

Whole Lung Tissue vs. BAL Fluid 16S Sequencing Analysis

Jenn Baker

June 24, 2020

This report was created with:

R version 4.0.2 (2020-06-22)

Platform: x86_64-apple-darwin17.0 (64-bit)

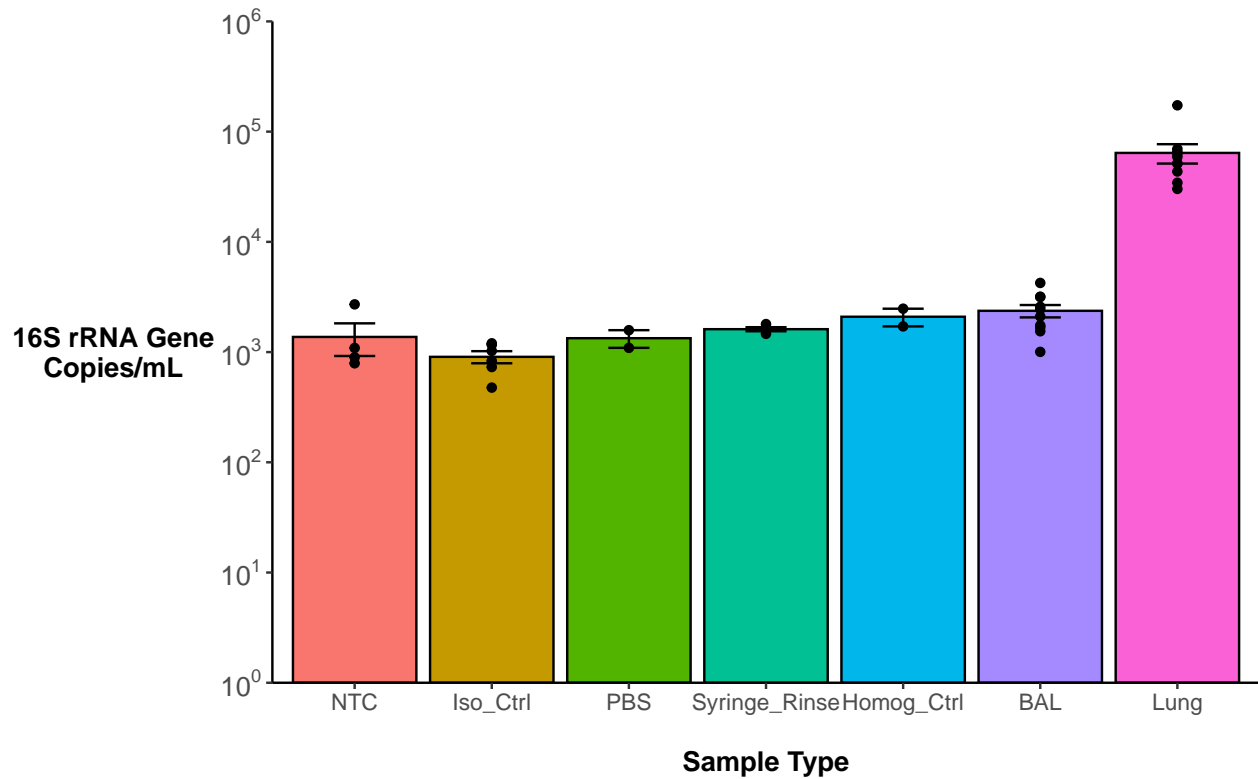
locale: en_US.UTF-8|en_US.UTF-8|en_US.UTF-8|C|en_US.UTF-8|en_US.UTF-8

attached base packages: *stats*, *graphics*, *grDevices*, *utils*, *datasets*, *methods* and *base*

other attached packages: *tadaatoolbox(v.0.17.0)*, *pander(v.0.6.3)*, *cbmbtools(v.0.0.0.9034)*, *vegan(v.2.5-6)*, *lattice(v.0.20-41)*, *permute(v.0.9-5)*, *ggpubr(v.0.4.0)*, *RColorBrewer(v.1.1-2)*, *knitr(v.1.29)*, *scales(v.1.1.1)*, *tidyselect(v.1.1.0)*, *forcats(v.0.5.0)*, *stringr(v.1.4.0)*, *dplyr(v.1.0.1)*, *purrr(v.0.3.4)*, *readr(v.1.3.1)*, *tidyr(v.1.1.1)*, *tibble(v.3.0.3)*, *ggplot2(v.3.3.2)* and *tidyverse(v.1.3.0)*

loaded via a namespace (and not attached): *nlme(v.3.1-148)*, *fs(v.1.5.0)*, *lubridate(v.1.7.9)*, *httr(v.1.4.2)*, *tools(v.4.0.2)*, *backports(v.1.1.8)*, *R6(v.2.4.1)*, *DBI(v.1.1.0)*, *mgcv(v.1.8-31)*, *colorspace(v.1.4-1)*, *withr(v.2.2.0)*, *Exact(v.2.0)*, *gridExtra(v.2.3)*, *pixiedust(v.0.9.0)*, *curl(v.4.3)*, *compiler(v.4.0.2)*, *cli(v.2.0.2)*, *rvest(v.0.3.6)*, *expm(v.0.999-5)*, *xml2(v.1.3.2)*, *mvtnorm(v.1.1-1)*, *digest(v.0.6.25)*, *foreign(v.0.8-80)*, *rmarkdown(v.2.3)*, *rio(v.0.5.16)*, *pkgconfig(v.2.0.3)*, *htmltools(v.0.5.0)*, *dbplyr(v.1.4.4)*, *rlang(v.0.4.7)*, *readxl(v.1.3.1)*, *rstudioapi(v.0.11)*, *generics(v.0.0.2)*, *jsonlite(v.1.7.0)*, *zip(v.2.0.4)*, *car(v.3.0-9)*, *magrittr(v.1.5)*, *Matrix(v.1.2-18)*, *Rcpp(v.1.0.5)*, *DescTools(v.0.99.37)*, *munsell(v.0.5.0)*, *fansi(v.0.4.1)*, *abind(v.1.4-5)*, *viridis(v.0.5.1)*, *lifecycle(v.0.2.0)*, *stringi(v.1.4.6)*, *yaml(v.2.2.1)*, *carData(v.3.0-4)*, *MASS(v.7.3-51.6)*, *grid(v.4.0.2)*, *blob(v.1.2.1)*, *parallel(v.4.0.2)*, *crayon(v.1.3.4)*, *haven(v.2.3.1)*, *splines(v.4.0.2)*, *hms(v.0.5.3)*, *pillar(v.1.4.6)*, *boot(v.1.3-25)*, *ggsignif(v.0.6.0)*, *codetools(v.0.2-16)*, *reprex(v.0.3.0)*, *glue(v.1.4.1)*, *evaluate(v.0.14)*, *data.table(v.1.13.0)*, *modelr(v.0.1.8)*, *vctrs(v.0.3.2)*, *cellranger(v.1.1.0)*, *gtable(v.0.3.0)*, *assertthat(v.0.2.1)*, *xfun(v.0.16)*, *openxlsx(v.4.1.5)*, *broom(v.0.7.0)*, *rstatix(v.0.6.0)*, *viridisLite(v.0.3.0)*, *cluster(v.2.1.0)* and *ellipsis(v.0.3.1)*

16S rRNA gene amplicon quantification



Hypothesis testing - 16S rRNA gene amplicon quantification

```
#overall difference between groups - kruskal wallis test for multiple group comparisons
kruskal.test(Gene_16S_copies_per_mL ~ Sample_Type, data=ddpcr)
# Kruskal-Wallis rank sum test
# data: Gene_16S_copies_per_mL by Sample_Type
# Kruskal-Wallis chi-squared = 30.451, df = 6, p-value = 3.227e-05

#pairwise difference between groups - Wilcoxon rank sum test, adjusted for multiple comparisons
pairwise.wilcox.test(ddpcr$Gene_16S_copies_per_mL, ddpcr$Sample_Type, p.adjust.method = "BH")
# Pairwise comparisons using Wilcoxon rank sum test
# data: ddpcr$Gene_16S_copies_per_mL and ddpcr$Sample_Type
```

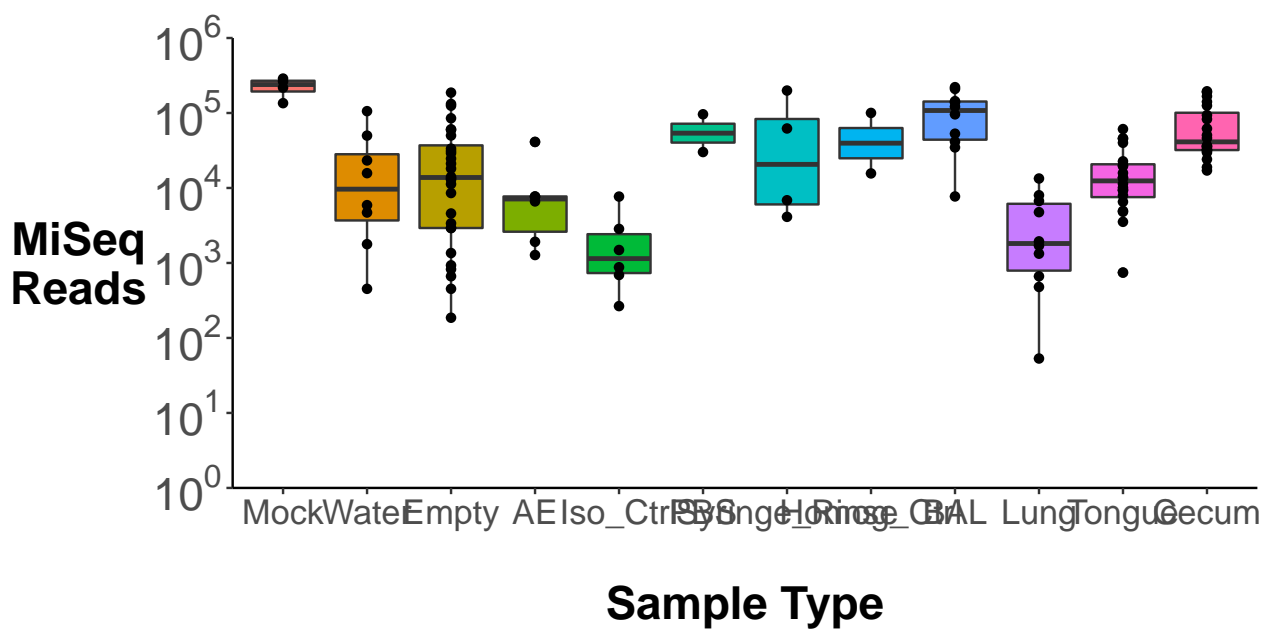
#	NTC	Iso_Ctrl	PBS	Syringe_Rinse	Homog_Ctrl	BAL
# Iso_Ctrl	0.64000	-	-	-	-	-
# PBS	0.58947	0.40000	-	-	-	-
# Syringe_Rinse	0.42353	0.03333	0.40000	-	-	-
# Homog_Ctrl	0.58947	0.16667	0.42353	0.40000	-	-
# BAL	0.22238	0.00734	0.31818	0.27082	0.90909	-
# Lung	0.00734	0.00170	0.06731	0.00734	0.06731	0.00012

P value adjustment method: BH

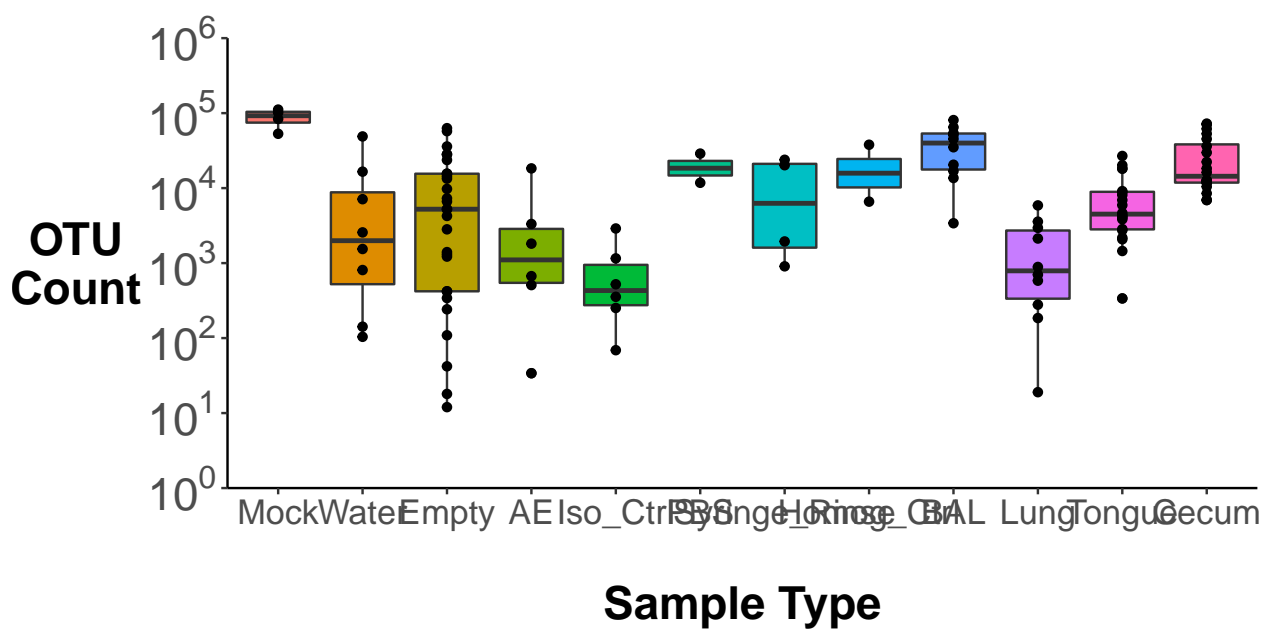
Quality Checks

III. Confirmation of Sufficient Reads

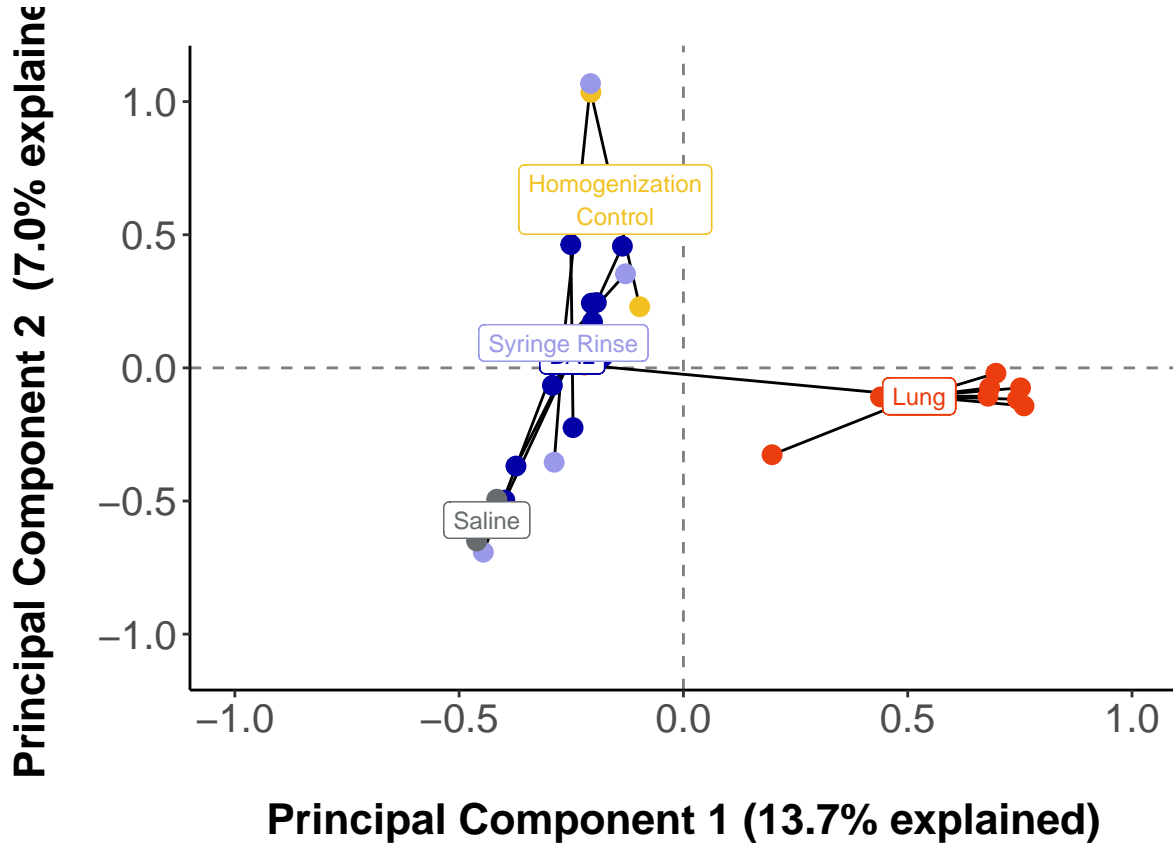
3.1 MiSeq Reads by Sample Type



3.2 OTU Counts by Sample Type



IV. Principal Component Analysis

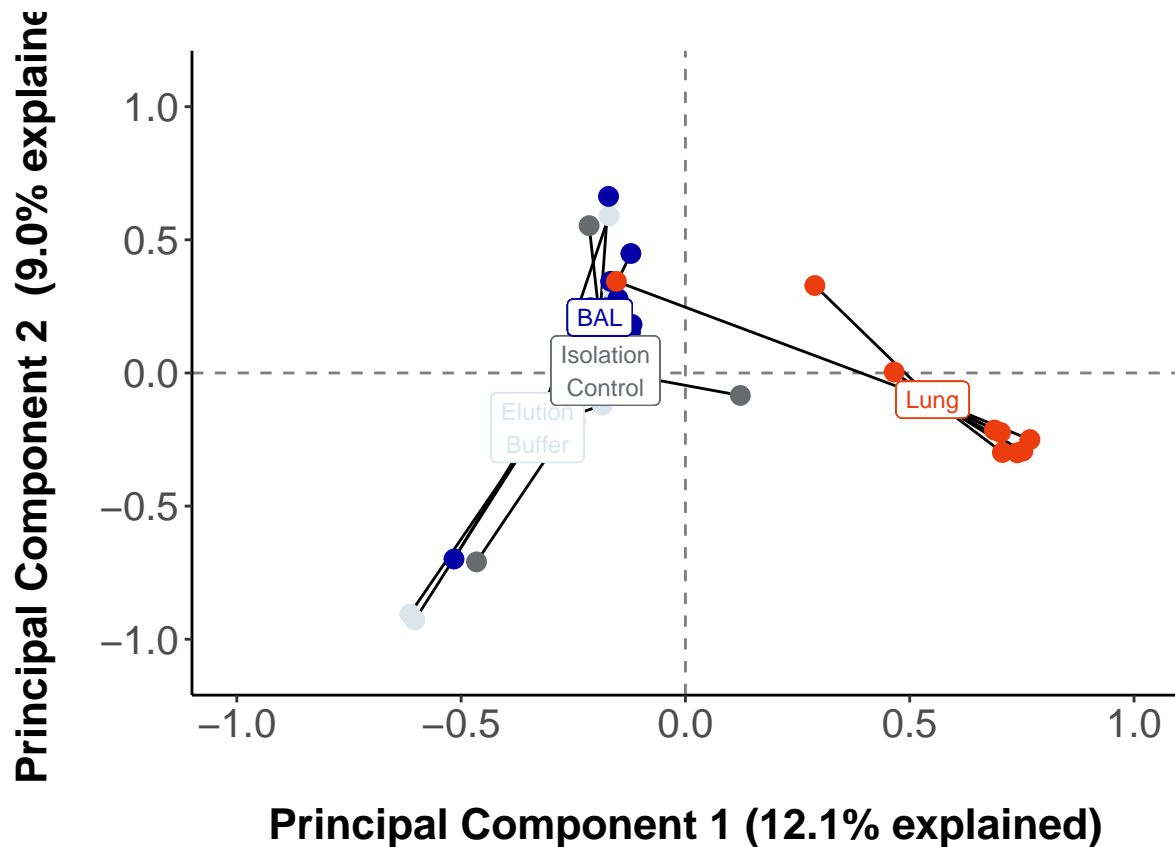


```
FALSE
FALSE Call:
FALSE adonis(formula = adonis.hel.df.wbn ~ otu.df.lungsamp.sampctrls$RA_Groups, permutations = 10000)
FALSE
FALSE Permutation: free
FALSE Number of permutations: 10000
FALSE
FALSE Terms added sequentially (first to last)
FALSE
FALSE      Df SumsOfSqs MeanSqs F.Model    R2
FALSE otu.df.lungsamp.sampctrls$RA_Groups  2    3.8879 1.94394  2.3806 0.16554
FALSE Residuals                        24   19.5980 0.81658    0.83446
FALSE Total                            26   23.4859          1.00000
FALSE      Pr(>F)
FALSE otu.df.lungsamp.sampctrls$RA_Groups 9.999e-05 ***
FALSE Residuals
FALSE Total
FALSE ---
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE Call:
FALSE adonis(formula = adonis.hel.bl ~ adonis.otudf.bl$Sample_Type, permutations = 10000, method = "bray")
FALSE
FALSE Permutation: free
```

```

FALSE Number of permutations: 10000
FALSE
FALSE Terms added sequentially (first to last)
FALSE
FALSE
      Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
FALSE adonis.otudf.bl$Sample_Type  1    2.4484  2.4484  3.1334 0.15563 2e-04 ***
FALSE Residuals                    17   13.2838  0.7814      0.84437
FALSE Total                        18   15.7323      1.00000
FALSE ---
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE Call:
FALSE adonis(formula = adonis.hel.wn ~ adonis.otudf.wn$RA_Groups, permutations = 10000,      method = "eucl
FALSE
FALSE Permutation: free
FALSE Number of permutations: 10000
FALSE
FALSE Terms added sequentially (first to last)
FALSE
      Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
FALSE adonis.otudf.wn$RA_Groups  1    2.4698  2.46976  3.3064 0.18061 3e-04 ***
FALSE Residuals                    15   11.2046  0.74697      0.81939
FALSE Total                        16   13.6744      1.00000
FALSE ---
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE Call:
FALSE adonis(formula = adonis.hel.bn ~ adonis.otudf.bn$Organ, permutations = 10000,      method = "eucl
FALSE
FALSE Permutation: free
FALSE Number of permutations: 10000
FALSE
FALSE Terms added sequentially (first to last)
FALSE
      Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
FALSE adonis.otudf.bn$Organ  1    0.9148  0.91483  0.99522 0.05856 0.4606
FALSE Residuals              16   14.7076  0.91922      0.94144
FALSE Total                  17   15.6224      1.00000

```



FALSE

FALSE Call:

FALSE `adonis(formula = otu.good.lungsamp.isoctrls_hel ~ otu.df.lungsamp.isoctrls$Sample_Type,`

`perm`

FALSE

FALSE Permutation: free

FALSE Number of permutations: 10000

FALSE

FALSE Terms added sequentially (first to last)

FALSE

	Df	SumsOfSqs	MeanSqs	F.Model	R2
otu.df.lungsamp.isoctrls\$Sample_Type	3	5.0693	1.6898	2.0376	0.1846
Residuals	27	22.3910	0.8293		0.8154
Total	30	27.4603			1.0000

FALSE Pr(>F)

FALSE `otu.df.lungsamp.isoctrls$Sample_Type` 9.999e-05 ***

FALSE Residuals

FALSE Total

FALSE ---

FALSE Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

FALSE

FALSE Call:

FALSE `adonis(formula = adonis.hel.il ~ adonis.otudf.il$RA_Groups, permutations = 10000,`

`method = "`

FALSE

FALSE Permutation: free

FALSE Number of permutations: 10000

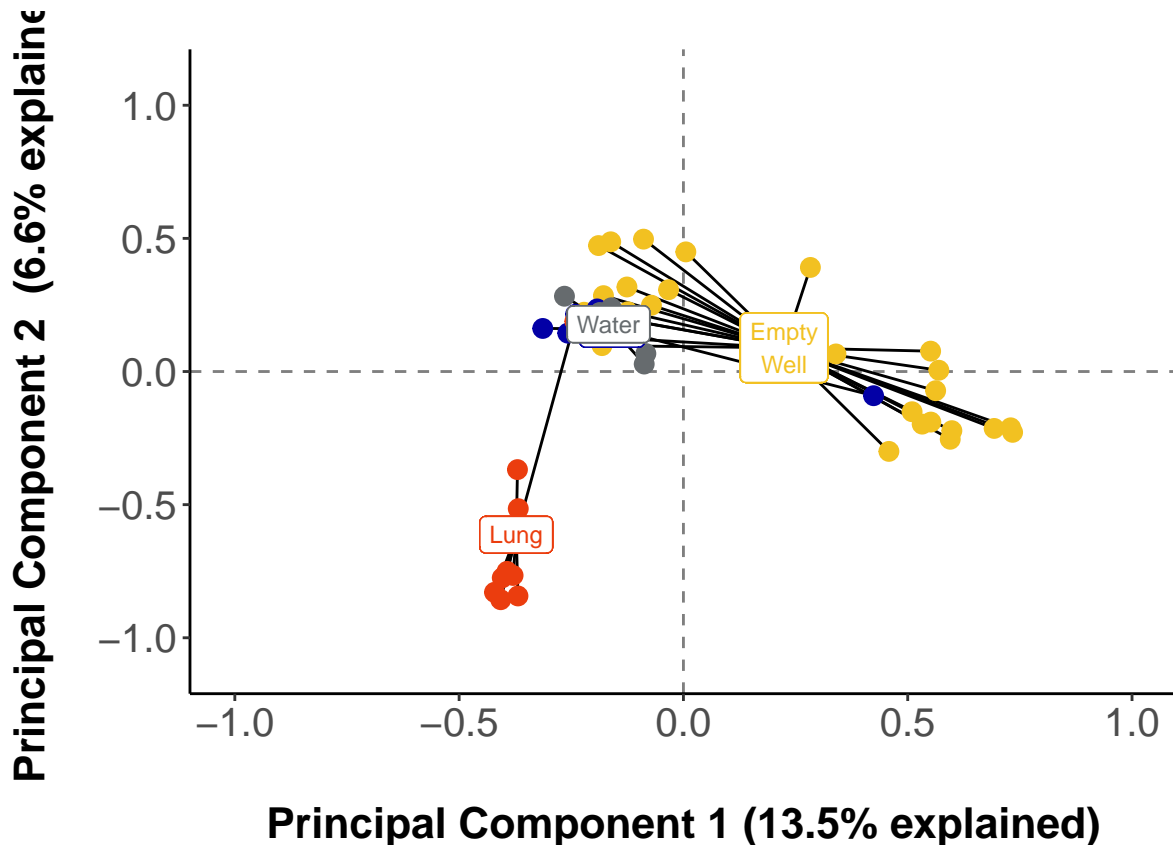
FALSE

FALSE Terms added sequentially (first to last)

```

FALSE
FALSE
FALSE      Df SumsOfSqs MeanSqs F.Model      R2      Pr(>F)
FALSE adonis.otudf.il$RA_Groups  1      2.6804 2.68040  3.4196 0.15253 9.999e-05 ***
FALSE Residuals      19      14.8927 0.78383      0.84747
FALSE Total          20      17.5731      1.00000
FALSE ---
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE Call:
FALSE adonis(formula = adonis.hel.ib ~ adonis.otudf.ib$RA_Groups, permutations = 10000,      method = "
FALSE
FALSE Permutation: free
FALSE Number of permutations: 10000
FALSE
FALSE Terms added sequentially (first to last)
FALSE
FALSE      Df SumsOfSqs MeanSqs F.Model      R2      Pr(>F)
FALSE adonis.otudf.ib$RA_Groups  1      1.2087 1.20866  1.3141 0.06165 0.07229 .
FALSE Residuals      20      18.3957 0.91978      0.93835
FALSE Total          21      19.6044      1.00000
FALSE ---
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```



```

FALSE
FALSE Call:
FALSE adonis(formula = otu.good.lungsamp.seqctrls_hel ~ otu.df.lungsamp.seqctrls$Sample_Type,      permu
FALSE

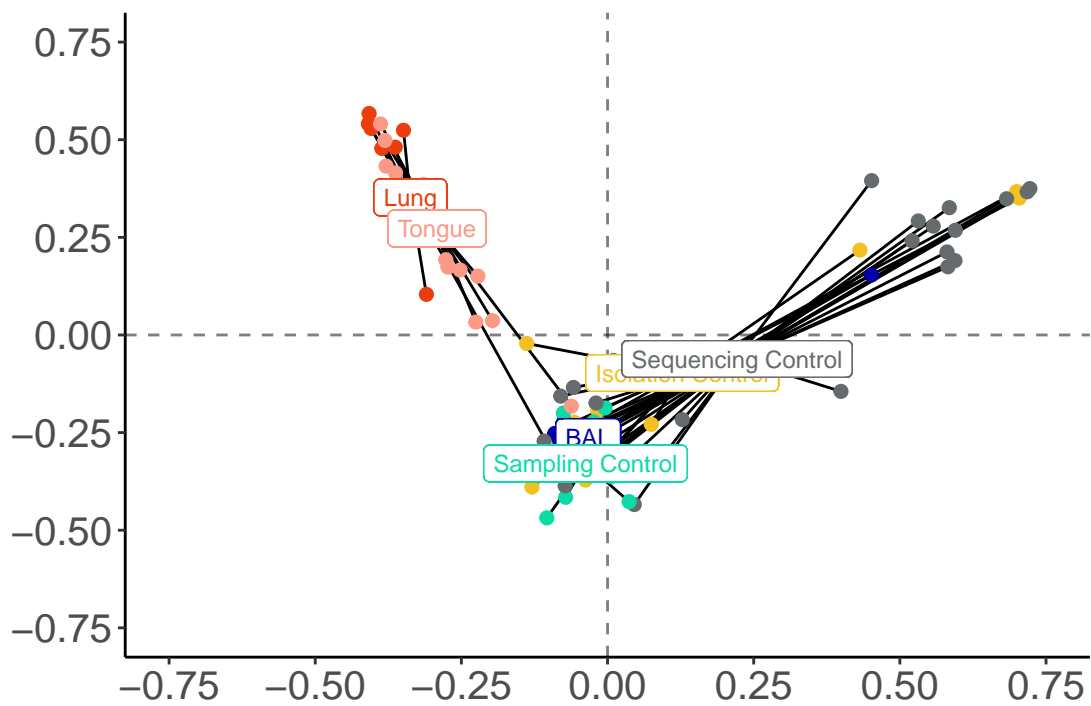
```

```

FALSE Permutation: free
FALSE Number of permutations: 10000
FALSE
FALSE Terms added sequentially (first to last)
FALSE
FALSE
FALSE Df SumsOfSqs MeanSqs F.Model R2
FALSE otu.df.lungsamp.seqctrls$Sample_Type 3 7.919 2.63959 3.2847 0.16193
FALSE Residuals 51 40.984 0.80361 0.83807
FALSE Total 54 48.903 1.00000
FALSE Pr(>F)
FALSE otu.df.lungsamp.seqctrls$Sample_Type 9.999e-05 ***
FALSE Residuals
FALSE Total
FALSE ---
FALSE Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE Call:
FALSE adonis(formula = adonis.hel.sl ~ adonis.otudf.sl$RA_Groups, permutations = 10000, method = "
FALSE
FALSE Permutation: free
FALSE Number of permutations: 10000
FALSE
FALSE Terms added sequentially (first to last)
FALSE
FALSE Df SumsOfSqs MeanSqs F.Model R2 Pr(>F)
FALSE adonis.otudf.sl$RA_Groups 1 4.146 4.1459 5.1613 0.10717 9.999e-05 ***
FALSE Residuals 43 34.541 0.8033 0.89283
FALSE Total 44 38.687 1.00000
FALSE ---
FALSE Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE Call:
FALSE adonis(formula = adonis.hel.sb ~ adonis.otudf.sb$RA_Groups, permutations = 10000, method = "
FALSE
FALSE Permutation: free
FALSE Number of permutations: 10000
FALSE
FALSE Terms added sequentially (first to last)
FALSE
FALSE Df SumsOfSqs MeanSqs F.Model R2 Pr(>F)
FALSE adonis.otudf.sb$RA_Groups 1 2.069 2.06931 2.3933 0.05159 0.0011 **
FALSE Residuals 44 38.044 0.86463 0.94841
FALSE Total 45 40.113 1.00000
FALSE ---
FALSE Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

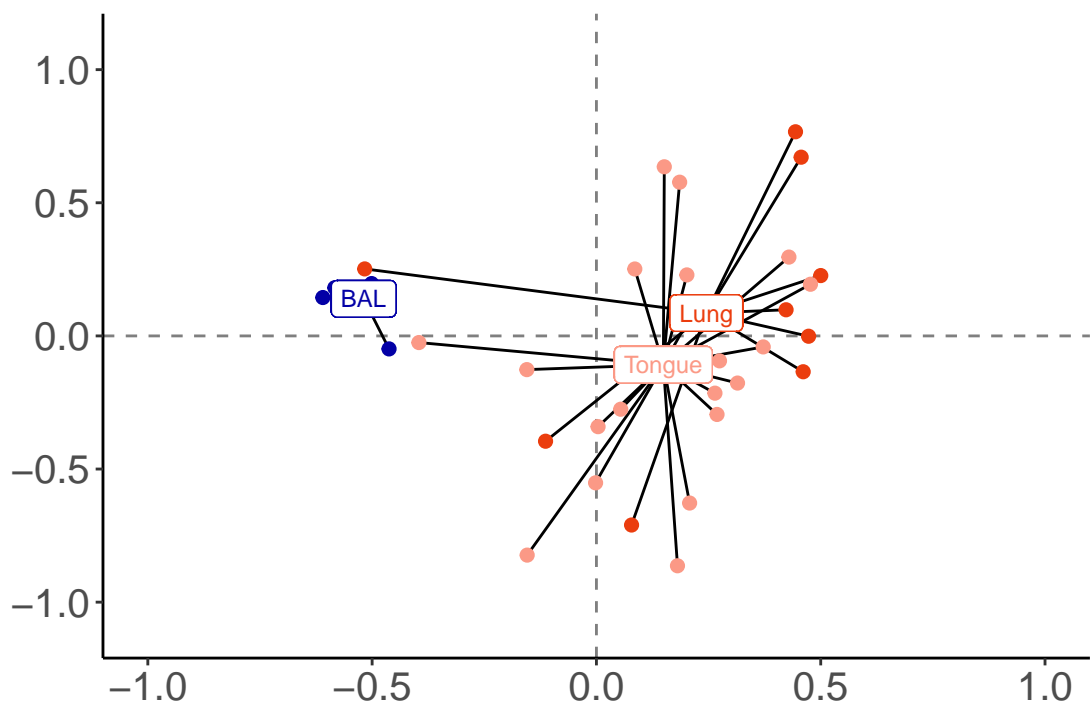
```


Principal Component 2 (6.3% explained)



Principal Component 1 (11.6% explained)

Principal Component 2 (4.6% explained)



Principal Component 1 (10.5% explained)

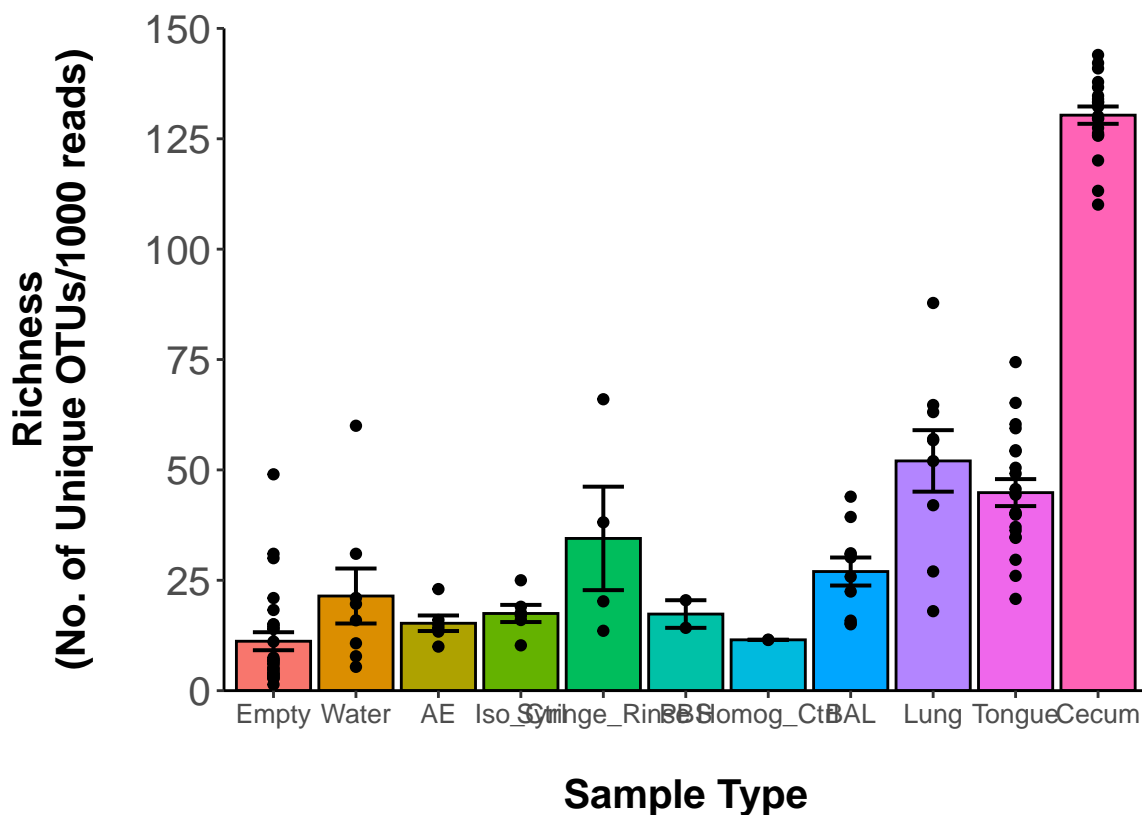
```

FALSE
FALSE Call:
FALSE adonis(formula = otu.good.tong.lungsamp_hel ~ otu.df.tong.lungsamp$Sample_Type,      permutations
FALSE
FALSE Permutation: free
FALSE Number of permutations: 10000
FALSE
FALSE Terms added sequentially (first to last)
FALSE
FALSE
FALSE      Df SumsOfSqs MeanSqs F.Model      R2      Pr(>F)
FALSE otu.df.tong.lungsamp$Sample_Type  2      3.9663 1.98317  2.6355 0.12772 9.999e-05
FALSE Residuals      36      27.0894 0.75248      0.87228
FALSE Total      38      31.0557      1.00000
FALSE
FALSE otu.df.tong.lungsamp$Sample_Type ***
FALSE Residuals
FALSE Total
FALSE ---
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE Call:
FALSE adonis(formula = adonis.hel.tl ~ adonis.otudf.tl$Sample_Type,      permutations = 10000, method =
FALSE
FALSE Permutation: free
FALSE Number of permutations: 10000
FALSE
FALSE Terms added sequentially (first to last)
FALSE
FALSE      Df SumsOfSqs MeanSqs F.Model      R2 Pr(>F)
FALSE adonis.otudf.tl$Sample_Type  1      1.039 1.03903  1.5005 0.05265 0.0138 *
FALSE Residuals      27      18.696 0.69244      0.94735
FALSE Total      28      19.735      1.00000
FALSE ---
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
FALSE
FALSE Call:
FALSE adonis(formula = adonis.hel.tb ~ adonis.otudf.tb$Sample_Type,      permutations = 10000, method =
FALSE
FALSE Permutation: free
FALSE Number of permutations: 10000
FALSE
FALSE Terms added sequentially (first to last)
FALSE
FALSE      Df SumsOfSqs MeanSqs F.Model      R2      Pr(>F)
FALSE adonis.otudf.tb$Sample_Type  1      2.6012 2.60116  3.2809 0.10489 9.999e-05 ***
FALSE Residuals      28      22.1989 0.79282      0.89511
FALSE Total      29      24.8001      1.00000
FALSE ---
FALSE Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

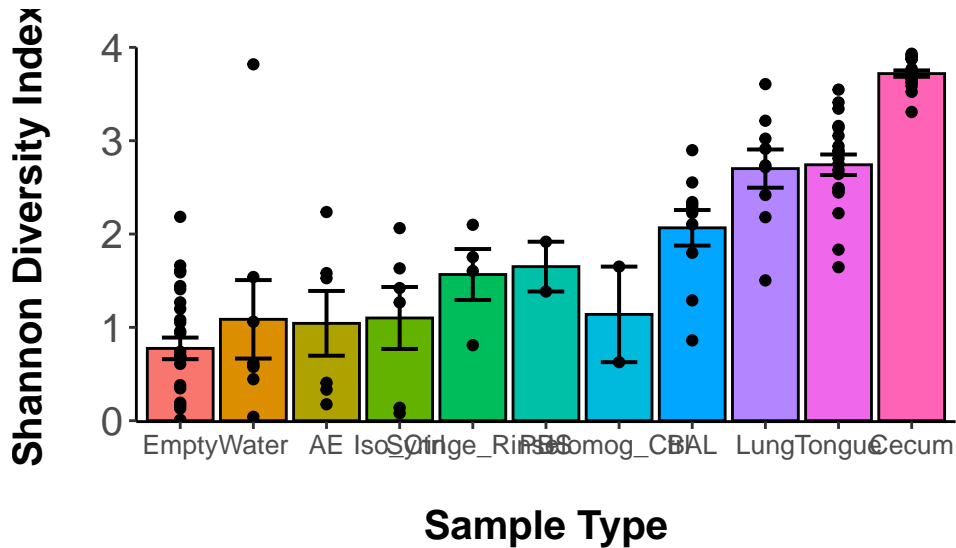
V. Diversity Indices

5.1 Alpha Diversity



```
FALSE Tukey multiple comparisons of means
FALSE 95% family-wise confidence level
FALSE
FALSE Fit: aov(formula = tukey_otu_df[, "Unique_Otus_per_1k_reads"] ~ tukey_otu_df[, "Alpha_Div_Groups"]
FALSE
FALSE $`tukey_otu_df[, "Alpha_Div_Groups"]`
FALSE      diff      lwr      upr
FALSE Isolation Control-BAL -10.612291 -27.006526  5.7819429
FALSE Lung-BAL 25.049970  7.457534 42.6424053
FALSE Sampling Control-BAL -2.528455 -20.690372 15.6334619
FALSE Sequencing Control-BAL -13.515401 -27.202081  0.1712778
FALSE Lung-Isolation Control 35.662261 18.778539 52.5459833
FALSE Sampling Control-Isolation Control 8.083837 -9.392476 25.5601491
FALSE Sequencing Control-Isolation Control -2.903110 -15.666004  9.8597843
FALSE Sampling Control-Lung -27.578425 -46.183380 -8.9734692
FALSE Sequencing Control-Lung -38.565371 -52.834721 -24.2960216
FALSE Sequencing Control-Sampling Control -10.986946 -25.952766  3.9788736
FALSE      p adj
FALSE Isolation Control-BAL 0.3746804
FALSE Lung-BAL 0.0014826
FALSE Sampling Control-BAL 0.9950044
FALSE Sequencing Control-BAL 0.0545862
FALSE Lung-Isolation Control 0.0000011
```

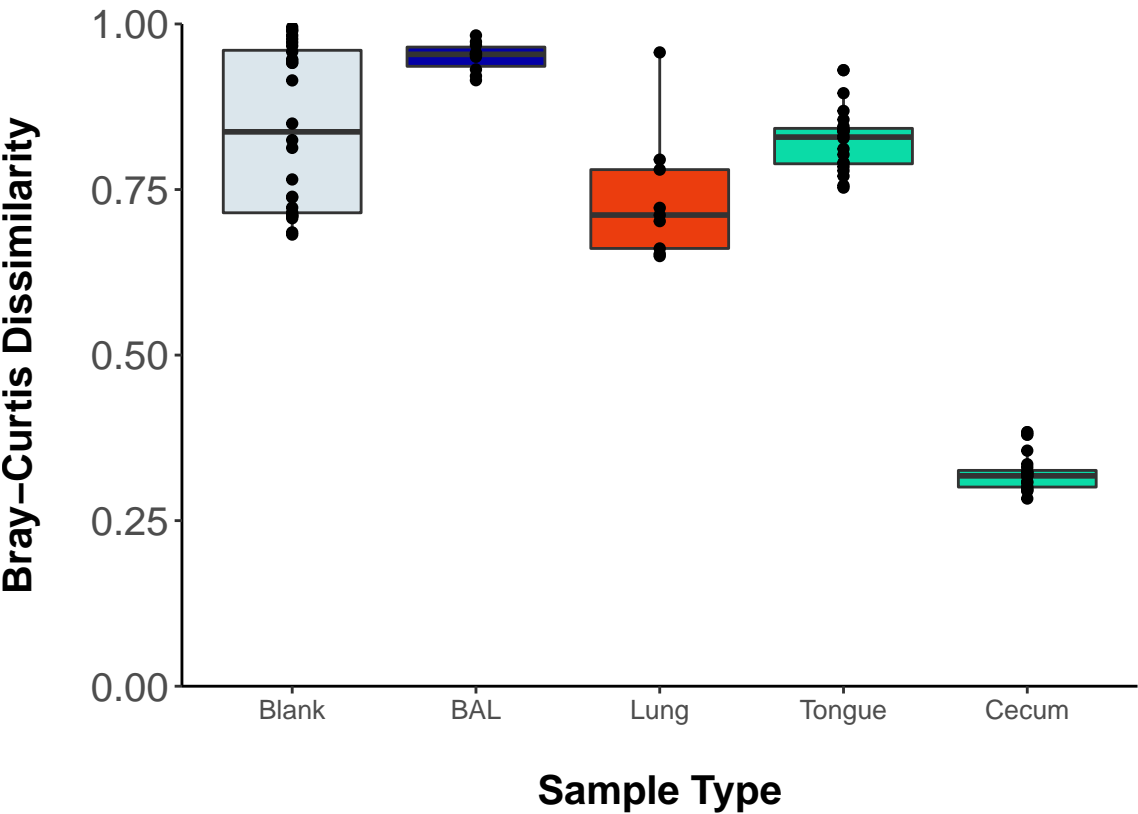
```
FALSE Sampling Control-Isolation Control 0.6949822
FALSE Sequencing Control-Isolation Control 0.9684857
FALSE Sampling Control-Lung 0.0008519
FALSE Sequencing Control-Lung 0.0000000
FALSE Sequencing Control-Sampling Control 0.2510366
```



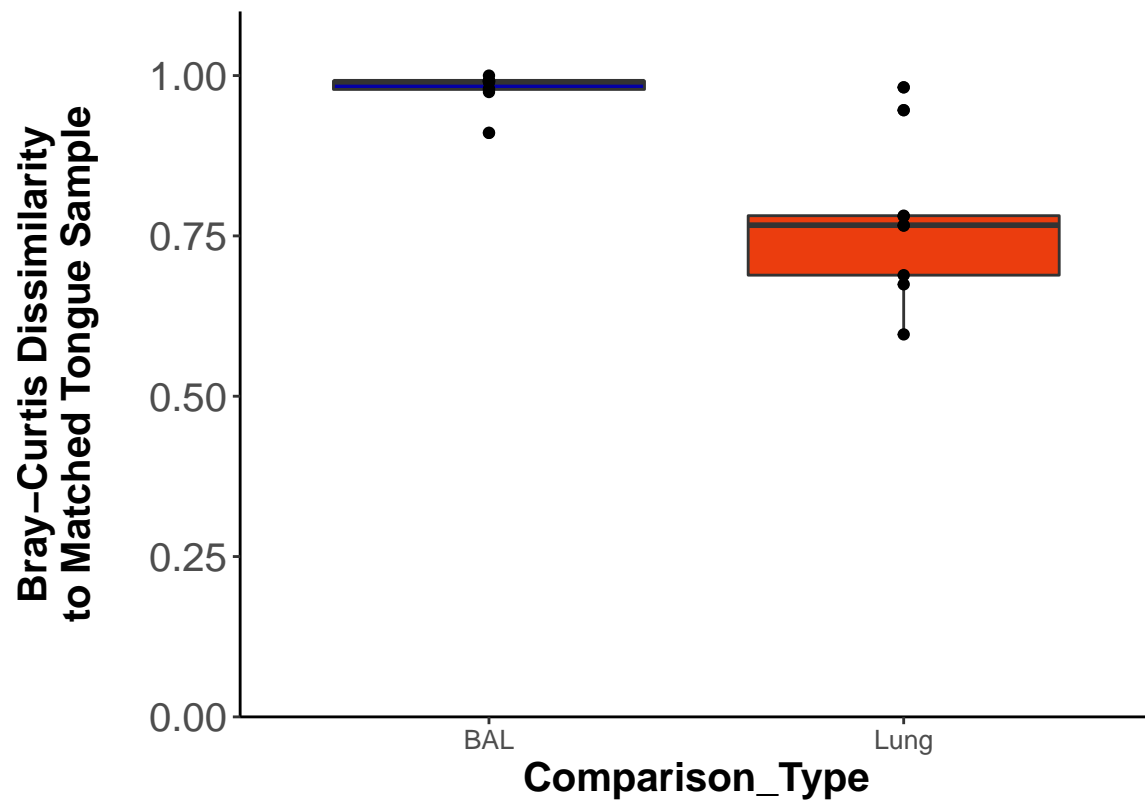
```
FALSE Tukey multiple comparisons of means
FALSE 95% family-wise confidence level
FALSE
FALSE Fit: aov(formula = tukey_otu_df[, "Shannon"] ~ tukey_otu_df[, "Alpha_Div_Groups"])
FALSE
FALSE $`tukey_otu_df[, "Alpha_Div_Groups"]`
FALSE
```

	diff	lwr	upr	p adj
FALSE Isolation Control-BAL	-0.9947478	-1.8514868	-0.1380087	0.0147639
FALSE Lung-BAL	0.6346623	-0.2846930	1.5540176	0.3098538
FALSE Sampling Control-BAL	-0.5859416	-1.5350572	0.3631740	0.4234102
FALSE Sequencing Control-BAL	-1.2226705	-1.9379166	-0.5074244	0.0000876
FALSE Lung-Isolation Control	1.6294101	0.7470911	2.5117290	0.0000205
FALSE Sampling Control-Isolation Control	0.4088062	-0.5044807	1.3220931	0.7201786
FALSE Sequencing Control-Isolation Control	-0.2279227	-0.8948932	0.4390477	0.8732743
FALSE Sampling Control-Lung	-1.2206039	-2.1928720	-0.2483357	0.0067338
FALSE Sequencing Control-Lung	-1.8573328	-2.6030284	-1.1116372	0.0000000
FALSE Sequencing Control-Sampling Control	-0.6367289	-1.4188211	0.1453633	0.1637833

5.2 Beta Diversity



```
FALSE
FALSE Pairwise comparisons using Wilcoxon rank sum test with continuity correction
FALSE
FALSE data: bray_dist_lbtce_long_mut_filt$BC_Index and bray_dist_lbtce_long_mut_filt$Sample_Type
FALSE
FALSE      BAL      Blank      Cecum      Lung
FALSE Blank 0.0074 - - -
FALSE Cecum < 2e-16 < 2e-16 - -
FALSE Lung < 2e-16 3.7e-11 < 2e-16 -
FALSE Tongue < 2e-16 < 2e-16 < 2e-16 0.0001
FALSE
FALSE P value adjustment method: BH
```



```
FALSE
FALSE Wilcoxon rank sum exact test
FALSE
FALSE data: BC_Index by Comparison_Type
FALSE W = 85, p-value = 0.0004114
FALSE alternative hypothesis: true location shift is not equal to 0
```

VI. Relative Abundance

