

- Introduction to bioinformatic analysis of SARS-CoV-2 amplicon sequencing data

Illumina Basespace and DRAGEN app

Illumina Basespace

- Free user profile
- Free apps
- Illumina still the biggest sequencing company (availability of data)

Illumina BaseSpace



Start using BaseSpace Sequence Hub

Log In

New User?
Register

Start a free trial*

*The 30-day free trial begins automatically when a user first creates an account and logs in. The trial grants users with 250 iCredits (which expire after 30 days) that can be applied to compute fees for any BaseSpace app and data storage above 1 TB. After the trial expires, the user can no longer run apps, but can continue to monitor sequencing runs, automatically demultiplex uploaded runs, store a limited amount of data, share & transfer with collaborators, and download data.

Product Highlights

As a key component of the BaseSpace Suite, BaseSpace Sequence Hub is a direct extension of your Illumina instruments. Encrypted data flow from the instrument into BaseSpace Sequence Hub, enabling you to manage and analyze your data easily with a curated set of analysis apps.

BaseSpace Sequence Hub, powered by Amazon Web Services (AWS):

- Offers a security-first environment
- Enables you to set up runs and monitor instrument run quality

- Promotes efficiency by converting sequencing data to a standard format and streaming directly to the cloud
- Provides access to computing resources without the capital expenditure of in-house infrastructure
- Increases organizational productivity with easy access to a multitude of genomic analysis apps (provided by you, Illumina, or third parties)

[See All BaseSpace Data Analysis Apps](#)

Illumina Basespace

Projects



FILE

NEW

EDIT

COPY

MERGE

DOWNLOAD

UPLOAD

Projects



NAME

OWNER

CovidSeq

Ivan Barilar

CovidSeq

SUMMARY BIOSAMPLES SAMPLES FASTQS OTHER DATASETS



Owner
Ivan Barilar

Size
487 MB

Last Updated
2023-09-24 11:13

Description
GenEpiBio train anal...

illumina

Register

Already have an account? [Sign in](#)

Email address

Not a valid email address

First name

Last name

Password

[Guidelines](#)

Not a valid password

Confirm password

Country/Region

☐ I'm not a robot



Create Account

Cancel

Illumina BaseSpace

CovidSeq

SUMMARY BIOSAMPLES SAMPLES FASTQS OTHER DATASETS



FILE

NEW

EDIT

COPY

MERGE

DOWNLOAD

UPLOAD

Size
487 MB

BIOSAMPLE WORKFLOW

FILES

CovidSeq

01. Choose File Type

- ☒ FASTQ
- ☐ VCF
- ☐ Manifest
- ☐ Other

02. Select and Review

BROWSE FILES



DRAG & DROP YOUR FILES HERE

16 files and 250 GB maximum upload

[Additional upload guidance](#)

Total to Upload:

0 Completed ...

--

Illumina Basespace

CovidSeq



CANCEL

01. Choose File Type

- ☒ FASTQ
- ☐ VCF
- ☐ Manifest
- ☐ Other

02. Select and Review

BROWSE FILES



DROP & DROP YOUR FILES HERE

16 files and 250 GB maximum upload

[Additional upload guidance](#)

COV-3126-A3080-lib47553_S56_L002_R1...	13.80 MB	×
COV-3126-A3080-lib47553_S56_L002_R2...	14.33 MB	×
COV-3126-A3080-lib47553_S56_L003_R1...	14.10 MB	×
COV-3126-A3080-lib47553_S56_L003_R2...	14.68 MB	×
COV-3126-A3080-lib47553_S56_L004_R1...	13.89 MB	×
COV-3126-A3080-lib47553_S56_L004_R2...	14.42 MB	×
COV-3126-A3080-lib47553_S56_L001_R1...	13.86 MB	×
COV-3126-A3080-lib47553_S56_L001_R2...	14.39 MB	×

Total to Upload:

Uploading 1 of 8... ⋮

113.47 MB

03. Complete Upload

Save upload to:

SAMPLE



BIOSAMPLE

SELECT BIOSAMPLE

Library Name

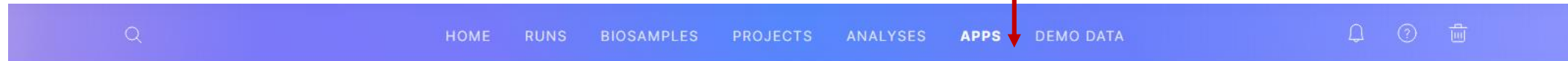
(No special characters, 60 max)

Library Prep Kit

SELECT

COMPLETE UPLOAD

Illumina DRAGEN COVID Lineage App



Accurate, comprehensive, and efficient secondary analysis with DRAGEN

Illumina DRAGEN apps provide pushbutton capabilities within BaseSpace Sequence Hub's intuitive and secure environment.



DRAGEN Germline
Illumina, Inc.



DRAGEN Enrichment
Illumina, Inc.



DRAGEN Somatic
Illumina, Inc.



DRAGEN RNA
Illumina, Inc.

Search Apps



DRAGEN (29)



DRAGEN Amplicon
Illumina, Inc.



DRAGEN Baseline Builder
Illumina, Inc.



DRAGEN CNV Baseline Builder
BaseSpace Labs



DRAGEN COVID Lineage
BaseSpace Labs



DRAGEN COVIDSeq Test (RUO)
Illumina, Inc.



DRAGEN Differential Expression
BaseSpace Labs

[App Developer Portal](#)



Bookmarks Empty

Get started by bookmarking apps you regularly use.

Illumina DRAGEN COVID Lineage App



SEQUENCE HUB

Thomas Kohl

HOME

RUNS

BIOSAMPLES

PROJECTS

ANALYSES

APPS

DEMO DATA



DRAGEN COVID Lineage

BaseSpace Labs

[Bookmark this app](#) [Help](#)

FOR RESEARCH USE ONLY — NOT FOR USE IN DIAGNOSTIC PROCEDURES

A new parameter, "Keep Scrubbed FASTQs" is ON by default, and will result in a 60% increase in the size of the output files. If you do NOT intend to submit to a public repository or your desired repository does not require it, please deselect the checkbox. Keeping scrubbed FASTQs facilitates public repository submission. If submitting to a public repository, please confirm all repository requirements have been met (including removal of human reads).

Overview

The Illumina® DRAGEN COVID Lineage App uses a customized version of the DRAGEN DNA pipeline to perform Kmer-based detection of SARS-CoV-2. It then aligns reads to a reference genome, calls variants, and generates a consensus genome sequence. It then optionally performs lineage/clade analysis using Pangolin and NextClade. Both DNA and RNA libraries are supported. This app works best with data generated using the ARTIC V3, V4, or V4.1* primer panels, but may still be useful for other data types.

[READ MORE](#)

What's New

Version



A newer version of this app (v4.0.3) is available.

3.5.9

LAUNCH APPLICATION

Illumina DRAGEN COVID Lineage App



Input Data

Please select your preferred method for input data to launch the analysis ⁱ

- ☒ Biosamples - Launch with FASTQs across multiple sequencing runs
- ☐ Samples - Launch with FASTQs from a single sequencing run

Configuration

Analysis Name

DRAGEN COVID Lineage 09/24/2023 2:45:43

Save Results To

[SELECT PROJECT](#)

Warning

Large numbers of samples (>1536) or high read depths may cause running time app to be aborted. In such a case, split the samples into multiple Analysis runs

Input Type ⁱ

- ☒ Samples
- ☐ Project

Biosample ⁱ

[SELECT BIOSAMPLE\(S\):](#)

Input Project ⁱ

[SELECT PROJECT](#)

Reference genome / PCR primer set ⁱ

Not Selected

Workflow Control

Enable or disable optional steps in the analysis.

Pangolin ⁱ

☒ Enable Pangolin

NextClade ⁱ

☒ Enable NextClade

+ Advanced Settings (Map/Align)

+ Advanced Settings (Variant Calling)

+ Advanced Settings (Consensus Sequence)

+ Advanced Settings (Pangolin & NextClade)

+ Advanced Settings (Output Files)

BaseSpace Labs

☐ I acknowledge and agree that (i) this is a BaseSpace Lab. Illumina has no obligation to provide any technical support including without limitation, any loss of data, incorrect results. App.

[LAUNCH APPLICATION](#)

Illumina Basespace



illumina

SEQUENCE HUB

Ivan Barilar

HOME

RUNS

BIOSAMPLES

PROJECTS

ANALYSES

APPS

DEMO DATA

CovidSeq

SUMMARY BIOSAMPLES SAMPLES FASTQS OTHER DATASETS



Owner	Size	Last Updated	Collaborators
Ivan Barilar	2 GB	2023-09-24 14:32	
Description	GenEpiBio train anal...		

Analyses

1 - 6 of 6 Show 25

☐ STATUS	ANALYSIS NAME	APPLICATION	SIZE	LAST MODIFIED	COMMENTS
☐ Complete	DRAGEN COVID Lineage 09/24	DRAGEN COVID Lineage (v4.0.3)	915 MB	2023-09-24 14:35	
☐ Complete	DRAGEN COVID Lineage 09/24	DRAGEN COVID Lineage (v4.0.3)	307 MB	2023-09-24 14:48	

Exercise

GENEPI-BIOTRAIN - VIRTUAL TRAINING 3:
INTRODUCTION TO BIOINFORMATIC ANALYSIS OF SARS-
COV-2 AMPLICON SEQUENCING DATA

► Participants

🏆 Badges

📊 Grades

► Introduction to bioinformatic analysis of
SARS-CoV...

► Aim and Objectives

► Restricted reference documentation

► Software requirements

► Session 1 (25 September): Introduction to
SARS-CoV...

▼ Practical exercise

📄 exercise_description_summary

📁 NGS data

► Session 2 (26 September): Steps of
bioinformatics ...

► Evaluation and certificate

► Contacts

GENEPI-BIOTRAIN - VIRTUAL TRAINING 3:
INTRODUCTION TO BIOINFORMATIC ANALYSIS OF SARS-
COV-2 AMPLICON SEQUENCING DATA

► Participants

🏆 Badges

📊 Grades

► Introduction to bioinformatic analysis of
SARS-CoV...

► Aim and Objectives

► Restricted reference documentation

► Software requirements

► Session 1 (25 September): Introduction to
SARS-CoV...

▼ Practical exercise

📁 NGS data

► Session 2 (26 September): SARSCOV2seq
bioinformati...

► Evaluation and certificate

► Contacts

ADMINISTRATION

▼ Folder administration

■ Edit settings

■ Locally assigned roles

■ Permissions

■ Check permissions

■ Filters

🔔 Notifications

■ Logs

■ Backup

NGS data

Manually mark this activity when complete

- 📁 COV-2074-A1664-lib41292_S5_L001_R1_001.fastq.gz
- 📁 COV-2074-A1664-lib41292_S5_L001_R2_001.fastq.gz
- 📁 COV-2074-A1664-lib41292_S5_L002_R1_001.fastq.gz
- 📁 COV-2074-A1664-lib41292_S5_L002_R2_001.fastq.gz
- 📁 COV-2074-A1664-lib41292_S5_L003_R1_001.fastq.gz
- 📁 COV-2074-A1664-lib41292_S5_L003_R2_001.fastq.gz
- 📁 COV-2074-A1664-lib41292_S5_L004_R1_001.fastq.gz
- 📁 COV-2074-A1664-lib41292_S5_L004_R2_001.fastq.gz
- 📁 COV-3126-A3080-lib47553_S56_L001_R1_001.fastq.gz
- 📁 COV-3126-A3080-lib47553_S56_L001_R2_001.fastq.gz
- 📁 COV-3126-A3080-lib47553_S56_L002_R1_001.fastq.gz
- 📁 COV-3126-A3080-lib47553_S56_L002_R2_001.fastq.gz
- 📁 COV-3126-A3080-lib47553_S56_L003_R1_001.fastq.gz
- 📁 COV-3126-A3080-lib47553_S56_L003_R2_001.fastq.gz
- 📁 COV-3126-A3080-lib47553_S56_L004_R1_001.fastq.gz
- 📁 COV-3126-A3080-lib47553_S56_L004_R2_001.fastq.gz
- 📁 COV-3126-A3086-lib47559_S62_L001_R1_001.fastq.gz
- 📁 COV-3126-A3086-lib47559_S62_L001_R2_001.fastq.gz
- 📁 COV-3126-A3086-lib47559_S62_L002_R1_001.fastq.gz
- 📁 COV-3126-A3086-lib47559_S62_L002_R2_001.fastq.gz
- 📁 COV-3126-A3086-lib47559_S62_L003_R1_001.fastq.gz
- 📁 COV-3126-A3086-lib47559_S62_L003_R2_001.fastq.gz
- 📁 COV-3126-A3086-lib47559_S62_L004_R1_001.fastq.gz
- 📁 COV-3126-A3086-lib47559_S62_L004_R2_001.fastq.gz

Download folder

Edit

blue: artic v3; orange: artic v4.1

Exercise

- Get an overview about the summary report and critically evaluate the results of the app.
 - Have a look on coverage and amplicon coverage. Do you observe any peculiarities?
 - Which SARS-CoV-2 variants are identified?
 - Which mutations in the S-gene are identified?
 - Would you continue with all datasets if you would like to analyze S-gene mutations?
- Also have a look at covariants.org and cov-lineages.org webpages for variant and mutation information