Monkeypox virus Submissions

Please see below for instructions on how to submit Monkeypox virus related data. If you have any queries or require assistance with your submission please contact us at: virus-dataflow@ebi.ac.uk.

Monkeypox Data Drag and Drop Tool

To facilitate rapid data sharing we are offering a drag and drop tool for Monkeypox virus data, whether you wish to submit reads or genomes. If you wish to submit Monkeypox virus data via this route, please send an email to virus-dataflow@ebi.ac.uk or follow the steps in the tool documentation.

Private Data Sharing

For groups or consortia looking to share data in a more restricted manner please contact virus-dataflow@ebi.ac.uk, we can offer a private data hub for Monkeypox data for which you and your nominated collaborators would have controlled access

Registering Studies

Data submissions to the ENA require that you register a study to contextualise and group your data. Details of how to do this can be found in our Study Registration Guide. Please ensure you describe your study adequately, as well as provide an informative title.

Data Claiming and Ownership

Alldata remain within the submitter's ownership. In addition, ENA Monkeypox virus studies can now be "claimed" using your ORCID ID and/or assigned a DOI. Please see here for ORCID claiming.

We can now offer DOI issuing for Monkeypox projects. Digital Object Identifiers (DOIs) are persistent identifiers that can be assigned to any type of entity. From the DOI handbook:

> A DOI name is an identifier (not a location) of an entity on digital networks. It provides a system for persistent and actionable identification and interoperable exchange of managed information on digital networks. A DOI name can be assigned to any entity — physical, digital or abstract — primarily for sharing with an interested user community or managing as intellectual property. The DOI system is designed for interoperability; that is to use, or work with, existing identifier and metadata schemes. DOI names may also be expressed as URLs (URIs).

DOI issuing for ENA records is performed by creating a BioStudies record containing all relevant ENA projects (https://www.ebi.ac.uk/biostudies/about). We will generate this BioStudies record on your behalf and it will hold pointers to the ENA project(s) of your choosing.

Registering Samples

Having registered a study, please proceed to register your samples. These are metadata objects that describe the source biological material of your experiments. Following this, the sequence data can be registered (as described in later sections).

Instructions for sample registration can be found in our Sample Registration Guide. As part of this process, you must select a sample checklist to describe metadata. If you require any support regarding sample metadata, please contact virus-dataflow@ebi.ac.uk.

Viral Samples

The most appropriate checklist for Monkeypox virus submissions is the "ENA virus pathogen reporting standard checklist" - ERC000033. This presents 9 mandatory, 15 recommended and 11 optional fields (along with any additional user-defined fields).

Please use the organism name "Monkeypox virus" and taxonomic ID 10244. It is recommended, as a minimum, that collection date and geographic location (e.g. country) are specified.

Submitting Reads

After registering your study and samples, you can submit your read files along with experimental (library-related) metadata. See our Read Submission Guide for detailed instructions on submitting reads.

We encourage submissions to include information on specific protocols used for the experiment. This should be provided in the library description. This can be, for example, the name and/or URL to a specific protocol. View our listing of the available full experimental metadata dictionaries.

Note: Submitted reads to ENA should not contain human identifiable reads. Please filter out human reads prior to submission - if required, here is a tool which can be used.

Submitting Assemblies

If submitting assemblies, you must have registered a study and a sample beforehand. It is also advised that the associated read data is also submitted. For instructions on assembly submission view our Assembly Submission Guide.

Assemblies can only be submitted using Webin-CLI program

Submitting Monkeypox assembled sequences with Webin-CLI

In case of the Webin-CLI program -context genome should be used. During the process, you must define metadata in the manifest file(s). Please specify 'clone or isolate' as the 'ASSEMBLY_TYPE'.

Each assembly/consensus sequence should also be submitted with a **chromosome list file** (see here), which should be gzipped and referenced in the assembly manifest file, with 'CHROMOSOME_LIST'.

For Monkeypox virus submissions, the chromosome list file should contain the following tab-separated columns (with no column header line):

- fasta header
- chromosome number (arbitrary value, set to 1)
- chromosome type (Monopartite for monkeypox virus)

e.g:

```
LR991698 1 Monopartite
```

Any assembly annotations, where provided, are captured according to INSDC Feature Table Definitions.

In alignment with INSDC partners, Monkeypox virus assemblies will **not** be assigned a GCA accession. However, sequence accessions will continue to be assigned, alongside ERZ analysis accessions which are the point of access for the submitted file(s). For more details on accessioning, view our Accessions Guide. To cite data, top-level project accessions (PRJ…) should be used as these are the most stable long-term accessions. View our guide to cite data for further details.

Submitting Targeted Sequences

If submitting targeted or annotated sequences, you must register a study as described above. See our Targeted Sequence Submission Guide for submission instructions. When submitting annotated sequences, you must select an appropriate Annotation Checklist. There are several virus-specific annotation checklists, with "Single Viral CDS" the most appropriate for complete or partial coding sequences from a viral gene. If your sequences do not fit the annotation checklists above please contact us at virus-dataflow@ebi.ac.uk.

Any annotations, where provided, are captured according to INSDC Feature Table Definitions.

If submitting single contig assemblies, or for any other support or issues around Monkeypox virus submissions please contact virus-dataflow@ebi.ac.uk.