CalBug

Digitizing California Arthropod Collections

A project supported by the National Science Foundation

Scanning Microscope Slides

Peter T. Oboyski
Essig Museum of Entomology
University of California, Berkeley
http://calbug.berkeley.edu

What is CalBug?

NSF - ADBC grant

Collaboration among the eight major entomology collections in California

Digitize 1.2 million specimens

Essig Museum of Entomology

California Academy of Sciences

California State Collection of Arthropods

Bohart Museum, UC Davis

Entomology Research Museum, UC Riverside

San Diego Natural History Museum

Santa Barbara Museum of Natural History

LA County Museum

California Arthropod Diversity Online

Museum of Entomology Collections

matches: 2050

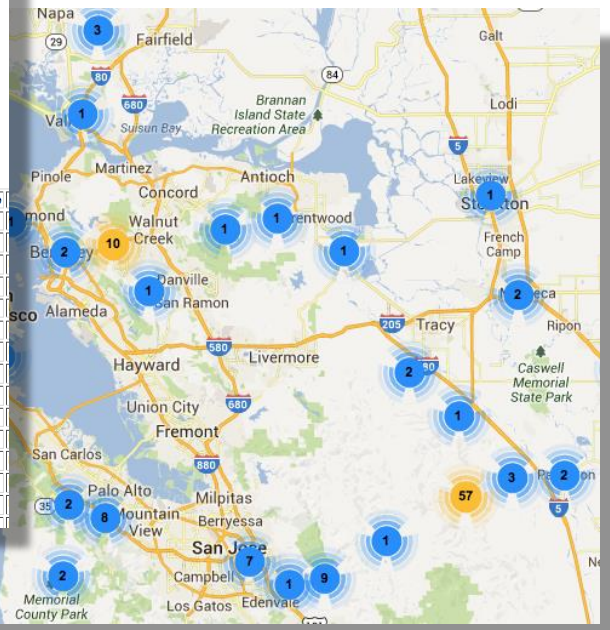
your results (tab-delimited text file, max 2050 lines, file size = 1115.5 K)
Windows users: 1) Add a .txt extension to save file, or 2) tell your browser to open the file in Excel.)

georeferenced records on a map

T FROM eme WHERE ScientificName like "%Libellula%" ORDER BY Family,ScientificName

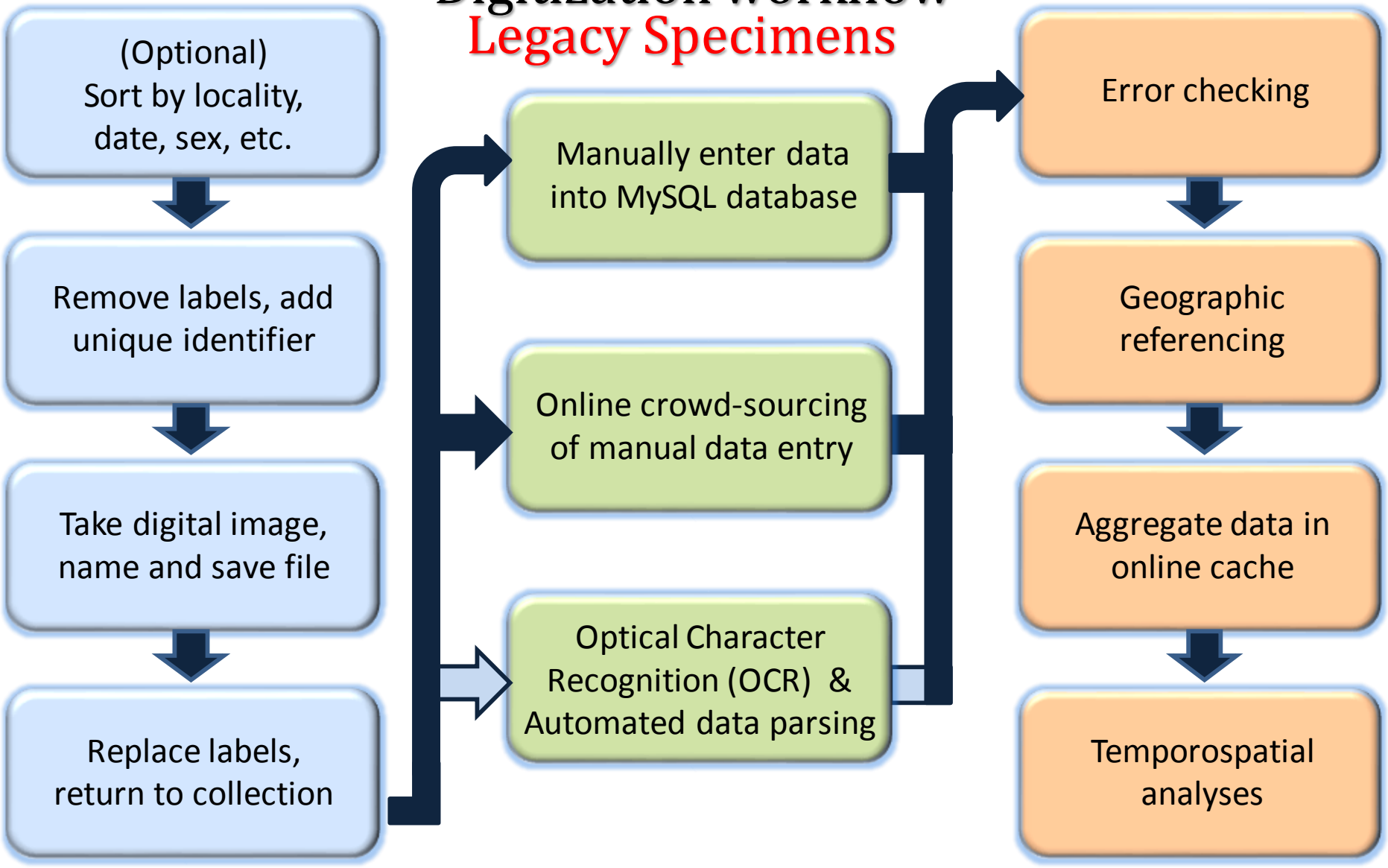
to see the full specimen record

	Order	Family	Genus Species ssp	Collector	Year	State/Province
	Odonata	Libellulidae	Libellula auripennis	Jerry A. Powell	1951	Florida
	Odonata	Libellulidae	Libellula auripennis		1939	Texas
	Odonata	Libellulidae	Libellula axilena	Jerry A. Powell	1951	Louisiana
	Odonata	Libellulidae	Libellula comanche	John W. MacSwain	1947	California
	Odonata	Libellulidae	Libellula comanche	John W. MacSwain & Ray F. Smith	1951	California
	Odonata	Libellulidae	Libellula comanche	John W. MacSwain	1947	California
	Odonata	Libellulidae	Libellula comanche	John W. MacSwain & William F. Barr	1947	California
	Odonata	Libellulidae	Libellula comanche	John W. MacSwain & William F. Barr	1947	California
	Odonata	Libellulidae	Libellula comanche	John W. MacSwain & William F. Barr	1947	California
	Odonata	Libellulidae	Libellula comanche	John W. MacSwain	1947	California
	Odonata	Libellulidae	Libellula comanche	John W. MacSwain & William F. Barr	1947	California
EMEC81906	Odonata	Libellulidae	Libellula comanche	John M. Burns	1957	California
EMEC81907	Odonata	Libellulidae	Libellula comanche	John M. Burns	1957	California
EMEC81908	Odonata	Libellulidae	Libellula comanche	John M. Burns	1957	California
EMEC81910	Odonata	Libellulidae	Libellula comanche	Catherine A. Toschi	1963	Arizona



Digitization workflow

Legacy Specimens



Digital camera tethered to computer

Filename = EMEC218958 Paracotalpa ursina.jpg

Why Image Specimens/Labels?

- Magnify difficult to read labels
- Verbatim archive of label data
  - Essential for proofing data
  - Useful for taxonomists interested in label data
- Data capture can be done remotely

EMEC332214 Aetheca wagneri.jpg

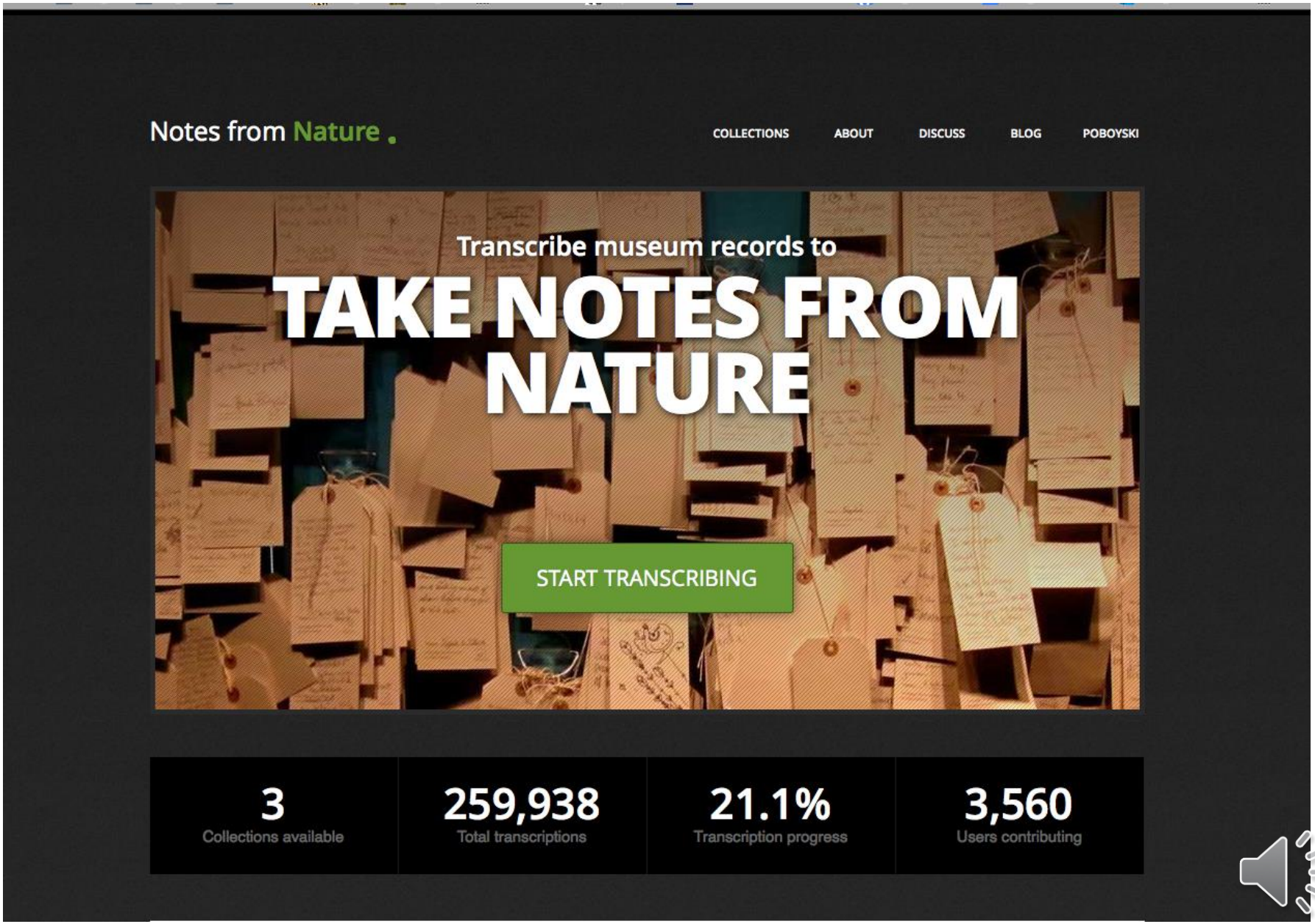
Notes from Nature

Citizen Science data transcription

ZOONIVERSE

REAL SCIENCE ONLINE

http://www.notesfromnature.org/ Launched April 22, 2013



Progress

Essig Museum …

Fully databased & georeferenced

- 7,891 Odonata (dragonflies, damselflies)
  - 753 Ephemeroptera (mayflies)
- 2,950 Trichoptera (caddisflies)
- 6,394 aquatic Coleoptera (beetles)
- 9,271 Tephritidae (fruit flies)
- 27,259 Total

Progress

Essig Museum …

Databased, but not georeferenced

5,666 Carabidae (ground beetles)

1,203 Siphonaptera (fleas)

27,225 Halictidae (sweat bees)

1,155 Andrenidae (burrowing bees)

753 Mecoptera & Neuroptera

36,002 Total (over 63,000 databased)

Progress

Essig Museum …

Imaged, but not databased

8,800 Hesperiidae (skippers)

4,580 Sphingidae (hawk moths)

2,274 Apis mellifera (honey bees)

7,586 Chrysididae (cuckoo wasps)

79,306 Sphecidae (sphecid wasps)

102,546 Total

Being databased by Notes from Nature

Progress

Essig Museum	>165,000 specimens
Cal Academy	>105,000
Cal Arthro Coll.(CDFA)	> 13,000
UC Davis	>110,000
UC Riverside	> 52,000
Santa Barbara	> 13,000
San Diego MNH	> 400
LA County	> 2,200

>460,000 Total specimens processed

Slide Scanning

EMEC332214 Aetheca wagneri.jpg

How to batch scan slides

Equipment

- ✓ Flat bed scanner
- ✓ Jig for arranging slides
- ✓ Unique ID labels

Software

- ✓ Scanner software
- ✓ Photoshop
- ✓ IrfanView

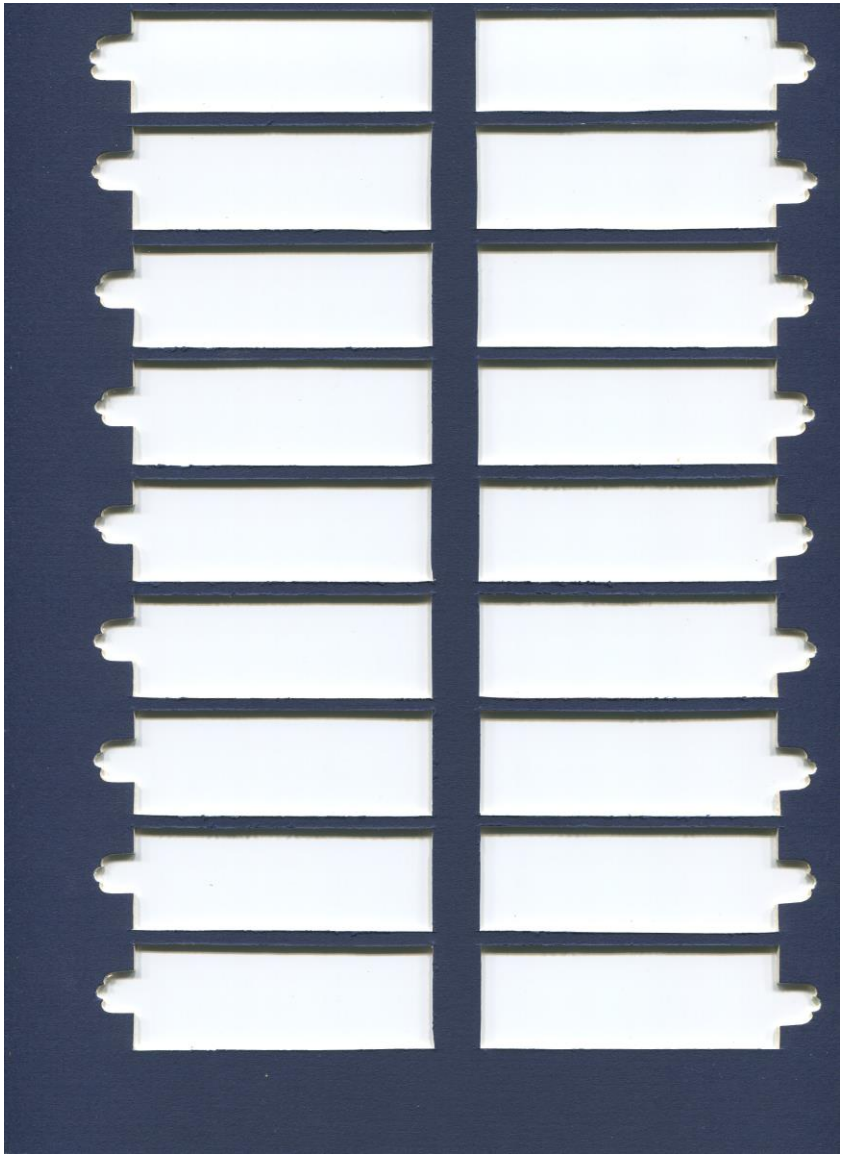
Process

1. Sort slides (optional)
2. Lay slides on scanner bed
3. Add unique ID labels
4. Scan slide set (at 400 DPI)
5. Open scan in Photoshop
6. Use an “Action” to
   Divide scan -> individual slides
7. “Save for Web”
8. Rename with IrfanView

Equipment

- Any flatbed scanner will do
- Use software that comes with the scanner
- Create a jig (we use matting for framing)

Jig cut from mat board



Jig filled with Aphid slides

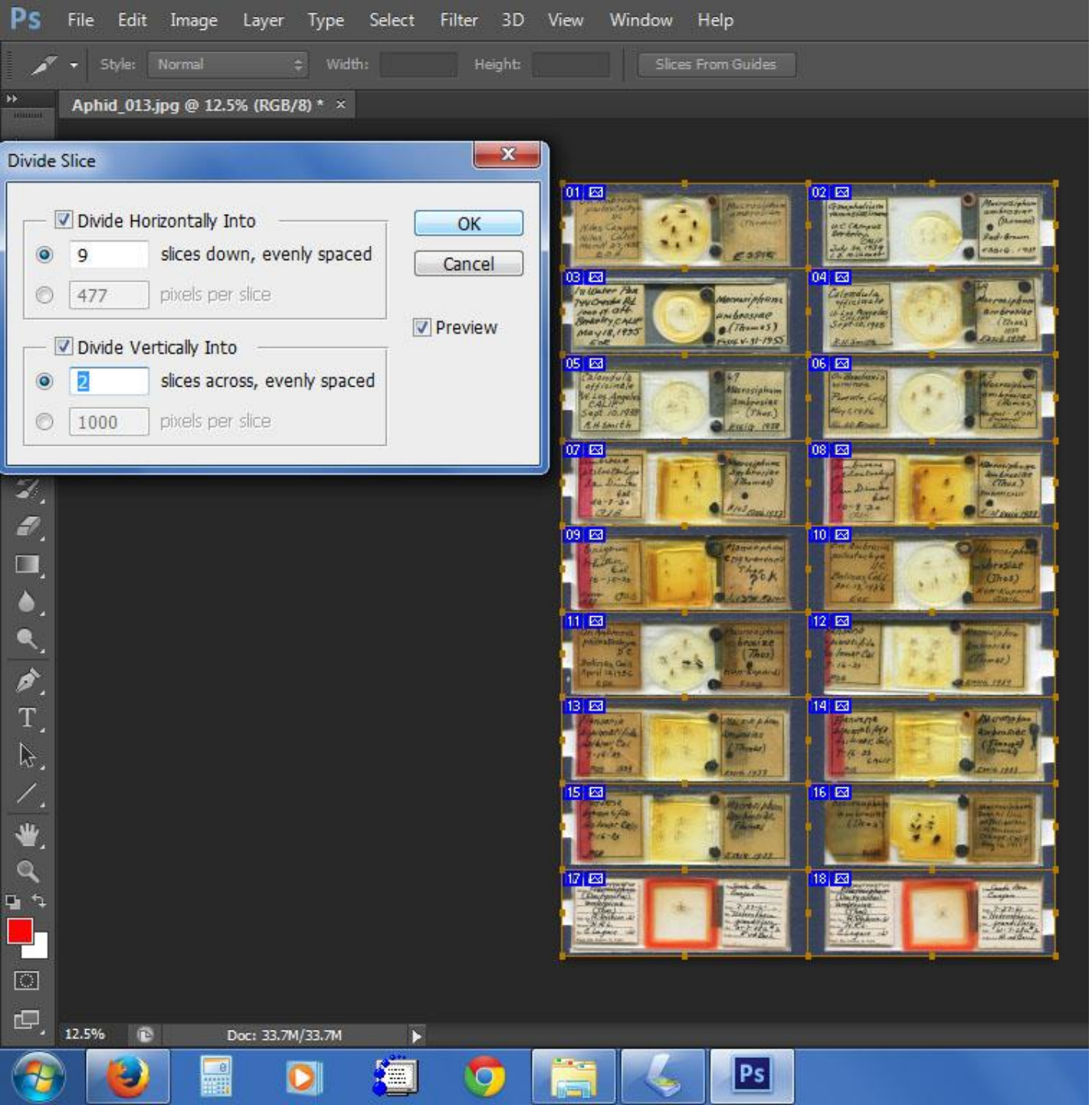
Use the

Crop Tool

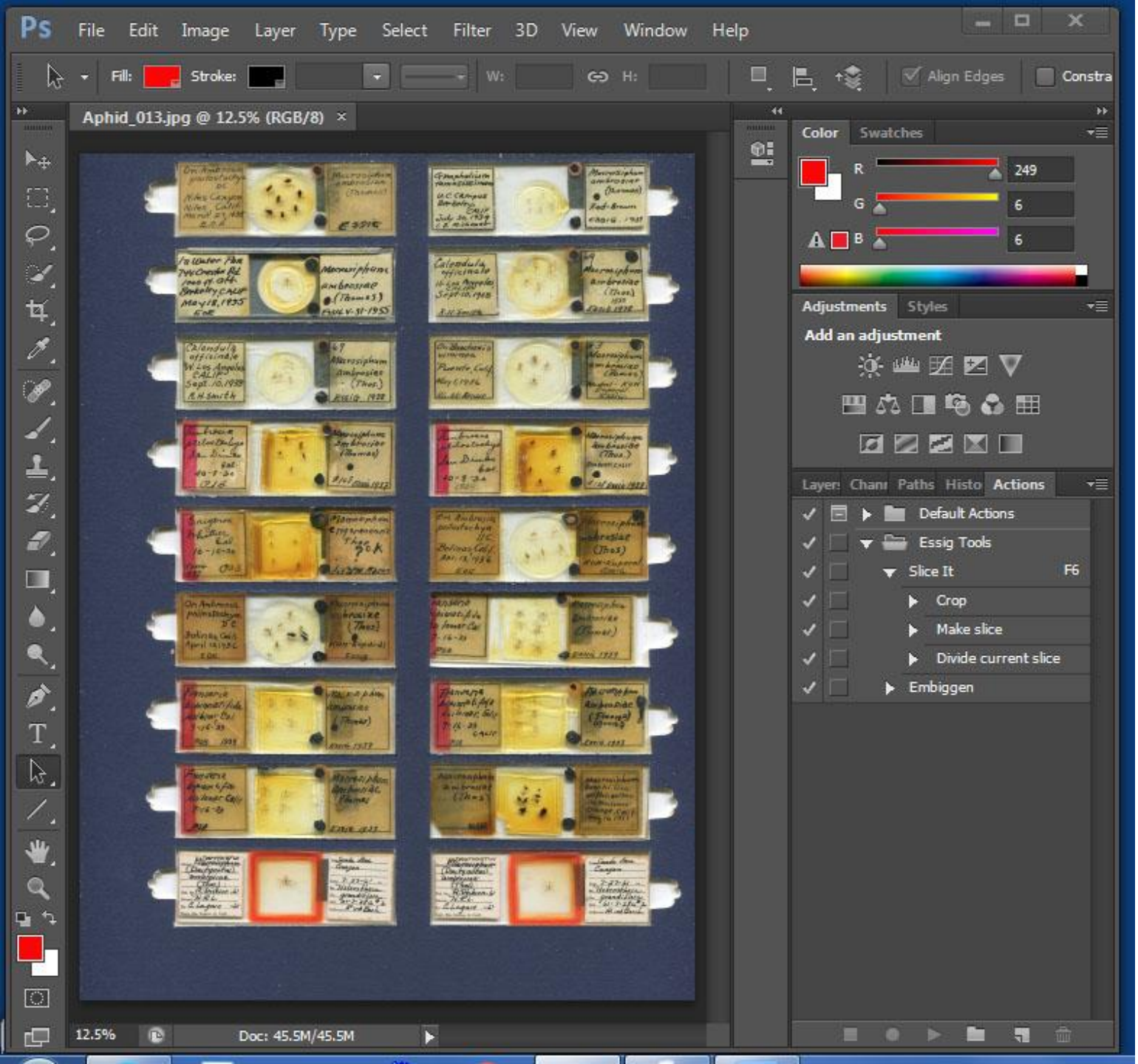
And

Slice Tool

To define the individual slides



This scan has already been cropped and is now being divided into 18 equal rectangles



Save the
- Crop
- Make Slice
- Divide Slice
Steps as an "Action"

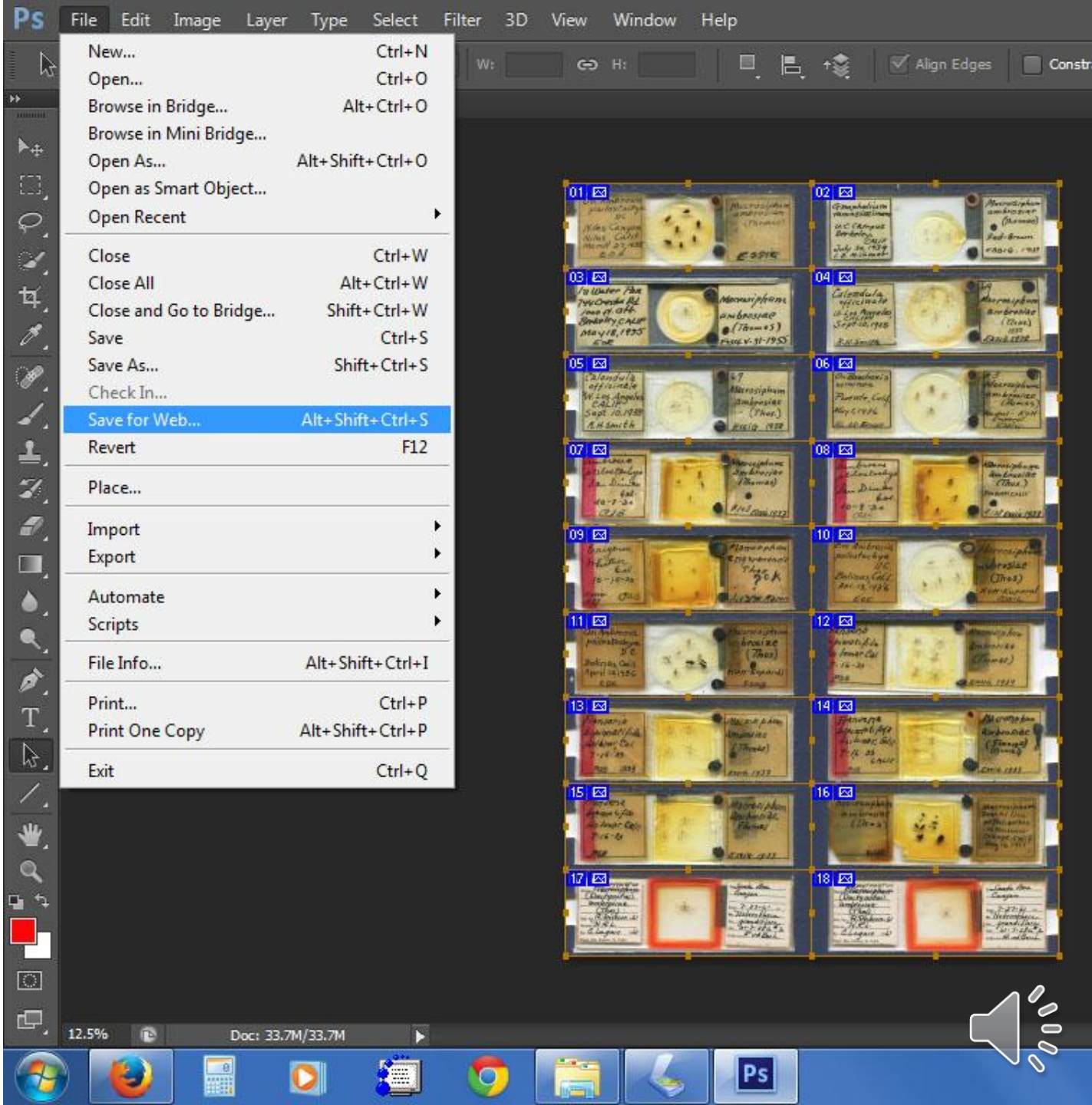
I named my Action "Slice It"

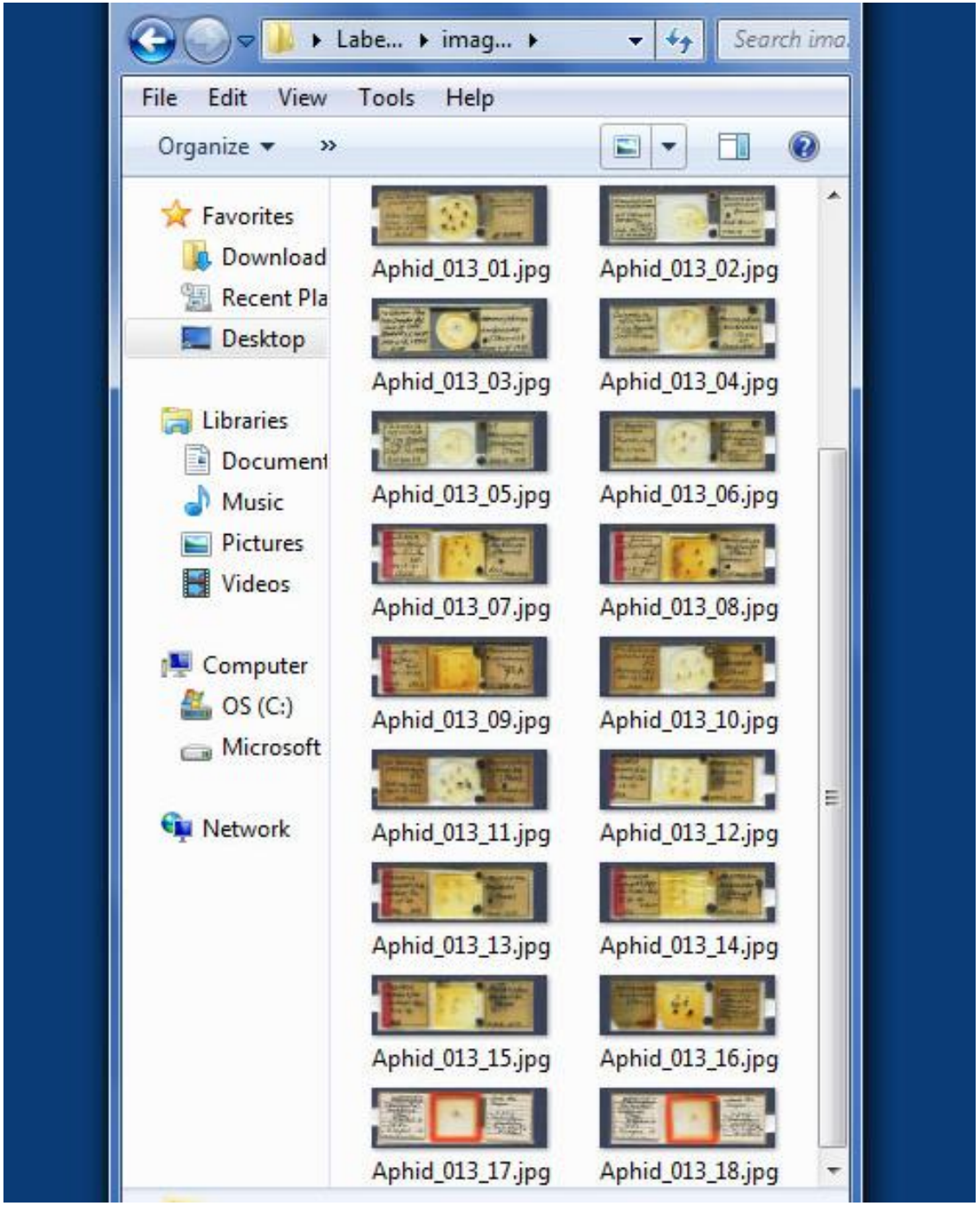
Just "play" the Action for each scan.

Process takes as long as it takes to click the mouse.

Choose the “Save for Web” option in the File menu

Aside:
This feature was created for faster loading of large images on web sites by cutting them into smaller pieces to be loaded simultaneously.

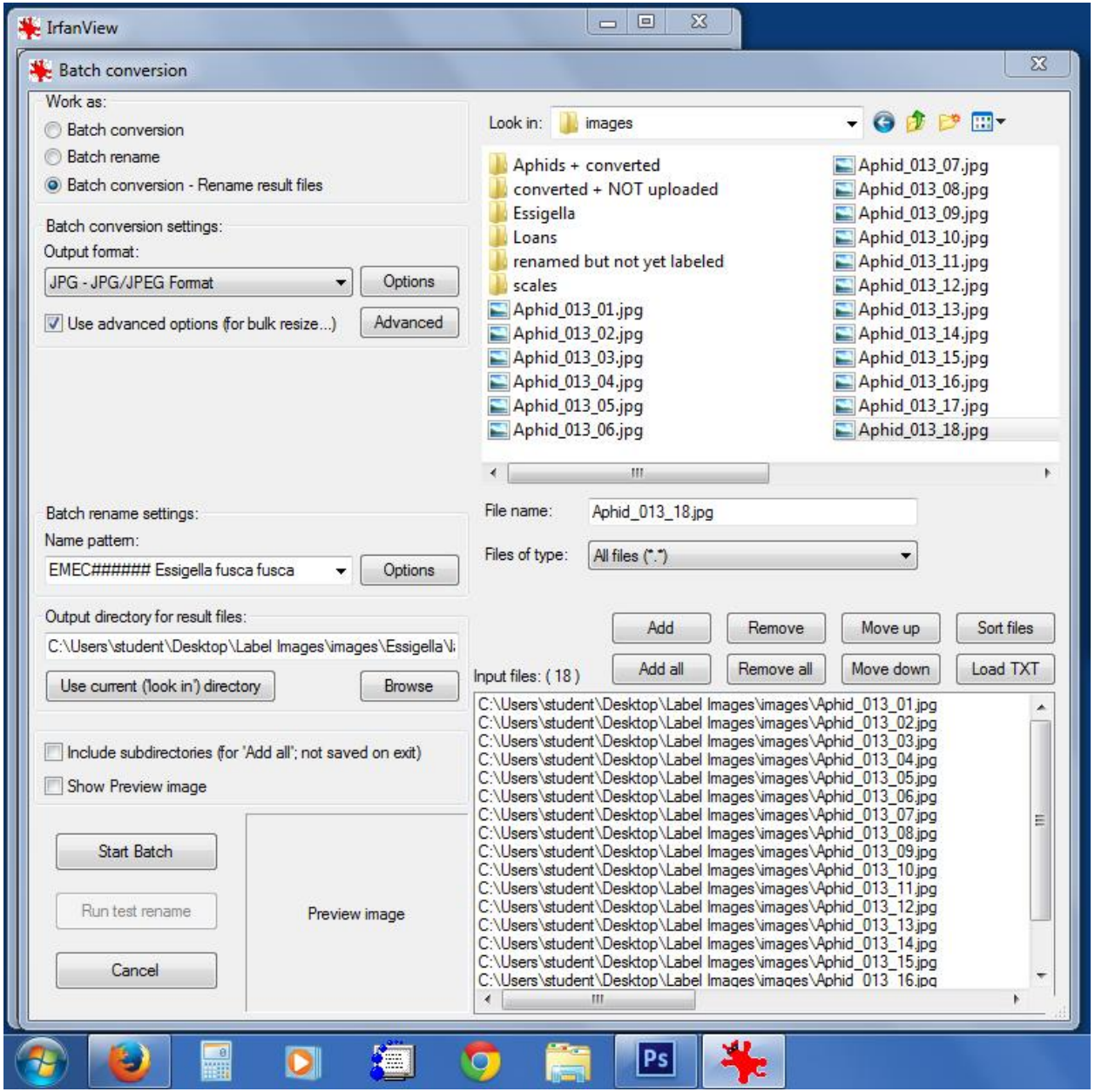




The result of the “Save for Web” option is 18 individual files ending in _01, _02, _03, etc.

The next step is to rename each of the individual files with the unique ID, genus, and species. We use IrfanView software by creating a pattern to batch process all the files at once.

Provide filename template



Select files

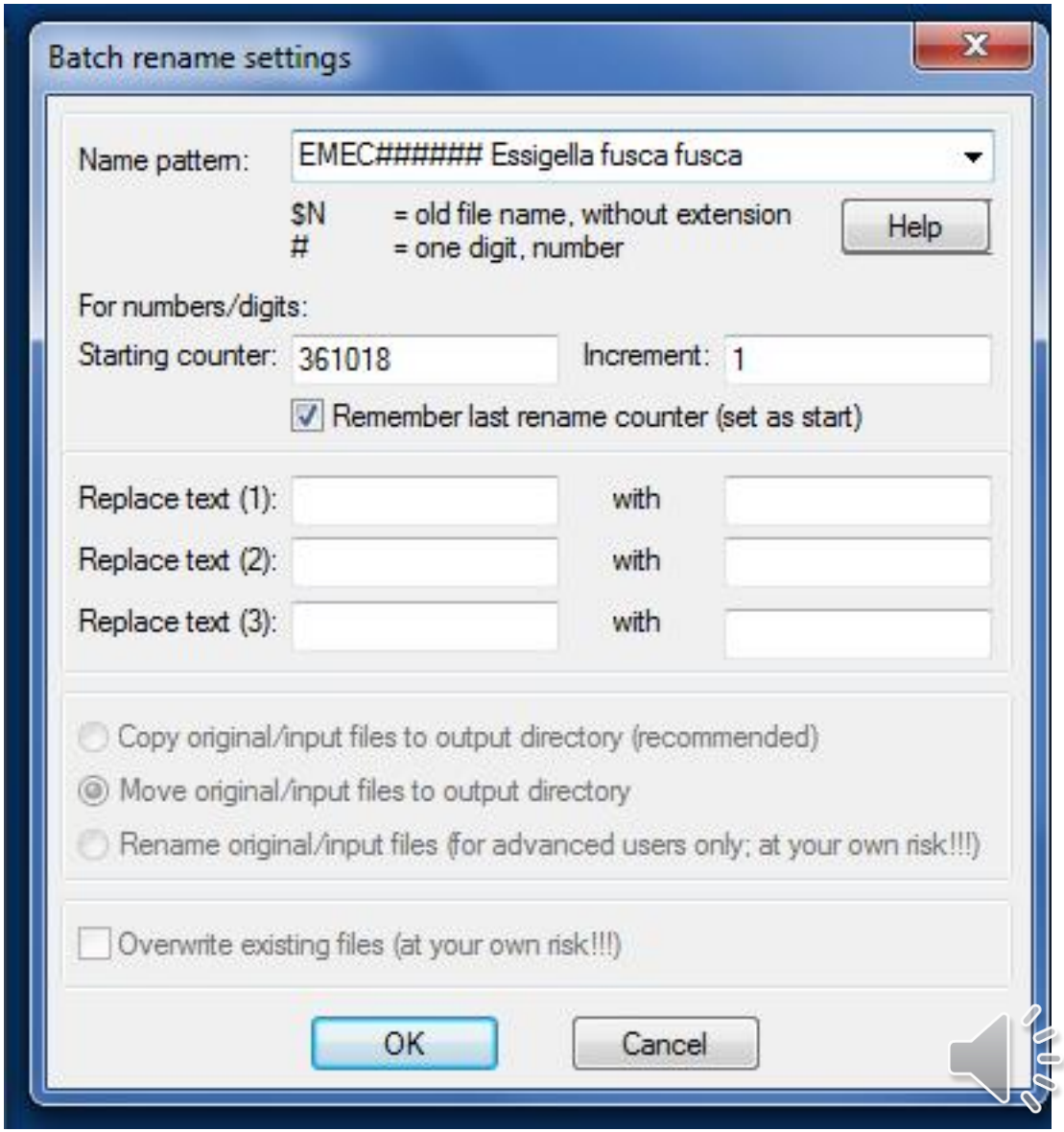
Process one species at a time!

Batch file renaming with IrfanView

Create name pattern

Use # symbol for digits

Provide the starting number




Aphid_013_01.jpg


Aphid_013_02.jpg


Aphid_013_03.jpg


Aphid_013_04.jpg


Aphid_013_05.jpg


Aphid_013_06.jpg


Aphid_013_07.jpg


Aphid_013_08.jpg


Aphid_013_09.jpg


Aphid_013_10.jpg


Aphid_013_11.jpg


Aphid_013_12.jpg


Aphid_013_13.jpg


Aphid_013_14.jpg


Aphid_013_15.jpg


Aphid_013_16.jpg


Aphid_013_17.jpg


Aphid_013_18.jpg


EMEC361018
Essigella fusca
fusca.jpg


EMEC361019
Essigella fusca
fusca.jpg


EMEC361020
Essigella fusca
fusca.jpg


EMEC361021
Essigella fusca
fusca.jpg


EMEC361022
Essigella fusca
fusca.jpg


EMEC361023
Essigella fusca
fusca.jpg


EMEC361024
Essigella fusca
fusca.jpg


EMEC361025
Essigella fusca
fusca.jpg


EMEC361026
Essigella fusca
fusca.jpg


EMEC361027
Essigella fusca
fusca.jpg


EMEC361028
Essigella fusca
fusca.jpg


EMEC361029
Essigella fusca
fusca.jpg


EMEC361030
Essigella fusca
fusca.jpg

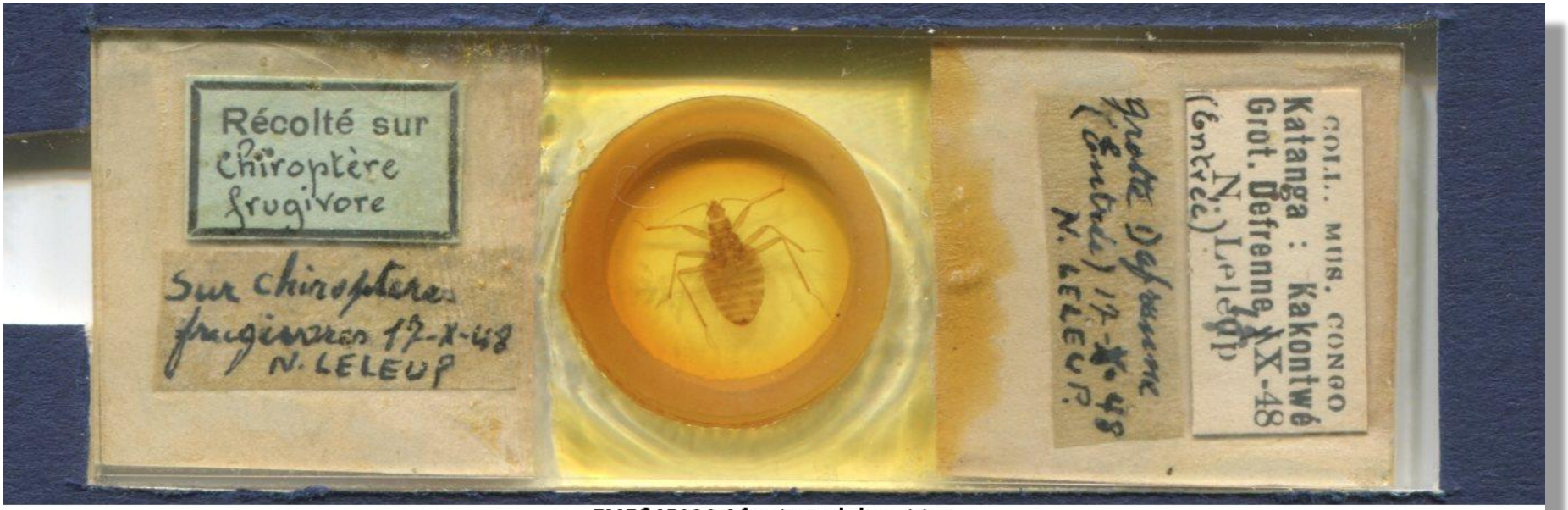

EMEC361031
Essigella fusca
fusca.jpg


EMEC361032
Essigella fusca
fusca.jpg


EMEC361033
Essigella fusca
fusca.jpg


EMEC361034
Essigella fusca
fusca.jpg

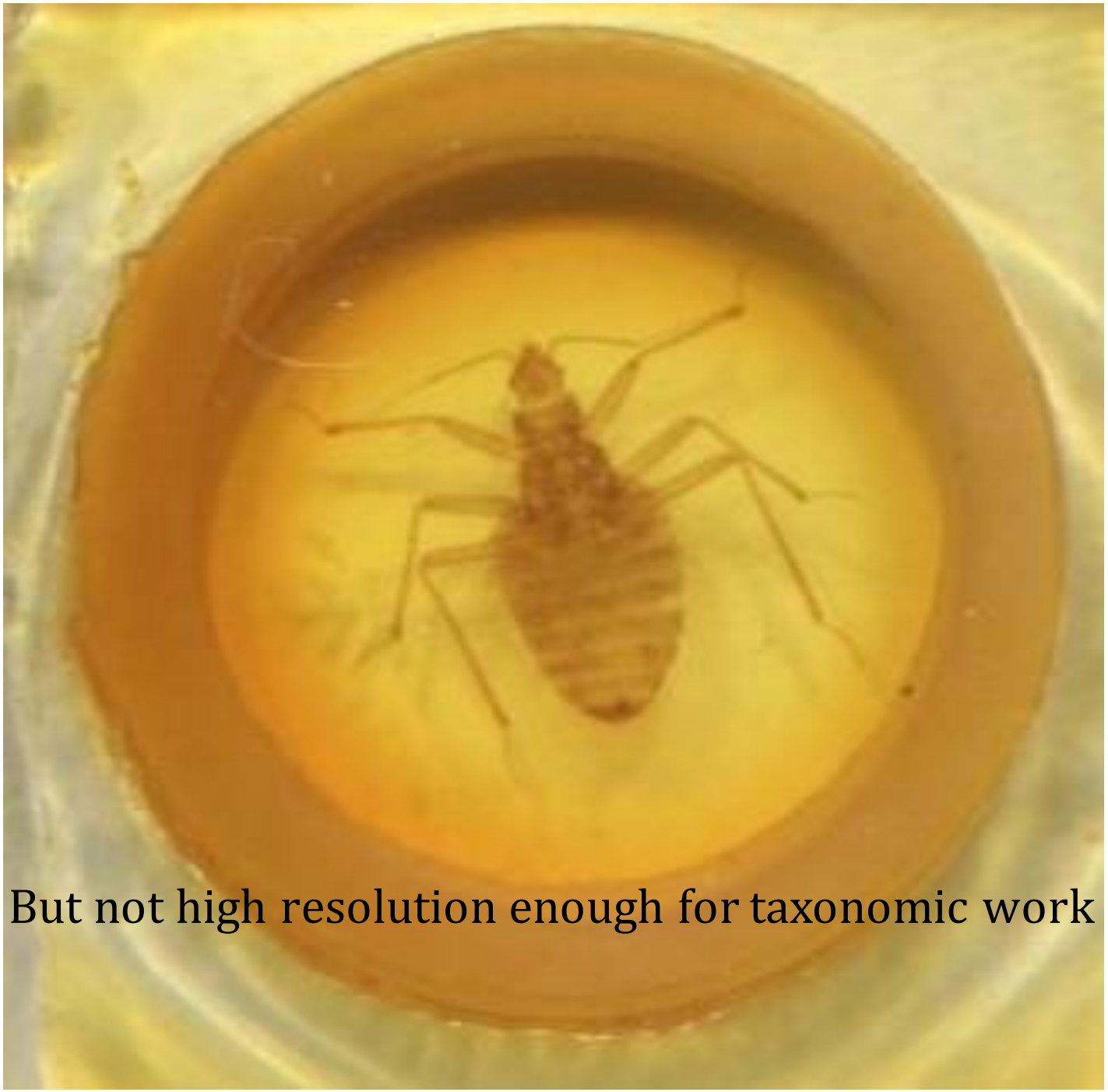

EMEC361035
Essigella fusca
fusca.jpg



EMEC45020 Afrocimex leleupi.jpg

We also use IrfanView to overlay the unique ID, genus, and species on the bottom of the image, since the unique ID is on the back of the slide and therefore not visible in the scan.

400 DPI
Seems to provide high enough resolution for difficult to read labels while keeping file size relatively small



But not high resolution enough for taxonomic work

Slide Scanning Speed

Our students average 135 to 170 slides scanned, processed, and renamed per hour.

The slowest step in the process is cutting out and sticking on the Unique ID labels (we are working on getting pre-cut archival quality stickers)

This does not include the actual databasing (ie. Transcribing the data from the images into our database)

Slide Scanning Cost

Scanner: $50 - $80

Jig (mat board): $2 - $10

Photoshop: $300 - $500 depending on version

- Photoshop freely available at some institutions
- freeware alternatives available [*GIMP*, *Image Cut*]

IrfanView: free

How to batch scan slides

Equipment

- ✓ Flat bed scanner
- ✓ Jig for arranging slides
- ✓ Unique ID labels

Software

- ✓ Scanner software
- ✓ Photoshop
- ✓ IrfanView

Process

1. Sort slides (optional)
2. Lay slides on scanner bed
3. Add unique ID labels
4. Scan slide set (at 400 DPI)
5. Open scan in Photoshop
6. Create an “Action” to
   Divide scan -> individual slides
7. “Save for Web”
8. Rename with IrfanView