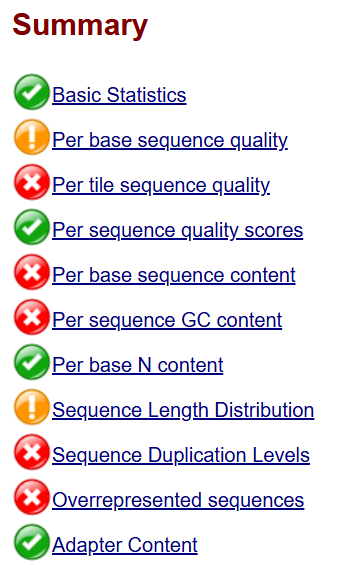
# Отчёт. Домашнее задание 11

## 1)Анализ качества до обработки

Был проведён анализ качества исходных ридов с помощью FastQC:  
  
fastqc WGSB\_46\_01\_PTV291209\_325278\_2\_1.fq.gz  
fastqc WGSB\_46\_01\_PTV291209\_325278\_2\_2.fq.gz  
  
По отчётам были выявлены следующие проблемы:  
- Снижение качества нуклеотидов к 3'-концу ридов  
- Незначительное смещение GC/AT в начале  
  


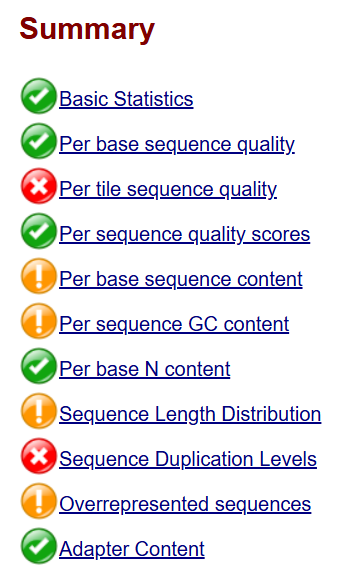
Эти проблемы потенциально влияют на результат сборки, поэтому был выполнен тримминг.

## 2)Тримминг адаптеров и фильтрация

Для очистки данных использовалась программа Trimmomatic.  
  
**Command**:

trimmomatic PE \  
WGSB\_46\_01\_PTV291209\_325278\_2\_1.fq.gz WGSB\_46\_01\_PTV291209\_325278\_2\_2.fq.gz \  
new\_R1.fastq.gz new\_R1\_unpaired.fastq.gz \  
new\_R2.fastq.gz new\_R2\_unpaired.fastq.gz \  
ILLUMINACLIP:adapters.fa:2:30:10 \  
LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36

Повторный анализ FastQC показал, что качество ридов значительно улучшилось.



## 3)Сборка с помощью SPAdes

Сборка проводилась с использованием SPAdes, подходящего для вирусных и бактериальных геномов.

**Command**::

**spades.py \  
-1 new\_R1.fastq.gz \  
-2 new\_R2.fastq.gz \  
--careful \ # снижает количество ошибок  
-k 80 \ #** несколько графов для разных длин k, чтобы объединять **разные уровни точности и длины  
-o spades\_output**

## 4)Оценка сборки через QUAST

Для анализа качества сборки использовалась программа quast:  
  
quast.py -o quast\_output spades\_output/scaffolds.fasta

Анализ включал в первую очередь - N50  
  
  
Дополнительно - Сборка после тримминга улучшилась: уменьшилось количество контигов, повысилось значение N50.  
Таким образом, применённые параметры обеспечили более качественную и точную сборку по сравнению с результатами, полученными на семинаре.