

## Задание по сборке de novo (15 практикум)

Мой код доступа: SRR4240379

Команда для скачивания:

```
wget ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR424/009/SRR4240379/SRR4240379.fastq.gz
```

### 1. Подготовка чтений программой trimmomatic

-Я сделала папку adapters.fasta , куда скопировала содержимое файлов папки /NGS/adapters.

-И удаляю адаптеры командой:

```
TrimmomaticSE -phred33 ../reads/SRR4240379.fastq.gz output.fq.gz  
ILLUMINACLIP:adapters.fasta:2:7:7
```

Результат : Input Reads: 7400155 Surviving: 7269852 (98.24%) Dropped: 130303 (1.76%)

Следовательно 1.76% чтений являлись адаптерами.

-Теперь удаляем нуклеотиды с качеством ниже 20 и последовательности длины меньше 32

```
команда: TrimmomaticSE -phred33 noad.fq.gz trimmed.fq.gz TRAILING:20  
MINLEN:32
```

результат: Input Reads: 7269852 Surviving: 6974267 (95.93%) Dropped: 295585 (4.07%)

### 2. Подготовка K-меров

Команда : velveth Assem 31 -short -fastq trimmed.fq.gz

### 3. Запуск velvetg

Запускаем velvetg : velvetg ./Assem

N50 of 25646

команда : cut -f6 stats.txt | sort --human-numeric-sort | less -S

результат(самые крупные континги):

5	33085	0	1	0.000000	36.259030	36.259030
9	49262	0	0	0.000000	34.772177	34.772177
6	49912	0	0	0.000000	35.907237	35.907237

Проверим , есть ли континги аномальными покрытиями , еще раз отсортировав  
Результат :

Континги с аномально большими покрытиями:

133	1	3	3	0.000000	474299.000000
105	1	3	3	0.000000	2694.000000

Континги с аномально маленькими покрытиями:

20	26061	0	0	0.000000	35.436821	35.436821
31	11443	0	1	0.000000	32.358560	32.358560

#### 4. Анализ

Сравнение программой megablast каждый из 3-х самых длинных контингов с хромосомой *Buchnera aphidicola* (GenBank/EMBL AC — CP009253)

-Разделила contigs.fa на отдельные fasta-файлы:

seqretsplit -filter contigs.fa dir/name.format

-Перенесла 3 самых длинных континга в новую папку newconting.fa

1)node\_5\_length\_33085

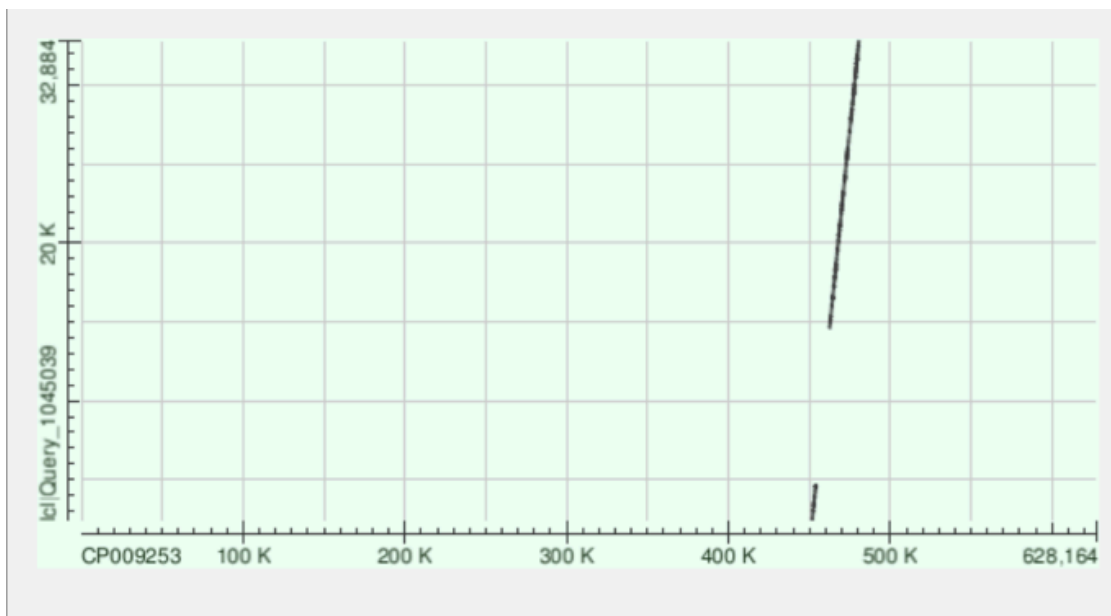


Рис1.Заметим крупную делению

2)node\_6\_length\_49912



Рис2.Заметим несколько делеций

3)NODE\_9\_length\_49262

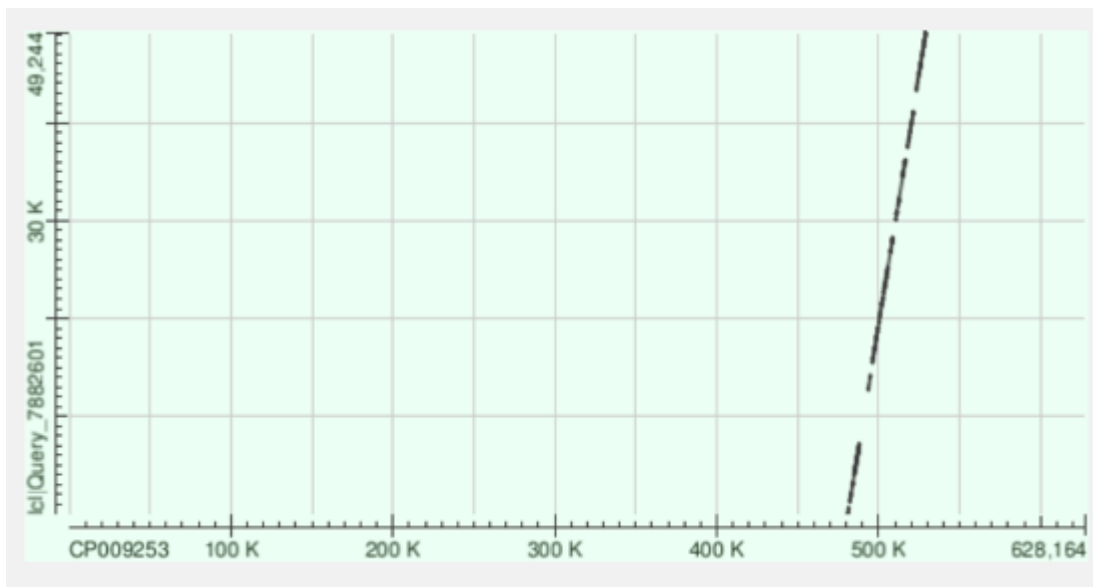


Рис3.Заметим несколько делеций

