Collecting experiences in Melanesia - best practices for efficient digitization

Shelley A. James
iDigBio, Florida Museum of Natural History
sjames@flmnh.ufl.edu

iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This work was also supported by NSF grants DEB-0950207 and DBI-1057453, and CEPF.

Island Biology, 2016

Field to database born digital – important?

- Making data accessible, discoverable, useable sooner
- Improve collections management efficiency
- Sustainability
- Expedition and collection tracking
- Publishing and citation

Minimum data fields needed for biodiversity aggregators?

Darwin Core CC0

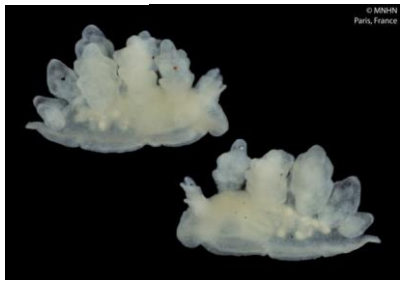
- recordID
- occurrenceID (unique!)
- scientificName
- **eventDate**
- **recordedBy** (ORCID!)
- **Field/collector number**
- **Locality information**
  - **eventID**
- catalogNumber
- institutionID
- collectionID
- Geological Context

Audubon Core CC BY

- recordID
- occurrenceID of specimen
- URL
- Camera EXIF
- photographer

Metadata

- Institution
- Collection
- Contact & info
- Description
- URL



Biodiversity Information Standards TDWG

Collect data with a particular level of reuse in mind!

Valevahalo (base) camp, Guadalcanal

Locality (incl. state/region)

Lat.	*N / S*	**Long.**	*E / W*	**Error**	**m**
Frequency				**Altitude**	**m**
				Aspect	
Habitat (substrate/host/assoc. species)				**GPS** ○ WGS84 (=GDA94) ○ Other _____	

Habit (bark, wood)

Habit (leaves)

Flowers/Sori

Fruits

Notes/Local names

Field det.

DNA

Images

Seed

Live coll.

Alcohol

Not pressed

Unicate

Collector **Date** . . 20

Collection team **No. SAJ**

Locality (incl. state/region)

Lat.	*N / S*	**Long.**	*E / W*	**Error**	**m**
Frequency				**Altitude**	**m**
				Aspect	
Habitat (substrate/host/assoc. species)				**GPS** ○ WGS84 (=GDA94) ○ Other _____	

Habit (bark, wood)

Habit (leaves)

Flowers/Sori

DNA

Images

Seed

Live coll.

Alcohol

Not pressed

Unicate

Date . . 20

No. SAJ

The Scenario

Botanical collecting expeditions

> 250 flowering and fruiting specimens

3+ voucher specimens

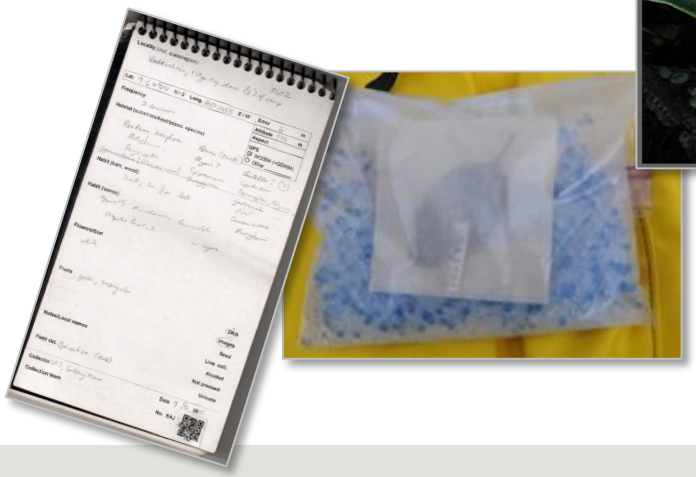
1 tissue sample → products

Alcohol collection fruits, flowers

10+ images per collection (RAW + JPG)

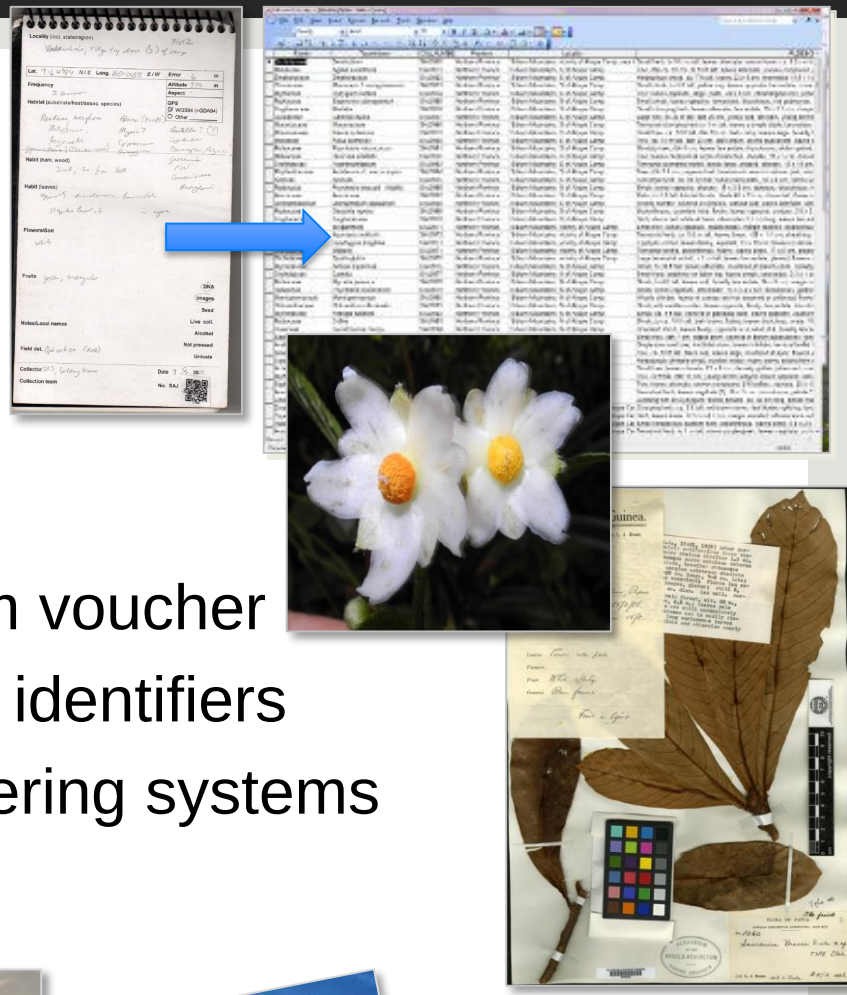
Living collection

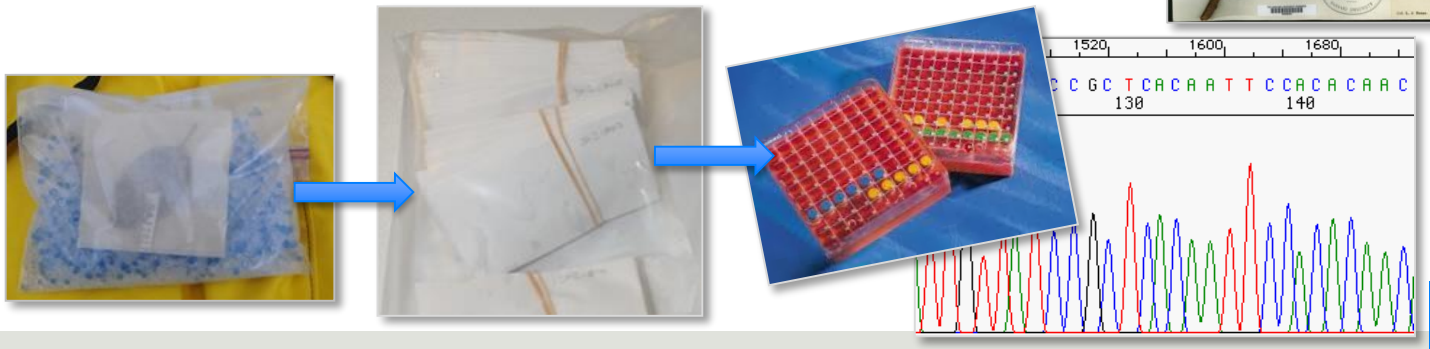
Notebook entry



Collection Tracking!

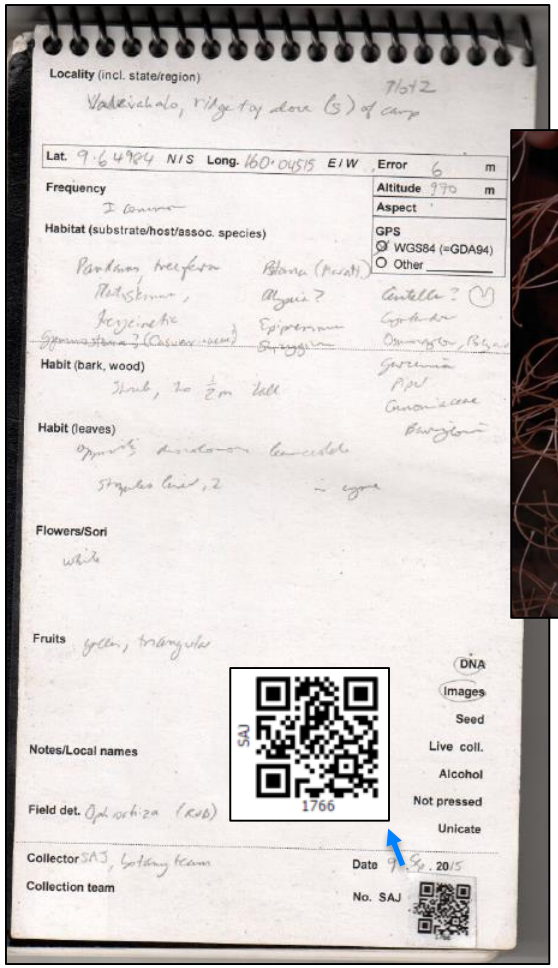
- Field number
- Expedition metadata
- Collection specimen numbers
- Tissue specimen number
- Field images, digitized herbarium voucher
- Genbank and other downstream identifiers
- Other duplicate specimen numbering systems
- etc.

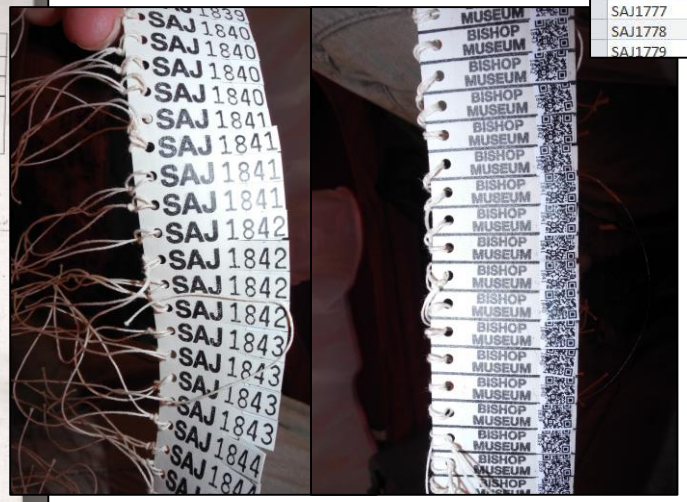




QR codes and collections data

dwc:recordN	dwc:family	dwc:scientificNa	dwc:eventID	dwc:recorde
SAJ1765	Rutaceae	Murraya	4BABCA2BC22A4A0	Chanel, S.
SAJ1766	Rubiaceae	Ophiorrhiza	916E0E076A344FA7	James, S.A.
SAJ1767	Rubiaceae	Oldenlandia	BD7305B8HH204BA	James, S.A.
SAJ1768	Urticaceae	Leucosyke	1F95B1E89C184877	James, S.A.
SAJ1769	Lamiaceae	Plectranthus scute	B655C1F8A0634D2C	James, S.A.
SAJ1770	Gesneriaceae	Cyrtandra cf. filibra	529843C14B31486B	James, S.A.
SAJ1771	Cunoniaceae		4A4D117D633B42B	James, S.A.
SAJ1772	Gentianaceae	Neubergia corynoc	34CDD225D7A34AF	Chanel, S.
SAJ1773	Piperaceae	Piper	A1AAD96EA9544B3	James, S.A.
SAJ1774	Orchidaceae	Bulbophyllum	5906567C3EE94F9F	Chanel, S.
SAJ1775	Asteraceae	Erigeron	C8A3B7166A324F2E	James, S.A.
SAJ1776	Sapindaceae	Harpullia	7C6C65BB015E49EC	James, S.A.
SAJ1777	Cunoniaceae	Spiraeanthemum	49BB1372D4F04D5E	James, S.A.
SAJ1778	Actinidiaceae	Saurauia	0BC9A336CB164CA0	Chanel, S.
SAJ1779	Melastomatace	Medinilla	8BBB0A36C9144A0E	Chanel, S.



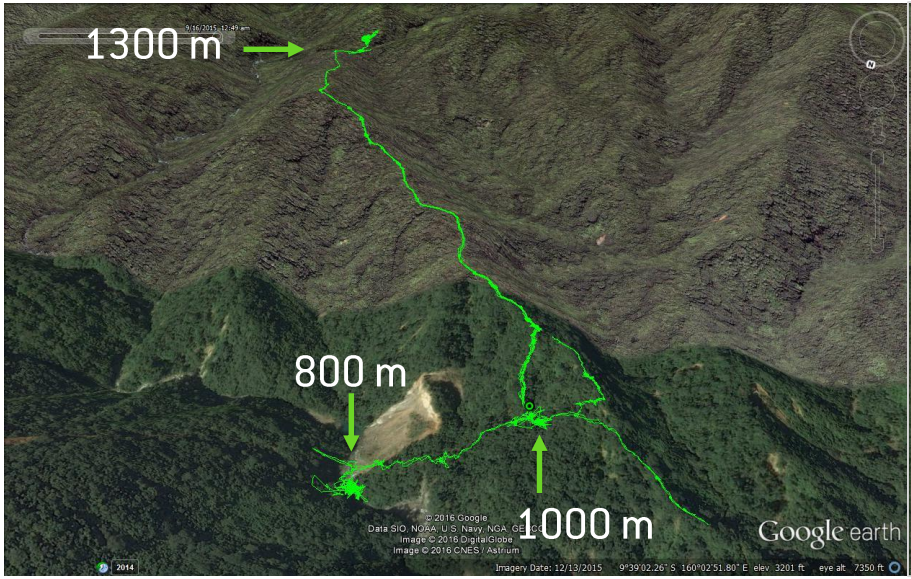


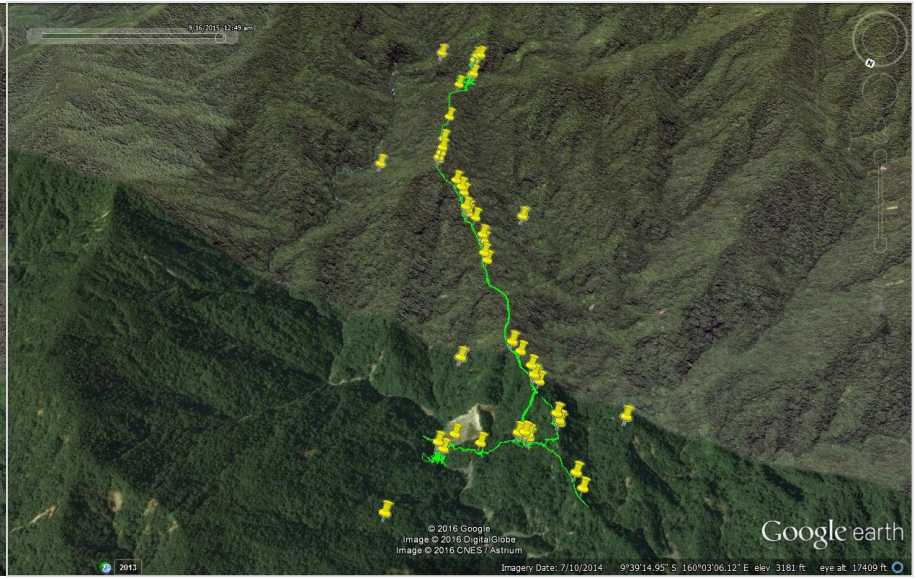
916E0E076A344FA7BEC4084443F3CC66

dwc:eventID

Other useful tips

- Keep GPS track; Geotag images
- Photograph tag number before each specimen
- Don’t delay on transcribing/cleaning/integrating data!





More tricks and tips!

- Mobile technology and apps
- Tools for cleaning and standardizing data

  - Data collection apps
  - Georeferencing tools
  - Photo geotagging
  - Measurement tools (height, direction, distance)
  - Audio collection tools

  - OpenRefine
  - Google Earth, other visualization software
  - Taxonomic name etc. validation services

More tricks and tips!

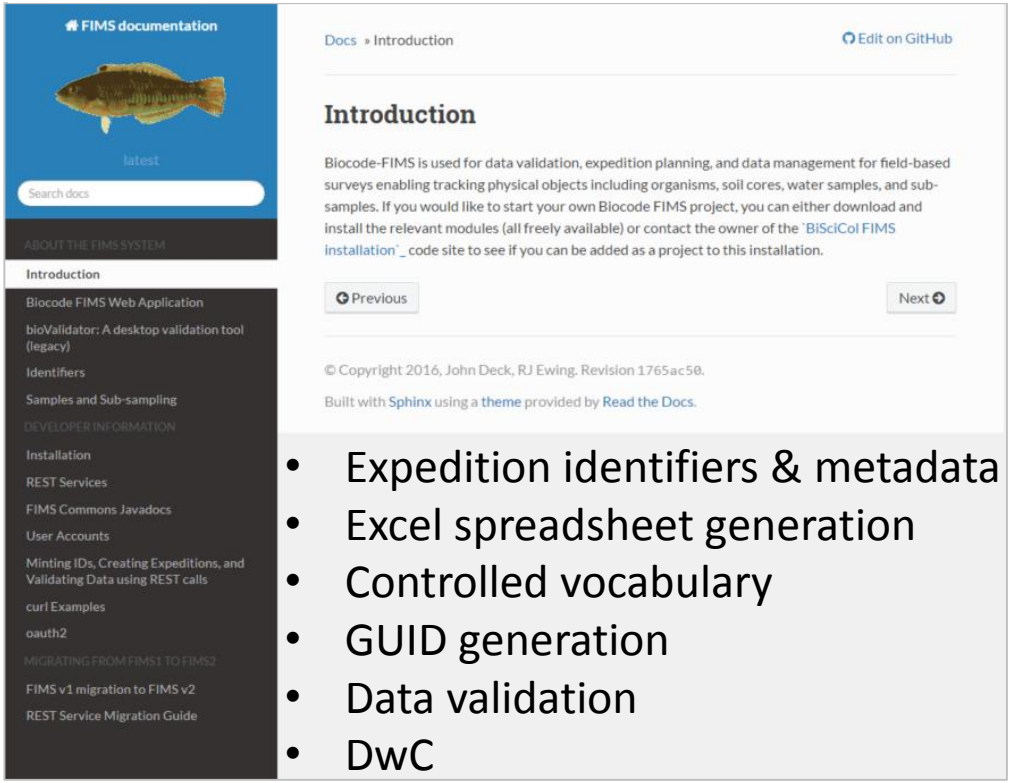
- Mobile technology and apps
- Tools for cleaning and standardizing data
- Field information management system
- Develop a sustainable workflow
- Publish using identifiers!
- Reach out to iDigBio for information:
  - Field to Database Wiki
  - Glossary of Terms

Biocode Field Information Management System

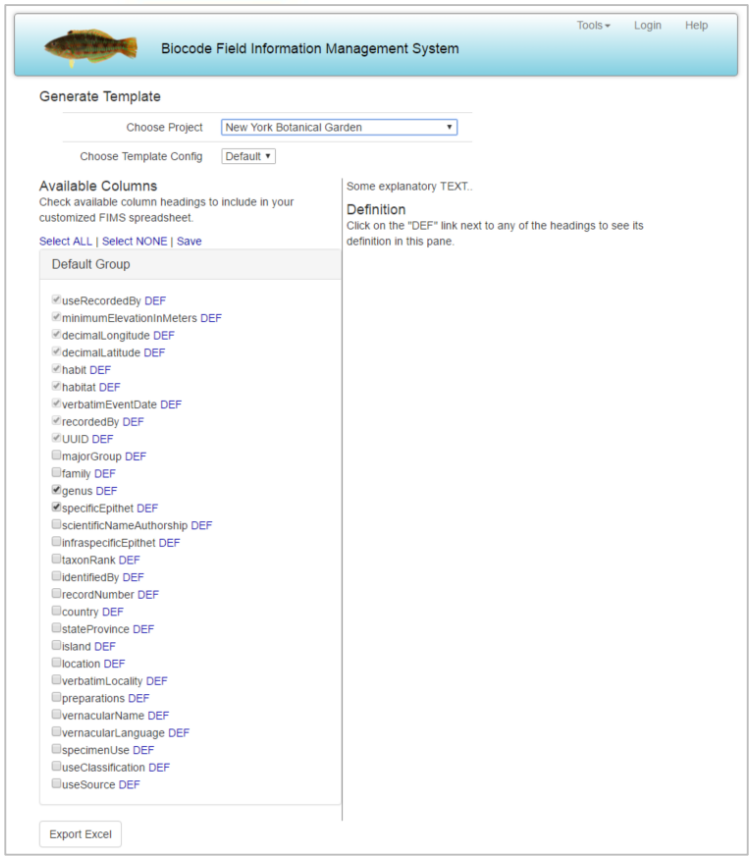
http://www.biscicol.org/

Tools Login Help

A Field Information Management System (FIMS) enables data collection at the source (in the field) by generating spreadsheet templates, validating data, and assigning persistent identifiers to collected samples. The following diagram shows how the system works. The most typical functions are the **Generate Template** and **Validate and Load Data** options, both of which can be found under the Tools menu.



- Expedition identifiers & metadata
- Excel spreadsheet generation
- Controlled vocabulary
- GUID generation
- Data validation
- DwC



More tricks and tips!

- Mobile technology and apps
- Tools for cleaning and standardizing data
- Field information management system
- Develop a sustainable workflow
- Publish & archive data using appropriate identifiers!

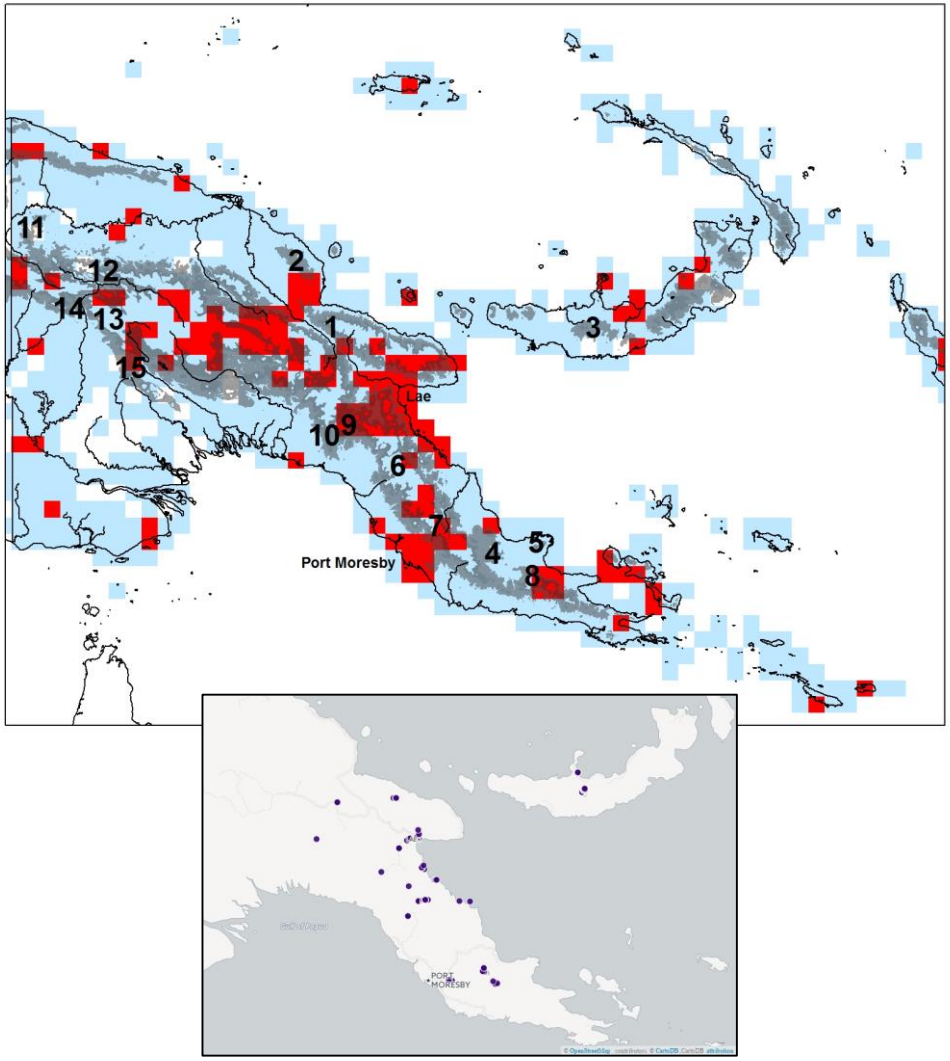
PENSOFT
GBIF
DRYAD
NCBI
iDigBio
Integrated Digitized Biocollections
PLAZI
TAKING CARE OF FREEDOM
figshare

More tricks and tips!

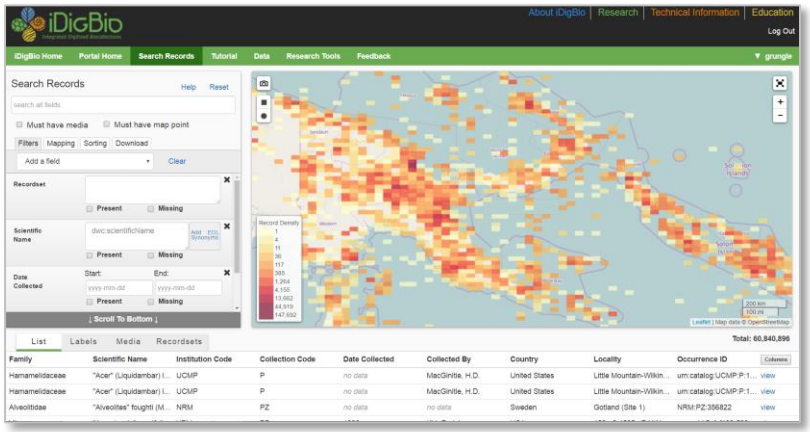
- Mobile technology and apps
- Tools for cleaning and standardizing data
- Field information management system
- Develop a sustainable workflow
- Publish & archive data using identifiers!
- Reach out to iDigBio for information:
  - Field to Database Wiki
  - Glossary of Terms

www.iDigBio.org

Using biodiversity data to plan expeditions



- iDigBio
- GBIF
- Other biodiversity data sources
- Physical specimens
- Previous survey documentation
- Literature



www.iDigBio.org

Shelley James, sjames@flmnh.ufl.edu

UF UNIVERSITY of FLORIDA
FLORIDA STATE UNIVERSITY 1851
FLORIDA MUSEUM OF NATURAL HISTORY
NSF

Support

- National Science Foundation
- iDigBio, Florida Museum of Natural History & University of Florida
- Bishop Museum
- experiment.com supporters
- Critical Ecosystem Partnership Fund
- University of South Pacific & American Museum of Natural History
- MacArthur Foundation, Mellon Foundation, other private foundations
- PNG Department of Environment and Conservation
- PNG National Museum, PNG Forest Research Institute, PNG Department of Forestry
- Solomon Islands Government
- Solomon Islands Community Conservation Partnership (SICCP)
- Uluna-Sutahuri Tribe, Kamiali Village, & many PNG villagers who provided land access and contributed knowledge, time and resources
- Many collaborators