# **Order of Operations for R scripts files**

Shows order of running the scripts to get

* raw data into useable form
* volume and biomass calculations for all data
* other columns useful for organizing and grouping the data
* See ~/Documents/Thesis/Microplankton/R Work/MicroplanktonAnalysis/scripts
* Master raw data counts are from the file DiffDimAll\_Detailed\_Final\_07\_11.xlsx, in /Users/allisonadams/My files/Thesis/Microplankton/MicroplanktonAnalysis/data7\_24/MasterFiles (originally in, but moved from, /Users/allisonadams/My files/Thesis/Microplankton/Microzoop Data Sheets/Data Sheets Detailed/Micropl Data Detailed)

Key:

File Name.R

optional description

the scripts and functions it requires

data files that it creates and where they are saved

02\_dataclean\_propCntd.R

converts the Proportion Counted file to long format

data/MasterFiles/PropCntd\_100\_R.csv

data/MasterFiles/PropCntd\_400.csv

data/MasterFiles/propCntd100\_final.Rdata

data/MasterFiles/propCntd400\_final.Rdata

02\_dataclean\_100\_400.R

Converts the raw data file to long format and separates the sampling event column into sampling event, experiment, replicate

data/MasterFiles/100x\_RawCount\_R.csv

data/MasterFiles/400x\_RawCount\_R.csv

data/MasterFiles/propCntd100\_final.Rdata

data/MasterFiles/propCntd400\_final.Rdata

data/MasterFiles/MasterRFiles/raw100\_final.Rdata

data/MasterFiles/MasterRFiles/raw400\_final.Rdata

03\_calcs\_volume.R

Calculates volume of organisms per cell, and total volume of all those cells in the sample

scripts/01\_function\_volume.R

data /MasterFiles/MasterRFiles/raw100\_final.Rdata

data/MasterFiles/MasterRFiles/raw400\_final.Rdata

data/Calculations/vol100.Rdata

data/Calculations/vol400.Rdata

03\_calcs\_biomass.R

Calculates biomas of organisms per cell, and total biomass of all those cells in the sample

scripts/01\_function\_biomass.R

data/Calculations/vol100.Rdata

data/Calculations/vol400.Rdata

data/Calculations/volbio100.Rdata

data/Calculations/volbio400.Rdata

03\_calcs\_volbio\_100400.R

Combines the 100x and 400x volume/biomass data; adds biomass, pgC mL-1, and

µgC L-1; anc counts mL-1 (cpm)

Later additions include:

* Changing experiment names, T24 -> E; FC -> C; IC -> I; site -> S
* Adding tintinnids to ciliate group; changing agglutinated tintinnides and hyline tintinnids so they would add to the cilate group
* Making centric diatoms a separate category
* Changing all the other diatom groups to pennateDiatom
* Adding ochrophyte to flagellates
* Adding the 15 group\_size names and their abbreviations that combins the names of the group column and the size column

scripts/01\_function\_esd.R

data/Calculations/volbio100.Rdata

data/Calculations/volbio400.Rdata

data/MasterFiles/MasterRFiles/volbio\_all.Rdata

data/Clearance Rates/volbio\_all\_cr.Rdata

!!!! NEW AND UPDATED !!!

**7/24/23 !!!! NEW AND UPDATED !!! AGAIN, 7/24/23 per notes, p 76 of my project notebook, with the same name, volbio\_all\_cr.Rdata. This is the official, latest, updated version as of 7/24/23**

# **Other R scripts and what they calculate**

## **Calculations**

**03\_calcs\_CR\_FR.R**

Calculates clearance rate and ingestion rate of the taxa in their small and large size groups

data/Clearance Rates/volbio\_all\_cr.Rdata

scripts/01\_function\_clearanceRates.R

scripts/01\_function\_feedingRate.R

Clearance Rate:

data/Clearance Rates/sumCpm\_CRmn.Rdata

mean CR for the sampling events, cells per mL

\*\*data/Clearance Rates/**CR\_Rep\_Mn.Rdata**—use this for CR plots

CR, cells per mL, for the individual reps as well as mean CR

Ingestion Rate:

data/Clearance Rates/Feeding Rates/CpmFRMn\_GrpSz.Rdata:

ingestion rate means, cells

data/Clearance Rates/ **FRCells\_Rep\_Mn.Rdata**—use this for FR cells plots

ingestion rates means and reps, cells

data/Clearance Rates/Feeding Rates**/FRBio\_Rep\_Mn.Rdata**—use this for FR biomass plots

ingestion rates means and reps, biomass in pgC copepod-1 d-1 and µgC copepod-1 d-1

data/Clearance Rates/Feeding Rates/FR\_bio\_Size.Rdata

ingestion rates with taxa ordered by large and small

03\_calcs\_Mn\_CR\_FR.R (located in the scripts < Clearance Rates folder)

Creates files that calculate the clearance rates, and ingestion rates for both cell counts and biomass

data/Clearance Rates/volbio\_all\_cr.Rdata

scripts/01\_function\_clearanceRates.R

scripts/01\_function\_feedingRate.R

data/Clearance Rates/cr\_cpm.Rdata (Clearance Rates)

data/Clearance Rates/crmn\_cpm.Rdata (Clearance Rates means)

03\_calcs\_IntlMean\_Fr.R

Calculates means of the initial samples cpm and bpm

data/Clearance Rates/volbio\_all\_cr.Rdata

scripts/01\_function\_feedingRate.R

**~~03\_calcs\_CrFr\_bySizeSmLg.R \* Preferred-~~ ~~the updated version is 03\_calcs\_CR\_FR.R~~**

**Now the updated version is ~/My files/Thesis/Microplankton/MicroplanktonAnalysis/scripts/03\_calcs\_CR\_IR\_New\_06\_03.R**

Calculates clearance rate and ingestion rate of the taxa in their small and large size groups

data/Clearance Rates/volbio\_all\_cr.Rdata

scripts/01\_function\_clearanceRates.R

scripts/01\_function\_feedingRate.R

Clearance Rate:

data/Clearance Rates/crmnsize\_sumAllCpm\_cr.Rdata

mean CR for the sampling events, cells per mL

\*\*data/Clearance Rates/CR\_Rep\_Mn.Rdata

CR, cells per mL, for the individual reps as well as mean CR

data/Clearance Rates/CrGrps.Rdata:

CR means with the group\_size abbreviations column

data/Clearance Rates/CrGrpsReps.Rdata

CR **means** and **reps** with the **group\_size** abbreviations column

Ingestion Rate:

data/Clearance Rates/Feeding Rates/CpmFrMn.Rdata:

ingestion rate means, cells

data/Clearance Rates/Feeding Rates/BpmFrMn.Rdata:

ingestion rate means, biomass

data/Clearance Rates/Feeding Rates/FRcpmBpm.Rdata:

ingestion rate means of both cells per mL and biomass pgC per mL and µg C per L

**data/Clearance Rates/Feeding Rates/FrGrps.Rdata**:

**preferred** df because it includes biomass in ug L-1, and has a column of group\_size names

03\_calcs\_SiteWaterCpm.R

Calculates cpm of site water samples

**New and Improved as of 8/23/23**

03\_calcs\_IrTop5\_0821.R

New ingestion rate calculations for top 5 + other taxa groups

data7\_24/FinalAnalysis/baseTop5.Rdata

source("scripts/01\_function\_feedingRate.R”)

Final Final/Clearance/CrMnTop5.Rdata

Final Final/Ingestion/IrTop5.Rdata

03\_calcs\_Top5.R

Use volbio\_all\_cr.Rdata to make a new taxa group df that lumps together all the low IR bio taxa groups into "other"

The main groups will be "CilLg", "CenDiaLg", "CilSm", "FlagSm", "CenDiaSm"

data/Clearance Rates/volbio\_all\_cr.Rdata

data7\_24/Biomass Analysis 06\_13/baseTop5kept.Rdata

data7\_24/Clearance Rates 2/IrTotAll.Rdata

data7\_24/Clearance Rates 2/IrTotAllTaxaKeptProp.Rdata

03\_calcs\_Proportion of TotIR per Taxa.R

Total Biomass ingestion rate for all taxa groups together, and also for only the taxa groups I kept after excluding some for low total counts or low or neg CR IR

data7\_24/Clearance Rates 2/CrIrCntRepMnTots.Rdata

data7\_24/Clearance Rates 2/IrTotAll.Rdata

and other files, see script

03\_calcs\_Top5PropIRbio.

7/26/23 Copied and pasted the code from 03\_calcs\_Proportion of TotIR per Taxa.R to use for Top 5. No editing done yet, need to do tomorrow

NOTE!!!! I stopped at line 59, can't figure it out. Will come back to it tomorrow.

data7\_24/Clearance Rates 2/IrTotAllTaxaKeptProp.Rdata

data/FinalAnalysis/baseTop5.Rdata

data/FinalAnalysis/CrIrAb5.Rdata: Top 5 taxa groups, plus “Other” group, had CR, FR, Total Cpm initial samples (abundance) for each sampling event

03\_calcs\_Abun\_Other.R

Several Calculations for looking at IR, CR, abundance, etc., to be able

to whittle down groups for final analysis

data/Clearance Rates/volbio\_all\_cr.Rdata

data/FinalAnalysis/baseTop5.Rdata

data/Clearance Rates/CR\_IRbio\_mn.Rdata

data/Abundance/abundance.Rdata

data/FinalAnalysis/CrIrAbCell.Rdata

data/FinalAnalysis/IRtotByEvent.Rdata

data/FinalAnalysis/CrIrTotAbCell.Rdata

data/Abundance/over200.Rdata

data/Abundance/abundance5.Rdata

data/FinalAnalysis/CrIrAb5Cell.Rdata

03\_calcs\_BiomassAbundance.R

Biomass, pgC ml-1 and µgC L-1

Includes plots as well

data7\_24/FinalAnalysis/baseTop5.Rdata

Final Final/Abundance/AImnAgg5.Rdata

Final Final/Abundance/AISumAgg5.Rdata

Final Final/Abundance/AImnAgg17.Rdata

Final Final/Abundance/AISumAgg17.Rdata

Final Final/Abundance/AI5TotProp.Rdata

## **Plots**

04\_plots\_IrRepMns.R

Creates plots of biomass ingestion rate, 17 taxa groups, reps with means dots

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

load("data7\_24/Clearance Rates 2/CrIrCntRepMn.Rdata")

Final Final/Ingestion/plot\_IR17all.pdf

04\_plots\_siteWater.R

Creates plots of site water samples abundance, counts or biomass

scripts/01\_function\_wimGraph and Palettes.R

data/SiteWater/siteCntBio.Rdata

04\_plots\_IrBioRelAbunRainbow.R

Relative abundance/ingestion rate, made by Odinga K on fiverr.com

data/Clearance Rates 2/IrTotAllTaxaKeptProp.csv, but look in data7\_24 for an Rdata file

04\_plots\_Prop\_12\_07\_23.R

Proportion of whole, ir bio, by taxa and by events, single bar

I entered by hand the data from

data7\_24/Clearance Rates 2/IrTotAllTaxaKeptProp.Rdata,

created in 03\_calcs\_Proportion of TotIR per Taxa.R

04\_plots\_AbunBio.R

Abundance plots, August 2023

Plot Abundance in biomass, absolute and proportionate by event, Top 5 + Other taxa groups

Final Final/Abundance/AI5TotProp.Rdata, For proportions and absolute, from 03\_calcs\_BiomassAbundance.R

Plots saved in MicroplanktonAnalysis/Final Final/Abundance/

~~04\_plots\_FRgroups.R~~

~~Creates plots of ingestion rates by cell counts and biomass in small/large taxa groups~~

~~scripts/01\_function\_wimGraph and Palettes.R~~

~~data/Clearance Rates/Feeding Rates/FRcpmBpm.Rdata~~

~~data/Clearance Rates/Feeding Rates/FrGrps.Rdata~~

~~AllGroupsPlot\_IngtsRt\_Bio.pdf~~

~~AllGroupsPlot\_IngstRt\_Cpm.pdf~~

~~04\_plots\_CRFR\_Samp\_Ev.R~~

~~Creates plots of clearance rates and ingestion rates as in 04\_plots\_FRgroups.R, but with individual sampling events~~

~~data/Clearance Rates/Feeding Rates/FrGrps.Rdata~~

~~data/Clearance Rates/CrcpmBpl.Rdata~~

04\_plots\_Abundance.R

04\_plots\_IR\_Various.R

04\_plots IR Rep Means.R

04\_plots\_IR\_Top5.R

04\_plots\_IR\_Reps.R

04\_plots\_FRgroups.R

04\_plots\_test 04\_12.R

04\_plots CR RepMeans.R

Plot\_Templates.R

Code for various types of plots

scripts/01\_function\_wimGraph and Palettes.R

scripts/01\_function\_Scale Break.R

# **Master Source Files**

Raw data excel spreadsheet

DiffDimAll\_Detailed\_Final\_07\_11.xlsx < Micropl Data Detailed < Data Sheets Detailed < Microzoop Data Sheets < Microplankton < Thesis < Documents

Base files for loading into R

Raw Counts:

100x\_RawCount\_R.xlsx

400x\_RawCount\_R.xlsx

Proportion of slide counted (for calculating counts per ml)

PropCntd\_100\_R.xlsx

PropCntd\_400\_R.xlsx

< MasterFiles < data< MicroplanktonAnalysis < R Work < Microplankton < Thesis < Documents

Order of Operations for getting to FrGrps.Rdata

which is the master df with ingestion rate calculations with the 8 taxa groups

**03\_calcs\_CrFr\_bySizeSmLg.R**

Goes backwards from FrGrps.Rdata

* FrGrps.Rdata
  + Feeding rate, cells, and biomass in pgC mL-1 and µgC L-1, with taxa group\_size abbreviations
* FRcpmBpm.Rdata
  + Same data as FrGrps.Rdata, except it does not have the 8 taxa group\_size abbreviations
  + This has the biomass in µgC per L, others below only have pgC per mL
* CpmFrMn and BpmFrMn, joined together
  + Each contains the ingestion rate means, in either cells or biomass
* sumAllBpm\_fr and sumAllCpm\_fr
  + Feeing rate calculated
* sumAllBpm\_Imn and sumAllCpm\_Imn
  + Means of initial samples, needed for the ingestion rate equation
* sumAllBpm\_I and sumAllCpm\_I
  + Only initial samples in this df, for calculating the initial sample means, above
* sumAllCpm\_cr
  + Clearance rates, to be used to calculate ingestion rates for both cells and biomass

Extra scripts

* FRcpm\_Rep\_Mn
  + Feeding rate, cells, means and replicates
* FRBugAllMn and FrAllOverall\_ug
  + Feeding rate of taxa group\_size abbreviation, all sampling event data grouped together, biomass in µgC L-1
* FrCpmOverall
  + Feeding rate of taxa group\_size abbreviation, all sampling event data grouped together, cells mL-1