taxaPen <- subset(volbio\_all,

select = c(samp\_ev, exp, rep, mag, grp\_sz, esd, szGroup, counts\_per\_ml, bio\_per\_vol\_pgc\_ml)) %>% mutate(totalMnCPM=counts\_per\_ml,totalMnBPM= bio\_per\_vol\_pgc\_ml)

taxaPen$Group <- paste(taxaPen$grp\_sz, taxaPen$esd)

taxaPen <- filter(taxaPen, grepl('pennate', Group))

taxaPen <- subset(taxaPen,totalMnCPM !=0)

taxaPen <- subset(taxaPen,select = c(samp\_ev, exp, rep, mag, Group, grp\_sz, esd, szGroup,

totalMnCPM, totalMnBPM))

taxaPen$szesd <- paste(taxaPen$grp\_sz, taxaPen$esd)

taxaPen$totalMnCPM<-formattable(taxaPen$totalMnCPM,

format="f",digits=2)

taxaPen$totalMnBPM<-formattable(taxaPen$totalMnBPM,

format="f",digits=2)

### Add up the counts per ml for each distinct pennate

## size/esd name but keep the esd and biomass columns

taxaPenlumpC <- aggregate(totalMnCPM ~ szesd +esd,

data = taxaPen, FUN = sum, na.rm =TRUE)

taxaPenlumpB <- aggregate(totalMnBPM ~ szesd + esd, data = taxaPen, FUN = sum, na.rm =TRUE)

taxaPenlump <- merge(taxaPenlumpC, taxaPenlumpB, by="szesd")

taxaPenlump<- subset(taxaPenlump,

select = c(szesd, esd.x, totalMnCPM, totalMnBPM))

colnames(taxaPenlump)[1] = "Group"

colnames(taxaPenlump)[2] = "esd"

### Put the pennates in small, medium and large

taxaPenlump$szGroup <- with(taxaPenlump, ifelse(esd < 12, 'small',ifelse(esd > 24, 'large', ifelse(esd >= 12 & esd <= 24, "medium", "WHAT?"))))

taxaPenlumpAgg <- aggregate(totalMnCPM ~ szGroup,

data = taxaPenlump, FUN = sum, na.rm =TRUE)

# Compare to the size groups I made just for the pennates by looking at the plot

taxaPenlump2<- taxaPenlump

taxaPenlump2$szGroup <- with(taxaPenlump, ifelse(esd < 10, 'small', ifelse(esd > 22, 'large', ifelse(esd >= 11 & esd <= 22, "medium", "WHAT?"))))

taxaPenlump2 <- aggregate(totalMnCPM ~ szGroup,

data = taxaPenlump2, FUN = sum, na.rm =TRUE)

write\_xlsx(taxaPenLump, "data/TopTen/taxaPenLump.xlsx")

save(taxaPenLump, file = "data/TopTen/taxaPenLump.Rdata")