

Class 13: UNIX For Bioinformatics

Tom Quach (A15549142)

11/13/2021

Storing our blast result into b

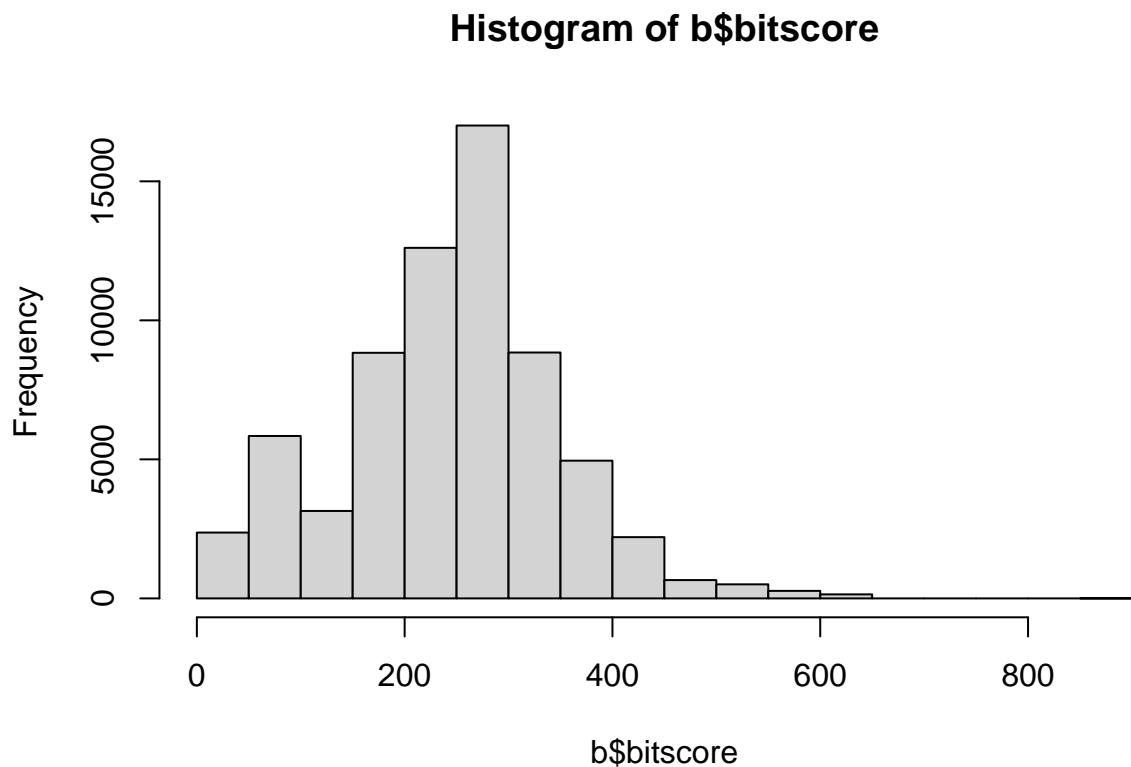
```
b <- read.delim("mm-second.x.zebrafish.tsv")
```

Setting the column names for the file

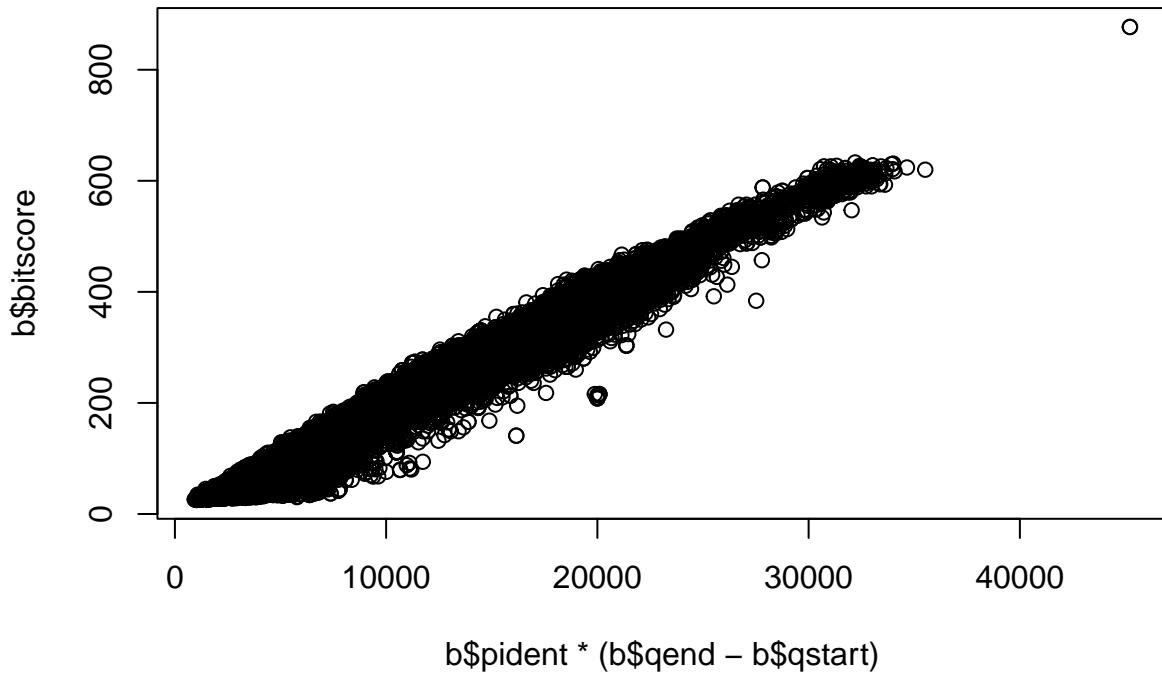
```
colnames(b) <- c("qseqid", "sseqid", "pident", "length", "mismatch", "gapopen", "qstart", "qend", "sstart", "sseqid", "strand", "qlen", "slen", "evalue", "bitscore")
```

Making a histogram of the \$bitscore values

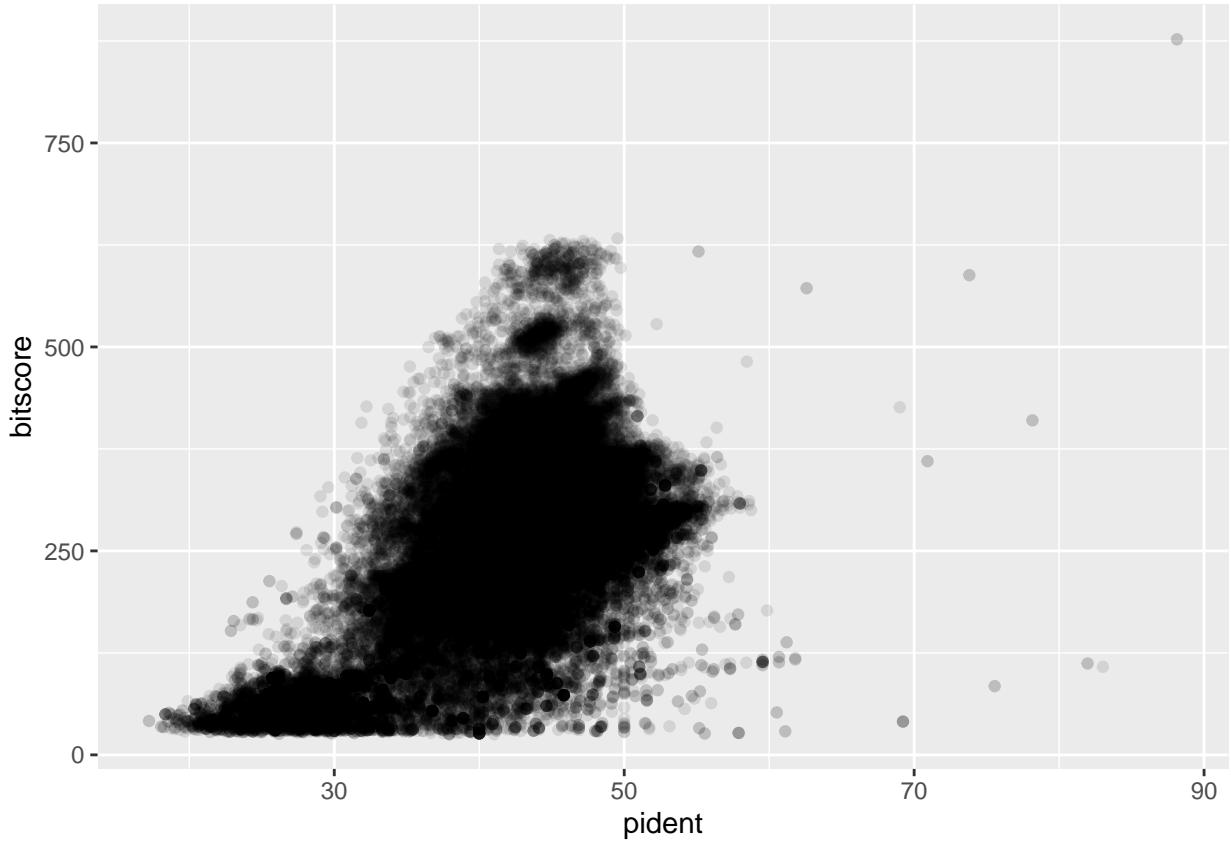
```
hist(b$bitscore, breaks = 30)
```



```
## Assuming your blast results are stored in an object called 'b'  
plot(b$pident * (b$qend - b$qstart), b$bitscore)
```



```
library(ggplot2)  
ggplot(b, aes(pident, bitscore)) + geom_point(alpha=0.1)
```



```
ggplot(b, aes((b$pident * (b$qend - b$qstart)), bitscore)) + geom_point(alpha=0.1) + geom_smooth()  
  
## Warning: Use of 'b$pident' is discouraged. Use 'pident' instead.  
  
## Warning: Use of 'b$qend' is discouraged. Use 'qend' instead.  
  
## Warning: Use of 'b$qstart' is discouraged. Use 'qstart' instead.  
  
## Warning: Use of 'b$pident' is discouraged. Use 'pident' instead.  
  
## Warning: Use of 'b$qend' is discouraged. Use 'qend' instead.  
  
## Warning: Use of 'b$qstart' is discouraged. Use 'qstart' instead.  
  
## 'geom_smooth()' using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

