

INVITED SPECIAL ARTICLE
For the Special Issue: Green Digitization: Online Botanical Collections Data Answering Real-World Questions

Digitization protocol for scoring reproductive phenology from herbarium specimens of seed plants

Jennifer M. Yost^{1,24}, Patrick W. Sweeney², Ed Gilbert³, Gil Nelson⁴, Robert Guralnick⁵, Amanda S. Gallinat⁶, Elizabeth R. Ellwood⁷, Natalie Rossington⁸, Charles G. Willis^{9,10}, Stanley D. Blum¹¹, Ramona L. Walls¹², Elspeth M. Haston¹³, Michael W. Denslow^{5,14}, Constantin M. Zohner¹⁵, Ashley B. Morris¹⁶, Brian J. Stucky⁵, J. Richard Carter¹⁷, David G. Baxter¹⁸, Kjell Bolmgren¹⁹, Ellen G. Denny²⁰, Ellen Dean²¹, Katalin D. Pearson²², Charles C. Davis⁹, Brent D. Mishler^{18,23}, Pamela S. Soltis⁵, and Susan J. Mazer⁸

Manuscript received 6 September 2017; revision accepted 2 January 2018.

¹ Department of Biological Sciences, California Polytechnic State University, 1 Grand Avenue, San Luis Obispo, California 93407, USA

² Division of Botany, Peabody Museum of Natural History, Yale University, P.O. Box 208118, New Haven, Connecticut 06520, USA

³ Arizona State University, School of Life Sciences, P.O. Box 874501, Tempe, Arizona 85287-4501, USA

⁴ iDigBio, College of Communication and Information, Florida State University, Tallahassee, Florida 32306, USA

⁵ Florida Museum of Natural History and Biodiversity Institute, University of Florida, Gainesville, Florida 32611, USA

⁶ Boston University, Department of Biology, 5 Cummington Mall, Boston, Massachusetts 02215, USA

⁷ La Brea Tar Pits and Museum, 5801 Wilshire Boulevard, Los Angeles, California 90036, USA

⁸ Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, California 93106-9620, USA

⁹ Department of Organismic and Evolutionary Biology, Harvard University Herbaria, 22 Divinity Avenue, Cambridge, Massachusetts 02138, USA

¹⁰ University of Minnesota, Department of Biology Teaching and Learning, 515 Delaware Street SE, Minneapolis, Minnesota 55455, USA

¹¹ Biodiversity Information Standards (TDWG), 1342 34th Avenue, San Francisco, California 94122, USA

¹² CyVerse, University of Arizona, 1657 East Helen Street, Tucson, Arizona 85721, USA

¹³ Royal Botanic Garden Edinburgh, 20a Inverleith Row, Edinburgh, EH3 5LR, United Kingdom

¹⁴ Department of Biology, Appalachian State University, Boone, North Carolina 28608, USA

¹⁵ Systematic Botany and Mycology, Department of Biology, Munich University (LMU), 80638, Munich, Germany

¹⁶ Department of Biology, Middle Tennessee State University, Murfreesboro, Tennessee 37138, USA

¹⁷ Biology Department, Valdosta State University, Valdosta, Georgia 31698, USA

¹⁸ University and Jepson Herbaria, University of California Berkeley, 1001 Valley Life Sciences Building, Berkeley, California 94720, USA

¹⁹ Swedish University of Agricultural Sciences, Unit for Field-based Forest Research, 360 30, Lammhult, Sweden

²⁰ USA National Phenology Network, University of Arizona, Tucson, Arizona 85721, USA

²¹ UC Davis Center for Plant Diversity, Plant Sciences M.S. 7, One Shields Avenue, Davis, California 95616, USA

²² Department of Biological Science, Florida State University, Tallahassee, Florida 32304, USA

²³ Department of Integrative Biology, University of California, Berkeley, California 94720-2465, USA

²⁴ Author for correspondence: jyost@calpoly.edu

Citation: Yost, J. M., P. W. Sweeney, E. Gilbert, G. Nelson, R. Guralnick, A. S. Gallinat, E. R. Ellwood, et al. 2018. Digitization protocol for scoring reproductive phenology from herbarium specimens of seed plants. *Applications in Plant Sciences* 6(2): e1022.

doi:10.1002/aps3.1022

Applications in Plant Sciences 2018 6(2): e1022; <http://www.wileyonlinelibrary.com/journal/AppsPlantSci> © 2018 Yost et al. *Applications in Plant Sciences* is published by Wiley Periodicals, Inc. on behalf of the Botanical Society of America. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

The digitization of herbarium specimens and their associated data have advanced our ability to understand complex and changing biological systems (Johnson, 2011; Pearse et al., 2017; Willis et al., 2017a). Digitizing herbarium records (capturing the taxon name, date of collection, location, and/or digital image) has advanced our ability to track changes in the distributions of organisms (Lavoie, 2013), but herbarium specimens are rich with additional information regarding plant health, reproductive condition, and morphology that is generally not captured in current digitization workflows (Nelson et al., 2015). Because the utility of specimens for research is accelerating, it is essential that we structure digital data collection in ways that best facilitate longevity and integration across data sources.

Of particular interest is the enormous potential of herbarium specimens as a resource for information on plant phenology (the timing of seasonal events such as flowering or fruiting). Plant phenology has complex, cascading effects on multiple levels of biological organization from individuals to ecosystems (Bertin, 1982). Temporal mismatches between plants and pollinators can quickly drive populations extinct, cause rapid evolutionary shifts, and result in billions of dollars of agricultural losses (Visser and Both, 2005; Both et al., 2006; Körner and Basler, 2010; Miller-Rumpff et al., 2010; Ozgul et al., 2010; Struttmann et al., 2015). Phenology has also been used to study the impact of climate change in a range of organisms and vegetation types (Bowers, 2007; Houle, 2007; Anderson et al., 2012; Lavoie, 2013). Consequently, maximizing the use of herbarium specimens for phenological research is not only important for improving our understanding of evolutionary change, it is also a matter of great practical concern for addressing environmental problems.

Recent studies have demonstrated the potential of herbarium specimens to be used in evaluating temporal and spatial variation in plant phenology (see Willis et al., 2017a for a review of these studies) despite known biases of herbarium records (Meyer et al., 2016; Pearse et al., 2017; Daru et al., 2018). These studies have provided three valuable outcomes. First, for several species, we now have a quantitative historical understanding of their phenological change over time (Rivera and Borchert, 2001; Primack et al., 2004; Miller-Rushing et al., 2006; Gallagher et al., 2009; Robbirt et al., 2011; Park 2012; Davis et al., 2015). Second, for some species, relationships between temporal or spatial variation in phenology and climate (e.g., local temperature and/or precipitation) have been detected; these relationships, in turn, provide a basis for forecasting the effects of ongoing climate change on the seasonal cycles of these taxa (Badeck et al., 2004; Franks et al., 2007; Matthews and Mazer, 2016; Prevéy et al., 2017). Third, we have an improved understanding of the specific advantages of herbarium specimens for phenological research, such as filling gaps in long-term or observational data sets, either for a period of time (Panchen et al., 2012) or for underrepresented regions (Gallagher et al., 2009; Panchen and Gorelick, 2017).

Given the ecological importance of phenology, the demonstrated value of herbarium specimens for phenological research, and the potential for digitization efforts to maximize herbarium records as a resource, it is necessary to develop robust standards for how phenological data are captured during or after the digitization process. There are currently two principal limitations to accessing and using phenological data from herbarium specimens: (1) the paucity of high-quality images accompanying digitized specimen records and (2) the lack of standardized methodology for capturing specimens' reproductive traits and sharing the resulting data. If any phenological information is present on a label or visible on a specimen, it

is parsed in numerous—and often arbitrary—ways during digitization. For example, phenological data embedded in a label might be “on a south facing slope in full flower,” but this information might be digitally captured in the ‘habitat,’ ‘notes,’ ‘plant description,’ or other field of a local database.

Even if a local database does contain a field explicitly for phenological characters, each institution independently decides how to record the states present on the sheet. For example, in the SEINet collaborative (<http://swbiodiversity.org/seinet/>), which consists of 251 U.S. collections and 11.8 million records, there are 2.6 million (22%) records with text present in a database field called ‘reproductiveCondition.’ The majority of terms found within this field specify flowering, fruiting, sterile, spores, and/or cones; however, these terms are expressed in over 4000 unique text strings (Table 1). Some collections specify “flowering,” while other records state “flws.” Some are ambiguous (12,000 are coded as merely “u,” presumably meaning “unknown” or “unrecorded”). The lack of a controlled vocabulary for this field makes aggregating these data for research purposes onerous. Local databases often share their data with data aggregators such as iDigBio (<https://www.idigbio.org/>) or directly with users as a suite of Darwin Core Archive files, an exchange standard described more fully below (Wieczorek et al., 2012; <http://rs.tdwg.org/dwc/>). However, the relevant Darwin Core fields are equally diverse, with most phenological traits being placed into the fields ‘occurrenceRemarks,’ ‘organismRemarks,’ ‘dynamicProperties,’ ‘reproductiveCondition,’ or ‘fieldNotes.’

It is clear that there is a huge potential for using phenological data from herbarium specimens (Willis et al., 2017a). We propose a method here to (1) broaden the scope and longevity of digitization efforts through a standardized methodology for scoring reproductive characters from herbarium specimens and (2) provide a means of sharing the resulting data in a Darwin Core format. The protocol we describe here will unlock the potential of herbaria for phenological research by facilitating comparability among herbaria, research groups, and other methodologies used to collect phenological data (e.g., citizen science observations, satellite imagery, and stationary camera images).

METHODS AND RESULTS

To make progress toward developing standards that reflect community-wide goals and feasible implementation, iDigBio sponsored a working group called Coding Phenological Data from Herbarium Sheets in March 2016 at the University of California, Berkeley (https://www.idigbio.org/wiki/index.php/Coding_Phenological_Data_from_Herbarium_Sheets).

This workshop brought together 37 participants from the United States, Scotland, England, Sweden, Canada, Germany, and Australia. The group represented a range of phenological interests, including phenological researchers (those who use the downstream data products obtained from specimens), herbarium collection personnel (those responsible for preparing and curating the specimens and their data), in situ phenological observers who record the phenological status of living plants, and data standards experts. A participant list can be found on the iDigBio wiki page (https://www.idigbio.org/wiki/index.php/Coding_Phenological_Data_from_Herbarium_Sheets) and in Appendix S1.

Prior to the workshop, we developed a survey to assess needs of the phenological community and herbarium data users and to

TABLE 1. Most frequently found text strings in the Darwin Core field 'reproductiveCondition' and the number of specimens in SEINet with that exact string. Nearly all of these can be quickly scored according to the protocol proposed here using the Attribute Mining Tool in Symbiota-based databases.

'reproductiveCondition' text samples	Specimen count
Flowering*	434,637
Flowering and fruiting	285,865
flower*	174,751
Fruiting	160,853
fl*	136,544
fr	97,098
flowering = early*	90,578
fl-fr	90,132
flowers*	86,372
fruit	85,260
flowering + fruiting = mid	75,128
vegetative	47,785
flr*	44,529
fertile	39,309
Flo*	36,867
veg	36,258
fl,fr	35,032
Flr & Frt	33,199
spores	31,824
sterile	30,478
Flower: Y Fruit: N* Vegetation: Y Bud: N*	27,508
Flower Fruit	27,254
flowers & fruit	26,301
fru	19,833

*Records that refer only to open flowers or flowering; these can be scored simultaneously with the Attribute Mining Tool, resulting in over 1 million new phenological records from a single scoring effort.

review the current ways phenological data were being captured. We received 76 responses to the survey, and the respondents identified themselves as being from collections, monitoring, or research areas (not mutually exclusive). With this survey and input from participants at the workshop, we reviewed the ways in which herbaria currently capture phenological traits. The two most-scored traits from specimens are the presence of open flowers and the presence of fruit (14 million specimens represented in this survey). Most respondents also felt that of all possible traits, open flowers and fruits were the most important traits to score on a specimen. Participants of the workshop echoed this sentiment. We reviewed previous phenological research that was based on data derived from herbarium specimens in order to identify the types of raw data necessary and sufficient to achieve a variety of research goals. These findings are summarized in Willis et al. (2017a).

When developing a scoring protocol, we considered the challenges and limitations to scoring specimens (*de novo* scoring and as part of regular digitization workflows) and the potential solutions to these limitations. We considered hard-to-see floral parts, trained vs. untrained scorers, the limited resources of most herbaria, and the likelihood of community-wide adoption. We also considered the costs and benefits of recording qualitative data (e.g., "open flowers present/absent") vs. quantitative data (e.g., counts or proportions of unopened flowers, open flowers, fruits).

One of our primary concerns is that any resulting data from attempts to score phenological traits should be shareable in Darwin Core-formatted files to help ensure the usefulness and longevity of these data. Representatives from the data standards community,

Biodiversity Information Standards (TDWG), including Darwin Core and Apple Core, provided input for representing phenological stages using current biodiversity standards.

Finally, to ensure that phenological traits from specimens can be integrated with other sources, participants included members of the USA National Phenology Network, the California Phenology Network, the National Ecological Observatory Network, the Royal Botanic Garden Edinburgh, the Pan-European Phenology Network, and the Plant Phenology Ontology.

We propose that reproductive traits for specimens of seed plants be scored according to the following hierarchical categories/questions (Tables 2 and 3). Our protocol uses terminology from the Plant Ontology to represent plant parts (e.g., flower, fruit) (<http://www.obofoundry.org/ontology/po.html>) and traits that correspond to plant phenological traits in the Plant Phenology Ontology (PPO; <http://htmlpreview.github.io/?https://github.com/PlantPhenoOntology/ppo/blob/master/documentation/ppo.html>) (e.g., reproductive structures present, unopened flowers present, open flowers present) (Walls et al., 2014; Stucky et al., 2016). By using a vocabulary that directly maps to ontologies, data collected with this method can be easily ingested into data stores using those ontologies and thereby integrated with other sources of phenological data such as direct observations *in situ* or remote sensing (www.plantphenology.org).

First-order scoring

The question "Are 'reproductive structures' present? (yes/no/not scorable)," while the broadest question, was still determined to have value for scoring specimen records. Having this information allows researchers to filter millions of records quickly to find those that contribute to phenological research. It is also relatively easy for users with different levels of botanical training (e.g., curators, volunteers, and citizen scientists) to score. A "yes" means that some reproductive structures of some kind (e.g., flowers, fruits, or cones) are present. A "no" means that the specimen is sterile and strictly vegetative. It is important to note that this first-order scoring can apply to all taxonomic groups, even beyond seed plants. Some taxonomic groups may exhibit specialized structures that make it more difficult for non-experts to complete this process (i.e., vegetative propagules that look like fruits), but we anticipate that this challenge will be limited. Minimally, first-order scoring will allow for records to be filtered and then subsequently scored in more detail.

Second-order scoring

For specimens that are scored as having reproductive structures present, it is valuable to characterize which reproductive structures are present. Most research thus far has used specimens with open flowers. For flowering plants, we propose the following second-order, non-mutually exclusive questions: "Are 'unopened flowers' present?," "Are 'open flowers' present?," "Are 'fruits' present?." For gymnosperms the questions are: "Are 'pollen cones' present?" and "Are 'seed cones' present?" (Tables 2 and 3). The term "bud/s" can confuse floral buds with leaf buds; therefore, the PPO and this protocol refer to unopened flowers only.

The second-order questions are not mutually exclusive. If unopened flowers, open flowers, and fruits are all present on a sheet, all questions can be answered in the affirmative. Having these data allows researchers to quickly identify the records that pertain to

TABLE 2. Proposed phenological scoring protocol for angiosperms. Single quote terms are defined in the Plant Ontology for plant parts (e.g., flower, fruit) and for traits that correspond to plant phenological traits in the Plant Phenology Ontology (e.g., reproductive structures present, unopened flowers present, open flowers present). Second-order questions are not mutually exclusive and are answered in the affirmative or left blank. Third-order questions define a specimen's position in the phenological cycle. Third-order categories are mutually exclusive with one another.

First-order	Second-order	Third-order
Are 'reproductive structures' present? (yes/no/not scorable)	Are 'unopen flowers' present?	Mostly 'unopen flowers?' (or counts)
	Are 'open flowers' present?	Mostly 'open flowers?' (or counts)
	Are 'fruits' present?	Mostly 'post-mature flowers?' (or counts) Mostly 'immature fruits?' (or counts) Mostly 'mature fruits?' (or counts) Mostly 'post-mature fruits?' (or counts)

TABLE 3. Proposed phenological scoring protocol for gymnosperms. Words in single quotes are defined in the Plant Ontology for plant parts (e.g., pollen cone) and for traits that correspond to plant phenological traits in the Plant Phenology Ontology (e.g., reproductive structures present, mature seed cones). Second-order questions are not mutually exclusive and are answered in the affirmative or left blank. Third-order questions define a specimen's position in the phenological cycle. Third-order categories in this example are not mutually exclusive with one another because pollen and seed cones develop independently.

First-order	Second-order	Third-order
Are 'reproductive structures' present? (yes/no/not scorable)	Are 'pollen cones' present?	Mostly 'immature pollen cones?' (or counts) Mostly 'mature pollen cones?' (or counts) Mostly 'post-mature pollen cones?' (or counts)
	Are 'seed cones' present?	Mostly 'immature seed cones?' (or counts) Mostly 'mature seed cones?' (or counts) Mostly 'post-mature seed cones' present?

their individual research questions. The second-order questions require greater training for personnel to accurately discriminate unopened flowers, open flowers, and fruits. For many taxa (e.g., grasses, sedges, rushes), floral structures are small and distinguishing between unopened flowers, open flowers, and fruits can be challenging. Additionally, it is important that scorers are trained to distinguish between leaf buds and flower buds (which contain unopened flowers). As training materials are developed for various plant groups they should be shared widely across the community.

Third-order scoring

Third-order scoring further subdivides the categories of the second-order scorings. While second-order scorings will determine which specimens should be included in phenological research, it is often valuable to know the specimen's specific phenophase. Analyses can

be more precise if we can distinguish between specimens in full flower from those specimens at the beginning or end of the flowering cycle. The third-order scorings are intended to place individual specimens at a specific point in phenological development. As such, these subcategories and the units used to report them may vary depending on the institutional or research priorities that generate them. We do not specify exactly what the third-order categories should be, as these will be determined by research priorities and staff time, but rather we explain how these questions are most commonly expressed or could be expressed within our proposed framework. Although we do not specify third-order categories, we do strongly recommend that researchers clearly define their categories and make their definitions broadly accessible, along with pertinent metadata. For example, the Simple Knowledge Organization System (SKOS) provides a framework for representing controlled vocabularies that easily lends itself to being shared online. The New England Vascular Plant (NEVP) project has devised a vocabulary following these guidelines and has published their vocabulary using SKOS (http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#00) (Table 4). Furthermore, regardless of the nature of third-order categories, we strongly recommend that researchers share these data via the Darwin Core extensions explained below. For quantitative phenological data, some research groups require count data from a specimen, such as the numbers of unopened flowers, open flowers, and fruits on each sheet. Counts would be considered third-order scoring. For some analyses, raw count data may be transformed to express the proportions of reproductive organs represented by unopened flowers, open flowers, and fruits, thereby distinguishing specimens that represent early-, peak-, or late-flowering individuals. Even if count data are not precisely recorded from a sheet, the degree of flowering can be binned into categories representing early, peak, and late flowering (see NEVP project third-order categories in Table 4). Note that in the example shown in Table 4, the third-order scorings, if categorical, are mutually exclusive. A sheet cannot simultaneously be 'mostly unopened flowers' and 'mostly open flowers.' Finally, in order to integrate third-order scorings with other sources of phenological data, we recommend use of the PPO. Counts, binned data, or other data representations can be mapped to the PPO and by so doing, assure interoperability with field monitoring data.

There are some taxonomic groups that have very small floral structures that either require an onerous amount of time to score or require expertise to determine what kinds of organs are actually present on the sheet. Members of Poaceae, Cyperaceae, and Juncaceae, as well as certain Asteraceae, are a few examples for which second-order scoring may be more challenging. However, it should be relatively easy to apply first-order scorings to these groups, thereby greatly increasing the utility of these specimens for phenological research. Our protocol does not address the presence/absence or abundance of male vs. female flowers, or distinguish between perfect and imperfect flowers in gynodioecious, gynomonoecious, or andromonoecious species, largely due to the fact that these categories have seldom been included in phenological research.

The timing of reproduction is not the only important phenological event of interest to be tracked in plants. Leaf bud break and leaf-out are important phenomena for deciduous forests, as is autumn senescence. These vegetative characters are often tracked via satellite imagery and in situ monitoring efforts. Scoring phenological leaf traits on herbarium specimens is rare (but see Zohner and Renner, 2014; Gallinat et al., 2015), but it provides

valuable insights into the effects of climate change (Chmielewski and Rotzer, 2001; Everill et al., 2014). A similar scoring protocol is recommended for foliar structures, although we do not specify a protocol here.

Sharing of phenological scorings

Darwin Core has emerged as a key standard for describing species occurrences. Online documentation, including definitions and examples, is provided for each term used in the Darwin Core (TDWG website: <http://rs.tdwg.org/dwc/terms/>). Any attempts to share phenological or other trait data from specimens should utilize Darwin Core fields to assure that the basic specimen occurrence information is standardized. However, phenological trait descriptions are

not part of the Darwin Core and therefore other mechanisms are needed to support narrower or broader data-sharing approaches.

We are proposing to share phenological data using the Darwin Core Extended MeasurementOrFact (eMoF) extension (Table 5) (<https://tools.gbif.org/dwca-validator/extension.do?id=http://rs.iobis.org/obis/terms/ExtendedMeasurementOrFact>). This extension provides a mechanism whereby many measurements or facts can be shared for each specimen record in a Darwin Core Archive. This extension allows for sharing of metadata associated with each phenological (or any other trait) scoring. For example, when evaluating data quality, it can be useful to know when, how, and by whom scorings were recorded. Accuracy may be affected by whether the specimen was scored from a web-based image or the physical specimen. An eMoF record can contain a definition of the type of

TABLE 4. An example of the proposed scoring protocol applied by the New England Vascular Plant (NEVP) thematic collections network. Definitions of reproductive phenological terms for flowering plants are displayed (draft version 1.2).

Scoring	NEVP definition	URI identifier
First-order		
Is the material on the sheet sterile?	No reproductive structures present (no unopened flowers, open flowers, or fruits)	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#02
Is there reproductive material present on the sheet?	At least one reproductive structure of any kind is present (unopened flowers, open flowers, or fruits)	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#03
Not scorable	Not possible to score reproductive condition using material present	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#16
Second-order		
Unopened flowers present?	At least one unopened flower is present	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#04
Open flowers present?	At least one open flower is present	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#06
Fruits present?	At least one immature or mature fruit is present	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#11
Third-order		
Mostly unopened flowers	Mostly unopened flowers (less than half open) but with at least one open flower; this category is mutually exclusive with mostly open and mostly old flowering stages and with mostly young, mostly mature, and past maturity fruiting stages	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#08
Mostly open flowers	Mostly open flowers (more than half open) with few unopened flowers or old flowers that have lost their petals; this category is mutually exclusive with mostly unopened and mostly old flowering stages and with mostly young, mostly mature, and past maturity fruiting stages	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#09
Mostly old flowers	Mostly old flowers (less than half open) that have lost their petals, but at least one flower still open; this category is mutually exclusive with mostly unopened and mostly open flowering stages and with mostly young, mostly mature, and past maturity fruiting stages	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#10
Mostly young fruit	Mostly immature fruits present (less than half mature) but at least one mature fruit present, mutually exclusive with mostly mature and past maturity fruiting stages and with mostly unopened, mostly open, and mostly old flowering stages	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#13
Mostly mature fruit	Mostly mature fruits present (more than half mature), mutually exclusive with mostly young and past maturity fruiting stages and with mostly unopened, mostly open, and mostly old flowering stages	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#14
Mostly past mature fruit	Fruits have fallen from stalks, withered, or dehisced and lacking seeds (less than half mature) but at least one mature fruit present, mutually exclusive with mostly young and mostly mature fruiting stages and with mostly unopened, mostly open, and mostly old flowering stages	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#15
Number of open flowers present	Number of open flowers present on specimen: 0 is an acceptable value	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#07
Number of fruits present	Number of mature fruits, 0 is an acceptable value	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#12

measurement, the value and units of the measurement, the method of measurement, and by whom and when it was measured.

Table 5 shows an example Darwin Core Archive file of the eMoF extension of two herbarium sheets that were scored using our protocol (measurementType called Phenology [ver.1.0]). The first record was scored with a measurementValue = 'Reproductive.' Additionally that same record is scored as measurementValue = 'Open flowers.' The second record, united by catalog number, is scored 'Reproductive,' 'Open flowers,' and 'Fruiting.'

Using the eMoF extension has a potential disadvantage, namely that it does not allow the measurement to be rigorously tied to a particular aspect of the core record. This means that any user can define a new and non-standard 'measurementType' and 'measurementValue' (e.g., potentially called "Flowering Time" and "having flowers" or "Flws"), which could lead to difficulty compiling data. Unless various measurementTypes and measurementValues are rigorously defined, an excessive number of unique text strings could be generated. To address this, we are working toward defining these terms within Apple Core. Apple Core is a set of best practice guidelines for publishing botanical specimen information for herbaria. A goal of the guidelines is to mitigate the generality of Darwin Core by providing detailed guidelines for publishing botanical specimen information in Darwin Core. These guidelines will include recommended terms, specific definitions, multiple examples, common issues, and controlled vocabularies where appropriate that are specific to herbarium specimens. Apple Core is a community-curated

resource that is still being refined, and interaction with phenological researchers will help to strengthen this resource. Finally, use of this approach is complementary with broader sharing initiatives that utilize ontologies, such as the Plant Phenology Ontology.

In the near future, using the eMoF extension will allow for phenological scorings to be published in iDigBio, the Global Biodiversity Information Facility (GBIF), and other public repositories. Darwin Core Archive publishing services are available within all Symbiota portals (Gries et al., 2014; <http://symbiota.org>) and form the basis from which iDigBio harvests specimen data from these portals (<http://swbiodiversity.org/seinet/collections/datasets/datapublisher.php> and <http://sernecportal.org/portal/collections/datasets/datapublisher.php>). Adherence to our protocol at local institutions will facilitate the search functions provided and developed by large aggregators such as iDigBio and GBIF.

CONCLUSIONS

Advantages of the proposed protocol

The questions presented here provide important data for researchers while also requiring minimal effort from herbarium curators. Phenological questions are easily integrated into standard label digitization workflows or could be subsequently scored from images. Due to the nested nature of the questions, a third-order question can be

TABLE 5. Phenological scorings^a in the extended MeasurementOrFact Darwin Core extension: an example from Arizona State University Fabaceae specimen records.

coreid	measurement-Type	measurementTypeID	measurementValue	measurement-ValueID	measurementUnit	measurementDeterminedDate	measurement-DeterminedBy	measurement-Remarks
652438	Phenology (ver 1.2)	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#01	Reproductive	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#03		2017-04-15T01:05:07Z	egbot	
652438	Phenology: reproductive	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#03	Open Flowers	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#06		2017-04-15T01:05:08Z	egbot	
652439	Phenology (ver 1.2)	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#01	Reproductive	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#03		2017-04-15T01:10:54Z	egbot	
652439	Phenology: reproductive	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#03	Open Flowers	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#06		2017-04-15T01:10:55Z	egbot	
652439	Phenology: reproductive	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#03	Fruiting	http://purl.org/nevp/vocabulary/reproductive-phenology_1.2#11		2017-04-15T01:10:56Z	egbot	

Note: measurementType = the nature of the measurement, fact, characteristic, or assertion; measurementTypeID = an identifier for the measurementType; measurementValue = the value of the measurement, fact, characteristic, or assertion; measurementValueID = an identifier for facts stored in the column measurementValue; measurementUnit = the units associated with the measurementValue, not used in this example; measurementRemarks = not used in this example, but could contain comments or notes accompanying the MeasurementOrFact; measurementDeterminedDate and measurementDeterminedBy = the person, date, and time that the scoring was applied.

^aIn the near future, these phenological mappings will be aggregated by iDigBio, GBIF, and other public repositories through Symbiota's Darwin Core Archive publishing services. The Darwin Core Archive publishing services are available within all Symbiota portals and are the central mechanized archive used by iDigBio to harvest and maintain updates of specimen data published from a Symbiota portal instance.

scored initially, with the appropriate second- and first-order questions automatically populated. For example, a report of “fruits present” on a specimen would automatically score a “yes” for the first-order question, indicating that reproductive structures are present. To answer first-order questions, the person who is performing the initial data entry for a specimen need only look at the sheet and check a box indicating whether reproductive structures of any kind are present. For databases that do not have the infrastructure to accommodate this type of scoring, a few alternatives are presented below.

Workflow implementation

Phenological scores can be recorded at a number of steps in a digitization workflow. In the case of an object-to-data workflow, scores could be made directly from the sheet as label data are being captured. With an image-based workflow, the scoring of specimens can be achieved by visual inspection of their images. The latter approach provides the option of making the image available online where the public (e.g., citizen scientists using Notes from Nature, CrowdCurio, or other platforms) can record phenological observations. Machine learning approaches are likely to facilitate our ability to score images at scale in the near future. Database fields in local databases need to be modified to accommodate the proposed structure. Implementation of controlled vocabularies can be facilitated with drop-down menus or pick-lists (see Figs. 1 and 2 for an example from a Symbiota portal); however, providing such functionality might require changes to database management software. Fortunately, a number of tools (described below) have been developed for scoring the phenological status of specimens.

For curators who do not have a database with Symbiota-type functionality that provides phenological checkboxes corresponding to our proposed protocol, we suggest that users enter phenological information into an appropriate text field within their existing database with the expectation that new tools will enable users to search these text fields and score the specimens appropriately (see Tools that facilitate scoring: Symbiota’s “Attribute Mining Tool” below). Ideally, every institution’s home database will include a text field dedicated exclusively to information pertaining to phenology. However, including phenological information as text anywhere within a given specimen’s label data is better than not capturing any phenological traits. To choose the best text field within a local database, it is important to know how the specimen data appear when shared using a Darwin Core Archive. If, for example, one’s local database conforms to Darwin Core, reproductive traits should be included in the ‘reproductiveCondition’ field. The words

entered into the text field should be unambiguous and should correspond to the protocol above (e.g., “unopened flowers,” “open flowers,” “fruit”). This is an action that all curators can immediately integrate into their current digitization workflows.

Conclusions for curators

Those managing or implementing digitization workflows should consider incorporating the scoring of phenological data into their workflows. At the very least, first- or second-order phenological data (as described above) should be considered for capture. Doing so will facilitate future scoring of the specimens. If time does not permit training herbarium personnel to record challenging second-order

Arizona State University Vascular Plant Herbarium (ASU-Plants)

Harvesting Filter

Occurrence trait: **Phenology (ver 1.0)** ▾
 Verbatim text source: **Reproductive Condition** ▾
 Filter by text (optional): **fl**
 Filter by taxon (optional): **Fabaceae** Get Field Values

Reproductive Condition

Select Source Field Value(s) - hold down control or shift buttons to select more than one value

flower, fruit - [23]
 flower: n fruit: y vegetation: y bud: n - [1]
 flowering & fruiting - [37]
 flowering + fruiting=mid - [13]
 flowering + fruiting=mid - [2]
 flowering - [83]
 flowering and fruiting - [19]
 flowering or fruiting - [1]
 flowering=early - [8]
 flowers & fruit - [5]
 flowers & leaves - [6]
 flowers - [56]
 flowers, fruit & leaves - [2]
 flr & frt - [2]
 flr - f91

Reproductive
 Flowering
 Fruiting
 Budding
 Sterile
 Not scorable

Notes:

Status: ▾

Batch Assign State(s)

FIGURE 1. Example of Symbiota’s Attribute Mining Tool. Here, a local database’s text field ‘Reproductive Condition’ was searched for all text strings containing “fl” in the Fabaceae. Highlighted references are text strings referring to both flowers and fruits. These were selected, and second-order scorings of “open flowers present” and “fruit present” were then applied to all specimens simultaneously.

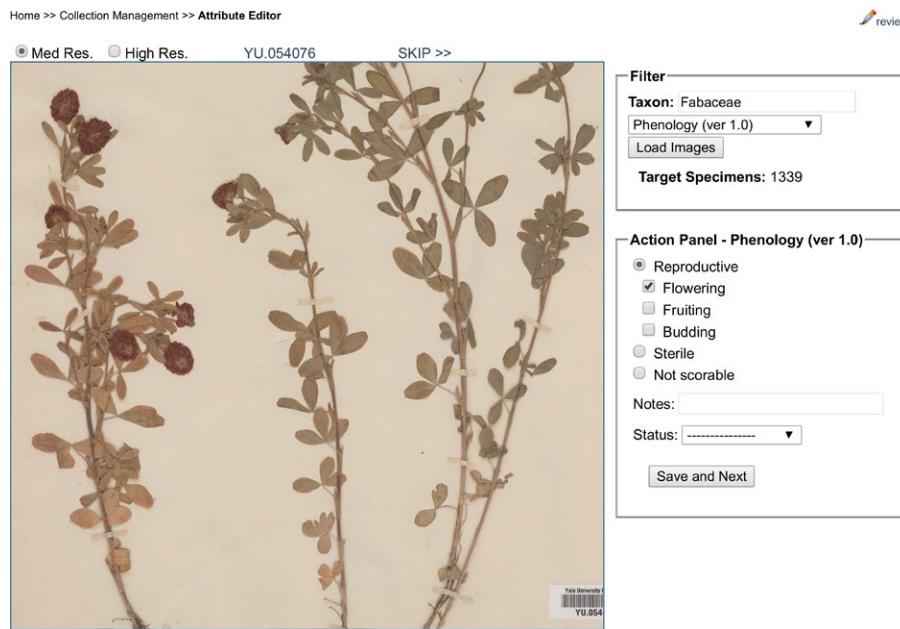


FIGURE 2. Example of Symbiota's Image Scoring Tool. Images of Fabaceae specimens were searched. The user can apply the desired level of scoring to each image that appears.

scorings, then simply adding the word “reproductive” somewhere in a relevant database field will aid future work and research use.

Tools that facilitate scoring

Although it is not the primary focus of this paper, we think it is useful to readers to make brief mention of how our protocol could be implemented and what tools are available for doing so.

Symbiota's “Attribute Mining Tool”—Part of the NEVP project was the development of a tool to score phenological traits using digitized label text (Fig. 1). This tool allows a user to search for specific words in database fields and map these to the proposed vocabulary. For example, using this tool to search the field ‘reproductiveCondition’ within SEINet resulted in over 4000 unique text strings (Table 1). The Attribute Mining Tool allows one to select all records containing text that refer solely to a single scoring category. For example, if a user were scoring “open flowers present” only, the user could select all the highlighted rows in Table 1 and click “Open flowers present.” In the example from SEINet presented in Table 1, this single scoring event would result in the selection and scoring of 1,031,786 records. In a separate scoring event, the user could select all records that make reference to both open flowers and fruits and then select “open flowers present” and “fruits present.” Because a curator is responsible for mapping free text strings from the database to a controlled vocabulary, this method does not rely on computerized inference. The ability to apply phenological scoring to any specimen within a Symbiota portal is highly efficient, and these types of tools should be developed within other database platforms.

Scoring images

Many platforms have been developed for remotely scoring images of specimens, and we review them below. It is vital that future

scoring platforms conform as closely as possible to the proposed protocol to facilitate data integration. Furthermore, it is vital that specimen trait data, even when scored outside of the local database, remain associated with the original specimen record. This will allow trait data and occurrence data to travel together through the data aggregation process, preventing duplicated scoring efforts.

Symbiota's “Image Scoring Tool”—The new Image Scoring Tool, developed as part of the NEVP project, allows Symbiota network users to filter images and apply a phenological score to them (Fig. 2). This approach has facilitated the scoring of over 240,000 images of New England specimens to date. Phenological scorings are being shared with end users through the Consortium of Northeastern Herbaria portal via the Darwin Core Extended MeasurementOrFact extension and Darwin Core Archives, as outlined above. This functionality will soon be available to all Symbiota-based databases.

Notes from Nature—Notes from Nature is an online citizen science platform (Hill et al., 2012) originally developed to support the transcription of specimen labels, but it has expanded to include phenological classifications. Notes from Nature extends the Zooniverse (<https://www.zooniverse.org/>) model by providing a simple way for curators or researchers to bundle and upload images, set targets for transcriptions or scoring, launch new expeditions, and engage volunteers (Fig. 3). Notes from Nature addresses data quality by requiring a minimum of three replicated classifications (provided by different participants) for each imaged specimen. When expeditions are completed, a suite of tools are available for reporting outcomes of efforts, and automated reconciliation occurs to produce a “best classification.” This includes data for phenological categories and for counts of reproductive structures.

Notes from Nature phenology expeditions have so far solicited reports of flowering and fruiting as well as counts of reproductive structures for *Quercus* L., *Coreopsis* L., and *Cakile* Mill. Notes from Nature has launched expeditions asking for simple annotations of open flowers or fruits present, to more complex expeditions where users are asked to count numbers of unopened flowers, open flowers, and fruits. Asking users to report first- and second-order scorings generated large volumes of accurate phenological data, whereas expeditions asking for more complex scorings, such as counts, had lower participation from the community of citizen science annotators and took much longer.

CrowdCurio—CrowdCurio is a new online platform designed to give researchers the ability to design and implement crowdsourcing projects tailored to their specific interests and data sources (Willis et al., 2017b; <https://www.crowdcurio.com/>). Most recently, a CrowdCurio project, titled “Thoreau's Field Notes,” demonstrated that the platform was an effective tool for crowdsourcing the collection of phenological data from digitized herbarium specimens (Willis et al., 2017b). Participants are presented with an image of

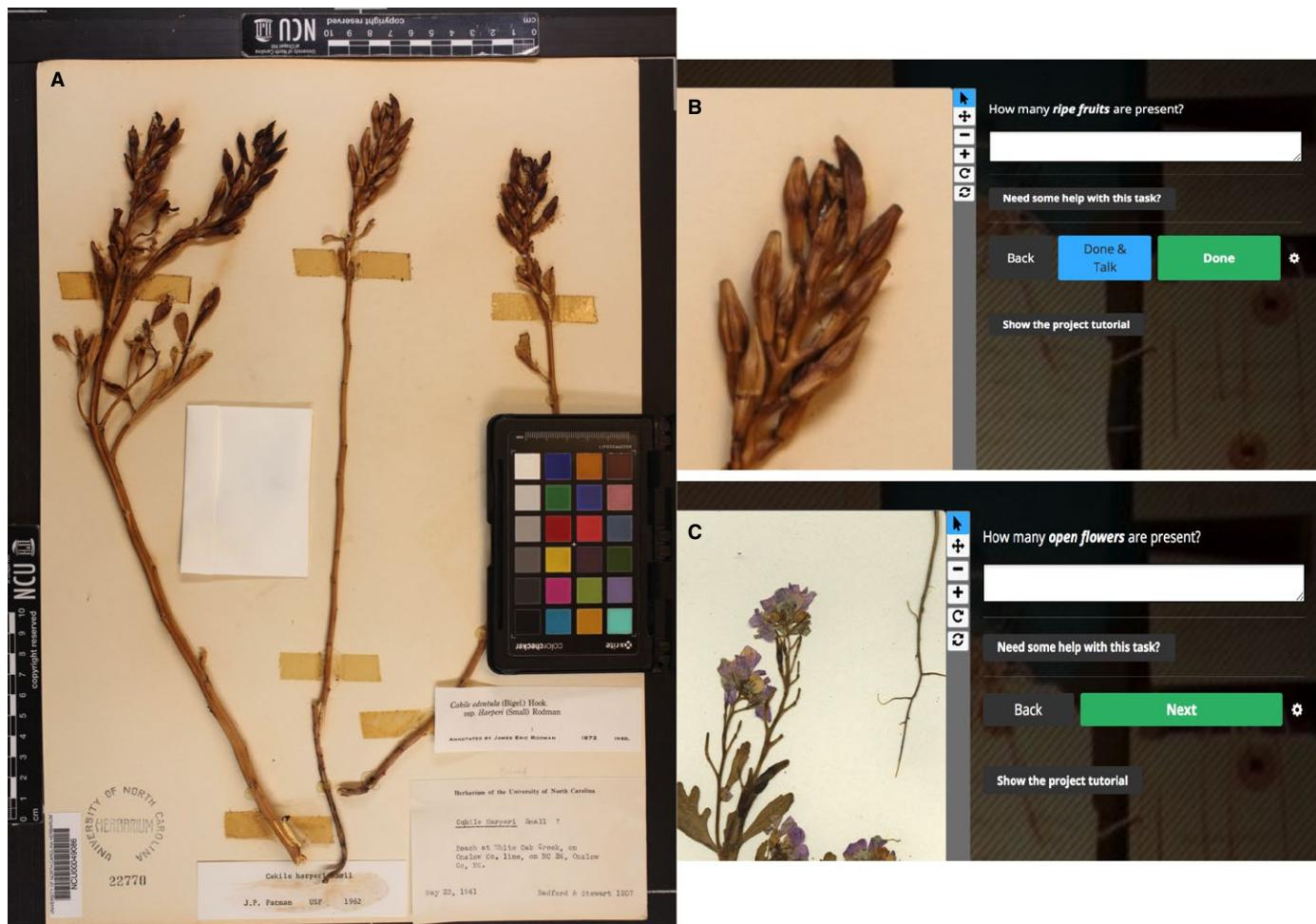


FIGURE 3. (A) A typical specimen image presented as part of a phenology expedition on Notes from Nature. (B) Classification task requesting that volunteers record the number of fruits that are visible on the herbarium specimen. (C) Classification task requesting that volunteers record the number of open flowers visible on the herbarium specimen. Note that parts B and C display tools that help volunteers to complete the task (e.g., pan, zoom, rotate, tutorial, and help).

a herbarium specimen and asked to annotate the image by clicking on each visible unopened flower, open flower, and fruit. These annotations are then transformed into counts that can be used to approximate the phenological stage of a given individual specimen.

In a preliminary study of the efficiency and quality of CrowdCurio data collection, Willis et al. (2017b) compared data collected by expert (herbarium curators) and non-expert (anonymous Amazon Mechanical Turk [Mturk] workers) participants for two common New England species: greater celandine (*Chelidonium majus* L.) and lowbush blueberry (*Vaccinium angustifolium* Aiton). They found that non-expert counts were similar to expert counts, but that non-experts were able to record nearly twice as much data at less cost over the same amount of time.

Data collected via crowdsourcing, however, are not without limitations. Although Willis et al. (2017b) found no difference in average counts between experts and non-experts, non-expert counts tended to be more variable per specimen. This in part depended on the specimen being assessed—specimens with more objects to count had higher error rates. As with any crowdsourcing project, care should be taken when choosing which specimens and taxa

to include (e.g., are the flowers easy to identify?). Additionally, CrowdCurio is in the process of implementing additional features to improve data quality, such as filtering users based on their ability to repeat the same task. The phenological data generated within CrowdCurio can be expressed according to the protocol outlined in this paper and shared via Darwin Core Archives.

Integration with other data sources

One ultimate goal is to combine herbarium specimen data with other sources of phenological data to make possible the detection of phenological changes across geographic, temporal, and taxonomic scales. The PPO provides an opportunity for herbarium data to be combined with disparate data sources, such as in situ phenological monitoring or satellite imagery. The PPO is a common vocabulary for describing plant phenological traits and was designed to provide a means to support global-scale integration of phenological data. Ontologies provide highly structured, controlled vocabularies for data annotation and are particularly useful for standardization, because they not only establish a common terminology but also

formalize logical relationships between terms such that they can be analyzed using machine reasoning. For example, logical term relationships in the PPO specify that any plant with “expanding leaves” must necessarily also have “non-senescing leaves.” This logical structure means that data can be integrated at different levels of detail and software can be used to establish new facts about the data that were not expressed in the original data sets. This structure in turn enables large-scale integration among a wide range of study types, including: (i) studies addressing similar phenophases but using different methodologies, (ii) studies involving different phenophases, and (iii) studies not specifically addressing phenology but producing other types of data (e.g., trait or climatic data). Thus, the PPO empowers researchers to aggregate larger data sets, at the global scale, and to address broader questions involving the interplay of phenology and other factors. Accordingly, the PPO is already being used to integrate data resources such as those from the USA National Phenology Network (Denny et al., 2014), the Pan-European Phenology Network (<http://www.pep725.eu/>), and herbarium digitization efforts. The PPO and associated integration tools are compatible with the Darwin Core and Apple Core standards and associated data-sharing tools discussed above.

Never before has an understanding of phenology been so important to humans. We are in a time of massive environmental change, and the organisms upon which we depend will have to adapt or migrate if they are to avoid local or global extinction. Herbarium specimens are critical to understanding and mitigating those changes. We need phenological data from specimens now more than ever, and researchers are ready and eager to analyze high-quality data sets, particularly those comprising high taxonomic diversity, temporal depth, and a broad geographic range. With minimal additional efforts during or post-digitization, specimens can be scored quickly and easily and contribute to our understanding of our changing planet and the flora that sustains it.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (grants DBI-1547229 [P.S.S.], DBI-0735191 and DBI-1265383 [R.L.W.], DBI-1458550 [R.P.G.], DBI-1410087 [A.B.M.], DBI-EF1208835 [C.C.D.], DEB-1556768 [S.J.M.], DBI-1458264 [J.R.C.], and DBI-1209149 [P.W.S.]). Additional support was provided by the Andrew W. Mellon Foundation, the Sibbald Trust, and the Scottish Government. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

AUTHOR CONTRIBUTIONS

All authors contributed equally to the development of the protocol. J.M.Y managed the writing of the paper with contributions from all authors. C.G.W, P.W.S, E.G., and M.W.D provided figures and text from related projects.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

LITERATURE CITED

Anderson, J. T., D. W. Inouye, A. M. McKinney, R. I. Colautti, and T. Mitchell-Olds. 2012. Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. *Proceedings of the Royal Society of London B: Biological Sciences* 279: 3843–3852.

Badeck, F.-W., A. Bondeau, K. Böttcher, D. Doktor, W. Lucht, J. Schaber, and S. Sitch. 2004. Responses of spring phenology to climate change. *New Phytologist* 162: 295–309.

Bertin, R. I. 1982. The ruby-throated hummingbird and its major food plants: Ranges, flowering phenology, and migration. *Canadian Journal of Zoology* 60: 210–219.

Both, C., S. Bouwhuis, C. M. Lessells, and M. E. Visser. 2006. Climate change and population declines in a long-distance migratory bird. *Nature* 441: 81–83.

Bowers, J. E. 2007. Has climatic warming altered spring flowering date of Sonoran desert shrubs? *Southwestern Naturalist* 52: 347–355.

Chmielewski, F.-M., and T. Rötzer. 2001. Response of tree phenology to climate change across Europe. *Agricultural and Forest Meteorology* 108: 101–112.

Daru, B. H., D. S. Park, R. B. Primack, C. G. Willis, D. S. Barrington, T. J. S. Whitfeld, T. G. Seidler, et al. 2018. Widespread sampling biases in herbaria revealed from large-scale digitization. *New Phytologist* 217: 939–955.

Davis, C. C., C. G. Willis, B. Connolly, C. Kelly, and A. M. Ellison. 2015. Herbarium records are reliable sources of phenological change driven by climate and provide novel insights into species phenological cueing mechanisms. *American Journal of Botany* 102: 1599–1609.

Denny, E. G., K. L. Gerst, A. J. Miller-Rushing, G. L. Tierney, T. M. Crimmins, C. A. Enquist, P. Guertin, et al. 2014. Standardized phenology monitoring methods to track plant and animal activity for science and resource management applications. *International Journal of Biometeorology* 58: 591–601.

Everill, P. H., R. B. Primack, E. R. Ellwood, and E. K. Melaas. 2014. Determining past leaf-out times of New England’s deciduous forests from herbarium specimens. *American Journal of Botany* 101: 1293–1300.

Franks, S. J., S. Sim, and A. E. Weis. 2007. Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proceedings of the National Academy of Sciences* 104: 1278–1282.

Gallagher, R. V., L. Hughes, and M. R. Leishman. 2009. Phenological trends among Australian alpine species: Using herbarium records to identify climate-change indicators. *Australian Journal of Botany* 57: 1–9.

Gallinat, A. S., R. B. Primack, and D. L. Wagner. 2015. Autumn, the neglected season in climate change research. *Trends in Ecology & Evolution* 30: 169–176.

Gries, C., E. Gilbert, and N. Franz. 2014. Symbiota—A virtual platform for creating voucher-based biodiversity information communities. *Biodiversity Data Journal* 2: e1114.

Hill, A., R. Guralnick, A. Smith, A. Sallans, R. Gillespie, M. Denslow, J. Gross, et al. 2012. The notes from nature tool for unlocking biodiversity records from museum records through citizen science. *ZooKeys* 209: 219–223.

Houle, G. 2007. Spring-flowering herbaceous plant species of the deciduous forests of eastern Canada and 20th century climate warming. *Canadian Journal of Forest Research* 37: 505–551.

Johnson, K. G., S. J. Brooks, P. B. Fenberg, A. G. Glover, K. E. James, A. M. Lister, E. Michel, et al. 2011. Climate change and biosphere response: Unlocking the collections vault. *BioScience* 61: 147–153.

Körner, C., and D. Basler. 2010. Phenology under global warming. *Science* 327: 1461–1462.

Lavoie, C. 2013. Biological collections in an ever changing world: Herbaria as tools for biogeographical and environmental studies. *Perspectives in Plant Ecology, Evolution and Systematics* 15: 68–76.

Matthews, E. R., and S. J. Mazer. 2016. Historical changes in flowering phenology are governed by temperature \times precipitation interactions in a widespread perennial herb in western North America. *New Phytologist* 210: 157–167.

Meyer, C., P. Weigelt, and H. Kreft. 2016. Multidimensional biases, gaps and uncertainties in global plant occurrence information. *Ecology Letters* 19: 992–1006.

Miller-Rushing, A. J., R. B. Primack, D. Primack, and S. Mukunda. 2006. Photographs and herbarium specimens as tools to document phenological

changes in response to global warming. *American Journal of Botany* 93: 1667–1674.

Miller-Struttmann, N. E., J. C. Geib, J. D. Franklin, P. G. Kevan, R. M. Holdo, D. Ebert-May, A. M. Lynn, et al. 2015. Functional mismatch in a bumble bee pollination mutualism under climate change. *Science* 349: 1541–1544.

Nelson, G., P. Sweeney, L. E. Wallace, R. K. Rabeler, D. Allard, H. Brown, J. R. Carter, et al. 2015. Digitization workflows for flat sheets and packets of plants, algae, and fungi. *Applications in Plant Sciences* 3: 1500065.

Ozgul, A., D. Z. Childs, M. K. Oli, K. B. Armitage, D. T. Blumstein, L. E. Olson, S. Tuljapurkar, and T. Coulson. 2010. Coupled dynamics of body mass and population growth in response to environmental change. *Nature* 466: 482–485.

Panchen, Z. A., and R. Gorelick. 2017. Prediction of Arctic plant phenological sensitivity to climate change from historical records. *Ecology and Evolution* 7: 1325–1338.

Panchen, Z. A., R. B. Primack, T. Aniško, and R. E. Lyons. 2012. Herbarium specimens, photographs, and field observations show Philadelphia area plants are responding to climate change. *American Journal of Botany* 99: 751–756.

Park, I. W. 2012. Digital herbarium archives as a spatially extensive, taxonomically discriminate phenological record; a comparison to MODIS satellite imagery. *International Journal of Biometeorology* 56: 1179–1182.

Pearse, W. D., C. C. Davis, D. W. Inouye, R. B. Primack, and T. J. Davies. 2017. A statistical estimator for determining the limits of contemporary and historic phenology. *Nature Ecology & Evolution* 1: 1876–1882.

Prevéy, J., M. Vellend, N. Rüger, R. D. Hollister, A. D. Bjorkman, I. H. Myers-Smith, S. C. Elmendorf, et al. 2017. Greater temperature sensitivity of plant phenology at colder sites: Implications for convergence across northern latitudes. *Global Change Biology* 23: 2660–2671.

Primack, D., C. Imbres, R. B. Primack, A. J. Miller-Rushing, and P. Del Tredici. 2004. Herbarium specimens demonstrate earlier flowering times in response to warming in Boston. *American Journal of Botany* 91: 1260–1264.

Rivera, G., and R. Borchert. 2001. Induction of flowering in tropical trees by a 30-min reduction in photoperiod: Evidence from field observations and herbarium specimens. *Tree Physiology* 21: 201–212.

Robbitt, K. M., A. J. Davy, M. J. Hutchings, and D. L. Roberts. 2011. Validation of biological collections as a source of phenological data for use in climate change studies: A case study with the orchid *Ophrys sphegodes*. *Journal of Ecology* 99: 235–241.

Rumpff, L., F. Coates, and J. W. Morgan. 2010. Biological indicators of climate change: Evidence from long-term flowering records of plants along the Victorian coast, Australia. *Australian Journal of Botany* 58: 428–439.

Stucky, B. J., J. Deck, E. Denny, R. P. Guralnick, R. L. Walls, and J. Yost. 2016. W12-04. The Plant Phenology Ontology for Phenological Data Integration. Proceedings of the International Conference on Biomedical Ontology and BioCreative (ICBO BioCreative 2016). CEUR-ws.org, Volume 1747. Available at http://icbo2016.cgrb.oregonstate.edu/sites/default/files/W14-03_ICBO2016.pdf [accessed 21 January 2018].

Visser, M. E., and C. Both. 2005. Shifts in phenology due to global climate change: The need for a yardstick. *Proceedings of the Royal Society of London B: Biological Sciences* 272: 2561–2569.

Walls, R. L., J. Deck, R. Guralnick, S. Baskauf, R. Beaman, S. Blum, S. Bowers, et al. 2014. Semantics in support of biodiversity knowledge discovery: An introduction to the biological collections ontology and related ontologies. *PloS One* 9: e89606.

Wieczorek, J., D. Bloom, R. Guralnick, S. Blum, M. Döring, R. Giovanni, T. Robertson, and D. Vieglais. 2012. Darwin Core: An evolving community-developed biodiversity data standard. *PloS One* 7: e29715.

Willis, C. G., E. R. Ellwood, R. B. Primack, C. C. Davis, K. D. Pearson, A. S. Gallinat, J. M. Yost, et al. 2017a. Old plants, new tricks: Phenological research using herbarium specimens. *Trends in Ecology & Evolution* 32: 531–546.

Willis, C. G., E. Law, A. C. Williams, B. F. Franzone, R. Bernardos, L. Bruno, C. Hopkins, et al. 2017b. CrowdCurio: An online crowdsourcing platform to facilitate climate change studies using herbarium specimens. *New Phytologist* 215: 479–488.

Zohner, C. M., and S. S. Renner. 2014. Common garden comparison of the leaf-out phenology of woody species from different native climates, combined with herbarium records, forecasts long-term change. *Ecology Letters* 17: 1016–1025.