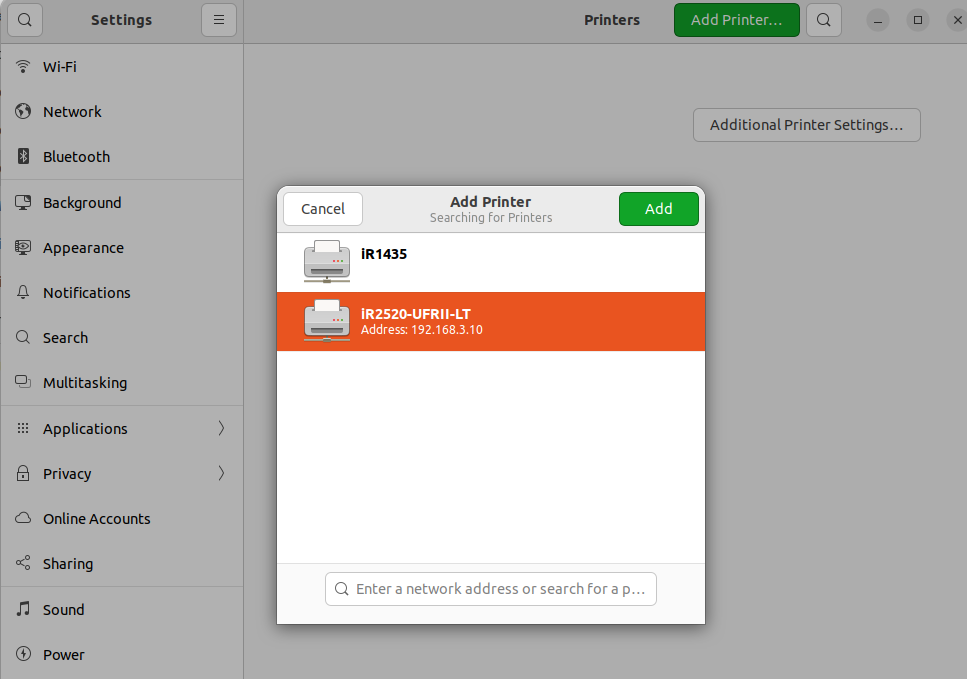
# 0. Unlock Directory from lock sign;

sudo chown -R username directory name/

# 0.1 Printer Setup:

1. Google search https://asia.canon/en/support/0100924010
2. Download
3. Extract the tar.gz file and install : sudo /linux-UFRII-drv-v560-m17n-10/linux-UFRII-drv-v560-m17n/install.sh
4. Now add printer below picture: IP: 192.168.3.10

# 0.2 Creating alias:

### When suppose you are using a long command or software to open from its original folder location, to remove the long path moving you create a small name and run its from anywhere at your scripts or terminal

#### Create a bin folder at home

#### If Software

Move the folder of your software - copy the install path location

#### 2. If file

copy the file path

#### Save the file path at bashrc by a small name

Alias small\_name=’file path’

# 1.scp : for file transfer

### Example Format:

scp -r file\_name\_that\_will\_transfer user\_name\_destination\_pc@ip\_destination\_pc**:**destination\_path

#### Example:

1. Example for **office laptop** to **office server**:

scp -r **file\_name** chrf@192.168.3.112:destination\_path

Password: md5(chrf)

1. Example for **office server** to **office laptop**:

scp -r **file\_name** chrf@192.168.3.58:/home/chrf/Desktop/CHRF\_Project

Password:meta

1. Example for **office laptop** to **my laptop**:

rf

scp -r **file\_name** preonath@192.168.3.35:destination\_path

Password:shu

1. Example for **my laptop** to **office laptop**:

scp -r **file\_name** chrf@192.168.3.109:destination\_path

Password:meta

# 3.Covid Sample upload from Basespace to CZID:

## Loin to CZID(shorturl.at/EQUXZ)

## Select Project

#### Upload option

* + - 1. Project > CHRF (Private)

## Analysis Type

* + - 1. SARS-CoV-2 Consensus Genome option
      2. Illumina
      3. COVIDseq

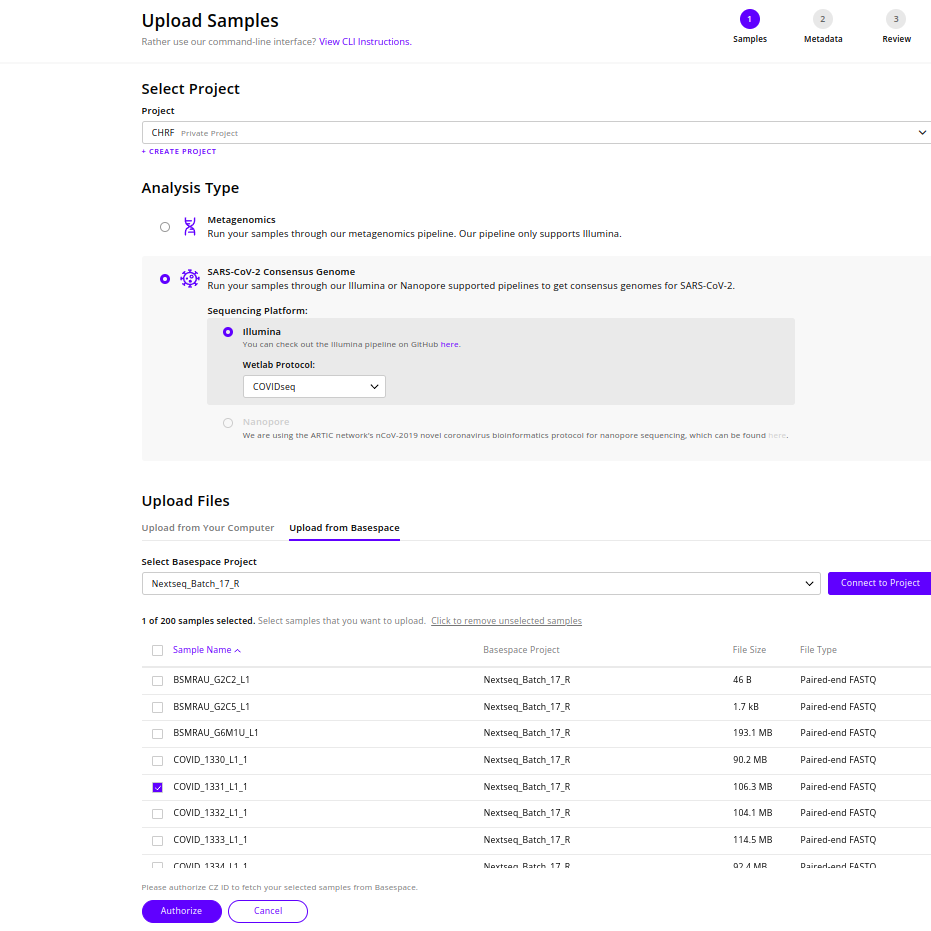
## Upload Files

#### Upload from Basespace

#### Connect to Basespace

#### Select Basespace Project

* + - * 1. Nextseq\_Batch\_17\_R(R=repeats)
        2. Unselect all Sample
        3. Select only Covid Sample
        4. Authorization



# 3.Nextstrain

### Example data

#### 2.1 **Installation :**

* Documents

<https://docs.nextstrain.org/en/latest/install.html>

Or

* Scripts: https://drive.google.com/file/d/15JUrOwgLtjyUXffBLCj8gzhauDzTSF22/view?usp=sharing

#### 2.2 Nextstrain build for novel coronavirus SARA-CoV2(ncov)

* Git clone from Nextstarin Github:

git clone <https://github.com/nextstrain/ncov.git>

This cloning folder of ncov has missed my-profiles material. So download this

material from my drive.

* Open build.yaml file and change the input file path ( metadata and sequences)

**Format:**

**inputs:**

**- name: local-data**

**metadata: data/example\_metadata.tsv**

**sequences: data/example\_sequences.fasta.gz**

* Open terminal = into ncov

nextstrain build . –cores all –configfiles .my-profile/example/builds.yaml

* **Output = auspice folder**

#### 

#### 

#### 2.3 Visualization:

* Open **auspice.us** from chrome & minimize
* Open **auspice folder** and take ncov\_bangladesh.json file on >
  + - * + Drag & Drop a Dataset on Here to view

### Real data

#### Data Download:

* Download All world Data:
  + GISAID > Login > search option >
    - Collection
      * Start Date: Default
      * End Date: 15 Nov 2021 then (16 Nov 2021 to Today date)
    - CompleteDefinition: GISAID considers genomes >29,000nt as complete and further assigns labels of high coverage <1% Ns (undefined bases) and low coverage >5% Ns.
    - Collection date completeDefinition: Only entries with complete collection date
  + Now Download as input augur pipeline format
  + Merge them
    - Merge Metadata: Scripts > https://www.biostars.org/p/485568/
    - Merge Fasta Files: cat \*.fasta > sequences.fasta

* Download Asia Data:
  + GISAID > Login > download option > nextregions > Asia

#### Main Folder:

1. Data Folder:
   1. sequences.fasta
   2. metadata.tsv
   3. hcov\_asia.fasta
   4. hcov\_asia.json
   5. hcov\_asia.tsv
   6. Readme.txt
2. Nextstrain-data-sampling-main: My Drive (shorturl.at/syX02) Or CHRF github
3. ncov : Download from Nextstrain github(<https://github.com/nextstrain/ncov>) or My drive

#### Activation Conda Environment :

* conda activate nextstrain

#### Run Pipeline = pipeline.sh

* Input :
  + Data Folder:
* Output
  + metadata\_before\_sampling.tsv
  + Stat\_before\_sampling.txt

**Run This step**

cd 6-2-2022

grep -v "CHRF-0369" metadata\_before\_sampling.tsv > sampler\_input.tsv

cd -

python3 new\_sampler.py \

--input ./6-2-2022/sampler\_input.tsv \

--samples 1121 \

--recent\_days 90 --output ./6-2-2022/metadata\_after\_sampling.tsv

######################

cd 6-2-2022

tail +2 metadata\_after\_sampling.tsv | cut -f 1 > bd\_seq\_list.txt

cd -

python3 BD\_sequence\_select.py --id\_file ./6-2-2022/bd\_seq\_list.txt --input ./6-2-2022/sequences.fasta\

--output ./6-2-2022/bd\_sequences.fasta

##############################

cd 6-2-2022

cat metadata\_after\_sampling.tsv hcov\_bd\_from\_asia.tsv > bd\_meta.tsv

cd -

########################

python3 dis\_div\_fixer.py --input ./6-2-2022/bd\_meta.tsv --output ./6-2-2022/sampled\_with\_division\_fixed.tsv

##########################

python3 post\_processer.py --input ./6-2-2022/sampled\_with\_division\_fixed.tsv \

--ref\_origin\_lab ./list\_of\_originating\_lab.txt --ref\_submit\_lab ./list\_of\_submatting\_lab.txt --output ./6-2-2022/BD\_final.tsv

#### Run Pipeline\_asia = pipeline\_asia.sh

* Input :
  + BD files using GISIAD
* Output
  + asia\_bd\_merged\_metadata.tsv
  + Asia\_bd\_merged\_sequences.fasta

#### 

#### Run Nextstrain from ncov

* cd ncov/
* Copy files at ncov/data folder from Pipeline\_asia.sh script output
  + ncov/data
    - asia\_bd\_merged\_metadata.tsv
    - asia\_bd\_merged\_sequences.fasta
* nextstrain build . --cores all --configfiles ./my\_profiles/example/builds.yaml
* Generate ncov\_bangladesh.jason file into ncov/auspice

#### 

#### To visualization the json file

* Do Google: auspice.us(<https://auspice.us/>)
* Drag & Drop a dataset on here to view

# 4.Nanopore assembly Sequence uploading at ENA

#### Create a metadata.csv file from Sample\_sheet.xlsx: manually

| sample \_sheet. | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Serial** | **ID** | **ONT\_StudyID** | **ONT\_SampleID** | **ONT\_RunID** | **Genome coverage** |  |  |  |
| 1 | E119 | ERP132884 | ERS8226991 | ERR9837010 | 84.11 |  |  |  |
|  |  |  |  |  |  |  |  |  |

| metadata | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| STUDY | SAMPLE | ASSEMBLYNAME | ASSEMBLY\_TYPE | COVERAGE | PROGRAM | PLATFORM | MOLECULETYPE | FASTA |
| ERP132884 | ERS8226991 | E119\_contigs\_MinION | isolate | 84.11 | Unicyler | MinION | genomic DNA | E119\_contigs\_MinION.fasta.gz |
| STUDY | SAMPLE | ASSEMBLYNAME | ASSEMBLY\_TYPE | COVERAGE | PROGRAM | PLATFORM | MOLECULETYPE | FASTA |

#### Convert metadata.csv to manifest.txt file:

* Script: metadat\_to\_menifest\_for\_assembly\_of\_nanopore.ipynb > at my drive https://drive.google.com/file/d/1TVeBhmblGa4yWO7u0rknezV2HJ9rInZE/view?usp=share\_link

##### manifest.txt

STUDY ERP132884

SAMPLE ERS8226991

ASSEMBLYNAME E119\_contigs\_MinION

ASSEMBLY\_TYPE isolate

COVERAGE 84.11

PROGRAM Unicyler

PLATFORM MinION

MOLECULETYPE genomic DNA

FASTA E119\_contigs\_MinION.fasta.gz

#### Folder(same) materials for uploadding:

* All assembly files
* All manifest files
* input.txt
* Webin-cli-4.3.0.jar
  + <https://ena-docs.readthedocs.io/en/latest/submit/general-guide/webin-cli.html>
* upload.sh and validate.sh script
  + https://drive.google.com/drive/folders/1fzurguasx41i5st4ZDLGE9\_aIoS6MCaL?usp=sharing

#### Run validate.sh & upload script

* https://drive.google.com/drive/folders/1XVedEZD0f0477xEDH2jNaVO\_\_nBys7co

# 5.Nanopore raw Sequence uploading at ENA

#### Create a metadata.csv file from Sample\_sheet.xlsx: manually

| sample \_sheet. | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | accession | lanename | Nanpore batch | RLB Barcodes | lanename | RLB BarcodesBioSample |  | Illumina run accession |
| SEAP7506789 | ERS13444798 | 30134\_1#127 | 3 | 1 | 30134\_1#127 | 1 | SAMEA111347248 | ERR4362458 |
| SEAP7506905 | ERS13444799 | 30134\_1#148 | 3 | 2 | 30134\_1#148 | 2 | SAMEA111347249 | ERR4362479 |

| metadata | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| STUDY | SAMPLE | NAME | PLATFORM | INSTRUMENT | LIBRARY\_SOURCE | LIBRARY\_SELECTION | LIBRARY\_STRATEGY | FASTQ |
| ERP132884 | ERS8227002 | E131\_MinION | OXFORD\_NANOPORE | MinION | GENOMIC | RANDOM | WGS | E131\_MinION.fastq.gz |
| ERP132884 | ERS8227005 | E134\_MinION | OXFORD\_NANOPORE | MinION | GENOMIC | RANDOM | WGS | E134\_MinION.fastq.gz |

#### Convert metadata.csv to manifest.txt file:

* Script: metadat\_to\_menifest\_for\_reads\_of\_nanopore.ipynb > at my drive https://drive.google.com/file/d/1yakR38bafzyHyIwAyAxOsiN0gtNWgESi/view?usp=share\_link

##### manifest.txt

STUDY ERP132884

SAMPLE ERS8227002

NAME E131\_MinION

PLATFORM OXFORD\_NANOPORE

INSTRUMENT MinION

LIBRARY\_SOURCE GENOMIC

LIBRARY\_SELECTION RANDOM

LIBRARY\_STRATEGY WGS

FASTQ E131\_MinION.fastq.gz

#### Folder(same) materials for uploadding:

* All raw read files
* All manifest files
* Input.txt
* Webin-cli-4.3.0.jar
  + <https://ena-docs.readthedocs.io/en/latest/submit/general-guide/webin-cli.html>
* upload.sh and validate.sh script
  + https://drive.google.com/drive/folders/1fzurguasx41i5st4ZDLGE9\_aIoS6MCaL?usp=sharing

#### Run validate.sh & upload script

* https://drive.google.com/file/d/1ifuEvBQhFGP7EmDH8fAnypAJ-p1a9434/view?usp=share\_link

ERR

Prcessing E342\_MinION...

2022-09-25 17:00:46.939282: W tensorflow/stream\_executor/platform/default/dso\_loader.cc:64] Could not load dynamic library 'libcudart.so.11.0'; dlerror: libcudart.so.11.0: cannot open shared object file: No such file or directory

2022-09-25 17:00:46.939303: I tensorflow/stream\_executor/cuda/cudart\_stub.cc:29] Ignore above cudart dlerror if you do not have a GPU set up on your machine.

Checking program versions

This is medaka 1.6.0

Program Version Required Pass

bcftools 1.9 1.11 False

bgzip 1.9 1.11 False

minimap2 2.24 2.11 True

samtools 1.6 1.11 False

tabix 1.9 1.11 False

(hybrid\_assembly)

# 6. CZID metagenomics:

I built-in pipeline to analysis metagenomics data

#### Manually

\*Upload: Project Name

Analysis Type > Metagenomics > All metagenomics fastq.gz file Drag & Drop

Continue

Manual Input:

| Sample Name | Host Organism | Sample Type | Water Control | Nucleotide Type | Collection Location | Collection Date |
| --- | --- | --- | --- | --- | --- | --- |
| Sample Name | Human | Pleural Fluid | If EC/NC  Then will Yes without NO | RNA/DNA | Bangladesh | date |

Continue > agree > Start Upload

# 

# 7. Tetraparvovirus(TPRV) upload to IDseq for generate consensus genome:

### **5.1 Within a folder contain:**

#### fastq\_files(raw data) folder

#### Contain metadata.csv

#### Contain Primer.bed

#### Reference fasta file

#### Script

#### 

#### metadata.csv:

| Sample Name | Host Organism | Sample Type | Nucleotide Type | Collection Date | Water Control | Collection Location |
| --- | --- | --- | --- | --- | --- | --- |
| CSF\_0044\_TP4\_S169 | Human | CSF | DNA | 2022-05 | No | Bangladesh |

#### Primer.bed

KM390024.1 86 105 Trpv4\_v3\_1\_LEFT 20 +

KM390024.1 464 482 Trpv4\_v3\_1\_RIGHT 19 -

#### CZID login from terminal:

Open terminal:

#### Script

***# For single sample***

czid consensus-genome upload-sample \

--project 'TPRV4\_trial' \

--sample-name 'TRPV\_sample\_1' \

--metadata-csv 'metadata.csv' \

--sequencing-platform 'Illumina' \

--reference-fasta 'NC\_007018.1.fasta' \

--primer-bed 'Primer.bed' \

'CSF\_0037\_TP4\_S8\_L001\_R1\_001.fastq.gz' 'CSF\_0037\_TP4\_S8\_L001\_R2\_001.fastq.gz'

***# For multiple sample***

czid consensus-genome upload-samples \

--project 'TPRV4\_trial' \

--metadata-csv 'metadata\_tprv\_KM.csv' \

--sequencing-platform 'Illumina' \

--reference-fasta 'KM390024.1.fasta' \

--primer-bed 'Primer\_K.bed' \

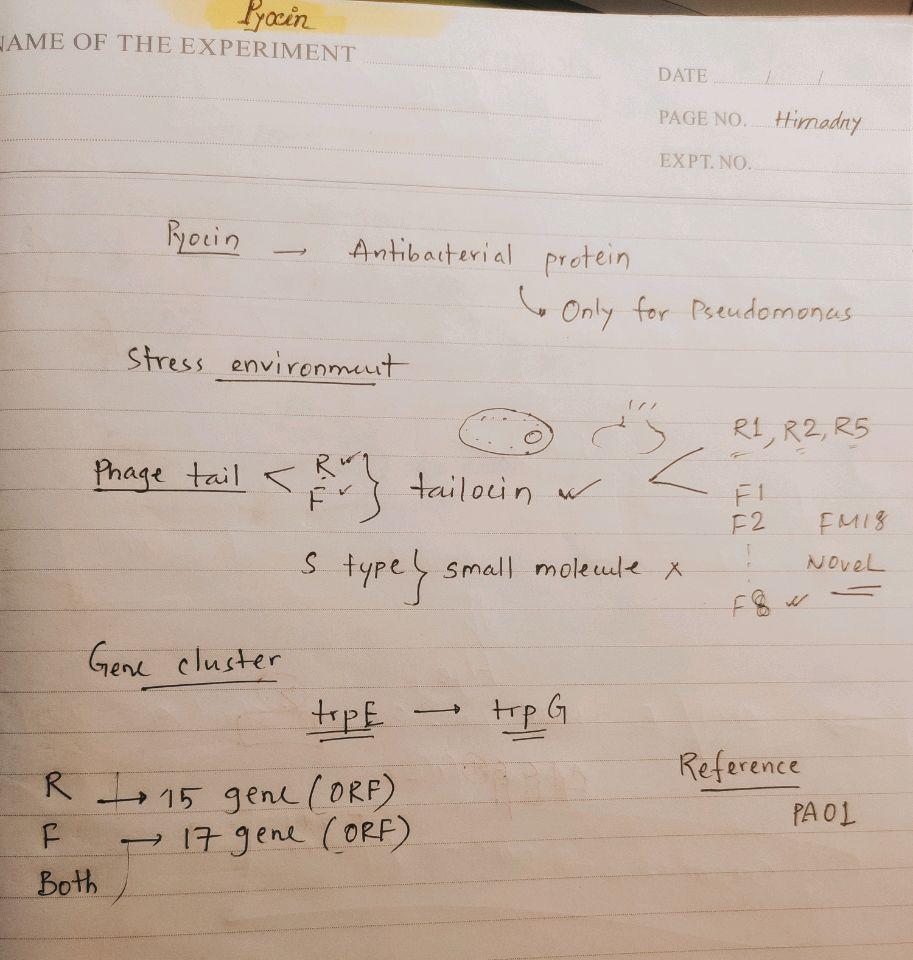
'/home/chrf/Desktop/czid\_cli/fastq\_files\_2/'

#### Now run the script

# 9. Serotype 19A

1. \*Population of sero 19 Change
2. \*Change from 19 to 19AF
3. \*Why this change
4. \*19A phenotype to 19AF genotypic

# 10. Pyocin



# 

# 11. Genome assembly using Unicycler:

| **rawpath="/media/chrf/CHRF\_Genomics\_03/GBS/Nextseq\_13"**  **header="GBS\_"**  **# The Nextseq\_13 folder contains only paired end reads. Here we will create a folder with their base name.**  **for files in $rawpath\/\*.fastq.gz**  **do**  **name=`basename $files| cut -f 1,2 -d '\_'`**  **echo $name**  **mkdir -p $rawpath/$name/1\_RawData**  **mv $rawpath\/$name\*\\_R1\_001.fastq.gz $rawpath\/$name\*\\_R2\_001.fastq.gz $rawpath/$name/1\_RawData/**  **done**  **# Create a directory (3\_Unicycler) within every base folder and keep the assembly output here**  **for dir in $rawpath\/$header\***  **do**  **name=`basename $dir`**  **echo $name**  **Mkdir -p -v 3\_Unicycler**  **unicycler -1 $dir\/1\_RawData/$name\*\\_R1\_001.fastq.gz -2 $dir\/1\_RawData/$name\*\\_R2\_001.fastq.gz --min\_fasta\_length 300 -t 12 --keep 1 -o $dir\/3\_Unicycler/**  **echo "Rename Contig file"**  **mv $dir\/3\_Unicycler/assembly.fasta $dir\/3\_Unicycler/$name\\_contigs.fasta**    **done** |
| --- |

# 12. Fastq\_file\_name\_that\_less\_than\_50\_M\_command

find . -type f -size -50M | cut -f 2 -d '.' | sed 's./..' | sed 's.\_R1\_001..' | sed 's.\_R2\_001..' | uniq | sort

# 

# 

# 

# 13. Counting\_contigs\_from\_fasta\_file

#### This background black because its copy from VS code

rawpath="/media/chrf/Home03/GBS/Nextseq\_13/Done"

header="GBS\_"

for dir in $rawpath\/$header\*

do

name=`basename $dir`

echo "Contig number of $name "

grep -c "^>" $rawpath\/$name/3\_Unicycler/$name\\_contigs.fasta

echo ""

Done

# 14. Coverage depth from Bam file

samtools depth -a -q 20 sorted\_indexed\_bam | awk '{sum+=$3} END {print "Average = ",sum/NR}'

# 15. N50 and Contig count from assembly file for check the Raw data & Assembly quality

path=$1

header=$2

for dir in $path$header\*

do

name=`basename $dir`

n50=`assembly\_stats $dir\/3\_Unicycler/$name\\_contigs.fasta | grep "N50" | head -1 | cut -f 2 -d ":"| cut -f 1 -d "," | sed -e 's/ //g'`

count=`assembly\_stats $dir\/3\_Unicycler/$name\\_contigs.fasta | grep "count" | head -1 | cut -f 2 -d ":"| cut -f 1 -d "," | sed -e 's/ //g'`

echo $name $n50 $count

done

# 

# 16. Single Cell RNA-Seq Analysis:

## **Documents:**

* 1. <https://www.singlecellcourse.org/introduction-to-single-cell-rna-seq.html>

## Seminar:

* 1. https://youtu.be/48wfaeg6MU0

## Datasets:

* 1. <https://cellxgene.cziscience.com/collections/35d0b748-3eed-43a5-a1c4-1dade5ec5ca0>
  2. https://cellxgene.cziscience.com/e/13b61a7d-5605-4948-ba48-02c588960143.cxg/

## Concept and **Coding:**

* 1. <https://www.embopress.org/doi/full/10.15252/msb.20188746>
  2. <http://www2.stat.duke.edu/~sayan/Sta613/2018/singlecellrnaseq-170131050320.pdf>
  3. <https://satijalab.org/seurat/>

### Video tutorial Full Concept :- <https://www.youtube.com/watch?v=qUjBmCQoKhU&list=PLEEE2A91B09B77B4A&index=47>

## **Paper:**

* 1. <https://cellxgene.cziscience.com/collections/35d0b748-3eed-43a5-a1c4-1dade5ec5ca0>

## Setup Single-cell RNA sequencing at CHRF:

* 1. <https://www.youtube.com/watch?v=tacuQtH_hgw>

## Deep Learning for Single cell RNA seq analysis:

* 1. https://www.youtube.com/watch?v=0jWOZoTsYzI&list=PLypiXJdtIca5sxV7aE3-PS9fYX3vUdIOX

# 17. GISAID:

## GISAID Folder contain

#### 13.1. EpicCoV\_BulkUpload\_COVIDSeq\_Batch\_number\_Date.fa

A merge fasta file of metadata’s id came from **Nextclade**

* Goes to the czid and CHRF Project
* Select the Sample (Covid)
* Select the Consensus Genome
* Download the sample by arrow button
* Select the Sample Metadata and Consensus Genome

> Rename the Consensus Genome Contig File by

>hCoV-19/Bangladesh/CHRF-Smaple\_id/year

>Sort the contig by its header

#### 13.2. EpicCoV\_BulkUpload\_COVIDSeq\_Batch\_number\_Date

Check the Format of my google drive file

>shorturl.at/fimA8

#### 13.3. EpicCoV\_BulkUpload\_COVIDSeq\_Batch\_number\_Date\_Indel\_list

>Merged fasta file Drag and Drop into Nextclade

> Take output file

> Download as xls format

Rearrange by checking format of my google drive file

>shorturl.at/esAST

# 18. NextClaid:

Summery

We have sequenced 13 samples. 11 samples are XBB. The other two samples are BM.1.1.1 and BM.4.1.1

All Sample minimum genomes call 97.5%. Except one, all sequence quality is good and "OK" to upload

One sample has too many private mutations(unique mutation) which causes frameshift mutation in the sequence which is not reliable. So that why we do not want to upload that sequences

# 19. Microbiology:

https://drive.google.com/file/d/1Cz1zRviA3UkqkHn1yl6wjJ13aBK7viS1/view?usp=share\_link

# 20. Sample Sheet Preparation for NextSeq 2000:

## Master Files:

#### 1.1. IDT\_14464944\_45.xlsx



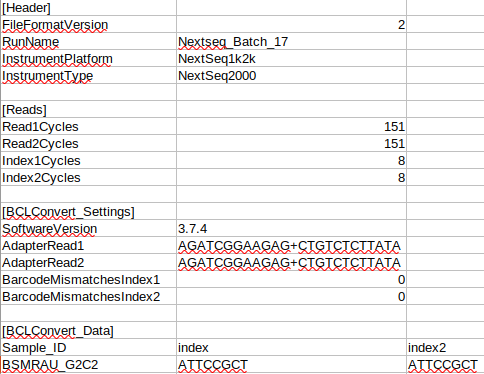
#### 1.2. IDT\_15156265\_15155023.xlsx



#### 1.3. Illumina\_Set\_1\_2\_3.xlsx

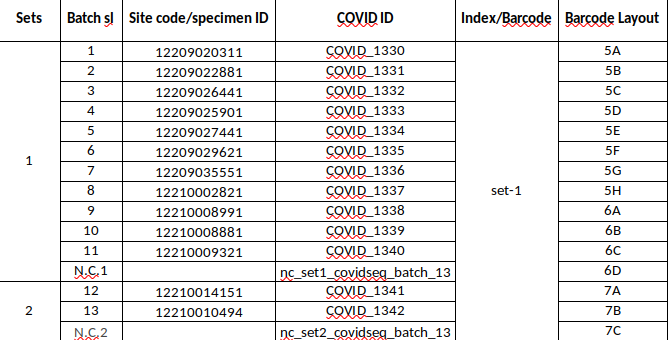


## NextSeq\_Batch\_template\_SampleSheet\_local\_24thDec2022.csv File

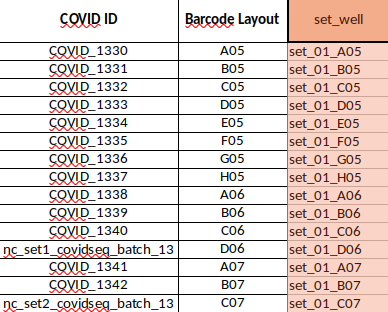


## Sample Files:

#### 3.1. Sample\_sheet\_COVIDSeq\_13.xlsx

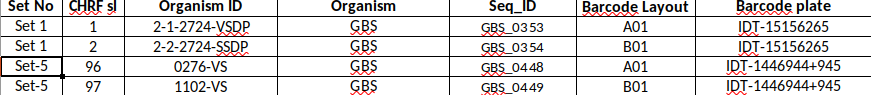


**Edit to below structure**

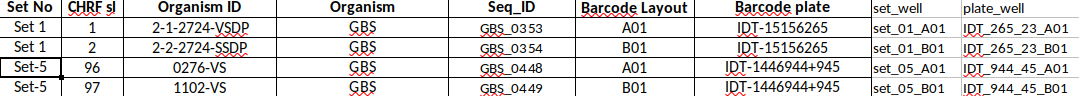


#### 

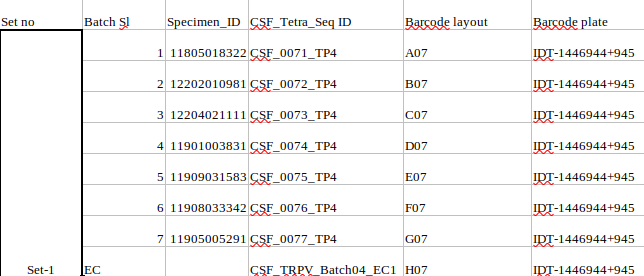
#### 3.2. Sample\_Sheet\_NextSeq\_17\_V1.xlsx



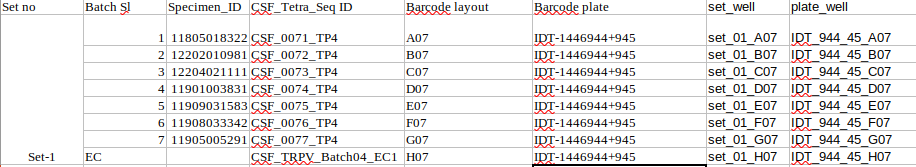
**Edit to below structure**

****

#### 3.3. TRPV\_Batch\_04\_Barcode\_layout.xlsx



**Edit to below structure**



#### 

#### Now Run Scripts that will match common plate\_well/set\_well column between master files and sample files

#### Extracted Index, Index2 from 4 no step (scripts run> files generate hobe) insert into the **NextSeq\_Batch\_template\_SampleSheet\_local\_24thDec2022.csv File according to its ID-Index-Index2**

# 21. Sample Sheet Preparation for iSeq 100:

BLAST:

Database Create:

$ blastn -db ./KPN -query 12899\_contigs.fasta -out KPN.blsn -outfmt "6 qseqid sseqid slen qstart qend length mismatch gapopen gaps sseq" -word\_size 5 -perc\_identity

Searching at my created database:

$ blastn -db ./KPN -query 12899\_contigs.fasta -out KPN.blsn -outfmt "6 qseqid sseqid slen qstart qend length mismatch gapopen gaps sseq" -word\_size 5 -perc\_identity 80

# 21. Install JupyterLab Desktop : Search google

22. Mega

# 22. Prevent Conda from activating the base environment by default

#### conda config --set auto\_activate\_base false

# 23. Different Kernel of Jupyter Notebook:

**Bash:**

**##------------------------------------------##**

###### # Make sure you are in the base env #

###### ##------------------------------------------##

conda activate base

**##------------------------------------------##**

###### # Install bash kernel in the base env #

###### ##------------------------------------------##

pip install bash\_kernel

python -m bash\_kernel.install

**##------------------------------------------------------------------------------------------------------------------------------------##**

###### # This time the Jupyter kernel is always the base kernel. To run code from another environment you # # should activate. the environment from Jupyter cell. #

###### ##------------------------------------------------------------------------------------------------------------------------------------##

**Python:**

***https://stackoverflow.com/questions/74611535/new-conda-environment-with-latest-python-version-for-jupyter-notebook***

###### ##------------------------------------------##

###### # Make sure you are in the base env #

###### ##------------------------------------------##

conda activate base

###### ##----------------##

###### # Update conda#

###### ##----------------##

*conda update conda*

###### ##----------------------------------------##

###### # To allow support for powershell #

###### ##----------------------------------------##

conda init --all

###### ##-----------------------------------------------------------------------------------------------------------------------------------##

###### # The conda-forge repository seems to have at least the latest stable Python version, so we will get #

###### # Python from there. Add conda-forge to channels of conda. #

###### ##-----------------------------------------------------------------------------------------------------------------------------------##

conda config --add channels conda-forge

###### ##-----------------------------------------------------------------------------------------------------------------------------------##

###### #To fix 500 internal server errors when trying to open a notebook later. Primarily, the nbconvert tool # #allows you to convert a Jupyter .ipynb notebook document file into another static format including # #HTML, LaTeX, PDF, Markdown, reStructuredText, and more. nbconvert can also add productivity to # #your workflow when used to execute notebooks programmatically. #

###### ##-----------------------------------------------------------------------------------------------------------------------------------##

pip3 install --upgrade --user nbconvert

###### ##-----------------------------------------------------------------------------------------------------------------------------------##

###### nb\_conda\_kernels enables a Jupyter Notebook or JupyterLab. Application in one conda environment to access kernels for Python, R, and other languages found in other environments.

###### ##-----------------------------------------------------------------------------------------------------------------------------------##

conda install nb\_conda\_kernels -y

###### ##------------------------------------------------------------------------------------------------------##

###### # I will now create a new conda env for Python 3.11 and environment name is ipy #

###### ##------------------------------------------------------------------------------------------------------##

conda create -n ipy python=3.11 -y

conda activate ipy

###### ##---------------------------------------------------------------------------------------------------------------------------##

**# Once installed, I need to install ipykernel. So Jupyter notebook can see the new environment ipy.#**

###### ##---------------------------------------------------------------------------------------------------------------------------##

conda install -n ipy ipykernel -y

###### ##------------------------------------------------##

**# Setup display Name of Jupyter Kernel #**

###### ##------------------------------------------------##

python -m ipykernel install --user --name ipy --display-name "IPython"

###### 

###### 

###### ##------------------------------------------------------------------------##

**# Install ipywidgets as well for some useful functionalities #**

###### ##------------------------------------------------------------------------##

conda install -n ipy ipywidgets -y

**R:**

###### 

###### 

###### 

###### ##----------------------------------------------------------------------------------------------------------------------------##

###### # Created new R environment and install r-base and some important packages all together#

###### ##----------------------------------------------------------------------------------------------------------------------------##

conda create -n r-env r-essentials r-base

***You should add r-essentials along with r-base as it will not only install all the important r packages but also help during IRkernel installation.***

###### ##---------------------------------------------------------------##

###### # To install a specific R version, let’s say R 4.2 #

###### ##---------------------------------------------------------------##

conda create -n ir r-essentials r-base=4.2

###### ##-------------------------------------------------------------------##

###### # Add IRkernel for jupyter notebook or jupyterlab #

###### ##--------------------------------------------------------------------##

**##---------------------------------------##**

###### # Make sure you are in the ir env #

###### ##---------------------------------------##

conda activate ir

**After installing R you may want to run your R codes in jupyter notebooks. In that case, you have to install the R kernel. This kernel will make your R available for jupyter notebook or lab.**

**Let’s say you install R version 4.2 in the ir environment. So, to make this version of R accessible for jupyter notebook/lab you have to install IRkernel in this environment.**

**##--------------------------------------------##**

###### # To install necessary dependencies #

###### ##--------------------------------------------##

sudo apt-get install libzmq3-dev libcurl4-openssl-dev libssl-dev jupyter-core jupyter-client

***You have run this command once. If you create a new conda environment and want to install IRkernel for a new R version inside that environment you don’t have to run this command again.***

**##--------------------------------##**

###### # Now, run R in the terminal #

###### ##---------------------------------##

**R**

**##-----------------------------------------------------------------------------##**

###### # Install IRkernel with some necessary packages from consol #

###### ##-----------------------------------------------------------------------------##

conda install r-irdisplay r-irkernel

***Obviously run this command in the R console, not in the bash terminal.***

**##--------------------------------------------##**

###### # Make kernel is available to Jupyter #

###### ##--------------------------------------------##

IRkernel::installspec(name = 'ir', displayname = 'IR')

**\*\*\*ir-env > kernel name of r env and its display name R 4.0.1\*\***

# 24. SRA Data Download Using Bioproject id and accession list:

#### Take Bioproject id

#### Open google and ncbi > select Bioproject option and search by my bioproject ID

#### 

#### 

#### Select > SRA from Related information

#### 

#### Show your Items

#### If needed any filter > Select : Results by taxon (organism name< select) > will show filter box

#### If multiple filter > Filter by writing AND for each option

#### 

#### Now : Click to Send to option > File> format(accession list) > Create File

#### 

#### Now run this script

#### 

a=$(cat accession\_list.txt)

for sra in $a

do

echo "Started Download "${sra}""

fastq-dump --split-files --outdir ./data --gzip "${sra}"

echo "Download Complete "${sra}""

done

# 25. Check Two Lists Intersected ID

grep -Fxf list\_1 list\_2

8. R and Rstudio install and uninstall:

#### 7.1 Install R

* https://computingforgeeks.com/how-to-install-r-and-rstudio-on-ubuntu-debian-mint/

#### 7.2 Install Rstudio:

wget https://download1.rstudio.org/desktop/bionic/amd64/rstudio-2022.02.2-485-amd64.deb

sudo apt install -f ./rstudio-2022.02.2-485-amd64.deb

#### 7.3 R in Google colab

1. Open your favorite browser.
2. Go to this URL: <https://colab.research.google.com/#create=true&language=r>, or this short URL <https://colab.to/r>

# 26. Ubuntu Install:

Setup after Installation :

sudo apt-get update

sudo apt-get upgrade

#Install Chrome:

wget<https://dl.google.com/linux/direct/google-chrome-stable_current_amd64.deb>

sudo dpkg -i google-chrome-stable\_current\_amd64.deb

#Install curl:

sudo apt-get install curl -y

curl <https://data.biostarhandbook.com/install.sh> | bash

Install Jupyter Lab:

wget https://github.com/jupyterlab/jupyterlab-desktop/releases/latest/download/JupyterLab-Setup-Debian.deb

sudo dpkg -i JupyterLab-Setup-Debian.deb

#Conda default base remove:

conda config –set auto\_activate\_base false