

В этом задании рекомендуется создавать виртуальные среды в conda или другом менеджере практически для каждой программы, т.к. возможны конфликты.

Не забывайте создавать отдельные виртуальные среды! Для этого ДЗ и следующего семинара нам понадобятся:

megahit:

```
conda install -c bioconda megahit
```

quast:

```
conda install -c bioconda quast
```

pilon:

```
conda install -c bioconda pilon
```

bowtie2:

```
conda install -c bioconda bowtie2
```

samtools:

```
conda install -c bioconda samtools
```

gatk:

```
conda install -c bioconda gatk4
```

bwa:

```
conda install -c bioconda bwa
```

Загрузите данные для анализа:

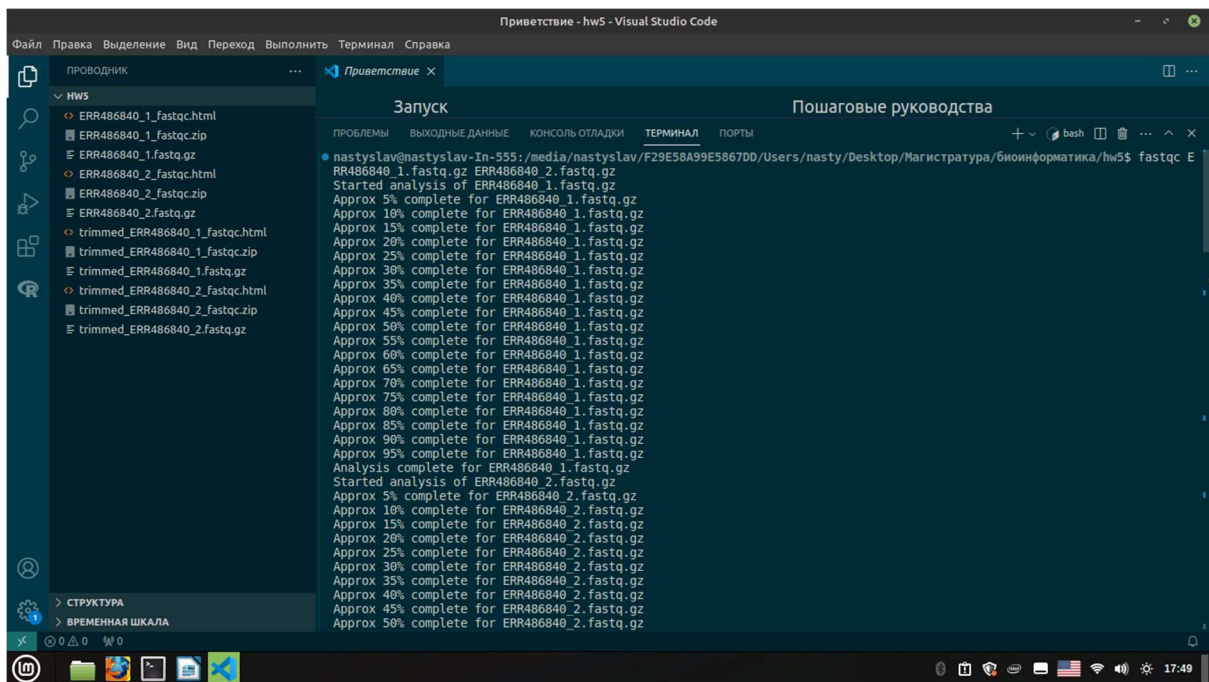
wget

```
ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR486/ERR486840/ERR486840_1.fastq.gz
```

wget

```
ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR486/ERR486840/ERR486840_2.fastq.gz
```

Примените fastqc.



```
Приветствие - hw5 - Visual Studio Code

Файл  Правка  Выделение  Вид  Переход  Выполнить  Терминал  Справка

PROVODNIK

HW5
  ERR486840_1_fastqc.html
  ERR486840_1_fastqc.zip
  ERR486840_1_fastqc.gz
  ERR486840_2_fastqc.html
  ERR486840_2_fastqc.zip
  ERR486840_2_fastqc.gz
  trimmed_ERR486840_1_fastqc.html
  trimmed_ERR486840_1_fastqc.zip
  trimmed_ERR486840_1_fastqc.gz
  trimmed_ERR486840_2_fastqc.html
  trimmed_ERR486840_2_fastqc.zip
  trimmed_ERR486840_2_fastqc.gz

СТРУКТУРА
ВРЕМЕННАЯ ШКАЛА

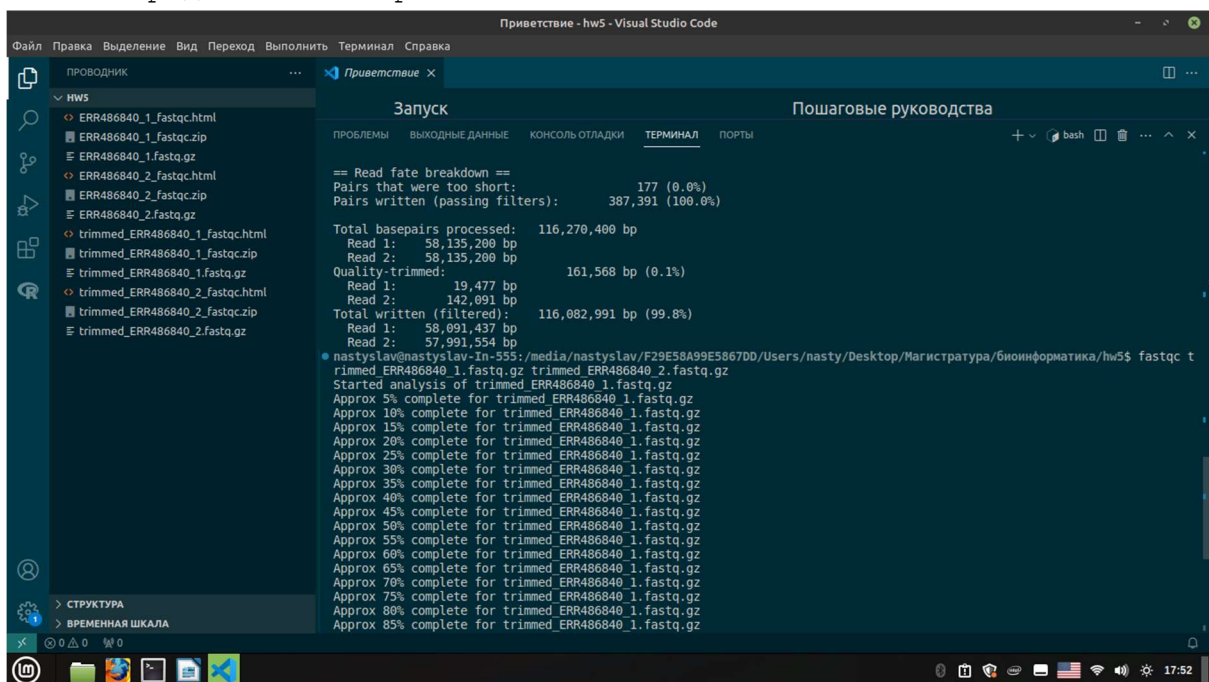
Запуск

ПРОБЛЕМЫ  ВЫХОДНЫЕ ДАННЫЕ  КОНСОЛЬ ОТЛАДКИ  ТЕРМИНАЛ  ПОРТЫ

nastyslav@nastyslav-In-S55: /media/nastyslav/F29E58A99E5867DD/Users/nasty/Desktop/Магистратура/биоинформатика/hw5$ fastqc ERR486840.1.fastq.gz ERR486840.2.fastq.gz
Started analysis of ERR486840.1.fastq.gz
Approx 5% complete for ERR486840.1.fastq.gz
Approx 10% complete for ERR486840.1.fastq.gz
Approx 15% complete for ERR486840.1.fastq.gz
Approx 20% complete for ERR486840.1.fastq.gz
Approx 25% complete for ERR486840.1.fastq.gz
Approx 30% complete for ERR486840.1.fastq.gz
Approx 35% complete for ERR486840.1.fastq.gz
Approx 40% complete for ERR486840.1.fastq.gz
Approx 45% complete for ERR486840.1.fastq.gz
Approx 50% complete for ERR486840.1.fastq.gz
Approx 55% complete for ERR486840.1.fastq.gz
Approx 60% complete for ERR486840.1.fastq.gz
Approx 65% complete for ERR486840.1.fastq.gz
Approx 70% complete for ERR486840.1.fastq.gz
Approx 75% complete for ERR486840.1.fastq.gz
Approx 80% complete for ERR486840.1.fastq.gz
Approx 85% complete for ERR486840.1.fastq.gz
Approx 90% complete for ERR486840.1.fastq.gz
Approx 95% complete for ERR486840.1.fastq.gz
Analysis complete for ERR486840.1.fastq.gz
Started analysis of ERR486840.2.fastq.gz
Approx 5% complete for ERR486840.2.fastq.gz
Approx 10% complete for ERR486840.2.fastq.gz
Approx 15% complete for ERR486840.2.fastq.gz
Approx 20% complete for ERR486840.2.fastq.gz
Approx 25% complete for ERR486840.2.fastq.gz
Approx 30% complete for ERR486840.2.fastq.gz
Approx 35% complete for ERR486840.2.fastq.gz
Approx 40% complete for ERR486840.2.fastq.gz
Approx 45% complete for ERR486840.2.fastq.gz
Approx 50% complete for ERR486840.2.fastq.gz
```

Адаптеры удалять не требуется. Удалите риды малой длины (менее 30), а также отсеки концы ридов с неудовлетворительным качеством прочтения (20).

Какой % ридов был потерян?



```
Приветствие - hw5 - Visual Studio Code

Файл  Правка  Выделение  Вид  Переход  Выполнить  Терминал  Справка

PROVODNIK

HW5
  ERR486840_1_fastqc.html
  ERR486840_1_fastqc.zip
  ERR486840_1_fastqc.gz
  ERR486840_2_fastqc.html
  ERR486840_2_fastqc.zip
  ERR486840_2_fastqc.gz
  trimmed_ERR486840_1_fastqc.html
  trimmed_ERR486840_1_fastqc.zip
  trimmed_ERR486840_1_fastqc.gz
  trimmed_ERR486840_2_fastqc.html
  trimmed_ERR486840_2_fastqc.zip
  trimmed_ERR486840_2_fastqc.gz

СТРУКТУРА
ВРЕМЕННАЯ ШКАЛА

Запуск

ПРОБЛЕМЫ  ВЫХОДНЫЕ ДАННЫЕ  КОНСОЛЬ ОТЛАДКИ  ТЕРМИНАЛ  ПОРТЫ

== Read fate breakdown ==
Pairs that were too short:      177 (0.0%)
Pairs written (passing filters): 387,391 (100.0%)

Total basepairs processed: 116,270,400 bp
  Read 1: 58,135,200 bp
  Read 2: 58,135,200 bp
Quality-trimmed: 161,568 bp (0.1%)
  Read 1: 19,477 bp
  Read 2: 142,091 bp
Total written (filtered): 116,082,991 bp (99.8%)
  Read 1: 58,091,437 bp
  Read 2: 57,991,554 bp

nastyslav@nastyslav-In-S55: /media/nastyslav/F29E58A99E5867DD/Users/nasty/Desktop/Магистратура/биоинформатика/hw5$ fastqc trimmed_ERR486840.1.fastq.gz trimmed_ERR486840.2.fastq.gz
Started analysis of trimmed_ERR486840.1.fastq.gz
Approx 5% complete for trimmed_ERR486840.1.fastq.gz
Approx 10% complete for trimmed_ERR486840.1.fastq.gz
Approx 15% complete for trimmed_ERR486840.1.fastq.gz
Approx 20% complete for trimmed_ERR486840.1.fastq.gz
Approx 25% complete for trimmed_ERR486840.1.fastq.gz
Approx 30% complete for trimmed_ERR486840.1.fastq.gz
Approx 35% complete for trimmed_ERR486840.1.fastq.gz
Approx 40% complete for trimmed_ERR486840.1.fastq.gz
Approx 45% complete for trimmed_ERR486840.1.fastq.gz
Approx 50% complete for trimmed_ERR486840.1.fastq.gz
Approx 55% complete for trimmed_ERR486840.1.fastq.gz
Approx 60% complete for trimmed_ERR486840.1.fastq.gz
Approx 65% complete for trimmed_ERR486840.1.fastq.gz
Approx 70% complete for trimmed_ERR486840.1.fastq.gz
Approx 75% complete for trimmed_ERR486840.1.fastq.gz
Approx 80% complete for trimmed_ERR486840.1.fastq.gz
Approx 85% complete for trimmed_ERR486840.1.fastq.gz
```

Из вывода программы видно, что обработаны все 387,568 пар ридов. Потеряно было только 177 пар (менее 0.1%). Таким образом, процент потерянных ридов составляет менее 0.1%, что является очень небольшим процентом.

Примените fastqc повторно.

Установите ассемблер megahit. Примените его к скачанным файлам (в обоих случаях воспользуйтесь параметрами -1, -2, -o).

```

Приветствие - hw5 - Visual Studio Code
Файл Плавка Выделение Вид Переход Выполнить Терминал Справка
Приветствие x
Запуск Пошаговые руководства
ПРОБЛЕМЫ ВЫХОДНЫЕ ДАННЫЕ КОНСОЛЬ ОТЛАДКИ ТЕРМИНАЛ ПОРТЫ
2024-04-07 18:02:31 - Build graph for k = 21
2024-04-07 18:02:32 - Assemble contigs from SDBG for k = 21
2024-04-07 18:02:33 - Local assembly for k = 21
2024-04-07 18:02:36 - Extract iterative edges from k = 21 to 29
2024-04-07 18:02:38 - Build graph for k = 29
2024-04-07 18:02:38 - Assemble contigs from SDBG for k = 29
2024-04-07 18:02:39 - Local assembly for k = 29
2024-04-07 18:02:42 - Extract iterative edges from k = 29 to 39
2024-04-07 18:02:44 - Build graph for k = 39
2024-04-07 18:02:44 - Assemble contigs from SDBG for k = 39
2024-04-07 18:02:45 - Local assembly for k = 39
2024-04-07 18:02:50 - Extract iterative edges from k = 39 to 59
2024-04-07 18:02:52 - Build graph for k = 59
2024-04-07 18:02:52 - Assemble contigs from SDBG for k = 59
2024-04-07 18:02:53 - Local assembly for k = 59
2024-04-07 18:03:00 - Extract iterative edges from k = 59 to 79
2024-04-07 18:03:02 - Build graph for k = 79
2024-04-07 18:03:02 - Assemble contigs from SDBG for k = 79
2024-04-07 18:03:03 - Local assembly for k = 79
2024-04-07 18:03:10 - Extract iterative edges from k = 79 to 99
2024-04-07 18:03:12 - Build graph for k = 99
2024-04-07 18:03:12 - Assemble contigs from SDBG for k = 99
2024-04-07 18:03:13 - Local assembly for k = 99
2024-04-07 18:03:18 - Extract iterative edges from k = 99 to 119
2024-04-07 18:03:20 - Build graph for k = 119
2024-04-07 18:03:21 - Assemble contigs from SDBG for k = 119
2024-04-07 18:03:21 - Local assembly for k = 119
2024-04-07 18:03:25 - Extract iterative edges from k = 119 to 141
2024-04-07 18:03:26 - Build graph for k = 141
2024-04-07 18:03:26 - Assemble contigs from SDBG for k = 141
2024-04-07 18:03:27 - Merging to output final contigs
2024-04-07 18:03:27 - 20 contigs, total 584402 bp, min 223 bp, max 399763 bp, avg 29220 bp, N50 399763 bp
2024-04-07 18:03:27 - ALL DONE. Time elapsed: 67.024646 seconds
nastyslav@nastyslav-In-535: /media/nastyslav/F29E58A99E3867DD/Users/nasty/Desktop/Магистратура/биоинформатика/hw5$
  
```

Оцените контаминацию. Проверьте, относится ли к другим видам какой-либо из коротких контигов: проведите поиск некоторых контигов в BLASTn против nr/nt базы данных и оцените результаты.

NCBI Blast:Nucleotide Sequence — Mozilla Firefox

Job Title: Nucleotide Sequence

RID: 15HGGK24013 Search expires on 04-09 00:13 am Download All

Program: BLASTN Citation

Database: nt See details

Query ID: lcl|Query\_5008819

Description: None

Molecule type: dna

Query Length: 752

Other reports: Distance tree of results MSA viewer

Filter Results

Organism: only top 20 will appear exclude

Type common name, binomial, taxid or group name

+ Add organism

Percent Identity: to E value: to Query Coverage: to

Filter Reset

Descriptions Graphic Summary Alignments Taxonomy

Sequences producing significant alignments Download Select columns Show 100

select all 100 sequences selected

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Mycoplasma genitalium isolate M6283 MgPa operon complete sequence	Mycoplasma genitalium...	1389	1389	100%	0.0	100.00%	9129	GU226200.1
Mycoplasma genitalium strain 1125.5h adhesion protein MG192 (mgpC) gene, partial cds	Mycoplasma genitalium...	935	935	100%	0.0	89.15%	1471	JX869124.1
Mycoplasma genitalium strain 1125.5h adhesion protein MG192 (mgpC) gene, partial cds	Mycoplasma genitalium...	915	915	100%	0.0	88.68%	1468	JX869121.1

NCBI Blast:Nucleotide Sequence — Mozilla Firefox

https://blast.ncbi.nlm.nih.gov/Blast.cgi

Job Title: Nucleotide Sequence

RID: 15HNWXED013 Search expires on 04-09 00:15 am Download All

Program: BLASTN Citation

Database: nt See details

Query ID: IclQuery\_1669083

Description: None

Molecule type: dna

Query Length: 227

Other reports: Distance tree of results MSA viewer

Filter Results

Organism: only top 20 will appear exclude

Type common name, binomial, taxid or group name

+ Add organism

Percent Identity: to E value: to Query Coverage: to

Filter Reset

Descriptions Graphic Summary Alignments Taxonomy

Sequences producing significant alignments

Download Select columns Show 100

select all 100 sequences selected

GenBank Graphics Distance tree of results MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> Mycoplasma genitalium isolate M6283.MgPa operon _complete sequence	Mycoplasma genitalium isolate M6283.MgPa operon _complete sequence	414	414	100%	4e-111	99.56%	9129	GU226200.1
<input checked="" type="checkbox"/> Mycoplasma genitalium strain 1125.5e adhesion protein MG192 (mgpC) gene _partial cds	Mycoplasma genitalium strain 1125.5e adhesion protein MG192 (mgpC) gene _partial cds	350	350	97%	1e-91	95.05%	1468	JX869121.1
<input checked="" type="checkbox"/> Mycoplasma genitalium strain A52.5B adhesion protein MG192 (mgpC) gene _partial cds	Mycoplasma genitalium strain A52.5B adhesion protein MG192 (mgpC) gene _partial cds	350	350	97%	1e-91	95.05%	1465	JX869103.1

NCBI Blast:Nucleotide Sequence — Mozilla Firefox

https://blast.ncbi.nlm.nih.gov/Blast.cgi

Job Title: Nucleotide Sequence

RID: 15HS8UKK016 Search expires on 04-09 00:17 am Download All

Program: BLASTN Citation

Database: nt See details

Query ID: IclQuery\_5170525

Description: None

Molecule type: dna

Query Length: 277

Other reports: Distance tree of results MSA viewer

Filter Results

Organism: only top 20 will appear exclude

Type common name, binomial, taxid or group name

+ Add organism

Percent Identity: to E value: to Query Coverage: to

Filter Reset

Descriptions Graphic Summary Alignments Taxonomy

Sequences producing significant alignments

Download Select columns Show 100

select all 45 sequences selected

GenBank Graphics Distance tree of results MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> Mycoplasma genitalium strain 73697 chromosome	Mycoplasma genitalium strain 73697 chromosome	268	1171	100%	4e-67	99.32%	579953	CP145106.1
<input checked="" type="checkbox"/> Mycoplasma genitalium strain 75956 chromosome _complete genome	Mycoplasma genitalium strain 75956 chromosome _complete genome	268	1013	100%	4e-67	99.32%	579908	CP145105.1
<input checked="" type="checkbox"/> Mycoplasma genitalium M2288 _complete genome	Mycoplasma genitalium M2288 _complete genome	268	755	96%	4e-67	99.32%	579558	CP003773.1



NCBI Blast:Nucleotide Sequence — Mozilla Firefox

https://blast.ncbi.nlm.nih.gov/Blast.cgi

Job Title: Nucleotide Sequence

RID: 15HUMK89016 Search expires on 04-09 00:18 am Download All

Program: BLASTN Citation

Database: nt See details

Query ID: lcl|Query\_5222743

Description: None

Molecule type: dna

Query Length: 279

Other reports: Distance tree of results MSA viewer

Filter Results

Organism: only top 20 will appear exclude

Type common name, binomial, taxid or group name

+ Add organism

Percent Identity: to E value: to Query Coverage: to

Filter Reset

Descriptions Graphic Summary Alignments Taxonomy

Sequences producing significant alignments

Download Select columns Show 100

select all 100 sequences selected

GenBank Graphics Distance tree of results MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> Mycoplasma genitalium strain M2282 MgPar 1 region genomic sequence	Mycoplasma genitalium	235	235	49%	5e-57	97.12%	1101	FJ872570.1
<input checked="" type="checkbox"/> Mycoplasma genitalium strain 64 MgPar 2 region genomic sequence	Mycoplasma genitalium	213	213	100%	2e-50	82.08%	2427	FJ872593.1
<input checked="" type="checkbox"/> Mycoplasma genitalium clone 1534EF90 MgPa adhesin protein (mgpB) gene, partial cds	Mycoplasma genitalium	206	325	93%	4e-48	94.16%	681	MT439579.1

NCBI Blast:Nucleotide Sequence — Mozilla Firefox

https://blast.ncbi.nlm.nih.gov/Blast.cgi

Job Title: Nucleotide Sequence

RID: 15HXTKFT013 Search expires on 04-09 00:20 am Download All

Program: BLASTN Citation

Database: nt See details

Query ID: lcl|Query\_5293299

Description: None

Molecule type: dna

Query Length: 221

Other reports: Distance tree of results MSA viewer

Filter Results

Organism: only top 20 will appear exclude

Type common name, binomial, taxid or group name

+ Add organism

Percent Identity: to E value: to Query Coverage: to

Filter Reset

Descriptions Graphic Summary Alignments Taxonomy

Sequences producing significant alignments

Download Select columns Show 100

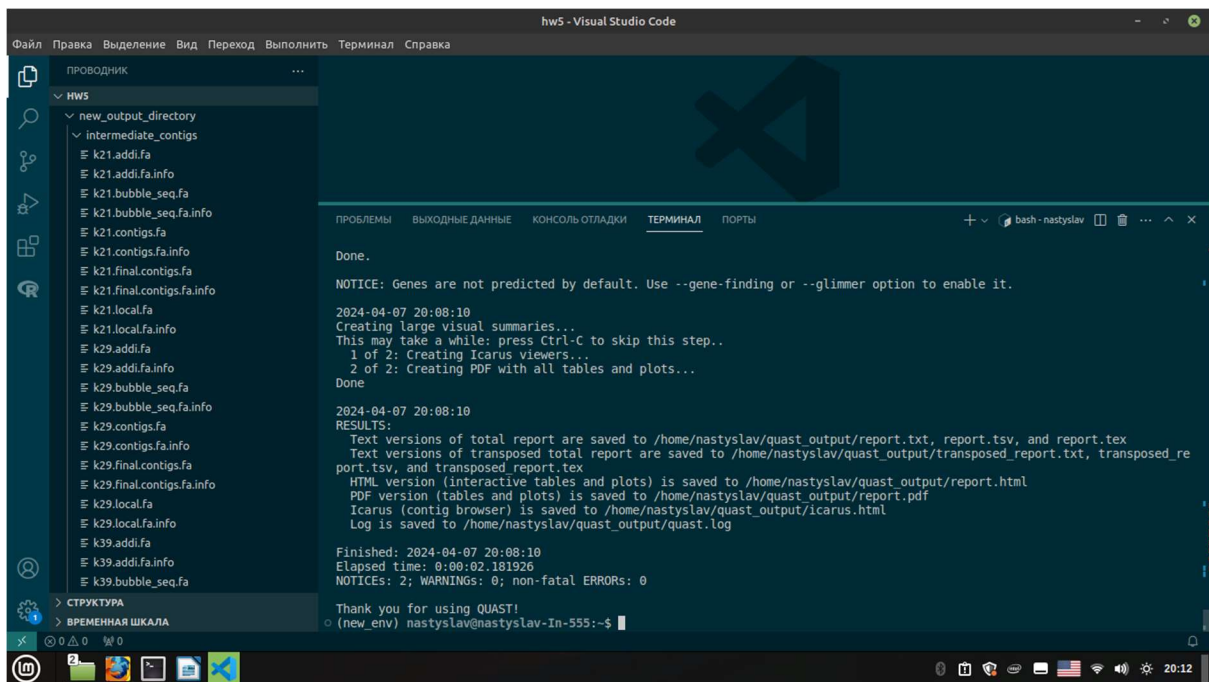
select all 100 sequences selected

GenBank Graphics Distance tree of results MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> Mycoplasma genitalium isolate M6283 MgPa operon, complete sequence	Mycoplasma genitalium	409	409	100%	2e-109	100.00%	9129	GU226200.1
<input checked="" type="checkbox"/> Mycoplasma genitalium isolate 111.2f adhesion protein MG192 (mgpC) gene, partial cds	Mycoplasma genitalium	392	392	100%	2e-104	98.64%	1476	JX857887.1
<input checked="" type="checkbox"/> Mycoplasma genitalium isolate 111.2e adhesion protein MG192 (mgpC) gene, partial cds	Mycoplasma genitalium	392	392	100%	2e-104	98.64%	1470	JX857886.1

Проверила несколько контигов в BLASTn. Данные последовательности принадлежат *Mycoplasma genitalium*. Контаминации не обнаружено т.к все эти последовательности получены от одного вида.

Проверьте качество полученного генома при помощи QUAST.



1) Сколько контигов получилось?

**Всего в результате сборки получилось 20 контигов. 8 контигов длиной не менее 1000 п.н., 5 контигов длиной не менее 5000 п.н., 5 контигов длиной не менее 10 000 п.н., 5 контигов длиной не менее 25 000 п.н. и 1 контиг длиной не менее 50000 п.н.**

<b>Assembly</b>	<b>final.contigs</b>
# contigs (>= 0 bp)	20
# contigs (>= 1000 bp)	8
# contigs (>= 5000 bp)	5
# contigs (>= 10000 bp)	5
# contigs (>= 25000 bp)	5
# contigs (>= 50000 bp)	1
Total length (>= 0 bp)	584402
Total length (>= 1000 bp)	579589
Total length (>= 5000 bp)	571952
Total length (>= 10000 bp)	571952
Total length (>= 25000 bp)	571952
Total length (>= 50000 bp)	399763
# contigs	11
Largest contig	399763
Total length	581884
GC (%)	31.70
N50	399763
N75	49179
L50	1
L75	2
# N's per 100 kbp	0.00

2) Каков показатель N50? Что это означает?

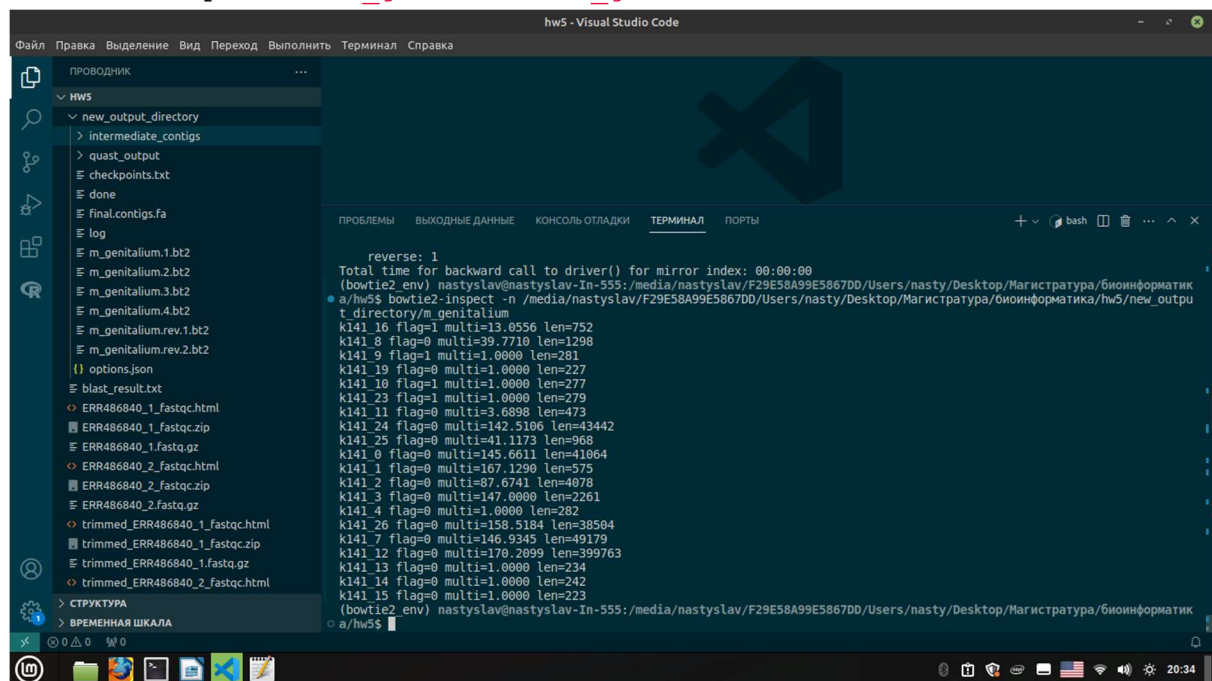
**399763. Показатель N50 представляет собой длину самого длинного контига, такую что суммарная длина всех контигов равна или больше половины общей длины генома.**

Исправим ошибки сборки при помощи Pilon: для этого сначала нужно создать базу данных для выравнивания bowtie2:

```
bowtie2-build m_genitalium/final.contigs.fa  
m_genitalium/m_genitalium
```

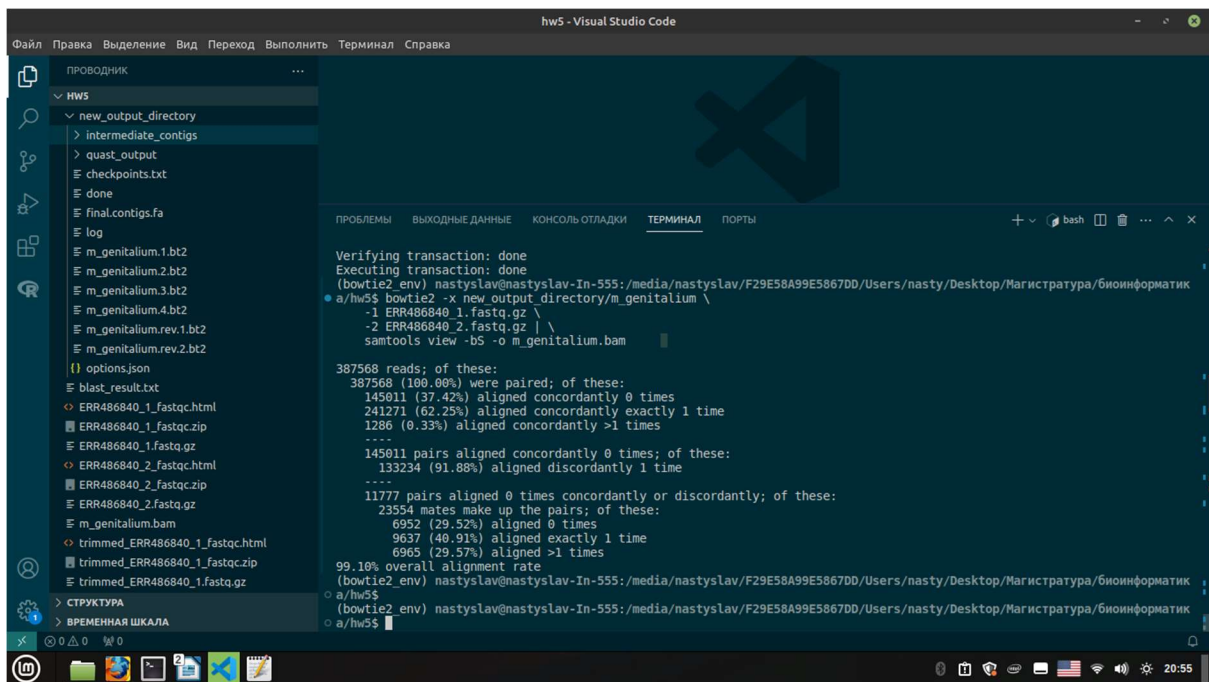
Можно посмотреть, что получилось:

```
bowtie2-inspect -n m_genitalium/m_genitalium
```



Затем выравниваем исходные риды на геном и сразу через pipe переводим в сжатый bam-формат:

```
bowtie2 -x m_genitalium/m_genitalium \  
-1 ERR486840_1.fastq.gz \  
-2 ERR486840_2.fastq.gz | \  
samtools view -bS -o m_genitalium.bam
```

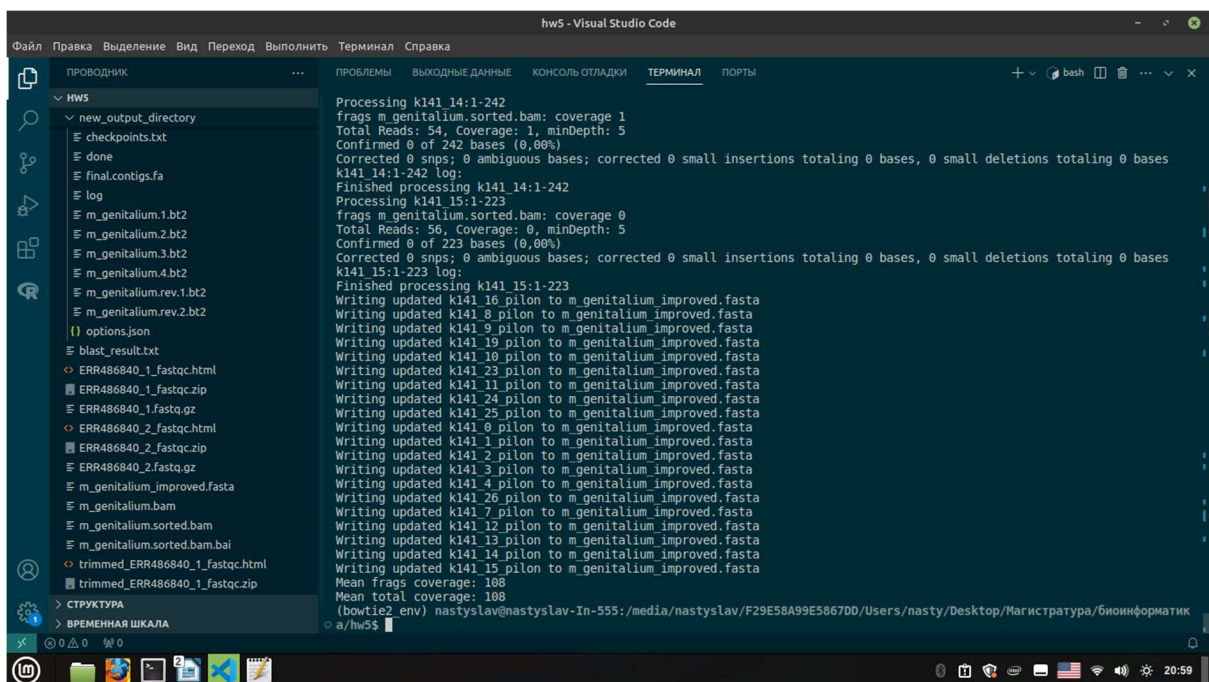


Осуществляем сортировку и индексирование полученного файла:

```
samtools sort m_genitalium.bam -o m_genitalium.sorted.bam
samtools index m_genitalium.sorted.bam
```

Наконец, применим Pilon:

```
pilon --genome m_genitalium/final.contigs.fa --frags
m_genitalium.sorted.bam --output m_genitalium_improved
```





Изменилось ли что-то? Что именно? Разберите вывод работы программы.

```
pilon --genome new_output_directory/final.contigs.fa --frags  
m_genitalium.sorted.bam --output m_genitalium_improve  
d
```

Pilon version [COMMIT] [DATE]

Genome: new\_output\_directory/final.contigs.fa

Fixing snps, indels, gaps, local *Pilon применяет исправления для  
однонуклеотидных полиморфизмов (SNP), вставок и делеций (indels),  
пробелов и локальных проблем.*

Input genome size: 584402 *Размер исходного генома до исправлений.*

Scanning BAMs

m\_genitalium.sorted.bam: 775136 reads, 0 filtered, 768184 mapped,  
485114 proper, 277074 stray, FR 100% 384+/-69, max 592 *Информация о  
выравнивании ридов на геном.*

*Далее идет обработка каждого участка генома*

Processing k141\_16:1-752

frags m\_genitalium.sorted.bam: coverage 61

Total Reads: 698, Coverage: 61, minDepth: 6

Confirmed 739 of 752 bases (98,27%)

Corrected 0 snps; 0 ambiguous bases; corrected 0 small insertions  
totaling 0 bases, 0 small deletions totaling 0 bases

k141\_16:1-752 log:

Finished processing k141\_16:1-752

Processing k141\_8:1-1298

frags m\_genitalium.sorted.bam: coverage 68

Total Reads: 1303, Coverage: 68, minDepth: 7

Confirmed 1282 of 1298 bases (98,77%)

Corrected 0 snps; 0 ambiguous bases; corrected 0 small insertions  
totaling 0 bases, 0 small deletions totaling 0 bases

k141\_8:1-1298 log:

Finished processing k141\_8:1-1298

Processing k141\_9:1-281

frags m\_genitalium.sorted.bam: coverage 0

Total Reads: 107, Coverage: 0, minDepth: 5

Confirmed 0 of 281 bases (0,00%)

Corrected 0 snps; 0 ambiguous bases; corrected 0 small insertions  
totaling 0 bases, 0 small deletions totaling 0 bases

k141\_9:1-281 log:

Finished processing k141\_9:1-281

Processing k141\_19:1-227

frags m\_genitalium.sorted.bam: coverage 0

Total Reads: 3, Coverage: 0, minDepth: 5

Confirmed 0 of 227 bases (0,00%)

Corrected 0 snps; 0 ambiguous bases; corrected 0 small insertions  
totaling 0 bases, 0 small deletions totaling 0 bases

k141\_19:1-227 log:

Finished processing k141\_19:1-227

Processing k141\_10:1-277

```
frags m_genitalium.sorted.bam: coverage 0
Total Reads: 1, Coverage: 0, minDepth: 5
Confirmed 0 of 277 bases (0,00%)
Corrected 0 snps; 0 ambiguous bases; corrected 0 small insertions
totaling 0 bases, 0 small deletions totaling 0 bases
k141_10:1-277 log:
Finished processing k141_10:1-277
Processing k141_23:1-279
frags m_genitalium.sorted.bam: coverage 1
Total Reads: 115, Coverage: 1, minDepth: 5
Confirmed 0 of 279 bases (0,00%)
Corrected 0 snps; 0 ambiguous bases; corrected 0 small insertions
totaling 0 bases, 0 small deletions totaling 0 bases
k141_23:1-279 log:
Finished processing k141_23:1-279
Processing k141_11:1-473
frags m_genitalium.sorted.bam: coverage 35
Total Reads: 421, Coverage: 35, minDepth: 5
Confirmed 465 of 473 bases (98,31%)
Corrected 0 snps; 0 ambiguous bases; corrected 0 small insertions
totaling 0 bases, 0 small deletions totaling 0 bases
k141_11:1-473 log:
Finished processing k141_11:1-473
Processing k141_24:1-43442
frags m_genitalium.sorted.bam: coverage 97
Total Reads: 52320, Coverage: 97, minDepth: 10
Confirmed 43408 of 43442 bases (99,92%)
Corrected 3 snps; 0 ambiguous bases; corrected 0 small insertions
totaling 0 bases, 0 small deletions totaling 0 bases
k141_24:1-43442 log:
Finished processing k141_24:1-43442
Processing k141_25:1-968
frags m_genitalium.sorted.bam: coverage 58
Total Reads: 983, Coverage: 58, minDepth: 6
Confirmed 928 of 968 bases (95,87%)
Corrected 0 snps; 0 ambiguous bases; corrected 0 small insertions
totaling 0 bases, 0 small deletions totaling 0 bases
k141_25:1-968 log:
Finished processing k141_25:1-968
Processing k141_0:1-41064
frags m_genitalium.sorted.bam: coverage 99
Total Reads: 49338, Coverage: 99, minDepth: 10
Confirmed 41031 of 41064 bases (99,92%)
Corrected 0 snps; 0 ambiguous bases; corrected 0 small insertions
totaling 0 bases, 0 small deletions totaling 0 bases
# Attempting to fix local continuity breaks
k141_0:1-41064 log:
Finished processing k141_0:1-41064
```

```
Processing k141_1:1-575
frags m_genitalium.sorted.bam: coverage 43
Total Reads: 535, Coverage: 43, minDepth: 5
Confirmed 566 of 575 bases (98,43%)
Corrected 0 snps; 0 ambiguous bases; corrected 0 small insertions
totaling 0 bases, 0 small deletions totaling 0 bases
k141_1:1-575 log:
Finished processing k141_1:1-575
Processing k141_2:1-4078
frags m_genitalium.sorted.bam: coverage 91
Total Reads: 4673, Coverage: 91, minDepth: 9
Confirmed 4062 of 4078 bases  Количество баз, подтвержденных после
исправлений.
```

*Можно сделать вывод о том, что Pilon исправил или подтвердил большинство баз в геноме. Так же программа осуществила изменения в нескольких контигах, где были обнаружены ошибки или несоответствия между выравниванием ридов. Большинство контигов осталось без изменений.*