

Lab 15

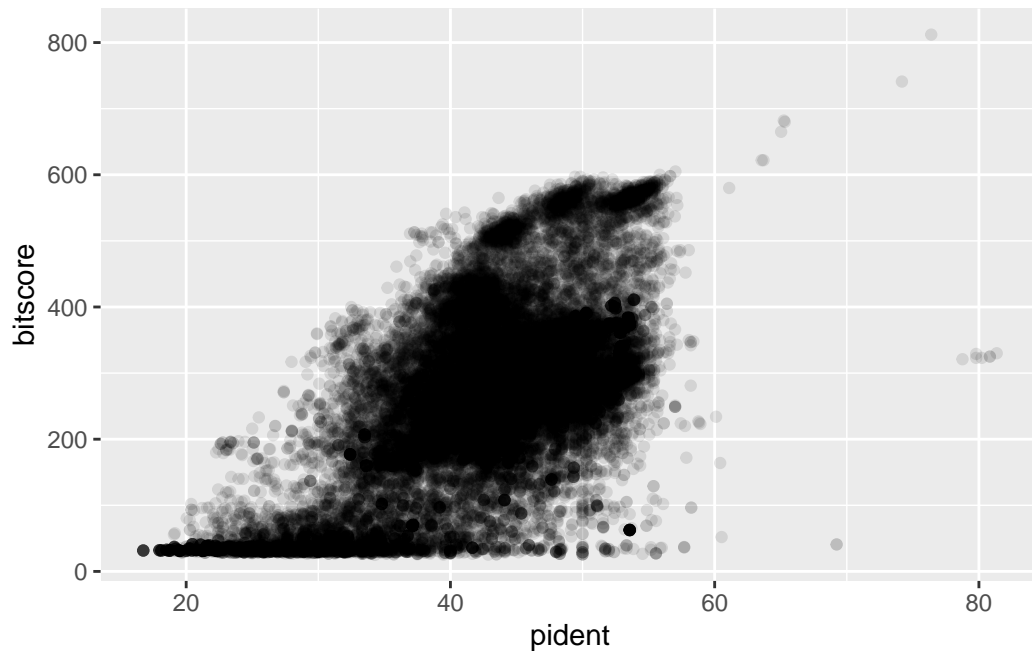
Bomin Xie A16144147

UNIX for bioinformatics

1. The star character “*” is used as a wildcard character to represent any sequence of characters. Right here, it is used to match and expand all the files in the current directory that have a “.faa.gz” extension.
2. There are 41 sequences in `mouse.1.protein.faa`.
3. Running without `> mm-first.faa` part, it will print out the first 11 line of the file without saving it.
4. It will save the first 11 line of the file into `mm-first.faa` file.
5. I use the command `grep -c . mm-second.faa` to determine the number of sequences.

```
blast_zebrafish <- read.csv("mm-second.x.zebrafish.tsv", sep="\t")
colnames(blast_zebrafish) = c("qseqid", "sseqid", "pident", "length", "mismatch", "gapopen")
```

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



```
ggplot(blast_zebrafish, aes((blast_zebrafish$pident * (blast_zebrafish$qend - blast_zebrafish$qstart) + blast_zebrafish$qstart), blast_zebrafish$bitscore)) +
  geom_point(alpha=0.1) + geom_smooth()
```

Warning: Use of `blast_zebrafish\$pident` is discouraged.
i Use `pident` instead.

Warning: Use of `blast_zebrafish\$qend` is discouraged.
i Use `qend` instead.

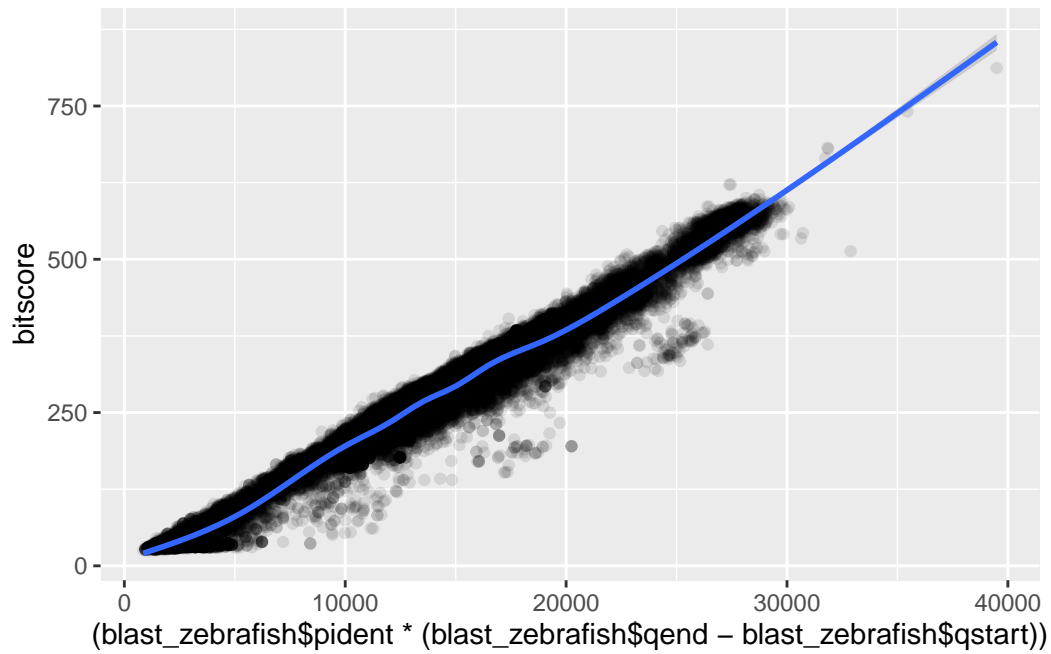
Warning: Use of `blast_zebrafish\$qstart` is discouraged.
i Use `qstart` instead.

Warning: Use of `blast_zebrafish\$pident` is discouraged.
i Use `pident` instead.

Warning: Use of `blast_zebrafish\$qend` is discouraged.
i Use `qend` instead.

Warning: Use of `blast_zebrafish\$qstart` is discouraged.
i Use `qstart` instead.

```
`geom_smooth()` using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```



6. The -r option stands for “recursive,” that allow scp to copy directories and their contents recursively, ensuring that all files and subdirectories within the specified directory are copied to the destination. The * character is used as a wildcard to represent any sequence of characters in the filename, which allows the match of multiple files or directories.