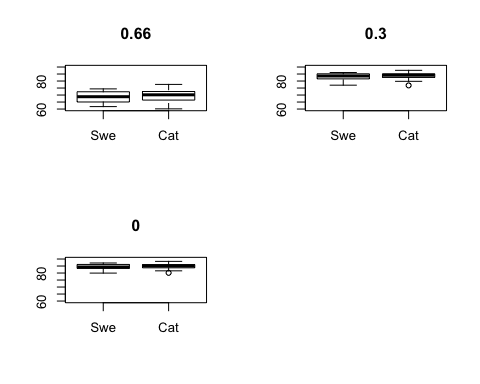
STAR\_mapping\_comp

KN

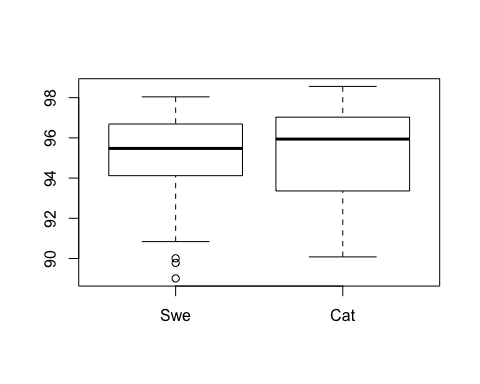
8 april 2020

## R Markdown

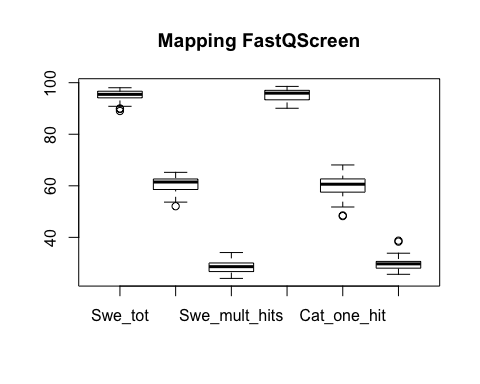
mapping\_results <- read.table("2b\_Mapping\_STAR\_STEP2\_comp\_settings\_cor.txt", header = T)  
  
cat\_subset <- mapping\_results[1:41, ]  
swe\_subset <- mapping\_results[42:66, ]  
par(mfrow=c(2,2))  
boxplot(mapping\_results[42:66, ]$uniq\_mapped\_reads\_066, mapping\_results[1:41, ]$uniq\_mapped\_reads\_066, ylim=c(60,90), names = c("Swe", "Cat"), main="0.66")  
boxplot(mapping\_results[42:66, ]$uniq\_mapped\_reads\_03, mapping\_results[1:41, ]$uniq\_mapped\_reads\_03, ylim=c(60,90), names = c("Swe", "Cat"), main="0.3")  
boxplot(mapping\_results[42:66, ]$uniq\_mapped\_reads\_0, mapping\_results[1:41, ]$uniq\_mapped\_reads\_0, ylim=c(60,90), names = c("Swe", "Cat"), main="0")



filtering\_sumstats <- read.table("filtering\_sumstats.txt", header = T)  
#excluding the gut samples 67-72 (both reads, rows 131-142)  
filtering\_sumstats <- filtering\_sumstats[1:130, ]  
   
boxplot(filtering\_sumstats[83:130, ]$Mapped\_lept, filtering\_sumstats[1:82, ]$Mapped\_lept, names = c("Swe", "Cat"))



boxplot(filtering\_sumstats[83:130, ]$Mapped\_lept, filtering\_sumstats[83:130, ]$one\_hit, filtering\_sumstats[83:130, ]$mult\_hits, filtering\_sumstats[1:82, ]$Mapped\_lept, filtering\_sumstats[1:82, ]$one\_hit, filtering\_sumstats[1:82, ]$mult\_hits, names = c("Swe\_tot", "Swe\_one\_hit", "Swe\_mult\_hits", "Cat\_tot", "Cat\_one\_hit", "Cat\_multiple"), main="Mapping FastQScreen")



summary(filtering\_sumstats)

## Sample\_ID Raw\_reads TrimGalore   
## P11917\_101\_S49\_R1: 1 Min. : 8321308 Min. : 7676231   
## P11917\_101\_S49\_R2: 1 1st Qu.: 9819672 1st Qu.: 9000224   
## P11917\_102\_S50\_R1: 1 Median :12021534 Median :11249409   
## P11917\_102\_S50\_R2: 1 Mean :12150053 Mean :11312275   
## P11917\_103\_S51\_R1: 1 3rd Qu.:13970952 3rd Qu.:12864385   
## P11917\_103\_S51\_R2: 1 Max. :18440161 Max. :17621986   
## (Other) :124   
## Prinseq CutAdapt Condetri   
## Min. : 6727947 Min. : 6566045 Min. : 6442542   
## 1st Qu.: 7963397 1st Qu.: 7754868 1st Qu.: 7620653   
## Median :10075653 Median : 9771315 Median : 9600375   
## Mean :10151854 Mean : 9900921 Mean : 9730005   
## 3rd Qu.:11666888 3rd Qu.:11422534 3rd Qu.:11203335   
## Max. :16433969 Max. :16142258 Max. :15894538   
##   
## SortMeRNA Map\_lept\_rep Mapped\_lept one\_hit   
## Min. : 5416797 Min. :0.3900 Min. :89.01 Min. :48.25   
## 1st Qu.: 7358189 1st Qu.:0.5125 1st Qu.:93.89 1st Qu.:57.78   
## Median : 8977748 Median :0.6400 Median :95.68 Median :60.93   
## Mean : 9011092 Mean :0.7128 Mean :95.20 Mean :60.26   
## 3rd Qu.:10350433 3rd Qu.:0.8675 3rd Qu.:96.97 3rd Qu.:62.65   
## Max. :15433681 Max. :1.4200 Max. :98.56 Max. :68.09   
##   
## mult\_hits   
## Min. :24.08   
## 1st Qu.:27.80   
## Median :29.18   
## Mean :29.35   
## 3rd Qu.:30.54   
## Max. :38.72   
##

colMeans(filtering\_sumstats[ ,2:7])

## Raw\_reads TrimGalore Prinseq CutAdapt Condetri SortMeRNA   
## 12150053 11312275 10151854 9900921 9730005 9011092

colSums(filtering\_sumstats[ ,2:7])

## Raw\_reads TrimGalore Prinseq CutAdapt Condetri SortMeRNA   
## 1579506858 1470595814 1319740998 1287119692 1264900712 1171441934

Summary Sample\_ID Raw\_reads TrimGalore Prinseq CutAdapt Condetri SortMeRNA  
P11917\_101\_S49\_R1: 1 Min. : 8321308 Min. : 7676231 Min. : 6727947 Min. : 6566045 Min. : 6442542 Min. : 5416797  
P11917\_101\_S49\_R2: 1 1st Qu.: 9819672 1st Qu.: 9000224 1st Qu.: 7963397 1st Qu.: 7754868 1st Qu.: 7620653 1st Qu.: 7358189  
P11917\_102\_S50\_R1: 1 Median :12021534 Median :11249409 Median :10075653 Median : 9771315 Median : 9600375 Median : 8977748  
P11917\_102\_S50\_R2: 1 Mean :12150053 Mean :11312275 Mean :10151854 Mean : 9900921 Mean : 9730005 Mean : 9011092  
P11917\_103\_S51\_R1: 1 3rd Qu.:13970952 3rd Qu.:12864385 3rd Qu.:11666888 3rd Qu.:11422534 3rd Qu.:11203335 3rd Qu.:10350433  
P11917\_103\_S51\_R2: 1 Max. :18440161 Max. :17621986 Max. :16433969 Max. :16142258 Max. :15894538 Max. :15433681  
(Other) :124  
Map\_lept\_rep Mapped\_lept one\_hit mult\_hits  
Min. :0.3900 Min. :89.01 Min. :48.25 Min. :24.08  
1st Qu.:0.5125 1st Qu.:93.89 1st Qu.:57.78 1st Qu.:27.80  
Median :0.6400 Median :95.68 Median :60.93 Median :29.18  
Mean :0.7128 Mean :95.20 Mean :60.26 Mean :29.35  
3rd Qu.:0.8675 3rd Qu.:96.97 3rd Qu.:62.65 3rd Qu.:30.54  
Max. :1.4200 Max. :98.56 Max. :68.09 Max. :38.72

colMeans Raw\_reads TrimGalore Prinseq CutAdapt Condetri SortMeRNA 12150053 11312275 10151854 9900921 9730005 9011092

ColSums Raw\_reads TrimGalore Prinseq CutAdapt Condetri SortMeRNA 1579506858 1470595814 1319740998 1287119692 1264900712 1171441934