**Brenda I. Cesar Assignment 1**

Question 1

1. samtools faidx ref.fa This created a file named ref.fa.fai

cat ref.fa.fai

Total length of reference sequence (in bases) = 233806

1. fastqc frag180.1.fq

There is a total of 35178 sequences divided in sequences of 100 bp, making it have ~351 reads

fastqc frag180.2.fq

There is a total of 35178 sequences divided in sequences of 100 bp, making it have ~351 reads

fastqc jump2k.1.fq

There is a total of 70355 sequences divided in sequences of 50 bp, making it have ~1,407 reads

fastqc jump2k.2.fq

There is a total of 70355 sequences divided in sequences of 50 bp, making it have ~1,407 reads

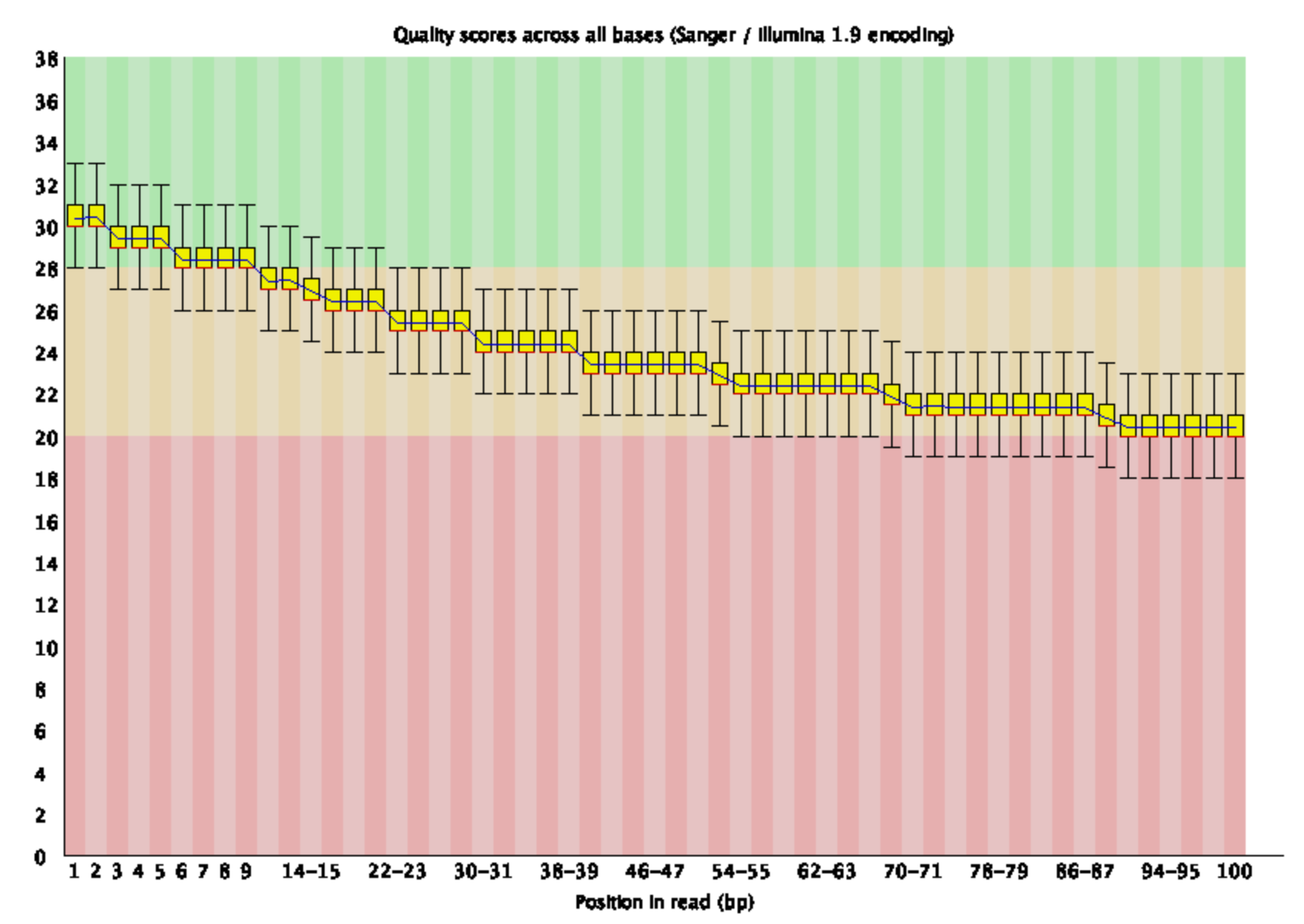
1. fastqc frag180.1.fq coverage = 35178 \* (2\*100) / 233806 = 30.09

fastqc frag180.2.fq coverage = 35178 \* (2\*100) / 233806 = 30.09

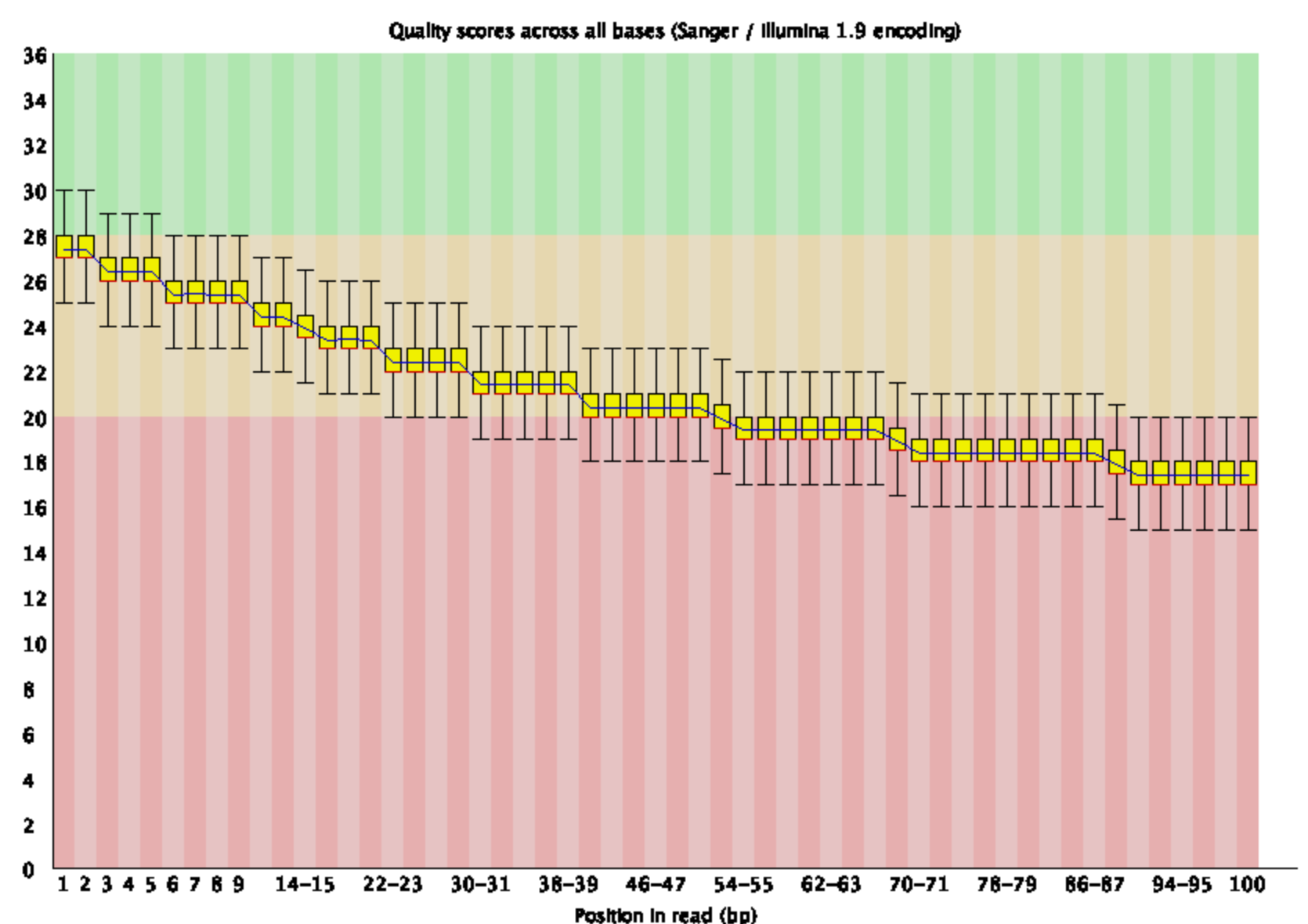
fastqc jump2k.1.fq coverage = 70355 \* (2\*50) / 233806 = 30.09

fastqc jump2k.2.fq coverage = 70355 \* (2\*50) / 233806 = 30.09

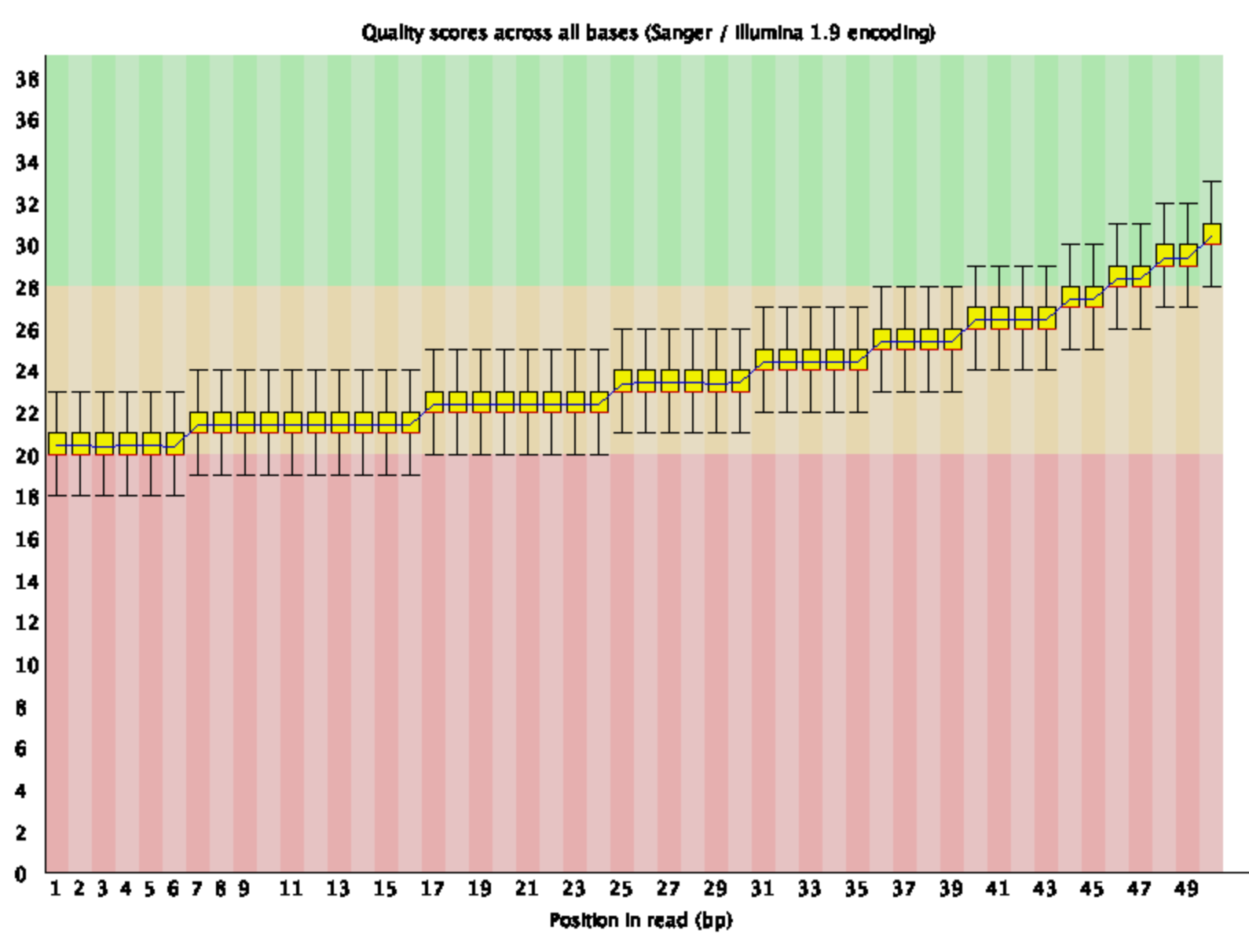
1. frag180.1



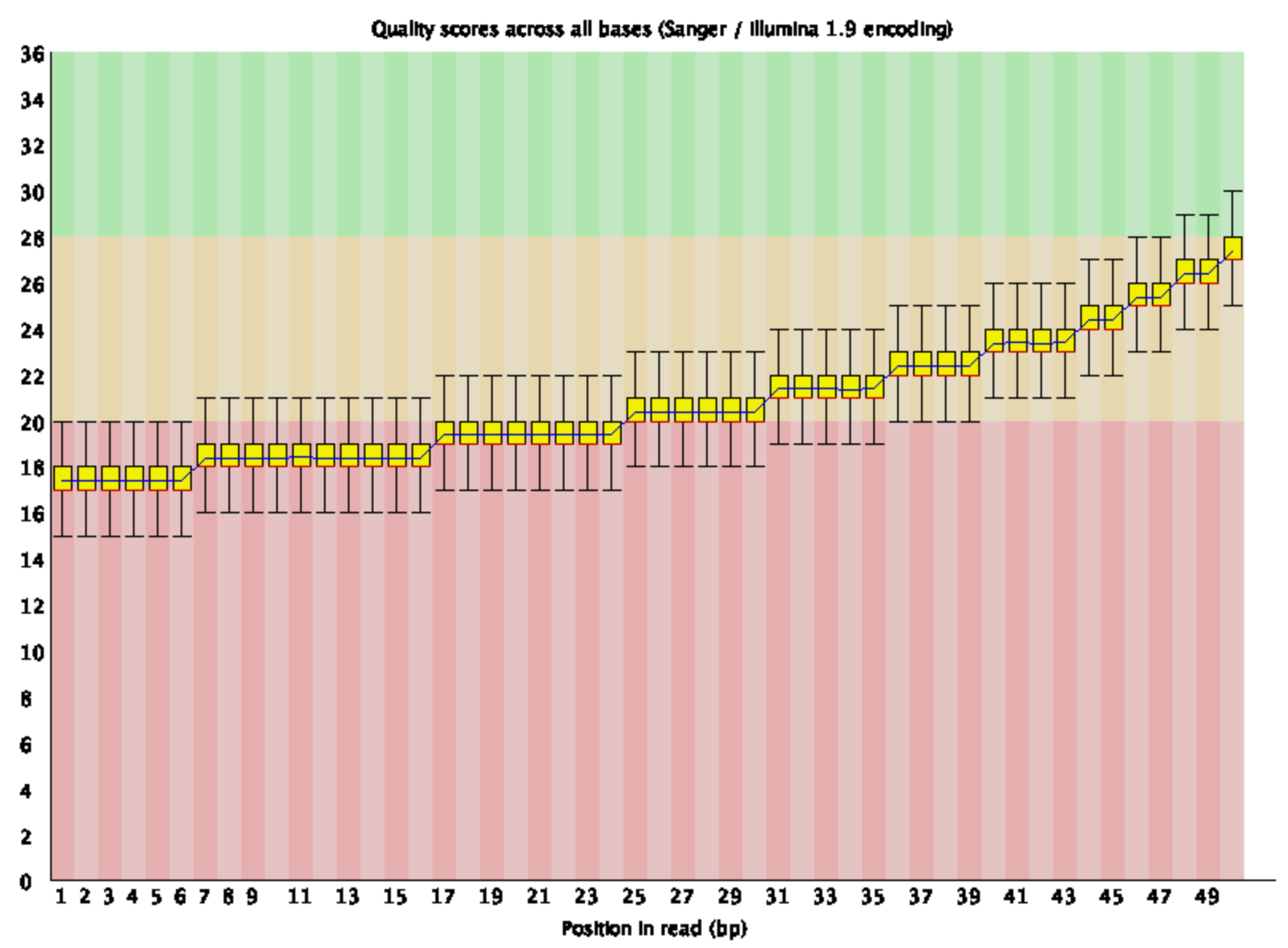
frag180.2



jump2k.1



jump2k.2



Question 2

1. jellyfish count -m 21 -C -s 1000000 frag180.1.fq

mv mer\_counts.jf mer\_countsfrag1.jf

jellyfish histo mer\_countsfrag1.jf > readsfrag1.histo

jellyfish histo mer\_countsfrag1.jf

There are 0 21-mers that occur 50 times in frag180.1.fq

jellyfish count -m 21 -C -s 1000000 frag180.2.fq

mv mer\_counts.jf mer\_countsfrag2.jf

jellyfish histo mer\_countsfrag2.jf > readsfrag2.histo

jellyfish histo mer\_countsfrag2.jf

There are 0 21-mers that occur 50 times in frag180.2.fq

jellyfish count -m 21 -C -s 1000000 jump2K.1.fq

mv mer\_counts.jf mer\_countsjump1.jf

jellyfish histo mer\_countsjump1.jf > readsjump1.histo

jellyfish histo mer\_countsjump1.jf

There are 0 21-mers that occur 50 times in frag180.1.fq

jellyfish count -m 21 -C -s 1000000 jump2K.2.fq

mv mer\_counts.jf mer\_countsjump2.jf

jellyfish histo mer\_countsjump2.jf > readsjump2.histo

jellyfish histo mer\_countsjump2.jf

There are 0 21-mers that occur 50 times in frag180.2.fq

Question 3

1. spades.py --pe1-1 frag180.1.fq --pe1-2 frag180.2.fq --mp1-1 jump2k.1.fq --mp1-2 jump2k.2.fq -o asm -t 4 -k 31

Grep -c ‘>’ asm/contigs.fasta

4 contigs were produced

1. samtools faidx asm/contigs.fasta

head -n 100 asm/contigs.fasta.fai

output: NODE\_1\_length\_105831\_cov\_20.671371 105831 36 60 61  
NODE\_2\_length\_47861\_cov\_20.231319 47861 107666 60 61  
NODE\_3\_length\_41352\_cov\_20.588756 41352 156360 60 61  
NODE\_4\_length\_39423\_cov\_20.384723 39423 198437 60 61

1. NODE\_1\_length\_105831\_cov\_20.671371 105831 36 60 61