- 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

*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.

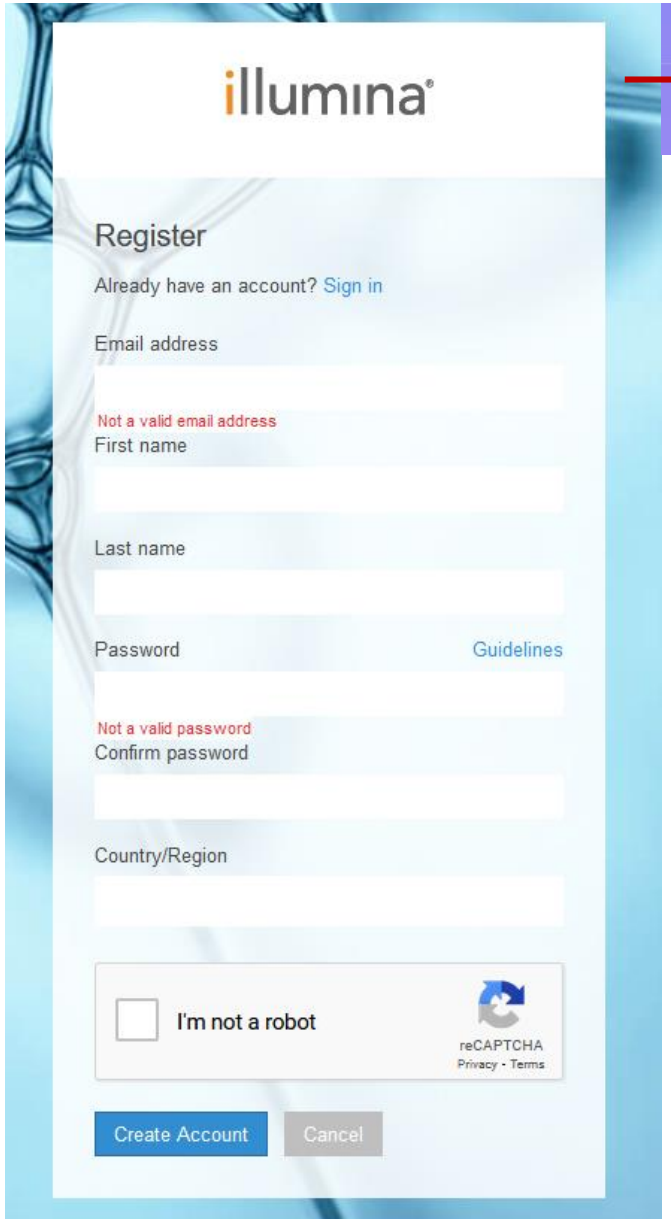
Product Highlights

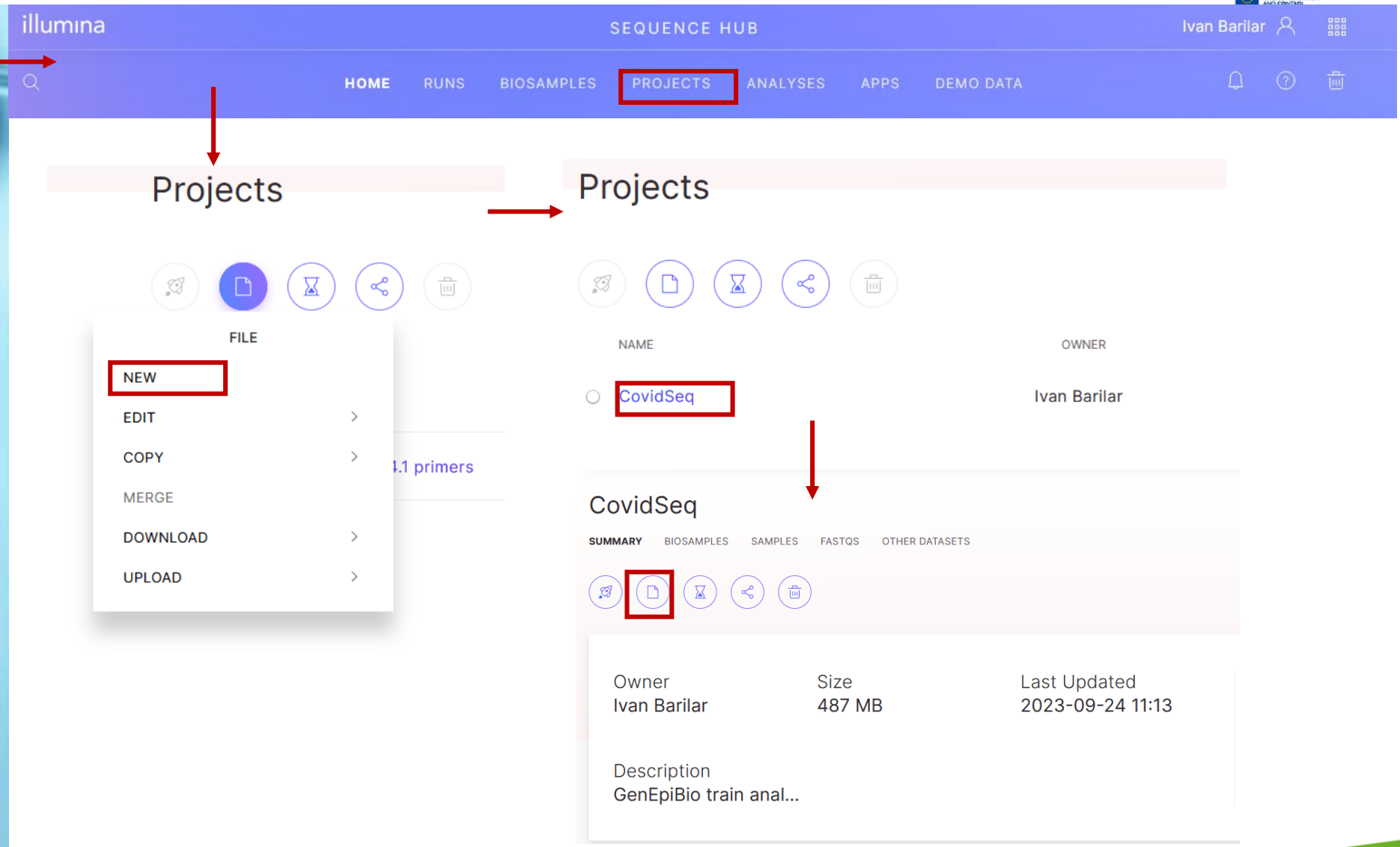
As a key component of the BaseSpace Suite, BaseSpace Sequence Hub is a direct extension of your Illumina instruments. Encrypted data flow from the instrument into BaseSpace Sequence Hub, enabling you to manage and analyze your data easily with a curated set of analysis apps. BaseSpace Sequence Hub, powered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

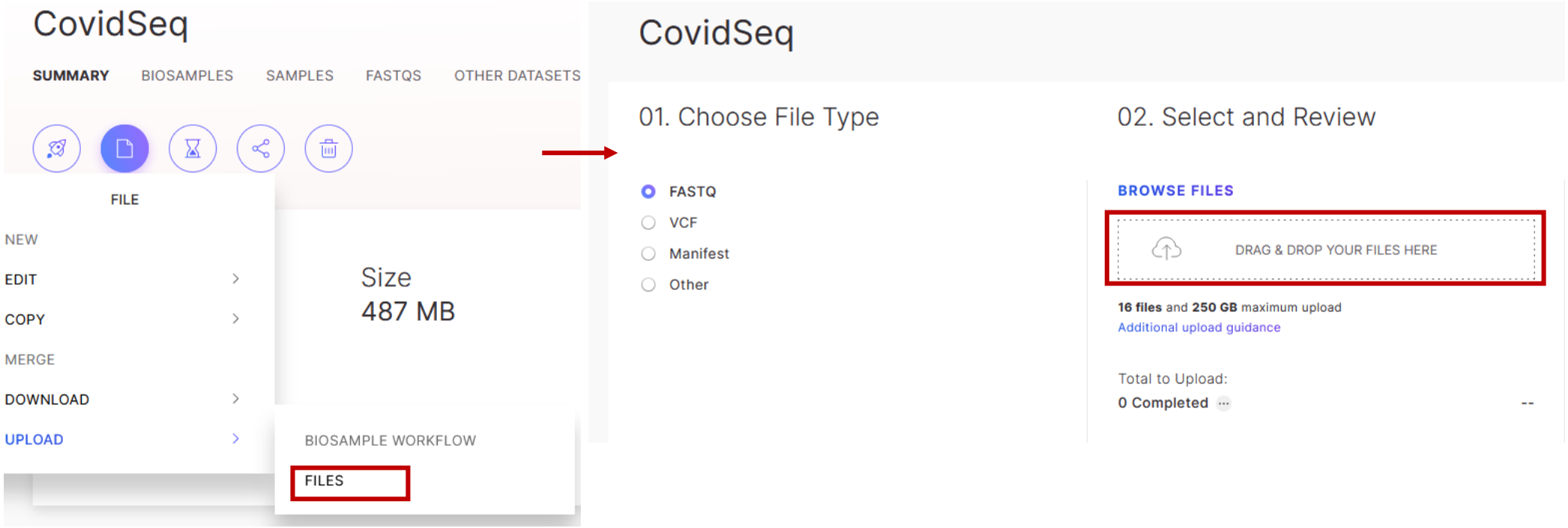
See All BaseSpace Data Analysis Apps

Illumina Basespace

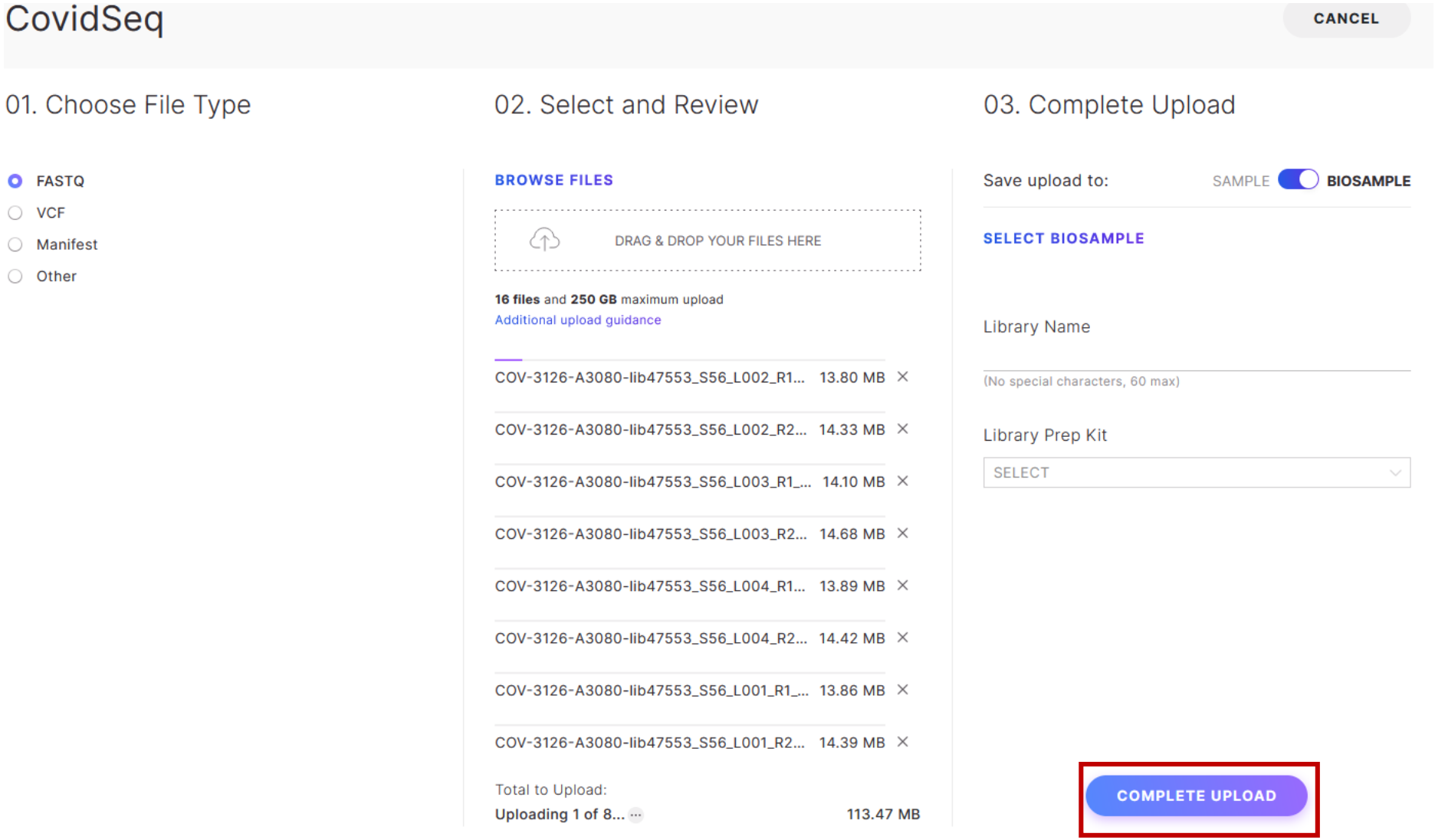




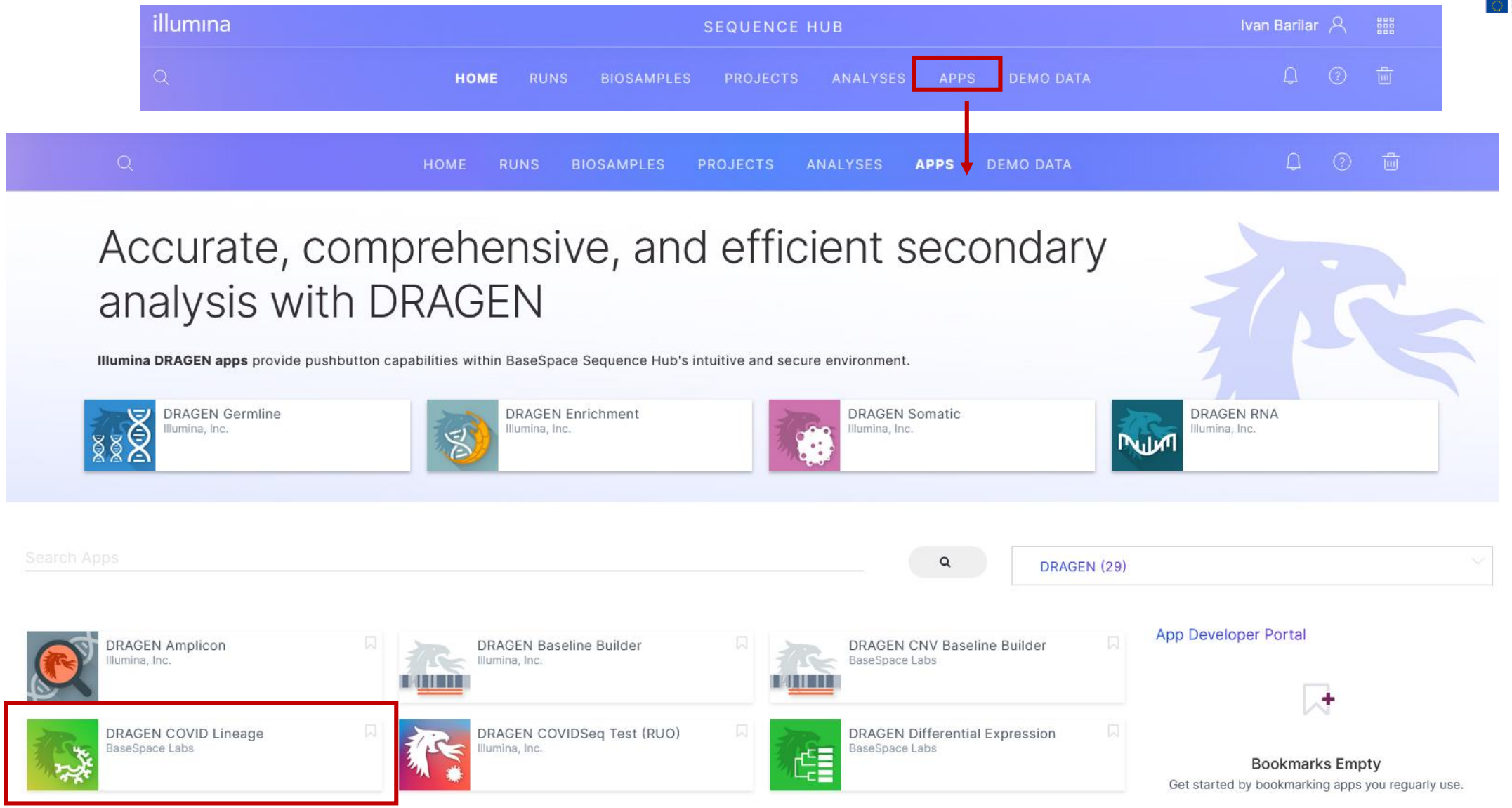
Illumina Basespace



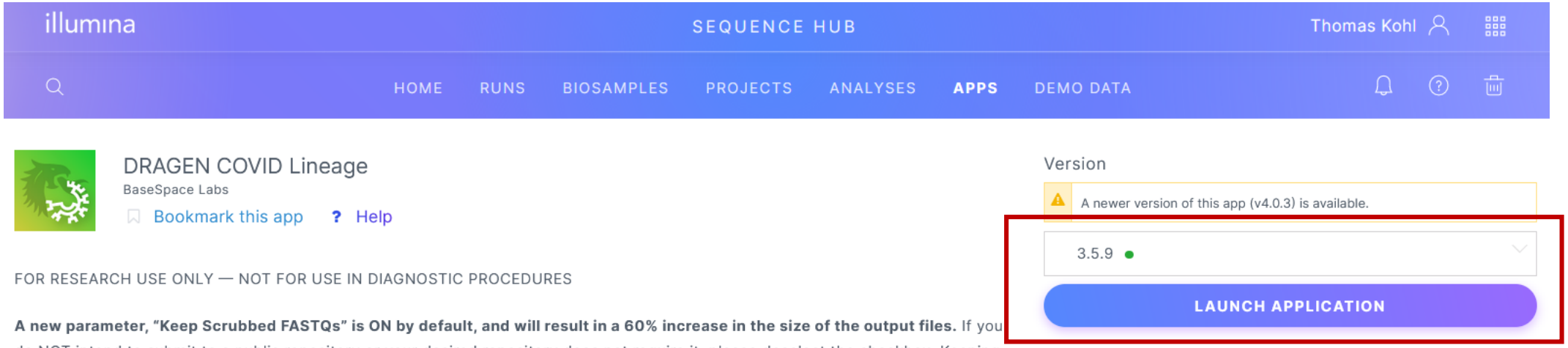
Illumina Basespace



Illumina DRAGEN COVID Lineage App



Illumina DRAGEN COVID Lineage App



Illumina DRAGEN COVID Lineage App

Input Data

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

- (•) 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 tim
app to be aborted. In such a case, split the samples into multiple Analysis runs

Input Type

- (•) Samples
- ( ) Project

Biosample
SELECT BIOSAMPLE(S):

Input Project
SELECT PROJECT

Reference genome / PCR primer set
Not Selected

Workflow Control
Enable or disable optional steps in the analysis.

Pangolin
[x] Enable Pangolin

NextClade
[x] Enable NextClade

+ Advanced Settings (Map/Align)

+ Advanced Settings (Variant Calling)

+ Advanced Settings (Consensus Sequence)

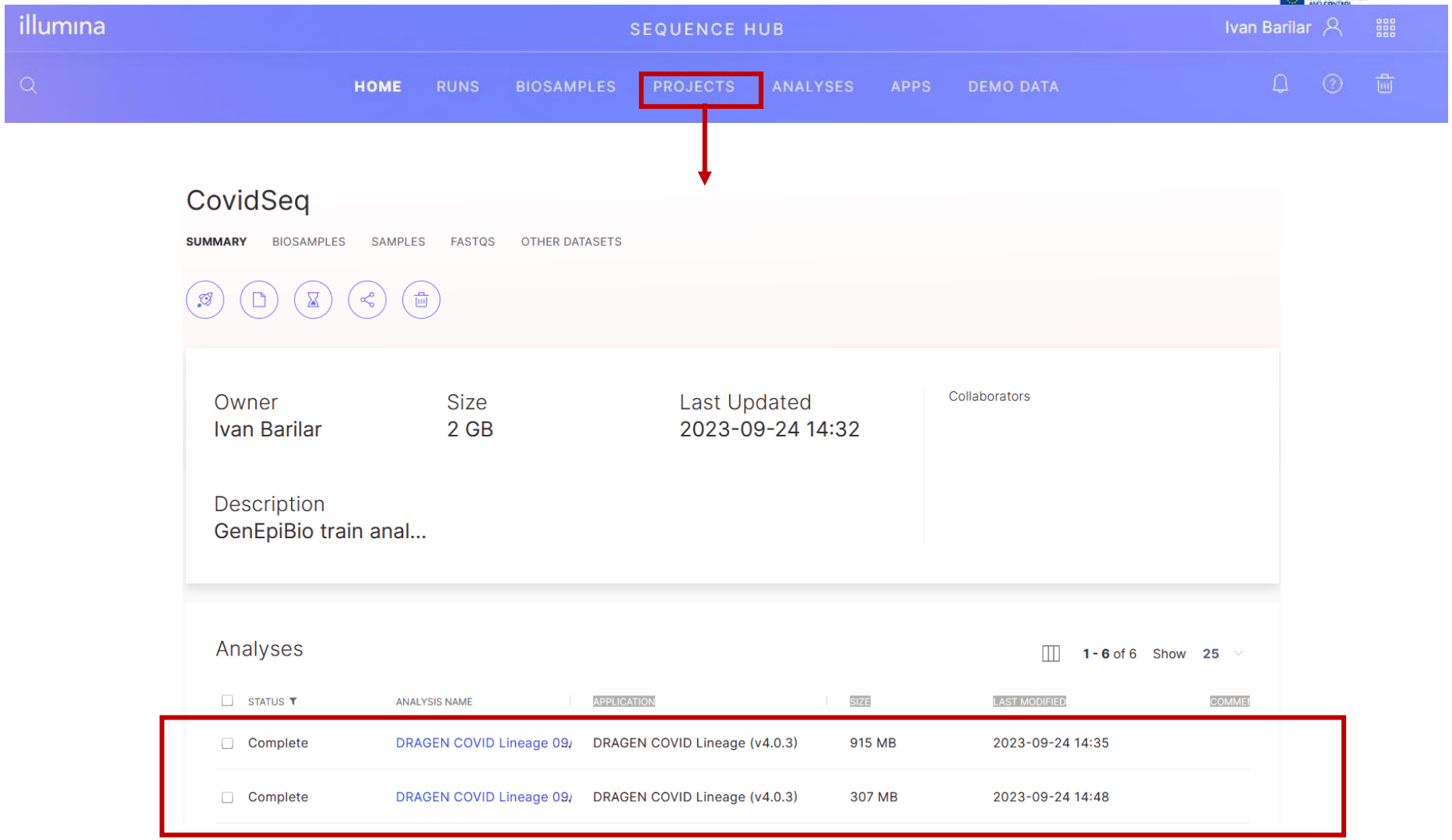
+ Advanced Settings (Pangolin & NextClade)

+ Advanced Settings (Output Files)

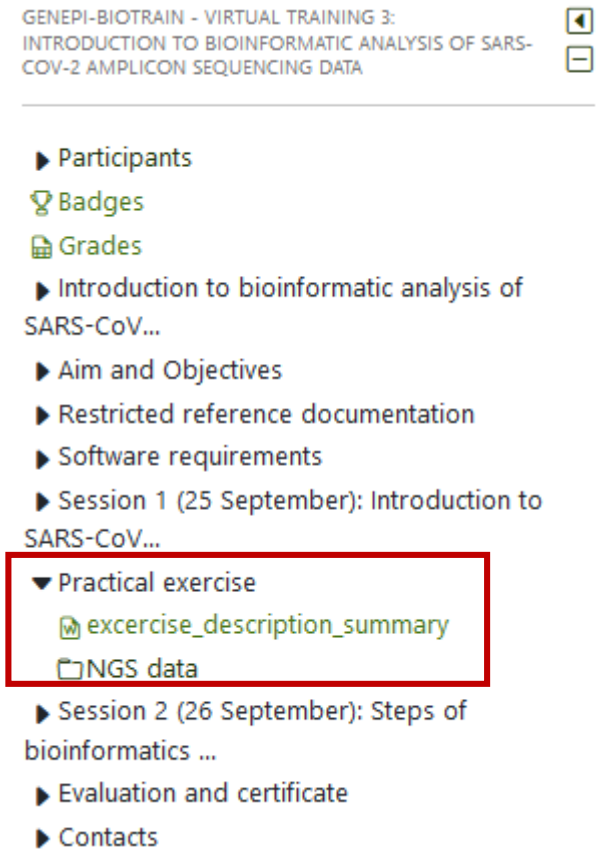
BaseSpace Labs
[ ] I acknowledge and agree that (i) this is a BaseSpace Lab
Illumina has no obligation to provide any technical suppo
including without limitation, any loss of data, incorrect re
App.

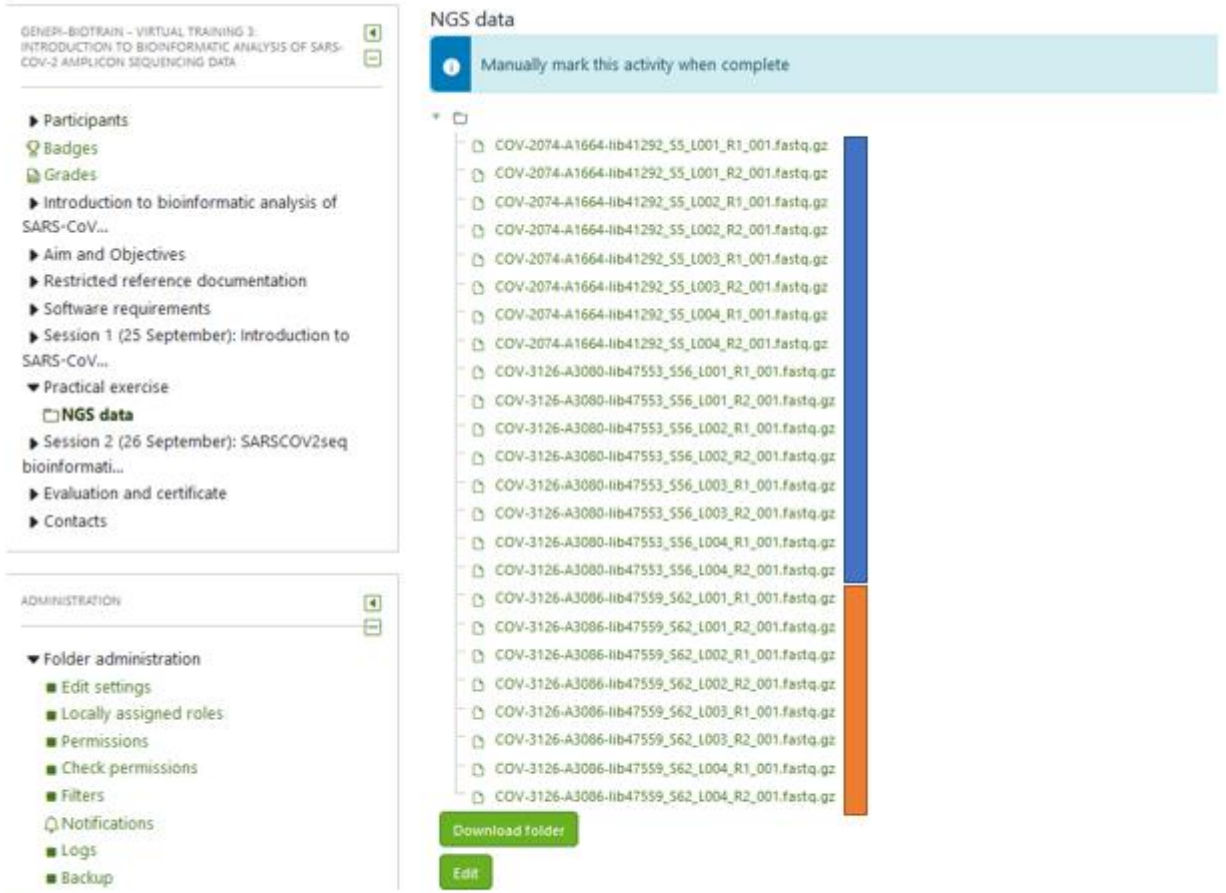
LAUNCH APPLICATION

Illumina Basespace



Exercise





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