NHM DToL Partner Sample Submission for DNA Barcoding Standard Operating Procedure

Version: 2.2

Published date: November 2020

Authors: Lyndall Pereira and Ian Barnes

This Standard Operating Procedure (SOP) contains guidance for other DToL partners on their submission of specimens and dissected tissue destined for DNA barcoding at the Natural History Museum, London. The guidance provided here covers the submission process once the tissue for barcoding has already been dissected from a specimen destined for whole genome sequencing (WGS) or from a "barcode exemplar" of another specimen.

Purpose: The aim of this document is to outline the procedures for submission of animal and fungal tissue samples to the NHM for DNA barcoding by other DToL partners. When a specimen has been dissected, individual pieces of tissue are placed into FluidX tubes which are sent to Sanger for whole genome sequencing (WGS) (see Sanger DToL Sample Submission SOP link below for details). One further piece of tissue should be reserved and sent to the NHM for DNA barcoding following the guidance below. In some cases, a specimen will act as a barcode exemplar for another specimen because the entire organisms must be used for WGS and it is not possible to take a sample for DNA barcoding (please refer to the RELATIONSHIP column in the **Recording Sample Metadata for the Darwin Tree of Life Project** SOP).

Future plans for this SOP: This SOP will be reviewed (on a quarterly/*ad hoc* basis) by NHM team members to incorporate feedback from the community.

Raising issues: We are still developing best practice and elements of this SOP are subject to change. We anticipate and welcome questions, comments and suggestions and are happy to share lessons learned. If you would like to share any advice, comments, techniques or lessons learned, please do not hesitate to email L.Pereira@nhm.ac.uk

Please see the links below to the taxon-specific sample processing SOPs for guidance on dissection, and for volumes of tissue needed for DNA barcoding. [links to be added]

Links to SOPs relating to this document:

- Sanger Sample Submission SOP v2.2
- Sanger Sample Metadata SOP v2.2

Document History

Version	***Date***	***Major changes***	***Contributors***
1.0	2020-05-11	First version	Lyndall Pereira, Ian Barnes, Gavin Broad, Ben Price
2.0	2020-05-18	DNA barcoding of whole specimens (including manifest for partners) added	Lyndall Pereira, Charlotte Barclay, Heather Allen, Darren Chooneea, Ben Price
2.1	2020-07-07	Integration with main manifest. More detailed instruction	Lyndall Pereira, Darren Chooneea, Gavin Broad, Ian Barnes, Mara Lawniczak
2.2	2020-11-10	Instructions on what to do with duplicate/whole specimens as vouchers for curation have been removed from this SOP.	Lyndall Pereira, Clementine Geeves, Darren Chooneea, Silvia Salatino

Specimens and Tissue Samples for DNA Barcoding and Vouchering

Overview

All specimens and tissue samples referred to in this SOP are to be sent to the Natural History Museum, London.

All specimens and tissue samples destined for DNA barcoding require entry into the main Sample Manifest [link]

Guidance on preparing taxon-specific samples relating to this SOP can be found in the Taxon-specific SOPS [links].

Prior to Dissection

Please ensure that each specimen has been photographed prior to dissection (please follow the guidance under "*Photographs of Specimen or Sample*" in the DToL Metadata SOP v2.2). In many cases, this photograph will be the only morphological voucher as the rest of the specimen will be processed for WGS and DNA barcoding.

Please refer to the taxon-specific SOPs (found in the shared folder of each Working Group) for guidance on the details of the photography (e.g. resolution, viewpoint) and the tissue to take for each taxon group.

Sample Collection Vessels

Covid-19 Note: *Number-barcoded tissue plates can be provided by the NHM to each DToL partner that will be sending tissue for DNA barcoding. Please contact Lyndall at L.Pereira@nhm.ac.uk for details, once staff are permitted back on site, we will be able to send out barcoded tissue plates.*

For DNA barcoding, the tissue **must** be placed into a semi-skirted (half-skirted) 96-well plate filled with 70% ethanol and ideally kept in a freezer (-20 ℃) or as cold as possible (keeping specimens/tissue in ethanol and in the fridge is fine for short time periods up to a month). The plate **must** be filled in columns (first tissue in A1, second in B1, third in C1 etc). Each tissue plate used should have a **unique label**. This can be a unique barcode which should be read with a scanner (to prevent errors), or any sticky label with a unique code written in ethanol proof ink. Each plate should be sealed before submission, either with caps or with adhesive PCR plate seals.

Fig 1. Example of a number-barcoded 96-well tissue plates filled with tissue samples from the latest DNA bioblitz (modified photo by NHM Molecular Labs)

In the absence of tissue plates, i.e. before you have received a tissue plate from NHM; or you will not be sending as many as 96 samples; or you wish to send a whole specimen/large tissue as a voucher, then barcoded tubes or any tube with a unique identifier can be used. NHM can provide unique identifier labels if needed, please contact L.Pereira@nhm.ac.uk.

Sample Data Template

Until COPO is ready to receive manifests we are operating the system set out below. When COPO is ready to receive manifests, this SOP will be updated.

[for bioinformatics part]

The input file for our analysis pipeline is a comma-separated CSV file containing the following headers in this specific order:

"Plate_ID,Specimen_ID,Well_ID,Order,Family,Genus,Species".

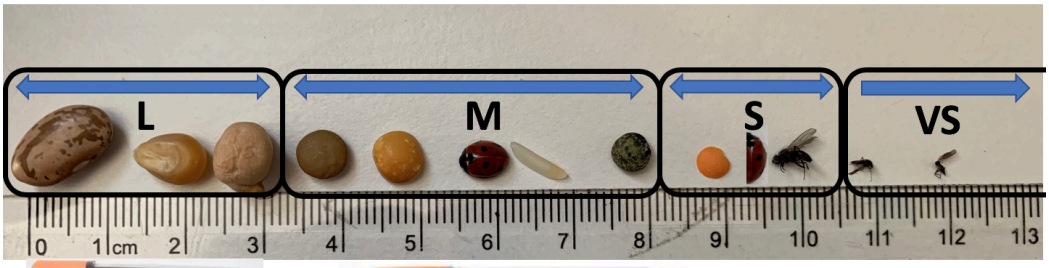
Here's an example of how a CSV file containing just two samples would look like:

Plate_ID,Specimen_ID,Well_ID,Order,Family,Genus,Species
MOZZ00000522A,Ox0490,A1,Lepidoptera,Zygaenidae,Zygaena,filipendulae
MOZZ00000522A,Ox0491,A2,Coleoptera,Lucanidae,Dorcus,Parallelipipedus

Importantly, the entries in this file should not contain any symbols (i.e. white spaces, "-", "+", "?", etc.).

Quantity of Tissue Required

As a general rule, a single lentil-sized (ca. 5mm) piece of tissue is the minimum requirement. Different tissue types are acceptable. For small specimens, six legs from the same individual should be used. In cases where it is impossible to collect the minimum required amount of tissue, then an exemplar specimen should be submitted Please see the taxon-specific SOPs for details. If the specimen being processed is large enough that there is tissue leftover from dissection for WGS and DNA barcoding, then this tissue should be sent to the NHM for deposition in the collections, either as a morphological or tissue voucher. The NHM collections, including Molecular Collections Facility (MCF), have the capacity for long term storage of vouchers.



0.5 mL FluidX tube
29.8 mm with cap

1.9 mL FluidX tube
43.9 mm with cap

Guidance for "Size of tissue in tube"
L = popcorn kernel or dried chickpea sized and larger
M = green, yellow lentil sized, whole ladybird size
S = red lentil, half a ladybird size
VS = smaller than half a red lentil
SC = single cell

Sample Data Linkage

This tissue plate barcode (or unique identifier) and the plate well number (A1,B1, C1, etc.) should be recorded on the Sanger Manifest in order to keep the data linked with the original specimen collected for WGS.

Please refer to the **Recording Sample Metadata for Darwin Tree of Life SOP** for column instructions for **Metadata Entry** into the Sanger Manifest. Columns that are specific to DNA barcoding samples are repeated here:

R. RELATIONSHIP: (ENA_submission) This is a free text field to permit declaration of any known parental, child, or sibling relationship between the specimen and any other specimens that are submitted for the DToL project, OR to declare if the specimen is a "barcode exemplar" for another specimen.

- If there are known genetic relationships between submitted specimens, please concisely state the relationship: "Full sibling to SPECIMEN_ID1", "Mother to SPECIMEN_ID2", "Maternal half sibling to SPECIMEN_ID1, SPECIMEN_ID2, and SPECIMEN_ID3", or "Trio child of SPECIMEN_ID1

and SPECIMEN_ID2". If knowledge of the relationships is not confident but suspected, do not add anything here and instead add this information to the "OTHER_INFORMATION" field (e.g., "suspected full or half sibling to SPECIMEN_ID2").

- If the specimen is acting as a barcoding exemplar for another specimen because the entire organism must be used for reference genome sequencing and it is not possible to take a sample for DNA barcoding (e.g., midges from the same swarm where one is submitted for sequencing and 5 are submitted individually for DNA barcoding), then add "barcode exemplar for SPECIMEN_IDx" and insert the SPECIMEN_ID for the specimen that is going for reference genome sequencing, potentially without its own DNA barcoding.

- If there is no relationship to note, this field can be left blank.

TT. **TISSUE_REMOVED_FOR_BARCODING**: State "**Y**" or "**N**". See the appropriate Molecular Barcoding SOPs for detailed instructions, noting that barcoding requires materials in specific tube or plate types so the SOP must be referred to. If you are collecting across different taxonomic groups, ensure you know which GALs will receive material so that you allocate your samples into different plates depending on their destination (as of October 2020, marine fungi and seaweeds go to MBA, plants go to RBGE, and everything else goes to NHM).

UU. **PLATE_ID_FOR_BARCODING**: This is the barcode number on the side of the tissue plate. Barcoding sites will provide pre-labelled plates and tubes. If you are submitting plant tissue, these will not be submitted in plates, so this is not necessary and you can put NOT_APPLICABLE.

VV. **TUBE_OR_WELL_ID_FOR_BARCODING**: This is either the well number on a plate (there are 96 wells per tissue plate) OR the barcode/unique identifier on the tube containing the tissue sample.

WW. **TISSUE_FOR_BARCODING**: Please state what part of the organism was dissected for DNA barcoding (e.g. leg, soft-body tissue etc.). Muscle tissue is ideal for barcoding. This list is a repeat of the attributes available for "ORGANISM_PART" with one addition of "DNA_EXTRACT"

XX. **BARCODE_PLATE_PRESERVATIVE**: Guidance is found in the barcoding SOPs. Typically, animal samples will be submerged in 70% ethanol, plant tissue will be preserved in silica gel, and fungal tissue will be frozen or lyophilized. Record the volume, concentration, and type of preservative/method of preservation used here.

YY. **PURPOSE_OF_SPECIMEN:**
The majority of specimens will be for "REFERENCE_GENOME". All samples listed for REFERENCE_GENOME sequencing are assumed to also need DNA BARCODING and RNA-SEQUENCING, and the term "REFERENCE GENOME" encompasses all three things (reference genome, barcoding, rna-seq) wherever samples allow. Please use REFERENCE GENOME for all specimens / samples of a particular species unless they should be destined for an alternative use only.

- If a particular tissue is needed solely for RNAseq use “RNASEQUENCING”
- If the specimen is intended for population genetics or resequencing please use “SHORT_READ_SEQUENCING.
- The drop-down option for DNA_BARCODING_ONLY is reserved for those specimens submitted solely for DNA barcoding (e.g., when the sample is too small to provide material for both reference genome and barcoding and genome paratype / other specimens must be used as proxies, or when the specimen was identified to species level but died before being preserved, or is otherwise unsuitable for HMW DNA, but the material is valuable for barcoding).

Shipping samples to NHM

Please follow this checklist before sending samples to the NHM for DNA barcoding:

- ❏ **Complete the Sample Data Template** and send or share it with the email DarwinTreeofLife@nhm.ac.uk before sending the samples - this **does not** need to be the Sanger-validated copy of the manifest, and can be the partner working-copy before submitted to Sanger.
- ❏ **Upload** the specimen images to the shared Google Drive folder with image name as the SPECIMEN_ID from the manifest (see instructions for photographs in the **Recording Sample Metadata for the Darwin Tree of Life Project** SOP)
- ❏ Ensure the plate is properly **sealed** and packaged with bubble-wrap and/or tissue as padding and clearly address the package as follows:

***Covid-19 Note*: Please do not send any samples to the NHM during the covid-19 lockdown and until staff are permitted back on site. Please keep all samples as cold as possible (fridge / freezer) until partners are notified that it is safe to send samples.**

Shipping address:

Samples for Darwin Tree of Life (DToL),
℅ Darren Chooneea/Clementine Geeves
Core Research Laboratories
The Natural History Museum
Cromwell Road
London
SW7 5BD

DNA barcode results

Until COPO is ready to receive manifests we are operating the system set out below. When COPO is ready to receive manifests, this SOP will be updated.