European Nucleotide Archive

01/03/2023 ena-announce

Dear ENA users,

In order to align the taxonomic restrictions placed on submissions of assemblies and samples to the ENA with that of other INSDC nodes, we intend on deploying some changes to what taxon IDs are considered submittable.

Currently, for all submissions to the ENA we require a species level taxon id to be used, with a full binomial name, or a placeholder taxon ID.

Current system example:

Mycobacterium sp. is not a submittable taxon ID, because although it has a species level rank, it does not possess a full binomial name.

Mycobacterium tuberculosis is a submittable taxon id, because it has both a species level rank and is also a full binomial name.

Changes to current system:

We plan on making the following changes to our validation system.

- We will be expanding the taxon requirements such that you will be able to submit data without a full binomial name, provided the taxon id has a species level rank. For example, *Mycobacterium sp.* will be considered a submittable taxon id for ENA samples after this change is made.
- This change will not extend to "clone or isolate" type genome assembly submissions. We will still require that users submit a taxon id with a full binomial name for these submissions.
- In the case of environmental assembly types such as MAGs, binned metagenomes or environmental SAGs, the more permissive taxon rule will apply to these assembly types.
- Transcriptomes, metatranscriptomes, targeted sequences and template sequences will also have the more permissive taxon ID requirements applied.

Planned Change Date:

Clear guidance will be provided in the form of update documentation, as well as new parameters in the taxonomy API to help adapt to these changes. Please keep an eye out in ena-announce and via Twitter for guidance around these changes closer to release.

We plan on making this change on the 2nd May 2023. The newest major release of Webin-CLI (6.0.0 and later) will be compatible with this change. Please upgrade to the latest version to ensure compatibility with the new validation.

Kind Regards,

The ENA Team