



Guidelines for submitting MCZ Catalog Numbers to NCBI BioProject & BioSample Records

When submitting next-generation sequencing data as a NCBI BioProject, researchers must include links to MCZ specimens on the **BioProject page**. After this is complete, each specimen should additionally be listed on a **BioSample page**.

Record the link to the MCZ specimen on the BioProject page

- Submit links to MCZ Catalog numbers within the “**Related Resources**” section of the **BioProject**. The URL address should be entered in the following required format:
http://mczbase.mcz.harvard.edu/guid/MCZ:collection_code:specimen_id (e.g., <http://mczbase.mcz.harvard.edu/guid/MCZ:IZ:19673>). All MCZ Catalog numbers should be complete and include any prefixes and original formatting (e.g., MCZ:Herp:R-2244).
- If multiple MCZ specimens are included in a single BioProject, multiple links should be listed on the BioProject page.

Record the MCZ Catalog Number on the BioSample page

- MCZ Catalog numbers and other identifiers are recorded on the BioSample page, where a description of the biological source materials used in experimental assays is included. Each physically unique specimen included in the study should be registered as a single BioSample with a unique set of attributes, and multiple BioProjects can be linked to a single BioSample.
- Submit the MCZ Catalog number as a “**specimen voucher**” within the “**Attributes**” section of a **BioSample**. The specimen voucher describes the *primary* identifier of the physical specimen from which the sequence was obtained, and is intended for use in cases where the voucher/sample is still available (e.g., museums, herbaria, frozen tissue collections). The specimen voucher should be entered in the following required format:
institution_code:collection_code:specimen_id (e.g., MCZ:IZ:19673). All MCZ Catalog numbers should be complete and include any prefixes and original formatting (e.g., MCZ:Herp:R-2244).

Below is the list of relevant Collection Codes for the different MCZ Collections. Note that the first letter of each collection must be capitalized:

| | |
|----------------------|------|
| Cryogenic | Cryo |
| Entomology | Ent |
| Ichthyology | Ich |
| Herpetology | Herp |
| Invertebrate Zoology | IZ |
| Malacology | Mala |
| Mammalogy | Mamm |
| Ornithology | Orn |

NCBI records for material submitted using the proper MCZ Catalog number format as noted above will be automatically hot-linked to the proper MCZbase specimen record (and the MCZbase record will be hot-linked back).



Additional Data

- It is suggested that information regarding type of tissue used to obtain the sequence data (e.g., heart, muscle, liver, ground whole organism) be recorded in the “**tissue**” field on the **BioSample** page.
- It is suggested that information regarding type of sample used to obtain the sequence data (e.g., cell culture, tissue sample, whole organism) be recorded as a source modifier using the “**sample_type**” field on the **BioSample** page.
- Refer to the relevant NCBI submission tool documentation for further instructions and a glossary of all available modifiers.