22.Metabarkodlama veri tabanı oluşturma

```
sudo apt-get update
sudo apt-get upgrade
sudo apt install python-dev
sudo apt install python-pip
sudo apt install gcc
sudo apt install subversion
pip install -U virtualenv
pip install -U sphinx
pip install -U cython
sudo apt install default-jre
sudo apt install perl
sudo apt-get install libz-dev
sudo apt-get install p7zip-full
sudo apt-get install screen

cd ~
mkdir ngs
```

```
cd ngs
mkdir software
cd software
```

OBITools (https://pythonhosted.org/OBITools/welcome.html#)

```
wget https://pythonhosted.org/OBITools/_downloads/get-obitools.py
python get-obitools.py
./obitools
cd OBITools-*/src
python setup.py --serenity install
cd ~/ngs/software
```

TagCleaner (http://tagcleaner.sourceforge.net/)

* check for latest version

```
wget
https://sourceforge.net/projects/tagcleaner/files/standalone/tagcleaner-standalone-0.16.tar.gz/download
tar -zxvf tagcleaner-standalone-*.tar.gz
```

FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/)

** check for latest version*

```
wget https://www.bioinformatics.babraham.ac.uk/projects/fastqc/fastqc_v0.11.9.zip
unzip fastqc_*
```

ecoPCR (https://git.metabarcoding.org/obitools/ecopcr/wikis/home):

** check for latest version*

```
wget https://git.metabarcoding.org/obitools/ecopcr/uploads/aa3828c196570ea156ce6d4baac22b10/ecopcr-1.0.1.tar.gz
tar -zxvf ecopcr-*.tar.gz
cd ecopcr/src/
make
```

ncbi standalone (https://www.ncbi.nlm.nih.gov/books/NBK52640/)

** check for latest version*

```
cd ~/ngs/software/
wget ftp://ftp.ncbi.nlm.nih.gov/blast/executables/LATEST/ncbi-blast-2.10.0+-x64-linux.tar.gz
tar -zxvpf  ncbi-blast-*.tar.gz
sudo apt install ncbi-blast+
export PATH=$PATH:$HOME/NGS/software/ncbi-blast-2.10.0+/bin
echo $PATH
blastn -h
```

Copying Data

```
cd ..
mkdir data
cd data
mkdir sampleID
cd sampleID
mkdir raw
cd raw
```

*copy your *.fastq.gz files here*

```
ls -l
gunzip *.gz
```

** faster alternative but should run for each file: 7z x rel_std_vrt_(dosya adı)*

```
ls -l
```

Analyzing fastq data ((http://www.wiki-zero.com/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvRkFTVFFfZm9ybWF0)

```
more sample.fastq
cat sample.fastq | awk 'NR%4==1{printf ">%s\n",
substr($0,2)}NR%4==2{print}' > sample.fasta
more sample.fasta
rm sample.fasta
```

FastQC Analysis

```
cd ..
mkdir obi_process
mkdir obi_process/0_fastqc
cd ~/ngs/software/FastQC/
fastqc -h|more
fastqc ~/ngs/data/sampleID/raw/sample.fastq -o
~/NGS/data/sampleID/obi_process/0_fastqc
```

```
cd ~/ngs/
mkdir database
cd database
mkdir embl
mkdir embl/raw
cd embl/raw
wget
ftp://ftp.ebi.ac.uk/pub/databases/ena/sequence/release/std/rel_std_vrt*.dat.gz
gunzip *.gz
* 7z x rel_std_vrt_file name
cd ../
```

NCBI taxdump

```
mkdir ../../taxo
cd ../../taxo
wget ftp://ftp.ncbi.nlm.nih.gov/pub/taxonomy/taxdump.tar.gz
tar -xzvf taxdump.tar.gz
more readme.txt
```

Converting to ecoPCR Format

(http://metabarcoding.org/obitools/doc/scripts/obiconvert.html)

```
cd ../embl
mkdir ecopcr
cd ecopcr
obiconvert --embl -t ../../taxo/ --ecopcrdb-output=vrt_r142 --
skip-on-error ../raw/*.dat
```

in silico PCR on ecoPCR Database

(http://metabarcoding.org/obitools/doc/scripts/ecoPCR.html)

```
cd ../
mkdir ecopcr_primerID
ecoPCR -d ~/ngs/database/embl/ecopcr/vrt_r142 -e 3 -l 100 -L
300 GTCGGTAAAACTCGTGCCAGC
CATAGTGGGGTATCTAATCCCAGTTTG >
~/ngs/database/embl/ecopcr_primerID/12S.vrt.r142.ecopcr
cd  ecopcr_primerID
more 12S.vrt.r142.ecopcr
cut -d "|" -f6 12S.vrt.r142.ecopcr | sort | uniq
cut -d "|" -f6 12S.vrt.r142.ecopcr | sort | uniq | wc -l
```

Converting Reference Database to FASTA

(https://pythonhosted.org/OBITools/wolves.html#clean-the-database)

```
obigrep -d ~/ngs/database/embl/ecopcr/vrt_r142 --require-rank=family --require-rank=genus --require-rank=species 12S.vrt.r142.ecopcr > 12S.vrt.r142.clean.fasta
obicount 12S.vrt.r142.clean.fasta
obiuniq -d ~/ngs/database/embl/ecopcr/vrt_r142 12S.vrt.r142.clean.fasta > 12S.vrt.r142.clean.uniq.fasta
obicount 12S.vrt.r142.clean.uniq.fasta
obiannotate --uniq-id 12S.vrt.r142.clean.uniq.fasta > db.fasta
```

Taxonomy database for BLASTn (https://www.ncbi.nlm.nih.gov/books/NBK279680/)

```
cd ~/ngs/database
mkdir blastdb
cd blastdb
wget ftp://ftp.ncbi.nlm.nih.gov/blast/db/taxdb.tar.gz
tar -zxvf taxdb.tar.gz
export BLASTDB=$HOME/ngs/database/blastdb
echo $BLASTDB
```

Building Local Reference Database

(https://www.ncbi.nlm.nih.gov/books/NBK279688/)

```
cd ~/ngs/database/blastdb
makeblastdb -h
makeblastdb -in localdb.fasta -parse_seqids -dbtype nucl
sed ‘s/ //g’ localdb.fasta
sed -i ‘s/ //g’ localdb.fasta
makeblastdb -in localdb.fasta -parse_seqids -dbtype nucl
```

```
cd ~/ngs/data/mifish/raw/
head sample.fastq
obihead sample.fastq
obitail sample.fastq
grep "GATCTGGG" sample.fastq
obigrep -s "GATCTGGG" sample.fastq
obicount sample.fastq
```