* Summary .csv file that contains the following, organized by sample type (media, control, infected)
  + Number of peaks per file
  + Percentage of peaks assigned for each file
  + Range, mean, and standard deviation for percentage of peaks assigned per sample type
  + Total number of assigned unique peaks detected per sample type across all files
  + Total number of assigned peaks that occur in 50% or greater of the files per sample type (≥18 for infected, ≥18 for control, ≥12 for media)
* For each sample type (3), create a csv file that contains the assigned peaks that occur in 50% or greater of the files per sample type (≥18 for infected, ≥18 for control, ≥12 for media)
  + For ‘Peak Height’, ‘Resolving Power’, ‘m/z error (ppm)’, ‘Confidence Score’, report averages. Add one new column for each in final summary files to list standard deviations as well
  + Remove columns ‘Calibrated m/z’, ‘m/z Error Score’, ‘Isotopologue Similarity’,
* Create background (media) subtracted files for control and infected
  + Create two varieties: one that removes samples features that are also present in the blank, and one that removes sample features if their peak heights are less than 3x greater than that of the blank
* For the final, background subtracted control and infected files, align them to create one file that contains features common to both (For Peak Height’, ‘Resolving Power’, ‘m/z error (ppm)’, ‘Confidence Score’, do not average. E.g., have two columns for peak height. One that lists the peak height for the control, one that lists the peak height for the infected), one file for unique infected, and one file for unique control (do this for both kinds of blank subtracted files).