

Smithsonian FIMS and Geneious LIMS Training

Week 1 - July 12, 2016

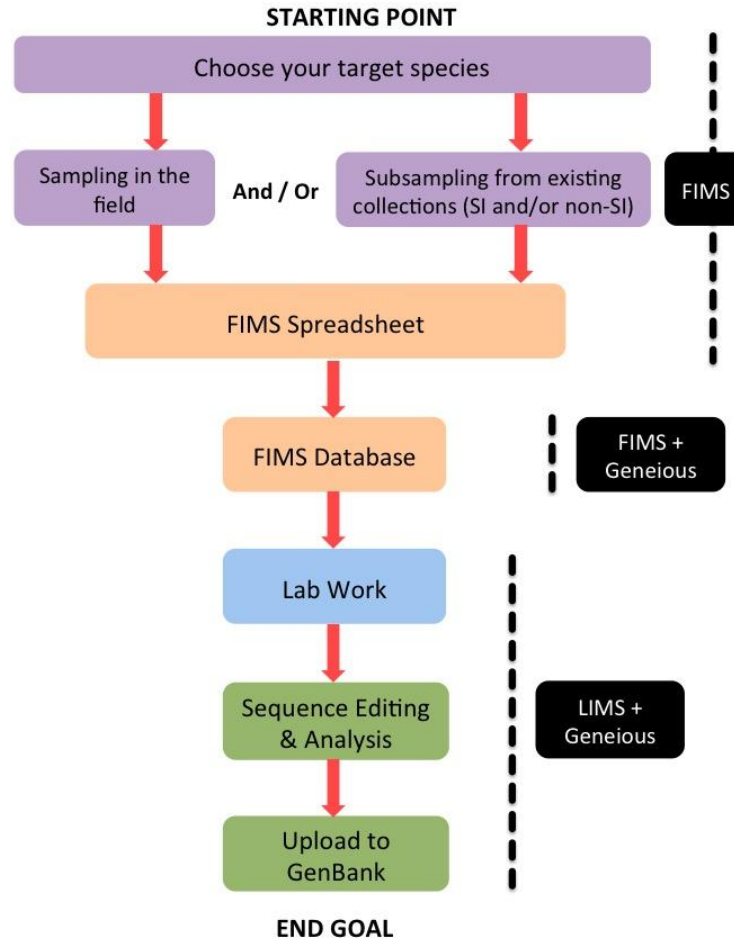
Niamh Redmond (RedmondN@si.edu)

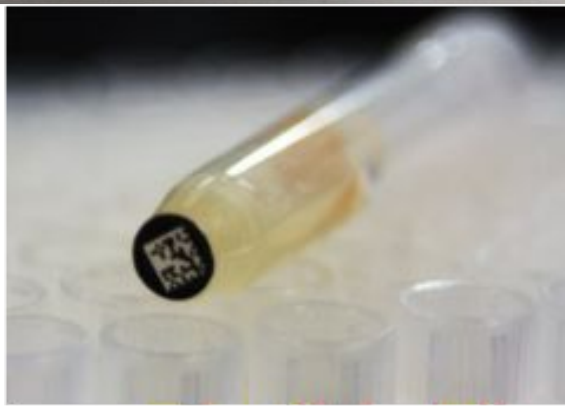
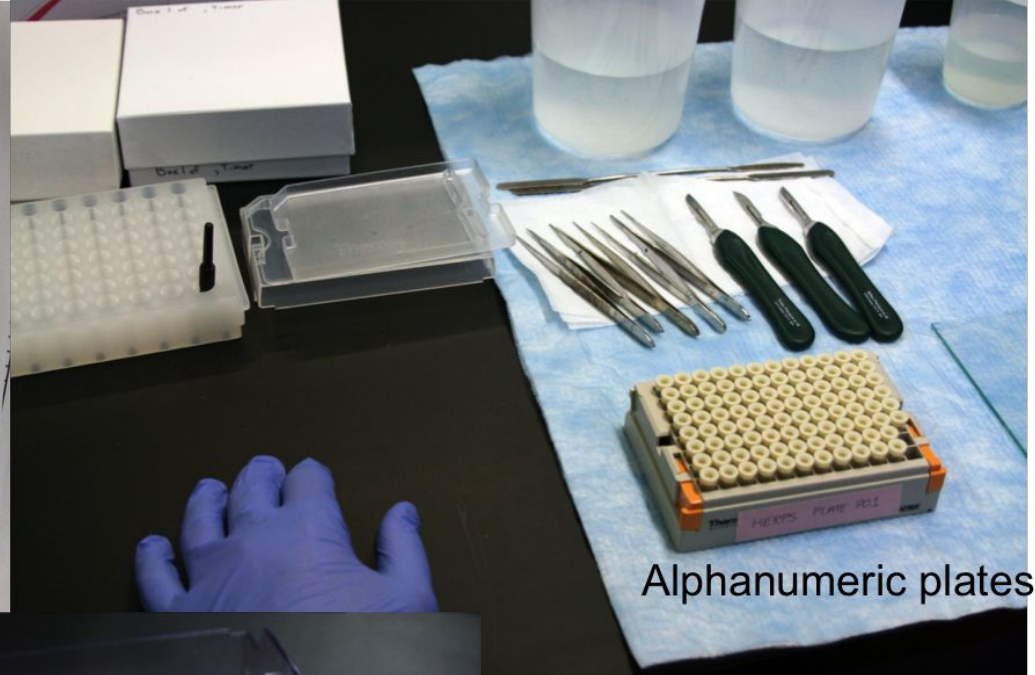
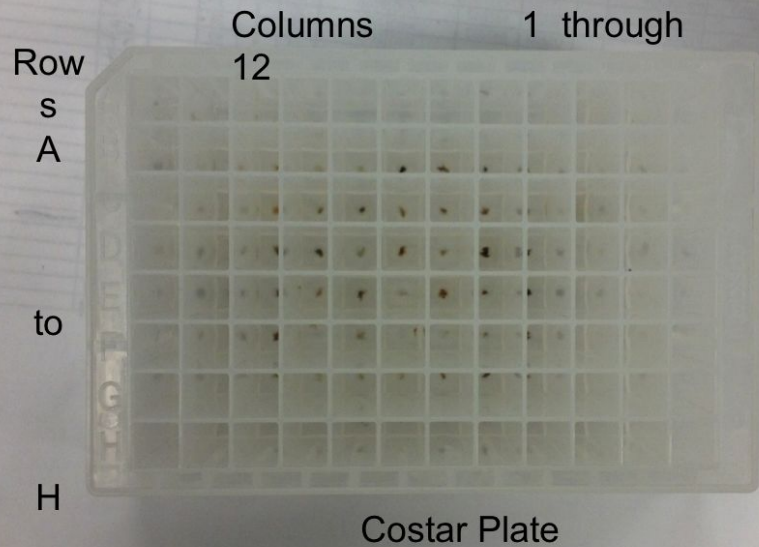
Mike Trizna (TriznaM@si.edu)

Session 1 Outline

- Pipeline Introduction
- Barcode data standard
- FIMS intro and spreadsheet validation
- Geneious and Biocode plugin Intro

Overview of SI Barcoding Pipeline

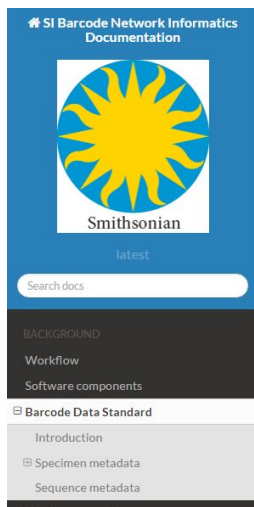




DNA Barcode Data Standard

Follow along at the SI Barcode Network documentation page on the BARCODE Data Standard:

http://sibarcodenetwork.readthedocs.io/en/latest/barcode_data_standard.html



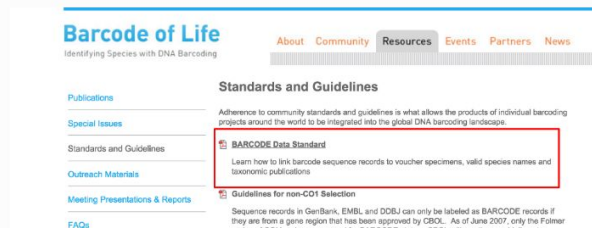
Docs » Barcode Data Standard

[Edit on GitHub](#)

Barcode Data Standard

Introduction

The Barcode Data Standard was established by the Consortium of the Barcode of Life soon after the first scientific paper by Dr. Paul Hebert was published that proposed the method of DNA Barcoding. You can find a link to the official data standard document on the CBOL website at: http://www.barcodeoflife.org/sites/default/files/DWG_data_standards-Final.pdf.



Exercise 1: Assessing GenBank Records

We want you to look at some real GenBank records. None of these records have earned the BARCODE keyword, but we want you to tell us what parts of the DNA Barcode Data Standard each record is missing.

Why did the following GenBank records not receive the BARCODE keyword?

1. <http://www.ncbi.nlm.nih.gov/nuccore/JF786290>
2. <http://www.ncbi.nlm.nih.gov/nuccore/DQ107705>
3. <http://www.ncbi.nlm.nih.gov/nuccore/HQ122000>
4. <http://www.ncbi.nlm.nih.gov/nuccore/EU679269>

FIMS intro and spreadsheet validation

- Differentiate all FIMS - NMNH, SI Barcode Network, etc. etc.
- Walk through spreadsheet (flows from minimum data above)
- Validation rules
- Demonstrate validation for SI Barcode Network project (BiSciCol), but no upload

Exercise 2: Completing & validating FIMS Spreadsheet

Download the FIMS spreadsheet: <https://github.com/MikeTrizna/SIBarcodeNetwork>

Click on "SI Barcoding Specimen Spreadsheet.xlsx", and then click "Raw"

Select any three of the following EMu records and fill the associated specimen data into the FIMS spreadsheet. (Search: <http://collections.nmnh.si.edu/search/>)

| | | |
|---------------|----------------|-------------|
| Botany:149519 | Mammals:123424 | IZ:1264996 |
| Herps:268949 | Fish:268949 | Birds:42676 |

Make up values for other columns e.g. extractionPlateID

Once complete, go to www.biscicol.org and validate your spreadsheet.

Do you receive any warnings or errors? If so, attempt to fix them until you successfully validate the FIMS spreadsheet.

Geneious

Geneious is a cross-platform (works equally well on Windows, Mac, and Linux) program that provides a simple GUI (Graphical User Interface) to many classic molecular biology tools in a layout similar to an e-mail client.

- [Assembly and Reference Mapping](#)
- [Sequence Analysis, Annotation and Prediction](#)
- [Molecular Cloning](#)
- [Microsatellite Analysis](#)

Find more information about Geneious at their website: <http://www.geneious.com/>

Menu bar

Toolbar

Local Directory

Document Table

Services

Document Viewer

Sources Panel

Help and Tutorial Panel

Options Panel

Geneious File Edit View Tools Sequence Annotate & Predict Help

9.1.5

Back Forward BLAST Agents Align/Assemble Tree Primers (restricted) Cloning (restricted) Biocode New Reaction (restricted) Back Up Search

Sources

- Local (0)
 - Sample Documents (0)
 - Alignments (8)
 - Plasmapper Features (314)
 - Plasmids from NEB (27)
 - Primers (12)
 - Protein Documents (6)
 - Tree Documents (4)
 - Deleted Items (0)
- Shared Databases
- Operations
- Biocode
 - Reporting
- NCBI
 - Gene
 - Genome
 - Nucleotide
 - PopSet
 - Protein
 - PubMed
 - SNP
 - Structure
 - Taxonomy
 - UniProt

1 of 8 selected

| Name | Description | Sequence Length | # Sequences | %GC | % Pairwise Id... | % Ident... |
|---------------------|---|-----------------|-------------|-----|------------------|------------|
| COXII CDS | Multiple alignment of 50 Cytochrome C O... | 705 | 50 | - | 72.6% | 35.6% |
| HIV env | 20 HIV envelope genes | - | 20 | - | - | - |
| Ig variable region | 5 immunoglobulin heavy chain variable r... | - | 5 | - | - | - |
| LysR family | 7 LysR protein sequences | - | 7 | - | - | - |
| Pairwise protein | Pairwise alignment of peptidase from kiw...221 | - | 2 | - | 47.1% | - |
| People | Multiple alignment of 5 sequences from f...426 | - | 5 | - | 85.3% | 72.1% |
| Three Kingdoms | Multiple alignment of Alanyl-tRNA synthetase... | 1,156 | 35 | - | 37.0% | 7.7% |
| Transcript variants | Multiple alignment of 4 variants of MAPK... | 3,106 | 4 | - | 88.7% | 74.3% |

Alignment View Annotations Distances Text View Info

Extract R.C. Translate Add/Edit Annotation (restricted) Allow Editing (restricted)

91%

Colors: ACCG - Edit

Options Panel

- ☒ Graphs
- ☒ Annotations
- ☒ Consensus
- ☒ Highlighting
- ☐ Complement
- ☐ Translation
- ☒ Linear View
- ☐ Wrap
- ☒ Show Names
- ☒ Show Description
- ☒ Show sequence numbers

Help

Tutorial Help

Contents

geneious

Welcome to Geneious

This tutorial will give you an introductory tour of some of the features of Geneious.

The top of the Geneious window has a tool bar presenting a row of icons for various functions.

There are a few small differences in how Apple Mac, Windows and Linux users use Geneious. We will explain these as they arise.

At any stage, to display help about the current Services folder or document viewer in Geneious click on the Help tab at the top of this panel.

Click the link below to proceed through the tutorial. Alternatively, click the contents link at the top of the page to choose a topic to jump to.

For many topics you will find a link to additional documentation.

Using 131 / 996 MB memory

Alt click on a sequence position or annotation, or select a region to zoom in. Alt-shift click to zoom out.

Plugins

Plugins are downloadable modules that can add new functionality to the program.

Biocode LIMS plugin: Laboratory Information Management System (LIMS) allows you to manage your laboratory and sequence editing workflows.

GenBank Submission plugin: Submits completed sequences to GenBank. It does all of the hard work of bundling together the various parts of a GenBank submission -- sequence data, specimen metadata, trace files, etc.

Other plugins include:

MAFFT, Microsatellite, PHYML, MIRA, Tophat

Installing Biocode Plugin

Exercise 3:

Download and install the Biocode plugin

Once complete, login

LIMS DB Info:

Server Address: db.bwplims.info

Database Name si_barcode_training

Username: si_training

Password: Sm!thsonian

Resources

- BWP FIMS Term Definitions: <http://barcodeofwildlife.org/fims/terms.html>

Note: We will update this link to the SI Barcode Network FIMS Definitions shortly

- SI Barcode Network GitHub page: <https://github.com/MikeTrizna/SIBarcodeNetwork>
- LIMS wiki containing information on the plugin and an online manual: http://software.mooreabiocode.org/index.php?title=Main_Page
- Geneious Forums: <https://support.geneious.com/home>