

# Code to produce figures for MVD poster

This document last updated on Wed Aug 2 10:17:14 2017

```
library(dplyr); library(reshape2); library(stringr); library(tidyr); library(ggplot2); library(cowplot)

##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:stats':
##
##   filter, lag
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
##
## Attaching package: 'tidyr'
##
## The following object is masked from 'package:reshape2':
##
##   smiths
##
## Attaching package: 'cowplot'
##
## The following object is masked from 'package:ggplot2':
##
##   ggsave
##
# Import measurements made in field
field_measurements <- read.csv("~/Dropbox/2017-traits/2017-trait-measurements.csv", stringsAsFactors = F)
tbl_df(field_measurements)

## # A tibble: 3,003 x 17
##   sequence plot species plant_number leaf_number leaf_length_cm
##   <int> <chr>   <chr>         <dbl>         <int>         <dbl>
## 1         1   748    chgl             1             1             3.3
## 2         2   748    chgl             1             2             3.5
## 3         3   748    chgl             1             3             3.5
## 4         4   748    chgl             2             1             3.5
## 5         5   748    chgl             2             2             3.5
## 6         6   748    chgl             2             3             2.8
## 7         7   748    chgl             3             1             4.4
## 8         8   748    chgl             3             2             4.9
## 9         9   748    chgl             3             3             4.5
## 10        10   748    chgl             4             1             4.5
## # ... with 2,993 more rows, and 11 more variables: leaf_width_cm <dbl>,
## #   wet_mass_g <dbl>, petiole_mass_g <dbl>, dry_mass_g <dbl>,
## #   dry_mass_petiole_g <dbl>, date_collected <chr>, notes <chr>, X <chr>,
## #   X.1 <chr>, X.2 <chr>, X.3 <chr>
##
# Convert a few of the columns into factors
field_measurements <- field_measurements %>% mutate_at(c(2, 3, 4, 5), funs(factor))
# let's uncapitalize all of the species names ... ugh
field_measurements <- field_measurements %>% mutate(species = tolower(species))
```

```

# Let's subset this dataframe so that it only has plants from plots that we care about for the current
# Let's also subset it to not include the grasses
field_measurements <- field_measurements %>% filter(plot %in% c("740", "741", "742", "746", "749", "753")
  filter(!(species %in% c("brma", "vuma", "homu"))))

# Let's set aside the issue of petioles for now.
field_measurements <- field_measurements %>% mutate(wet_mass_g = rowSums(cbind(wet_mass_g, petiole_mass_g)),
  mutate(dry_mass_g = rowSums(cbind(dry_mass_g, dry_mass_petiole_g), na.rm = T)) %>% select(-dry_mass_p

```

Now let's work on importing and cleaning the leaf area dataset

```

# Import leaf areas
leaf_areas <- read.csv("~/Dropbox/2017-traits/leaf_areas.csv", stringsAsFactors = F)
tbl_df(leaf_areas)

```

```

## # A tibble: 1,023 x 7
##       Name Count Total.Area Average.Size X.Area Mean      X
##       <chr> <int>      <dbl>      <dbl> <dbl> <dbl> <chr>
## 1 744-pler-p1-l1      1      1.282      1.282  1.815  255
## 2 744-pler-p1-l2      1      1.465      1.465  2.074  255
## 3 744-pler-p1-l3      1      0.974      0.974  1.380  255
## 4 744-pler-p2-l1      1      1.471      1.471  1.933  255
## 5 744-pler-p2-l2      1      1.179      1.179  1.548  255
## 6 744-pler-p2-l3      1      1.112      1.112  1.460  255
## 7 744-pler-p3-l1      2      1.128      0.564  1.420  255
## 8 744-pler-p3-l2      3      0.949      0.316  1.194  255
## 9 744-pler-p3-l3      2      0.921      0.460  1.159  255
## 10 744-pler-p4-l1      2      1.151      0.575  2.031  255
## # ... with 1,013 more rows

```

*# Let's get rid of the cases where petiole area was done separately- for now, let's do all together.*

```

leaf_areas <- leaf_areas %>% group_by(Name) %>% mutate(Total.Area = sum(Total.Area)) %>%
  filter(!(X == "petiole")) %>% select(-X)

```

*# Evidently there are some non-plot species in this dataframe:*

```

leaf_areas %>% filter(grepl("lebo", Name))

```

```

## # A tibble: 24 x 6
## # Groups:   Name [24]
##       Name Count Total.Area Average.Size X.Area Mean
##       <chr> <int>      <dbl>      <dbl> <dbl> <dbl>
## 1 lebo-p1-l1      1      2.209      2.209  7.209  255
## 2 lebo-p1-l2      1      2.617      2.617  8.539  255
## 3 lebo-p1-l3      1      1.024      1.024  3.342  255
## 4 lebo-p2-l1      1      1.466      1.466  5.485  255
## 5 lebo-p2-l2      1      1.402      1.402  5.247  255
## 6 lebo-p2-l3      1      0.914      0.914  3.421  255
## 7 lebo-p3-l1      1      1.672      1.672  2.795  255
## 8 lebo-p3-l2      1      1.446      1.446  2.417  255
## 9 lebo-p3-l3      1      1.204      1.204  2.013  255
## 10 lebo-p4-l1      1      2.019      2.019  2.844  255
## # ... with 14 more rows

```

```
# Let's remove those columns for now
```

```
leaf_areas <- leaf_areas %>% filter(!(grepl("lebo", Name)))
```

We need to split up the “Name” column of `leaf_areas` before we merge it with the field measurements document. While we’re at it, let’s also make sure that we only have the plots and the species we want (i.e. no grasses, only non-hummock candi valley plots)

```
leaf_areas <- separate(leaf_areas, Name, sep = "-", into = c("plot", "species", "plant_number", "leaf_n  
leaf_areas <- leaf_areas %>%  
  filter(plot %in% c("740", "741", "742", "746", "749", "753", "754", "comp", "cage_1")) %>%  
  filter(!(species %in% c("brma", "vuma", "homu")))
```

Now we need to remove some columns (e.g. we don’t need `count`, `average.size`, `x.area`) and rename others (`Total.Area` should just be `leaf_area_cm2`). We also need to convert some columns that are currently treated as characters into being treated as factors (eg. `plot`, `species`, `plant_number`, `leaf_number`). Finally, we need to remove the letter “p” from the `plant_number` column and the letter “l” from the `leaf_number` column. Finally, there are a few cases where the original `Name` column contained `.jpeg`; let’s get rid of that as well.

```
leaf_areas <- leaf_areas %>%  
  select(plot:leaf_number, leaf_area_cm2 = Total.Area) # select the columns we want  
leaf_areas <- leaf_areas %>%  
  mutate(plant_number = str_replace(plant_number, "p", "")) %>% # remove the p  
  mutate(leaf_number = str_replace(leaf_number, "l", "")) %>% # remove the l  
  mutate(leaf_number = str_replace(leaf_number, ".jpeg", "")) %>%  
  mutate(leaf_number = str_replace(leaf_number, ".jpg", "")) %>% # remove the .jpeg  
  mutate_at(c(1:4), funs(factor))  
leaf_areas %>% filter(plot %in% c("740", "741", "742", "746", "749", "753", "754", "comp", "cage_1")) %>%  
  filter(!(species %in% c("brma", "vuma", "homu")))
```

```
## # A tibble: 663 x 5  
##   plot species plant_number leaf_number leaf_area_cm2  
##   <fctr> <fctr>      <fctr>      <fctr>      <dbl>  
## 1    740   pler          1          1        1.148  
## 2    740   pler          1          2        1.486  
## 3    740   pler          1          3        1.205  
## 4    740   pler          2          1        0.564  
## 5    740   pler          2          2        0.592  
## 6    740   pler          2          3        0.386  
## 7    741  eupe          1          1        0.570  
## 8    741  eupe          1          2        0.812  
## 9    741  eupe          1          3        0.754  
## 10   741  eupe          2          1        1.105  
## # ... with 653 more rows
```

Now, we are ready to merge the `leaf_areas` dataframe with the `field_measurements` dataframe:

```
merged_measurements <- (left_join(field_measurements, leaf_areas, by = c("plot", "species", "plant_number"))
```

```
## Warning: Column `plot` joining factors with different levels, coercing to  
## character vector
```

```
## Warning: Column `species` joining character vector and factor, coercing  
## into character vector
```

```
## Warning: Column `plant_number` joining factors with different levels,
```

```
## coercing to character vector
```

```
missing_areas <- merged_measurements %>% filter(is.na(leaf_area_cm2))
```

```
missing_mass <- merged_measurements %>% filter(is.na(dry_mass_g))
```

Finally, there are a few species for which we did not get both L and C (namely, chgl, clpu, clbo - we only got Comp plot traits for these :( )

```
merged_measurements <- merged_measurements %>% filter(!(species %in% c("chgl", "clpu", "clbo")))
```

## Calculating traits

Finally, we can calculate some traits from these numbers:

```
merged_measurements <- merged_measurements %>%
```

```
  mutate(ldmc_mg_g = (dry_mass_g*1000)/wet_mass_g) %>%
```

```
  mutate(sla_cm2_g = leaf_area_cm2/dry_mass_g)
```

```
merged_measurements <- merged_measurements %>% mutate(plot_type = ifelse(plot == "comp", "comp", "lambda"))
```

```
merged_measurements <- merged_measurements %>% select(19, 2:9, 16:18, 10:15)
```

## Averaging to the individual level

Our dataframe currently has 1 row per leaf; we would like to do all our analyses at the individual level.

```
individual_traits <- merged_measurements %>% group_by(plot_type, plot, species, plant_number) %>% summarise()
```

```
# Let's do some exploratory plotting
```

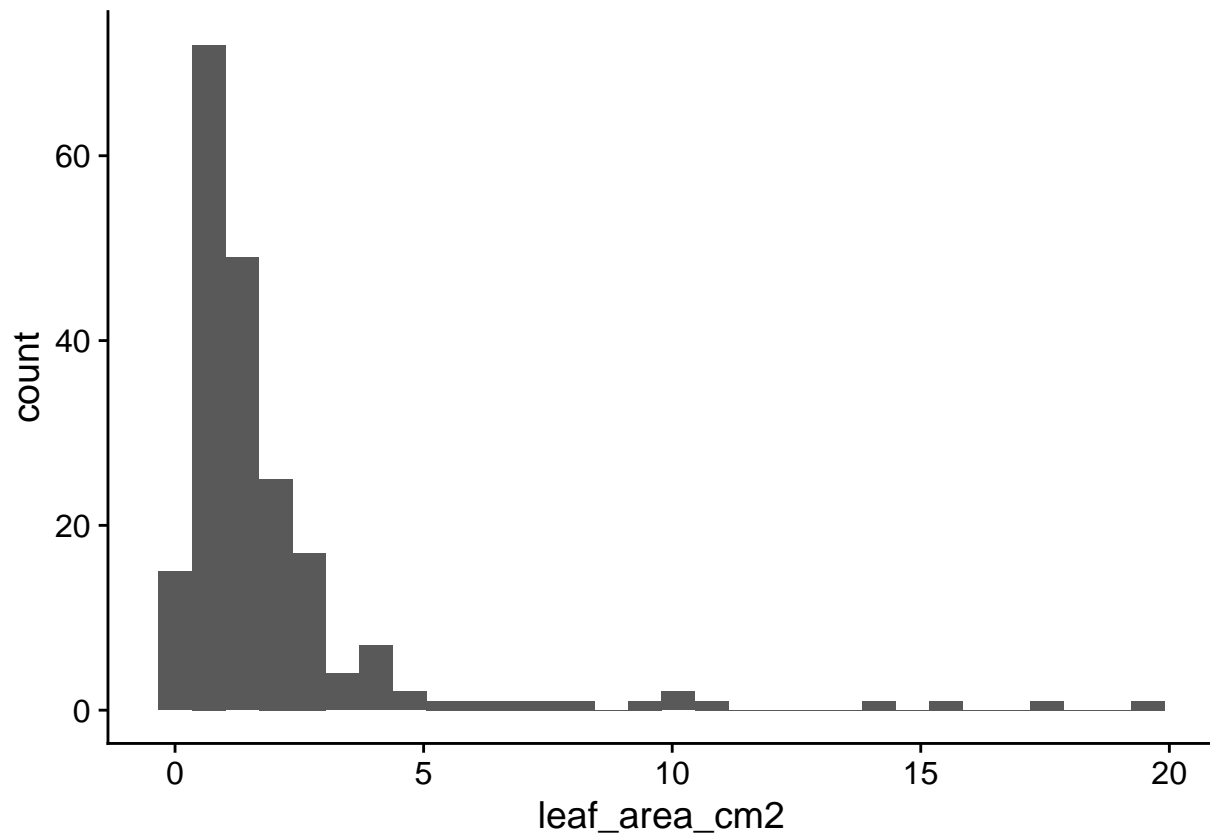
```
# Histogram of leaf areas
```

```
g_ind_traits <- ggplot(data = individual_traits)
```

```
g_ind_traits + geom_histogram(aes(x = leaf_area_cm2))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

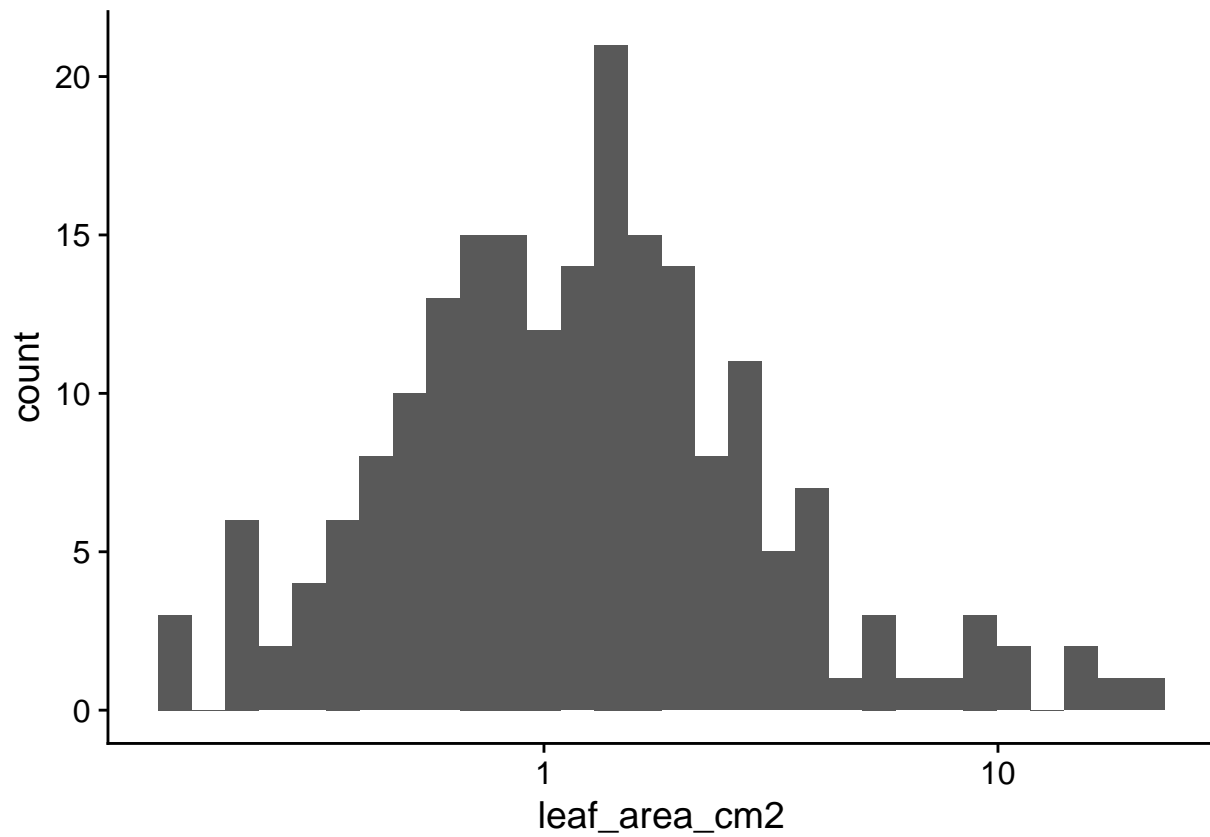
```
## Warning: Removed 1 rows containing non-finite values (stat_bin).
```



```
g_ind_traits + geom_histogram(aes(x = leaf_area_cm2)) + scale_x_log10()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

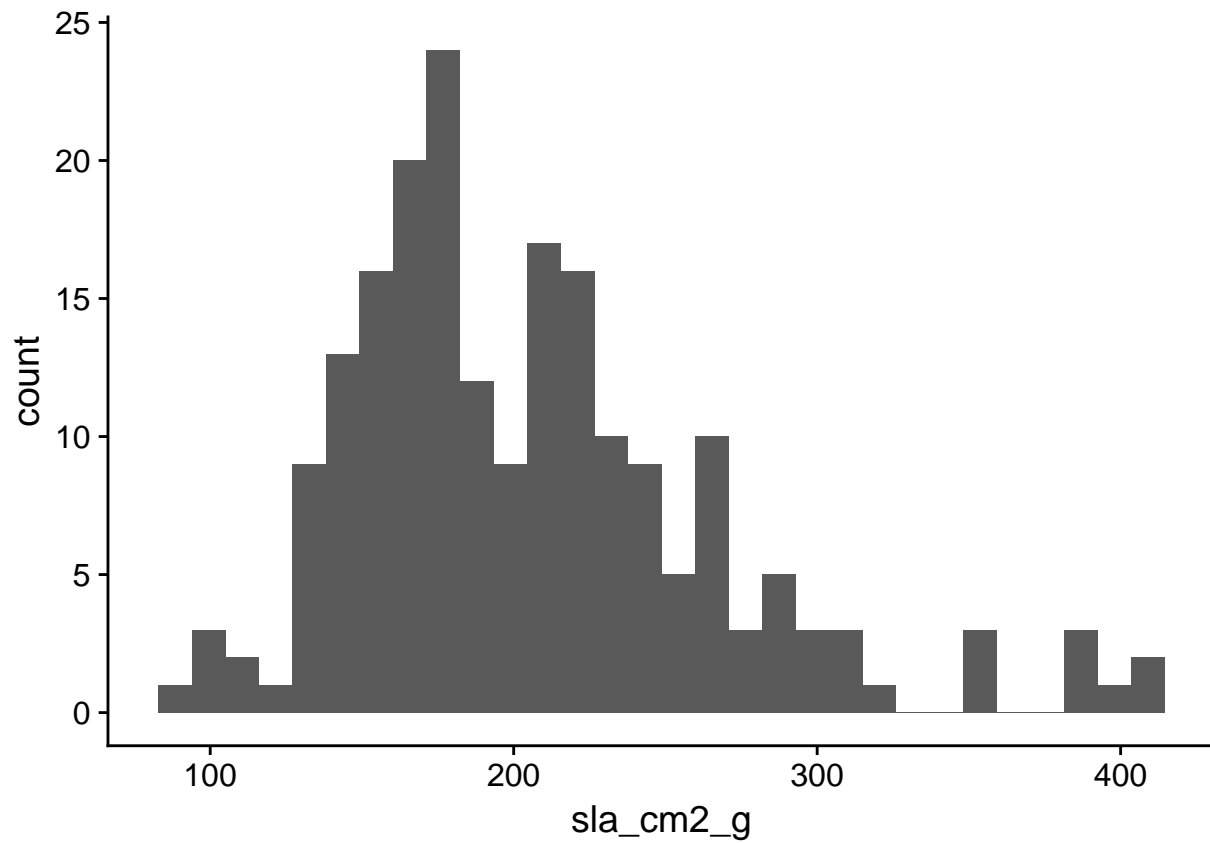
```
## Warning: Removed 1 rows containing non-finite values (stat_bin).
```



```
g_ind_traits + geom_histogram(aes(x = sla_cm2_g))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

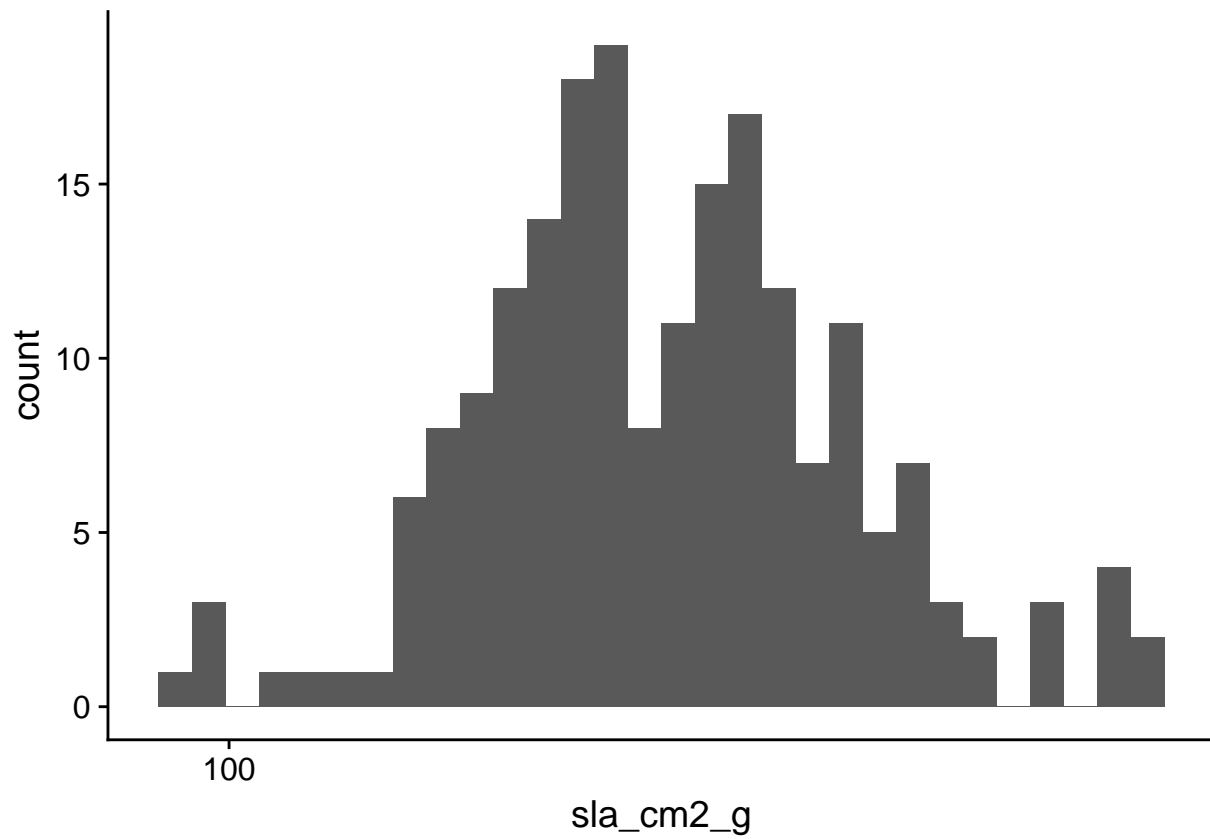
```
## Warning: Removed 4 rows containing non-finite values (stat_bin).
```



```
g_ind_traits + geom_histogram(aes(x = sla_cm2_g)) + scale_x_log10()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

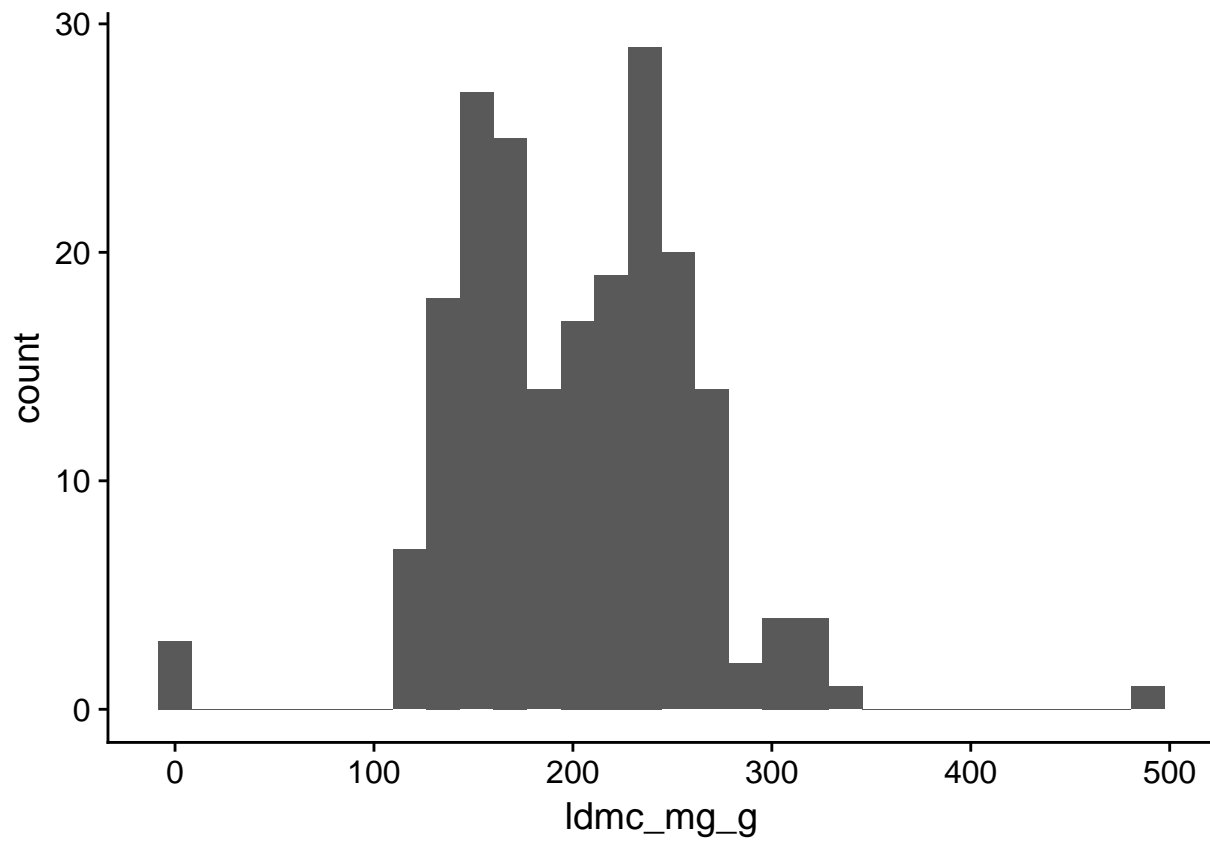
```
## Warning: Removed 4 rows containing non-finite values (stat_bin).
```



```
g_ind_traits + geom_histogram(aes(x = ldmc_mg_g))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



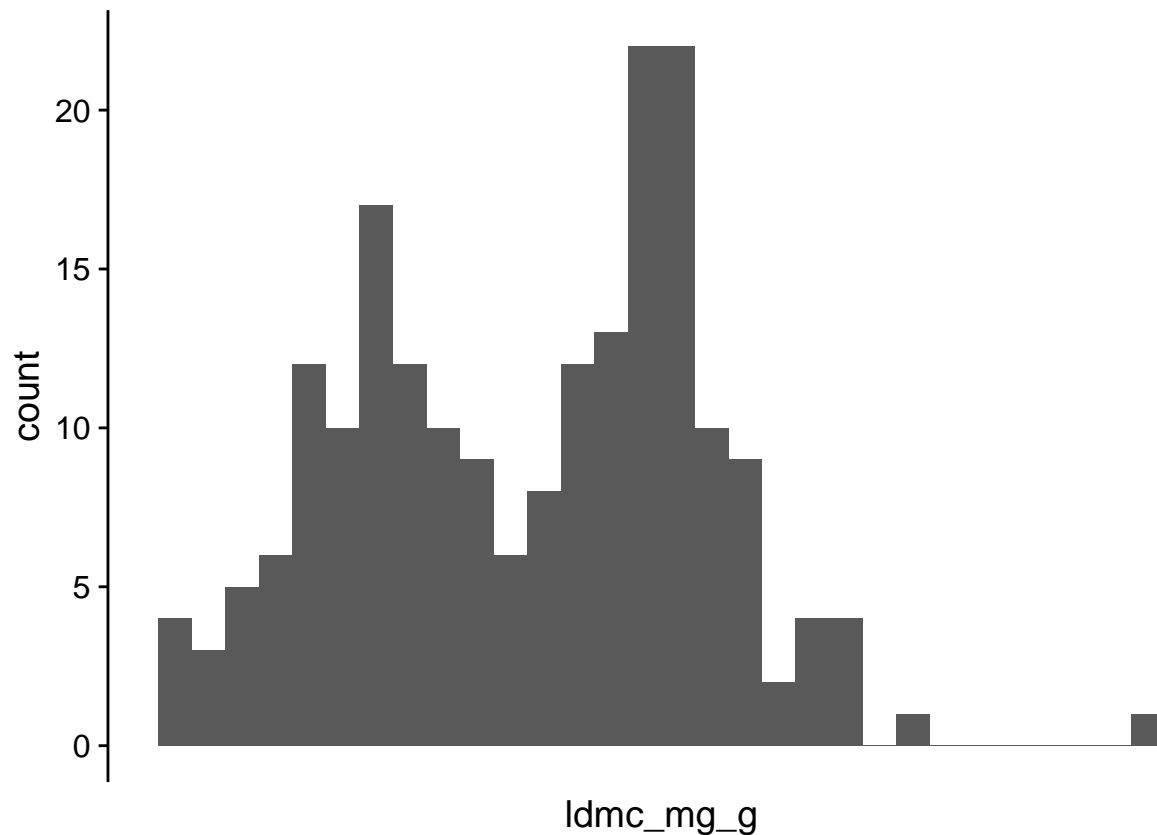


```
g_ind_traits + geom_histogram(aes(x = ldmc_mg_g)) + scale_x_log10()
```

```
## Warning: Transformation introduced infinite values in continuous x-axis
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 3 rows containing non-finite values (stat_bin).
```



```
individual_traits <- individual_traits %>%
  mutate(log_leaf_area_cm2 = log10(leaf_area_cm2)) %>%
  mutate(log_sla_cm2_g = log10(sla_cm2_g)) %>%
  mutate(log_ldmc_mg_g = log10(ldmc_mg_g))
```

## Let's see if 2012 traits fit in here...

The 2012 traits are, unfortunately, in a slightly different format, so we must do some more data munging first... Ugh.

```
comptraits_2012 <- read.csv("~/Dropbox/2017-traits/old-data/tapioca_traits.csv", stringsAsFactors = F)
# First let's select the columns we want
comptraits_2012 <- comptraits_2012 %>% select(species, plant_number = plant, leaf_area_cm2 = aveLeafSize)
# replace the infernal "SACA" in the 2012 df...
comptraits_2012 <- comptraits_2012 %>% mutate(species = ifelse(species == "SACA", "SACO", species))

# and to keep things clean let's remove the species for which we lack 2017 data
comptraits_2012 <- comptraits_2012 %>% filter(!(species %in% c("ANAR", "AGRE", "ERBO", "SIGA", "GECA")))
# and lowercase all the species names
comptraits_2012$species <- tolower(comptraits_2012$species)

# and get rid of rows with NAs...
comptraits_2012 <- comptraits_2012 %>% filter(!(is.na(leaf_area_cm2)))
comptraits_2012 <- comptraits_2012 %>%
  mutate(log_leaf_area_cm2 = log10(leaf_area_cm2)) %>%
  mutate(log_sla_cm2_g = log10(sla_cm2_g)) %>%
```

```
mutate(log_ldmc_mg_g = log10(ldmc_mg_g))

comptraits_2012$year_measured = as.factor(2012)
```

## Compare 2012 competition traits to 2017 competition traits

```
comptraits_2017 <- individual_traits %>% filter(plot_type == "comp") %>% ungroup() %>% select(-c(1:2))
comptraits_2017$year_measured = as.factor(2017)

comptraits <- rbind(comptraits_2012, comptraits_2017)
```

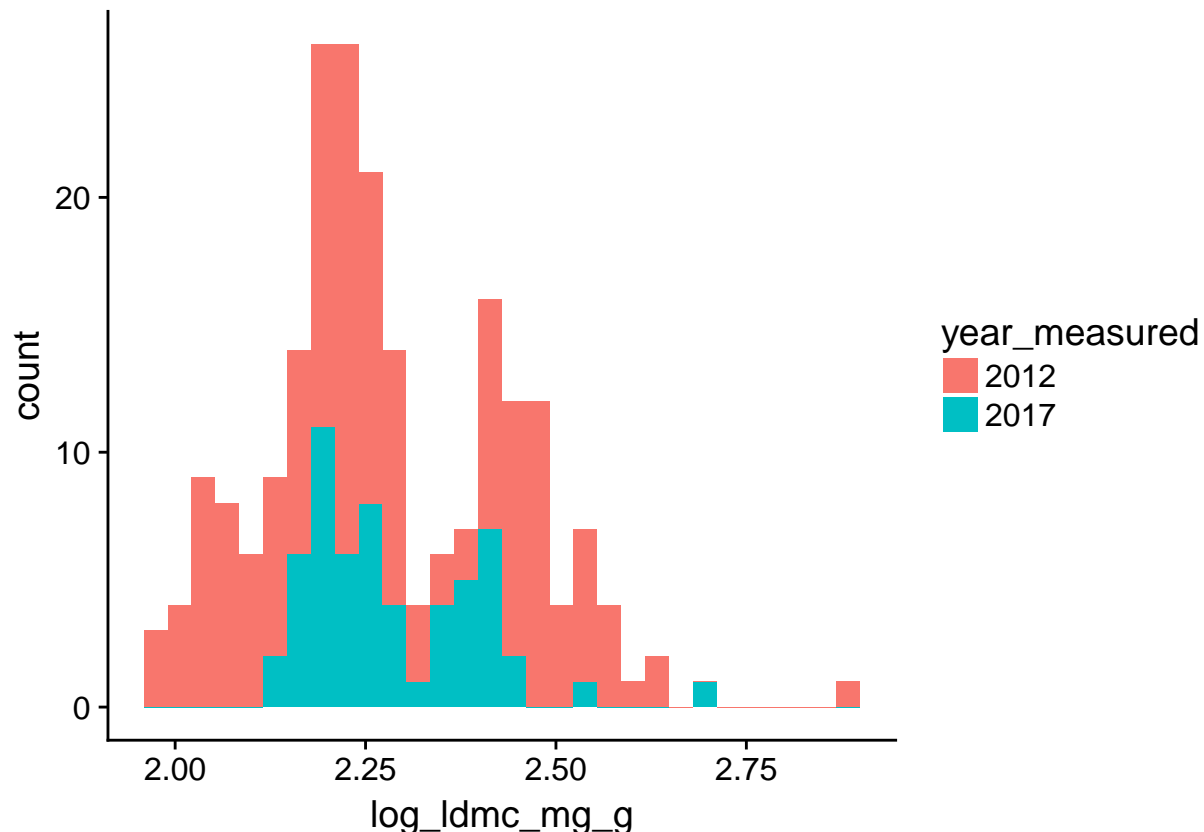
```
# Actually, we first need to remove some species that we can't really compare

comptraits <- comptraits %>% filter(!(species %in% c("mica", "clbo", "clpu", "lopu")))

comp_comparison <- ggplot(comptraits)

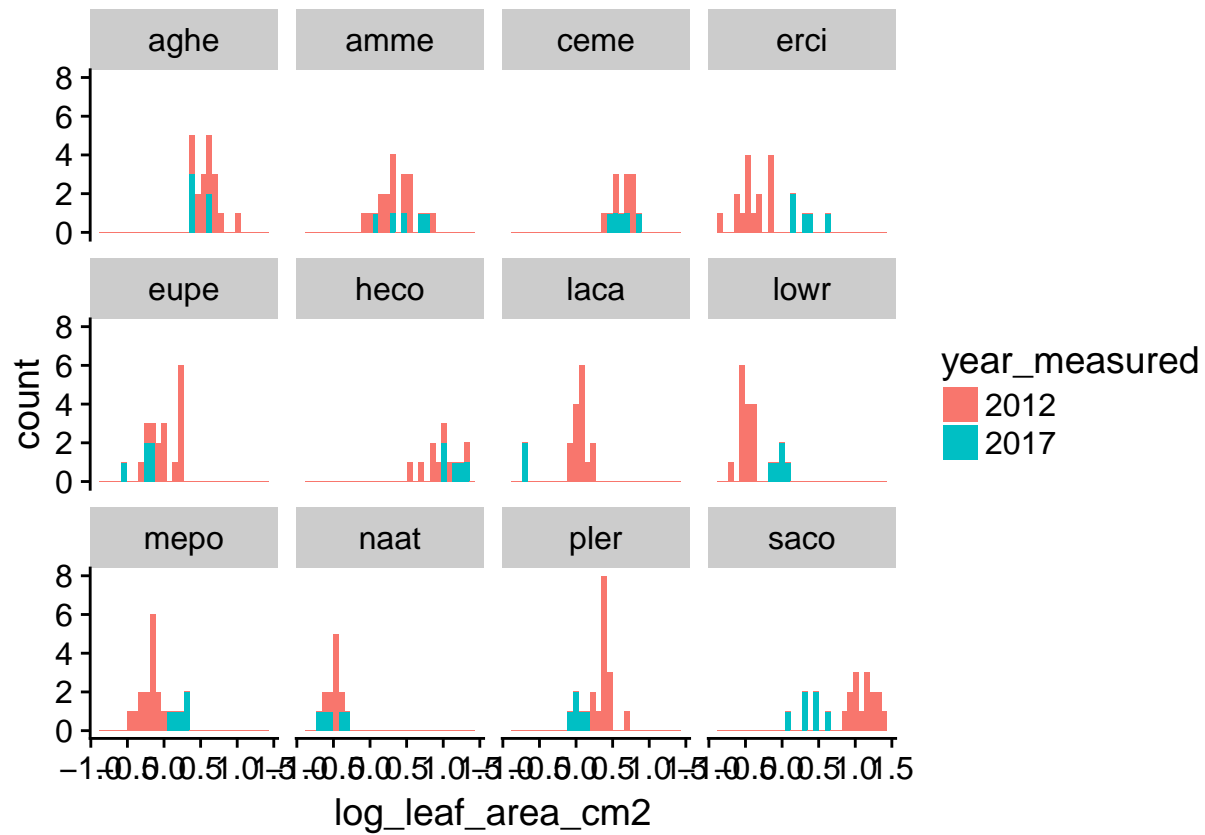
comp_comparison + geom_histogram(aes(x = log_ldmc_mg_g, fill = year_measured))

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



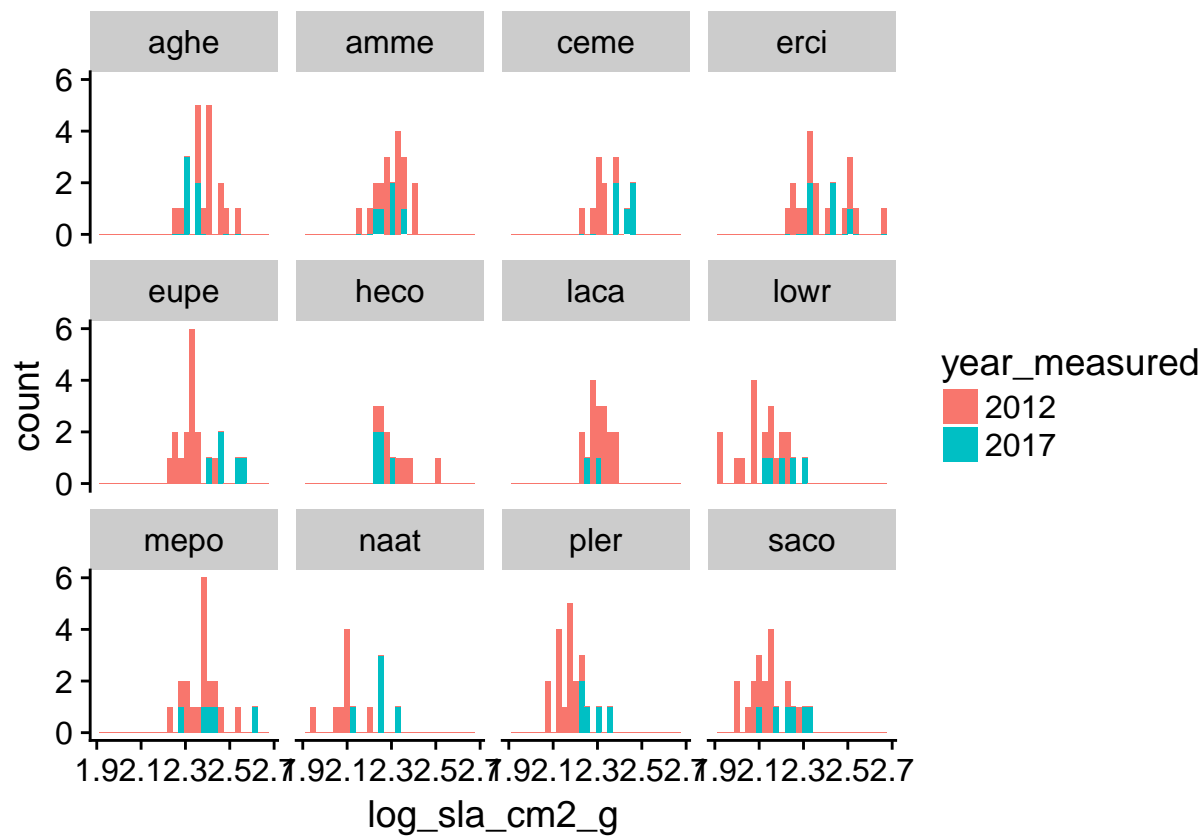
```
comp_comparison + geom_histogram(aes(x = log_leaf_area_cm2, fill = year_measured)) + facet_wrap("species")

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



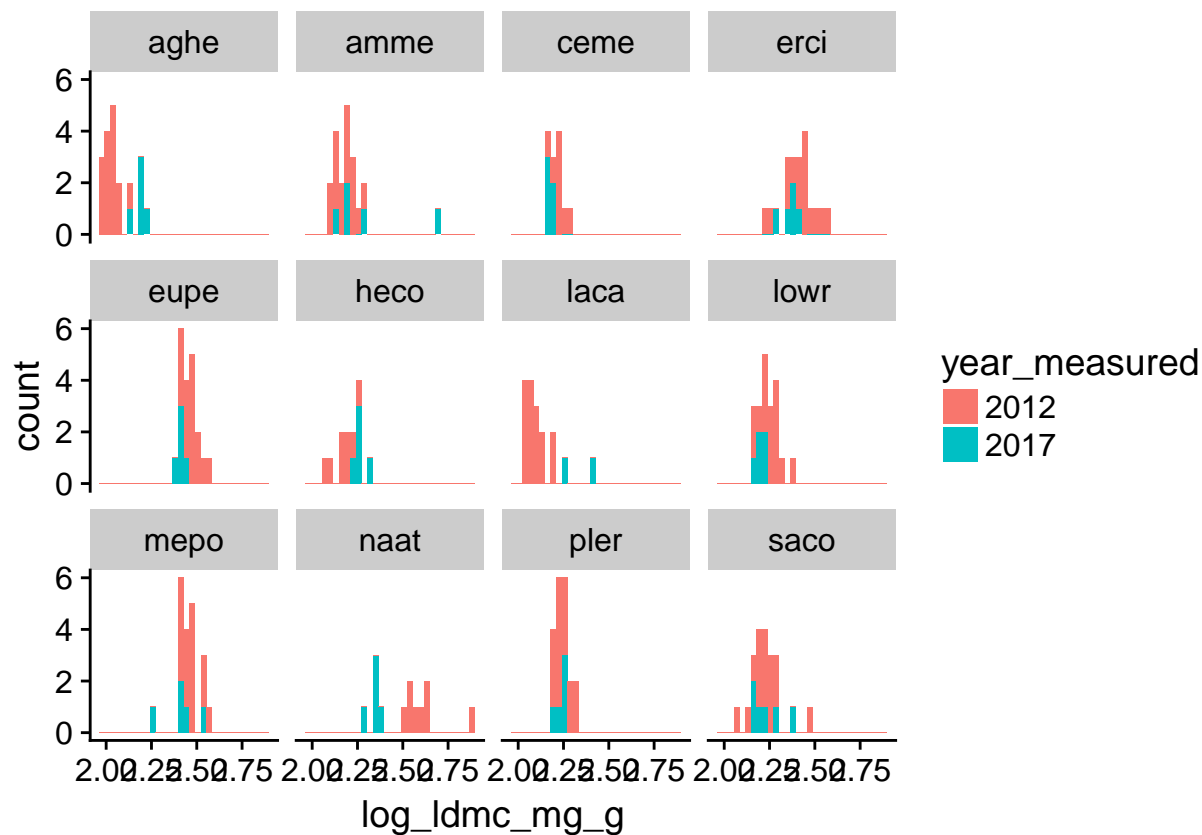
```
comp_comparison + geom_histogram(aes(x = log_sla_cm2_g, fill = year_measured)) + facet_wrap("species")

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
comp_comparison + geom_histogram(aes(x = log_ldmc_mg_g, fill = year_measured)) + facet_wrap("species")

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
summary(aov(data = comptraits, log_leaf_area_cm2~species+year_measured))
```

```
##               Df Sum Sq Mean Sq F value Pr(>F)
## species       11  45.08   4.098   63.781 <2e-16 ***
## year_measured  1   0.01   0.007    0.115  0.734
## Residuals     204  13.11   0.064
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(aov(data = comptraits, log_sla_cm2_g~species+year_measured))
```

```
##               Df Sum Sq Mean Sq F value    Pr(>F)
## species       11 2.1430  0.19482   29.19 < 2e-16 ***
## year_measured  1 0.1591  0.15906   23.83 2.12e-06 ***
## Residuals     204 1.3615  0.00667
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(aov(data = comptraits, log_ldmc_mg_g~species+year_measured))
```

```
##               Df Sum Sq Mean Sq F value Pr(>F)
## species       11  3.870   0.3518  47.062 <2e-16 ***
## year_measured  1  0.000   0.0000   0.005  0.942
## Residuals     204  1.525   0.0075
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

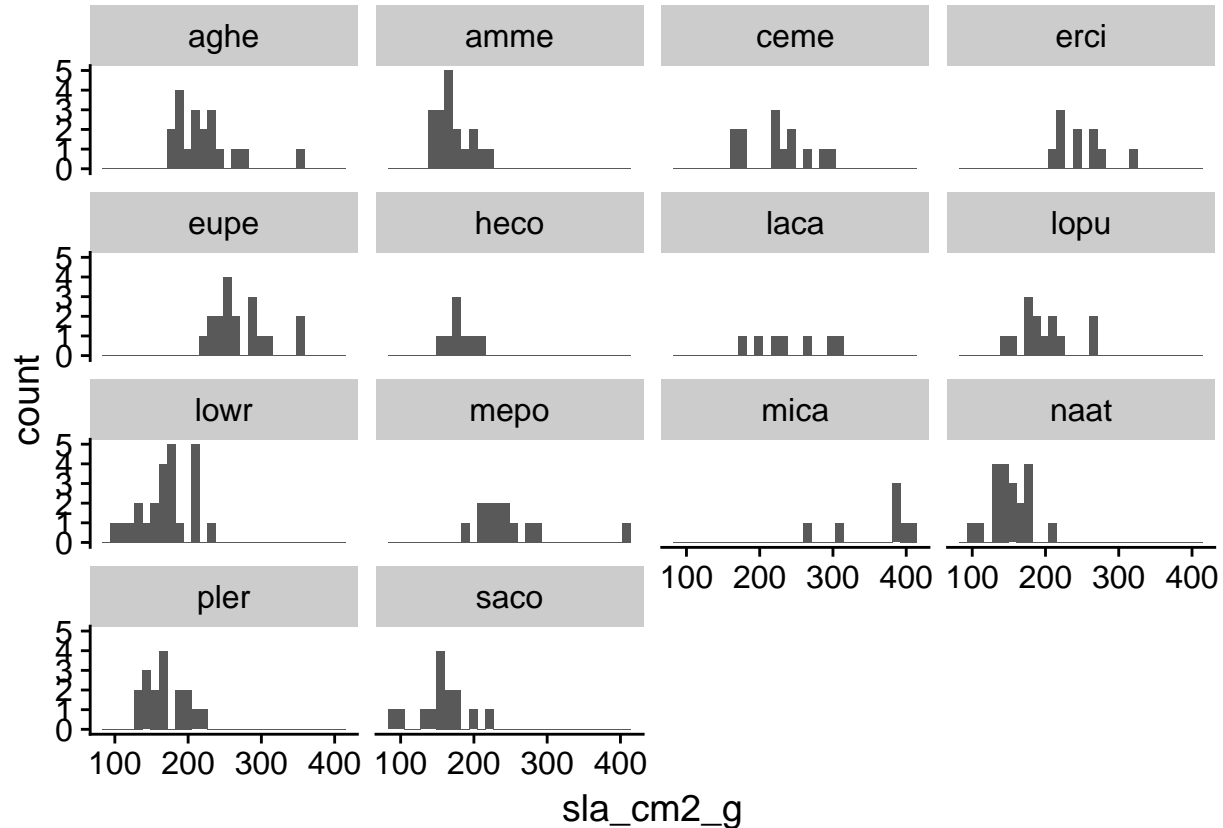
## Let's analyse!

*# first, let's just look at the distribution of the traits*

```
pl <- ggplot(data= individual_traits)
pl + geom_histogram(aes(x = sla_cm2_g)) + facet_wrap("species")
```

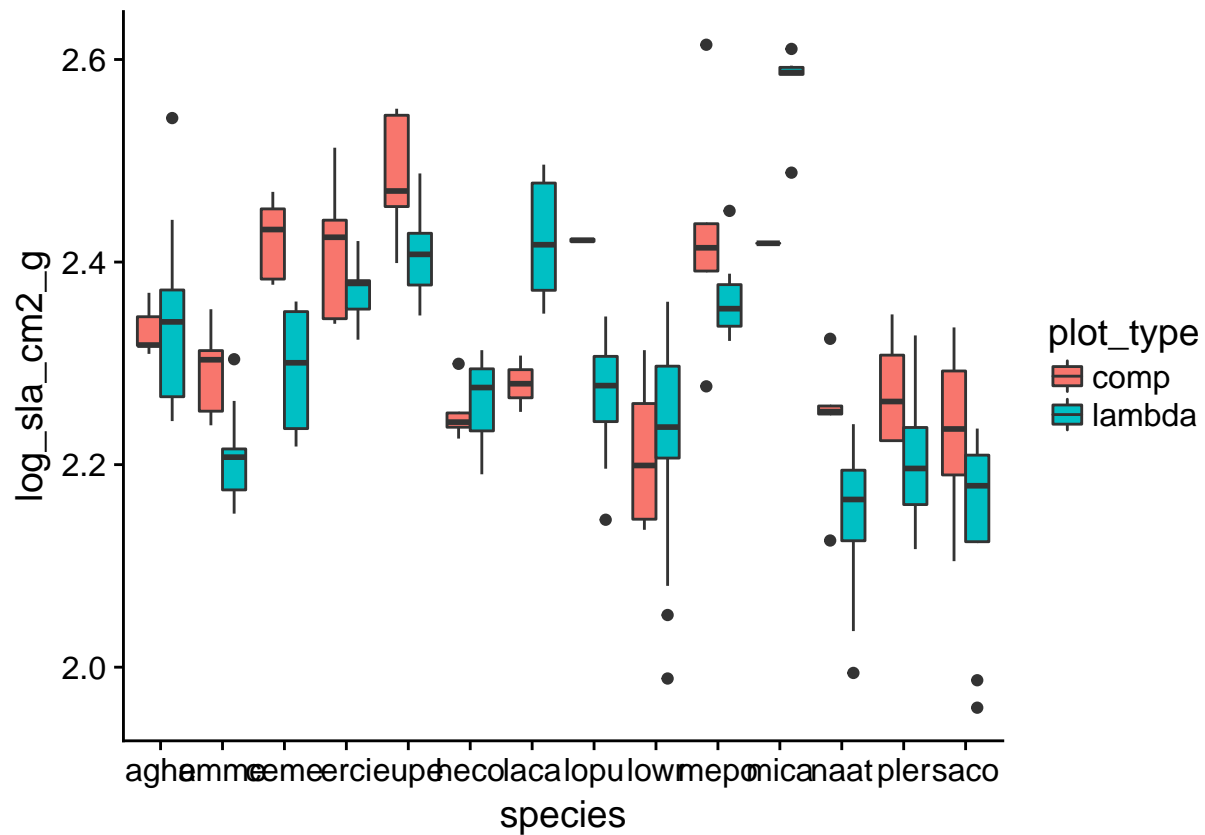
## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

## Warning: Removed 4 rows containing non-finite values (stat\_bin).



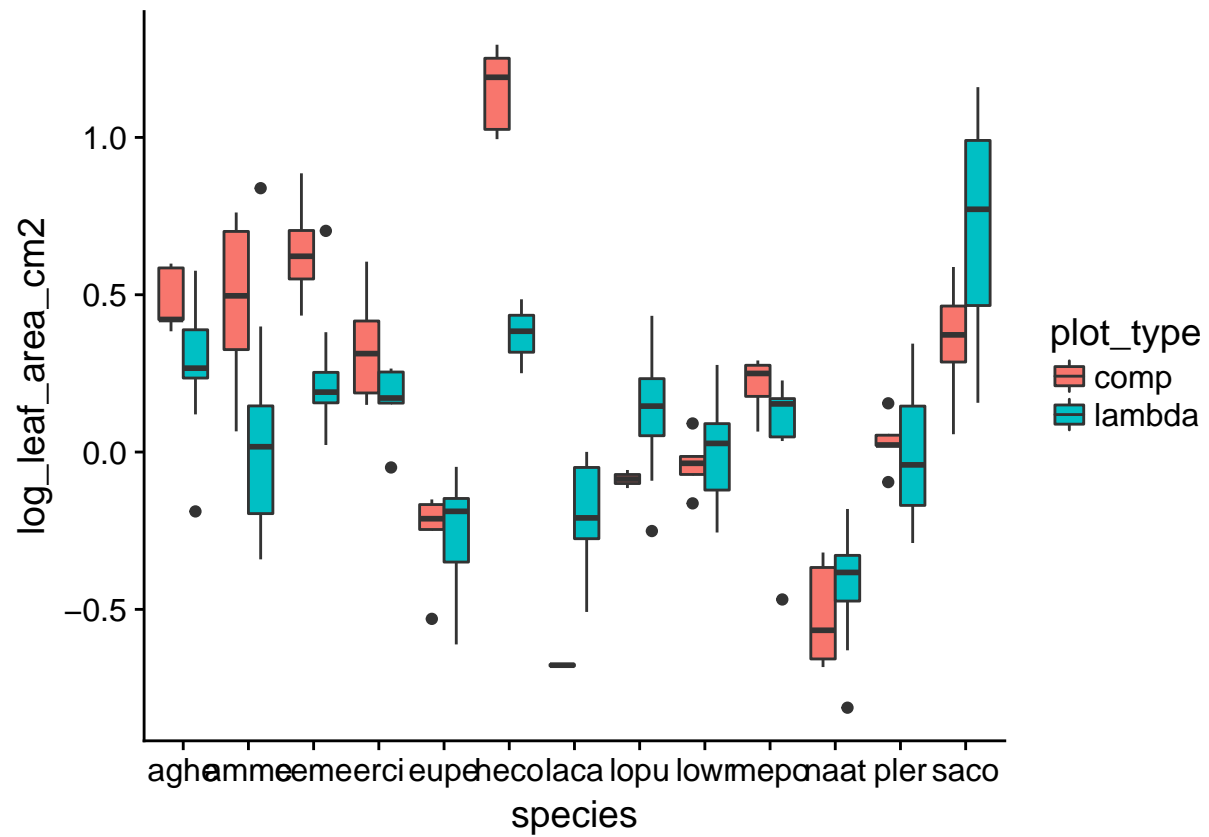
```
pl + geom_boxplot(aes(y = log_sla_cm2_g, x = species, fill = plot_type))
```

## Warning: Removed 4 rows containing non-finite values (stat\_boxplot).

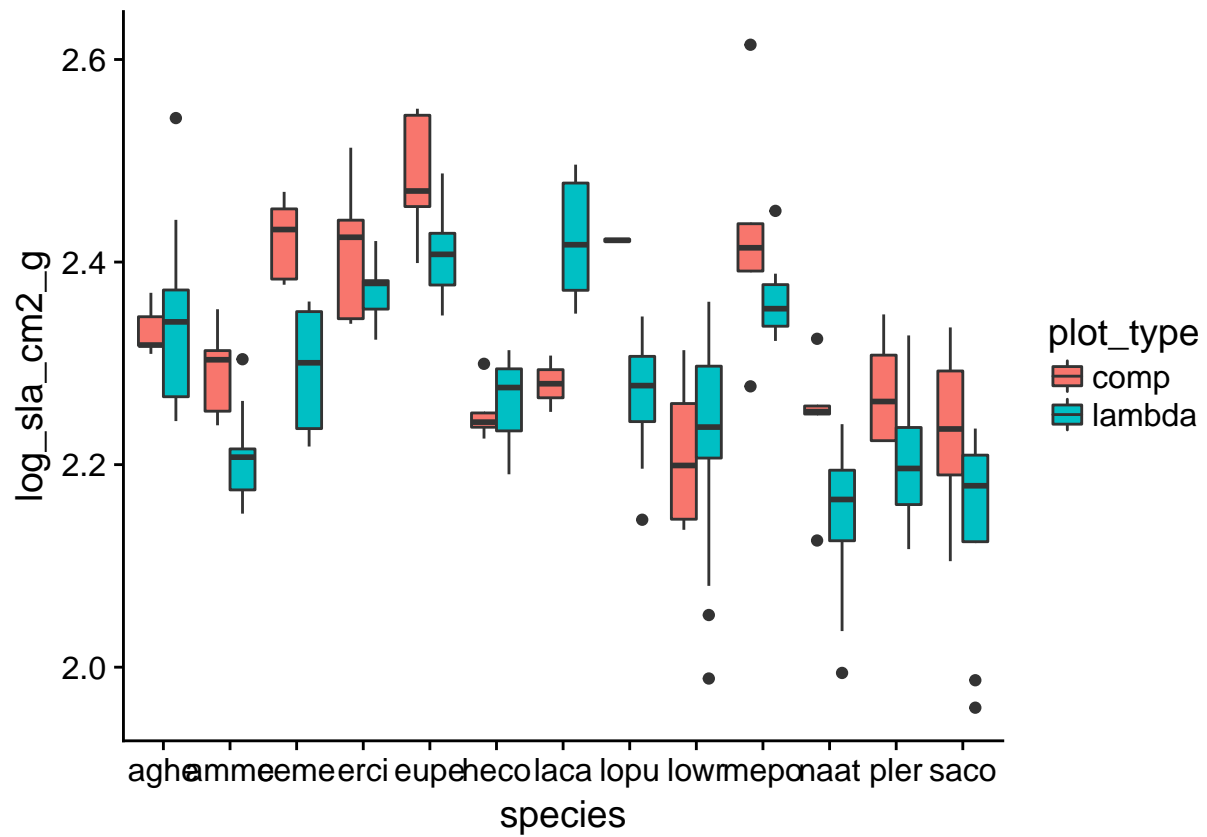


```
# I think we need to get rid of MICA for now :-)
individual_traits <- individual_traits %>% filter(!(species %in% c("mica")))
individual_traits <- individual_traits %>% filter(is.finite(sla_cm2_g))
pl <- ggplot(data= individual_traits)
pl + geom_boxplot(aes(y = log_leaf_area_cm2, x = species, fill = plot_type))
```

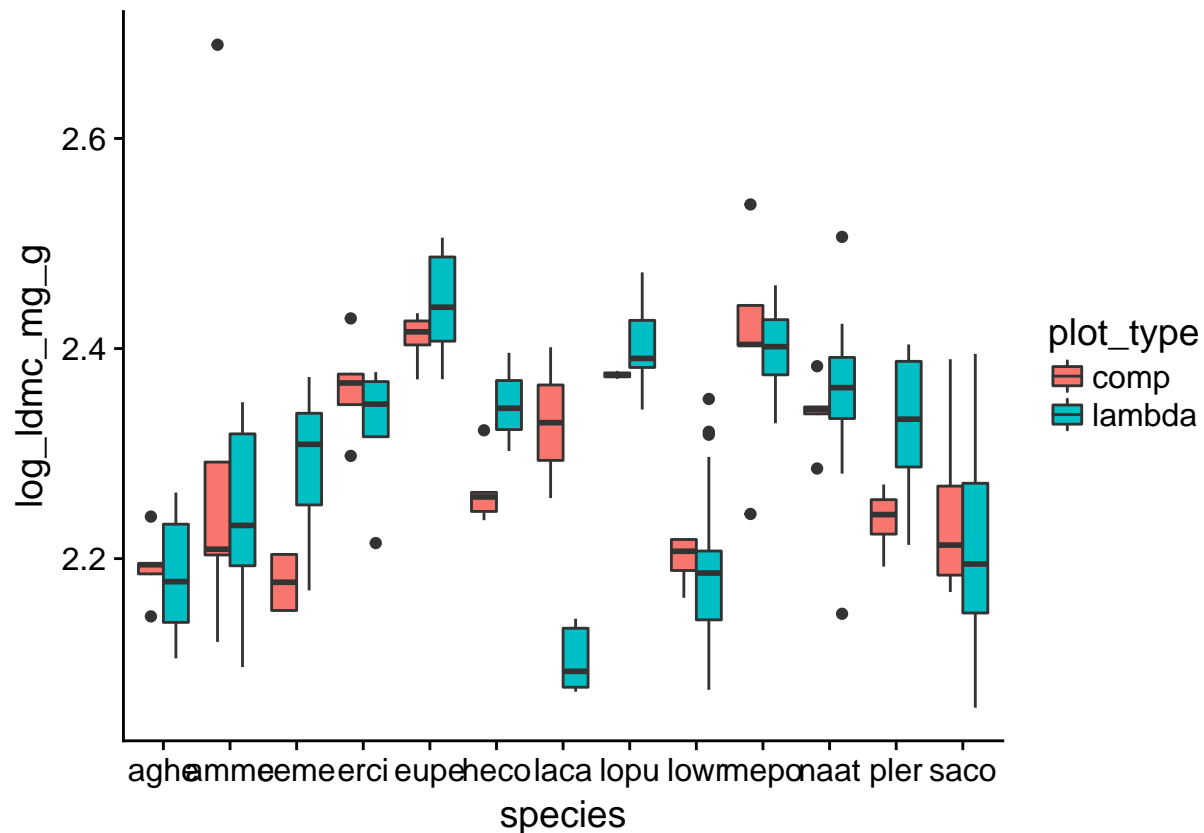




```
pl + geom_boxplot(aes(y = log_sla_cm2_g, x = species, fill = plot_type))
```



```
pl + geom_boxplot(aes(y = log_ldmc_mg_g, x = species, fill = plot_type))
```



```
individual_traits
```

```
## # A tibble: 194 x 10
## # Groups:   plot_type, plot, species [51]
##   plot_type plot species plant_number leaf_area_cm2 ldmc_mg_g sla_cm2_g
##   <chr> <chr> <chr> <chr> <dbl> <dbl> <dbl>
## 1 comp comp aghe 1 3.847667 139.6463 234.2334
## 2 comp comp aghe 2 3.971000 153.2782 207.7736
## 3 comp comp aghe 3 2.636333 156.7932 208.0037
## 4 comp comp aghe 4 2.420000 156.3852 221.8313
## 5 comp comp aghe 5 2.638333 173.7939 203.8769
## 6 comp comp amme 1 5.773000 132.0554 225.6465
## 7 comp comp amme 2 1.163667 195.8444 178.9707
## 8 comp comp amme 3 3.137333 488.9206 173.3157
## 9 comp comp amme 4 2.116333 161.7953 201.1727
## 10 comp comp amme 5 5.026000 159.7637 205.3974
## # ... with 184 more rows, and 3 more variables: log_leaf_area_cm2 <dbl>,
## # log_sla_cm2_g <dbl>, log_ldmc_mg_g <dbl>
```

```
sla_aov <- lm(data = individual_traits, log_sla_cm2_g ~ species * plot_type)
anova(sla_aov)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: log_sla_cm2_g
```

```
##      Df Sum Sq Mean Sq F value    Pr(>F)
## species    12  1.40600  0.117166  23.6190 < 2.2e-16 ***
## plot_type     1  0.11712  0.117121  23.6098 2.689e-06 ***
```

```
## species:plot_type 12 0.14504 0.012086 2.4364 0.006037 **
## Residuals        168 0.83340 0.004961
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

ldmc_aov <- lm(data = individual_traits, log_ldmc_mg_g~species*plot_type)
anova(ldmc_aov)

## Analysis of Variance Table
##
## Response: log_ldmc_mg_g
##           Df Sum Sq Mean Sq F value    Pr(>F)
## species      12  1.40891  0.117409  21.8165 < 2.2e-16 ***
## plot_type      1  0.00211  0.002105   0.3912  0.532522
## species:plot_type 12  0.17680  0.014734   2.7378  0.002054 **
## Residuals    168  0.90412  0.005382
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

la_aov <- lm(data = individual_traits, log_leaf_area_cm2~species*plot_type)
anova(la_aov)

## Analysis of Variance Table
##
## Response: log_leaf_area_cm2
##           Df Sum Sq Mean Sq F value    Pr(>F)
## species      12 20.0379  1.66982  42.7008 < 2.2e-16 ***
## plot_type      1  0.2459  0.24587   6.2874  0.01311 *
## species:plot_type 12  3.2068  0.26724   6.8338 5.624e-10 ***
## Residuals    168  6.5697  0.03911
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Let's ask some simple questions

On average, did trait values go up or down or stay equal?

```
species_summary <- individual_traits %>% group_by(species, plot_type) %>% summarize_at(vars(leaf_area_cm2, log_ldmc_mg_g, log_sla_cm2_g))

species_summary2 <- species_summary %>% gather(variable, value, -(plot_type:species)) %>% unite(temp, plot_type, species)

species_summary2 <- species_summary2 %>% mutate(diff_area_LminusC = lambda_log_leaf_area_cm2 - comp_log_leaf_area_cm2,
diff_ldmc_LminusC = lambda_log_ldmc_mg_g - comp_log_ldmc_mg_g,
diff_sla_LminusC = lambda_log_sla_cm2_g - comp_log_sla_cm2_g)

range(species_summary$log_sla_cm2_g)

## [1] 142.3007 307.4973
range(species_summary$log_ldmc_mg_g)

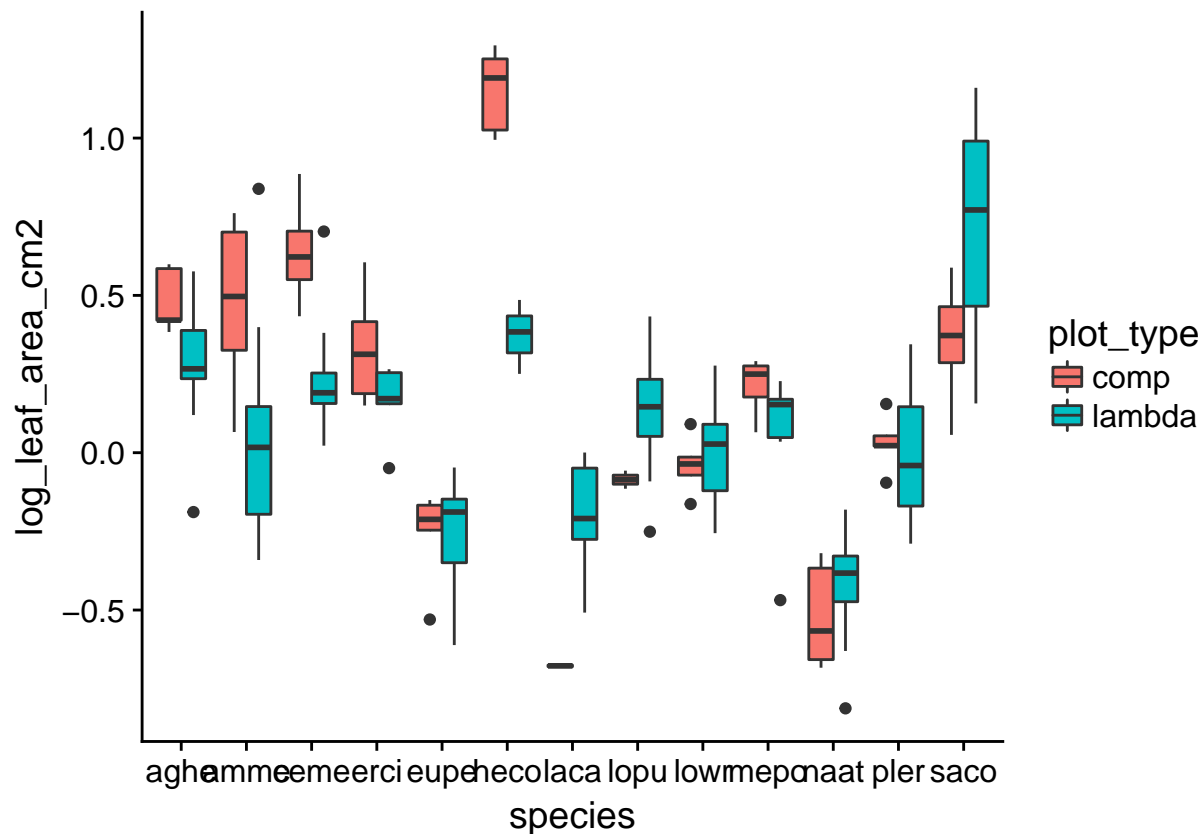
## [1] 127.3479 279.2649
range(individual_traits$leaf_area_cm2)
```

```
## [1] 0.15400 19.70567
```

```
View(species_summary2)
```

## Figures for poster

```
cbPalette <- c("#999999", "#E69F00", "#56B4E9", "#009E73", "#F0E442", "#0072B2", "#D55E00", "#CC79A7")
give.n <- function(x){ return(c(ymin = median, label = length(x))) }
pl + geom_boxplot(aes(y = log_leaf_area_cm2, x = species, fill = plot_type))
```



```
theme = theme_set(theme_minimal())
theme = theme_update(legend.position="top", legend.title=element_blank(), panel.grid.major.x=element_blank(),
  slaplot <- pl + geom_boxplot(aes(y = log_sla_cm2_g, x = species, fill = plot_type), position=position_dodge(),
    labs(x = "", y = expression("log"[10]*"(Specific Leaf Area) (cm" ^2* "/g)")) +
    ggtitle("Trait varies across species and competitive background") +
    theme(plot.title = element_text(face = "bold", size = 25, vjust=2, hjust = 0.5), axis.title = element_text(size = 14, face = "plain"),
    guides(fill=FALSE) + scale_fill_manual(values=c("skyblue", "orange"))

ldmcpplot <- pl + geom_boxplot(aes(y = log_ldmc_mg_g, x = species, fill = plot_type), position=position_dodge(),
  labs(x = "", y = expression("log"[10]*"(Leaf Dry Matter Content) (mg/g)")) +
  theme(axis.title = element_text(size = 14, face = "plain"), legend.position = "bottom") +
  scale_fill_manual(values=c("skyblue", "orange")) + guides(fill=FALSE)
  # guides(fill=guide_legend(title=NULL), plot_type = guide_legend(override.aes = list(lwd = .1)))

sizeplot <- pl + geom_boxplot(aes(y = log_leaf_area_cm2, x = species, fill = plot_type), position=position_dodge(),
  labs(x = "Species", y = expression("log"[10]*"(Leaf Size) (cm" ^2* ")")) +
  theme(axis.title = element_text(size = 14, face = "plain"), legend.position = "bottom") +
```

```
scale_fill_manual(values=c("skyblue", "orange")) +  
guides(fill=guide_legend(title=NULL), plot_type = guide_legend(override.aes = list(lwd = .1)))  
  
to_save <- plot_grid(slaplot, ldmcplot, sizeplot, align = 'v', ncol = 1, labels = c("A: SLA", "B: LDMC"  
save_plot("~/grad/conferences/esa2018/vanDyke-poster/alltraits.png", to_save, base_height = 18, base_as
```