Liberating Arctic botanical biodiversity data at the Canadian Museum of Nature



Jennifer C. Doubt, Lisa C. Gualtieri, Cassandra M. Robillard, Lyndsey A. Sharp, Paul C. Sokoloff, Jeffery M. Saarela

Botany Section and Centre for Arctic Knowledge and Exploration, Canadian Museum of Nature, Ottawa, ON, Canada. Contact: jdoubt@nature.ca

Introduction

Core to the polar research information spectrum are the millions of specimens in natural history collections. These specimens document the distribution of species in time and space, and often include additional ecological information, providing a wealth of biodiversity data. They also serve as vouchers for the datasets that underpin scientific conclusions, allowing future workers to confirm or revise identifications. Finally, these specimens are ongoing sources of new data – including genetic information - as they are continually re-used and re-interpreted in addressing new questions over the course of centuries.

Natural history museums face a massive challenge in databasing and imaging their collections, in order for them to be widely discovered, shared, used and reused in research and outreach. This challenge is commonly augmented by unprocessed specimen backlogs: specimens collected and stored, sometimes many decades ago, that have never been accessioned or prepared for long-term use, and that are essentially invisible to collection users, and unavailable for study.

In 2017, a philanthropic donation by the Sitka Foundation to the Canadian Museum of Nature (CMN), made possible the large scale digitization of Canadian Arctic (Yukon Territory, Northwest Territories, Nunavut) holdings.

National Herbarium of Canada



The National Herbarium of Canada, at the CMN, houses a growing collection of 1.25 million specimens of wild bryophytes, lichens, vascular plants and algae collected between 1766 and the present day. About 20% of the collection is databased so far. Staff estimate the Arctic portion (Canadian Arctic, as well as Alaska and Greenland) of the collection to represent about 100 000 specimens.

CMN scientists currently lead the collaborative production of an Arctic Flora of Canada and Alaska (http://arcticplants.myspecies.info/), continuing a tradition of Arctic botanical research and collection that dates to the earliest incarnations of the institution. Consequently, the National Herbarium of Canada provides a key international resource on Arctic plants over the past 200+ years.





Herbarium technician, Lyndsey Sharp (right) and student assistant Junyi Meng prepare a specimen.

Student assistant Rachel Bergeron files specimens in the National Herbarium of Canada.



Student assistant Annie Dicaire prepares to put a specimen into the light box, which will minimize shadows when she captures an image with the camera above.

Handling tens of thousands of specimens for digitization provided the opportunity to upgrade them with tracking barcodes (as shown) and repairs.

Still in the newsprint in which they were pressed decades or more ago, these backlog specimens are safe but inaccessible to herbarium users.

Digitization Project Goals and Staff

With the understanding that available resources could not support full imaging and databasing for all Arctic holdings, project staff identified the need to, at minimum, make all specimen objects "discoverable" both by those who manage the collection and by those who use it. Accordingly, the following goals (in priority order) were set for digitizing Arctic herbarium specimens:

- 1. Capture specimen images and link them to complete or partial database records, and make them publicly available on the internet, via the Global Biodiversity Information Facility (https://www.gbif.org).
- 2. Implement specimen barcodes and taxonomic updates to facilitate collection management and use
- 3. Release (accession, mount) 5000 backlogged Arctic specimens collected by Sylvia Edlund (1980s) and Margaret Oldenburg (1940s) for digitization and use
- 4. Enter detailed specimen data (from images) and georeferenced collecting localities

Project planning and problem-solving is shared among three collection staff with broad ranging collection management responsibilities, and two full-time project staff. Digitization is carried out by the two project staff who are supported by two full-time new undergraduate students each semester. In concert with digitization, a third undergraduate student is hired to process backlogged specimens, and guide them into the digitization workflow.

Workflow and Equipment

Like all collections embarking on digitization projects, CMN staff designed and continually refine a setup (Table 1) and workflow (Figure 1) tailored to priorities, physical configuration, staffing, and equipment unique to their collection.

Table 1: Digitization Equipment for each of two workstations

Copy Stand	Kaiser Base Board only, 800x600mm (31.5 x 23.6 x 1.3")
	Kaiser Extra Tall Column for RSX
	Kaiser Extension Arm
	Kaiser Copy/Camera Arm RA1
Lighting	Orte LED Pro Box Plus 1419 NY Botanical Garden Modified [discontinued]
Camera	Sony A7R II
Lens	Sony Sonnar T* FE 55mm F/1.8 ZA
Camera accessories	Sony ACPW20 AC ADAPTER
	Micro USB cable (6') tether
Other accessories	Tiffen Q-13 colour separation guide (20 cm)
	Scale-bar
Computer*	Dell Optiplex 9020 D07S Workstation (Intel Core i5, Gen 4, i5-4590, Quad-core, 3.30 GHz, 8 GB
	DDR3 SDRAM, 500 GB HDD)
Monitor	BenQ BL2711U 27" LED LCD
Camera control	CaptureOne Pro Sony 10
software	
Image processing	CaptureOne Pro Sony 10
software	
Data management	Barcodefiler
software	Microsoft Excel
	Windows PowerShell

*In addition to the computers associated with the imaging stations, one Lenovo ThinkStation P320 30BH005SUS Workstation (Intel Core i7, Gen 7, i7-7700K Quad-core, 4.20 GHz, 16 GB DDR4 SDRAM, 512 GB SSD) is used to speed the processing of large batches of images and associated data.

Figure 1: Workflow Summary for digitization of Arctic herbarium specimens at the National Herbarium of Canada, Canadian Museum of Nature





Student assistants Rachel Bergeron (right) and Lyn Vakulenko and sort willow specimens.

Arctic specimen of River Beauty, from 1991, made accessible in 2018.



re-discovered in 2018.

Digitization and Project Benefits

Convenient, worldwide stakeholder access

- Stakeholders can search and browse holdings, and download data any time over the internet (https://www.gbif.org)
- Specimen image remains available even when specimen is in use

Responsible management of public assets

- For the first time, an inventory of constituent objects in a major part of the herbarium exists
- Systematic handling for digitization facilitates collection upgrades (e.g., annotations, barcoding)
- CMN newly holds digital copies of assets previously available only as unique hard copies
- 'Crowd' power becomes accessible: more errors can be noticed and reported; duplicate records at different herbaria can be created/upgraded using shared digital assets, rather than duplicate effort
- Staff time shifts increasingly from service (e.g., consulting the collection to respond to external inquiries) to other, neglected priorities (e.g., collection development, conservation)

Experience and employment for students in biology and museum studies

• Students, handling thousands of specimens and scrutinizing their labels, gain knowledge of Arctic plant taxonomy, geography and history, and experience in collection management

WITH THANKS to the Sitka Foundation, and to undergraduate project staff Rachel Bergeron, Brigid Christison, Emily Dang, Annie Dicaire, Erin Johnston, Jasmine Lai, Junyi Meng, Magean Ng, Galyna Vakulenko, Wilson Wang; CMN digitization leads Shannon Asencio and David Shorthouse; CMN Advancement team; CMN Research and Collections team.