

# SRA-TOOLKIT

---

## DEMONSTRATION

---



# SRA TOOLKIT

## CATEGORIES

Data retrieval  
and config

Convert FASTQ,  
SAM, FASTA files

Format-specific  
import/export

Validation  
and QC

Utilities for data  
management

Encryption  
controls

# SRA TOOLKIT

---

## ALL TOOLS

---

abi-dump  
abi-load  
align-info  
bam-load  
blastn\_vdb  
cache-mgr  
cg-load  
dump-ref-fasta  
fasterq-dump  
fastq-dump  
fastq-load  
helicos-load

illumina-dump  
illumina-load  
kar  
kdbmeta  
latf-load  
pacbio-load  
prefetch  
rcexplain  
remote-fuser  
sam-dump  
sff-dump  
sff-load

sra-blastn  
sropath  
sra-pileup  
sra-search  
sra-sort  
sra-sort-cg  
sra-stat  
sratools  
srf-load  
test-sra  
vdb-config  
vdb-copy

vdb-decrypt  
vdb-dump  
vdb-encrypt  
vdb-lock  
vdb-passwd  
vdb-unlock  
vdb-validate

# SRA TOOLKIT

---

**FOR THE EVERYDAY USER**

---

abi-dump  
abi-load  
align-info  
bam-load  
blastn\_vdb  
cache-mgr  
cg-load  
dump-ref-fasta  
**fasterq-dump**  
fastq-dump  
fastq-load  
helicos-load

illumina-dump  
illumina-load  
kar  
kdbmeta  
latf-load  
pacbio-load  
**prefetch**  
rceexplain  
remote-fuser  
**sam-dump**  
sff-dump  
sff-load

sra-blastn  
sropath  
sra-pileup  
sra-search  
sra-sort  
sra-sort-cg  
**sra-stat**  
sratools  
srf-load  
test-sra  
**vdb-config**  
vdb-copy

vdb-decrypt  
**vdb-dump**  
vdb-encrypt  
vdb-lock  
vdb-passwd  
vdb-unlock  
vdb-validate

# SRA TOOLKIT

---

## fastq-dump (Legacy)

Converts SRA accessions into FASTQ files.  
**Slow** on modern datasets, but supports --stdout (useful for piping into other tools)

## fasterq-dump

Converts SRA accessions into FASTQ files. **Fast** and scalable, but requires more temporary disk space (10x FASTQ size)

## prefetch

Downloads SRA accessions from NCBI servers

## sam-dump

Converts SRA accessions into SAM format

## sra-stat

Reports statistics about SRA accessions (ie. such bases, read layout, and overall size)

## vdb-config

Opens and edits SRA Toolkit configuration menu

## vdb-dump

Dumps raw data directly from the VDB database tables (or specific columns)

INSTALL

CONFIG

USE

# SRA-TOOKIT

DEMONSTRATION

# FOLLOW ALONG

The screenshot shows a web browser window with the URL [github.com/ncbi/sra-tools/wiki](https://github.com/ncbi/sra-tools/wiki). The main content area displays the following:

- 01. Downloading SRA Toolkit**
- NCBI SRA Toolkit**
- Below are the latest releases of various tools and release checksum file.
- SRA Toolkit**
- Compiled binaries/install scripts of March 18, 2025, version 3.2.1:
  - [AlmaLinux 64 bit architecture](#) - non-sudo tar archive
  - [Ubuntu Linux 64 bit architecture](#) - non-sudo tar archive

A sidebar on the right contains the following text and an illustration:

Instructions on how to install the **SRA Toolkit** are available in the official **sra-tools** Github repo.

**SITE HERE**

# FOLLOW ALONG

A digital composite image illustrating the "Follow Along" section. On the left, a yellow rectangular callout box contains the text: "But... we have a convenient **starter guide** in our workshop repository." Below this text is a cartoon illustration of a scientist in a white lab coat and glasses, cheering with arms raised. To the right of the scientist is a white button labeled "GUIDE HERE" with a hand cursor icon pointing at it. On the right side of the image is a screenshot of a GitHub repository page for "workshop-sra-hands-on". The repository name is displayed at the top. Below it, a file icon and the filename "Starter\_Guide.md" are shown. The GitHub interface includes standard navigation elements like back, forward, and search buttons.

< > ⌂

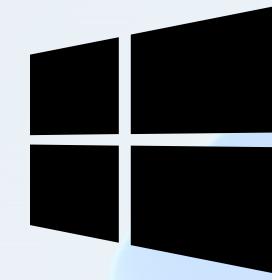
github.com / ncbi / workshop-sra-hands-on

workshop-sra-hands-on

GUIDE HERE

Starter\_Guide.md

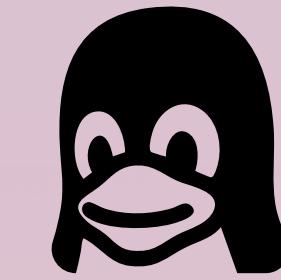
# INSTALLING THE SRA TOOLKIT



**Windows**



**macOS**



**Linux**

*Ubuntu & AlmaLinux*

# INSTALLING THE SRA TOOLKIT



①

If you're on **Windows**,  
download the installation files.

Visit the *SRA Tools*  
GitHub repository at  
the following link.

[github.com/ncbi/sra-tools/wiki](https://github.com/ncbi/sra-tools/wiki)

02. Installing SRA Toolkit

[Jump to bottom](#)

Andrew Klymenko edited this page on Mar 19 · 15 revisions

The SRA Toolkit provides 64-bit binary installations for the  
Ubuntu and Alma Linux distributions, for Mac OS X, and for  
Windows.

Download the current  
**Windows** installation file.

OS	are available here
Windows	<a href="#">srato toolkit.current-win64.zip</a>



srato toolkit-  
win64.zip

Right click  
& extract.

If you're on **macOS**  
or **Linux**, please wait!

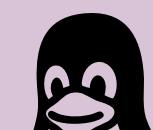
We will install directly  
from the Terminal on  
the next slides.



# INSTALLING THE SRA TOOLKIT

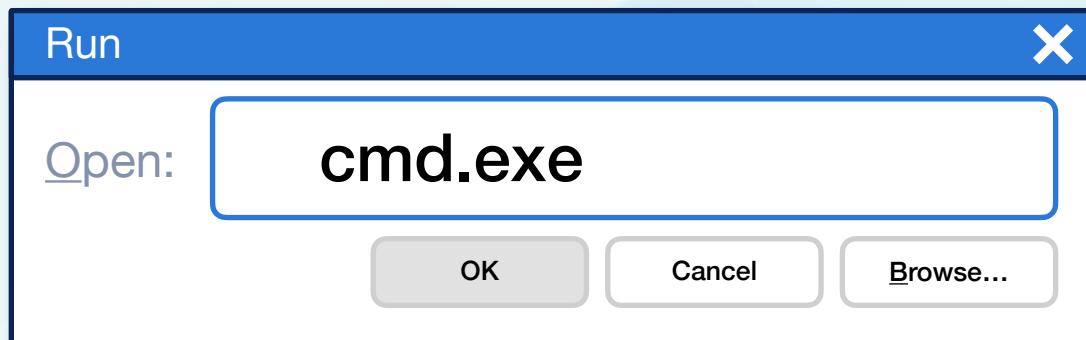
 Windows

 macOS

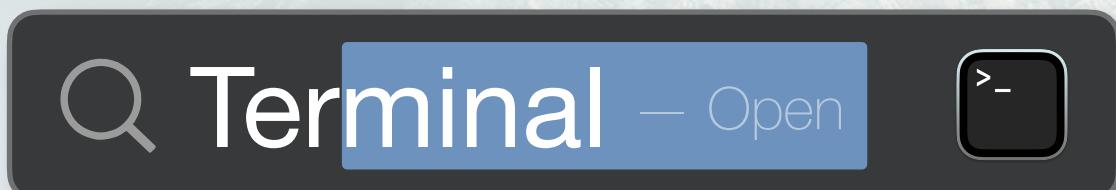
 Linux

- ② Open a new session in your Command Prompt or Terminal.

Press 



Press 



cmd.exe

C:\...\> |

Press 

terminal

terminal

user@ubuntu:~\$ |

# INSTALLING THE SRA TOOLKIT



**Bonus:** Installation for Mac/Linux is  
on *Homebrew* for at [brew.sh](#)

cmd.exe - □ ×

C:\...\> |

**Sorry!** Homebrew  
is not available  
for Windows!



Mac:~user\$ - Terminal

```
Mac:~user$ /bin/bash -c
"$(curl -fsSL https://
raw.githubusercontent.com/
Homebrew/install/HEAD/
install.sh)"

Mac:~user$ brew install
sratoolkit |
```

[brew.sh](#)



New releases typically  
lag a few weeks behind  
the official source.

# INSTALLING THE SRA TOOLKIT



```
cmd.exe - □ ×  
C:\...\> cd Desktop  
C:\...\> cd sratoolkit.current-  
win64  
C:\...\> cd bin  
C:\...\>
```



③

Extract the toolkit files you accessed from the NCBI server.

```
Terminal  
Mac:~user$ curl --output  
sratoolkit.tar.gz  
https://ftp-  
trace.ncbi.nlm.nih.gov/sra/sdk/  
current/sratoolkit.current-  
mac64.tar.gz  
  
Mac:~user$ tar -vxzf  
sratoolkit.tar.gz  
  
Mac:~user$ export PATH=$PWD/  
sratoolkit.<ver>/bin:$PATH
```



```
terminal - □ ×  
user@ubuntu:~$ wget --output-  
document sratoolkit.tar.gz  
https://ftp-  
trace.ncbi.nlm.nih.gov/sra/sdk/  
current/sratoolkit.current-  
ubuntu64.tar.gz  
  
user@ubuntu:~$ tar -vxzf  
sratoolkit.tar.gz  
  
user@ubuntu:~$ export  
PATH=$PWD/sratoolkit.<ver>/bin:  
$PATH
```

For AlmaLinux, instead use [-alma\\_linux64](#)

# LET'S REVIEW!

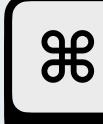
- ① For  **Windows** users:  
download the correct SRA  
installation file from *GitHub*.

github.com/ncbi/  
sra-tools/wiki

Right click and extract it to  
the Desktop.

- ② Search for your Terminal  
or command prompt, and  
create a new session.

**Windows:**  + 

**macOS:**  +  Space

**Linux:** 

- ③ Extract the toolkit files.



**Windows**

> `cd Desktop/sratookit.current-win64/bin`



**macOS**

```
$ curl --output sratookit.tar.gz
https://ftp-trace.ncbi.nlm.nih.gov/sra/
sdk/current/sratookit.current-
mac64.tar.gz
$ tar -vxzf sratookit.tar.gz
$ export PATH=$PWD/sratookit.<ver>/bin:
$PATH
```



**Linux**

```
$ wget --output-document sratookit.tar.gz
https://ftp-trace.ncbi.nlm.nih.gov/sra/
sdk/current/sratookit.current-
ubuntu64.tar.gz
$ tar -vxzf sratookit.tar.gz
$ export PATH=$PWD/sratookit.<ver>/bin:
$PATH
```

INSTALL

CONFIG

USE

# SRA-TOOLKIT

---

DEMONSTRATION

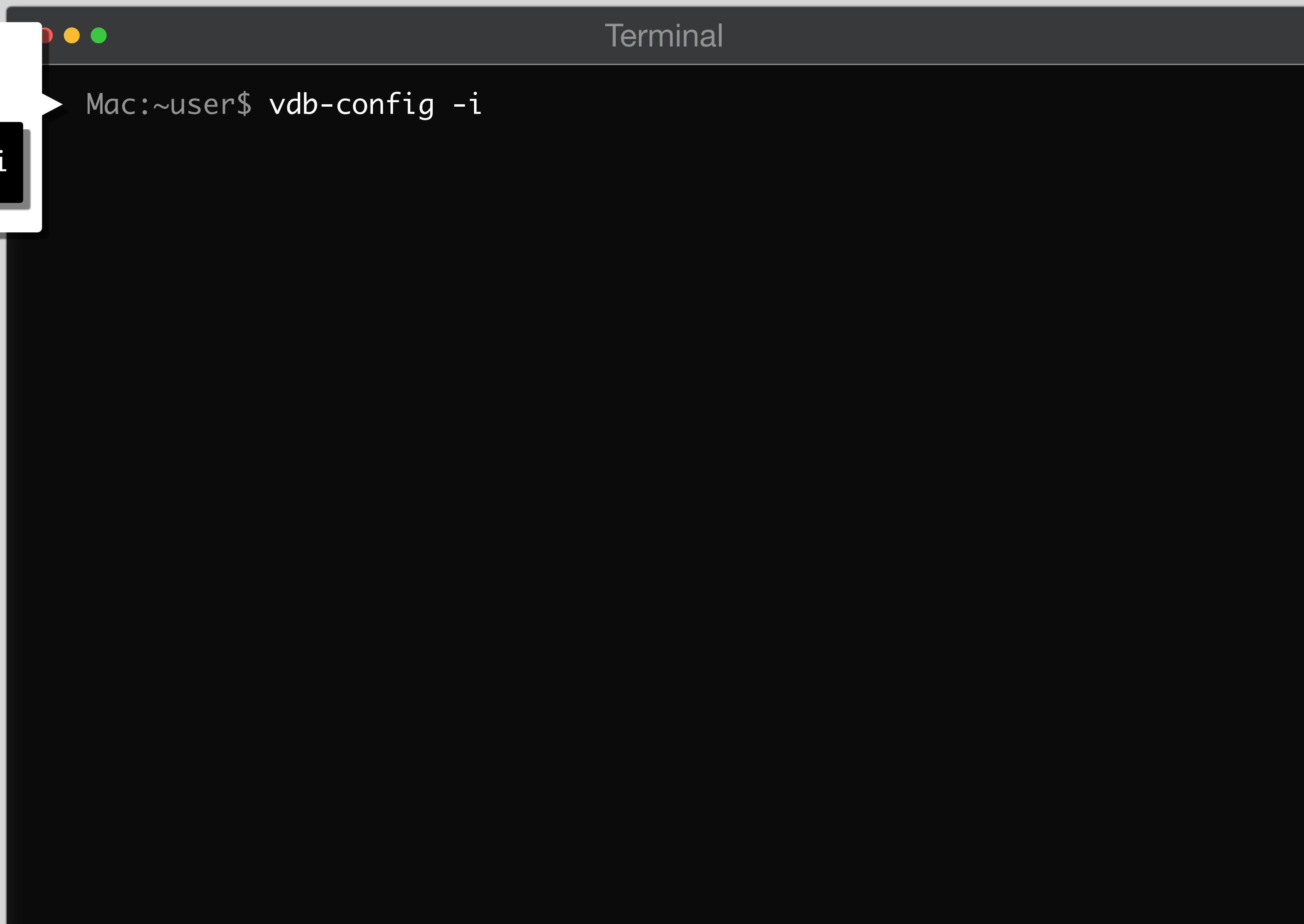
---



# CONFIGURING THE SRA TOOLKIT

Run this command

```
vdb-config -i
```



Mac:~user\$ vdb-config -i

# CONFIGURING THE SRA TOOLKIT

Run this command  
`vdb-config -i`

Press **TAB** to move the red cursor and **ENTER** to select...  
...or just type the underlined Red letters.

Exit by typing '**x**'.

Mac:~user\$ vdb-config -i

Terminal

SRA configuration

[ save ] [ exit ] [ discard ] [ default ]

**MAIN** CACHE AWS GCP NET ITOOLS

[X] Enable Remote Access

[ ] Prefer SRA Lite files with simplified base Quality scores

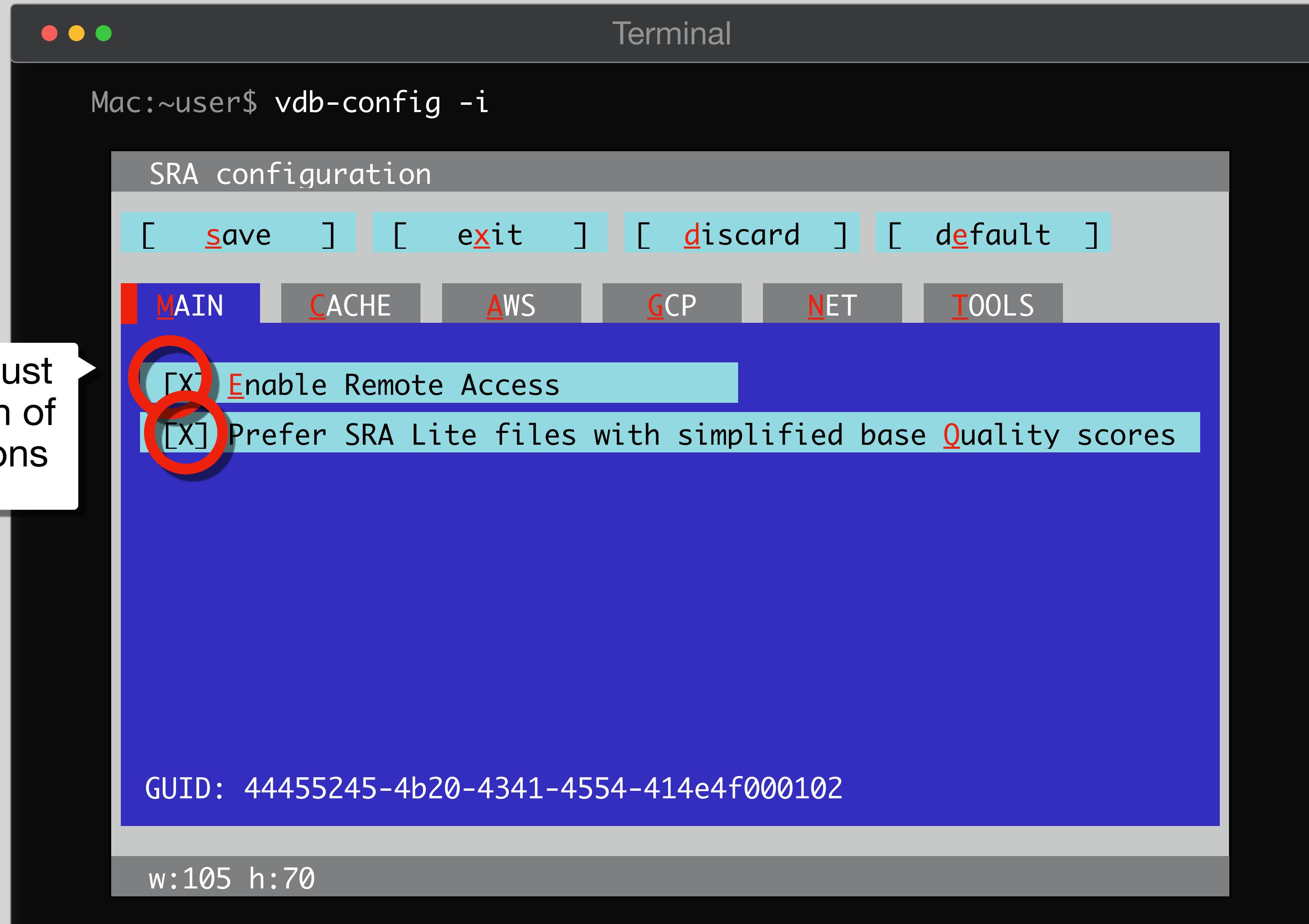
GUID: 44455245-4b20-4341-4554-414e4f000102

w:105 h:70

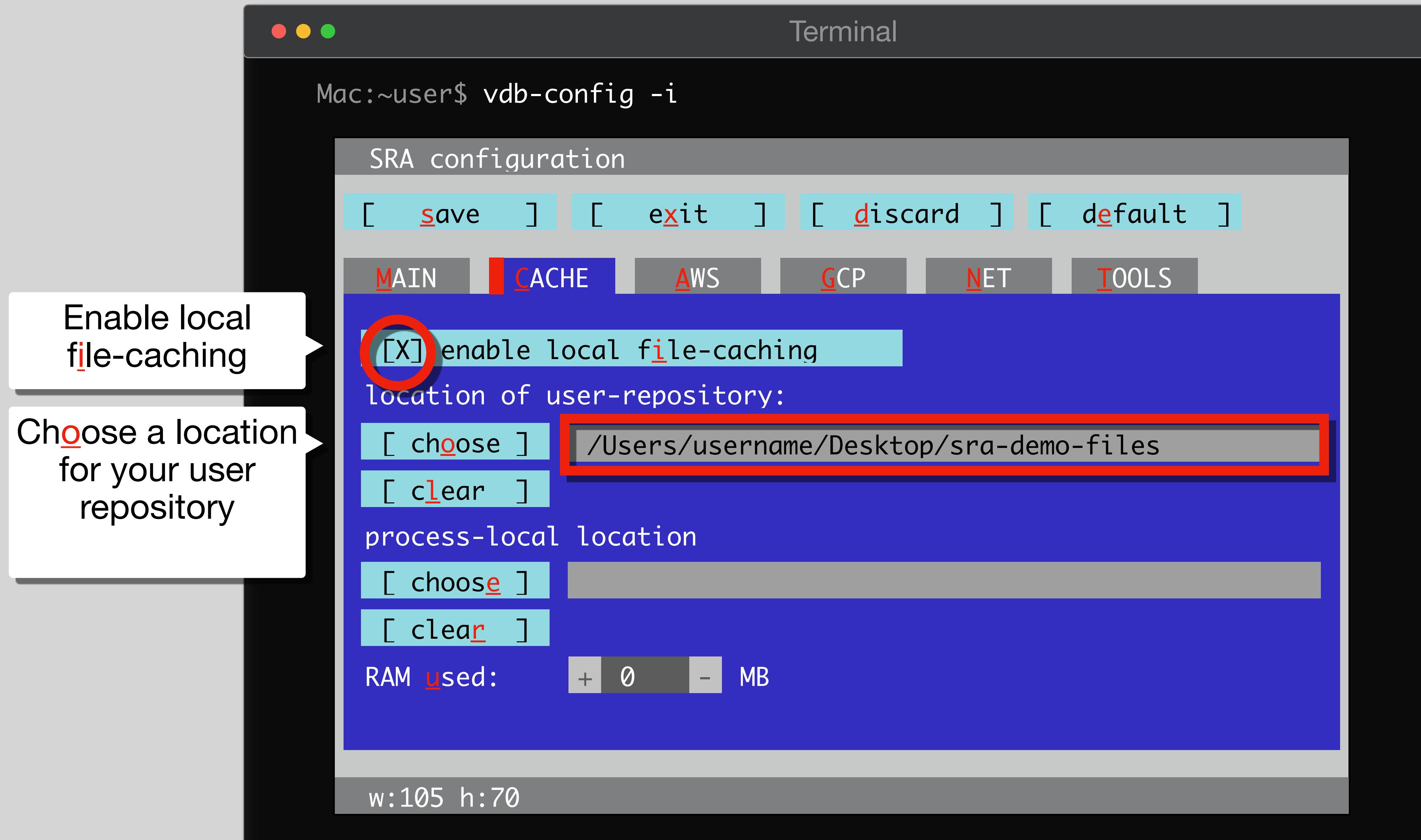
Next we will use this menu to configure our settings in order to access public or controlled-access data.



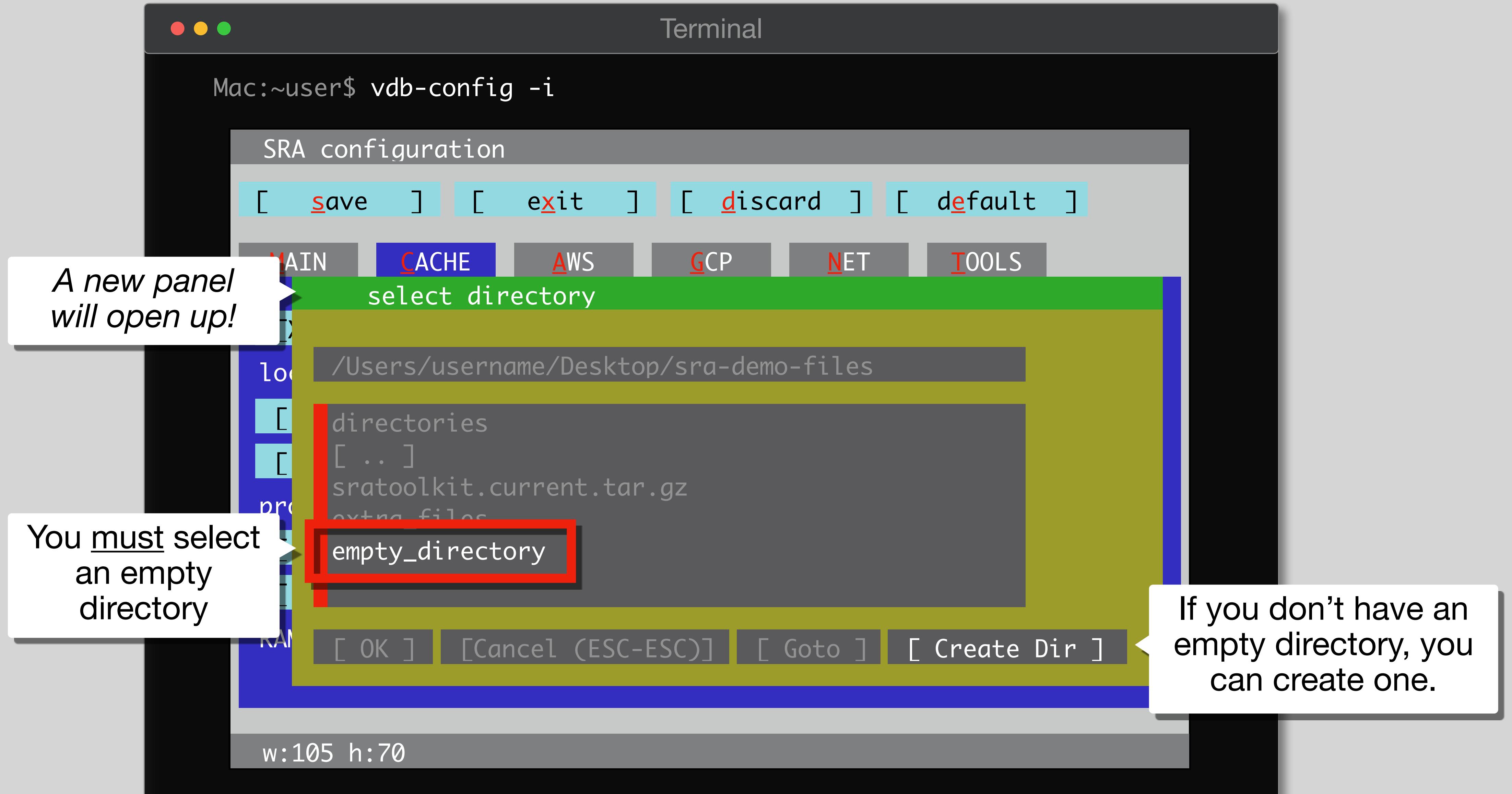
# CONFIGURING THE SRA TOOLKIT



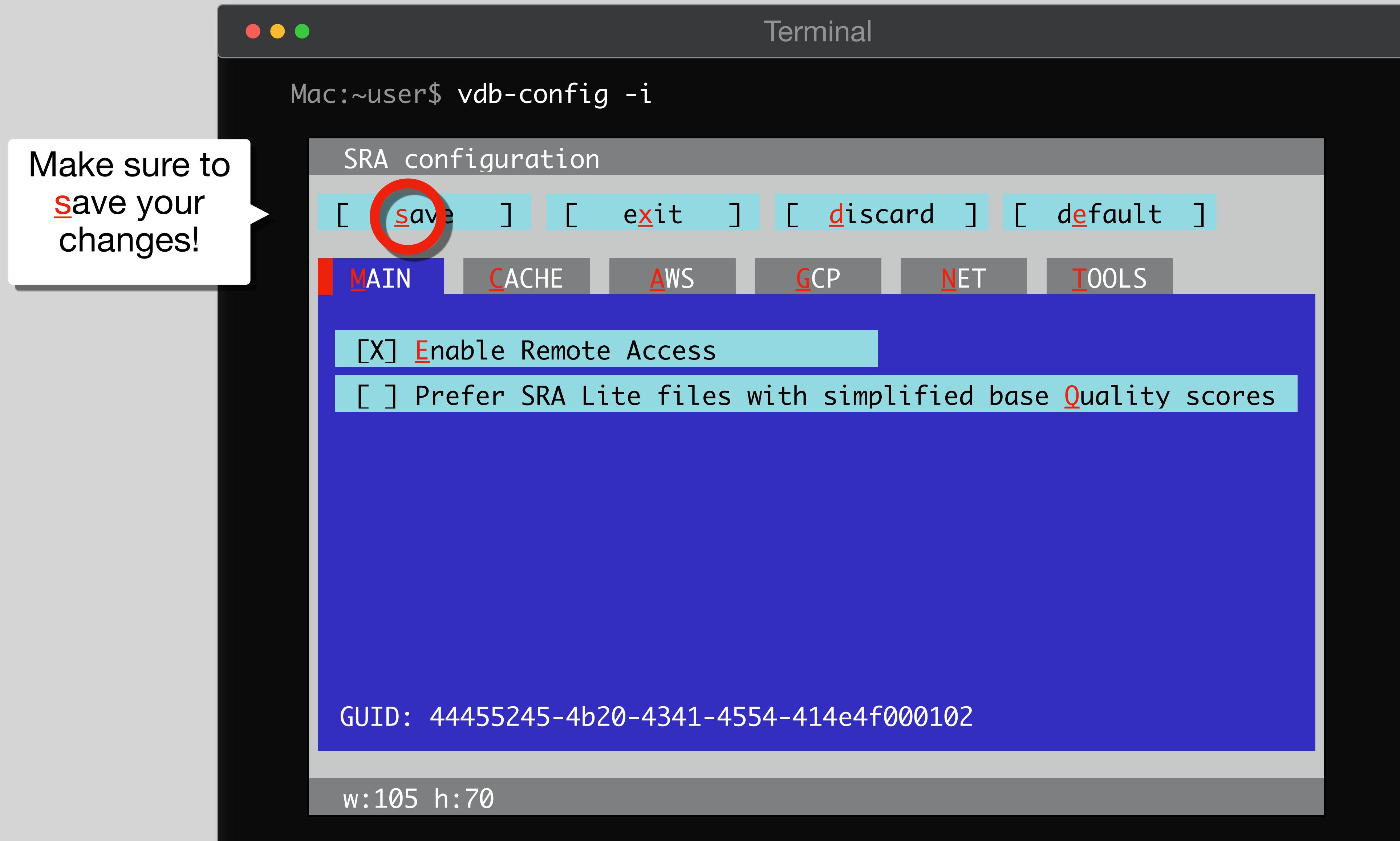
# CONFIGURING THE SRA TOOLKIT



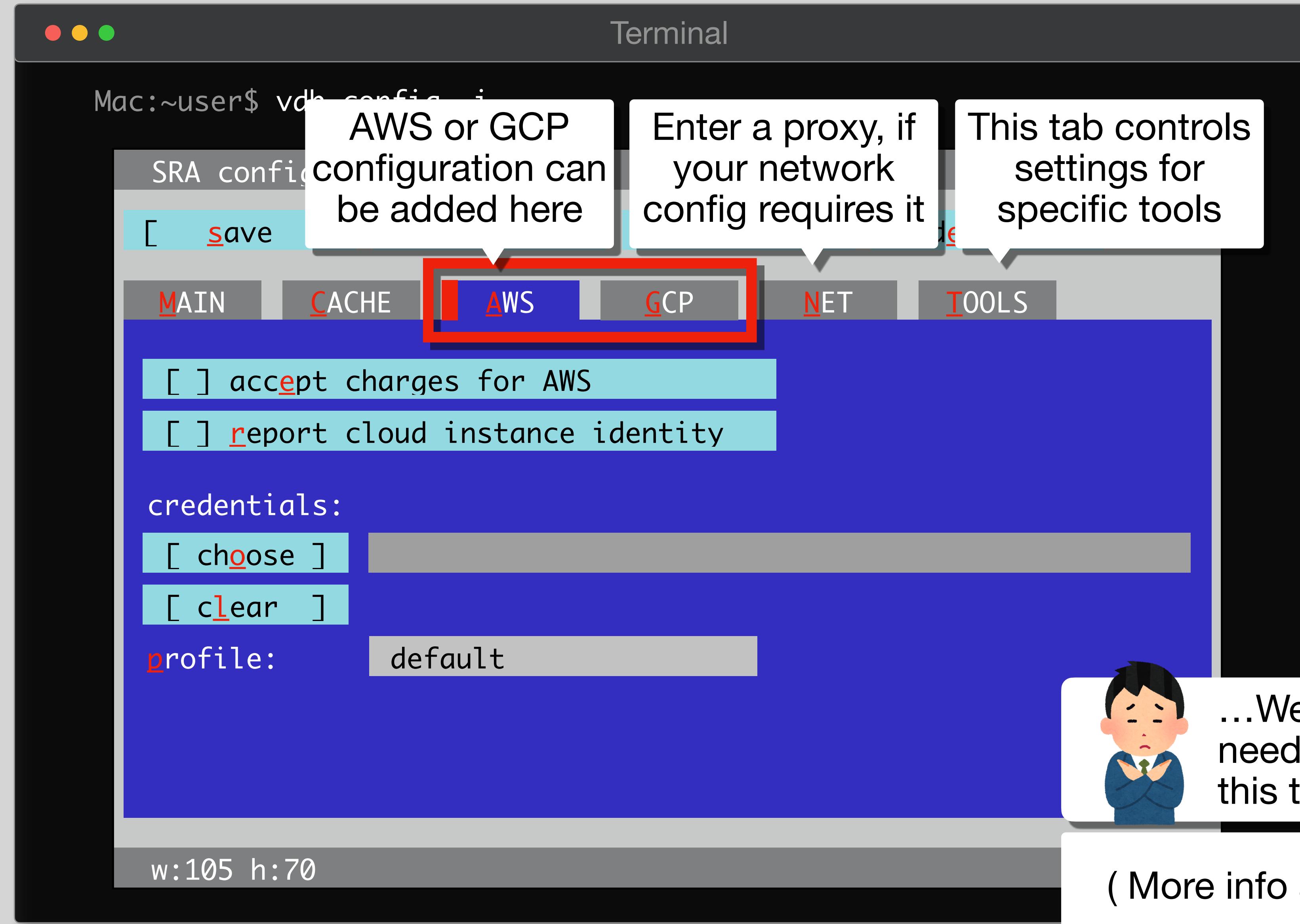
# CONFIGURING THE SRA TOOLKIT



# CONFIGURING THE SRA TOOLKIT



# CONFIGURING THE SRA TOOLKIT



# LET'S REVIEW!

- ① In the **TERMINAL**, run the following command:

```
vdb-config -i
```

This opens the “*SRA Configuration*” menu.

Use **Tab** to navigate the options, and **Enter** to select.

- ② In the **MAIN** menu tab, toggle “**Enable Remote Access**”.

- ③ In the **CACHE** menu tab, make sure that “**Enable Local File Caching**” is toggled.

- ④ Select “**Choose**” under the “*Location of user-repository*” section.

You must select an empty directory, so create one if you do not have one

- ⑤ “**Save**” your changes. Then “**Exit**” the config menu.

INSTALL

CONFIG

USE

# SRA-TOOLKIT

---

DEMONSTRATION

---

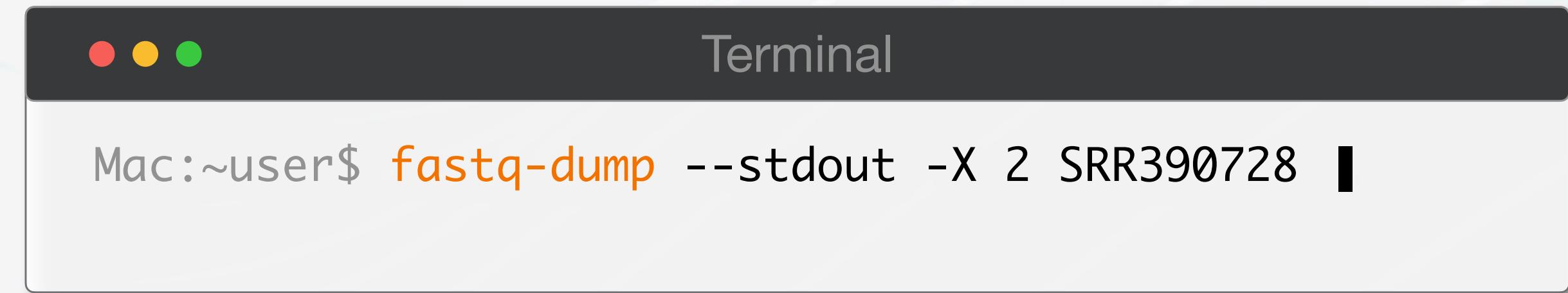


National Library of Medicine  
National Center for Biotechnology Information



# Run commands directly

If the *sratoolkit* has been installed into your *user bin* folder, then just run your commands — it's easy!



A screenshot of a Mac OS X terminal window titled "Terminal". The window has the characteristic red, yellow, and green close buttons at the top left. The title bar is dark grey with the word "Terminal" in a light grey font. The main pane is white and contains the command: "Mac:~user\$ fastq-dump --stdout -X 2 SRR390728 |". The command is written in black text, while the file identifier "SRR390728" is in orange.

## **Possible *bin* locations**

---

 **Windows** `C:\Program Files\sratoolkit\bin\`

 **macOS** `/usr/local/bin/`

 **Linux** `/usr/local/bin/`

# Run commands from a directory

Just ‘`cd`’ into the dir with the extracted `sra toolkit` file.

Run any command inside the `/bin` subdir directly from the command line.

cmd.exe

```
C:\...\> cd Desktop\sra-demo-files\  
  
C:\...\> dir  
  
01/09/2025 03:24 PM <DIR> .  
01/09/2025 03:24 PM <DIR> ..  
01/09/2025 03:24 PM <DIR> sra toolkit.3.2.1-win64  
01/09/2025 03:24 PM           sra toolkit.current-win64.zip  
    2 File(s)        4,352,000 bytes  
    3 Dir(s)   116,170,997,760 bytes free  
  
C:\...\> sra toolkit.3.2.1-win64\bin\fastq-dump --stdout  
-X 2 SRR390728 |
```

Add `\bin\` to your path!

Terminal

```
Mac:~user$ cd Desktop/sra-demo-files/  
  
Mac:~user$ ls  
  
sra toolkit.3.2.1-mac-x86_64  
sra toolkit.3.2.1-mac-x86_64.tar.gz  
  
Mac:~user$ sra toolkit.3.2.1-mac-x86_64/bin/fastq-dump  
--stdout -X 2 SRR390728 |
```

# Testing out our own installation



Let's do a quick test that our SRA Toolkit is functional!

```
C:\...\sratoolkit.3.2.1-win64\bin\ fastq-dump space --stdout space -X 2 space SRR390728
```

*path to the tool  
on your OS  
if necessary*

*download  
sequence  
reads*

*writes  
output  
to the  
Terminal*

*limits  
output  
to first  
2 reads*

*SRA  
accession  
number*

# Testing out our own installation

Let's run that command in the Terminal.



A screenshot of a Mac OS X terminal window titled "Terminal". The window shows the command "fastq-dump --stdout -X 2 SRR390728" being run, followed by its output. The output includes sequencing data for two reads from sample SRR390728. A callout box on the right contains the command text, and another callout box below it provides instructions about the expected output.

```
Mac:~user$ fastq-dump --stdout -X 2 SRR390728
Read 2 spots for SRR390728
Written 2 spots for SRR390728
@SRR390728.1 1 length=72
CATTCTTCACGTAGTTCTCGAGCCTGGTTTCAGC
GATGGAGAATGACTTGACAAGCTGAGAGAAGNTNC
+SRR390728.1 1 length=72
?????????????????????????????????????
?????????????????????????????????????
@SRR390728.2 2 length=72
AAGTAGGTCTCGTCTGTGTTTCTACGAGCTTGTGT
TCCAGCTGACCCACTCCCTGGGTGGGGGACTGGGT
+SRR390728.2 2 length=72
?????????????????????????????????????
?????????????????????????????????????
Mac:~user$ |
```

Run this command

```
fastq-dump  
--stdout  
-X 2  
SRR390728
```

The output should look like this.

# HELP! (What do these tools do?)

Run this option...

toolname --help

...or this one:

toolname -h

The `--help` option (also known as `-h`) provides a quick overview of a tool.

Includes

- usage
- syntax
- inputs/outputs
- options

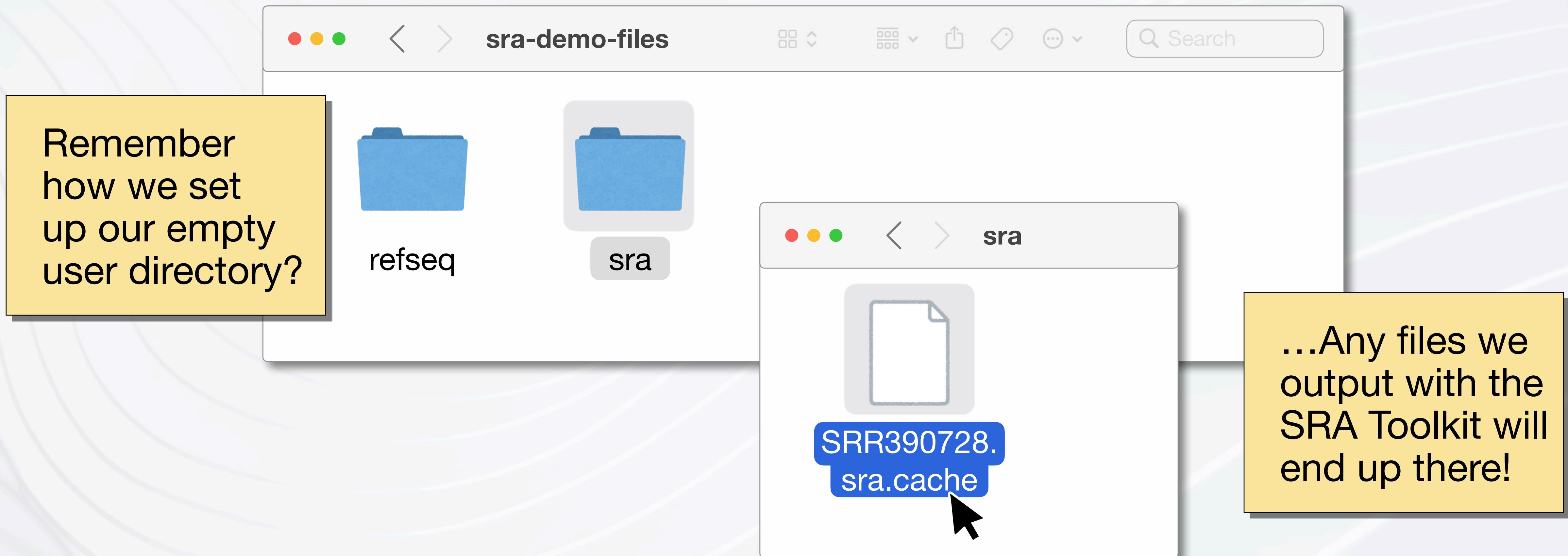
```
Mac:~user$ fastq-dump --help
```

Usage:  
`fastq-dump [options] <path> [<path>...]`  
`fastq-dump [options] <accession>`

INPUT  
`-A<accession>` Replaces accession derived from <path> in filename(s) and deflines (only for single table dump)  
`--table <table-name>` Table name within cSRA object, default is "SEQUENCE"

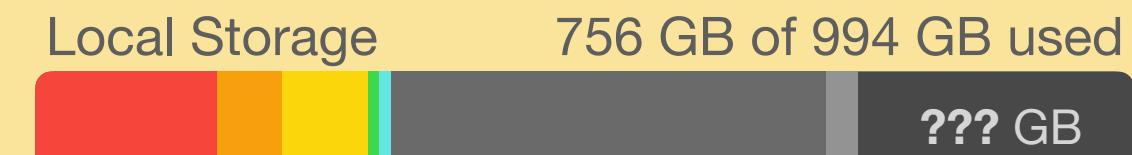
PROCESSING  
Read Splitting  
`--split-spot` Sequence data may be used in raw form or split into individual reads  
Split spots into individual reads  
Full Spot Filters  
Applied to the full spot independently of `--split-spot`

# Where do our files go?



# How much disk space do you have?

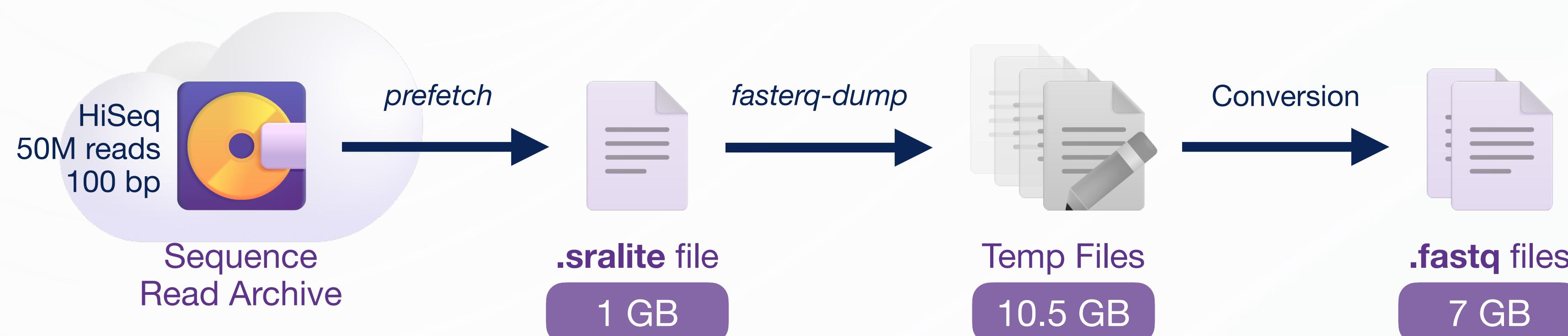
The SRA Toolkit  
requires significant  
disk space!



`df -h`

Checks your  
file system  
disk space

You will need ~18x the  
size of your  
original accession.



**18 GB**

Total space needed for conversion

# Check how much space we need

vdb-dump

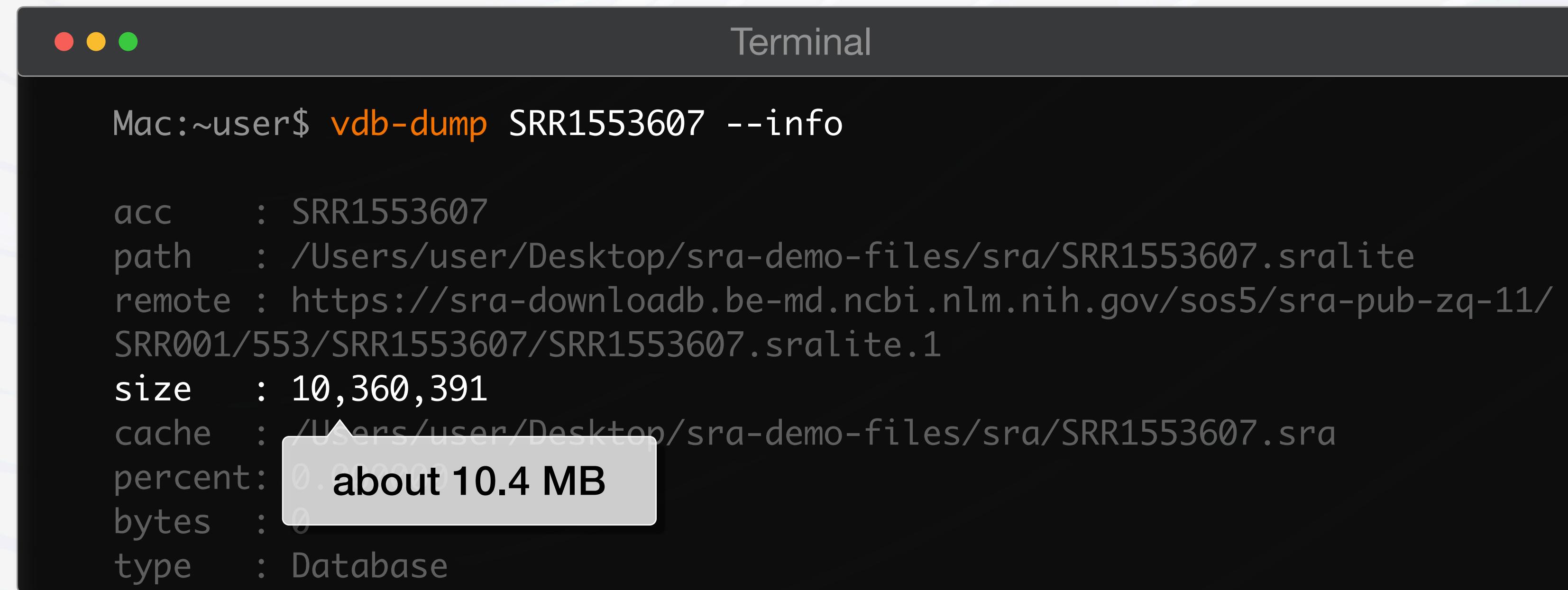
*Output the native  
VDB format of SRA  
data*

SRR1553607

*SRA  
accession  
number*

--info

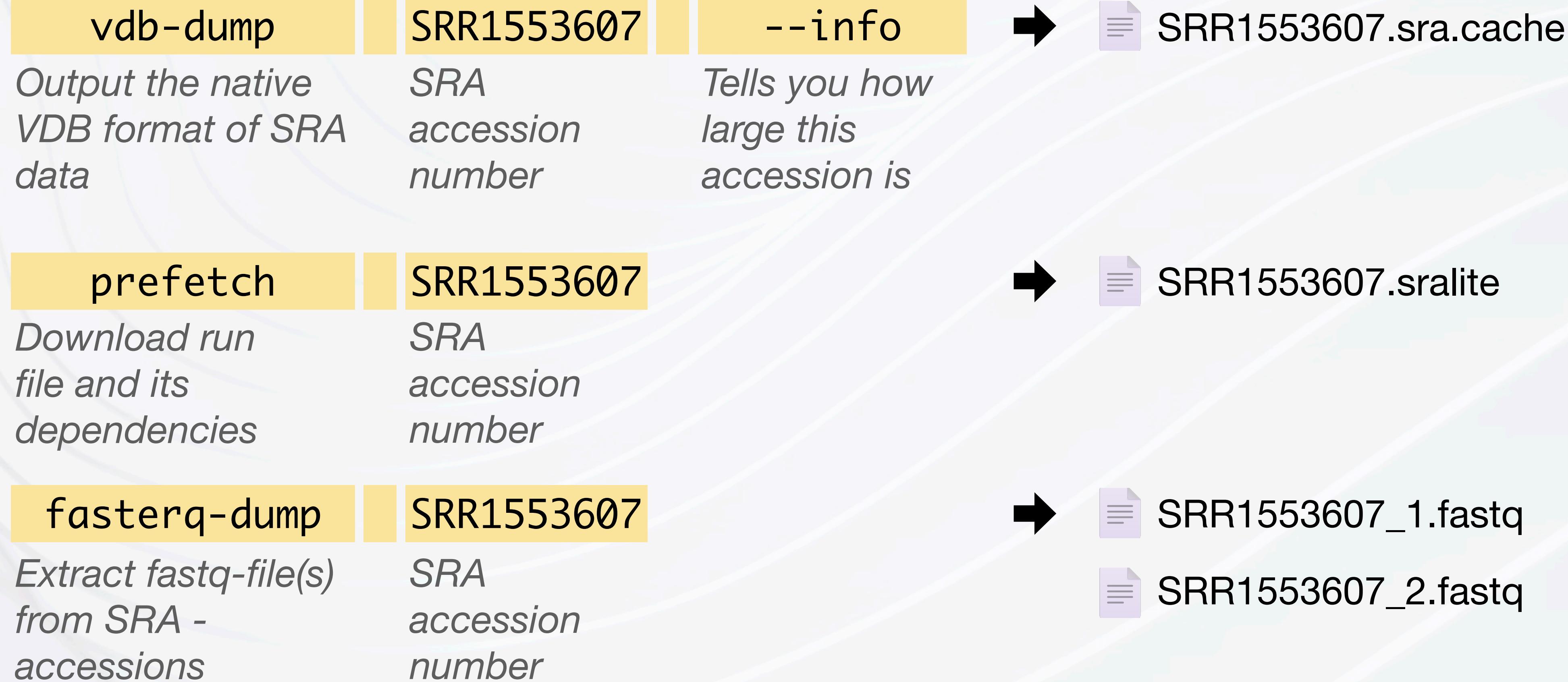
*Tells you how  
large this  
accession is*



Mac:~user\$ vdb-dump SRR1553607 --info

```
acc      : SRR1553607
path     : /Users/user/Desktop/sra-demo-files/sra/SRR1553607.sralite
remote   : https://sra-downloadb.be-md.ncbi.nlm.nih.gov/sos5/sra-pub-zq-11/
           SRR001/553/SRR1553607/SRR1553607.sralite.1
size     : 10,360,391
cache    : /Users/user/Desktop/sra-demo-files/sra/SRR1553607.sra
percent  : 0. about 10.4 MB
bytes    : 0
type     : Database
```

# Testing out our own installation



# LET'S REVIEW!

- ① In the **TERMINAL**, run the following command:

```
fastq-dump --stdout  
-X 2 SRR390728
```

This is used for downloading sequence reads.

- ② Run the following command:

```
vdb-dump SRR1553607  
--info
```

Note the info presented in the Terminal.

- ③ Run the following command:

```
prefetch SRR1553607
```

This will download the run file as an **.sralite** file.

- ③ Run the following command:

```
fasterq-dump SRR1553607
```

This will extract the fastq-file to your current dir.

# HOW TO SUBMIT



# SRA SUBMISSION

ncbi.nlm.nih.gov

**National Library of Medicine**  
National Center for Biotechnology Information

All Databases  Search  Log in

**NCBI Home**

- Resource List (A-Z)
- All Resources
- Chemicals & Bioassays
- Data & Software
- DNA & RNA
- Domains & Structures
- Genes & Expression
- Genetics & Medicine
- Genomes & Maps
- Homology
- Literature
- Proteins
- Sequence Analysis
- Taxonomy
- Training & Tutorials
- Variation

**Welcome to NCBI**

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

[About the NCBI](#) | [Mission](#) | [Organization](#) | [NCBI News & Blog](#)

**Submit**  
Deposit data or manuscripts into NCBI databases 

**Download**  
Transfer NCBI data to your computer 

**Learn**  
Find help documents, attend a class or watch a tutorial 

**Develop**  
Use NCBI APIs and code libraries to build applications 

**Analyze**  
Identify an NCBI tool for your data analysis task 

**Research**  
Explore NCBI research and collaborative projects 

**Popular Resources**

- PubMed
- Bookshelf
- PubMed Central
- BLAST
- Nucleotide
- Genome
- SNP
- Gene
- Protein
- PubChem

**The Submission Portal** can be found on the main NCBI page.

**NCBI News & Blog**

- An Updated Bacterial and Archaeal Reference Genome Collection is Available! 02 Sep 2025 [Download the updated bacterial and archaeal reference genome collection](#)
- GenBank Release 268.0 is Available! 26 Aug 2025 [View details about GenBank Release 268.0](#)
- GenBank release 268.0 (8/18/2025) is now available on the NCBI FTP site. This release has 47.01 trillion bases and 5.90 [Upcoming Changes to GenBank Project Lists and Symlinks](#)

# SRA SUBMISSION

The screenshot shows the NCBI Submission Portal interface for Sequence Read Archive (SRA) submissions. At the top, it says "Submission Portal" and "Submit to the world's largest public repository of biological and scientific information". Below this, there is a search bar with placeholder text "Enter sequence type" and a "Suggest tool" button. A row of buttons includes "SARS-CoV-2", "16S rRNA", "genome", "ITS", and "SRA". The main content area is titled "Sequence Read Archive (SRA)" and describes it as the largest publicly-available repository of high throughput sequencing data. It accepts data from all branches of life and environmental surveys. A callout box points to the "Learn more" button, which is highlighted with a red border. Another callout box points to the "Submit" button, also highlighted with a red border.

ncbi.nlm.nih.gov

**Submission Portal**  
Submit to the world's largest public repository of biological and scientific information

Type a few words about the sequence data you are submitting and select an option to learn more. You can also browse submission information below.

**What do you want to submit?**

Enter a few words about your sequence data.

Enter sequence type  Suggest tool

SARS-CoV-2 16S rRNA genome ITS SRA

**Sequence Read Archive (SRA)**

SRA is the largest publicly-available repository of high throughput sequencing data. The archive accepts data from all branches of life as well as metagenomic and environmental surveys.

Submit unassembled, high throughput sequencing reads

[SARS-CoV-2 submission instructions](#)

**Learn more** **Submit**

See the “**Learn more**” button?  
Click it for detailed instructions on what you’ll need to submit!

Scroll down for Sequence Read Archive submission.

Click the “**Submit**” button to proceed.

# SRA SUBMISSION

ncbi.nlm.nih.gov

**Submission Portal**

**Sequence Read Archive (SRA)**

The SRA accepts genetic data and the associated quality scores produced by next generation sequencing technologies. Please refer to the [File Format Guide](#).

- Files can be compressed using **gzip** or **bzip2**, and may be submitted in a tar archive but archiving and/or compressing your files is not required. **Do not use zip!**
- All file names must be **unique** and **not contain any sensitive information**. File names as submitted appear publicly in the Google and AWS clouds.
- **Each file must be listed** in the SRA metadata table. If you are uploading a tar archive, list each file name, not the archive name.
- Use the **preload** option if you are uploading files **>10 GB** or **>300 files**. All files for a submission must be uploaded into a **single folder that is associated only with a single submission**.

**Submission Portal**

**Sequence Read Archive (SRA) submission: SUB15607954**

New

1 SUBMITTER > 2 GENERAL INFO > 3 SRA METADATA > 4 FILES > 5 REVIEW & SUBMIT

**Submitter**

\* First (given) name Middle name \* Last (family) name

\* Email (primary) Email (secondary)

Required fields are marked with \* asterisk

Click on “New Submission”.

Then follow the on-screen instructions!

It couldn't be any easier.

# SRA SUBMISSION

ncbi.nlm.nih.gov

**Submission Portal**

Home My submissions Manage data Templates My profile

**Sequence Read Archive (SRA) submission: SUB15607954**

New

1 SUBMITTER 2 GENERAL INFO 3 SRA METADATA 4 FILES 5 REVIEW & SUBMIT

**General Information**

Required fields are marked with \* asterisk

**BioProject**

i BioProject describes the goal of your research effort.

\* Do you already have a BioProject accession number for this research?

Yes  No (a BioProject will be created within this submission)

\* Existing BioProject

PRJNAXXXXX

**BioSample**

i The BioSample records the detailed biological and physical properties of the sample that was sequenced. A BioSample can be used in more than one BioProject since it should be used for all the data that were obtained from that sample. Usually SRA data sets are generated from more than one sample.

\* Do you already have BioSample accession numbers for these samples?

Yes  No (BioSamples will be created within this submission)

**Release date**

2024-01-01

You'll be asked to provide info like your **BioProject** (if you have one), your **BioSample** (if you already have one), and even the date you want your submission to be made public.

# SRA SUBMISSION

ncbi.nlm.nih.gov

Submission Portal

Home My submissions Manage data Templates My profile

Sequence Read Archive (SRA) submission: SUB15607954

Delete submission

New

1 SUBMITTER > 2 GENERAL INFO > 3 PROJECT INFO > 4 BIOSAMPLE TYPE > 5 BIOSAMPLE ATTRIBUTES > 6 SRA METADATA > 7 FILES >

8 REVIEW & SUBMIT

Project Info

Required fields are marked with \* asterisk

\* Project title ?

\* Public description ?  
4000 characters allowed

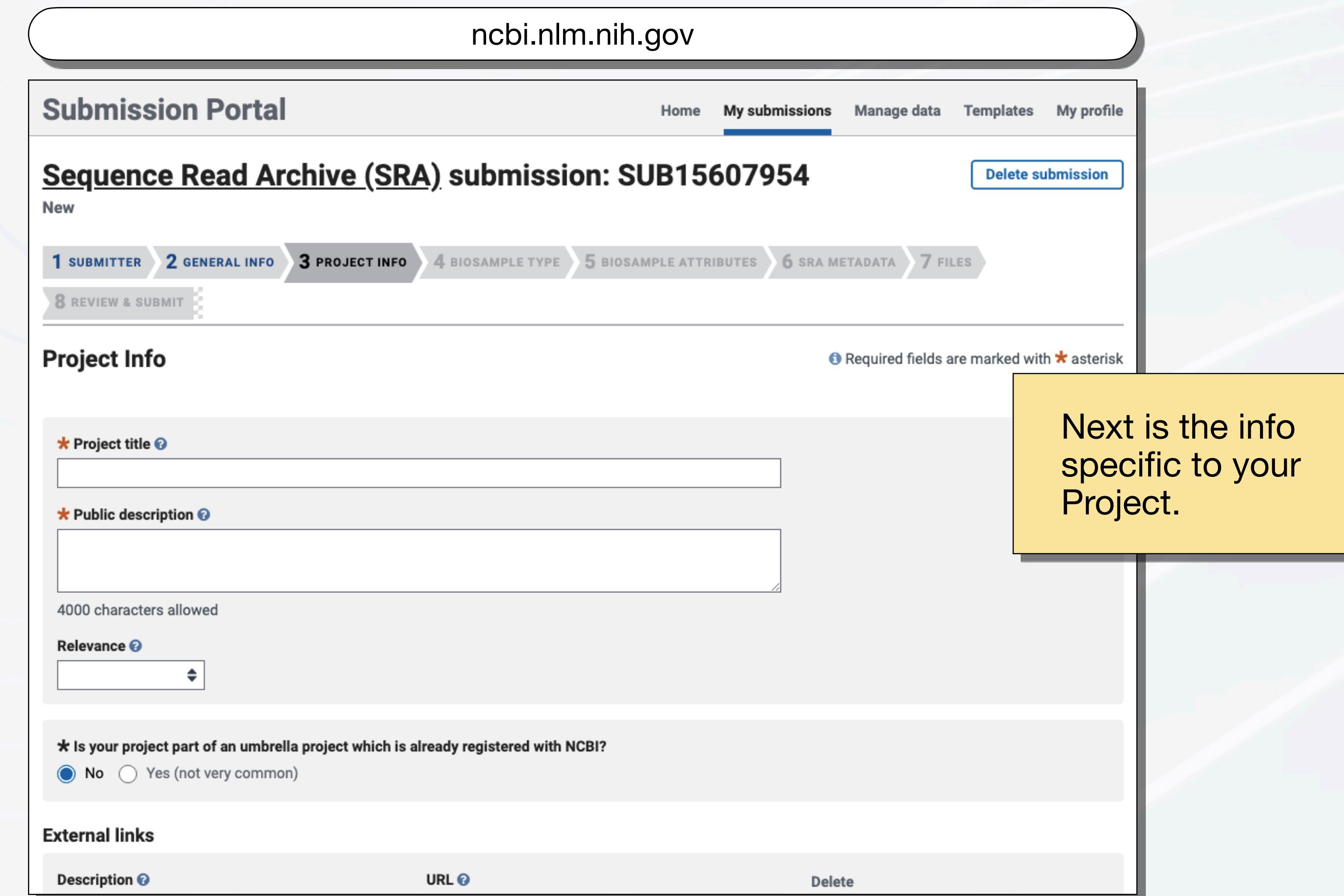
Relevance ?

\* Is your project part of an umbrella project which is already registered with NCBI?  
 No  Yes (not very common)

External links

Description ? URL ? Delete

Next is the info specific to your Project.



# SRA SUBMISSION

ncbi.nlm.nih.gov

**Sample Type** \* Required fields are marked with \* asterisk

\* Select the package that best describes your samples.

All packages   Packages for MAG submitters   Packages for metagenome submitters

(Optional) Filter packages by organism name  
Enter the full scientific name of your samples, e.g., Escherichia coli Reset and show all packages

Want BioSample packages:  
scientific name of the organism of your samples.  
Samples are derived from a species **not represented in NCBI's Taxonomy database**, enter the genus-level name, e.g., *Escherichia*.  
Samples are derived from **more than one organism**, enter the common species, genus, or family, e.g., *Enterobacteriaceae*.  
Samples are **metagenomic/environmental**, or **metagenome-assembled genomes (MAG)**, select the appropriate tab above.  
For more information about organism names, see [Organism information](#).

**NCBI packages** [More...](#)

- SARS-CoV-2: clinical or host-associated**  
Use for SARS-CoV-2 samples that are relevant to public health. Required attributes include those considered useful for the rapid analysis and trace back of SARS-CoV-2 cases.
- SARS-CoV-2: wastewater surveillance**  
Use for SARS-CoV-2 wastewater surveillance samples that are relevant to public health. Required attributes include those considered useful for the rapid analysis and trace back of SARS-CoV-2 cases.
- Pathogen**  
Use for pathogen samples that are relevant to public health. Required attributes include those considered useful for the rapid analysis and trace back of pathogens.
- One Health Enteric**  
Use for microbial isolates that are collected for genomic

**GSC MiS packages for genomes, metagenomes, and marker sequences** [More...](#)

- MiGS Cultured Bacterial/Archaeal**  
Use for cultured bacterial or archaeal genomic sequences. Organism must have lineage [Bacteria](#) or [Archaea](#).
- MiGS Eukaryotic**  
Use for eukaryotic genomic sequences. Organism must have lineage [Eukaryota](#).
- MiGS Viral**  
Use for virus genomic sequences. Organism must have lineage [Viruses](#).
- MIMAG Metagenome-assembled Genome**  
Use for metagenome-assembled genome sequences produced using computational binning tools that group sequences into individual organism genome assemblies starting from metagenomic data sets. Organism cannot contain the term 'metagenome'. Use the [MILVIC package](#) for virus genomes.

Note that some samples require very specific requirements outlined in their description.

# SRA SUBMISSION

ncbi.nlm.nih.gov

**Submission Portal**

Home My submissions Manage data Templates My profile

**Sequence Read Archive (SRA) submission: SUB15607954**

Test, Sep 06 '25

**1 SUBMITTER** > **2 GENERAL INFO** > **3 PROJECT INFO** > **4 BIOSAMPLE TYPE** > **5 BIOSAMPLE ATTRIBUTES** > **6 SRA METADATA** > **7 FILES**

**8 REVIEW & SUBMIT**

**Attributes**

Package Invertebrate; version 1.0

Required fields are marked with **\*** asterisk.  
At least one of the fields marked with **\*\***, **††** or **‡‡** is required.

**\* How do you want to provide your BioSample attributes?**

Use built-in table editor  
 Upload a file using Excel or text format (tab-delimited) that includes the attributes for each of your BioSamples

**Continue**

**FOLLOW NCBI**

X f in GitHub RSS

Connect with NLM National Library of Medicine Web Policies FOIA Help Accessibility

# SRA SUBMISSION

ncbi.nlm.nih.gov

## Submission Portal

Submit to the world's largest public repository of biological and scientific information

Type a few words about the sequence data you are submitting and select an option to learn more. You can also browse submission information below.

### What do you want to submit?

Enter a few words about your sequence data.

🔍 Suggest tool

[SARS-CoV-2](#) [16S rRNA](#) [genome](#) [ITS](#) [SRA](#)

## Sequence Read Archive (SRA)

SRA is the largest publicly-available repository of high throughput sequencing data. The archive accepts data from all branches of life as well as metagenomic and environmental surveys.

See the “**Learn more**” button?

Submit unassembled, high throughput sequencing reads

[SARS-CoV-2 submission instructions](#)

Learn more Submit

Click it for detailed instructions on what you'll need to submit!

# RESOURCES



# GET MORE INFO

## Sequence Read Archive – Main Page

ncbi.nlm.nih.gov/sra

**SRA - Now available on the cloud**

Sequence Read Archive (SRA) data, available through multiple cloud providers and NCBI servers, is the largest publicly available repository of high throughput sequencing data. The archive accepts data from all branches of life as well as metagenomic and environmental surveys. SRA stores raw sequencing data and alignment information to enhance reproducibility and facilitate new discoveries.

**Getting Started**

- [Documentation](#)
- [How to submit](#)
- [How to search and download](#)
- [How to use SRA in the cloud](#)
- [Submit to SRA](#)

**Tools and Software**

- [Download SRA Toolkit](#)
- [SRA Toolkit Documentation](#)
- [SRA-BLAST](#)
- [SRA Run Browser](#)
- [SRA Run Selector](#)

**Related Resources**

- [Submission Portal](#)
- [dbGaP Home](#)
- [BioProject](#)
- [BioSample](#)

## SRA Data Formats Manual

doi.org/10.5281/zenodo.15677383

**Sequence Read Archive (SRA) Data**  
Technical User Manual

Document: v1.1  
SRA Toolkit: v3.2.1  
Last updated: June 16, 2025

[www.ncbi.nlm.nih.gov/sra](#)

# Acknowledgements

Anne Apuzzo  
Inna Belaia  
Anatoly Boshkin  
Rodney Brister  
Daniel Broderick  
Derek Caetano-Anollés  
Olivia Carl  
Chad Charowhas  
Regan Checcio  
Kristen Cleveland  
Ryan Connor  
Yamini Damodaran  
Amanda Davis  
Daniel Deriy  
Lauren Desmond  
Marcin Domagalski  
Michael Domrachev  
Kenneth Durbrow

Alexandre Efremov  
Ravinder Eskandary  
Matthew Garafalo  
Priyanka Ghosh  
George Godynskiy  
Ivan Goranov  
Kathleen Griswold  
Jonathan Gunti  
Derek Hart  
Denise Hicks  
Svetlana Iazvovskaia  
Laura Iwig  
Kenneth Katz  
Fedor Karmanov  
Roman Katargin  
Michael Kimelman  
Dmitrii Kishchukov  
Andrew Klymenko

This research was supported by the Intramural Research Program of the NIH, National Library of Medicine.

Andrey Kochergin  
Anatoliy Kuznetsov  
Richard Lapoint  
Martin Leach  
Connor Lee  
Sarib Mahmood  
Anatoly Mnev  
Chris O'Sullivan  
Kevin Palis  
Sergey Ponomarov  
Wolfgang Raetz  
Diamond Ringo  
Juan Riveros  
Charlie Seto  
Andrei Shkeda  
Yuriy Skripchenko  
Jennifer Smith  
Adam Stine

Nicholas Stoler  
Kamen Todorov  
Jonathan Trow  
Alexandr Ukrainchik  
Mike Vartanian  
Oleg Volonsevich  
Moyo Williams  
Eugene Yaschenko  
Kaipeng Yu  
Vadim Zalunin

## Additional Support

*NIH Capital Investment Fund*  
*NIH STRIDES Initiative*

# GIVE US YOUR FEEDBACK!

This 3-min feedback form  
is **anonymous** and goes  
directly to the instructors.

3:00



Scan it.

Or type it.

Or click it.