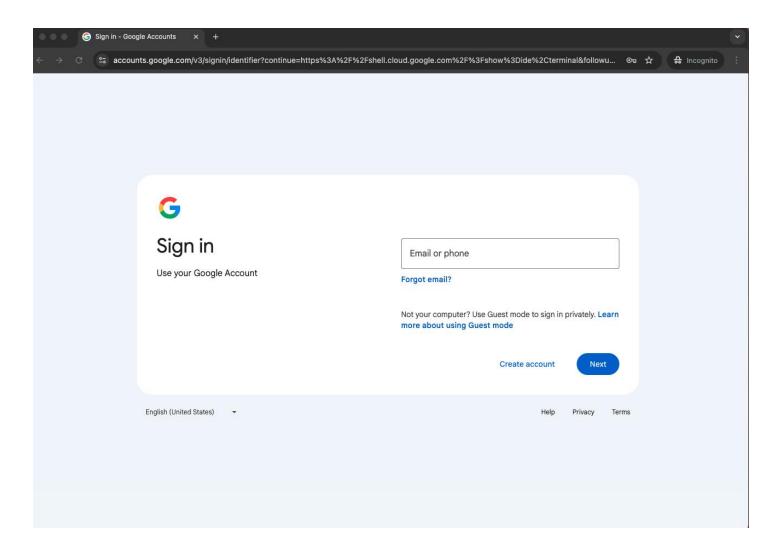
Sommersemester 2025 Vertiefte Bestimmungsübungen an Tieren (MEES003/C3)

Environmental DNA (eDNA) Metabarcoding analysis

Day 0

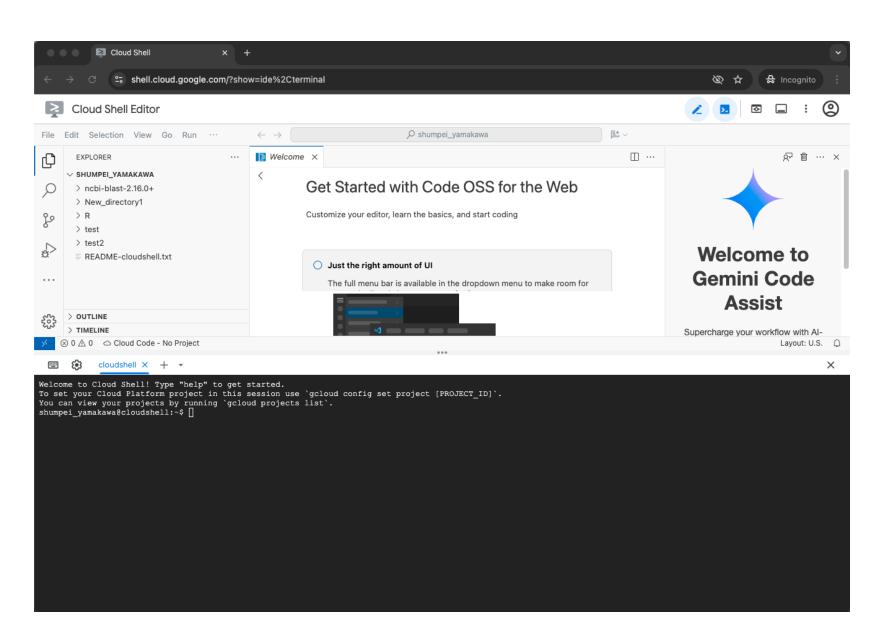
Go to https://shell.cloud.google.com/?show=ide%2Cterminal



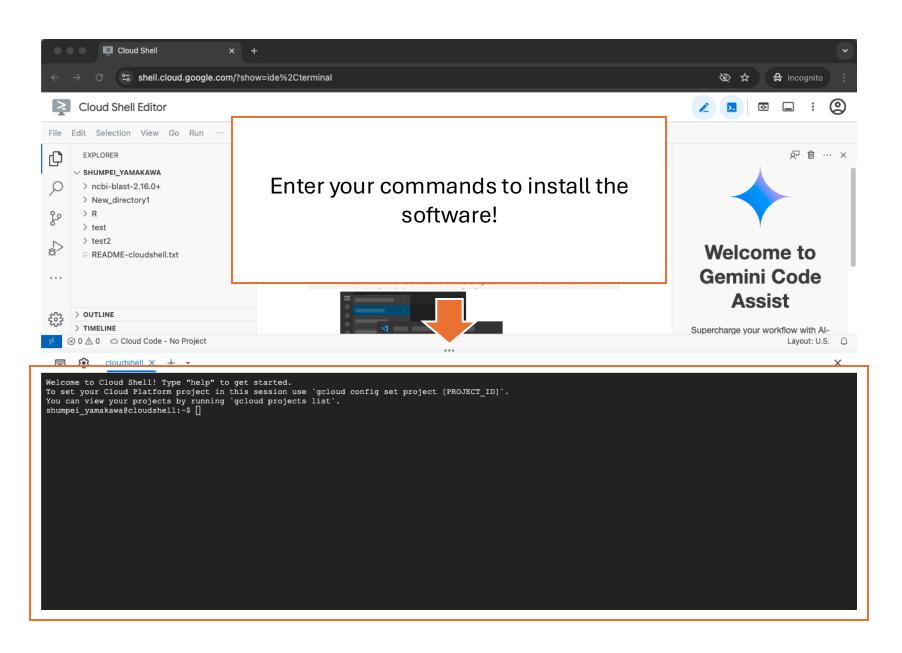
Sign in if you have an account

If you don't want to create an account, set up the environment on your PC (see Day 0.md [Optional]).

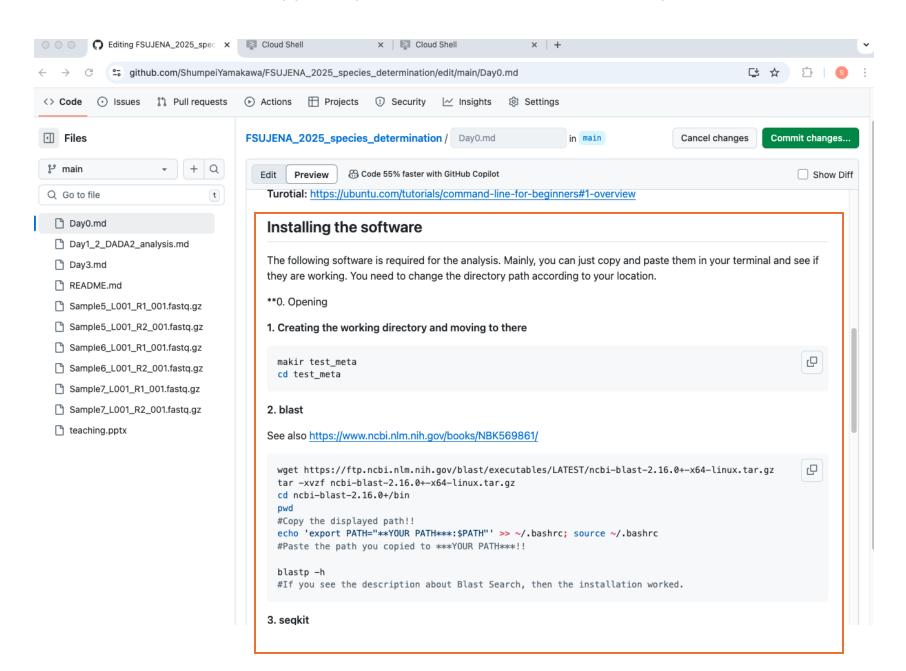
The following cloud shell will appear:



The following cloud shell will appear:



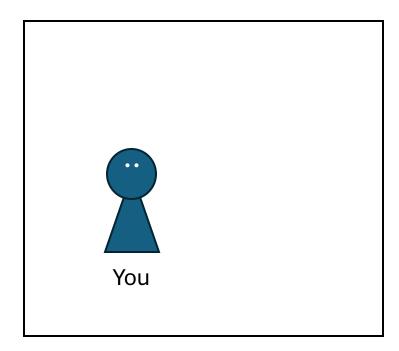
Copy and paste the commands from Day0.md.



1. Creating the working directory and moving to there



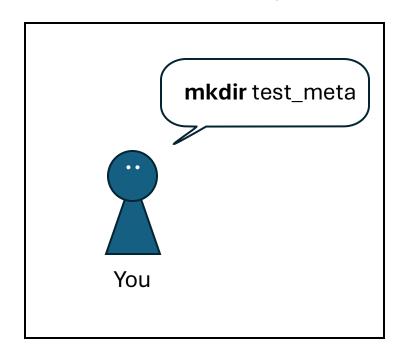
Home directory



First, you are at "home directory"

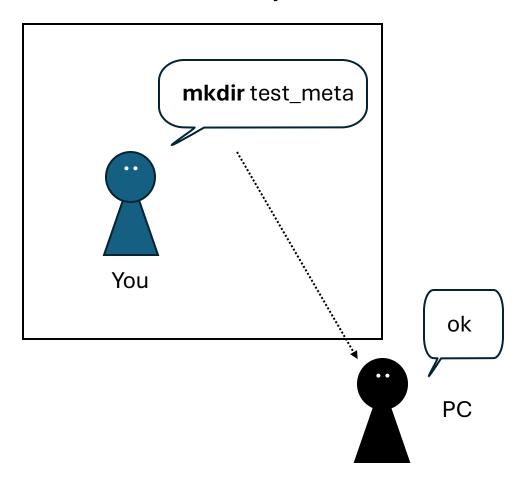
(If you wan to know the current your location/path, type "pdw")

Home directory

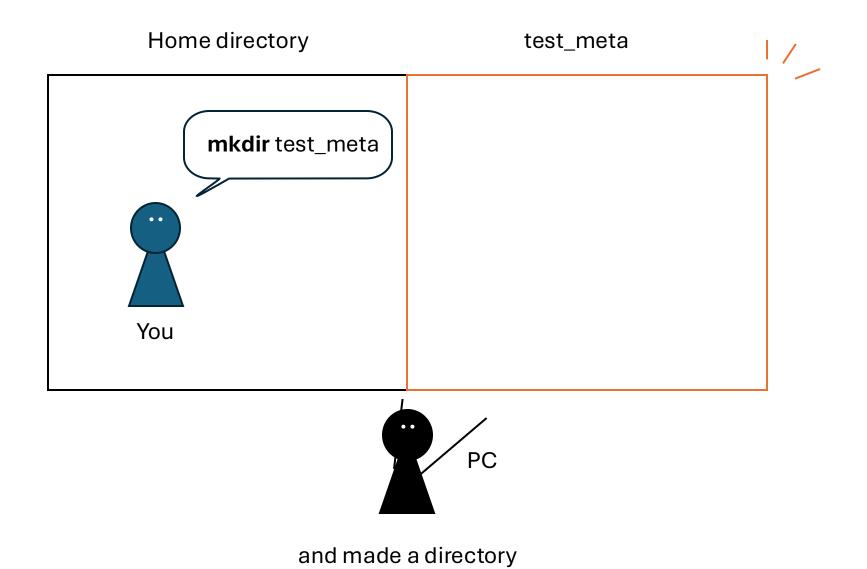


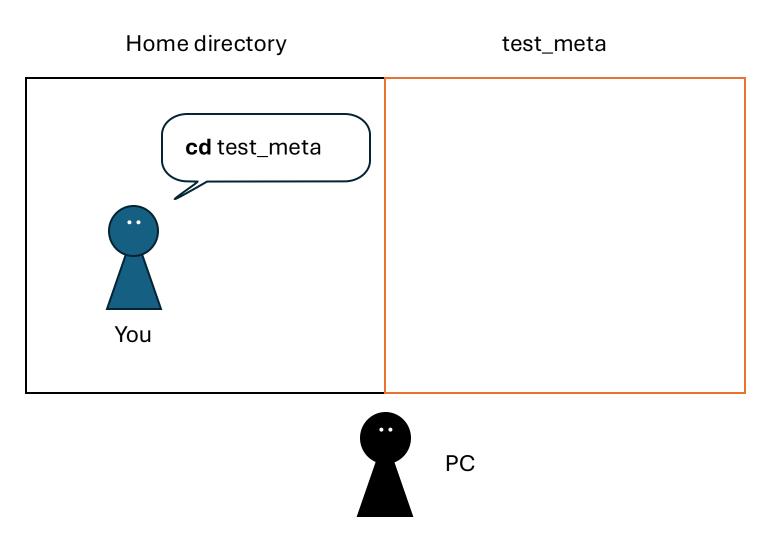
Then, you typed "mkdir test_meta" mkdir is a command to make a directory

Home directory

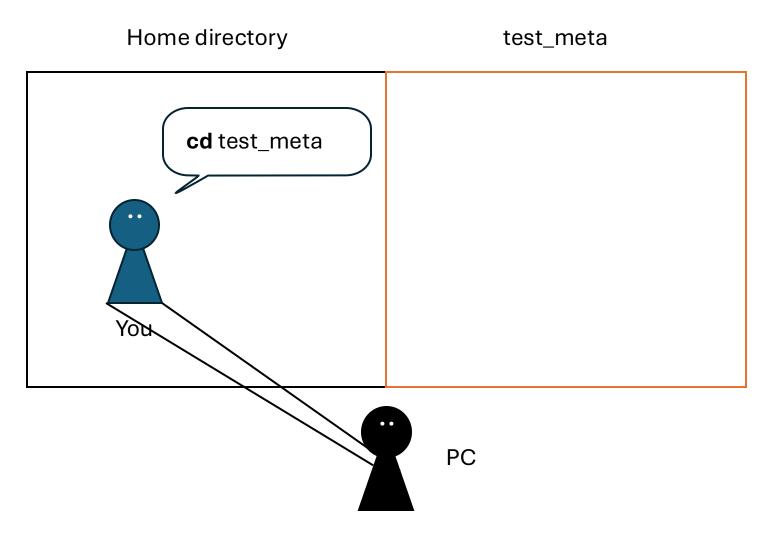


PC recognizes this command

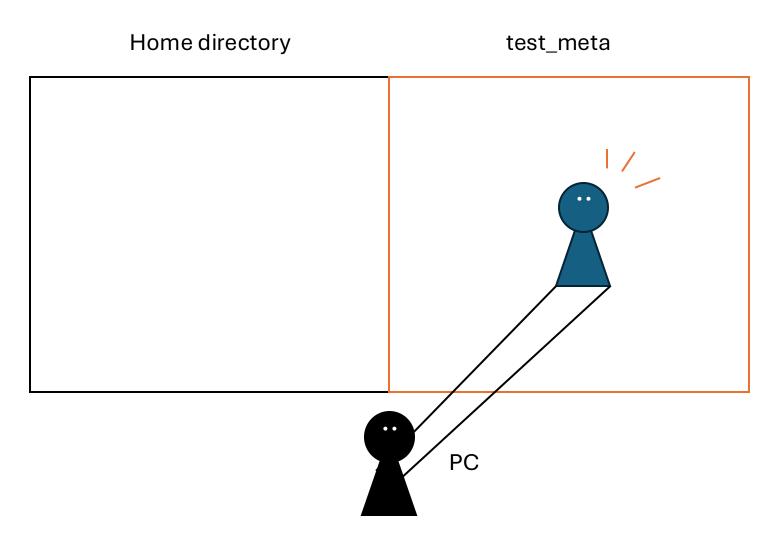




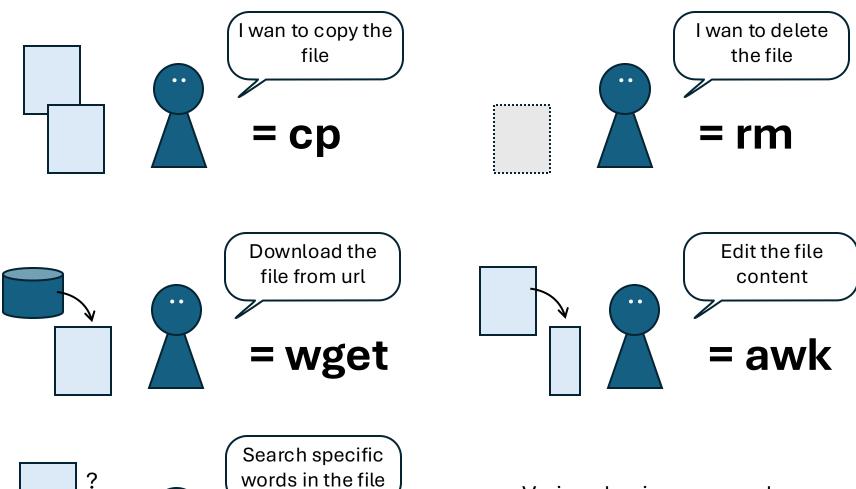
"cd" is "change directory"



"cd" is "change directory"



"cd" is "change directory"



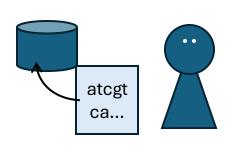
= grep

"abc"

Various basic commands are already prepared

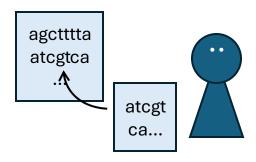
See the official tutorial:

https://ubuntu.com/tutorials/command-line-forbeginners#1-overview



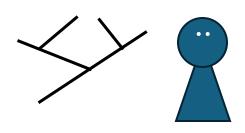
I want to make a Blast database

= makeblastdb



Search specific sequences

= seqkit grep



Make a phylogenic tree

= raxml

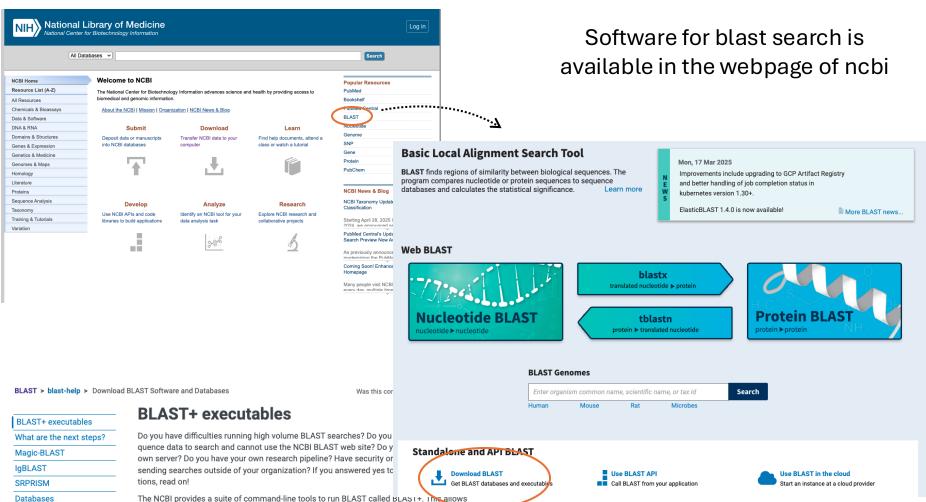
Need to download specific software for bioinformatic analysis

2. blast

See also https://www.ncbi.nlm.nih.gov/books/NBK569861/

```
wget https://ftp.ncbi.nlm.nih.gov/blast/executables/LATEST/ncbi-blast-2.16.0+-x64-linux.tar.gz
tar -xvzf ncbi-blast-2.16.0+-x64-linux.tar.gz
cd ncbi-blast-2.16.0+/bin
pwd
#Copy the displayed path!!
echo 'export PATH="**YOUR PATH***:$PATH"' >> ~/.bashrc; source ~/.bashrc
#Paste the path you copied to ***YOUR PATH***!!

blastp -h
#If you see the description about Blast Search, then the installation worked.
```



The NCBI provides a suite of command-line tools to run BLAST called BLAST+. This allows users to perform BLAST searches on their own server without size, volume and database restrictions. BLAST+ can be used with a command line so it can be integrated directly into your workflow.

What are the next steps?

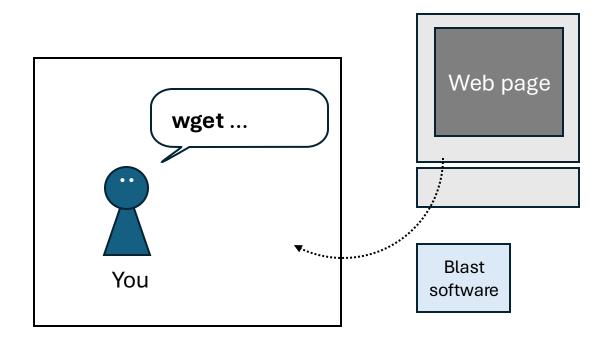
Download and install BLAST+. Installers and source code are available from https://ft-p.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/. Download the databases you need, (see database section below), or create your own. Start searching.

For more details, please see the BLAST+ user manual, the BLAST Help manual, the BLAST releases notes, and the article in BMC Bioinformatics (PubMed link). See our versioning policy.

The BLAST+ suite is the currently supported package. The older C toolkit executables are no longer supported. See our versioning policy.

We are always listening and welcome your feedback at BLAST Support Center.







Type "ls" (see the contents in the current directory)

```
shumpei_yamakawa@cloudshell:~/test_meta$ ls
ncbi-blast-2.16.0+-x64-linux.tar.gz
```



Type "ls" (see the contents in the current directory)





Type "ls" (see the contents in the current directory)

```
shumpei_yamakawa@cloudshell:~/test_meta$ ls
ncbi-blast-2.16.0+-x64-linux.tar.gz

Compressed archive file
```

tar -xvzf ncbi-blast-2.16.0+-x64-linux.tar.gz

Command for decompressing

```
ncbi-blast-2.16.0+-x64-linux.tar.gz
shumpei yamakawa@cloudshell:~/test meta$ tar -xvzf ncbi-blast-2.16.0+-x64-linux.tar.qz
ncbi-blast-2.16.0+/
ncbi-blast-2.16.0+/ChangeLog
ncbi-blast-2.16.0+/bin/
ncbi-blast-2.16.0+/bin/rpsblast
ncbi-blast-2.16.0+/bin/blastn vdb
ncbi-blast-2.16.0+/bin/blast vdb cmd
ncbi-blast-2.16.0+/bin/makeprofiledb
ncbi-blast-2.16.0+/bin/blastp
ncbi-blast-2.16.0+/bin/deltablast
ncbi-blast-2.16.0+/bin/blastx
ncbi-blast-2.16.0+/bin/rpstblastn
ncbi-blast-2.16.0+/bin/psiblast
ncbi-blast-2.16.0+/bin/tblastx
ncbi-blast-2.16.0+/bin/tblastn vdb
ncbi-blast-2.16.0+/bin/blast formatter vdb
ncbi-blast-2.16.0+/bin/segmasker
ncbi-blast-2.16.0+/bin/convert2blastmask
ncbi-blast-2.16.0+/bin/blastdbcmd
ncbi-blast-2.16.0+/bin/get species taxids.sh
ncbi-blast-2.16.0+/bin/dustmasker
ncbi-blast-2.16.0+/bin/cleanup-blastdb-volumes.py
ncbi-blast-2.16.0+/bin/blastn
ncbi-blast-2.16.0+/bin/windowmasker
ncbi-blast-2.16.0+/bin/makeblastdb
ncbi-blast-2.16.0+/bin/blastdb aliastool
ncbi-blast-2.16.0+/bin/update blastdb.pl
ncbi-blast-2.16.0+/bin/makembindex
ncbi-blast-2.16.0+/bin/blast formatter
ncbi-blast-2.16.0+/bin/legacy blast.pl
ncbi-blast-2.16.0+/bin/blastdbcheck
ncbi-blast-2.16.0+/bin/tblastn
ncbi-blast-2.16.0+/LICENSE
ncbi-blast-2.16.0+/BLAST PRIVACY
ncbi-blast-2.16.0+/ncbi package info
ncbi-blast-2.16.0+/README
ncbi-blast-2.16.0+/doc/
ncbi-blast-2.16.0+/doc/README.txt
shumpei yamakawa@cloudshell:~/test meta$ ls
ncbi-blast-2.16.0+ ncbi-blast-2.16.0+-x64-linux.tar.gz
shumpei yamakawa@cloudshell:~/test meta$ [
```

```
Path set*
  cd ncbi-blast-2.16.0+/bin
2 pwd
  #Copy the displayed path!!
(3) echo 'export PATH="**YOUR PATH***:$PATH"' >> ~/.bashrc; source ~/.bashrc
  #Paste the path you copied to ***YOUR PATH***!!
  shumpei yamakawa@cloudshell:~/test meta$ cd ncbi-blast-2.16.0+/bin
  2
```

shumpei yamakawa@cloudshell:~/test meta/ncbi-blast-2.16.0+/bin\$ pwd

^{*} If you want to know more about path setting, see the attached slides

```
Path set
① cd ncbi-blast-2.16.0+/bin
2 pwd
  #Copy the displayed path!!
(3) echo 'export PATH="**YOUR PATH***:$PATH"' >> ~/.bashrc; source ~/.bashrc
  #Paste the path you copied to ***YOUR PATH***!!
  shumpei yamakawa@cloudshell:~/test meta$ cd ncbi-blast-2.16.0+/bin
  2
  shumpei yamakawa@cloudshell:~/test meta/ncbi-blast-2.16.0+/bin$ pwd
  /home/shumpei yamakawa/test meta/ncbi-blast-2.16.0+/bin
```

YOUR PATH

```
Path set
① cd ncbi-blast-2.16.0+/bin
(2) pwd
  #Copy the displayed path!!
(3) echo 'export PATH="**YOUR PATH***: $PATH"' >> ~/.bashrc; source ~/.bashrc
  #Paste the path you copied to ***YOUR PATH***!!
  (1)
  shumpei yamakawa@cloudshell:~/test meta$ cd ncbi-blast-2.16.0+/bin
  2
  shumpei yamakawa@cloudshell:~/test meta/ncbi-blast-2.16.0+/bin$ pwd
  /home/shumpei yamakawa/test meta/ncbi-blast-2.16.0+/bin
```

echo 'export PATH="**YOUR PATH***: \$PATH"' >> ~/.bashrc; source ~/.bashrc

```
Path set
  cd ncbi-blast-2.16.0+/bin
(2) pwd
  #Copy the displayed path!!
(3) echo 'export PATH="**YOUR PATH***: $PATH"' >> ~/.bashrc; source ~/.bashrc
  #Paste the path you copied to ***YOUR PATH***!!
  shumpei yamakawa@cloudshell:~/test meta$ cd ncbi-blast-2.16.0+/bin
  (2)
  shumpei yamakawa@cloudshell:~/test meta/ncbi-blast-2.16.0+/bin$ pwd
  /home/shumpei yamakawa/test meta/ncbi-blast-2.16.0+/bin
  echo 'export PATH="**YOUR PATH***: $PATH"' >> ~/.bashrc; source ~/.bashrc
                                  Substitute your path and remove ***
                       /home/shu.../bin :$PATH"' >> ~/.bashrc; source ~/.bashrc
   echo 'export PATH="
```

```
Path set
  cd ncbi-blast-2.16.0+/bin
2 pwd
  #Copy the displayed path!!
(3) echo 'export PATH="**YOUR PATH***:$PATH"' >> ~/.bashrc; source ~/.bashrc
  #Paste the path you copied to ***YOUR PATH***!!
  shumpei yamakawa@cloudshell:~/test meta$ cd ncbi-blast-2.16.0+/bin
  (2)
  shumpei yamakawa@cloudshell:~/test meta/ncbi-blast-2.16.0+/bin$ pwd
  /home/shumpei yamakawa/test meta/ncbi-blast-2.16.0+/bin
```

export PATH="/home/shumpei yamakawa/test meta/ncbi-blast-2.16.0+/bin \$PATH"' >> ~/.bashrc; source ~/.bashr

#If you see the description about Blast Search, then the installation worked.

```
shumpei yamakawa@cloudshell:~/test meta/ncbi-blast-2.16.0+/bin$ blastp -h
USAGE
 blastp [-h] [-help] [-import search strategy filename]
    [-export search strategy filename] [-task task name] [-db database name]
    [-dbsize num letters] [-gilist filename] [-segidlist filename]
    [-negative gilist filename] [-negative segidlist filename]
    [-taxids taxids] [-negative taxids taxids] [-taxidlist filename]
    [-negative taxidlist filename] [-no taxid expansion] [-ipglist filename]
    [-negative ipglist filename] [-entrez query entrez query]
    [-db soft mask filtering algorithm] [-db hard mask filtering algorithm]
    [-subject subject input file] [-subject loc range] [-query input file]
    [-out output file] [-evalue evalue] [-word size int value]
    [-gapopen open penalty] [-gapextend extend penalty]
    [-qcov hsp perc float value] [-max hsps int value]
    [-xdrop ungap float value] [-xdrop gap float value]
    [-xdrop gap final float value] [-searchsp int value] [-seg SEG options]
    [-soft masking soft masking] [-matrix matrix name]
    [-threshold float value] [-culling limit int value]
    [-best hit overhang float value] [-best hit score edge float value]
    [-subject besthit] [-window size int value] [-lcase masking]
    [-query loc range] [-parse deflines] [-outfmt format] [-show gis]
    [-num descriptions int value] [-num alignments int value]
    [-line length line length] [-html] [-sorthits sort hits]
    [-sorthsps sort hsps] [-max target segs num sequences]
    [-num threads int value] [-mt mode int value] [-ungapped] [-remote]
    [-comp based stats compo] [-use sw tback] [-version]
DESCRIPTION
   Protein-Protein BLAST 2.16.0+
Use '-help' to print detailed descriptions of command line arguments
shumpei yamakawa@cloudshell:~/test meta/ncbi-blast-2.16.0+/bin$
```

3. segkit

The software for manipulating the sequence files. See also https://bioinf.shenwei.me/seqkit/

```
cd ~/test_meta
#move to the test directory
wget https://github.com/shenwei356/seqkit/releases/download/v2.10.0/seqkit_linux_amd64.tar.gz
tar -xvzf seqkit_linux_amd64.tar.gz
mkdir seqkit_dir
mv seqkit seqkit_dir/
cd seqkit_dir
pwd
#Copy the displayed path!!
echo 'export PATH="**YOUR PATH***: $PATH"' >> ~/.bashrc; source ~/.bashrc
#Paste the path you copied to ***YOUR PATH***!!
segkit -h
#If you see the description about segkit Search, then the installation worked.
cd ~/
#Go back to the home directory
                                                             Repeat the steps above
```

4. update/developer environments

```
Q
sudo apt update
sudo apt install −y \
    liblapack-dev \
    libblas-dev \
   gfortran \
    libpng-dev \
   libgsl-dev \
   libtiff5-dev \
   libjpeg-dev \
   libxml2-dev \
    libcurl4-openssl-dev \
    libxt-dev \
    libfftw3-dev \
    libreadline-dev \
   zlib1g-dev \
   libx11-dev \
    libncurses-dev \
    libbz2-dev \
    liblzma-dev
```

R is a programming language and software environment for statistical computing, data analysis, and visualization. See also https://www.r-project.org/

```
#move to the test directory

# update indices
sudo apt update -qq
# install two helper packages we need
sudo apt install --no-install-recommends software-properties-common dirmngr
# add the signing key (by Michael Rutter) for these repos
# To verify key, run gpg --show-keys /etc/apt/trusted.gpg.d/cran_ubuntu_key.asc
# Fingerprint: E298A3A825C0D65DFD57CBB651716619E084DAB9
wget -q0- https://cloud.r-project.org/bin/linux/ubuntu/marutter_pubkey.asc | sudo tee -a /etc/apt/trusted.gpg.d/cran_
# add the repo from CRAN -- lsb_release adjusts to 'noble' or 'jammy' or ... as needed
sudo add-apt-repository "deb https://cloud.r-project.org/bin/linux/ubuntu $(lsb_release -cs)-cran40/"
# install R itself
sudo apt install --no-install-recommends r-base
```

During installing, you may see the following...

Enter -> go to "R console"

```
shumpei_yamakawa@cloudshell:~$ R

R version 4.5.0 (2025-04-11) -- "How About a Twenty-Six"
Copyright (C) 2025 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> []
```

Start installing R packages

6. dada2: the R package for metabarcoding analysis

R provides a lot of the "packages" which are designed for specific analysis, and the package "dada2" performs to "infer exact amplicon sequence variants (ASVs) from amplicon data (https://benjineb.github.io/dada2/)". This package needs to be installed in R console.

```
# install the softwares
install.packages(c("RcppEigen", "png", "deldir", "interp", "latticeExtra"))
if (!requireNamespace("BiocManager", quietly = TRUE))
    install.packages("BiocManager")
BiocManager::install("ShortRead")
BiocManager::install("dada2")
library(dada2)
packageVersion("dada2")
#Can you see the version of dada2? If so, dada2 was correctly installed!
```

shumpei_yamakawa@cloudshell:~\$ R

Enter -> go to "R console"

```
shumpei_yamakawa@cloudshell:~$ R

R version 4.5.0 (2025-04-11) -- "How About a Twenty-Six"
Copyright (C) 2025 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> []
```

Start installing R packages

6. dada2: the R package for metabarcoding analysis

R provides a lot of the "packages" which are designed for specific analysis, and the packag sequence variants (ASVs) from amplicon data (https://benjjneb.github.io/dada2/)". This packages

During installation, you may be prompted to select "Yes" or "No."

Choose "Yes."

```
# install the softwares
install.packages(c("RcppEigen", "png", "deldir", "interp", "latti
if (!requireNamespace("BiocManager", quietly = TRUE))
    install.packages("BiocManager")
BiocManager::install("ShortRead")
BiocManager::install("dada2")
library(dada2)
packageVersion("dada2")
#Can you see the version of dada2? If so, dada2 was correctly installed!
```

```
shumpei yamakawa@cloudshell:~$ R
R version 4.5.0 (2025-04-11) -- "How About a Twenty-Six"
Copyright (C) 2025 The R Foundation for Statistical Computing
Platform: x86 64-pc-linux-gnu
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
  Natural language support but running in an English locale
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
> q()
Save workspace image? [y/n/c]: n
shumpei yamakawa@cloudshell:~$
```

Type "q()" to escape from R console

Appendix

Public tutorial

https://ubuntu.com/tutorials/command-line-for-beginners#1-overview

https://opensource.com/article/17/6/set-path-linux

Illustrated guides about basic usage of linux

- 1. Location and directory structure
- 2. Installing software

1. Location and directory structure

```
Welcome to Cloud Shell! Type "help" to get started.

To set your Cloud Platform project in this session use `gcloud config set project [PROJECT_ID]`.

You can view your projects by running `gcloud projects list`.

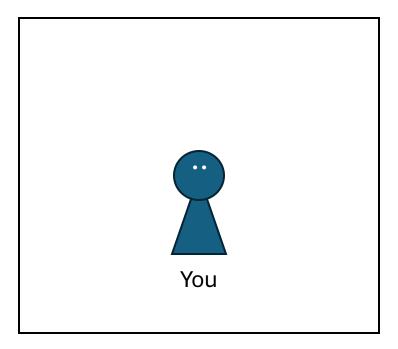
Your home disk usage is at 95%.

You can find suggestions to clear space at https://cloud.google.com/shell/docs/quotas-limits#clearing_disk_space.

Upgrade to Cloud Workstations for a secure development environment without these limits: https://cloud.google.com/workstations.

shumpei_yamakawa@cloudshell:~$ []
```

Home directory



"type pwd -> enter"

```
Welcome to Cloud Shell! Type "help" to get started.

To set your Cloud Platform project in this session use `gcloud config set project [PROJECT_ID]`.

You can view your projects by running `gcloud projects list`.

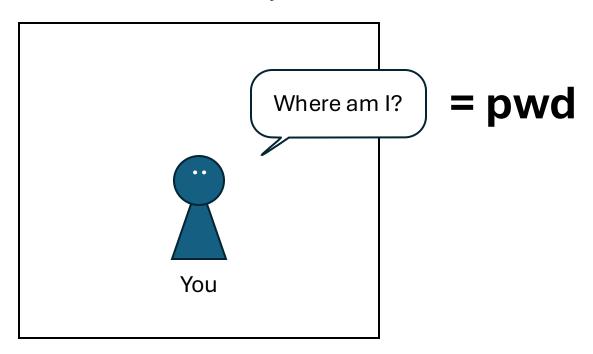
Your home disk usage is at 95%.

You can find suggestions to clear space at https://cloud.google.com/shell/docs/quotas-limits#clearing_disk_space.

Upgrade to Cloud Workstations for a secure development environment without these limits: https://cloud.google.com/workstations.

shumpei_yamakawa@cloudshell:~$ pwd]
```

Home directory



```
Welcome to Cloud Shell! Type "help" to get started.

To set your Cloud Platform project in this session use `gcloud config set project [PROJECT_ID]`.

You can view your projects by running `gcloud projects list`.

Your home disk usage is at 95%.

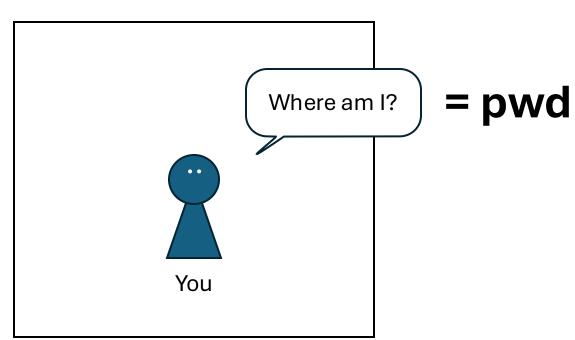
You can find suggestions to clear space at https://cloud.google.com/shell/docs/quotas-limits#clearing_disk_space.

Upgrade to Cloud Workstations for a secure development environment without these limits: https://cloud.google.com/workstations.

shumpei_yamakawa@cloudshell:~$ pwd
/home/shumpei_yamakawa
Shumpei_vamakawa@cloudshell:~$ [
```

your current address (PATH)

Home directory



```
Welcome to Cloud Shell! Type "help" to get started.

To set your Cloud Platform project in this session use `gcloud config set project [PROJECT_ID]`.

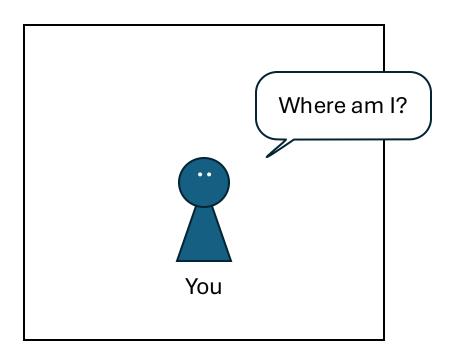
You can view your projects by running `gcloud projects list`.

Your home disk usage is at 95%.

You can find suggestions to clear space at https://cloud.google.com/shell/docs/quotas-limits#clearing_disk_space.

Upgrade to Cloud Workstations for a secure development environment without these limits: https://cloud.google.com/workstations.

shumpei_yamakawa@cloudshell:~$ pwd
/home/shumpei_yamakawa
shumpei_yamakawa@cloudshell:~$ []
```



```
Welcome to Cloud Shell! Type "help" to get started.

To set your Cloud Platform project in this session use `gcloud config set project [PROJECT_ID]`.

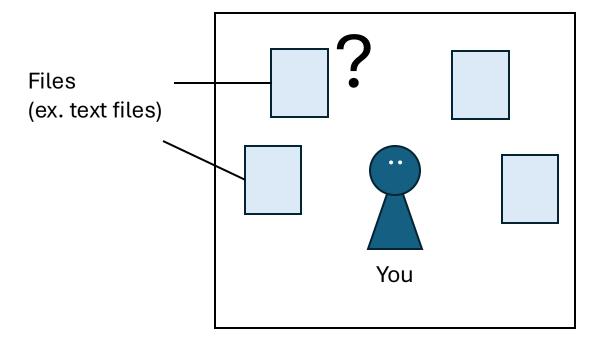
You can view your projects by running `gcloud projects list`.

Your home disk usage is at 95%.

You can find suggestions to clear space at https://cloud.google.com/shell/docs/quotas-limits#clearing_disk_space.

Upgrade to Cloud Workstations for a secure development environment without these limits: https://cloud.google.com/workstations.

shumpei_yamakawa@cloudshell:~$ pwd
/home/shumpei_yamakawa@cloudshell:~$ []
```



"type ls -> enter"

```
Welcome to Cloud Shell! Type "help" to get started.

To set your Cloud Platform project in this session use `gcloud config set project [PROJECT_ID]`.

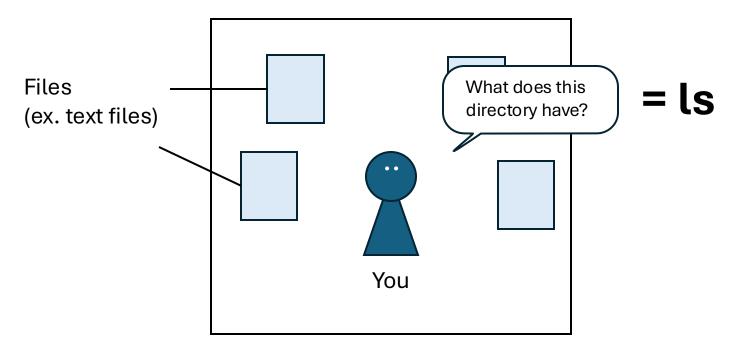
You can view your projects by running `gcloud projects list`.

Your home disk usage is at 95%.

You can find suggestions to clear space at https://cloud.google.com/shell/docs/quotas-limits#clearing_disk_space.

Upgrade to Cloud Workstations for a secure development environment without these limits: https://cloud.google.com/workstations.

shumpei_yamakawa@cloudshell:~$ pwd
/home/shumpei_yamakawa
shumpei_yamakawa@cloudshell:~$ ls[]
```



```
Welcome to Cloud Shell! Type "help" to get started.

To set your Cloud Platform project in this session use `gcloud config set project [PROJECT_ID]`.

You can view your projects by running `gcloud projects list`.

Your home disk usage is at 95%.

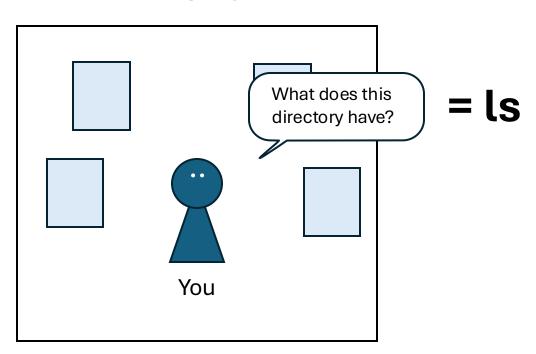
You can find suggestions to clear space at https://cloud.google.com/shell/docs/quotas-limits#clearing_disk_space.

Upgrade to Cloud Workstations for a secure development environment without these limits: https://cloud.google.com/workstations.

shumpei_yamakawa@cloudshell:~$ pwd
/home/shumpei_yamakawa
shumpei_yamakawa@cloudshell:~$ ls

README-cloudshell.txt
shumpei_yamakawa@cloudshell:~$ |
```

You have "README-cloudshell.txt"



```
Welcome to Cloud Shell! Type "help" to get started.

To set your Cloud Platform project in this session use `gcloud config set project [PROJECT_ID]`.

You can view your projects by running `gcloud projects list`.

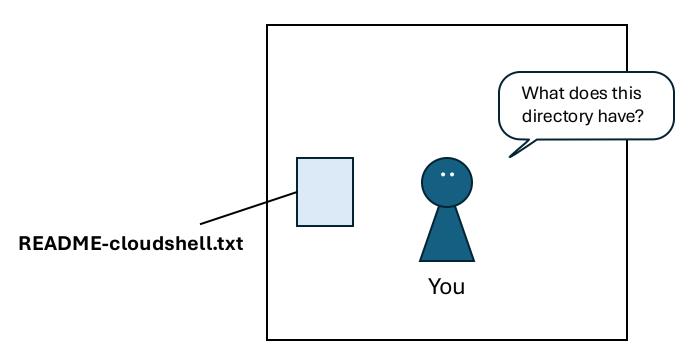
Your home disk usage is at 95%.

You can find suggestions to clear space at https://cloud.google.com/shell/docs/quotas-limits#clearing_disk_space.

Upgrade to Cloud Workstations for a secure development environment without these limits: https://cloud.google.com/workstations.

shumpei_yamakawa@cloudshell:~$ pwd
/home/shumpei_yamakawa
shumpei_yamakawa@cloudshell:~$ ls

README-cloudshell.txt
shumpei_yamakawa@cloudshell:~$ []
```



```
Welcome to Cloud Shell! Type "help" to get started.

To set your Cloud Platform project in this session use `gcloud config set project [PROJECT_ID]`.

You can view your projects by running `gcloud projects list`.

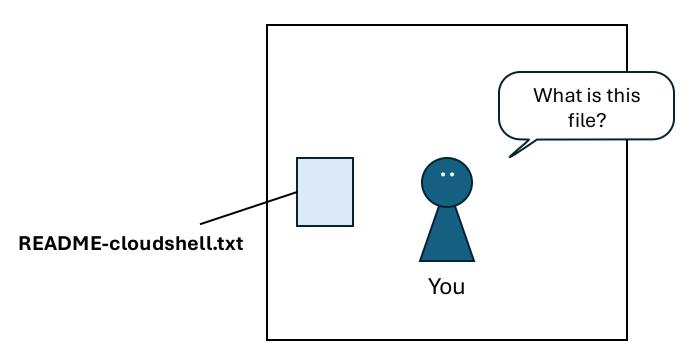
Your home disk usage is at 95%.

You can find suggestions to clear space at https://cloud.google.com/shell/docs/quotas-limits#clearing_disk_space.

Upgrade to Cloud Workstations for a secure development environment without these limits: https://cloud.google.com/workstations.

shumpei_yamakawa@cloudshell:~$ pwd
/home/shumpei_yamakawa
shumpei_yamakawa@cloudshell:~$ ls

README-cloudshell.txt
shumpei_yamakawa@cloudshell:~$ |
```



Type "less {file name}" and enter

```
Welcome to Cloud Shell! Type "help" to get started.

To set your Cloud Platform project in this session use `gcloud config set project [PROJECT_ID]`.

You can view your projects by running `gcloud projects list`.

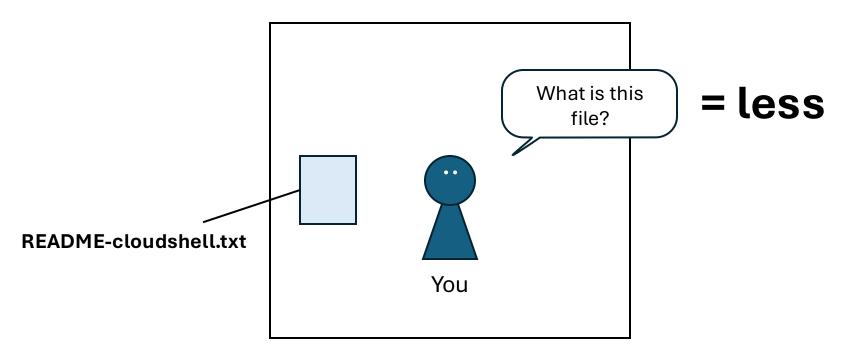
Your home disk usage is at 95%.

You can find suggestions to clear space at https://cloud.google.com/shell/docs/quotas-limits#clearing_disk_space.

Upgrade to Cloud Workstations for a secure development environment without these limits: https://cloud.google.com/workstations.

shumpei_yamakawa@cloudshell:~$ pwd
/home/shumpei_yamakawa
shumpei_yamakawa@cloudshell:~$ ls

README-cloudshell.txt test
shumpei_yamakawa@cloudshell:~$ less README-cloudshell.txt [
```



```
Welcome to Cloud Shell! Type "help" to get started.

To set your Cloud Platform project in this session use `gcloud config set project [PROJECT_ID]`.

You can view your projects by running `gcloud projects list`.

Your home disk usage is at 95%.

You can find suggestions to clear space at https://cloud.google.com/shell/docs/quotas-limits#clearing_disk_space.

Upgrade to Cloud Workstations for a secure development environment without these limits: https://cloud.google.com/workstations.

shumpei_yamakawa@cloudshell:~$ pwd
/home/shumpei_yamakawa
shumpei_yamakawa@cloudshell:~$ ls

README-cloudshell.txt test
shumpei_yamakawa@cloudshell:~$ less README-cloudshell.txt [
```

enter

Welcome to Google Cloud Shell, a tool for managing resources hosted on Google Cloud Platform! The machine comes pre-installed with the Google Cloud SDK and other popular developer tools.

Your 5GB home directory will persist across sessions, but the VM is ephemeral and will be reset approximately 20 minutes after your session ends. No system-wide change will persist beyond that.

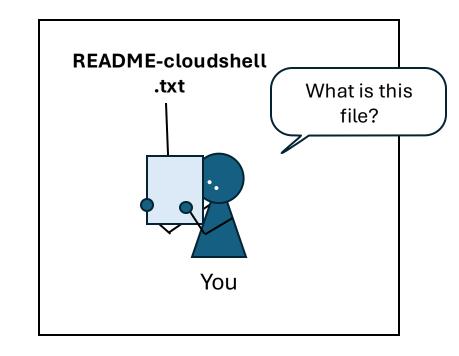
Type "gcloud help" to get help on using Cloud SDK. For more examples, visit https://cloud.google.com/shell/docs/examples

Type "cloudshell help" to get help on using the "cloudshell" utility. Common functionality is aliased to short commands in your shell, for example, you can type "dl <filename>" at Bash prompt to download a file. Type "cloudshell aliases" to see these commands.

Type "help" to see this message any time. Type "builtin help" to see Bash interpreter help.

README-cloudshell.txt (END)

Type "q"



shumpei_yamakawa@cloudshell:~\$ less README-cloudshell.txt shumpei_yamakawa@cloudshell:~\$ [

"cat" is also an option to look at the file content

Shumpei_yamakawa@cloudshell:~\$ cat README-cloudshell.txt

Welcome to Google Cloud Shell, a tool for managing resources hosted on Google Cloud Platform!
The machine comes pre-installed with the Google Cloud SDK and other popular developer tools.

Your 5GB home directory will persist across sessions, but the VM is ephemeral and will be reset approximately 20 minutes after your session ends. No system-wide change will persist beyond that.

Type "gcloud help" to get help on using Cloud SDK. For more examples, visit https://cloud.google.com/shell/docs/quickstart and https://cloud.google.com/shell/docs/examples

Type "cloudshell help" to get help on using the "cloudshell" utility. Common functionality is aliased to short commands in your shell, for example, you can type "dl <filename>" at Bash prompt to download a file. Type "cloudshell aliases" to see these commands.

Type "help" to see this message any time. Type "builtin help" to see Bash interpreter help.

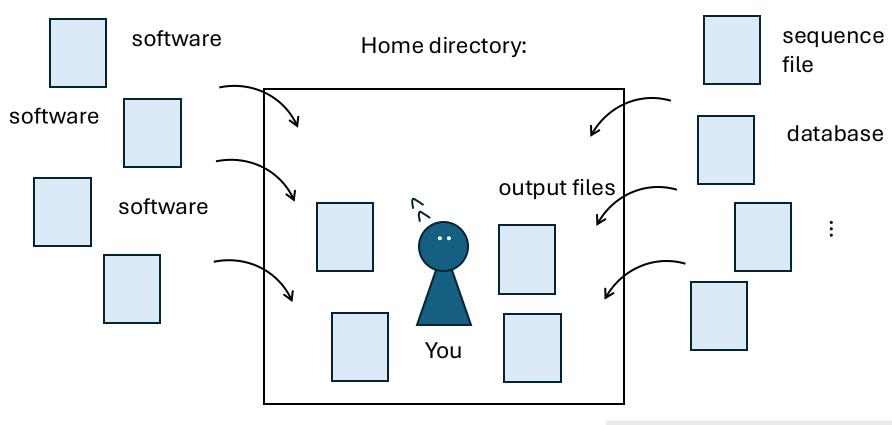
shumpei_yamakawa@cloudshell:~\$ []

Less vs cat

· less: in different output page For big data

· cat: in standard output For small data and to "concatenate"

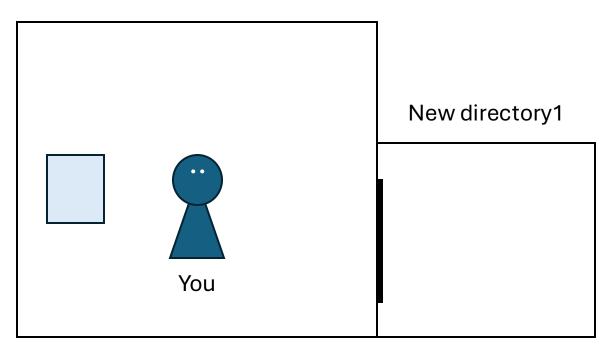
Do not stack too many files in the home directly





Home directory

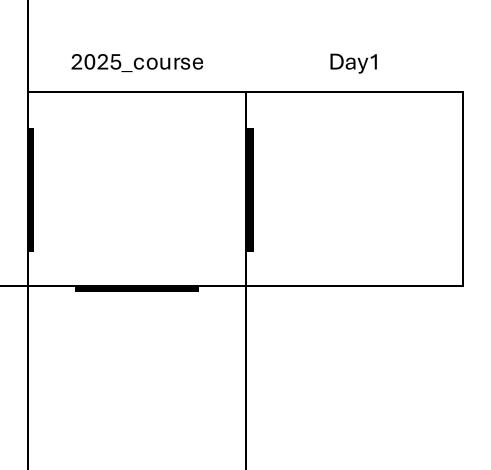
Hierarchical structure is ideal



Hierarchical structure is ideal Home directory New directory1 New directory2 You New directory3

Home directory You

Hierarchical structure is ideal

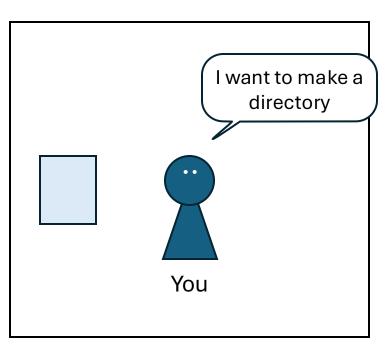


Hierarchical structure is ideal Home directory Ex) Day1 2025_course Soft Soft Day1 ware ware Day1 Day1 analysis in Day1 directory Day2 Day2 Day2

shumpei_yamakawa@cloudshell:~\$ mkdir New_directory1

Home directory

Any name is okay

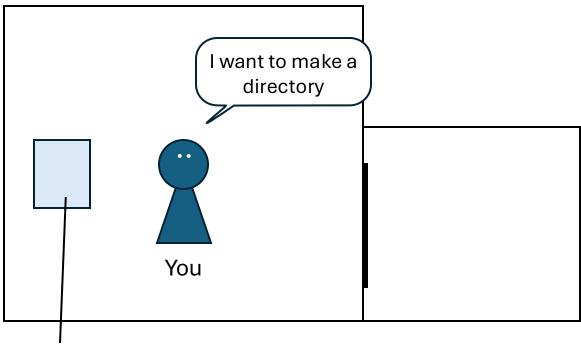


= mkdir

Check the directory by the command ls

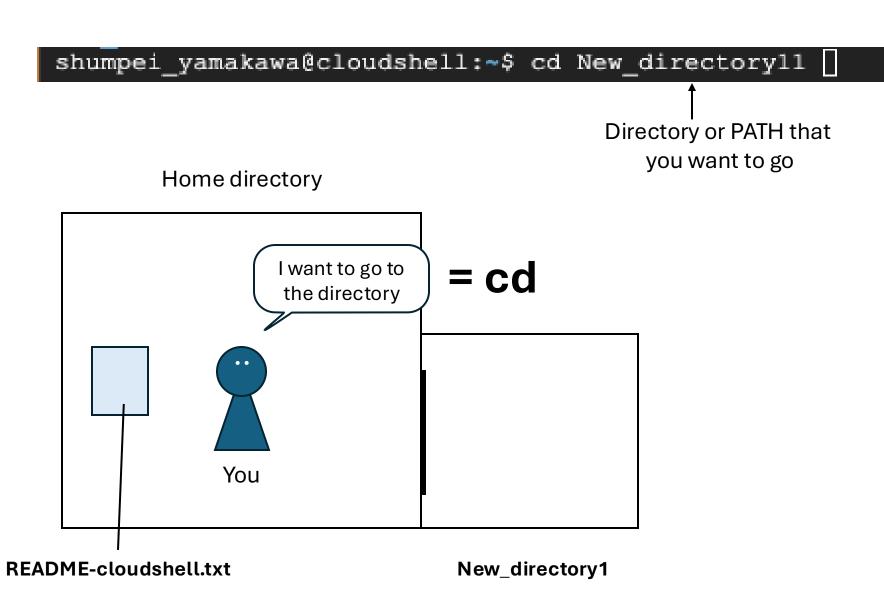
```
shumpei_yamakawa@cloudshell:~$ mkdir New_directory1
shumpei_yamakawa@cloudshell:~$ ls
New_directory1 README-cloudshell.txt
```

Home directory



README-cloudshell.txt

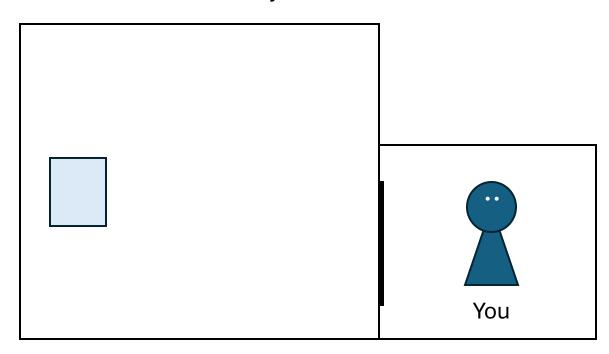
Check the directory by the command ls



Check the directory by the command ls

```
shumpei_yamakawa@cloudshell:~$ cd New_directory1/
shumpei_yamakawa@cloudshell:~/New_directory1$ pwd
/home/shumpei_yamakawa/New_directory1
```

Home directory

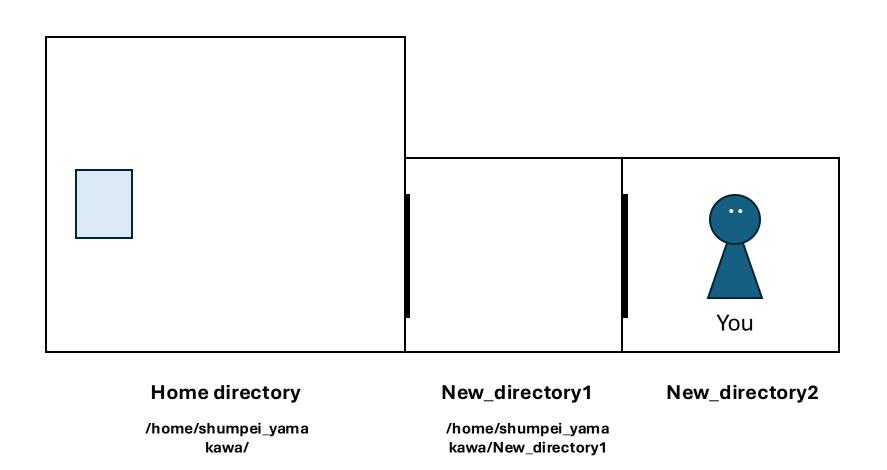


New_directory1

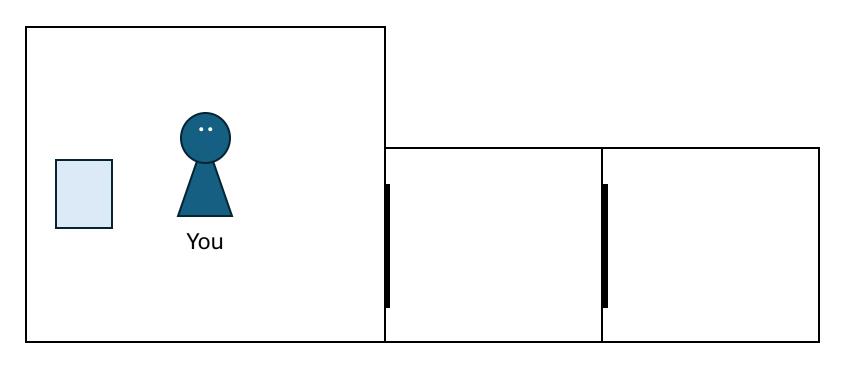
/home/shumpei_yama kawa/

/home/shumpei_yama kawa/New_directory1

Exercise1 Make a new directory and move into there



Exercise2 Go back to the home directory



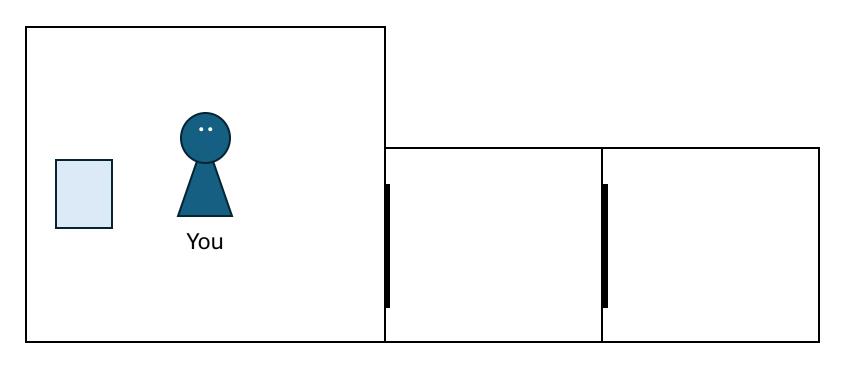
Home directory

/home/shumpei_yama kawa/

New_directory1

/home/shumpei_yama kawa/New_directory1

Exercise2 Go back to the home directory



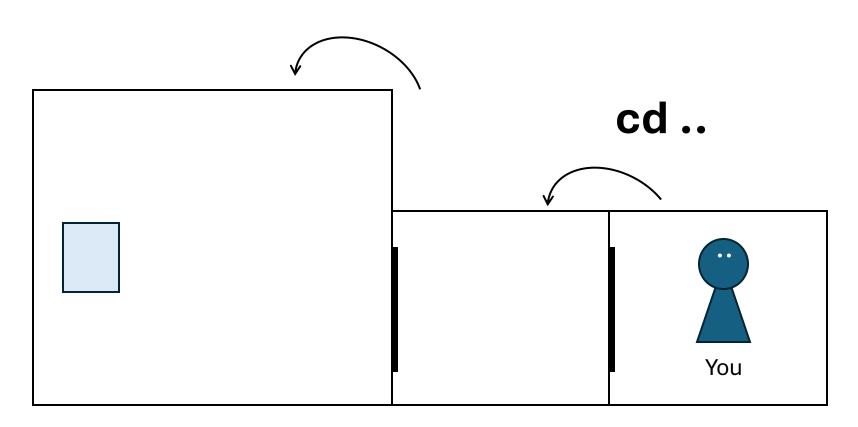
Home directory

/home/shumpei_yama kawa/

New_directory1

/home/shumpei_yama kawa/New_directory1

cd ..



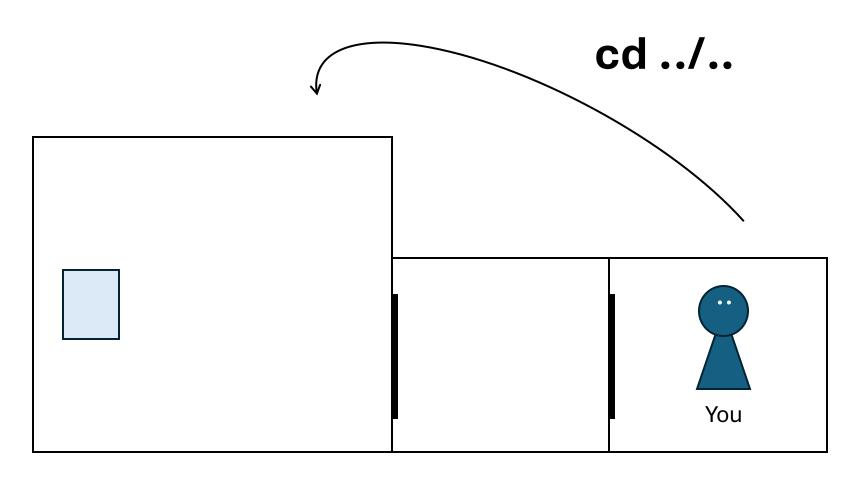
Home directory

/home/shumpei_yama kawa/

New_directory1

/home/shumpei_yama kawa/New_directory1

New_directory2



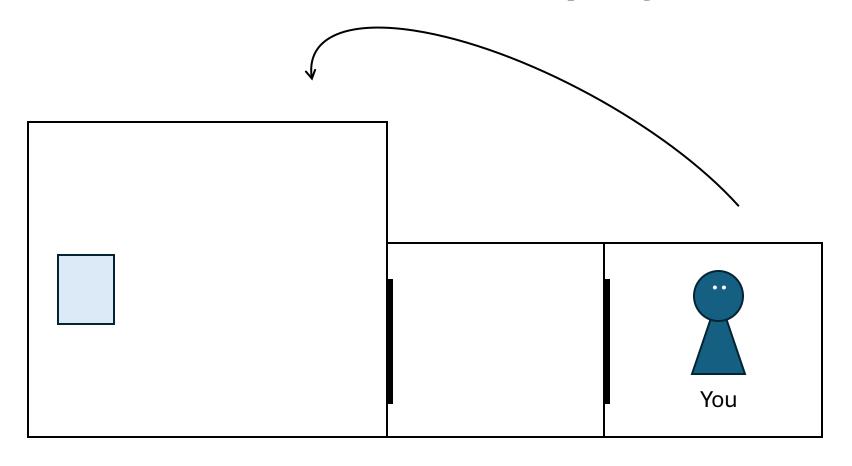
Home directory

/home/shumpei_yama kawa/

New_directory1

/home/shumpei_yama kawa/New_directory1

cd /home/shumpe_yamakawa



Home directory

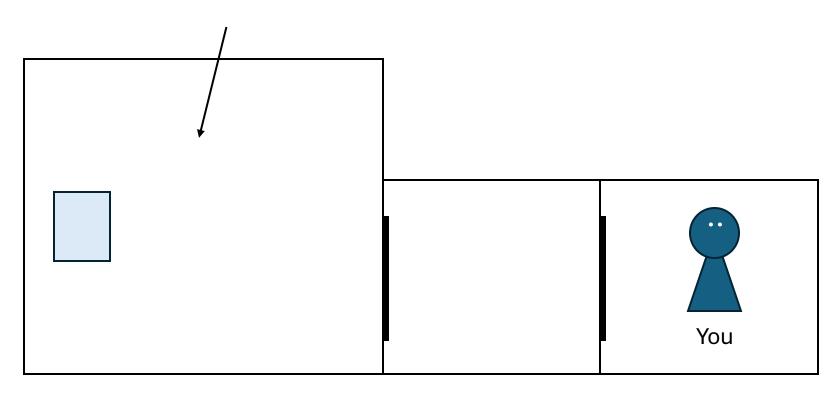
/home/shumpei_yama kawa/

New_directory1

/home/shumpei_yama kawa/New_directory1

Absolute path ../../

Relative path /home/shumpe_yamakawa



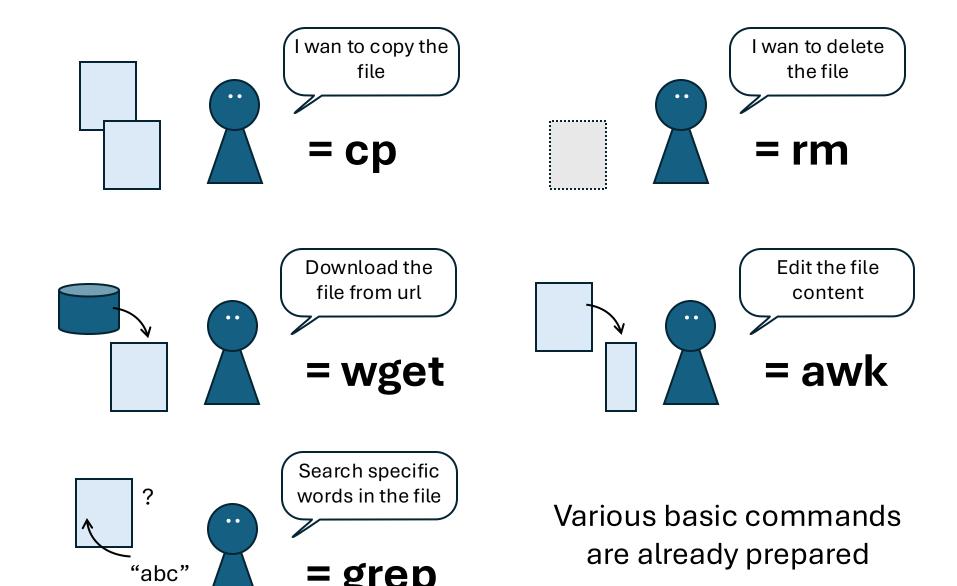
Home directory

/home/shumpei_yama kawa/

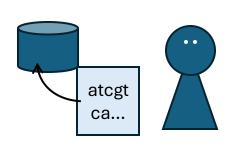
New_directory1

/home/shumpei_yama kawa/New_directory1

2. Basic commands and specific analysis software

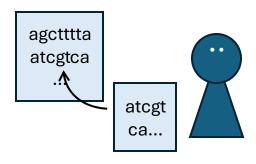


= grep



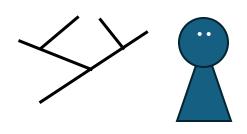
I want to make a Blast database

= makeblastdb



Search specific sequences

= seqkit grep



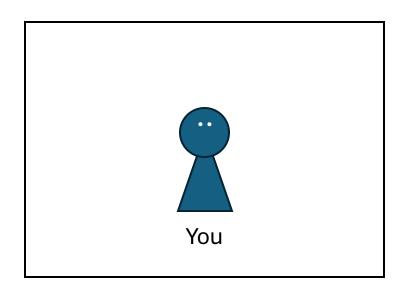
Make a phylogenic tree

= raxml

Need to download specific software for bioinformatic analysis

/usr/bin/

Executable files (binary)



/usr/bin/home/

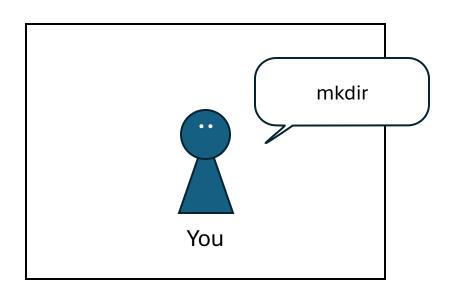


Files for the basic commands such as Cat, less, makdir, pwd etc are included in this directory

/usr/bin/

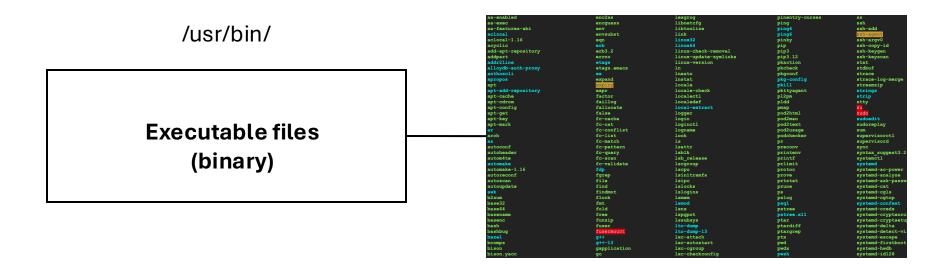
Executable files (binary)

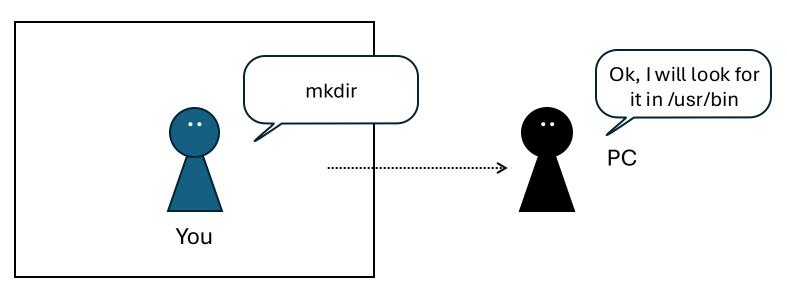




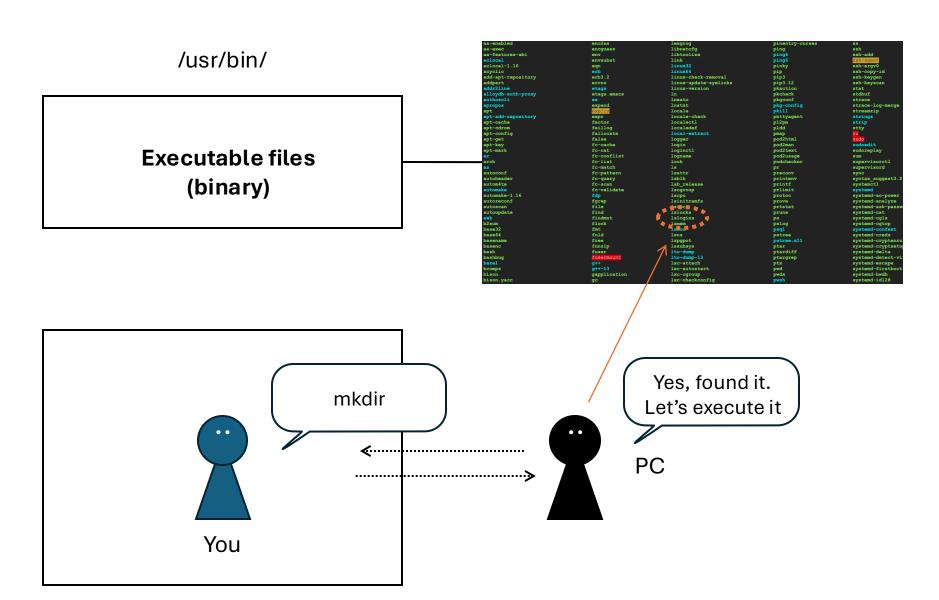
/usr/bin/home/

shumpei_yamakawa@cloudshell:~\$ mkdir

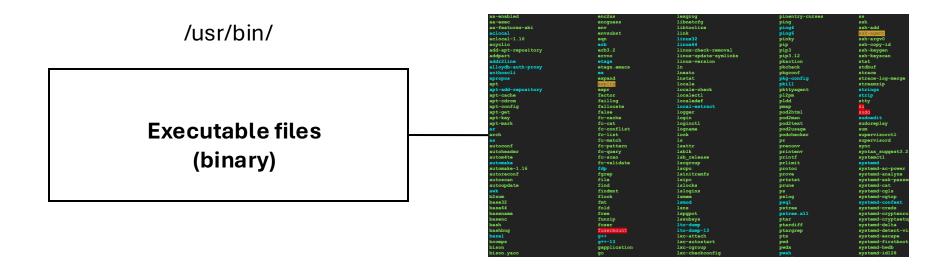


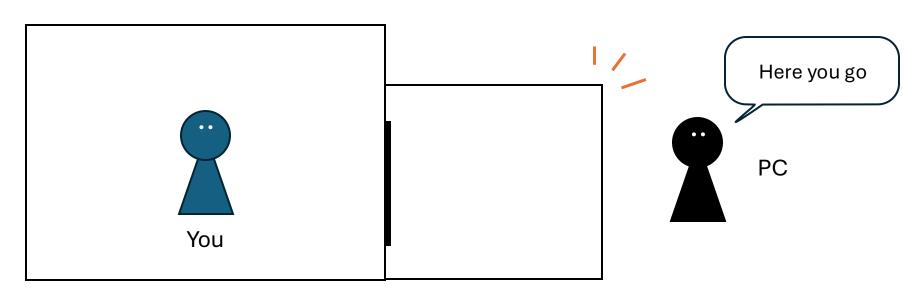


/usr/bin/home/

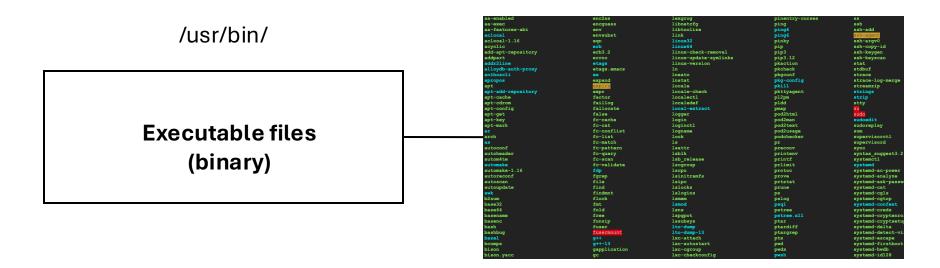


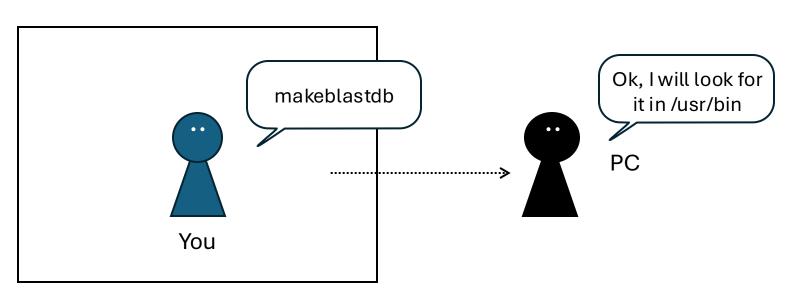
/usr/bin/home/



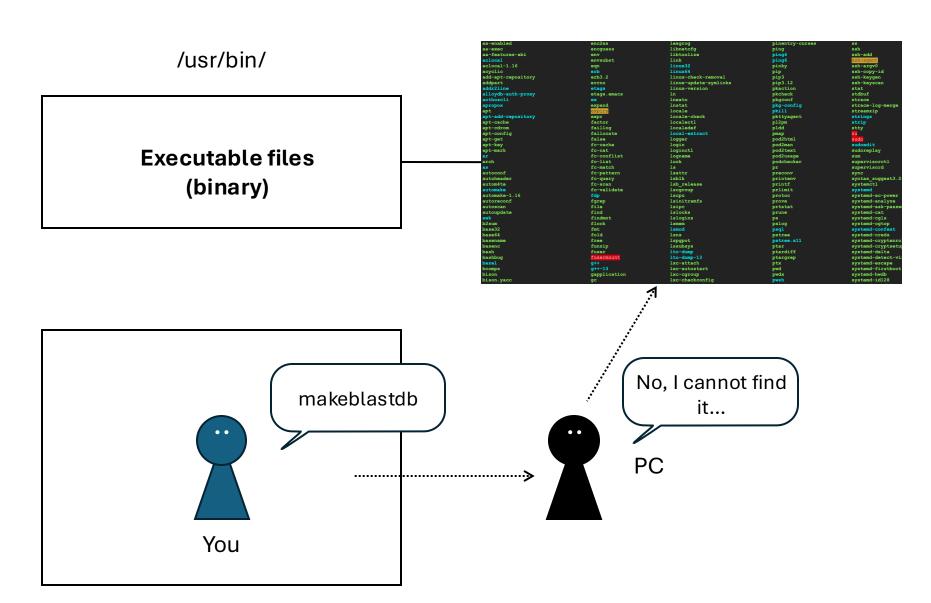


/usr/bin/home/

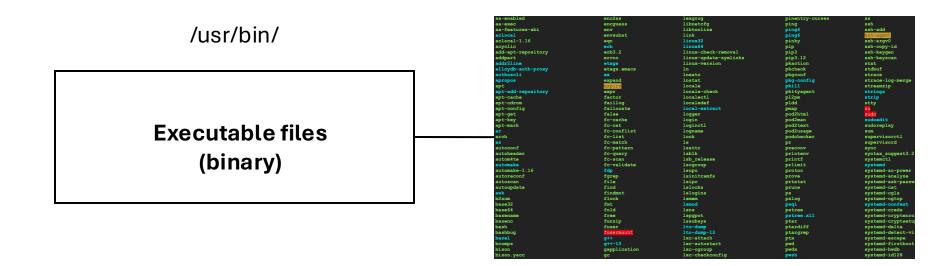


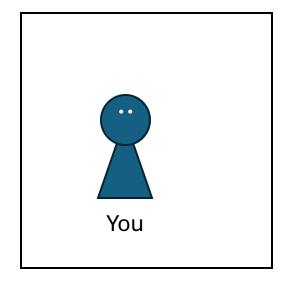


/usr/bin/home/

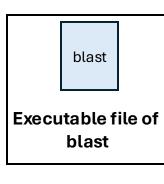


/usr/bin/home/





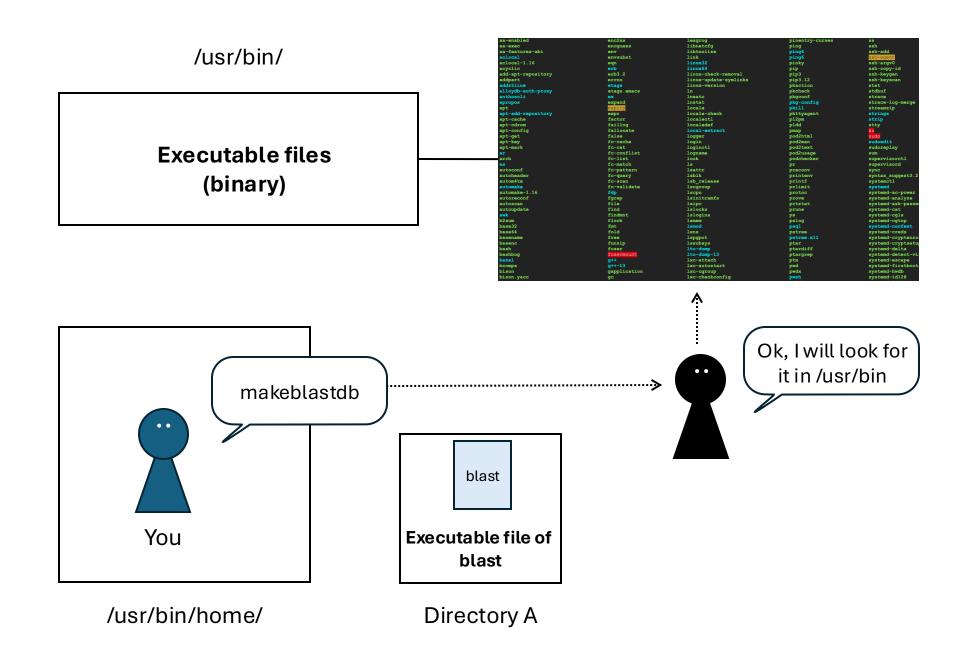


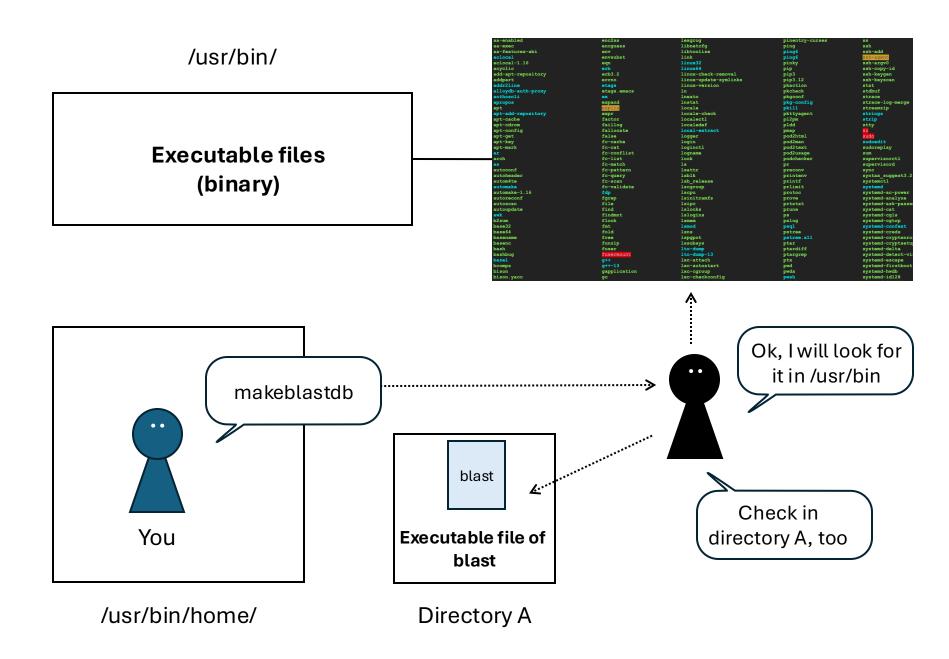


Directory A

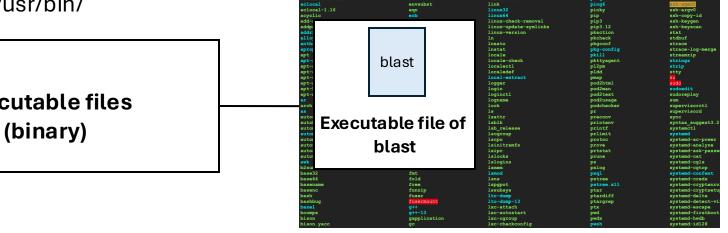


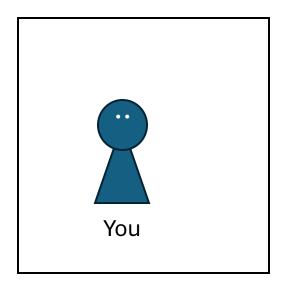
Place the file into different directory





/usr/bin/ Executable files





/usr/bin/home/

This is also possible in principle

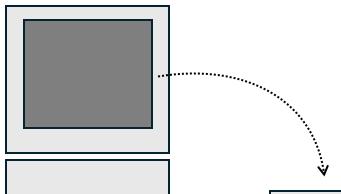
→ but, the directory is for the basic command lines and may cause conflict with them...

"not recommended"

3. Installing the software

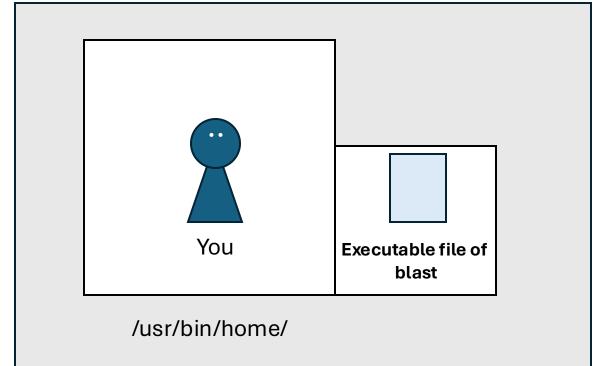
Installing the software



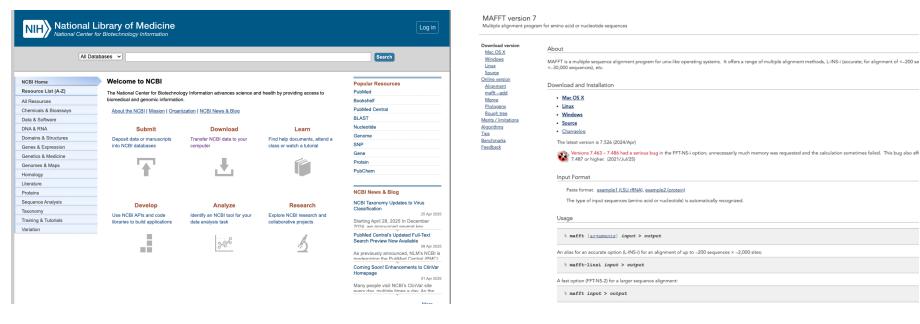


- 1. Download the software
- 2. Add a directory to PATH

Your local PC (=Cloud Shell)



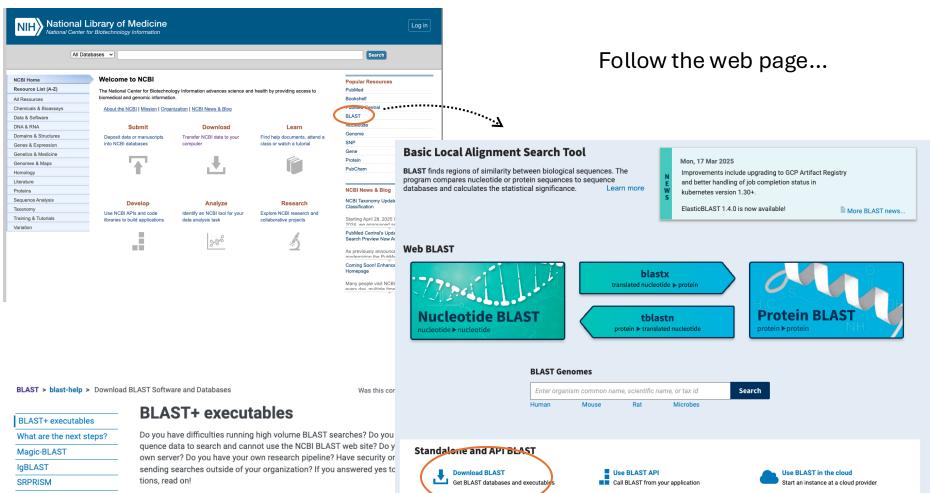
1. Download the software



Ex) NCBI provides a lot of the tools such as blast, sra-tools, and so on...

Ex) The webpage of MAFFT, an alignment tool

Go to the webpage of the software



Databases

The NCBI provides a suite of command-line tools to run BLAST called BLAST+. The allows users to perform BLAST searches on their own server without size, volume and database restrictions. BLAST+ can be used with a command line so it can be integrated directly into your workflow.

What are the next steps?

Download and install BLAST+. Installers and source code are available from https://ft-p.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/. Download the databases you need, (see database section below), or create your own. Start searching.

For more details, please see the BLAST+ user manual, the BLAST Help manual, the BLAST releases notes, and the article in BMC Bioinformatics (PubMed link). See our versioning policy.

The BLAST+ suite is the currently supported package. The older C toolkit executables are no longer supported. See our versioning policy.

We are always listening and welcome your feedback at BLAST Support Center.

Index of /blast/executables/blast+/LATEST

Last modifi	ied	Size	
		_	
2024-06-25	14:34	85	
2024-06-25	14:31	21M	
2024-06-25	14:35	63	FTP
2024-06-25	14:31	202M	
2024-06-25	14:35	66	(File Transfer Protocol)
2024-07-30	11:04	225M	(1110 114110101 1 1010001)
2024-07-30	11:04	74	
2024-06-25	14:33	191M	
2024-06-25	14:35	75	
2024-06-25	14:33	193M	
2024-06-25	14:35	65	
2024-06-25	14:35	27M	
2024-06-25	14:35	64	
2024-06-25	14:35		
2024-06-25	14:35		
		76	
2024-06-25	14:43	400M	
2024-06-25	14:30	129M	
		63	
2024-06-25	14:33		
		206M	
		71	
		70	
		208M	
2024-06-25	14:35	64	
	2024-06-25 2024-06-25 2024-06-25 2024-06-25 2024-07-30 2024-07-30 2024-06-25	2024-06-25 14:34 2024-06-25 14:31 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:35 2024-07-30 11:04 2024-06-25 14:33 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:44 2024-06-25 14:44 2024-06-25 14:44 2024-06-25 14:44 2024-06-25 14:44 2024-06-25 14:35 2024-06-25 14:35	2024-06-25 14:34 85 2024-06-25 14:31 21M 2024-06-25 14:35 63 2024-06-25 14:35 66 2024-07-30 11:04 225M 2024-06-25 14:33 191M 2024-06-25 14:35 75 2024-06-25 14:35 65 2024-06-25 14:35 65 2024-06-25 14:35 65 2024-06-25 14:35 65 2024-06-25 14:35 65 2024-06-25 14:35 65 2024-06-25 14:35 61 2024-06-25 14:35 61 2024-06-25 14:44 398M 2024-06-25 14:44 76 2024-06-25 14:44 76 2024-06-25 14:44 66 2024-06-25 14:44 66 2024-06-25 14:45 63 2024-06-25 14:35 63 2024-06-25 14:35 63 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70

HHS Vulnerability Disclosure

Index of /blast/executables/blast+/LATEST

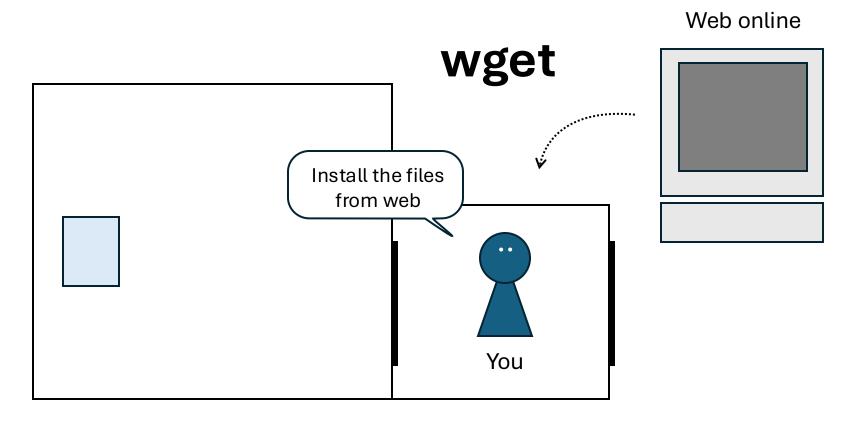
Last modifi	ied	Size	
		_	
2024-06-25	14:34	85	
2024-06-25	14:31	21M	
2024-06-25	14:35	63	FTP
2024-06-25	14:31	202M	
2024-06-25	14:35	66	(File Transfer Protocol)
2024-07-30	11:04	225M	(1110 114110101 1 1010001)
2024-07-30	11:04	74	
2024-06-25	14:33	191M	
2024-06-25	14:35	75	
2024-06-25	14:33	193M	
2024-06-25	14:35	65	
2024-06-25	14:35	27M	
2024-06-25	14:35	64	
2024-06-25	14:35		
2024-06-25	14:35		
		76	
2024-06-25	14:43	400M	
2024-06-25	14:30	129M	
		63	
2024-06-25	14:33		
		206M	
		71	
		208M	
2024-06-25	14:35	64	
	2024-06-25 2024-06-25 2024-06-25 2024-06-25 2024-07-30 2024-07-30 2024-06-25	2024-06-25 14:34 2024-06-25 14:31 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:35 2024-07-30 11:04 2024-07-30 11:04 2024-06-25 14:33 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:35 2024-06-25 14:44 2024-06-25 14:44 2024-06-25 14:44 2024-06-25 14:44 2024-06-25 14:44 2024-06-25 14:35 2024-06-25 14:35	2024-06-25 14:34 85 2024-06-25 14:31 21M 2024-06-25 14:35 63 2024-06-25 14:35 66 2024-07-30 11:04 225M 2024-06-25 14:33 191M 2024-06-25 14:35 75 2024-06-25 14:35 65 2024-06-25 14:35 65 2024-06-25 14:35 65 2024-06-25 14:35 65 2024-06-25 14:35 65 2024-06-25 14:35 65 2024-06-25 14:35 61 2024-06-25 14:35 61 2024-06-25 14:44 398M 2024-06-25 14:44 76 2024-06-25 14:44 76 2024-06-25 14:44 66 2024-06-25 14:44 66 2024-06-25 14:45 63 2024-06-25 14:35 63 2024-06-25 14:35 63 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70 2024-06-25 14:35 70

HHS Vulnerability Disclosure

Index of /blast/executables/blast+/LATEST

Name	Last modified	Size	
Parent Directory		-	
ChangeLog	2024-06-25 14:34	85	
ncbi-blast-2.16.0+-1.src.rpm	2024-06-25 14:31	21M	
ncbi-blast-2.16.0+-1.src.rpm.md5	2024-06-25 14:35	63	
ncbi-blast-2.16.0+-1.x86 64.rpm	2024-06-25 14:31	202M	
ncbi-blast-2.16.0+-1.x86 64.rpm.md5	2024-06-25 14:35	66	
<pre>ncbi-blast-2.16.0+-aarch64-linux.tar.gz</pre>	2024-07-30 11:04	225M	
<pre>ncbi-blast-2.16.0+-aarch64-linux.tar.gz.md5</pre>	2024-07-30 11:04	74	
<pre>ncbi-blast-2.16.0+-aarch64-macosx.tar.gz</pre>	2024-06-25 14:33	191M	
<pre>ncbi-blast-2.16.0+-aarch64-macosx.tar.gz.md5</pre>	2024-06-25 14:35	75	
ncbi-blast-2.16.0+-aarch64.dmg	2024-06-25 14:33	193M	
<u>ncbi-blast-2.16.0+-aarch64.dmg.md5</u>	2024-06-25 14:35	65	
<u>ncbi-blast-2.16.0+-src.tar.gz</u>	2024-06-25 14:35	27M	
<u>ncbi-blast-2.16.0+-src.tar.gz.md5</u>	2024-06-25 14:35	64	
<u>ncbi-blast-2.16.0+-src.zip</u>	2024-06-25 14:35	31M	Copy link address
<pre>ncbi-blast-2.16.0+-src.zip.md5</pre>	2024-06-25 14:35	61	oopy and address
<u>ncbi-blast-2.16.0+-universal-macosx.tar.gz</u>	2024-06-25 14:44	398M	
<pre>ncbi-blast-2.16.0+-universal-macosx.tar.gz.md5</pre>		76	
<u>ncbi-blast-2.16.0+-universal.dmg</u>	2024-06-25 14:43	400M	https://ftp.ncbi.nlm.nih.gov/blast/e
<pre>ncbi-blast-2.16.0+-universal.dmg.md5</pre>	2024-06-25 14:44	66	xecutables/blast+/LATEST/ncbi-
ncbi-blast-2.16.0+-win64.exe	2024-06-25 14:30	129M	blast-2.16.0+-x64-linux.tar.gz
ncbi-blast-2.16.0+-win64.exe.md5	2024-06-25 14:35	63	5td 5t 2:10:01 X04 tilldx.tdl.62
<u>ncbi-blast-2.16.0+-x64-linux.tar.gz</u>	2024-06-25 14:33	246M	
ncbi-blast-2.16.0+-x64-linux.tar.gz.md5	2024-06-25 14:35	70	
<u>ncbi-blast-2.16.0+-x64-macosx.tar.gz</u>	2024-06-25 14:35	206M	
<u>ncbi-blast-2.16.0+-x64-macosx.tar.gz.md5</u>	2024-06-25 14:35	71	
<u>ncbi-blast-2.16.0+-x64-win64.tar.gz</u>	2024-06-25 14:31	133M	
<pre>ncbi-blast-2.16.0+-x64-win64.tar.gz.md5</pre>	2024-06-25 14:35	70	
<u>ncbi-blast-2.16.0+-x86_64.dmg</u>	2024-06-25 14:34	208M	
ncbi-blast-2.16.0+-x86_64.dmg.md5	2024-06-25 14:35	64	

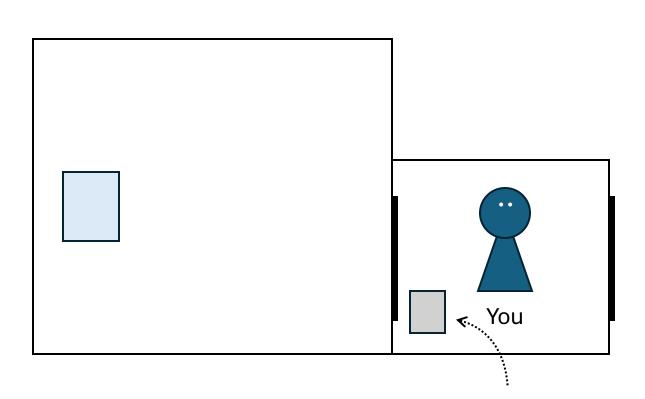
HHS Vulnerability Disclosure



Home directory

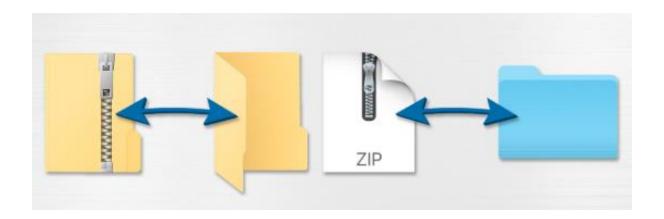
/home/shumpei_yama kawa/ New_directory1

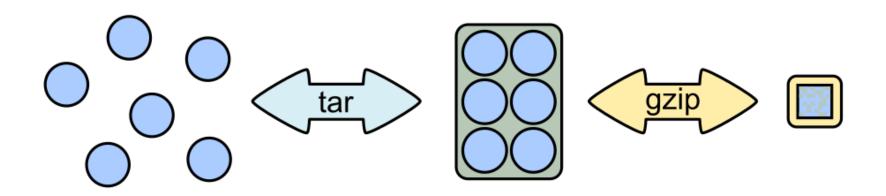
/home/shumpei_yama kawa/New_directory1



ncbi-blast-2.16.0+x64-linux.<u>tar.gz</u>

Archive and compress

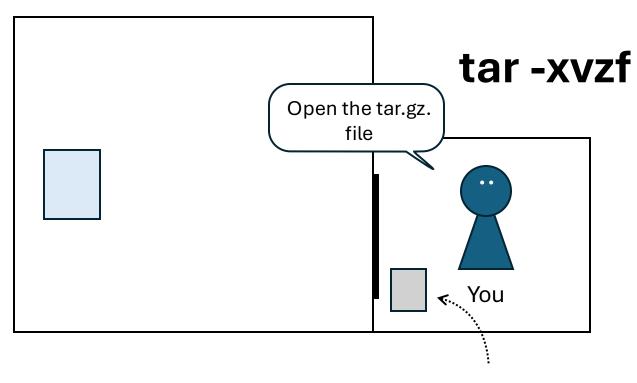




archive

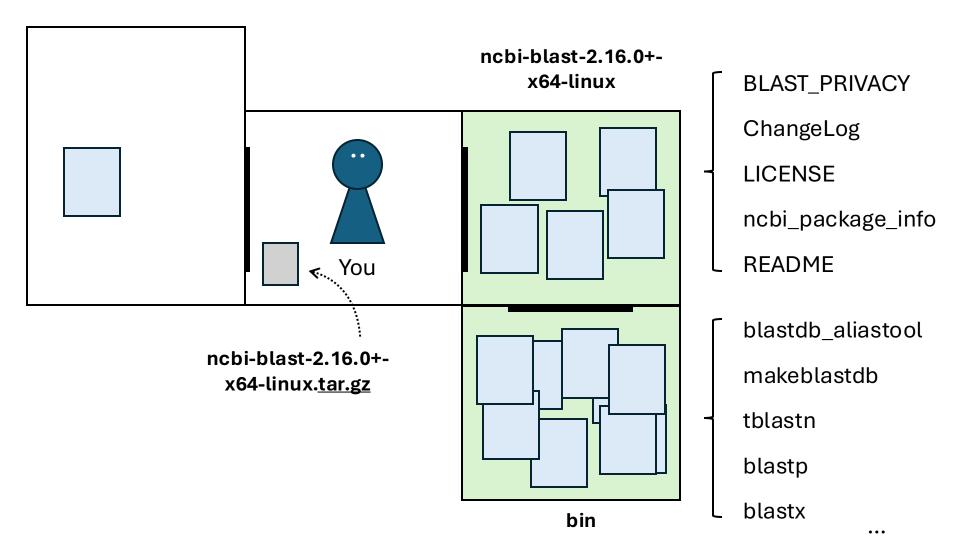
compress

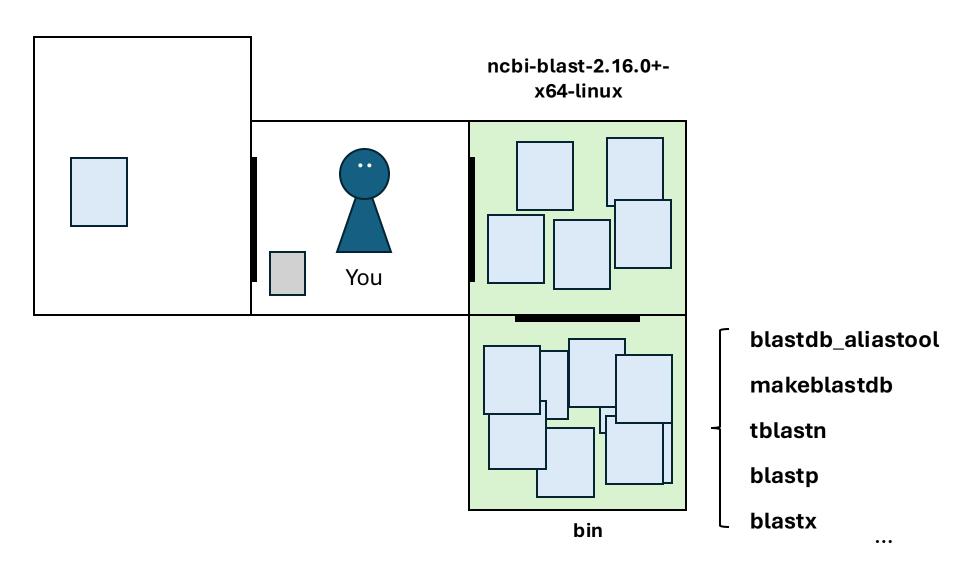
~\$ tar -xvzf ncbi-blast-2.16.0+-x64-linux.tar.gz



ncbi-blast-2.16.0+x64-linux.tar.gz

```
shumpei_yamakawa@cloudshell:~$ tar -xvzf ncbi-blast-2.16.0+-x64-linux.tar.gz
ncbi-blast-2.16.0+/
ncbi-blast-2.16.0+/ChangeLog
ncbi-blast-2.16.0+/bin/
ncbi-blast-2.16.0+/bin/rpsblast
tar: ncbi-blast-2.16.0+/bin/rpsblast: Wrote only 6656 of 10240 bytes
```





Executable files

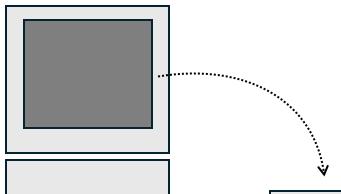
Not always so easy....

RAxML

```
wget https://github.com/stamatak/standard-RAxML.git
unzip master.zip
cd standard-RAxML-master/
make -f Makefile.gcc
rm *.0
make -f Makefile.SSE3.gcc
rm *.0
make -f Makefile.PTHREADS.gcc
rm *.0
make -f Makefile.SSE3.PTHREADS.gcc
echo 'export PATH="/home/shumpei_yamakawa/test/standard-RAxML-master:$PATH"' >> ~/.bashrc; source ~/.bashrc
```

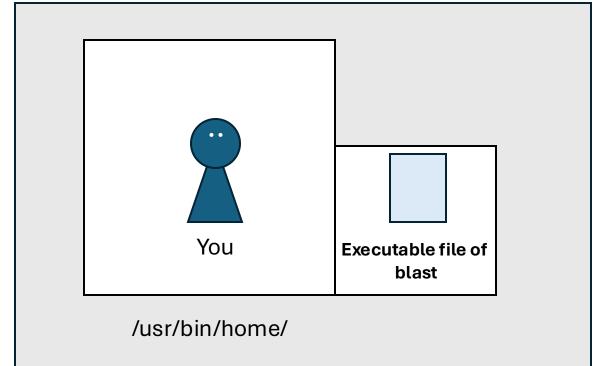
Installing the software

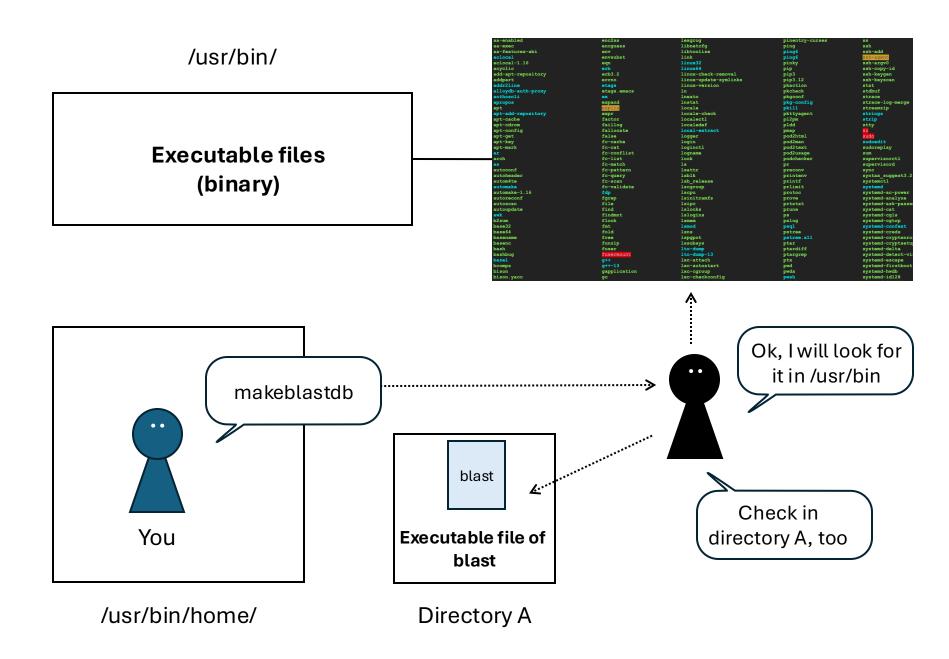


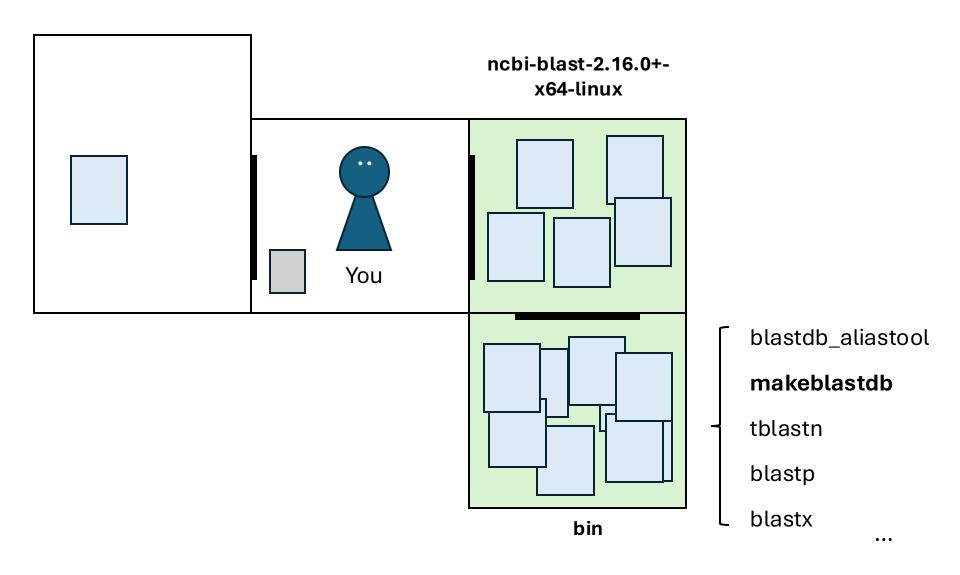


- 1. Download the software
- 2. Add a directory to PATH

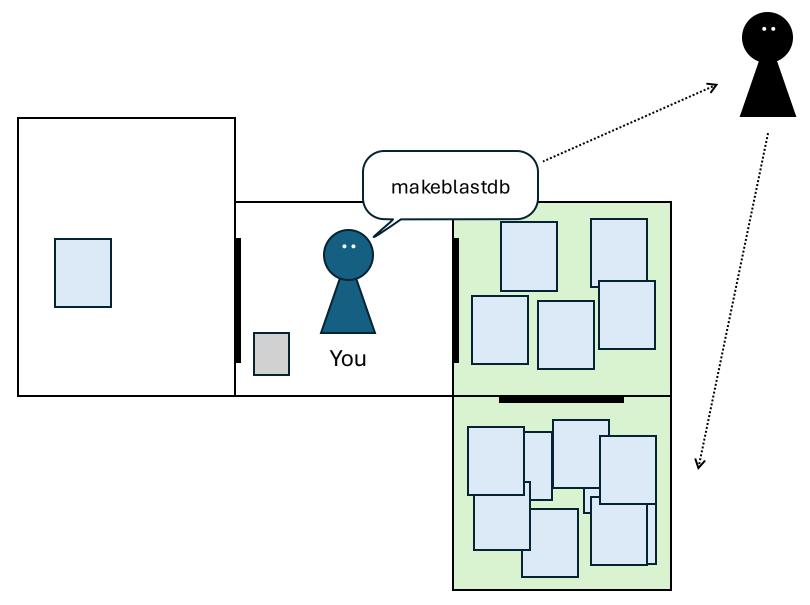
Your local PC (=Cloud Shell)





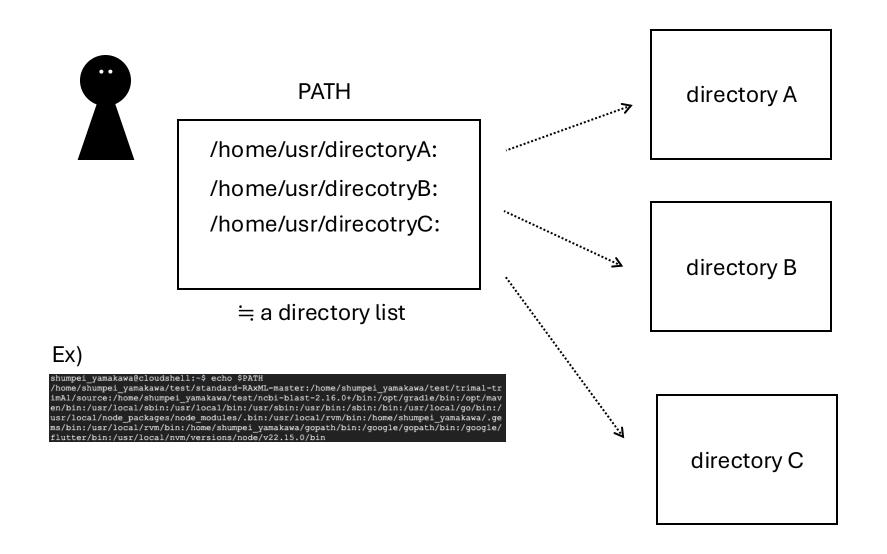


Executable files

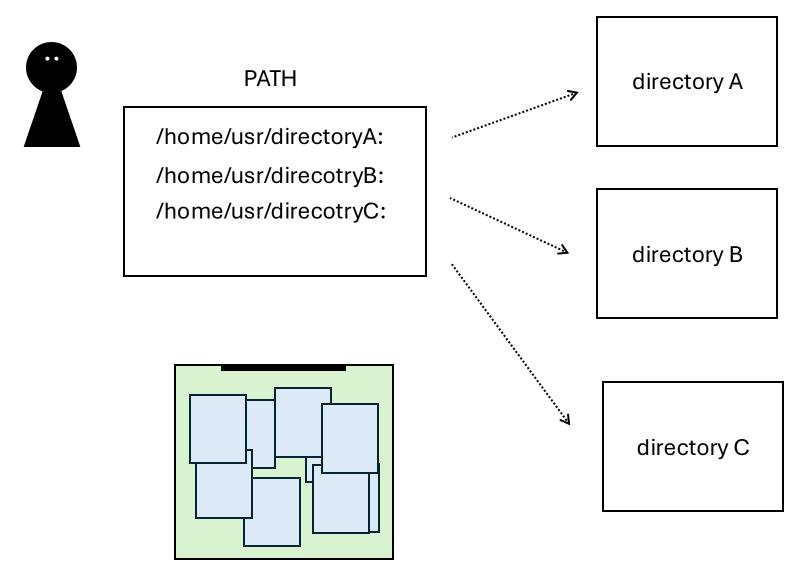


/home/shumpei_yamakawa/n cbi-blast-2.16.0+/bin

PATH: Search path for executable commands

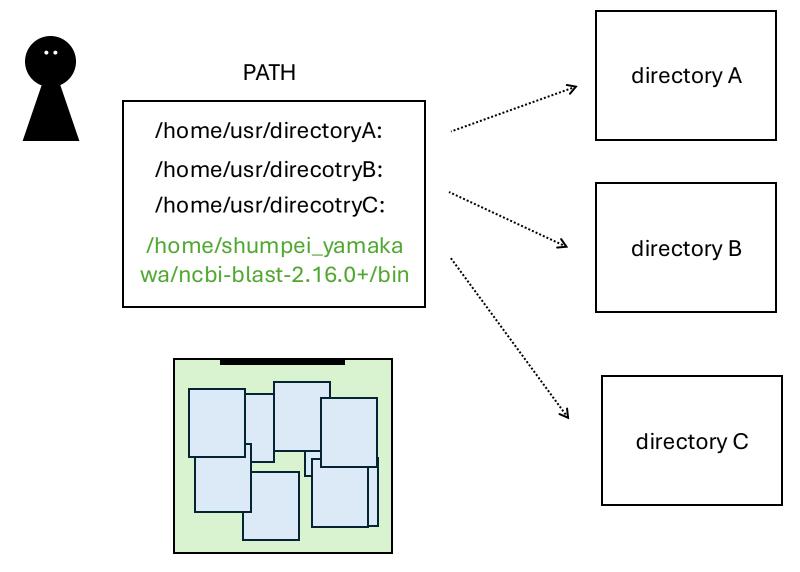


PATH: Search path for executable commands

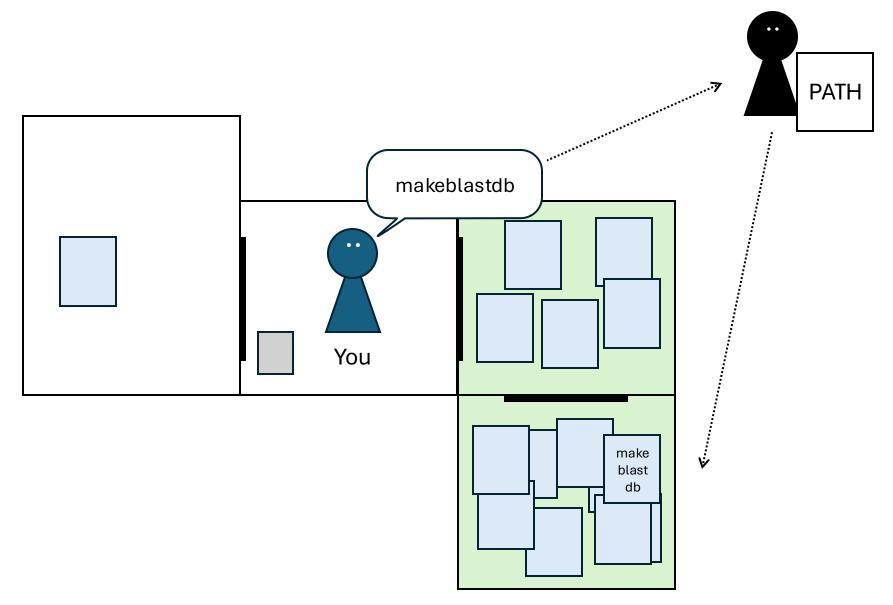


/home/shumpei_yamakawa/ncbi-blast-2.16.0+/bin

PATH: Search path for executable commands

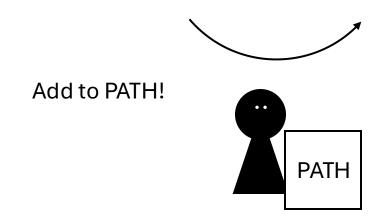


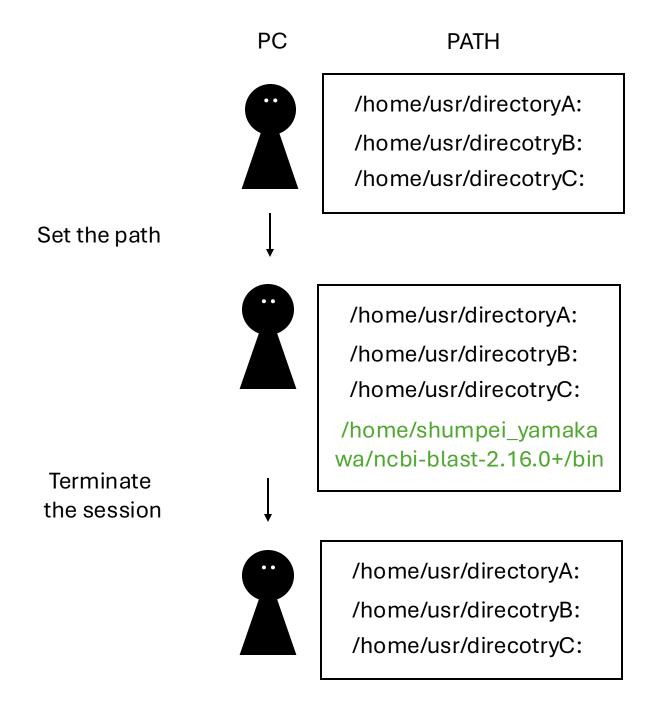
/home/shumpei_yamakawa/ncbi-blast-2.16.0+/bin



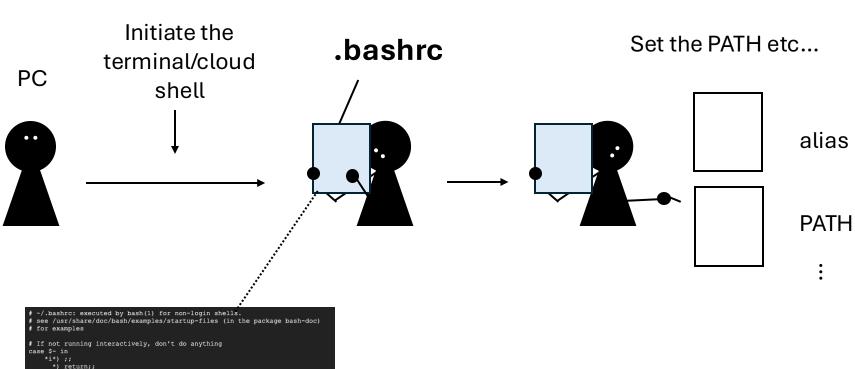
/home/shumpei_yamakawa/n cbi-blast-2.16.0+/bin

export PATH="/home/shumpei_yamakawa/test/ncbi-blast-2.16.0+/bin:\$PATH"





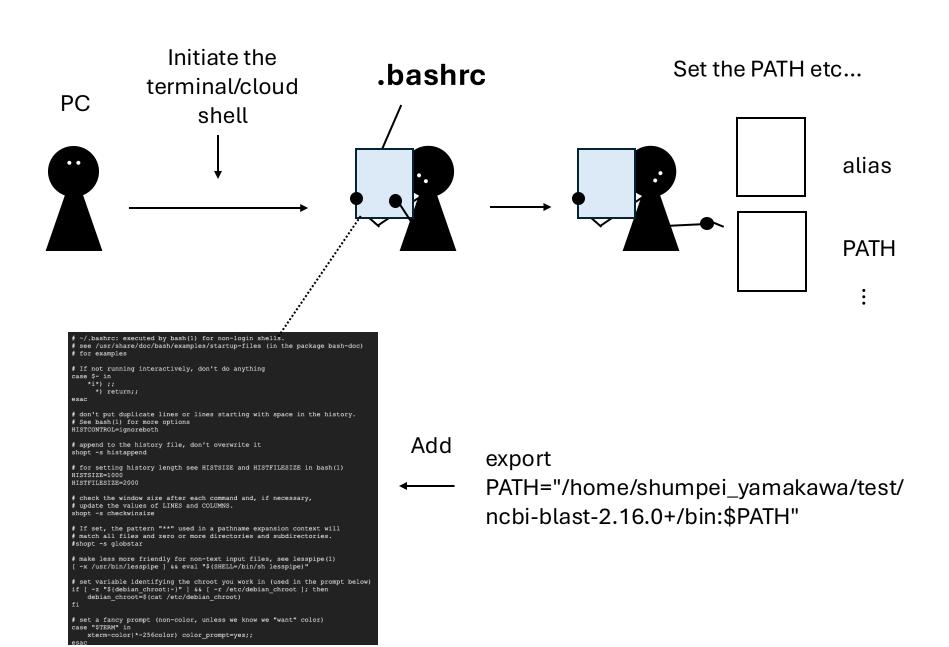
PATH setting is only temporal in a session

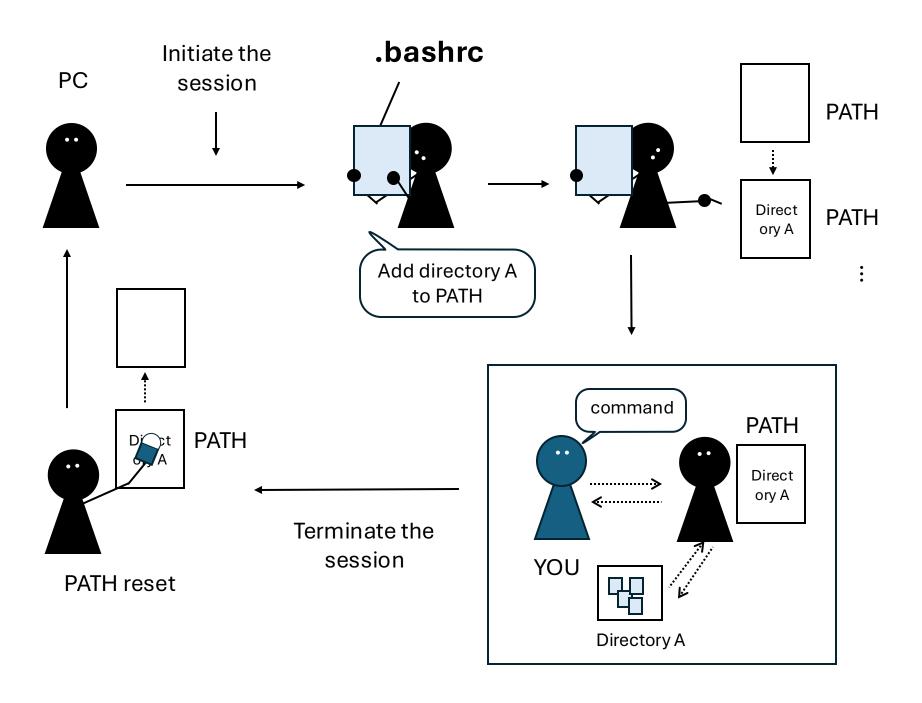


*) return;; # don't put duplicate lines or lines starting with space in the history. # See bash(1) for more options HISTCONTROL=ignoreboth # append to the history file, don't overwrite it shopt -s histappend $\mbox{\$ for setting history length see HISTSIZE}$ and HISTFILESIZE in bash(1) HISTSIZE=1000 HISTFILESIZE=2000 # check the window size after each command and, if necessary, # update the values of LINES and COLUMNS. shopt -s checkwinsize # If set, the pattern "**" used in a pathname expansion context will # match all files and zero or more directories and subdirectories. #shopt -s globstar # make less more friendly for non-text input files, see lesspipe(1)
[-x /usr/bin/lesspipe] && eval "\$(SHELL=/bin/sh lesspipe)" # set variable identifying the chroot you work in (used in the prompt below)
if [-z "\${debian_chroot:-}"] && [-r /etc/debian_chroot]; then debian chroot=\$(cat /etc/debian chroot) # set a fancy prompt (non-color, unless we know we "want" color) xterm-color|*-256color) color prompt=yes;;

.bashrc =

A file automatically executed whenever a Bash shell session starts





```
# ~/.bashrc: executed by bash(1) for non-login shells.
# see /usr/share/doc/bash/examples/startup-files (in the package bash-doc)
# for examples
# If not running interactively, don't do anything
case S- in
    *i*) ;;
      *) return;;
esac
# don't put duplicate lines or lines starting with space in the history.
# See bash(1) for more options
HISTCONTROL=ignoreboth
# append to the history file, don't overwrite it
shopt -s histappend
# for setting history length see HISTSIZE and HISTFILESIZE in bash(1)
HISTSIZE=1000
HISTFILESIZE=2000
# check the window size after each command and, if necessary,
# update the values of LINES and COLUMNS.
shopt -s checkwinsize
# If set, the pattern "**" used in a pathname expansion context will
# match all files and zero or more directories and subdirectories.
#shopt -s globstar
# make less more friendly for non-text input files, see lesspipe(1)
[ -x /usr/bin/lesspipe ] && eval "$(SHELL=/bin/sh lesspipe)"
# set variable identifying the chroot you work in (used in the prompt below)
if [ -z "${debian chroot:-}" ] && [ -r /etc/debian chroot ]; then
    debian chroot=$(cat /etc/debian chroot)
# set a fancy prompt (non-color, unless we know we "want" color)
case "$TERM" in
    xterm-color|*-256color) color prompt=yes;;
```

~/.bashrc

:

```
alias grep='grep --color=auto'
    alias fgrep='fgrep --color=auto'
    alias egrep='egrep --color=auto'
# colored GCC warnings and errors
#export GCC COLORS='error=01;31:warning=01;35:note=01;36:caret=01;32:locus=01
# some more ls aliases
alias ll='ls -alF'
alias la='ls -A'
alias l='ls -CF'
# Add an "alert" alias for long running commands. Use like so:
    sleep 10; alert
alias alert='notify-send --urgency=low -i "$([ $? = 0 ] && echo terminal || ec
# Alias definitions.
# You may want to put all your additions into a separate file like
# ~/.bash aliases, instead of adding them here directly.
# See /usr/share/doc/bash-doc/examples in the bash-doc package.
if [ -f ~/.bash aliases ]; then
   . ~/.bash aliases
# enable programmable completion features (you don't need to enable
# this, if it's already enabled in /etc/bash.bashrc and /etc/profile
# sources /etc/bash.bashrc).
if ! shopt -oq posix; then
 if [ -f /usr/share/bash-completion/bash completion ]; then
    . /usr/share/bash-completion/bash completion
 elif [ -f /etc/bash completion ]; then
    . /etc/bash completion
source /google/devshell/bashrc.google
export PATH="/home/shumpei yamakawa/test/ncbi-blast-2.16.0+/bin:$PATH"
```