**Script file path that has the counts and biomass data needed for making graphs; shows where to find data frames and plots**

Raw data excel spreadsheet

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

1. **Master Raw Data**
   1. 02\_dataclean\_100\_400.R
      1. Uses
      2. libraries: tidyverse, writexl
         1. data/MasterRawData/100x\_RawCount\_R.csv
         2. data/MasterRawData/400x\_RawCount\_R.csv
      3. Creates
         1. data/MasterRawData/MasterRFiles/raw400\_final.Rdata
         2. data/MasterRawData/MasterRFiles/raw100\_final.Rdata
   2. 02\_dataclean\_samp\_ev\_separate.R
      1. Shows how I separated the sampling event names from the experiment types and replicates, used in 02\_dataclean\_100\_400.R
2. **Functions**
   1. Volume: scripts/01\_function\_volume.R
   2. Biomass: scripts/01\_function\_biomass.R
   3. Equivalent Spherical Diameter : scripts/01\_function\_esd.R
   4. Clearance Rate: scripts/01\_function\_clearanceRates.R
   5. Feeding Rate: scripts/01\_function\_feedingRate.R
3. **Calculations**
   1. Volume; 03\_calcs\_volume.R
      1. Uses
         1. libraries: tidyverse, writexl
         2. "scripts/01\_function\_volume.R"
         3. data/MasterRawData/MasterRFiles/raw100\_final.Rdata
         4. data/MasterRawData/MasterRFiles/raw400\_final.Rdata
      2. Creates
         1. data/Calculations/vol100.Rdata
         2. data/Calculations/vol100\_no0.Rdata
         3. data/Calculations/vol400.Rdata
         4. data/Calculations/vol400\_no0.Rdata
      3. Also adds a column that calculates the mean volume per organism
   2. Biomass; 03\_calcs\_biomass.R
      1. Uses
         1. libraries: tidyverse, writexl
         2. scripts/01\_function\_biomass.R
         3. data/Calculations/vol100.Rdata
         4. data/Calculations/vol400.Rdata
      2. Creates
         * 1. data/Calculations/volbio100.Rdata
           2. data/Calculations/volbio100\_no0.Rdata
           3. data/Calculations/volbio400.Rdata
           4. data/Calculations/volbio400\_no0.Rdata
      3. Also adds a column that calculates the mean biomass per organism
   3. 100x and 400x Volume and Biomass combined; 03\_calcs\_volbio\_100400.R
      1. Uses
         1. libraries: tidyverse, writexl, formattable
         2. data/Calculations/volbio100.Rdata
         3. data/Calculations/volbio100\_no0.Rdata
         4. data/Calculations/volbio400.Rdata
         5. data/Calculations/volbio400\_no0.Rdata
         6. scripts/01\_function\_esd.R
      2. Creates
         1. data/MasterFiles/MasterRFiles/volbio\_all.Rdata
      3. Adds the following columns: total volume µm3; mean volume per organism µm3, total biomass pgC; biomass per volume pgC ml-1; total counts per ml; esd.
   4. Biomass and Counts totals per sample; 03\_Biomass\_Means\_Final\_Sz\_2022\_11\_09.R
      1. Uses data/MasterFiles/MasterRFiles/volbio\_all.Rdata
      2. Creates
         1. Updates volbio\_all with four more columns (I added these to the orginal volbio\_all script in 03\_calcs\_volbio\_100400.R.
            1. grp\_typ: combines the group column with the type column
            2. grp\_sz: combines grp\_typ with size to add dimensions
            3. grp\_esd: combines grp\_typ with esd
         2. Changes the total counts column to counts per mL, since the volumes of samples settled isn’t identical for each 100x and 400x sample.
         3. count\_sum\_rep\_fin\_sz : total counts per ml of all samp\_ev, exp, and rep, for grp\_sz
         4. count\_sum\_rep\_fin\_grptyp: total counts per ml of all samp\_ev, exp, and rep, for grp\_typ
         5. count\_sum\_rep\_fin\_grpesd: total counts per ml of all samp\_ev, exp, and rep, for grp\_esd
         6. ctsum\_rep\_mn\_esd and ctsum\_rep\_mnesd\_no0: calculates the means of the totalCPM (counts per ml) across the replicates; n0 removes the rows that have a mean of 0, since that means there were no counts in those replicates; includes columns, samp\_ev, exp, grp\_esd and mn\_ct\_ml.
         7. ctsum\_rep\_sz\_fct24: as above, with sizes only, and removes the Initials and site water samples
         8. ctsum\_rep\_esd\_fct24: as above, but with esd instead of sz
         9. Contains several other R data files that I ended up not using.
   5. See what the most common taxa groups are: 03\_calcs\_Taxa\_Grp\_Exploration.R
      1. Uses
         1. data/MasterFiles/MasterRFiles/volbio\_all.Rdata
         2. 01\_function\_wimGraph and Palettes.R
      2. Creates
         1. taxapic.Rdata: selects samp\_ev, grp\_typ, counts\_per\_ml and adds them up per each sampling event
         2. taxaPicAll.Rdata: lists all the different Group names and their totalCPM
         3. taxa\_list.Rdata: lists all the taxa/size groups, and adds up their totalCPM
         4. Other data frames that single out a taxon, such as taxaListDolicho.Rdata, etc.
   6. Single Organism Analysis to Help Decide how to Lump different sizes into larger Groups: 03\_calcs\_GroupCollapse.R
      1. Uses
         1. libraries: tidyverse, writexl, formattable
         2. Need to load wimGraph and wimPalettes
         3. volbio.all as source file
      2. Creates data files and plot code for multiple organism groups such as pennate diatoms, centric diatoms, flagellate cryptomonas, etc. All organism groups have their own versions of the following example of pennate diatoms:
         1. taxaPen: only pennate diatoms; includes samp\_ev, exp, rep, mag, Group (group, type, size, esd), esd, counts\_per\_ml, bio\_per\_org\_pgC
         2. pennateLook: pretty much the same as taxaPen, but separates out grp\_sz and esd, and includes vol\_per\_org\_um3
         3. taxaPenLump: collapses all matching Group entries into one, with total mean biomass per organism. I couldn’t make R add the volume as well, due to some sort of typo, presumably, which I hope to fix.
         4. penPlot: only the esd and the totalCPM (counts per mL)
      3. Saved in the folder data/TopTen

**To Be Continued…**

1. Mnct\_GrpTyp\_Graphs.R
   1. Grpsz\_reps.Rdata
2. ctbm\_grptyp.Rdata
   1. 03\_Taxa\_Groups.R
   2. 03\_Bmean\_Ct\_GrpTyp.R (Script that combines group and type into one category, “grp\_typ”, and also adds a column for mean count across the replicates)
   3. ct\_bmn\_fin.Rdata
   4. ctsum\_rep\_mn\_no0.Rdata
      1. 03\_Biomass\_Means\_Final\_2022\_11\_09.R
      2. 03\_Biomass\_Means\_Final\_10\_25\_22.R (Script that calculates the total counts and biomass means of each organism per sampling event and experiment type (site, FC, IC or T24)
         1. volbio\_all\_no0.Rdata (Rdata file that has the biomass of 100 and 400, but removed the rows with 0 zero counts): *Contains everything except the means*
         2. OR volbio\_all.Rdata (Rdata file that has the biomass of 100 and 400, buthas rows with 0 zero counts) *Contains everything except the means; also includes ESD calculations now (11/18)*
            1. 03\_calcs\_volbio\_100400.R (R script that combines the 100x biomass with the 400x biomass)

volbio100.Rdata (Rdata file that has calculated the biomass)

AND volbio400.Rdata (same as above)

03\_calcs\_biomass.R (Contains the function to calculate the biomass of the organisms)

vol100.Rdata (Rdata file that has calculated the volume)

AND vol400.Rdata (same as above)

03\_calcs\_volume.R (Contains the function to calculate the volume of the organisms)

I had to start over with the list because Word wouldn’t make more indents, but the list continues below

1. 03\_calcs\_volume.R (Contains the function to calculate the volume of the organisms)
   1. raw100\_final.Rdata (Rdata file that has rearranged 100x\_RawCount\_R.csv to put it in a usable format)
   2. AND raw400\_final.Rdata (same as above)
      1. 02\_dataclean\_100\_400.R
         1. 100x\_RawCount\_R.csv. (Raw data file, also in xlsx)
         2. 400x\_RawCount\_R.csv (Raw data file, also in xlsx)

**KEY**

|  |  |  |
| --- | --- | --- |
| Column name | What it means | An example |
|  |  |  |
| samp\_ev | Sampling Event | SJR1, WLD2, etc |
| samp\_date | Date that the sample was taken | 2019-09-03 |
| exp | Type of experiment sample | FC (controls), IC (initials), T24 (experimenal), site (site water sample) |
| Group or group | Organism group (I’m not identifying them by species or genus, just broad “groups”) | chlorohyte, ciliate, diatom, etc. |
| type | more specific type of organism | actinastrum, as in chlorophyte actinastrum; or cone, as in ciliate cone |
| shp | Shape of the organism- I made these up and used them in the volume calculations so they correspond to the correct volume equations for the shape of the organism | cone1, prispar (prism parallelogram) |
| sa | short axis measurment in micrometers | 8 or other numbers |
| la | long axis measurement in micrometers | same as above |
| wi | width measurement in micrometers, or equivalent third dimensional measurment. Not all organisms have this | same as above |
| counts | how many individual organisms I counted in that sample | same as above |
| pres\_fact | preservative factor of the sample, it’s a calculation that factors in to the total volume of the water sample, the volume of the preservative that was added to the sample | .0922 or something like that |
| vol\_set\_ml | volume of sample settled in milliliters- the volume I took out of the main sample bottle and settled the microplankton on the bottom of the utermohl chamber | 27.0 or something like that |
| n\_bugs | number of bugs, or copepods that were put in the incubation bottle so they could eat the plankton in it. | 24, always 24 |
| time\_d | time in days that the incubation bottles were incubated |  |
| volume  also tot\_vol\_um3 | the total volume of all the organisms in the count for that row, in units of cubic micrometers | varies |
| vol\_per\_org\_um3 | the volume of each organism, if there were multiple counts of them in that row, in units of cubic micormeters |  |
| biomass\_pgC  also tot\_biomass\_pgc | biomass in units of picograms of carbon | varies |
| bio\_per\_vol\_pgc\_ml | biomass per volume in picograms of carbon per milliliter | varies |
| bio\_per\_vol\_ugl | biomass per volume in micrograms per liter | varies |
| counts\_per\_ml | number of organisms per milliliter in that sample | varies |
| bmn\_pgml | mean biomass in picograms of carbon per milliliter | varies |
| tot\_ct | total counts of that organism, across the experimentals samples |  |