

```
cd exercise_6
```

```
#####
# show contents of folder
#
ls -lah
total 24K
drwxr-xr-x 2 dinkel gibson 4.0K Aug 20 10:46 .
drwxr-xr-x 7 dinkel gibson 4.0K Aug 31 11:17 ..
-rw-r--r-- 1 dinkel gibson 4.1K Jul 31 18:26 EPSINS.fasta
-rw-r--r-- 1 dinkel gibson 628 Aug 6 17:18 FYN_HUMAN.fasta
-rw-r--r-- 1 dinkel gibson 643 Aug 6 17:23 SRC_HUMAN.fasta
#####

#####
# cat (=catenate) the file SRC_HUMAN.fasta
#
cat SRC_HUMAN.fasta
>sp|P12931|SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens GN=SRC PE=1 SV=3
MGSNKS KPKDASQRRRSLEPAENVHGAGGGAFFPASQTPSKPASADGHRGPSAAAFAPAAAE
PKLFGGFNSSDVTSTPQRAGPLAGGVTTFFVALYDYESRTETDLSFKKGERLQIVNNTTEGD
WWLAHSLSTGQTGYIPSNYVAPSDSIQAEWYFGKITRRESERLLLNAENPRGTFVLVRES
ETTKGAYCLSVSDFDNAKGLNVKHYKIRKLDSGGFYITSRTQFNSLQQLVAYYSKHADGL
CHRLTTVCPTSKPQTQGLAKDAWEIPRESLRLEVKLGGCGFGEVWMGTWNGTTRVAIKTL
KPGTMSPEAFLQEAQVMKKLRHEKLVQLYAVVSEEPYIVTEYMSKGSLLDFLKGETGKY
LRLPQLVDMAAQIASGMAYVERMNYVHRDLRAANILVGENLVCKVADFGLARLIEDNEYT
ARQGAKFPIKWTAPAAALYGRFTIKSDVWSFGILLTELTTKGRVPYPGMVNREVLDQVER
GYRMPCPECPESLHDLMCQCWRKEPEERPTFEYLQAFLEDYFTSTEPQYQPGENL
#####

#####
# cat the file SRC_HUMAN.fasta, showing line-numbers, Tabs, line-ends and non-printing characters
#
cat -n -A SRC_HUMAN.fasta
1 >sp|P12931|SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens GN=SRC PE=1 SV=3$
2 MGSNKS KPKDASQRRRSLEPAENVHGAGGGAFFPASQTPSKPASADGHRGPSAAAFAPAAAE$
3 PKLFGGFNSSDVTSTPQRAGPLAGGVTTFFVALYDYESRTETDLSFKKGERLQIVNNTTEGD^I$
4 WWLAHSLSTGQTGYIPSNYVAPSDSIQAEWYFGKITRRESERLLLNAENPRGTFVLVRES$
5 ETTKGAYCLSVSDFDNAKGLNVKHYKIRKLDSGGFYITSRTQFNSLQQLVAYYSKHADGL$
6 CHRLTTVCPTSKPQTQGLAKDAWEIPRESLRLEVKLGGCGFGEVWMGTWNGTTRVAIKTL$
7 KPGTMSPEAFLQEAQVMKKLRHEKLVQLYAVVSEEPYIVTEYMSKGSLLDFLKGETGKY$
8 LRLPQLVDMAAQIASGMAYVERMNYVHRDLRAANILVGENLVCKVADFGLARLIEDNEYT$
9 ARQGAKFPIKWTAPAAALYGRFTIKSDVWSFGILLTELTTKGRVPYPGMVNREVLDQVER$
10 GYRMPCPECPESLHDLMCQCWRKEPEERPTFEYLQAFLEDYFTSTEPQYQPGENL$
#####

#####
# cat the file SRC_HUMAN.fasta, showing line-numbers
#
cat -n SRC_HUMAN.fasta
1 >sp|P12931|SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens GN=SRC PE=1 SV=3
2 MGSNKS KPKDASQRRRSLEPAENVHGAGGGAFFPASQTPSKPASADGHRGPSAAAFAPAAAE
3 PKLFGGFNSSDVTSTPQRAGPLAGGVTTFFVALYDYESRTETDLSFKKGERLQIVNNTTEGD
4 WWLAHSLSTGQTGYIPSNYVAPSDSIQAEWYFGKITRRESERLLLNAENPRGTFVLVRES
5 ETTKGAYCLSVSDFDNAKGLNVKHYKIRKLDSGGFYITSRTQFNSLQQLVAYYSKHADGL
6 CHRLTTVCPTSKPQTQGLAKDAWEIPRESLRLEVKLGGCGFGEVWMGTWNGTTRVAIKTL
7 KPGTMSPEAFLQEAQVMKKLRHEKLVQLYAVVSEEPYIVTEYMSKGSLLDFLKGETGKY
8 LRLPQLVDMAAQIASGMAYVERMNYVHRDLRAANILVGENLVCKVADFGLARLIEDNEYT
9 ARQGAKFPIKWTAPAAALYGRFTIKSDVWSFGILLTELTTKGRVPYPGMVNREVLDQVER
10 GYRMPCPECPESLHDLMCQCWRKEPEERPTFEYLQAFLEDYFTSTEPQYQPGENL
#####

#####
# cat two files
#
cat SRC_HUMAN.fasta FYN_HUMAN.fasta
>sp|P12931|SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens GN=SRC PE=1 SV=3
MGSNKS KPKDASQRRRSLEPAENVHGAGGGAFFPASQTPSKPASADGHRGPSAAAFAPAAAE
```

```
PKLFGGFNSSDVTVTSPQRAGPLAGGVTTTFVALYDYESRTETDLSFKKGERLQIVNTEGD
WWLAHSLSTGQTGYIPSNYVAPSDSIQAEWYFGKITRRESERLLLNAENPRGTFILVRES
ETTKGAYCLSVSDFDNAKGLNVKHYKIRKLDSSGGFYITSRTQFNSLQQLVAYYSKHADGL
CHRLTTVCPTSKPQTQGLAKDAWEIPRESLRLEVKLGGCFGEVWMGTWNGTTRVAIKTL
KPGTMSPEAFLEAQAVMVKLRHEKLVQLYAVVSEEPYIYIVTEYMSKGSLLDFLKGETGKY
LRLPQLVDMAAQIASGMAYVERMNYVHRDLRAANILVGENLVCKVADFGARLIEDNEYT
ARQGAKFPIKWTAPAAALYGRFTIKSDVWSFGILLTELTTKGRVPYPGMVNREVLQDQVER
GYRMPCPPECPESLHDLMCQCWRKEPEERPTFEYLLQAFLEDYFTSTEPQYQPGENL
>sp|P06241|FYN_HUMAN Tyrosine-protein kinase Fyn OS=Homo sapiens GN=FYN PE=1 SV=3
MGCVCQCKDKEATKLTEERDGSNLQSSGYRYGTDPTPQHYPSTFVTSIPNYYNFHAAGGQG
LTVFGGVNSSSHTGTLRTRGGTGVTFLFVALYDYEARTEDDLSFHKGEKFQILNSSEGDDWW
EARSLLTTGETGYIPSNYVAPVDSIQAEWYFGKLGRKDAERQLLSFGNPRGTFILIRESET
TKGAYSLSIRDWDDMKGDHVKHYKIRKLDNGGYYITTRAQFETLQQLVQHYSERAAAGLCC
RLVVPCHKGMPLRLTDLVSKTKDVWEIPRESLQLIKRLGNGQFGEVWMGTWNGNTKVAIKT
LKPGTMSPEFLEEAIQIMKKLKHDKLVQLYAVVSEEPYIYIVTEYMNKGSLLDFLKDGEGR
ALKLPNLVDMAAQVAAGMAYIERMNYIHRDLRSANILVGNGLICKIADFGARLIEDNEY
TARQGAKFPIKWTAPAAALYGRFTIKSDVWSFGILLTELVTKGRVPYPGMNNREVLEQVE
RGYRMPCPQDCPISLHELMIHCKWKDPEERPTFEYLLQSFLLEDYFTATEPQYQPGENL
#####

#####
# pause
#
sleep 2
#####

#####
# just view the first 10 lines of the file EPSINS.fasta
#
head EPSINS.fasta
>sp|Q96N21|AP4AT_HUMAN AP-4 complex accessory subunit tepsin OS=Homo sapiens GN=C17o
rf56 PE=1 SV=1
MAAAPPLRDRLSFLHRLPILLKGTSDDDVPCPGYLFEEIAKISHESPGSSQCLLEYLLSR
LHSSSGHGKLVKILLYLCSHGSSFFLLILKRNSAFIQEAFAAGPPDPLHGNSLYQKV
RAAAQDLGSTLFDSTVPLAPSPQLGTPPATGMGSQARPHSTLQGFQYSGKEHGRTAVRHQ
PGQAGGGWDELDSGPSSQNSQNSDLRSVSDSGSHSGSDSHSGASREPGLAERVEVVAL
SDCQQELSLVRTVTRGPRAFLSREEAQHFIAKAGLLNCEAVLQLLTCHLRGTSECTQLRA
LCAIASLGSSDLLPQEHILLRTRPWLQELSMGSPGPVTNKATKILRHFEASCGQLSPARG
TSAEPGPPTAALPGPSDLLTDAVPLPGSQVFLQPLSSTPVSSRSPAPSSGMPSSPVPTPPP
DASPIAPAGDPSEAEARLAESRRWRPERIPGGTDSKRGPPSSCAWSRDSLFAGMELVACP
RLVGAGAAAGSCPDAPRAPQTSSQRTAAKEPPGSEPSAFAFLNA
#####

#####
# just view the last 10 lines of the file EPSINS.fasta
#
tail EPSINS.fasta
MNMLWSRMLKDNKKNWRRVYKSLLLLAYLIRNGSERVVTSAREHIYDLRSLNHYHFVDEH
GKDQGINIRQKVVELVEFAQDDRLREERKKAKKNKDKYVGVSSDSVGGFRYSERYDPEP
KSKWDEEWDKNKSAFPFSDKLGELSDKIGSTIDDTISKFRKDRDSDPERCSDSDEEKKA
RRGRSPKGEFKDEEETVTTKHIHITQATETTTTRHKRTANPSKTIDLGAAHYTGDKASP
DQNASTHTPQSSVKTSVPSSKSSGDLVDLFDGTSQSTGGSADLFGGFADFGSAAASGSP
SQVTATSGNGDFGDWSAFNQAPSGPVASSGEFFGSASQPAVELVSGSQSALGPPPAASNS
SDLFDLMGSSQATMTSSQSMNFSMMSTNTVGLGLPMSRSQNTDMVQKSVSKTLPSTWSDP
SVNISLDNLLPGMQPSKPQQPSLNTMIQQQNMQQPMNMTQSFQAVNLSSPSNMLPVRPQ
TNALIGGPMPMSMPNMTGTMGMAPLGNTPMNQSMGMNMNIGMSAAGMGLTGTMGMMGM
PNIAMTSGTVQPKQDAFANFANFSK
#####

#####
# view the first line of each fasta file in the current directory:
#
head -n1 *.fasta -q
>sp|Q96N21|AP4AT_HUMAN AP-4 complex accessory subunit tepsin OS=Homo sapiens GN=C17o
rf56 PE=1 SV=1
>sp|P06241|FYN_HUMAN Tyrosine-protein kinase Fyn OS=Homo sapiens GN=FYN PE=1 SV=3
>sp|P12931|SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens GN=S
RC PE=1 SV=3
#####
```

```
#####
# pause
#
sleep 2
#####

#####
# Use two different ways to display only the headerline (first line, starting with ">")
# of the two files SRC_HUMAN.fasta FYN_HUMAN.fasta
#
head -n1 SRC_HUMAN.fasta FYN_HUMAN.fasta
==> SRC_HUMAN.fasta <==
>sp|P12931|SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens GN=SRC PE=1 SV=3

==> FYN_HUMAN.fasta <==
>sp|P06241|FYN_HUMAN Tyrosine-protein kinase Fyn OS=Homo sapiens GN=FYN PE=1 SV=3
grep ">" SRC_HUMAN.fasta FYN_HUMAN.fasta
SRC_HUMAN.fasta:>sp|P12931|SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens GN=SRC PE=1 SV=3
FYN_HUMAN.fasta:>sp|P06241|FYN_HUMAN Tyrosine-protein kinase Fyn OS=Homo sapiens GN=FYN PE=1 SV=3
cat SRC_HUMAN.fasta FYN_HUMAN.fasta | grep ">"
>sp|P12931|SRC_HUMAN Proto-oncogene tyrosine-protein kinase Src OS=Homo sapiens GN=SRC PE=1 SV=3
>sp|P06241|FYN_HUMAN Tyrosine-protein kinase Fyn OS=Homo sapiens GN=FYN PE=1 SV=3
#####

cd -
/home/dinkel/exercises
```