Group Project - Math 4339

Anuska Nair

Madeline Luong

Phy Tran

*The Influence of Elevalation on Eight Alpine Species and their Endophyte Microbial Diversity*

1. **Introduction:**

Alpine plant species are pioneer plants, which colonize previously biodiverse ecosystems. In these ecosystems, abiotic stresses can influence how plants adapt to the new environment. We seek to measure plants’ adaptation by exploring the species richness and diversity. We do that by exploring if elevation significantly influences diversity (fungal diversity, bacterial diversity, and plant diversity) and/or if elevation significantly influence species richness (morphology richness, fungal richness, morphology count). We begin with a dataset of seeds collected from eight different plant species along an elevational gradient at Nederland, Colorado; then OTU (Fungi [ITS] diversity and Bacteria Diversity) data are extracted from each elevation-plant species pair. We will use multivariate hypothesis testing (MANOVA/ ANOVAS, LINEAR MIXED MODELS, AND PCAs) to evaluate if the influences of the new environment are significant. Finally, we will use a chi- square analysis to confirm.

1. **Hypothesis Testing**
2. *Normality of the Variables*:

We test the normality of the variables as well as the multivariate normality in R. Bacterial Richness is the only variable whose data are normally distributed. Neither the multivariate data nor the univariate analysis of the rest of the variables show normal distribution.

$multivariateNormality

Test Statistic p value Result

1 Mardia Skewness 198.008023195746 3.37017046737806e-11 NO

2 Mardia Kurtosis 2.94747218979628 0.00320383568738225 NO

3 MVN <NA> <NA> NO

$univariateNormality

Test Variable Statistic p value Normality

1 Shapiro-Wilk Plant div 0.7949 1e-04 NO

2 Shapiro-Wilk morphct 0.6221 <0.001 NO

3 Shapiro-Wilk morphrich 0.8284 2e-04 NO

4 Shapiro-Wilk fungi rich 0.9222 0.0306 NO

5 Shapiro-Wilk Bacter rich 0.9639 0.3885 YES

6 Shapiro-Wilk ITS Div 0.8949 0.0063 NO

7 Shapiro-Wilk Bac DIV 0.8646 0.0013 NO

$Descriptives

n Mean Std.Dev Median Min Max 25th 75th Skew

Plant div 30 3.1300000 0.2533908 3.1200 2.83 3.440 2.830000 3.44000 0.05716240

morphct 30 2.4333333 3.9539010 1.5000 0.00 20.000 0.000000 3.00000 2.95903939

morphrich 30 1.2666667 1.3113124 1.0000 0.00 4.000 0.000000 2.00000 0.49578629

fungi rich 30 4.9666667 3.3060377 4.5000 0.00 11.000 3.000000 7.50000 0.29952610

Bacter rich 30 24.2333333 15.1513817 21.0000 0.00 61.000 12.750000 33.25000 0.53167525

ITS Div 30 0.8274533 0.6408670 1.0275 0.00 1.919 0.093625 1.32125 -0.03690153

Bac DIV 30 2.4256333 0.9065750 2.6570 0.00 3.558 2.085500 2.98525 -1.17172385

Kurtosis

Plant div -1.5983333

morphct 10.2176204

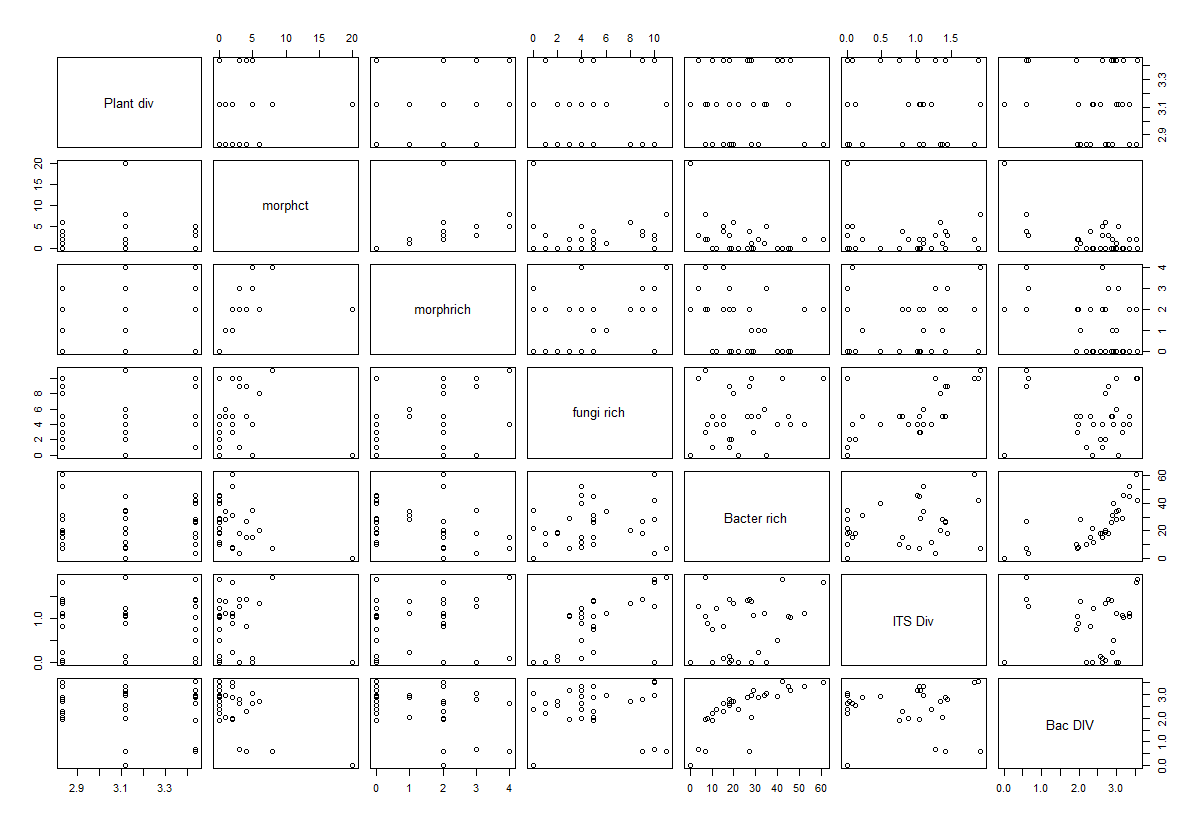
morphrich -1.0294336

fungi rich -1.0417323

Bacter rich -0.5317956

ITS Div -1.3743602

Bac DIV 0.5640672

We can also observe the data distribution in the following scatterplot:

1. *Individual ANOVA*

We perform an ANOVA analysis on individual variables in R.

In relation to types of plants and levels of elevation, the ANOVA analysis shows significant correlation between morphology richness and the types of plants, but not with the level of elevation. The morphology count, bacterial richness, fungal richness, ITS diversity and bacterial diversity, however, are not affected by either types of plants or the levels of elevation. The ANOVA analysis below show that we are not able to reject the null hypothesis.

**Morphology Richness:**

Analysis of Variance Table

Response: morphrich

Df Sum Sq Mean Sq F value Pr(>F)

plant 9 35.200 3.9111 5.1014 0.001632 \*\*

elv 2 0.867 0.4333 0.5652 0.578002

Residuals 18 13.800 0.7667

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**Morphology Count**

Analysis of Variance Table

Response: morphct

Df Sum Sq Mean Sq F value Pr(>F)

plant 9 214.700 23.856 1.9898 0.1024

elv 2 22.867 11.433 0.9537 0.4040

Residuals 18 215.800 11.989

**Fungal Richness**

Analysis of Variance Table

Response: alp$`fungi rich`

Df Sum Sq Mean Sq F value Pr(>F)

plant 9 142.300 15.8111 2.0713 0.09013 .

elv 2 37.267 18.6333 2.4410 0.11534

Residuals 18 137.400 7.6333

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**Bacterial Richness**

Analysis of Variance Table

Response: alp$`Bacter rich`

Df Sum Sq Mean Sq F value Pr(>F)

plant 9 3050.7 338.97 1.7692 0.1448

elv 2 158.1 79.03 0.4125 0.6681

Residuals 18 3448.6 191.59

**Fungal Diversity**

Analysis of Variance Table

Response: alp$`ITS Div`

Df Sum Sq Mean Sq F value Pr(>F)

plant 9 3.3732 0.37480 0.8061 0.6167

elv 2 0.1679 0.08394 0.1805 0.8363

Residuals 18 8.3696 0.46498

**Bacterial Diversity**

Analysis of Variance Table

Response: alp$`Bac DIV`

Df Sum Sq Mean Sq F value Pr(>F)

plant 9 8.4742 0.94158 1.1666 0.3712

elv 2 0.8318 0.41590 0.5153 0.6059

Residuals 18 14.5285 0.80714

1. *Fixing Linear Mixed-Effects Model*

Using lme4 package for R, we attempt to fit restricted maximum likelihood (REML) estimates of our parameters in linear mixed-effects models. Fitting the morphology richness relations to the levels of elevation and using Chi-square hypothesis testing, we are unable to reject the null hypothesis. After removing outlier terms in our data, we still are not able to reject the null hypothesis for morphology richness. The same result is observed for all the other variables except for fungal richness, where we see a significance in p- value of 0.09 after removing the outliers in the data.

**Fitting morphology richness with levels of elevation**

Analysis of Deviance Table (Type II Wald chisquare tests)

Response: alp$morphrich

Chisq Df Pr(>Chisq)

elv 1.1304 2 0.5682

We compute all the single terms in the scope of the argument to decide what can be added to or dropped from the model. Then we fit the new data in the model and also compute a table of the changes in fit.

**Rerun the model after deleting terms**

Single term deletions

Model:

alp$morphrich ~ elv + (1 | plant)

Df AIC LRT Pr(Chi)

<none> 100.300

elv 2 97.518 1.2182 0.5438

**Fitting bacterial richness with levels of elevation**

Linear mixed model fit by REML ['lmerMod']

Formula: alp$`Bacter rich` ~ elv + (1 | plant)

Data: alp

REML criterion at convergence: 230.5599

Random effects:

Groups Name Std.Dev.

plant (Intercept) 7.009

Residual 13.842

Number of obs: 30, groups: plant, 10

Fixed Effects:

(Intercept) elvglvl elvnwt

21.0 4.6 5.1

**Chi-square test for statistical significance**

Analysis of Variance Table

Df Sum Sq Mean Sq F value

elv 2 158.07 79.033 0.4125

We continue with the same term-deletion process discussed above:

**Fitting bacterial richness with levels of elevation**

Model:

alp$`Bacter rich` ~ elv + (1 | plant)

Df AIC LRT Pr(Chi)

<none> 255.34

elv 2 252.24 0.89631 0.6388

**Fitting fungi richness with levels of elevation**

Linear mixed model fit by REML ['lmerMod']

Formula: alp$`fungi rich` ~ elv + (1 | plant)

Data: alp

REML criterion at convergence: 144.9623

Random effects:

Groups Name Std.Dev.

plant (Intercept) 1.651

Residual 2.763

Number of obs: 30, groups: plant, 10

Fixed Effects:

(Intercept) elvglvl elvnwt

3.5 2.7 1.7

**Chi-square test for statistical significance**

Analysis of Variance Table

Df Sum Sq Mean Sq F value

elv 2 37.267 18.633 2.441

We continue with the same term-deletion process discussed above for fungi richness:

**Fitting fungi richness with levels of elevation**

Model:

alp$`fungi rich` ~ elv + (1 | plant)

Df AIC LRT Pr(Chi)

<none> 160.23

elv 2 161.03 4.7997 0.09073 .

---

Signif. codes:

0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Linear mixed model fit by REML ['lmerMod']

Formula: alp$`fungi rich` ~ elv + (1 | plant)

Data: alp

REML criterion at convergence: 144.9623

Random effects:

Groups Name Std.Dev.

plant (Intercept) 1.651

Residual 2.763

Number of obs: 30, groups: plant, 10

Fixed Effects:

(Intercept) elvglvl elvnwt

3.5 2.7 1.7

**Chi-square test for statistical significance**

Analysis of Variance Table

Df Sum Sq Mean Sq F value

elv 2 22.867 11.433 0.9537

**Fitting morphology count with levels of elevation**

Linear mixed model fit by REML ['lmerMod']

Formula: alp$morphct ~ elv + (1 | plant)

Data: alp

REML criterion at convergence: 156.7902

Random effects:

Groups Name Std.Dev.

plant (Intercept) 1.989

Residual 3.462

Number of obs: 30, groups: plant, 10

Fixed Effects:

(Intercept) elvglvl elvnwt

3.6 -2.1 -1.4

**Chi-square test for statistical significance**

Analysis of Variance Table

Df Sum Sq Mean Sq F value

elv 2 22.867 11.433 0.9537

We continue with the same term-deletion process discussed above for morphology count:

Model:

alp$morphct ~ elv + (1 | plant)

Df AIC LRT Pr(Chi)

<none> 173.38

elv 2 171.39 2.0143 0.3653

**Fitting fungal diversity with levels of elevation**

> modITS<- lmer(alp$`ITS Div`~ elv + (1|plant), data=alp)**SHOULD THERE BE A RESULT FOR THIS? I couldn’t run it in my r.**

**Chi-square test for statistical significance**

Analysis of Variance Table

Df Sum Sq Mean Sq F value

elv 2 0.16788 0.083942 0.193

We continue with the same term-deletion process discussed above for fungal diversity:

Model:

alp$`ITS Div` ~ elv + (1 | plant)

Df AIC LRT Pr(Chi)

<none> 66.997

elv 2 63.423 0.42587 0.8082

**Fitting bacterial diversity with levels of elevation**

Linear mixed model fit by REML ['lmerMod']

Formula: alp$`Bac DIV` ~ elv + (1 | plant)

Data: alp

REML criterion at convergence: 79.1319

Random effects:

Groups Name Std.Dev.

plant (Intercept) 0.2117

Residual 0.8984

Number of obs: 30, groups: plant, 10

Fixed Effects:

(Intercept) elvglvl elvnwt

2.2393 0.1548 0.4042

**Chi-square test for statistical significance**

Analysis of Variance Table

Df Sum Sq Mean Sq F value

elv 2 0.8318 0.4159 0.5153

We continue with the same term-deletion process discussed above for bacteria diversity:

Single term deletions

Model:

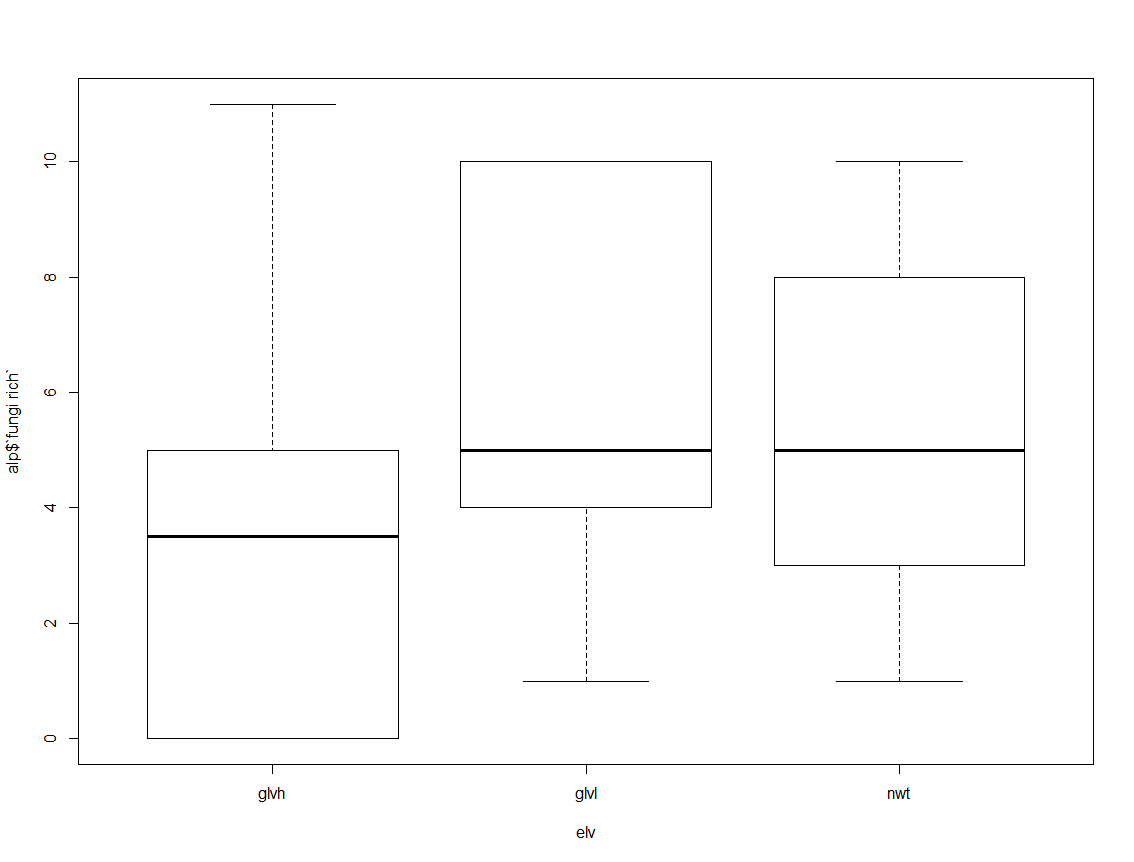
alp$`Bac DIV` ~ elv + (1 | plant)

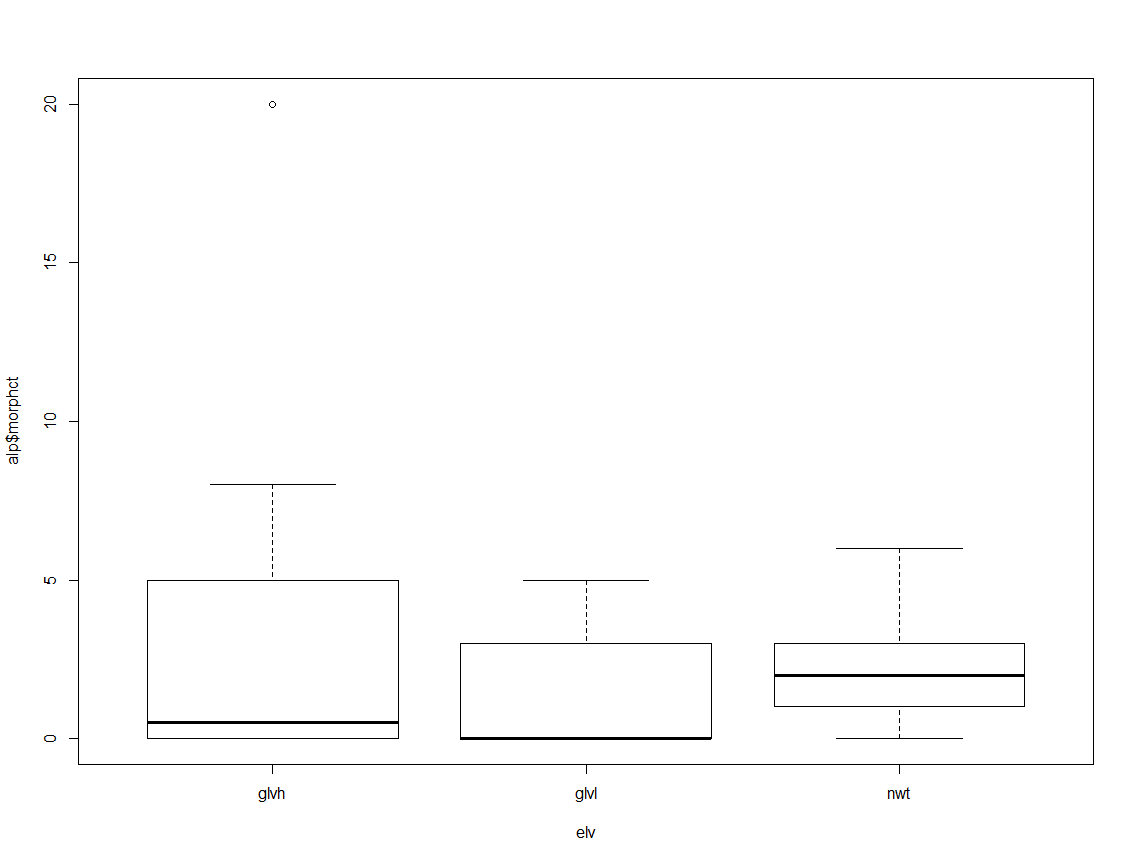
Df AIC LRT Pr(Chi)

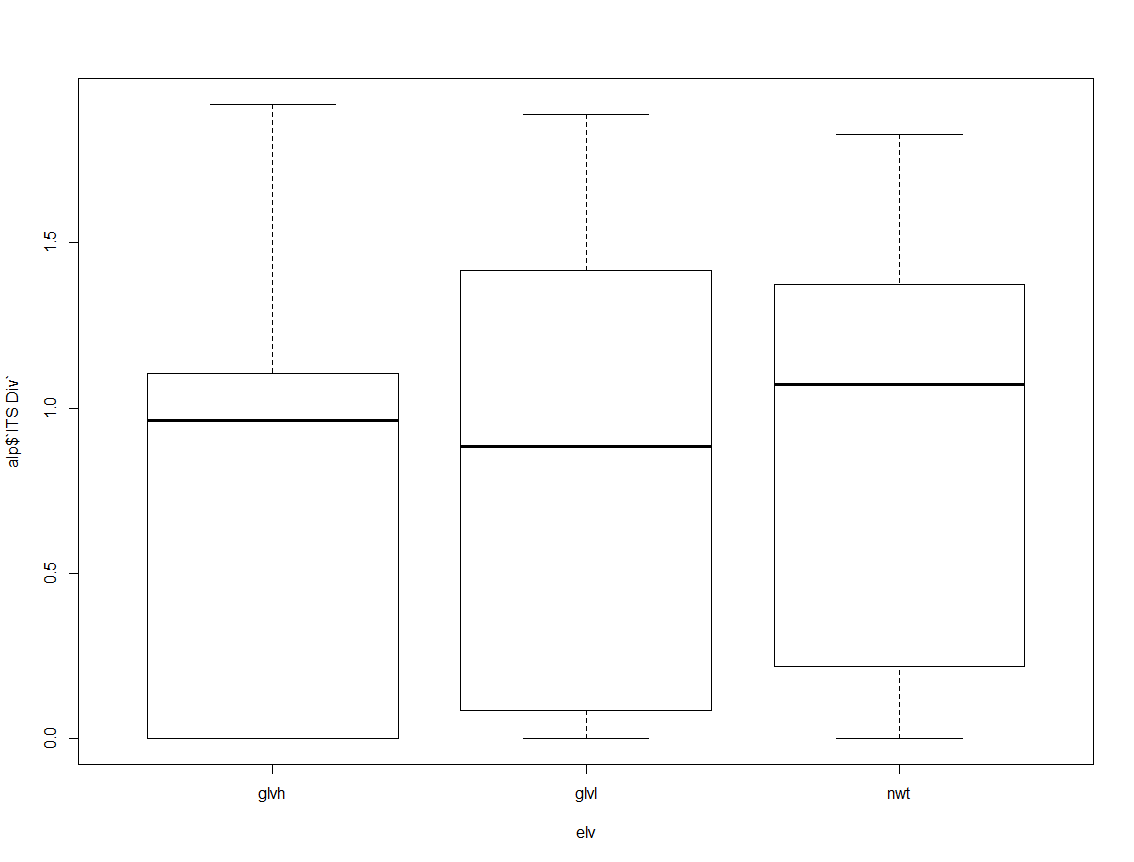
<none> 87.088

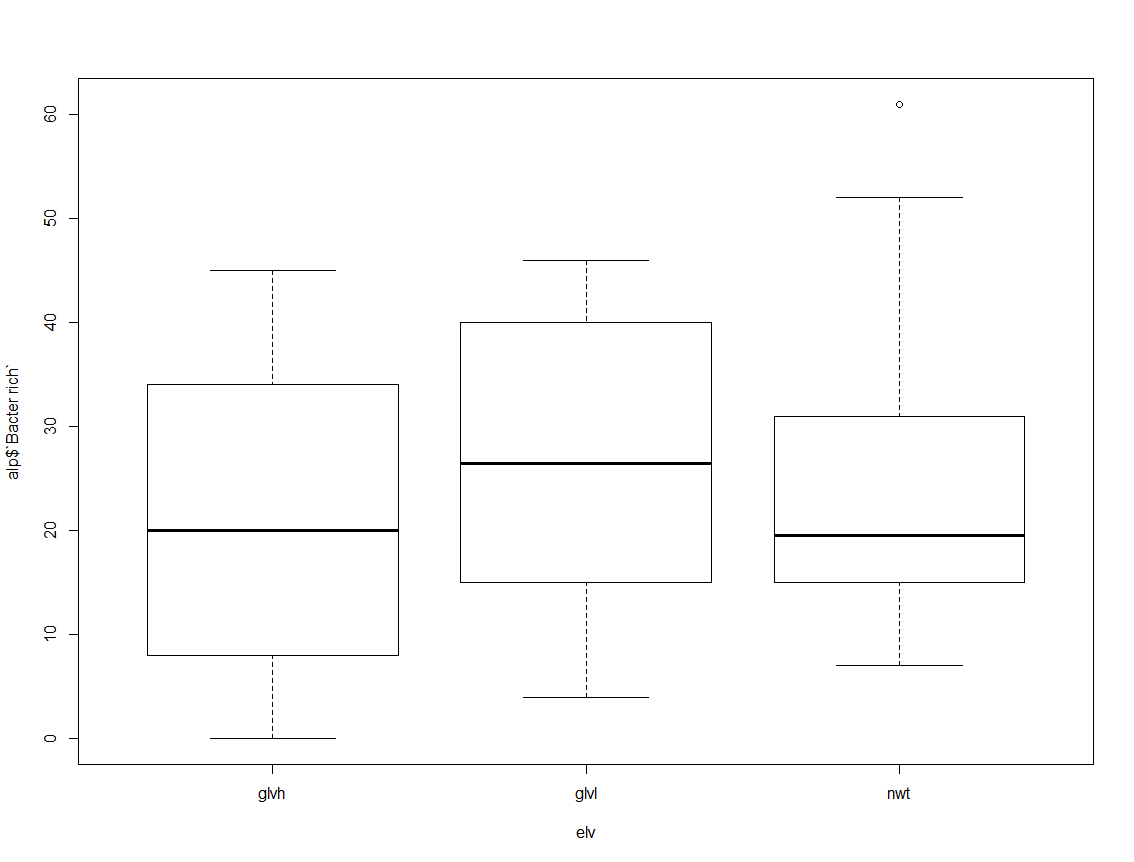
elv 2 84.202 1.1135 0.5731

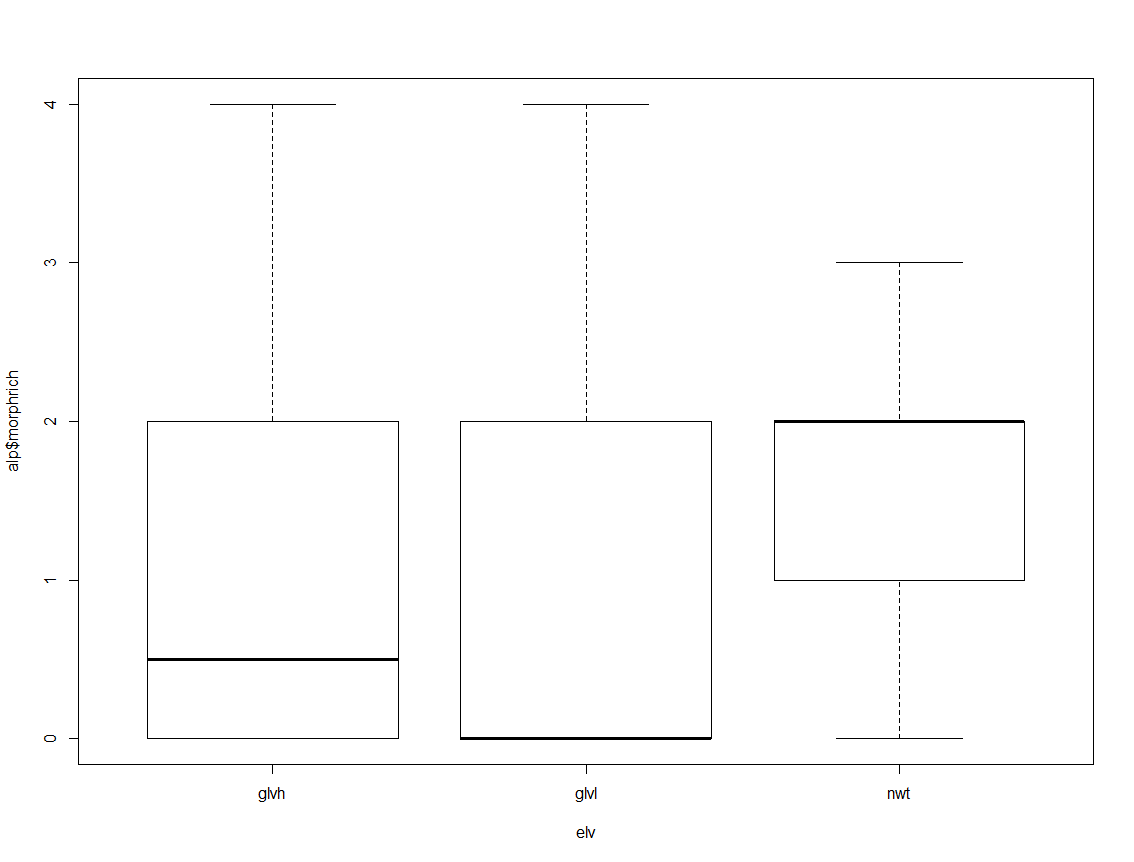
The boxplots below demonstrate the distribution of each variable in terms of levels of elevation; fungi richness is shown to have statistical significance.

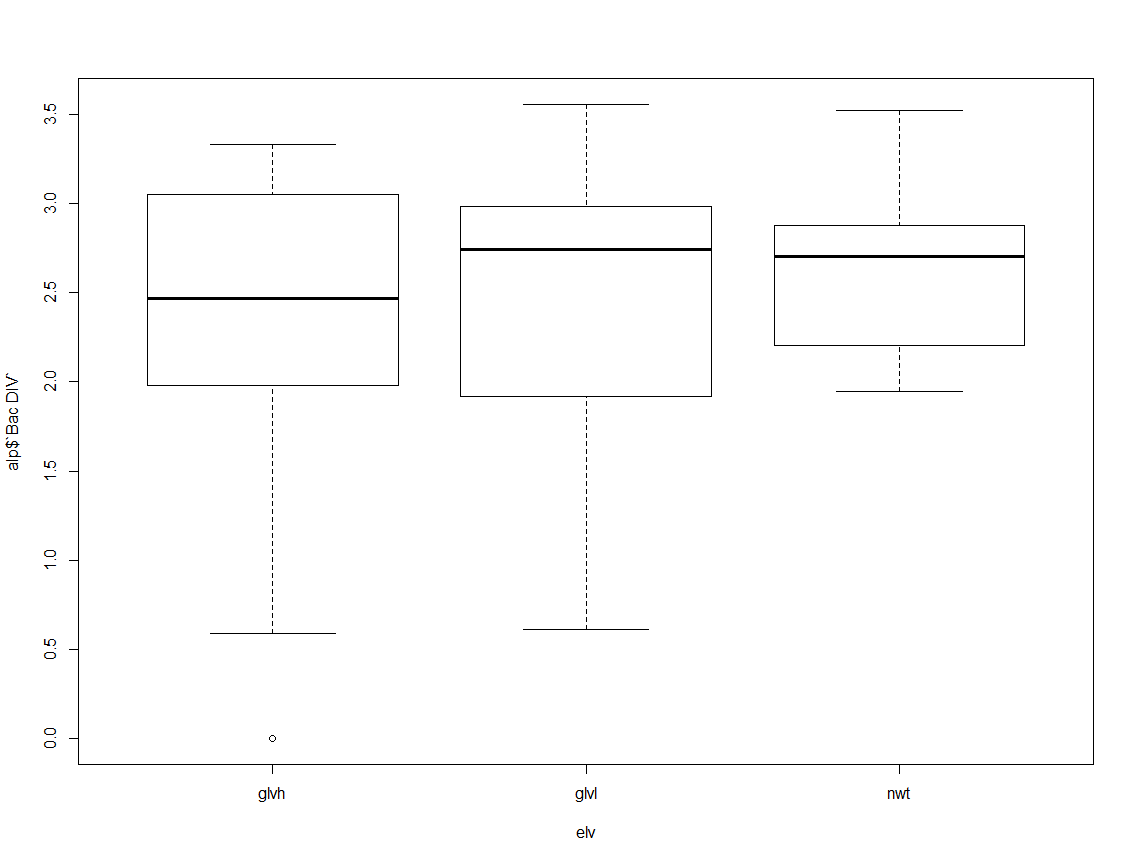




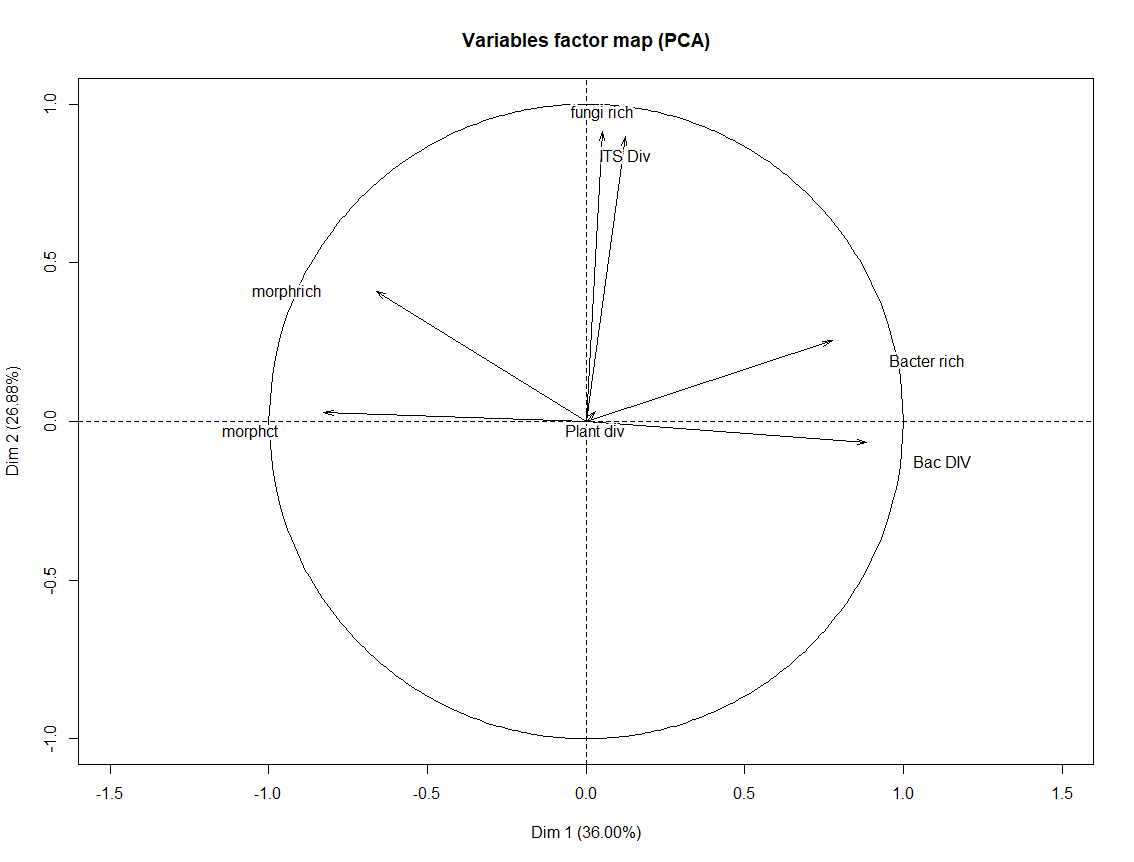






 *4) Principal component analysis (PCA)*

We run a PCA test to see the correlations between each variable. The results show that fungi richness and ITS diversity are highly correlated, as ITS diveristy is DNA sequences from Fungi. This relationship applies to bacteria richness and diversity as well. Then, we plotted PCA results for fungi diversity and different plant species. The graph below shows that two plant species, carpyr and silica, have different fungi diversity. However, in terms of bacteria diversity, we can conclude that there is no difference observed across plant species. Lastly, we plot plant species with elevation. The plot shows that plant species do not differ as the levels of elevation increase.



**Square-root of Eigenvalues – Correlations for PCAs**

Dim.1 Dim.2 Dim.3 Dim.4 Dim.5

Plant div 0.02958358 0.03209319 0.93473428 0.3377101 -0.02631865

morphct -0.82667094 0.02872291 -0.17309202 0.3476817 0.36107613

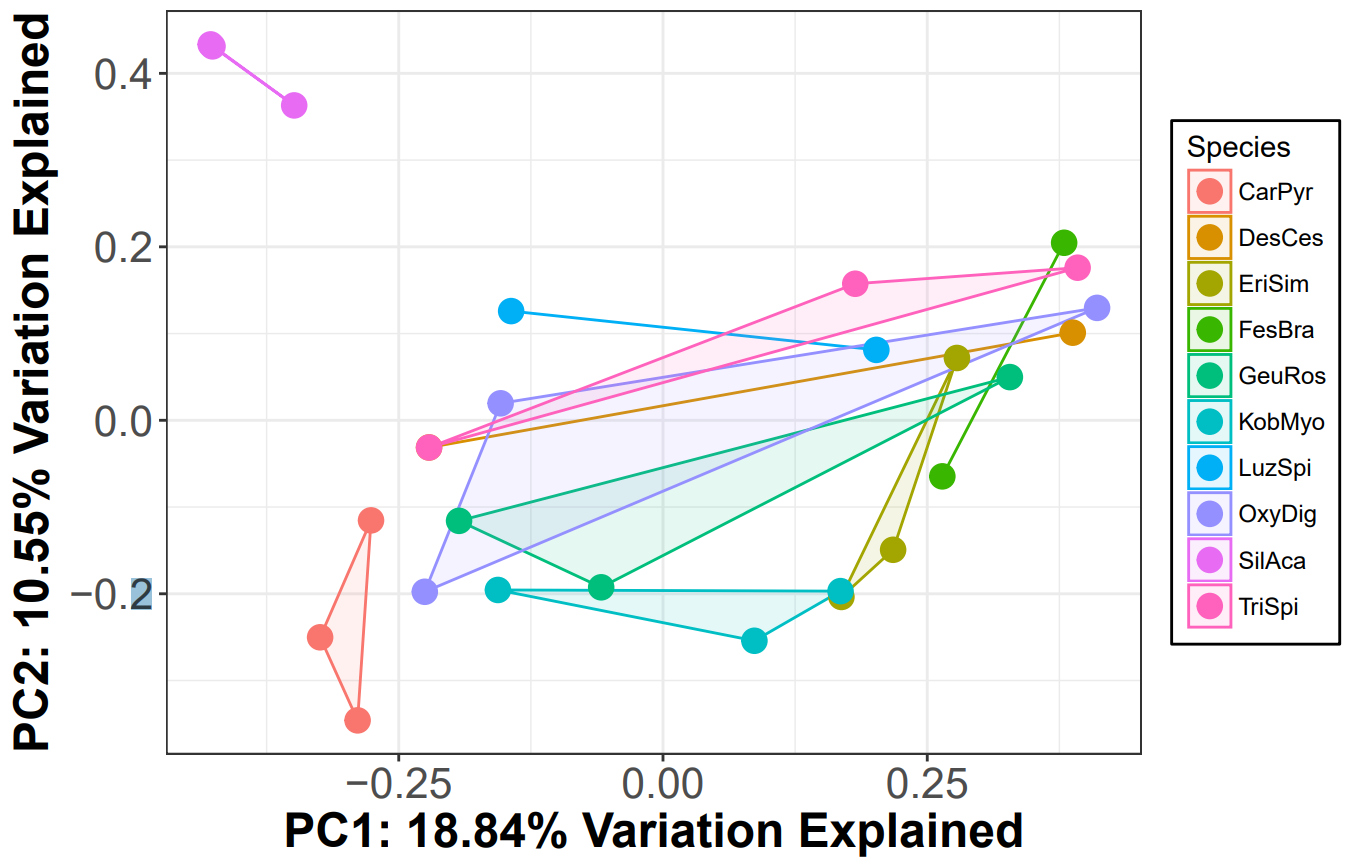
morphrich -0.66170553 0.41035095 -0.28362638 0.3750912 -0.39104858

fungi rich 0.05070126 0.91418595 0.16942430 -0.0804331 -0.09641161

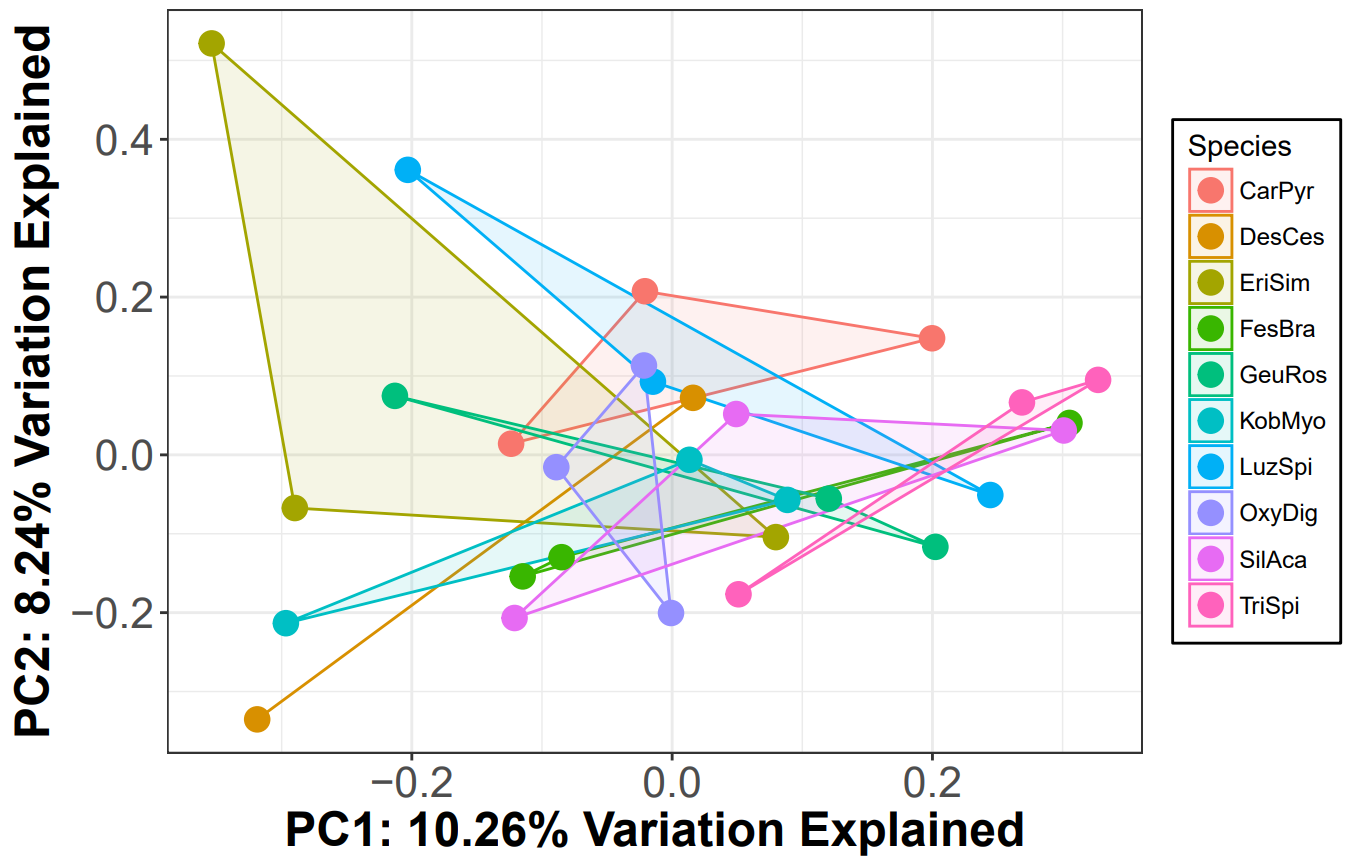
Bacter rich 0.77474722 0.25350463 -0.19095768 0.4535710 0.23329968

ITS Div 0.12301760 0.89825031 -0.03489109 -0.2233794 0.18700178

Bac DIV 0.88322420 -0.06671753 -0.24317037 0.2329900 -0.17929068

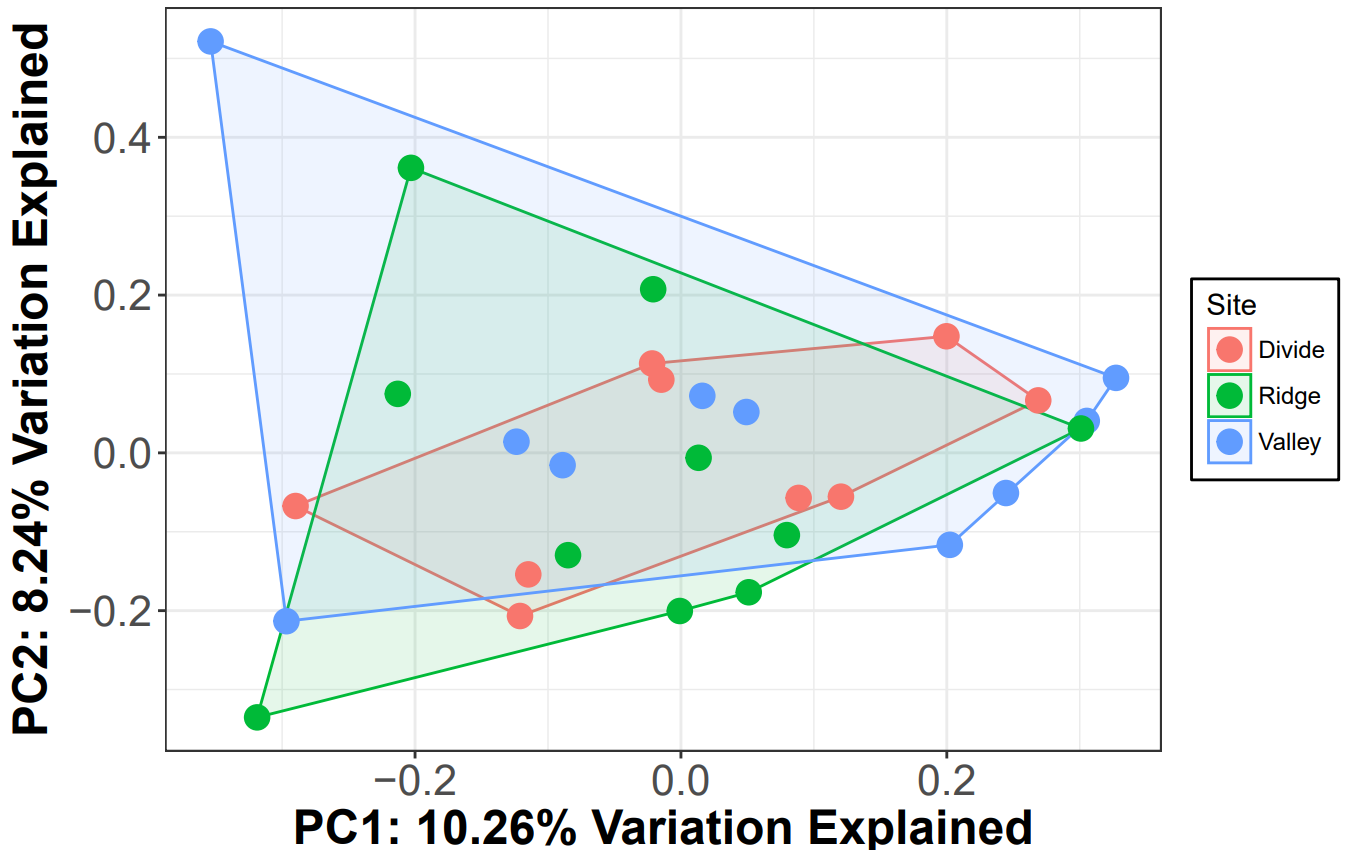
***Variation for ITS Diversity:***

***Variations for Bacterial Diversity:***



***Variations of Plant Species with Levels of Elevation:***

Each dot on the graph below represents a plant species.



1. **Conclusion**

From the statistical analysis results (MANOVA/ANOVAs, LINEAR MIXED MODELS, AND PCAs), we can conclude statistical significance in two variables. From the Linear mixed models test, we are able to conclude statistical significance with a p-value of 0.09 for fungi richness, which is almost in the range of significance. The boxplots confirm our conclusion, indicating that as the levels of elevation increase, fungi richness decreases. From the PCA tests, we are able to visualize the correlations of each variable. Here, we can conclude that two plant species have different fungi diversity. However, we can only reject our null hypothesis for one variable, fungi richness.