Field Collections to Digital Data: A Workflow for Fossils and the Use of Digital Data for Reconstructing Ancient Forests

Dori L. Contreras
University of California Museum of Paleontology
New collections: meeting multiple needs

Research project goals:
- Collect new fossils
- Identify/describe
- Collect data/measure
- Analyze data
- Publish research
  - (Wait! Need repository info and specimen numbers!)

Museum & archival goals:
- Care and house specimens
- Manage and store associated data
- Accessibility
  - *Digitization of collection records and specimen photos
Research based on new collections: integrating short and long-term needs

1) Collection
   - Finding and collecting fossils
   - Locality data and samples
   - Field data (e.g. census counts)

2) Identification/description, etc.

3) Curation
   - Specimen preparation
   - Labels & specimen numbers
   - Organize and storage

4) Digitization
   - Photography
   - Data entry

5) Digital data
   - Management
   - Further research use
Research based on new collections: integrating short and long-term needs

Collection → Curation → Digitization

Identification/description
Census data (counts)

Analyses
Science!
Digital measurement

Photography
Data entry

Data portal
Data aggregator

Available to public and researchers
An example from ongoing project

Reconstructing a Cretaceous forest

1) Project description

2) Methods & Workflow
   - roadblocks and solutions

3) Research and use of digital data
AIM: to address fundamental questions about the structure, diversity, and functioning of forests during the Late Cretaceous ecological radiation of angiosperms

Redrawn from Niklas 1986
Current Project:

Reconstructing structure & functional diversity of a Cretaceous forest

“Dori’s Tuff” deposit

• South-Central New Mexico; McRae Fm.
• Late Campanian; 74.7 Ma (U-Pb, Amato et al. 2017)
• Ashfall bed with little to no transport
  • Single depositional event on stable floodplain
• ~1.2 km!
Project Overview:

Reconstructing structure & functional diversity of a Cretaceous forest

Phase 1. “Build” a forest
- describe taxonomic diversity
- relative abundance of taxa
- spatial structure of community

Phase 2. Measure the forest: quantify functional diversity
- measure functional leaf traits of all taxa
- reconstruct trait diversity across transect

*quantitative and spatial explicit sampling scheme
*large sample sizes
*digital measurements of georeferenced specimens
A workflow from field to digital

Field
- Collection
- Identification/description
- Census data

Museum
- Curation
- Digitization
- Trait measurement
- Photography
- Data entry
- Data portal
- Data aggregator
Established 26 quarry sites spanning deposit:
- each treated as separate locality
- collected relevant site data for each:
  - GPS coordinates
  - sedimentological info and samples, quarry dimensions
  - photographs
Field Work – collections and census

At each quarry:
- Bust out rock and expose fossils
  *each rock gets a unique field ID number*
Field Work – collections and census

At each quarry:
- Bust out rock and expose fossils
  *each rock gets a unique field ID number*
- Census: Identify and record leaf morphotypes by:
  1) Relative abundance: number of leaf specimens
  2) Percent cover: # of 2-cm line increments crossed (Wing et al. 1993, 2012)

> 6,350 specimens
> 61,718 cm² of rock surface

> 1,945 rocks brought to UCMP

“Keepers” are wrapped and labeled
Conifers & Cycads

Geinitzia type

Brachyphyllum sp.

Redwood relative

Zamiod cycad
Angiosperms - monocots

- Sabalites
- Orontium mackii
- Zingerberopsis
- Pandanites
Angiosperms – toothed “dicots”

scale = 5 mm
Angiosperms – pinnate “dicots”

scale = 5 mm
The spoils of field work

**Spoils:**
- Locality info
- Census databooks
- Boxes of fossils

**Collection** → **Curation** → **Digitization**

- Identification/description
- Census data

- Trait measurement
- Photography
- Data entry
- Data portal
- Data aggregator
The transition

**Spoils:**
- Locality info
- Census databooks
- Boxes of fossils

**Collection** → **Curation** → **Digitization** → **Museum**

- Identification/description
- Census data

- Trait measurement
- Photography
- Data entry
- Data portal
- Data aggregator
Museum and Lab: the transition

1) **Data entry**: census data

2) **Unpacking fossils**
   - Matching parts and counterparts
   - Boxing and organizing into drawers
   - Curatorial log
     *Rock ID*
Museum and Lab: the transition

1) **Data entry:** census data

2) **Unpacking fossils**
   - Matching parts and counterparts
   - Boxing and organizing into drawers
   - Curatorial log
     - *Rock ID
     - *Reconcile for printing labels with species identifications
101 Drawers to process....

Field

Spoils:
- Locality info
- Census databooks
- Boxes of fossils

Collection → Curation → Digitization

Identification/description
- Census data

Museum

The un-censused
- specimens that could not be identified in field
  - new morphotypes

Trait measurement

Photography

Data entry

The censused
- specimens ID’ed & well-preserved enough for trait measurements
What needs to be done?

1) Identify and census the difficult specimens

2) Prep specimens if necessary

3) Complete descriptions of taxa

4) Curation: assign # to each specimen on each rock and label; fill out identification cards; enter data

5) Photograph specimens
People Power
**Processing: Two Methods**

**Batch processing**
- one type of task at a time
- flag specimens for different processes
- let specimens build up until have a “batch” to process

**Integration of tasks**
- do everything at once!
- drawer by drawer, processing each specimen completely
Re-organization and customization of workspace
Re-organization and customization of workspace

Prep station
Descriptions

Photography station

White-stripe and census station
Data entry, labels
Re-organization and customization of workspace

Prep station
Descriptions

Photography station

White-stripe and census station
Data entry
Re-organization and customization of workspace
Shifting pinch-points in the workflow

Batch processing

Field work → Curation → Digitization

preparation

Identification/description
Census data

Analyses

Science!

Trait measurement

Photography
Data entry

Data portal
Data aggregator

Available to public
and researchers
Shifting pinch-points in the workflow

Integrated processing

- Field work
- Curation
- Digitization

Identification/description
Census data

preparation

Analyses
Science!

Trait measurement

Photography
Data entry

Data portal
Data aggregator

Available to public
and researchers
Project Overview:

Reconstructing functional diversity of a Cretaceous forest

**Phase 1. “Build” a forest**
- describe taxonomic diversity
- relative abundance of taxa
- spatial structure of community

**Phase 2. Measure the forest: quantify functional diversity**
- measure functional leaf traits of all taxa
- reconstruct trait diversity across transect

*quantitative and spatial explicit sampling scheme
*large sample sizes
*digital measurements of georeferenced specimens
Taxonomic diversity

From census:

➢ 158 leaf morphotypes

➢ Angiosperms: ~89% of diversity (141 morphotypes)

Species accumulation curves
Community composition

Relative abundance and % cover

![Bar graph showing relative abundance and % cover for different taxonomic groups.](chart.png)
Community composition

Dominance and rank abundance
Community composition

Dominance and rank abundance

Morphotype

Number of specimens

TaxonGroup
- Conifer
- Cycad
- Dicot
- Fern
- Monocot
- Other
Community composition

Dominance and rank abundance

Morphotype

Number of specimens

TaxonGroup
Conifer
Cycad
Dicot
Fern
Monocot
Other

Redwood relative
tree

Zingerberopsis
ground cover
Community composition

Dominance and rank abundance

- Brachyphyllum sp.
- Dryophyllum sp.
- Redwood relative
- Zingerberopsis
- Sabalites sp.

Ground cover
Spatial heterogeneity

*Also beta diversity; Sorenson’s index by quarry distance
Project Overview:

Reconstructing functional diversity of a Cretaceous forest

Phase 1. “Build” a forest
-describe taxonomic diversity
-relative abundance of taxa
-spatial structure of community

Phase 2. Measure the forest: quantify functional diversity
-measure functional leaf traits of all taxa
-reconstruct trait diversity across transect

*quantitative and spatial explicit sampling scheme
*large sample sizes
*digital measurements of georeferenced specimens
Using the digital data

Field work → Curation → Digitization

Identification/description
Census data

preparation

Analyses

Science!

Trait measurement

Photography
Data entry

Data portal
Data aggregator

Available to public and researchers
Digital Trait Measurements

Digitized specimens piped into Adobe Photoshop for image analysis:

- Leaf length
- Leaf width
- Tooth spacing
- Leaf area
- Leaf perimeter
- Specific leaf area
- Petiole length
- Petiole width
- Petiole area
Using the digital data

Field work → Curation

Identification/description
Census data

preparation

Digitization

Analyses

Trait measurement

Science!

Photography
Data entry

Data portal
Data aggregator

Available to public and researchers
The data

Specimen data

Images

Census & Trait data

Morphotype descriptions

**COLLECTIONS**

UCMP has the largest paleontological collection of any university museum in the world. These well-curated and comprehensively documented collections of fossils and modern organisms representing vertebrates to vertebrate collected from all continents. The Museum serves the University community in research and education, and provides support for instruction at Berkeley and other UC campuses. In addition, the collections are used by paleontologists, biologists, and geologists throughout the world.

**POLICIES & CONTACTS**

UCMP collections and facilities are available for use by qualified researchers. Follow the link for policies on using the collections, or for contact information.

**UCMP COLLECTIONS**

Read more about UCMP’s paleontological collections.

- Mollusks
- Insects
- Vertebrates

**ON-LINE COLLECTIONS DATABASE**

Visitors can search the online database for information about specimens housed at UCMP.

**RELATED ITEMS**

- UCB Archival/Supplementary Data Collection
- Research reports and supplementary data sets

- UCMP-Archival Collections
- Information about library holdings, images, and memorabilia

- UCMP Data Model
- About modeling data for collections management.

**Pacific Rim Catalog**

Database on Pacific Rim biodiversity.

**CalPhotos**

- [CalPhotos](https://calphotos.berkeley.edu)

**iDigBio**

- [iDigBio](https://idigbio.org)

**DRYAD**

- [DRYAD](https://dryad.org)

**.CSV**
Image Data: Adobe Bridge

Images saved as Raw and JPEG

Metadata entered using Adobe Bridge

GPS
- Latitude: 33.10.299N
- Longitude: 107.7.0937W
- Altitude: Manual capture

Description: UCMP F115.01
Keywords: B115; Peltate; Entire; Palmate

IPTC Subject Code

Date Created: 5/30/17, 12:06:45 PM

Sublocation: TT92-3 Quarry 4
The data

Specimen data

Images

Census & Trait data

Morphotype descriptions

CalPhotos

.CSV
The data

Specimen data

Images

Census & Trait data

Morphotype descriptions

A new database?
**Morphotype Descriptions: Filemaker Pro**

**Section I. Leaf Characters**

<table>
<thead>
<tr>
<th>Character</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaf attachment</td>
<td>petiolate 1.1</td>
</tr>
<tr>
<td>Leaf arrangement</td>
<td>not visible 88</td>
</tr>
<tr>
<td>Leaf organization</td>
<td>not visible 88</td>
</tr>
<tr>
<td>Leaflet organization</td>
<td>not visible 88</td>
</tr>
<tr>
<td>Leaflet attachment</td>
<td>absent 0</td>
</tr>
<tr>
<td>Blade</td>
<td>marginal 7.1</td>
</tr>
<tr>
<td>Lamina attachment</td>
<td>submarginal 8.5</td>
</tr>
<tr>
<td>Laminar size</td>
<td>ovate 10.3</td>
</tr>
<tr>
<td>Laminar L/W ratio</td>
<td>submarginal 11.1</td>
</tr>
<tr>
<td>Laminar shape</td>
<td>submarginal 12</td>
</tr>
<tr>
<td>Median symmetry</td>
<td>submarginal 13</td>
</tr>
<tr>
<td>Lobation</td>
<td>submarginal 14.23</td>
</tr>
<tr>
<td>Margin type</td>
<td>submarginal 15</td>
</tr>
<tr>
<td>Margin features</td>
<td>submarginal 16</td>
</tr>
<tr>
<td>Apex angle</td>
<td>obtuse 16.2</td>
</tr>
<tr>
<td>Apex shape</td>
<td>rounded 17.21</td>
</tr>
<tr>
<td>Apex shape b</td>
<td>straight 17.1</td>
</tr>
<tr>
<td>Base angle</td>
<td>reflex 18.3</td>
</tr>
<tr>
<td>Base shape</td>
<td>straight 19</td>
</tr>
<tr>
<td>Term apex features</td>
<td>not visible 88</td>
</tr>
</tbody>
</table>

**Section II. Veneration**

<table>
<thead>
<tr>
<th>Character</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary vein framework</td>
<td>basal</td>
</tr>
<tr>
<td>Naked basal veins</td>
<td>24</td>
</tr>
<tr>
<td>Number of basal veins</td>
<td>9-10</td>
</tr>
<tr>
<td>Apogonic veins</td>
<td>absent 0</td>
</tr>
<tr>
<td>Perimarginal veins</td>
<td>30</td>
</tr>
<tr>
<td>Major 2nd spacing</td>
<td>major 31</td>
</tr>
<tr>
<td>Variation of 2nd angle</td>
<td>major 32</td>
</tr>
<tr>
<td>Major 2nd attachment</td>
<td>major 33</td>
</tr>
<tr>
<td>Intersecondary</td>
<td>proximal 34.1</td>
</tr>
<tr>
<td>Proximal length</td>
<td>absent 0</td>
</tr>
<tr>
<td>Distal course</td>
<td>34.2</td>
</tr>
<tr>
<td>Vein frequency</td>
<td>34.3</td>
</tr>
<tr>
<td>Angle of percutum 3rd</td>
<td>35.1.2</td>
</tr>
<tr>
<td>Vein angle variability</td>
<td>36</td>
</tr>
<tr>
<td>Epimediastal terrains</td>
<td>37.1</td>
</tr>
<tr>
<td>Inter secondary</td>
<td>proximal 37.2</td>
</tr>
<tr>
<td>Distal course</td>
<td>37.2.2</td>
</tr>
<tr>
<td>Extent 3rd course</td>
<td>38</td>
</tr>
</tbody>
</table>

**Section III. Teeth**

<table>
<thead>
<tr>
<th>Character</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tooth spacing</td>
<td>regular 44.2</td>
</tr>
<tr>
<td>Number of orders of teeth</td>
<td>one 45.1</td>
</tr>
<tr>
<td>Teeth per cm</td>
<td>46</td>
</tr>
<tr>
<td>Shape</td>
<td>moudn 47.2</td>
</tr>
<tr>
<td>Distal tooth shape</td>
<td>concave 48-CC</td>
</tr>
<tr>
<td>Proximal tooth shape</td>
<td>concave 48-CC</td>
</tr>
<tr>
<td>Principal vein termination</td>
<td>50</td>
</tr>
<tr>
<td>Course of accessory vein</td>
<td>51</td>
</tr>
<tr>
<td>Features of the tooth apex</td>
<td>52</td>
</tr>
</tbody>
</table>

**Section I. Leaf Description**

Leaf arrangement not visible with petiole attachment. Leaf petiole features absent. Leaf arrangement not visible with not visible attachment. Leaf organization not visible. Blade attachment: marginal, laminal size: mesophyll, laminal L/W ratio: symmetrical blade, blade mediately symmetrical with symmetrical blade, blade united, and margin crenate with absent angle features. Apex angle obtuse, apex shape rounded to straight, blade base straight to shape, blade base straight, blade base apex 19. Surface texture absent, and surficial glands absent.

**Section II. Veneration Description**

Primary venation basal acrothalamous with 9-10 basal veins and apogonic absent. Secondary veins: fasted torbidothalamous with anerien secondary veins present, minor secondary course simple torbidothalamous, and absent perimarginal veins. Major secondary venation with variation of secondary veins and major attachment decurrent. Intersecondary proximal course absent, angle: distal course, and frequency: intersecondary tertiary ven fabric convex opposite percutent with the angle of percutent tertiary veins and vein angle variability. Epimediastal territories with proximal course and distal course. The exterior tertiary course is a: Quaternary ven fabric is: Quaternary ven fabric not visible, and with: Marginal ultimate.
Morphotype Descriptions: Filemaker Pro
Going public!

Field work → Curation → Digitization

preparation

Identification/description
Census data

Analyses
Science!
Trait measurement

Photography
Data entry

Data portal
Data aggregator

Available to public and researchers
Take-home points

• Workflows should integrate project-based data collection with curation and digitization

• Use of effective links (field-assigned Rock IDs) to bridge field data with museum specimens

• Reorganize/customize workspace for integration of tasks to increase efficiency

• Person-power! Importance of involving students and volunteers in research
Funding:
Lewis and Clark Fund for Exploration and Field Research (2015); UCMP Graduate Student Awards (2013, 2014, & 2015); Integrative Biology Graduate Research Fund: The Reshetko Family Scholarship (2015); Geological Society of America Graduate Student Research Grant (2014); Integrative Biology Graduate Research Fund (2014 & 2015); Sigma Xi Grants-in-Aid of Research UC-Berkeley Chapter (2014); Mid-American Paleontological Society (MAPS) Outstanding Student Research Award (2013); GRAC Research Funds, UC-Berkeley Integrative Biology Dept. (2013 & 2015); UCMP Graduate Student Fellowship; NSF Graduate Research Fellowship

Acknowledgements

IDigBio and NSF

Amazing students and volunteers!

Looy Lab
UC Museum of Paleontology
Armendaris Ranch, Tom Waddell, Ted Turner

Diane Erwin, Cindy Looy, Pat Holroyd, Gary Upchurch, Greg Mack

Funding: