##############################################################################

############### INTIAL TOP 5+OTHER CPM TOTALS #########################

##############################################################################

### Code copied from 03\_calcs\_SiteWaterAbund.R, then modified to include initials

## top 5 + other abundance and relative abundance

library(tidyverse)

library(writexl)

############## INITIALS, TOP 5 + OTHER, WITH PROPORTIONS #####################

##############################################################################

### Use baseTop5.Rdata as the base. It's the same as volbio\_all\_cr.Rdata, but

## has a column for the top 5 + other taxa group categories

load("data7\_24/FinalAnalysis/baseTop5.Rdata")

initials5cpm <- baseTop5 %>%

filter(exp == "I") %>%

select(samp\_ev, exp, rep, grp\_sz, counts, cpm, group\_size, taxaGroup)

### take the mean of the 3 replicates of each taxaGroup

I5cpmTot <- initials5cpm %>%

group\_by(samp\_ev, grp\_sz) %>%

mutate(cpmTotMnofReps = mean(cpm))

### Remove the unneeded columns

I5cpmTot2 <- I5cpmTot %>%

subset(select = c(samp\_ev, grp\_sz, taxaGroup, cpmTotMnofReps))

### Remove the rows that are "duplicates" because I only want to keep one

## mean value, not the mean value for all three replicates

I5cpmTot3 <- I5cpmTot2 %>% distinct()

### Sum the counts that make up each of the 5+other taxa groups, sum

## cpmTotMnofReps per event per taxa group

I5cpmTot4 <- I5cpmTot3 %>%

group\_by(samp\_ev, taxaGroup) %>%

mutate(totCpmTxEv = sum(cpmTotMnofReps)) %>%

ungroup

### Remove grp\_sz and cpmMnTot... columns, then remove the "duplicates"

## so I have just one entry per event per taxa group, and not the extra rows

## that were the reps

I5cpmTot5<- I5cpmTot4 %>%

subset(select = c(samp\_ev, taxaGroup, totCpmTxEv)) %>%

distinct()

### Add a column of total cpm, all taxa, per sampling event

I5cpmTotEverything <- I5cpmTot5 %>%

group\_by(samp\_ev) %>%

mutate(cpmTotallTxperEv = sum(totCpmTxEv)) %>%

ungroup

### Calculate proportions

### Calculate the proportion of the total cpm that each top5+other taxa group

## contributed to total cpm per event

I5cpmProp <- I5cpmTotEverything %>%

mutate(IcpmProp = totCpmTxEv/cpmTotallTxperEv)

### Test IcpmProp

testI5cpmProp <- I5cpmProp %>%

filter(samp\_ev =="YBP2")

### summed the prop column, it checks out, sums to 1

save(I5cpmProp, file = "Final Final/Abundance/Cells/I5cpmProp.Rdata")

write\_xlsx(I5cpmProp, "Final Final/Abundance/Cells/I5cpmProp.xlsx")

load("Final Final/Abundance/Cells/I5cpmProp.Rdata")

### Plot

source("scripts/01\_function\_wimGraph and Palettes.R")

plot\_I5cpmProp <-ggplot(I5cpmProp, aes(fill=taxaGroup, y=IcpmProp, x=samp\_ev)) +

geom\_bar(position="fill", stat="identity")+

scale\_fill\_manual(values =

c("CenDiaLg" = "cornflowerblue", "CenDiaSm" = "lightskyblue",

"CilLg" = "salmon3", "CilSm" = "salmon1",

"FlagSm" = "#85B22C", "Other" = "peachpuff"),

limits = c("CenDiaLg", "CenDiaSm", "CilLg", "CilSm", "FlagSm",

"Other"),

name = "Taxa Group")+

xlab(NULL)+

ylab(NULL)+

ggtitle("Taxa Group Relative Counts per mL, Initial Samples")+

theme(plot.title = element\_text(hjust = 0.5),

axis.title.y = element\_text(size = 10),

axis.text.y = element\_text(size = 6))+

wimGraph()

plot\_I5cpmProp

load("/Users/allisonadams/My files/Thesis/Microplankton/MicroplanktonAnalysis/Final Final/Abundance/Cells/IcpmProp.Rdata")