AllAsia: Bringing Asian plant diversity to digital life

Charles C. Davis (lead)
Harvard University Herbaria
cdavis@oeb.harvard.edu
Core Objective

- **Digitize** 3 million Asian vascular plant specimens from 19 U.S. herbaria with special attention on Southeast Asia and the Himalaya-Hengduan mountains.
Rationale

- Our main goal is to mobilize data to support basic biodiversity discovery and facilitate ecological and evolutionary investigations especially in hyperdiverse regions of Asia.
Largest continent on Earth;
Complex terrain;
90,000–100,000 species of vascular plants;
Astonishing functional diversity.
54 million herbarium specimens in iDigBio;
4.5% are from Asia and 2% from US institutions;
GBIF = 78 million; 12% from Asia very few with images.
Rationale: **AIMS**

- **Digitize** 3 million vascular plant specimens from US herbaria (Aim 1);

- Develop novel informatics tools and high-throughput digitization **methods** (Aim 2);

- Aggregate and link all digitized records into an **All Asia** portal, 15 million specimens (Aim 3);

- Enhance opportunities for grades 9–12 STEM learners and **early career** scientists (Aim 4).
Species discovery and biogeography in iconic biodiversity hotspots;
Elucidating eco-evolutionary drivers of diversification in contrasting temperate and tropical biomes;
Investigating species phenological response to climate;
Protecting and forecasting biodiversity across imperiled biomes;
Creating next-generation herbarium digitization via innovative design solutions.
Digitization Plan: Consortium Organization

- Harvard University (HUH)
- Bishop Museum (BISH)
- Brown University (BRU)
- U. of Alaska (ALA)
- Botanical Research Inst. Texas (BRIT)
- Chicago Botanic Garden (CHIC)
- U. of Cincinnati (CINC)
- Cleveland Museum of Natural History (CLM)
- U. of Colorado (COLO)
- U. of Massachusetts Amherst (MASS)
- U. of Michigan (MICH)
- Missouri Botanical Garden (MO)
- Miami University (MU)
- U. of New Hampshire (NHA)
- New York Bot, Garden (NY)
- Ohio State Uni. (OS)
- California Bot. Garden (RSA)
- U. of Vermont (VT)
- Smithsonian (US; not funded)
Digitization Plan: Organization

East: VT, BU, NY, US, MASS, HUH, NHA, BRU

Midwest: MO, MU, CLM, OS, CINC, MICH, CHIC

West: RSA, BISH, ALA, COLO, BRIT

Hedrick et al. (2020)
Innovation:
HUH Conveyor Belt

Sweeney et al. (2018)
Back to the future: next generation imaging

• Single station workflow
• Ergonomic design customized for herbarium sheets
• Separate imaging from transcription

8 seconds per image! (down from 20 sec)

• Transcription virtual, separated from imaging (40% reduction)

Key design features:
1. Camera and mount
2. LED lighting
3. Tablet-sized monitor with adjustable arm
4. Adjustable queue tray
5. Receiving area
6. Imaging surface
7. Cubby
8. Height-adjustable legs
9. Barcode dispenser
10. Easy-fire button

Charles C. Davis | AllAsia: Bringing Asian plant diversity to digital life | cdavis@oeb.harvard.edu

C. Grassa | J. Kennedy
Innovation: Machine learning and label transcription

- Label transcription still **largely completed** by human workers; plenty of room for automation
- Goal: Use the LSTM-RNN tool to generate **automate transcription** of handwritten herbarium specimens labels
- Code will be made available through an **opensource license** for broad use
Innovation: Rapid data entry

- Enhanced Symbiota for rapid data entry workflow
- Leveraging lessons learned from HUH rapid entry tool
- Automatic label detection
- Text suggestions from AI-based automatic transcription
Mobilization

- All data and images will be available to public through All Asia Symbiota portal
- 12 million records contributed by international partners (including P, France; L, Netherlands; MW, Russia; CNH, China);
- 15 million records total
Outreach: Bringing biodiversity and computer scientists together

Planned Workshops:

- Plant Biodiversity in Asia: promises and challenges (Boston, MA)
- Collections and the digital herbarium (Ann Arbor, MI)
- Novel applications of digital collections (Fairbanks, AK)

Overarching question: What does the herbarium of the future look like?

Charles C. Davis  |  AllAsia: Bringing Asian plant diversity to digital life  |  cdavis@oeb.harvard.edu
OUTREACH: HACKATHONS

FROM PROBLEMS TO PROTOTYPES TO PRODUCTION

ideathons, hackathons, and experiential learning centered around interdisciplinary collaboration and innovation

1. Identify problems to be solved to advance research and on-the-ground impact

2. Match teams to problems & mentors to build prototypes through virtual and in-person hackathons

3. Continue development on promising prototypes with student development teams & mentors

Fostering diversity and inclusion through interdisciplinarity

Charles C. Davis | AllAsia: Bringing Asian plant diversity to digital life | cdavis@oeb.harvard.edu
Acknowledgements

NSF
National Science Foundation

iDigBio
Integrated Digitized Biocollections

Symbiota²
Mobilizing Biodiversity Data

BUSpark!

iPlant Collaborative™

Charles C. Davis             |  AllAsia: Bringing Asian plant diversity to digital life | cdavis@oeb.harvard.edu