TCN Quarterly Progress Report

TCN Name

Building a global consortium of bryophytes and lichens: keystones of cryptobiotic communities (GLOBAL)[1]

Person Completing the Report

Miranda Zwingelberg (GLOBAL Project Manager)

Share Progress in Digitization Efforts

This report covers progress completed during the period of January 1 - March 31, 2021.

Digitization progress at all GLOBAL institutions continued to be limited by COVID restrictions during 2021-Q1. A few collaborators saw slight increases in access during this period, allowing more collections to begin some digitization activities. Several others are still waiting to start GLOBAL work until conditions are more favorable.

Imaging Equipment

Additional progress was reported in purchasing and setting up imaging equipment during 2021-Q1. COLO purchased basic equipment (color square, blackout paper, etc.) for use with an existing light box. NY purchased a new camera and lens. MSC bought a new PC and began researching camera options. TENN tested out a new camera and lens to use in a second imaging station. F set up two remote label imaging stations and one on-site light box. YU finalized their imaging set up and CINC & MU set up a second imaging station.

---

[1] Throughout this report, herbaria are referred to by their Index Herbariorum acronyms, which correspond to institutional names as follows: ALA = University of Alaska, Fairbanks, ASU = Arizona State University, BRY = Brigham Young University, CINC & MU = University of Cincinnati & Miami University, COLO = University of Colorado, DUKE = Duke University, F = The Field Museum, FLAS = University of Florida, ILL & ILLS = University of Illinois at Urbana-Champaign & Illinois Natural History Survey, LSU = Louisiana State University, MICH = University of Michigan, MIN = University of Minnesota, MO = Missouri Botanical Garden, MSC = Michigan State University, NY = New York Botanical Garden, OSC = Oregon State University, PH = The Academy of Natural Sciences of Drexel University, TENN = University of Tennessee, Knoxville, UC = University of California, Berkeley, WIS = University of Wisconsin, YU = Yale University

Workflow Development

Many institutions spent time in 2021-Q1 continuing to draft, update, and refine their imaging workflows (ASU, F, LSU, MO, MSC, NY, TENN, WIS, and YU). CINC & MU were able to take advantage of a student's photography background to improve their imaging set up. Additional activities included cleaning data (DUKE, MO, and UC) and pre-curation of specimens (F). ASU worked on developing a new lichen and bryophyte label format that can be used to print specimen labels directly onto packets from within the portals. MO reestablished their IPT to allow updates to the Bryophyte Portal. The GLOBAL leadership team organized a meeting of the Specimen Imaging Working Group in February to facilitate sharing between the collaborators.

Essential progress was also made in mapping out centralized georeferencing workflows, especially for those collections shared to the portals as "snapshots." The GLOBAL Project Manager, Portal Manager, and Georeferencing Manager met with the ASU team to review options in February. They then began to schedule individual consultations with collaborators to discuss the flow of coordinate data between the portals and internal databases (DUKE, F, LSU, and MIN). These meetings will continue into 2021-Q2.

Image Uploading

Important progress was made by the IT team at ASU to establish image hosting services and workflows for the submission of GLOBAL images to the Bryophyte and Lichen Portals. External storage was made available to the portals via a mounted directory. An Apache server and the portals were configured to utilize the server mount. A Dropbox group was established to manage requests for each institution and collection type. A manual workflow was developed for transferring images from Dropbox request folders to processing directories within the ASU server system. Scripts were setup and configured to create image derivatives, transfer files to web server space, and link the images to occurrence records within databases. Lichen and bryophyte processing profiles were set up for each institution. Trial uploading of images with this manual work-around were completed successfully (ASU, COLO, DUKE, TENN, WIS) and the option was made available to the entire GLOBAL group in March.

Personnel

The GLOBAL Portal Manager (ASU) and a Lead Digitizer at NY started working on the project in January 2021. F trained employees for their remote stations and an on-site technician. PH hired a full-time co-op student for barcoding and imaging work, who started at

the end of March. MSC and YU made progress hiring undergraduates for summer work. FLAS brought on 4 work-study students.

Digitization

Twelve institutions (BRY, CINC & MU, COLO, DUKE, F, FLAS, ILL & ILLS, LSU, NY, PH, TENN, and WIS) reported progress on digitization deliverables, with a total of 20,062 specimens barcoded (6,680 bryophytes and 13,202 lichens), 10,209 labels imaged (4,009 bryophytes and 6,200 lichens), 5,956 specimens imaged (1,559 bryophytes and 4,397 lichens), 5,881 specimen records uploaded to the portal (2,356 bryophytes and 3,525 lichens), 15,361 skeletal records created (2,618 bryophytes and 12,743 lichens), 14,494 labels fully transcribed (12,831 bryophytes and 1,663 lichens), and 16,702 specimens georeferenced (1,649 bryophytes and 15,053 lichens).

Table 1: Digitization progress by GLOBAL collaborators in 2021-Q1, separated by Bryophyte (B) and Lichen (L) specimens.

|	# Barcodes Added		# Labels Imaged		# Specimens Imaged		# Uploaded to Portal		# Skeletal Records Created		# Fully Transcribed		# Georeferenced	
	B	L	B	L	B	L	B	L	B	L	B	L	B	L
ALA														
ASU														
BRY		350												
CINC & MU	254						254				6,138	45	125	3,250
COLO		2,279		2,279				3,048		3,048		3		
DUKE	1,016		1,231		668	1,834	572	50	329		89		10	
F	1,500	2,624	339	2,624					1,500	1,663				
FLAS	3,765	577	1,551	461			1,258	367	367		62	20		
ILL & ILLS											5,000			
LSU	47	8			3		47	8	8			503		483
MICH														
MIN														
MO														
MSC														
NY	157	7,042	542	514	542	514			157	7,042	1,289	101	757	28
OSC														
PH	121		121		121									
TENN			225		225		225	52	257		253	1	62	
UC														
WIS		322		322		2,049				990		990	695	11,292
YU														
Totals	6,860	13,202	4,009	6,200	1,559	4,397	2,356	3,525	2,618	12,743	12,831	1,663	1,649	15,053
B+L Totals		20,062		10,209		5,956		5,881		15,361		14,494		16,702
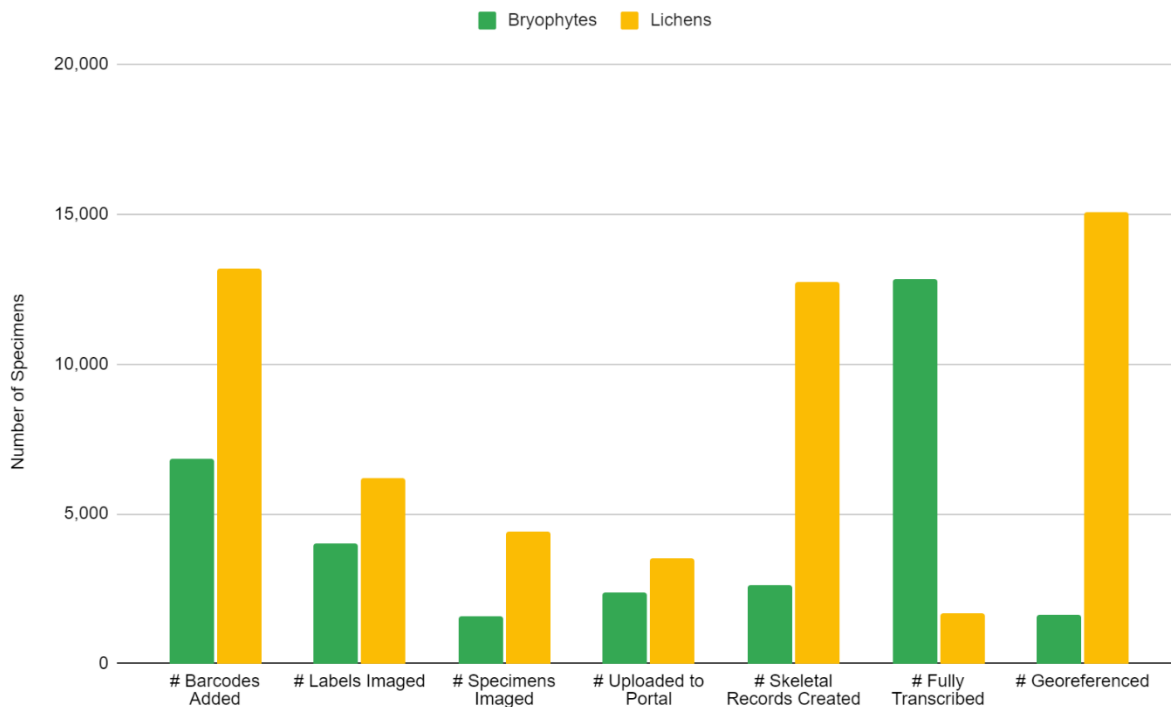
Figure 1: Digitization progress for the GLOBAL collaboration in 2021-Q1, separated by Bryophyte and Lichen specimens.

Share Best Practices, Standards, and Lessons Learned

Flexible Workflows

COVID restrictions continue to require flexible digitization workflows. Some institutions are still using remote imaging stations (F, DUKE), and others are making use of off-site activities like label transcription (COLO, CINC & MU, DUKE, F, and LSU). COLO and F continue to image only labels in order to provide remote transcribers with material. NY is prioritizing barcoding and skeletal data capture along with imaging to take advantage of limited time in their imaging lab. Those collaborators who are imaging specimens have also demonstrated the importance of flexibility. Two main strategies have emerged: photographing the label and specimen separately or removing the specimen and capturing both label and specimen in a single image. CINC & MU originally chose to take two images, but changed to removing the specimen onto archival paper for a single image, which takes less time overall for them. Alternately, PH began

by taking one image, but found that leaving messy specimens with loose dirt inside their packets and photographing them separately from the labels was more efficient. TENN continues to image specimens and labels separately to decrease handling of delicate specimens. DUKE has updated their original workflow for multiple packets per sheet to take one full-sheet photograph and use post-processing to crop to individual specimen labels. NY uses black felt and paper to isolate individual packets on the same sheet for imaging and makes use of an auto-cropping function afterwards. This range of solutions shows that the "optimal workflow" will vary depending on the institution's set up and the storage, composition, and fragility of the specimens. Rates and workflows are expected to continue to be updated as experience and COVID restrictions progress.

Collaboration

Basecamp, Zoom, and email were used to collaboratively problem-solve workflow challenges and share alternative methods and tools across institutions. The working group for Specimen Imaging held one Zoom meeting in 2021-Q1 where participants demonstrated additional workflows and discussed OCR, post-processing, and file formats. There was a preliminary meeting to begin discussion of Skeletal Data standards and tools. The Georeferencing Working Group met twice, once with a smaller focused group and a second time open to all GLOBAL collaborators to discuss tools, resources, and plans. These meetings allow the group to ask questions, offer feedback, and discuss their own challenges.

The first Management Committee Meeting was held in February open to all GLOBAL team members to review quarterly and cumulative grant progress. The GLOBAL Project Manager (TENN) completed initial meetings with several collaborators in early 2021-Q1 (ILL/ILLS, MIN, NY, OSC, YU) to meet the teams and discuss plans and challenges. She also met weekly with the Portal Manager (ASU) to coordinate project activities and needs.

Leveraging Institutional Expertise

Collaborators continued to post test images of specimens on Basecamp for feedback on image content and quality (CINC & MU, LSU, MSC, NY, PH, WIS, and YU). The GLOBAL Portal Manager assessed shared images for appropriate quality and size and GLOBAL participants offered feedback and advice.

Duplicate Matching

Transcription and georeferencing are time consuming steps of digitization. Duplicate specimens, duplicate collecting events, and exsiccate series are common within and across GLOBAL collections. Utilizing these existing records can reduce the number of specimens that need individual attention. In order to take advantage of this resource, all GLOBAL collaborators have been encouraged to use the duplicate matching tools available in the portals. The GLOBAL Portal Manager also shared a script she developed during the Georeferencing Meeting in March. The script matches possible duplicates in the portal and produces a list of duplicate records with coordinates that could be transferred to those that have not yet been georeferenced. The Portal and Georeferencing Managers began discussing use of the script with each collaborator to confirm those interested in this resource.

Another source of possible duplicates that was explored during 2021-Q1 was international collections data. The Project Manager, Lead PI, and Portal Manager worked with the GLOBAL team to compile a list of publically shared international collections that could include transcribed and geolocated specimens duplicated in GLOBAL collections. Seventeen international collections were successfully added into the Bryophyte and Lichen Portals during 2021-Q1, providing over 780,000 bryophyte and 540,000 lichen specimens as possible duplicates.

Exsiccate sets are common in bryophyte and lichen collections, so an accurate list of titles is important to help record and match specimens. Lead PI Budke (TENN) worked to update the Bryophyte Portal list and ASU PI Bungartz cleaned up the library of exsiccate in the Lichen Portal, linking exsiccate titles with IndExs (http://www.botanischestaatssammlung.de/DatabaseClients/IndExs/index.jsp). Best practices were shared on how to link exsiccate specimen records during image and skeletal metadata upload. Exsiccate management tools were reconfigured and improved for advanced management and presentation of the exsiccate datasets (ASU).

Share Identified Gaps in Digitization Areas and Technology

Imaging Uploading

Batch image uploading was flagged as the highest priority during 2020-Q4. In 2021-Q1, ASU continued developing a tool to batch upload images into the portals and a method to automate the transfer of images from Dropbox to the ASU server. The method is expected to

use the Dropbox API and portal controls are being developed to automate or semi-automate the transfer of images into the ASU system. ASU is also developing use of a third party, open source JavaScript library to establish an additional method for batch submission of images directly within the user interface. An interim image uploading workflow was put in place and tested with five institutions during 2021-Q1. It was made available to the whole GLOBAL team in March and will remain in place until the automated version is available.

Barcode Renaming

Another challenge identified was renaming image files with the specimen barcode. Automated renaming is problematic for images taken inside a specimen packet, in which the barcode may not be visible. Additional improvements were made to the "BarcodeRenamer" tool in 2021-Q1 to improve efficiency with different barcode formats and orientations. Development continued on the "PhotoWatcher," a program that will rename images as part of the image capturing routine and can be used to harvest skeletal occurrence metadata. The program is being beta-tested and, once a stable and reliable version is available, it will be shared with all other institutions participating in the TCN (ASU).

Georeferencing

Connecting new specimen records with existing coordinates from duplicate specimens, collecting events, or localities offers a wealth of data to save time on georeferencing for the GLOBAL project. However, making those connections, and assessing the quality of existing georeferences, poses challenges to efficient and effective use of this data resource. In response, ASU worked on developing a georeferencing tool that will harness GEOLocate web services to rapidly georeference batches of specimens with little human effort, as well as a tool to optimize the harvesting of data from specimen duplicates and specimens from duplicate collection events. WIS continued to strategize georeferencing actions across multiple institutions in various stages of progress. Difficulties include inconsistent global geography, completeness of label transcription, and data management among institutions. Questions remain on standards for some geographic fields.

Taxonomic Filters

Specimen data outside of the Bryophyte and Lichen Portals offers the opportunity for additional information as well as possible duplicates to update incomplete records on existing specimens. One challenge is separating the particular records of interest from a larger collection of data. The GLOBAL Portal Manager and ASU team worked to develop taxonomic

filters to isolate bryophyte records from plant data and separate lichenized, lichenicolous and allied fungi from occurrence records of non-lichenized fungi.

Share Opportunities to Enhance Training Efforts

Symbiota

ASU created new versions of the Symbiota checklist tutorial video and a tutorial on using the new label generation tools in Symbiota portals. All videos now available on the Symbiota YouTube channel (https://www.youtube.com/channel/UC7glMVLRnTA6ES3VTsci7iQ). A guide for use of the portals' Skeletal Data Entry tool was developed between the GLOBAL Portal Manager (ASU) and the Project Management team at TENN.

Transcription

The GLOBAL Project Manager began compiling transcription resources during 2021-Q1 to share on Basecamp and the project website. Resources were submitted by several collaborators including CINC & MU, COLO, LSU, and WIS. A student worker at CINC recorded a video on techniques for deciphering and translating non-English, hand-written herbarium labels (https://www.youtube.com/watch?v=M_rPFzPfZTM&feature=youtu.be). LSU shared a guide to Brazilian road markers and TENN uploaded a copy of a guide to Japanese geography developed by a former student. COLO project manager Allen recorded a series of videos on how to install and use the Imaging Workflow Application from the LBCC TCN. Several institutions under the project are reviewing this system for the project. He also recorded a video to help students transcribe from images in the portal.

Georeferencing

The GLOBAL Management team led a Georeferencing Meeting with the GLOBAL group to review available georeferencing resources and tools. WIS continued to post resources on shared pages for georeferencing help sites and training manuals and began meeting with participating institutions to discuss georeferencing readiness and strategies.

Share Collaborations with other TCNs, Institutions, and/or Organizations

F connected with members of the TCN's External Advisory Committee to confirm their invitations, roles, and availability to participate. All responses were positive and the Executive Committee plans to meet to discuss how to best engage this resource.

ASU, in collaboration with the International Union for the Conservation of Nature (IUCN), established an agreement to share a Global Checklist of Red-Listed Lichen Species through the portal. That list is now online at: https://lichenportal.org/cnalh/checklists/checklist.php?clid=1430&pid=0. Species listed as part of the Global Fungal Red-List Initiative will be added as soon as their assessments are published. Detailed occurrence records of red-listed species in the portal are protected and accessible only to a select group of data managers and conservation specialists.

Lead PI Budke (TENN) sent emails out to contacts at twenty international collections chosen to utilize as data snapshots for duplicate matching in the portals. The correspondence introduced the GLOBAL project and made some initial connections between the project and the broader international community. All feedback received was very positive and collections will be acknowledged on the GLOBAL project website.

The GLOBAL Project Manager (TENN) reached out to the NSF Program Officer to discuss and confirm reporting requirements and COVID related extension protocols. She also participated in the BIO-wide Virtual Office Hours: COVID-19 Response on March 2.

The GLOBAL Lead PI and Project Manager (TENN) participated in the February iDigBio Quarterly IAC meeting to connect with other active TCN's.

Share Opportunities and Strategies for Sustainability

Portal Management

Katie Pearson was hired as the GLOBAL Portal Manager and was introduced to the community. She was provided with necessary logins and data management access and reached out to existing portal contributors. She will be responsible for providing support and monitoring the help emails of the Lichen and Bryophyte Portals (ASU).

ASU established SSL security certificates for the Bryophyte and Lichen Portals and three associated portals (Arctic Lichen Flora, Frullania, and Líquenes en América Latina), which involved remapping some of the internal images to use the https protocols. Various image URLs were also remapped to https protocol for institution that have recently installed SSL certificates. They will continue to encourage and assist other institutions hosting external images mapped into the portal to install SSL certificates.

ILL & ILLS transitioned the MyCo Portal to the same code base, Symbiota-light, that is used for the Bryophyte and Lichen Portals.

Taxonomy

A meeting to discuss plans for Bryophyte taxonomy was held in February involving participants from ASU, DUKE, F, MO, and TENN. An updated version of the nomenclature backbone for the Bryophyte Portal was completed by MO and is in the process of being integrated by ASU. Data maintenance was completed on the taxonomic thesaurus in the lichen consortium, and new names continued to be added and linked with synonyms and basionyms (ASU).

Share Education, Outreach, Diversity, & Inclusion (EODI) Activities

F and TENN worked on an abstract (on behalf of all TCN participants and institutions) to be submitted in April 2021 to the online conference: Bryophytes, lichens, and northern ecosystems in a changing world (July 6-9, 2021). The Québec B(ryophytes) and L(ichens) (BL2021) will bring together four major bryological, lichenological and botanical societies: the International Association of Bryologists (IAB), the American Bryological and Lichenological Society (ABLS), the Canadian Botanical Association (CBA-ABC) and the Société québécoise de bryologie (SQB).

ASU participated in discussions on how to best present the new portal to the international research community (with the goal to have more international communities join). Data management agreements were developed with international collections joining the Lichen Portal, particularly collections from South America. ASU PI Bungartz continued outreach to lichenologists from Latin America by organizing a Symbiota workshop (in Spanish) for the Consorcio de Herbarios de Líquenes en América Latina during the 9th Symposium of the

International Association for Lichenology, in Brazil August 1-6; the two day workshop is being planned as post-congress event: https://doity.com.br/ial9/blog/workshops.

The GLOBAL Project Manager met with TENN OIT staff for training on website development and management in February. She began building the GLOBAL WordPress website in March. Project information, progress, and resources have been added and the site is to be launched publicly in April.

Share Information About Your Website and/or Portal Usage

The GLOBAL Project Manager built the GLOBAL WordPress project website in March, to be launched publicly in April.

Share Other Activities and/or Progress

Imaging Tagging

ASU made the first steps towards developing an online glossary and image library as part of comprehensive revision of character matrix that drives identification keys, with the goal of tagging images with controlled vocabulary of key characters.