Jan 19, 2023

DNA Barcoding Standard Operating Protocol, Plants and Lichens at RBGE, Sample Data

Forked from Collections Standard Operating Protocol, Plant group: Bryophytes

In 2 collections

DOI

dx.doi.org/10.17504/protocols.io.e6nvwjzr9lmk/v1

Laura L Forrest[1], David Bell[1], Michelle Hart[1]

[1]Royal Botanic Garden Edinburgh

Darwin Tree of Life

Laura L Forrest

Royal Botanic Garden Edinburgh

DOI: https://dx.doi.org/10.17504/protocols.io.e6nvwjzr9lmk/v1

Protocol Citation: Laura L Forrest, David Bell, Michelle Hart 2023. DNA Barcoding Standard Operating Protocol, Plants and Lichens at RBGE, Sample Data. **protocols.io** **https://dx.doi.org/10.17504/protocols.io.e6nvwjzr9lmk/v1**

Protocol status: Working
We use this protocol and it's working

Created: January 17, 2023

Last Modified: January 19, 2023

Protocol Integer ID: 75362

Keywords: DToL, RBGE, Plant Barcoding, Darwin Tree of Life, BOLD, DNA, DNA bank, EDNA, lichen sample data for the darwin tree, tissue samples for dna barcoding, dna barcoding, dna barcoding standard operating protocol, part of the species identification process, lichen sample data, darwin tree of life project, species identification process, various land plant taxa, various land plant taxa within the scope, darwin tree, wellcome sanger institute for whole genome sequencing, genome sequencing, part of the collection dtol taxon, dtol taxon, lichens at rbge, whole genome sequencing, land plant, genome, dna, genome sequence, collection dtol taxon, royal botanic garden, plant, whole genome, barcoding sop, lichen, place at the royal botanic garden, sample data, lycopodiophyta, embryophyta, processing of plant, royal botanic garden edinburgh, species, cycadophyta, tissue sample, bryophyta, anthocerotophyta, fungi

Abstract

This is part of the collection **DToL Taxon-specific Standard Operating Procedures for the Plant Working Group (protocols.io)**. The SOP collection contains guidance on how to process the various land plant taxa within the scope of the Darwin Tree of Life project. The guidance specifically refers to the metadata associated with tissue samples for DNA barcoding (which takes place at the Royal Botanic Garden (RBGE)). Every specimen is submitted for DNA barcoding first before potentially being sent to the Wellcome Sanger institute for whole genome sequencing.

This DNA barcoding SOP outlines the processing of plant and lichen sample data for the Darwin Tree of Life project (DToL) at the Royal Botanic Garden Edinburgh (RBGE).

DNA barcoding is used as part of the species identification process AND sample tracking (to check that the genome sequence corresponds to the material that was sent and that there have been no sample mix-ups).

Definition: Land plants (Embryophyta) and lichens
Including: Bryophyta, Marchantiophyta, Anthocerotophyta, Lycopodiophyta, Polypodiophyta, Pinophyta, Cycadophyta, Ginkgophyta, Gnetophyta, Magnoliophyta, lichenized fungi

Excluding: all non-lichenized fungi

Guidelines

Including: Bryophyta, Marchantiophyta, Anthocerotophyta, Lycopodiophyta, Polypodiophyta, Pinophyta, Cycadophyta, Ginkgophyta, Gnetophyta, Magnoliophyta, lichenized fungi

Sample Tissue
Plant material (typically leaf or thallus material, but may be other plant tissue such as flowers) or lichen thallus tissue is collected and submitted to the Barcoding Centre following Standard Operation Protocols in the collection **DToL Taxon-specific Standard Operating Procedures for the Plant Working Group (protocols.io)**

Note

Previous versions of the Plant Working Group SOPs can be found here:
RBGE DToL Sample collection Standard Operating Procedure Vascular Plants
RBGE DToL Sample collection Standard Operating Procedure Bryophytes
SOP RBGE Plant DNA Barcoding sample submission

Required sample metadata
The collector or Genome Acquisition Lab emails a completed **DToL SAMPLE MANIFEST** for the collections that are to be barcoded to the Barcoding centre. This is then adapted to include one line per collection number, with all the Genome Acquisition Lab numbers (e.g. EDTOL or KDTOL numbers) that apply to that collection concatenated into a single cell (e.g. EDTOL00053,EDTOL00054,EDTOL00055).

The data in this manifest is then used to produce:

1. EDNA numbers: These are in-house accession numbers that are given to RBGE DNA extractions.

Note

A DNA banking form EDNA_submission_sheet_211209.xlsx is completed at RBGE for each batch of extractions, and uploaded to the RBGE in-house DNA database, **EDNA**. This spreadsheet comprises a mixture of compulsory and optional cells for each extraction, many of which are **defined**.

Most of the EDNA information can be obtained from the barcode-adapted sample manifest. The changes to the way that this form is usually filled in that are specific for DToL samples are:

1. The 'Sample Note' column (M) is filled in with all the DToL numbers that apply to each DNA extraction, separated by commas (no spaces). As a limited number of characters are allowed in each field, DToL numbers should be condensed by hyphenating ranges (e.g. EDTOL00053-00055).
2. The DNA extraction numbers follow the DToL letter series described below.
3. "Darwin Tree of Life sample" is added in the note column (T).
4. The samples on the spreadsheet are uploaded to EDNA under the Darwin Tree of Life project.
5. The EDNA upload form is saved in a DToL barcode folder on the RBGE DNA server.

Once EDNA numbers have been assigned, a list of them can be downloaded from the EDNA database. The list will contain the EDNA numbers and collector numbers for each collection, associated with the DToL numbers for each sample.

The EDNA numbers are used to update:

- The autofill lists we use to generate sequence trace file names (DToL Sequencing Request Form.xlsx ; Molecular Lab Registration Forms folder on RBGE DNA server).
- Our Barcode lab book spreadsheets (Data folder / DToL folder / barcode folder / lab books on the RBGE DNA server).
- Our DNA banking forms (Raw folder / EDNA folder on the RBGE DNA server: the plate list, the plate positions list, and the EDbank numbers list).

2. BOLD specimens (with Sample IDs and Process IDs): this is the data about each DToL DNA barcoding collection that is stored in the open access Barcode of Life database. Each specimen contains sample metadata, and can include image files (e.g. habitat photos, close-ups of fronds, flowers, leaves etc.) and a collection map.

Note

1. We use our barcode-adapted DToL sample manifest and the EDNA.csv download to populate a BOLD **Version 3.1 Spreadsheet** (download from BOLD each time as the versions can change), which is uploaded to BOLD to create **BOLD specimens**. We use our EDNA number as the **BOLD sample ID number**, while the complete DToL specimen ID number/string of numbers from the barcode-adapted manifest is added into the Notes field. We also include the sample's major lineage in the Extra Info column, allowing us to sort the project and datasets into broad taxonomic groups (angiosperms, conifers, ferns, lycophytes, mosses, hornworts, liverworts, lichens).
2. Once the samples have been uploaded to our BOLD EDToL project, we add them to either our RBGE or RBGK datasets **(DS-DTOLRBGE; DS-DTOLRBGK)**, which team members from the Genome Acquisition Labs have access to, as well as to relevant taxon datasets (**DS-FERN**, **DS-LICHEN**, **DS-LIVERWRT**, **DS-LYCOPHYT**, **DS-MOSS**, **DS-SEEDPLNT**) for ease of analysis.
3. After creating them, the samples are downloaded from BOLD in order to get the **BOLD ProcessID number** that will be needed to upload raw trace files to the BOLD EDTOL project. (BOLD downloads can be saved to the DToL barcode folder on the RBGE DNA server.)
4. After BOLD specimens have been created, photographs of the plants can be uploaded to BOLD in a zipped folder including the Specimen Image form **Version 3.0** (download an empty form from the BOLD site) and the images. This has to be completed by the submitting Genome Acquisition Lab (or by the Barcode Hub if the Genome Acquisition Lab submits photos to them along with a Specimen Image form). In order to upload images to BOLD, the submitter needs edit rights to the BOLD EDTOL project, not to the individual Genome Acquisition Lab's datasets.
5. Images should be named in a standard manner, as described in the Collection SOPs, preferably by the collector number as it appears in the manifest, a one-word description, and a number if there are several images in a series, e.g. MR204_flower_1.jpg, MR204_flower_2.jpg; MR204_leaf.jpg, MR204_habitat.jpg.
6. Copies of the submitted zipped image folder should be saved, as if there are any problems in the future, e.g. a set of images needing deleted, BOLD will remove the entire submitted batch; it will be important to know exactly what was in the batch in order to resubmit it.

For DToL barocding at RBGE, all the relevant collector, EDNA and BOLD numbers for each sample are pasted into a shared google sheet, **RBGE DToL Barcoding**. This sheet must NOT be sorted, as it contains several linked sheets. (To sort, copy and paste the relevant columns into a new spreadsheet).

	A	B	C	D
		Date	***Changes***	***Contributors***
	1.0	August 2020	First draft	Laura L Forrest, Michelle L Hart

	A	B	C	D
	1.1	January 2021	Revisions	Laura L Forrest

Previous Version History, RBGE DToL DNA Barcoding Standard Operating Procedure

Working SOP, checked by experts

Troubleshooting