**Dlugosch Lab RADseq submissions to NCBI**

**Log in to the NCBI submission portal to submit data:**

* <https://submit.ncbi.nlm.nih.gov/>
* Login via the NCBI PDA
  + user: DlugoschLab
  + password: MiraTheCat119

**1) All projects need to be entered as BioProjects**

Click on BioProjects to see what is already entered or add one (check with KMD) <https://submit.ncbi.nlm.nih.gov/subs/bioproject>

* PRJNA275994 : Population genomics of Eustenopus villosus
* PRJNA275993 : Population genomics of Silene acaulis
* PRJNA275992 : Population genomics of Centaurea nicaeensis
* PRJNA275988 : Population genomics of Centaurea melitensis
* PRJNA275986 : Population genomics of Centaurea solstitialis
* PRJNA275979 : Sources of introduction of Matthiola parviflora to North America

**2) All samples need to be entered as BioSamples**

Check the lab google doc “DLab NCBI submission records” to find existing and record new entries.

Where to submit: https://submit.ncbi.nlm.nih.gov/subs/biosample/

* Click on New submission
* (verify that KMD is the official submitter)

General information

* Yes, the biosamples can be released immediately.
* Here you can choose a batch upload for large sets of individuals.
  + Please don't mix species that belong to different Projects.

Info for a single sample

* “Plant sample” or “Invertebrate” or “Model organism or animal sample” as appropriate
* Sample name is a code for the species (see below) AND

the sample code that identifies individuals, e.g. “MAPA TH-32”

PLEASE make all letters capitals for uniformity.

**CODE: SPECIES**

MAPA: Matthiola parviflora (plant)

CESO: Centaurea solstitialis (plant: yellow starthistle)

CEME: Centaurea melitensis (plant: Malta starthistle)

CENI: Centaurea nicaeensis (plant)

SIAC: Silene acaulis (plant: moss campion)

EUVI: Eustenopus villosus (insect: hairy weevil)

PADO: Passer domesticus (bird: house sparrow)

* Enter the scientific name of the species - it should come up online for single entries.
* Enter the continent of collection (e.g. “North America”) for the “ecotype”
* developmental stage:

enter “seedling” or “adult” as appropriate for the plants

enter “adult” for the weevils

enter “adult” for the birds (or other stage as appropriate)

* For geographical location, use the name of the country only, as per the list they provide
* For tissue: plants are all “leaf”, weevils are all “whole individual” and birds are “blood”
* Something in the 'attributes' field needs to be unique to every sample, so for “isolation source” enter the sample name
* Lat/Long
* BioProject (exact title)
* Sample Title (under Comments): please use the sample name from the Attributes tab

**3) To submit sequence data, log in to the Sequence Read Archive**

* You must have BioProjects and BioSamples created already.
* You will need your final de-multiplexed files ready, and an MD5 checksum calculated for each.
* Note that you should submit data from each run separately, even if the same individual had more than one run of data.

Where to submit: <http://www.ncbi.nlm.nih.gov/Traces/sra_sub/sub.cgi>

* Under Submit> Submissions you should see a list of species.
* Data submissions have a release date that you set. There should be one submission for each release of data that you anticipate (e.g. new data to be published in a paper). You can add to a submission at any time, so only make a new one if you have a new set of data that you expect to release separately.
* Note that only one BioProject can be associated with a data submission.

Brittany Please expand here how to go about sequence submissions now