## **Order of Calculations for Clearance Rate and Ingestion Rate calculations in R**

and how to combine them

See ~/Documents/Thesis/Microplankton/R Work/MicroplanktonAnalysis/scripts/03\_calcs\_CrFr\_bySizeSmLg.R

## Counts per mL (cpm)

### see Equations for Microplankton Analysis.docx

cpm = counts/(propCntd\*pres\_fact\*vol\_set\_ml)

## Unit Conversions

1 pg = 0.000001 µg

1 µg = 1,000,000 pg

1 mL = 0.001 L

1 L = 1,000 mL

1 pg mL-1 = 0.001 µg L-1 \* \*

µg C L-1 = pg C m L-1 \* .001

## Volume

### see volume dimensions-to µm3-conversions in Equations for Microplankton Analysis.docx

## Biomass

### see biomass volume-to-carbon conversions in Equations for Microplankton Analysis.docx

bio\_pgC\_ml = tot\_biomass\_pgC/(propCntd\*pres\_fact\*vol\_set\_ml)

bio\_ugC\_l = bio\_pgC\_ml\*.001

When calculating straight pgC to µgC , not per mL or L, as in pgC c-1d-1 to µg C c-1d-1, then use the conversion of:

1 pg = 0.000001 µg

1 µg = 1,000,000 pg

## Clearance Rate, ml copepod-1 d-1

(ml of water with that organism in it; i.e., concentration of that organism, cells per mL)

Use the cell count data to calculate clearance rate, since “determination of clearance rate depends upon measured changes in cell concentrations caused by grazers in experimental chambers.” Marin et al., 1986

V /T \* ( lnC -lnE)/n

V= volume of experimental container, in our case, 595 mL

T= time of experiment, in our case, 1 day

C = means of control samples, counts (cells) ml-1

E = experimental samples, counts (cells) ml-1

## Ingestion Rate (aka feeding rate or consumption rate), quantity (biomass pgC, or cell counts) copepod-1 d-1

### Note: Use Initial means in units of pgC mL-1 to calculate the Ingestion Rate, because the Clearance Rate is in units of ml copepod-1 d-1. If I want Ingestion Rate in terms of µg C copepod-1 d-1, then I can do the conversion later by dividing the FR pgC copepod-1 d-1 by 1,000,000, since 1 µg = 1,000,000 pg.

CR x mean I

CR = Clearance rate, in ml copepod-1 d-1

I = initial samples, either in pgC mL-1 or counts mL-1

## Basic Calculations Preparation Steps

1. Start with the basic raw data. Each row contains the following main data points:
   1. Sampling event
   2. Organism type
   3. Organsim dimensions
   4. Number of cells counted
   5. Replicates 1, 2 and 3, for the initial, experimental and control samples
   6. One site water sample for each sampling event
2. Use this basic raw data to calculate
   1. Clearance Rates
      1. Need to calculate the means of the replicates of the control samples counts mL-1.
      2. Use the individual replicates of the experimental samples counts mL-1 in the clearance rate equation.
   2. Ingestion Rates
      1. Need to calculate the means of the replicates of the initial samples biomass pgC mL-1.
      2. Use the clearance rates for the other part of the equation
   3. Abundance, in biomass pgC mL-1 or µgC L-1, means of the replicates of the initial samples, pgC mL-1 or µgC L-1.
3. Combine the organisms into major taxa groups, each with a small and large grouping
   1. Combine the rates and abundance data of the individual organisms into the major groups by summing the data
4. Based on data results, keep the top 5 taxa groups and lump all the others into a group named, “Other”.
   1. Combine the rates and abundance data of the major taxa groups into the top 5 + other by summing the data

## Steps in R

### Clearance Rate:

1. Create the base data frame from the main data frame that has everything, volbio\_all\_cr.Rdata
   1. columns include: samp\_ev, group\_size, exp, rep, cpm, bio\_pgC\_ml
2. Sum up the cpm for the 15 group\_size groups, adding all cpm for organisms that fall into those 15 categories, such as, all the small centric diatoms in a sampling event, experimental bottle
   1. 15 major taxa groups abbreviations: ChlSm, ChlLg, CilSm, CilLg, CyanoSm, CyanoLg, CenDiaLg, CenDiaSm, PenDiaSm, PenDiaLg, DinoLg, FlagSm, FlagLg, UnidSm, UnidLg
3. Create a data frame with only the C and E cpm columns (the control samples and the experimental samples).
4. Create a data frame from the above with only the control samples and another one with only the experimental samples
5. Using the control samples data frame, apply the mean function to counts per mL to get the control means across the three replicates. When writing the group\_by argument, leave out the rep column so that what remains in the data frame is one row for each individiual organism/size and the mean of the control sample counts per ml.
6. Since clearance rate needs the mean control samples and the three replicates experimental samples, join those two data frames into one. This will necessarily include the rep column, since we need the experimental samples individucal replicate counts or biomass for the calculation.
7. Then, with the joined data frame, remove the unneeded columns and rename as needed.
8. Calculate the clearance rate with the clearance rate function
9. The resulting data frame includes all the replicates since the clearance rate was calculated for each replicate.
10. Take the mean CR
11. Replace the NAs with 0 because if not, when R calculates the mean, if one of the reps has NA, but the other one or two have a number, R will not calculate it and will return NA
12. Take the mean of the CR
13. Make a df that combines rep CR with mean CR, to use for plotting the CR means with the reps

### Ingestion Rate

1. Create a data frame with only the I column (the initial samples).
2. As in #3, using the initial samples data frame, apply the mean function to biomass per ml (bio\_pgC\_ml), to get the control means across the three replicates. When writing the group\_by argument, leave out the rep column so that what remains in the data frame is one row for each individual organism/size and the mean of the control sample counts per ml or biomass per ml.
3. Join the initial means data frame with the clearance rate data frame.
4. Remove unneeded columns and rename as needed.
5. Calculate ingestion rate with the feeding rate function.
   1. Ingestion rates were calculated from cell counts of the control beaker and separately for each beaker with grazers. Frost 1972
6. Calculate the mean FR.
7. Make a df that combines rep FR with mean FR.
8. For the biomass ingestion rates, add a column that converts FR from pgC copepod-1 d-1 to ugC copepod-1 d-1, since I want to plot in ug. Divide the FR pgC copepod-1 d-1 by 1,000,000, since 1 µg = 1,000,000 pg.

## Other version of steps to calculating **clearance rates** for each sampling event, by taxa groups

1. Sum the counts per ml (cpm), grouping by sampling event, taxa group, experiment (Control, Experimental, Initial, Site water) and replicate
2. Take the mean of the Control cpm three replicates
3. Calculate clearance rates.

### Other calculations using these clearance rates:

#### Mean of the clearance rates of each taxa group across all sampling events

1. Take the mean of the clearance rates of each taxa group across all sampling events, but in the summarize command, include na.rm = TRUE so that it will calculate the means of the real numbers and ignore the NAs.

## Other version of steps to calculating **ingestion rates** for each sampling event, by taxa groups

## Abundance Calculations

### Use baseTop5 as the base data frame, for calculating with the Top 5 + Other taxa groups.

For Initial Samples

1. Create a df with just the I samples:   
   abundanceI <- baseTop5 %>% filter(exp == "I")
2. Calculate the means of the replicates of biomass pgC mL-1 per entry (per organism + dimensions), then sum those replicate means.
   1. group\_by(samp\_ev, taxaGroup) %>%
   2. summarise(mnBpmItxEv = mean(bio\_pgC\_ml),
   3. .groups = 'drop') %>%
   4. as.data.frame() %>%
   5. rename(event = samp\_ev)

For Site Water Samples

1. Create a df with just the S samples:   
   abundanceS <- baseTop5 %>% filter(exp == "S")
2. Calculate the means of the replicates of biomass pgC mL-1.
   1. group\_by(samp\_ev, taxaGroup) %>%
   2. summarise(mnBpmStxEv = sum(bio\_pgC\_ml), # sum the biomass, pgC per ml by taxaGroup, per samp\_ev. This will sum all the various entries that make up the “Other” taxa group, and all the different sizes of each of the top 5 taxa groups
   3. .groups = 'drop') %>%
   4. as.data.frame() %>%
   5. rename(event = samp\_ev)

Extra step: to calculate micrograms per L-1:

A few things I tried 8/1/23, but didn’t quite work yet:

### Filter for the intial samples only

abundanceI <- baseTop5 %>%

filter(exp == "I")

### take the mean of all the biomass in pgC per ml, by event and taxaGroup

AImnAgg <- aggregate(bio\_pgC\_ml ~ samp\_ev + taxaGroup,

data = abundanceI, mean)

### sum up the counts per ml by event and taxaGroup

AItotCpmAgg <- aggregate(cpm ~ samp\_ev + taxaGroup,

data = abundanceI, sum)

### Join the two data sets together and rename samp\_ev and mean biomass

AIbioMnCpm <- left\_join(AImnAgg, AItotCpmAgg) %>%

rename(mnBioPgMl=bio\_pgC\_ml) %>%

rename(event = samp\_ev)

### Add a column of biomass in ugC per Liter

AIbioMnCpm <- AIbioMnCpm %>%

mutate(mnBioUgL = mnBioPgMl\*.001)

### This above worked, but need to take the means of the replicates,

## not all the biomass data

### Filter for the intial samples only

abundanceI <- baseTop5 %>%

filter(exp == "I")

### take the mean of all the biomass in pgC per ml, by event and taxaGroup

AImnRepsAgg <- aggregate(bio\_pgC\_ml ~ samp\_ev + szesd,

data = abundanceI, mean)

### sum up the counts per ml by event and taxaGroup

AItotCpmAgg <- aggregate(cpm ~ samp\_ev + taxaGroup,

data = abundanceI, sum)

### Join the two data sets together and rename samp\_ev and mean biomass

AIbioMnCpm <- left\_join(AImnAgg, AItotCpmAgg) %>%

rename(mnBioPgMl=bio\_pgC\_ml) %>%

rename(event = samp\_ev)

### Add a column of biomass in ugC per Liter

AIbioMnCpm <- AIbioMnCpm %>%

mutate(mnBioUgL = mnBioPgMl\*.001)