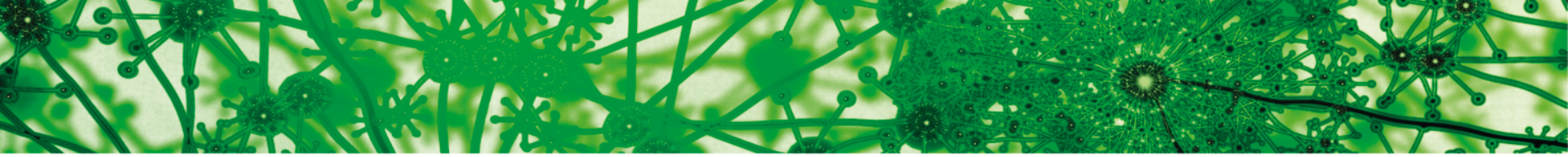
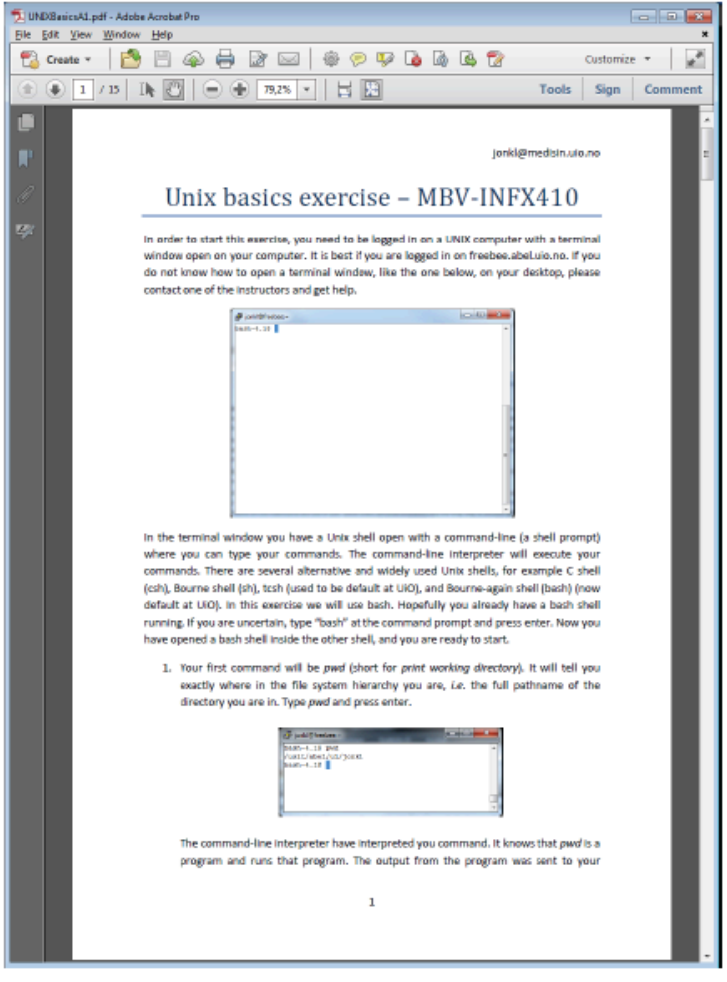


UiO : Department of Biosciences
University of Oslo

**November 16. More Unix**

# Bioinformatics for Molecular Biology




Jon K. Lærdahl,
Structural Bioinformatics

Very important to be able to use basic Unix

Make your own Unix scripts is *not* very central

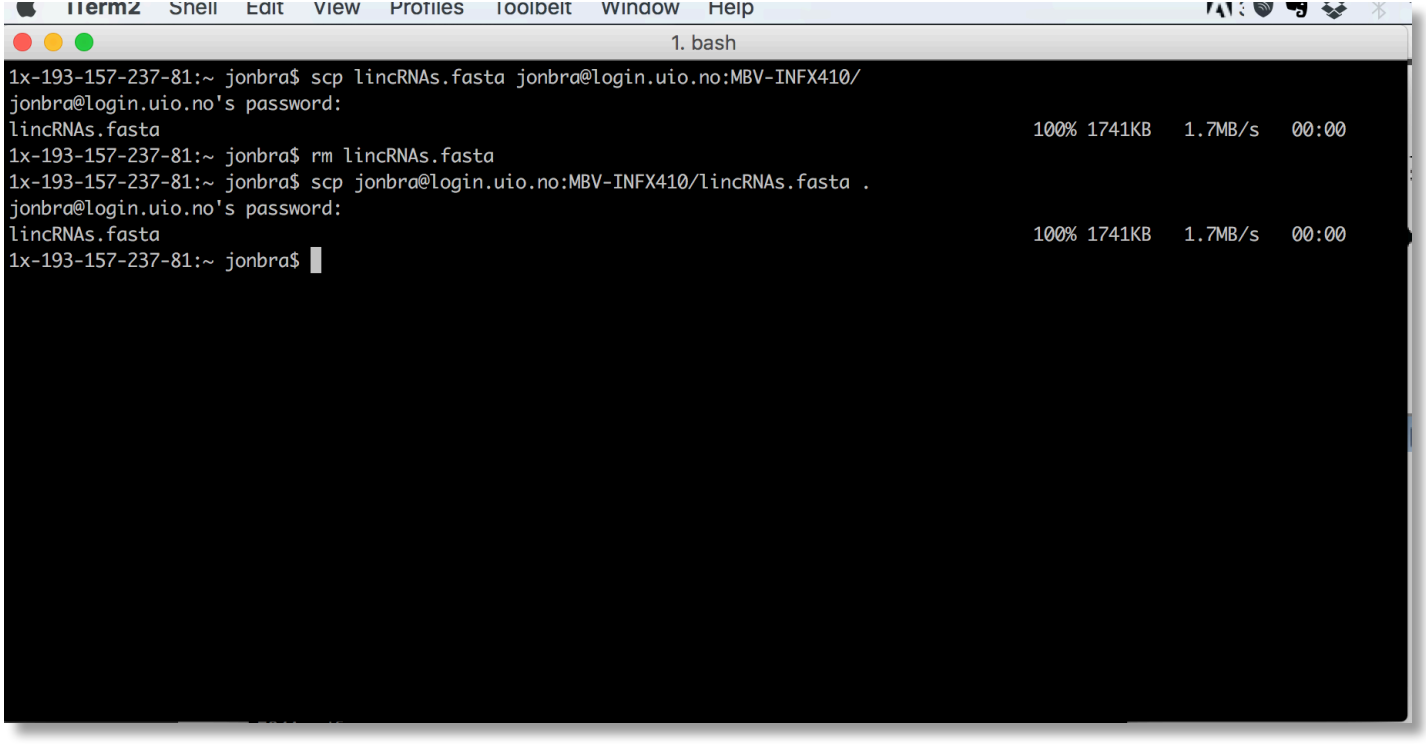
Most likely you will instead use Python for writing small programs/scripts

Jon K. Lærdahl,
Structural Bioinformatics

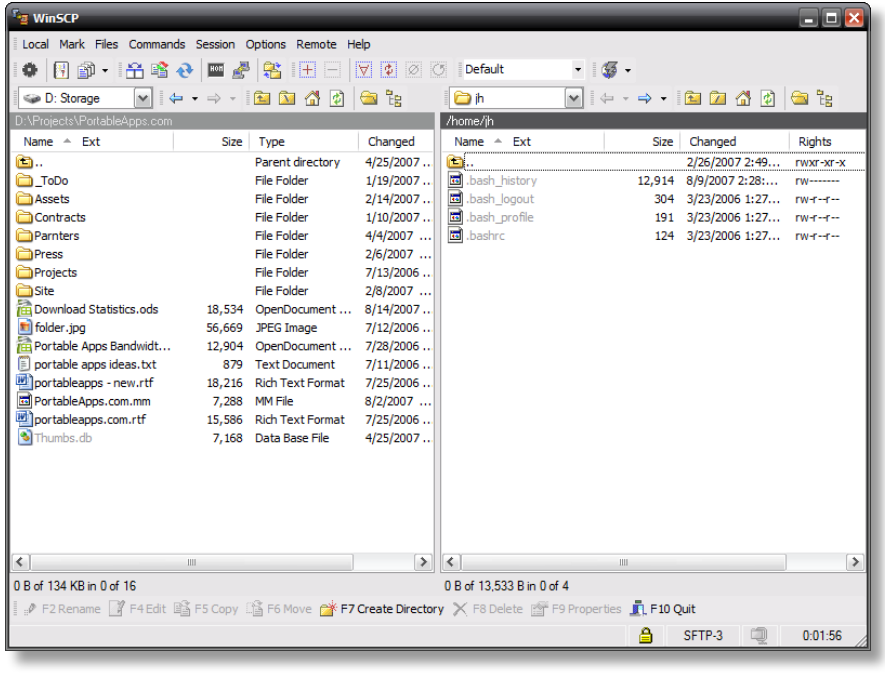
# More Unix stuff

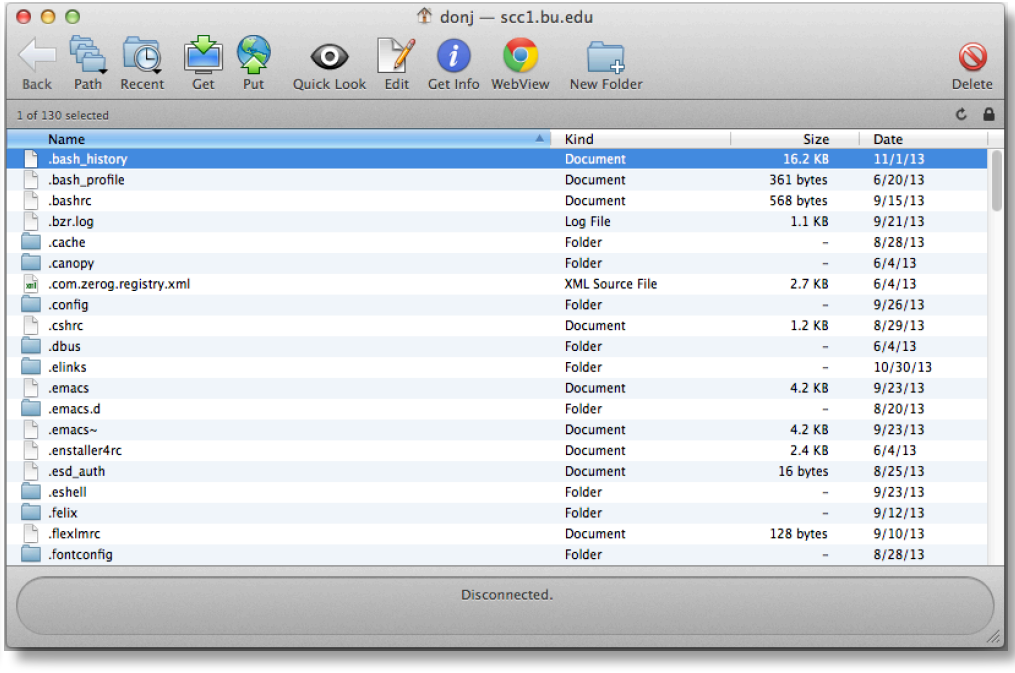
- You ***must*** be able to move files between laptop and a Unix server such as freebee.abel.uio.no or login.uio.no!
- That is, between the harddisk on your laptop and your UiO home area (M:)
- Most likely you use ”UiO Programkiosk”
  - Or WebDrive, Cyberduck etc.

# Copy files to and from UiO home area (Terminal/Cygwin etc.)

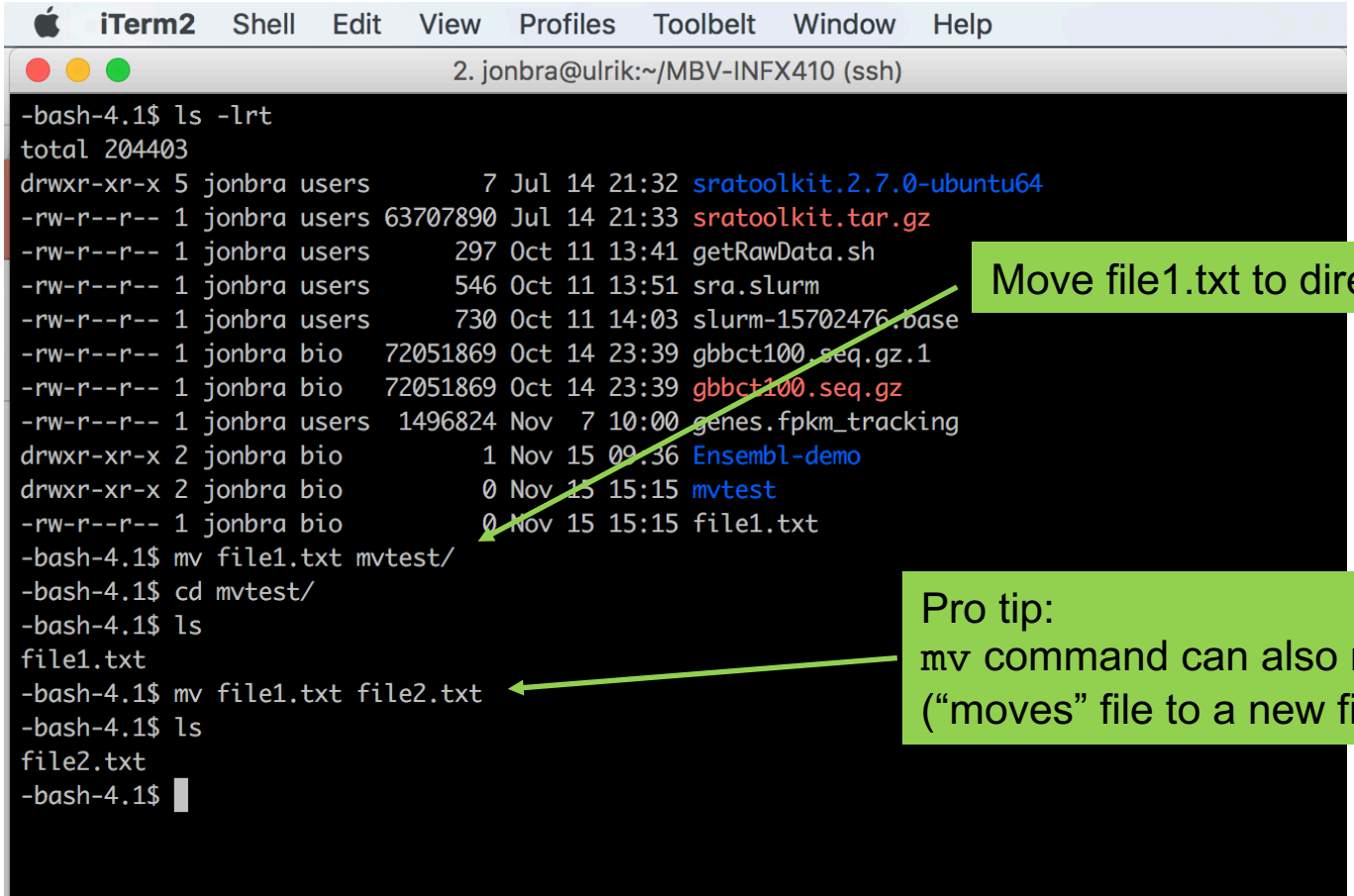


# Copy files to and from UiO home area (graphical)



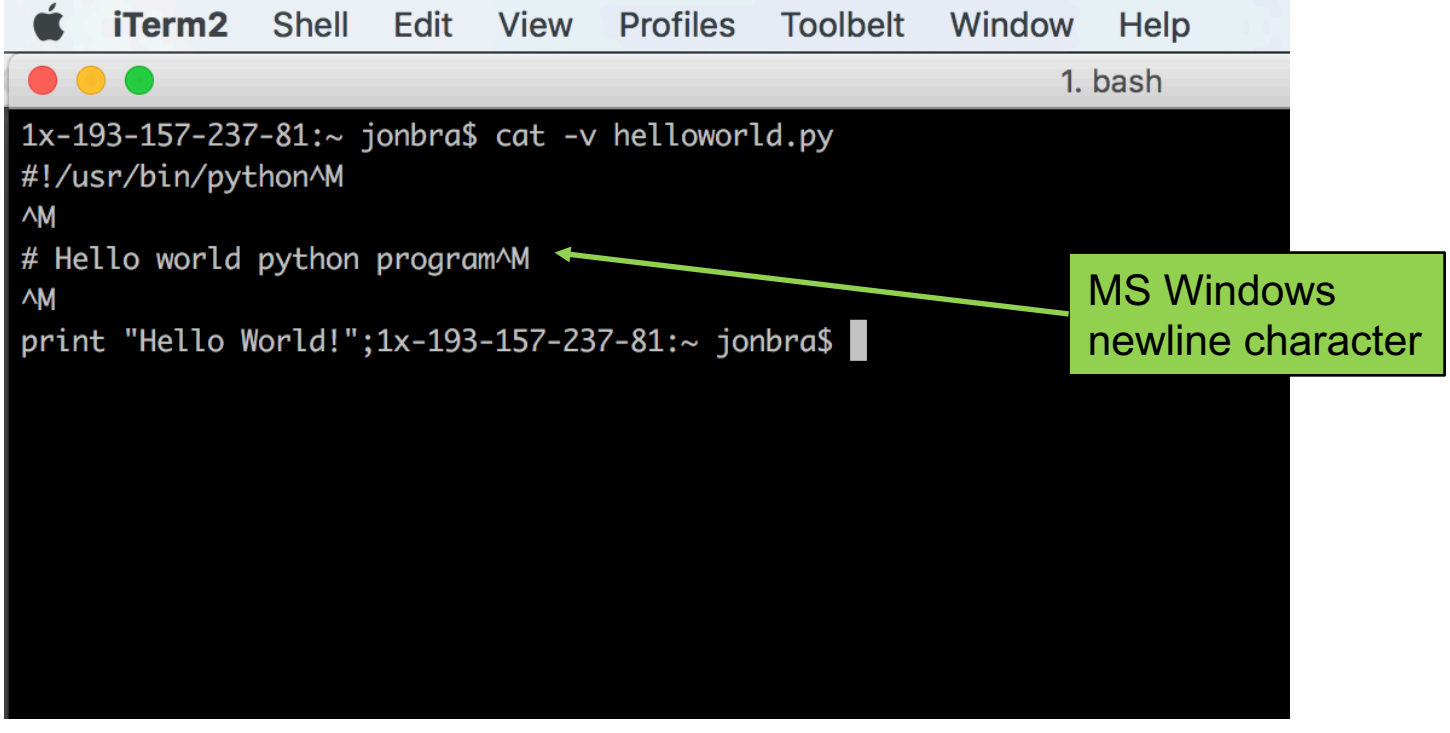



# Moving files with the `mv` command



# Non-printing characters

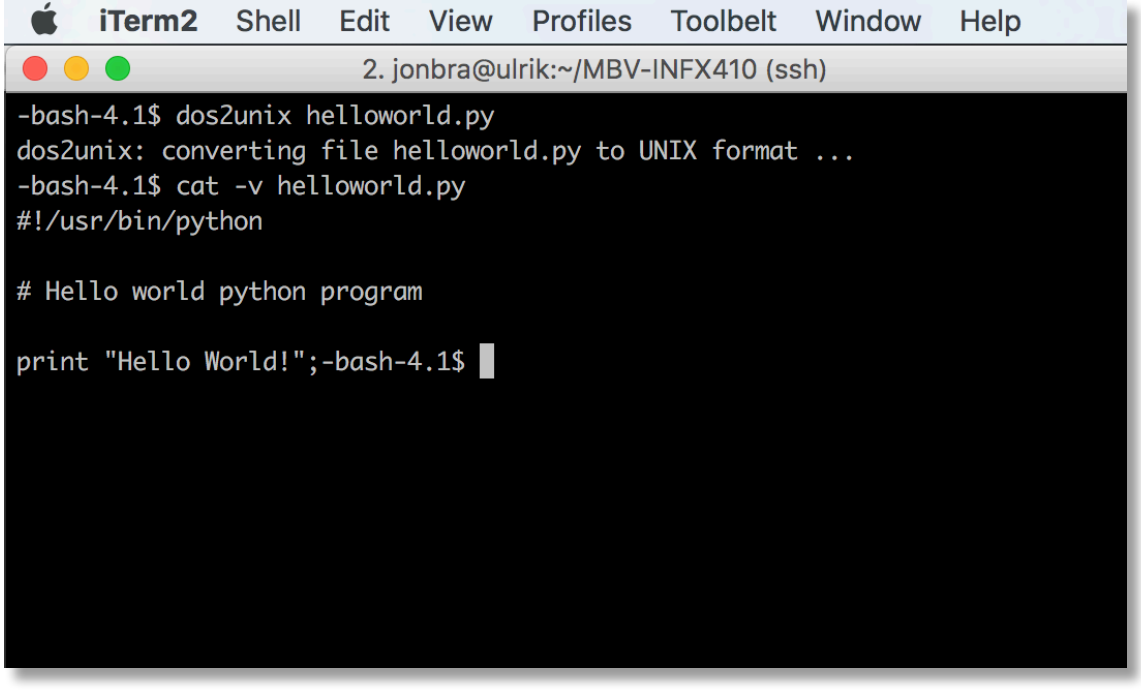
Use `cat -v` to see “invisible characters”. Typically End of Line (EOL) characters

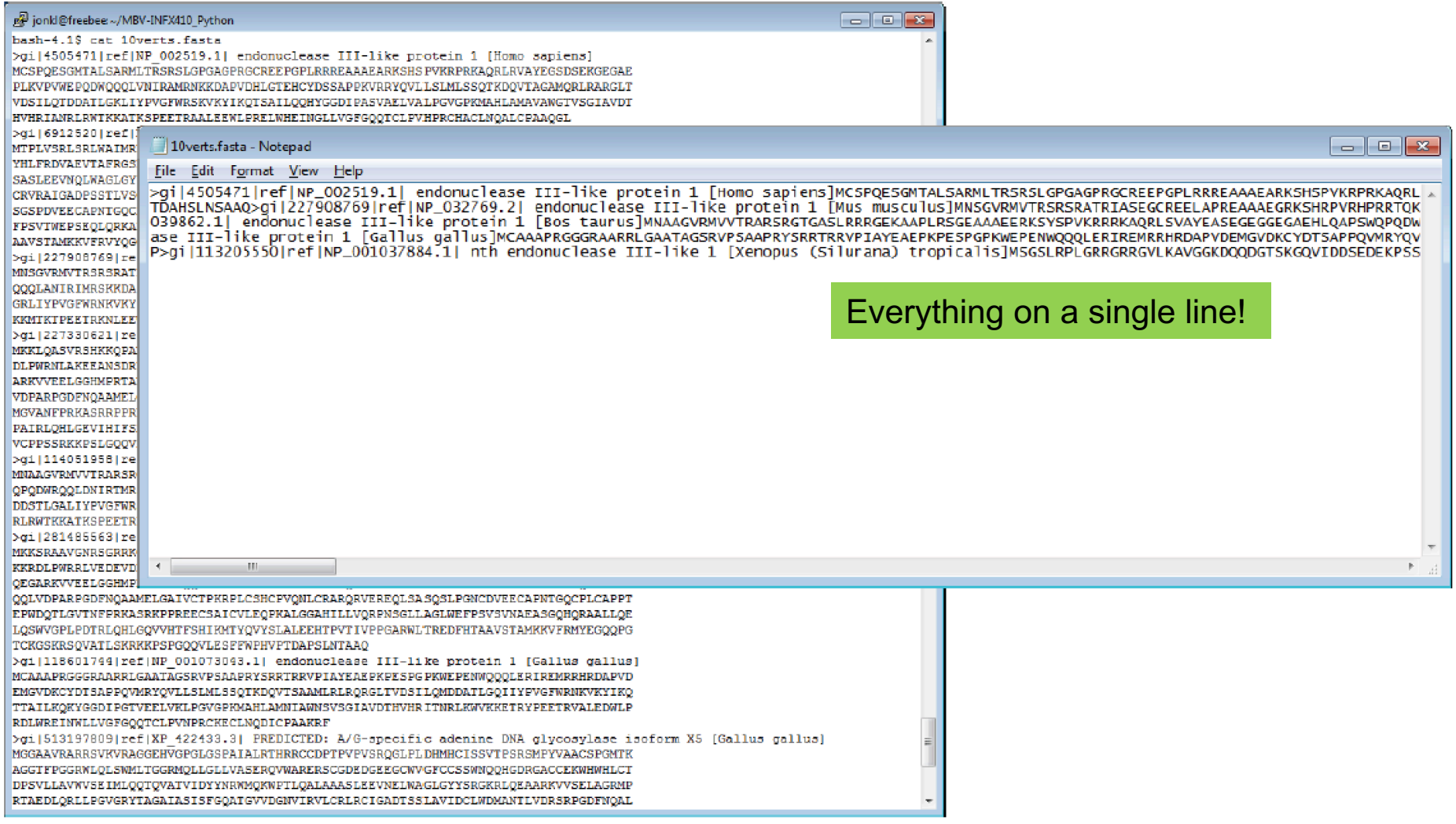


NB! `dos2unix` might have to be installed. Already installed on freebee

# Non-printing characters

EOLs differ between Windows and Unix (including Mac) languages
Use `dos2unix` command to go from MS Windows to Unix
Use `unix2dos` command to do the reverse



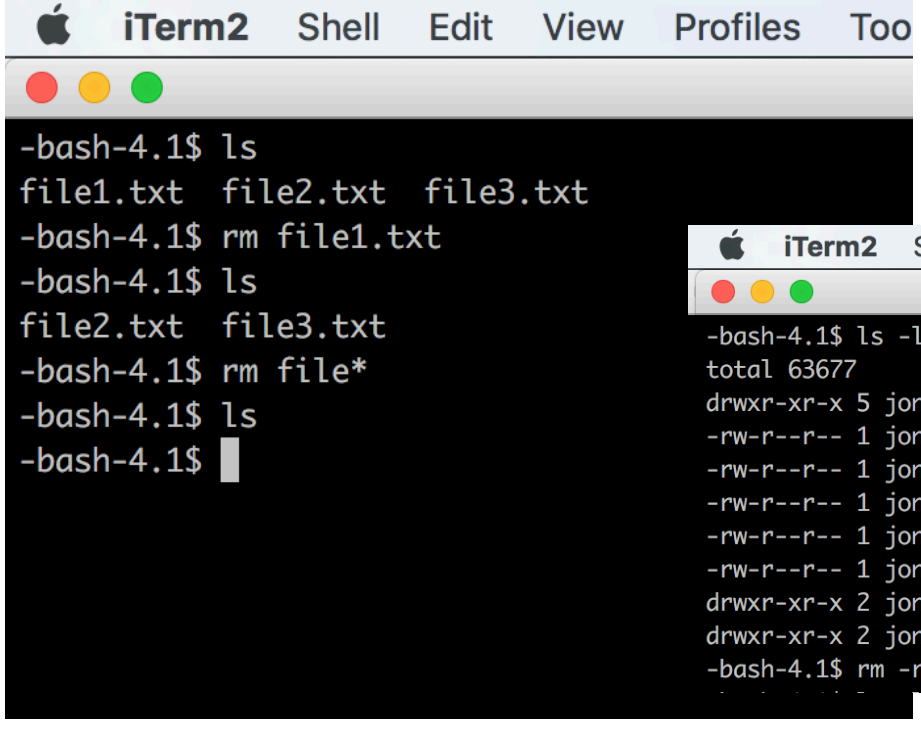

jonkl@freebee ~/MBV-INFX410_Python
bash-4.1$ cat 10verts.fasta
>gi|4505471|ref|NP_002519.1| endonuclease III-like protein 1 [Homo sapiens]
MCSPQESGMTALSARMLTRSRSLGPGAGPRGCREEPGPLRRREAAAEARKSHSPVKRPRKAQRLRVAYEGSDSEKGEGAE
PLKVPVWEPQDWQQQLVNIRAMRNKKDAPVDHLGTEHCYDSSAPPKVRRYQVLLSLMLSSQTKDQVTAGAMQRLRARGLT
VDSILQTDDAILGKLIYPVGFWRSKVKYIKQTSAILQQHYGGDIPASVAELVALPGVGPKMAHLAMAVAWGTVSGIAVDT
HVHRIANRLRWTKKATKSPEETRAALEEWLPRELWHEINGLLVGFGQQTCLPVHPRCHACLNQALCPAAQGL
>gi|6912520|ref|
MTPLVSRLSRLWAIMR
YHLFRDVAEVTAFRGS
SASLEEVNQLWAGLGY
CRVRAIGADPSSTLVS
SGSPDVEECAPNTGQC
FPSVIWEPSEQLQRKA
AAVSTAMKKVFRVYQG
>gi|227908769|re
MNSGVRMVTRSRSRAT
QQQLANIRIHRSKKDA
GRLIYPVGFWRNKVKY
KKMTKIPEEIRKNLEE
>gi|227330621|re
MKKLQASVRSHKKQPA
DLPWRNLAKEEANSDR
ARKVVEELGGHMPRTA
VDPARPGDFNQAAMEL
MGVANFPRKASRRPPR
PAIRLQHLGEVIHIFS
VCPPSSRKKPSLGQQV
>gi|114051958|re
MNAAGVRMVVTRARSR
QPQDWRQQLDNIRTMR
DDSTLGALIYPVGFWR
RLRWTKKATKSPEETR
>gi|281485563|re
MKKSRAAVGNRSGRRK
KKRDLPWRRLVEDEVD
QEGARKVVEELGGHMP
QQLVDPARPGDFNQAAMELGAIVCTPKRPLCSHCPVQNLCRARQRVEREQLSASQSLPGNCDVEECAPNTGQCPLCAPPT
EPWDQTLGVTNFPRKASRRPPREECSAICVLEQPKALGGAHILLVQRPNSGLLAGLWEFPSVSVNAEASGQHQRAALLQE
LQSWVGPLPDTRLQHLGQVVHTFSHIKMTYQVYSLALEEHTPVTIVPPGARWLTREDFHTAAVSTAMKKVFRMYEGQQPG
TCKGSKRSQVAILSKRKKPSPGQQVLESFFWPHVPTDAPSLNTAAQ
>gi|118601744|ref|NP_001073063.1| endonuclease III-like protein 1 [Gallus gallus]
MCAAAPRGGGRAARRLGAATAGSRVPSAAPRYSRRTRRVPIAYEAEPKPESPGPKWEPENWQQQLERIREMRRHRDAPVD
EMGVDKCYDTSAPPQVMRYQVLLSLMLSSQTKDQVTSAAMLRLRQRGLTVDSILQMDDATLGQIIYFVGFWRNKVKYIKQ
TTAILKQKYGGDIPGTVEELVKLPGVGPKMAHLAMNIAWNSVSGIAVDTHVHRITNRLKWVKKETRYPEETRVALEDWLP
RDLWREINWLLVGFGQQTCLPVNPRCKECLNQDICPAAKRF
>gi|513197809|ref|XP_422433.3| PREDICTED: A/G-specific adenine DNA glycosylase isoform X5 [Gallus gallus]
MGGAAVRARRSVKVRAGGEHVGPGLGSPAIALRTHRRCCDPTPVPVSRQGLPLDHMHCISSVTPSRSMPYVAACSPGMTK
AGGTFPGGRWLQLSWMLTGGRMQLLGLLVASERQVWARERSGGDEDGEEGCWVGFCCSSWNQQHGDRGACCEKWHWHLGI
DPSVLLAVWVSEIMLQQTQVATVIDYYNRWMQKWPTLQALAAASLEEVNELWAGLGYYSRGKRLQEAARKVVSELAGRMP
RTAEDLQRLLPGVGRYTAGAIASISFGQAIGVVDGNVIRVLCRLRCIGADTSSLAVIDCLWDMANTLVDRSRPGDFNQAL
10verts.fasta - Notepad
File Edit Format View Help
>gi|4505471|ref|NP_002519.1| endonuclease III-like protein 1 [Homo sapiens]MCSPQESGMTALSARMLTRSRSLGPGAGPRGCREEPGPLRRREAAAEARKSHSPVKRPRKAQRL
TDAHSLNSAAQ>gi|227908769|ref|NP_032769.2| endonuclease III-like protein 1 [Mus musculus]MNSGVRMVTRSRSRATRIASEGCREELAPREAAAEGRKSHRPVRHPRRTQK
039862.1| endonuclease III-like protein 1 [Bos taurus]MNAAGVRMVVTRARSRGTGASLRRRGEKAAPLRSGEAAAEERKSYSPVKRRRKAQRLSVAYEASEGEGGEGAEHLQAPSWQPQDW
ase III-like protein 1 [Gallus gallus]MCAAAPRGGGRAARRLGAATAGSRVPSAAPRYSRRTRRVPIAYEAEPKPESPGPKWEPENWQQQLERIREMRRHRDAPVDEMGVDKCYDTSAPPQVMRYQV
P>gi|113205550|ref|NP_001037884.1| nth endonuclease III-like 1 [Xenopus (Silurana) tropicalis]MSGSLRPLGRRGRRGVLKAVGGKDQQDGTSKGQVIDDSEDEKPSS
Everything on a single line!

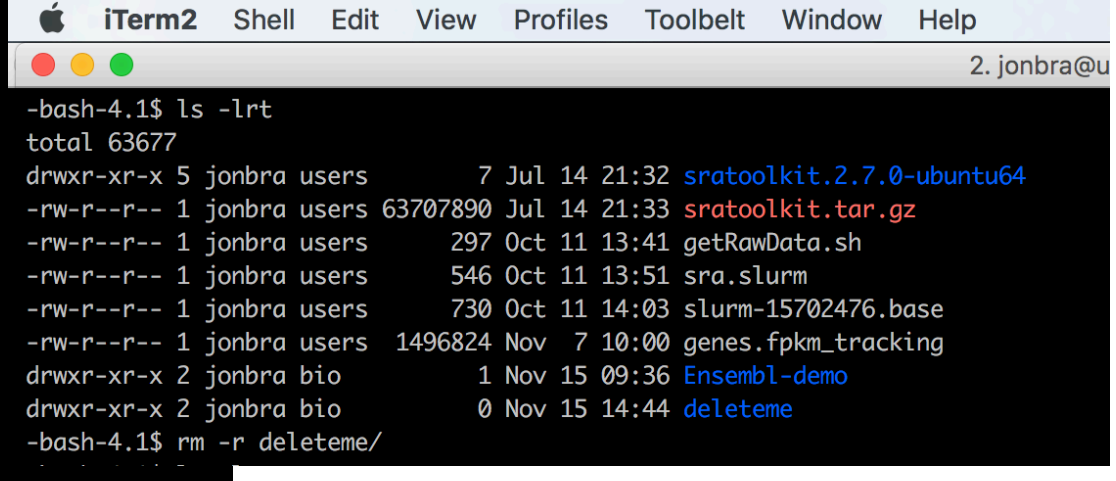
NB! There's no trash can on Unix!
Your files are immediately deleted!

# Delete files with the `rm` command

Delete empty directories with `rm –r`

Delete directories AND
files within with `rm –rf` Use with caution!!

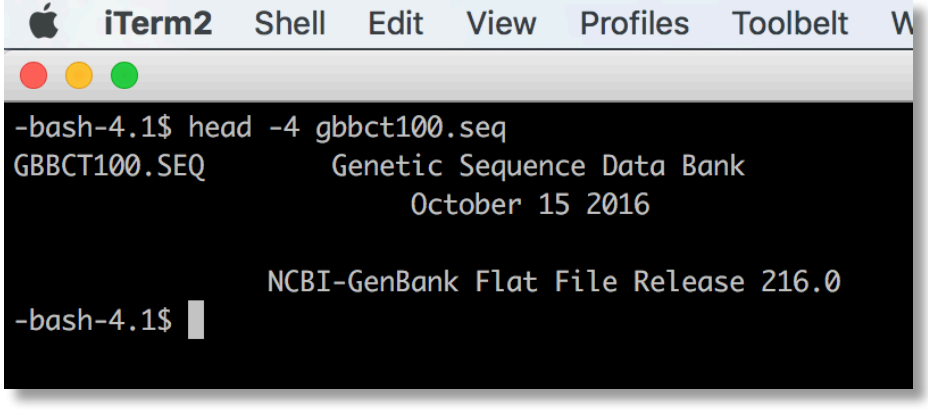


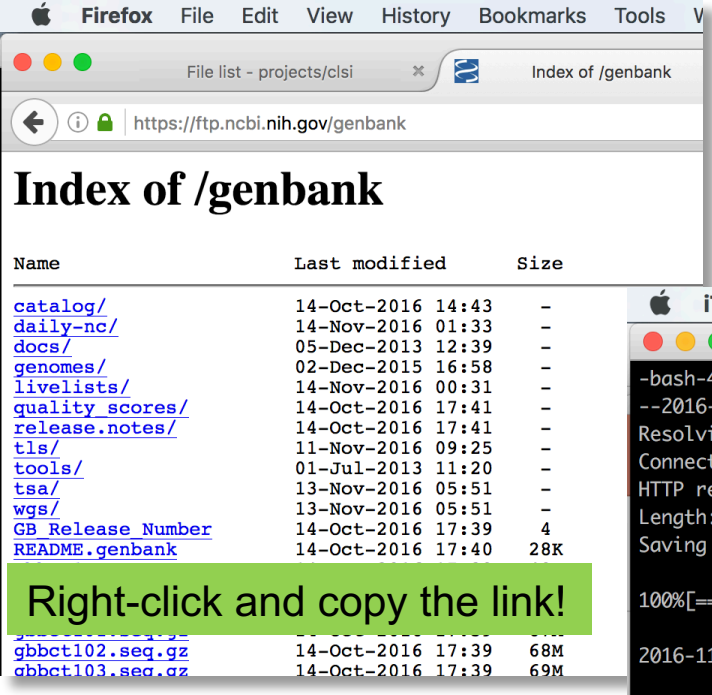
# Viewing files

Pro tip:
Use the different
options to tweak the output

- `cat` displays the entire content on the screen
- `less` displays the content on the screen one file/line at a time (press ”q” to exit).
- `head/tail` displays the first or last lines of a file



# Downloading files from the internet




The `wget` command should work on freebee
Other options are `curl` or `ftp`

# Back up your data!

- Don’t store only on your laptop
- Use M: on UiO machines
- Same as ~/ in Unix environment
- ~/pc/Desktop on login or freebee is the same as M: on UiO machines