

QC of Single-Cell RNA-seq samples

Dan Spakowicz

September 28, 2015

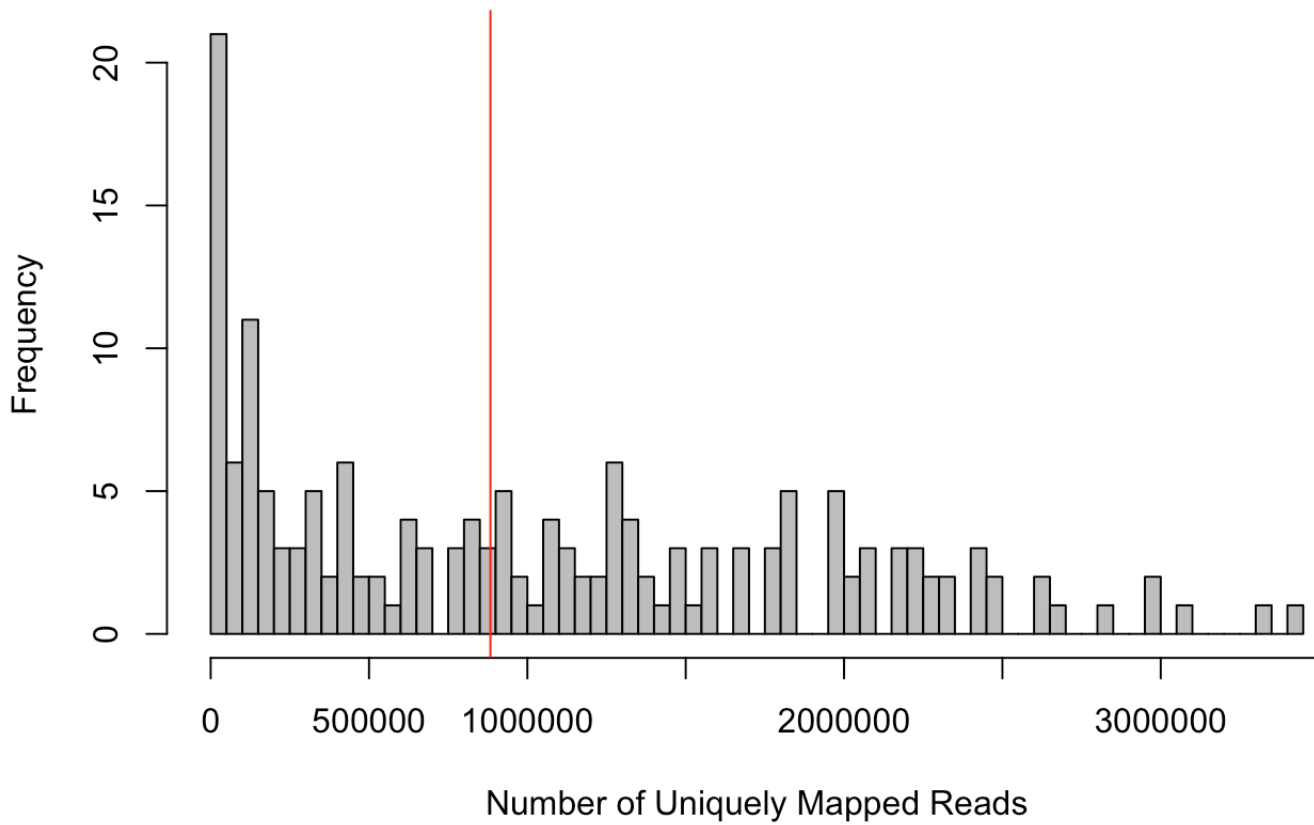
These samples were mapped using the ENCODE STAR pipeline.

```
setwd("~/Box Sync/projects/chupp/")
df <- read.table("QC.all")
colnames(df) <- c("Sample", "NumInputReads", "UniquelyMappedReadsNum", "UniquelyMappedReadsPerc")
summary(df)
```

```
##           Sample      NumInputReads      UniquelyMappedReadsNum
## Sample_790-SC-0_C81:  1      Min.      :    4800      Min.      :    613
## Sample_790-SC-1_C01:  1      1st Qu.: 939030      1st Qu.: 198268
## Sample_790-SC-1_C02:  1      Median :1677638      Median : 883306
## Sample_790-SC-1_C03:  1      Mean    :1785843      Mean    :1031480
## Sample_790-SC-1_C04:  1      3rd Qu.:2507301      3rd Qu.:1697007
## Sample_790-SC-1_C05:  1      Max.    :4000000      Max.    :3434389
## (Other)                :162
## UniquelyMappedReadsPerc
## Min.      : 0.62
## 1st Qu.:28.95
## Median   :64.63
## Mean     :52.28
## 3rd Qu.:77.03
## Max.     :85.86
##
```

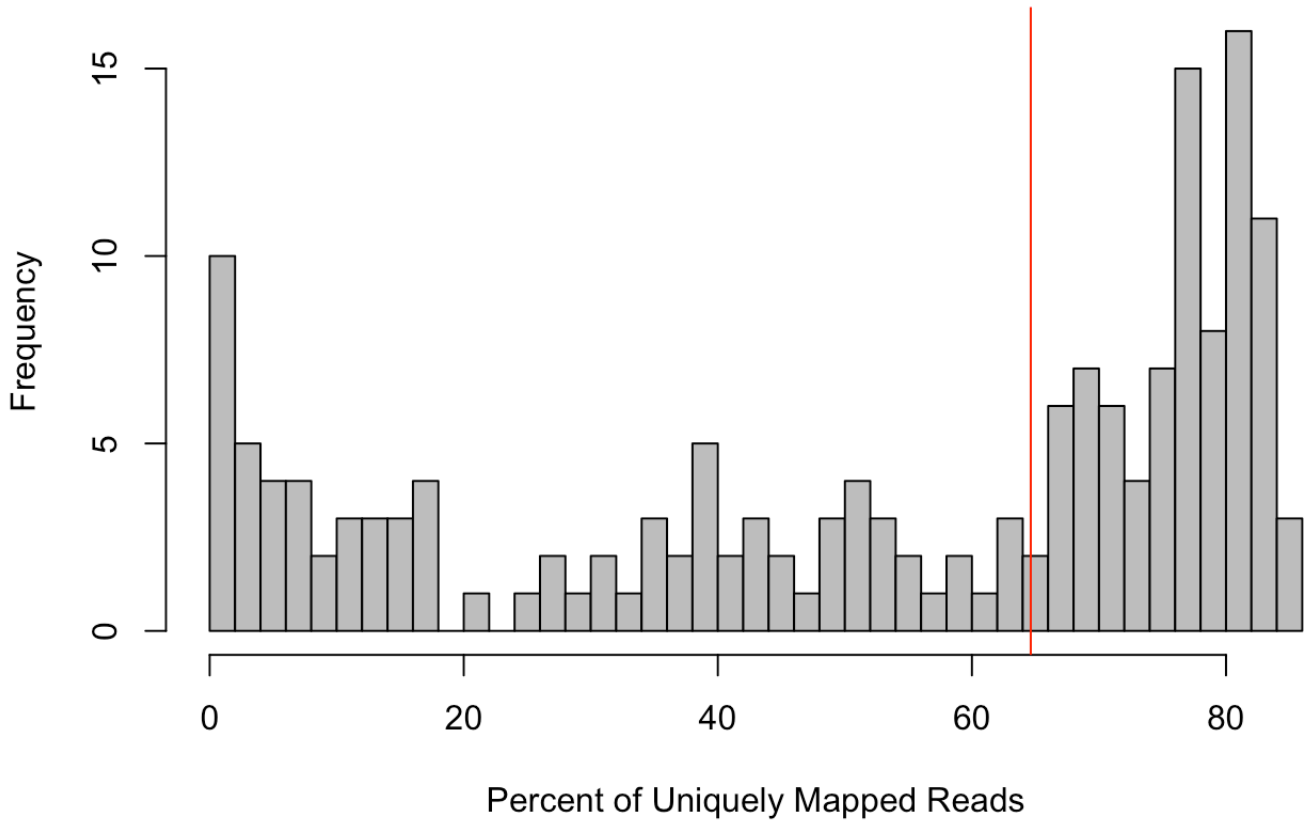
```
hist(df$UniquelyMappedReadsNum, col = "grey", xlab = "Number of Uniquely Mapped Reads", main = "Histogram of the Number of Uniquely Mapped Reads", 100)
abline(v=(median(df$UniquelyMappedReadsNum)), col = "red")
```

Histogram of the Number of Uniquely Mapped Reads



```
hist(df$UniquelyMappedReadsPerc, 50, col = "grey", xlab = "Percent of Uniquely Mapped Reads", main = "Histogram of the Percent of Uniquely Mapped Reads")
abline(v=(median(df$UniquelyMappedReadsPerc)), col = "red")
```

Histogram of the Percent of Uniquely Mapped Reads



The vertical red lines indicate the median values.