

Alignment Stats

Lucas Michel Todó, Cristina Bancells
Alfred Cortes and Juan R. Gonzalez

Barcelona Global Health Institute (ISGlobal), Campus PRBB

March 3, 2017

Contents

1	Importing Data	2
2	Plots	3
2.1	MAPQ	3
2.2	Fragment Length	4

1 Importing Data

First we import the data for every alignment:

```
#### Fetch files and create iterables ####
file_names <- list.files("./inputs")
file_names <- gsub("_read[12]_sample.fastq", "", file_names)
file_names <- unique(file_names)
dir = "./params_%/"
params <- c(1:3)
for(element in params){
  gsub("%%", element, dir)
}

mapq <- read.csv2(file = "./params_1/A7K9_14456_TTAGGC_MAPQ.csv", sep = "\t", header = TRUE)

## Warning in file(file, "rt"): cannot open file './params_1/A7K9_14456_TTAGGC_MAPQ.csv': No such file or directory
## Error in file(file, "rt"): cannot open the connection

lens <- read.csv2(file = "./params_1/A7K9_14456_TTAGGC_lengths.csv", sep = "\t", header = TRUE)

## Warning in file(file, "rt"): cannot open file './params_1/A7K9_14456_TTAGGC_lengths.csv': No such file or directory
## Error in file(file, "rt"): cannot open the connection

lens_something <- read.csv2(file = "./params_1/A7K9_14456_TTAGGC_somethingunusual.csv", sep = "\t", header = TRUE)

## Warning in file(file, "rt"): cannot open file './params_1/A7K9_14456_TTAGGC_somethingunusual.csv': No such file or directory
## Error in file(file, "rt"): cannot open the connection

abs_lens <- abs(as.numeric(lens))

## Error in eval(expr, envir, enclos): object 'lens' not found

abs_lens_something <- abs(as.numeric(lens_something))

## Error in eval(expr, envir, enclos): object 'lens_something' not found

max(as.numeric(lens))

## Error in eval(expr, envir, enclos): object 'lens' not found
```

```

which(lens == 1057118)

## Error in which(lens == 1057118): object 'lens' not found

purged_lens <- abs_lens

## Error in eval(expr, envir, enclos): object 'abs_lens' not found

purged_lens[86703:86704] <- c(0,0)

## Error in purged_lens[86703:86704] <- c(0, 0): object 'purged_lens'
not found

```

2 Plots

2.1 MAPQ

```

x <- as.data.frame(as.numeric(mapq))

## Error in as.data.frame(as.numeric(mapq)): object 'mapq' not found

colnames(x) <- "MAPQ"

## Error in colnames(x) <- "MAPQ": object 'x' not found

ggplot(x, aes(x = MAPQ)) +
  geom_histogram(binwidth = 1) +
  labs(x = "MAPQ", y = "Count") +
  scale_x_continuous(breaks = seq(0, 45, by = 2), limits = c(0,48)) +
  scale_y_continuous(breaks = seq(0,35000, by = 2500))

## Error in ggplot(x, aes(x = MAPQ)): object 'x' not found

table(x > 5)

## Error in eval(expr, envir, enclos): object 'x' not found

```

2.2 Fragment Length

```
x <- as.data.frame(purged_lens[purged_lens != 0])  
  
## Error in as.data.frame(purged_lens[purged_lens != 0]): object  
'purged_lens' not found  
  
colnames(x) <- "len"  
  
## Error in colnames(x) <- "len": object 'x' not found  
  
ggplot(x, aes(x = len)) +  
  geom_histogram(bins = 100) +  
  labs(x = "Lengths", y = "Count") +  
  scale_x_continuous(breaks = seq(75, 550, by = 25), limits = c(75, 550))  
  
## Error in ggplot(x, aes(x = len)): object 'x' not found
```

And some text after.