## **Order of Calculations from Raw Data to Clearance Rate, Ingestion Rate and Abundance**

1. Prepare the raw data, 02\_dataclean\_100\_400.R
   1. Raw data
      1. These operations are done separately for the 100x raw data and the 400x raw data, and later combined into a master data frame.
      2. data/MasterFiles/100x\_RawCount\_R.csv and
      3. data/MasterFiles/400x\_RawCount\_R.csv
   2. Convert to long format
   3. Rearrange the column orders to have the station/sample column first, using the select function, and then change the NAs in the counts column to 0s using the mutate function.
   4. Add the headers column to include the other data
      1. samp\_date, cnt\_date, sample, pres\_fact, vol\_set\_ml, n\_bugs, time\_d
      2. headers\_100 <- read\_csv("data/MasterFiles/100xHeaders.csv
   5. Join headers df to raw100\_fin df
   6. Separate the sampling event names into three columns: sampling event, experimental and replicate
      1. First, make a separate column for the sampling event
      2. Then, using the above file, make a separate column for the experiment type
      3. Then using the above file, make a separate column for the replicate; tN/A will be filled in where the experiment was "site" and there were no replicates
      4. Then, separate the experiment type from the replicate number
   7. Reorder the columns and remove sample, Organism, name, count dates, since those columns were replaced or revised; also add a column that indicates if it was in the 100x or 400x raw data files.
   8. Correct the date format
   9. Add the proportion counted column, created in 02\_dataclean\_propCntd.R
   10. Repeat all the above for the 400x raw data.1
2. Calculate Volume, 03\_calcs\_volume.R
   1. Uses data files raw100\_final and raw400\_final from 02\_dataclean\_100\_400.R
      1. These operations are done separately for the 100x data and the 400x data, and later combined into a master data frame.
   2. Uses volume function, 01\_function\_volume.R
   3. Load the raw counts data, file created in 02\_dataclean\_100\_400.R
   4. Run the volume function
      1. Volume function returns volume per cell in cubic micrometers (um^3), vol\_per\_cell\_um3
      2. tot\_volume\_um3 is the volume per cell times the counts (number of organisms counted)
      3. Also, add a column that calculates the mean volume per organism
3. Calculate the biomass, 03\_calcs\_biomass.R
   1. Uses data files vol100 and vol400, which were created in the above file, 03\_calcs\_volume.R
      1. These operations are done separately for the 100x data and the 400x data, and later combined into a master data frame.
   2. Uses biomass function, 01\_function\_biomass.R
   3. Load vol100, which is the data file with the volume calculations, created in 03\_calcs\_volume.R
   4. Run the biomass function
      1. Units of the biomass equation results are pgC per cell
      2. biomass\_cell\_pgC is the biomass per individual cell
      3. tot\_biomass\_pgC is the total biomass of all the cells counted
4. Join the 100x and 400x volume and biomass data into one data frame, 03\_calcs\_volbio\_100400.R
   1. To be used as a base data file for other calculations
   2. Calculations include
      1. volume per organism, um^3
      2. total volume, um^3
      3. Biomass per cell pgC
      4. Total biomass, pgC
   3. Load the 100x and 400x volume/biomass data files, created in 03\_calcs\_biomass.R
   4. Join the two datasets; add columns as follows:
      1. biomass per milliliter, pgC per ml:
         1. total biomass in pgC divided by volume, which is the proportion of sample counted, times preservative factor, times volume of sample settled: Total biomass: tot\_biomass\_pgC/(propCntd\*pres\_fact\*vol\_set\_ml) and then name it bio\_pgC\_ml
      2. Convert pgC per mL to ugC per L (bio\_pgC\_ml\*.001)
      3. Counts per milliliter (counts/(propCntd\*pres\_fact\*vol\_set\_ml)
      4. Equivalent Spherical Diameter
         1. Use the equation for the volume of a sphere, V= 4/3\*pi\*r3, and solve for r, and then double to get the diameter.
         2. r = the cube root of 0.75V/pi, so the function will be that the column esd = 2\* the cube root of (0.75V/pi), where V= volume
      5. Several columns that put group categories together
         1. grp-typ so that the organism group and type are together
         2. grp\_sz, so that the organism group and type and size are together
         3. grp\_esd, so tht the organsim group and type and esd are together
      6. Size categories, esd < 15 is small, esd >= 15 is large
      7. Change the experiment names:
         1. T24 = E, for experimental samples
         2. FC = C, for control samples
         3. IC = I for initial samples
         4. site = S, for site water samples
      8. Change organism group names as follows, so that I can consolidate them for calculating main group CR and IR, and save as a separate data frame, volbio\_all\_cr.Rdata
         1. Note that chlorophytes, cyanobacteria and unidentified don't need consolidating, as groups will be calclulated all together as is.
         2. Add tintinnids to the ciliate Group
         3. Make centric diatoms their own category
         4. All other diatoms are changed to Group, "pennateDiatom", and keep their type category
         5. Change Group "Ochrophyte" to flagellate, because what I was calling Ochrophytes were what looked to me like synura, which are flagellates, but I couldn’t identify them as such definitively.
         6. Reorder the columns
5. Calculate clearance rates and ingestion rates for both cell counts and biomass, 03\_calcs\_Mn\_CR\_FR.R
   1. See Order of Calculations CR and IR.docx for the full order of operations for both clearance rates and ingestion rates