Journal Pre-proof

Best practices for digitizing a wood slide collection: The Bailey-Wetmore Wood Collection of the Harvard University Herbaria

Madelynn von Baeyer, John M. Marston

PII: S1040-6182(20)30518-8

DOI: https://doi.org/10.1016/j.quaint.2020.08.053

Reference: JQI 8487

To appear in: *Quaternary International*

Received Date: 12 June 2020

Revised Date: 18 August 2020

Accepted Date: 30 August 2020

Please cite this article as: von Baeyer, M., Marston, J.M., Best practices for digitizing a wood slide collection: The Bailey-Wetmore Wood Collection of the Harvard University Herbaria, *Quaternary International* (2020), doi: https://doi.org/10.1016/j.quaint.2020.08.053.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier Ltd.

Journal Pre-proof

Best Practices for Digitizing a Wood Slide Collection: the Bailey-Wetmore Wood Collection of the Harvard University Herbaria

Madelynn von Baeyer[a]* and John M. Marston[b,c]

a Harvard University Herbaria, 22 Divinity Avenue, Cambridge, MA 02138, USA. madelynn_vonbaeyer@fas.harvard.edu

b Program in Archaeology, Boston University, 675 Commonwealth Avenue, Boston, MA 02215, USA

c Department of Anthropology, Boston University, 232 Bay State Road, Boston, MA 02215, USA

* Corresponding author

Abstract

As herbaria move to digitize their collections, the question remains of how to efficiently digitize collections other than standard herbarium sheets, such as wood slide collections. Beginning in September 2018, the Harvard University Herbaria began a project to image and digitize the wood slides contained in the Bailey-Wetmore Wood Collection. The primary goal of this project was to produce images of the wood tissue that could be used for specimen-level research and to make them available on the internet for remote scholarship. A secondary goal was to establish best practices for digitizing and imaging a microscope slide collection of tissue sections. Due to the size of the wood slide collection (approximately 30,000 slides), a medical histology scanner and virtual microscopy software were used to image these slides. This article outlines the workflow used to create these images and compares the results with digital resources currently available for wood anatomy research. Prior to this project, the very little of the Bailey-Wetmore Wood Collection was cataloged digitally and none of it was imaged, which made access to this unique collection difficult. By imaging and digitizing 6,605 slides in the collection,

Journal Pre-proof

this project has demonstrated how other institutions can make similar slide collections available to the broader scientific community.

Keywords: Wood slides, virtual microscopy, slide digitization, slide imaging, herbarium collections

Journal Pre-proof

Journal Pre-proof

1. Introduction

In the past 10 to 15 years, herbaria around the world have committed themselves to digitizing and imaging their collections to make them more accessible to the public (Nelson et al., 2015; Sweeney et al., 2018; Tegelberg et al., 2014; Thiers et al., 2016; Tulig et al., 2012). This has opened up collections that previously were only accessible on-site to researchers across the globe. The bulk of these digitization projects have focused on developing efficient digitizing and imaging protocols for herbarium specimen sheets to maximize the number of specimens digitized as quickly as possible with the highest accuracy (Nelson et al., 2015; Sweeney et al., 2018; Tegelberg et al., 2014; Thiers et al., 2016; Tulig et al., 2012). Herbarium collections, however, are not solely composed of herbarium sheets: they contain numerous subcollections that have unique scientific value, e.g. economic botany, historical wood slide, and paleobotanical collections. Such subcollections are valuable in scientific research outside of botany and ecology, the fields that have traditionally made the most use of herbaria collections. For example, the field of anthracology, the study of ancient wood charcoal, requires access to reference collections of wood tissues, both historical and modern. Unfortunately, only a few herbaria digitization and imaging projects have included collections besides herbarium sheets, such as slides; moreover, when slides are digitized, they may not be imaged in a way that allows for specimen level research, prohibiting further scientific use of the digitized collection (Allan et al., 2019; Decker et al., 2018; Heerlien et al., 2015; Musson et al., 2020).

To address this problem, a project was initiated to digitally image and publish wood slides in the Bailey-Wetmore Wood Collection at the Harvard University Herbaria (HUH). The goal of this project was to design a workflow and a product that preserved the ability to conduct specimen-level research using digital slides. Very little has been published on how to digitize

Journal Pre-proof

and image microscope slides efficiently and at a high resolution, so our workflow was created in part as a blueprint for other institutions that may be considering a similar project. In order to create an efficient workflow, this project relies on a high-throughput histology scanner than uses virtual microscopy to image slides. Virtual microscopy is a technique that creates a digital image of a microscope slide that allows for the same high-resolution microscopy work possible using a normal bright-field microscope with the original slide. Virtual microscopy was primarily developed for medical uses and has been shown to be an effective tool in research and education (Kumar et al., 2004; Lundin et al., 2004; Mione et al., 2013). With this being a new technique in slide imaging, very few other herbaria have yet employed this approach, with the exception being the Royal Botanic Gardens, Kew in the United Kingdom (Musson et al., 2020). To date, however, no institution has published an in-depth workflow using this technique. In this article, we demonstrate that virtual microscopy is also beneficial in digitizing and imaging herbarium slide collections, thereby preserving the scientific value of herbarium collections in digital form.

The Bailey-Wetmore Wood Collection provides an ideal test case for virtual microscopy of wood slides. The collection has a worldwide geographic distribution, making it an important resource for researchers around the globe. Furthermore, increased industrialization in many parts of the world since the late 1800s has impacted the wood ecology of areas sampled in this collection. It is probable that the Bailey-Wetmore Wood Collection contains specimens of currently rare, and even extinct, species that were more commonly available to people living hundreds or thousands of years ago, so is of particular interest to anthracologists. Finally, while the HUH has an open access policy for its collections, access to the wood collection was still in-house only. Therefore, this project increases accessibility of an under-utilized, difficult-to-access resource of significant utility to a wide variety of researchers.

Journal Pre-proof

2. Digitized Wood Slide Resources

Relatively few wood collections, and wood slides specifically, have been digitized to date, especially in comparison to the number of global wood collections (Lens et al., 2016). The primary resource for digital microscopic wood anatomy is the InsideWood database (InsideWood, 2004–onwards; Wheeler, 2011). InsideWood has over 36,000 images of modern and fossil wood compiled from over 15 different institutions, although the bulk of the collection is from the David A. Kribs wood collection of North Carolina State University. Accompanying these images are over 5,800 anatomical descriptions representing at least 7,000 species worldwide (Wheeler, 2011). The anatomical descriptions are keyed using the "IAWA List of Microscopic Features for Hardwood Identification" (IAWA Committee, 1989) and the "IAWA List of Microscopic Features for Softwood Identification IAWA" (IAWA Committee, 2004), and the database is searchable based on taxonomic name and anatomical features. This is a powerful tool for identifying unknown wood specimens. The database, however, does have some limitations. The images are a combination of older 35-mm film photographs from negatives and more recent digital images. In most of the images, only a portion of the entire slide is shown which may or may not show all the anatomical features associated with the species. Furthermore, while InsideWood has considerable geographic breadth, the collection is mainly centered on North American species.

Several other digitized photographic atlases and identification keys exist online as well and are linked through the InsideWood (2004–onwards) webpage. The most similar database to InsideWood is the Dendrochronological Database, a database primarily of European wood species (Schmatz and Heller-Kellenberger, as of March 3, 2005) and the Database of Japanese Woods by the Forestry and Forest Products Research Institute (FFPRI Wood Identification

Journal Pre-proof

Database Team, as of October 15, 2019), both of which are searchable by anatomical features, as defined by the IAWA, and taxonomic names. Otherwise, most of these photographic atlases are digitized copies of photographic plates found in previously published books that are accompanied by anatomical descriptions. A few are searchable using taxonomic names. A few herbaria with large wood collections, like the Turvuren Xylarium, Oxford University, the Smithsonian, and the USDA Forest Products Laboratory have digitized and selectively imaged their xylaria (wood collections) as well. The goal of this digitization is primarily to make the contents of these collections public, not to create a digital platform for research. Therefore, the images of these specimen are a mix of macroscopic images of wood sections, herbarium sheets, and microscopic images with no real system of what type of image, or images, are associated with what specimen.

This lack of systematically imaged wood collections is a result of many constraints. All herbarium imaging projects have to balance the importance of three different variables: cost, efficiency (i.e., scanning the largest number of slides in the shortest time, including the amount of post-processing required for each image), and quality (specifically the suitability of the final image to be used for specimen-level research). When considering these factors for imaging herbarium sheets, the solution is relatively straightforward since the most cost-effective imaging technique, macro-photography, can also produce images suitable for non-destructive specimen-level research; therefore the main constraints become efficiency of both imaging and post-processing, a challenge for which several research groups have developed clever solutions (Nelson et al., 2015; Sweeney et al., 2018; Tegelberg et al., 2014; Thiers et al., 2016; Tulig et al., 2012). When imaging a wood collection, however, it is necessary to capture both macroscopic and microscopic images to display the full array of diagnostic features which requires both

Journal Pre-proof

129 macro- and micro-photography (IAWA Committee, 1989, 2004; Ruffinatto and Crivellaro,
130 2020). This shifts consideration to whether or not it is cost effective and/or efficient to create
131 images suitable for specimen-level research.

132 For slide imaging projects, there have been a variety of solutions that weigh each of these
133 constraints differently. The slide imaging projects from the Naturalis Biodiversity Center,
134 Netherlands (Heerlien et al., 2015) and the Natural History Museum, London (Allan et al., 2019)
135 prioritized efficiency and cost by digitizing hundreds of thousands of slides in a short amount of
136 time, with a limited number of staff, by taking overhead photos of batches of 100 slides and
137 using optical character recognition (OCR) to automate metadata capture. This approach relies on
138 strict staging and label protocols to digitize the record of each specimen. The final output is a
139 comprehensive digital record of a large portion of the collection, but does not include images of
140 slide contents sufficient for specimen-level research. In contrast, the Virtual Microscope Slide
141 Collection (VIRMISCO) project at the Senckenberg Museum of Natural History Görlitz (Decker
142 et al., 2018) prioritized creating a collection of images that can be used for specimen level
143 research using virtual microscopy. Their project was not designed to be efficient since their
144 digitization protocols do not seem to be designed for large batch digitization and Decker et al.
145 (2018) caution researchers that the time to create a set of z-stacked images for virtual microscopy
146 takes around an hour per slide. In terms of output, the VIRMISCO collection contains hundreds
147 of specimens instead of hundreds of thousands (Decker et al., 2018).

148 Creating digital images for specimen-level research from slides will never be as fast as
149 overhead macro-photography, but there are ways to increase the efficiency of this approach, such
150 as through the use of a medical histology scanner for bulk slide imaging. This is the approach
151 that Kew Gardens, UK (Musson et al., 2020) took to image their 130,000+ wood slide collection.

Journal Pre-proof

The Kew project scanned 100 slides at a time using a Zeiss Axio Scan microscope to create a z-stacked set of images that were subsequently converted into an image with an extended depth-of-field (EDF) during post-processing. This approach creates flat TIFF images that were uploaded into the Kew database (Musson et al., 2020). These images do display some larger diagnostic anatomical features, but like photographic atlases, these static EDF images obscure the smallest diagnostic anatomical features. To capitalize on creating as many images as possible that can be used for specimen-level research, the approach adopted to image the wood slides in the Bailey-Wetmore Wood Collection combines aspects of the VIRMISCO project and the project at Kew by using a histology scanner to scan batches of slides at a time to increase throughput while using virtual microscopy to allow for specimen-level research.

3. Description of Collection

The Harvard University Herbaria Bailey-Wetmore Wood Collection is comprised of approximately 30,000 dry wood block specimens; 30,000 microscope slides of thin sections of wood tissue; 15,000 to 20,000 slides of pollen, nodal preparations, leaves, and seeds; and approximately 1,000 wood remains preserved, or embalmed, in spirits. The specimens in the wood collection were amassed primarily by Dr. Irving W. Bailey and Dr. Ralph H. Wetmore and their students between 1912, when Bailey was first appointed at the Harvard School of Forestry, until 1982, the last year for which a recorded gift is accessioned (Wetmore, 1974; Wetmore et al., 1974). The main purpose of this collection was to provide evidence that evolutionary and taxonomic trends are reflected in wood anatomy (Torrey, 1994; Wetmore, 1974; Wetmore et al., 1974). To this end, the wood specimen collection is large and cosmopolitan. However, metadata records exist only for part of the collection, primarily the dried wood blocks, in a paper-based card catalog with a unique numbering system that was never formally associated with accession

Journal Pre-proof

numbers used by the HUH (Wetmore et al., 1974). No inventory of the slides was available in Herbaria records and not all slides that are in the Bailey-Wetmore Wood Collection received a wood collection number, especially slides that were donated as part of a personal herbarium. We estimate that only 85% of the slides in the collection are entered into the card catalog.

Accordingly, in 2016 a survey of genera present in the wood slide collection was conducted by co-author Marston and members of his Environmental Archaeology Laboratory at Boston University, finding that the slide collection contains 30,826 slides representing 2,766 genera. The bulk of the slide collection, which was created primarily by sectioning dried samples in the collection, includes species from the Americas (a maximum of 19,882 slides and 1,353 genera; as many genera appear on multiple continents, these figures represent maximums) and Asia (a maximum of 17,979 slides and 1,118 genera). Collections from Africa (a maximum of 10,500 slides and 769 genera) and Oceania (a maximum of 8,016 slides and 532 genera) are smaller, while Europe (a maximum of 4,860 slides and 197 genera) is the smallest assemblage. Ecologically, the collection is strongest in genera from the tropics (a maximum of 24,295 slides and 2,096 genera) and subtropics (a maximum of 15,912 slides and 1,273 genera), with the fewest genera from temperate regions (a maximum of 10,079 slides and 582 genera). Similar statistics are not available for the dried wood specimens, but as many of the slides were produced from these specimens, we hypothesize that the dried wood specimens likely share a similar geographic dispersion.

This diversity of the wood collection was a deliberate goal of both Bailey and Wetmore. Their research questions on the relationship between evolutionary and anatomical traits were broadly applied and therefore a global wood collection was necessary (Wetmore, 1974). To this end, wood specimens were collected from around the world through: 1) collection trips, such as

Journal Pre-proof

198 Edmund W. Sinnott and Arthur J. Eames' trip to Australasia in 1910–1911 and Wetmore's
199 collection trips with R. H. Woodward, a graduate assistant, to Soledad, Cuba and the Canal Zone
200 Biological Area on Barro Colorado Island, Panama in 1929 (Wetmore et al., 1974); and 2)
201 formalized or semi-formalized purchasing relationships between institutions and individuals,
202 such as the relationship Bailey and Wetmore had with Professor Samuel J. Record at the Yale
203 School of Forestry in the 1930s and 1940s (Wetmore, 1974; Wetmore et al., 1974), and the
204 relationship between the Arnold Arboretum, Syracuse University, and the North Carolina State
205 University School of Forestry in the 1960s and 1970s that made new microscope slides of
206 samples of American woods from the Bailey-Wetmore wood collection and Syracuse University.
207 The wood slide portion of the Bailey-Wetmore Wood Collection grew through two other
208 methods as well: 1) gifts of complete collections from specific collectors, notably H. F. Munroe
209 of Chicago, whose slides have dates from the 1890s, and Albert Hanford Moore, who was a
210 student at Harvard at the beginning of the 20th century (A.B. 1905, A.M. 1906, Ph.D. student
211 until 1909); and 2) by purchasing slide sets from forestry laboratories such as that of Sul Ross
212 State University and the Forest Products Laboratories of Canada, at McGill University in
213 Montreal.

214 The result of these varying collection methods is a historical, cosmopolitan wood slide
215 collection that spans almost a century from the late 19th century into the late 20th century.
216 Conventions for the production of microscope slides of wood thin sections did not change
217 appreciably during these 100 years. Most slides in the collection consist of glass 1 in × 3 in
218 microscope slides with a permanently mounted glass or plastic coverslip. The slides have been
219 stored vertically in metal trays that are designed to hold four slides each and fit into the drawers
220 of a slide cabinet. Some differences in size and shape are apparent, however. In particular, the

Journal Pre-proof

thickness of both slides and coverslips varies substantially. Slides vary in their contents, but the vast majority contain at least one section of wood: the transverse section, radial section, or tangential section, either alone or mounted with one or (more often) the two other sections. A small number of slides contain macerated wood preparations. Most tissues on the slides (especially those prepared after 1900) have been stained, although the staining media were not recorded for the vast majority of slides. Pre-1900 slides have a higher percentage of unstained or very lightly stained tissue. While the exact mounting medium is unknown for most of the slides, it is assumed based on the number of slides that have yellowed with age that Canada balsam was the most frequently used mounting medium (Brown, 1997). The slides mounted with balsam have aged relatively well. Significant yellow discoloration can be observed on many slides, but it does not impede one's ability to study or image the tissue anatomy. For the slides mounted between approximately 1968 and 1972, the same is not true. During this time, the mounting medium changed to a solution, probably a gum chloral medium (Brown, 1997), that has since crystalized grossly, impeding one's ability to observe anatomical characteristics. Luckily, the bulk of the collection is made up of slides that were prepared using Canada balsam. There is no standard layout for label placement or information recorded on the label across the collection, though slides sourced from a particular collector or vendor are generally internally consistent.

4. Methods

Below, we discuss the approach for slide digitization. To clarify these methods, our discussion of the methodology is divided into task clusters (modified from Nelson et al. (2012)): 1) sample selection, 2) pre-digitization curation and staging (referred to as curation and staging), 3) specimen image capture (slide imaging), 4) data capture (metadata digitization), and 5) specimen image processing (image processing). Each of these task clusters is discussed in detail

Journal Pre-proof

below to illustrate the applicability of each cluster for future projects of this nature.

4.1. Sample Selection

This first phase of the digitization project was limited to two years, making it clear that not all 30,000 slides in the collection could be imaged. Thus, we targeted discrete taxonomic groups for initial imaging to facilitate future slide image searches and to provide a solid framework for the future digitization and curation of the rest of the collection. The gymnosperms (2,827 slides) and monocots (341 slides) were imaged in entirety (Table 1). The dicot angiosperms, however, comprise the remainder of the collection and were too numerous to scan entirely, so a family-by-family approach was adopted with an initial focus on families found in southwest Asia, prioritizing large cosmopolitan families to increase the usefulness of this project for scholars who work in all regions. Paving the way for future work, it is now simple to tell whether or not a slide has been imaged, as all imaged slides were given QR codes—a new addition to the wood slide curation.

Table 1: List of families imaged with slide totals

Gymnosperms **2,827 slides**	**Monocots** **341 slides**	**Angiosperms** **3,437 slides**
Abietaecea 1,439 slides	Alismaceae 2 slides	Fabaceae (Caesalpinaceae) 463 slides
Araucariaceae 32 slides	Amarylliadaceae 15 slides	Fabaceae (Mimosaceae) 304 slides
Cephalotaxaceae 21 slides	Araceae 44 slides	Fabaceae (Papilionaceae) 754 slides
Cupressaceae 431 slides	Commelinaceae 3 slides	Fagaceae 536 slides
Cycadaceae 28 slides	Cyperaceae 1 slide	Linaceae 180 slides
Ginkoaceae 22 slides	Dioscoreaceae 8 slides	Platanaceae 42 slides
Gnetaceae 182 slides	Flagellariaceae 1 slide	Polygonaceae 75 slides
Taxaceae	Graminae	Ranunculaceae

Journal Pre-proof

61 slides	29 slides	97 slides
Taxodiaceae 611 slides	Iridaceae 3 slides	Rhamnaceae 102 slides
	Juncaceae 26 slides	Rosaceae 335 slides
	Liliaceae 153 slides	Thymalaeaceae 17 slides
	Musaceae 3 slides	Ulmaceae 188 slides
	Orchidaceae 1 slide	Umbelliferae 10 slides
	Palmae 26 slides	Urticaceae 118 slides
	Pandaceae 7 slides	Verbenaceae 159 slides
	Restionaceae 6 slides	Vitaceae 57 slides
	Typhaceae 4 slides	
	Zingberaceae 9 slides	

4.2. Curation and Staging

Slides were transported in their file cabinet drawers to a curation and staging area to be cleaned, curated, and staged. The materials used in the curation and staging included cardboard slide folders that hold 20 slides, microfiber cloths, very thin forceps, preprinted 2D QR code labels, jade glue, and temporary slide tags.

Journal Pre-proof

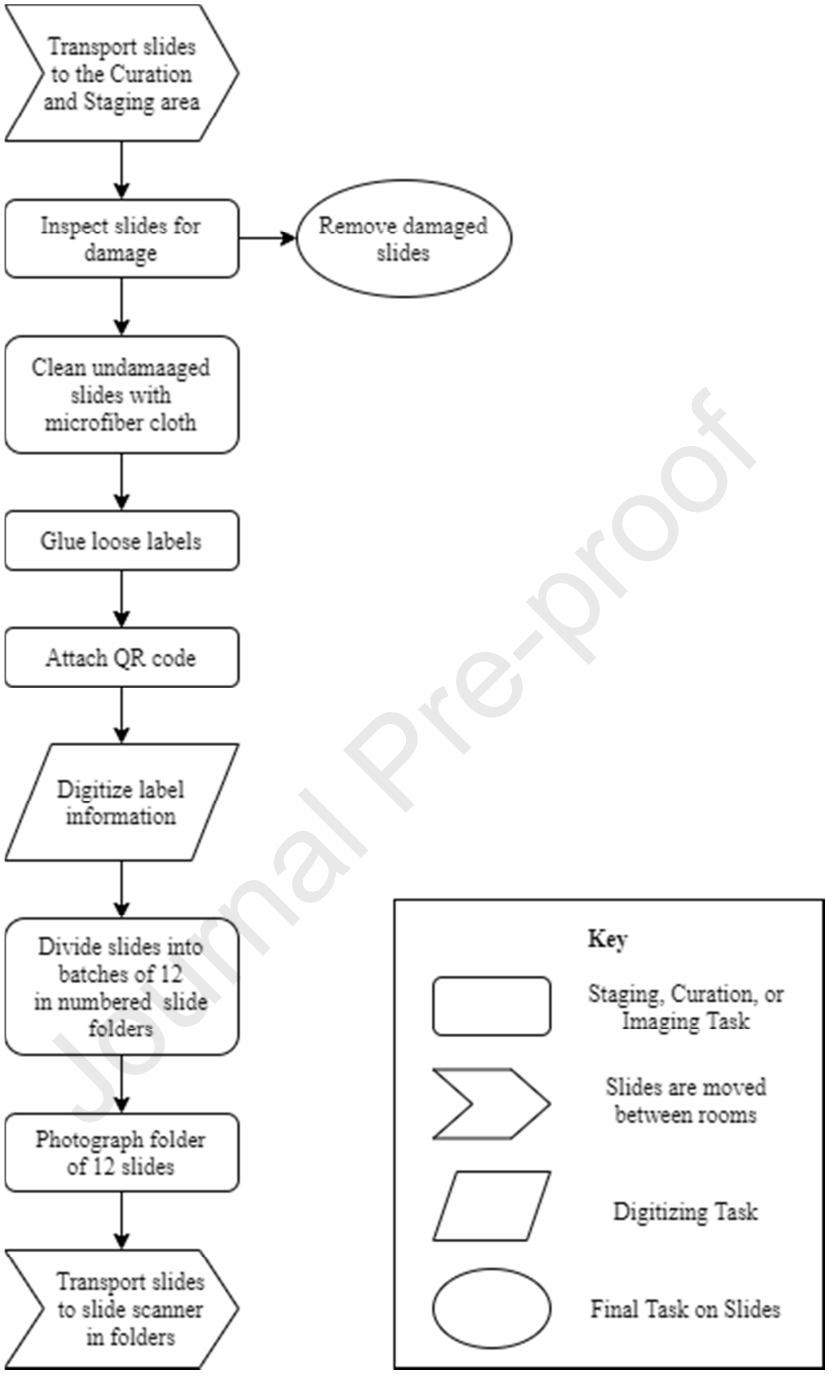

Figure 1: Workflow for the curation and staging task cluster of the project

To prepare a batch of slides for imaging, the corresponding slide holders were removed from the drawer for slide cleaning and staging to keep as much of the original organization of the

Journal Pre-proof

Journal Pre-proof

drawer as possible. Each individual slide was inspected to see if the coverslip or slide itself was broken. If the slide was broken in a way that compromised the tissue or made it impossible to place in a tray for imaging, the slide was placed in a broken slide folder, while a colored tag with the folder and slot number of the cardboard folder replaced the slide in the drawer. Each slide was then cleaned with a microfiber cloth to remove dust and accumulated dirt, with particular attention to the surfaces around the mounted tissue. Accumulations of Canada balsam on the slide surfaces were only removed (mechanically) if they impeded the placement of the QR code label.

Once the slide was cleaned, the original paper label was checked to ensure secure attachment. If the label was loose, a small amount of jade glue was applied to the label to reattach it using a superfine microbrush. Finally, a 4.8 mm × 4.8 mm label with a 2D QR code and a numeric printing of the QR code was applied to the slide with forceps. The QR codes are archival quality with a 1 mil matte polypropylene lamination on a 3 mil matte white polypropylene facestock and a 1 mil acrylic special adhesive designed to adhere to slides and coverslips. Since there is no standard layout for the slides in the collection, it was impractical to designate a standard placement of the QR code. QR codes were never placed on the original paper label. Whenever possible, the QR code was placed directly on to the glass slide, or alternatively on the coverslip. Ideally, the QR code was adhered to an area that was visible when the slide was in its vertical slide holder in the drawer to reduce the amount of handling necessary to read the QR code when the slide is in the drawer.

Journal Pre-proof

Figure 2: The tools used for slide cleaning and QR code label attachment; on left is a roll of pre-printed QR codes.

Once the QR code was attached to the slides, groups of 12 slides were staged in numbered cardboard folders and photographed in order to create a digital image of the full slide and label. Each folder held up to 12 slides from only one family, yet may hold multiple genera and preserved the order found in the drawer. A temporary slide tag with the folder and slide slot number was inserted into the slide holder from the drawer to mark the slide's placement in staging. These tags were not replaced until the slide had finished being imaged. To mitigate

Journal Pre-proof

298 delay caused by longer-than-average imaging, a second set of folders and slide tags were created
299 to use when necessary.

300 Once the batches of slides were in the folders, an overhead photo of the folder with the
301 slides was taken with a DSLR Canon EOS 5D Mark III with a Canon EF 50 mm f/2.5 Compact
302 Macro Lens in a stationary photo box, the NYBG Modified MK Direct Photo ebox 1410, to
303 ensure that the label information on the slide was captured digitally. This was necessary because
304 the variability of slide layouts in the collection often made it impossible to capture the slide label
305 information with the histology scanner, which was designed to capture information from modern,
306 uniform medical slides. Furthermore, the overhead image created an important reference of the
307 order of each slide in a batch, which became a necessary tool in the post-processing stage.

308 These overhead images were then uploaded into the HUH Wood Slide Image Collection
309 in Slide Atlas 3.0.5, an open source data management platform that is powered by Girder 3.1.0,
310 built on the Resonant platform by Kitware (Grauer et al., 2016; Mullen, 2016). In Slide Atlas, the
311 HUH Wood Slide Collection folders were organized by family, scanning date, and folder number
312 per scanning date. The overhead folder images were uploaded into the appropriate folder number
313 by date to allow for easy reference during subsequent imaging and post-processing stages.

Journal Pre-proof

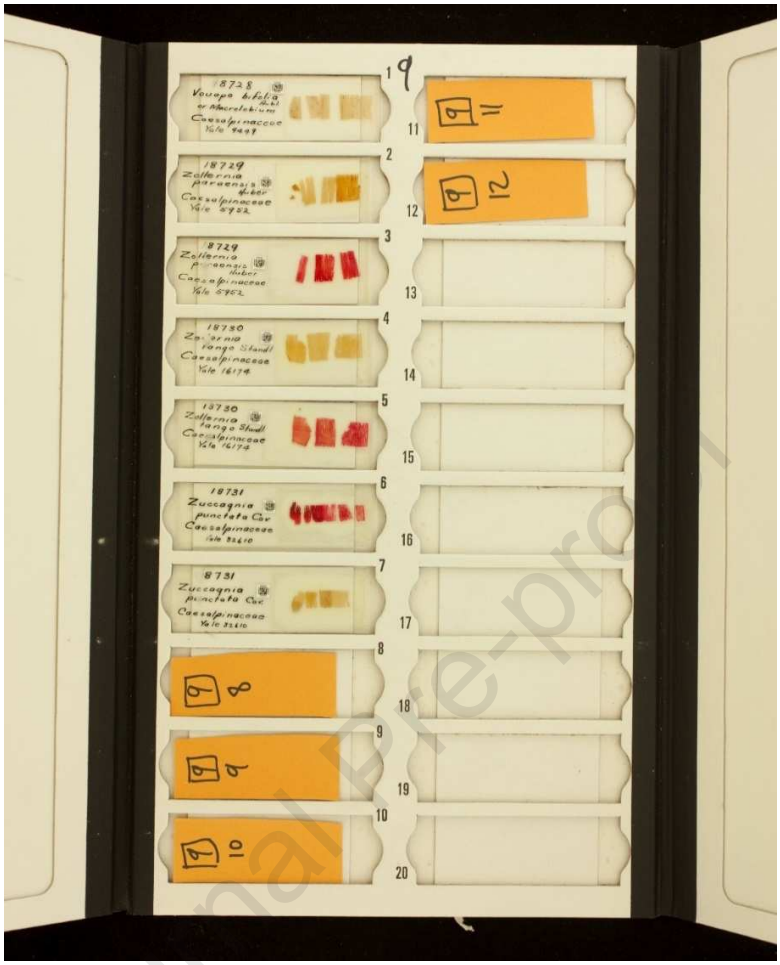

Figure 3: Image of a staging folder with the slides of a batch and the temporary slide tags labeled with the folder and slide slot numbers.

4.3. Slide imaging

The scanner used to image the Bailey-Wetmore Wood Collection slides was the Huron Digital Pathology Model LE 120 scanner and associated software that was purchased jointly by the Harvard University Herbaria, the Harvard Museum of Comparative Zoology, and the Harvard Center for Brain Science. The Huron LE 120 has the capacity to scan 120 (10 trays of 12) 1 in × 3 in microscope slides in one session using brightfield microscopy. The slides can be scanned at 10x, 20x, and 40x magnifications with a 0.4 μm/pixel resolution at 20x and 0.2

Journal Pre-proof

μm/pixel resolution at 40x. The scanner scans bands of the identified tissue and then stitches the bands together to create the whole slide scan in a non-proprietary 24-bit RGB pyramidal big TIFF format (Huron Digital Pathologies, 2018). This file format allows researchers to view the images in most image management software, including with the accompanying Windows-based HuronViewer version 1.3.1 and the web-based Slide Atlas. Both platforms can dynamically zoom up to 160x and offer simple image annotations.

Journal Pre-proof

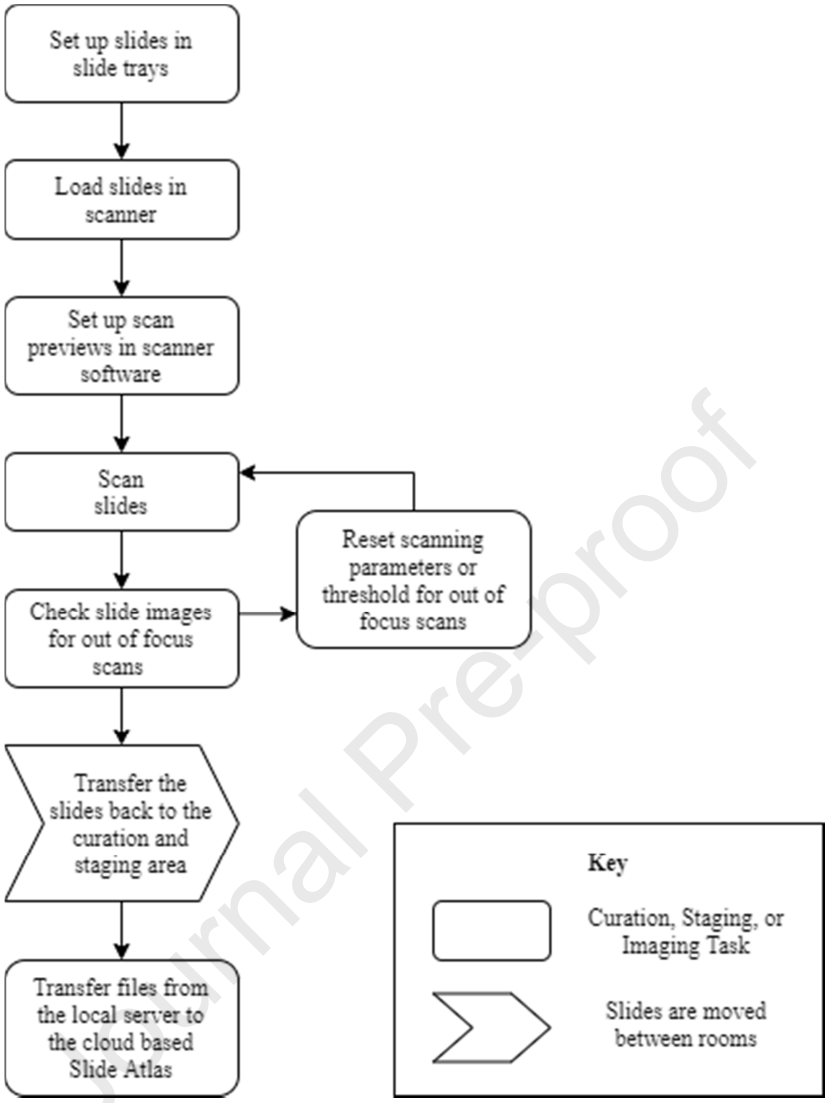

Figure 4: Workflow for the slide imaging task cluster.

Imaging the Wetmore-Bailey Wood Collection slides starts by transferring the batch of 12 slides from a folder to a slide tray that accommodates 12 1 in x 3 in slides. Each slide in the tray was checked to see the entire tissue was visible or if some part of the tray cut off portions of the tissue and if the slide thickness was too large for the standard tray. If either of these were true, the slide was moved out of the standard tray and placed on a custom holder that Stephen Turney from the Center for Brain Science designed specifically for this purpose. The custom tray

Journal Pre-proof

holder had a clear base and could hold four slides with the dimensions up to 1.5 in wide and 3.5 in tall. To create these custom holders, which were 3D printed on-site at the Center for Brain Science, Huron supplied Turney with the original tray designs and the RFID tags necessary for the scanner to recognize new trays. Turney was able to design a new tray style in CAD, order the metal parts online, and 3D print the frame of a new tray holder. Turney then added a new slide tray definition to the Huron scanner software that defined the new dimensions to scan. In order to preserve the same numbering within a batch from staging to imaging, if a slide did not fit within the standard 12 slide tray, that slot was skipped in the tray. This ensured that when the tray was scanned, the slide number of the image files generated corresponded to the slide numbers in the overhead batch photo.

Once all the slides were loaded into trays, the trays were loaded into the scanner and the three scan settings were applied to all the trays: file prefix, magnification, and z-stacking. Each slide was scanned with the prefix HUH. This prefix can be set to any value, but for this project a simple prefix was chosen to increase flexibility for the project. Each slide was scanned using 10x brightfield microscopy to create a set of z-stacked images. A set of z-stacked images are images of the same object taken at slightly different focal planes along the z-axis, the object-to-camera axis (El-Gabry et al., 2014; Olkowicz et al., 2019). This was necessary because most wood tissue slides have uneven topology such that the entirety of the tissue is not in focus along any one plane. By experimenting with sets of z-stacks from exemplar slides from the Bailey-Wetmore Wood Collection, it was determined that 5 scans with 4 μm spacing between each scan, for a total depth of 16 μm, captured the majority of variability on most slides in the collection.

A standard scanning magnification and z-stack setting was chosen for this project instead of employing multiple objectives and z-stack configurations for the slides to increase the

Journal Pre-proof

efficiency of the project. The scan time per slide in the Huron LE 120, and in all digital pathology scanners, is dependent on the objective as well as the number of z-stack slices scanned for each slide. Because z-stacking was necessary to produce images suitable for specimen level research, the slides were scanned on the lowest objective, 10x, to decrease scanning time. A set number of z-stacks was used to streamline the scanning process and to reduce the scanning time of each slide as much as possible. By using the lowest objective and a minimum number of z-stack slices, the amount of digital storage needed for each set of files was reduced as well.

It is important to note that because this project utilizes virtual microscopy software, these standardized scan settings did not compromise the ability to capture minute anatomical features of wood anatomy. The digital microscopy software built into Slide Atlas allows for further magnification of the image, up to 160x, which allows researchers to magnify the wood anatomy up to 1600x. Therefore, minute anatomical features necessary for genus and species level identification that are observable only at magnifications higher than 500x are observable in the scanned images (see Figure 6 for an example).

Once these settings were applied, the scanning area of each slide was previewed in the scanner software on the local Windows 7 PC and focal points and white balance for each slide were applied onto a preview image generated by the scanner. Focal points were distributed fairly evenly within the boundaries of the tissue section. No focal points were placed in areas of air bubbles, areas where the mounting medium had thickened, or areas under dried mounting medium over the coverslip. On slides where the mounting medium had crystalized in part, the slide was still imaged and focal points were applied to areas not under crystalized medium. On slides where the entire medium had crystalized, the focal points were placed following the normal protocol. Finally, the white balance point for each slide was set at a point with the

Journal Pre-proof

lightest shade and the least contrast of the area covered by the coverslip. Once these focal points and the white balance were set for all slides in a tray, the tray was added to the scanning queue. The queue to scan was started once the focal points and white balance was set for the first tray.

The time it took to image a slide varied depending on the size of tissue being scanned, but typically a full batch of 60 slides took about 15 hours to scan. After this initial scan, the images were reviewed to see if they were in focus with minimal stitching problems. Each slide that was not adequately scanned was rescanned with an adjusted scanning threshold or with different focal points, as needed to correct the image. The selective rescanning process was highly variable in duration, but on average took around 6 to 8 hours. Once the rescanning process was over, the slides were removed from the scanner, returned to their cardboard staging folders, and returned to the curation and staging area to return to their metal slide holders in the drawers.

The final step in the imaging workflow was to move the scanning images from storage on a local server onto the data management platform Slide Atlas. Slide images from standard-sized trays were uploaded into the batch folder created during the staging process, while slides from custom-sized trays needed to be placed back into position in their original (digital) location. Using the overhead batch photo and matching the slide numbers of the image files, it was possible to identify where each slide from the custom trays belonged; those slide images were then moved into the correct folders and renamed to match the rest of the tray. The z-stack of images created for each slide was named using the built-in naming convention for the scanner that included the HUH prefix, the tray number (assigned sequentially for every HUH tray), the slide number, and the slice number.

Once the images were moved to Slide Atlas, they were stored on Harvard University's Research Computing cloud that can accommodate the large data storage needs of this project.

Journal Pre-proof

Two terabytes of storage were reserved for these images. Using Slide Atlas, or another similar visualization software, is necessary for these images because they are created as Pyramid TIFF files. Pyramid TIFF files use jpeg compression to store multiple bitmaps of the same image at different spatial resolutions, requiring specific visualization software to view (Library of Congress, as of February 22, 2017). This file format allows for digital microscopy and dynamic zooming, and significant compression without compromising resolution that allows these files to remain under a gigabyte (GB) each. When these files are transformed into flat TIFFs that can be read by simpler image software, the size of the file jumps to 7-10 GB per slice, per z-stack.

4.4. Metadata digitization

Metadata, here defined as including accession number, QR code, taxon name, collector and collection information, geographical information, and other details about the plant collected, was digitized in two stages. The first stage directly followed slide cleaning and the attachment of the QR code and involved digitization of the information on the slide label into a tab-delimited spreadsheet. The information found on the physical slide typically included some or all of the following: the QR code number, taxon name, a Harvard or Arnold Arboretum Wood number, a previous collection number from other collections (such as Yale or the New York State College of Forestry), initials of the slide preparator, and date of slide preparation.

The second phase of digitization occurred after imaging and consisted of cross-referencing the Harvard or Arnold Arboretum Wood number (for all slides where one was present) against the card catalog. If a record was present for that specimen in the card catalog, any additional metadata present on the card was transcribed into the spreadsheet. This record frequently included details on the collector, geographic location of collection, and plant specimen information. Cross-referencing the Wood number also served as a quality control

Journal Pre-proof

check on the label data, reducing human error in transcription. We did consider using optical character recognition (OCR) to automate entry of slide labels and card catalog entries, but ultimately decided against it, since the lack of standardization for information included in those records, and inconsistencies in the position of information and in handwriting, would have necessitated close manual reinspection of every record, thus taking more time than simple transcription.

Ideally, metadata digitization would have been entered directly into the existing database of records at HUH, Specify. However, each new Specify record requires a minimum set of metadata to create an entry: taxon name, collector, and geographic region. Since there is no full inventory of the slide collection or a complete record of metadata for each slide, and it was clear that collector and geographic region information were not available for every slide, it was decided to bypass Specify as the primary database for this project. Instead, all metadata available for each slide in the project was recorded in a tab-delimited spreadsheet, which can be directly integrated into Specify or its successor database in the future.

4.5. Image processing

The final step of this project was to process both the overhead images taken in staging and the z-stack images to create seven views of the slides: an overhead, macro-photographic view of the slide, five micro-photographic z-stacked images, and one composite extended-depth-of-field (EDF) micro-photographic image.

The overhead slide image was a cropped image taken from the batch image taken during staging. The cropped image was made using Inselect, a desktop application designed specifically for natural history collections by the Natural History Museum, London that automates cropping a single image of multiple specimens into multiple images of individual specimens (Hudson et al.,

Journal Pre-proof

2015). In Inselect, one can delineate bounding boxes for cropping around each specimen and save that configuration of bounding boxes to apply to other photos with a similar layout. Furthermore, Inselect can read QR codes and allows for automated (with OCR) or manual input of metadata that is then associated with the cropped image (Hudson et al., 2015). For this project, the full batch photo was uploaded into Inselect and each slide was cropped. The QR code was read and embedded into the metadata for the cropped photo. The cropped slide photo was then renamed with the acronym of the herbarium with which the slide is associated, the QR code number, and a view designator: e.g., A12345678_b wherein A is the herbarium acronym, 12345678 is the QR code, and _b is the view designator.

The EDF image was created using a custom python script written by Stephen Turney that uses the Girder 3.1.0 API (Kitware, 2014-2018) and the "focusstack" python script by Charles McGuiness (2015) to pull all the z-stack slice images in a batch folder on Slide Atlas, convert the Pyramid TIFFs to flat TIFFs, create the EDF image from each set of z-stacks (up to 12 in a batch), and upload the new image back into the batch folder. This process was designed to run on a local PC server with 128 GB of RAM and is run in two stages, combining the first three slices and then the output and the last two slices to create an EDF image. When all five slices were combined at once, there was not enough RAM available to create an EDF for many slides. EDF images can sometimes contain processing artifacts, e.g., shadows that do not appear on the original photos or halos around elements. These processing artifacts could obscure anatomical features of the wood tissue. To mitigate this issue, the original, uncombined slices that do not contain processing artifacts are also made available to view.

The final step of image post-processing was to change the file names of the z-stacked and EDF images from the names assigned by the scanner software to names designated by HUH.

480 This was accomplished with another python script designed by Turney in conjunction with the
481 Girder API to replace the old image names with the new names composed of the herbarium
482 acronym, QR code number, and view designator (as described above); this information was
483 automatically pulled from a tab-delimited spreadsheet.

484

485 **5.** Results

486

487

Journal Pre-proof

Journal Pre-proof

Figure 5: Screenshot of EDF image of a slide in Slide Atlas showing the full scan of the slide *Larix europaea*, A01913288_b_composite, HW 10896 from the Bailey Collection, collected in Europe. A) the full extended-depth-of-field image in Slide Atlas; B–D) details in Slide Atlas of B) tangential section, c) radial section, and d) cross section. ***Print in color***

Journal Pre-proof

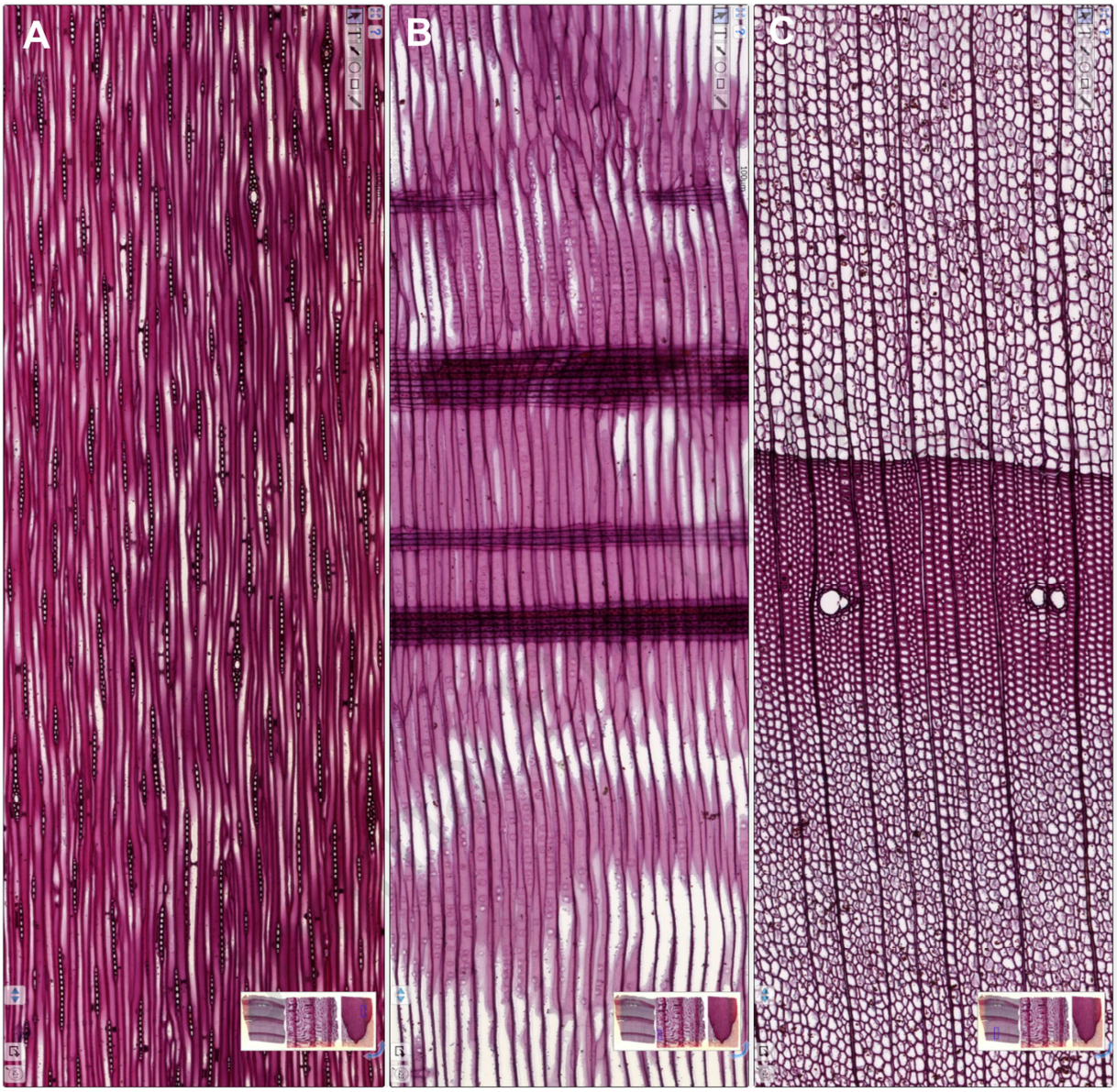

Figure 6: Multiple EDF images of slide *Larix europaea*, A01913288_b_composite, HW 10896 from the Bailey Collection, collected in Europe. A) tangential section, B) radial section, and C) cross section. ***Print in color***

The workflow outlined above has resulted in 6,605 slides imaged between January 7, 2019 and March 6, 2020, about 20% of the total collection. Seven views of each slide are available. The slides are organized by the taxon name listed on the slide, with no effort to update

Journal Pre-proof

these to follow contemporary nomenclature, so a knowledge of taxonomic synonyms is necessary to use the collection. A future stage of this project, however, will explore assigning current synonyms automatically through cross-referencing with an authoritative taxonomic source, such as IPNI, the International Plant Names Index (IPNI, 2020). The scanning output described above represents the full-time work of one person, with a part-time student assistant, and sporadic but significant contributions from a biodiversity informatics specialist and an imaging specialist. Beginning Fall 2020, all of the imaged slides will be available through the Harvard University Herbaria website with a link to the metadata; an example is shown in Figure 5 and 6.

Our workflow was primarily focused on creating digitized images that permit specimen-level research, rendering a previously little-known collection accessible to researchers worldwide. The methods developed here demonstrate that is possible to scan microscope slides efficiently to create high-resolution images suitable for specimen-level research. While we anticipate further efficiency improvements in metadata digitization and image processing workflows, we do not believe that it is possible to fully automate this process for a historical, heterogeneous collection like the Bailey-Wetmore collection. This workflow was particularly efficient due to the ability to use custom slide trays in the histology scanner. Without this feature, the approximately 10% of the collection with slides that did not fit into the standard trays would have been poorly imaged or not imaged at all.

The level of efficiency developed here was dependent on the purchase of a costly high-throughput histology scanner. Not only are these scanners quite expensive to purchase, but they require a technician on site to maintain them, which creates a very high initial cost for projects using this workflow. This project was also dependent on the availability of a powerful personal

Journal Pre-proof

computer, server space to host the necessary imaging software, and substantial amounts of remote storage for the hosted images.

A tradeoff was also made in image quality. The final images created by this workflow were designed to be sufficient for research, but not all images are perfect. Each image is oriented based on how it was mounted on the slide which usually does not conform to correct anatomical orientation. No time was taken to edit the images to remove dirt captured by the scan, or to ensure perfect band stitching, although obviously mis-scanned slides were rescanned. This decision was made because since the entire slide was scanned, a relatively large amount of tissue is available for each specimen so the goal was to ensure that some portion of the entire scan is clean and focused, and thus appropriate for research and publication, barring any other slide-wide quality issues like crystalized mounting medium.

6. Discussion and Conclusions

The primary goal of this workflow was to create a digital product that researchers can use in a very similar way to physical slides, thus increasing accessibility of the collection. We argue that digitizing slides without digital microscopy is a wasted opportunity, as low-resolution images of slides do not improve collection access. Because of the large size of the slide collection in the Bailey-Wetmore Wood Collection, it was most efficient to create a workflow that uses a high-throughput slide scanner to digitize and image these slides. This scanner was quite costly, but worth considering for every institution with a large slide collection. Additionally, as the scanner resides in a shared facility, multiple campus units contributed to its purchase and upkeep costs, making the budget more feasible for HUH. Since this workflow does not need constant access to the scanner, the corresponding time sharing did not prove a

Journal Pre-proof

significant bottleneck for this project. There are other options that enable virtual microscopy, however, for institutions with smaller wood slide collections, namely benchtop slide scanners that will scan one or two slides at time. This approach is not as efficient on a large scale, and it may not have the same flexibility in the size of slide trays and holders, but it is several orders of magnitude less expensive to acquire such equipment. We would only recommend such an approach for collections that hold fewer than 1000 slides.

One limitation of this project, however, is the limited metadata available with each specimen. When compared to other digital wood resources available, especially Inside Wood (InsideWood, 2004–onwards; Wheeler, 2011), the Dendrochronological Database (Schmatz and Heller-Kellenberger, as of March 3, 2005), or the Database of Japanese Woods (FFPRI Wood Identification Database Team, as of October 15, 2019), the Bailey-Wetmore Wood Collection slide images are more difficult to use to identify unknown wood species, since the images are not searchable by anatomical feature. In this way, the digital Bailey-Wetmore Wood Collection is comparable to the digital imaging projects of the Turvuren Xylarium of the Royal Museum of Central Africa, Oxford University, the Smithsonian, and the USDA Forest Products Laboratory. This project, however, has a few key advantages over those peers. First, the number of imaged slides is much larger than these four databases. Second, the entire slide was imaged, rather than only small areas of tissue available in most digital wood resources, including InsideWood. Imaging the entire slide increases the possibility of capturing rare and hard-to-see anatomical features, which are not often present in cropped photos used in other digital repositories. Finally, efforts were made to preserve the geographic diversity of the collection, even in this first digitization and imaging phase. The Bailey-Wetmore Wood Collection is designed to be a comprehensive, global collection and so families were imaged in full, thereby capturing the

Journal Pre-proof

species diversity of every family as captured in the collection. This is especially important because this is a historical collection that includes species that are rare today.

We designed the workflow described here to demonstrate the technology necessary to digitize and image slide collections in a way that maximizes the scientific potential of the collection, while being efficient and adhering to the curation standards of the Harvard University Herbaria. The project was successful on all three of these counts. In future stages of this project, we aim to expand imaging to the entire slide collection, to cross-reference archaic taxonomic designations with current taxonomic standards, and ultimately to implement a searchable anatomical key based on IAWA standards. The Bailey-Wetmore Wood Collection is a worldwide collection of significant scientific value that we aim to make accessible to the worldwide community of scholars, dependent only on access to a computer and internet connection.

7. Acknowledgements

The authors would like to acknowledge Michaela Schmull, Jonathan Kennedy, Stephen Turney, and HUH Directors Charles C. Davis and Elena Kramer for their help and guidance on this project. This project would not have been possible without the work of Emily Brown and Kathleen Depina. The funding for this project was provided by a Research Fellowship from the Harvard University Herbaria awarded to Madelynn von Baeyer.

Journal Pre-proof

Journal Pre-proof

590 **8.** References

591 Allan, E.L., Livermore, L., Price, B., Shchedrina, O., Smith, V., 2019. A Novel Automated Mass
592 Digitisation Workflow for Natural History Microscope Slides. Biodiversity Data Journal 7, 1–15.
593 https://doi.org/10.3897/BDJ.7.e32342.
594
595 Brown, P., 1997. A review of Techniques used in the preparation, curation and conservation of
596 Microsciope slides at the Natural History Museum, London. The Biology Curator 10, 1–33
597
598 Decker, P., Christian, A., Xylander, W.E.R., 2018. VIRMISCO—The Virtual Microscope Slide
599 Collection. ZooKeys 741, 271–282. https://doi.org/10.3897/zookeys.741.22284.
600
601 El-Gabry, E.A., Parwani, A.V., Pantanowitz, L., 2014. Whole-slide imaging: widening the scope
602 of cytopathology. Diagnostic Histopathology 20, 456–461.
603 https://doi.org/10.1016/j.mpdhp.2014.10.006.
604
605 FFPRI Wood Identification Database Team, as of October 15, 2019. Accessed on April 17,
606 2020, http://www.ffpri.affrc.go.jp/en/database.html.
607
608 Grauer, M., Rose, L., Choudhury, R., 2016. Understanding the Resonant Platform, April 20,
609 2020, https://blog.kitware.com/the-resonant-platform/.
610
611 Heerlien, M., Van Leusen, J., Schnörr, S., De Jong-Kole, S., Raes, N., Van Hulsen, K., 2015.
612 The Natural History Production Line: An Industrial Approach to the Digitization of Scientific
613 Collections. Journal on Computing and Cultural Heritage (JOCCH) 8, 1–11.
614 https://doi.org/10.1145/2644822.
615
616 Hudson, L.N., Blagoderov, V., Heaton, A., Holtzhausen, P., Livermore, L., Price, B.W., van Der
617 Walt, S., Smith, V.S., 2015. Inselect: Automating the Digitization of Natural History Collections.
618 PloS one 10, e0143402. https://doi.org/10.1371/journal.pone.0143402.
619
620 Huron Digital Pathologies, 2018. TissueScope LE120 Slide Scanner,
621 http://www.hurondigitalpathology.com/wp-content/uploads/2018/10/LE120-Brochure-October-
622 2018-1.pdf.
623
624 IAWA Committee, 1989. IAWA list of microscopic features for hardwood identification. IAWA
625 Bulletin n.s. 10, 219–332
626
627 IAWA Committee, 2004. IAWA list of microscopic features for softwood identification. IAWA
628 Journal 25, 1–70. https://doi.org/10.1163/22941932-90000349.
629
630 InsideWood, 2004–onwards. Published on the Internet. Accessed on April 20, 2020,
631 http://insidewood.lib.ncsu.edu/search.
632

Journal Pre-proof

IPNI, 2020. International Plant Names Index. Published on the Internet. The Royal Botanic Gardens, Kew, Harvard University Herbaria & Libraries and Australian National Botanic Gardens, Accessed on June 1, 2020, http://www.ipni.org.

Kitware, 2014-2018. API Documentation, Accessed on April 24, 2020, https://girder.readthedocs.io/en/stable/api-docs.html.

Kumar, R.K., Velan, G.M., Korell, S.O., Kandara, M., Dee, F.R., Wakefield, D., 2004. Virtual microscopy for learning and assessment in pathology. Journal of Pathology 204, 613–618. https://doi.org/10.1002/path.1658.

Lens, F., Lynch, A.H., Gasson, P.E., 2016. Index Xylariorum 4.1, Accessed on April 16, 2020, https://globaltimbertrackingnetwork.org/products/iawa-index-xylariorum/.

Library of Congress, as of February 22, 2017. TIFF, Pyramid, Accessed on April 21, 2020, https://www.loc.gov/preservation/digital/formats/fdd/fdd000237.shtml.

Lundin, M., Lundin, J., Isola, J., 2004. Virtual microscopy. Journal of Clinical Pathology 57, 1250–1251

McGuiness, C., 2015. focusstack, Accessed on April 24, 2020, https://github.com/cmcguinness/focusstack.

Mione, S., Valcke, M., Cornelissen, M., 2013. Evaluation of virtual microscopy in medical histology teaching. Anatomical Sciences Education 6, 307–315. https://doi.org/10.1002/ase.1353.

Mullen, Z., 2016. Girder 2.0 officially released, Accessed on April 20, 2020, https://blog.kitware.com/girder-2-0-officially-released/.

Musson, A., Reed, L., Bojarska, M., Fulcher, T., 2020. Digitising Kew's microscope slide collection, Accessed on March 27, 2020, https://www.kew.org/read-and-watch/digitising-microscope-slide.

Nelson, G., Paul, D., Riccardi, G., Mast, A., 2012. Five task clusters that enable efficient and effective digitization of biological collections. ZooKeys 209, 19–45. https://doi.org/10.3897/zookeys.209.3135.

Nelson, G., Sweeney, P., Wallace, L.E., Rabeler, R.K., Allard, D., Brown, H., Carter, J.R., Denslow, M.W., Ellwood, E.R., Germain-Aubrey, C.C., Gilbert, E., Gillespie, E., Goertzen, L.R., Legler, B., Marchant, D.B., Marsico, T.D., Morris, A.B., Murrell, Z., Nazaire, M., Neefus, C., Oberreiter, S., Paul, D., Ruhfel, B.R., Sasek, T., Shaw, J., Soltis, P.S., Watson, K., Weeks, A., Mast, A.R., 2015. Digitization workflows for flat sheets and packets of plants, algae, and fungi. Applications in Plant Sciences 3, 1500065. https://doi.org/10.3732/apps.1500065.

Journal Pre-proof

679 Olkowicz, M., Dabrowski, M., Pluymakers, A., 2019. Focus stacking photogrammetry for micro-
680 scale roughness reconstruction: a methodological study. Photogrammetric Record 34, 11–35.
681 https://doi.org/10.1111/phor.12270.
682
683 Ruffinatto, F., Crivellaro, A., 2020. Atlas of Macroscopic Wood Identification: With a Special
684 Focus on Timbers Used in Europe and CITES-Listed Species. Springer International Publishing
685 AG, Cham, Switzerland. https://doi.org/10.1007/978-3-030-23566-6.
686
687 Schmatz, D., Heller-Kellenberger, I., as of March 3, 2005. Dendrochronological Database,
688 Accessed on April 16, 2020,
689 https://www.waldwissen.net/waldwirtschaft/waldbau/wachstum/wsl_dendrochronological_datab
690 ase/index_EN.
691
692 Sweeney, P.W., Starly, B., Morris, P.J., Xu, Y., Jones, A., Radharkrishnan, S., Grassa, C.J.,
693 Davis, C.C., 2018. Large-scale digitization of herbarium specimens: Development and usage of
694 an automated, high-throughput conveyor system. Taxon 67, 165–178.
695 https://doi.org/10.12705/671.9.
696
697 Tegelberg, R., Mononen, T., Saarenmaa, H., 2014. High-performance digitization of natural
698 history collections: Automated imaging lines for herbarium and insect specimens. Taxon 63,
699 1307–1313. https://doi.org/10.12705/636.13.
700
701 Thiers, B., Tulig, M., Watson, K., 2016. Digitization of The New York Botanical Garden
702 Herbarium. Brittonia 68, 324–333. https://doi.org/10.1007/s12228-016-9423-7.
703
704 Torrey, J.G., 1994. Ralph H. Wetmore 1892–1989; a biographical memoir, Bibliographical
705 Memoirs. National Academy of Sciences of the U.S.A., Washington, pp. 421–436
706 Tulig, M., Tarnowsky, N., Bevans, M., Anthony Kirchgessner, B.M., Thiers, M., 2012.
707
708 Increasing the efficiency of digitization workflows for herbarium specimens. ZooKeys 209, 103–
709 113. https://doi.org/10.3897/zookeys.209.3125.
710
711 Wetmore, R.H., 1974. Irving Widmer Bailey, 1884–1967; a biographical memoir, Biographical
712 Memoirs. National Academy of Sciences of the U.S.A., Washington, pp. 21–56
713 Wetmore, R.H., Barghoorn, E.S., Stern, W.L., 1974. The Harvard University Wood Collection in
714 the Rejuvenation of Systematic Wood Anatomy. Taxon 23, 739–745.
715 https://doi.org/10.2307/1218435.
716

717 Wheeler, E.A., 2011. InsideWood—A Web Resource for Hardwood Anatomy. Iawa Journal 32,
718 199–211

719

Journal Pre-proof

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proof