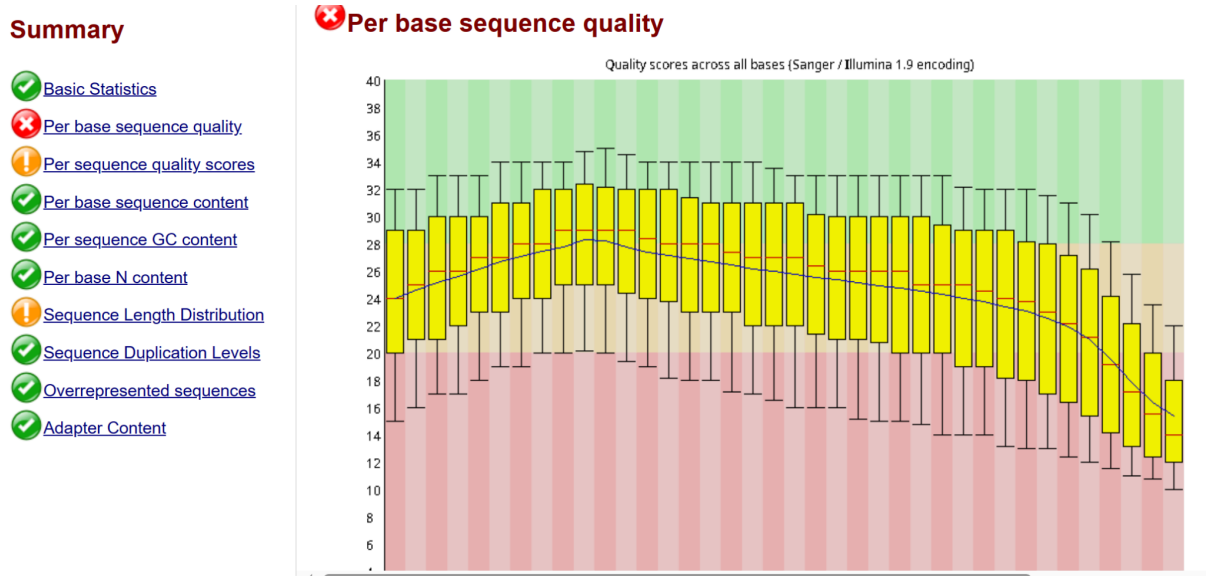
Результаты fastqc:



Команда для тримминга:

trimmomatic PE ILLUMINACLIP:./Sequencing_adaptors.fasta:2:30:10
SLIDINGWINDOW:4:17 LEADING:20 TRAILING:20

```
1314447 reads; of these:
  1314447 (100.00%) were paired; of these:
    2560 (0.19%) aligned concordantly 0 times
    1217284 (92.61%) aligned concordantly exactly 1 time
    94603 (7.20%) aligned concordantly >1 times
    ----
    2560 pairs aligned concordantly 0 times; of these:
      1055 (41.21%) aligned discordantly 1 time
    ----
    1505 pairs aligned 0 times concordantly or discordantly; of these:
      3010 mates make up the pairs; of these:
        1902 (63.19%) aligned 0 times
        155 (5.15%) aligned exactly 1 time
        953 (31.66%) aligned >1 times
99.93% overall alignment rate
```

Общий скрипт:

```
#!/bin/bash

echo 'Write the path to bowtie2 index: $1'
read BOWTIE2_INDEX
echo "write the path to your fastq file 1"
read FASTQ_FILE_1
echo "write the path to your fastq file 2"
read FASTQ_FILE_2
echo "write number of threads for bowtie2"
read THREADS
echo "write path to output_dir "
read OUTPUT_DIR
echo "write the name for output SORTED BAM file (without .bam at the end, just name)"
read OUTPUT_FILE_NAME

echo "write the path to your ref sequence fasta"
read REFERENCE_FASTA

bowtie2 --threads $THREADS -x $BOWTIE2_INDEX --sensitive -1 $FASTQ_FILE_1 -2
$FASTQ_FILE_2 |
samtools view -@ $THREADS -S -b  |


samtools sort -@ $THREADS > "${OUTPUT_DIR}/${OUTPUT_FILE_NAME}_sorted.bam"


samtools index -@ $THREADS "${OUTPUT_DIR}/${OUTPUT_FILE_NAME}_sorted.bam"




mkdir "${OUTPUT_DIR}/bam_file_stats"

samtools stats -@ $THREADS "${OUTPUT_DIR}/${OUTPUT_FILE_NAME}_sorted.bam"  >
"${OUTPUT_DIR}/bam_file_stats/${OUTPUT_FILE_NAME}_stats_file.stats"

plot-bamstats -p
"${OUTPUT_DIR}/bam_file_stats/${OUTPUT_FILE_NAME}_stats_file_graph"
"${OUTPUT_DIR}/bam_file_stats/${OUTPUT_FILE_NAME}_stats_file"

samtools view -@ $THREADS -b -F 4
"${OUTPUT_DIR}/${OUTPUT_FILE_NAME}_sorted.bam" |
```

samtools sort -@ $THREADS >
"${OUTPUT_DIR}/${OUTPUT_FILE_NAME}_sorted_F4_flag.bam"

#надо установить gnuplot
samtools coverage "${OUTPUT_DIR}/${OUTPUT_FILE_NAME}_sorted.bam" >
"${OUTPUT_DIR}/bam_file_stats/${OUTPUT_FILE_NAME}_coverage_file.txt"

Samtools stats output:

SN raw total sequences: 2626992 # excluding supplementary and secondary reads

SN filtered sequences: 0

SN sequences: 2626992

SN is sorted: 1

SN 1st fragments: 1313491

SN last fragments: 1313501

SN reads mapped: 2626992

SN reads mapped and paired: 2626064 # paired-end technology bit set + both mates mapped

SN reads unmapped: 0

SN reads properly paired: 2623774 # proper-pair bit set

SN reads paired: 2626992 # paired-end technology bit set

SN reads duplicated: 0 # PCR or optical duplicate bit set

SN reads MQ0: 425 # mapped and MQ=0

(mapped&paired-MQ0)/total reads = 0,99

Coverage file:

```
#rname  startpos    endpos  numreads    covbases    coverage    meandepth   meanbaseq   meanmapq
NC_000964.3 1   4215606 2626992 4215600 99.9999 47.8217 27.7    40.3
```