
library(qvalue)
qtest1=read.table("q-value.txt")
data.frame(qtest1)
qq1=qtest1[,2]
qqobj1=qvalue(qq1)
qplot(qqobj1)
final=cbind(qtest1[,1:3],qqobj1$qvalues)
qq2=qvalue(qq1, fdr.level=0.1)
final2=cbind(final[1:4],qq2$significant)
write.csv(final2, file = "q-value.csv")
```