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

############# CLEARANCE RATE CALCULATIONS ############################

#################### FOR TOP 5 + OTHER #############################

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

### 11/17/23 Edited from the original df, 03\_calcs\_CrTop5\_0821.R in Final Final/

## Clearance/R and Rdata files:

## I needed to use the df that has a "1" added to the cpmE of

## YBP1 CenDiaLg because it was a zero, which results in an NA for CR, and that

## results in IR non-existent. I re-ran the code in a new R script, Final Final/

## 03\_calcs\_BaseTop5.R.

### Now will re-run CR and IR

### DFs created here:

load("Final Final/Clearance/R and Rdata files/CR\_Reps\_Top5.Rdata")

load("Final Final/Clearance/R and Rdata files/CrMnTop5.Rdata")

load("Final Final/Clearance/R and Rdata files/CR\_Rep\_Mn\_Top5.Rdata")

### 8/21/23

library(tidyverse)

library(writexl)

load("Final Final/baseTop5.Rdata")

source("scripts/01\_function\_clearanceRates.R")

### DF created here: ("Final Final/Clearance/CrMnTop5.Rdata")

### Use baseTop5.Rdata as the base file, because it has the taxaGroup column with

## just top 5 + other

### Select only the columns needed for CR

baseTop5CR <- baseTop5 %>%

select(samp\_ev, exp, rep, taxaGroup, cpm)

### Sum up the cpm for the top 5 + other taxa groups, adding all cpm for organisms

## that fall into those top 5 + other taxa groups, such as, all the small centric diatoms

## in a sampling event, experimental bottle

sumCpm <- baseTop5CR %>%

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

summarise(TotalCpm=sum(cpm),

.groups = 'drop') %>%

as.data.frame()

## Create the base data frame that has only the control and experimental cpm

sumCpm\_CE <- sumCpm %>%

filter( str\_detect(exp, "C|E"))

### Create another df from the above, with only the control samples

sumCpm\_C <- sumCpm\_CE%>%

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

rename(cpmC = TotalCpm)

### Create a df with only experimental samples

sumCpm\_E <- sumCpm\_CE%>%

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

rename(cpmE = TotalCpm)

### Create a df with only initial samples (for the ingestion rates by counts, for later)

sumCpm\_I <- sumCpm%>%

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

rename(cpmI = TotalCpm)

### Apply the mean function to the controls df to get control mean counts

## per ml across the three replicates. Leave out the rep column so

## that what remains in the df is one row for each individual organism/size

## and the mean of the control sample counts per ml or biomass per ml

sumCpm\_Cmn <- sumCpm\_C %>%

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

summarise(CmnCpm=mean(cpmC),

.groups = 'drop') %>%

as.data.frame()

### Since CR needs the mean control samples and the three replicates experimental

## samples, join the experimental sample df with the control means df. This will

## necessarily include the rep column, since we need the experimental samples

## individual replicate counts or biomass for the calculation.

sumCpmE\_Cmn <- left\_join(sumCpm\_E, sumCpm\_Cmn,

by = c("samp\_ev", "taxaGroup"))

### Remove unneeded columns, and rename and re-order remaining columns

## Do I need the exp column?--Can't keep it since E is one column and C

## is another column

names(sumCpmE\_Cmn)

sumCpmE\_Cmn <- select(sumCpmE\_Cmn,

event = samp\_ev,

taxaGroup, rep, cpmE,

CmnCpm)

### Calculate clearance rates. The resulting data frame includes all the replicates

## since the CR was calculated for each replicate.

source("scripts/01\_function\_clearanceRates.R")

CR\_Reps\_Top5 <- rowwise(sumCpmE\_Cmn) %>%

mutate(CRmlcd = cr\_func(controlMnCt = CmnCpm, expCt = cpmE))

write\_xlsx(CR\_Rep\_Mn\_Top5, "Final Final/Clearance/Excel docs/CR\_Reps\_Top5.xlsx")

save(CR\_Rep\_Mn\_Top5,file = "Final Final/Clearance/R and Rdata files/CR\_Reps\_Top5.Rdata")

### Take the mean of the CR per taxaGroup per event

CrMnTop5 <- CR\_Reps\_Top5 %>%

group\_by(event, taxaGroup) %>%

summarize(CrMNmlcd = mean(CRmlcd, na.rm = TRUE))

write\_xlsx(CrMnTop5, "Final Final/Clearance/Excel docs/CrMnTop5.xlsx")

save(CrMnTop5,file = "Final Final/Clearance/R and Rdata files/CrMnTop5.Rdata")

### Join the above two dfs so the CR reps and means are in the same df

CR\_Rep\_Mn\_Top5 <- left\_join(CR\_Reps\_Top5, CrMnTop5,

by = c("event", "taxaGroup"))

write\_xlsx(CR\_Rep\_Mn\_Top5, "Final Final/Clearance/Excel docs/CR\_Rep\_Mn\_Top5.xlsx")

save(CR\_Rep\_Mn\_Top5,file = "Final Final/Clearance/R and Rdata files/CR\_Rep\_Mn\_Top5.Rdata")