Vulture Bee Diversity Statistics on 16S Illumina Sequences of the Gut Microbiome

Jessica Maccaro

10/4/2021

Loading vegan for beta diversity analysis then displaying R version

```
library(vegan)
```

```
## Loading required package: permute
```

```
## Loading required package: lattice
```

```
## This is vegan 2.5-7
```

```
version
```

```
##                    _
## platform       x86_64-apple-darwin17.0
## arch           x86_64
## os             darwin17.0
## system         x86_64, darwin17.0
## status
## major          4
## minor          0.3
## year           2020
## month          10
## day            10
## svn rev        79318
## language       R
## version.string R version 4.0.3 (2020-10-10)
## nickname       Bunny-Wunnies Freak Out
```

Beta Diversity Stats

####Loading Files

I had to make sure my files were nice and clean in excel. For the mapping file I made sure to remove any blanks, I removed sample 161 because we saw previously it was a weird outlier, I removed all the wasps and baits and I sorted in assending order. For downstream plotting purposes I also had to add a "Code" column where I numerically coded the necrophagy. 1 = Obligate, 2 = Facultative, 3 = Absent.

The OTU table was generated in qiime2 after eliminating the same samples as the meta data file above through filtering. BUT also we lost sample IDs when we specified our sequencing depth. We retained

1,053,000 (44.90%) sequences in 135 (85.44%) samples at sampling depth of 7800. So I went through the metadata file again in excel and removed any samples lost in this step as well so the mapping file matched the OTU table. Note: before doing this I had to transpose the OTU table so that the ASVs were the columns and the sampleIDs were the rows. I also had to make sure to remove the '#' that is in the OTU header so R ould read the header.

The species mapping file only includes species with at least 3 representatives.The necrophagy mapping file includes all species.

```
map.necro <- read.table("Map_nowasp.b.txt", header = T, row.names = 1, sep='\t')
attach(map.necro)

map.species <- read.table("Map_R_species.txt", header = T, row.names = 1, sep='\t')
attach(map.species)
```

```
## The following objects are masked from map.necro:
##
##      BarcodeSequence, Code, CollectionPoint, Date, Ex., Genus,
##      LinkerPrimerSeq, Location, Necrophagy, species
```

```
OTU.necrophagy <-read.table("rarif_feat_tab_ordered_transposed.txt", header = T,row.names = 1, sep = "\
OTU.mn <- as.matrix(OTU.necrophagy)


OTU.species <-read.table("OTU_R_Species.txt", header = T,row.names = 1, sep = "\t")
OTU.ms <- as.matrix(OTU.species)
```

NMDS Stress

For necrophagy: after making sure the mapping file and OTU table have the sample number of observations (135) I can move on. So we are looking at 135 samples after all the filtering and decontamination steps. After checking lots of different k (from 2-5). k= 5 had a stress below 0.1.

```
BCn.nmds = metaMDS(OTU.mn, distance="bray", k=5, trymax=1000)
```

```
## Square root transformation
## Wisconsin double standardization
## Run 0 stress 0.09009076
## Run 1 stress 0.09027115
## ... Procrustes: rmse 0.01426934  max resid 0.08948022
## Run 2 stress 0.09075229
## Run 3 stress 0.09027881
## ... Procrustes: rmse 0.0147123   max resid 0.08637618
## Run 4 stress 0.09168814
## Run 5 stress 0.09022915
## ... Procrustes: rmse 0.01223685  max resid 0.05262672
## Run 6 stress 0.09182476
## Run 7 stress 0.09027572
## ... Procrustes: rmse 0.01365909  max resid 0.08690296
## Run 8 stress 0.09144226
## Run 9 stress 0.09033105
## ... Procrustes: rmse 0.01286924  max resid 0.08592655
## Run 10 stress 0.09021405
```

```
## ... Procrustes: rmse 0.008715193  max resid 0.08303864
## Run 11 stress 0.09123162
## Run 12 stress 0.09018981
## ... Procrustes: rmse 0.01182279  max resid 0.05311512
## Run 13 stress 0.09219263
## Run 14 stress 0.09029114
## ... Procrustes: rmse 0.01516508  max resid 0.09270769
## Run 15 stress 0.09017889
## ... Procrustes: rmse 0.01066256  max resid 0.08405584
## Run 16 stress 0.09022897
## ... Procrustes: rmse 0.008927588  max resid 0.04896286
## Run 17 stress 0.09043534
## ... Procrustes: rmse 0.01519879  max resid 0.09112179
## Run 18 stress 0.09116759
## Run 19 stress 0.09028769
## ... Procrustes: rmse 0.009649336  max resid 0.08091445
## Run 20 stress 0.09048604
## ... Procrustes: rmse 0.01549117  max resid 0.06164568
## Run 21 stress 0.09020208
## ... Procrustes: rmse 0.008586382  max resid 0.08333296
## Run 22 stress 0.09019539
## ... Procrustes: rmse 0.008817916  max resid 0.03165584
## Run 23 stress 0.09124963
## Run 24 stress 0.09234101
## Run 25 stress 0.09010252
## ... Procrustes: rmse 0.009985668  max resid 0.03462441
## Run 26 stress 0.09027219
## ... Procrustes: rmse 0.014437  max resid 0.08893377
## Run 27 stress 0.0916602
## Run 28 stress 0.090081
## ... New best solution
## ... Procrustes: rmse 0.002508616  max resid 0.007098521
## ... Similar to previous best
## *** Solution reached
```

BCn.nmds

```
##
## Call:
## metaMDS(comm = OTU.mn, distance = "bray", k = 5, trymax = 1000)
##
## global Multidimensional Scaling using monoMDS
##
## Data:      wisconsin(sqrt(OTU.mn))
## Distance: bray
##
## Dimensions: 5
## Stress:      0.090081
## Stress type 1, weak ties
## Two convergent solutions found after 28 tries
## Scaling: centring, PC rotation, halfchange scaling
## Species: expanded scores based on 'wisconsin(sqrt(OTU.mn))'
```

For species: the OTU and species have the same number of samples 128.I stuck with a k = 5 because it was

below 0.1. When I tried k = 4, it was 0.1056

```
BCs.nmds = metaMDS(OTU.species, distance="bray", k=5, trymax=1000)
```

```
## Square root transformation
## Wisconsin double standardization
## Run 0 stress 0.08152214
## Run 1 stress 0.0818473
## ... Procrustes: rmse 0.04365016  max resid 0.1192911
## Run 2 stress 0.08170859
## ... Procrustes: rmse 0.006854799  max resid 0.06226793
## Run 3 stress 0.08155536
## ... Procrustes: rmse 0.007497374  max resid 0.04137497
## Run 4 stress 0.08309753
## Run 5 stress 0.08167224
## ... Procrustes: rmse 0.0445661  max resid 0.1331438
## Run 6 stress 0.08152202
## ... New best solution
## ... Procrustes: rmse 0.000131615  max resid 0.0004794039
## ... Similar to previous best
## Run 7 stress 0.08239046
## Run 8 stress 0.08155771
## ... Procrustes: rmse 0.04636376  max resid 0.1452321
## Run 9 stress 0.08153664
## ... Procrustes: rmse 0.044981  max resid 0.1346306
## Run 10 stress 0.08201505
## ... Procrustes: rmse 0.04076937  max resid 0.1217434
## Run 11 stress 0.08321832
## Run 12 stress 0.08157364
## ... Procrustes: rmse 0.0448362  max resid 0.1195802
## Run 13 stress 0.08167918
## ... Procrustes: rmse 0.006508764  max resid 0.06418044
## Run 14 stress 0.08172934
## ... Procrustes: rmse 0.04448119  max resid 0.1377995
## Run 15 stress 0.08155549
## ... Procrustes: rmse 0.007486585  max resid 0.04064591
## Run 16 stress 0.0815731
## ... Procrustes: rmse 0.04657057  max resid 0.1237769
## Run 17 stress 0.08167958
## ... Procrustes: rmse 0.006423543  max resid 0.06404409
## Run 18 stress 0.08151408
## ... New best solution
## ... Procrustes: rmse 0.04552337  max resid 0.1247215
## Run 19 stress 0.08171972
## ... Procrustes: rmse 0.02323494  max resid 0.1299883
## Run 20 stress 0.0813944
## ... New best solution
## ... Procrustes: rmse 0.01029658  max resid 0.0639726
## Run 21 stress 0.08364581
## Run 22 stress 0.08155562
## ... Procrustes: rmse 0.04612278  max resid 0.1245482
## Run 23 stress 0.08151843
## ... Procrustes: rmse 0.01113046  max resid 0.06109605
## Run 24 stress 0.08156071
```

```
## ... Procrustes: rmse 0.01349852  max resid 0.08922834
## Run 25 stress 0.08326015
## Run 26 stress 0.08153667
## ... Procrustes: rmse 0.01797559  max resid 0.1095523
## Run 27 stress 0.0815542
## ... Procrustes: rmse 0.04602549  max resid 0.1242109
## Run 28 stress 0.08152228
## ... Procrustes: rmse 0.04551684  max resid 0.1247562
## Run 29 stress 0.08156905
## ... Procrustes: rmse 0.01365536  max resid 0.08985773
## Run 30 stress 0.08297403
## Run 31 stress 0.08198131
## Run 32 stress 0.08155428
## ... Procrustes: rmse 0.04602772  max resid 0.1244037
## Run 33 stress 0.08138821
## ... New best solution
## ... Procrustes: rmse 0.00569004  max resid 0.05619351
## Run 34 stress 0.08335884
## Run 35 stress 0.08155406
## ... Procrustes: rmse 0.04621458  max resid 0.1254153
## Run 36 stress 0.08252549
## Run 37 stress 0.08168776
## ... Procrustes: rmse 0.04524198  max resid 0.1254999
## Run 38 stress 0.08185572
## ... Procrustes: rmse 0.02290885  max resid 0.1286179
## Run 39 stress 0.08505478
## Run 40 stress 0.08149506
## ... Procrustes: rmse 0.01800624  max resid 0.1098081
## Run 41 stress 0.08144811
## ... Procrustes: rmse 0.008274652  max resid 0.04121367
## Run 42 stress 0.08155056
## ... Procrustes: rmse 0.01325007  max resid 0.1052847
## Run 43 stress 0.08275533
## Run 44 stress 0.08174471
## ... Procrustes: rmse 0.02333401  max resid 0.1281757
## Run 45 stress 0.08198186
## Run 46 stress 0.08350317
## Run 47 stress 0.08174373
## ... Procrustes: rmse 0.04579235  max resid 0.1258682
## Run 48 stress 0.08189532
## Run 49 stress 0.08168042
## ... Procrustes: rmse 0.04541624  max resid 0.1265786
## Run 50 stress 0.08139766
## ... Procrustes: rmse 0.007038205  max resid 0.03549682
## Run 51 stress 0.08140265
## ... Procrustes: rmse 0.00698366  max resid 0.0353969
## Run 52 stress 0.0822617
## Run 53 stress 0.08168058
## ... Procrustes: rmse 0.04538033  max resid 0.1266659
## Run 54 stress 0.0814952
## ... Procrustes: rmse 0.01813026  max resid 0.1111139
## Run 55 stress 0.08161671
## ... Procrustes: rmse 0.04585238  max resid 0.1259878
## Run 56 stress 0.08174748
```

```
## ... Procrustes: rmse 0.01757192  max resid 0.1181077
## Run 57 stress 0.08176807
## ... Procrustes: rmse 0.02408046  max resid 0.1293775
## Run 58 stress 0.08138605
## ... New best solution
## ... Procrustes: rmse 0.002184199  max resid 0.01278841
## Run 59 stress 0.08196649
## Run 60 stress 0.08174281
## ... Procrustes: rmse 0.0458092  max resid 0.124189
## Run 61 stress 0.0815542
## ... Procrustes: rmse 0.04608965  max resid 0.1230418
## Run 62 stress 0.08155528
## ... Procrustes: rmse 0.04617167  max resid 0.1233056
## Run 63 stress 0.08167966
## ... Procrustes: rmse 0.04526487  max resid 0.1240667
## Run 64 stress 0.08139746
## ... Procrustes: rmse 0.006728545  max resid 0.03623781
## Run 65 stress 0.08157385
## ... Procrustes: rmse 0.01367538  max resid 0.08779209
## Run 66 stress 0.08213941
## Run 67 stress 0.08152323
## ... Procrustes: rmse 0.04561562  max resid 0.1236008
## Run 68 stress 0.08152252
## ... Procrustes: rmse 0.045565  max resid 0.1236761
## Run 69 stress 0.08155053
## ... Procrustes: rmse 0.01809547  max resid 0.1157671
## Run 70 stress 0.081743
## ... Procrustes: rmse 0.04575645  max resid 0.123857
## Run 71 stress 0.08155438
## ... Procrustes: rmse 0.04610044  max resid 0.1233833
## Run 72 stress 0.08155465
## ... Procrustes: rmse 0.04610373  max resid 0.1234829
## Run 73 stress 0.08186536
## ... Procrustes: rmse 0.0262857  max resid 0.1326898
## Run 74 stress 0.08152207
## ... Procrustes: rmse 0.04557099  max resid 0.1235015
## Run 75 stress 0.08139746
## ... Procrustes: rmse 0.005582594  max resid 0.0554406
## Run 76 stress 0.08174284
## ... Procrustes: rmse 0.04574564  max resid 0.123832
## Run 77 stress 0.08203245
## Run 78 stress 0.08155543
## ... Procrustes: rmse 0.04613062  max resid 0.1231246
## Run 79 stress 0.08155563
## ... Procrustes: rmse 0.01373508  max resid 0.1082174
## Run 80 stress 0.0813862
## ... Procrustes: rmse 0.00884026  max resid 0.05607939
## Run 81 stress 0.08139336
## ... Procrustes: rmse 0.00545155  max resid 0.05597435
## Run 82 stress 0.08330109
## Run 83 stress 0.08174188
## ... Procrustes: rmse 0.04573714  max resid 0.1241153
## Run 84 stress 0.08179605
## ... Procrustes: rmse 0.02498539  max resid 0.1317269
```

```
## Run 85 stress 0.08149465
## ... Procrustes: rmse 0.01863619  max resid 0.1134672
## Run 86 stress 0.08175724
## ... Procrustes: rmse 0.02408304  max resid 0.1306686
## Run 87 stress 0.08229856
## Run 88 stress 0.08190593
## Run 89 stress 0.08150076
## ... Procrustes: rmse 0.01893237  max resid 0.1136092
## Run 90 stress 0.08152305
## ... Procrustes: rmse 0.0456131  max resid 0.1235995
## Run 91 stress 0.08152228
## ... Procrustes: rmse 0.04559437  max resid 0.1237827
## Run 92 stress 0.08152271
## ... Procrustes: rmse 0.04556943  max resid 0.1235018
## Run 93 stress 0.08151142
## ... Procrustes: rmse 0.01932758  max resid 0.1171804
## Run 94 stress 0.08159131
## ... Procrustes: rmse 0.04594608  max resid 0.1237757
## Run 95 stress 0.08139938
## ... Procrustes: rmse 0.006831357  max resid 0.03626741
## Run 96 stress 0.08152239
## ... Procrustes: rmse 0.04556912  max resid 0.1237
## Run 97 stress 0.08139392
## ... Procrustes: rmse 0.005356846  max resid 0.05499289
## Run 98 stress 0.08175094
## ... Procrustes: rmse 0.0456571  max resid 0.1238168
## Run 99 stress 0.08153437
## ... Procrustes: rmse 0.01590137  max resid 0.1135841
## Run 100 stress 0.08163998
## ... Procrustes: rmse 0.01522377  max resid 0.1052028
## Run 101 stress 0.08168414
## ... Procrustes: rmse 0.04642134  max resid 0.1240674
## Run 102 stress 0.08156902
## ... Procrustes: rmse 0.01292477  max resid 0.08402157
## Run 103 stress 0.08153707
## ... Procrustes: rmse 0.01720784  max resid 0.1109039
## Run 104 stress 0.08391653
## Run 105 stress 0.08188256
## ... Procrustes: rmse 0.02382939  max resid 0.1311142
## Run 106 stress 0.08155542
## ... Procrustes: rmse 0.04618045  max resid 0.1234443
## Run 107 stress 0.08226643
## Run 108 stress 0.08149557
## ... Procrustes: rmse 0.01872491  max resid 0.1141256
## Run 109 stress 0.0835731
## Run 110 stress 0.0823844
## Run 111 stress 0.082096
## Run 112 stress 0.08167955
## ... Procrustes: rmse 0.04525086  max resid 0.1241161
## Run 113 stress 0.08177477
## ... Procrustes: rmse 0.0460416  max resid 0.1237237
## Run 114 stress 0.08154138
## ... Procrustes: rmse 0.01773832  max resid 0.1137751
## Run 115 stress 0.08196575
```

```
## Run 116 stress 0.08179682
## ... Procrustes: rmse 0.0219546  max resid 0.1287829
## Run 117 stress 0.0815372
## ... Procrustes: rmse 0.01739545  max resid 0.1109437
## Run 118 stress 0.0817546
## ... Procrustes: rmse 0.04566236  max resid 0.1242335
## Run 119 stress 0.08183546
## ... Procrustes: rmse 0.04554386  max resid 0.1234786
## Run 120 stress 0.08152258
## ... Procrustes: rmse 0.0456129  max resid 0.1236741
## Run 121 stress 0.08194733
## Run 122 stress 0.08174265
## ... Procrustes: rmse 0.04571746  max resid 0.1238924
## Run 123 stress 0.08157121
## ... Procrustes: rmse 0.01685035  max resid 0.1149984
## Run 124 stress 0.08183973
## ... Procrustes: rmse 0.0226702  max resid 0.1323361
## Run 125 stress 0.0816797
## ... Procrustes: rmse 0.04528295  max resid 0.1244133
## Run 126 stress 0.08152181
## ... Procrustes: rmse 0.04558441  max resid 0.1236718
## Run 127 stress 0.08174217
## ... Procrustes: rmse 0.0457142  max resid 0.1239853
## Run 128 stress 0.08168082
## ... Procrustes: rmse 0.04520866  max resid 0.1239874
## Run 129 stress 0.08162453
## ... Procrustes: rmse 0.01672564  max resid 0.1091467
## Run 130 stress 0.08168489
## ... Procrustes: rmse 0.04531787  max resid 0.1243097
## Run 131 stress 0.08162452
## ... Procrustes: rmse 0.01676063  max resid 0.1093987
## Run 132 stress 0.0817729
## ... Procrustes: rmse 0.02444594  max resid 0.1312588
## Run 133 stress 0.08174332
## ... Procrustes: rmse 0.04577077  max resid 0.1236832
## Run 134 stress 0.08155763
## ... Procrustes: rmse 0.04607712  max resid 0.1234645
## Run 135 stress 0.08152214
## ... Procrustes: rmse 0.04558064  max resid 0.1235203
## Run 136 stress 0.08169326
## ... Procrustes: rmse 0.01729774  max resid 0.1168132
## Run 137 stress 0.08174848
## ... Procrustes: rmse 0.04496933  max resid 0.1244322
## Run 138 stress 0.08171602
## ... Procrustes: rmse 0.01486223  max resid 0.110905
## Run 139 stress 0.08180015
## ... Procrustes: rmse 0.02234427  max resid 0.1289124
## Run 140 stress 0.08149977
## ... Procrustes: rmse 0.01893804  max resid 0.113626
## Run 141 stress 0.08174213
## ... Procrustes: rmse 0.04576283  max resid 0.1238462
## Run 142 stress 0.08208488
## Run 143 stress 0.08202189
## Run 144 stress 0.0815086
```

```
## ... Procrustes: rmse 0.01534946  max resid 0.1109449
## Run 145 stress 0.08149536
## ... Procrustes: rmse 0.01847063  max resid 0.1132745
## Run 146 stress 0.08155467
## ... Procrustes: rmse 0.04610337  max resid 0.1235263
## Run 147 stress 0.0816802
## ... Procrustes: rmse 0.04526949  max resid 0.1243231
## Run 148 stress 0.0815113
## ... Procrustes: rmse 0.009622574  max resid 0.05638424
## Run 149 stress 0.08155463
## ... Procrustes: rmse 0.04608548  max resid 0.1232481
## Run 150 stress 0.08152246
## ... Procrustes: rmse 0.04556221  max resid 0.1238686
## Run 151 stress 0.08218281
## Run 152 stress 0.08140057
## ... Procrustes: rmse 0.006615775  max resid 0.03449005
## Run 153 stress 0.08139643
## ... Procrustes: rmse 0.005486408  max resid 0.05563264
## Run 154 stress 0.08152205
## ... Procrustes: rmse 0.04559283  max resid 0.1236648
## Run 155 stress 0.0816802
## ... Procrustes: rmse 0.04521983  max resid 0.1244625
## Run 156 stress 0.08160524
## ... Procrustes: rmse 0.01938035  max resid 0.1214077
## Run 157 stress 0.08169769
## ... Procrustes: rmse 0.04530217  max resid 0.1242941
## Run 158 stress 0.08173154
## ... Procrustes: rmse 0.04579356  max resid 0.123971
## Run 159 stress 0.08155434
## ... Procrustes: rmse 0.04614379  max resid 0.1232383
## Run 160 stress 0.0815225
## ... Procrustes: rmse 0.04560969  max resid 0.1239129
## Run 161 stress 0.08208479
## Run 162 stress 0.08139622
## ... Procrustes: rmse 0.005474554  max resid 0.05581463
## Run 163 stress 0.08155932
## ... Procrustes: rmse 0.04621794  max resid 0.1233779
## Run 164 stress 0.0815547
## ... Procrustes: rmse 0.04609293  max resid 0.123492
## Run 165 stress 0.08138621
## ... Procrustes: rmse 0.001057665  max resid 0.006744383
## ... Similar to previous best
## *** Solution reached
```

```
BCs.nmds
```

```
##
## Call:
## metaMDS(comm = OTU.species, distance = "bray", k = 5, trymax = 1000)
##
## global Multidimensional Scaling using monoMDS
##
## Data:     wisconsin(sqrt(OTU.species))
## Distance: bray
```

```
## 
## Dimensions: 5
## Stress:      0.08138605
## Stress type 1, weak ties
## Two convergent solutions found after 165 tries
## Scaling: centring, PC rotation, halfchange scaling
## Species: expanded scores based on 'wisconsin(sqrt(OTU.species))'
```

####Figure 3: Microbial Community Composition by Host Species and Diet

I have to add a bunch of colors for species so I'm using this cool package to see which palletes are available to use for color blind folks.
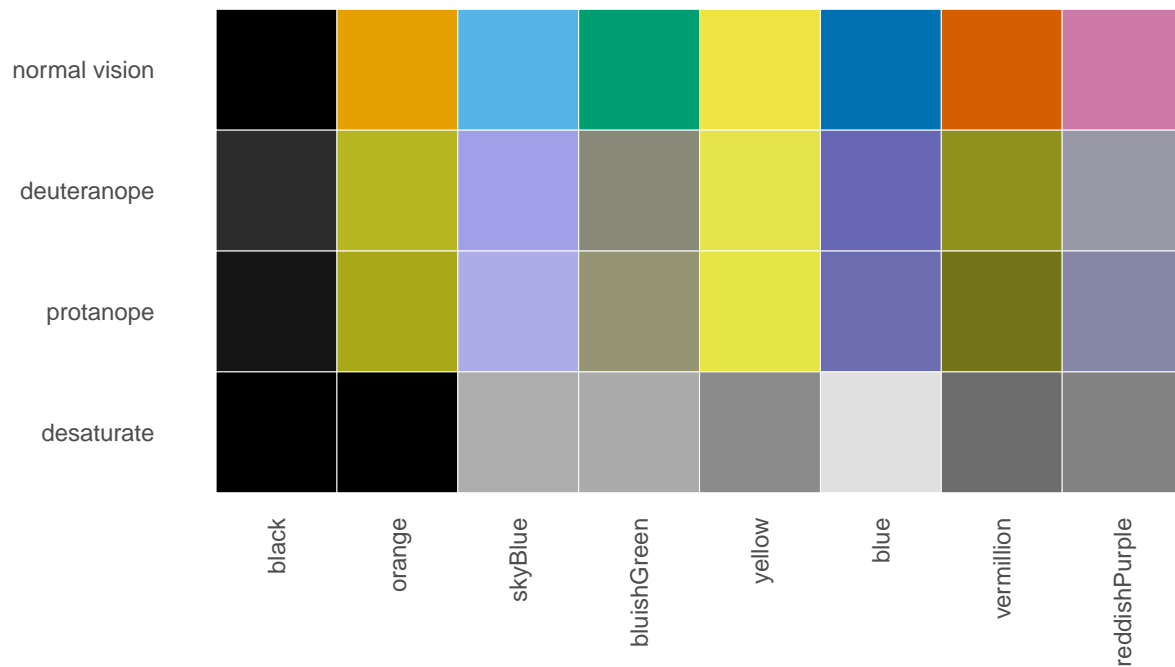
```
#install.packages("colorBlindness")
library(colorBlindness)
displayAvailablePalette(color="black")
```

```
## Warning: It is deprecated to specify 'guide = FALSE' to remove a guide. Please
## use 'guide = "none"' instead.
```
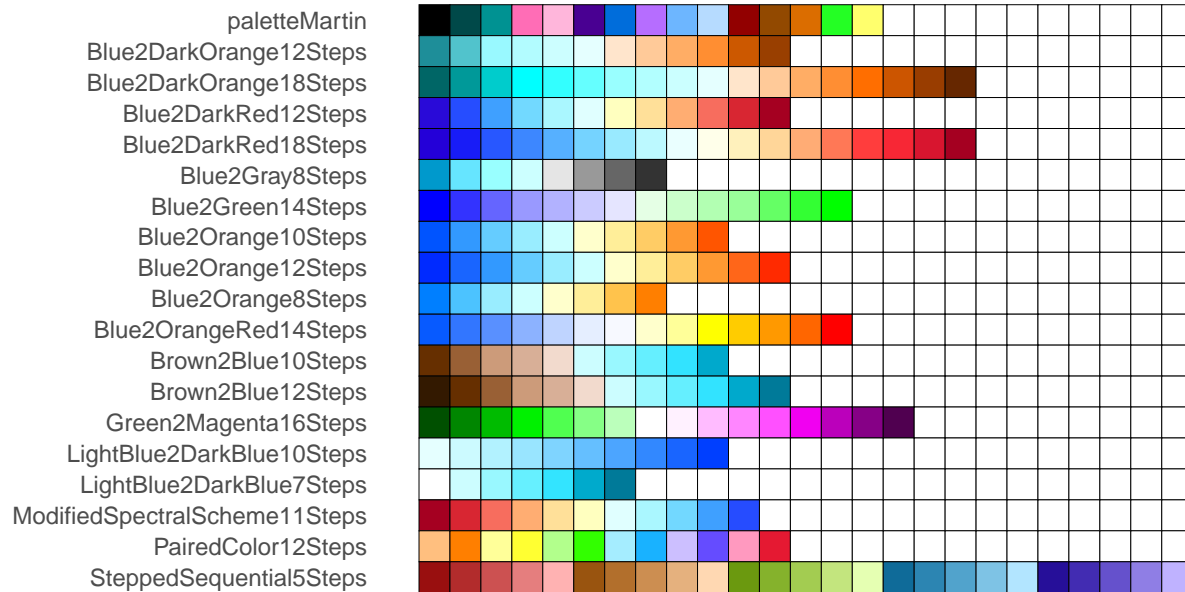


```
displayAllColors(safeColors, color="white")
```

```
## Warning: It is deprecated to specify 'guide = FALSE' to remove a guide. Please
## use 'guide = "none"' instead.
```

We are using the necrophagy mapping file which includes species with n < 3. I also arranged them from pollenivores to necrophages so the blue are pollen feeding and the most red is obligate necrophage. The ellipses represent the 95% confidence intervals for the diet type.

```
plot(BCn.nmds, display = "sites")
orditorp(BCn.nmds, "sites")
```
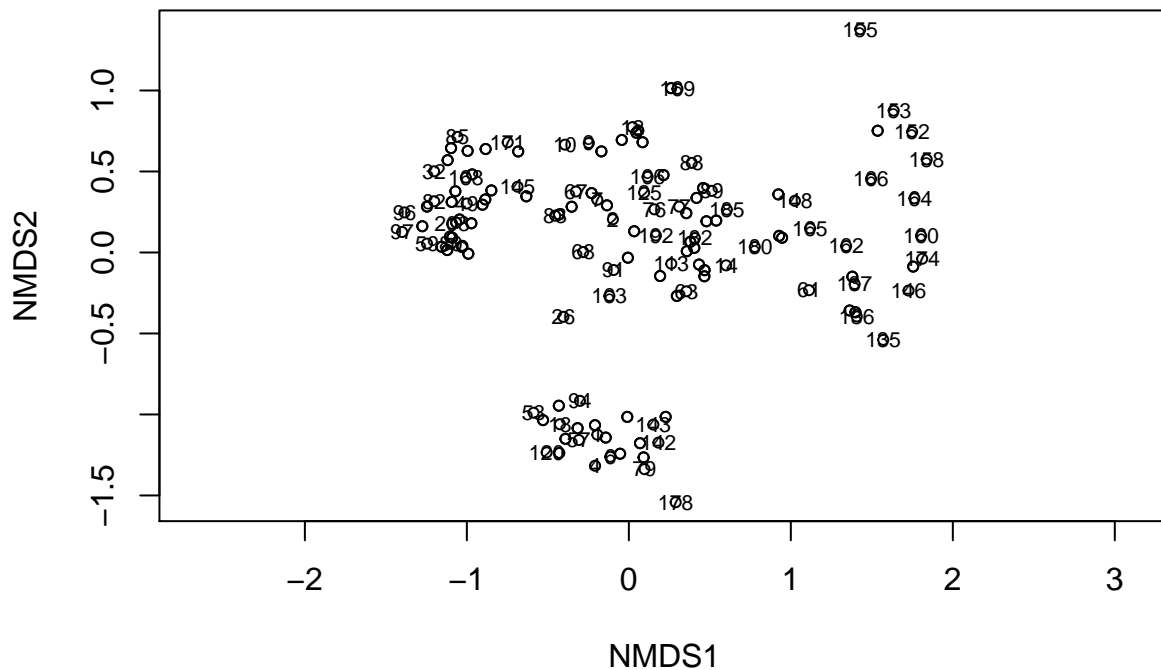
```
colvec_s <- c(Blue2DarkRed18Steps)

sp.map.df <- as.data.frame(map.necro)
sp.map.ord <- sp.map.df
sp.map.ord$species <- factor(sp.map.ord$species,
                             levels = c("angustula",
                                        "buchwaldi",
                                        "wheeleri",
                                        "argyrea",
                                        "sp.",
                                        "opaca",
                                        "flavoscutellata",
                                        "panamica",
                                        "costaricensis",
                                        "mellicolor",
                                        "musarum",
                                        "orizabaensis",
                                        "corvina",
                                        "ferricauda",
                                        "fulviventris",
                                        "silvestriana",
                                        "necrophaga"))

plot(BCn.nmds, "sites",
     main ="Microbial Community Composition by Host Species and Diet")
with(sp.map.ord,
```

```
    points(BCn.nmds,
           display = "sites",
           pch = 20,
           col = colvec_s[sp.map.ord$species]))

legend("topleft", legend=c("Te. angustula","Te. buchwaldi", "S.wheeleri", "Sc.argyrea", "Scaptotrigona s

legend("bottomright", legend=c("Obligate","Facultative","Absent"), col=c("darkorchid","goldenrod1","darl

ordiellipse(
  BCn.nmds,
  map.necro$Code,
  display = "sites",
  conf = 0.95,
  col = c("darkorchid","goldenrod1","darkolivegreen"),
  lty = c(1, 2, 3),
  lwd = 2.5
)
```
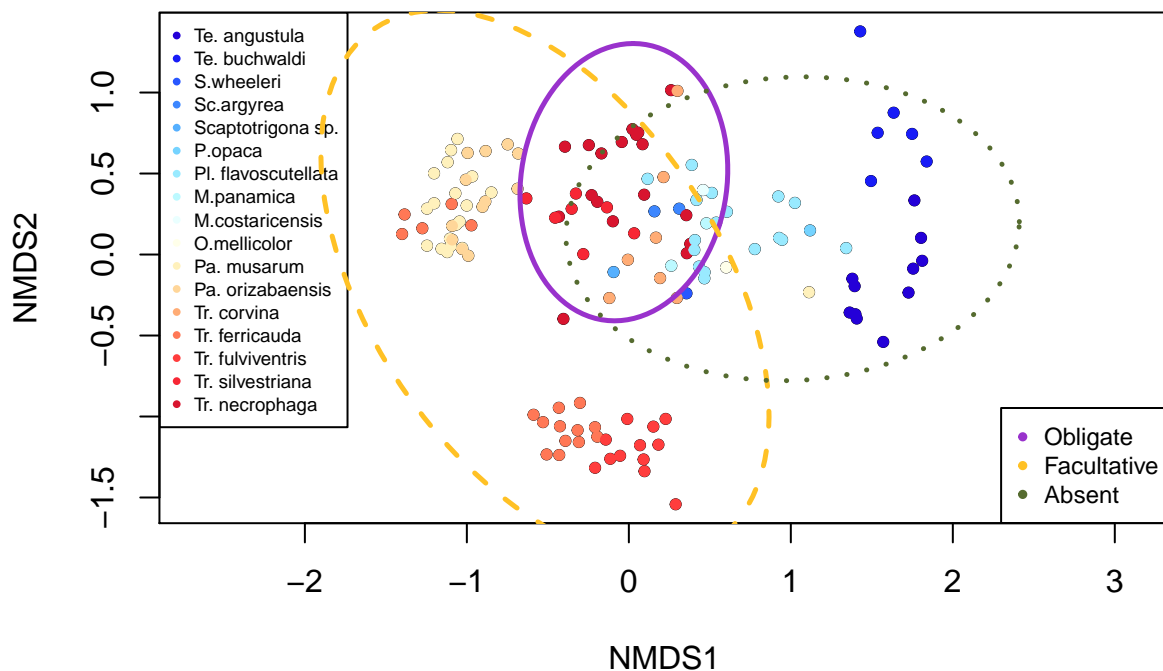

Microbial Community Composition by Host Species and Diet

```
dev.print(pdf,          # copies the plot to a the PDF file
          "FinalPlots/Species+NecroNMDS.pdf")
```

```
## pdf
##   2
```

####ADONIS

This is to get significance values for the beta diversity (differences in communities by sample).I ran a seperate Adonis for species and necrophagy because they are co-linear so I could not include them in the same model. I used Location which is the specific baits as a strata, but the adonis2 notation is "blocks".The idea is to account for the possibility that the same colony could be at different baits.

I first calculate bray curtis distance matrices for both species and necrophagy.

```r
BCn.dist <- vegdist(OTU.mn, distance="bray")
BCs.dist <- vegdist(OTU.ms, distance="bray")
```

Then I run PERMANOVA on distances for necrophagy:

```r
perms.n <- with(map.necro, how(nperm = 1000, blocks = Location))
adonis2(formula = BCn.dist ~ Necrophagy, permutations = perms.n, data = map.necro, pairwise = TRUE)
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Blocks:  Location
## Permutation: free
## Number of permutations: 1000
##
## adonis2(formula = BCn.dist ~ Necrophagy, data = map.necro, permutations = perms.n, pairwise = TRUE)
##             Df SumOfSqs      R2      F   Pr(>F)
## Necrophagy   2    7.473 0.12232 9.1981 0.000999 ***
## Residual   132   53.624 0.87768
## Total      134   61.098 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Run PERMANOVA on distances for species

```r
#species
perms.s <- with(map.species, how(nperm = 1000, blocks = Location))
adonis2(formula = BCs.dist ~ species, permutations = perms.s, data = map.species, pairwise= TRUE)
```

```
## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Blocks:  Location
## Permutation: free
## Number of permutations: 1000
##
## adonis2(formula = BCs.dist ~ species, data = map.species, permutations = perms.s, pairwise = TRUE)
##          Df SumOfSqs      R2      F   Pr(>F)
## species    9    8.417 0.14971 2.2497 0.000999 ***
## Residual 115   47.808 0.85029
## Total    124   56.226 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

###Beta Dispersion

This shows the different levels of dispersion between samples in each category. Again I run this for both necrophagy and species.The block is the locaton of specific baits like in the adonis.

Necrophagy:

```
disp.necro = betadisper(BCn.dist, map.necro$Necrophagy)
perms.bdn <- with(map.necro, how(nperm = 1000, blocks = Location))
p.w.bd_necro <- permutest(disp.necro, pairwise=TRUE, permutations=perms.bdn)
p.w.bd_necro
```

```
##
## Permutation test for homogeneity of multivariate dispersions
## Blocks:  Location
## Permutation: free
## Number of permutations: 1000
##
## Response: Distances
##           Df  Sum Sq  Mean Sq      F N.Perm   Pr(>F)
## Groups      2 0.11887 0.059434 11.972   1000 0.000999 ***
## Residuals 132 0.65532 0.004965
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Pairwise comparisons:
## (Observed p-value below diagonal, permuted p-value above diagonal)
##               Absent Facultative Obligate
## Absent                1.4985e-02   0.0859
## Facultative 7.7516e-02              0.0030
## Obligate    1.2159e-02  3.5691e-06
```

Species:

```
disp.species = betadisper(BCs.dist, map.species$species)
perms.bds <- with(map.species, how(nperm = 1000, blocks = Location))
permutest(disp.species, pairwise=TRUE, permutations=perms.bds)
```

```
##
## Permutation test for homogeneity of multivariate dispersions
## Blocks:  Location
## Permutation: free
## Number of permutations: 1000
##
## Response: Distances
##           Df  Sum Sq   Mean Sq      F N.Perm Pr(>F)
## Groups      9 0.04869 0.0054104 0.9647   1000 0.7393
## Residuals 115 0.64493 0.0056081
##
## Pairwise comparisons:
## (Observed p-value below diagonal, permuted p-value above diagonal)
##           angustula buchwaldi  corvina ferricauda flavoscutellata
## angustula            0.958042 0.308691   0.088911        0.521478
## buchwaldi  0.882939           0.243756   0.162837        0.232767
## corvina    0.318675  0.149729            0.970030        0.601399
```

```
## ferricauda       0.131847  0.049849 0.976319                0.555445
## flavoscutellata  0.238672  0.174457 0.594157    0.437155
## fulviventris     0.154680  0.038451 0.830873    0.807552      0.338866
## musarum          0.219123  0.290944 0.929313    0.875922      0.723343
## necrophaga       0.063938  0.031389 0.587495    0.466737      0.159497
## orizabaensis     0.338925  0.367739 0.817279    0.716993      0.896666
## silvestriana     0.367947  0.198462 0.846081    0.772038      0.769109
##               fulviventris  musarum necrophaga orizabaensis silvestriana
## angustula         0.091908 0.136863   0.006993     0.236763       0.1668
## buchwaldi         0.168831 0.848152   0.406593     0.488511       0.5255
## corvina           0.824176 0.976024   0.736264     0.752248       0.8761
## ferricauda        0.751249 0.973027   0.652348     0.576424       0.7233
## flavoscutellata   0.497503 1.000000   0.741259     0.906094       0.9251
## fulviventris               0.892108   0.690310     0.544455       0.5634
## musarum           0.780757            0.472527     0.928072       0.9401
## necrophaga        0.640165 0.511363                0.471529       0.3966
## orizabaensis      0.642415 0.871037   0.410313                    0.9291
## silvestriana      0.637563 0.952767   0.443360     0.931088
```

##Alpha Diversity Stats

####Load files

I load in observed_otu alpha diversity files generated in qiime2 with raw ASVs. They went through dada2 in qiime2 so they are ASVs not OTUs. This file only includes species that have 3 or more representatives.

```
K_W_n <- read.table("alpha_div_species.txt", header = T, row.names = 1, sep='\t')
AlDiv.df <- as.data.frame(K_W_n)
```

####Linear Mixed Model

I am just loading the packages I will need. Also making sure the data is loaded as factors and verifying that they are. Plus checking the distribution of the data, which looks fairly normally distributed.

```
#Load and install packages
#install.packages("DHARma")
library(lme4)
```

```
## Loading required package: Matrix
```

```
packageVersion("lme4")
```

```
## [1] '1.1.27.1'
```

```
library(DHARMa)
```

```
## This is DHARMa 0.4.3. For overview type '?DHARMa'. For recent changes, type news(package = 'DHARMa')
```

```
packageVersion("DHARma")
```

```
## [1] '0.4.3'
```

```r
library(car)
```

```
## Loading required package: carData
```

```r
packageVersion("car")
```

```
## [1] '3.0.11'
```

```r
#Load data as factors and check that they are
res_ASV <- AlDiv.df$observed_otus
AlDiv.df$Species <- as.factor(AlDiv.df$Species)
AlDiv.df$Collection <- as.factor(AlDiv.df$Collection)
AlDiv.df$Location <- as.factor(AlDiv.df$Location)
AlDiv.df$Necrophagy <- as.factor(AlDiv.df$Necrophagy)
summary(AlDiv.df)
```

```
##   observed_otus                Species     Species_num              Necrophagy
##  Min.   : 14.00   musarum        :19   Min.   : 1.000   Absent     :34
##  1st Qu.: 33.25   necrophaga     :18   1st Qu.: 3.000   Facultative:74
##  Median : 47.00   ferricauda     :17   Median : 6.000   Obligate   :18
##  Mean   : 50.98   flavoscutellata:17   Mean   : 5.802
##  3rd Qu.: 70.50   fulviventris   :12   3rd Qu.: 8.000
##  Max.   :108.00   orizabaensis   :12   Max.   :10.000
##                   (Other)        :31
##    Necro_num          Collection Collection_num          Location
##  Min.   :1.000   Chicken  :52   Min.   :1.000   LasCrucesDining:32
##  1st Qu.:1.000   Flower   :52   1st Qu.:2.000   LaSelvaDining  :14
##  Median :2.000   Patrolling:22  Median :2.000   Bait6          :13
##  Mean   :1.873                  Mean   :2.238   Bait1          : 9
##  3rd Qu.:2.000                  3rd Qu.:3.000   Bait9          : 7
##  Max.   :3.000                  Max.   :3.000   LaSelvaLab     : 7
##                                                 (Other)        :44
##   Location_num    Species_plot       Coll_Patrol
##  Min.   : 1.00   Length:126        Length:126
##  1st Qu.: 7.00   Class :character  Class :character
##  Median :13.00   Mode  :character  Mode  :character
##  Mean   :10.78
##  3rd Qu.:14.00
##  Max.   :21.00
##
```

```r
#Check the distribution. Looks like a normal distribution so we can use an lmm
hist(AlDiv.df$observed_otus)
```

Histogram of AlDiv.df$observed_otus
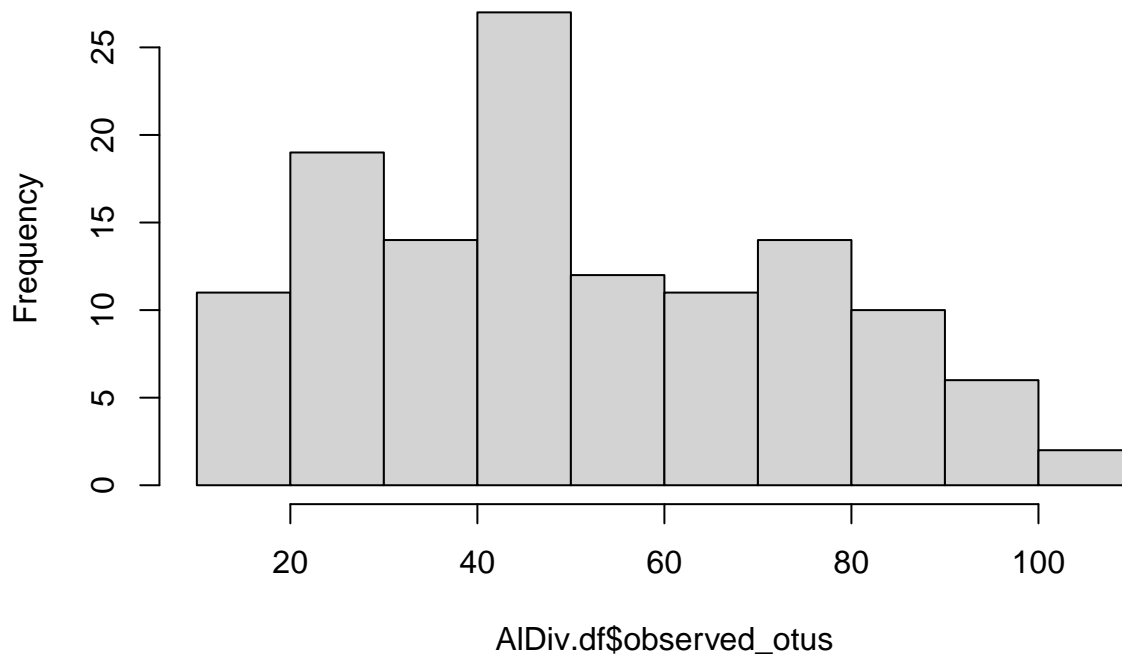
Necro model

To evaluate the role of necrophagy (absent, facultative, and obligate) we constructed a LMM and then will run a likelihood ratio test to evaluate its importance.

```
m1<-lmer(observed_otus~ Necrophagy + (1|Location) + (1|Species), data=AlDiv.df, na.action="na.fail", REI
Anova(m1)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: observed_otus
##            Chisq Df Pr(>Chisq)
## Necrophagy 8.2039  2    0.01654 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(m1)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: observed_otus ~ Necrophagy + (1 | Location) + (1 | Species)
##    Data: AlDiv.df
##
##     AIC      BIC   logLik deviance df.resid
##   1032.8   1049.9   -510.4   1020.8      120
##
## Scaled residuals:
```

```
##      Min      1Q   Median      3Q      Max
## -2.25522 -0.60608 -0.03491  0.53792  2.09455
##
## Random effects:
##  Groups    Name         Variance Std.Dev.
##  Location (Intercept)  55.76     7.467
##  Species  (Intercept) 164.79    12.837
##  Residual              134.56    11.600
## Number of obs: 126, groups:  Location, 21; Species, 10
##
## Fixed effects:
##                       Estimate Std. Error t value
## (Intercept)            32.727       8.173   4.004
## NecrophagyFacultative  27.058       9.645   2.805
## NecrophagyObligate     10.662      15.567   0.685
##
## Correlation of Fixed Effects:
##             (Intr) NcrphF
## NcrphgyFclt -0.802
## NcrphgyOblg -0.507  0.430
```
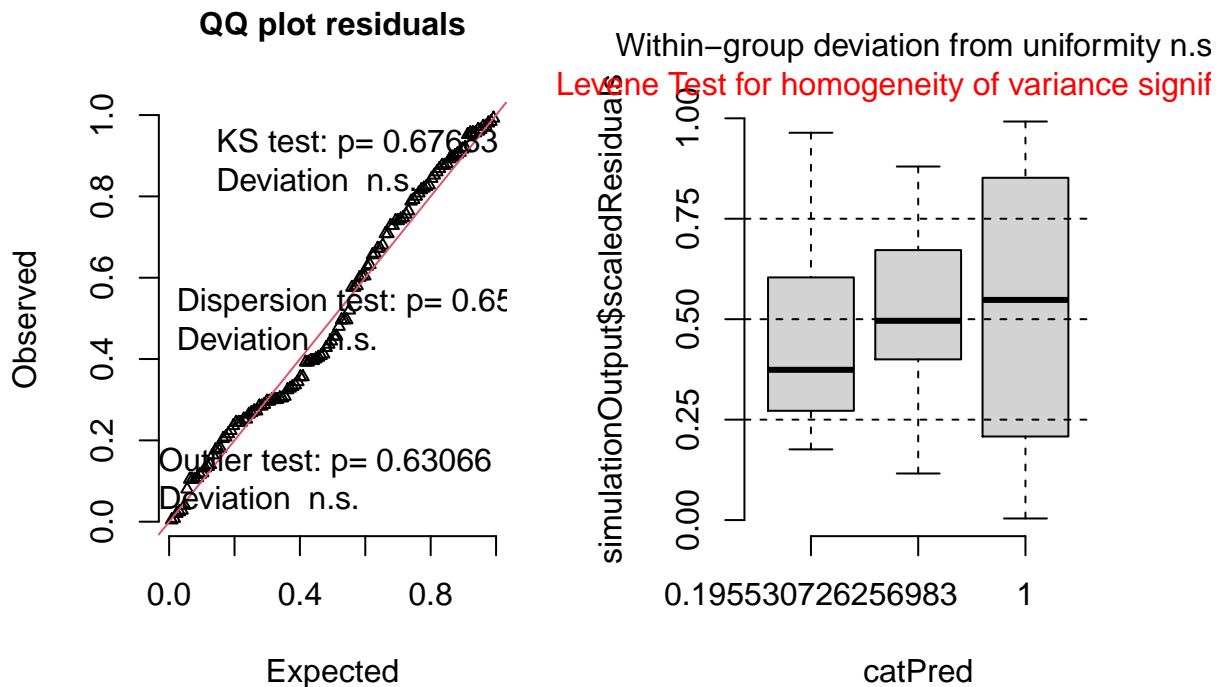
```r
plot(m1)
```



```r
qqnorm(resid(m1))
```

19
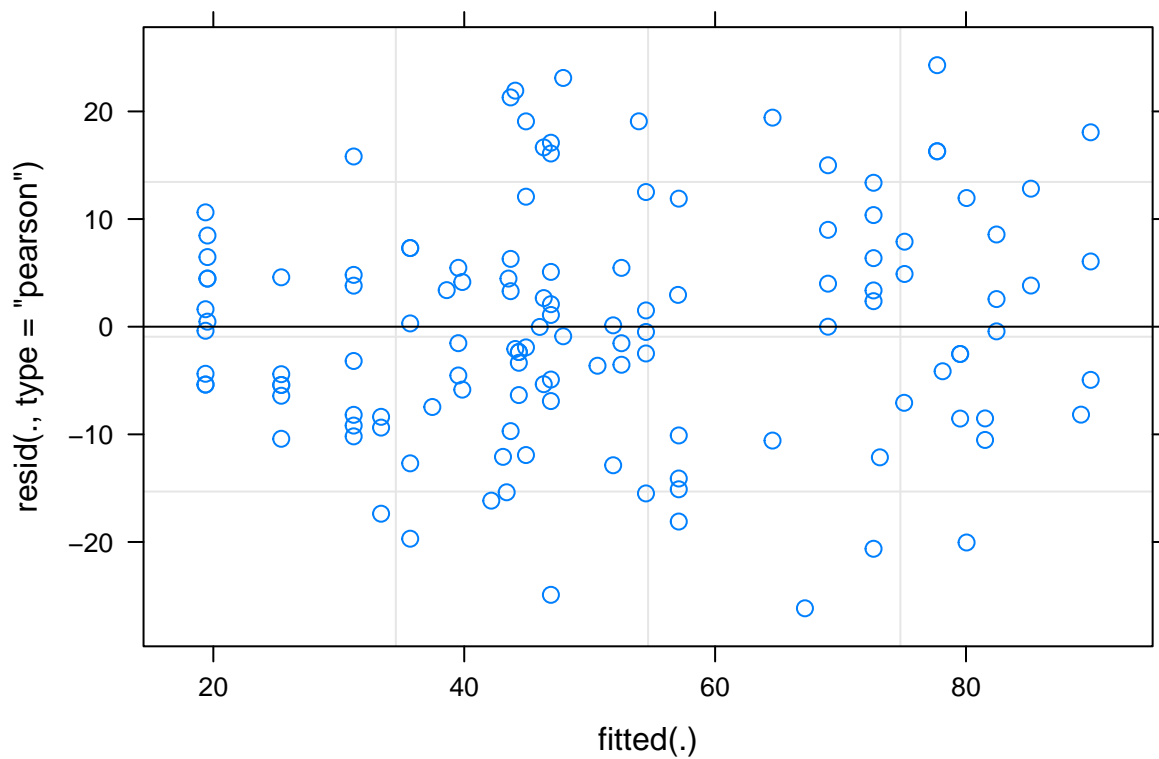
Normal Q–Q Plot
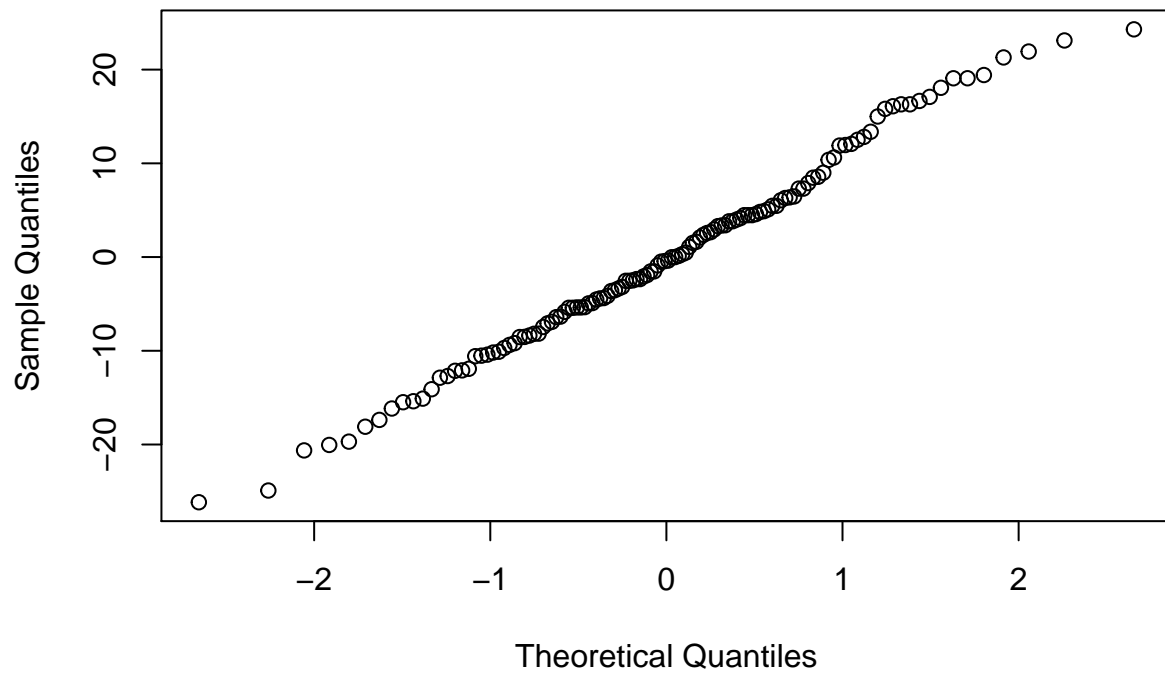

```
simoutbin500.1<-simulateResiduals(fittedModel=m1, n=250)
plot(simoutbin500.1)
```

DHARMa residual diagnostics

QQ plot residuals

KS test: p= 0.67683
Deviation n.s.

Dispersion test: p= 0.65
Deviation n.s.

Outlier test: p= 0.63066
Deviation n.s.

Observed / Expected

Within−group deviation from uniformity n.s
Levene Test for homogeneity of variance signif

simulationOutput$scaledResiduals / catPred

```
m1.1<- lmer(observed_otus ~ 1 + (1|Location) + (1|Species), data=AlDiv.df, na.action="na.fail", REML=F)
anova(m1,m1.1) # value for necrophagy (diet)
```

```
## Data: AlDiv.df
## Models:
## m1.1: observed_otus ~ 1 + (1 | Location) + (1 | Species)
## m1: observed_otus ~ Necrophagy + (1 | Location) + (1 | Species)
##       npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## m1.1    4 1034.8 1046.2 -513.41   1026.8
## m1      6 1032.8 1049.9 -510.42   1020.8 5.9849  2    0.05016 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Species model

Since diet and species are colinear, we have to evaluate them in two different models

```
#Run models
m2<-lmer(observed_otus~ Species + Collection + (1|Location), data=AlDiv.df, na.action="na.fail", REML=F)
vif(m2)
```

```
##               GVIF Df GVIF^(1/(2*Df))
## Species    2.449038  9        1.051020
## Collection 2.449038  2        1.250976
```

```r
Anova(m2)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: observed_otus
##             Chisq Df Pr(>Chisq)
## Species    214.8300  9   < 2e-16 ***
## Collection   6.5214  2    0.03836 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(m2)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: observed_otus ~ Species + Collection + (1 | Location)
##    Data: AlDiv.df
##
##      AIC      BIC   logLik deviance df.resid
##   1010.6   1050.3   -491.3    982.6      112
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.28848 -0.69793  0.03845  0.63591  2.12805
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Location (Intercept)  14.89    3.859
##  Residual             132.34   11.504
## Number of obs: 126, groups:  Location, 21
##
## Fixed effects:
##                        Estimate Std. Error t value
## (Intercept)              32.182      5.502   5.849
## Speciesbuchwaldi          0.427      6.313   0.068
## Speciescorvina           14.569      6.314   2.307
## Speciesferricauda        28.406      5.954   4.771
## Speciesflavoscutellata   14.612      4.896   2.985
## Speciesfulviventris      15.433      5.458   2.828
## Speciesmusarum           48.242      6.036   7.993
## Speciesnecrophaga        14.725      5.906   2.493
## Speciesorizabaensis      56.185      5.392  10.420
## Speciessilvestriana      27.574      6.464   4.266
## CollectionFlower        -10.135      4.005  -2.531
## CollectionPatrolling     -5.603      4.818  -1.163
##
## Correlation of Fixed Effects:
##            (Intr) Spcsbc Spcscr Spcsfr Spcsflvs Spcsflvv Spcsms Spcsnc Spcsrz
## Specsbchwld -0.456
## Speciescrvn -0.533  0.430
## Specisfrrcd -0.835  0.418  0.489
## Spcsflvsctl -0.534  0.471  0.480  0.483
## Spcsflvvntr -0.674  0.481  0.523  0.609  0.555
## Speciesmsrm -0.873  0.421  0.496  0.762  0.492    0.614
```

```
## Specsncrphg -0.821   0.433   0.541   0.713   0.521     0.614     0.761
## Specsrzbnss -0.662   0.480   0.496   0.586   0.563     0.582     0.599   0.590
## Spcsslvstrn -0.516   0.377   0.513   0.474   0.488     0.500     0.490   0.549   0.465
## CollctnFlwr -0.500  -0.037  -0.049   0.319  -0.033     0.085     0.381   0.277   0.121
## CllctnPtrll -0.635   0.135   0.193   0.506  -0.070     0.271     0.543   0.469   0.241
##            Spcssl CllctF
## Specsbchwld
## Speciescrvn
## Specisfrrcd
## Spcsflvsctl
## Spcsflvvntr
## Speciesmsrm
## Specsncrphg
## Specsrzbnss
## Spcsslvstrn
## CollctnFlwr -0.055
## CllctnPtrll  0.109   0.571
```
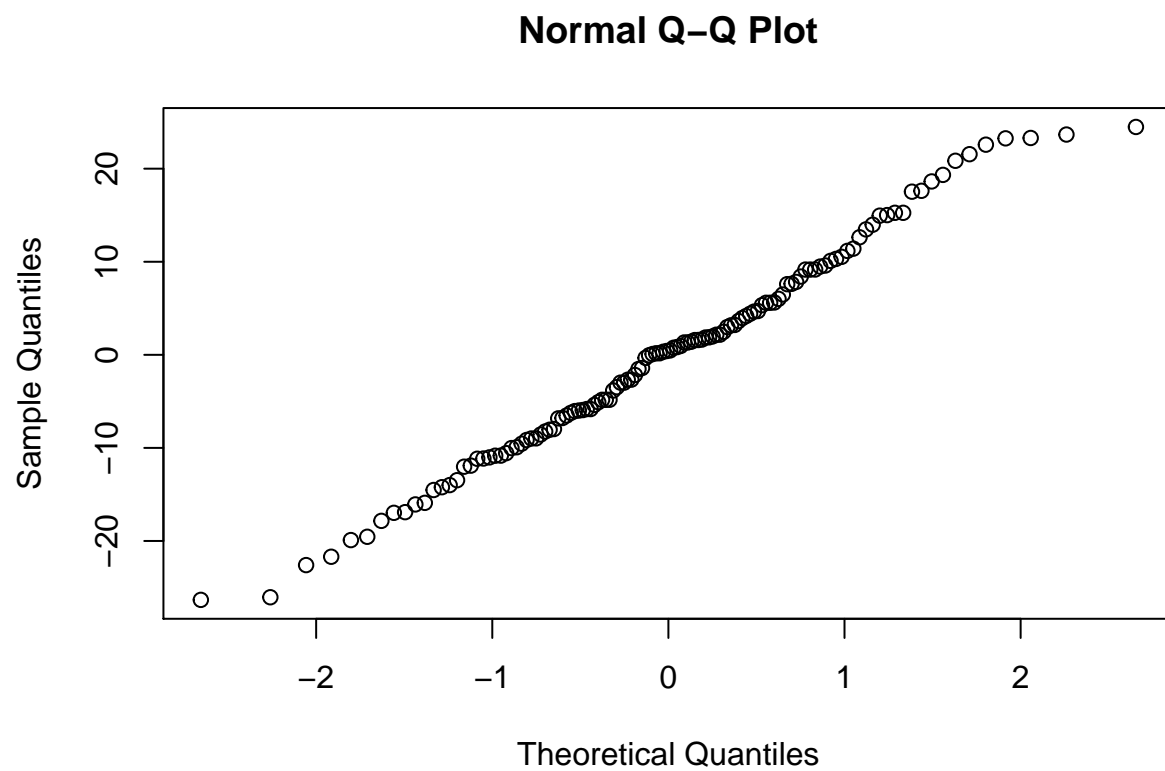
We included collection substrate as a fixed effect in the species model as its inclusion improved model fit, likely by explaining some of the differences in diet

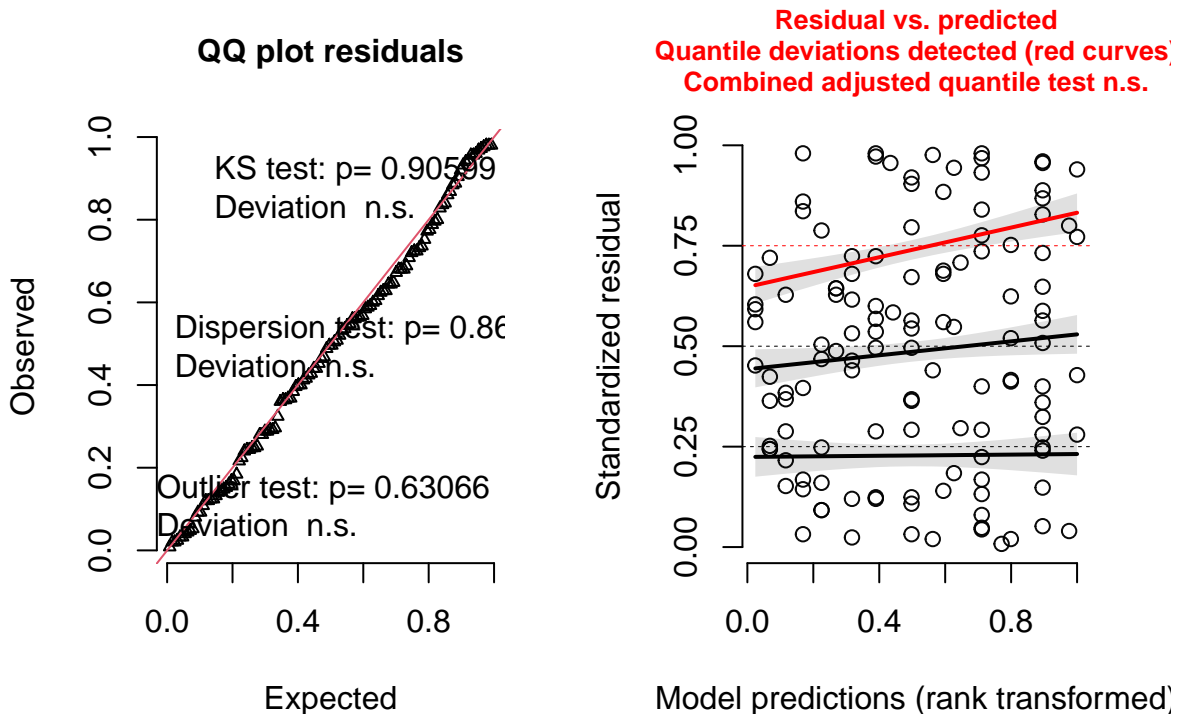Then checked the model assumptions:

```r
plot(m2)
```

```
qqnorm(resid(m2))
```

Normal Q–Q Plot
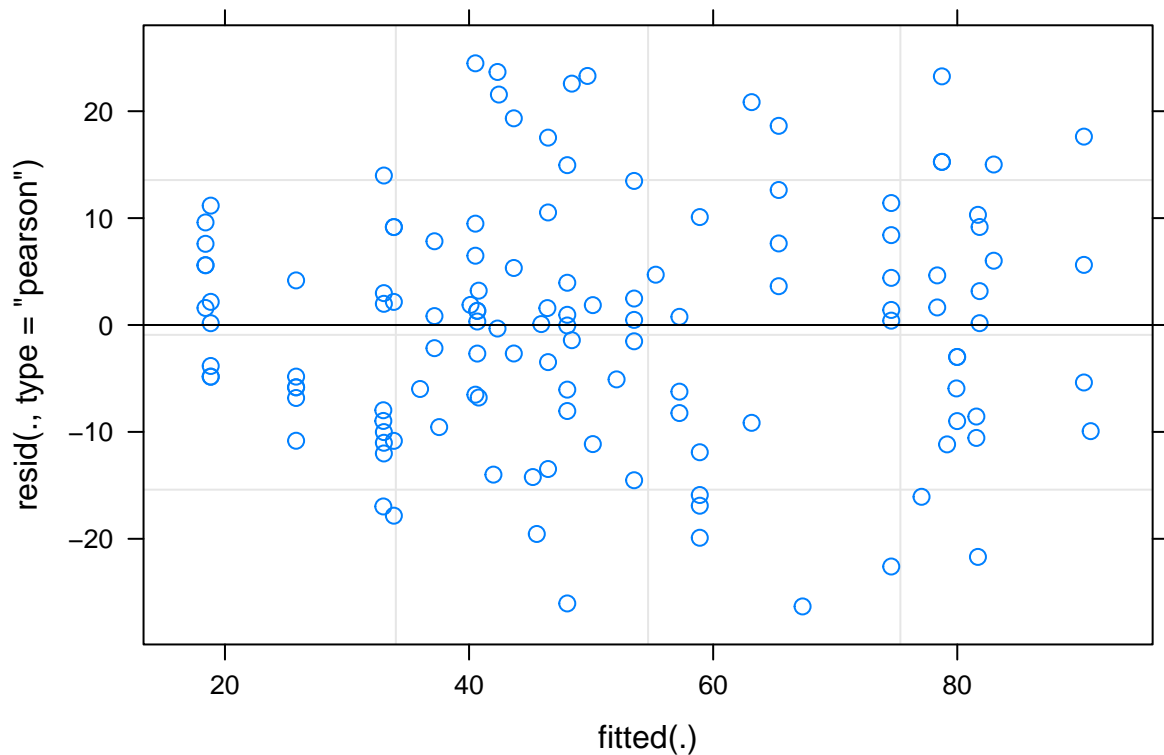

```
simoutbin500.2<-simulateResiduals(fittedModel=m2, n=250)
plot(simoutbin500.2)
```

DHARMa residual diagnostics

QQ plot residuals

KS test: p= 0.90599
Deviation n.s.

Dispersion test: p= 0.86
Deviation n.s.

Outlier test: p= 0.63066
Deviation n.s.

Observed (y-axis) / Expected (x-axis)

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test n.s.

Standardized residual (y-axis) / Model predictions (rank transformed) (x-axis)


```r
m2.1<-lmer(observed_otus ~ Collection + (1|Location), data=AlDiv.df, na.action="na.fail", REML=F)
anova(m2,m2.1) # value for species
```

```
## Data: AlDiv.df
## Models:
## m2.1: observed_otus ~ Collection + (1 | Location)
## m2: observed_otus ~ Species + Collection + (1 | Location)
##      npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## m2.1    5 1117.9 1132.1 -553.95  1107.89
## m2     14 1010.6 1050.3 -491.28   982.56 125.34  9  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
m2.2 <-lmer(observed_otus ~ Species + (1|Location), data=AlDiv.df, na.action="na.fail", REML=F)
anova(m2, m2.2) # value for collection substrate
```

```
## Data: AlDiv.df
## Models:
## m2.2: observed_otus ~ Species + (1 | Location)
## m2: observed_otus ~ Species + Collection + (1 | Location)
##      npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## m2.2   12 1011.1 1045.1 -493.54   987.07
## m2     14 1010.6 1050.3 -491.28   982.56 4.5198  2     0.1044
```

####Tukey Tests

Next I ran a tukey test to see which diet significantly different from eachother.

```
#install.packages("emmeans")
library(emmeans)
packageVersion("emmeans")
```

```
## [1] '1.6.2.1'
```

```
cox1.lm1<-lsmeans(m1, specs="Necrophagy")
collection_p.h <- pairs(cox1.lm1)
collection_p.h
```

```
##  contrast              estimate   SE   df t.ratio p.value
##  Absent - Facultative     -27.1 11.4 14.1  -2.368  0.0786
##  Absent - Obligate        -10.7 18.5 13.8  -0.576  0.8348
##  Facultative - Obligate    16.4 17.1 13.5   0.958  0.6145
##
## Degrees-of-freedom method: kenward-roger
## P value adjustment: tukey method for comparing a family of 3 estimates
```

```
write.csv(collection_p.h, "FinalPlots/revised_necro_p_hoc_updated.tsv")
```

No significant differences in diet.

I ran a Tukey test on species to get pairwise comparisons.

```
cox3.lm2<-lsmeans(m2, specs="Species")
species_p.h <- pairs(cox3.lm2)
species_p.h
```

```
##  contrast                     estimate   SE    df t.ratio p.value
##  angustula - buchwaldi         -0.4270 6.69 133.2  -0.064  1.0000
##  angustula - corvina          -14.5687 6.89 141.8  -2.113  0.5216
##  angustula - ferricauda       -28.4060 6.59 133.3  -4.311  0.0013
##  angustula - flavoscutellata  -14.6120 5.47 121.4  -2.673  0.1967
##  angustula - fulviventris     -15.4327 5.98 138.9  -2.583  0.2359
##  angustula - musarum          -48.2417 6.64 134.8  -7.261  <.0001
##  angustula - necrophaga       -14.7246 6.68 121.2  -2.204  0.4598
##  angustula - orizabaensis     -56.1853 5.89 137.5  -9.540  <.0001
##  angustula - silvestriana     -27.5738 7.31 124.7  -3.772  0.0091
##  buchwaldi - corvina          -14.1417 7.19 132.7  -1.967  0.6231
##  buchwaldi - ferricauda       -27.9790 7.22 141.3  -3.876  0.0061
##  buchwaldi - flavoscutellata  -14.1850 6.15 122.6  -2.306  0.3926
##  buchwaldi - fulviventris     -15.0057 6.40 131.6  -2.346  0.3674
##  buchwaldi - musarum          -47.8147 7.16 140.6  -6.677  <.0001
##  buchwaldi - necrophaga       -14.2975 7.14 142.1  -2.002  0.5990
##  buchwaldi - orizabaensis     -55.7583 6.32 126.7  -8.826  <.0001
##  buchwaldi - silvestriana     -27.1468 7.80 142.6  -3.479  0.0227
##  corvina - ferricauda         -13.8372 6.73 140.8  -2.055  0.5624
##  corvina - flavoscutellata     -0.0433 6.27 138.4  -0.007  1.0000
##  corvina - fulviventris        -0.8640 6.20 137.6  -0.139  1.0000
##  corvina - musarum            -33.6730 6.59 136.6  -5.110  <.0001
```

```
##  corvina - necrophaga          -0.1558 6.22 132.4  -0.025  1.0000
##  corvina - orizabaensis       -41.6166 6.38 140.1  -6.525  <.0001
##  corvina - silvestriana       -13.0051 6.58 119.9  -1.977  0.6167
##  ferricauda - flavoscutellata  13.7939 6.20 124.0   2.224  0.4465
##  ferricauda - fulviventris     12.9732 5.53 136.3   2.348  0.3658
##  ferricauda - musarum         -19.8358 4.53 130.2  -4.375  0.0010
##  ferricauda - necrophaga       13.6814 5.15  97.8   2.659  0.2056
##  ferricauda - orizabaensis    -27.7793 5.83 106.3  -4.767  0.0002
##  ferricauda - silvestriana      0.8321 7.04 134.7   0.118  1.0000
##  flavoscutellata - fulviventris -0.8207 5.25 139.7  -0.156  1.0000
##  flavoscutellata - musarum    -33.6297 6.10 135.4  -5.513  <.0001
##  flavoscutellata - necrophaga  -0.1125 5.97 129.4  -0.019  1.0000
##  flavoscutellata - orizabaensis -41.5733 5.16 139.7 -8.057  <.0001
##  flavoscutellata - silvestriana -12.9618 6.54 137.6 -1.981  0.6137
##  fulviventris - musarum       -32.8090 5.47 139.5  -6.001  <.0001
##  fulviventris - necrophaga      0.7082 5.56 132.5   0.127  1.0000
##  fulviventris - orizabaensis  -40.7526 5.36 139.2  -7.610  <.0001
##  fulviventris - silvestriana  -12.1411 6.59 141.3  -1.843  0.7070
##  musarum - necrophaga          33.5172 4.50 140.9   7.455  <.0001
##  musarum - orizabaensis        -7.9436 5.70 117.8  -1.393  0.9271
##  musarum - silvestriana        20.6679 6.80 140.2   3.041  0.0805
##  necrophaga - orizabaensis    -41.4607 5.84 104.2  -7.103  <.0001
##  necrophaga - silvestriana    -12.8493 6.26 134.8  -2.053  0.5636
##  orizabaensis - silvestriana   28.6115 6.88 134.8   4.161  0.0022
##
## Results are averaged over the levels of: Collection
## Degrees-of-freedom method: kenward-roger
## P value adjustment: tukey method for comparing a family of 10 estimates
```

```r
write.csv(species_p.h, "FinalPlots/revised_species_post_hoc_updated.tsv")
```

There are several significant differences between species.

####Figure 2 ASVs by Species and Diet

```r
library(ggplot2)

#Order the Species by necrophagy
Sp.ord <- AlDiv.df
Sp.ord$Species_plot <- factor(Sp.ord$Species_plot,
                              levels = c("Te. angustula (n=11)",
                                         "Te. buchwaldi (n=6)",
                                         "Pl. flavoscutellata (n=17)",
                                         "Pa. musarum (n=19)",
                                         "Pa. orizabaensis (n=12)",
                                         "T. corvina (n=7)",
                                         "T. ferricauda (n=17)",
                                         "T. fulviventris (n=12)",
                                         "T. silvestriana (n=7)",
                                         "T. necrophaga (n=18)"))

#Plot by species
bp.sp <- ggplot(Sp.ord, aes(x=Species_plot, y= observed_otus, fill=Necrophagy)) +
        geom_boxplot() +
```

```
        theme_classic()+
        scale_fill_manual(values=c("darkolivegreen","goldenrod1","darkorchid"))+
        theme(axis.text.x = element_text(angle = 45, hjust= 1.0, size = 8)) +
        labs(y="ASVs", x="Species")+
        theme(legend.position="right")+
        ggtitle("ASVs by Species and Diet")+
        theme(plot.title = element_text(hjust = 0.5))
bp.sp
```
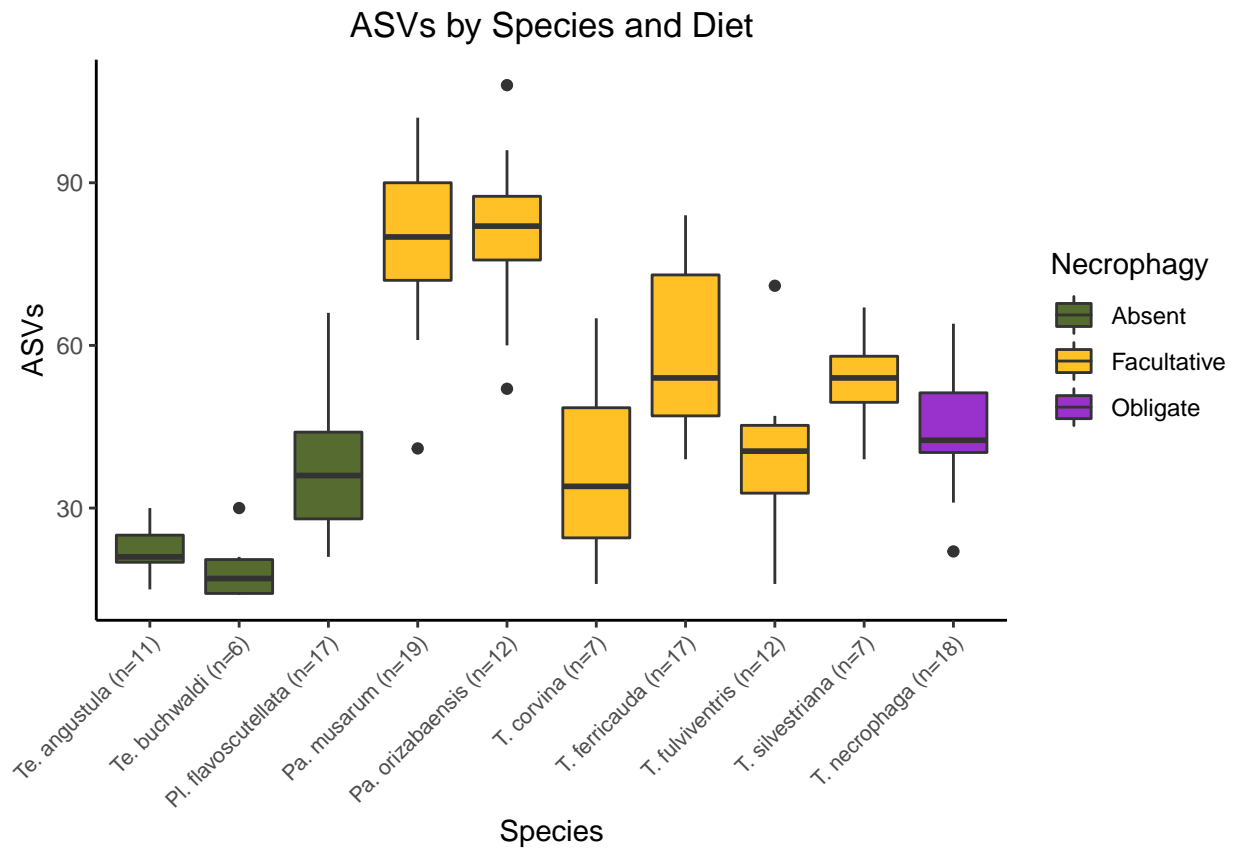


```
dev.print(pdf,          # copies the plot to a the PDF file
          "FinalPlots/K_W_plot_species.pdf")
```

```
## pdf
##   2
```