

Lauryn Angell

John Collet

Tripti Rathi

Color Me Impressed

Plants, for the sake of simplicity, are green. Although it is well established that the presence of chlorophyll contributes to the green hue of leaves, the causes for variation in the green color of leaves have yet to be extensively explored. The relationship between color and other physical, genotypic, and environmental factors is largely unknown.

Leaf color is dependent on chlorophyll density. Indeed, chlorophyll is a poor absorber of green and near green portions of the visible light spectrum (Atwell *et al.* 1999). In short, the denser the chlorophyll content, the darker the leaf. Leaf color also indicates tree sickness or environmental stresses (Schaber *et al.* 2009). Similarly, small leaves are associated with harsh conditions such as water constraints and nutrient deficiency or even heat loss (Nicotra *et al.* 2011), indicating that size and color could have some connection. Analyzing leaves in regards to this prior research could help establish a concrete relationship between leaf color and variables such as location, ploidy, and leaf size.

The quaking aspens (*Populus tremuloides*) northwest of Silver Lake in the Big Cottonwood Canyon are an excellent population to determine the ways in which leaf color is affected by various factors. Aspen leaves are unique in a variety of aspects including shape, position on the tree, and their tendency to quake in even the slightest breeze, hence their name. More importantly, aspens' reproduction system includes both sexual and asexual reproduction, the latter of which results in large populations of clones. While most sexually produced aspens are diploids, a small number are triploids. Triploids seem to have some differences including lower wood density (Yanchuk *et al.* 1983) and larger clonal populations (Mock *et al.* 2012), indicating that they may differ in other respects as well, such as leaf color. In addition, genome size in plants correlates to cell size (Bennett, 2005), so the triploids larger genome would lead to larger cell size as well.

We are applying these factors to the silver lake aspen population. The aspens in this area lie on a slope rising from east to west with a variety of different diploid and triploid clones. We will be analyzing factors within and among clones to determine how leaf color varies among ploidy, clonality, and size of leaf.

To reiterate, we hypothesize the following:

1. Since triploid leaves have bigger cells, they will be lighter because the chlorophyll will be more diffused throughout the leaf.

2. Leaves that have a smaller size in terms of length and width will be darker and therefore have higher chlorophyll density in order to maintain the same level of energy absorption as leaves with a larger size.
3. Leaf color will vary in some way in relation to latitude and longitude.

Ultimately, the goal is to determine whether substantial variations in leaf color exist and the cause behind these variations.

To compare leaf color versus three different factors (tree clonality, ploidy, and leaf size), three different comparisons were conducted. To ensure that the data was consistent among the clones, only leaves from 9/4/13 and 9/8/14 were examined. Groups in Jon Seger's Biology of Variation class collected leaves from Big Cottonwood, then pasted the leaves on sheets of printer paper (around 20 or so per sheet and 3 sheets per tree). The leaf sheets were scanned and uploaded onto a server.

To measure area, the leaf sheets were uploaded onto Image J, an image processing system. Length and width measurements of each leaf were taken that were then multiplied together to get a general area measurement. The leaves were also genotyped to identify the trees, separate them into clones, and determine ploidy.

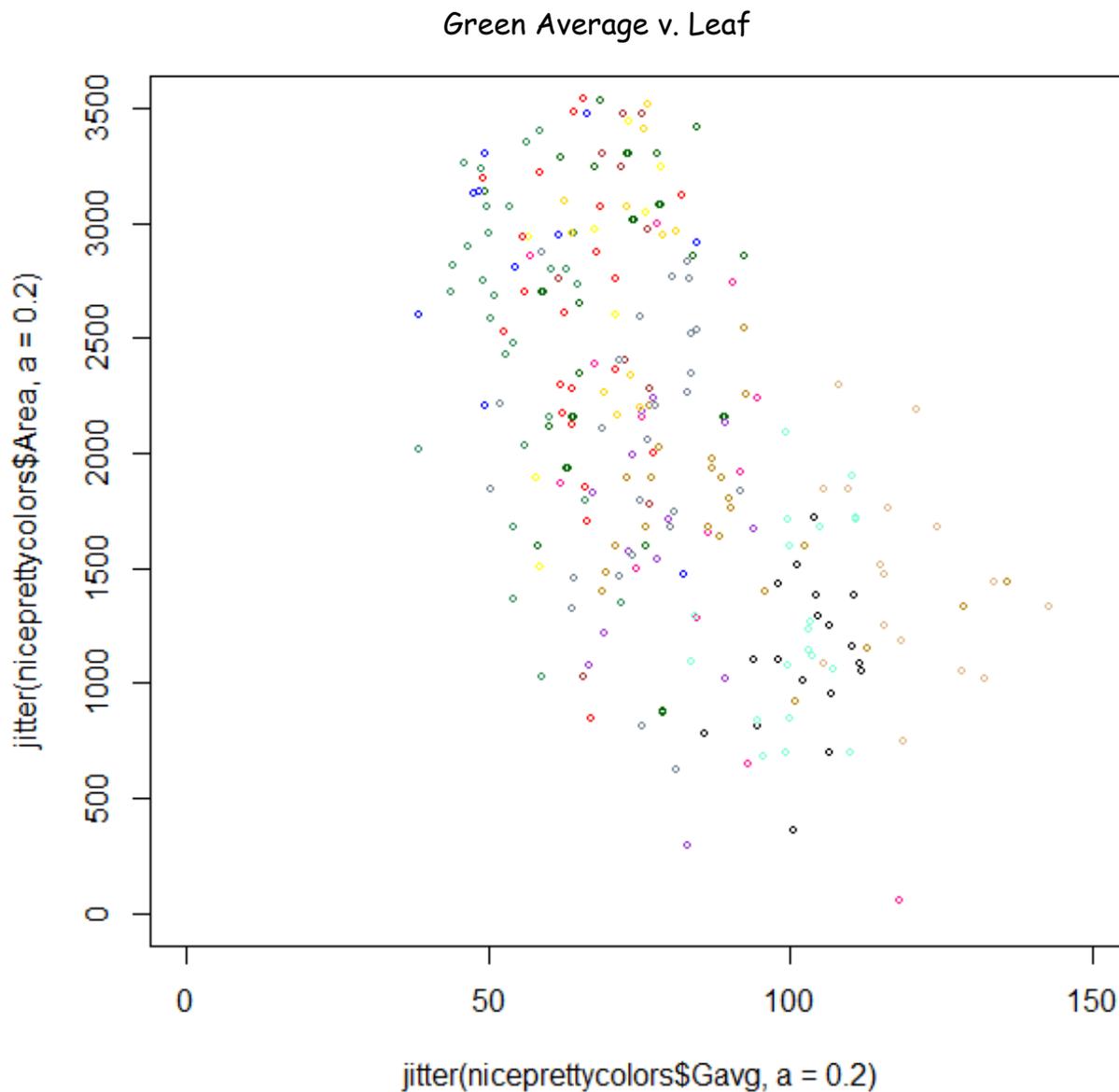
The biggest challenge was determining the scale for measuring or quantifying color. Photoshop, an image analysis software, was utilized to determine the saturation and hue of the leaf. On the leaf, two places were evaluated for color. With the marquee tool, a section from the right hemisphere was cut, blurred, and then quantified with the eyedropper tool. This process produced three values (red, green, blue). The same was done to the left side and the color values were averaged.

The data gathered (tree, clone, ploidy, leaf area, color values) was uploaded to an excel spreadsheet which was used for an analysis of variance or ANOVA in the program R. After constraining the data set to remove any outliers (blue1-b2 values greater than 15), the leaf color data was compared against leaf area, clone, and ploidy to yield results.

Results:

Legend:

7_7 pink	7_2 green
7_8 tan	3_5 slate gray
8_1 yellow	3_2 gold
8_2 violet	4_1 sea green
5_5 brown	4_2 blue
5_6 purple	5_2 red
7_1 aquamarine	7_3 black



```
#Comparing Gavg to Area
```

```
Analysis of Variance Table
```

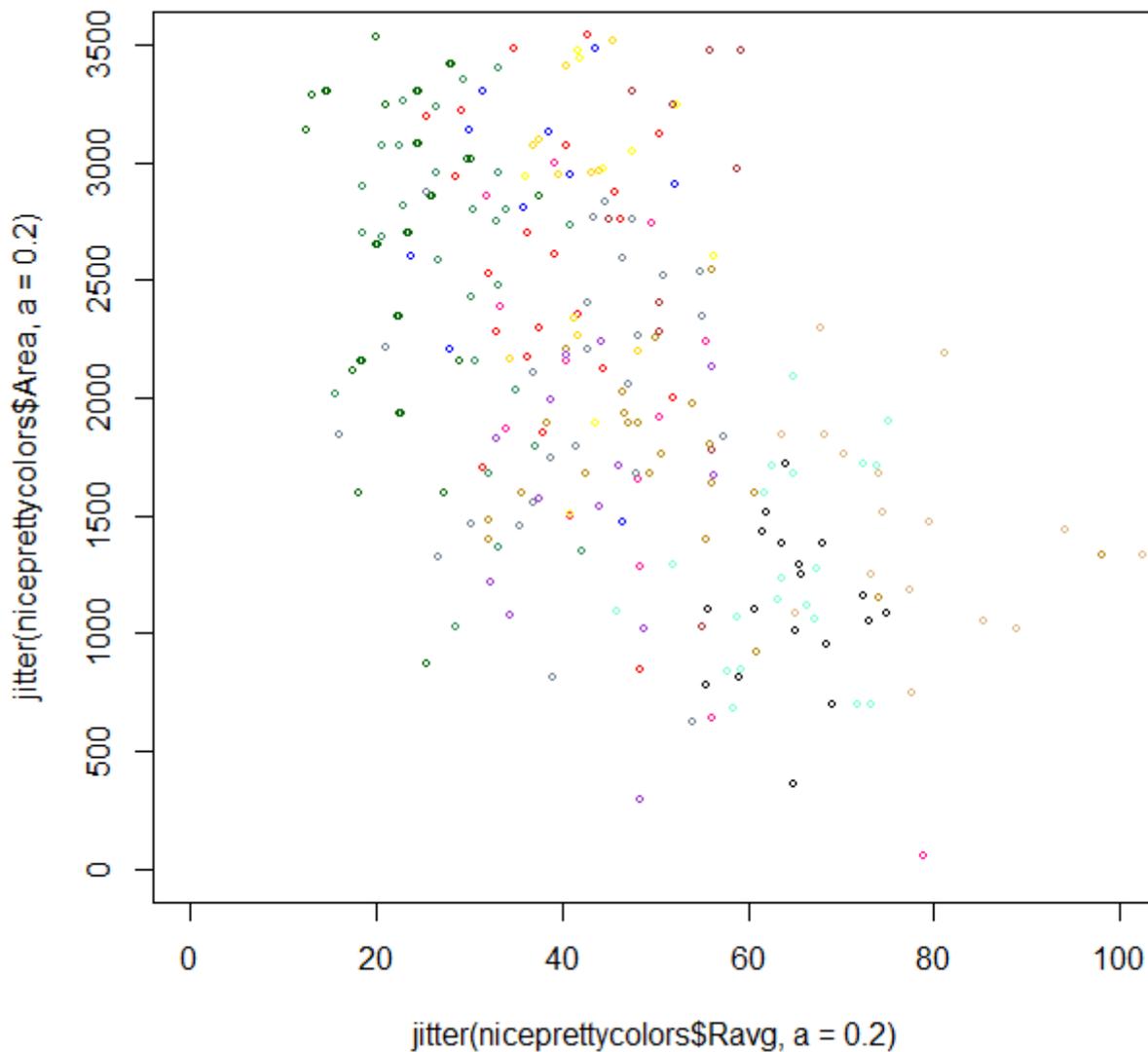
R	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Area	1	41080	41080	144.63	< 2.2e-16 ***
Residuals	313	88906	284		

```
summary(G2a)
```

```
Residual standard error: 16.85 on 313 degrees of freedom
```

```
Multiple R-squared: 0.316, Adjusted R-squared: 0.3139
```

F-statistic: 144.6 on 1 Red Average v. Area and 313 DF, p-value: <



2.2e-16

#Comparing Ravg to Area

Analysis of Variance Table

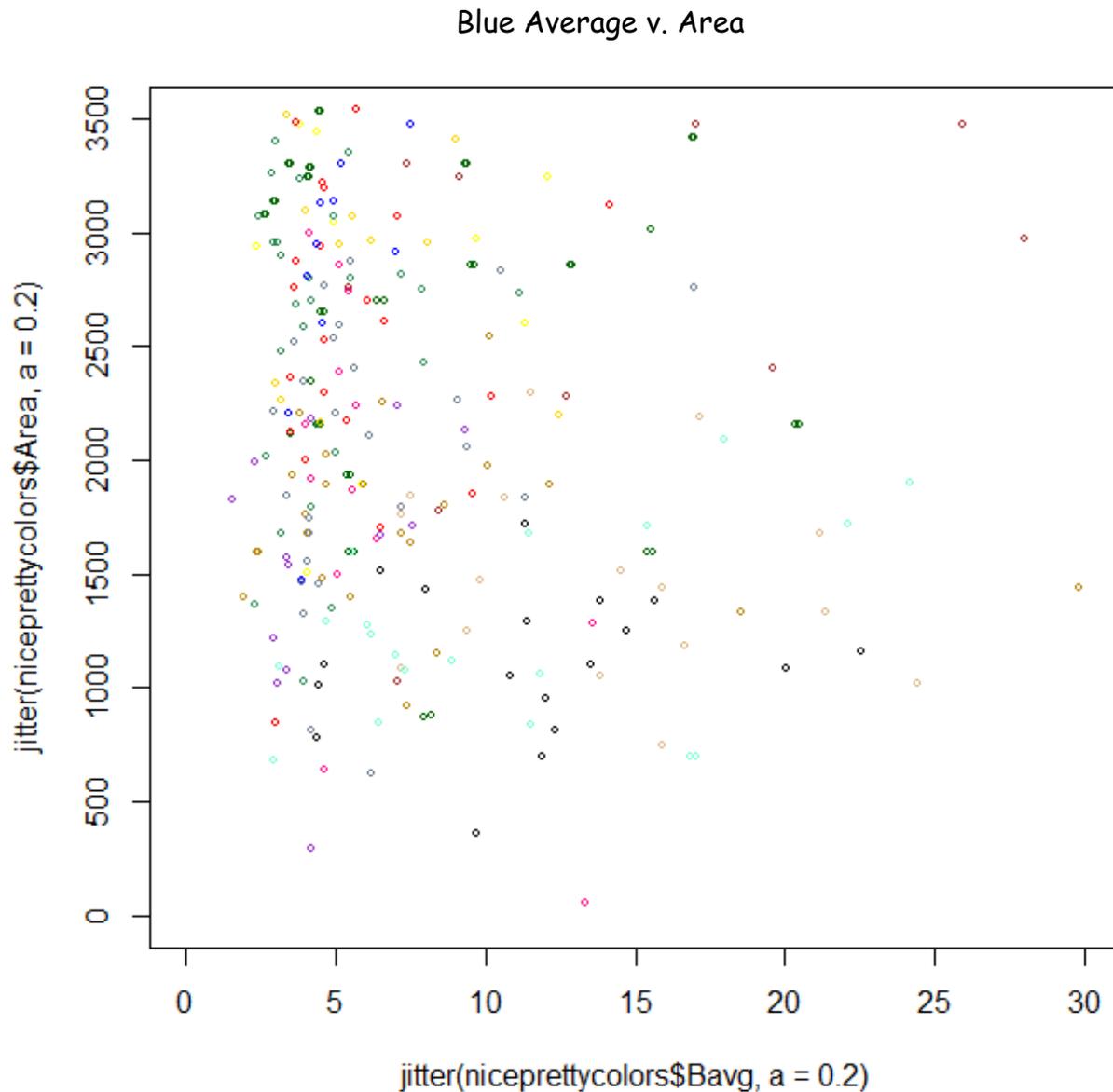
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Area	1	24187	24186.8	103.54	< 2.2e-16 ***
Residuals	313	73115	233.6		

>summary(R2a)

Residual standard error: 15.28 on 313 degrees of freedom

Multiple R-squared: 0.2486, Adjusted R-squared: 0.2462

F-statistic: 103.5 on 1 and 313 DF, p-value: $< 2.2e-16$



#Comparing Bavg to Area

Analysis of Variance Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Area	1	216.2	216.245	8.3608	0.004102 **
Residuals	313	8095.5	25.864		

> summary(B2a)

Residual standard error: 5.086 on 313 degrees of freedom

Multiple R-squared: 0.02602, Adjusted R-squared: 0.02291

F-statistic: 8.361 on 1 and 313 DF, p-value: 0.004102

Green vs. Tree

```
> G2tree <- lm(Gavg ~ tree, data=niceprettycolors)
```

```
> anova(G2tree)
```

Analysis of Variance Table

Response: Gavg

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
tree	13	97191	7476.3	68.619	< 2.2e-16

Residuals	301	32795	109.0		
-----------	-----	-------	-------	--	--

```
> summary(G2tree)
```

Residual standard error: 10.44 on 301 degrees of freedom

Multiple R-squared: 0.7477, Adjusted R-squared: 0.7368

F-statistic: 68.62 on 13 and 301 DF, p-value: < 2.2e-16

Green vs. tree and area

```
> G2tree <- lm(Gavg ~ tree + Area, data=niceprettycolors)
```

```
> anova(G2tree)
```

Analysis of Variance Table

Response: Gavg

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
tree	13	97191	7476.3	69.5221	< 2e-16 ***
Area	1	534	533.7	4.9625	0.02664 *
Residuals	300	32261	107.5		

```
> summary(G2tree)
```

Residual standard error: 10.37 on 300 degrees of freedom

Multiple R-squared: 0.7518, Adjusted R-squared: 0.7402

F-statistic: 64.91 on 14 and 300 DF, p-value: < 2.2e-16

There's hardly any increase in the R-squared value between the Gavg vs. tree and Gavg vs. tree and Area, so Area has little to do with color. Most of it can be explained because area varies by tree as shown by:

```
> A2t <- lm(Area ~ tree, data=niceprettycolors)
```

```
> anova(A2t)
```

Analysis of Variance Table

Response: Area

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
tree	13	198201119	15246240	23.183	< 2.2e-16 ***
Residuals	301	197947346	657632		

> summary(A2t)

Residual standard error: 810.9 on 301 degrees of freedom

Multiple R-squared: 0.5003, Adjusted R-squared: 0.4787

F-statistic: 23.18 on 13 and 301 DF, p-value: < 2.2e-16

Therefore, area can be ruled out.

Because "tree" is a subset of clone which is a subset of ploidy, we cannot perform a test that accounts for all three simultaneously as anova only accounts for the most specific condition, namely tree. Taking the averages for each tree produces unreliable results due to our relatively small sample of leaves for each tree. The same is especially true for clones and ploidy. As such, we only accounted for one genetic variable at a time, and looking at the graphs, our values seem to make sense (tree has the tightest relationship, then clone, then ploidy).

Gavg vs. clone

```
> G2clone <- lm(Gavg ~ clone, data=niceprettycolors)
> anova(G2clone)
```

Analysis of Variance Table

Response: Gavg

		Df	Sum Sq	Mean Sq	F value	Pr(>F)
clone	4	67992	16998		84.997	< 2.2e-16

Residuals 310 61995 200

> summary(G2clone)

Residual standard error: 14.14 on 310 degrees of freedom

Multiple R-squared: 0.5231, Adjusted R-squared: 0.5169

F-statistic: 85 on 4 and 310 DF, p-value: < 2.2e-16

#Gavg vs. ploidy

```
> G2ploidy <- lm(Gavg ~ ploidy, data=niceprettycolors)
> anova(G2ploidy)
```

Analysis of Variance Table

Response: Gavg

		Df	Sum Sq	Mean Sq	F value	Pr(>F)
ploidy	1	37984	37984	129.23		< 2.2e-16 ***
Residuals	313	92002	294			

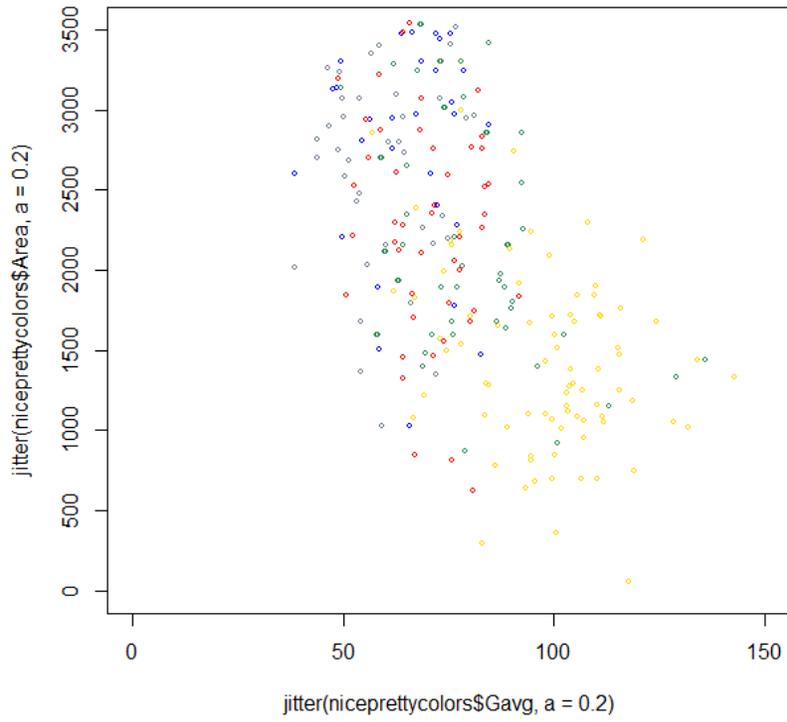
> summary(G2ploidy)

Residual standard error: 17.14 on 313 degrees of freedom

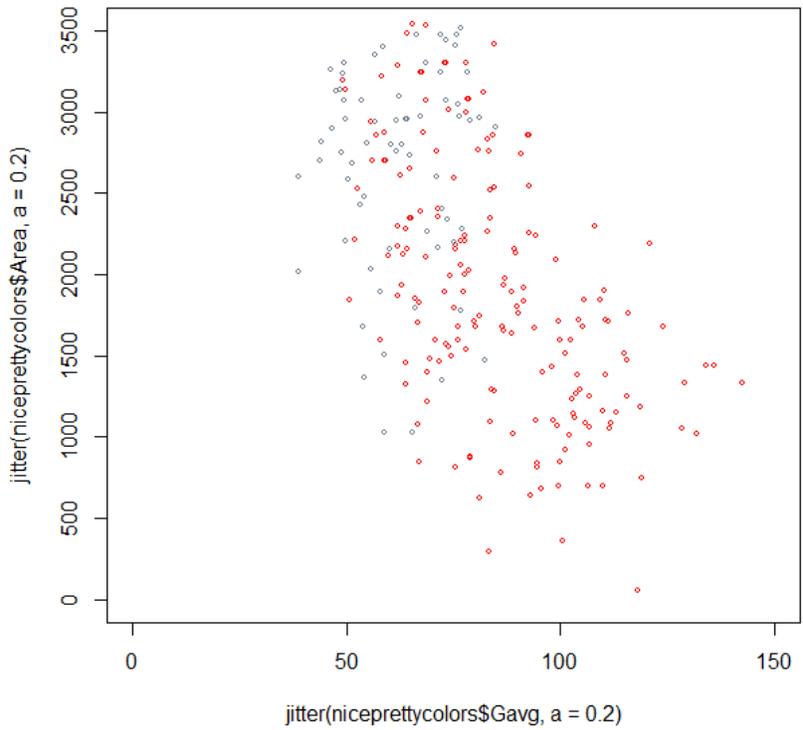
Multiple R-squared: 0.2922, Adjusted R-squared: 0.29

F-statistic: 129.2 on 1 and 313 DF, p-value: < 2.2e-16

Green Average vs. Clone



Green Average vs. Ploidy



RED-diploid

BLUE-triploid

Discussion/Conclusion:

We compared leaf color values to area, tree, and ploidy. According to the ANOVA analysis, the multiple R-squared value only rose to 75.18% from 74.77% when area was accounted for, indicating that area is not a factor in leaf color, contrary to our hypothesis. In contrast, when color was compared to clone, the multiple R-squared value was 52.31% and the P value was less than $2.2e-16$, indicating huge significance. Similar results were gathered with ploidy since the R-squared value was 29.22% and the P value was $2.2e-16$, showing significance. Furthermore, the diploid trees were lighter in color than the triploid trees. This could be explained by the fact that bigger cells, indicating more chlorophyll might be in the leaves, causing a darker leaf.

Several limitations exist in our research. First, the sample set is considerably small and inadequate to establish a certain relationship between color and various factors. The data was also prone to pseudoreplication or inflation of data set. Second, several the inconsistencies in color analysis i.e. human error, analysis on different parts of the leaf, differences in the size of leaf evaluated certainly impact the final outcome. Finally, time constraints and potential human error affected the final research.

Our research raises several questions. First, does leaf color indicate anything about leaf fitness? Second, does a tree's location influence its leaf color? Also, how does leaf color relate to a tree's health/fitness? These questions can further cement a relationship between leaf color and tree factors.