Biodiversity survey and specimen data

1. Some important points about surveys – Carol Poole – BioGaps project

2. Templates for data capture - collection specimens and survey data

Introduction

- Most foundational biodiversity information projects involve some aspect of sampling or collecting specimens or data capture from specimens in collections.

- Data that have come in from funded projects have been problematic and there have been many queries about the template– lack of understanding about what the template means / what data are required.

- Proposals for sampling / surveying biodiversity have lacked important details about methods and approach.

- This session provides an example of a survey project that does consider critical elements.

- And goes through template and the metadata requirements with an opportunity for comments and questions.

- Input from participants welcome – the template can be modified, and we will all learn from others' experiences / suggestions

BioGaps project

- Commended by the panel and by some top scientists for the survey design

Some important considerations for surveys

- Selection of sites – must be based on scientific principles – think about the questions that need to be answered, statistical analyses.
- Sampling must be quantified to allow spatial and temporal comparisons, and allows an indication of absence.
- Feasible – think about time required to sort mixed samples, travelling between sites, time required to do the sampling.
- Identification of material by experts must be considered. Material identified only to family level has limited use.
- Data to be recorded must be carefully planned beforehand and implemented in the field.
- GPS co-ordinates for all samples taken are critical – even if the whole specimen is not collected.
- Specimens must be deposited in a recognised collection institution.
- Permits, land owner permission, impacts on populations, ethics – all must be considered.

Comments, questions or suggestions for surveys??

FBIP data submission requirements

- For FBIP-funded projects there is a data template that should be used for submitting all survey data, for collection specimen data capture, and for samples that are sequenced / DNA barcoded.

- We don't prescribe what software must be used – the template is an Excel spreadsheet and we would prefer that data are submitted in this format (for now at least).

Why are we providing a template?

- Ensure that we get critical components of the data (what it is, where and when it was collected, where the specimen is, what were the circumstances of its collection – for example)
- Ensure that data sets can be merged without reworking all the data – nationally and globally
- Overall – improving the quality and quantity of data because surveys, sampling, data capture are all expensive activities so we must maximise the outputs from investments.

What is the template?

- Based on the "Simple Darwin Core" which is the GBIF and global standard for biodiversity data – simple because it's a flat structure, rather than a relational structure.

- The Darwin Core is primarily based on taxa, their occurrence in nature as documented by observations, specimens, and samples.

- The template provided has been somewhat simplified and condensed – I have tried to make it accessible and useable by people (like me) who are not data / informatics specialists.

- If anyone wants to use the full set of fields or is more competent – they are welcome to add fields from the list provided for the Darwin Core, or have extensions.

- The full list of fields with definitions is available at http://rs.tdwg.org/dwc/terms/index.htm

The template fields

- Some of them are recommended as requiring "controlled vocabulary" – pick from a standard set of terms; others – free text.

- The options for some aspects of aquatic samples are a bit different to those for terrestrial samples (include depth rather than elevation, waterbody)

Questions, comments, suggestions for templates?

Metadata

- Metadata is the information about your data. It describes who collected it, why it was collected, how it can be used and the methods used to collect it, amongst other things.

- Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is often called data about data or information about information.

FBIP data sets and metadata

- Previously not required – which has meant that I am going back to compile the metadata – after seeing the importance of this for managing and tracking data sets.

- Future data sets – will require metadata.

Questions, comments, suggestions for metadata?

General discussion