are typical of eukaryotes [8]. They are abundant in oceans, freshwater, and soils. They also possess diverse metabolism – such as aerobic chemoheterotrophs – and autotrophic anaerobes that oxidize ammonia to nitrogen (anammox process) [8]. These organisms facilitate the anammox process, contributing approximately 50% of the atmospheric nitrogen molecules in the global nitrogen cycle [8]. Additionally, the anammox process is important in the nitrogen-rich-wastewater remediation technology [8]. Thus, this report further addresses the change in the abundance of *Planctomyces* with depth and oxygen concentration, the richness (OTUs/ASVs) of *Planctomyces*, and the change in the abundance of OTUs/ASVs within *Planctomyces* genus with depth and oxygen concentration in Saanich Inlet.

2 Methods

For both Mothur and QIIME2 pipeline, sequences from Saanich Inlet were amplified using 515F and 808R primers. Sequences were generated on MiSeq with Phred33 quality scores [1]. For both the Mothur and QIIME2 pipelines, the data were cleaned up by filtering out various sequences (Low quality, chimeric sequences, etc.), aligned and classified with SILVA database, and trimmed by their start and end sites [9,10]. The end results for both pipelines were then included in a phyloseq object which contains the following: OTU/ASV Table, Taxonomy, and Sample Metadata [9,10].

The samples were normalized to 100,000 sequences per sample to facilitate comparisons between samples. The normalized counts were then converted to relative abundance percentages. Next, a series of filtering was applied according to two rules: i) Exclude OTUs that are not observed in more than 4 samples; ii) Prune samples and OTUs with unknown values, such as `unclassified` value. This has resulted in thirdMTaxa and thirdQTaxa taxa from Mothur and QIIME2, respectively. No other pre-processing was applied. The implementations are done entirely using R (v3.4.3) and relied on some efficient third-party libraries, such as `phyloseq`, `tidyverse`, `gridExtra`, and `magrittr` [11–16].

```
load("data/mothur_phyloseq.RData")
load("data/qiime2_phyloseq.RData")
set.seed(4832)
rarefiedM <- rarefy_even_depth(mothur, sample.size = 1e+05)
rarefiedQ <- rarefy_even_depth(qiime2, sample.size = 1e+05)
rarefiedMPer = transform_sample_counts(rarefiedM, function(x) 100 *
    x/sum(x))
rarefiedQPer = transform_sample_counts(rarefiedQ, function(x) 100 *
    x/sum(x))
```

```
# First rule
firstMTaxa <- filter_taxa(rarefiedMPer, function(x) sum(x ==
    0) <= 4, TRUE)
firstQTaxa <- filter_taxa(rarefiedQPer, function(x) sum(x ==
    0) <= 4, TRUE)

# Second rule
basedOnGenus <- as.data.frame(tax_table(firstMTaxa)) %>% filter(!str_detect(Genus,
    "uncultured|unclassified"))
secondMTaxa = subset_taxa(firstMTaxa, Genus %in% basedOnGenus$Genus)
```