Title: DrawerDissect: Whole-drawer insect imaging, segmentation, and transcription using AI

Authors: *Elizabeth G. Postema[1], Leah Briscoe[1,2], Chloe Harder[1,3], George R. A. Hancock[4], Lucy D. Guarnieri[5], Tony Eisel[1,6], Kelton Welch[7], Nicole Fisher[8], Christine Johnson[9], Diego Souza[1], Dexter Phillip[1], Rebekah Baquiran[1], Tatiana Sepulveda[1], *Bruno A. S. de Medeiros[1]

1. Field Museum, 1400 S DuSable Lake Shore Drive, Chicago, IL 60605, USA
2. Lake Forest College, 555 N Sheridan Rd, Lake Forest, IL 60045, USA
3. Lane Tech College Prep High School, 2501 W Addison St, Chicago, IL 60618, USA
4. Centre for Ecology and Conservation, University of Exeter, Penryn, TR10 9FE United Kingdom
5. The Ohio State University, Department of Entomology, 2021 Coffey Road, Columbus, Ohio, 43201
6. Roosevelt University, 430 S Michigan Ave, Chicago, IL 60605
7. Ecdysis Foundation, Estelline, SD 57234, USA
8. National Research Collections Australia, Clunis Ross St, Canberra, ACT 2601, Australia
9. American Museum of Natural History, 200 Central Park West at 79th Street, NY 10024, USA

*Corresponding authors

AUTHOR CONTRIBUTIONS

EGP led the pipeline development, imaged drawers, and analyzed data. BASM conceived the pipeline, assisted in the pipeline development, and funded the research. EGP and BASM led the writing of the manuscript. LB, CH, LDG, and TE assisted in the pipeline development, imaged drawers, and analyzed data. LDG, KW, NF, CJ and BASM provided images for training AI models. GRAH developed the workflow for batch color analysis of masked specimens. DS and DP transcribed specimen metadata. RB and TS developed the workflow for batch-uploading data and images to KE EMu. All authors contributed critically to the drafts and gave final approval for publication.

DATA AVAILABILITY STATEMENT:

The source code for DrawerDissect and the identification model *Cicindel-ID* are available at github.com/EGPostema/DrawerDissect and github.com/de-Medeiros-insect-lab/Cicindelinae_ID, respectively. All images and annotations used to train FMNH roboflow models can be found at universe.roboflow.com/field-museum. Training weights for *Cicindel-ID* are available at huggingface.co/brunoasm/eva02_large_patch14_448.Cicindela_ID_FMNH. The ImageJ plugin used to extract color pattern data is available at github.com/GeorgeHancock471/DrawerDissect-ImageJ-Plugins. All data and R code used in this manuscript will be made available via Dryad upon acceptance.

ABSTRACT

1. Natural history museums often curate large collections of pinned insects. These collections represent invaluable records of biodiversity information, ecological patterns and phenotypic variation. A common goal of museums is to create digital versions of these records for curation and research purposes. However, traditional methods of specimen imaging and metadata transcription are prohibitively labor-intensive. High-throughput imaging of entire specimen drawers integrated with computer vision (CV) artificial intelligence (AI) models provides a potential solution.

2. Here we present DrawerDissect, a python-based pipeline for processing whole-drawer photographs, and a workflow to use it in entomological collections. By using custom CV models and large language models for text transcription, DrawerDissect can crop and segment specimens from images and extract metadata from specimen labels. DrawerDissect is flexible, customizable, and modular, allowing rapid downstream analyses of phenotypic features such as color, pattern, shape and size.

3. We used DrawerDissect to digitize the Field Museum's (FMNH's) entire tiger beetle (family Cicindelidae) collection, resulting in 13,484 high-resolution dorsal photographs, masked specimen images, and basic body measurements. All specimens are linked to taxonomic and biogeographic data. We also extracted specific location metadata for 3,648 specimens. We then provide an example of using DrawerDissect outputs with existing color analysis methods in ImageJ to investigate taxonomic and geographic differences in coloration. Finally, we trained an accurate species identification model, *Cicindel-ID*, using ~7,000 masked images of specimens in the genus *Cicindela*.

4. DrawerDissect's novel multi-model AI workflow provides an efficient and reproducible framework that meets the demands of high-throughput digitization of natural history museum collections, unlocking the research potential of large specimen collections.

DATA/CODE FOR PEER REVIEW: All python scripts can be found at github.com/EGPostema/DrawerDissect and github.com/de-Medeiros-insect-lab/Cicindelinae_ID. The ImageJ plugin used to extract color information is available at github.com/GeorgeHancock471/DrawerDissect-ImageJ-Plugins. All R code and datasheets used in this manuscript are available in the file *drawerdissect_analyses.zip*

KEYWORDS: artificial intelligence, computer vision, digitization, high-throughput imaging, insects, image segmentation, machine learning, museum specimens

1 INTRODUCTION

Natural history collections connect the past to the future. Each specimen preserves a unique ecological and evolutionary history, with some specimens maintaining their scientific value for centuries after their initial acquisition. Collections have been leveraged to track the impacts of climate change among plants and animals (Bates et al., 2023; Lister, 2011; Sanders et al., 2023); to elucidate the ecological drivers of diverse morphological traits (Crowell et al., 2024; Holmes et al., 2016); and to estimate the sizes, ranges and compositions of past populations (Davis et al., 2023). Additionally, specimens provide critical genomic data that can be used for species identification and evolutionary studies (de Medeiros et al., 2025; Ruane & Austin, 2017). Despite their immense scientific value, natural history collections face a persistent digitization bottleneck. Metadata digitization and specimen imaging are necessary for collections-based research, but they are time- and labor-intensive, as they traditionally involve handling specimens one at a time. The challenge is magnified for hyper-diverse groups such as insects. In North America, for example, only ~5% of arthropod specimens have data digitally available, and fewer than 2% have been imaged (Cobb et al., 2019). Given the ecological significance of insects (Losey & Vaughan, 2006), and their perilous position in the face of rapid global change (Wagner, 2020; Yang et al., 2021), there is an urgent need to mobilize insect specimen data and images.

Most digitization efforts in the past decades have focused on the transcription of specimen label metadata to standardized formats such as DarwinCore (Wieczorek et al., 2012), while imaging specimens has often been a secondary goal. For example, of the 576,502 specimens in the Field Museum (FMNH)'s insect collection with digital records, only 19,989 have associated images. The traditional focus on metadata is justifiable, as extracting information from specimen images normally requires labor intensive processing steps like outlining or landmarking (van den Berg et al., 2024; Watanabe, 2018). Advances in computer vision (CV) techniques have now made these steps automatable (Borowiec et al., 2022; Lürig, 2022; Weinstein, 2018; Wilson, 2023). With these new tools, images are not just byproducts of digitization with limited uses, but high-quality sources of data themselves.

While image analysis capabilities have advanced dramatically, efficient image acquisition remains a bottleneck in mass digitization workflows. There are many existing solutions for high-throughput imaging in collections, depending on the type of organism and method of preservation (Picturae, https://picturae.com/; Sys et al., 2022; Weeks et al., 2023). For pinned insects, mass-digitization solutions are under development (LightningBug, https://www.lightningbug.tech/prog-orig; Picturae; Steinke et al., 2024), but none are currently in widespread use. No matter how efficient individual specimen handling can be made, a faster solution is to photograph whole drawers (containing hundreds of specimens) at a time (Holovachov et al., 2014). While image resolution for individual specimens used to be a

limitation for this approach, modern imaging systems are capable of producing high-quality whole-drawer images (Fig. 1a).

High-throughput imaging is fast, but whole-drawer images create another challenge: extracting individual-level data. A possible solution to this processing bottleneck has been clear for over 10 years. Holovachov et al. (2014) presciently ask, "[c]an specimens in the [whole-drawer] image be analyzed and identified using a computer algorithm and machine learning?" Looking at the recent explosion of CV-based pipelines for specimen digitization, we can largely respond "yes" to this question in 2025 (Stenhouse et al., 2025). For example, LeafMachine2 uses a series of computer vision models to extract phenotypic data and transcribe labels from herbarium specimens (Weaver & Smith, 2023); Skelevision segments and measures bones from batch-imaged bird skeletons (Weeks et al., 2023); CollembolAI detects and classifies small invertebrates in photographs of soil samples (Sys et al., 2022); and DiversityScanner combines robotic sorting and imaging of bulk insect samples with a classification model to predict family-level identity (Wührl et al., 2022). Applying the same concept to pinned insect collections, we wanted to create an AI-driven, user-friendly pipeline for processing whole-drawer images, with two key outputs:

(1) Individual specimen images, linked to taxonomic identity and, when possible, sampling locations extracted from label text.
(2) Specimen images with censored backgrounds and pins ("masked specimens") suitable for downstream morphological analysis.

We call this multi-model pipeline "DrawerDissect." DrawerDissect relies on an AI service, Roboflow (Dwyer et al., 2014), for object detection and segmentation models. For text transcription, we use Anthropic's large language model (LLM), Claude (https://console.anthropic.com/). We designed DrawerDissect to be flexible, with built-in toggles for different desired outputs; customizable, with the option to swap between different models, update existing models, and edit LLM prompts; and modular, with steps that can be run independently or combined in unique workflows.

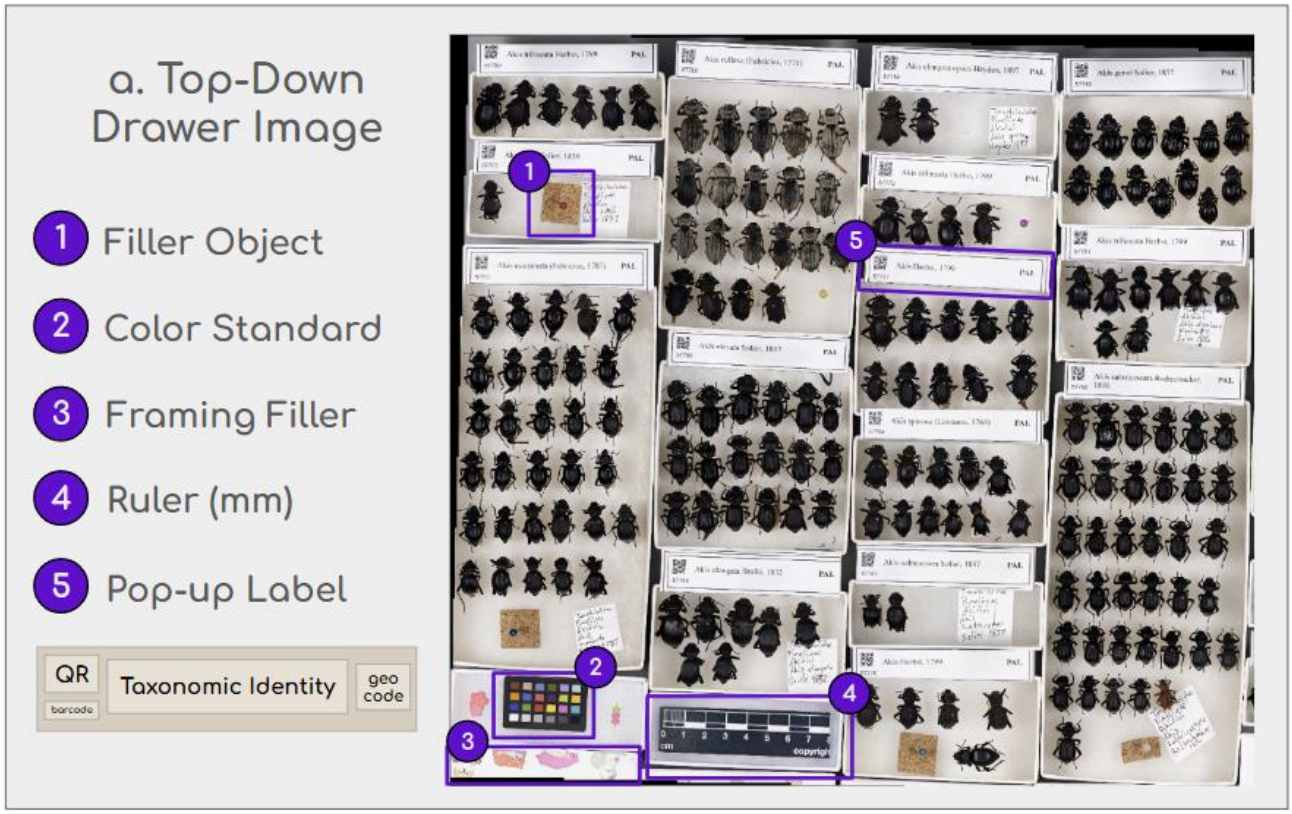

Figure 1. (a) A top-down whole-drawer image taken by the GIGAMacro Magnify2, with key objects labeled. (b) Standard FMNH unit tray header label. (c) A new pop-up unit tray label, here folded flat when stored.

2 MATERIALS AND METHOD

Our workflow involves three steps: (1) imaging drawers, (2) running images through DrawerDissect, and (3) optional post-curation. First, we describe our imaging set-up using a GIGAMacro Magnify2 system (Four Chambers Studio, https://gigamacro.com/). We then give a broad overview of DrawerDissect in terms of installation, configuration, and model training/selection. Finally, we describe the steps that occur during a standard DrawerDissect run. A streamlined user guide with quick-start instructions can be found at the DrawerDissect github: github.com/EGPostema/DrawerDissect.

To demonstrate the utility of DrawerDissect we photographed and processed the FMNH's entire 44-drawer collection of pinned tiger beetles (Coleoptera: Cicindelidae; Duran & Gough, 2020). Tiger beetles are colorful insects, often with intricately patterned and iridescent elytra. The ecology and evolution of this group's coloration has been the focus of much research, particularly with respect to background-matching, thermal physiology, and morphological convergence (French et al., 2021; Pearson & Vogler, 2001; Schultz & Bernard, 1989; Schultz & Hadley, 1987; Yamamoto & Sota, 2020). To test DrawerDissect's ability to integrate with existing batch color analysis pipelines (Hancock et al., 2025), we use the masked images generated by DrawerDissect to investigate geographic and climatic patterns of color diversity in two subspecies of a well-represented species in the collection, *Cicindela formosa*. Finally, we used over 7,000 masked images generated by DrawerDissect to train a taxonomic identification model, *Cicindel-ID,* for the speciose genus *Cicindela*.

2.1 Whole-Drawer Imaging

DrawerDissect was designed with images produced by a GIGAMacro Magnify2 imaging system, and uses FMNH conventions for insect drawers where specimens are grouped into trays ("unit trays") based on shared taxonomic identity and geographical origin (Fig. 1b). FMNH unit trays have header labels that display this information, typically affixed to the upper tray wall (Fig. 1b). To make the tray-level data visible for imaging, we developed pop-up headers that display the same information; the pop-up portion is visible during imaging (Fig. 1a), but folds down flat when trays are stored (Fig. 1c). The only pre-curation step needed for drawer imaging is to generate an inventory of unit trays for each drawer, replace the old labels with the new ones, and then arrange all unit trays for imaging on the GIGAMacro platform. For further imaging details (e.g. camera specifications, file formats), see supplemental materials Section S1.

While we describe our digitization methods based on this standard, DrawerDissect can be adapted to various configurations. At minimum, running the pipeline requires a reasonably clear, high-quality image of an insect drawer, achievable with a number of different imaging systems (Holovachov et al., 2014).

2.2 Overview of DrawerDissect

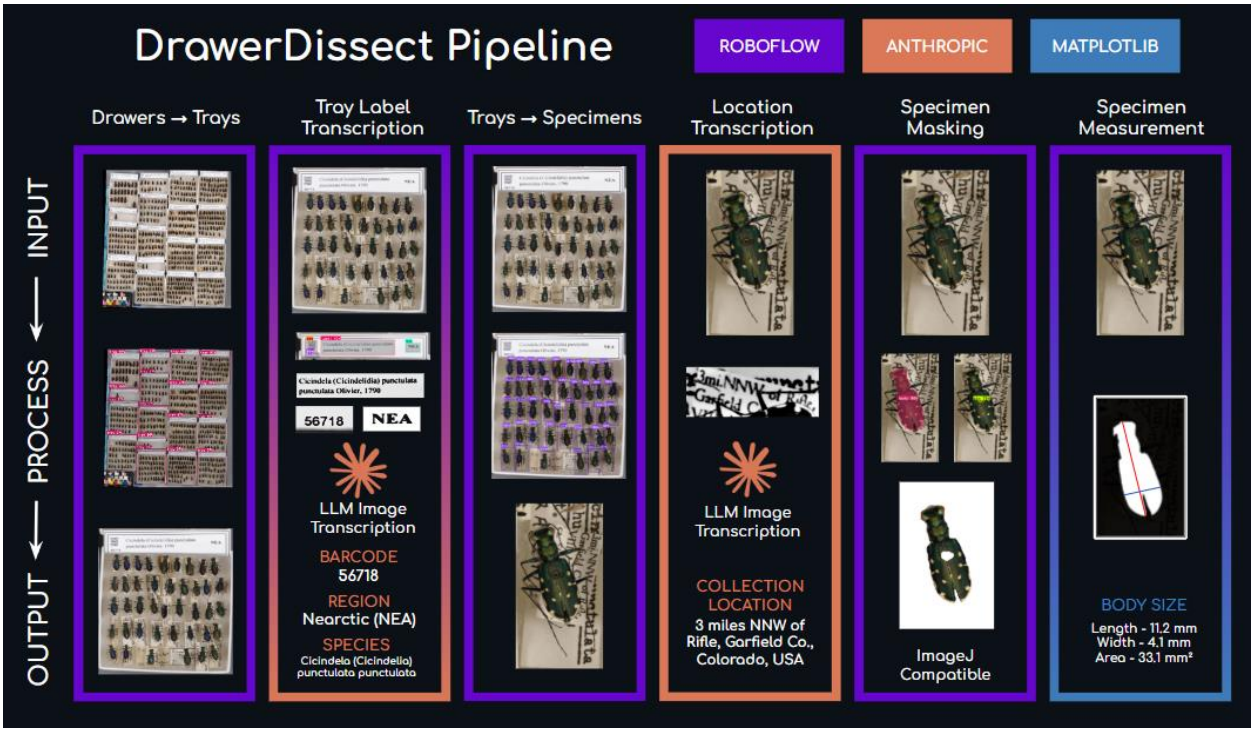

Figure 2. Overview of DrawerDissect, with key python packages highlighted.

DrawerDissect is a Python-based pipeline available from GitHub, including documentation on installation and usage at github.com/EGPostema/DrawerDissect. The repository includes test data to familiarize new users with the pipeline's steps and organizational structure. We aimed to make the pipeline accessible to users who are not expert programmers. Only a rudimentary familiarity with their computer's command-line interface is required to run the script. Overall, DrawerDissect can be installed and run in four steps:

1. Create an environment for the program to run (5 lines of command-line code)
2. Add whole-drawer images to a folder
3. Set up API keys from Roboflow and Anthropic (up to 2 lines of command-line code)
4. Run the full pipeline with a single command

2.2.1 Models Used in DrawerDissect

DrawerDissect relies on Roboflow models for all steps requiring vision-only models, such as object detection and segmentation. We found that general foundation models perform poorly for our tasks, and therefore trained our own specialized models for detection and segmentation. Reading labels using LLMs, however, is sufficiently accurate without specialized fine-tuning. Training the detection and segmentation models requires a set of images that have been labeled

with the desired output. We use Roboflow's labeling tools to make this step faster. Once annotated, we then split the images into training, validation, and test sets at a standard ratio of 70:20:10. We also employ image augmentations to increase model generalizability (Borowiec et al., 2022).

We trained our models using Roboflow servers. Training can take anywhere between a few minutes and several hours depending on the size and complexity of the model; 4-5 hours was typical for our largest models with datasets of >3,000 images. When the model is finished training, Roboflow automatically reports the model's precision and recall, as well as rates of false positives/negatives (Table 1). We trained and have provided public access to all the models used in our pipeline at the time of manuscript submission using the Roboflow Universe platform. See Table 1 for model-specific performance, and supplemental Table S1 for the taxonomic groups included in each of our six public models.

TABLE 1: Model Composition and Performance

model id (*version)	# of images (**train, valid, test)	precision	recall	false positive rate	false negative rate
trayfinder-base (5)	(180, 12, 6)	99.6%	99.1%	0.02%	0.01%
trayfinder-popup (17)	(162, 11, 6)	100.0%	100.0%	0.01%	0.00%
labelfinder (7)	(3555, 201, 103)	96.0%	99.0%	0.02%	0.01%
bugfinder-kdn9e (13)	(4235, 238, 120)	97.1%	96.7%	0.02%	0.02%
bugmasker-all (5)	(2035, 113, 62)	99.9%	97.3%	0.02%	0.02%
pinmasker (6)	(2210, 126, 67)	91.0%	93.0%	0.08%	0.07%

*The version used to produce the results in Section 3.1.

**Includes augmented images, which increases the size of the training split x5.

By default, users are set up with public FMNH models via Roboflow for image processing, while all transcription steps are done with an Anthropic LLM. We use these subscription-based services rather than open source models because they are more accessible to a broader user base that may lack the coding experience or computing infrastructure to run AI models locally (Heron et al., 2013). Running AI models locally requires powerful servers with advanced GPUs, which can cost tens of thousands of dollars to set up and high energy costs to run. Advanced users experienced in AI coding, or large institutional teams, could modify DrawerDissect code to use

local processing instead of API calls. For open-source model recommendations, see supplemental materials Section S2.

2.3 Running DrawerDissect

DrawerDissect consists of a series of steps (Fig. 2) that can be run automatically, in sequence, with a single command: '*python process_images.py all*'. The github documentation includes a full list of command-line steps, optional arguments for running specific step combinations, and how to process specific drawer(s).

2.3.1 Object Detection and Segmentation

DrawerDissect uses Roboflow models (3.0 Object Detection and 3.0 Instance Segmentation; Table 1) to find and outline objects in images (Fig. 3-4), described below. An asterisk marks models used by default in the current version of DrawerDissect.

trayfinder-base *
This model detects unit trays from drawer images (Fig. 3a).

trayfinder-popup
This model detects unit trays (with pop-up labels) from drawer images. This is a FMNH-specific model alternative to *trayfinder-base*.

labelfinder *
This model can detect the locations of taxonomic information, barcodes, and geocodes. This model can detect both pop-up label text (Fig. 3b) and taxonomic text within trays (Fig. 3c).

bugfinder-kdn9 *
This model detects pinned insects (Fig. 4a).

bugmasker-all *
This model outlines the main body of insects, excluding legs and antennae (Fig. 4b).

pinmasker *
This model outlines the specimen pin (Fig. 4b), if present.

Figure 3. (a) Trays in a drawer detected by *trayfinder-base*, (b) *labelfinder* detects multiple text classes for a pop-up label, (c) *labelfinder* detects a handwritten species name in a tray.

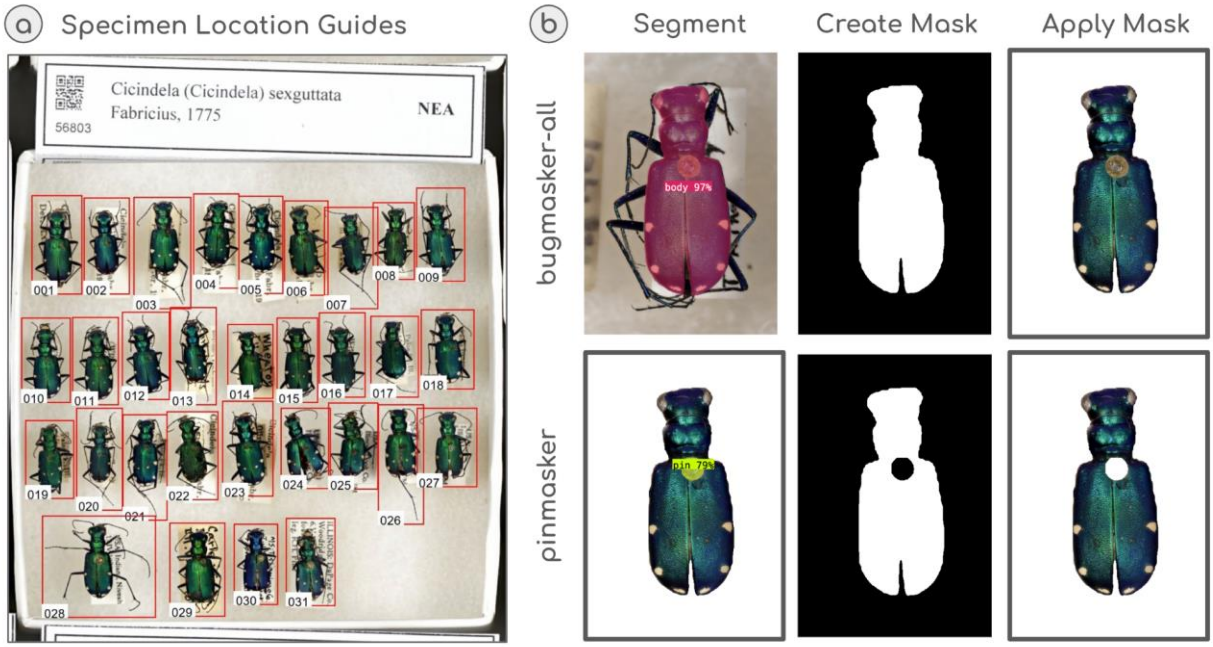

Figure 4. (a) Tray guides produced from *bugfinder-kdn9* coordinates. Specimens are automatically numbered from left to right, top to bottom. (b) Roboflow coordinates from *bugmasker-all* and *pinmasker* are translated into binary masks by DrawerDissect to create the final masked specimen images.

All CV models output coordinates in a JSON file that are used by DrawerDissect to crop, measure, or mask images. Due to file size and format constraints, all TIFs are temporarily converted to JPGs before processing, and large images are resized to fit within a 1000x1000px

square. During cropping, the coordinates from reduced-size images are always rescaled to the original, full-sized image.

2.3.2 Masking and Measuring Specimens

To create transparent versions of the masked specimens, coordinates from *bugmasker-all* and *pinmasker* are used to generate binary masks: black and white images (PNGs) where all background pixels are black and all body pixels are white (Fig. 4b). We then run the masks through a filtering step that fills in any partial segmentations. To create the final masked image, the white pixels clip out the full-color specimen, while the black pixels become transparent.

For each specimen mask, we measure len1, which is the greatest distance between any two points on the outline, and len2, the maximum distance perpendicular to len1. These measurements are good estimates for length and width, though the script is intentionally agnostic to orientation. We also calculate the area, in pixels squared, of the body. DrawerDissect can output a map of len1 and len2 for each specimen mask if desired.

2.3.3 LLM-Based Transcription

To transcribe different kinds of text from images, we use Anthropic's API with customizable prompts. For the results in Section 3.1, we used the model *claude-3-7-sonnet-20250219.* Before transcription, all images are preprocessed to enhance text visibility by converting to grayscale and increasing contrast. Both handwritten and typed material can be transcribed (Fig. 3b-c). We review all AI-generated transcriptions manually.

Unit tray label text is transcribed from cropped images generated by *labelfinder* (Fig. 3). We have three separate prompts for each type of transcription, which are tailored to how these different types of information are structured. These prompts can be edited by users for their specific label structure in the *config file*, though our default settings will work for most cases.

In many specimen images, partial text from the top label is visible. We tested a novel method of metadata extraction based solely on this fragmentary information. The outputs of this method are transcriptions of the verbatim text, an initial location estimate, the model's final location estimate, and an assessment of the estimation's quality (Fig. 5). The quality ranks are defined as: (a) verified, when the location estimate is logical given the verbatim text, (b) unreliable, when there is not sufficiently specific verbatim text to justify the location estimate, and (c) unknown, when there was not enough verbatim text to result in a location estimate.

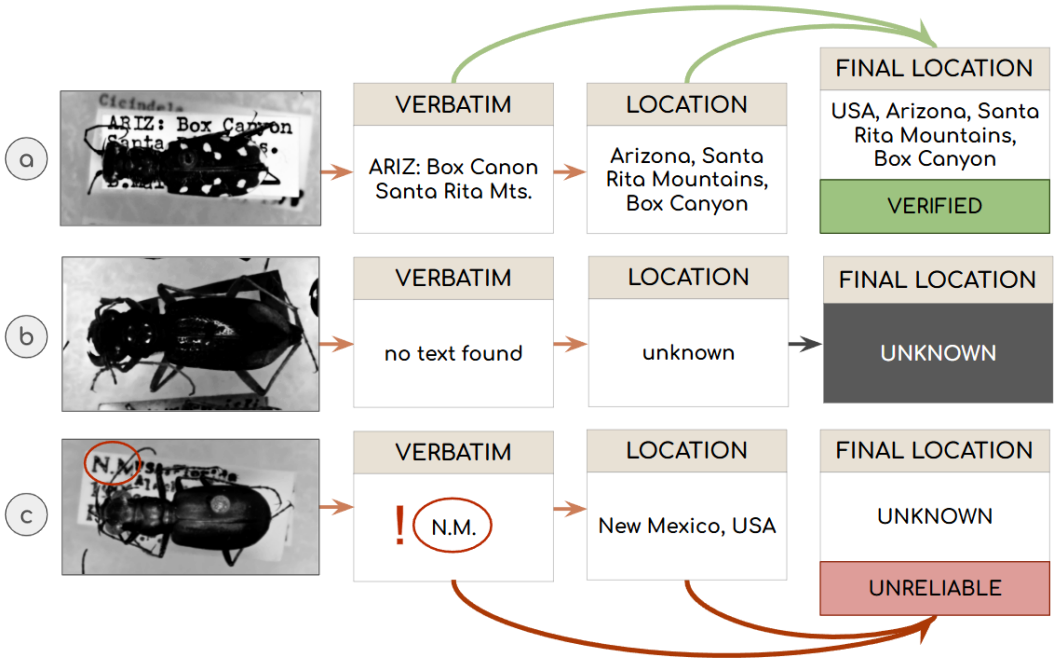

Figure 5. (a) A successful transcription, resulting in a verified location, despite a typo in the verbatim text. (b) An unknown location and final location from a specimen with no visible label. (c) An unreliable location due to the verbatim text's lack of geographic specificity.

2.3.4 Databasing

The final step of the pipeline compiles a time-stamped folder with data summaries at the drawer, tray, and specimen level for a given run. To upload the specimen images and their metadata to KE EMu, we print specimen labels with FMNH numbers and an associated QR code. We apply these labels to each pinned specimen, using the tray guides (Fig. 4b) to match the DrawerDissect specimen image (e.g. spec_001) to the FMNH number. Lastly, we run a Python script to reorganize data to fit the metadata structure expected by KE EMu, and batch-import the images plus metadata.

2.4 Use Case 1: Batch Color Analysis of *Cicindela formosa*

To test the utility of DrawerDissect's masked image outputs, we targeted a specimen- and data-rich taxa within the FMNH tiger beetle collection: the big sand tiger beetle (*Cicindela formosa*). *C. formosa* exhibits a variety of color patterns across its range, with multiple purported subspecies (Pearson et al., 2006). There are clear regional variations in this species' maculations, which are often used to justify subspecies delineations (French et al., 2021; Gaumer, 1977; Pearson et al., 2006). The FMNH collection includes more than 300 *C. formosa* collected across North America, with two semi-sympatric subspecies: *C. f. formosa* and *C. f. generosa*. Due to a previous digitizing effort, the FMNH has records of the collection locations for all specimens of

these taxa. Given the historical use of color and pattern characteristics to identify *C. f. formosa* versus *C. f. generosa*, we might expect the two subspecies to significantly differ in appearance, even in regions where they overlap. However, color differences between the subspecies could also be explained by local adaptation. In this case, we might expect sympatric individuals to be more phenotypically similar, and for color to vary along environmental gradients known to impact animal coloration (Postema et al., 2023). To analyze phenotypic variation between and among the two subspecies, we used an existing ImageJ batch-analysis pipeline (see Hancock et al., 2025) to quantitatively measure CIELab color and pattern geometry in the masked *C. formosa* specimens (n = 374). For a full description of our phenotyping method and statistical test, see supplemental materials Section S3.1-2.

2.5 Use Case 2: Training a Classification Model to Identify *Cicindela* Specimens

Batch-imaging and image processing tools like DrawerDissect can supply taxonomic identification models with the large amounts of training data they require (Spiesman et al., 2021; Sun et al., 2021; Truong & Van der Wal, 2024; Welch & Lundgren, 2024). In our collection of tiger beetles, the majority of specimens (7000+) belong to the genus *Cicindela*, representing 239 unique species and subspecies. There are also a small number of specimens in our collection (12) that were identified as *Cicindela* but lacked a species identification. As a proof of concept, here we train a model to identify species and subspecies of *Cicindela* based on the masked images produced with DrawerDissect, and apply this model to the unidentified specimens. To train and evaluate this model, we split all identified specimens into training, validation and test sets using a 70:20:10 ratio for each species or subspecies. Single-image taxa were only included in the training set. Our full training method (model architecture, augmentations, model weights, etc.) is described in supplemental materials Section S4. We evaluated the final model using the test set, with a confidence threshold of 0.5. We also predicted labels for the 12 unknown samples and compared results to manual identifications (Table S2) using guides to the Cicindelidae (Pearson et al., 2006; Shiyake, 2017; Trautner & Geigenmüller, 1987)

3 RESULTS

3.1 Digitizing the FMNH Tiger Beetle Collection

With DrawerDissect, we were able to produce a complete inventory of the FMNH tiger beetle collection. We processed the collection into 13,496 separate images of intact specimens obtained from 44 drawers of pinned tiger beetles from around the world (Fig. 7). All specimen photos are linked to taxonomic and biogeographic metadata. The majority of these specimens (13,484) were successfully masked and measured (length, width, and area, in mm). Over a quarter (3,627) of the specimen images are associated with more specific collection locations, as a result of automated transcription via DrawerDissect, manual transcriptions for one specimen per tray, and

additional metadata from a previous databasing effort (see Section 2.4). Our collection contains 56 unique genera and 663 unique species of tiger beetle. This represents nearly a third of the ~2,300 known species globally (Gough et al., 2019). FMNH tiger beetles were collected from at least 1,002 unique locations across 91 countries, and all continents excluding Antarctica (Fig. 7).

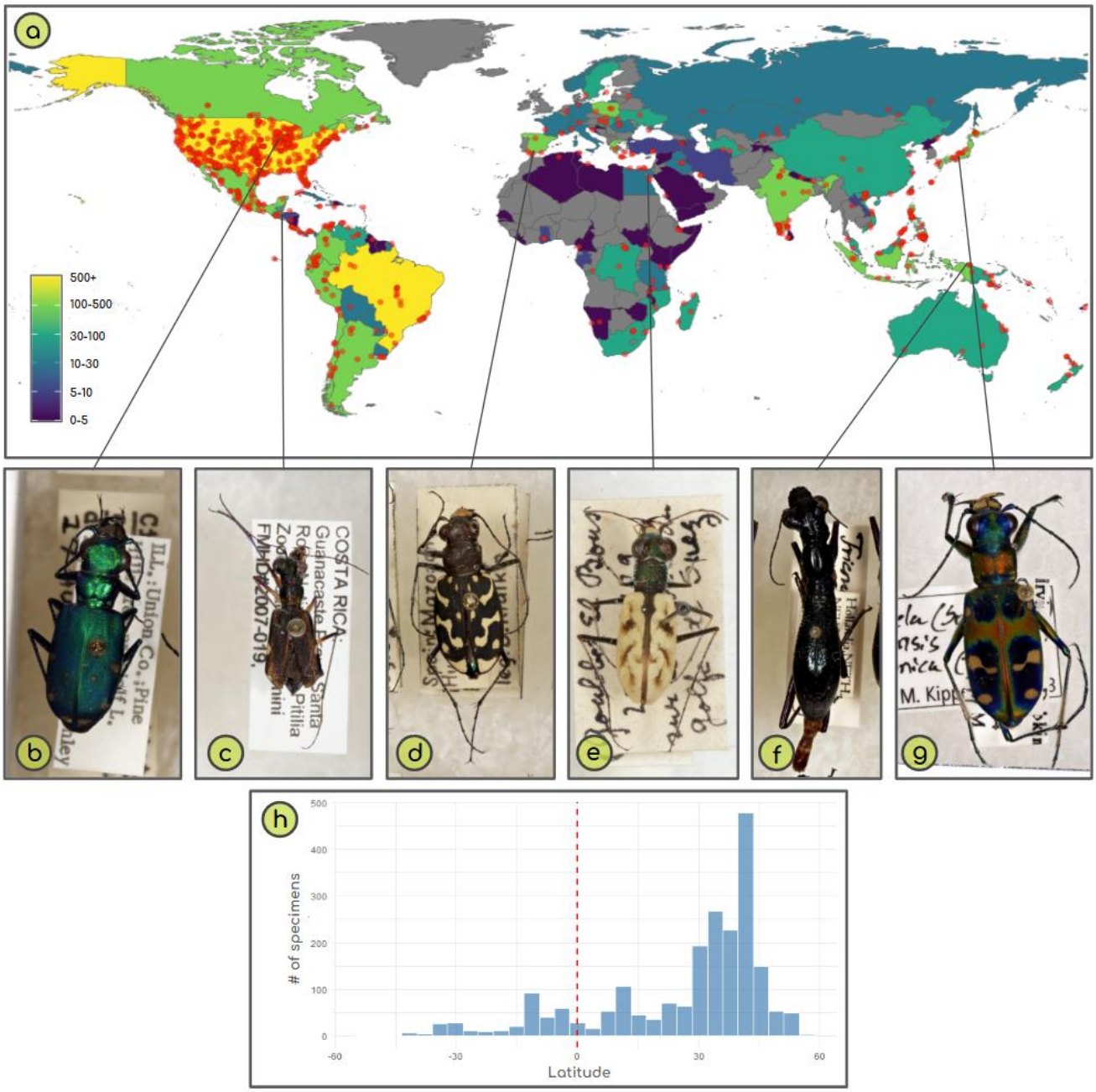

Figure 6. (a) A world map showing the approximate number of FMNH specimens per country. The per country counts are estimated from the number of specimens in a given tray and the countries of specimens from the same tray with known countries. Collection sites with known coordinates are marked with red points. Middle: (b) *Cicindela sexgutatta* from Pine Hill, IL, USA. (c) *Odontocheila iodopleura* from the Pitilla Zoological Station, Costa Rica. (d) *Lophyra flexuosa* from Mazagón, Spain. (e) *Hypaetha*

singularis from Ghoubbet-El-Bous, on the coast of the Gulf of Suez in Egypt. (f); *Tricondyla aptera aptera* from Jayapura, Indonesia; (g) *Cicindela chinensis japonica* from Ōdai, Japan. (h) Specimen distribution by latitude (n = 2159).

3.1.1 Unit Tray Cropping and Transcription

Out of 941 trays, *trayfinder-popup* made no detection errors. 36 trays (3.9%) required manual transcription of one or more pieces of tray label text. For the majority of trays, missed transcriptions occured when *labelfinder* failed to detect text. In four instances, label information was detected, but the bounding box did not fully encompass the text. We observed only two instances of taxonomic transcription typos: *Manticora* misspelled as *Mantichora*, and *Distipsidera* misspelled as *Distintipsidera*. For barcodes, there were three instances where the numbers 5 and 6 were erroneously switched. For geocodes, there were 9 transcription errors that resulted in the incorrect biogeographical realm. All tray-level text was typed, not handwritten.

3.1.2 Specimen Detection and Masking

Overall, *bugfinder-kdn9e* made 13,662 specimen detections. 153 (1.1%) of these were false positives, where non-specimens were erroneously detected. This is a higher rate of false positives than the model's roboflow test set (0.02%; Table 1). No specimens were missed by the model, meaning the rate of false negatives was 0% (lower than the roboflow-calculated rate of 0.02%; Table 1).

Out of the 13,662 cropped specimen photos generated from *bugfinder-kdn9e*, *bugmasker-all* found masks for 13,501 images (Fig. 8). Of the 161 images where no mask was found, 154 were true negatives, meaning they did not contain a specimen, or contained a damaged specimen. Therefore, we used masking to mitigate false positive errors in specimen detection. Only 7 were false negatives, where no mask was found despite the image containing a complete, intact specimen. In the masked set, we found 12 false positives, where a mask was applied to a non-specimen (3) or heavily damaged specimen (9). In 5 cases a mask did not fully outline an intact specimen. Our actual rates of false positives (0.09%) and false negatives (0.05%) for our set of tiger beetles were below the rates reported for the model by Roboflow (Table 1).

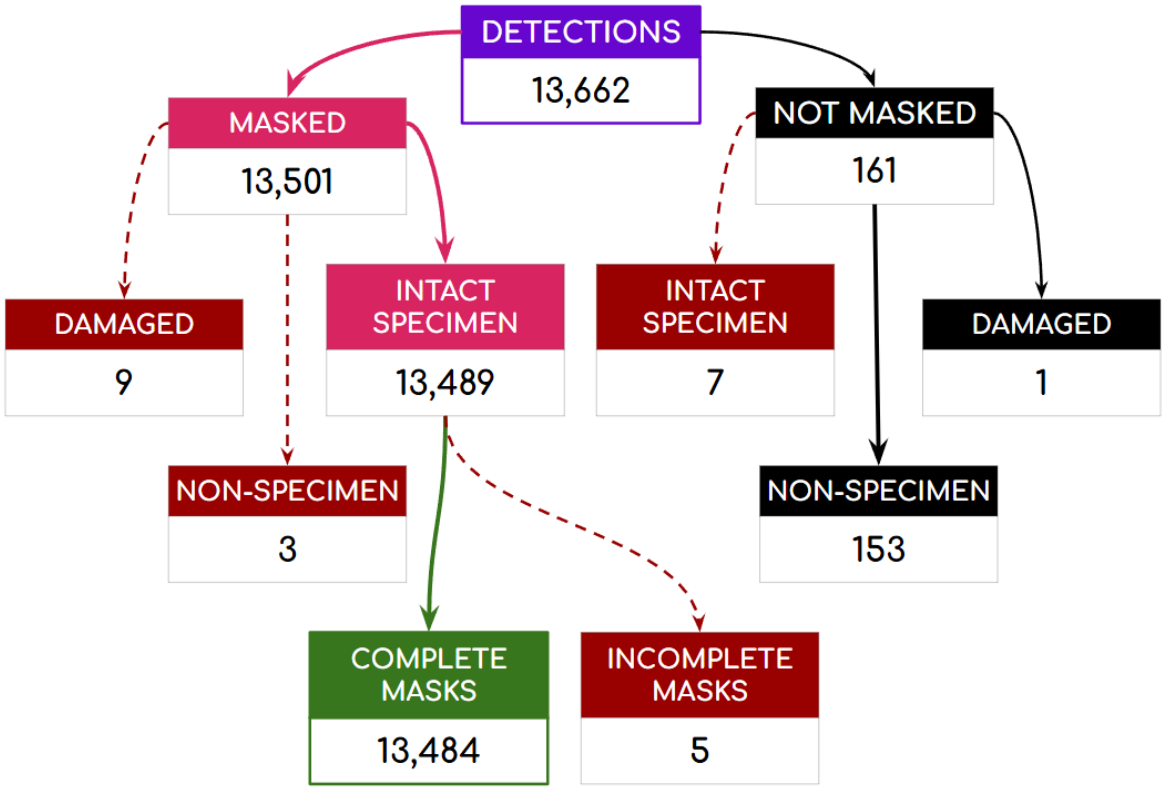

Figure 7. Flowchart showing the outcomes of DrawerDissect masking. Solid lines = correct responses, dotted lines = incorrect responses.

3.1.3 Specimen-level Label Transcription

For the majority of specimens (80.3%), no location could be determined from the image alone (Fig. 6). In total we were able to reconstruct locations for 2,475 specimens (18.2%). This includes correct estimates from the verbatim text that were judged as verified by the LLM (9.4% of images), and cases where there was enough context from the verbatim text that a location could be determined by a human reviewer even if the final location was judged as unknown/unreliable by the LLM (8.8% of images). Bad location estimates that were judged as reliable by the LLM (false positives) were infrequent, at 1.6% of all images. Of the 425 unique collection locations extracted from the specimen images by the LLM, most (69.6%) could be linked to a set of centroid coordinates (with up to a 20mi radius of uncertainty) on Google Maps.

3.1.4 Size Measurements

We were able to measure all masked specimens (13,484; Fig. 8), generating length, width, and area in mm/mm2. To assess the accuracy of the digital measurements, we compared them against a random subset of specimens (n = 79) that we hand-measured using calipers (Fig. 9a-b). We

found that the digital and manual measurements were highly correlated (length: $R^2 = 0.989$, width: $R^2 = 0.961$). Body area, length, and width varied both within and across genera (Fig. 9c).

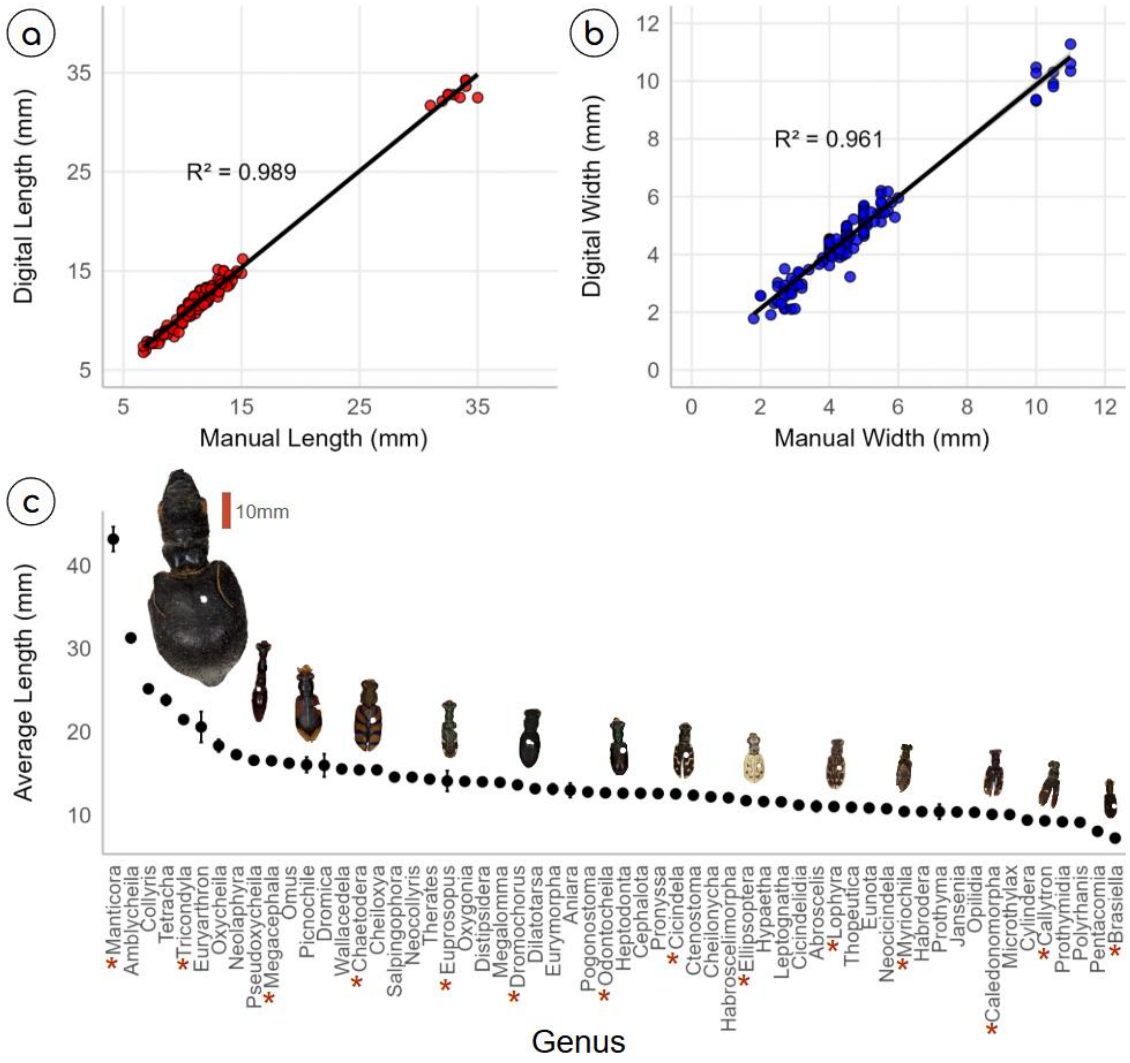

Figure 8. (a) The relationship between manual and digital length for a set of randomly selected specimens (n = 143), both in mm. (b) The relationship between manual and digital width for the same set of specimens. (c) The average body length, in mm, +/- SE of each genera in the FMNH collection (n = 13,484). *Indicates genus with an associated icon generated by DrawerDissect.

3.1.5 Total Time and Cost

Imaging all 44 drawers took ~70 hours, including moving drawers, inventorying unit tray labels, and printing and replacing header labels. On a Windows computer with an AMD RyzenTM 7 Processor 7800X3D CPU and 32GB of RAM, DrawerDissect is able to fully process an 8GB drawer image in 16 minutes. For 44 drawers it took ~10 hours to fully process, with the benefit of parallel processing. We spent an additional 40 hours on data validation, and another 35 hours

to manually transcribe location information for one specimen per unit tray - an optional step to guide our manual review of DrawerDissect's location estimates. In total, the full workflow (imaging, processing, and data-cleaning) for the tiger beetle collection took ~155 hours. This translates to ~42 seconds per specimen. This is significantly faster than manually photographing, outlining, measuring, and transcribing partial metadata for 13,484 specimens.

The costs for whole-drawer imaging and DrawerDissect digitization are human labor (for pre-curation, equipment operation, and data validation), the up-front cost of the imaging system (GIGAMacro Magnify2 with telecentric lens: ~$60k), annual software license fees (Adobe Lightroom, PTGui), and AI inference fees (Table 2). Anthropic API does not have a free tier, but Roboflow includes 360,000 annual inferences for users with free accounts.

TABLE 2: Time and Cost Estimates

Step	cost per specimen*	cost for all Cicindelidae*	processing time per specimen	processing time for all Cicindelidae
Transcription	$0.00757	~$100	1.6 seconds	6 hours
CV Image Processing	2.1 inferences	29,228 inferences	1.1 seconds	4 hours
DrawerDissect SUBTOTAL	*$0.00757 + 2.1 inferences*	*~$100 + 29,228 inferences*	*2.7 seconds*	*10 hours*
Pre-Curation, Imaging**	$0.31	$4,235.00	19 seconds	70 hours
Data Cleaning**	$0.34	$4,537.50	20 seconds	75 hours
TOTAL	**$0.66 + 2.1 inferences**	**$8,872.5 + 29,228 inferences**	**41.7 seconds**	**155 hours**

*Does not include the up-front cost of the imaging rig or image-processing software licenses.
**Costs for these steps were calculated from the hourly wages of one postdoc ($26/hr) and two paid interns ($17/hr and $17.50/hr).

3.2 <u>Use Case 1:</u> Phenotypic Variation in Subspecies of *Cicindela formosa*

Overall, we found that the two *C. formosa* subspecies are significantly different in appearance (Fig. 10a-b), with *C. f. formosa* being significantly redder than *C. f. generosa*. However, this distinction decreases in the overlapping parts of their ranges; sympatric individuals exhibit lighter, less red elytra regardless of subspecies (Fig. 10a). In these areas, color features may be less useful for distinguishing subspecies. Further genetic analysis would be required to determine

whether the sympatric similarity is due to hybridization and/or convergent evolution (French et al., 2021). In terms of climatic factors, there was a significant pattern of increased redness in hotter, more arid locations. This was mainly driven by the large populations of *C. f. formosa* in the western and southwestern United States (Fig. 10a). A thermoregulatory (heat-reducing) function of increased redness is unlikely, as metallic red portions of *C. f. formosa* elytra show increased IR absorbance compared to white portions (Schultz & Hadley, 1987). Elytral color may instead be driven by local background-matching camouflage, as in other tiger beetle species (Yamamoto & Sota, 2020). For the full statistical analysis of these differences, see supplemental materials Section S3.3.

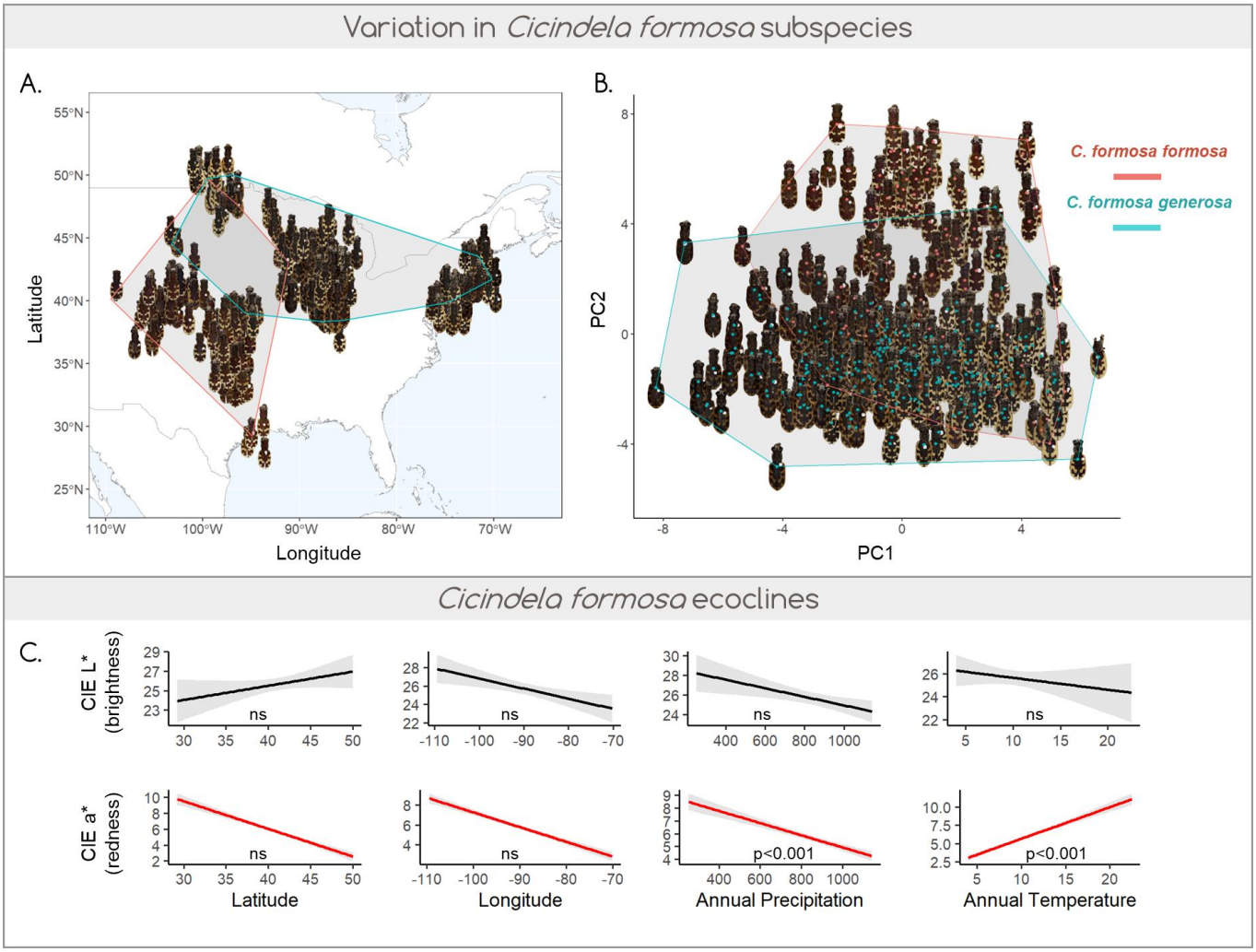

Figure 9. (a) A map of the collection locations of *C. f. formosa* and *generosa*. (b) Specimens of each subspecies plotted by PC1 (darker to lighter) and PC2 (greener to redder). (c) Relationships between phenotypic and abiotic variables for all *C. formosa* specimens (n=374). Significant relationships are labeled within plots; ns = non-significant.

3.3 Use Case 2: Results of *Cicindel-ID* Specimen Identification

Species and subspecies of *Cicindela* could be identified with high accuracy even for modest sample sizes in the training set (supplemental materials, Fig. S2a). The average F1 score, which balances precision and recall, increases with availability of samples in the training set (Fig. S2b), with high precision throughout (Fig. S2c). Specifically, we observed 97.0% precision and 96.4% recall for species when averaged over specimens (Fig. S2c). For subspecies, these numbers were somewhat smaller, at 85.0% precision and recall averaging over specimens (Fig. S2c). This means that most specimens can be predicted with very high accuracy, while the model may fail to recognize rare species and subspecies. Using *Cicindel-ID* to identify unknown samples reinforces this pattern. Among the 12 unidentified specimens, 10 were common species that *Cicindel-ID* predicted correctly (supplemental Table S2). Of the two samples that resulted in no prediction, one was a subspecies not in the FMNH collection and therefore not present in the training set (*C. transbaicalica japanensis*). The model correctly failed to make a prediction in this case. We identified the other unknown sample as *C. sylvicola*, which included 25 specimens in the training set. The trained model has a moderate accuracy in identifying this taxon (75% precision, 60% recall), so this result was a false negative.

4 DISCUSSION

DrawerDissect is the first usage of general-purpose equipment for mass digitization of whole pinned insect drawers yielding specimen-level images, masks, and data. Using DrawerDissect, we were able to successfully image, segment, and extract data (taxonomic, morphometric, and geographic) from 13,484 pinned insect specimens in about three weeks. This is a major increase in efficiency: in the last 20 years, fewer than 20,000 insect specimens were imaged in the entire FMNH pinned collection. Batch-imaging combined with AI processing is the clear way forward for large-scale digitization projects in natural history collections (Stenhouse et al., 2025; Weaver & Smith, 2023; Weeks et al., 2023).

For insects, few mainstream options for mass digitization exist. There are ongoing developments for bulk digitization of fluid-preserved insect collections based on bespoke imaging and robotic equipment (Wührl et al., 2022). Picturae (https://picturae.com/services/entomology-digitization/) also offers a solution for pinned insects including metadata digitization that can purportedly reach a rate of 5,000 specimens daily. However, it involves custom robotic equipment, a large installation footprint, and individual handling of each specimen. InSelect (Hudson et al., 2015) is another option for processing whole-drawer images into single-specimen images, but it lacks customizable models that perform well on variably-positioned and overlapping specimens. Additionally, InSelect does not include tools for segmentation or text transcription. Currently, DrawerDissect is the only bulk insect digitization method that (a) works with general-use

imaging equipment, (b) offers built-in, adjustable image-processing and text transcription steps, and (c) does not require handling individual specimens..

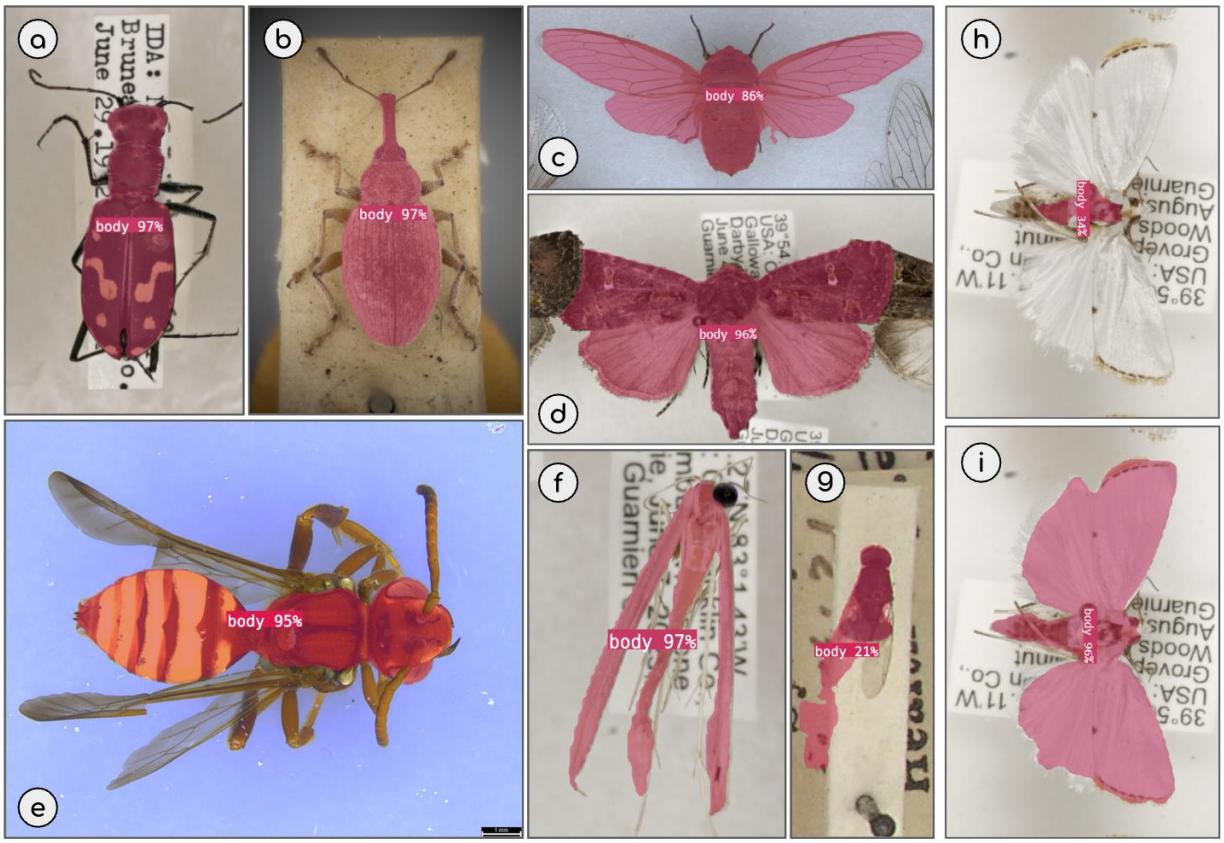

Figure 10. Example *bugmasker-all* predictions for various taxa. The pink overlay shows the location of the specimen body and model confidence is shown as a percent. (a) A tiger beetle from the FMNH collection, photo by EGP/LB/CH/BM, (b) a weevil, photo by BM, (c) a cicada from the Australian National Insect Collection (ANIC), (d) a noctuid moth photo by LDG/TE (e) a wasp, photo by KW, (f) a plume moth, photo by LDG/TE, (g) a fly (Diptera) from the ANIC. (h) Prediction made by *bugmasker-all* (v5). (i) Adding targeted training data significantly improved *bugmasker-all* (v8)'s performance.

Using multiple CV models (Table 1), DrawerDissect produces high-quality dorsal images of specimens that exclude the background and insect pin (Fig. 4b). This automates a time-consuming step that is necessary for most morphological analyses (Correa-Carmona, 2025; Van Belleghem et al., 2018; Weaver & Smith, 2023; Weeks et al., 2023; Weller et al., 2024). Notably, DrawerDissect's masking step (model id: *bugmasker-all*) is able to segment pinned insects of various shapes, sizes, colors, orientations, and degrees of overlap with other specimens, and does not require a standardized background (Fig. 12). Additionally, retraining existing models to identify new taxa (or to improve predictions for underperforming groups; Fig. 12h-i) is simple. For example, to train *bugfinder-all* to successfully detect moths, we only had to annotate ~100 new images. We anticipate that the models we or others develop for insect

segmentation (Correa-Carmona, 2025; Mráz et al., 2023) will only improve in taxonomic and morphological generalizability over time.

One limitation of our approach is that the metadata collected is fragmentary. For example, DrawerDissect was only able to automatically transcribe collection locations for ~18% of our tiger beetles, and these locations were often partial. However, there are ongoing efforts to optimize the transcription of museum labels in terms of prompt engineering, model selection, and cost effectiveness (Herbst et al., 2025). We also demonstrate that this kind of data can still have research uses. As insect collections tend to have more metadata records than specimen images (Cobb et al., 2019), a possible strategy for batch-imaging is to target previously digitized groups that lack images, as we did in the case of *Cicindela formosa*. The combination of AI-driven phenotype extraction and traditional label transcription has the potential to produce rich datasets to investigate the drivers of insect morphological diversity. For example, we were able to extract accurate length and width measurements from all 13,484 masked specimens (Fig. 9), as well as detailed color and pattern data from a subset of big sand tiger beetles (*C. f. formosa* and *C. f. generosa*; Fig. 10a-b). Using the latter dataset, combined with existing location records and climate data, we found significant differences both between subspecies and along environmental clines (Fig. 10b-c).

Species identification models (Borowiec et al., 2022; Spiesman et al., 2021; Sun et al., 2021; Welch & Lundgren, 2024) are another promising use case for DrawerDissect's masking step, as we demonstrate with the *Cicindel-ID* (Fig. 11). *Cicindel-ID* was able to achieve high precision (97%) and recall (96.4%) for species in the genus *Cicindela*. Masked museum specimens offer both quality and quantity as training data for species ID models: they are standardized, censored to avoid shortcut-learning (Geirhos et al., 2020; Weaver & Smith, 2023), taxonomically diverse, and intraspecifically varied. Backgroundless specimens can also be composited onto naturalistic backgrounds to simulate in situ photographs (Sun et al., 2021), potentially expanding training sets for field image classification.

AI, in the form of both tailored CV models and LLMs, is changing the scale and nature of biological research (Borowiec et al., 2022). Chaining these simple-but-specialized models together, in an assembly-line fashion, produces powerful workflows for processing batch-imaged specimens (Weaver & Smith, 2023). The outputs of these multi-model pipelines can then be fed into existing image analysis pipelines or used to train new CV models, opening the door to truly high-throughput digitization and analysis of preserved specimens (Lürig, 2022; Van Belleghem et al., 2018; van den Berg et al., 2024).

ACKNOWLEDGEMENTS:

We thank the Grainger Bioinformatics Center for funding this research and providing computational support, as well as the Collaborative Imaging Lab and its manager, Stephanie Ware. We also thank the FMNH for funding support through the Science Innovation Award. This work was made possible by several FMNH internship programs, including the Women in Science Internship and the Prince Fellowship. We thank Gene Cooper, Jacob Witt, Janeen Jones, and Matt Wadleigh for their invaluable technical assistance and programming advice. We are grateful to Rick Ree for his insightful comments on the manuscript, to Tom Schultz for his expertise on tiger beetle coloration, to Maureen Turcatel for help with collection management, to several FMNH volunteers (Francis Dela Pena, Ian Chiu, and Lily Bazan) for specimen barcoding, and to Matt Buffington for discussions on the value of whole-drawer imaging.

REFERENCES

Bates, J. M., Fidino, M., Nowak-Boyd, L., Strausberger, B. M., Schmidt, K. A., & Whelan, C. J. (2023). Climate change affects bird nesting phenology: Comparing contemporary field and historical museum nesting records. *Journal of Animal Ecology, 92*(2), 263–272. https://doi.org/10.1111/1365-2656.13683

Borowiec, M. L., Dikow, R. B., Frandsen, P. B., McKeeken, A., Valentini, G., & White, A. E. (2022). Deep learning as a tool for ecology and evolution. *Methods in Ecology and Evolution, 13*(8), Article 8. https://doi.org/10.1111/2041-210X.13901

Correa-Carmona, Y., Böttger, D., Korsch, D., Holzmann, K.L., Alonso-Alonso, P., Pinos, A., Yon, F., Keller, A., Steffan-Dewenter, I., Peters, M.K., & Brehm, G. (2025). LEPY: A Python-Based Pipeline for Automated Morphological Trait Analysis of Lepidoptera Images. *EcoEvoRxiv*. https://doi.org/10.32942/X2WS78

Crowell, H. L., Curlis, J. D., Weller, H. I., & Davis-Rabosky, A. R. (2024). Ecological drivers of ultraviolet colour evolution in snakes. *Nature Communications, 15*(1), 5213. https://doi.org/10.1038/s41467-024-49506-4

Davis, C. L., Guralnick, R. P., & Zipkin, E. F. (2023). Challenges and opportunities for using natural history collections to estimate insect population trends. *Journal of Animal Ecology*, *92*(2), 237–249. https://doi.org/10.1111/1365-2656.13763

de la Hidalga, A. N., Rosin, P. L., Sun, X., Livermore, L., Durrant, J., Turner, J., Dillen, M., Musson, A., Phillips, S., Groom, Q., & Hardisty, A. (2022). Cross-validation of a semantic segmentation network for natural history collection specimens. *Machine Vision and Applications*, 33(3), 39. https://doi.org/10.1007/s00138-022-01276-z

Duran, D. P., & Gough, H. M. (2020). Validation of tiger beetles as distinct family (Coleoptera: Cicindelidae), review and reclassification of tribal relationships. *Systematic Entomology*, 45(4), 723-729. https://doi.org/10.1111/syen.12440

Dwyer, B., Nelson, J., Hansen, T., et. al. (2024). Roboflow (Version 1.0) [Software]. Available from https://roboflow.com.

French, R., Bell, A., Calladine, K., Acorn, J., & Sperling, F. (2021). Genomic distinctness despite shared color patterns among threatened populations of a tiger beetle. *Conservation Genetics*, *22*, 1–16. https://doi.org/10.1007/s10592-021-01370-1

Gaumer, G. C. (1977). The variation and taxonomy of *Cicindela formosa* Say (Coleoptera: Cicindelidae) [Doctoral dissertation, Texas A&M University]. *ProQuest Dissertations & Theses Global* (7720372).

Gough, H. M., Duran, D. P., Kawahara, A. Y., & Toussaint, E. F. A. (2019). A comprehensive

molecular phylogeny of tiger beetles (Coleoptera, Carabidae, Cicindelinae). *Systematic Entomology*, 44(2), 305-321. https://doi.org/10.1111/syen.12324

Hancock, G. R. A., Cuthill, I. C., & Troscianko, J. (2025). Shining a light on camouflage evolution: Using genetic algorithms to determine the effects of geometry and lighting on optimal camouflage. bioRxiv. https://doi.org/10.1101/2025.03.01.640995

Herbst, R., Stille, D., Gwilliam, G. F., III, et al. (2025). Unlocking the past: The potential of large language models to revolutionize transcription of natural history collections. *Data Intelligence*, 7, 237–264. https://doi.org/10.3724/2096-7004.di.2025.0011

Holmes, M. W., Hammond, T. T., Wogan, G. O. U., Walsh, R. E., LaBarbera, K., Wommack, E. A., Martins, F. M., Crawford, J. C., Mack, K. L., Bloch, L. M., & Nachman, M. W. (2016). Natural history collections as windows on evolutionary processes. *Molecular Ecology*, *25*(4), Article 4. https://doi.org/10.1111/mec.13529

Holovachov, O., Zatushevsky, A., & Shydlovsky, I. (2014). Whole-Drawer Imaging of Entomological Collections: Benefits, Limitations and Alternative Applications. *Journal of Conservation and Museum Studies*, *12*(1). https://doi.org/10.5334/jcms.1021218

Hudson, L. N., Blagoderov, V., Heaton, A., Holtzhausen, P., Livermore, L., Price, B. W., Walt, S. van der, & Smith, V. S. (2015). Inselect: Automating the Digitization of Natural History Collections. *PLOS ONE*, *10*(11), e0143402. https://doi.org/10.1371/journal.pone.0143402

Lister, A. M. (2011). Natural history collections as sources of long-term datasets. *Trends in Ecology & Evolution*, *26*(4), 153–154. https://doi.org/10.1016/j.tree.2010.12.009

Losey, J. E., & Vaughan, M. (2006). The Economic Value of Ecological Services Provided by Insects. *BioScience*, *56*(4), 311–323. https://doi.org/10.1641/0006-3568(2006)56[311:TEVOES]2.0.CO;2

Lürig, M. D. (2022). phenopype: A phenotyping pipeline for Python. *Methods in Ecology and Evolution*, *13*(3), 569–576. https://doi.org/10.1111/2041-210X.13771

Medeiros, B. A. S. de, Cai, L., Flynn, P. J., Yan, Y., Duan, X., Marinho, L. C., Anderson, C., & Davis, C. (2025). A composite universal DNA signature for the Tree of Life. *Nature Ecology Evolution*. https://doi.org/10.1038/s41559-025-02752-1

Mráz, R., Štěpka, K., Pekár, M., Matula, P., & Pekar, S. (2023). MAPHIS—Measuring arthropod phenotypes using hierarchical image segmentations. *Methods in Ecology and Evolution*, *15*. https://doi.org/10.1111/2041-210X.14250

Pearson, D. L., Knisley, C. B., & Kazilek, C. J. (2006). A field guide to the tiger beetles of the United States and Canada. Oxford University Press.

Postema, E. G., Lippey, M. K., & Armstrong-Ingram, T. (2023). Color under pressure: How multiple factors shape defensive coloration. *Behavioral Ecology*, *34*(1), 1–13. https://doi.org/10.1093/beheco/arac056

Ruane, S., & Austin, C. C. (2017). Phylogenomics using formalin-fixed and 100+ year-old intractable natural history specimens. *Molecular Ecology Resources*, 17(5), 1003-1008. https://doi.org/10.1111/1755-0998.12655

Sanders, N. J., Cooper, N., Davis Rabosky, A. R., & Gibson, D. J. (2023). Leveraging natural history collections to understand the impacts of global change. *Journal of Animal Ecology*, 92(2), 232-236. https://doi.org/10.1111/1365-2656.13882

Schultz, T. D., & Hadley, N. F. (1987). Structural Colors of Tiger Beetles and Their Role in Heat Transfer through the Integument. *Physiological Zoology*, *60*(6), 737–745. https://doi.org/10.1086/physzool.60.6.30159990

Shiyake, S. (2017). How to identify Tiger Beetles. Osaka Museum of Natural History.

Spiesman, B. J., Gratton, C., Hatfield, R. G., Hsu, W. H., Jepsen, S., McCornack, B., Patel, K., & Wang, G. (2021). Assessing the potential for deep learning and computer vision to identify bumble bee species from images. *Scientific Reports, 11*(1), 7580. https://doi.org/10.1038/s41598-021-87210-1

Steinke, D., McKeown, J. T. A., Zyba, A., McLeod, J., Feng, C., & Hebert, P. D. N. (2024). Low-cost, high-volume imaging for entomological digitization. *ZooKeys, 1206*, 315–326. https://doi.org/10.3897/zookeys.1206.123670

Stenhouse, A., Fisher, N., Lepschi, B., Schmidt-Lebuhn, A., Rodriguez, J., Turco, F., Reeson, A., Paris, C., & Thrall, P. H. (2025). A vision of human–AI collaboration for enhanced biological collection curation and research. *BioScience,* biaf021. https://doi.org/10.1093/biosci/biaf021

Sun, J., Futahashi, R., & Yamanaka, T. (2021). Improving the Accuracy of Species Identification by Combining Deep Learning With Field Occurrence Records. *Frontiers in Ecology and Evolution*, *9*. https://doi.org/10.3389/fevo.2021.762173

Sys, S., Weißbach, S., Jakob, L., Gerber, S., & Schneider, C. (2022). CollembolAI, a macrophotography and computer vision workflow to digitize and characterize samples of soil invertebrate communities preserved in fluid. *Methods in Ecology and Evolution*, *13*(12), 2729–2742. https://doi.org/10.1111/2041-210X.14001

Trautner, J., & Geigenmüller, K. (1987). Tiger beetles ground beetles: Illustrated key to the Cicindelidae and Carabidae of Europe. Margraf.

Truong, M.-X. A., & Van der Wal, R. (2024). Exploring the landscape of automated species identification apps: Development, promise, and user appraisal. *Bioscience*, *74*(9), 601–613. https://doi.org/10.1093/biosci/biae077

Van Belleghem, S. M., Papa, R., Ortiz-Zuazaga, H., Hendrickx, F., Jiggins, C. D., Owen McMillan, W., & Counterman, B. A. (2018). patternize: An R package for quantifying colour pattern variation. *Methods in Ecology and Evolution*, *9*(2), 390–398. https://doi.org/10.1111/2041-210X.12853

van den Berg, C. P., Condon, N. D., Conradsen, C., White, T. E., & Cheney, K. L. (2024). Automated workflows using Quantitative Colour Pattern Analysis (QCPA): A guide to batch processing and downstream data analysis. *Evolutionary Ecology*, 38, 387–397. https://doi.org/10.1007/s10682-024-10291-7

Wagner, D. L. (2020). Insect Declines in the Anthropocene. *Annual Review of Entomology*, *65*, 457–480. https://doi.org/10.1146/annurev-ento-011019-025151

Watanabe, A. (2018). How many landmarks are enough to characterize shape and size variation? *PLOS ONE*, *13*(6), e0198341. https://doi.org/10.1371/journal.pone.0198341

Weaver, W. N., & Smith, S. A. (2023). From leaves to labels: Building modular machine learning networks for rapid herbarium specimen analysis with LeafMachine2. *Applications in Plant Sciences, 11*(5), Article 5. https://doi.org/10.1002/aps3.11548

Weeks, B. C., Zhou, Z., O'Brien, B. K., Darling, R., Dean, M., Dias, T., Hassena, G., Zhang, M., & Fouhey, D. F. (2023). A deep neural network for high-throughput measurement of functional traits on museum skeletal specimens. *Methods in Ecology and Evolution*, *14*(2), 347–359. https://doi.org/10.1111/2041-210X.13864

Weinstein, B. G. (2018). A computer vision for animal ecology. *Journal of Animal Ecology*, *87*(3), 533–545. https://doi.org/10.1111/1365-2656.12780

Welch, K. D., & Lundgren, J. G. (2024). Introducing BugBox: A Platform for AI-Assisted

Bioinventories of Arthropods. *American Entomologist*, *70*(3), 31–33. https://doi.org/10.1093/ae/tmae047

Weller, H. I., Hiller, A. E., Lord, N. P., & Van Belleghem, S. M. (2024). recolorize: An R package for flexible colour segmentation of biological images. *Ecology Letters*, *27*(2), e14378. https://doi.org/10.1111/ele.14378

Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T., & Vieglais, D. (2012). Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. *PLOS ONE*, *7*(1), e29715. https://doi.org/10.1371/journal.pone.0029715

Wiken, E., Jiménez Nava, F., & Griffith, G. (2011). North American terrestrial ecoregions — Level III. Commission for Environmental Cooperation.

Wilson, R. J., de Siqueira, A. F., Brooks, S. J., Price, B. W., Simon, L. M., van der Walt, S. J., & Fenberg, P. B. (2023). Applying computer vision to digitised natural history collections for climate change research: Temperature-size responses in British butterflies. *Methods in Ecology and Evolution*, *14*(2), 372–384. https://doi.org/10.1111/2041-210X.13844

Wu, S., Chang, C.-M., Mai, G.-S., Rubenstein, D. R., Yang, C.-M., Huang, Y.-T., Lin, H.-H., Shih, L.-C., Chen, S.-W., & Shen, S.-F. (2019). Artificial intelligence reveals environmental constraints on colour diversity in insects. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-12500-2

Wührl, L., Pylatiuk, C., Giersch, M., Lapp, F., von Rintelen, T., Balke, M., Schmidt, S., Cerretti, P., & Meier, R. (2022). DiversityScanner: Robotic handling of small invertebrates with machine learning methods. *Molecular Ecology Resources*, *22*(4), 1626–1638. https://doi.org/10.1111/1755-0998.13567

Yamamoto, N., & Sota, T. (2020). Evolutionary fine-tuning of background-matching camouflage among geographical populations in the sandy beach tiger beetle. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1941), 20202315. https://doi.org/10.1098/rspb.2020.2315

Yang, L. H., Postema, E. G., Hayes, T. E., Lippey, M. K., & MacArthur-Waltz, D. J. (2021). The complexity of global change and its effects on insects. *Current Opinion in Insect Science*, *47*, 90–102. https://doi.org/10.1016/j.cois.2021.05.001