Week 1: Foundations of Computing, UNIX Practical
MSc/MRes CMEE 2014-15

Samraat Pawar

Imperial College London

October 10, 2014

PRACTICAL: MAKE SURE THE BASICS WORK I

- In the directory `UNIX/Data/fasta` you find some FASTA files

- These files have an header starting with $>$ followed by the name of the sequence and other metadata

- Starting from the second line, we have the sequence data.

- Write a file called `UnixPrac1.txt` with UNIX shell commands that do the following (Number each command with a hashed comment):

  1. Count how many lines are in each file
  2. Print everything starting from the second line for the `E. coli` genome
  3. Count the sequence length of this genome
  4. Count the matches of a particular sequence, "ATGC" in the genome of *E. coli* (hint: Start by removing the first line and removing newline characters)
  5. Compute the AT/GC ratio