

Ancient Animals' Role Analysis Final Report

Keliang Xu, Shuting Li, Tao He

3/21/2022

Abstract

This study is an analysis of the role of animals in the ancient economy at Teotihuacan Mexico. The purpose of this report is to provide statistical analysis to this study with the main goal of analyzing whether there exist significant differences in animals' diets from different sites. We did some exploratory data analysis to see the distribution of the isotope levels and used two-sample t-tests and Mann-Whitney U Test to compare the average levels of isotopes in different sites. Finally, we concluded that there are no significant differences in animals' diets from BAOX and PATT.

Introduction

The isotopes we focus on are stable carbon and nitrogen isotopes, which can reflect the level of human-grown plants and protein in the diet respectively. Since isotopes can enter animals' bones by eating, by comparing the average levels of carbon and nitrogen isotopes from ancient animals' bones between two sites of ancient Mexico, we can get insight into whether animals from these two sites had similar diets.

Data Cleaning and Processing

Firstly, to clean this big data set, and also keep the same as the client's data, we dropped 4 useless observations, which IDs are MC381, MC389, MC393, MC76.

Secondly, to help focus on the main question, we removed all irrelevant variables, the final variables of interest are "Taxa", "Site", "d13C_VPDB", "ap_13C_VPDB" and "d15N_air". Then we combined "d13C_VPDB", "ap_13C_VPDB" and "d15N_air" into one variable which named "isotope level".

Data Description

Table1: Data Explanation

Variable	Explanation	Description
Taxa	Animal Species	turkey, cottontails, deer, hare
site	Ancient Mexico Cities	BAOX, PATT
isotope level	Carbon and Nitrogen level	d13C_VPDB, ap_13C_VPDB, d15N_air

EDA

Summary of Data:

Table 2: Summary of Data

Taxa	Site	variable	min	median	mean	max	sd	count
Cottontail	BAOX	d13C_VPDB	-20.49	-15.69	-16.26	-11.53	3.14	9
Cottontail	BAOX	d15N_air	3.59	6.58	6.99	13.60	2.86	9
Cottontail	BAOX	ap_13C_VPDB	-12.80	-7.69	-8.71	-5.50	2.54	9
Cottontail	Tlajinga	d13C_VPDB	-21.14	-16.92	-16.66	-9.90	3.80	8
Cottontail	Tlajinga	d15N_air	2.77	3.88	4.30	7.00	1.53	8
Cottontail	Tlajinga	ap_13C_VPDB	-13.26	-8.41	-9.25	-4.58	3.03	8
Deer	BAOX	d13C_VPDB	-20.63	-19.38	-18.83	-16.58	1.37	9
Deer	BAOX	d15N_air	3.50	4.97	5.27	6.91	1.22	9
Deer	BAOX	ap_13C_VPDB	-12.84	-10.43	-10.12	-6.51	1.78	9
Deer	Tlajinga	d13C_VPDB	-22.08	-18.92	-19.25	-15.79	2.47	9
Deer	Tlajinga	d15N_air	2.93	4.71	4.67	6.48	1.20	9
Deer	Tlajinga	ap_13C_VPDB	-15.67	-10.94	-10.96	-8.29	2.57	8
Hare	BAOX	d13C_VPDB	-19.51	-16.97	-16.16	-11.19	2.64	11
Hare	BAOX	d15N_air	3.40	6.17	6.28	8.44	1.58	11
Hare	BAOX	ap_13C_VPDB	-11.74	-8.38	-7.98	-4.52	2.38	11
Hare	Tlajinga	d13C_VPDB	-20.92	-14.02	-15.06	-11.85	2.89	8
Hare	Tlajinga	d15N_air	1.69	5.64	5.48	8.29	2.43	8
Hare	Tlajinga	ap_13C_VPDB	-13.05	-7.39	-8.16	-5.58	2.52	8
Turkey	BAOX	d13C_VPDB	-15.26	-12.68	-12.90	-10.35	1.86	8
Turkey	BAOX	d15N_air	5.50	8.26	8.32	10.75	1.63	8
Turkey	BAOX	ap_13C_VPDB	-7.13	-4.14	-4.49	-2.66	1.36	8
Turkey	Tlajinga	d13C_VPDB	-14.21	-11.95	-11.91	-9.81	1.73	6
Turkey	Tlajinga	d15N_air	5.27	7.61	7.31	8.53	1.26	6
Turkey	Tlajinga	ap_13C_VPDB	-7.14	-4.96	-5.39	-4.05	1.21	6

In this form, we can see the mean, median, maximum, minimum, and standard deviation values of isotopes for each kind of animal in each site. The result is exactly same with client's result.

We can see almost all average isotope levels are similar between different sites, except Nitrogen values in cottontail.

Boxplot

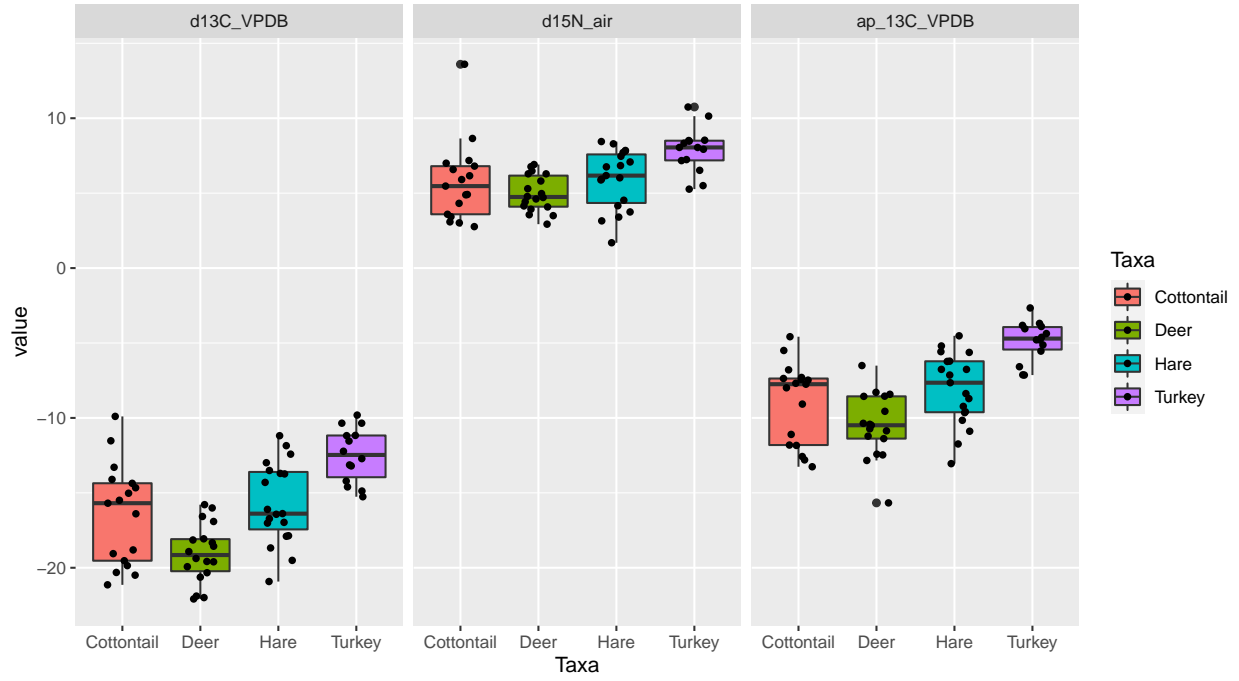


Figure 1: Isotope Levels vs. Taxa

In figure 1, we compared isotopes levels in 4 taxa. It showed the level of d15N_air is less diverse than the other 2 isotopes.

Then we compared isotopes in two sites for each kind of animal. Figure 2 and figure 3 showed that the median level of d13C_VPDB and d15N_air in BAOX is a little higher than Tlajinga of cottontail. Similarly, the median level of d13C_VPDB and d15N_air of hare also show differences in the boxplot.

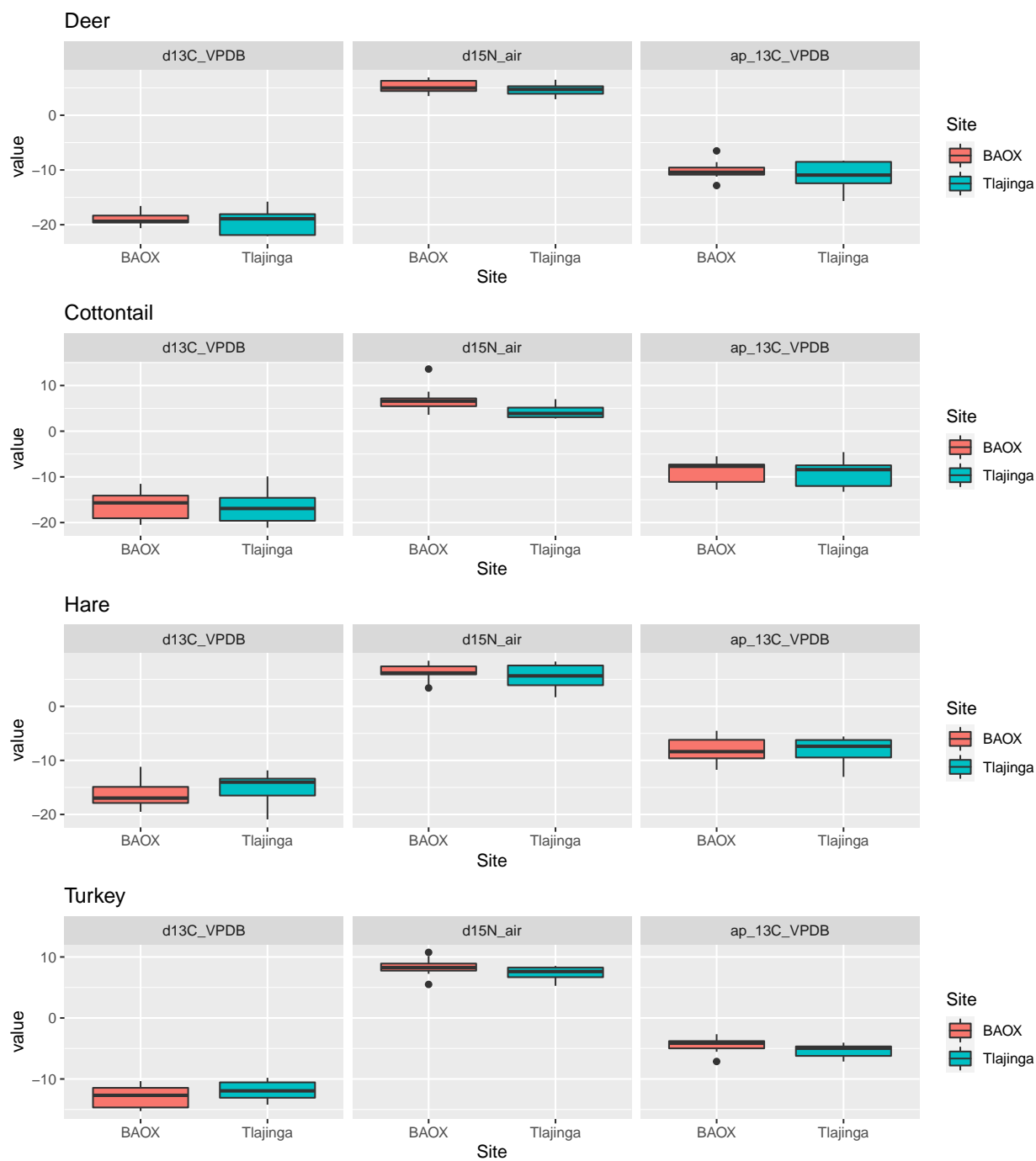


Figure 2: Isotope Levels vs. Site

Testing

Two-Sample T-tests

After finding We plan to try some statistic tests or some regression to compare the isotopes level in two sites more precisely. We would like to use the T-tests to find the difference of diets between two sites.

Firstly, we checked the normality.

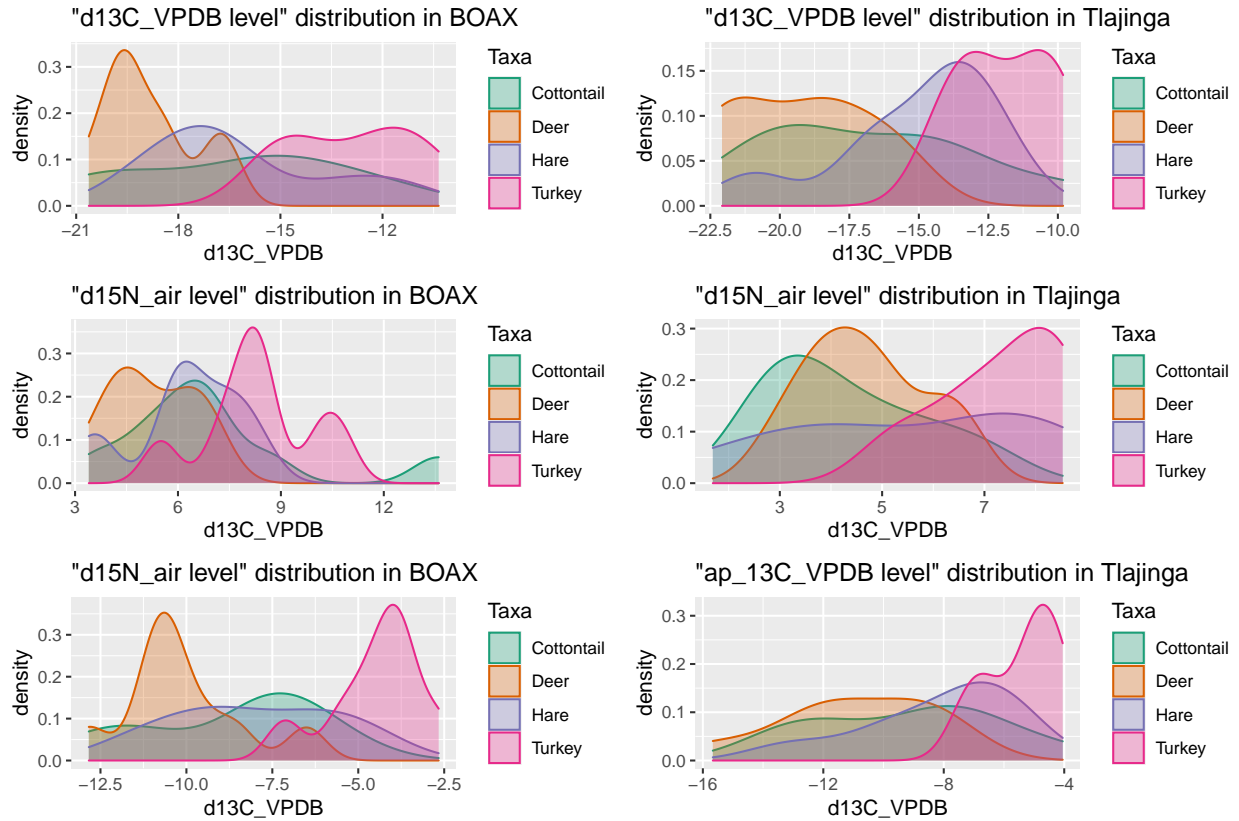


Figure 3: Distribution of Three Isotope Levels For Each Species In Each Place

It is difficult to observe from the figures whether they follow the normal distribution. But comparing the distributions of the two places, most of them follow different distributions. It is reasonable since our sample size is too small.

Therefore, we also plot the Q-Q plots for checking the normality. The purpose of the quantile-quantile (Q-Q) plot is to show if two data sets come from the same distribution. Plotting the first data set's quantiles along the x-axis and plotting the second data set's quantiles along the y-axis is how the plot constructed.

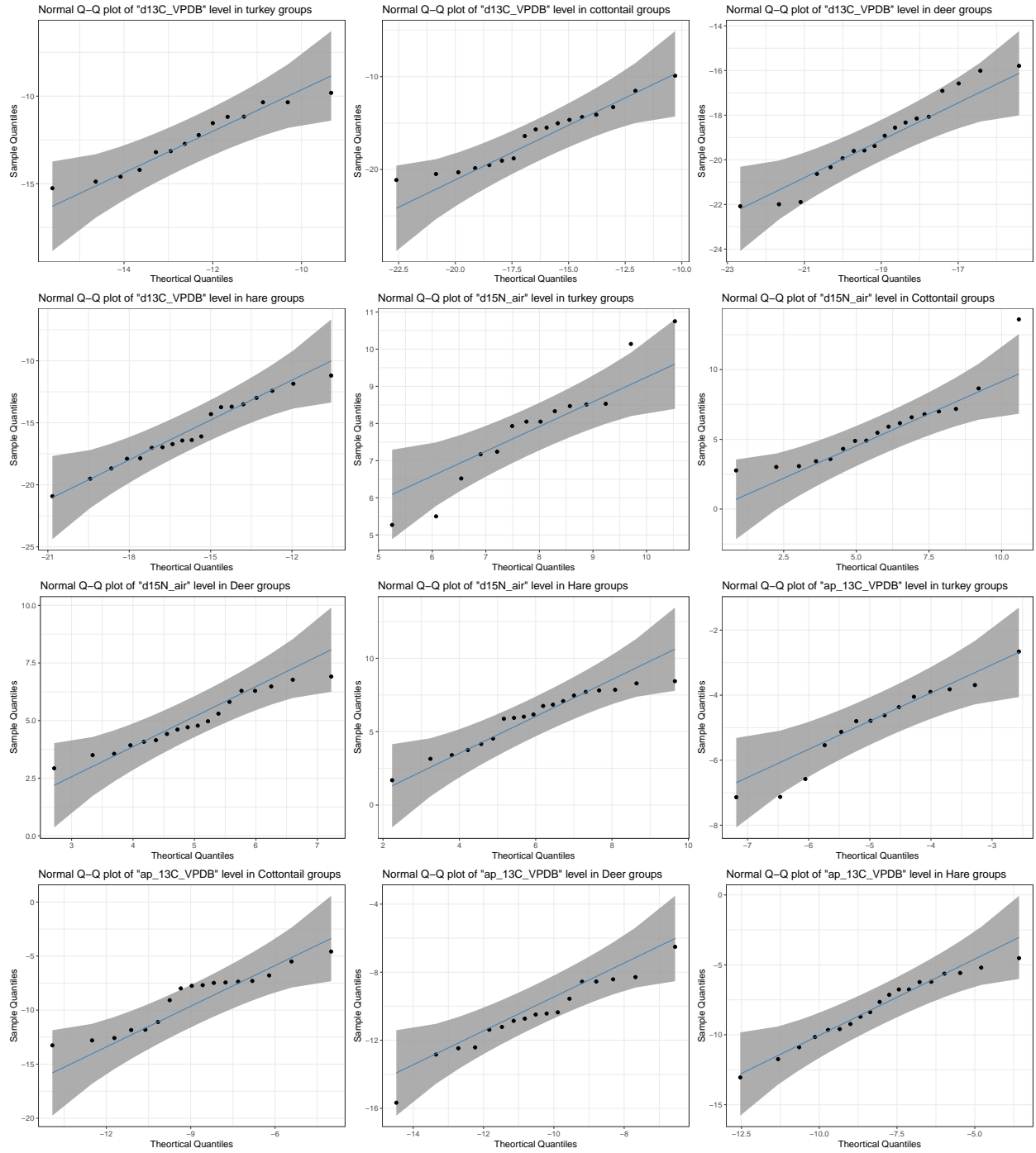


Figure 4: Normal Q-Q Plots in Each Species Group

From the normal Q-Q plots, except “d15N_{air}” level in turkey groups, as almost all the points fall approximately along this reference line, we can assume normality to some extent.

After we checked the distribution of isotope levels for each animal in each space, the sample size is too small to identify if they follow normal distributions. We decided to do some shapiro.test.

Shapiro.test is a test of normality. If the p-value from the result is greater than the significance level (e.g. 0.05), we can assume the normality. Then, from the results of shapiro.test(), we can regard all groups have the normal distribution at 95% confidence interval.

Since we have four different species and three isotope levels, we are supposed to have till 12 tests, which will exist the multiple testing problem. Therefore, we calculate the proper significant level so that the probability of observing at least one significant result due to chance remains below we desired significance level. The final significant level we set is 0.01. Moreover, We assume the null hypothesis is the isotope level for each species between the two places are the same, while the alternative hypothesis is the isotope level for each species between the two places are not the same.

At the 99% confidence interval, we can conclude that: all the isotope levels are the same between BAOX and PATT.

Mann Whitney U Test

Mann Whitney U Test with R Package

We found that description of R function wilcox.test from R package stats: Performs one- and two-sample Wilcoxon tests on vectors of data; the latter is also known as ‘Mann-Whitney’ test. That was exactly what we need. And the same as T-test, we use 99% confidence level and same null hypothesis to test.

We can carefully look at these results. Since these p-values are not less than 0.01, we fail to reject the null hypothesis.

Check Mann-Whitney U Test with theory

We are concerned that small data size (about n=10) may affect the accuracy of the R package function. Therefore, we wrote a program to test the data in full accordance with the Mann-Whitney U Test theory with a significance level of .01. In the program, if the return value is True we can reject the null hypothesis, and if the return value is False, we fail to reject the null hypothesis.

Based on all return values being false, we fail to reject the null hypothesis. Since we fail to reject the null hypothesis, we do not have sufficient evidence to say that the “Carbon and Nitrogen Level” of each animals in two sites are different.

Discussion

When we checked the normality of the data set, we use the shapiro.test() to confirm the distribution of those isotope levels. However, the disadvantage of this method is we can not suppose the distribution of isotope level has not the significant difference with normal distribution at 95% confidence interval. We can not make sure that they are exactly following the normal distribution. Similarly, when we do the t-tests, we conclude that the isotope levels for each species are the same between the two close places. We still can not say they are the same at 100% level. We can only say that at 95% confidence interval, the diets between BAOX and PATT are the same.

Actually, we firstly use the most common significant level with 0.05 to do the two-sample t-tests. However, the results show that there exists a significant one among the 12 tests. At first blush, this doesn’t seem like

a bad outcome. However, considering a case where we have 12 hypotheses to test, and a significance level of 0.05, there exists the probability of observing at least one significant result just due to chance.

$$\begin{aligned} p(\text{at least one significant result}) &= 1 - (\text{no significant results}) \\ &= 1 - (1 - \alpha)^n \\ &= 1 - (1 - \alpha)^{12} \end{aligned}$$

If we set the significant level with 0.05,

$$\begin{aligned} p(\text{at least one significant result}) &= 1 - (1 - 0.05)^{12} \\ &\approx 0.46, \end{aligned}$$

the probability of rejecting the null hypothesis by chance will be 0.46, almost a half. Then, we decided to set the significant level with 0.01,

$$\begin{aligned} p(\text{at least one significant result}) &= 1 - (1 - 0.01)^{12} \\ &\approx 0.11, \end{aligned}$$

which significantly reduces chance to fail to reject the null hypothesis. Therefore, we chose the 0.01 significant level in our final work.

Conclusion

We checked all client's works, including the summary of data and box plots. Our results are consistent with hers. Then, after we did two tests to statistically analyse the main problem we want to solve, we concluded that there were no significant differences in animals' diets from BAOX and PATT.

Citation

Somerville, A.D., Sugiyama, N., Manzanilla, L.R. et al. Leporid management and specialized food production at Teotihuacan: stable isotope data from cottontail and jackrabbit bone collagen. *Archaeol Anthropol Sci* 9, 83–97 (2017). <https://doi.org/10.1007/s12520-016-0420-2>

Appendix

Test code

Two-Sample T-tests are as followings:

“d13C_VPDB” level

```
# Assumption  
# H0 :  $\nu_1 = \nu_2$   
# Ha :  $\nu_1 \neq \nu_2$ 
```

```
# "d13C_VPDB" level in turkey groups  
with(test.turkey.d13C, shapiro.test(value[Site == "BAOX"])) # p = 0.4606
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: value[Site == "BAOX"]  
## W = 0.9237, p-value = 0.4606
```

```
with(test.turkey.d13C, shapiro.test(value[Site == "Tlajinga"])) # p = 0.7027
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: value[Site == "Tlajinga"]  
## W = 0.94538, p-value = 0.7027
```

```
# From the output, the two p-values are greater than the significance level 0.05  
# implying that the distribution of the data are not significantly different  
# from the normal distribution. In other words, we can assume the normality.
```

```
t.test(value ~ Site, data = test.turkey.d13C, conf.level = 0.99, var.equal = TRUE)
```

```
##  
## Two Sample t-test  
##  
## data: value by Site  
## t = -1.0102, df = 12, p-value = 0.3323  
## alternative hypothesis: true difference in means between group BAOX and group Tlajinga is not equal to 0  
## 99 percent confidence interval:  
## -3.968245 1.995745  
## sample estimates:  
## mean in group BAOX mean in group Tlajinga  
## -12.89625 -11.91000
```

```
# p = 0.3323  
# Not reject!
```

```

# "d13C_VPDB" level in Cottontail groups
with(test.Cottontail.d13C, shapiro.test(value[Site == "BAOX"])) # p = 0.5969

##
## Shapiro-Wilk normality test
##
## data: value[Site == "BAOX"]
## W = 0.94142, p-value = 0.5969

with(test.Cottontail.d13C, shapiro.test(value[Site == "Tlajinga"])) # p = 0.3941

##
## Shapiro-Wilk normality test
##
## data: value[Site == "Tlajinga"]
## W = 0.91546, p-value = 0.3941

# From the output, the two p-values are greater than the significance level 0.05
# implying that the distribution of the data are not significantly different
# from the normal distribution. In other words, we can assume the normality.
t.test(value ~ Site, data = test.Cottontail.d13C, conf.level = 0.99, var.equal = TRUE)

##
## Two Sample t-test
##
## data: value by Site
## t = 0.23657, df = 15, p-value = 0.8162
## alternative hypothesis: true difference in means between group BAOX and group Tlajinga is not equal
## 99 percent confidence interval:
## -4.558434 5.354267
## sample estimates:
## mean in group BAOX mean in group Tlajinga
## -16.26333 -16.66125

# p = 0.8162
# Not reject!

# "d13C_VPDB" level in Deer groups
with(test.Deer.d13C, shapiro.test(value[Site == "BAOX"])) # p = 0.3914

##
## Shapiro-Wilk normality test
##
## data: value[Site == "BAOX"]
## W = 0.9199, p-value = 0.3914

with(test.Deer.d13C, shapiro.test(value[Site == "Tlajinga"])) # p = 0.2177

##
## Shapiro-Wilk normality test
##
## data: value[Site == "Tlajinga"]
## W = 0.89371, p-value = 0.2177

```

```

# From the output, the two p-values are greater than the significance level 0.05
# implying that the distribution of the data are not significantly different
# from the normal distribution. In other words, we can assume the normality.
t.test(value ~ Site, data = test.Deer.d13C, conf.level = 0.99, var.equal = TRUE)

```

```

##
## Two Sample t-test
##
## data: value by Site
## t = 0.44023, df = 16, p-value = 0.6657
## alternative hypothesis: true difference in means between group BAOX and group Tlajinga is not equal
## 99 percent confidence interval:
## -2.335259 3.164148
## sample estimates:
## mean in group BAOX mean in group Tlajinga
## -18.83333 -19.24778

```

```

# p = 0.6657
# Not reject!

```

```

# "d13C_VPDB" level in Hare groups
with(test.Hare.d13C, shapiro.test(value[Site == "BAOX"])) # p = 0.2707

```

```

##
## Shapiro-Wilk normality test
##
## data: value[Site == "BAOX"]
## W = 0.91386, p-value = 0.2707

```

```

with(test.Hare.d13C, shapiro.test(value[Site == "Tlajinga"])) # p = 0.2447

```

```

##
## Shapiro-Wilk normality test
##
## data: value[Site == "Tlajinga"]
## W = 0.89208, p-value = 0.2447

```

```

# From the output, the two p-values are greater than the significance level 0.05
# implying that the distribution of the data are not significantly different
# from the normal distribution. In other words, we can assume the normality.
t.test(value ~ Site, data = test.Hare.d13C, conf.level = 0.99, var.equal = TRUE)

```

```

##
## Two Sample t-test
##
## data: value by Site
## t = -0.86338, df = 17, p-value = 0.3999
## alternative hypothesis: true difference in means between group BAOX and group Tlajinga is not equal
## 99 percent confidence interval:

```

```
## -4.795489 2.594125
## sample estimates:
## mean in group BAOX mean in group Tlajinga
## -16.15818 -15.05750
```

```
# p = 0.3999
# Not reject!
```

“d15N_{air}” level

```
# "d15Nair" level in turkey groups
with(test.turkey.d15N, shapiro.test(value[Site == "BAOX"])) # p = 0.7717
```

```
##
## Shapiro-Wilk normality test
##
## data: value[Site == "BAOX"]
## W = 0.95604, p-value = 0.7717
```

```
with(test.turkey.d15N, shapiro.test(value[Site == "Tlajinga"])) # p = 0.4532
```

```
##
## Shapiro-Wilk normality test
##
## data: value[Site == "Tlajinga"]
## W = 0.91253, p-value = 0.4532
```

```
# From the output, the two p-values are greater than the significance level 0.05
# implying that the distribution of the data are not significantly different
# from the normal distribution. In other words, we can assume the normality.
```

```
t.test(value ~ Site, data = test.turkey.d15N, conf.level = 0.99, var.equal = TRUE)
```

```
##
## Two Sample t-test
##
## data: value by Site
## t = 1.261, df = 12, p-value = 0.2313
## alternative hypothesis: true difference in means between group BAOX and group Tlajinga is not equal
## 99 percent confidence interval:
## -1.439417 3.463583
## sample estimates:
## mean in group BAOX mean in group Tlajinga
## 8.323750 7.311667
```

```
# p = 0.2313
# Not reject!
```

```

# "d15N_air" level in Cottontail groups
with(test.Cottontail.d15N, shapiro.test(value[Site == "BAOX"])) # p = 0.09437

##
## Shapiro-Wilk normality test
##
## data: value[Site == "BAOX"]
## W = 0.85937, p-value = 0.09437

with(test.Cottontail.d15N, shapiro.test(value[Site == "Tlajinga"])) # p = 0.2703

##
## Shapiro-Wilk normality test
##
## data: value[Site == "Tlajinga"]
## W = 0.89679, p-value = 0.2703

# From the output, the two p-values are greater than the significance level 0.05
# implying that the distribution of the data are not significantly different
#from the normal distribution. In other words, we can assume the normality.
t.test(value ~ Site, data = test.Cottontail.d15N, conf.level = 0.99, var.equal = TRUE)

##
## Two Sample t-test
##
## data: value by Site
## t = 2.3623, df = 15, p-value = 0.0321
## alternative hypothesis: true difference in means between group BAOX and group Tlajinga is not equal
## 99 percent confidence interval:
## -0.6644531 6.0366753
## sample estimates:
## mean in group BAOX mean in group Tlajinga
## 6.991111 4.305000

# p = 0.0321
# Not Reject!

# "d15N_air" level in Deer groups
with(test.Deer.d15N, shapiro.test(value[Site == "BAOX"])) # p = 0.5592

##
## Shapiro-Wilk normality test
##
## data: value[Site == "BAOX"]
## W = 0.93783, p-value = 0.5592

with(test.Deer.d15N, shapiro.test(value[Site == "Tlajinga"])) # p = 0.7696

##
## Shapiro-Wilk normality test
##
## data: value[Site == "Tlajinga"]
## W = 0.9573, p-value = 0.7696

```

```

# From the output, the two p-values are greater than the significance level 0.05
# implying that the distribution of the data are not significantly different
# from the normal distribution. In other words, we can assume the normality.
t.test(value ~ Site, data = test.Deer.d15N, conf.level = 0.99, var.equal = TRUE)

```

```

##
## Two Sample t-test
##
## data: value by Site
## t = 1.0472, df = 16, p-value = 0.3106
## alternative hypothesis: true difference in means between group BAOX and group Tlajinga is not equal
## 99 percent confidence interval:
## -1.067488 2.260822
## sample estimates:
## mean in group BAOX mean in group Tlajinga
## 5.270000 4.673333

```

```

# p = 0.3106
# Not reject!

```

```

# "d15N_air" level in Hare groups
with(test.Hare.d15N, shapiro.test(value[Site == "BAOX"])) # p = 0.3232

```

```

##
## Shapiro-Wilk normality test
##
## data: value[Site == "BAOX"]
## W = 0.92055, p-value = 0.3232

```

```

with(test.Hare.d15N, shapiro.test(value[Site == "Tlajinga"])) # p = 0.4432

```

```

##
## Shapiro-Wilk normality test
##
## data: value[Site == "Tlajinga"]
## W = 0.92163, p-value = 0.4432

```

```

# From the output, the two p-values are greater than the significance level 0.05
# implying that the distribution of the data are not significantly different
# from the normal distribution. In other words, we can assume the normality.
t.test(value ~ Site, data = test.Hare.d15N, conf.level = 0.99, var.equal = TRUE)

```

```

##
## Two Sample t-test
##
## data: value by Site
## t = 0.86085, df = 17, p-value = 0.4013
## alternative hypothesis: true difference in means between group BAOX and group Tlajinga is not equal
## 99 percent confidence interval:

```

```
## -1.872939 3.455666
## sample estimates:
##      mean in group BAOX mean in group Tlajinga
##              6.276364              5.485000
```

```
# p = 0.4013
# Not reject!
```

“ap_13C_VPDB” level

```
# "ap_13C_VPDB" level in turkey groups
with(test.turkey.ap13C, shapiro.test(value[Site == "BAOX"])) # p = 0.5876
```

```
##
## Shapiro-Wilk normality test
##
## data: value[Site == "BAOX"]
## W = 0.9376, p-value = 0.5876
```

```
with(test.turkey.ap13C, shapiro.test(value[Site == "Tlajinga"])) # p = 0.3992
```

```
##
## Shapiro-Wilk normality test
##
## data: value[Site == "Tlajinga"]
## W = 0.90418, p-value = 0.3992
```

From the output, the two p-values are greater than the significance level 0.05 implying that the distribution is normal

```
t.test(value ~ Site, data = test.turkey.ap13C, conf.level = 0.99, var.equal = TRUE)
```

```
##
## Two Sample t-test
##
## data: value by Site
## t = 1.2818, df = 12, p-value = 0.2241
## alternative hypothesis: true difference in means between group BAOX and group Tlajinga is not equal
## 99 percent confidence interval:
## -1.243518 3.041852
## sample estimates:
##      mean in group BAOX mean in group Tlajinga
##      -4.487500      -5.386667
```

```
# p = 0.2241
# Not reject!
```

```
# "ap_13C_VPDB" level in Cottontail groups
with(test.Cottontail.ap13C, shapiro.test(value[Site == "BAOX"])) # p = 0.1675
```

```
##
## Shapiro-Wilk normality test
##
## data:  value[Site == "BAOX"]
## W = 0.88268, p-value = 0.1675
```

```
with(test.Cottontail.ap13C, shapiro.test(value[Site == "Tlajinga"])) # p = 0.4795
```

```
##
## Shapiro-Wilk normality test
##
## data:  value[Site == "Tlajinga"]
## W = 0.9259, p-value = 0.4795
```

```
# From the output, the two p-values are greater than the significance level 0.05
# implying that the distribution of the data are not significantly different
# from the normal distribution. In other words, we can assume the normality.
t.test(value ~ Site, data = test.Cottontail.ap13C, conf.level = 0.99, var.equal = TRUE)
```

```
##
## Two Sample t-test
##
## data:  value by Site
## t = 0.39896, df = 15, p-value = 0.6955
## alternative hypothesis: true difference in means between group BAOX and group Tlajinga is not equal
## 99 percent confidence interval:
##  -3.439573  4.516795
## sample estimates:
##      mean in group BAOX mean in group Tlajinga
##      -8.708889          -9.247500
```

```
# p = 0.6955
# not Reject!
```

```
# "ap_13C_VPDB" level in Deer groups
with(test.Deer.ap13C, shapiro.test(value[Site == "BAOX"])) # p = 0.54
```

```
##
## Shapiro-Wilk normality test
##
## data:  value[Site == "BAOX"]
## W = 0.93596, p-value = 0.54
```

```
with(test.Deer.ap13C, shapiro.test(value[Site == "Tlajinga"])) # p = 0.3121
```

```
##
## Shapiro-Wilk normality test
##
## data:  value[Site == "Tlajinga"]
## W = 0.90374, p-value = 0.3121
```



```

# From the output, the two p-values are greater than the significance level 0.05
# implying that the distribution of the data are not significantly different
# from the normal distribution. In other words, we can assume the normality.
t.test(value ~ Site, data = test.Deer.ap13C, conf.level = 0.99, var.equal = TRUE)

```

```

##
## Two Sample t-test
##
## data: value by Site
## t = 0.79213, df = 15, p-value = 0.4406
## alternative hypothesis: true difference in means between group BAOX and group Tlajinga is not equal
## 99 percent confidence interval:
## -2.291227 3.975950
## sample estimates:
## mean in group BAOX mean in group Tlajinga
## -10.11889 -10.96125

```

```

# p = 0.4406
# Not reject!

```

```

# "ap_13C_VPDB" level in Hare groups
with(test.Hare.ap13C, shapiro.test(value[Site == "BAOX"])) # p = 0.7399

```

```

##
## Shapiro-Wilk normality test
##
## data: value[Site == "BAOX"]
## W = 0.95748, p-value = 0.7399

```

```

with(test.Hare.ap13C, shapiro.test(value[Site == "Tlajinga"])) # p = 0.2754

```

```

##
## Shapiro-Wilk normality test
##
## data: value[Site == "Tlajinga"]
## W = 0.8977, p-value = 0.2754

```

```

# From the output, the two p-values are greater than the significance level 0.05
# implying that the distribution of the data are not significantly different
# from the normal distribution. In other words, we can assume the normality.

```

```

t.test(value ~ Site, data = test.Hare.ap13C, conf.level = 0.99, var.equal = TRUE)

```

```

##
## Two Sample t-test
##
## data: value by Site
## t = 0.15129, df = 17, p-value = 0.8815
## alternative hypothesis: true difference in means between group BAOX and group Tlajinga is not equal
## 99 percent confidence interval:

```

```
## -3.117610 3.461019
## sample estimates:
##      mean in group BAOX mean in group Tlajinga
##      -7.984545      -8.156250
```

```
# p = 0.8815
# Not reject!
```

Mann Whitney U Test with R Package

Mann-Whitney U Test for deer.

```
##
## Wilcoxon rank sum exact test
##
## data: value by Site
## W = 43, p-value = 0.8633
## alternative hypothesis: true location shift is not equal to 0
```

```
##
## Wilcoxon rank sum exact test
##
## data: value by Site
## W = 41, p-value = 0.673
## alternative hypothesis: true location shift is not equal to 0
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: value by Site
## W = 51.5, p-value = 0.3536
## alternative hypothesis: true location shift is not equal to 0
```

Mann-Whitney U Test for Cottontail.

```
##
## Wilcoxon rank sum exact test
##
## data: value by Site
## W = 38, p-value = 0.8884
## alternative hypothesis: true location shift is not equal to 0
```

```
##
## Wilcoxon rank sum exact test
##
## data: value by Site
## W = 43, p-value = 0.5414
## alternative hypothesis: true location shift is not equal to 0
```

```
##
## Wilcoxon rank sum exact test
```

```
##
## data:  value by Site
## W = 60, p-value = 0.02065
## alternative hypothesis: true location shift is not equal to 0
```

Mann-Whitney U Test for Hare.

```
##
## Wilcoxon rank sum exact test
##
## data:  value by Site
## W = 32, p-value = 0.3511
## alternative hypothesis: true location shift is not equal to 0
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data:  value by Site
## W = 45, p-value = 0.9671
## alternative hypothesis: true location shift is not equal to 0
```

```
##
## Wilcoxon rank sum exact test
##
## data:  value by Site
## W = 50, p-value = 0.6574
## alternative hypothesis: true location shift is not equal to 0
```

Mann-Whitney U Test for Turkey.

```
##
## Wilcoxon rank sum test with continuity correction
##
## data:  value by Site
## W = 15.5, p-value = 0.3012
## alternative hypothesis: true location shift is not equal to 0
```

```
##
## Wilcoxon rank sum exact test
##
## data:  value by Site
## W = 36, p-value = 0.1419
## alternative hypothesis: true location shift is not equal to 0
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data:  value by Site
## W = 32.5, p-value = 0.3012
## alternative hypothesis: true location shift is not equal to 0
```

Check Mann-Whitney U Test with theory

Mann-Whitney U Test for deer.

```
## [1] "d13C_VPDB:"  
## [1] FALSE  
## [1] "d15N_air:"  
## [1] FALSE  
## [1] "ap_13C_VPDB:"  
## [1] FALSE
```

Mann-Whitney U Test for Cottontail.

```
## [1] "d13C_VPDB:"  
## [1] FALSE  
## [1] "d15N_air:"  
## [1] FALSE  
## [1] "ap_13C_VPDB:"  
## [1] FALSE
```

Mann-Whitney U Test for Hare.

```
## [1] "d13C_VPDB:"  
## [1] FALSE  
## [1] "d15N_air:"  
## [1] FALSE  
## [1] "ap_13C_VPDB:"  
## [1] FALSE
```

Mann-Whitney U Test for Turkey.

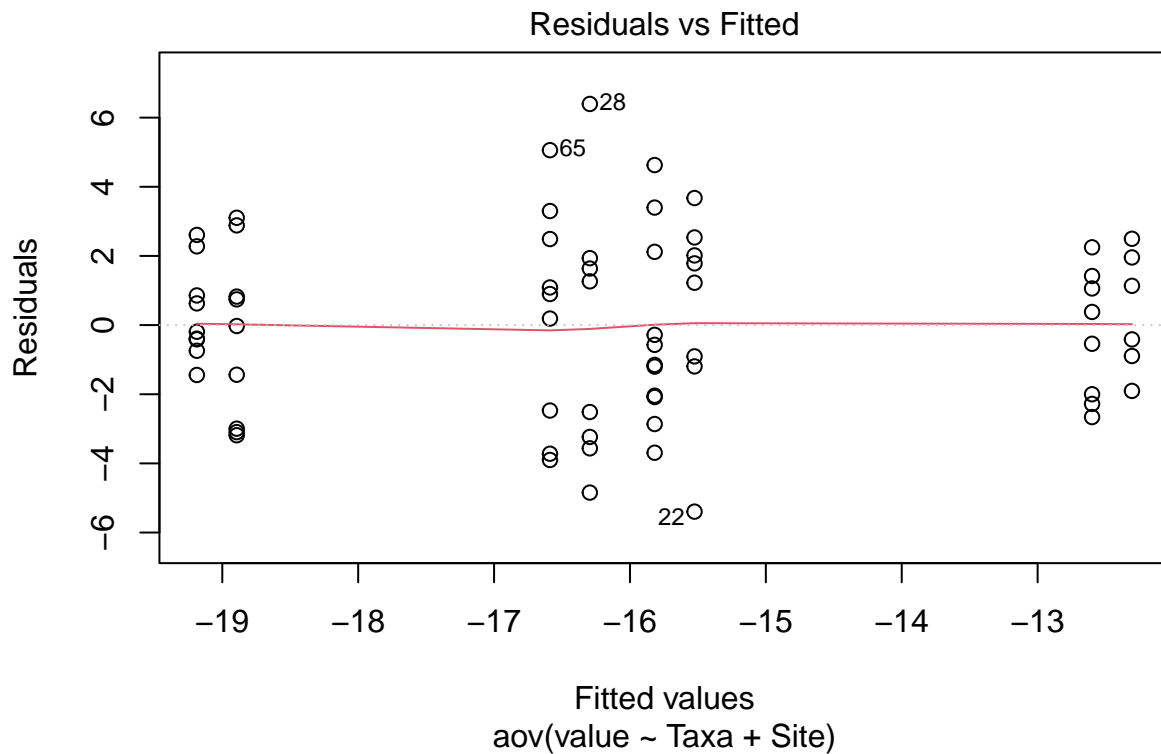
```
## [1] "d13C_VPDB:"  
## [1] FALSE  
## [1] "d15N_air:"  
## [1] FALSE  
## [1] "ap_13C_VPDB:"  
## [1] FALSE
```

Two-Way ANOVA Test

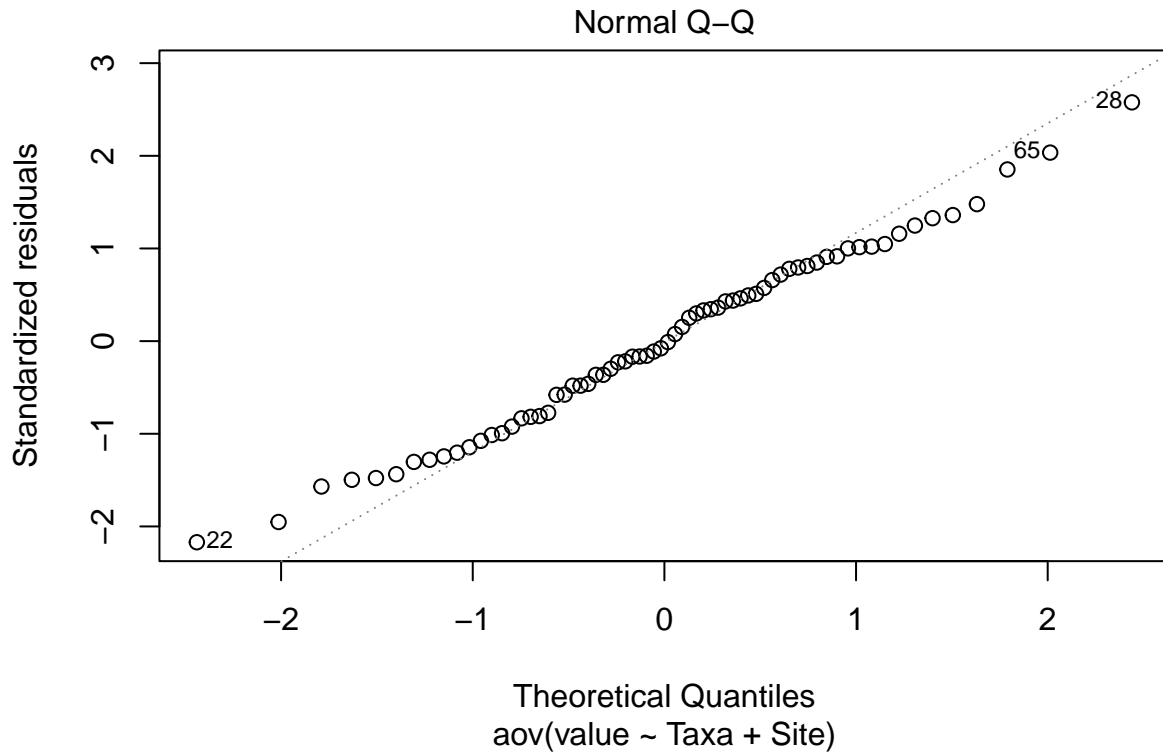
```
data_dc<-aggdata[which(aggdata$variable=="d13C_VPDB"),]  
data_n<-aggdata[which(aggdata$variable=="d15N_air"),]  
data_apc<-aggdata[which(aggdata$variable=="ap_13C_VPDB"),]  
  
res.aovdc<-aov(value ~ Taxa + Site, data = data_dc)  
summary(res.aovdc)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)        
## Taxa       3  345.0   114.99   17.262 2.72e-08 ***  
## Site       1    1.4     1.45    0.217   0.643        
## Residuals 63  419.7     6.66                
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
plot(res.aovdc, 1)
```



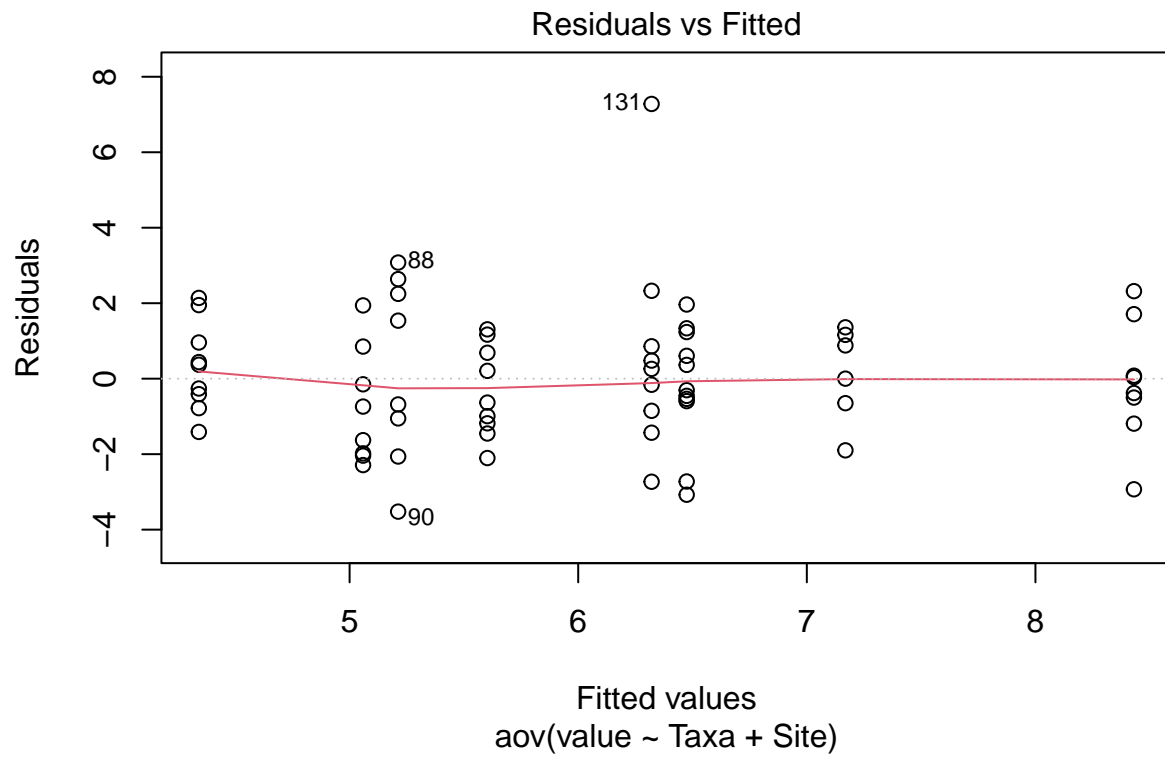
```
plot(res.aovdc, 2)
```



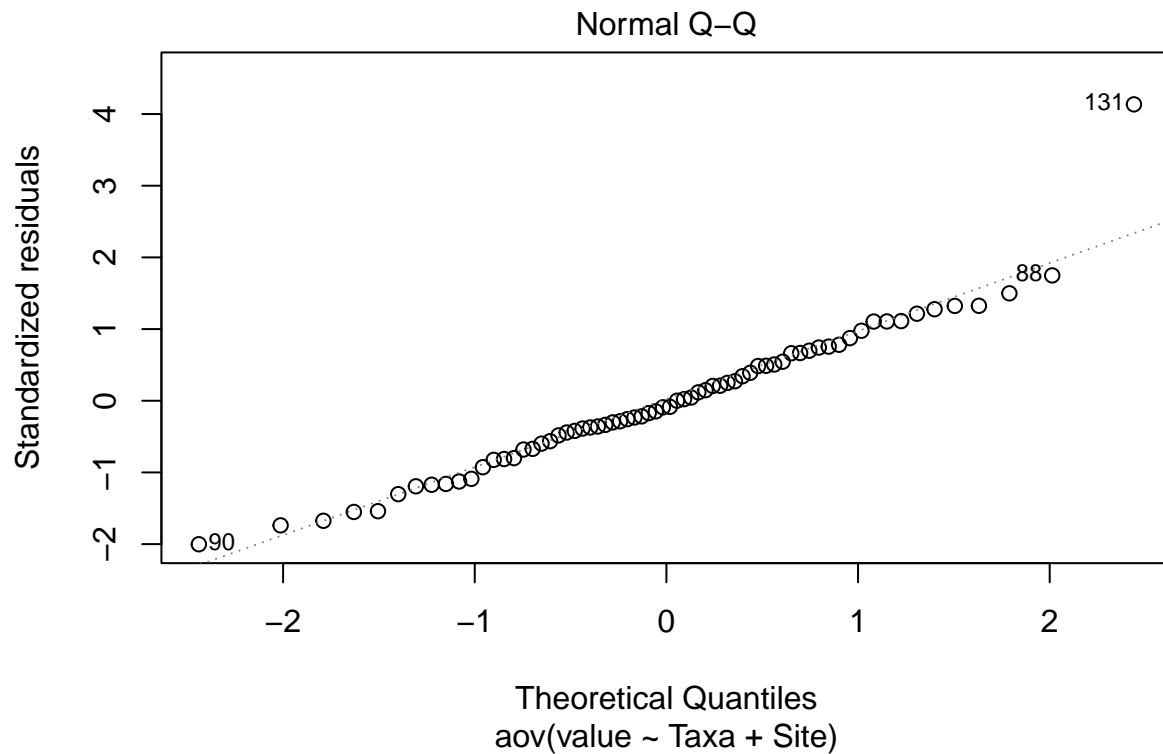
```
res.aovn<-aov(value ~ Taxa + Site, data = data_n)
summary(res.aovn)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Taxa       3  70.30   23.433    7.017 0.000381 ***
## Site       1  26.75   26.751    8.011 0.006231 **
## Residuals 63 210.37    3.339
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
plot(res.aovn, 1)
```



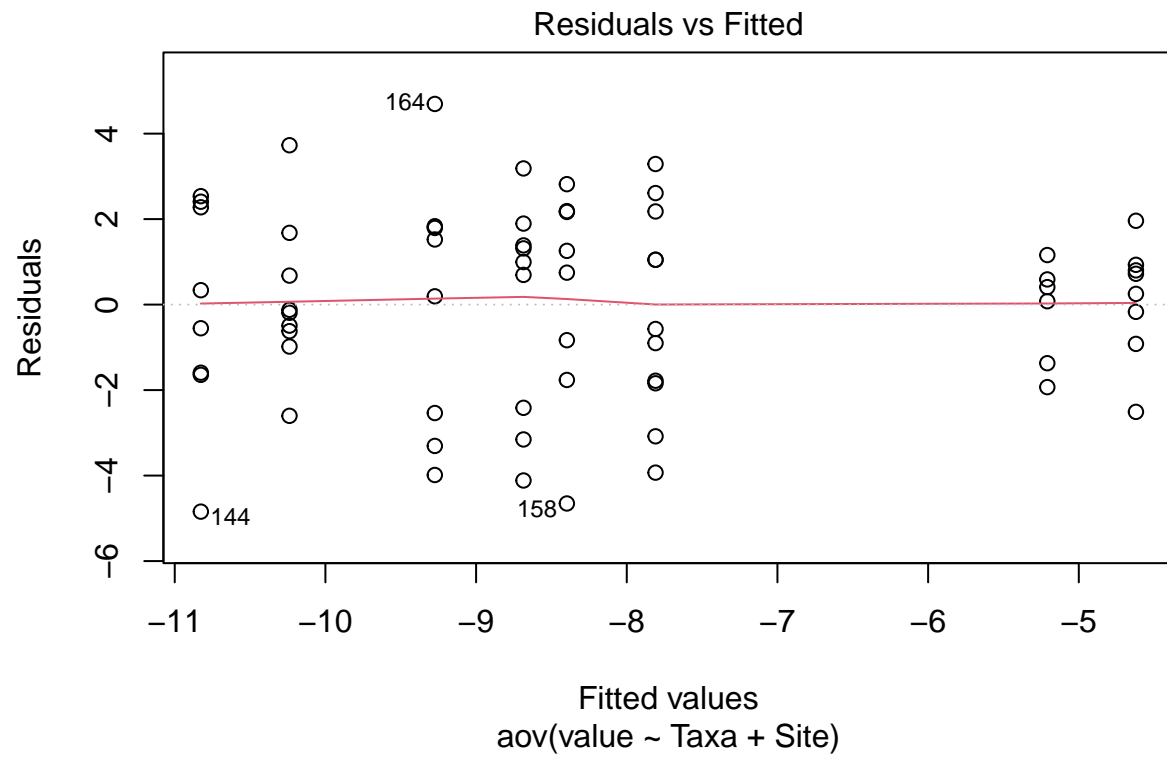
```
plot(res.aovn, 2)
```



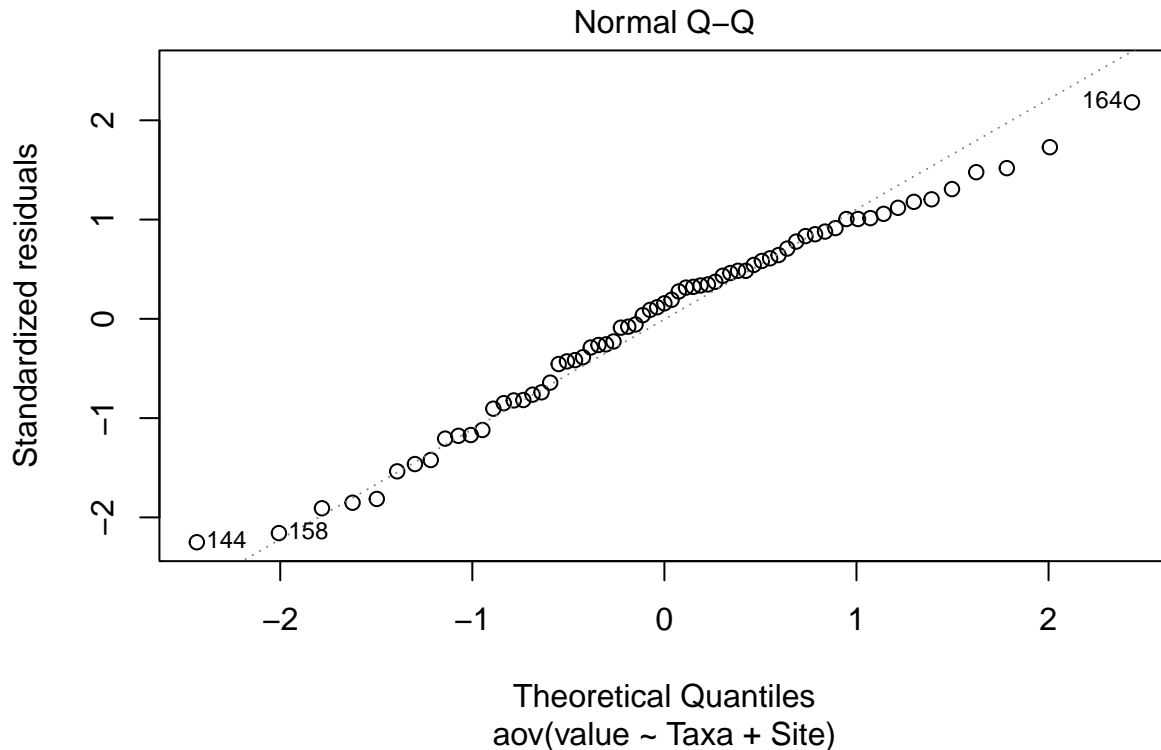
```
res.aovapc<-aov(value ~ Taxa + Site, data = data_apc)
summary(res.aovapc)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Taxa       3 256.24   85.41  17.043 3.48e-08 ***
## Site       1   5.72    5.72   1.142   0.289
## Residuals 62 310.72    5.01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 1 observation deleted due to missingness
```

```
plot(res.aovapc, 1)
```

```
plot(res.aovapc, 2)
```



```
TukeyHSD(res.aovdc)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = value ~ Taxa + Site, data = data_dc)
##
## $Taxa
##          diff          lwr          upr      p adj
## Deer-Cottontail -2.5899673 -4.8934550 -0.2864797 0.0215223
## Hare-Cottontail  0.7558514 -1.5180026  3.0297054 0.8166165
## Turkey-Cottontail 3.9770168  1.5188862  6.4351474 0.0003836
## Hare-Deer        3.3458187  1.1055486  5.5860889 0.0011540
## Turkey-Deer      6.5669841  4.1398862  8.9940821 0.0000000
## Turkey-Hare      3.2211654  0.8221738  5.6201571 0.0040810
##
## $Site
##          diff          lwr          upr      p adj
## Tlajinga-BA0X 0.2922179 -0.9635901  1.548026 0.6435346
```

```
TukeyHSD(res.aovn)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
```

```
##
## Fit: aov(formula = value ~ Taxa + Site, data = data_n)
##
## $Taxa
##           diff      lwr      upr      p adj
## Deer-Cottontail -0.7553922 -2.3862969 0.8755126 0.6151741
## Hare-Cottontail  0.2160991 -1.3938245 1.8260227 0.9846190
## Turkey-Cottontail 2.1629412  0.4225469 3.9033355 0.0089706
## Hare-Deer        0.9714912 -0.6146545 2.5576370 0.3769832
## Turkey-Deer      2.9183333  1.1999106 4.6367560 0.0001838
## Turkey-Hare      1.9468421  0.2483191 3.6453651 0.0184089
##
## $Site
##           diff      lwr      upr      p adj
## Tlajinga-BAOX -1.25665 -2.145782 -0.3675189 0.0063363
```

```
TukeyHSD(res.aovapc)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = value ~ Taxa + Site, data = data_apc)
##
## $Taxa
##           diff      lwr      upr      p adj
## Deer-Cottontail -1.5529412 -3.5801451 0.4742628 0.1909102
## Hare-Cottontail  0.9055108 -1.0676246 2.8786463 0.6219146
## Turkey-Cottontail 4.0894958  1.9564544 6.2225372 0.0000232
## Hare-Deer        2.4584520  0.4853166 4.4315875 0.0087752
## Turkey-Deer      5.6424370  3.5093956 7.7754784 0.0000000
## Turkey-Hare      3.1839850  1.1022614 5.2657085 0.0008498
##
## $Site
##           diff      lwr      upr      p adj
## Tlajinga-BAOX -0.5871218 -1.686554 0.5123107 0.2898869
```

```
library(AICcmodavg)

two.way<-aov(value ~ Taxa + Site, data = data_dc)
interaction<-aov(value ~ Taxa * Site, data = data_dc)

model.set<-list(two.way, interaction)
model.names <- c("two.way", "interaction")

aictab(model.set, modnames = model.names)
```

```
##
## Model selection based on AICc:
##
##           K   AICc Delta_AICc AICcWt Cum.Wt      LL
## two.way    6 330.11      0.00   0.96   0.96 -158.37
## interaction 9 336.37      6.26   0.04   1.00 -157.63
```

Resource

multiple testing

Megan Goldman, 2008, Statistics for Bioinformatics

Mann-Whitney U Test

Mann-Whitney U Test Theory

Mann-Whitney U Test table