Pacbio Reads

Locations:
Sample location:
Pacbio location: /storage/datasets/Tetrahymena_thermophila/SRX2635099/fastq

Tlr reference:
/storage/datasets/Tetrahymena_thermophila/analyses/virus/data/Tlr_complete.fasta

Creating pacbio_reads.sam
cat /storage/datasets/Tetrahymena_thermophila/SRX2635099/fastq/*.fastq > pacbio_reads.sam

Aligning the pacbio reads with minimap2 onto the Tlr_complete:
/home/jgjohns6/minimap2-2.12_x64-linux/minimap2 -a
/storage/datasets/Tetrahymena_thermophila/analyses/virus/data/Tlr_complete.fasta
/home/jgjohns6/Pacbio_reads.fastq > pacbio_minimap.sam

Creating sorted bam file from sam file:
samtools view -bS pacbio_minimap.sam | samtools sort -> pacbio_minimap_sorted.bam

Create index for bam file in order to filter:
samtools index pacbio_minimap_sorted.bam pacbio_minimap_ultra_sorted.bam.bai

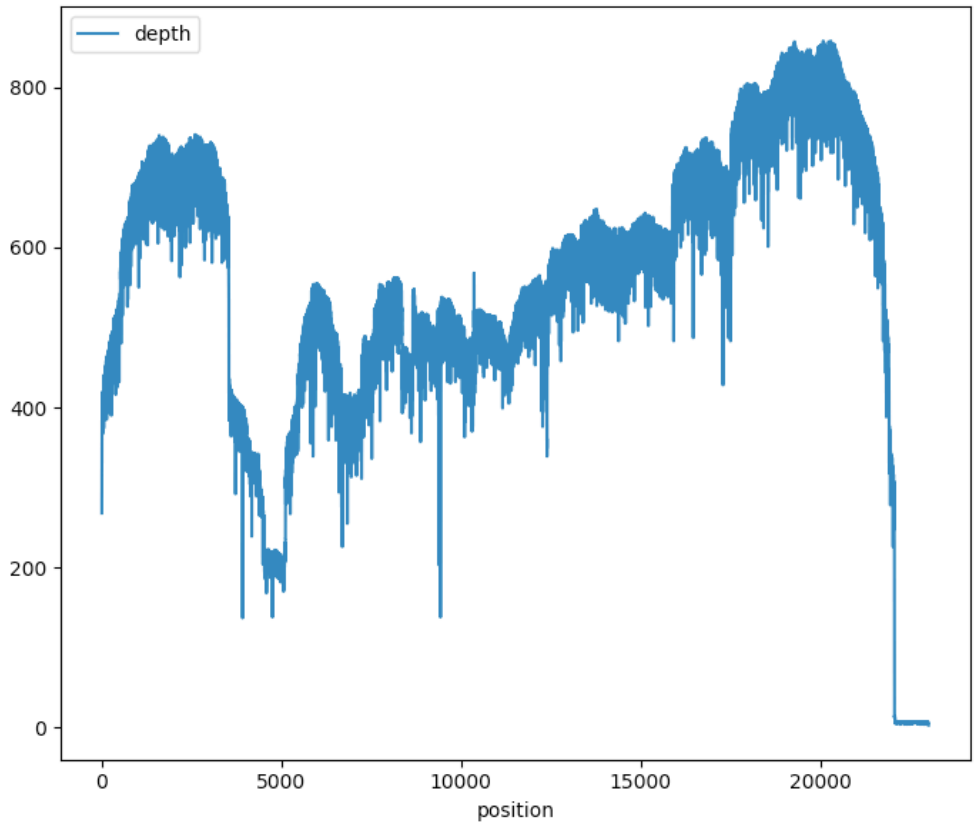
Filter by AF45 chromosome:
samtools view -h -b pacbio_minimap_ultra_sorted.bam AF451863.1 > pacbio_mm_uAF45.bam

Filter by sam flag 3852 and mapping quality 3:
samtools view -h -F3852 -q10 pacbio_minimap_sorted.bam > Pacbio_mm_uAF45_3852_3.bam

Creating coverage file from bam file:
samtools depth pacbio_mm_uAF45_3852_3.bam > pacbio_mm_uAF45_3852_3.coverage

Creating coverage histogram:

```
import pandas as pd
import matplotlib.pyplot as plt
import scipy
import numpy as np
 df=pd.read_csv("pacbio_ mm_uAF45_3852_3.coverage",header=None, sep="\t", names=['contig',
'position', 'depth'])
#print(df)
 df.plot(x='position', y='depth')
plt.show()
```


depth
800
600
400
200
0
0
5000
10000
15000
20000
position