Mapping Warmup: BLAT

BLAT is an incredible useful tool to have in ones bioinformatics tool kit. If you are already familiar with it then skip this exercise. If you have never used it then get some practice with it.

In

```
$ROOT45/Labs/2_Mapping/data/exercise_2_1_blat.fasta
```

are FASTA sequences from the Human genome. Determine what they are using BLAT (http://genome.ucsc.edu). It is under the tools menu. They sequences are cDNA from from build hg19. Note; UCSC has now switch over to hg38 as the default. You could probably BLAT these sequences against that build but practice setting the correct genome build hg19.

N.B., remember to source your `config.sh` file if you can not find that sequence file.

Extra credit:

For biologists

Answer the following questions (without using google):

- Which of the gene(s) identified are important in cancer biology? Why?

- Which gene(s) are least important in cancer biology? What disease are they likely to be relavant too?

- Which gene(s) are important in the immune system? Why?

- Which gene is missing in half (approximately) of you?

For computer scientists

Same questions, but you can use google.

Extra, extra credit

- Which of these genes is mostly like to be of interest to one group of mathematicians.