

# Blast Results Analysis

Ivy Yang

3/1/2022

Read the results TSV (Tab separated format) file from BLAST

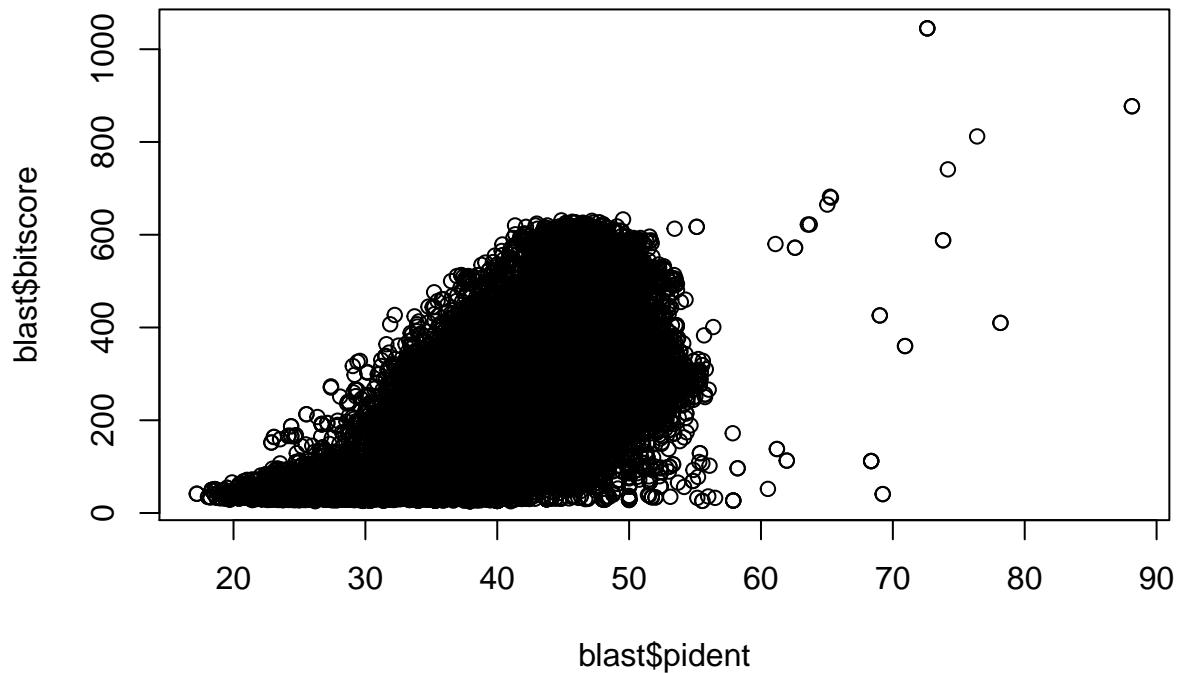
```
blast <- read.table("mm-second.x.zebrafish.tsv")
head(blast)
```

```
##          V1          V2          V3          V4          V5          V6          V7          V8          V9          V10          V11          V12
## 1 YP_220550.1  NP_059331.1  69.010      313       97       0       4      316      10      322  1.24e-150   426.0
## 2 YP_220551.1  NP_059332.1  44.509      346      188       3       1      344      1      344  8.62e-92   279.0
## 3 YP_220551.1  NP_059341.1  24.540      163      112       3      112      263      231      393  5.14e-06   49.7
## 4 YP_220551.1  NP_059340.1  26.804      97       65       2      98      188      200      296  1.00e-01   35.8
## 5 YP_220552.1  NP_059333.1  88.132      514      61       0       1      514      1      514  0.00e+00  877.0
## 6 YP_220552.1 XP_021326074.1 31.818      66       32       2      427      482      16      78  6.70e+00   29.3
```

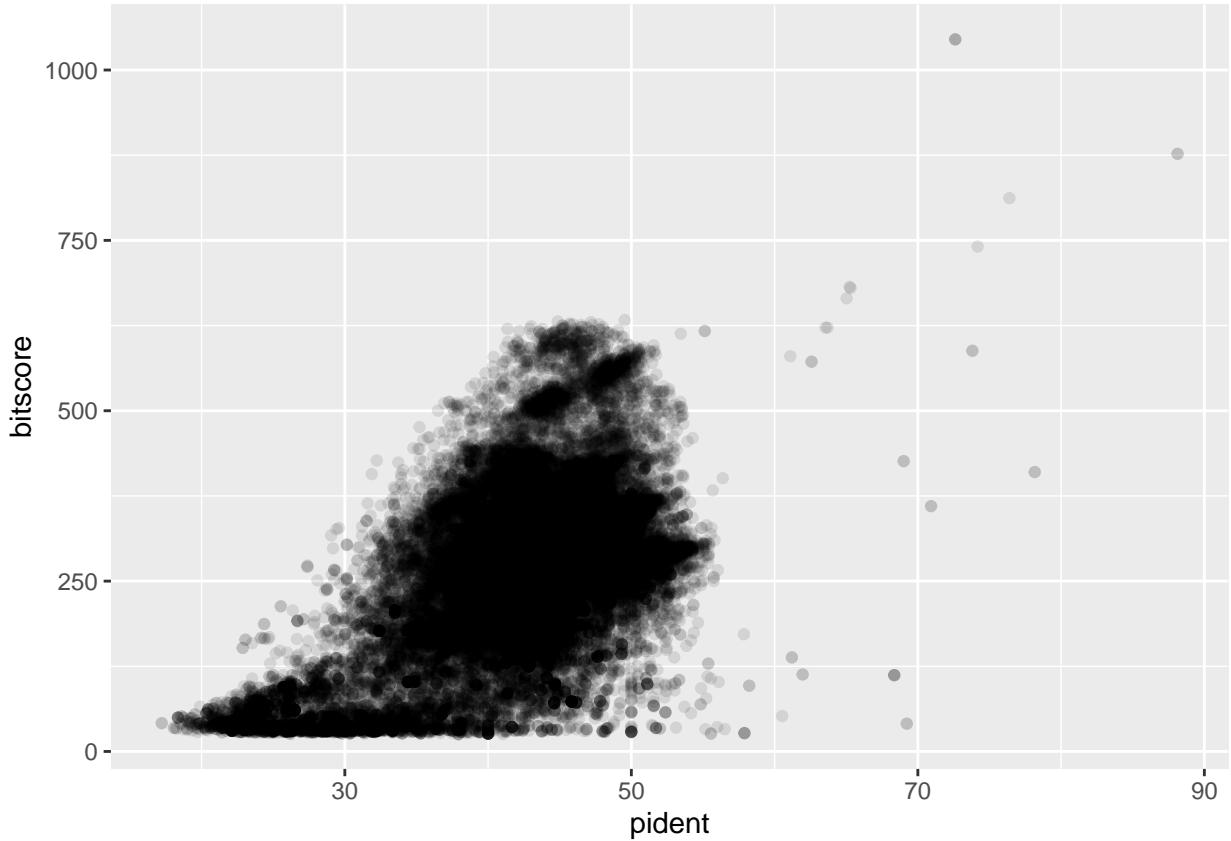
```
colnames(blast) <- c("qseqid", "sseqid", "pident", "length", "mismatch", "gapopen", "qstart", "qend", "tstart", "tend", "score", "evalue", "bitscore")
```

Plot some aspects

```
plot(blast$pident, blast$bitscore)
```



```
# Let's try using ggplot2
library(ggplot2)
ggplot(blast,aes(pident,bitscore)) + geom_point(alpha=0.1)
```



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

