

23. Metabarkodlama veri analizi

1. Illumina Paired End

(<http://metabarcoding.org/obitools/doc/scripts/illuminapairedend.html>)

```
cd ..
mkdir obi_process/1_align
illuminapairedend --score-min=40 -r raw/sample.fastq
raw/sample.fastq > obi_process/1_align/sample.align.fastq
cd obi_process
obihead -n 1 1_align/sample.align.fastq
obicount 1_align/sample.align.fastq
```

2. Filtering Unpaired

(<http://metabarcoding.org/obitools/doc/scripts/obigrep.html>)

```
mkdir 2_merge
obigrep -p 'mode!="joined"' 1_align/sample.align.fastq >
2_merge/sample.merge.fastq
obicount 2_merge/sample.merge.fastq
awk 'NR%4==1' 1_align/sample.align.fastq
awk 'NR%4==1' 1_align/sample.align.fastq | awk -F ";"
' {print$13}' | sort | uniq
```

Option 1:

3a. Extracting tags and removing primers from cleaned sequences

```
mkdir 3_tag  
ngsfilter -t sample.ngsfilter.txt -u unidentified.fastq  
2_merge/sample.merge.fastq > 3_tag/sample.tag.fastq  
obicount 3_tag/sample.tag.fastq
```

```
CAAAC TGGGATTAGATACCCCACTATG -mm5 3 -mm3 3
perl ~/ngs/software/tagcleaner-standalone-0.16/tagcleaner.pl -
fastq 2_merge/sample.merge.fastq -out 3_trim/sample.trim.info
-tag5 GTCGGTAAACTCGTGCCAGC -tag3
CAAAC TGGGATTAGATACCCCACTATG -mm5 3 -mm3 3 -
info
obihead -n 1 3_trim/sample.trim.fastq
obihead -n 1 3_trim/sample.trim.info.fastq
awk 'NR%4==1' 3_trim/sample.trim.info.fastq
awk 'NR%4==1' 3_trim/sample.trim.info.fastq | cut -d " " -f3
awk 'NR%4==1' 3_trim/sample.trim.info.fastq | cut -d " " -f3 |
sort -n | head
awk 'NR%4==1' 3_trim/sample.trim.info.fastq | cut -d " " -f3 |
sort -n | tail
awk 'NR%4==1' 3_trim/sample.trim.info.fastq | cut -d " " -f3 |
sort -n | uniq -c
awk 'NR%4==1' 3_trim/sample.trim.info.fastq | cut -d " " -f3 |
sort -n | uniq -c > 3_trim/sample.trim.stats
awk 'NR%4==1' 3_trim/sample.trim.info.fastq | cut -d " " -f2 |
sort -n | uniq -c > 3_trim/sample.untrim.stats
awk 'NR%4==1' 3_trim/sample.trim.info.fastq | cut -d " " -f6 |
sort -n | uniq -c > 3_trim/sample.5mismatch.stats
awk 'NR%4==1' 3_trim/sample.trim.info.fastq | cut -d " " -f7 |
sort -n | uniq -c > 3_trim/sample.3mismatch.stats
```

4. Uniq Sequences and Converting into FASTA

(<http://metabarcoding.org/obitools/doc/scripts/obiuniq.html>)

```
mkdir 4_uniq
obiuniq -m sample 3_tag/sample.tag.fastq >
4_uniq/sample.uniq.fasta
obihead -n 1 4_uniq/sample.uniq.fasta
obicount 4_uniq/sample.uniq.fasta
obiannotate -k count -k merged_sample
4_uniq/sample.uniq.fasta > $$ ; mv $$
4_uniq/sample.uniq.fasta
obihead -n 1 4_uniq/sample.uniq.fasta
```

OBISStats

(<http://metabarcoding.org/obitools/doc/scripts/obistat.html>)

```
obistat -c count 4_uniq/sample.uniq.fasta
obistat -c count 4_uniq/sample.uniq.fasta | sort -nk1 | head
obistat -c count 4_uniq/sample.uniq.fasta | sort -nk1 | tail
grep "count=111096" 4_uniq/sample.uniq.fasta
grep -A 5 "count=111096" 4_uniq/sample.uniq.fasta
```

5. OBICount

```
mkdir 5_filter
obigrep -l 150 -L 300 -p 'count>=10' 4_uniq/sample.uniq.fasta
> 5_filter/sample.c10.1150.L300.fasta
obigrep -l 150 -L 300 -p 'count<10 and count>=1'
4_uniq/sample.uniq.fasta > 5_filter/sample.c1-
10.1150.L300.fasta
obicount 5_filter/sample.c10.1150.L300.fasta
obicount 5_filter/sample.c5-10.1150.L300.fasta
```

6. Filtering PCR/Sequencing Errors ([filtering_errors.pdf](#)) (<http://metabarcoding.org/obitools/doc/scripts/obiclean.html>)

```
mkdir 6_clean
obiclean -r 0.05 -H 5_filter/sample.c10.1150.L300.fasta >
6_clean/sample.c10.1150.L300.clean.fasta
obiclean -r 0.05 -H 5_filter/sample.c1-10.1150.L300.fasta >
6_clean/sample.c1-10.1150.L300.clean.fasta
obihead -n 1 6_clean/sample.c10.1150.L300.clean.fasta
obicount 6_clean/sample.c10.1150.L300.clean.fasta
```

7. Taxonomic Assignment (offline)

(<http://metabarcoding.org/obitools/doc/scripts/ecotag.html>)

```
cd ~/ngs/data/sampleID/obi_process
mkdir 7_ecotag
ecotag -d ~/ngs/database/embl/ecopcr/vrt_r142 -R
~/NGS/database/embl/ecopcr_primerID/db.fasta
6_clean/sample.c10.1150.L300.clean.fasta >
7_ecotag/sample.c10.ecotag.fasta
ecotag -d ~/ngs/database/embl/ecopcr/vrt_r142 -R
~/NGS/database/embl/ecopcr_primerID/db.fasta
6_clean/sample.c1-10.1150.L300.clean.fasta >
7_ecotag/sample.c1-10.ecotag.fasta
```

Assignment using remote BLAST (online)

```
mkdir 11_rblast.results
```

** “ ” problematic, manually write the code*

```
blastn -query 6_clean/sample.c10.1150.L250.clean.fasta -db nt -  
remote -max_target_seqs 1 -outfmt “6 qseqid stitle length pident  
ssciname” -out 11_rblast.results/sample.c10.rblast.results
```

```
blastn -query 6_clean/sample.c1-10.1150.L250.clean.fasta -db nt  
-remote -max_target_seqs 1 -outfmt “6 qseqid stitle length  
pident ssciname” -out 11_rblast.results/sample.c1-  
10.rblast.results
```

Assignment using Local Database

```
cd ~/ngs/data/sampleID/obi_process
mkdir 12_lblast.results
blastn -h
blastn -query 6_clean/sample.c10.1150.L300.clean.fasta -db
~/ngs/database/blastdb/localdb.fasta -max_target_seqs 1 -outfmt
"6 qseqid sseqid length pident" -out
12_lblast.results/sample.c10.lblast.results
blastn -query 6_clean/sample.c1-10.1150.L300.clean.fasta -db
~/ngs/database/blastdb/localdb.fasta -max_target_seqs 1 -outfmt
"6 qseqid sseqid length pident" -out 12_lblast.results/sample.c1-
10.lblast.results
```

8. Annotating/Exporting to Spreadsheet

(<http://metabarcoding.org/obitools/doc/scripts/obitab.html>)

```
mkdir 8_ann
```

```
obihead 7_ecotag/sample.c10.ecotag.fasta
```

```
obihead 7_ecotag/sample.c1-10.ecotag.fasta
```

```
obiannotate --delete-tag=scientific_name_by_db --delete-  
tag=obiclean_samplecount --delete-tag=obiclean_count --  
delete-tag=obiclean_singletoncount --delete-  
tag=obiclean_cluster --delete-tag=obiclean_internalcount --  
delete-tag=obiclean_head --delete-tag=taxid_by_db --delete-  
tag=obiclean_headcount --delete-tag=id_status --delete-  
tag=rank_by_db --delete-tag=order_name --delete-tag=order  
7_ecotag/sample.c10.ecotag.fasta > 8_ann/sample.c10.ann.fasta
```

```
obiannotate --delete-tag=scientific_name_by_db --delete-  
tag=obiclean_samplecount --delete-tag=obiclean_count --  
delete-tag=obiclean_singletoncount --delete-  
tag=obiclean_cluster --delete-tag=obiclean_internalcount --  
delete-tag=obiclean_head --delete-tag=taxid_by_db --delete-  
tag=obiclean_headcount --delete-tag=id_status --delete-  
tag=rank_by_db --delete-tag=order_name --delete-tag=order  
7_ecotag/sample.c1-10.ecotag.fasta > 8_ann/sample.c1-  
10.ann.fasta
```

```
obihead 8_ann/sample.c10.ann.fasta
```

9. Sorting Read Numbers

```
mkdir 9_sort  
obisort -k count -r 8_ann/sample.c10.ann.fasta >  
9_sort/sample.c10.sort.fasta  
obisort -k count -r 8_ann/sample.c1-10.ann.fasta >  
9_sort/sample.c1-10.sort.fasta
```

10. Tabbing Results

```
mkdir 10_embl.results  
obitab -o 9_sort/sample.c10.sort.fasta >  
10_embl.results/sample.c10.results.tab  
obitab -o 9_sort/sample.c10.sort.fasta >  
10_embl.results/sample.c10.results.xls  
obitab -o 9_sort/sample.c1-10.sort.fasta >  
10_embl.results/sample.c1-10.results.tab  
obitab -o 9_sort/sample.c1-10.sort.fasta >  
10_embl.results/sample.c1-10.results.xls
```