

Whole Lung Tissue vs. BAL Fluid 16S Sequencing Analysis (Trimmed/QC'd Final Dataset)

Jennifer M. Baker

August 18, 2020 (updated)

This report was created with:

R version 4.0.2 (2020-06-22)

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

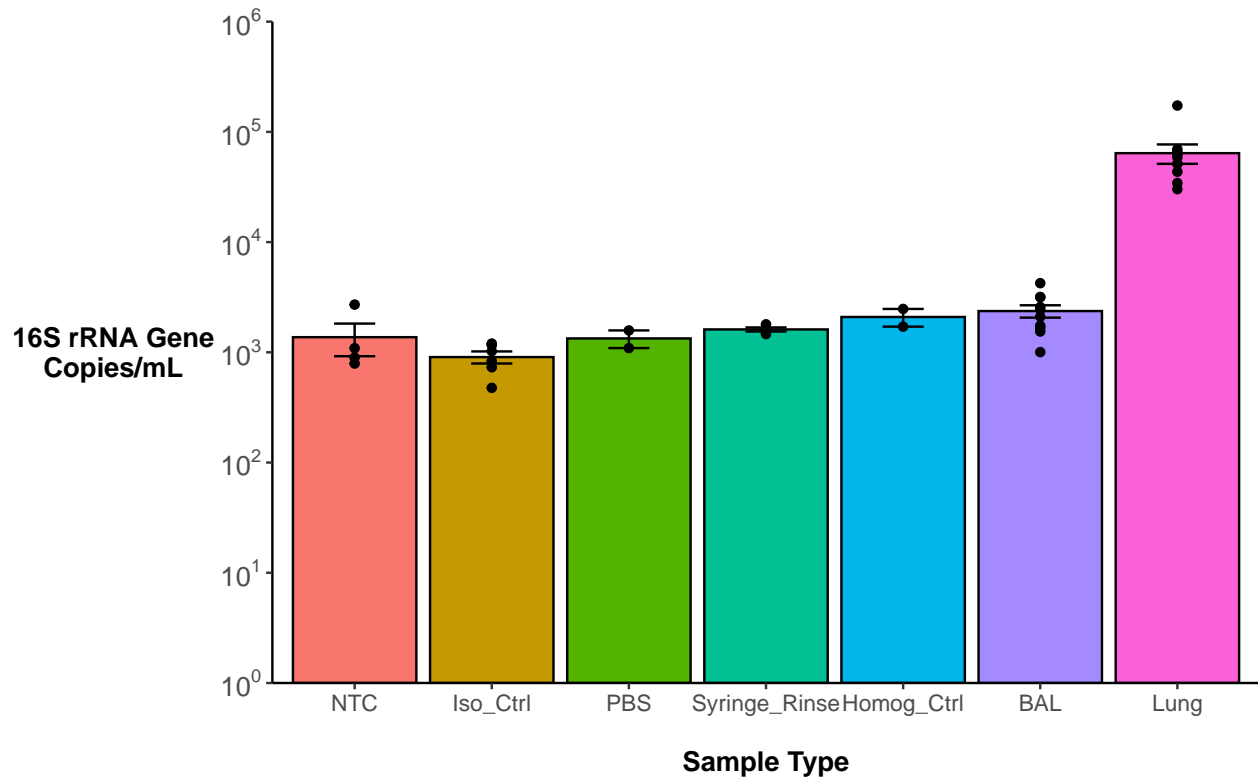
locale: en_US.UTF-8|en_US.UTF-8|en_US.UTF-8|C|en_US.UTF-8|en_US.UTF-8

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

other attached packages: *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)*, *knitr(v.1.29)*, *scales(v.1.1.1)*, *gtools(v.3.8.2)*, *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): *httr(v.1.4.2)*, *jsonlite(v.1.7.0)*, *splines(v.4.0.2)*, *carData(v.3.0-4)*, *modelr(v.0.1.8)*, *assertthat(v.0.2.1)*, *blob(v.1.2.1)*, *cellranger(v.1.1.0)*, *yaml(v.2.2.1)*, *pillar(v.1.4.6)*, *backports(v.1.1.8)*, *glue(v.1.4.1)*, *digest(v.0.6.25)*, *ggsignif(v.0.6.0)*, *rvest(v.0.3.6)*, *colorspace(v.1.4-1)*, *Matrix(v.1.2-18)*, *htmltools(v.0.5.0)*, *pkgconfig(v.2.0.3)*, *broom(v.0.7.0)*, *haven(v.2.3.1)*, *openxlsx(v.4.1.5)*, *rio(v.0.5.16)*, *mgcv(v.1.8-31)*, *generics(v.0.0.2)*, *car(v.3.0-9)*, *ellipsis(v.0.3.1)*, *withr(v.2.2.0)*, *cli(v.2.0.2)*, *magrittr(v.1.5)*, *crayon(v.1.3.4)*, *readxl(v.1.3.1)*, *evaluate(v.0.14)*, *fs(v.1.5.0)*, *fansi(v.0.4.1)*, *nlme(v.3.1-148)*, *MASS(v.7.3-51.6)*, *rstatix(v.0.6.0)*, *xmll2(v.1.3.2)*, *foreign(v.0.8-80)*, *tools(v.4.0.2)*, *data.table(v.1.13.0)*, *hms(v.0.5.3)*, *lifecycle(v.0.2.0)*, *munsell(v.0.5.0)*, *reprex(v.0.3.0)*, *cluster(v.2.1.0)*, *zip(v.2.0.4)*, *compiler(v.4.0.2)*, *rlang(v.0.4.7)*, *grid(v.4.0.2)*, *rstudioapi(v.0.11)*, *rmarkdown(v.2.3)*, *codetools(v.0.2-16)*, *gtable(v.0.3.0)*, *abind(v.1.4-5)*, *DBI(v.1.1.0)*, *curl(v.4.3)*, *R6(v.2.4.1)*, *lubridate(v.1.7.9)*, *stringi(v.1.4.6)*, *parallel(v.4.0.2)*, *Rcpp(v.1.0.5)*, *vctrs(v.0.3.2)*, *dbplyr(v.1.4.4)* and *xfun(v.0.16)*

I. 16S rRNA gene amplicon quantification



Hypothesis testing - 16S rRNA gene amplicon quantification

Kruskal-Wallis rank sum test

data: Gene_16S_copies_per_mL by Sample_Type
 Kruskal-Wallis chi-squared = 29.276, df = 6, p-value = 5.394e-05

Pairwise comparisons using Wilcoxon rank sum exact test

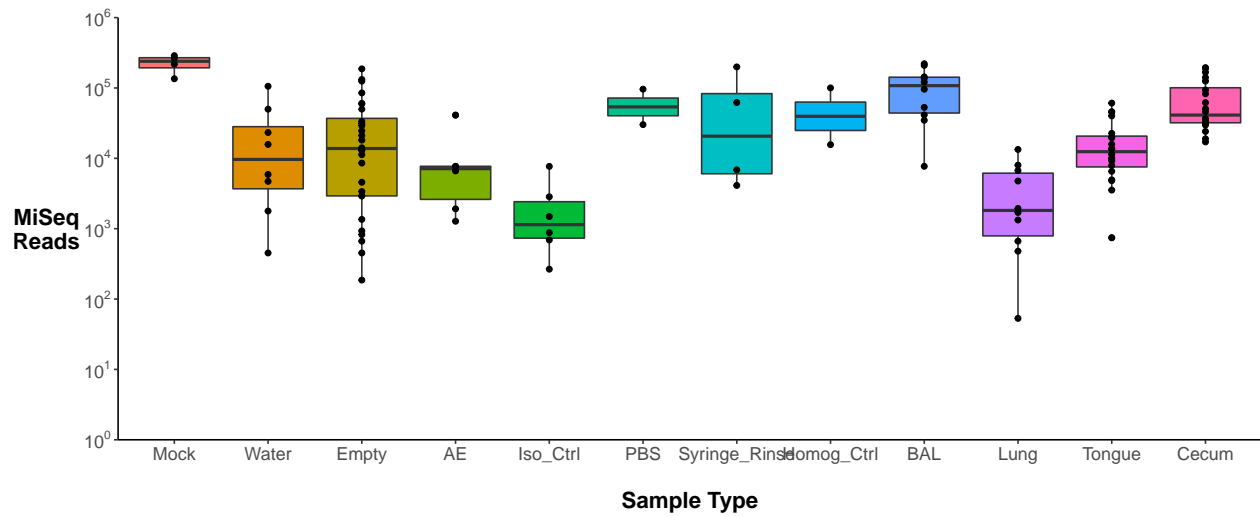
data: ddpcr\$Gene_16S_copies_per_mL and ddpcr\$Sample_Type

	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	0.00839	0.31818	0.27082	0.90909	-
Lung	0.00839	0.00262	0.07955	0.00839	0.07955	0.00023

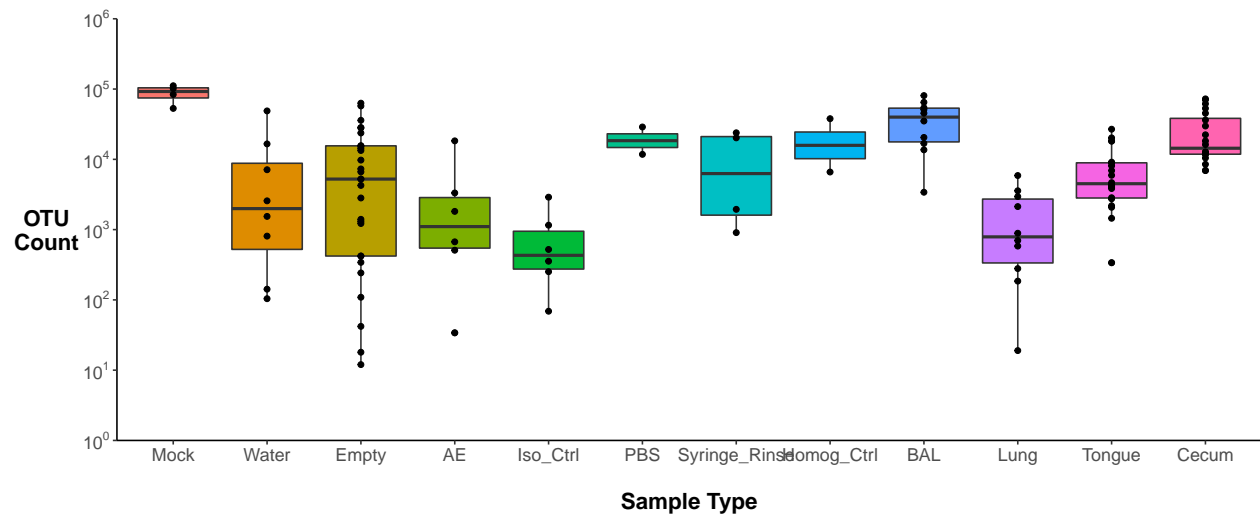
P value adjustment method: BH

II. Quality Checks - Confirmation of Sufficient Reads

2.1 MiSeq Reads by Sample Type

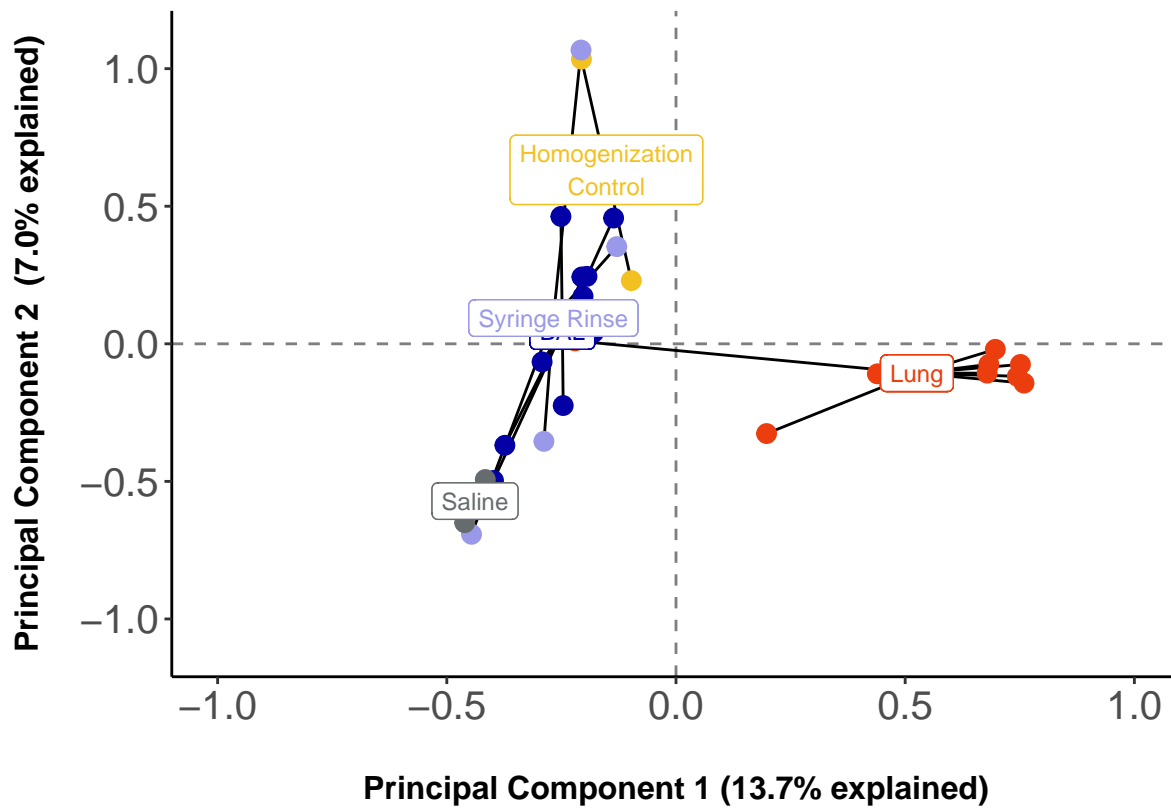


2.2 OTU Counts by Sample Type



III. Principal Component Analysis

3.1 Lung Samples vs. Sampling Controls



Call:

```
adonis(formula = adonis.hel.df.wbn ~ otu.df.lungsamp.sampctrls$RA_Groups, permutations = 10000, me
```

Permutation: free

Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2
otu.df.lungsamp.sampctrls\$RA_Groups	2	3.8879	1.94394	2.3806	0.16554
Residuals	24	19.5980	0.81658		0.83446
Total	26	23.4859			1.00000

Pr(>F)

```
otu.df.lungsamp.sampctrls$RA_Groups 9.999e-05 ***
```

Residuals

Total

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

Call:

```
adonis(formula = adonis.hel.bl ~ adonis.otudf.bl$Sample_Type, permutations = 10000, method = "eucl
```

Permutation: free

Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
adonis.otudf.bl\$Sample_Type	1	2.4484	2.4484	3.1334	0.15563	9.999e-05 ***
Residuals	17	13.2838	0.7814		0.84437	
Total	18	15.7323			1.00000	

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

Call:

adonis(formula = adonis.hel.wn ~ adonis.otudf.wn\$RA_Groups, permutations = 10000, method = "euclidean")

Permutation: free

Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
adonis.otudf.wn\$RA_Groups	1	2.4698	2.46976	3.3064	0.18061	4e-04 ***
Residuals	15	11.2046	0.74697		0.81939	
Total	16	13.6744			1.00000	

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

Call:

adonis(formula = adonis.hel.bn ~ adonis.otudf.bn\$Organ, permutations = 10000, method = "euclidean")

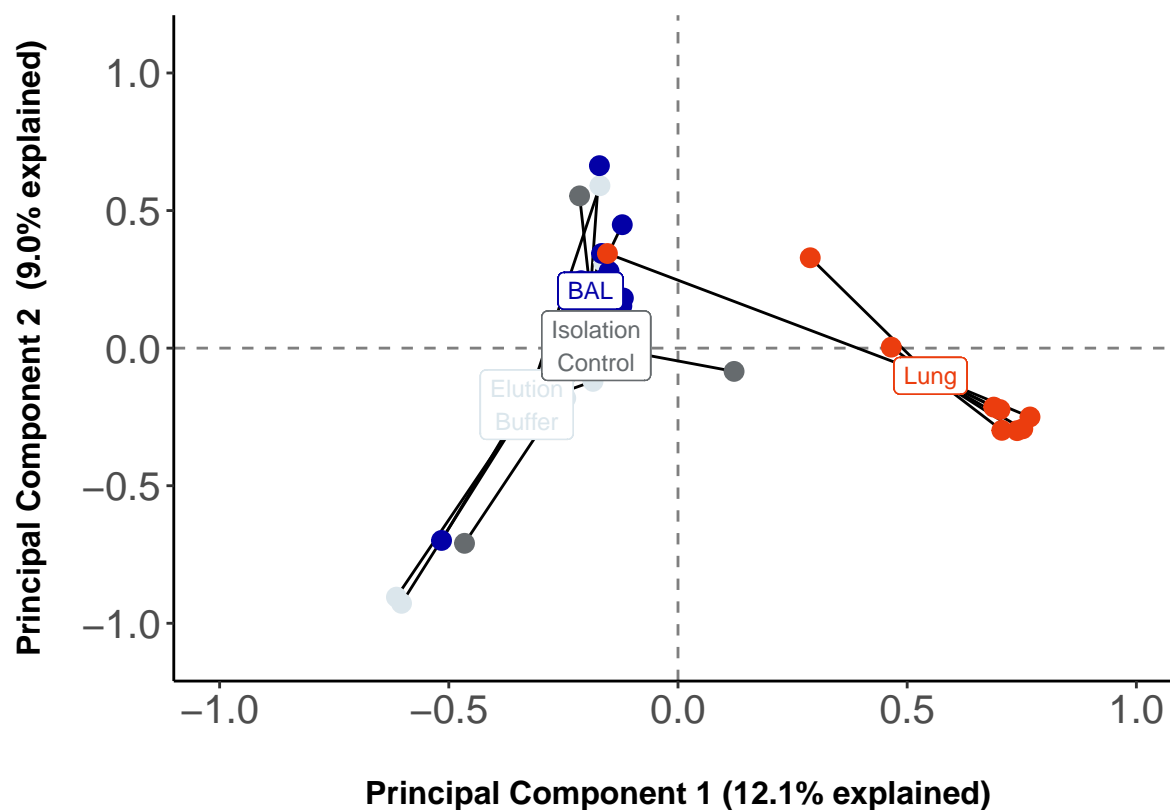
Permutation: free

Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
adonis.otudf.bn\$Organ	1	0.9148	0.91483	0.99522	0.05856	0.463
Residuals	16	14.7076	0.91922		0.94144	
Total	17	15.6224			1.00000	

3.2 Lung Samples v. Isolation Controls



Call:

```
adonis(formula = otu.good.lungsamp.isoctrls_hel ~ otu.df.lungsamp.isoctrls$Sample_Type,          permutation = 10000)
```

Permutation: free

Number of permutations: 10000

Terms added sequentially (first to last)

	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

Pr(>F)

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

Residuals

Total

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

Call:

```
adonis(formula = adonis.hel.il ~ adonis.otudf.il$RA_Groups, permutations = 10000,          method = "euclidean")
```

Permutation: free

Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
adonis.otudf.il\$RA_Groups	1	2.6804	2.68040	3.4196	0.15253	9.999e-05 ***
Residuals	19	14.8927	0.78383		0.84747	
Total	20	17.5731			1.00000	

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

Call:

adonis(formula = adonis.hel.ib ~ adonis.otudf.ib\$RA_Groups, permutations = 10000, method = "euclid

Permutation: free

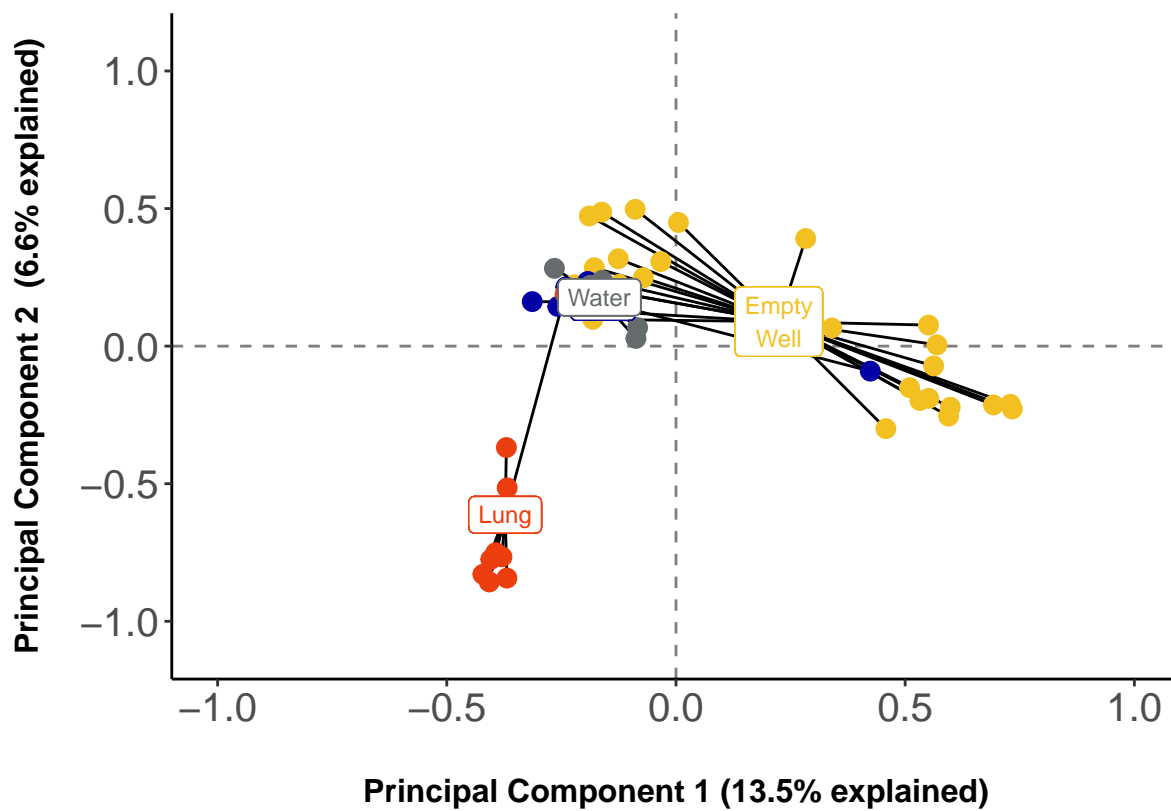
Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
adonis.otudf.ib\$RA_Groups	1	1.2087	1.20866	1.3141	0.06165	0.07149 .
Residuals	20	18.3957	0.91978		0.93835	
Total	21	19.6044			1.00000	

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

3.3 Lung Samples vs. Sequencing Controls



Call:

```
adonis(formula = otu.good.lungsamp.seqctrls_hel ~ otu.df.lungsamp.seqctrls$Sample_Type,          permutation = 10000)
```

Permutation: free

Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2
otu.df.lungsamp.seqctrls\$Sample_Type	3	7.919	2.63959	3.2847	0.16193
Residuals	51	40.984	0.80361		0.83807
Total	54	48.903			1.00000

Pr(>F)

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

Residuals

Total

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

Call:

```
adonis(formula = adonis.hel.sl ~ adonis.otudf.sl$RA_Groups, permutations = 10000,          method = "euclidean")
```

Permutation: free

Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
adonis.otudf.sl\$RA_Groups	1	4.146	4.1459	5.1613	0.10717	9.999e-05 ***
Residuals	43	34.541	0.8033		0.89283	
Total	44	38.687			1.00000	

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

Call:

adonis(formula = adonis.hel.sb ~ adonis.otudf.sb\$RA_Groups, permutations = 10000, method = "euclid

Permutation: free

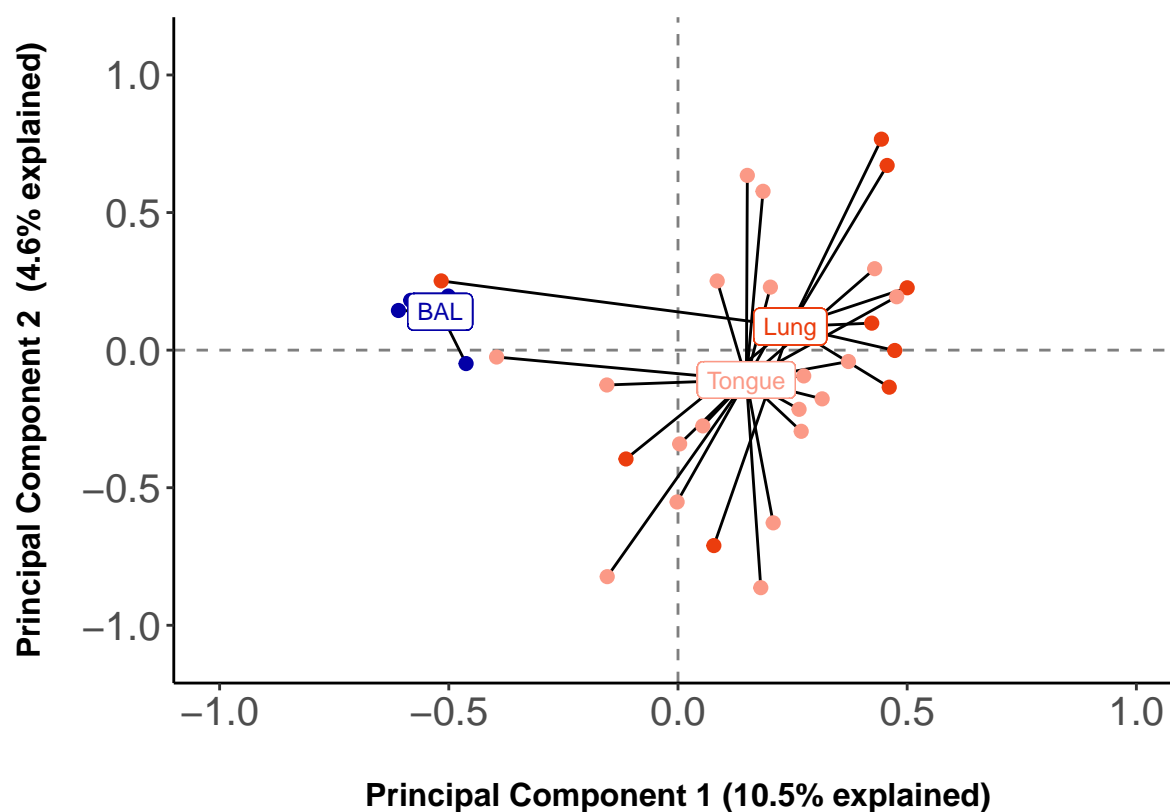
Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
adonis.otudf.sb\$RA_Groups	1	2.069	2.06931	2.3933	0.05159	4e-04 ***
Residuals	44	38.044	0.86463		0.94841	
Total	45	40.113			1.00000	

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

3.4 Lung Samples vs. Tongue



Call:
`adonis(formula = otu.good.tong.lungsamp_hel ~ otu.df.tong.lungsamp$Sample_Type, permutations = 10000)`

Permutation: free
 Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
otu.df.tong.lungsamp\$Sample_Type	2	3.9663	1.98317	2.6355	0.12772	9.999e-05
Residuals	36	27.0894	0.75248		0.87228	
Total	38	31.0557			1.00000	

otu.df.tong.lungsamp\$Sample_Type ***

Residuals

Total

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

Call:
`adonis(formula = adonis.hel.tl ~ adonis.otudf.tl$Sample_Type, permutations = 10000, method = "eucl")`

Permutation: free
 Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
adonis.otudf.tl\$Sample_Type	1	1.039	1.03903	1.5005	0.05265	0.0142 *
Residuals	27	18.696	0.69244		0.94735	
Total	28	19.735			1.00000	

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

Call:

adonis(formula = adonis.hel.tb ~ adonis.otudf.tb\$Sample_Type, permutations = 10000, method = "eucl

Permutation: free

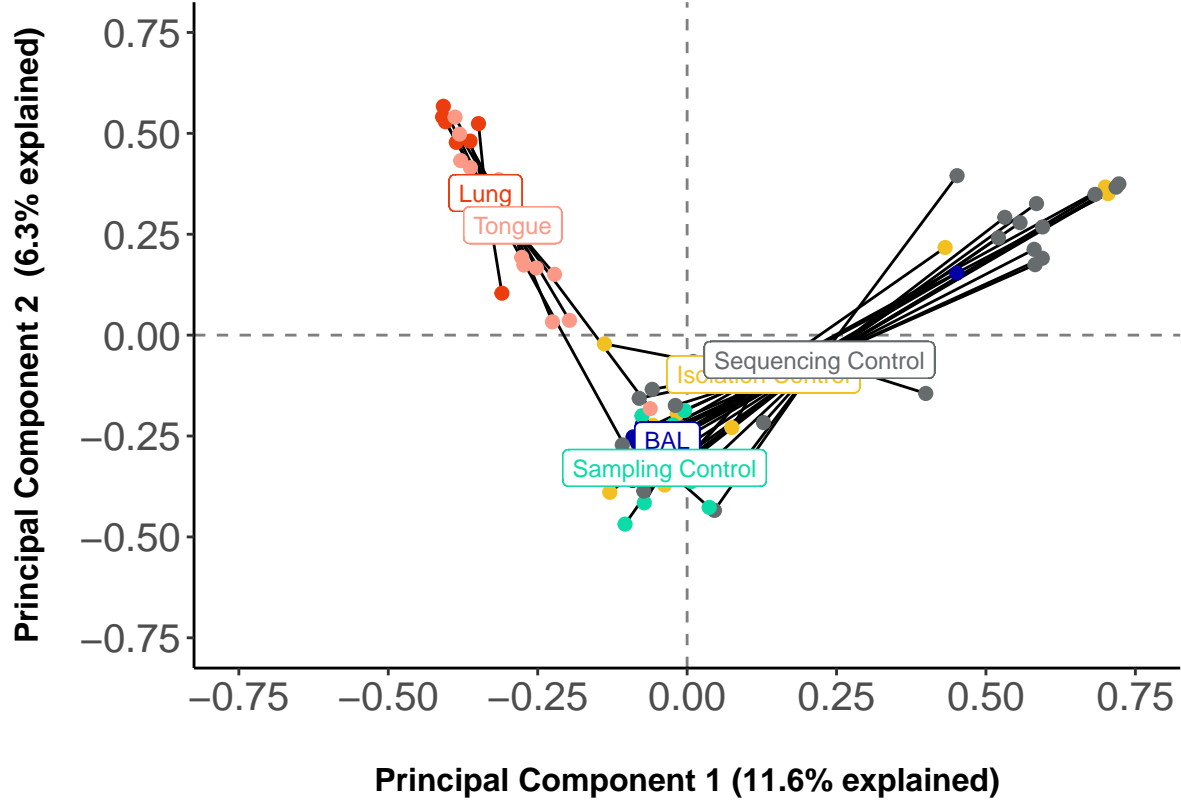
Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
adonis.otudf.tb\$Sample_Type	1	2.6012	2.60116	3.2809	0.10489	9.999e-05 ***
Residuals	28	22.1989	0.79282		0.89511	
Total	29	24.8001			1.00000	

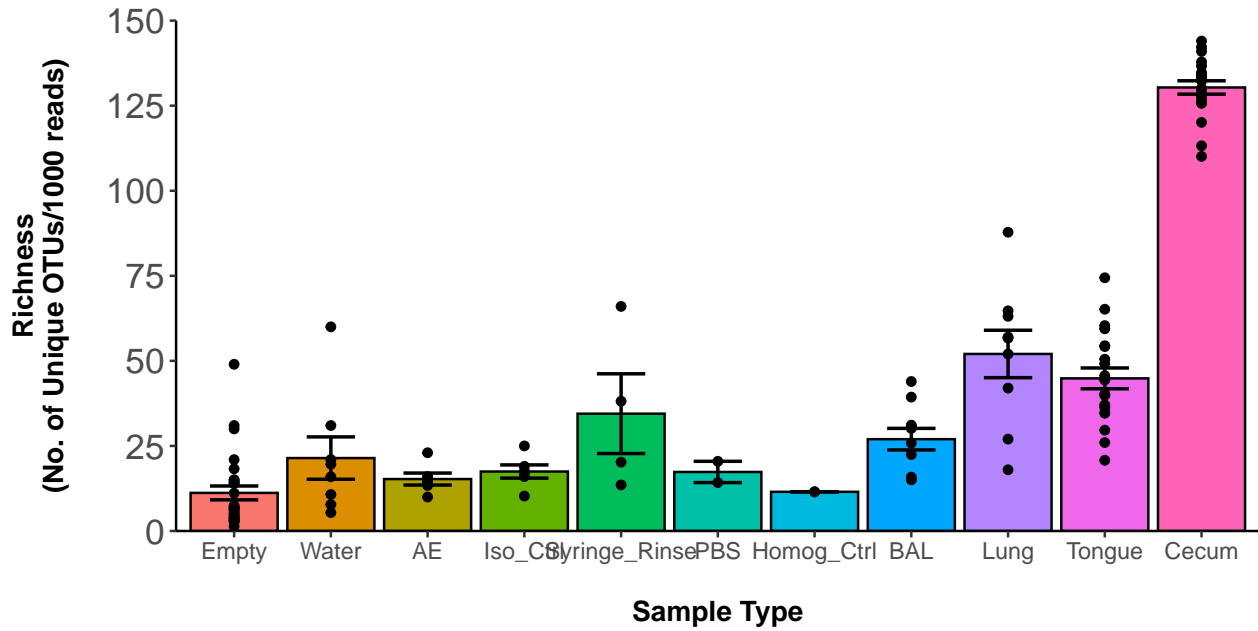
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

3.5 Lung vs. Tongue vs. Negative Controls



IV. Diversity Indices

4.1 Alpha Diversity - Rarified Richness



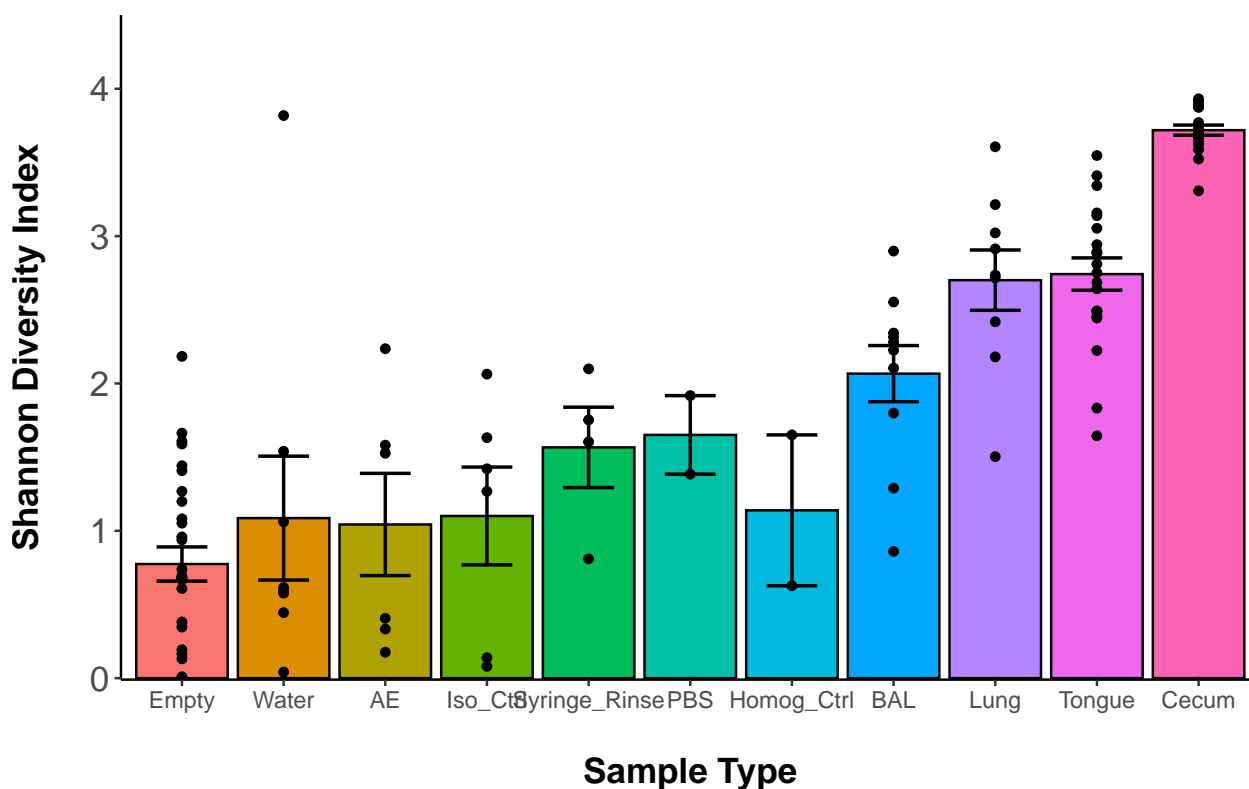
Tukey multiple comparisons of means
95% family-wise confidence level

```
Fit: aov(formula = tukey_otu_df[, "Unique_Otus_per_1k_reads"] ~ tukey_otu_df[, "Alpha_Div_Groups"])
```

```
$`tukey_otu_df[, "Alpha_Div_Groups"]`
```

	diff	lwr	upr
Isolation Control-BAL	-10.612291	-27.006526	5.7819429
Lung-BAL	25.049970	7.457534	42.6424053
Sampling Control-BAL	-2.528455	-20.690372	15.6334619
Sequencing Control-BAL	-13.515401	-27.202081	0.1712778
Lung-Isolation Control	35.662261	18.778539	52.5459833
Sampling Control-Isolation Control	8.083837	-9.392476	25.5601491
Sequencing Control-Isolation Control	-2.903110	-15.666004	9.8597843
Sampling Control-Lung	-27.578425	-46.183380	-8.9734692
Sequencing Control-Lung	-38.565371	-52.834721	-24.2960216
Sequencing Control-Sampling Control	-10.986946	-25.952766	3.9788736
	p adj		
Isolation Control-BAL	0.3746804		
Lung-BAL	0.0014826		
Sampling Control-BAL	0.9950044		
Sequencing Control-BAL	0.0545862		
Lung-Isolation Control	0.0000011		
Sampling Control-Isolation Control	0.6949822		
Sequencing Control-Isolation Control	0.9684857		
Sampling Control-Lung	0.0008519		
Sequencing Control-Lung	0.0000000		
Sequencing Control-Sampling Control	0.2510366		

4.2 Alpha Diversity - Shannon Diversity Index



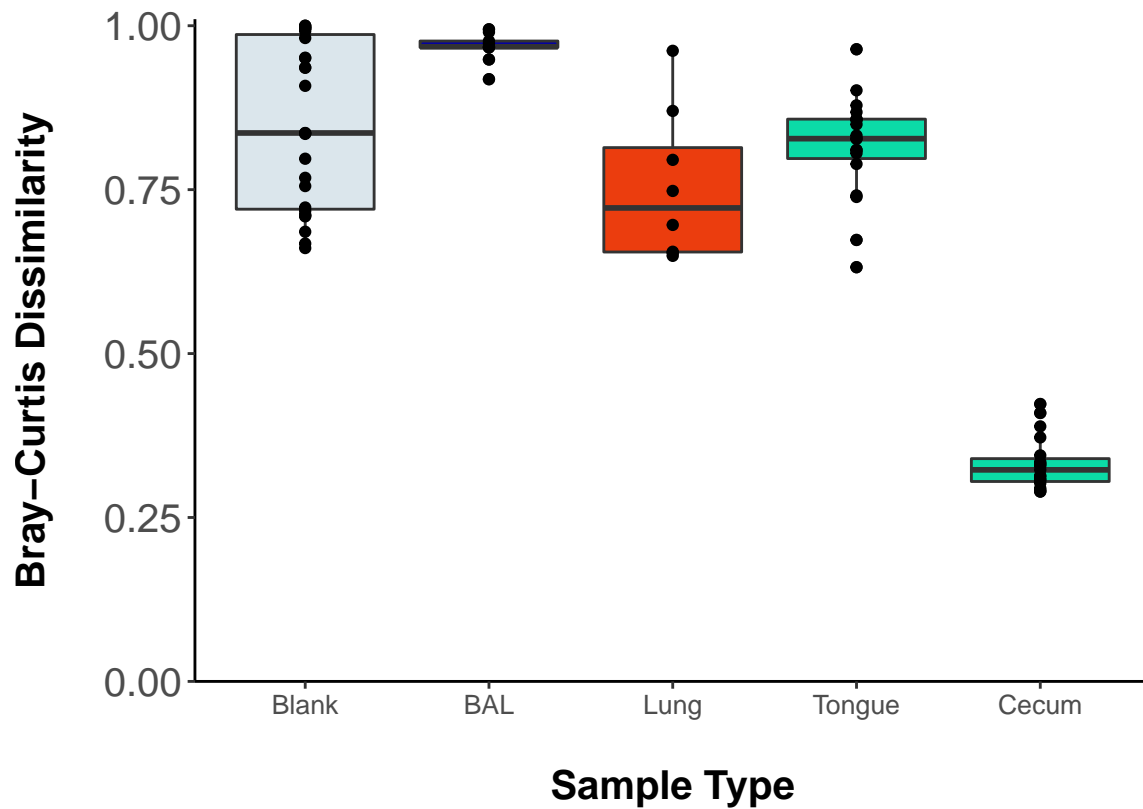
Tukey multiple comparisons of means
95% family-wise confidence level

```
Fit: aov(formula = tukey_otu_df[, "Shannon"] ~ tukey_otu_df[, "Alpha_Div_Groups"])
```

```
$`tukey_otu_df[, "Alpha_Div_Groups"]`
```

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

4.3 Beta Diversity - Bray-Curtis dissimilarity Index within Sample Type



[1] 87 2

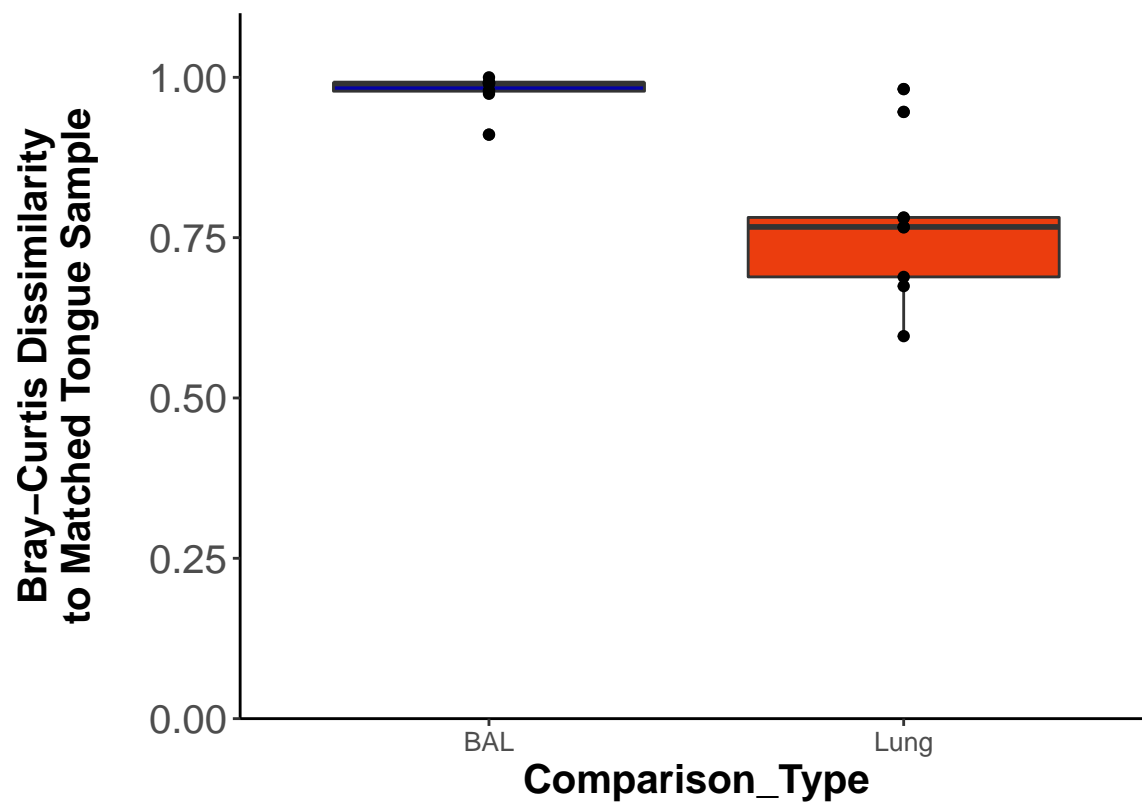
Pairwise comparisons using Wilcoxon rank sum test with continuity correction

data: bray_dist_lbtce_long_mut_filt_uniq\$BC_Index and bray_dist_lbtce_long_mut_filt_uniq\$Sample_Type

	BAL	Blank	Cecum	Lung
Blank	0.2730	-	-	-
Cecum	< 2e-16	< 2e-16	-	-
Lung	6.3e-09	7.4e-06	< 2e-16	-
Tongue	< 2e-16	8.5e-13	< 2e-16	0.0053

P value adjustment method: BH

4.4 Bray-Curtis dissimilarity - Lung Samples vs. Matched Tongue Samples



Wilcoxon rank sum exact test

data: BC_Index by Comparison_Type

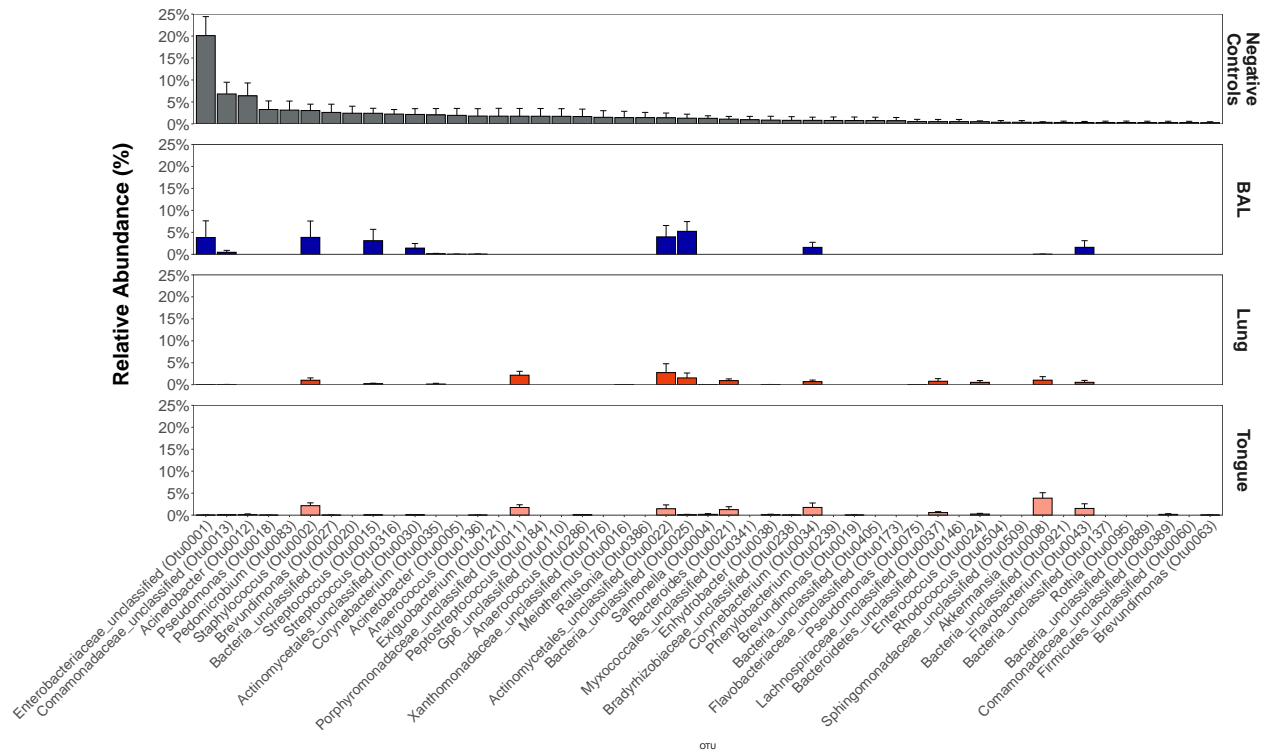
W = 85, p-value = 0.0004114

alternative hypothesis: true location shift is not equal to 0

5.1 Ordered by Lung OTUs



5.2 Ordered by Pooled Negative Control OTUs



5.3 Ordered by Tongue OTUs

