

**Оценка потенциала
библиотеки парноконцевых
ридов, полученных на
платформе Illumina, для
полногеномной и локальной
сборки**

*Научный руководитель:
Кулакова Милана Анатольевна, КБН СНС,
СПбГУ, кафедра эмбриологии, лаборатория
экспериментальной эмбриологии*

*Студент:
Зорина Наталья*

Задачи:

- 1.** Анализ качества рядов и репрезентативности библиотеки
- 2.** Оценка потенциала библиотеки для полногеномной сборки
- 3.** Оценка потенциала библиотеки для локальной сборки (Нох- локус)

Характеристики библиотеки:

1. Размер генома *Alitta virens* ~ 650 Мб
2. Источник ДНК - сперматозоиды, принадлежащие одной особи
3. Платформа для секвенирования - Illumina HiSeq 2500
4. Nextera DNA Library Preparation Kit
5. Размер вставки ~ 250 нуклеотидов
6. Длина парноконцевых ридов - 101 нуклеотид

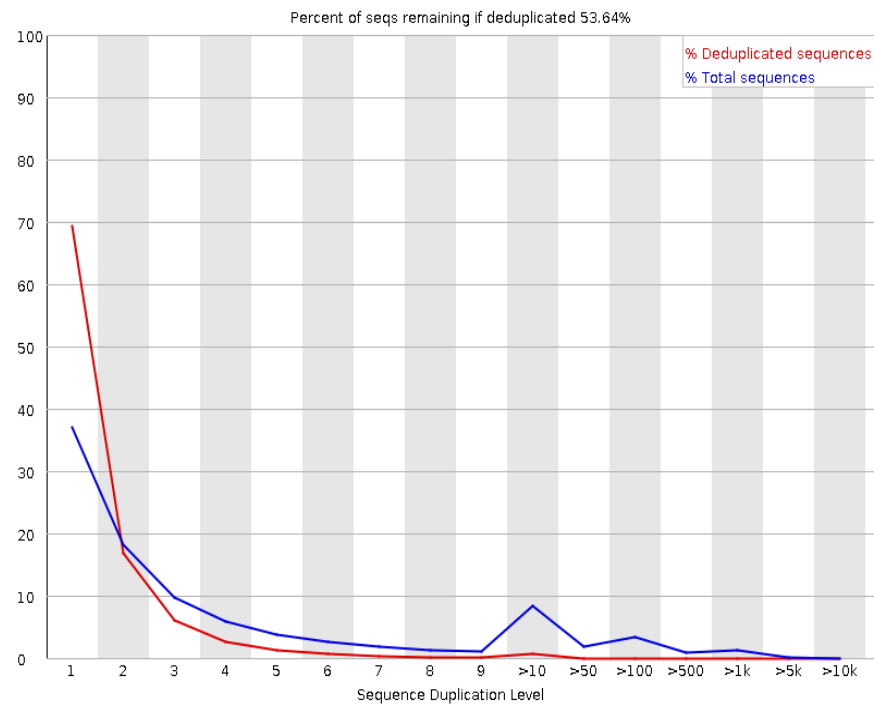
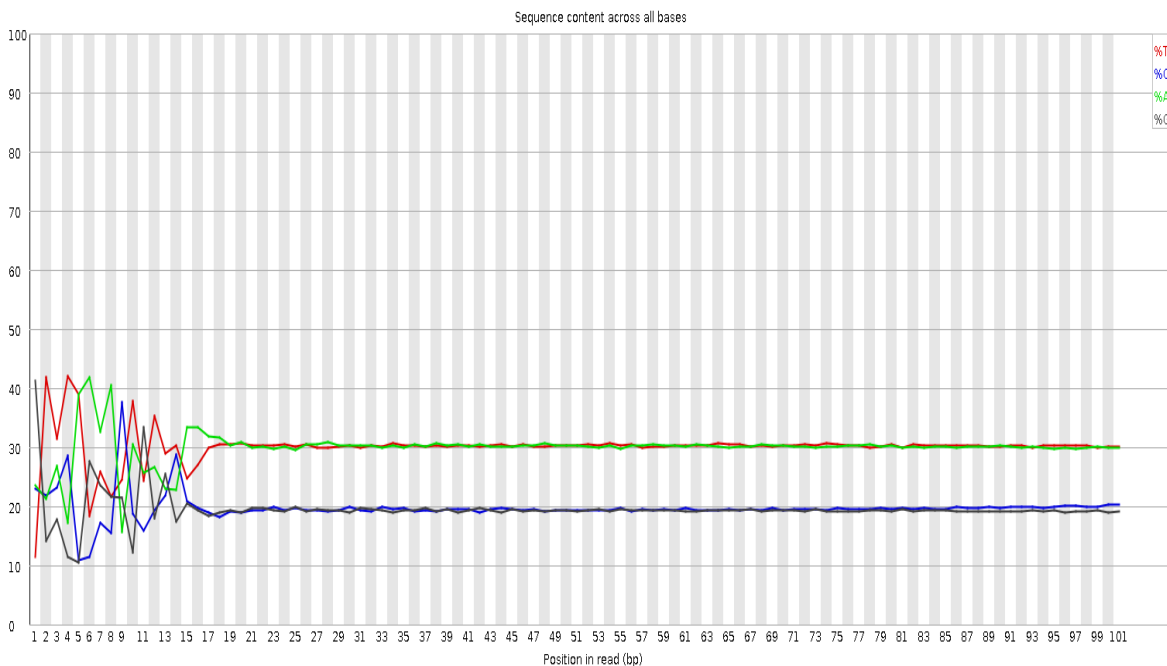
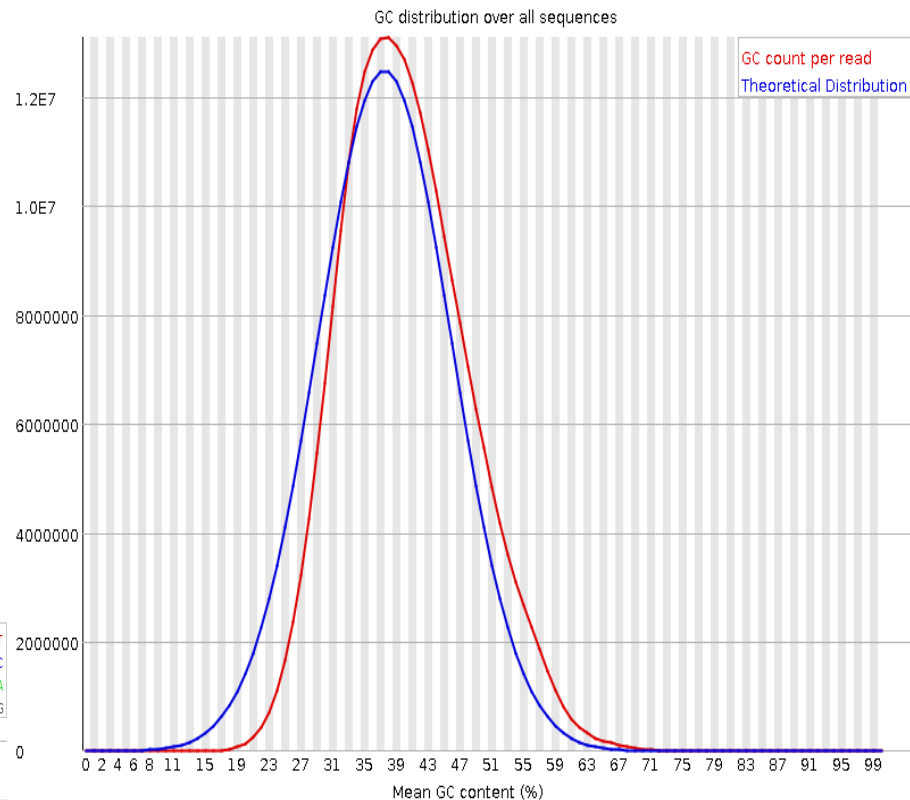
Quality Check

FastQC v0.10.1

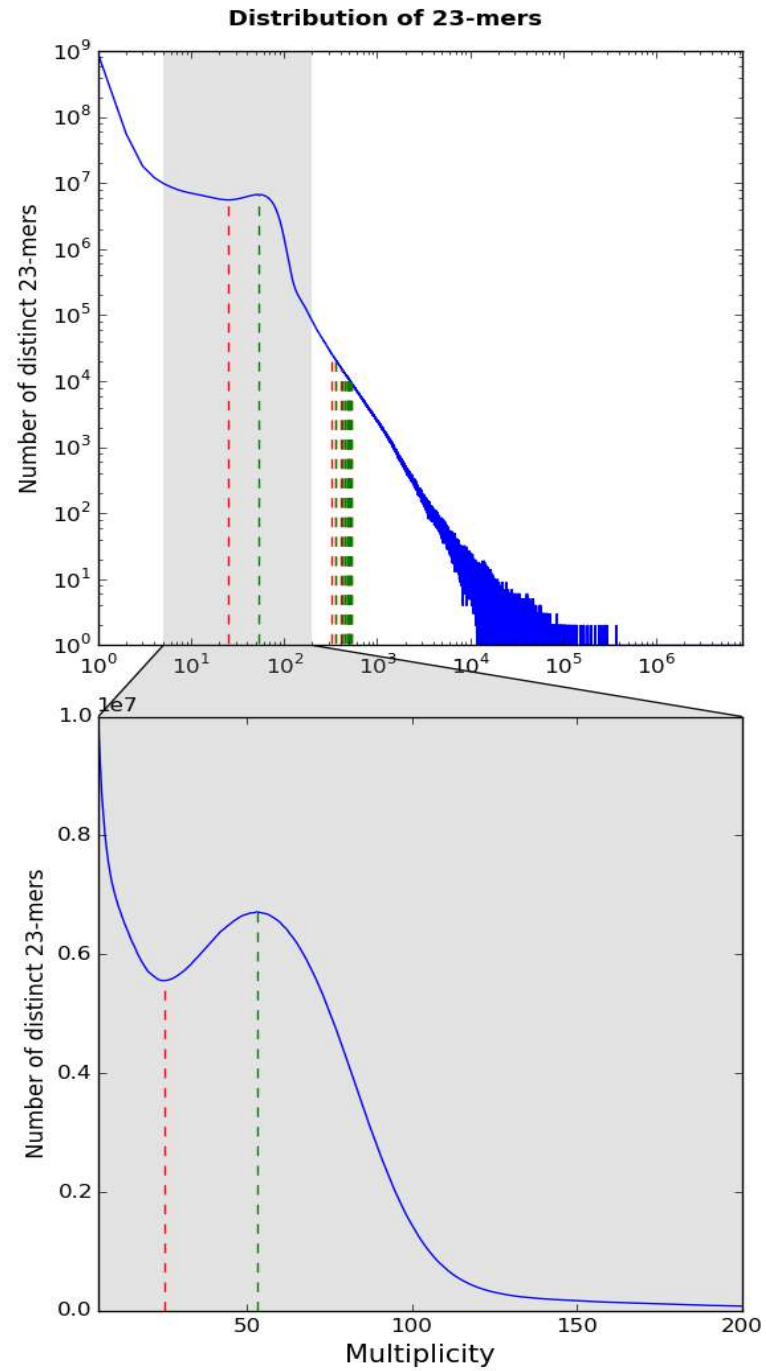
- Per base sequence quality
- Per sequence quality scores
- Per base sequence content
- Per base GC content
- Per sequence GC content
- Per base N content
- Sequence Length Distribution
- Sequence Duplication Levels
- Overrepresented sequences
- Kmer Content

FastQC

HICUP Pipeline for analysing Hi-C data

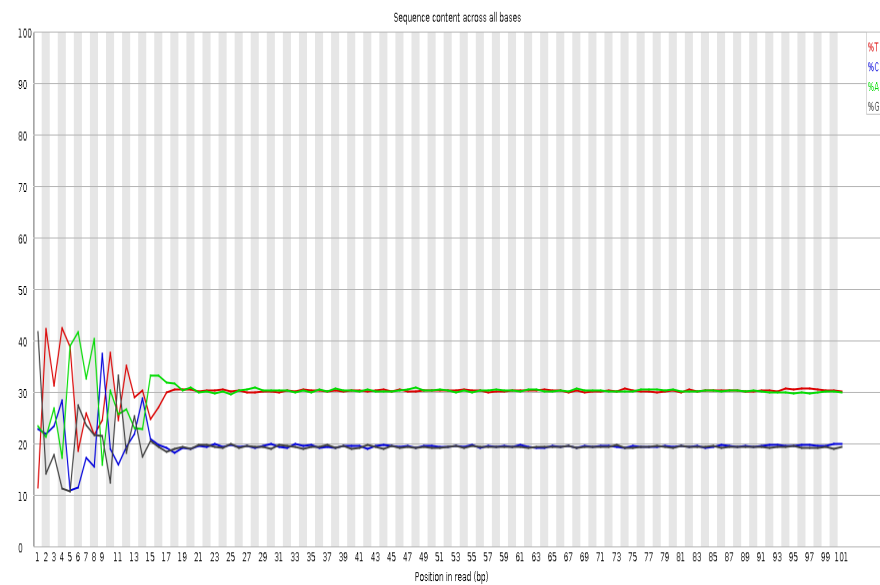
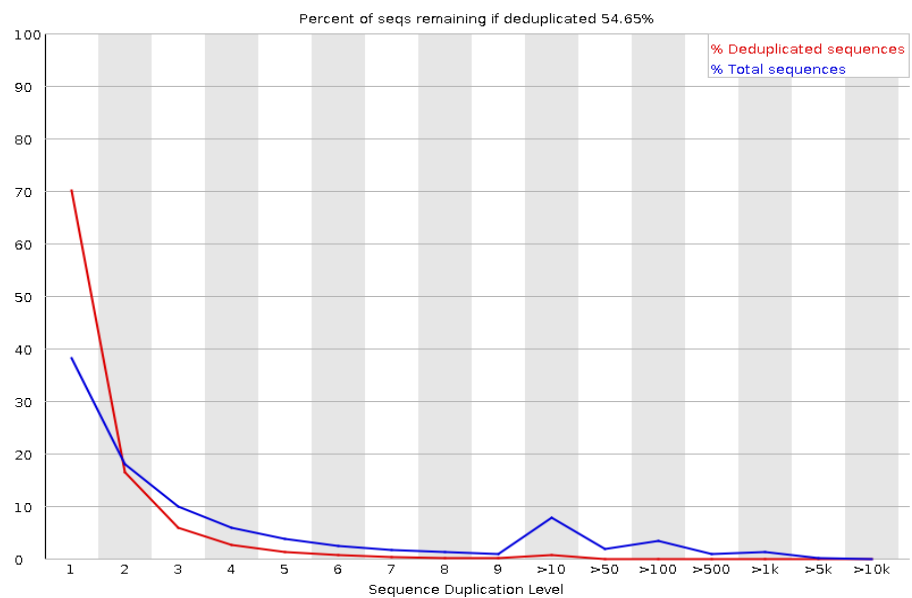
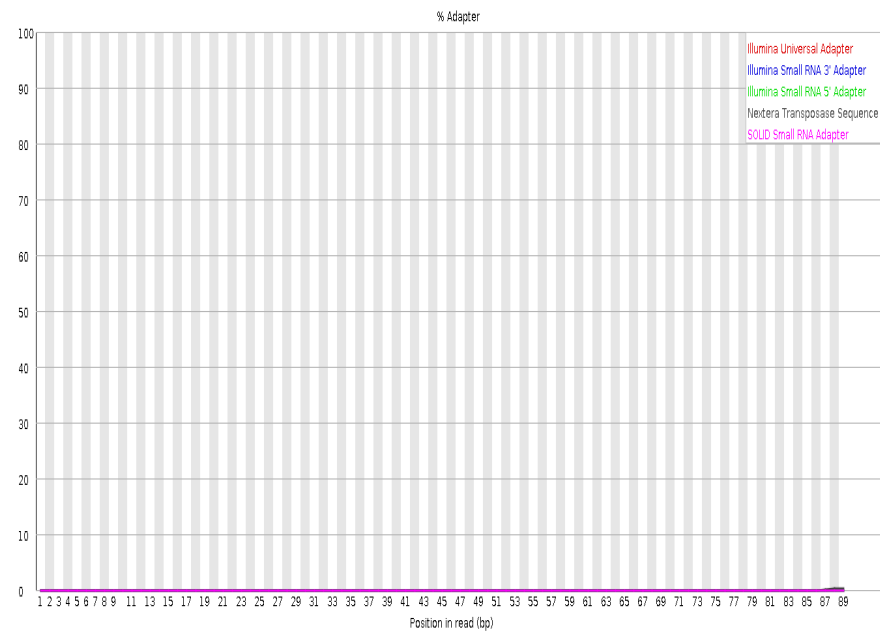
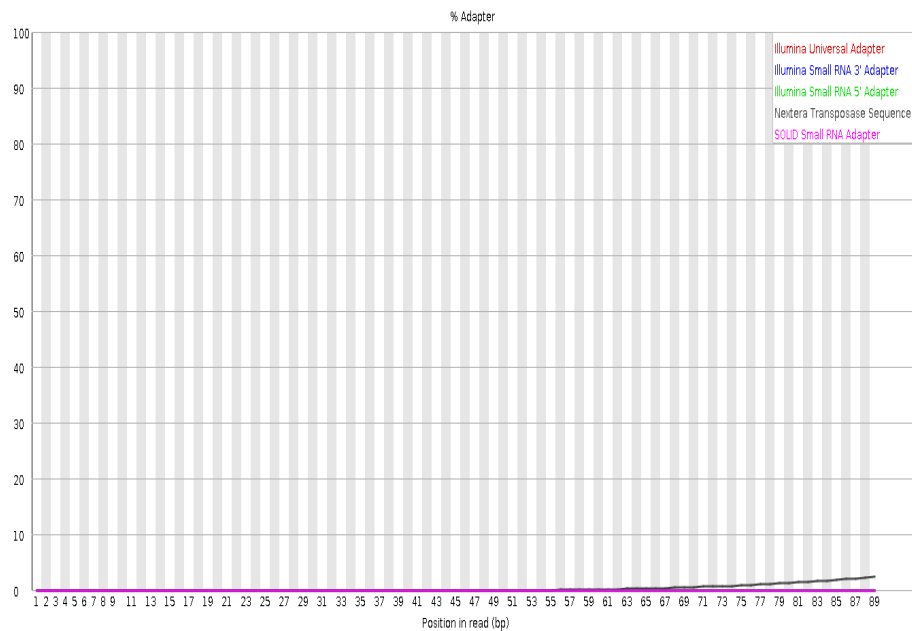


Распределение k -меров (Jellyfish)





Trimmomatic





Quast
MAKING QUALITY A HABIT

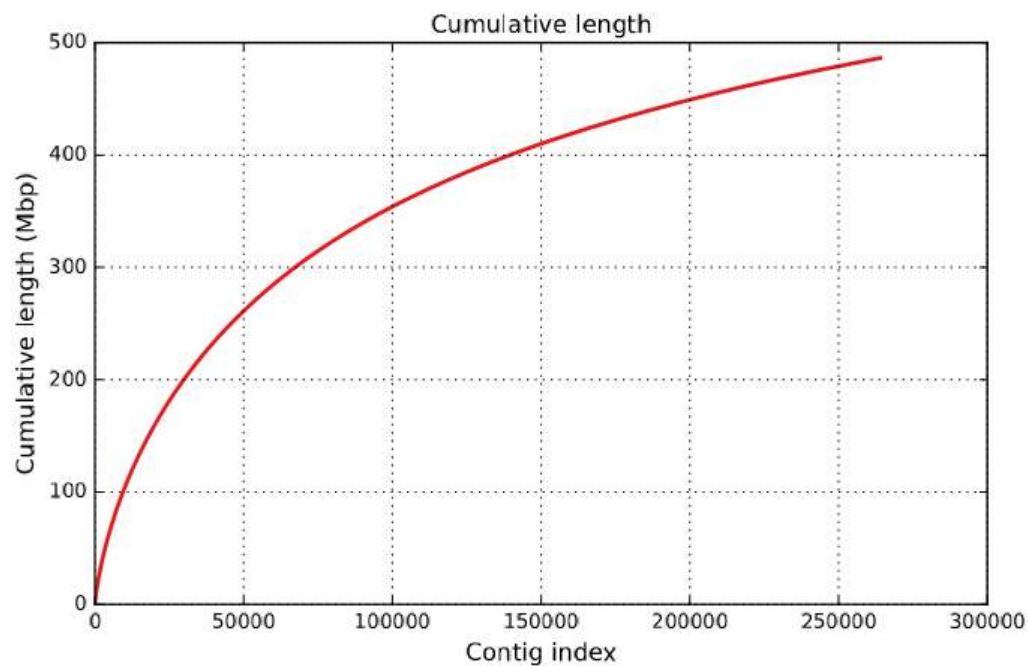
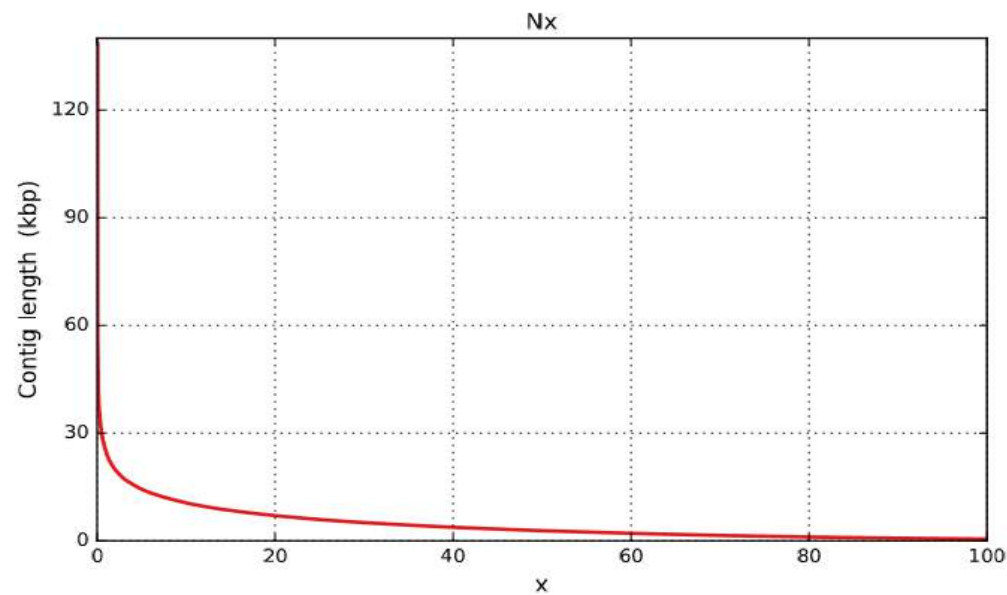
SOAPdenovo

Statistics without reference graph_prefix.scafSeq

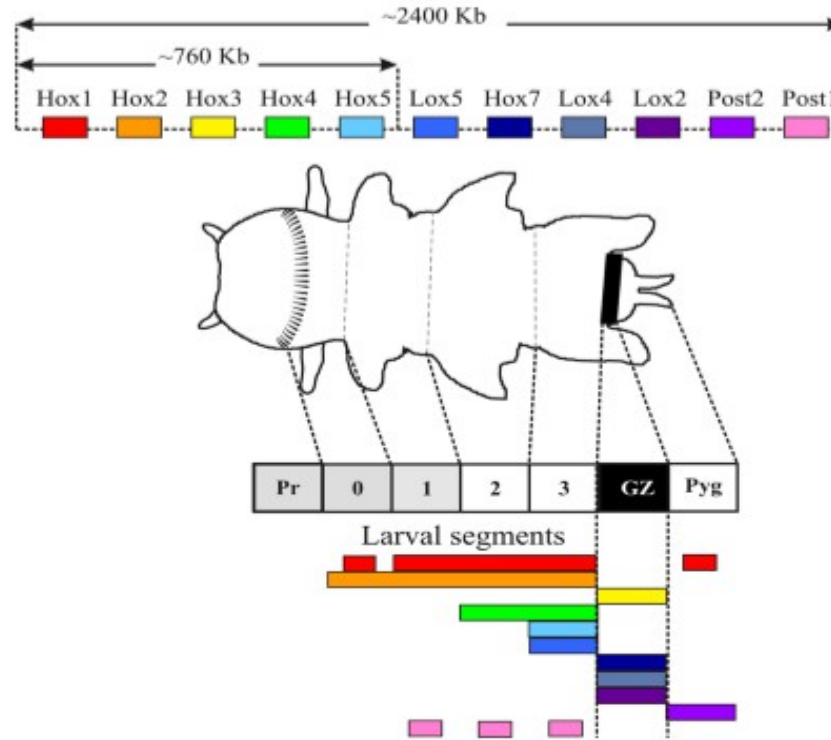
# contigs	264 012
# contigs (>= 0 bp)	2 468 785
# contigs (>= 1000 bp)	137 924
Largest contig	138 840
Total length	486 423 644
Total length (>= 0 bp)	836 563 483
Total length (>= 1000 bp)	398 529 769
N50	2833
N75	1277
L50	43 254
L75	108 038
GC (%)	37.31

Mismatches

# N's	38 741 557
# N's per 100 kbp	7964.57



Локальная сборка Нох - кластера



geneIOUS

Результаты:

1. В библиотеке присутствует большое количество ридов с ошибками, что делает ее не репрезентативной
2. Библиотека не пригодна для качественной полногеномной сборки и дальнейшего аннотирования
3. Библиотека имеет довольно хороший потенциал для локальных сборок

Спасибо за внимание!

