

# QAA

**Plot the trimmed read length distributions for both R1 and R2 reads (on the same plot).**

You can produce 2 different plots for your 2 different RNA-seq samples. There are a number of ways you could possibly do this. One useful thing your plot should show, for example, is whether R1s are trimmed more extensively than R2s, or vice versa. Comment on whether you expect R1s and R2s to be adapter-trimmed at different rates.

```
R1_29 <- as.data.frame(read_tsv('29_R1_trimmed.tsv', col_names =FALSE))

## Rows: 2041795 Columns: 1

## -- Column specification -----
## Delimiter: "\t"
## dbl (1): X1

##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
R2_29 <- as.data.frame(read_tsv("29_R2_trimmed.tsv", col_names =FALSE))

## Rows: 2041795 Columns: 1

## -- Column specification -----
## Delimiter: "\t"
## dbl (1): X1

##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
R_29 <- bind_rows(R1_29,R2_29)
R_29$Direction <- rep( c("Forward", "Reverse"), c(nrow(R1_29), nrow(R2_29)))

R1_8 <- as.data.frame(read_tsv("8_R1_trimmed.tsv", col_names=FALSE))
```

```
## Rows: 34791161 Columns: 1

## -- Column specification -----
## Delimiter: "\t"
## dbl (1): X1
```

```
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
R2_8 <- as.data.frame(read_tsv("8_R2_trimmed.tsv", col_names=FALSE))
```

```
## Rows: 34791161 Columns: 1
```

```
## -- Column specification -----
## Delimiter: "\t"
## dbl (1): X1
```

```
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

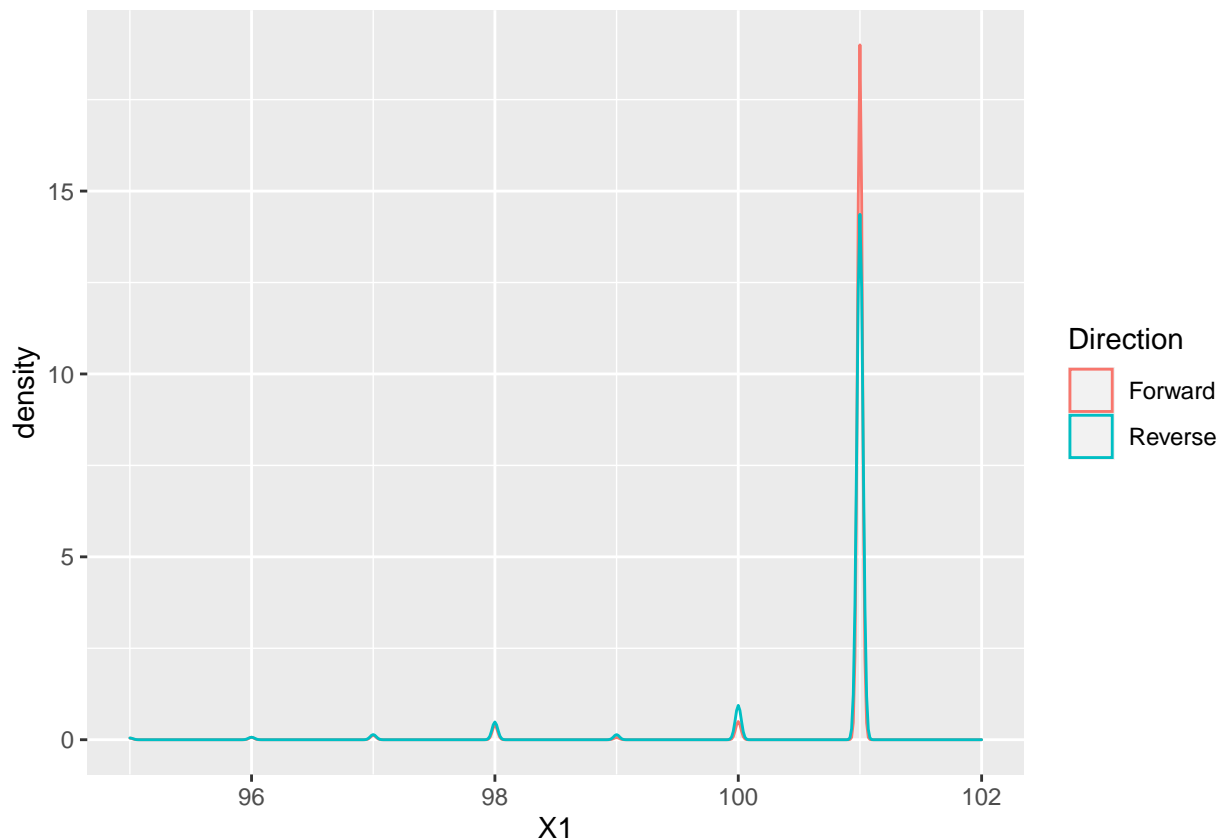
```
R_8 <- bind_rows(R1_8,R2_8)
```

```
R_8$Direction <- rep( c("Forward", "Reverse"), c(nrow(R1_8), nrow(R2_8)))
```

```
# str(R1_8)
```

```
ggplot(data=R_8, aes(x=X1, color=Direction)) + geom_density() + scale_x_continuous(limits = c(95,102))
```

```
## Warning: Removed 3968592 rows containing non-finite values (stat_density).
```



```
ggplot(data=R_29, aes(x=X1, color=Direction)) + geom_density() + scale_x_continuous(limits = c(95,102))
```

```
## Warning: Removed 294076 rows containing non-finite values (stat_density).
```

