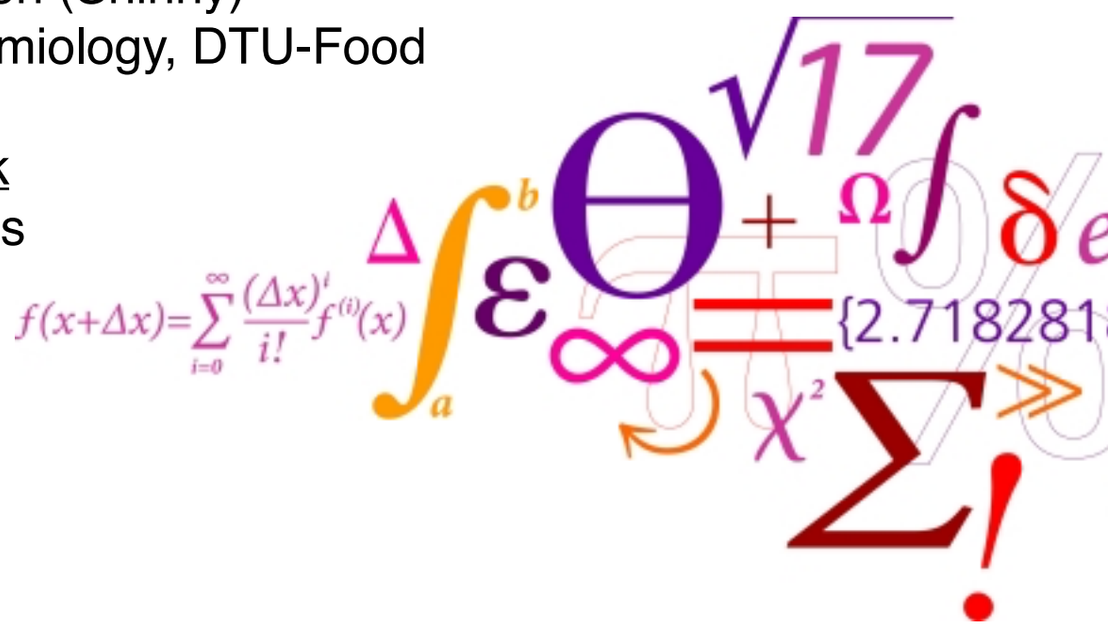
Sharing raw sequence data to ENA

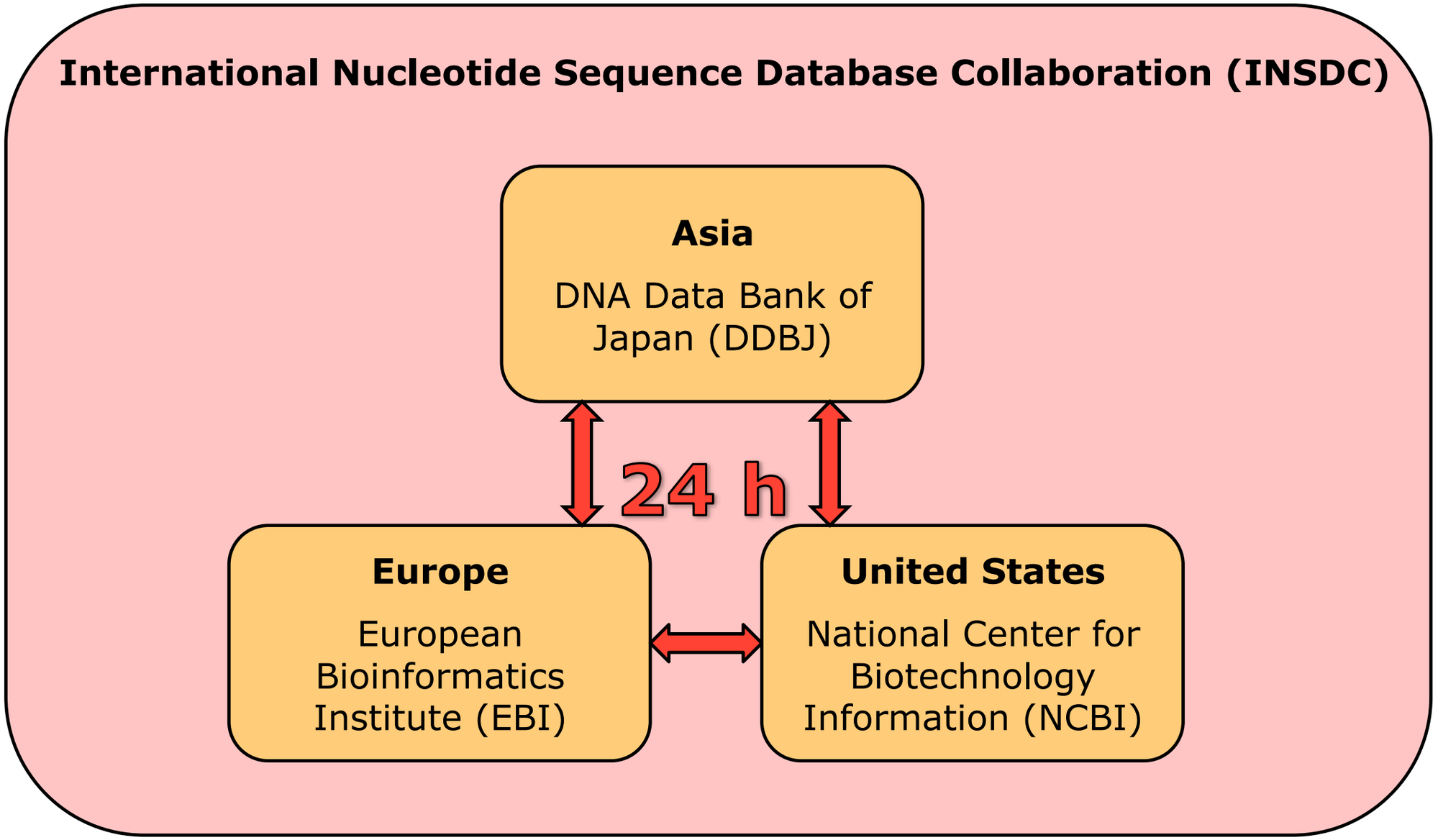
Pimlapas Leekitcharoenphon (Shinny)
Research Group of Genomic Epidemiology, DTU-Food

pile@food.dtu.dk
@ShinnyPimlapas



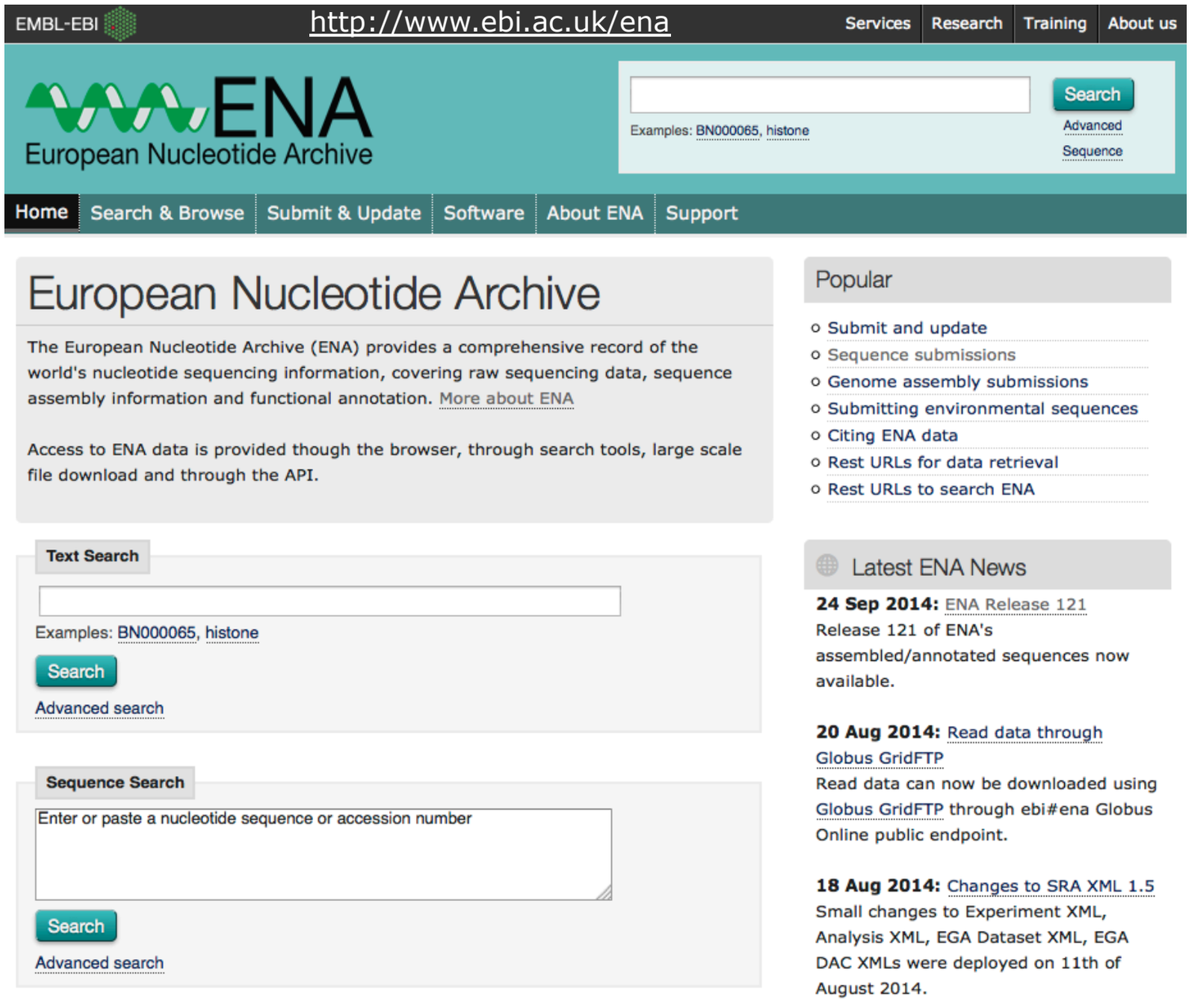
Infectious Disease Bioinformatics course 2022

Data storage & Access



Data storage & Access

European Bioinformatics Institute (EBI)



EMBL-EBI

Find | Help | Feedback

Databases | Tools | Research | Training | Industry | About Us | Help | Site Index

ENA

- ENA Home
- Search & Browse
- Submit & Update
- About ENA
- Contact

NEWS AND ANNOUNCEMENTS

CRAM toolkit 0.7 released
7 Mar 2012

CRAM toolkit 0.7 has been released. More information with download, installation and usage instructions are available here.

The future of sequence archiving
2 Mar 2012

Future of sequence archiving and the role of data compression is explored in a new paper from EBI to be published in Gigascience.

EBI Home » ENA Home ›

European Nucleotide Archive

The European Nucleotide Archive (ENA) provides a comprehensive record of the world's nucleotide sequencing information, covering raw sequencing data, sequence assembly information and functional annotation ... more

Access to ENA data is provided though the browser, through search tools, large scale file download and through the API.

ENA
European Nucleotide Archive

Text search

Enter search query, for example: BN000065 Search

Sequence Search Advanced Search

Enter or paste a nucleotide sequence or accession number

Search

www.ebi.ac.uk/ena/

EMBL-EBI | Services | Research | Training | About us | EMBL-EBI Hinxton

EBI Search

PRJEB22091 x

Examples: VAV_HUMAN , tpi1 , Sulston ...

Build Query

Help & Documentation | About EBI Search | Feedback

EMBL-EBI to be HTTPS by default from 1st October

On the 1st October the majority of services hosted on www.ebi.ac.uk will be served over HTTPS by default. Services that are becoming HTTPS by default will automatically redirect users accessing the site on insecure HTTP URLs to secure HTTPS URLs.
Users of EMBL-EBI services may wish to update links, bookmarks or API clients to use the HTTPS URLs.

Search results for ***PRJEB22091***

Showing **15** results out of **129** in All results

Filter your results

Source

All results (129)

Nucleotide sequences (129)

Nucleotide sequences (129 results)

PRJEB22091

Danmap 2016

Related data

Source: Study
ID: PRJEB22091

Navigation | Read Files | Portal | Attributes

Bulk Download Files ⚠ (Please use Firefox to launch the bulk downloader app.)

Download: 1 - 127 of 127 results in TEXT

Select columns

Showing results 1 - 10 of 127 results

Study accession	Sample accession	Secondary sample accession	Experiment accession	Run accession	Tax ID	Scientific name	Instrument model	Library layout	FASTQ files (FTP)	FASTQ files (Galaxy)	Submitted files (FTP)	Submitted files (Galaxy)	NCBI SRA file (FTP)	NCBI SRA file (Galaxy)	CRAM Index files (FTP)	CRAM Index files (Galaxy)
PRJEB22091	SAMEA104205896	ERS1864914	ERX2148098	ERR2091304	562	Escherichia coli	Illumina MiSeq	PAIRED	File 1 File 2	File 1 File 2	Fastq file 1 Fastq file 2	Fastq file 1 Fastq file 2	File 1	File 1		
PRJEB22091	SAMEA104205897	ERS1864915	ERX2148099	ERR2091305	562	Escherichia coli	Illumina MiSeq	PAIRED	File 1 File 2	File 1 File 2	Fastq file 1 Fastq file 2	Fastq file 1 Fastq file 2	File 1	File 1		
PRJEB22091	SAMEA104205898	ERS1864916	ERX2148100	ERR2091306	562	Escherichia coli	Illumina MiSeq	PAIRED	File 1 File 2	File 1 File 2	Fastq file 1 Fastq file 2	Fastq file 1 Fastq file 2	File 1	File 1		
PRJEB22091	SAMEA104205899	ERS1864917	ERX2148101	ERR2091307	562	Escherichia coli	Illumina MiSeq	PAIRED	File 1 File 2	File 1 File 2	Fastq file 1 Fastq file 2	Fastq file 1 Fastq file 2	File 1	File 1		
PRJEB22091	SAMEA104205900	ERS1864918	ERX2148102	ERR2091308	562	Escherichia coli	Illumina MiSeq	PAIRED	File 1 File 2	File 1 File 2	Fastq file 1 Fastq file 2	Fastq file 1 Fastq file 2	File 1	File 1		
PRJEB22091	SAMEA104205901	ERS1864919	ERX2148103	ERR2091309	562	Escherichia coli	Illumina MiSeq	PAIRED	File 1 File 2	File 1 File 2	Fastq file 1 Fastq file 2	Fastq file 1 Fastq file 2	File 1	File 1		
PRJEB22091	SAMEA104205902	ERS1864920	ERX2148104	ERR2091310	562	Escherichia coli	Illumina MiSeq	PAIRED	File 1 File 2	File 1 File 2	Fastq file 1 Fastq file 2	Fastq file 1 Fastq file 2	File 1	File 1		
PRJEB22091	SAMEA104205903	ERS1864921	ERX2148105	ERR2091311	562	Escherichia coli	Illumina MiSeq	PAIRED	File 1 File 2	File 1 File 2	Fastq file 1 Fastq file 2	Fastq file 1 Fastq file 2	File 1	File 1		

Submitting data in ENA

- Manual and automated

- Use Webin (interactive web submission system) for new sequencing projects, assembled sequences and annotation

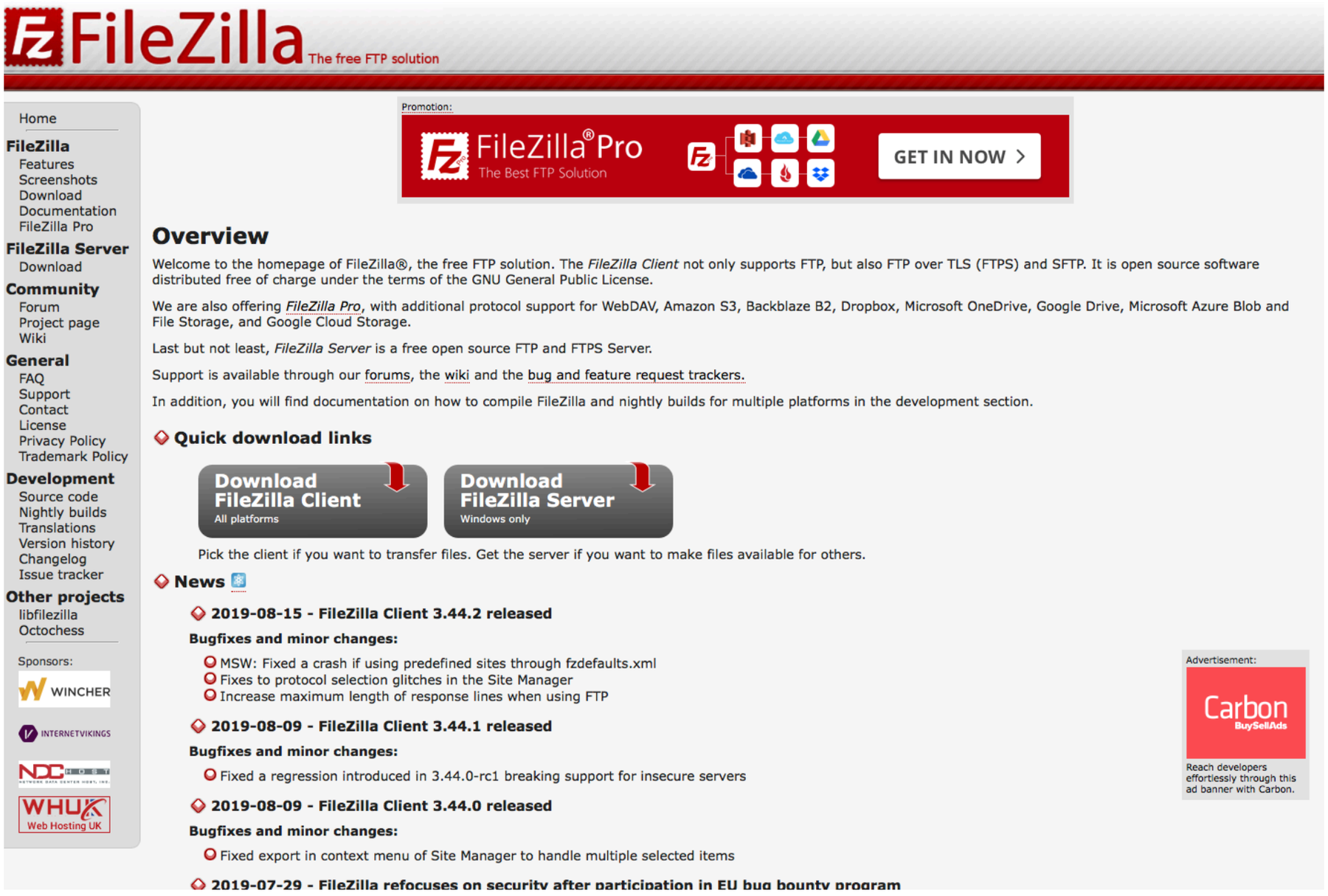
- For other types of data, different channels are available (e.g. datasubs@ebi.ac.uk, FTP, RESTful web-based service)

- www.ebi.ac.uk/ena/about/submit_and_update

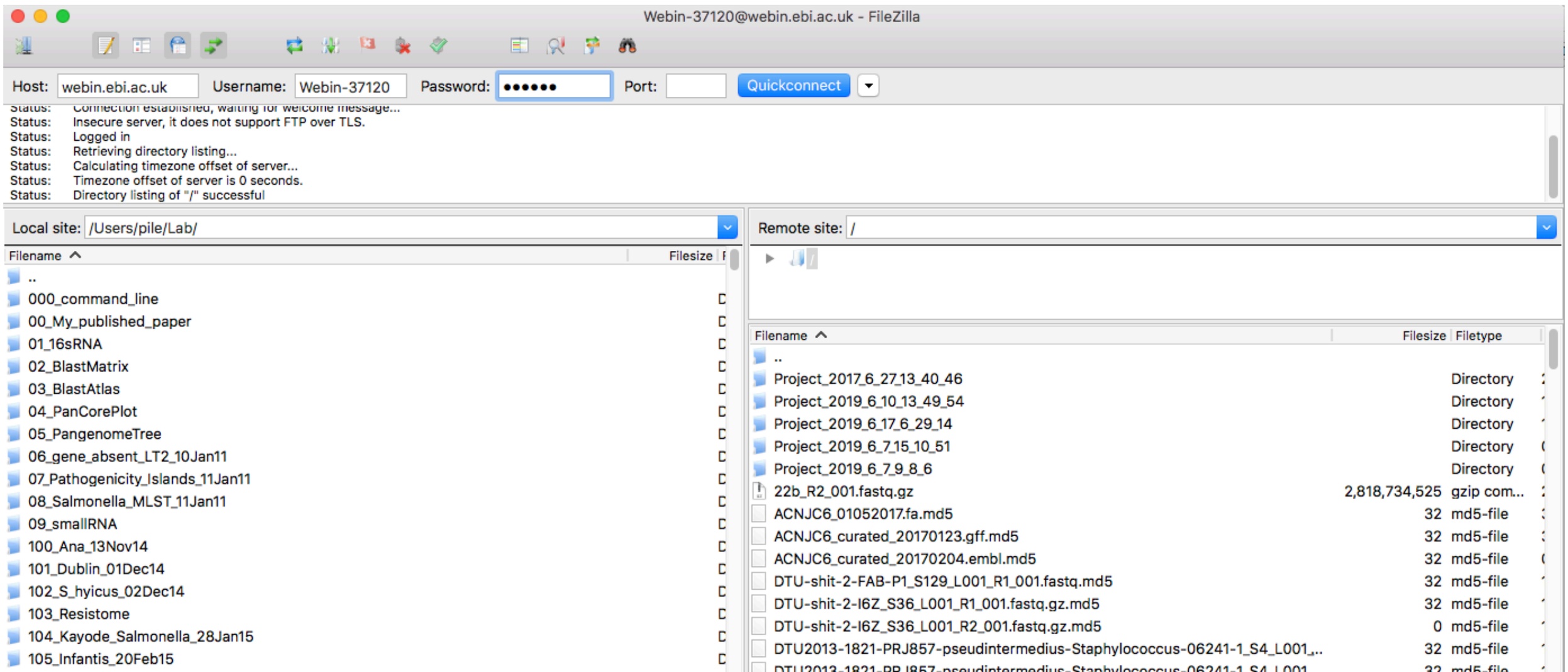
Submitting data in ENA

- 1) Install a program to connect to FTP

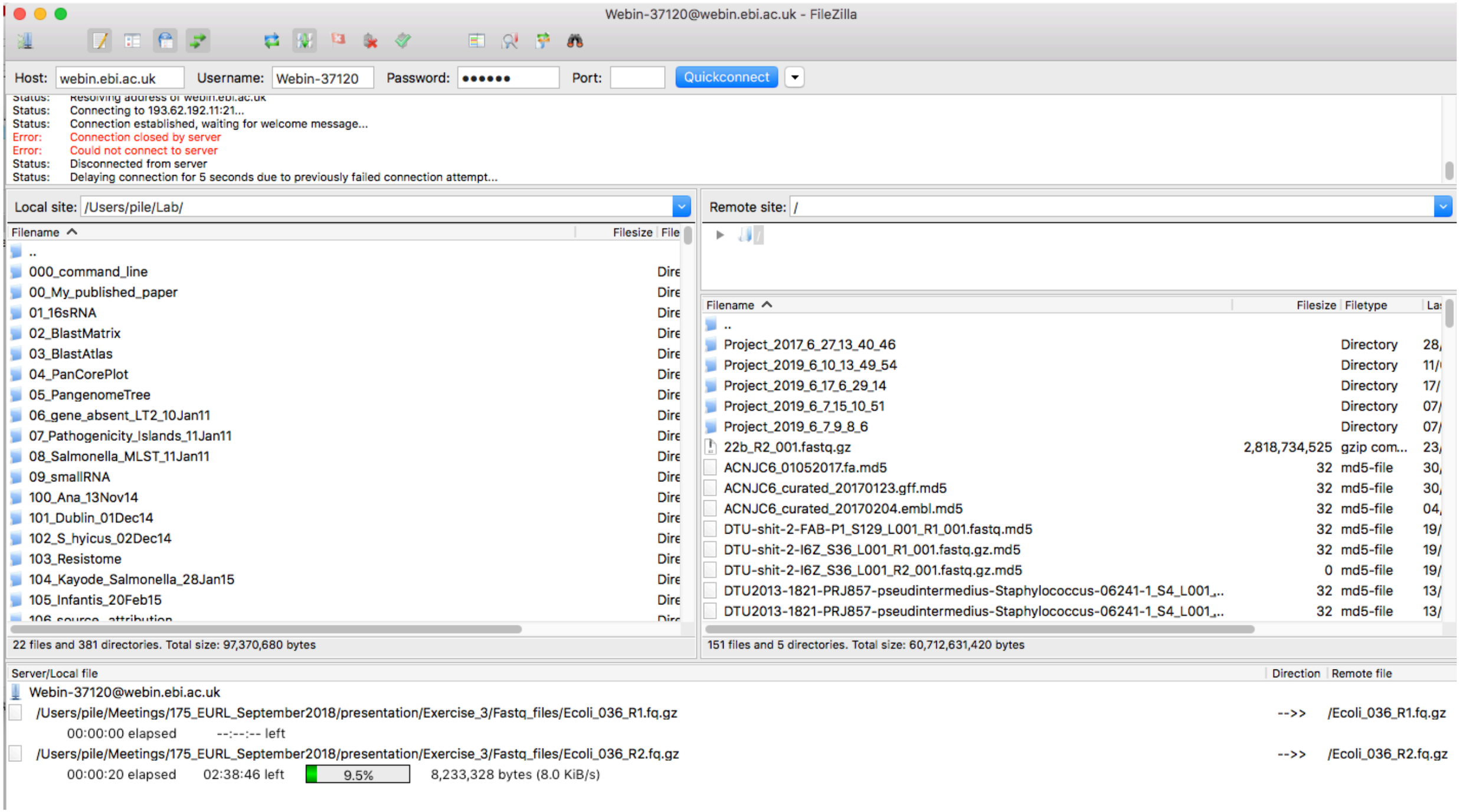
- https://filezilla-project.org



- 2) Connect to FTP
- Host: webin.ebi.ac.uk
- Username: Your own Webin ID
- Password: Your own password for your ENA account. Then click ‘Connect’



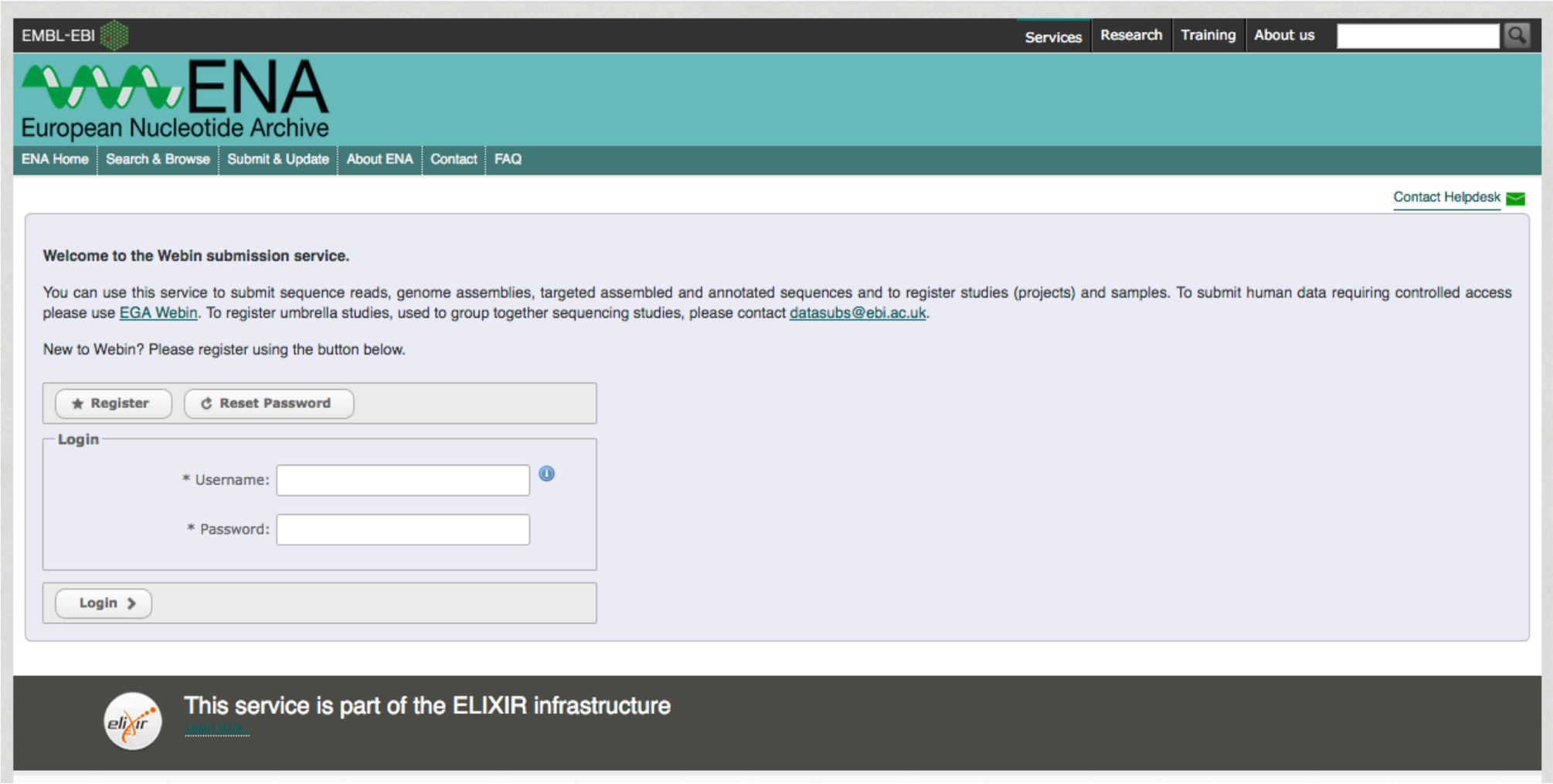
- 3) Transfer the genome files to ENA FTP



Submitting data in ENA

- 4) log in ENA with your own webin account

- https://www.ebi.ac.uk/ena/submit/sra/#home



EMBL-EBI Services Research Training About us

ENA
European Nucleotide Archive

ENA Home | Search & Browse | Submit & Update | About ENA | Contact | FAQ

Contact Helpdesk **Webin-37120** Logout

Home | New Submission | Studies | Samples | Runs | Analyses

Welcome to the Webin submission service.

You can use this service to submit sequence reads, genome assemblies, targeted assembled and annotated sequences and to register studies (projects) and samples. To register umbrella studies, used to group together sequencing studies, please contact datasubs@ebi.ac.uk.

Help about submitting read data

▶ My account details

This service is part of the ELIXIR infrastructure
Learn more ›

- 5) click ‘New Submission’ and add ‘Center Name’. Then click ‘Next’

EMBL-EBI Services Research Training About us

ENA
European Nucleotide Archive

ENA Home | Search & Browse | Submit & Update | About ENA | Contact | FAQ

Contact Helpdesk Webin-37120 Logout

Home | New Submission | Studies | Samples | Runs | Analyses

Start >> Study >> Sample >> Run >> Finish

You can use this service to submit sequence reads, genome assemblies, targeted assembled and annotated sequences and to register studies (projects) and samples. To register umbrella studies, used to group together sequencing studies, please contact datasubs@ebi.ac.uk.

new-submission-start-view-action-panel

Please select the type of submission you would like to make:

Submit sequence reads and experiments

We recommend that Fastq, BAM, and CRAM read files are submitted using Webin-CLI.

When using this interface instead of Webin-CLI, raw sequences must be uploaded in one of the supported data formats before they can be submitted. All data submitted in a single submission will be associated with the same study. Data for different studies must be submitted in separate submissions. The study and the sequenced samples can be either pre-registered or registered during the submission process. Please note that each individual study and sample should be registered only once. In addition, you will be asked to provide information about the sequencing libraries and instruments. Please quote the study accession number (ERP*) when citing data submitted to ENA.

Read here for more information on how submit your raw reads to ENA.

- Register study (project)
- Taxonomy Check/Request
- Register samples
- Submit genome assemblies
- Submit other assembled and annotated sequences [formerly EMBL-Bank]

The first step of your submission is to upload data files. Data files can be uploaded using FTP or Aspera, or using the Webin File Uploader. If you have already uploaded your data files into your Webin upload area please proceed directly to the next step. Please note that unsubmitted files that are older than 2 months will be deleted as explained in our Fair Use Policy.

Specify the name of the center you are submitting for.

Centre Name: (*) Technical University of Denm

Next >>

- 6) Click ‘Create a new study’

EMBL-EBI Services Research Training About us

ENA
European Nucleotide Archive

ENA Home | Search & Browse | Submit & Update | About ENA | Contact | FAQ

Contact Helpdesk Webin-37120 Logout

Home | New Submission | Studies | Samples | Runs | Analyses

Start ✔ >> **Study** >> Sample >> Run >> Finish

Select an existing study or Create a new study

Search by: Accession / Unique name: ERP...

Show: accession unique name Reset

	Primary Accession	Secondary Accession	Title	Submission Date	Status
○	PRJEB33587	ERP116393	Whole genome sequencing of non-typhoidal Salmonellae serovars isolated from humans and poultry in Ma [...]	19-Jul-2019	Public
○	PRJEB33169	ERP115939	EFSA_EURL-AR_Confirmatory_Testing_2017	24-Jun-2019	Confidential
○	PRJEB32144	ERP114785	Formalin-fixed paraffin-embedded (FFPE) ring trial	12-Apr-2019	Public
○	PRJEB32083	ERP114714	Bacterial alpha diversity trajectories and immune-microbial associations of the gut, oral, and nasal [...]	08-Apr-2019	Confidential
○	PRJEB31650	ERP114226	Metagenomic assessment of the effect of sample processing on different microbiomes	11-Mar-2019	Public
○	PRJEB30991	ERP113496	The ESBL/AmpC resistance pool in Escherichia coli in pigs and pig farmers Vietnam	28-Jan-2019	Confidential
○	PRJEB30894	ERP113384	The gut, oral, and nasal microbiota in pediatric allogeneic HSCT and microbial marker candidates for [...]	21-Jan-2019	Confidential
○	PRJEB30892	ERP113382	Food Metagenomic ring trial - smoked salmon spiked with a microbial mock community	21-Jan-2019	Confidential
○	PRJEB30604	ERP113078	Genomic of Vibrio Cholera from Tanzania	07-Jan-2019	Confidential
○	PRJEB27955	ERP110095	Comamonas testosteroni	27-Jul-2018	Public

1-10 of 67 Number of rows 100

<< Previous Next >>

- 7) Set release date, provide title, and abstract

EMBL-EBI Services Research Training About us

ENA
European Nucleotide Archive

ENA Home | Search & Browse | Submit & Update | About ENA | Contact | FAQ

Contact Helpdesk Webin-37120 Logout

Home | New Submission | Studies | Samples | Runs | Analyses

Start ✔ >> **Study** >> Sample >> Run >> Finish

Select an existing study or Create a new study

You can use this page to submit a new study into the European Nucleotide Archive (ENA).

Please answer the questions below. Mandatory fields are denoted by (*).

If you need to create umbrella studies to group several studies together, please contact us at datasubs@ebi.ac.uk. You can then link your umbrella study by using the 'Studies' report tab, click on the 'Edit' button of the study you want to edit and use the 'Edit ENA links' tab.

Please specify the release date of your study:
This is when your study will be made public.
30-Nov-2019

Please provide a short name for the study:

Please provide a short descriptive title for the study: (*)
E.coli

Please provide an abstract to describe the study in detail: (*)
E.coli

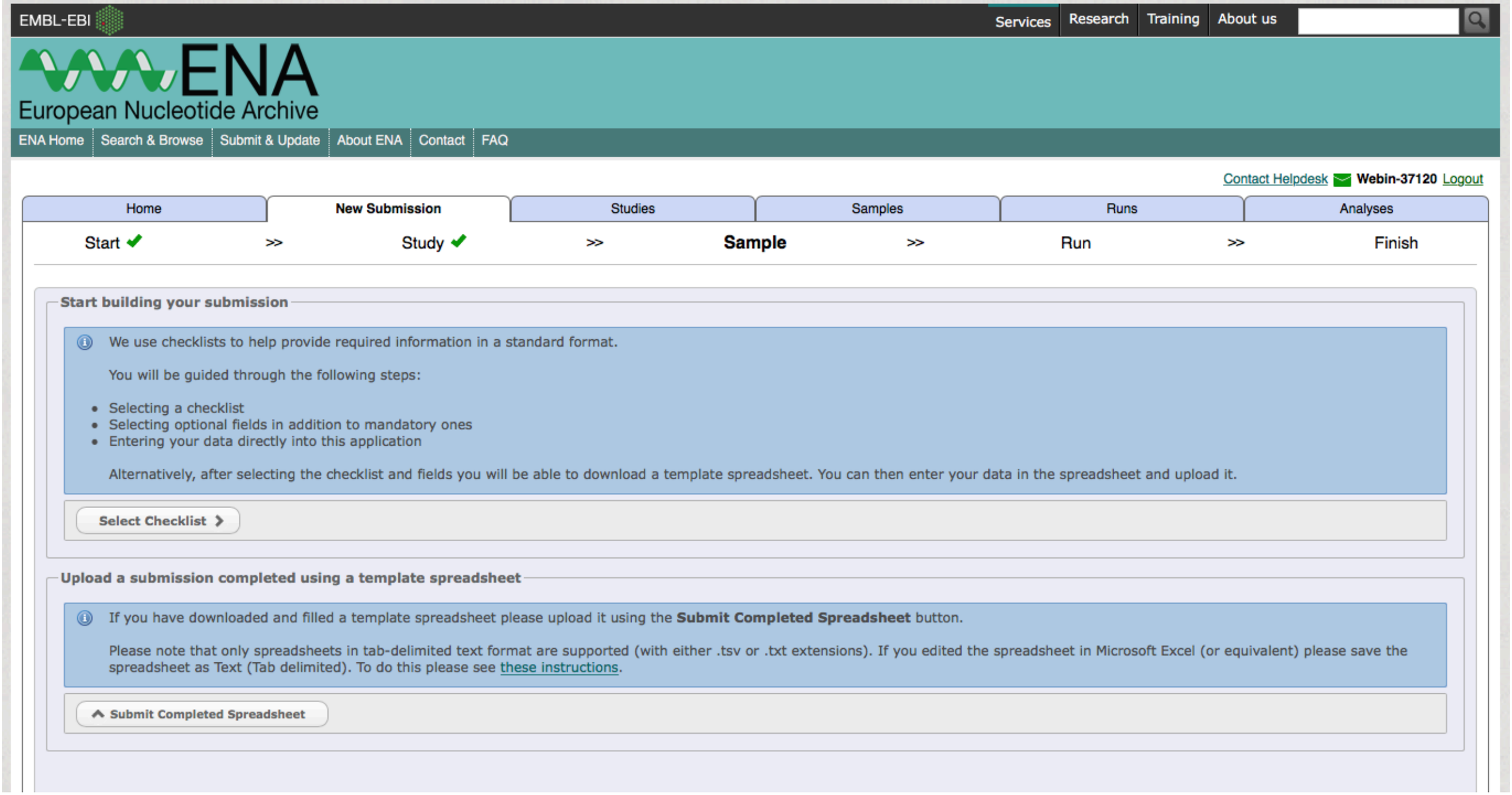
Please provide attributes to add a deeper description of the study:
Tag FieldType
Add

Please provide PubMed IDs of publications you want to associate with the study:
(numeric value)
PubMed IDs
Add

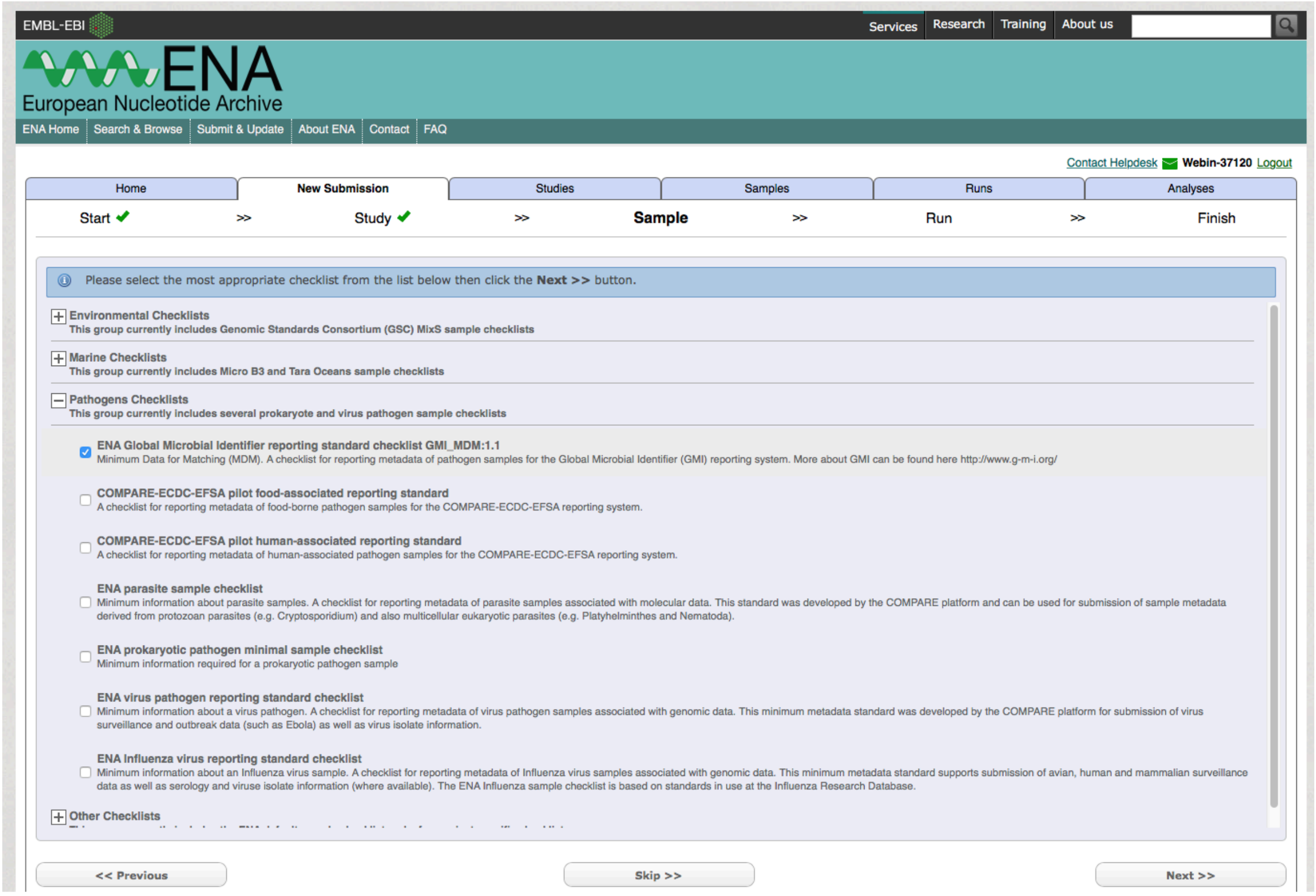
For genome assembly projects only: In this study, will you provide functional genome annotation? (*)
PLEASE ANSWER WITH YES IF YOU HAVE ANNOTATION: Locus tag prefixes are only associated to studies providing functional genome annotation.
○ Yes
◉ No

<< Previous Next >>

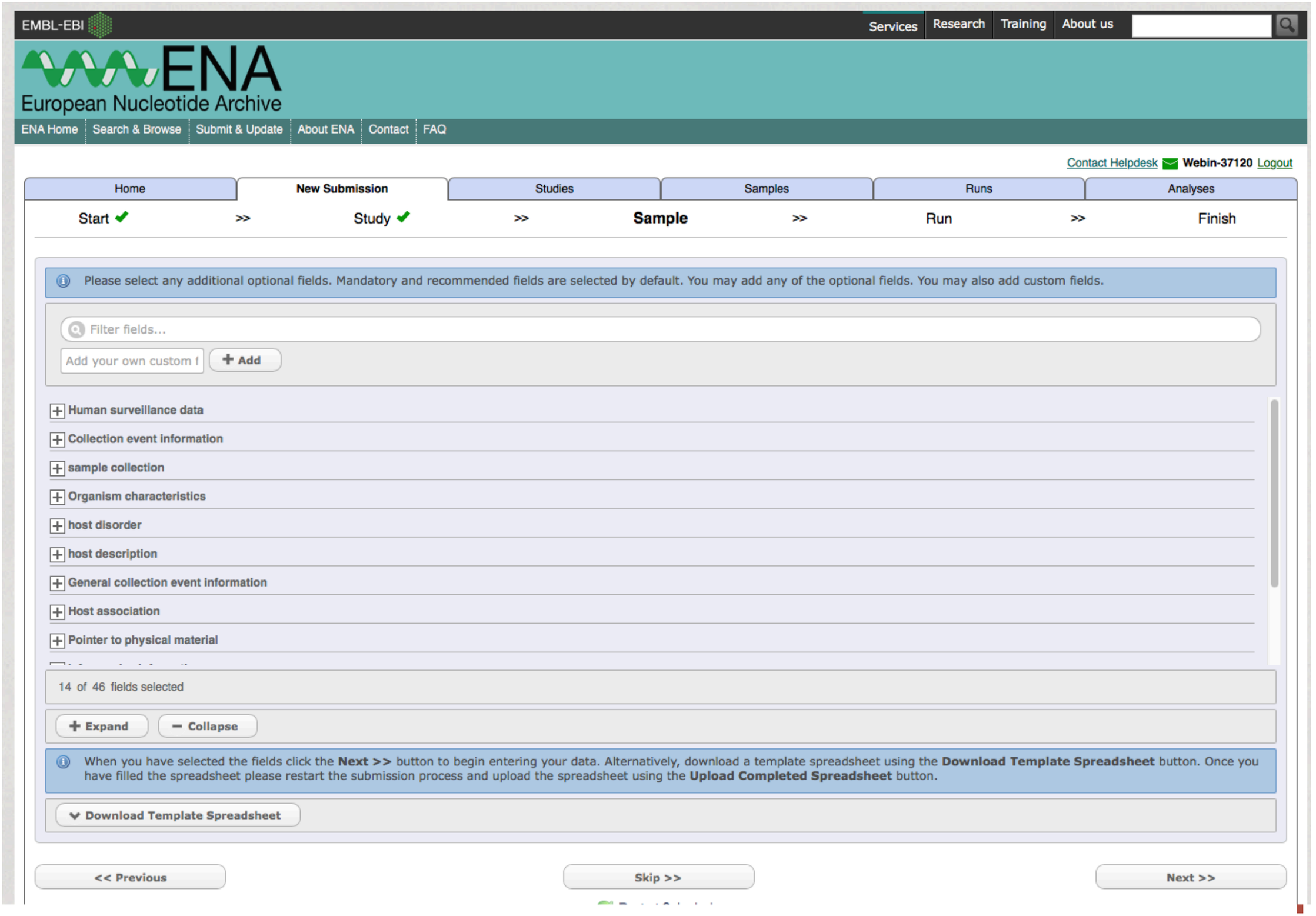
- 8) ‘Select Checklist’



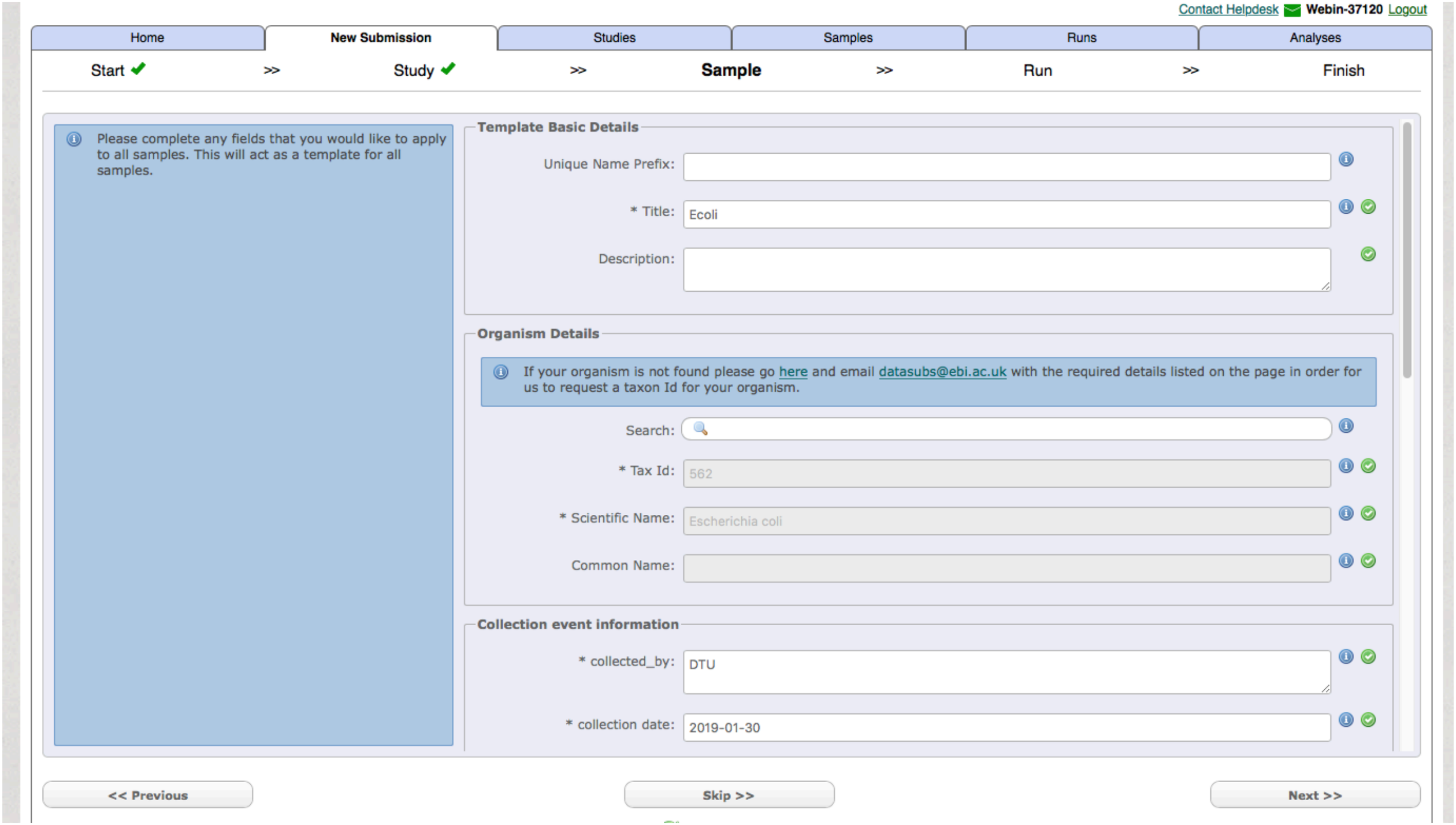
- 9) Select ‘Pathogens Checklists’ and then select ‘ENA GMI_MDK:1.1’



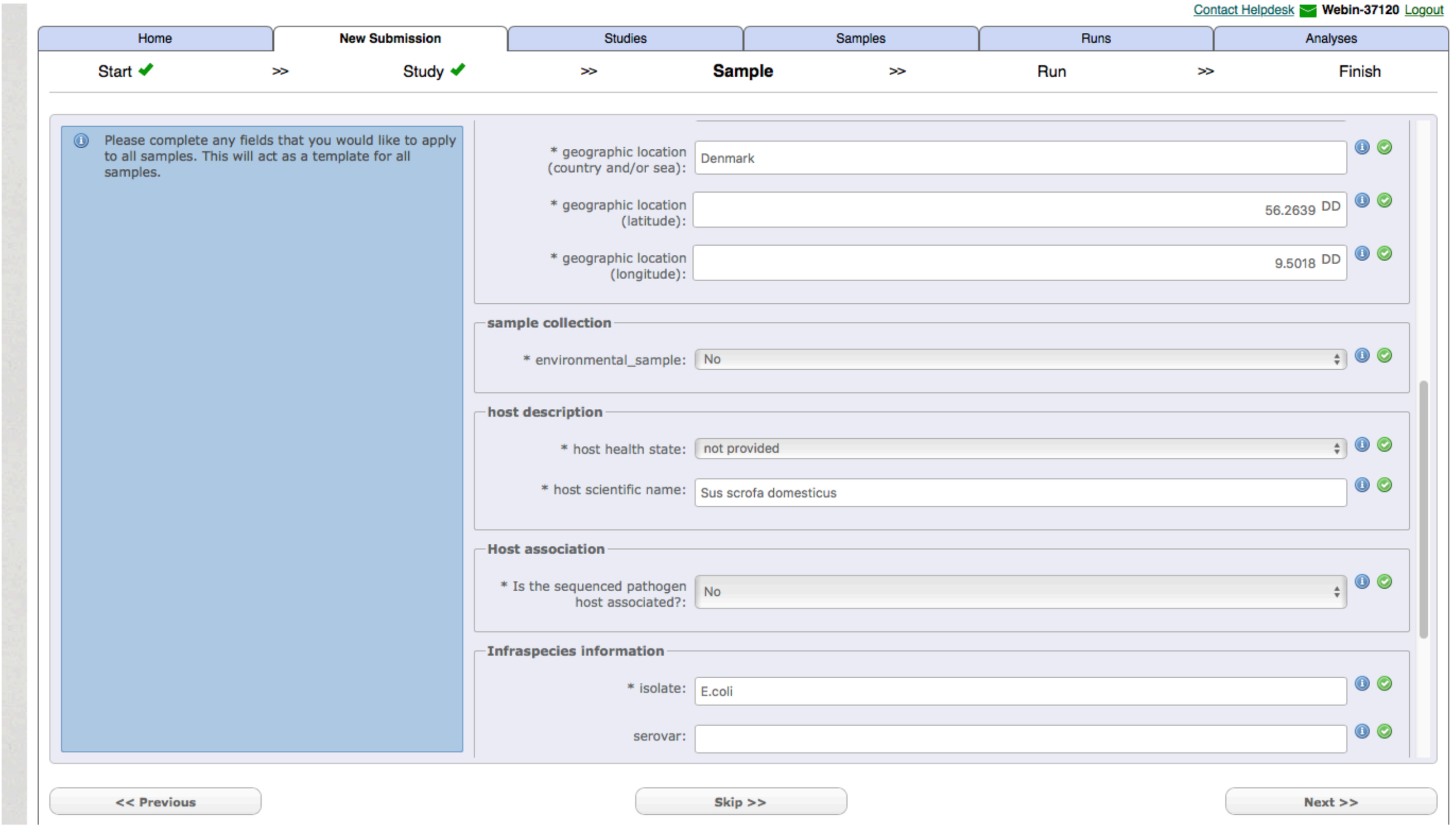
- 10) You have an option to download metadata as excel template or you just click ‘Next’ to fill out metadata via website



- 11) Add ‘title’, search ‘Escherichia coli’ (you have to wait until the system give you organism name to choose), add ‘collected_by’, add ‘collection date’



- 12) Add ‘country’, ‘latitude’, ‘longtitude’, ‘environmental_sample’, host health state’, ‘host scientific name’, ‘host associated’ and ‘isolate’



- 13) add number of sample. In this case, it is only 1 sample to add (click '+Add')

Contact Helpdesk Webin-37120 Logout

Home | New Submission | Studies | Samples | Runs | Analyses

Start ✔ >> Study ✔ >> **Sample** >> Run >> Finish

Please add samples to the submission. Multiple samples can be created by increasing the number by the add button

+ Add 1 samples

1

1-1 of 1

Please submit by clicking the **Submit** Button. Alternatively, download your data as a spreadsheet using the **Download Spreadsheet** button. Once you have filled the spreadsheet please restart the submission process and upload the spreadsheet using the **Upload Completed Spreadsheet** button.

Download Spreadsheet

There are no samples to submit

Previous Sample | Next Sample

Basic Details

* Unique Name: 1

* Title: Ecoli

Description:

Organism Details

If your organism is not found please go here and email datasubs@ebi.ac.uk with the required details listed on the page in order for us to request a taxon Id for your organism.

Search:

* Tax Id: 562

* Scientific Name: Escherichia coli

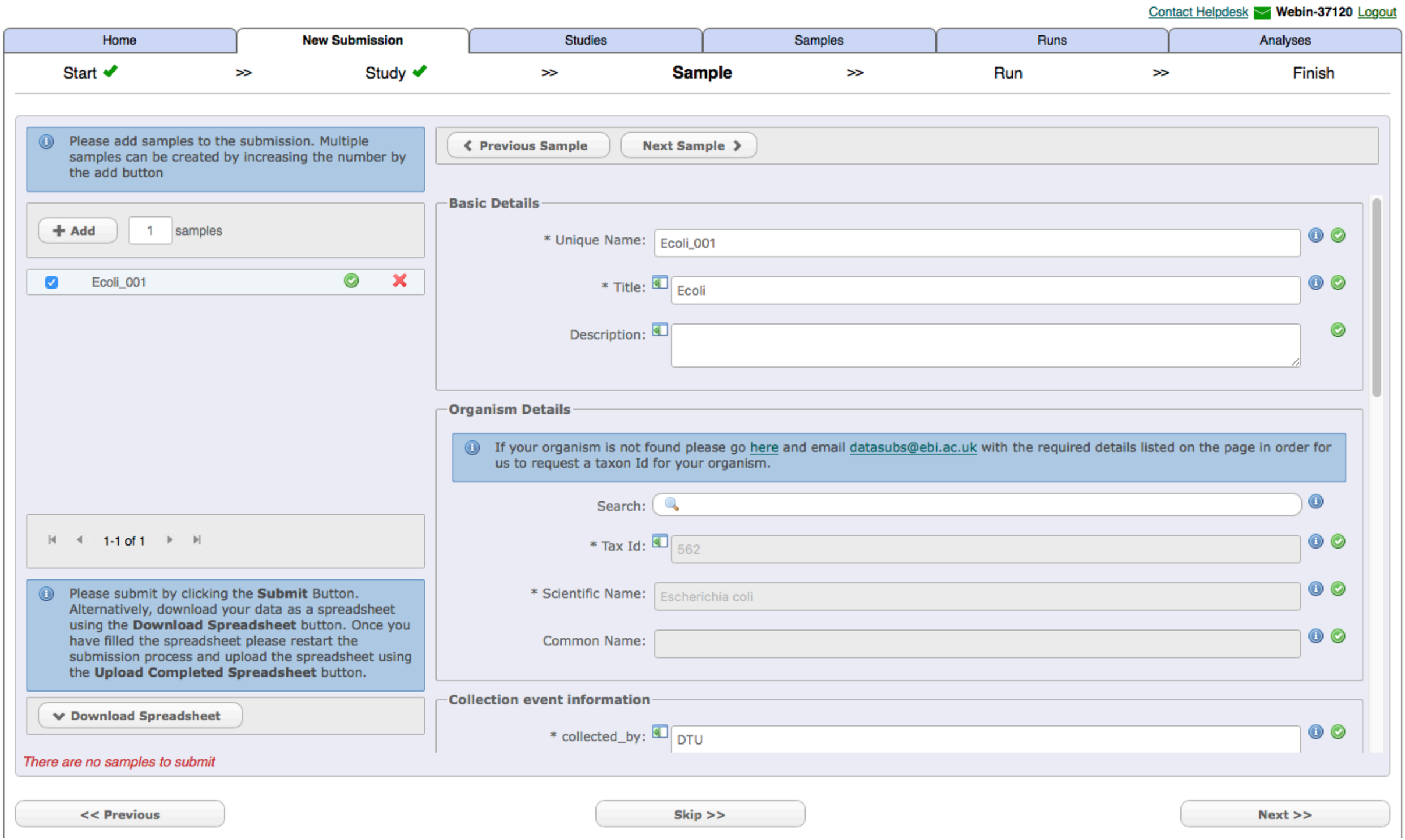
Common Name:

Collection event information

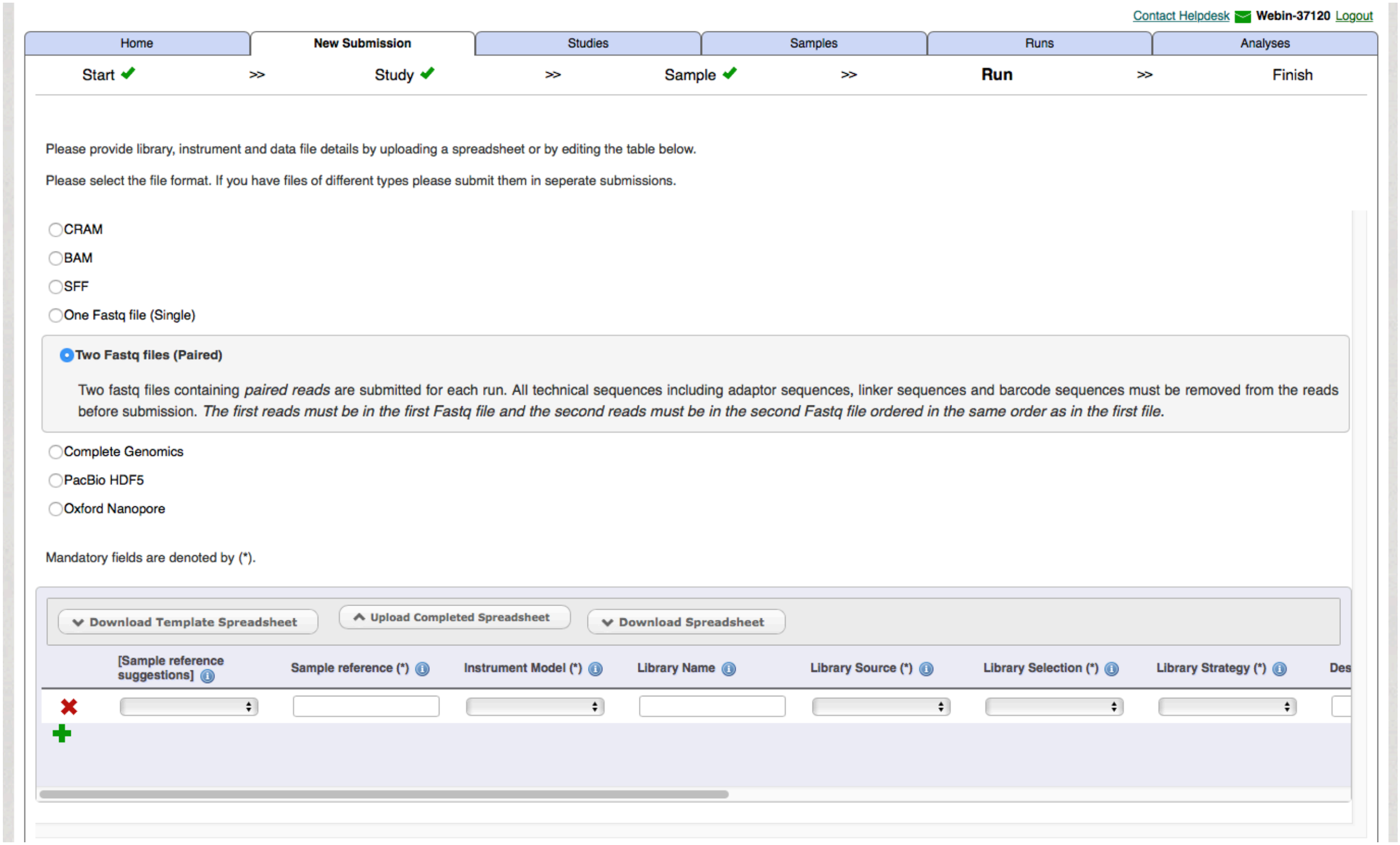
* collected_by: DTU

<< Previous | Skip >> | Next >>

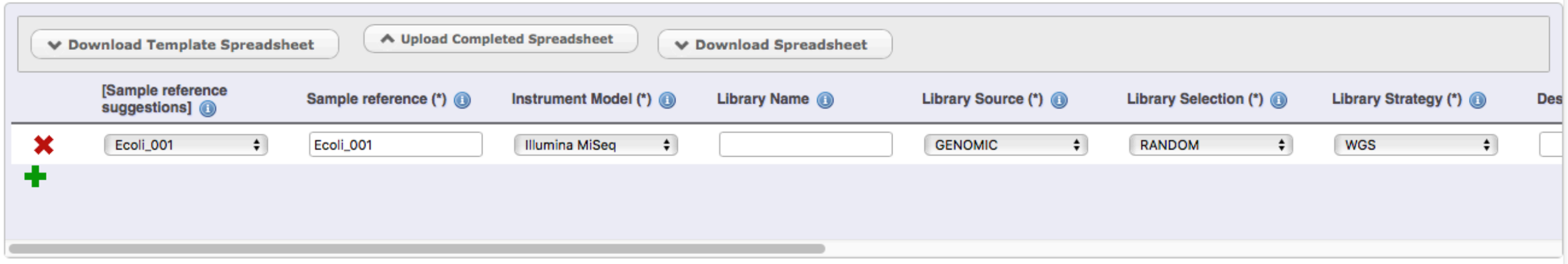
- 14) Add ‘Unique Name’ for the submitted sample. In this case, try to add ‘Ecoli_001’

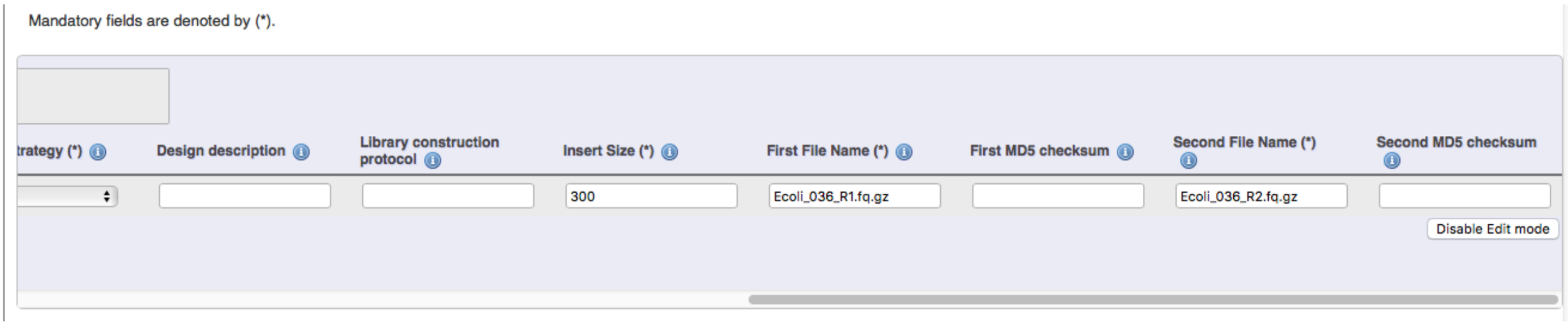


- 15) Click type of submitted files. In this case is ‘Two Fastq files (Paired)’

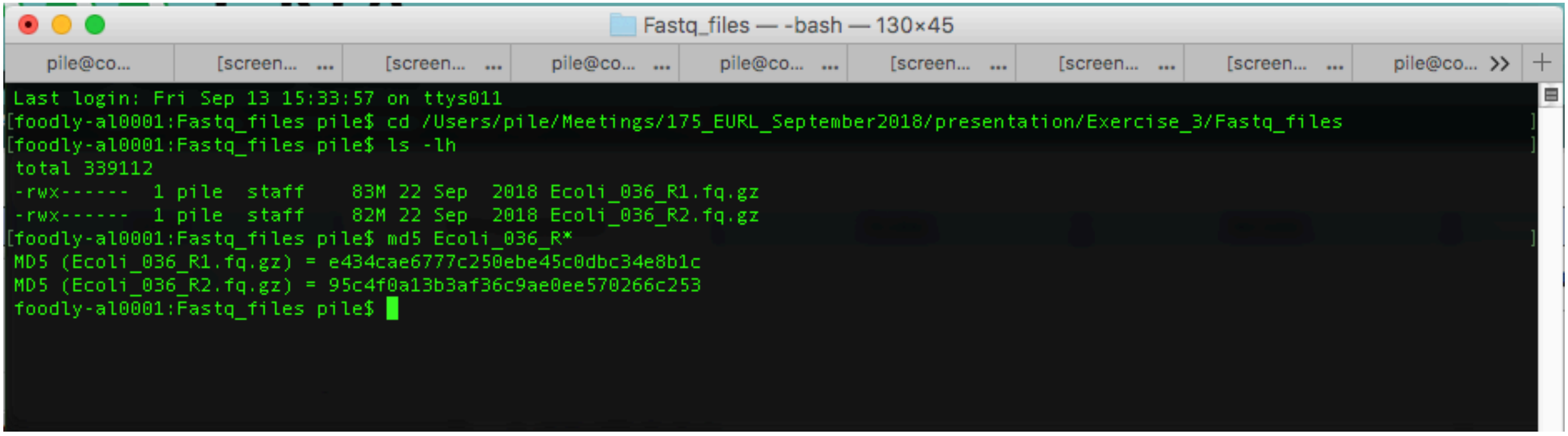
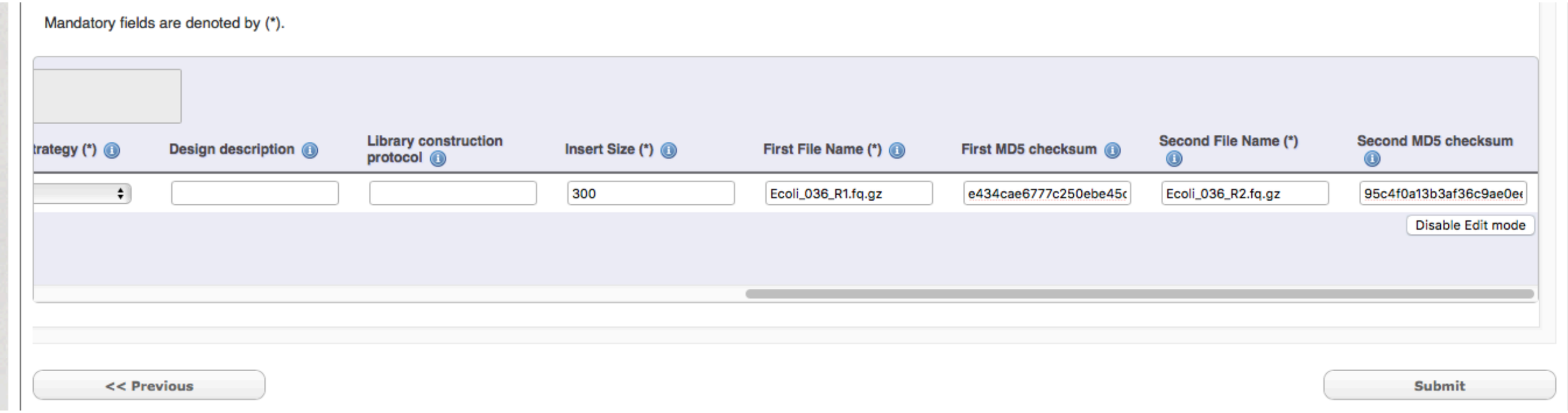


- 16) Add required metadata (*)





- 17) MD5 checksum is also required



- 18) The last step is to click ‘Submit’. If something goes wrong, the system will give you error message.