

# Анализ геномов дрожжей из Петергофской генетической коллекции

Антон Матиив

Руководитель:  
Юрий Барбитов  
Институт биоинформатики  
Кафедра генетики и биотехнологии СПбГУ

# Nanopore DNA sequencing



# Задачи:

- 1) “Полировка” собранного генома из длинных ридов (Nanopolish, Racon)
- 2) Функциональная аннотация на основании референсного штамма
- 3) Использование полученного “полированного” генома в качестве референса для поиска мутаций в других штаммах



Термин "полировка"  
применяется в случае  
коррекции при помощи  
других данных - либо  
"сырого сигнала"  
(**Nanopolish**), либо  
коротких прочтений  
(**Racon**).

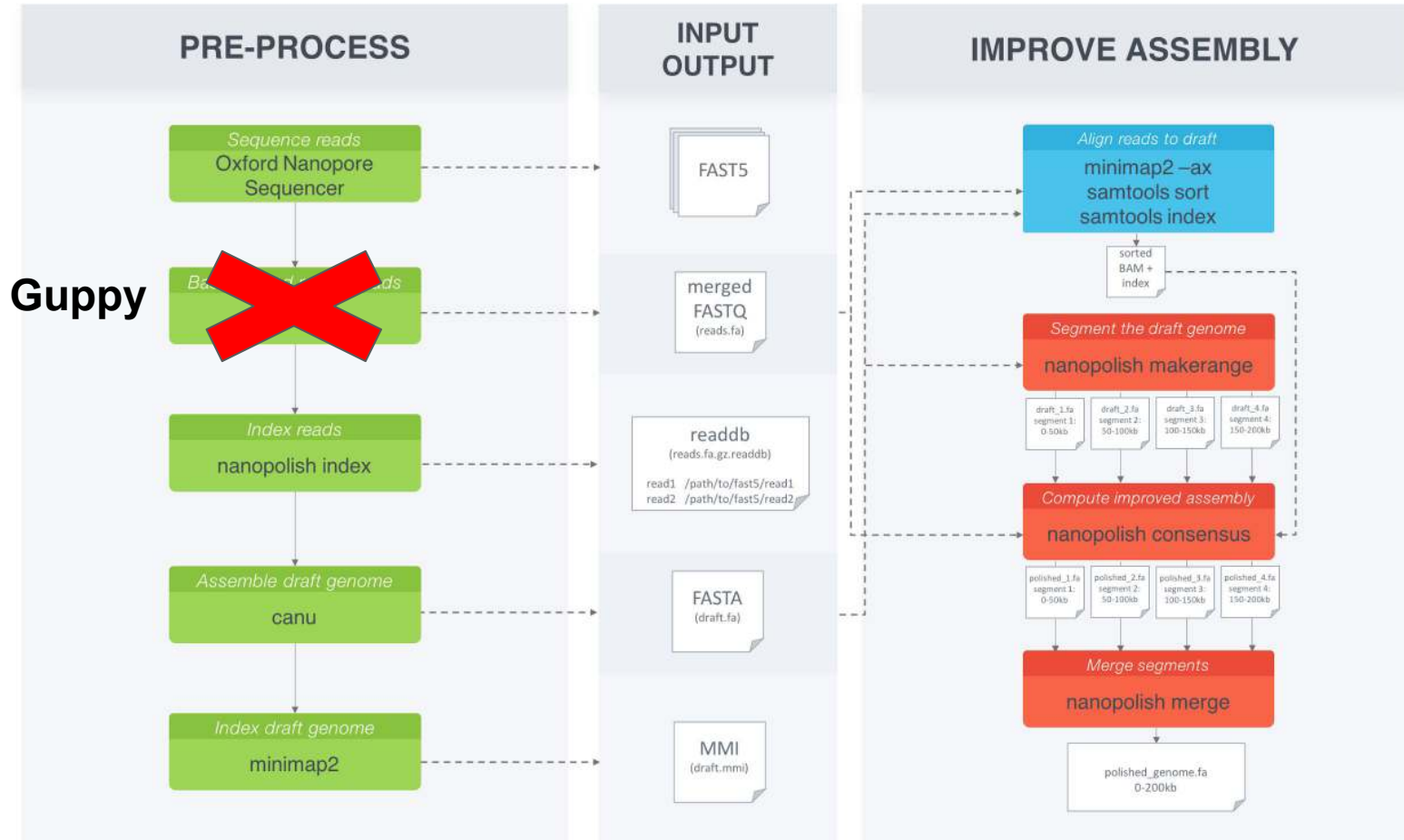


Draft genome assembly



Nanopolish + Racon

# Nanopolish pipeline



# Оценка сборок с помощью QUAST

	<b>yeast_1d.guppy_213. canu (Draft)</b>	<b>polished_genome</b>	<b>polished_genome _racon</b>
<b>Total length</b>	<b>12396504</b>	<b>12445256</b>	<b>12444850</b>
<b>Largest contig</b>	<b>1095012</b>	<b>1099417</b>	<b>1099192</b>
<b>N50</b>	<b>822302</b>	<b>825460</b>	<b>825513</b>
<b>mismatches per 100 kbp</b>	<b>212.37</b>	<b>207.85</b>	<b>205.12</b>
<b>indels per 100 kbp</b>	<b>343.76</b>	<b>73.09</b>	<b>25.04</b>

# Аннотация сборки с помощью Exonerate

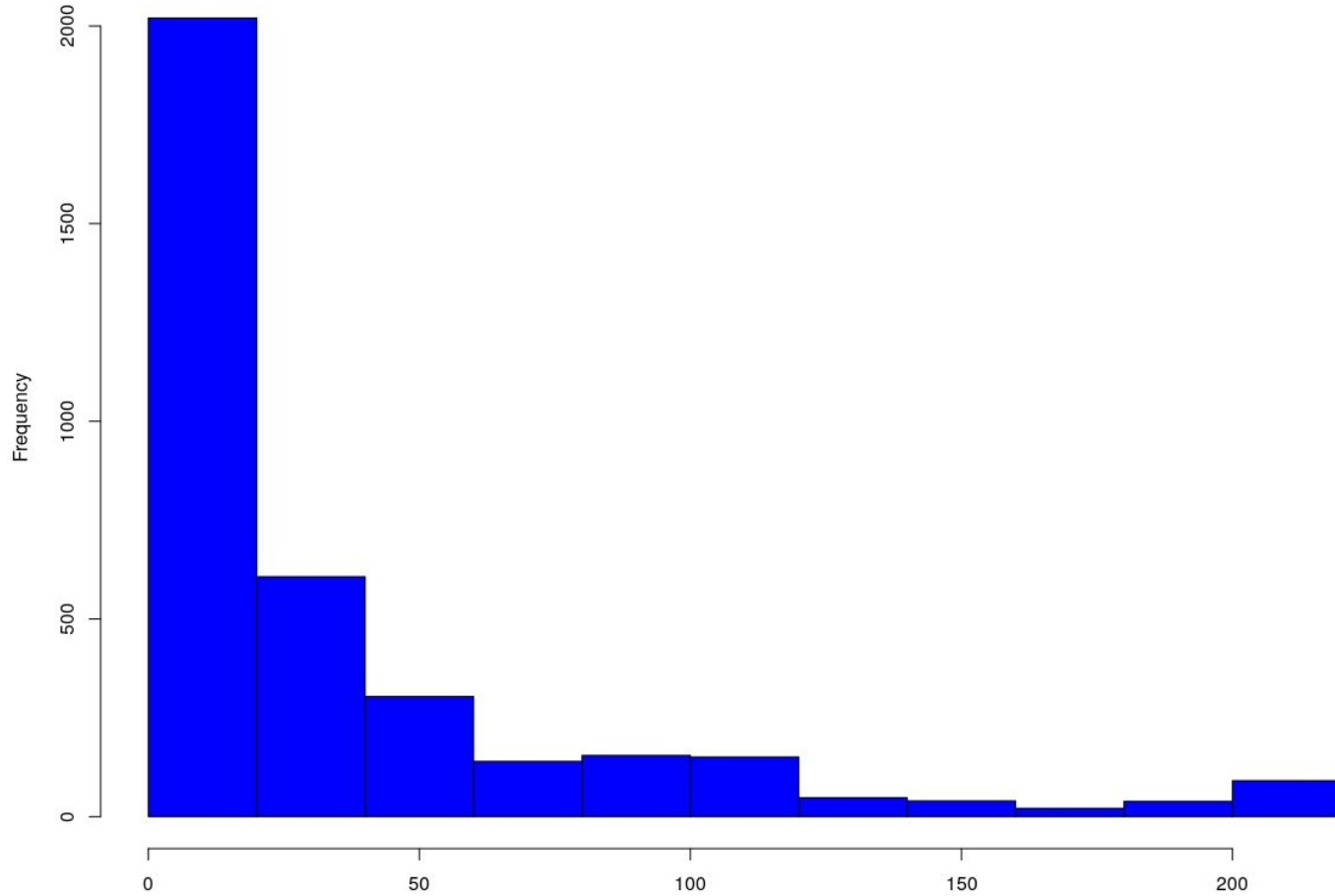
Использовали список аминокислотных последовательностей белков S288С из базы UniProt. Построили базу данных с помощью SnpEff

## **In progress:**

поиск мутаций в образцах с использованием собранного генома в качестве референса



# Распределение числа аллелей





Спасибо за внимание!