

Whole Lung Tissue vs. BAL Fluid 16S Sequencing Analysis (Untrimmed 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)

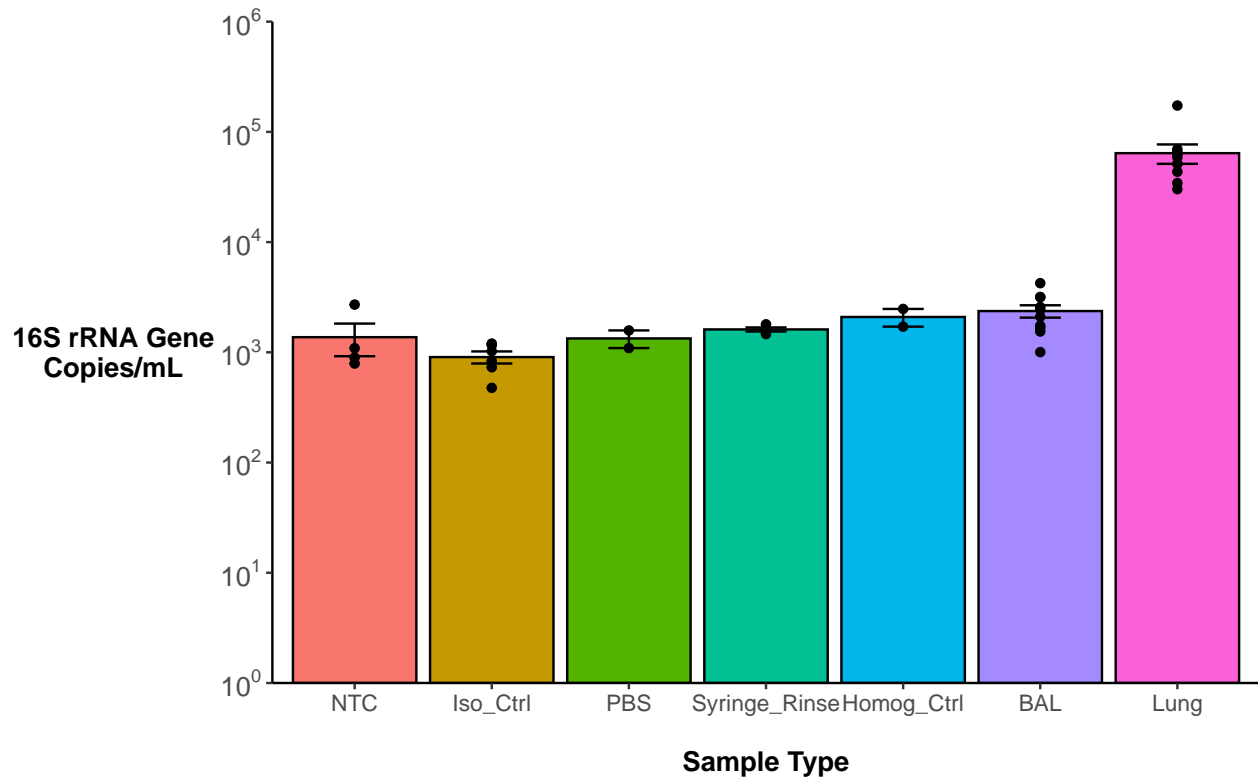
locale: en_US.UTF-8|en_US.UTF-8|en_US.UTF-8|C|en_US.UTF-8|en_US.UTF-8

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), *xml2*(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

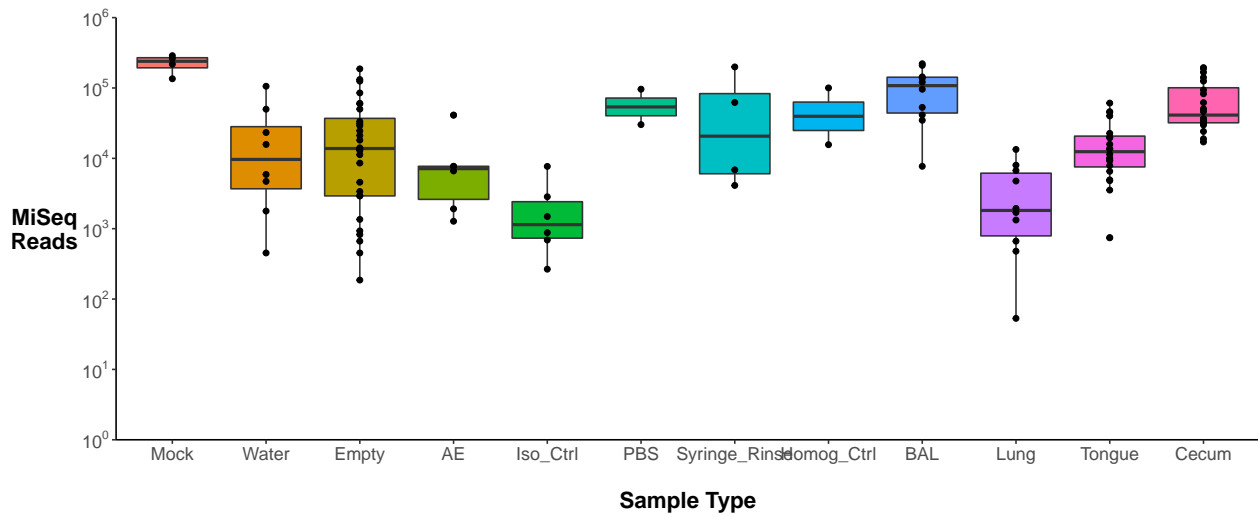
data: ddpcr\$Gene_16S_copies_per_mL and ddpcr\$Sample_Type

	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

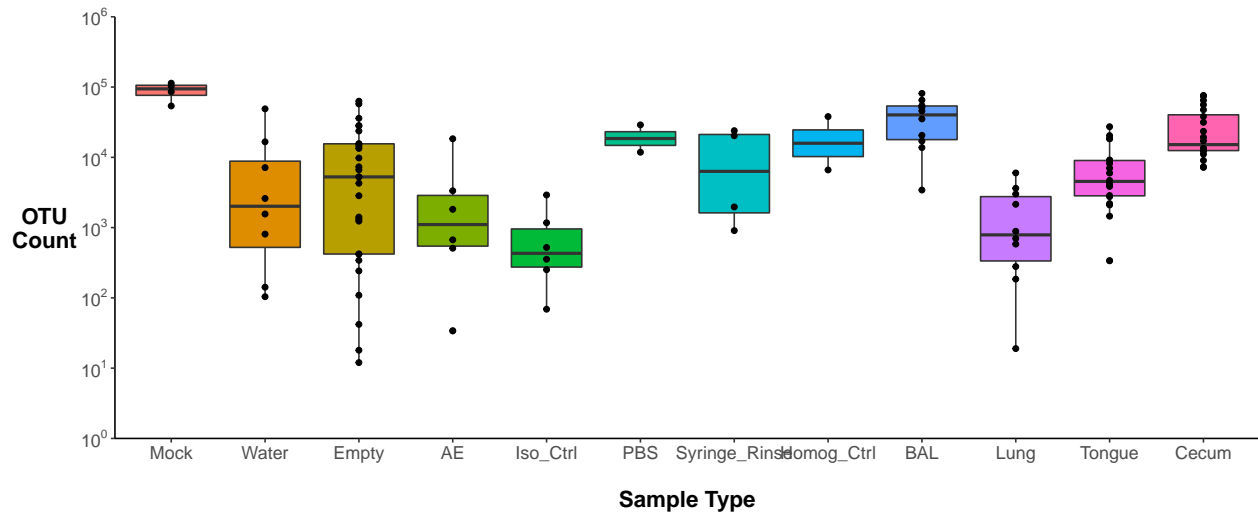
P value adjustment method: BH

II. Quality Checks - Confirmation of Sufficient Reads

2.1 MiSeq Reads by Sample Type

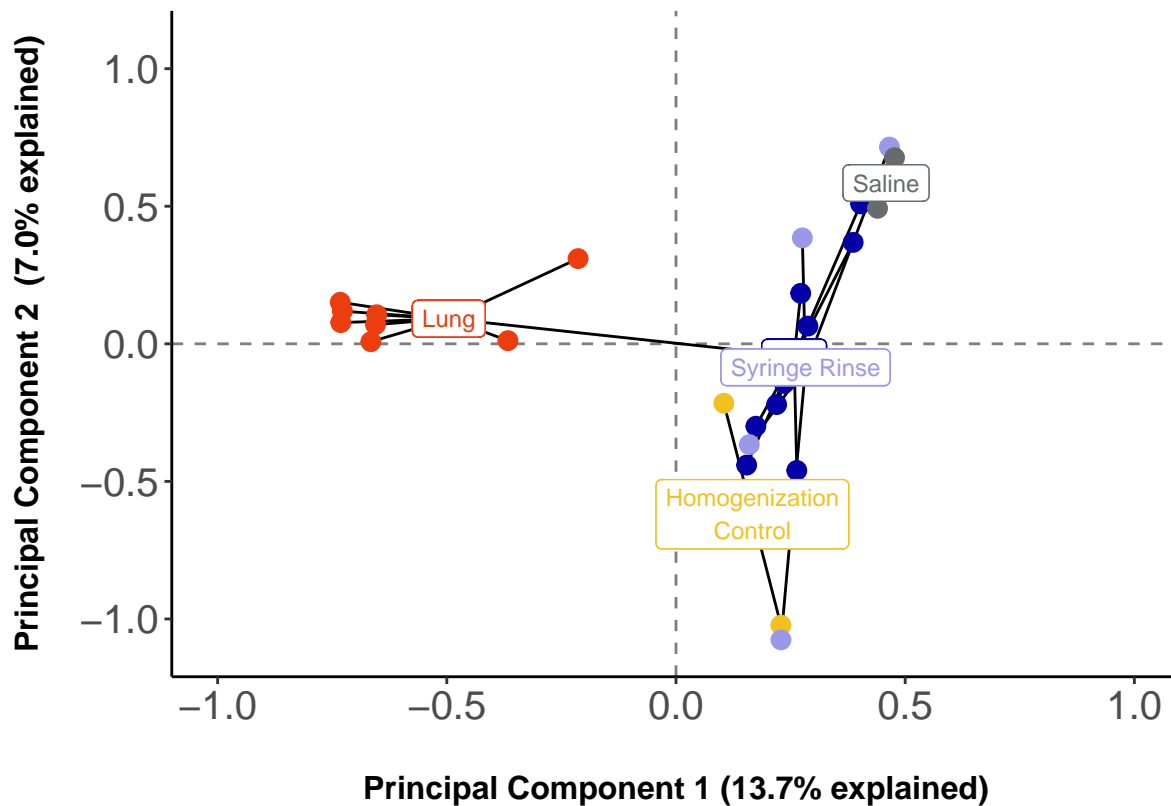


2.2 OTU Counts by Sample Type



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	3.9774	1.98870	2.4676	0.16486
Residuals	25	20.1478	0.80591		0.83514
Total	27	24.1252			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	2.469	2.46898	3.1997	0.15093	9.999e-05 ***
Residuals	18	13.889	0.77162		0.84907	
Total	19	16.358			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	2.5283	2.52828	3.423	0.17623	2e-04 ***
Residuals	16	11.8179	0.73862		0.82377	
Total	17	14.3462			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")

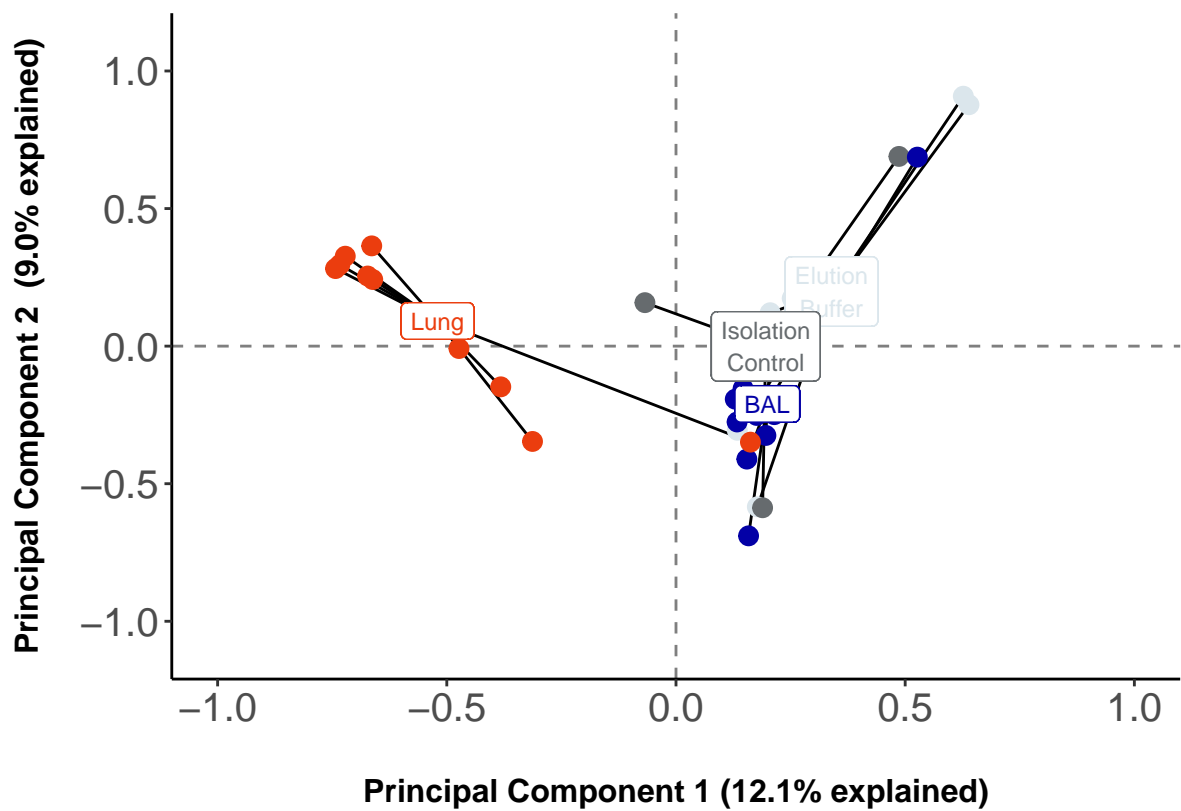
Permutation: free

Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
adonis.otudf.bn\$Organ	1	0.9155	0.91549	1.0041	0.05905	0.4382
Residuals	16	14.5885	0.91178		0.94095	
Total	17	15.5040			1.00000	

3.2 Lung Samples v. Isolation Controls



Call:
adonis(formula = otu.good.lungsamp.isoctrls_hel ~ otu.df.lungsamp.isoctrls\$Sample_Type, permutation

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	5.1677	1.7226	2.1035	0.18392
Residuals	28	22.9291	0.8189		0.81608
Total	31	28.0968			1.00000

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 = "euclid

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	2.7563	2.75633	3.5591	0.15107	9.999e-05 ***
Residuals	20	15.4890	0.77445		0.84893	
Total	21	18.2453			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

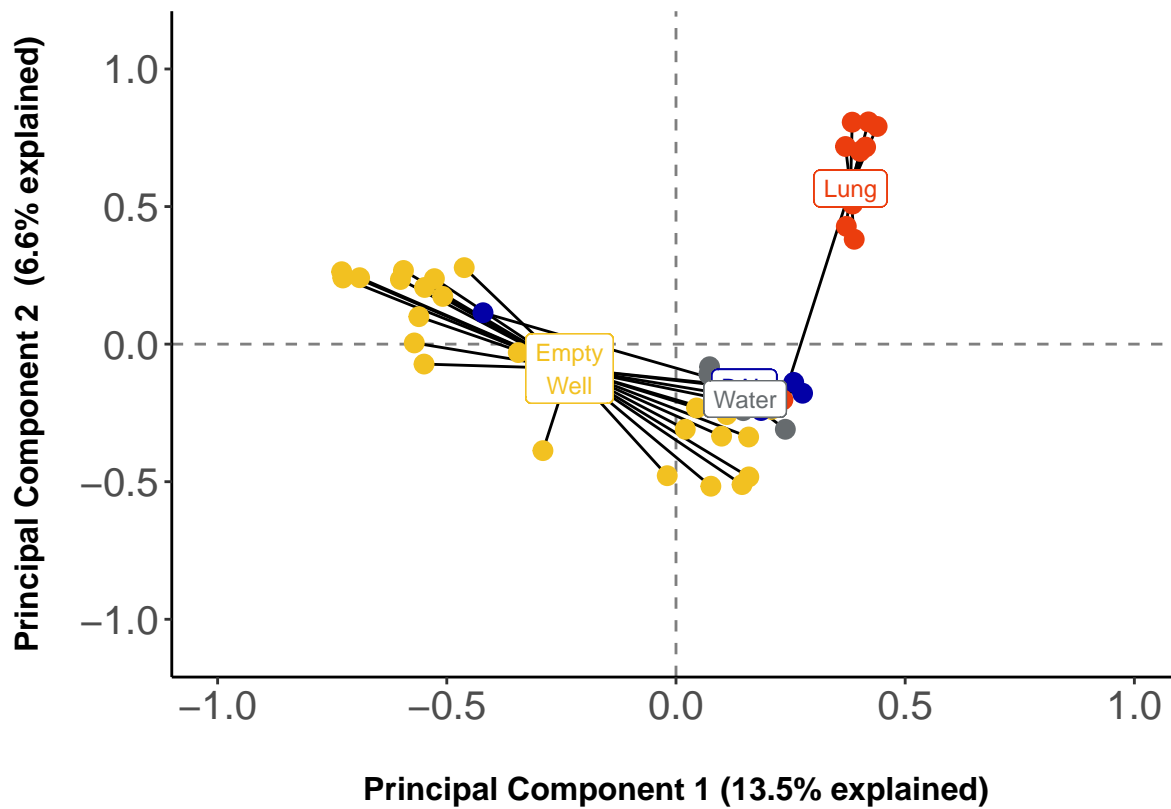
Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
adonis.otudf.ib\$RA_Groups	1	1.2216	1.22157	1.338	0.06271	0.05969 .
Residuals	20	18.2596	0.91298		0.93729	
Total	21	19.4812			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
```

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	8.155	2.71844	3.4176	0.1647
Residuals	52	41.362	0.79543		0.8353
Total	55	49.517			1.0000

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 = "euclid
```

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	4.402	4.4017	5.5369	0.11177	9.999e-05 ***
Residuals	44	34.979	0.7950		0.88823	
Total	45	39.381			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

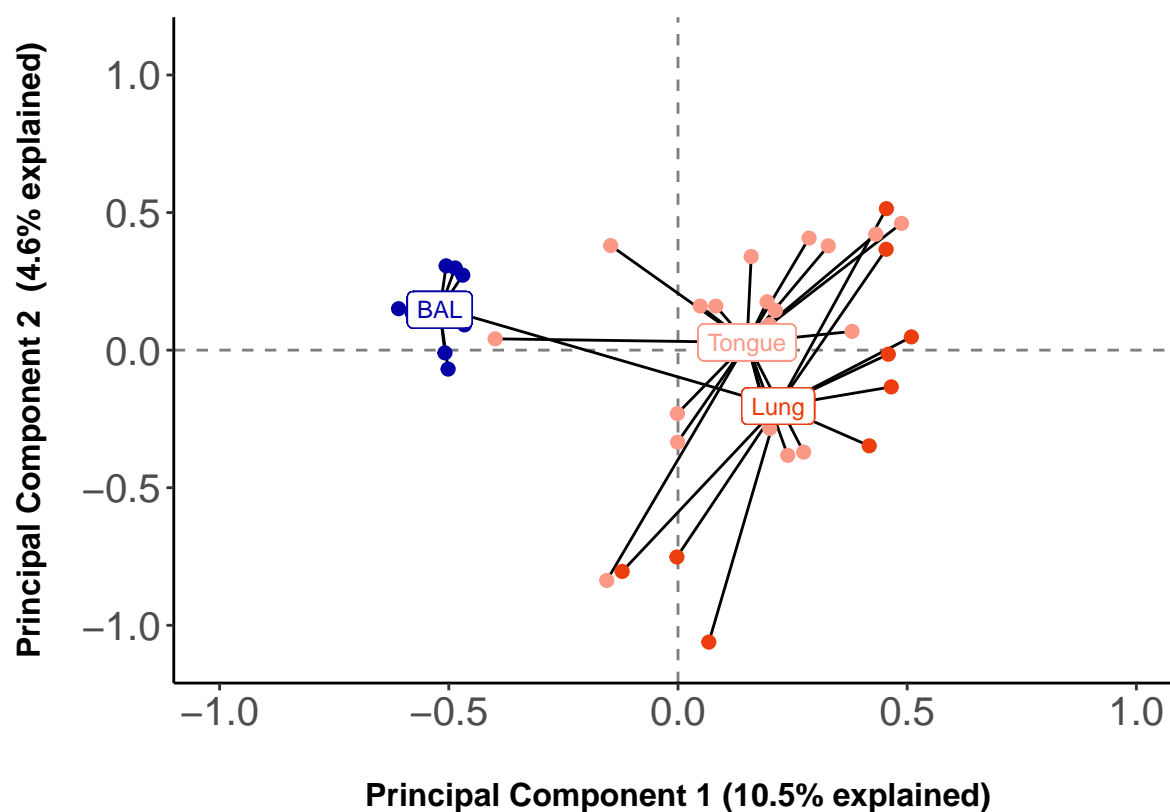
Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
adonis.otudf.sb\$RA_Groups	1	2.083	2.08261	2.4274	0.05228	4e-04 ***
Residuals	44	37.750	0.85795		0.94772	
Total	45	39.833			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	3.9905	1.9953	2.6829	0.12665	9.999e-05
Residuals	37	27.5168	0.7437		0.87335	
Total	39	31.5073			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	1.0559	1.05587	1.5409	0.05216	0.009199 **
Residuals	28	19.1868	0.68524		0.94784	
Total	29	20.2427			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

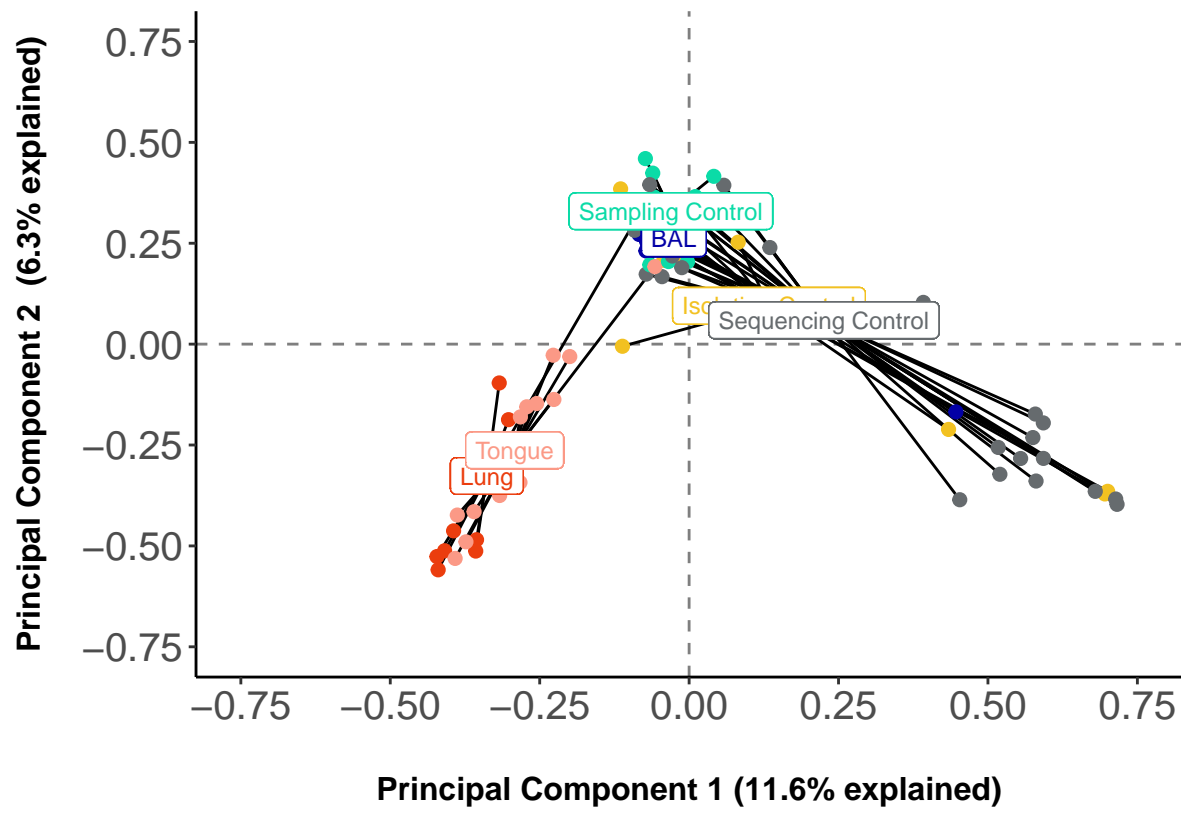
Number of permutations: 10000

Terms added sequentially (first to last)

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
adonis.otudf.tb\$Sample_Type	1	2.6188	2.61884	3.3395	0.10656	9.999e-05 ***
Residuals	28	21.9574	0.78419		0.89344	
Total	29	24.5763			1.00000	

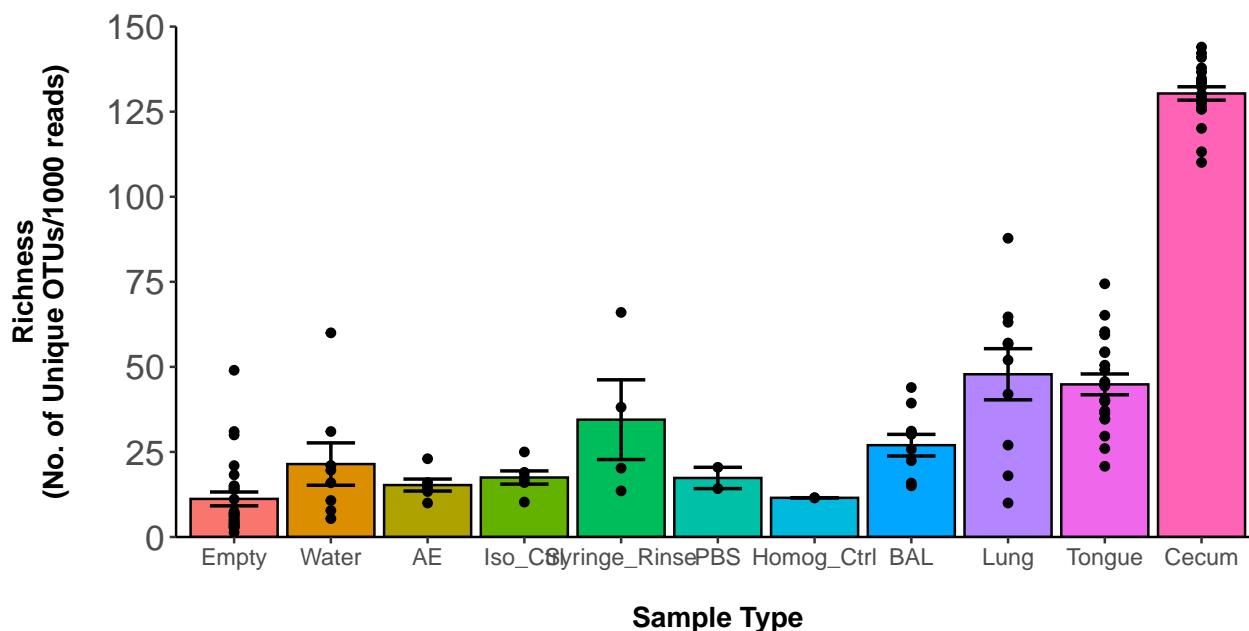
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

3.5 Lung vs. Tongue vs. Negative Controls



IV. Diversity Indices

4.1 Alpha Diversity - Rarified Richness



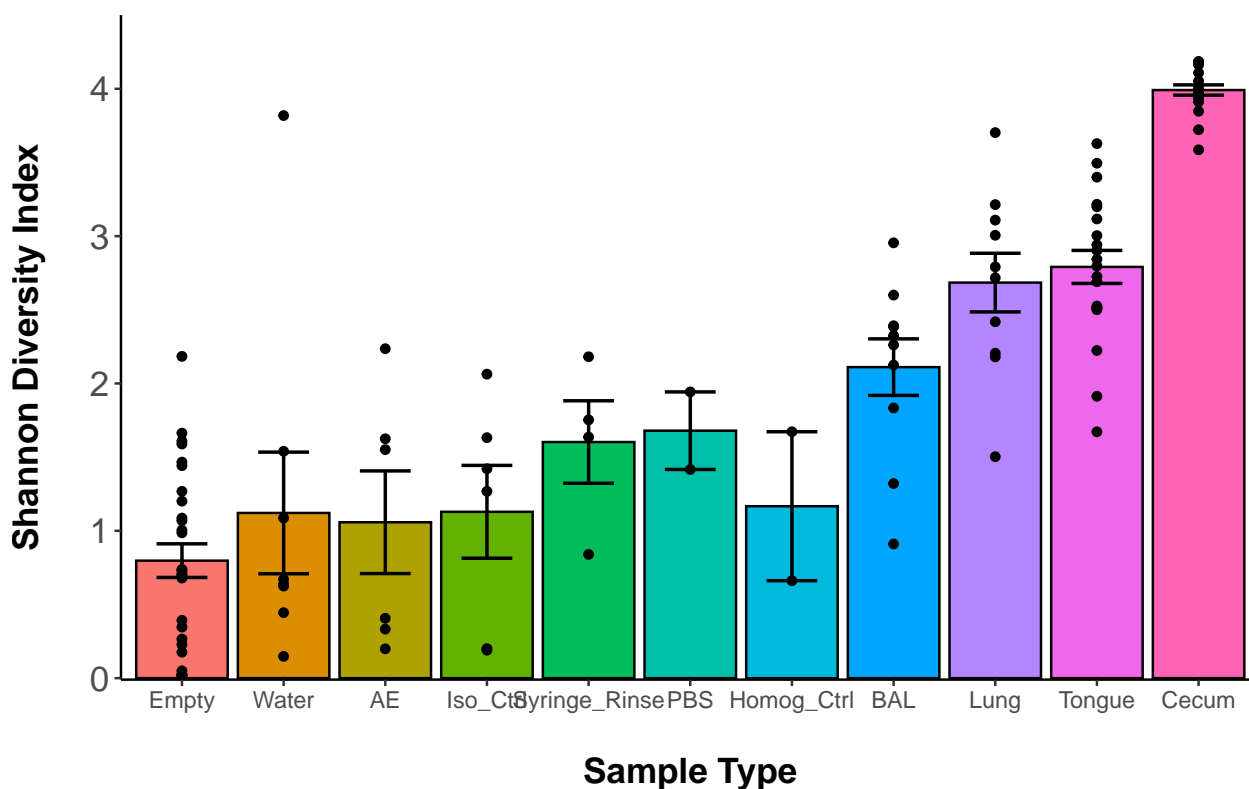
Tukey multiple comparisons of means
95% family-wise confidence level

```
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	-27.844847	6.6202640
Lung-BAL	20.846543	2.847725	38.8453597
Sampling Control-BAL	-2.528455	-21.619083	16.5621736
Sequencing Control-BAL	-13.515401	-27.901951	0.8711478
Lung-Isolation Control	31.458834	14.226279	48.6913895
Sampling Control-Isolation Control	8.083837	-10.286129	26.4538024
Sequencing Control-Isolation Control	-2.903110	-16.318636	10.5124164
Sampling Control-Lung	-23.374997	-42.465626	-4.2843690
Sequencing Control-Lung	-34.361944	-48.748493	-19.9753948
Sequencing Control-Sampling Control	-10.986946	-26.718045	4.7441525
	p adj		
Isolation Control-BAL	0.4262449		
Lung-BAL	0.0150895		
Sampling Control-BAL	0.9958863		
Sequencing Control-BAL	0.0757056		
Lung-Isolation Control	0.0000253		
Sampling Control-Isolation Control	0.7329503		
Sequencing Control-Isolation Control	0.9737499		
Sampling Control-Lung	0.0087438		
Sequencing Control-Lung	0.0000000		
Sequencing Control-Sampling Control	0.2986417		

4.2 Alpha Diversity - Shannon Diversity Index



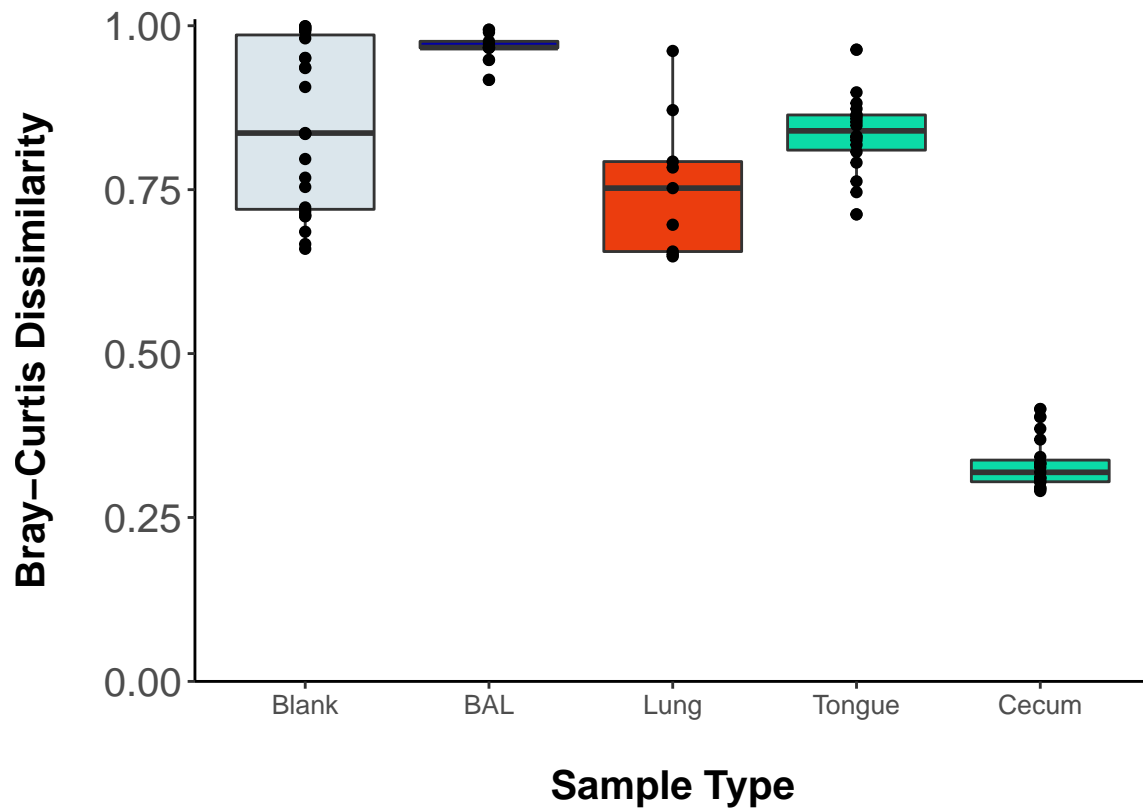
Tukey multiple comparisons of means
95% family-wise confidence level

```
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	-1.0172362	-1.8642930	-0.1701794	0.0106434
Lung-BAL	0.5738854	-0.3108366	1.4586074	0.3728207
Sampling Control-BAL	-0.5979371	-1.5363265	0.3404522	0.3911169
Sequencing Control-BAL	-1.2413229	-1.9484858	-0.5341600	0.0000534
Lung-Isolation Control	1.5911216	0.7440648	2.4381785	0.0000143
Sampling Control-Isolation Control	0.4192991	-0.4836665	1.3222647	0.6921918
Sequencing Control-Isolation Control	-0.2240867	-0.8835195	0.4353462	0.8756718
Sampling Control-Lung	-1.1718226	-2.1102119	-0.2334332	0.0071118
Sequencing Control-Lung	-1.8152083	-2.5223712	-1.1080454	0.0000000
Sequencing Control-Sampling Control	-0.6433858	-1.4166393	0.1298678	0.1478642

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



[1] 88 2

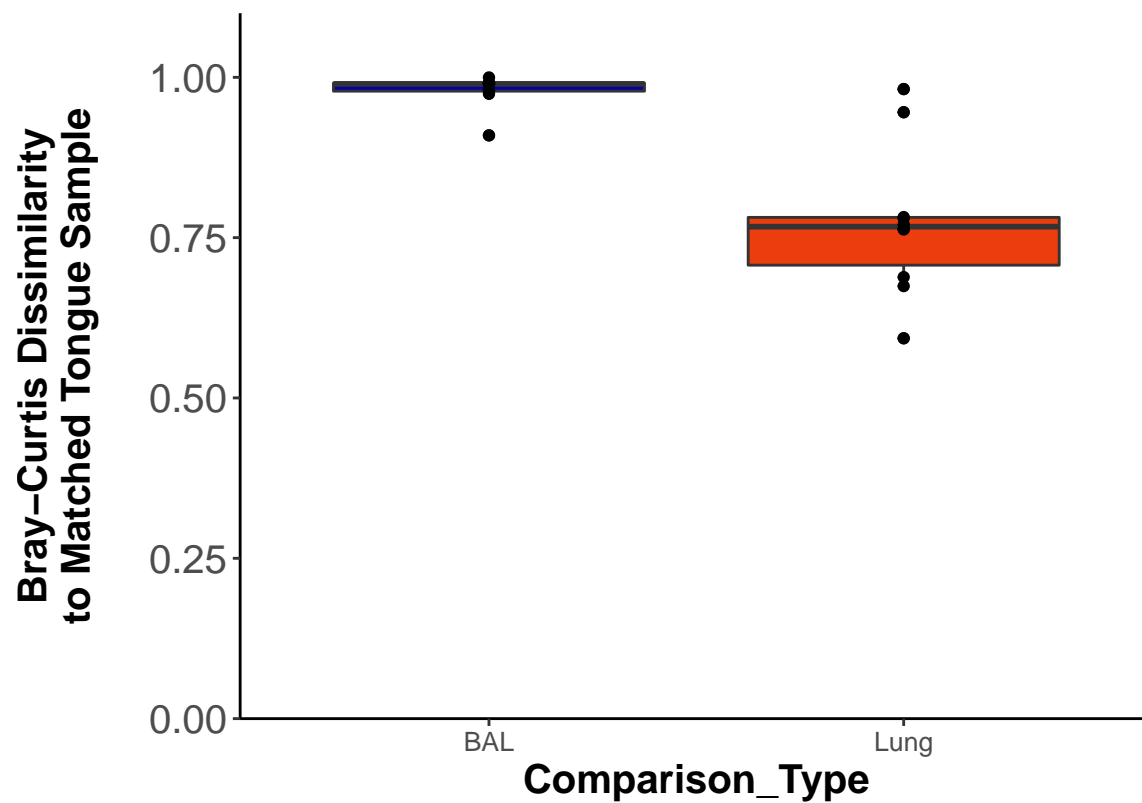
Pairwise comparisons using Wilcoxon rank sum test with continuity correction

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	0.29581	-	-	-
Cecum	< 2e-16	< 2e-16	-	-
Lung	3.8e-12	7.1e-07	< 2e-16	-
Tongue	< 2e-16	2.4e-11	< 2e-16	0.00028

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

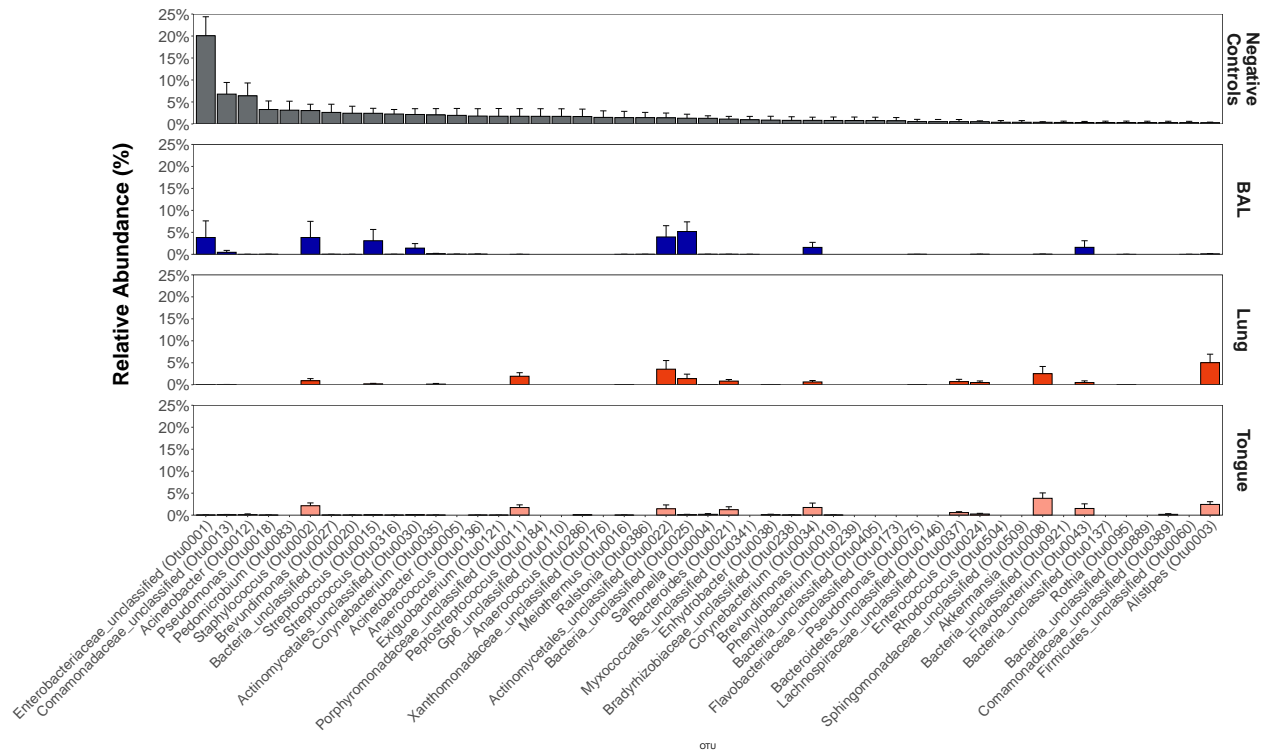
W = 95, p-value = 0.0002057

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

