Advanced Bioinformatics Workshop 1 Questions - Linux

1. Create a new directory to store the data and perform the exercises.

/mnt/c/Users/Mohammed Bilal$

1. How many sequences in the fasta file?

grep –c “^>” AB\_test\_dta.fasta

1. How many unique sequences?

grep “^>” AB\_ test.dta.fasta | sort | uniq | wc -1

1. Which sequence comes first alphabetically?

grep “^>” AB\_test\_dta.fasta | sort | head -1

1. How many sequences contain the motif “NNNNNNNNNN”?

Grep –c NNNNNNNNNN AB\_test\_dta.fasta (If you exclude –c then you will get all the sequences).

1. Which sequences are they?

grep B1 NNNNNNNNNN AB\_test\_dta.fasta | grep “^>”