Barcoding and Preparing to Image

Before specimens can be imaged, they need to be barcoded and pulled from the collection. Specimens from the targeted families will be given a barcode at the top right corner. If there is an obstruction on the top right, the next best place is the top left or along the top edge of the specimen.

After the specimens have been barcoded, they are removed from the collection and scanned into the appropriate family log to help keep track of progress and to create a pathway back to the specimen. To create the log, the “filed as name” is checked against the accepted names list in the Tropicos Excel file. If the name appears on the list, copy and paste it in the “Name” (filed as) field (Figure 1). If the species name does not appear on the list of accepted names, type “sans Tropicos” next to the species name. After the species name is added to the list, move the cursor into the “Barcode” (scanned) column, and then scan each barcode to create a link between the “filed as name” and the barcode. Lastly, after the specimen has been imaged, fill in the “Date Imaged and Imager” field with your initials. *This process and the renamming of specimens have been streamlined into the imaging process as of April 2013. Please see the section “Transitioning to using the TCN Imaging Workflow Application.”

Specimen folders will be labeled with either a single species (in which case the order within the folder is not significant) or in a folder with multiple species. If a folder contains multiple species (e.g., an A-Z folder or a genus folder with a letter like “Carex A or Quercus A-Z”) it will be in alphabetical order from the top to the bottom. As these specimens are logged double check to make sure they are in correct alphabetical order. In the process of logging the images the specimens will end up in reverse alphabetical order. Do not worry about reversing these as they will correct themselves as they are imaged.

Housekeeping

While the PhotoBox is warming up, please assess the conditions both inside and outside the PhotoBox. Debris from the specimens tends to accumulate on the velvet within the box and on the floor and mat beneath the box. Make a ring of tape from the tape roll kept near the PhotoBox to gently pickup plant material and dust from the velvet. Also, if necessary, please press down the ruler, color square and corners of the sheet so they lie flat. Please use a dustpan and brush to sweep up debris from the floor. These items are located at the base of one of the microscope tables in the back of the herbarium. These steps will help keep the images and imaging room looking their best and reduce dust which is bad for the collection space.
Figure 1. The specimen log is used to track our progress and to create a link between the barcode and the filed as name. This will help to facilitate linking the image back to the specimen after the images are processed using optical character recognition software.

**Imaging**

To begin an imaging session, log in to the computer using your Identikey. Turn on or wake up the camera (by pressing the shutter button) to bring up the Cannon EOS Utility software and flip the power switch on the PhotoBox. The EOS software should launch as soon as the computer detects that the camera is connected and powered on. If the software does not launch automatically, you can start the software by accessing the shortcut for the EOS Utility on the desktop. (If the software does not launch automatically go to Preferences > Basic Settings and click the checkbox for Start EOS Utility Automatically when camera is connected.) When the EOS Utility launches, select “Camera Settings/Remote Shooting”
Cannon EOS software

As of April 2013 COLO has changed its workflow. Pre-imaging setup now starts at the “Transitioning to using the TCN Imaging Workflow Application” header.

Settings for the Canon EOS Utility software for capturing pictures appears to be global to the computer (i.e. all users will use the same settings). The settings for viewing images are unique to each user and often do not carry over from session to session. For this reason each user’s settings need to be checked before each imaging session.

Before imaging, double check to make sure your name appears below the shutter button. If your name appears, your file has already been selected and you are ready to take pictures. If someone else’s name is there, open the folder underneath the shutter button (under the browse option) and navigate to your folder on the C drive within the public folder (Figure 2). Specifically your folder will be on the C drive under computer:

C:\Users\Public\Pictures\your name.

This folder stores the active work and the CR2 files before they are converted and processed in LightRoom. It is important to save your images in the public folder as they will not be accessible for post processing if they are in the wrong folder. If everything is set correctly your name will appear under the shutter button (Figure 3).

Figure 2. The pathway to the folder where you will store your images. It is very important to store your images in the public folder as they will not be viewable by anyone else if they are in your storage space.
Figure 3. If the folder pathway is selected correctly your name will appear under the shutter shown with the arrow above. If your Identikey name appears here images are not being stored to the public pictures area and will not be available for post processing.
**Digital Photo Professional**

Digital Photo Professional will open automatically when the first image is captured. If the software does not launch there is a shortcut on the desktop. Because Digital Photo Professional and EOS Utility are two separate programs you will need to navigate to your folder in the same manner as you did for the EOS Utility. If you have accessed the correct folder the header of the Digital Photo Professional should read: C:\Users\Public\Pictures\your name.

Setting the image thumbnails to the smallest size appears to be the most effective setting as it allows you to capture more images before having to use the scrollbar (Figure 4).

Also sorting the order images appear to “Shooting Date/Time” under the view and sort tab will keep the last image shot at the end of the series of thumbnails (Figure 5). (By default the images will be sorted by file name so they will move automatically when renamed).

Adjust the workstation to fit your needs! The less stretching and reaching for items you do the better. Moving the carts close in makes it easier to reach specimens without a lot of moving around. Some people find it easiest to stand others prefer to sit. The most important thing is to figure out what works best for you. We have also found it is best also to put the garbage can in front of your feet in case you need to clear debris from specimens.
Figure 4. Selecting small thumbnails allows for more images to be taken before having to scroll down.
Figure 5. Selecting sort by “Shooting Date/Time” will ensure the last image on the screen is the last image captured.
Transitioning to using the TCN Imaging Workflow Application

Starting in April 2013 COLO transitioned to using a metadata capture tool from the University of Wisconsin to log specimens and to rename images. COLO made this transition to help streamline the imaging process. The tool allows imagers to rename images quicker and capture taxonomy (and geography if desired) while imaging.

The main change in this transition is that COLO will no longer be logging specimens prior to imaging them. From an imaging standpoint there will not be much of a change except that we will need to point the tool to your folder at each imaging session.

In the first tab called Working Folder click on the Select New Working Folder and navigate to your folder (Figure 6). The tool will create a destination and processing folder automatically if they have not already been created.

The pathway should read

C:\Users\Public\Pictures\Your Name

The application looks in this folder to find the processing folder so it can rename the image when the barcode is scanned. The application moves renamed photos from the processing folder to the destination folder.
Figure 6. Prior to taking pictures, the Imaging Application must be directed to the folder where images will be stored.

When you set the destination on the Canon EOS software the destination will be your name folder and then the processing folder (Figure 7). This tells the software where to store the image after it has been captured. Images are stored in the processing folder until they are renamed and moved to the destination folder.

The pathway should read:

C:\Users\Public\Pictures\Your Name\processing
Figure 7. The destination on the Canon EOS software will be the processing folder within your folder. Make sure images are being sent to the public pictures folder so they are accessible by all users.

The select data source tab links the tool to the taxonomy source file (Figure 8). The data source only needs to be set when starting a new family. The source files are located in C:\ImagingWorkflow\Taxonomy Files for TCN Imaging Application

The file is a list of accepted names that populate a dropdown menu in the application. Using a dropdown list helps to prevent keystroke errors and can act as a guide for difficult to read hand written labels.
Figure 8. A new taxonomy file must be uploaded to the application each time a new family is started. This only has to be done at the start of each new family.

The Persist Metadata tab allows the imager to select what information will be applied to each image (Figure 9). Specific metadata will appear on every image across different sessions unless it is changed. Image recorder name will be set at the start of each imaging session and the Institution Code and Name will be applied to every specimen. The bottom half of the interface allows the imager to select what items will persist from image to image. For this project we will not be capturing geography so the only field that needs to be selected to persist is “Scientific Name”.
Figure 9. There are often multiple specimens collected from the same state and county. Setting the imaging tool to persist specific metadata between specimens allows for the capture of basic skeletal record information without the need to retype information for each specimen.

The “enter data tab” is the interface where all of the metadata is entered (Figure 10). Any item checked in the last tab will carry over from specimen to specimen but persisted information can be changed at any time by overwriting the information contained in the field. i.e. if the imager changes the county from Boulder to Clear Creek before submitting the image the next image would be prepopulated with Clear Creek for the county field. To start a new imaging session, click on the “New Session Button.” This will tell the application to look for new images in the processing folder. The enter data tab is used to populate the skeletal file. Since scientific name was set to persist from record to record the imager will only need to change this information as it changes.
Figure 10. Before imaging the first specimen the tool must be set up to look for new images. Clicking the “New Session” button will start an imaging session and tell the tool to start looking in the processing folder for new images.

Scan the barcode into the “Barcode Field“ and make sure you use the correct species name (Name from the outside of the folder) in the “Scientific Name” field and hit enter (Figure 11). The image will automatically move into a folder with today’s date in the destination folder. The tool will also create an Excel file and populate it with any metadata that was entered. If a name does not appear in the dropdown list the imager will need to open the skeletal file and type the name into the file after it is submitted (the best practice is to select the genus name without a species (i.e. Salvia) from the dropdown open the skeletal log and add the species. Any name entered in this way should be appended with (sans Tropicos). Our lists do not include species from Mexico (unless their range extends into the United States) so most specimens not appearing in the list will be from Mexico.
Taxonomy and the scanned barcode are required for each specimen.

Calibration

To begin imaging turn the PhotoBox on and give it time to warm up; it appears to take about 10-15 minutes for the lights in the PhotoBox to warm up so it is best to turn the PhotoBox on when logging into the computer. If the lights in the PhotoBox do not have enough time to warm up before calibrating the exposure, RGB and white balance images may not be properly calibrated!

After taking the first picture of the day the image file needs to be checked to make sure the exposure is within a tolerable range. To check exposure we will measure the RGB (Red Green Blue) values of the white square. RGB values should be within 3 of each other and all values should be between 240 and 246 (Figure 12). To test the RGB values, double-click on the color calibrator to zoom in so you can see the white square at the top of the image on the color calibrator. When you move the pointer over the white square the RGB values should be 243 for all three values which represent true white. It is important to calibrate the RGB as close to this value as possible but so long as they are within the acceptable range of +/- 3 they will be corrected during post processing using LightRoom. The initial color balance should be close to the settings seen in Figure 6. It is more important that the numbers be close together than it is for them to be exactly at 243 (i.e. 242, 242, 243 is fine but 245, 243, 240 should be adjusted as the highest and lowest value should be within 3 points of each other). Keep in mind that
this is not an exact process. You will notice that as you move across the white square the RGB values will fluctuate.

Figure 12. The RGB (Red Green Blue) values of a properly exposed image for the white square should be 243, 243, 243. It is more important that the numbers be close together than it is for them to be at 243 (i.e. 242, 242, 243 is fine but 245, 243, 240 should be adjusted as the highest and lowest value should be within 3 points of each other).

Remedying Exposure Issues

If the PhotoBox needs to be adjusted, both the Kelvin Setting and White Balance can be altered to achieve a proper exposure. Lowering the Kelvin number will create a bluer image and will be represented in the third value “B” conversely raising the Kelvin value will create a redder image
represented by the R value. Green adjustments can be made using the white balance shifting the image either nearer to or farther away from the green portion of the spectrum. When you are satisfied with your results for the RGB values, take another image to make sure that values remain relatively consistent keeping in mind that the RGB values may differ slightly as you move across the white square. Before moving on to the next image, check the current image to make sure that the image is in focus and appears sharp. To do this, double click the image and use the cursor to move around the image making sure that the image appears in focus across the entire specimen and the specimen label.

**Digitizing**

Each day specimens will be pulled from the collection and placed on a cart in alphabetical order based on genus and species names within the current family being processed. As a reminder, folders will be labeled with either a single species name (in which case the order within the folder is not significant) or in a folder with multiple species. If a folder contains multiple species (eg., an A-Z folder or a genus folder with a letter like “Carex A or Quercus A-Z”) it will be in alphabetical order from the top to the bottom when it is removed from the collection. During the logging process these specimens will end up in reverse order. The order will be righted after the images are taken. This step reduces sorting time over the course of the project. Please keep an eye out for specimens out of alphabetical order and also for those species for which we have at least three specimens. Species with three specimens should be removed from the A-Z folder and a separate manila folder should be created for that species. Keep in mind, for this project, only specimens from North America are being imaged. These should be the only specimens in the collections with barcodes. In the A-Z folders, most likely there will be a mix of specimens from North America as well as from other parts of the world. If a specimen outside of North America is barcoded, it should be skipped during imaging but the barcode should not be removed from the specimen as ultimately every specimen in the herbarium will receive a barcode.

To begin imaging, start by lining the specimen up against the guide sheet in the PhotoBox. Using the corners closest to the entrance of the box as a guide and “pinching” the corners together is a good way to align the specimen to the guide sheet (Figure 13). Specimen sheets will vary in size; older specimens are often not as wide and on thinner paper stock. Rather than trying to center these specimens align them along the top left corner of the guide sheet closest to the entrance of the PhotoBox (Figure 14). The specimen may not reach the bottom or the right side of the guide sheet but it will be properly oriented when the image is captured.
Figure 13. Aligning the specimen and “pinching” the corners of the guide and specimen sheets will help to ensure that the image captured is not askew. If the specimen is smaller or larger than the guide sheet align the specimen with the left side of the guide sheet at the corner closest to the entrance of the PhotoBox.
Figure 14. Older herbarium sheets are often not as wide as more contemporary ones. Do not try to center the specimen. Align the specimen with the guide sheet closest to the entrance of the Photobox and the image will be properly oriented.

The images are high resolution so any dust or debris on the specimen sheet will be visible in the image. Gently cleaning the specimen sheet to help remove dust and debris is advisable when necessary. Dust and debris will also accumulate inside the PhotoBox and will need to be removed periodically.
If you are using the settings suggested in the getting started area of this guide (using a small thumbnail and order by “Shooting Date/Time”) the last image taken will be the final one in the list of thumbnails in the Digital Photo Professional program. All images will be named with a default “IMG_” name when initially shot but we want each image to be named according to the barcode ID. To do this, select the image and hit the ALT key and “R” to rename the image and the tab key to move the cursor into the file name field. The tab key has to be hit twice for the first specimen of the day because the execute button will be available to select but will not be selectable until the specimen is renamed starting with the second specimen (this is because Digital Photo Professional will not allow two files to have the same name and by default the file name will populate with the last name entered or scanned). The tabs across the top of the window should be set to user generated (Figure 15) but this setting will carry across all sessions so it should not have to be selected each imaging session. Listen for the beep when the barcode is scanned as it will often sound slightly different (typically an elongated beep) if the barcode is not correctly read. If held too long under the scanner sometimes the barcode will be copied twice. Other times, when the barcode is not properly read, symbols will be substituted for some numbers. We have found that waiting to hit execute until you take the next image saves time and allows for a more streamlined process. *This process and the renaming of specimens have been streamlined into the imaging process as of April 2013. Please see the section “Transitioning to using the TCN Imaging Workflow Application”
Figure 15. The name field should be set to User-selected string but this will carry over from each session so it should already be set. After hitting Alt – R to bring up the renaming option the tab key will move the cursor and highlight the file name field.

**Finishing an Image**

After the barcode has been scanned, place the specimen on the left side of the folder and stamp the specimen imaged. Try to stamp the specimen directly under the barcode unless there is reason not to do so (i.e. if there is material or any other obstruction); if the image stamp cannot go under the barcode the next best place is to the left of the barcode or along the top of the specimen Figure 16. Aligning the stamp with the bottom of the barcode will help to place the Imaged stamp in the same place each time. These are fragile specimens; if there is a chance you will damage material trying to place the image stamp under the barcode find a better spot for the stamp. Once the specimen is stamped Imaged place the specimen on the right side of the folder. Do not stamp specimens on top of each other this will damage the specimen(s) under the one being stamped!
Figure 16. The imaged stamp should be placed below the barcode. If this area is not available place the stamp to the left of the barcode or along the top of the specimen sheet.

At this point you have completed one specimen and it is time to move onto the next one. Again do not hit execute (to rename the file) until you go to take the next picture. After the specimen is loaded into the PhotoBox click execute to finish naming the previous specimen and click the shutter to capture the next image.

Focus Failure Issues

From time to time the camera will be unable to focus on a particular specimen. Typically when this happens you will notice that there is an absence of material at the center of the specimen sheet. There is a dime on top of the PhotoBox to help with this process. To trick the camera into focusing on the specimen open the PhotoBox and place a dime in the center of the sheet and hover the mouse cursor over the shutter button. This will cause the camera to focus on the dime. Do not take an image yet. Move from the shutter button to the AF/MF button and switch the button to MF for manual focus (Figure 17). At this point the camera focus will be locked in place. Remove the dime and capture the image. Double check to make sure the image is in focus. When the focus seems correct and the image has been captured switch the focus back to AF and continue imaging.
If the camera has a focus failure it will need to be tricked into capturing the image. To do this, place a dime at the center of the sheet and hover over the shutter and switch the focus from auto to manual. Remove the dime and capture the image.

Recap of the Imaging process:

1) Set the pathway to your folder. Your name should appear below the shutter button.
2) Calibrate the exposure by checking the white square (it should be close to an RGB value of 243,243,243) for the first specimen.
3) Align the specimen with the guide sheet and capture the image.
4) Use ALT – R followed by Tab to bring up the renaming function and place the cursor in the name field, and lastly, scan the barcode.

5) Stamp the specimen as imaged and place it back in the folder.

End of Day

Before closing out the imaging software do a quick check to make sure that all of the images captured have been renamed. Any file that begins with “IMG_” has not been renamed. These files can be renamed by opening up the image writing down the 8 digit barcode and entering this as the file name. Ideally there should not be any files that have to be fixed at the end of the day. At the end of your day do not forget to record your progress in the Excel file. The date imaged field and your initials should be filled in each day next to the specimens you have imaged. Filling out the log helps us to maintain records and make it easier to generate progress reports.

If you are the final imager of the day turn off and replace the cover on the PhotoBox. Also replace the lens cap on the camera, turn it off and put the plastic sheet over the camera to help minimize the accumulation of dust on the imaging station.

Post Processing

Image Correction

Post process is performed using LightRoom presets to fix several items including: color, white balance, sharpness, resolution and exposure as well as adding metadata. The first step in this process is to double check to make sure the images have been renamed correctly and to look for any file that is named with the “IMG_” prefix. The simplest way to check for images that have not been named is to open the public folder where the images are stored and sort the files by name. Any specimen name that contains a letter will appear at the top of the list so any mistakes should appear at the top of the list. If there are any mistakes they should be corrected before post-processing or the change will have to be made three times (on the CR2, the DNG and JPEG). CR2 images will not be readable by the computer outside of Digital Photo Professional (or LightRoom) so you will have to open the image using Digital Photo Professional. Renaming images can be done in a similar manner as before but the barcode will need to be read off of the specimen and corrected in the renaming field. After the images have been checked for letters and IMG prefixes we will spot check the file name and the actual barcode. To do this zoom into the barcode on the first specimen and compare it to the file name that appears on the image. For this process check every tenth image for a match. To move easily through the images in order to capture every tenth one, place the cursor in the image located on the far right end of the tool bar at the bottom of the screen.

The second step in post process is to import the images and to associate metadata and copyright information with each image. This is done using the “Default Herbarium” preset (under the Library tab) and syncing the metadata to all of the imported images in LightRoom (Figure 18 and 19). Images are imported into LightRoom and then the “Default Herbarium” preset is associated with the file. Before
syncing the data across all images the name of the person who captured the image is entered into the Creator field (right below the copyright fields).

Figure 18. Apply Metadata to the images using the Default Herbarium preset which adds copyright and institution information to the image.
Figure 19. Close up of the Metadata process. Default Herbarium is associated with all files. Additionally the imagers name should be added to the Creator field shown in white.

Now that the metadata has been associated with the first image we want to sync the metadata to the rest of the images. Select the image that the metadata was added to (this should be the first image in the row of thumbnails) and hit Ctrl “A” to select all of the images. When all images are selected hit the Sync Metadata button. After the sync process, check a few images to make sure the metadata was applied to all specimens.

The third step is the actual post processing of the images. After the metadata has been associated with the images the first specimen needs to be white balanced (if the first specimen in the row is markedly different (i.e. on a different color card stock) from the rest of the images being processed it may not be the best one to use as a template). To perform the white balance click on the eye drop function and click the same white square used to check exposure during the calibration step. After clicking the white square scroll to the tonal curve and set it to linear. The exposure needs to be adjusted to get the white balance as close to 96.6 as possible. Once white balance is achieved we run the Batch Process user created process to fix sharpness, color, resolution etc. After white balance and the Batch Process have been performed on the first image the process needs to be applied to all of the images. Again select the first image in the sequence and press “Ctrl” “A” to select all images and then click the sync settings button. If all of the images have been synched a “+/−” symbol will appear on each image.
Creating and Storing Processed Images

Once we have added our metadata and run our batch process, our RAW CR2 images need to be converted into a digital negative (DNG) for archiving and a JPEG to be sent off for OCR (optical character recognition) software to database the specimen. We are using the DNG format to archive because it takes less storage space and it is not a proprietary format like the Cannon CR2.

To create both of these files, we need to navigate back to the “Library” tab in Light Room and click the export button. The DNG export process does not require extra settings just select the DNG option under File Settings and Format (Figure 20). The images should be stored in the family folder for that species. Each time the images are exported they should be saved in a folder with the date processed in case they need to be located.

Figure 20. DNG creation for the archive. Make sure the format is set to DNG before selecting export and the pathway is set to save the images in the processed_DNG/family folder.
To create the JPEG files for the OCR, select the JPEG format under “File Settings”. Under “quality select 100” and then under “Image Sizing”, set the resolution to 300 pixels per inch (Figure 21). Again the images need to be saved under the family folder for the species and placed in a subfolder with the date that they were exported.

Figure 21. JPEG export process. With the JPEG export process two settings (quality and resolution) need to be checked prior to export. Quality should read 100 and resolution should be 300 pixels per inch.
Recap order of post process:

1) Associate metadata using default herbarium and adding the imager’s name to the creator field.
2) White balance the first image. Don’t forget to change the tonal curve to linear before adjusting the exposure to reach 96.6.
3) Add the Batch Process on the first image.
4) Select all of the images and sync the white balance and batch process to all images.
5) Export the images to DNG and JPEG and move the old CR2s to the K drive folder containing the processed CR2 files.

Back Up and Data Protection

After images have been post processed and stored on the K Drive they are backed up twice a day as part of the server’s back up process. To further protect our data backups will be made in two different fashions:

1) Our back up drive contains a complete copy of our data. This drive will contain a DNG and a JPEG for each image created. All data are backed up to this drive once per week.
2) Our deep storage contains a copy of each family as they are completed or as sections are uploaded for large families. Our entire dataset will be backed up for the deep storage. This is done in case images become corrupted on the main drives and acts as a point in time to go back to should the need arise.

CR2 files for the time being will be moved to the K Drive but will not be backed up beyond what is being done on the K drive. This is due to their large size and the fact that our DNG files contain the same information.

Imaging protocol from the New York Botanical Gardens was used as the basis for this document and can be found at:

Imaging:

http://tcn.amnh.org/documents

Select Imaging Manual.pdf located under user manuals and help files.

Post Processing