

# Field to database to aggregator and beyond: documenting the flora of Melanesia

Shelley A. James

iDigBio, Florida Museum of Natural History



*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 CEPP.*

SPNHC, Berlin, 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



- recordID
- **occurrenceID** (unique!)
- scientificName
- eventDate
- recordedBy (ORCID!)
- Locality information
- catalogNumber
- **institutionID**
- **collectionID**
- Geological Context

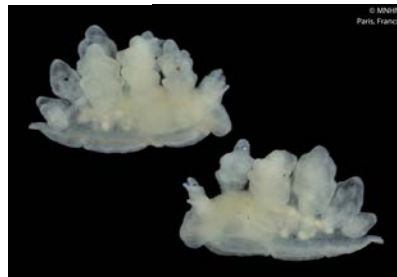
## Audubon Core



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

## Metadata

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



## Valevahalo (base) camp, Guadalcanal







Locality (incl. state/region)

Lat.	N / S	Long.	E / W	Error	m
Frequency				Altitude	m
Habitat (substrate/host/assoc. species)				Aspect	
GPS					
<input type="radio"/> WGS84 (=GDA94)					
<input type="radio"/> Other _____					

---

Habit (bark, wood)

Habit (leaves)

Flowers/Sori

Fruits

Notes/Local names

Field det.

---

Collector \_\_\_\_\_ Date . . . 20

Collection team \_\_\_\_\_ No. SAJ \_\_\_\_\_

DNA

Images

Seed

Live coll.

Alcohol

Not pressed

Unicate





Locality (incl. state/region)

Lat.	N / S	Long.	E / W	Error	m
Frequency				Altitude	m
Habitat (substrate/host/assoc. species)				Aspect	
GPS					
<input type="radio"/> WGS84 (=GDA94)					
<input type="radio"/> 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 taxa/specimens

3+ voucher specimens

1 tissue sample → products

Alcohol collection fruits, flowers

10+ images per collection (RAW + JPEG)

Living collection

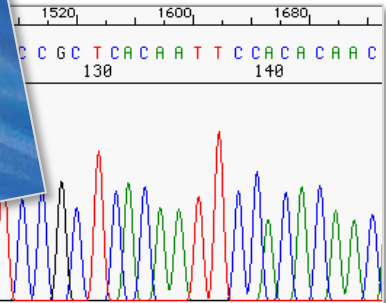
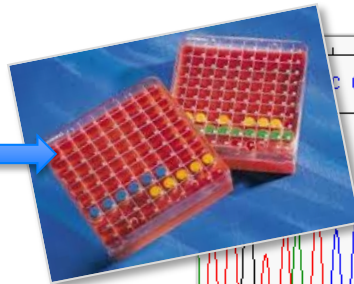
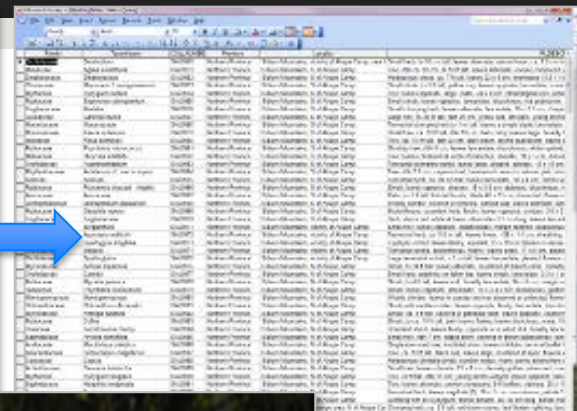
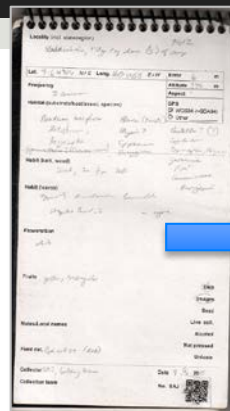
Notebook entry





# Collection Tracking!

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



# QR codes and collections data

dwc:recordID	dwc:family	dwc:scientificName	dwc:eventID	dwc:recordedBy
SAJ1765	Rutaceae	Murraya	4BABC2BC22A4AC	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. filibri	529843C14B31486B	James, S.A.
SAJ1771	Cunoniaceae		4A4D117D633B42BI	James, S.A.
SAJ1772	Gentianaceae	Neubergia corynoloba	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	0BC9A336CB164CAI	Chanel, S.
SAJ1779	Melastomataceae	Medinilla	8BB0A36C9144A0I	Chanel, S.

Locality (incl. state/region) *7/02*  
*Vadivelalo, ridge top above (S) of camp*

Lat. *9.64984* NIS Long. *160.04515* E/W Error *6* m  
 Frequency *I common* Altitude *970* m  
 Aspect  
 Habitat (substrate/host/assoc. species) GPS  
 WGS84 (=GDA94)  
 Other

Habit (bark, wood) *Parkia nelsonii* *Balan (Murr.)*  
*Protium*, *Albizia?* *Antella?*   
*Freycinetia* *Excoecaria* *Cordia*  
*Syntherisma?* *Casuarina* *Samanea* *Orangia* *Platanus*  
 Habit (leaves) *Shrub, to 2m tall* *Green*  
*Opposite, dark green, lanceolate* *Pubescent*  
*Stipules cordate, 2* *in pairs*

Flowers/Sori *white*

Fruits *yellow, triangular*

Notes/Local names

Field det. *Op. 10/1/20 (KSB)*

Collector *SAJ, Botany team* Date *9 Sep. 2015*  
 Collection team No. SAJ

DNA  
 Images  
 Seed  
 Live coll.  
 Alcohol  
 Not pressed  
 Unicate

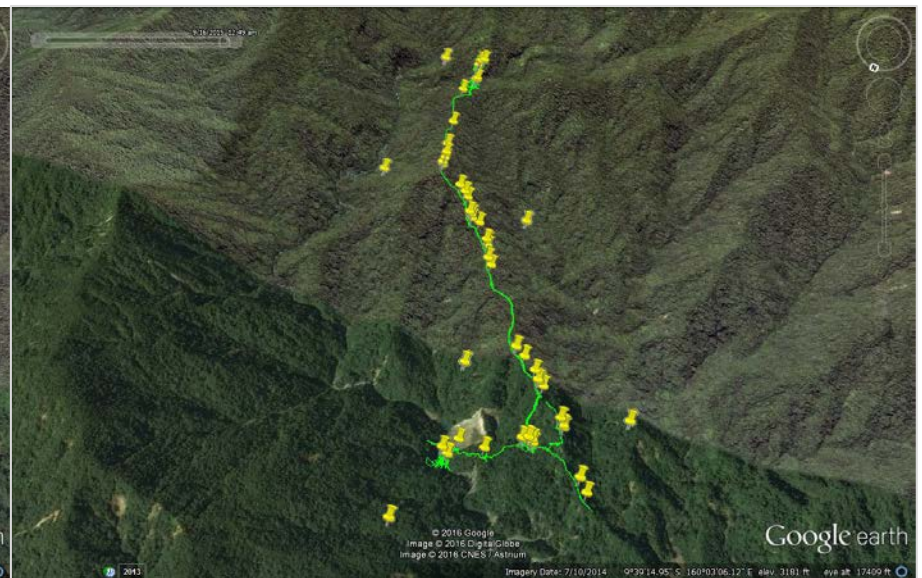
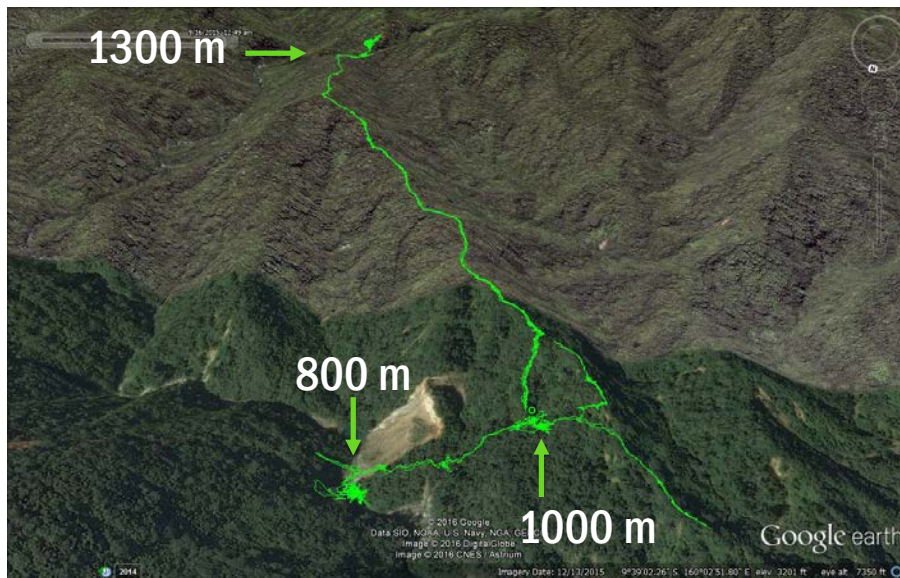
SAJ 1766



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  *Refine*
- 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



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.



**FIMS documentation**



latest

ABOUT THE FIMS SYSTEM

**Introduction**

- Biocode FIMS Web Application
- bioValidator: A desktop validation tool (legacy)
- Identifiers
- Samples and Sub-sampling

DEVELOPER INFORMATION

- Installation
- REST Services
- FIMS Commons Javadocs
- User Accounts
- Minting IDs, Creating Expeditions, and Validating Data using REST calls
- curl Examples
- oauth2

MIGRATING FROM FIMS1 TO FIMS2

- FIMS v1 migration to FIMS v2
- REST Service Migration Guide

Docs » Introduction [Edit on GitHub](#)

## Introduction

Biocode-FIMS is used for data validation, expedition planning, and data management for field-based surveys enabling tracking physical objects including organisms, soil cores, water samples, and sub-samples. If you would like to start your own Biocode FIMS project, you can either download and install the relevant modules (all freely available) or contact the owner of the 'BiSciCol FIMS installation' code site to see if you can be added as a project to this installation.

[Previous](#)
[Next](#)

© Copyright 2016, John Deck, RJ Ewing. Revision 1765ac58.  
Built with [Sphinx](#) using a theme provided by [Read the Docs](#).

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

**Biocode Field Information Management System**

Tools ▾ Login Help

### Generate Template

Choose Project: New York Botanical Garden ▾

Choose Template Config: Default ▾

**Available Columns**

Check available column headings to include in your customized FIMS spreadsheet.

Select ALL | Select NONE | Save

Default Group

- useRecordedBy DEF
- minimumElevationInMeters DEF
- decimalLongitude DEF
- decimalAltitude DEF
- habit DEF
- habitat DEF
- verbatimEventDate DEF
- recordedBy DEF
- UUID DEF
- majorGroup DEF
- family DEF
- genus DEF
- specificEpithet DEF
- scientificNameAuthorship DEF
- infraspecificEpithet DEF
- taxonRank DEF
- identifiedBy DEF
- recordNumber DEF
- country DEF
- stateProvince DEF
- island DEF
- location DEF
- verbatimLocality DEF
- preparations DEF
- vernacularName DEF
- vernacularLanguage DEF
- specimenUse DEF
- useClassification DEF
- useSource DEF

Some explanatory TEXT.

**Definition**

Click on the "DEF" link next to any of the headings to see its definition in this pane.

[Export Excel](#)

## 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!

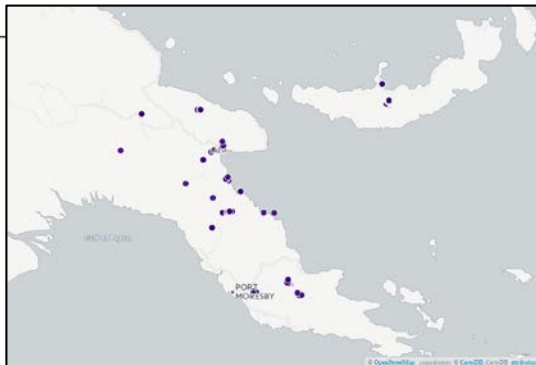
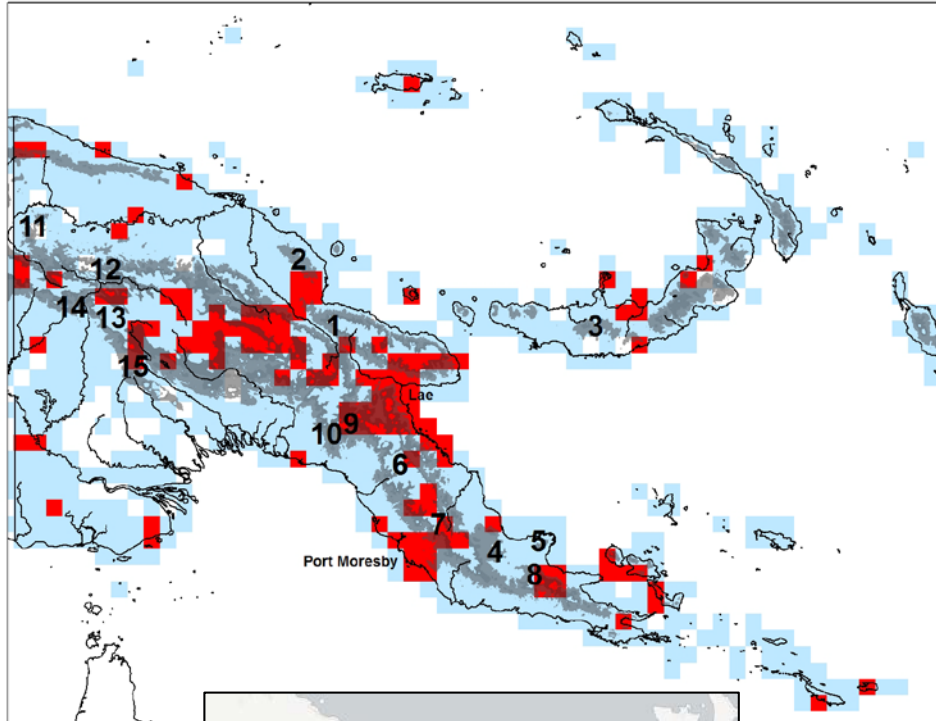


## More tricks and tips!

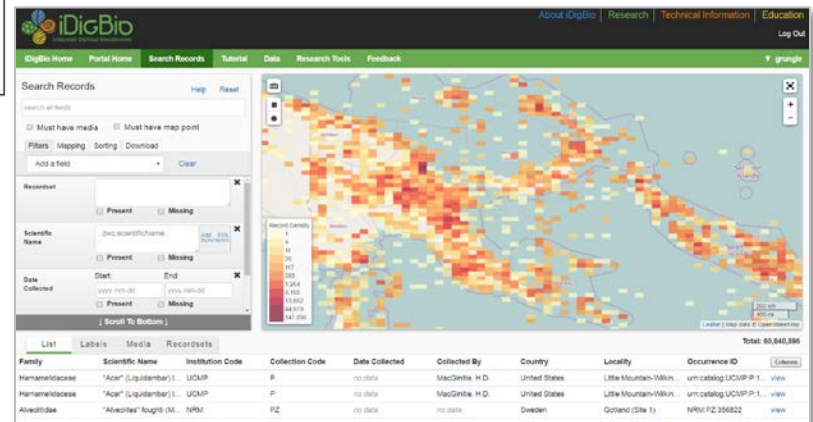
- Mobile technology and apps
- Tools for cleaning and standardizing data
- Field information management system
- Develop a sustainable workflow
- Publish & archive data using identifiers!
- Reach out to iDigBio for information:
  - Field to Database Wiki
  - Glossary of Terms



# Using biodiversity data to plan expeditions



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



**Search Results Table:**

Family	Scientific Name	Institution Code	Collection Code	Date Collected	Collected By	Country	Locality	Occurrence ID
Ranunculaceae	"Acp" (Liquidambar) L.	UCJRP	B	02/2004	MacGillivray, M.D.	United States	Little Mountain-NMUN	urn:catalog:UCJRP:11
Ranunculaceae	"Acp" (Liquidambar) L.	UCJRP	P	02/2004	MacGillivray, M.D.	United States	Little Mountain-NMUN	urn:catalog:UCJRP:11
Alveolidae	"Alveolites" (Siphon) M.	NM	PZ	02/2004	MacGillivray, M.D.	Sweden	Gotland (Site 1)	NR001PZ:200402

[www.iDigBio.org](http://www.iDigBio.org)

**I Dig Bio**  
**do you?**



Advancing Digitization of Biological Collections

Shelley James, [sjames@flmnh.ufl.edu](mailto:sjames@flmnh.ufl.edu)

