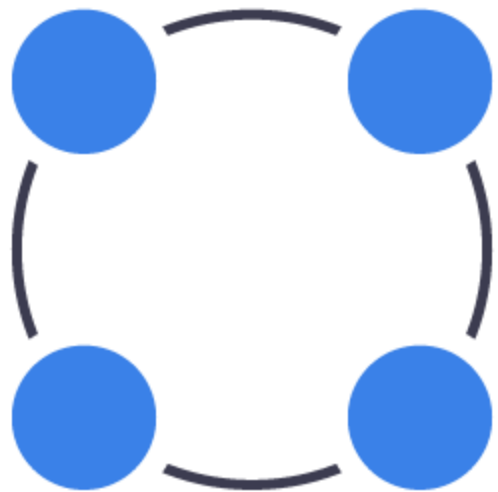


Анализ влияния гипоксии на
дифференциальную экспрессию генов в
клетках периферической крови в
зависимости от физической нагрузки

Выполнил: Маловичко Юрий

Руководитель: Готов А. С.,
ФГБНУ “НИИ АГиР им.Д.О.Отта”,
РЦ «Центр Биобанк» СПбГУ



Идентификация всех типов РНК с использованием технологии секвенирования Illumina и Nanopore

Выполнил: Маловичко Юрий
Руководитель: Готов А. С.,
ФГБНУ «НИИ АГиР им.Д.О.Отта»,
РЦ «Центр Биобанк» СПбГУ

Ждем руководителя месяц



Пайплайна нет



Напарника нет



Данных нет



Руководитель болеет



Хотя нет, он в отпуске

У нанопора сломана ячейка



shiiiet

Что изменилось?



```
[uvmalovichko@110-7198-609 Skaters_RNA_20180301]$ ls
1after_S2_L007_R1_001.fastq.gz      2after_S4_L007_R2_001.fastq.gz
1after_S2_L007_R2_001.fastq.gz      2before_S3_L007_R1_001.fastq.gz
1before_S1_L007_R1_001.fastq.gz      2before_S3_L007_R2_001.fastq.gz
1before_S1_L007_R2_001.fastq.gz      3after_S6_L007_R1_001.fastq.gz
2after_S4_L007_R1_001.fastq.gz      3after_S6_L007_R2_001.fastq.gz
```

Что изменилось? ❌



Что изменилось?



Исходные данные

- Материал – мононуклеарные лейкоциты периферической крови 7 членов Олимпийской конькобежной сборной
- Условия – среднегорье (Фонт-Роме, Франция, 1800 м над у. м.)
- Забор материала до и после тренировки
- Очистка материала от мРНК - GLOBINclear™ Kit, human for globin mRNA depletion
- Подготовка библиотек - TruSeq Stranded mRNA libraries
- Платформа - HiSeq 4000 (2x150 b.p.)

Выбранный пайплайн

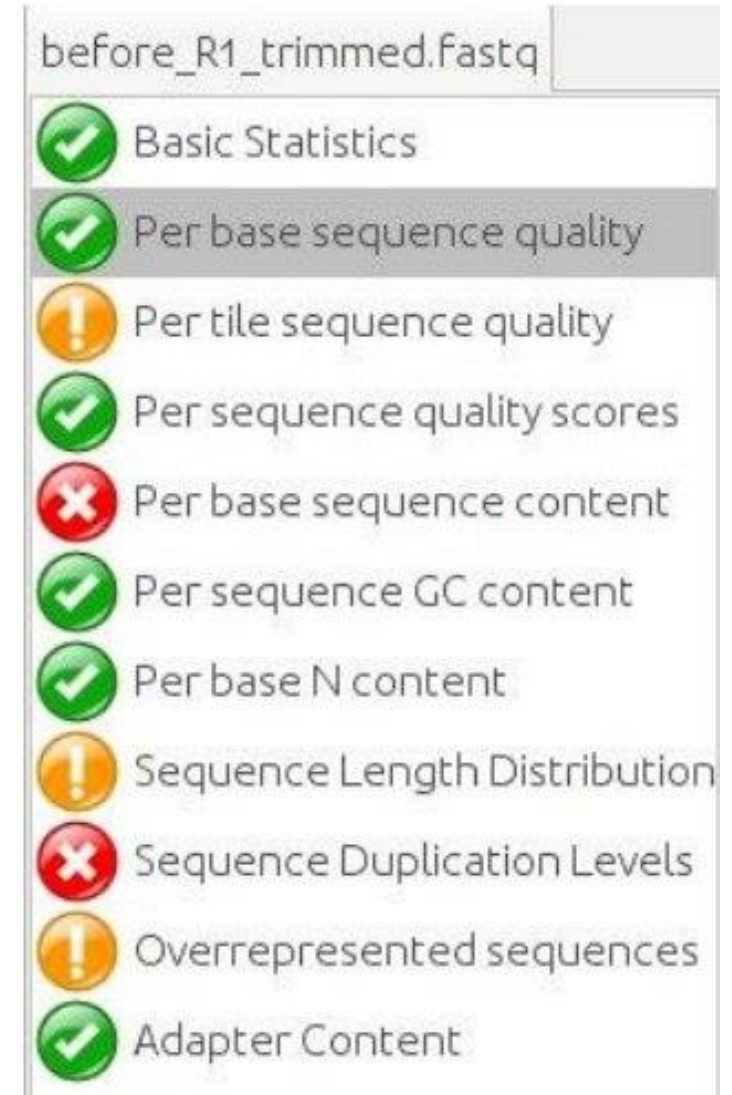
- Контроль качества в FastQC
- Псевдовыравнивание в Kallisto
- Анализ дифференциальной экспрессии при помощи Sleuth
- Визуализация данных в sleuth и Phantasus
- ???
- Красивые картинки
- Красивые выводы

На что не хватило времени/сил/желания

- Выравнивание на GRCh38 при помощи STAR*
- Анализ обогащения при помощи GSEA

Контроль качества ридов в FastQC

- Производилась очистка от последовательностей адаптеров при помощи Trimmomatic – никаких изменений
- Дальнейший тримминг не производился



Псевдовыравнивание и подсчет в kallisto

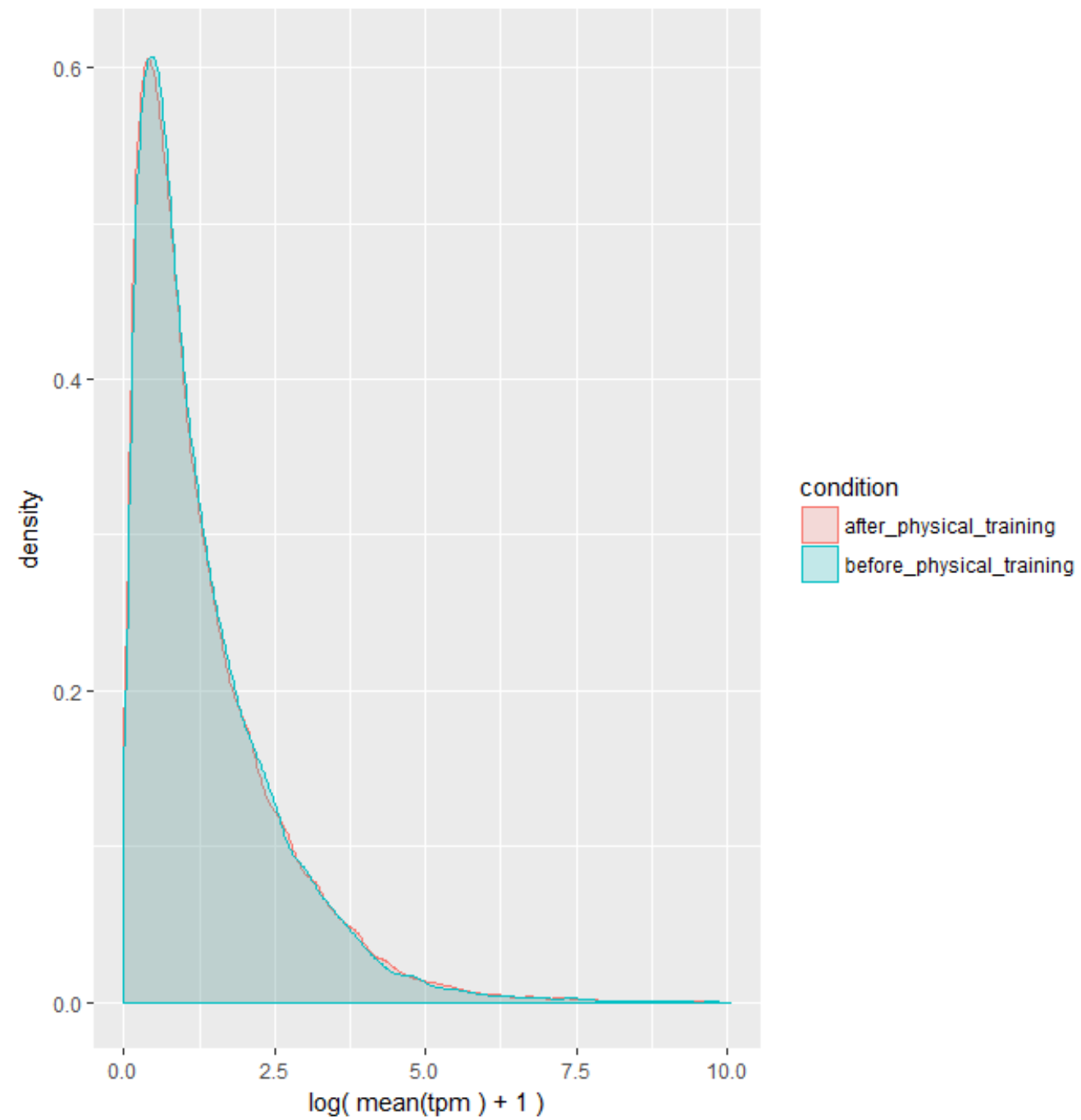
- Версия программы – kallisto-0.44.0
- Референс для kallisto index – hg19
- Kallisto quant выполнялся с 50 репликами бутстрепа
- Сабсеты Top 50

ENST00000547918.2	380	212.699	3.01868	0.685591
ENST00000390361.3	397	228.703	17.1343	3.61916
ENST00000426402.2	544	373.458	36.0179	4.65897
ENST00000390371.3	388	220.109	16.2577	3.56808
ENST00000390367.3	364	197.844	16.9062	4.12795
ENST00000390394.3	413	244.083	0	0
ENST00000426318.2	395	226.772	2.5	0.532554
ENST00000390396.1	347	182.237	0.5	0.13254
ENST00000390397.2	381	213.634	0	0
ENST00000390398.3	381	213.634	0	0
ENST00000620569.1	409	240.278	78.3309	15.7483
ENST00000417977.2	447	277.332	66	11.4963
ENST00000455382.2	421	251.81	121	23.2127
ENST00000390399.3	387	219.22	102.5	22.5869
ENST00000390392.3	455	285.19	16.7406	2.83563
ENST00000390430.2	473	303.073	59	9.40412
ENST00000390429.3	337	173.371	0	0
ENST00000390428.3	404	235.469	35	7.18037
ENST00000390378.1	343	178.81	0.5	0.13508
ENST00000390375.2	384	216.352	29	6.47517
ENST00000390368.2	410	241.181	89.2922	17.8848
ENST00000390454.2	403	234.517	54	11.1233
ENST00000390448.3	406	237.422	58	11.801
ENST00000390435.1	341	176.956	88	24.0232
ENST00000390373.2	344	179.571	0	0
ENST00000390372.3	350	184.918	95	24.8175
ENST00000390400.2	364	197.844	237	57.8679
ENST00000617639.1	406	237.422	34.5	7.01959
ENST00000616518.1	421	251.81	24.5	4.70009
ENST00000620773.1	347	182.237	1	0.265079

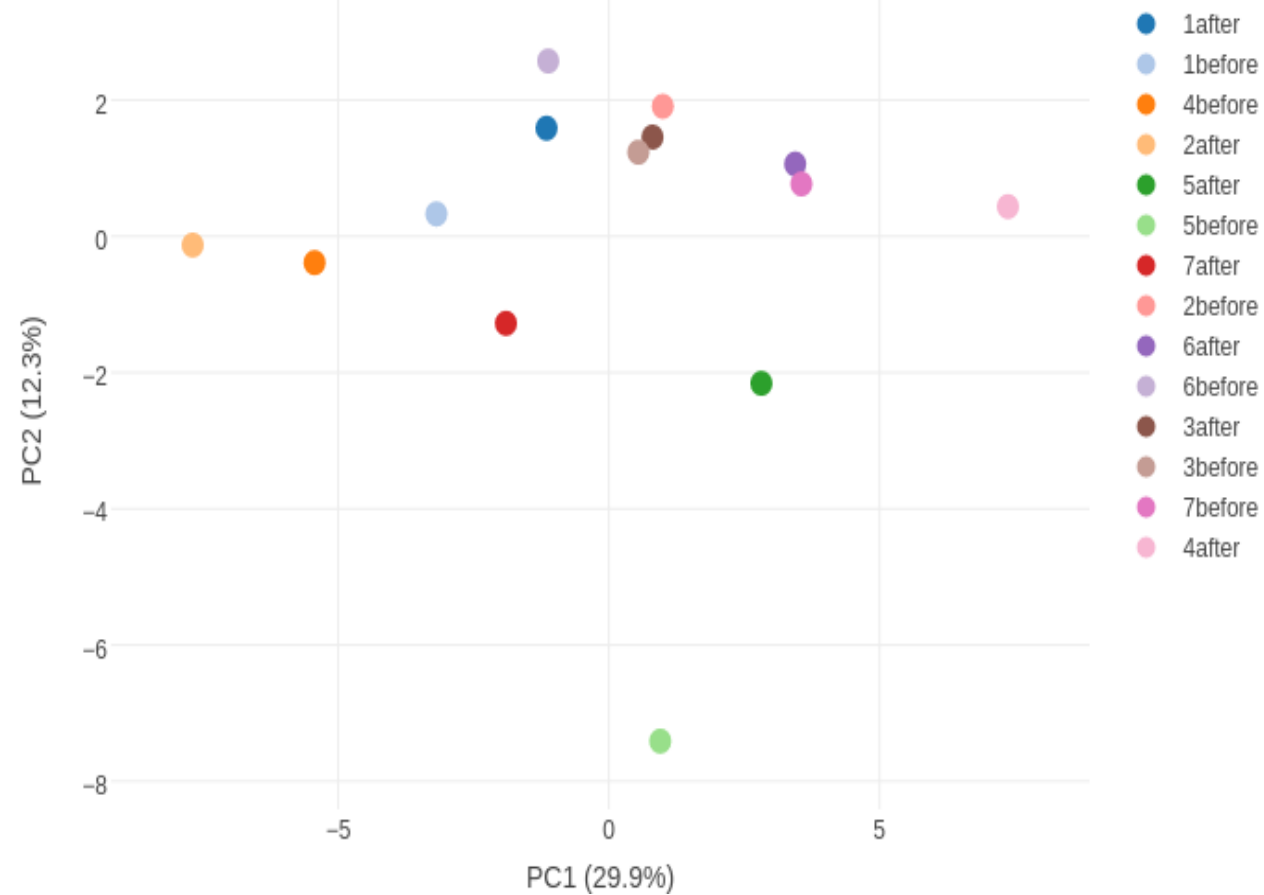
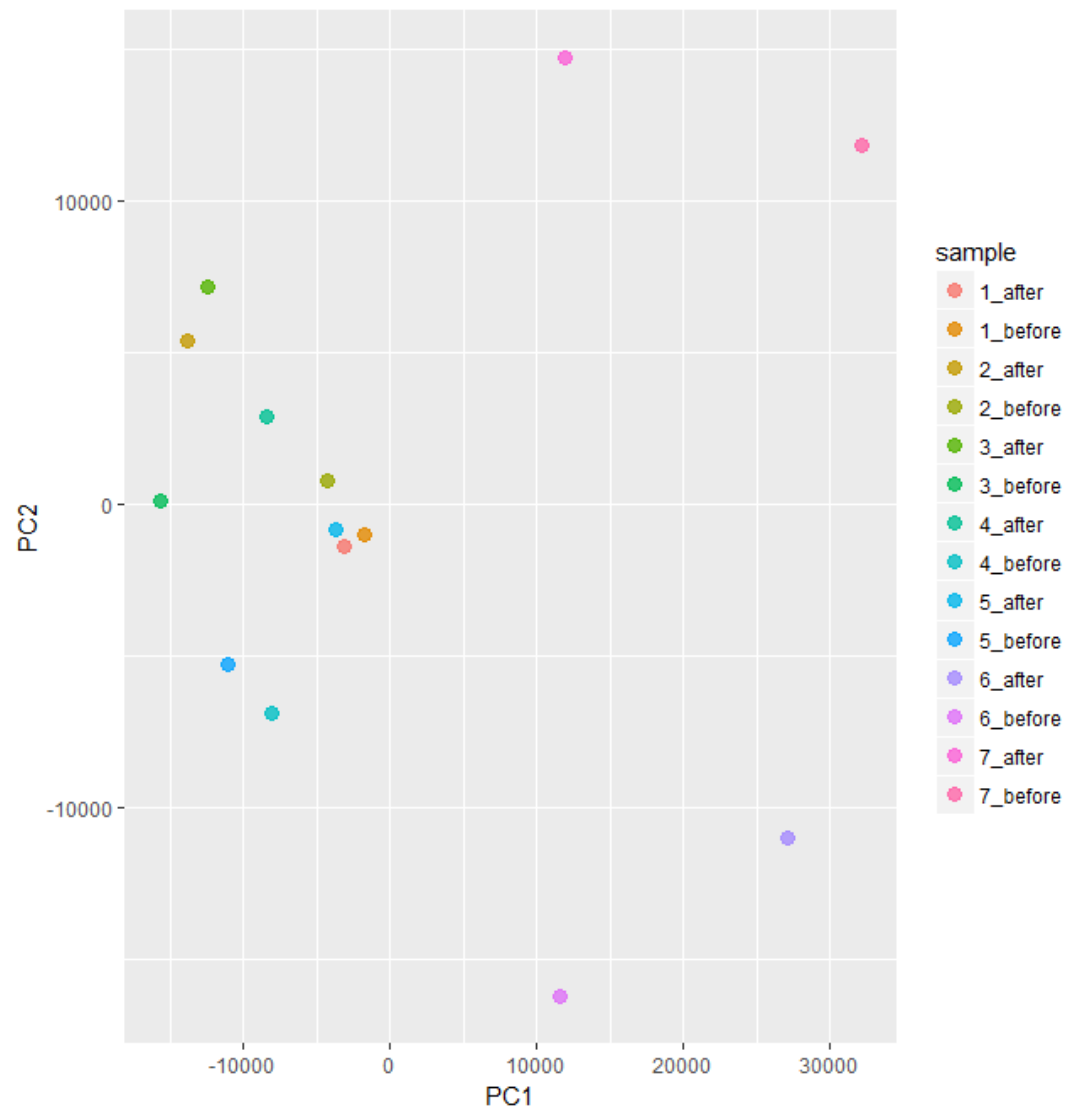
Анализ *kallisto quant* при помощи sleuth

- Всем транскриптам присваивались Ensembl Transcript ID, Ensembl Gene ID и тривиальное название при помощи biomaRt
- Из-за отсутствия каких-либо метаданных в качестве единственного фактора использовалось состояние («до тренировки/после тренировки»)
- Полученную таблицу отфильтровали по $p\text{-value} < 0.05$

Density plot



PCA: sleuth x Phantastus (top 50)



ИТОГ

- С транскриптом были проведены базовые операции: QC, псевдовыравнивание, анализ дифференциальной экспрессии
- Из этого можно что-то выжать в дальнейшем

Спасибо за внимание!