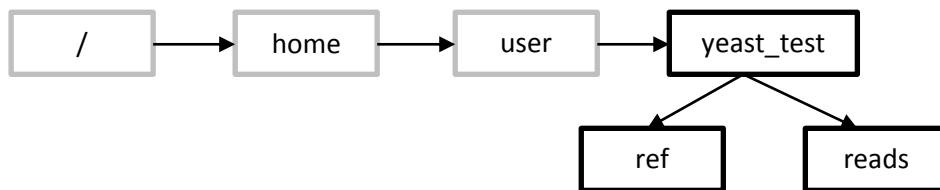


# Introduction to Linux: Exercises

- 1) Open a Terminal window.
- 2) Get the path of your working directory.
- 3) List the content of your home directory.
- 4) Create the following folder structure (The grey folders already exist).



- 5) Use "ls" to check if it worked. If you accidentally created a folder with an incorrect name use "rmdir" to remove it.
- 6) What are the commands "tar" and "wget" for? Check the man pages or the internet.

- 7) Download the yeast read sequences into the “yeast\_test” folder.  
(URL: [http://www.cibiv.at/~philipp\\_/files/yeast-reads.tar](http://www.cibiv.at/~philipp_/files/yeast-reads.tar) )
- 8) Move the “yeast-reads.tar” files to the “reads” folder.
- 9) Open the man pages for the command “tar” to see how to unzip the archive containing the reads “yeast-reads.tar” (look for examples).
- 10) Unzip “yeast-reads.tar”
- 11) List the contents of the “reads” folder.
- 12) If everything worked, delete the “yeast-reads.tar” file.
- 13) Combine the three read files into a single file “yeast-reads\_all.fq”.
- 14) Delete all read files except “yeast-reads\_all.fq”
- 15) Extract the first 10 reads from file “yeast-reads\_all.fq” and save them as a new file called “test\_reads.fq”.  
*The file is a fastq file. Each read consists of four lines in the FASTQ format.*
- 16) Extract the last 5 reads from the file “yeast-reads\_all.fq” and append them to the file “test\_reads.fq”.  
*The file is a fastq file. Each read consists of four lines in the FASTQ format.*
- 17) Open the file to see whether it worked. It should contain 15 reads (command: cat, head, less, nano).
- 18) Use “nano” to remove the read that consist only of “N” characters.  
*Use Ctrl + K to remove a line. Use Ctrl + O to save and Ctrl + X to exit nano.*

19) Count how many reads are contained in the files “yeast-reads\_all.fq” and “test\_reads.fq”

(command: wc, Try to find the parameter that tells wc to only count the lines)

*The file is a fastq file. Each read consists of four lines in the FASTQ format.*

20) Download the yeast reference and unzip it into the “ref” folder

URL: [http://www.cibiv.at/~philipp\\_/files/yeast-ref.tar.gz](http://www.cibiv.at/~philipp_/files/yeast-ref.tar.gz)

*Use “tar -xvfz” for unzipping*

21) Count the number of sequences present in “yeast-ref.fq” (command: grep)

*The file is in fasta format. Each sequence entry starts with an “>”*

22) Extract the names of all reference sequences and save them in a new file “ref\_names.txt”.

23) Open “ref\_names.txt” to see whether it worked.

24) Get the line number at which the sequence with the name “chromosome=IX” starts in the file “yeast-ref.fasta”.

*Check the grep man page.*

25) Download BWA in the “yeast\_test” folder and unzip it.

URL: [http://www.cibiv.at/~philipp\\_/files/bwa.tar.gz](http://www.cibiv.at/~philipp_/files/bwa.tar.gz)

*Use the command “tar -xvfz” to unzip the bwa.tar.gz file*

26) Try to execute BWA.

27) Make the file executable (command: chmod).

28) Run BWA.

29) Optional: Map the reads with the following commands:

`./bwa index refs/yeast-ref.fasta` (this might take 2 or 3 minutes)

`./bwa aln refs/yeast-ref.fasta reads/test_reads.fq > result.sai`

`./bwa samse refs/yeast-ref.fasta result.sai reads/yeast-reads_all.fq > result.sam`

The mapped reads are located in the “results.sam” file. We will discuss the SAM format later.