Homework day 4 – FastQC, Trimmomatic, Bowtie2 and IGV.

Authors: Mary Allen & Daniel Ramirez

FastQC & Trimmomatic

1. Edit the template.sbatch to run fastQC on the files:
Day4HW_R1.fastq
Day4HW_R2.fastq

found in the directory:
/scratch/Shares/public/sread2018/data_files/fastq_for_quality_check/

2. Edit the template.sbatch to run Trimmomatic on the files:
Day4HW_R1.fastq
Day4HW_R2.fastq

Save the output files as:
Day4HW_R1_paired_out.fastq
Day4HW_R1_unpaired_out.fastq
Day4HW_R2_paired_out.fastq
Day4HW_R2_unpaired_out.fastq

These are paired-end read files, with phred 33 quality scores and the library was made using Nextera adapters, whose fasta file is on /opt/trimmomatic/0.36/adapters

3. Run FastQC again on the output files:
Day4HW_R1_paired_out.fastq
Day4HW_R2_paired_out.fastq

Do these files have their adapters correctly trimmed?

Bowtie2 & IGV

1. Edit the template.sbatch to run bowtie2 (and samtools to create sorted bam files with their respective indexes) on one of the following files:
chr21EliGABPAchip.fastq chr21ElizabethGABPAchip.fastq
chr21EricGABPAchip.fastq chr21EthanGABPAchip.fastq

found in the directory: /scratch/Shares/public/sread2018/data_files/ChIP-seq/

2. Visualize the resulting sorted.bam files using IGV.
What happens to the mapping efficiency if you play with bowtie2 preset options? (e.g. changing --very-sensitive to –very-fast)