

leafArea__30June2015

Purpose: I just want the total area between genotypes. I am performing the analysis using the plants that were grown for the asymmetry data.

```
library(tidyr)
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.1.3
```

```
library(reshape2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
##
## The following object is masked from 'package:stats':
##
##   filter
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
##ggplot custom theme
```

```
cbPalette <- c("#56B4E9", "#999999", "#E69F00", "#009E73", "#F0E442", "#0072B2", "#D55E00", "#CC79A7")

customTheme <- list(theme(axis.title.x = element_text(face="bold", size=30),
  axis.title.y = element_text(face="bold", size=30),
  axis.text.x = element_text(size=16),
  axis.text.y = element_text(size=16),
  legend.title = element_text(size=30),
  legend.text = element_text(size=20),
  strip.text.x = element_text(size = 20),
  strip.text.y = element_text(size = 20),
  legend.position="none",
  strip.background = element_rect(fill="#b0c4de"))
)
```

```
leafArea <- read.csv("../data/leafArea_30June2015.csv")
```

```
head(leafArea)
```

```
##           Label      Area Perim. Circ.    AR Round Solidity
## 1 L1_EA10.bmp  0.00800  0.836 0.147 3.046 0.328    0.600
## 2 L1_EA10.bmp 13.05300 34.747 0.136 1.248 0.802    0.861
## 3 L1_EA11.bmp 11.40200 27.369 0.191 1.099 0.910    0.864
## 4 L1_EA11.bmp 11.40200 27.380 0.191 1.099 0.910    0.864
## 5 L1_EA12.bmp 10.75500 31.696 0.135 1.213 0.824    0.845
## 6 L1_EA12.bmp  0.00068  0.078 1.000 2.000 0.500    1.000
```

Clean Up. I used `getArea.ijm` macros to get the the areas from binary .bmp files, but in order to achieve this I had to click in multiple areas of the same file to ensure I was able to grab the leaf using the magic wand tool. Therefore I have two measurements per file. I need to first get rid of any measurement below 3 cm squared. Then get rid of any double measurements.

```
dim(leafArea)
```

```
## [1] 2854    7
```

```
leafArea2 <- subset(leafArea, Area >= 3.0)
```

```
leafArea2 <- subset(leafArea2, Area <= 200)
dim(leafArea2)
```

```
## [1] 1897    7
```

Now time to get rid of the duplicate names.

```
leafArea3 <- subset(leafArea2, !duplicated(leafArea2$Label))
dim(leafArea3)
```

```
## [1] 1223    7
```

We are left with 1230, and there are a total of 1,430 files. The other files the leaves were not centered properly or have holes and will not be used further in analysis.

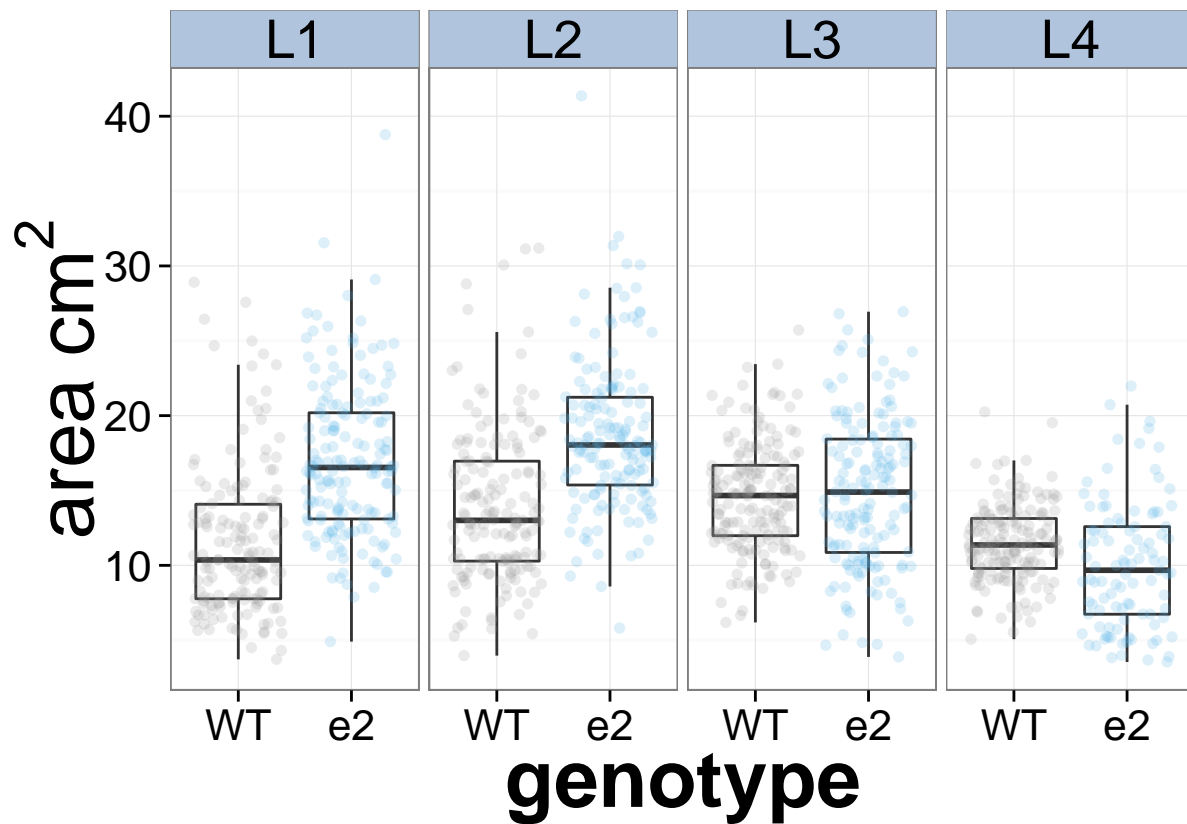
```
#Make Genotype Column
leafArea3$genotype <- ifelse(grepl("_E", leafArea3$Label, ignore.case = T), "e2",
                             ifelse(grepl("_R", leafArea3$Label, ignore.case = T), "WT", "WT"))

#Make ID column
leafArea4 <- separate(data = leafArea3, col = Label, into = c("Leaf", "ID"), sep = "_")
```

Area

```
ggplot(leafArea4, aes(genotype, Area)) +
  geom_boxplot(outlier.shape = NA) +
  geom_point(alpha = .2, position = "jitter", aes(color = genotype)) +
  scale_x_discrete(limits=c("WT", "e2")) +
  scale_color_manual(values=cbPalette) +
  facet_grid(.~Leaf) +
  theme_bw() +
  ylab(expression(paste("area ", "c", m2, sep=""))) +
  customTheme
```

```
## Warning: Removed 2 rows containing missing values (geom_point).
## Warning: Removed 6 rows containing missing values (geom_point).
## Warning: Removed 6 rows containing missing values (geom_point).
## Warning: Removed 5 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 2 rows containing missing values (geom_point).
```



Now I am going to just look at mature leaves

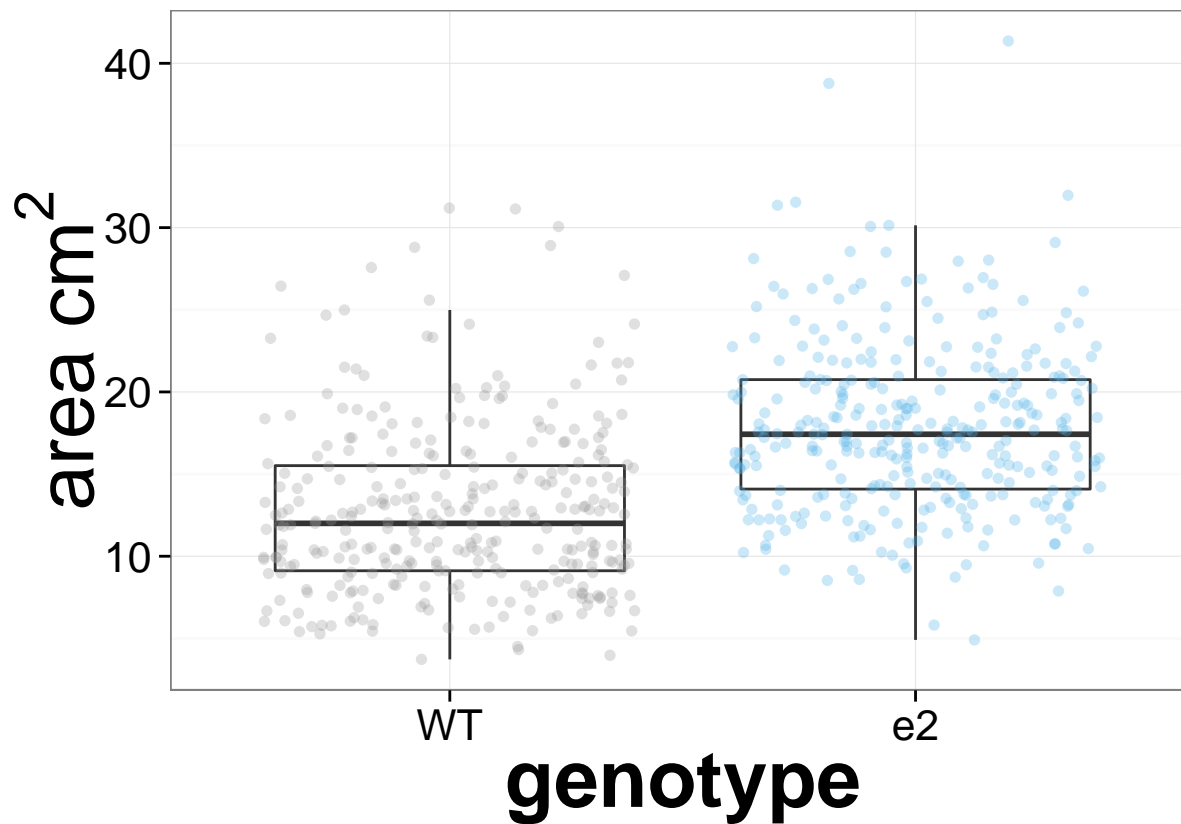
```
L1L2 <- subset(leafArea4, Leaf == "L1" | Leaf == "L2")
```

```
L1L2$genotype <- as.factor(L1L2$genotype)
summary(L1L2$genotype)
```

```
## e2 WT
## 316 315
```

```
ggplot(L1L2, aes(genotype, Area)) +
  geom_boxplot(outlier.shape = NA) +
  geom_point(alpha = .3, position = "jitter", aes(color = genotype)) +
  scale_x_discrete(limits=c("WT", "e2")) +
  scale_color_manual(values=cbPalette) +
  theme_bw() +
  ylab(expression(paste("area ", "cm", "²", sep=""))) +
  customTheme
```

```
## Warning: Removed 5 rows containing missing values (geom_point).
## Warning: Removed 9 rows containing missing values (geom_point).
```



```
t.test(L1L2$Area ~ L1L2$genotype)
```

```
##
##  Welch Two Sample t-test
##
## data:  L1L2$Area by L1L2$genotype
## t = 12.18, df = 628.9, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  4.241 5.872
## sample estimates:
## mean in group e2 mean in group WT
##           17.82           12.76
```

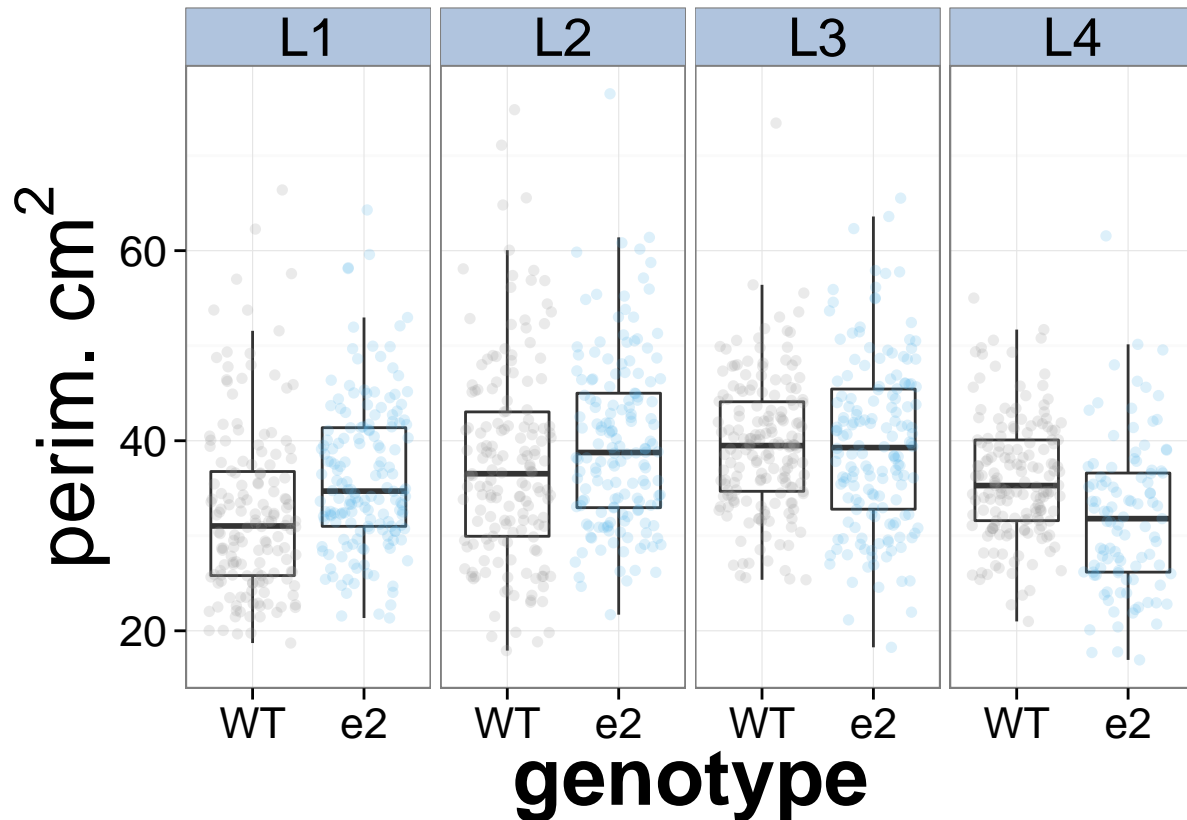
I want to make a table of the Area, Perm., AR, Round, and Solidity for the paper.

```
head(leafArea4)
```

| | Leaf | ID | Area | Perim. | Circ. | AR | Round | Solidity | genotype |
|-------|------|----------|-------|--------|-------|-------|-------|----------|----------|
| ## 2 | L1 | EA10.bmp | 13.05 | 34.75 | 0.136 | 1.248 | 0.802 | 0.861 | e2 |
| ## 3 | L1 | EA11.bmp | 11.40 | 27.37 | 0.191 | 1.099 | 0.910 | 0.864 | e2 |
| ## 5 | L1 | EA12.bmp | 10.76 | 31.70 | 0.135 | 1.213 | 0.824 | 0.845 | e2 |
| ## 7 | L1 | EA13.bmp | 13.04 | 25.42 | 0.254 | 1.213 | 0.824 | 0.935 | e2 |
| ## 10 | L1 | EA14.bmp | 15.66 | 43.12 | 0.106 | 1.438 | 0.695 | 0.815 | e2 |
| ## 11 | L1 | EA15.bmp | 12.45 | 37.22 | 0.113 | 1.280 | 0.782 | 0.846 | e2 |

```
ggplot(leafArea4, aes(genotype, Perim.)) +
  geom_boxplot(outlier.shape = NA) +
  geom_point(alpha = .2, position = "jitter", aes(color = genotype)) +
  scale_x_discrete(limits=c("WT", "e2")) +
  scale_color_manual(values=cbPalette) +
  facet_grid(.~Leaf) +
  theme_bw() +
  ylab(expression(paste("perim. ", "cm2", sep=""))) +
  customTheme
```

```
## Warning: Removed 4 rows containing missing values (geom_point).
## Warning: Removed 6 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 4 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
```



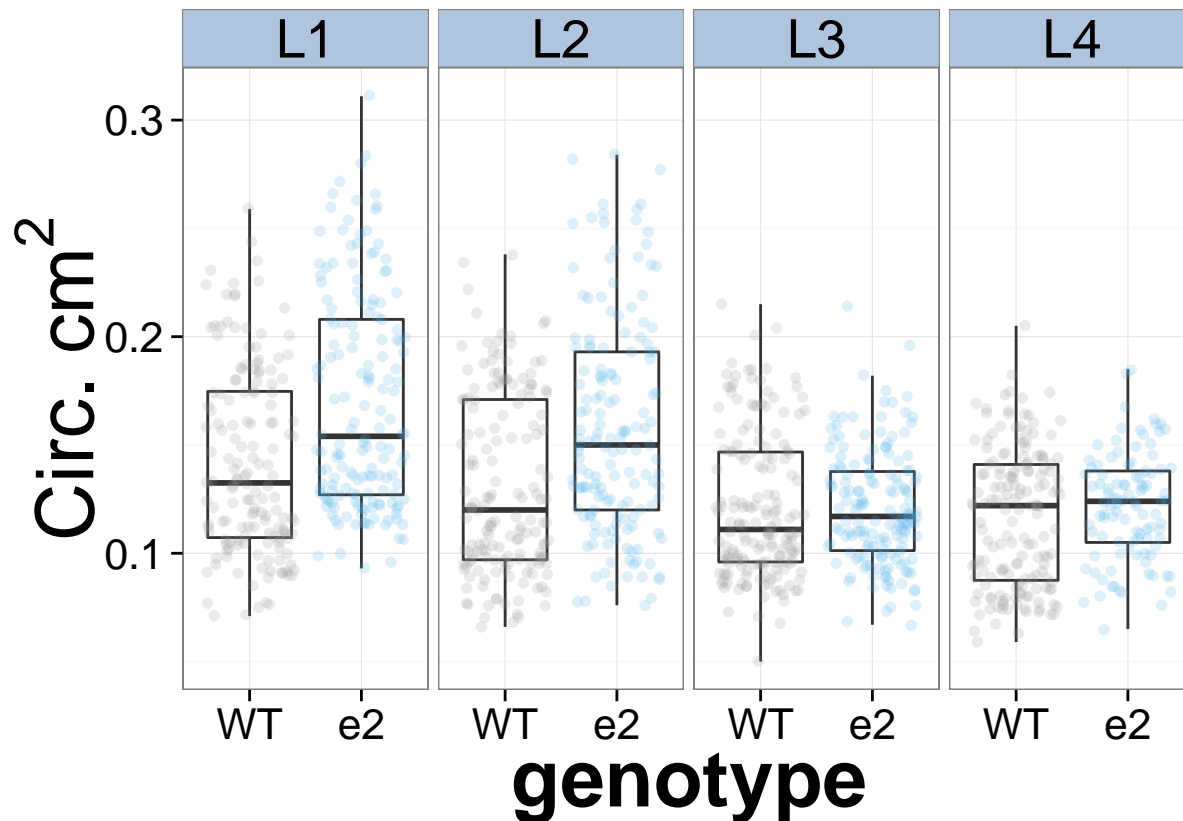
```
t.test(L1L2$Perim. ~ L1L2$genotype)
```

```
##
## Welch Two Sample t-test
##
## data: L1L2$Perim. by L1L2$genotype
## t = 3.762, df = 612.5, p-value = 0.0001848
```

```
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.353 4.310
## sample estimates:
## mean in group e2 mean in group WT
##           37.94           35.11
```

```
ggplot(leafArea4, aes(genotype, Circ.)) +
  geom_boxplot(outlier.shape = NA) +
  geom_point(alpha = .2, position = "jitter", aes(color = genotype)) +
  scale_x_discrete(limits=c("WT", "e2")) +
  scale_color_manual(values=cbPalette) +
  facet_grid(.~Leaf) +
  theme_bw() +
  ylab(expression(paste("Circ. ", "cm", "m^2", sep=""))) +
  customTheme
```

```
## Warning: Removed 2 rows containing missing values (geom_point).
```



```
t.test(L1L2$Circ. ~ L1L2$genotype)
```

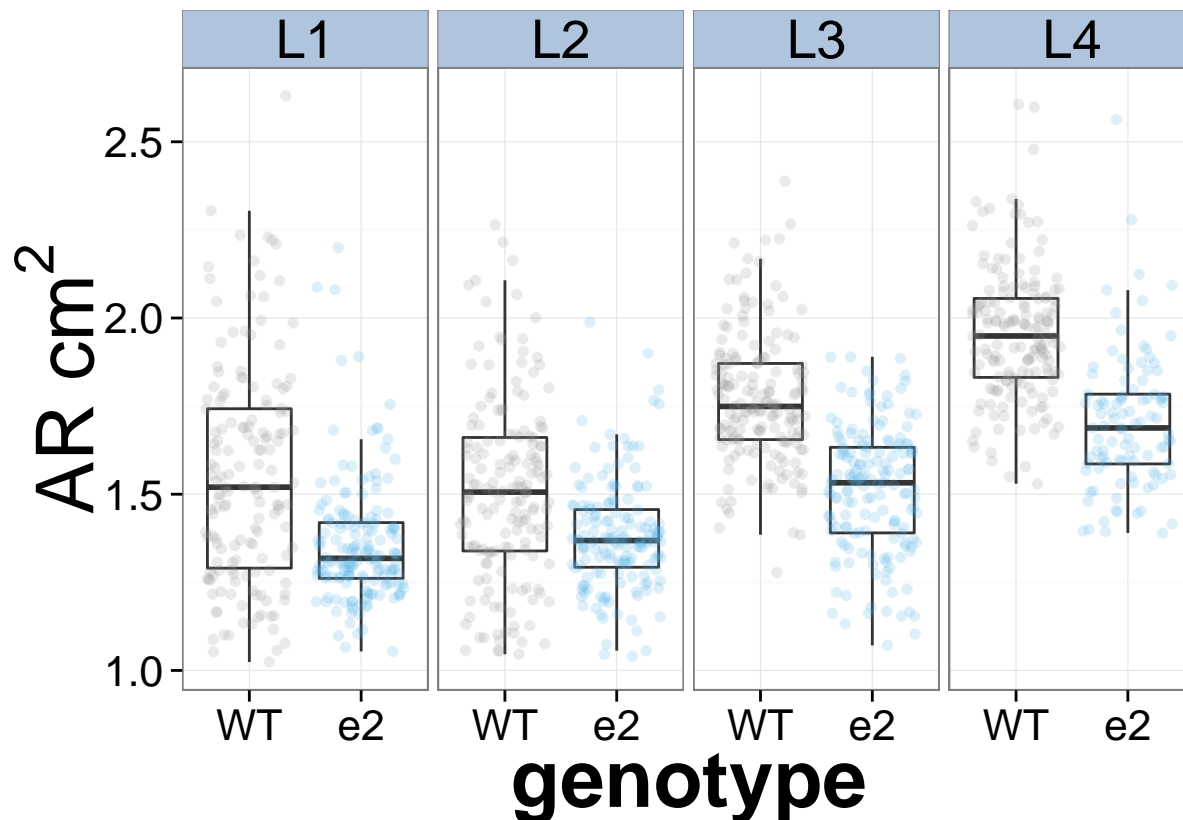
```
##
## Welch Two Sample t-test
##
## data:  L1L2$Circ. by L1L2$genotype
## t = 7.492, df = 612.4, p-value = 2.391e-13
```

```
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 0.02052 0.03510
## sample estimates:
## mean in group e2 mean in group WT
## 0.1641 0.1363
```

Aspect Ratio

```
ggplot(leafArea4, aes(genotype, AR)) +
  geom_boxplot(outlier.shape = NA) +
  geom_point(alpha = .2, position = "jitter", aes(color = genotype)) +
  scale_x_discrete(limits=c("WT", "e2")) +
  scale_color_manual(values=cbPalette) +
  facet_grid(.~Leaf) +
  theme_bw() +
  ylab(expression(paste("AR ", "cm", "m^2", sep=""))) +
  customTheme
```

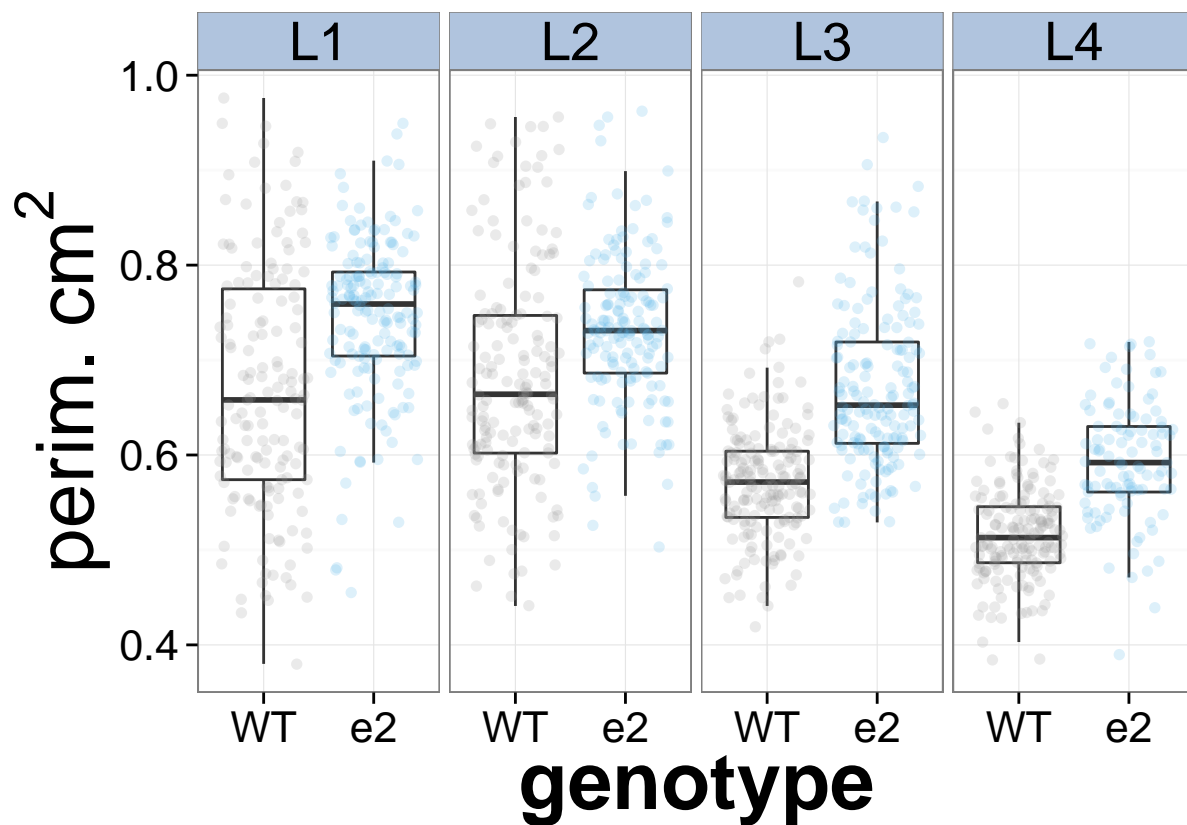
```
## Warning: Removed 10 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 8 rows containing missing values (geom_point).
## Warning: Removed 3 rows containing missing values (geom_point).
## Warning: Removed 6 rows containing missing values (geom_point).
## Warning: Removed 4 rows containing missing values (geom_point).
## Warning: Removed 3 rows containing missing values (geom_point).
```



Round

```
ggplot(leafArea4, aes(genotype, Round)) +  
  geom_boxplot(outlier.shape = NA) +  
  geom_point(alpha = .2, position = "jitter", aes(color = genotype)) +  
  scale_x_discrete(limits=c("WT", "e2")) +  
  scale_color_manual(values=cbPalette) +  
  facet_grid(.~Leaf) +  
  theme_bw() +  
  ylab(expression(paste("perim. ", "cm", "m^2", sep=""))) +  
  customTheme
```

```
## Warning: Removed 8 rows containing missing values (geom_point).  
## Warning: Removed 6 rows containing missing values (geom_point).  
## Warning: Removed 3 rows containing missing values (geom_point).  
## Warning: Removed 5 rows containing missing values (geom_point).  
## Warning: Removed 2 rows containing missing values (geom_point).  
## Warning: Removed 4 rows containing missing values (geom_point).
```



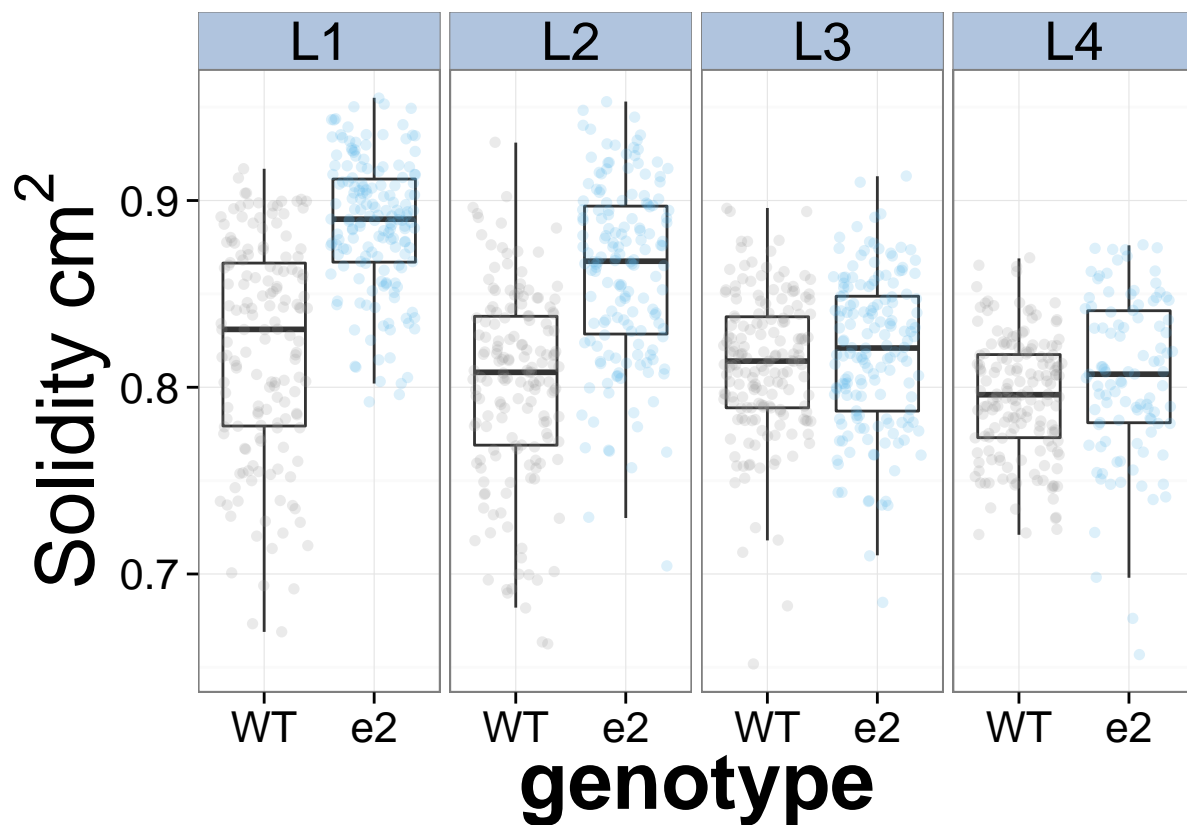
Solidity

```
ggplot(leafArea4, aes(genotype, Solidity)) +  
  geom_boxplot(outlier.shape = NA) +
```



```
geom_point(alpha = .2, position = "jitter", aes(color = genotype)) +
scale_x_discrete(limits=c("WT", "e2")) +
scale_color_manual(values=cbPalette) +
facet_grid(.~Leaf) +
theme_bw() +
ylab(expression(paste("Solidity ", "c", m^2, sep=""))) +
customTheme
```

```
## Warning: Removed 2 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 2 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 3 rows containing missing values (geom_point).
## Warning: Removed 2 rows containing missing values (geom_point).
```



From looking at the data I think the best way to proceed is to combine only L1 and L2, since they are more “fully expanded”, then make one large graph that summarizes all the values, by melting

Making the leaf characteristics table.

```
melted <- melt(leafArea4, id.vars = c("Leaf", "ID", "genotype"))

# by leaf number
summarise(group_by(melted, genotype, Leaf, variable),
           mean=mean(value), sd=sd(value))
```

```
## Source: local data frame [48 x 5]
## Groups: genotype, Leaf
##
##   genotype Leaf variable   mean    sd
## 1      WT   L1     Area 11.5752 5.02904
## 2      WT   L1  Perim. 32.5106 9.07932
## 3      WT   L1   Circ.  0.1418 0.04306
## 4      WT   L1     AR   1.5534 0.31143
## 5      WT   L1   Round  0.6689 0.12981
## 6      WT   L1 Solidity  0.8206 0.05777
## 7      WT   L2     Area 13.9003 5.18201
## 8      WT   L2  Perim. 37.5973 10.58549
## 9      WT   L2   Circ.  0.1310 0.04148
## 10     WT   L2     AR   1.5098 0.26108
## 11     WT   L2   Round  0.6823 0.11850
## 12     WT   L2 Solidity  0.8015 0.05163
## 13     WT   L3     Area 14.5384 3.64430
## 14     WT   L3  Perim. 39.5190 7.11193
## 15     WT   L3   Circ.  0.1213 0.03374
## 16     WT   L3     AR   1.7709 0.18574
## 17     WT   L3   Round  0.5707 0.05917
## 18     WT   L3 Solidity  0.8121 0.03783
## 19     WT   L4     Area 11.4935 2.59249
## 20     WT   L4  Perim. 35.9941 6.34834
## 21     WT   L4   Circ.  0.1164 0.03182
## 22     WT   L4     AR   1.9545 0.18912
## 23     WT   L4   Round  0.5163 0.04901
## 24     WT   L4 Solidity  0.7946 0.03242
## 25     e2   L1     Area 17.0035 5.18177
## 26     e2   L1  Perim. 36.2062 7.88959
## 27     e2   L1   Circ.  0.1698 0.04985
## 28     e2   L1     AR   1.3604 0.17844
## 29     e2   L1   Round  0.7459 0.08391
## 30     e2   L1 Solidity  0.8872 0.03555
## 31     e2   L2     Area 18.6795 5.09659
## 32     e2   L2  Perim. 39.7684 9.07122
## 33     e2   L2   Circ.  0.1581 0.05042
## 34     e2   L2     AR   1.3813 0.15674
## 35     e2   L2   Round  0.7329 0.08029
## 36     e2   L2 Solidity  0.8620 0.04756
## 37     e2   L3     Area 14.8799 5.03542
## 38     e2   L3  Perim. 39.5454 8.99519
## 39     e2   L3   Circ.  0.1209 0.02654
## 40     e2   L3     AR   1.5118 0.18039
## 41     e2   L3   Round  0.6714 0.08508
## 42     e2   L3 Solidity  0.8187 0.04084
## 43     e2   L4     Area 10.0180 4.24383
## 44     e2   L4  Perim. 32.0043 7.97196
## 45     e2   L4   Circ.  0.1215 0.02468
## 46     e2   L4     AR   1.7043 0.19430
## 47     e2   L4   Round  0.5938 0.06287
## 48     e2   L4 Solidity  0.8052 0.04379
```

```
# ignoring leaf number
summaryLeaf <- summarise(group_by(melted, genotype, variable),
  mean=mean(value), sd=sd(value))

castSumLeaf.sd <- dcast(summaryLeaf, genotype ~ variable, value.var = "sd")
castSumLeaf.mean <- dcast(summaryLeaf, genotype ~ variable, value.var = "mean")
```