

Module 3: Project 1 by Team 5

mothur vs. QIIME2 Microbiome Data Analysis

Karl Abuan

May Ho

Jonah Lin

Leilynaz Malekafzali

Tiffany Yang

Abdur Rahman M. A. Basher

March 20, 2018

Abstract

This is the abstract. It consists of two paragraphs.

Contents

1	Introduction	2
2	Problem Formulation	2
3	Materials and Experimental Configuration	2
3.1	Experimental Protocols	2
3.2	Dataset	2
3.3	Parameters Configuration	2
3.4	Data Preporocessing	2
4	Results	2
4.1	Analysis of microbial community structure along with depth and oxygen concentration	2
4.2	Analysis of abundance information of Planctomyces along with depth and/or oxygen concentration	8
4.3	Estimate richness (number of OTUs/ASVs) for Planctomyces	9
4.4	Interpretation of abundance information of OTUs/ASVs of Planctomyces along with depth and/or oxygen concentration	10
5	Discussion	14
	References	14

1 Introduction

2 Problem Formulation

3 Materials and Experimental Configuration

3.1 Experimental Protocols

P1. Analysis of microbial community structure along with depth and oxygen concentration.

P2. Analysis of abundance information of Planctomyces along with depth and/or oxygen concentration.

P3. Estimate richness (number of OTUs/ASVs) for Planctomyces.

P4. Interpretation of abundance information of OTUs/ASVs of Planctomyces along with depth and/or oxygen concentration.

3.2 Dataset

3.3 Parameters Configuration

3.4 Data Preprocessing

We use saanich inlet datasets that are propocessed using Mothur and QIIME2

Samples are then rarefied/normalized to 100,000 sequences per sample to facilitate comparisons between samples. A random seed was set to ensure reproducibility.

Rarefied counts were converted to relative abundance percentages.

Next, we perform a series of filterings according to three rules: i)- exclude OTUs that are not observed for more than 4 samples; ii)- prune samples and OTUs with unknown values, such as `unclassified` value; and iii)- any phylum fail to have more than 5 OTUs should be trimmed. The codes used for applying the three rules are:

4 Results

4.1 Analysis of microbial community structure along with depth and oxygen concentration

We first estimate the overall taxa diversity using Shannon’s diversity index as shown in [Figure 1](#).

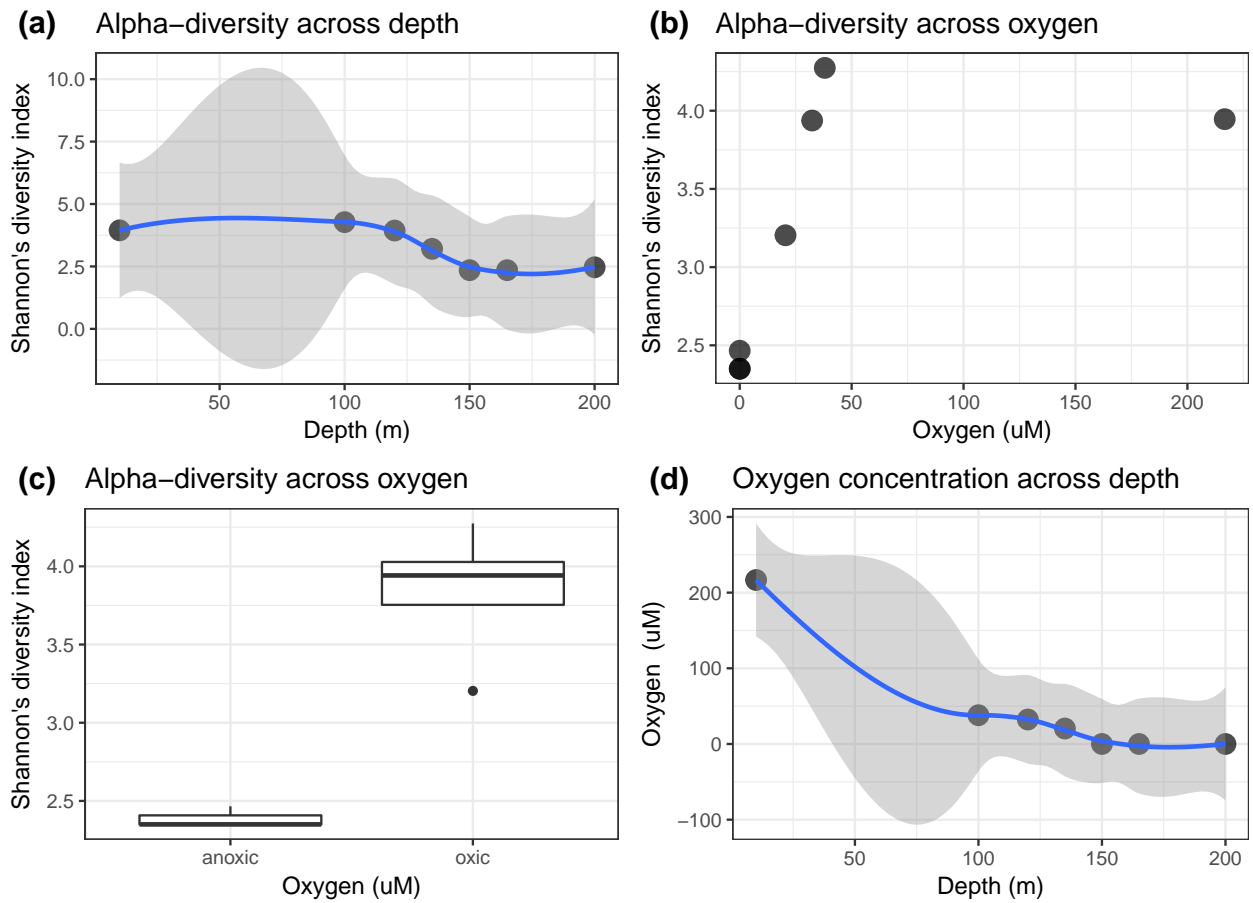


Figure 1: Mothur

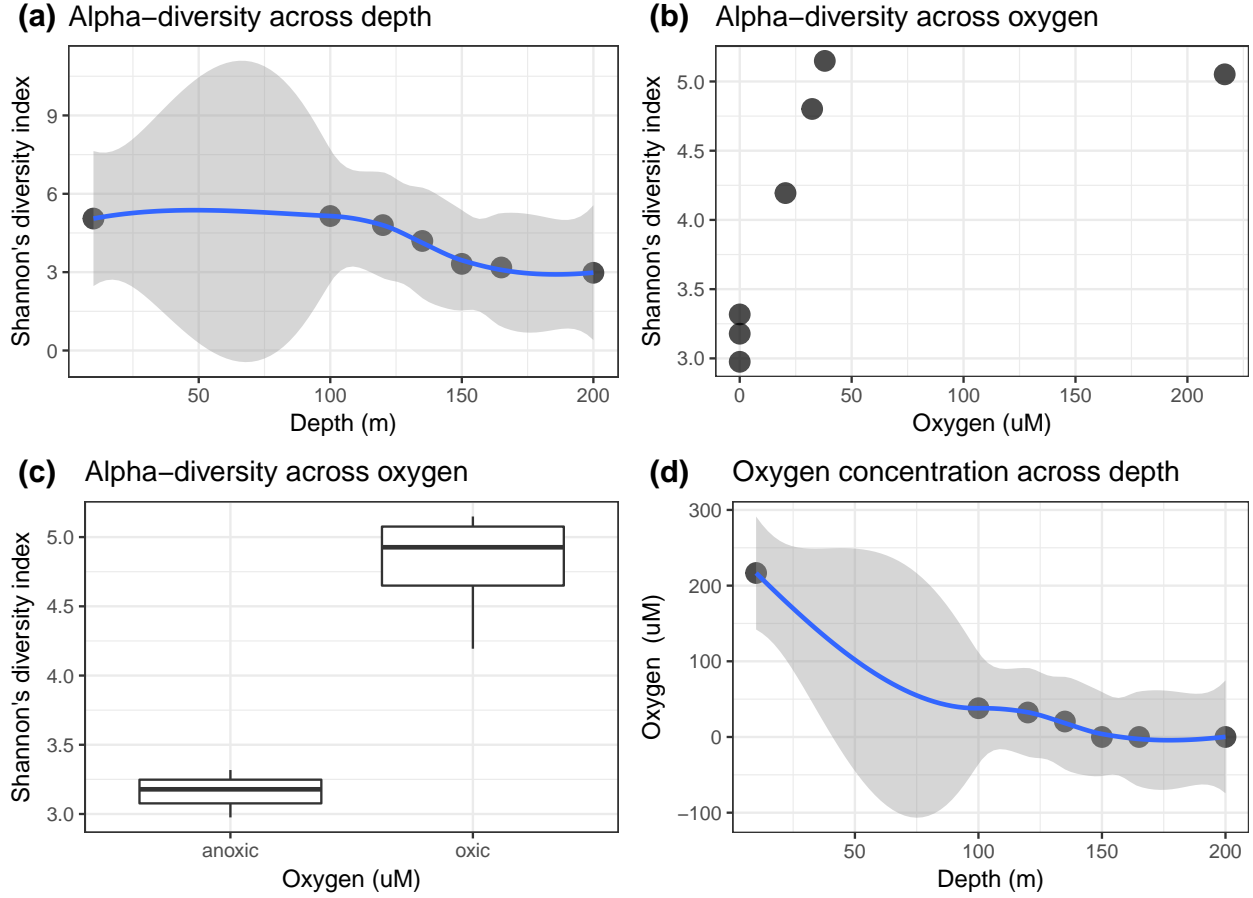


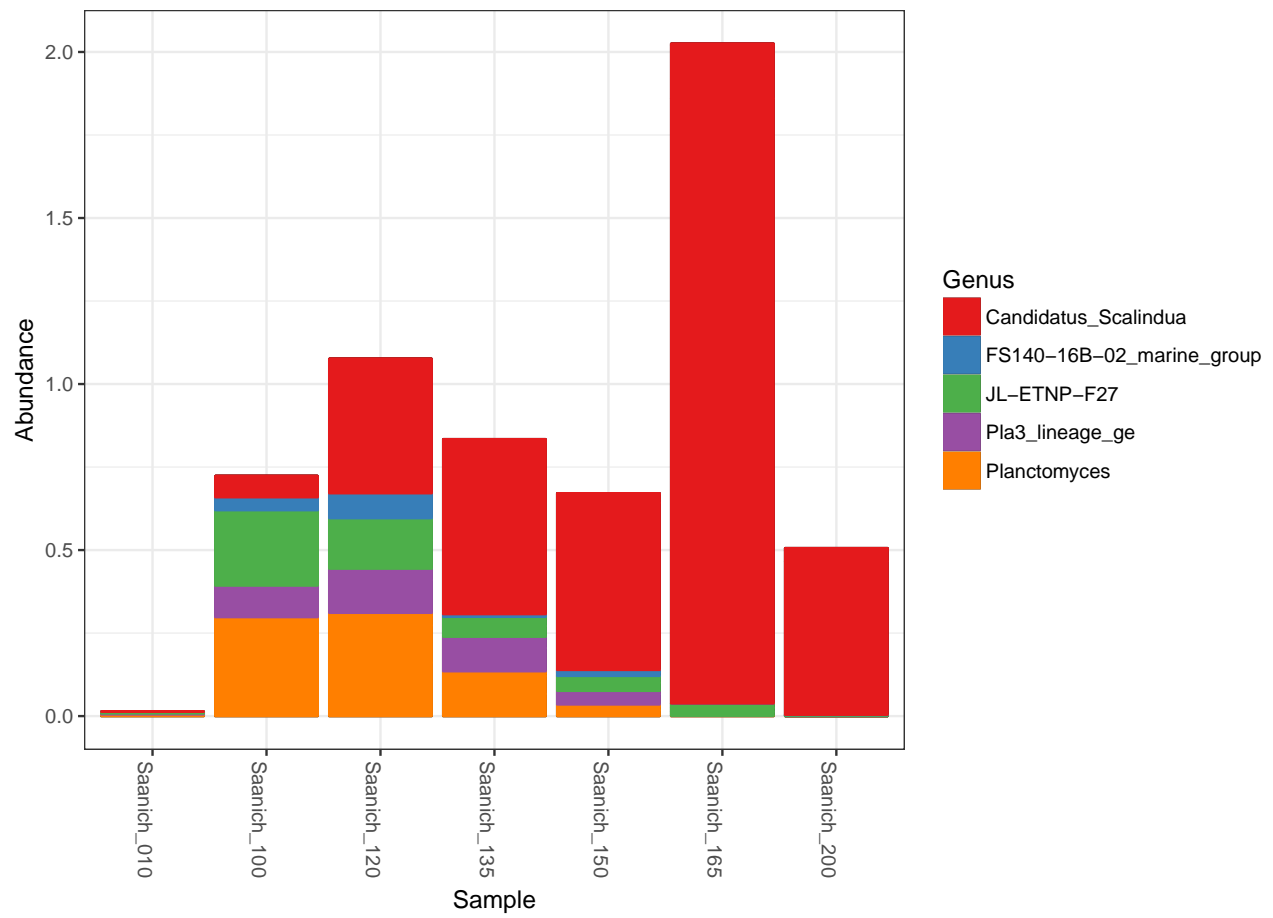
Figure 2: QIIME2

To understand the number of taxa associated with the Phylum rank, we found that there are around Proteobacteria, Bacteroidetes, Thaumarchaeota, Actinobacteria, Marinimicrobia_(SAR406_clade), Planctomycetes, Verrucomicrobia taxa in mothur in compare to Proteobacteria, Bacteroidetes, Thaumarchaeota, Actinobacteria, Marinimicrobia_(SAR406_clade), Planctomycetes, Verrucomicrobia taxa in QIIME2.

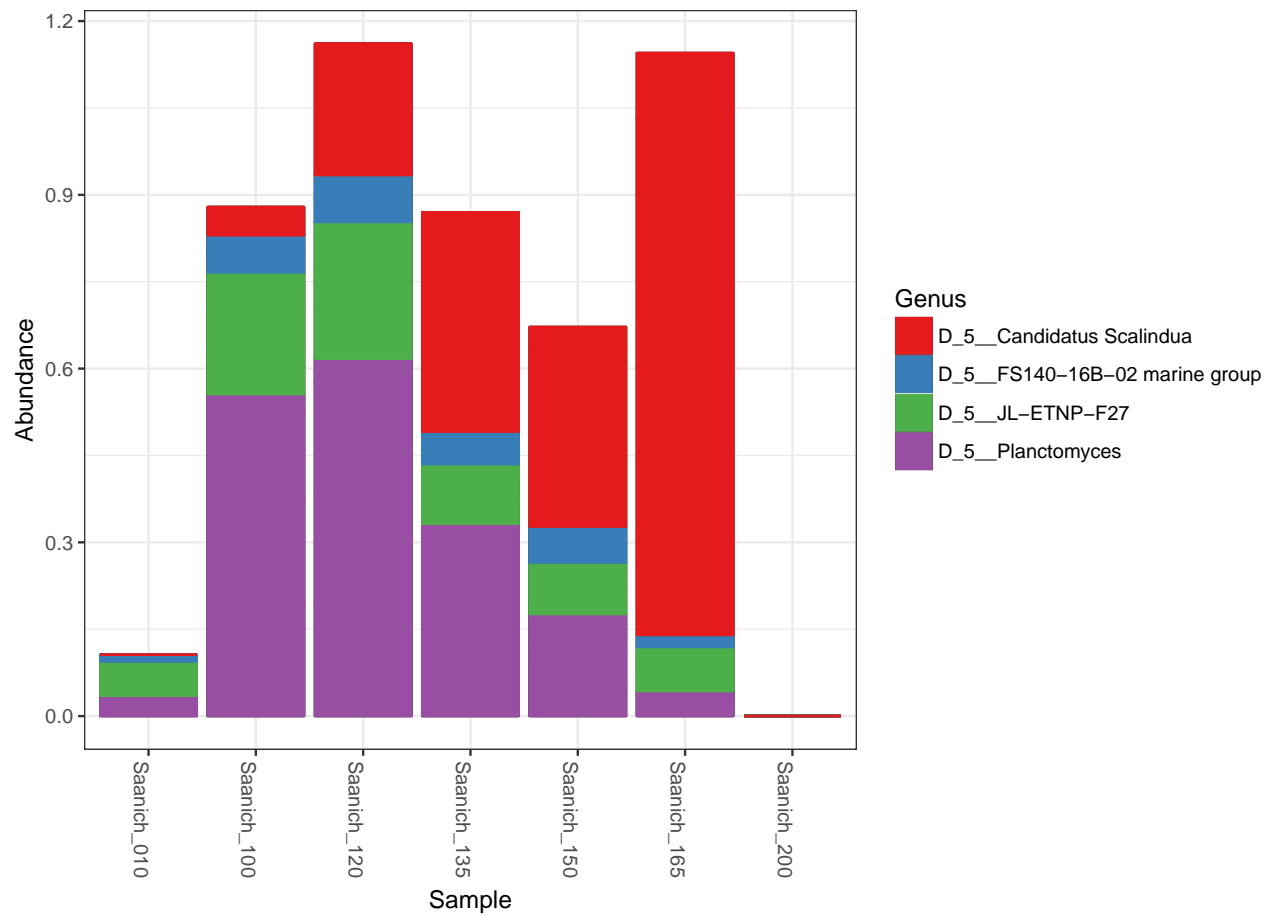
Table 1: Phylums from Mothur vs Phylums from QIIME2

Phylums from Mothur	Phylums from QIIME2
Proteobacteria	D_1__Proteobacteria
Bacteroidetes	D_1__Bacteroidetes
Thaumarchaeota	D_1__Planctomycetes
Actinobacteria	D_1__Thaumarchaeota
Marinimicrobia_(SAR406_clade)	NA
Planctomycetes	NA
Verrucomicrobia	NA

We choose the *Planctomycetes* phylum, and explored the distribution of genera of this phylum.

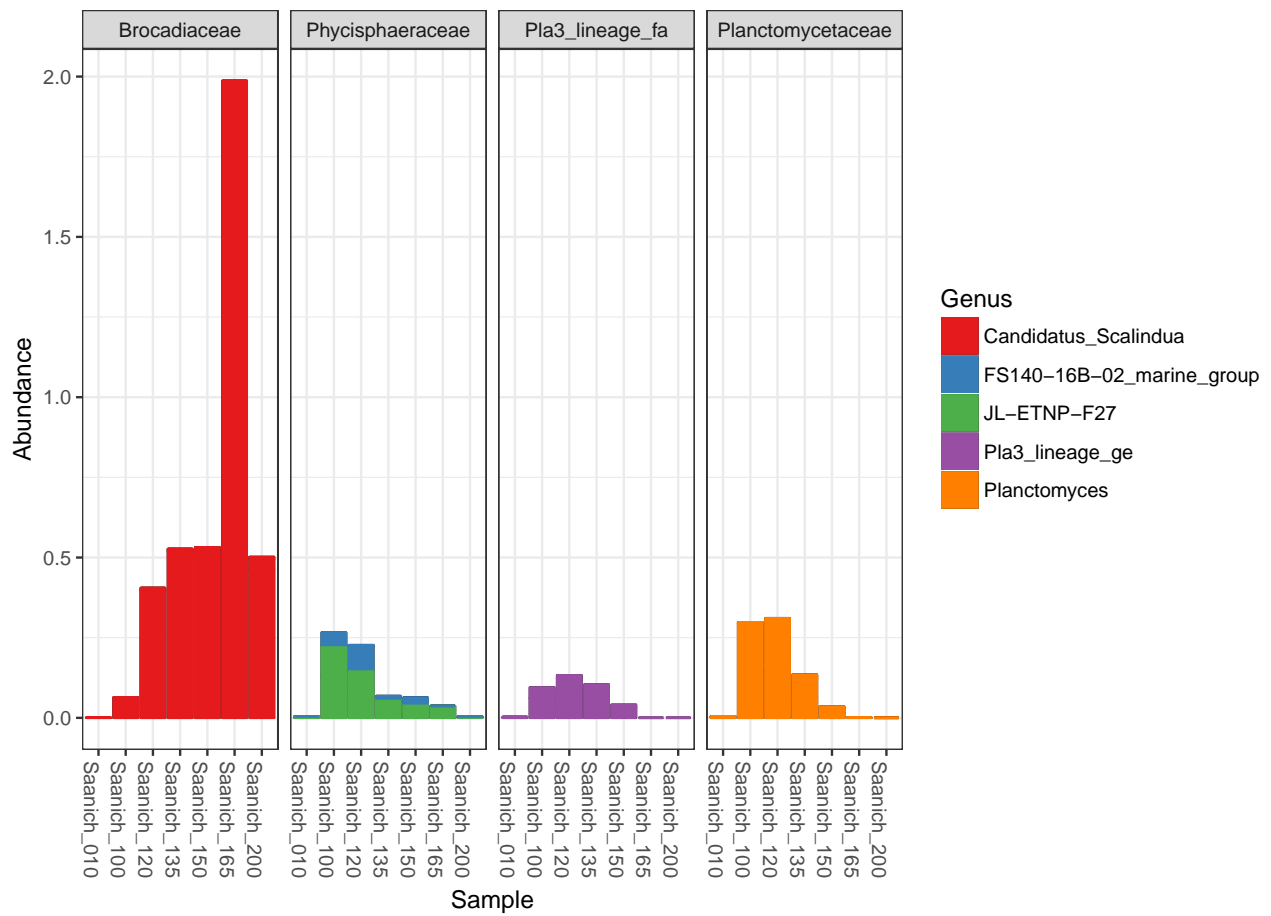


Genus distribution of Planctomycetes across samples from mothur

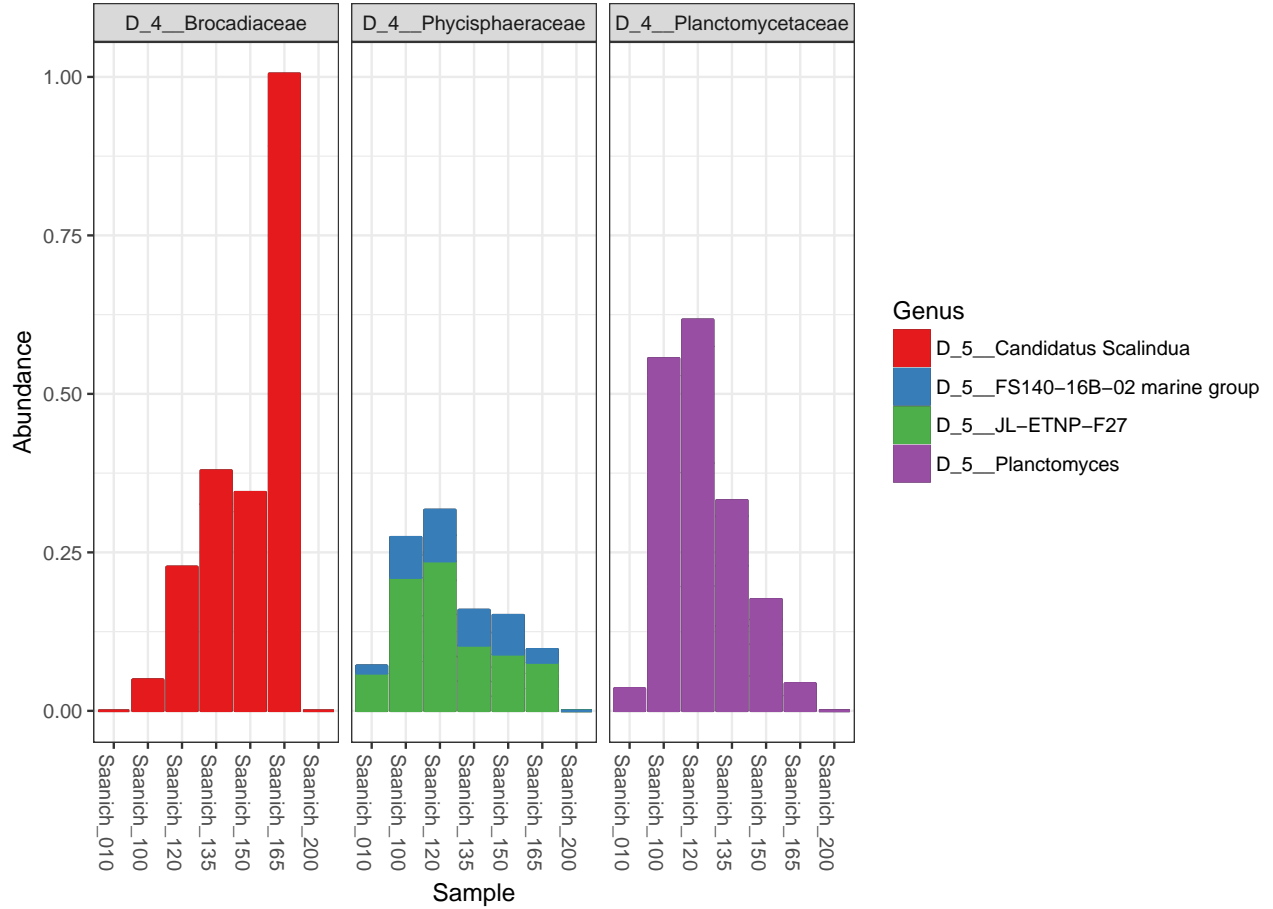


Genus distribution of Planctomycetes across samples from QIIME2

We further investigated the genus distribution of this phylum across samples grouped by family level.



Genus distribution of Planctomyces across samples grouped by family level from mothur



Genus distribution of Planctomyces across samples grouped by family level from QIIME2

Finally, we settled on performing experimental analysis at *Planctomyces* genus level and its associated OTUs.

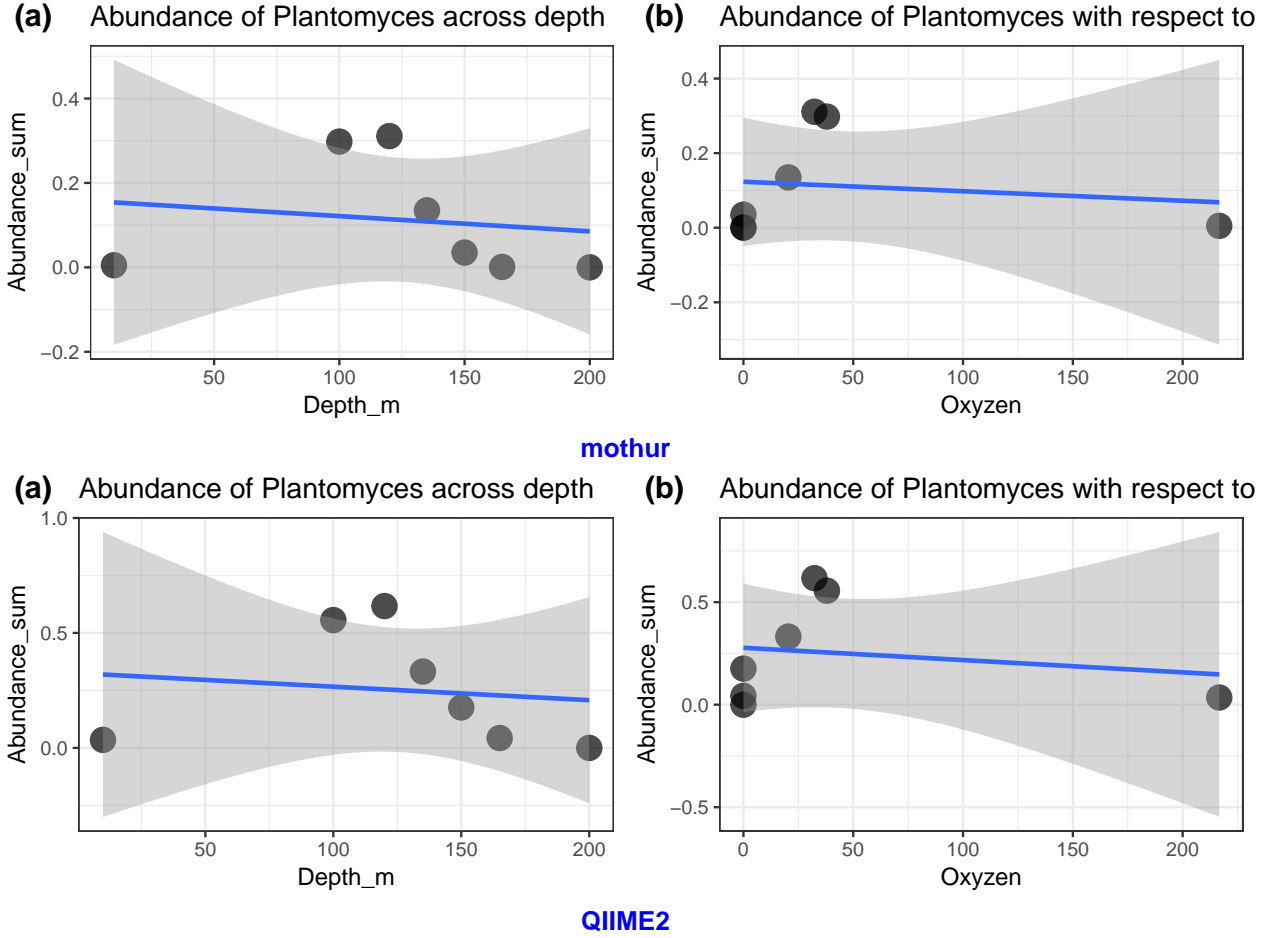
Table 2: OTUs from Mothur vs ASVs from QIIME2

OTUs from Mothur	ASVs from QIIME2
Otu0125	Asv232
Otu0144	Asv799
Otu0401	Asv1021
Otu0592	Asv1124

4.2 Analysis of abundance information of Planctomyces along with depth and/or oxygen concentration

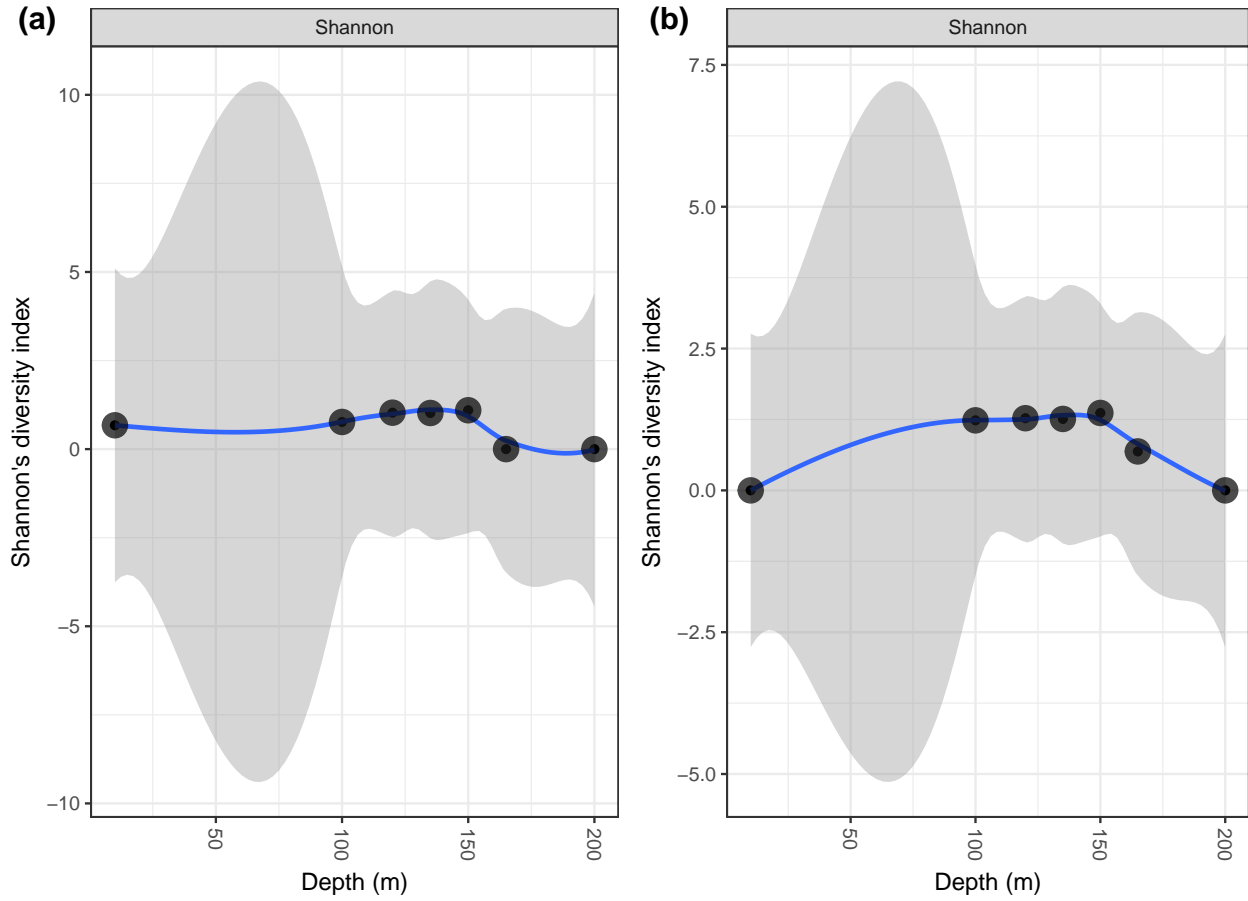
Table 3: Correlation data of OTUs within *Planctomyces* genus across depth and oxygen concentration from mothur and QIIME2

Covariates	Estimate	Std. Error	t-value	Pr(> t)
Depth (mothur)	-0.0003609	0.0010227	-0.3528908	0.7385598
O2_uM (mothur)	-0.0002544	0.0007941	-0.3203956	0.7616253
Depth (QIIME2)	-0.0005878	0.0018774	-0.3130933	0.7668485
O2_uM (QIIME2)	-0.0005997	0.0014441	-0.4152436	0.6951812



4.3 Estimate richness (number of OTUs/ASVs) for *Planctomyces*

We explore the diversity of *Planctomyces* across depth.



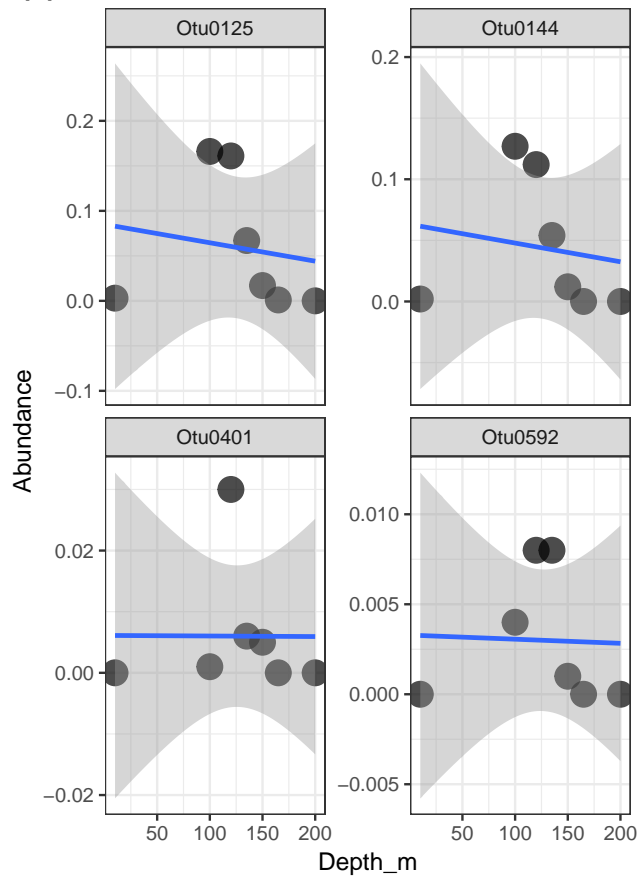
Alpha-diversity of Planctomyces across depth from (a) mothur and (b) QIIME2

4.4 Interpretation of abundance information of OTUs/ASVs of Planctomyces along with depth and/or oxygen concentration

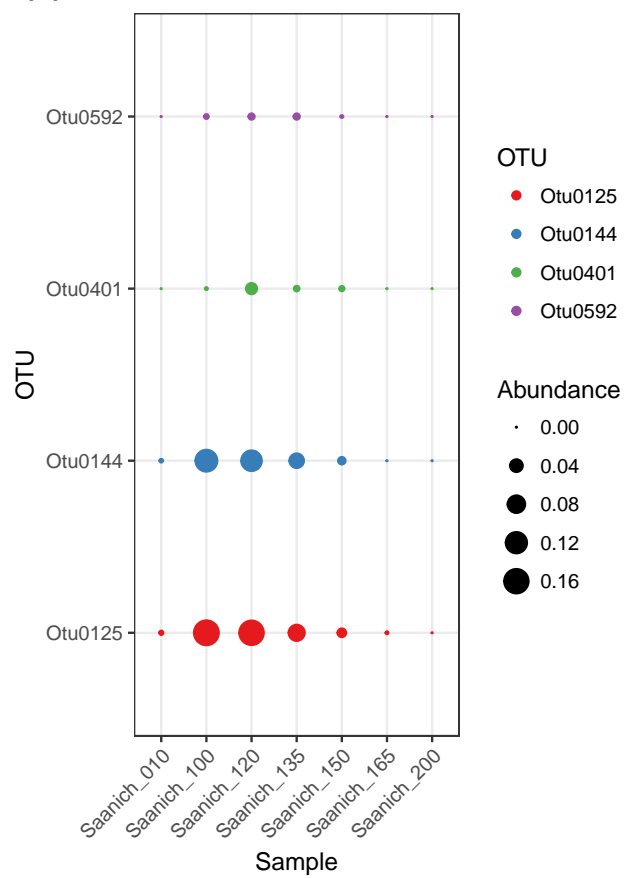
Table 4: Correlation data of OTUs within Planctomyces genus across depth

Covariates	Estimate	Std. Error	t-value	Pr(> t)
Otu0125 (mothur)	-0.0002045	0.0005479	-0.3731784	0.7243139
Otu0144 (mothur)	-0.0001533	0.0004035	-0.3798581	0.7196506
Otu0401 (mothur)	-0.0000009	0.0000807	-0.0113619	0.9913741
Otu0592 (mothur)	-0.0000023	0.0000274	-0.0836392	0.9365887
Asv232 (QIIME2)	-0.0001665	0.0006541	-0.2545201	0.8092302
Asv799 (QIIME2)	-0.0000953	0.0005682	-0.1676505	0.8734282
Asv1021 (QIIME2)	0.0000544	0.0001222	0.4454921	0.6745908
Asv1124 (QIIME2)	-0.0003805	0.0005859	-0.6493922	0.5447317

(a) Per OTU abundance information

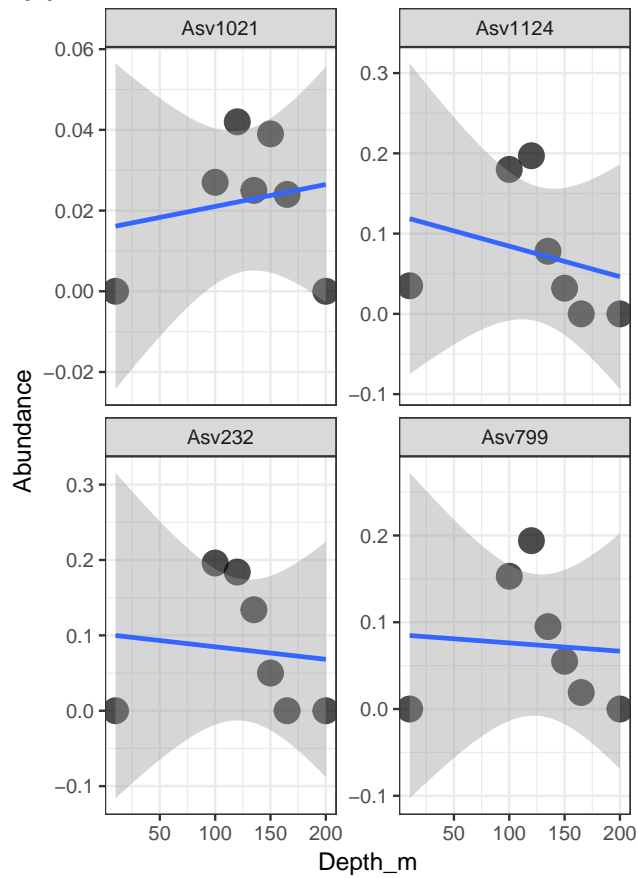


(b) Abundance of OTUs across samples

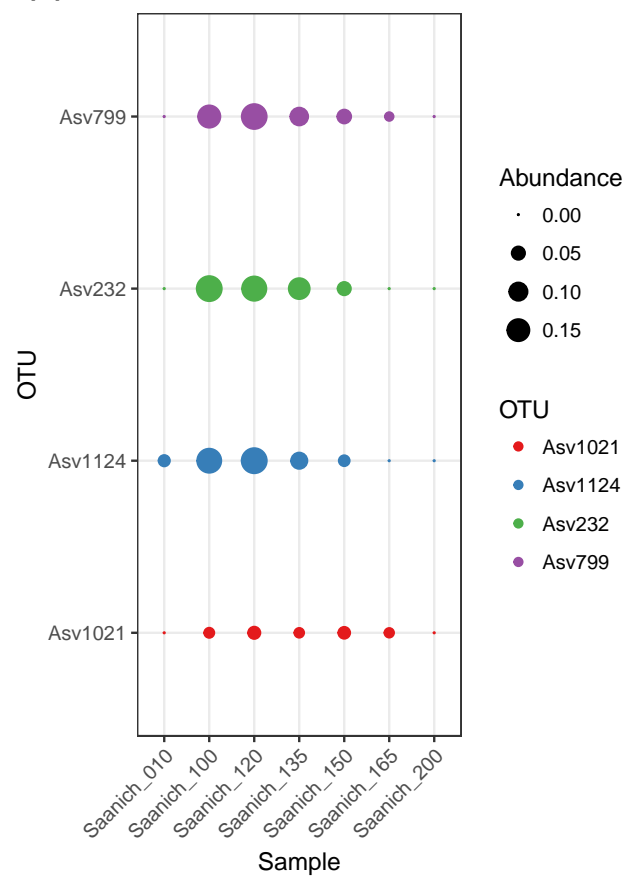


Abundance of OTUs within *Planctomyces* genus across depth from mothur

(a) Per ASV abundance information

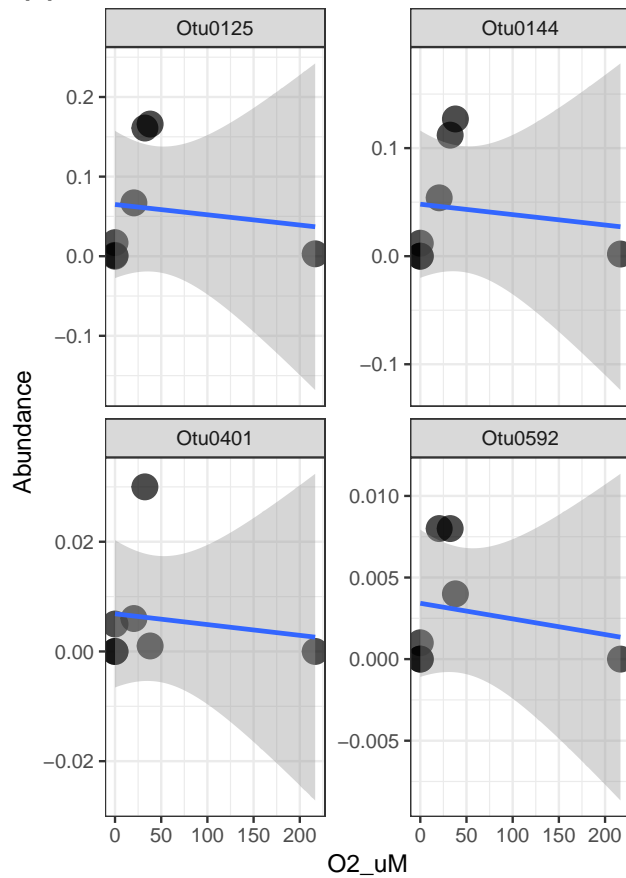


(b) Abundance of ASVs across samples

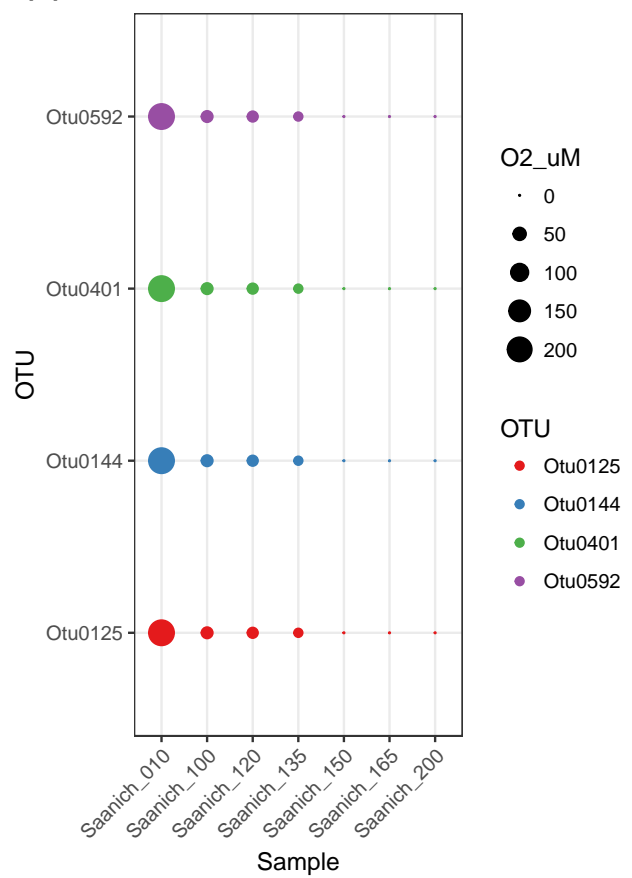


Abundance of ASVs within *Planctomyces* genus across depth from mothur

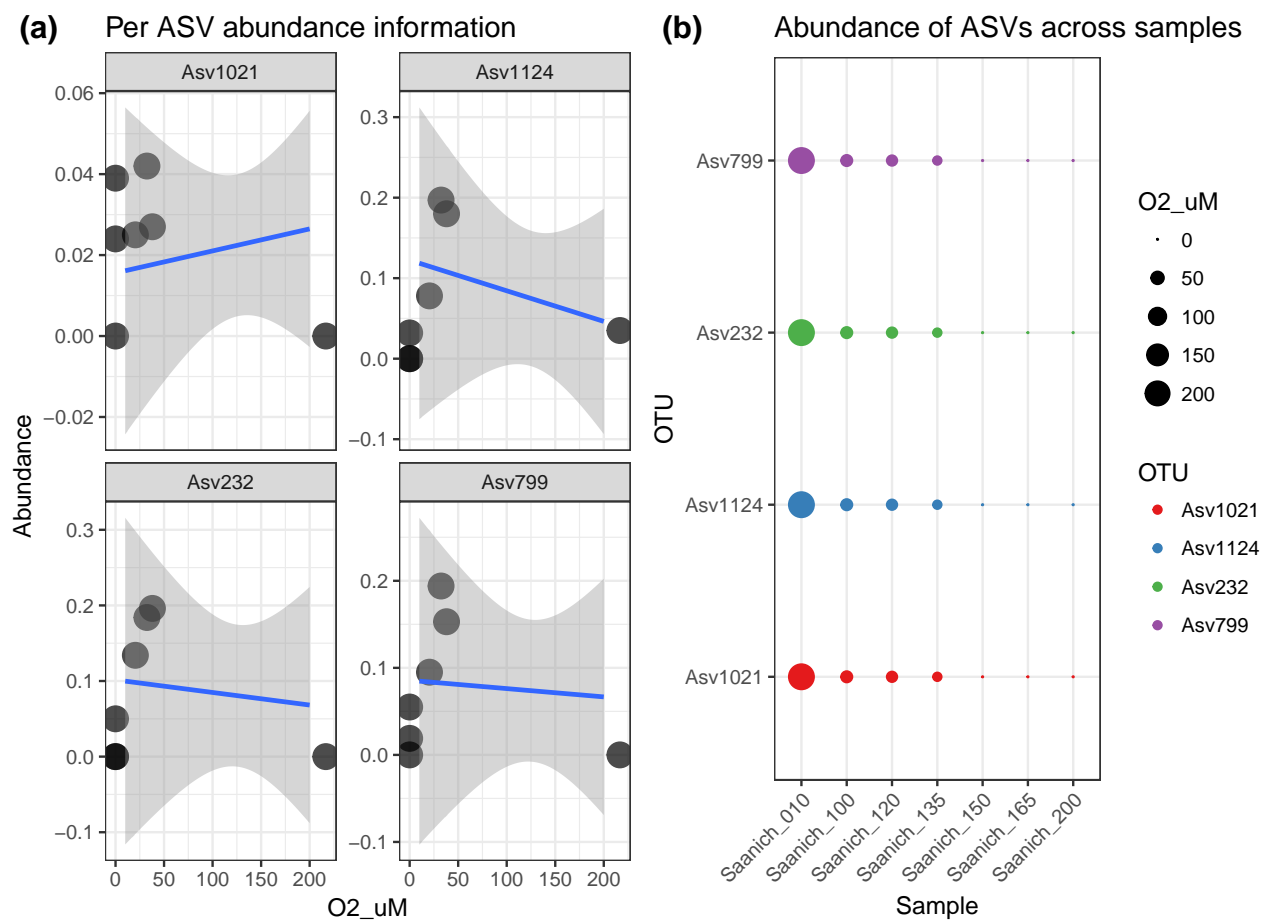
(a) Per OTU abundance information



(b) Abundance of OTUs across samples



Abundance of OTUs within *Planctomyces* genus across oxygen concentration from mothur



Abundance of ASVs within *Planctomyces* genus across oxygen concentration from QIIME2

5 Discussion

(Hawley et al. 2017; Torres-Beltrán et al. 2017)

References

- Hawley, Alyse K, Mónica Torres-Beltrán, Elena Zaikova, David A Walsh, Andreas Mueller, Melanie Scofield, Sam Kheirandish, et al. 2017. "A Compendium of Multi-Omic Sequence Information from the Saanich Inlet Water Column." *Scientific Data* 4. Nature Publishing Group: 170160.
- Torres-Beltrán, Mónica, Alyse K Hawley, David Capelle, Elena Zaikova, David A Walsh, Andreas Mueller, Melanie Scofield, et al. 2017. "A Compendium of Geochemical Information from the Saanich Inlet Water Column." *Scientific Data* 4. Nature Publishing Group: 170159.