# REVIEW

# The unrealized potential of herbaria for global change biology

EMILY K. MEINEKE,[1,2,3] CHARLES C. DAVIS,[1] AND T. JONATHAN DAVIES[2]

[1]*Department of Organismic and Evolutionary Biology, Harvard University Herbaria, 22 Divinity Avenue, Cambridge, Massachusetts 02138 USA*
[2]*Department of Biology, McGill University, 1205 Dr. Penfield Avenue, Montreal, Quebec H3A 1B1 Canada*

*Abstract.* Plant and fungal specimens in herbaria are becoming primary resources for investigating how plant phenology and geographic distributions shift with climate change, greatly expanding inferences across spatial, temporal, and phylogenetic dimensions. However, these specimens contain a wealth of additional data, including nutrients, defensive compounds, herbivore damage, disease lesions, and signatures of physiological processes, that capture ecological and evolutionary responses to the Anthropocene but which are less frequently utilized. Here, we outline the diversity of herbarium data, global change topics to which they have been applied, and new hypotheses they could inform. We find that herbarium data have been used extensively to study impacts of climate change and invasive species, but that such data are less commonly used to address other drivers of biodiversity loss, including habitat conversion, pollution, and overexploitation. In addition, we note that fungal specimens are under-explored relative to vascular plants. To facilitate broader application of plant and fungal specimens in global change research, we consider the limitations of these data and modern sampling and statistical tools that may be applied to surmount challenges they present. Using a case study of insect herbivory, we illustrate how novel herbarium data may be employed to test hypotheses for which few data exist. With the goal of positioning herbaria as hubs for global change research, we suggest future research directions and curation priorities.

*Key words: climate change; extinction; global change; habitat conversion; herbarium; historical data; invasive species; museum specimens.*

## INTRODUCTION

A key challenge for biologists today is to determine how species are responding to the major drivers of global change and biodiversity loss: habitat conversion and degradation, climate change, invasive species, pollution, and overexploitation (Millennium Ecosystem Assessment 2005). Over the past decades, field observations and experiments have informed much of our understanding of biological responses to these major drivers, particularly climate change. However, like all scientific approaches, they have limitations. Experiments are almost always at smaller spatial scales than inferences (e.g., Pelini et al. 2011) and field observations are often restricted to temperate biomes (Wolkovich et al. 2012). Experiments and observations typically only allow researchers to test hypotheses about a single driver of global change, while organisms are generally exposed to many. Perhaps most critically, the majority of experiments and field observations are short term. Experiments addressing global change are commonly conducted for a few years at most (e.g., Diamond et al. 2012), limited by grant timelines and funding cycles. Field observations are often designed to span latitude and altitude as a proxy for warming, relying on the assumption

that patterns across space will represent future patterns across time (e.g., Kozlov et al. 2013). While long-term field observations can span over a century in cases where people across generations collect the same observations, these data are available for few phenomena, though these include a variety measurements such as insect occurrence and plant diversity (e.g., Warren et al. 2001, Keeling and Whorf 2005, Aono and Kazui 2008, Parolo and Rossi 2008). Scientists have increasingly turned to biological collections to expand data across time, space, and taxonomy, thus better matching the scales at which recent global change is occurring (Pyke and Ehrlich 2010). In particular, herbarium specimens, preserved (often pressed) plants and fungi, have been the subject of a new wave of global change research.

The potential of herbarium data for global change biology stems in large part from its temporal extent. Since the 1700s, scientists, including Linnaeus and Darwin, have collected herbarium specimens to describe new species, aid taxonomic classifications, and as part of regional floristic treatments (e.g., Moffett 2014). Although collecting has slowed in recent years in many localities (Meyer et al. 2016, e.g., Fig. 1a), scientists and amateurs continue to collect. In many parts of the world, such as the northeastern United States, the density of sampling extends far beyond what is available from observations and experiments (Fig. 1b) and encompasses most lineages of vascular plants (Fig. 1c), fungi, diatoms, lichens, bryophytes, and groups variously classified as algae. Current
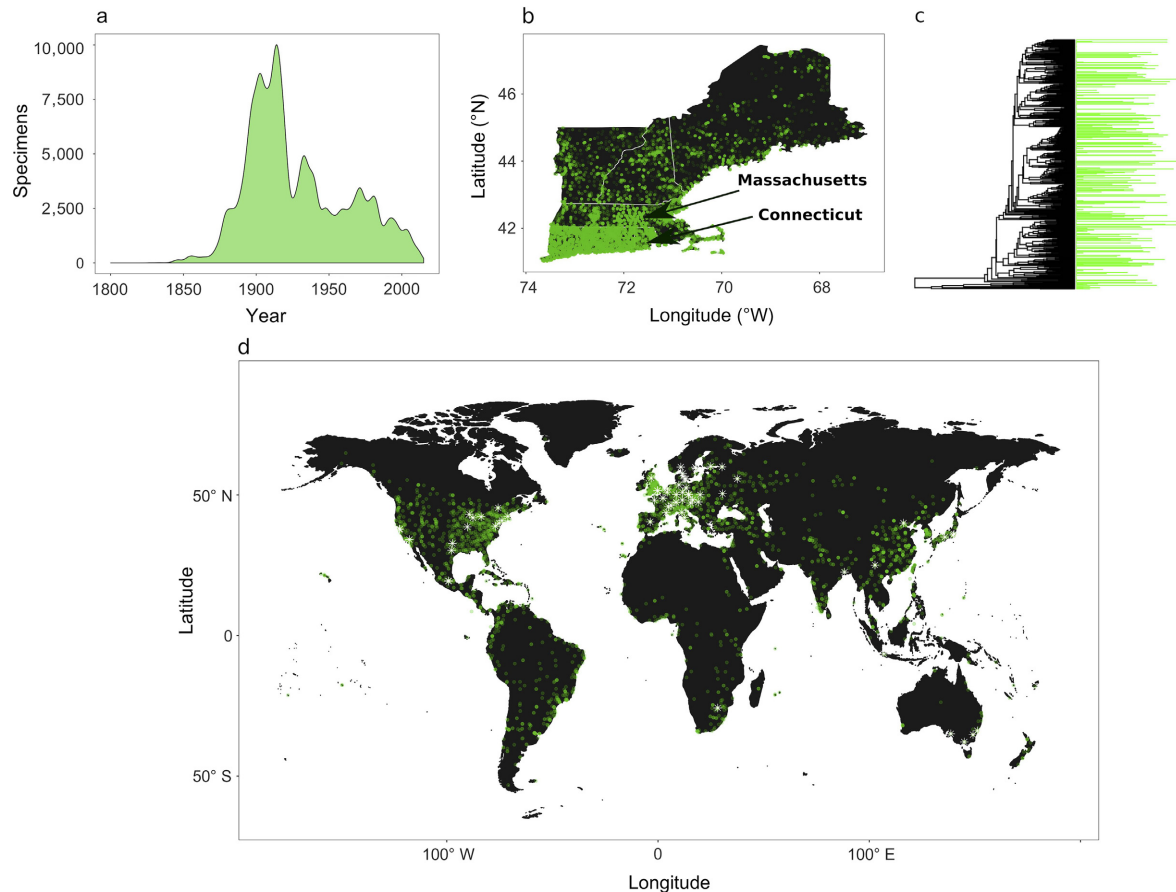
Fig. 1. The spatial, temporal, and phylogenetic extents of herbarium specimens. Here we explore the richness of herbarium data, with emphasis on the major herbaria in the area within the northeastern USA referred to as New England (http://neherbaria.org/). (a) The herbaria of New England house more than one-half million regional herbarium specimens collected as early as the 1800s, before industrialization and globalization, that can serve as baselines prior to anthropogenic change. The density plot represents the number of collections each year. (b) These specimens have spatial coverage across the region (shown here) and millions more specimens are available from other parts of the world. Though there is clear spatial bias, the spatial coverage exceeds that which is typically available in observations and experiments. (c) These specimens span most branches of the vascular plant tree of life. Here, bars represent log abundance of specimens in families, which are organized by evolutionary history (Harris and Davies 2016). (d) Herbaria are distributed worldwide. Many of the larger herbaria have amassed similar coverage to the New England herbaria across these axes. Here, symbols represent herbaria. Herbaria with more than one million specimens are indicated by white stars.

estimates indicate that herbaria house over 350,000,000 specimens (Thiers 2016), representing increasingly connected national networks of data (e.g., the Consortium of Northeastern Herbaria, Chinese Virtual Herbarium) and international efforts to aggregate these data (Fig. 1d; e.g., Australasian Virtual Herbarium, iDigBio, and GBIF).[4,5,6,7,8]

Today, we employ these specimens for purposes that could not have been imagined by their collectors. Evolutionary biologists extract and sequence ancient DNA from herbarium specimens to reconstruct phylogenetic relationships or infer population dynamics (Gugerli et al. 2005, Wandeler et al. 2007) and have documented selection on species traits through changes in plant morphology (Kavanagh et al. 2011, Burns et al. 2012). More recently, ecologists routinely

use herbarium specimens as occurrence records for determining species distributions (Graham et al. 2004, Newbold 2010) and as records of leafing, flowering, and fruiting phenology (Primack et al. 2004, Bolmgren and Lonnberg 2005, Miller-Rushing et al. 2006, Everill et al. 2014) to understand how these aspects of biology are influenced by rising global temperatures. The value of such data is now well recognized (Lavoie and Lachance 2006, Pyke and Ehrlich 2010, Johnson et al. 2011, Lavoie 2013, Vellend et al. 2013, Willis et al. 2017a). The unrealized potential of herbaria is perhaps most apparent in the opportunities presented by other types of data that are only rarely extracted from specimens but which also provide insight into the effects of global change, including signatures of pollination efficiency, pollution concentrations, physiological characteristics, nutrient concentrations, pathogen loads, morphological and anatomical traits, genotypes, endophytes, and herbivory (Table 1). These data could allow researchers to address diverse hypotheses about species and ecosystem responses to global change, from shifts at the species level to changing environmental

[4] http://neherbaria.org/
[5] http://www.cvh.ac.cn/news/8
[6] https://avh.chah.org.au/
[7] https://www.idigbio.org/
[8] https://www.gbif.org/

processes, including nutrient cycles, changes in air quality, and biological control (Table 1).

Given the potential value of herbaria to global change research, it is perhaps surprising that, with the exception of their use as records of species in occurrences and phenology, they have not been more widely used. One explanation is that herbarium data present distinct challenges not present in data gathered from experiments and field observations. Some of these are associated with the vast and dispersed nature of herbarium collections. For example, selecting appropriate focal taxa is critical in any ecological or evolutionary study. While collections contain millions of specimens, they often lack easily accessible digital records. This data gap makes it difficult to determine which taxa are well-represented within and across collections. Other challenges arise as a result of biases; plants are more frequently collected at certain times of year to capture flowering or fruiting, and collection effort has been uneven across space and time (Meyer et al. 2016, Daru et al. 2018). Ongoing digitization efforts may help reduce these biases because digital collections can be aggregated to provide a wider range of specimens than individual collections. However, additional challenges arise because of preservation artefacts; DNA degrades, specimens lose coloration over time, and insects often eat pressed plants housed within museums. Extracting credible data from collections thus requires overcoming significant and complex obstacles.

In contrast to previous perspectives, which have highlighted the potential of herbaria to inform our understanding of species phenology and geographic distributions, particularly with regard to climate change (Primack et al. 2004, Lavoie and Lachance 2006, Pyke and Ehrlich 2010, Johnson et al. 2011, Lavoie 2013, Vellend et al. 2013, Willis et al. 2017a), here, we explore the broader applications of herbarium specimens to global change research. First, we consider less common applications to global change biology, emphasizing novel methods. Next, we discuss current limitations of using herbarium collections for these purposes by identifying research gaps and challenges associated with collecting and analyzing herbarium data. To illustrate how some such challenges may be overcome, we present a case study focused on species interactions, an aspect of global change biology that is data poor, but for which herbarium data may be suited despite biases in the underlying collections. Finally, we outline novel future research directions and suggest curation priorities with the goal of positioning herbaria as primary data repositories for ecological and evolutionary research on the effects of global change.

## Contributions of Herbaria to Global Change Research

Interest in applying herbarium data for investigating global change has grown in recent decades. While climate change has been the focus of most studies, use of herbarium data is slowly permeating other areas of global change biology research. Here, we briefly review the ways in which these data have been employed to examine biological responses to global change and biodiversity loss. We categorize research addressing these drivers into three types of biotic responses: shifts in distributions and population sizes, changes in physiology or morphology, and altered ecological interactions.

Within these categories, we consider each of the five key drivers of change: climate change, habitat conversion and degradation, invasive species, and, where studies are available. pollution, and overexploitation, ordered by how thoroughly they have been investigated. We do not attempt a comprehensive review of all published studies, but rather focus on examples that broadly demonstrate innovative approaches for extracting herbarium data.

### Shifts in distributions and population sizes

Herbarium specimens typically include collection localities and therefore serve as occurrence records that are now widely used to parameterize species distribution models and to understand the effects of recent global change on species geographic ranges (D'Andrea et al. 2009, Feeley 2012, Calinger 2015). For many species, herbarium specimens have revealed plant species range shifts both upward in elevation and poleward in latitude in response to recent warming (Feeley et al. 2013), with some species' ranges contracting while others expand (Feeley 2012). The use of herbarium specimens in this context is now widespread. Though most studies have focused on terrestrial vascular plants, some have included algae (Riera et al. 2015) and species from aquatic systems (Wernberg et al. 2011, Yaakub et al. 2014), highlighting the taxonomic and functional diversity represented in herbaria.

Herbaria occurrence records have also revealed the effects of habitat conversion on species composition, ranges, and abundance. In developed countries, this research has focused on urbanization, arguably the most profound form of global change in these regions (United Nations 2008). Herbarium specimens are among the only floristic records for these areas prior to development and have revealed that urbanization drives native species declines across many cities (e.g., Bertin 2002, DeCandido et al. 2004, Dolan et al. 2011, Gregor et al. 2012, Celesti-Grapow et al. 2013). Herbaria can also provide insights into urban filters on plant communities and have demonstrated that urbanization reduces the occurrence of species associated with wetland habitats (Bertin 2002, DeCandido et al. 2004, Dolan et al. 2011, Gregor et al. 2012, Celesti-Grapow et al. 2013). Unfortunately, while habitat loss through urbanization and deforestation in remote, highly biodiverse regions, such as the wet tropics, is likely a major driver of extinction (e.g., see Wearn et al. 2012), herbarium specimens from many of these regions are relatively sparse (Meyer et al. 2016) and thus provide poor baselines from which to derive estimates of biodiversity change (Feeley and Silman 2011). Nonetheless, herbarium data can help identify species in decline and regions that may provide refuges (Farnsworth and Ogurcak 2006, Romeiras et al. 2014), as well as capture the interactions between climate change and habitat conversion, notably to demonstrate how human land use limits the extent to which species can track their climatic niches (Feeley and Silman 2010).

While many native species are declining, nonnative species are increasing in abundance, and a small proportion of these have become invasive (Mack et al. 2000). Herbaria provide historical records of invasive species spread through time and across space (Lavoie et al. 2007, Crawford and

REVIEW

TABLE 1. Key research questions and hypotheses from the global change literature for which herbaria may be well suited as a data source.

| Research questions | Herbarium data | Hypotheses | Keywords | Citations |
|---|---|---|---|---|
| **Climate change** | | | | |
| Are phenological shifts increasing or decreasing in magnitude? | flowering plant, fungal fruiting leaf-out locality, date collected | Plant flowering and leaf-out are becoming less sensitive to climate as species reach their tolerance thresholds. Fungal fruiting patterns are changing across the season due to the redistribution of rainfall. | "Climate change" & "Phenolog*" or "Flower*" or "Leaf-out" or "Fruit*" or "Seed*" | 95; Diez et al. (2013), Kauserud et al. (2010), Matthews and Mazer (2016) |
| How does plant geographic location affect phenological responses to climate change? | flowering plant, fungal fruiting leaf-out locality, date collected (latitude, longitude) | Within and between species, plant phenological cueing mechanisms vary across latitude. | same as row above | same as row above |
| Are species' ranges shifting with climate change? | locality, date collected (historical and current ranges) | Species are moving poleward and up elevations due to climate change. Human development limits plant species movement poleward. | "Climate change" & "Species distribution" or "Range" & "Latitude" or "Elevation" or "Urbanization" | 20; Feeley et al. (2013), Feeley and Silman (2010) |
| Does dispersal syndrome influence plant range shifts due to climate change? | fruit dispersal mode locality, date collected (historical and current ranges) | Bird-dispersed plants are able to migrate longer distances than those dispersed simply by gravity. | "Climate Change" & "Dispersal" | 19 |
| How does climate change affect pest/pathogen abundance/ diversity/community structure? | herbivory; insects and their damage with specimens that can be assigned species identity, e.g., leaf mines, galls (historical and current ranges); pathogen lesions, DNA, RNA (historical and current ranges; host shifts); locality, date collected (community structure) | Herbivore abundance and damage increases with warming, consistent with fossil evidence. Herbivore distributions are expanding north and/or retracting at southern limits due to climate warming. Warmer climates increase the geographic spread of pathogens/herbivores, facilitating shifts to novel hosts. In areas that are warming, communities shift and become less structurally even due to increasing abundance of species with high thermal tolerances. | "Climate Change" & "Pest" or "Herbiv*" or "Pathogen" or "Disease" | 12; Youngsteadt et al. (2015) |
| Have pollination rates changed through time? If so, are declines driven by climate change? Does climate change interact with other global change drivers, such as urbanization? | flower morphology, e.g., corolla length; metrics of fitness, such as seed size, set; pollen protein concentrations; pollen removal | Pollination has decreased due to a combination of drivers, including climate change. | "Climate Change" & "Pollin*" or "Pollen" | 25; Miller-Struttmann et al. (2015), Pauw and Hawkins (2011), Ziska et al. (2016) |
| How have climatic niches changed, and how well do current environmental niche models match past plant distributions? | Locality, date collected (realized historical and current niches) | Climate envelopes predict geographical distributions, such that species fill newly available niche space driven by climate change | "Climate Change" & "Niche" or "Species distribution" or "SDM" | 59 |
| Are climatic changes and plant range shifts associated with shifts in gas exchange rates? | gas exchange proxies, e.g., stomatal densities; isotope ratios, and guard cell lengths; water use efficiency proxies, e.g., stomatal density, molecular hydrogen isotope composition δD | Climate change has increased photosynthetic rates, except when plants are water stressed. | "Climate change" & "Photosynthe*" or "Gas exchange" or "Stomata" or "Stomatal conductance" | 7; Miller-Rushing et al. (2009) |

TABLE 1. (Continued)

| Research questions | Herbarium data | Hypotheses | Keywords | Citations |
|---|---|---|---|---|
| Are plants evolving in situ to climate change? Are cooler adapted plant genotypes being displaced by warmer adapted genotypes via migration? | plant DNA (allele frequencies); morphology, e.g., leaf size, shape, specific leaf area | Plant gene frequencies and phenotypes have shifted over time, and these shifts are consistent with changing climate, e.g., environments that have become drier select for more dry-tolerant phenotypes. | "Climate change" & "Evolution" or "Adaptation" | 36 |
| What roles do plant water relations play in responses to climate change across phylogeny and habitats? | water use efficiency proxies, e.g., stomatal density, molecular hydrogen isotope composition δD; gas exchange proxies, e.g., stomatal densities, isotope ratios, and guard cell lengths; morphology, e.g., leaf size, shape, specific leaf area; plant carbon content | Warming increases plant productivity at mid and upper latitudes, except when plants are water stressed. Plants worldwide are living close to their hydraulic limits. | "Climate change" & "Water-use efficiency" or "Water stress" or "Water potential" | 4 |
| Does climate change promote phenological asynchrony and/or ecological mismatch between plants and associated species? | flowering; flower morphology, e.g., corolla length; Plant, fungal fruiting; leaf-out; herbivory; pollen removal | Plant and insect phenology shift at similar rates with temperature, such that pollination and herbivory rates are constant despite climate change. Plants and their insect/mycorrhizal associates respond to different cues and thus will become/are less synchronized due to climate change. Asynchronies driven by climate change will be reduced over time by rapid selection for insects to synchronize with plants and in the case of pollination, and vice versa | "Climate change" & "Synchrony" or "Asynchrony" or "Ecological mismatch" | 5; Kharouba and Vellend (2015), Miller-Struttmann et al. (2015) |
| **Invasive species** | | | | |
| Is invasive plant spread facilitated by genomic change? | plant DNA (allele frequencies); locality, date collected (time of introduction, spread) | New mutations or gene combinations enable invasive species to overcome dispersal barriers, perhaps via gene surfing on expanding population fronts. | "Invasive" or "Non-native" & "Genome" & "Adaptation" or "Genomic change" | 12; Buswell et al. (2011), Vandepitte et al. (2014) |
| Have invasive plants demonstrated greater phenological advancement with climate warming than native species? | flowering; leaf-out | Greater phenological advancement of nonnative compared to native species facilitates invasions. | "Invasive" or "Non-native" & "Phenolog*" or "Flower*" or "Leaf-out" or "Fruit*" or "Seed*" | 46; Calinger (2015) |
| What are the physical pathways of invasive plant spread? | plant, leaf miner DNA; locality, date collected (time of introduction, spread) | Natural pathways, such as waterways, were historically more important for invasive plant and insect species spread, but increasingly roads and railroads are key. | "Invasive" or "Non-native" & "Spread" or "Railroad" or "Road" | 80; Barney (2006), Joly et al. (2011), Saltonstall (2002) |
| In novel habitats, does release from natural enemies promote invasive plant spread? | herbivory; insects and their damage that can be assigned species identity, e.g., leaf mines, galls; pathogen lesions, DNA, RNA; plant defensive compounds | One mechanism by which species become invasive is escape from co-evolved natural enemies. (Enemy Release Hypothesis). | "Invasive" or "Non-native" & "Natural enem*" or "Natural enemy release" | 2; Zangerl and Berenbaum (2005) |

REVIEW

REVIEW

TABLE 1. (Continued)

| Research questions | Hypotheses | Herbarium data | Keywords | Citations |
|---|---|---|---|---|
| What roles do plant diseases play in invasions? | Diseases carried by nonnative plants can facilitate their invasions via apparent competition. | pathogen lesions, DNA, RNA; locality, date collected (time of introduction, spread) | "Invasive" or Non-native" & "Pathogen" or "Disease" | 8; Malmstrom et al. (2007) |
| Does exotic plant relatedness to natives determine invasiveness? | Exotic insect herbivores and pathogens are more likely to establish on novel host plants closely related to their co-evolved host plants. Exotic plant/pathogen/ herbivore relatedness to native plants reduces the probability that they become invasive. (Darwin's Naturalization Hypothesis). | herbivory; insects and their damage that can be assigned species identity, e.g., leaf mines, galls (time of introduction, spread, host shifts); plant defensive compounds; locality, date collected (time of introduction, spread) | "Invasive" or "Non-native" & "Naturalization Hypothesis" | 0; though this search returns no references, see Park and Potter (2013) and Schaefer et al. (2011) |
| **Habitat conversion** | | | | |
| How do restored plant communities and their associates compare to pre-disturbance communities? | Effects of restoration on diversity and community structure depend on land use legacies. | herbivory; insects and their damage that can be assigned species identity, e.g., leaf mines, galls; pathogen/endophyte/ mycorrhizal DNA; locality, date collected (community metrics) | "Land use" or "Disturbance" & "Restoration" | 5; Bertin (2002), Celesti-Grapow et al. (2013), DeCandido et al. (2004), Dolan et al. (2011), Gregor et al. (2012) |
| Can we use plant species occurrences prior to habitat change to guide restoration efforts? | Herbaria capture historical diversity and thus could serve as blueprints for restoration. | locality, date collected (species composition prior to intensified anthropogenic change) | "Restoration" | 37 |
| Has global change led to no-analog plant communities? | Habitat conversion, trade, climate change, among forms of global change, have led to novel plant and fungal communities. | locality, date collected (historical community structure) | "No-analog communit*" or "Novel communit*" | 0 |
| Are some plant community structures more robust to disturbance and/or invasion? | More diverse plant communities are more resilient to herbivore/invasive plant/ pathogen pressure. (Biodiversity Insurance Hypothesis). | locality, date collected (historical community structure, diversity); plant DNA (phylogeny reconstructions) | "Insurance Hypothesis" or "Resilience" & "Diversity" or "Richness" | 2 |
| How do different types of habitat change filter plant species and their microbial/arthropod associates? | Agriculture, urbanization, and other types of human development have signatures, such that biota in these habitats worldwide share common traits, and communities include certain species with global distributions. (Biotic Homogenization Hypothesis). | herbivory; insects and their damage that can be assigned species identity, e.g., leaf mines, galls; pathogen/endophyte/ mycorrhizal DNA | "Ecological filter" or "Habitat filter" | 3 |
| What is the timescale of natural plant restoration? | Landscape connectivity increases the rate at which plants and their associates re-enter habitats. | herbivory; insects and their damage that can be assigned species identity, e.g., leaf mines, galls; pathogen/endophyte/ mycorrhizal DNA; locality, date collected time series of (community structure, diversity) | "Re-establish*" or "Brownfield" or "Succession" or "Regeneration" | 37 |

Table 1. (Continued)

| Research questions | Herbarium data | Hypotheses | Keywords | Citations |
| --- | --- | --- | --- | --- |
| How do plant associates respond to land use change? | herbivory; insects and their damage that can be assigned species identity, e.g., leaf mines, galls; pathogen/endophyte/mycorrhizal DNA; pollen protein concentrations; pollen removal | Biodiversity of plant associates has decreased over time due to habitat conversion. Urbanization increases/decreases insect abundance/diversity/herbivory/pollination. Development releases some insects from their co-evolved natural enemies, allowing them to become pests (i.e., Enemy Release Hypothesis). Reduced endophyte loads and diversity in cities affects plant interactions with pollinators and herbivores. | "Land use" or "Disturbance" & "Herbiv*" or "Pathogen" or "Disease" or "Pollin*" or "Pollen" or "Endophyte" or "Mutualis*" or "Commensal*" or "Competit*" or "Predat*" | 17; Pauw and Hawkins (2011), Youngsteadt et al. (2015) |
| What selection pressures does land use change impose on plants? | morphology, e.g., leaf size, shape, specific leaf area; plant DNA (allele frequencies) | Cities worldwide, except deserts, select for species that can withstand relatively hot, dry, and open conditions. | "Land use" or "Disturbance" & "Adapt*" or "Selection" or "Evolution" | 14; Dolan et al. (2011), Neil et al. (2010) |
| How has habitat conversion affected plants of cultural and economic importance, such as medicinal plants and wild crop relatives? | locality, date collected (abundance, historical and current ranges); morphology, e.g., leaf size, shape, specific leaf area; metrics of fitness, such as seed size, set; physiological characteristics, such as stomatal densities, isotope ratios, and guard cell lengths | Habitat conversion has reduced abundance and range sizes of many species. | "Land use" or "Disturbance" & "Economic" or "Crop" or "Medicinal" | 14; Farnsworth and Ogurcak (2006) |
| Pollution | | | | |
| How has atmospheric pollution altered community composition of plants and fungi? | pollutant concentrations; locality; date collected (community structure, diversity) | Pollution selects for resistant species and lineages, driving phylogenetic under-dispersion and reduced diversity. | "Pollut*" & "Communit*" or "Assembl*" | 8 |
| Does biotic diversity enhance bioremediation, i.e., rates of pollutant removal, from a system? | same as row above | More diverse communities remove pollutants more efficiently. | "Pollut*" & "Bioremediation" or "Recovery" & "Diversity" or "Richness" | 0 |
| How has exposure to atmospheric pollution varied through time and across regions? | same as row above | Humans began polluting environments early in our history. | "Pollut*" & "Histor*" & "Human" | 6 |
| How does atmospheric pollution affect plant associates? | pollutant concentrations; locality, date collected (community structure, phylogenetic diversity); pollutant concentrations; herbivory; insects and their damage that can be assigned species identity, e.g., leaf mines, galls; leaf miner DNA (allele frequencies); pathogen/endophyte/mycorrhizal DNA; pollen removal | Pollution reduces plant associate diversity by selecting resilient species and genotypes. | "Pollut*" & "Herbiv*" or "Pathogen" or "Disease" or "Pollin*" or "Pollen" or "Endophyte" or "Mutualis*" or "Commensal*" or "Competit*" or "Predat*" | 3 |

REVIEW

REVIEW

TABLE 1. (Continued)

| Research questions | Herbarium data | Hypotheses | Keywords | Citations |
|---|---|---|---|---|
| Are pesticides, e.g., neonicotinoids/ chlorpyrifos, responsible for the sudden decline of many insects? | insects and their damage that can be assigned species identity, e.g., leaf mines, galls; pesticide residues in/on pollen and leaves; pollen removal | Pesticide presence has increased in natural plant populations over time. Pesticides are present in a diversity of non-crop plants. Wild insect diversity/abundance/ herbivory has decreased due to pesticides. Pollen transfer has declined with pesticide use | "Pollu*" & "Pesticide" & "Insect" | 0 |
| *Exploitation* | | | | |
| Do protected areas truly protect threatened and endangered plant species? | locality, date collected (historical and current occurrence within protected areas) | Because the locations of protected areas are driven by human concerns, and rarely informed by diversity data, many species are not protected. The amount of biodiversity preserved over time depends on management regimes, and resulting habitat patch size and connectivity. | "Protected area" and "Threatened species" or "Endangered species" | 3; Romeiras et al. (2014) |
| Can we predict plant extinctions and populations declines due to exploitation? | metrics of fitness, such as seed size, set; morphology, e.g., leaf size, shape, specific leaf area; physiological characteristics, such as stomatal densities, isotope ratios, and guard cell lengths; plant DNA (allele frequencies) | Elevated rates of genotypic/phenotypic change precede population collapse. | "Harvest" and "Extinction" or "Extirpation" or "Decline" | 4 |
| What are the selection coefficients imposed on plants by human harvests? | Morphology, e.g., leaf size, shape, specific leaf area; plant DNA (allele frequencies) | Human harvests of wild plants reduce plant size by inducing selection pressure against larger individuals. | "Harvest" & "CITES" or "Exploitation" & "Selection" or "Adaptation" or "Morphology" or "Size" or "Height" or "Specific leaf area" | 36; Law and Salick (2005) |

*Notes:* Most of the research questions apply to multiple global change drivers and are grouped by the global change driver that has received the most interest. These suggestions are biased toward plants, for which more diverse herbarium data extraction methods have been developed, though we believe many could also be tested for fungi. We searched Web of Science to estimate the current interest in each research topic. Searches were performed with "Herbari*" &' the listed keywords. The citations column contains the number of citations retrieved and, when available, iconic, representative examples from the literature.

Hoagland 2009) and thus can help identify the mechanisms by which nonnative species have dispersed from continent-to-continent and expanded their geographic ranges. Analyses of plant occurrence records from herbaria have revealed that human transportation networks, notably paved roads and railroads, are important pathways for invasion (e.g., Barney 2006, Joly et al. 2011). With recent advances in molecular techniques, it is now possible to map the spread not just of species but also particular genotypes using herbarium data, as has been illustrated for the common reed, *Phragmites australis* (Saltonstall 2002). These fine-scale historical data can provide information on both the rate and direction of spread, revealing likely centers of introduction and regions of high vulnerability.

### Physiological and morphological change

Most species are unable to completely escape global change in space and thus must adapt or acclimate in situ or risk extinction. Herbarium specimens can capture physiological and morphological changes reflecting such responses, including shifts in morphology (Law and Salick 2005, Leger 2013), timing of life histories (Kharouba and Vellend 2015), and physiology (Miller-Rushing et al. 2009). In a small but growing number of studies, physiological and morphological data, such as on plant leaf size (Guerin et al. 2012), stomatal densities (Miller-Rushing et al. 2009), carbon and oxygen isotope measurements (Miller-Rushing et al. 2009, Bonal et al. 2011), and specific leaf area (Reef and Lovelock 2014), have been extracted and provide insights into changing photosynthetic rates and leaf palatability for herbivores, for example.

The most studied of these shifts with herbarium data are changes in plant phenology. Flower counts from herbarium specimens have revealed advances in peak flowering of approximately 2.4 d for each 1°C rise in temperatures (Calinger et al. 2013), and similar advances due to the urban heat island effect (Primack et al. 2004) that are particularly pronounced in ephemeral species (Neil et al. 2010). Researchers have also turned to herbaria to identify potential cues driving plant phenology, revealing interactive effects of temperature, precipitation, and latitude on the timing of flowering (Matthews and Mazer 2016). Such data have proven extremely valuable as both a biotic index of climate change, and as a record of biotic responses to climate warming. The extraction and application of phenological data from herbaria has been reviewed extensively elsewhere (Miller-Rushing et al. 2006, Willis et al. 2017a).

Plant phenology is just one response to changes in atmospheric chemistry. Concentrations of pollutants in the atmosphere, including heavy metals, anthropogenic nitrogen, carbon dioxide ($CO_2$), and other greenhouse gases have varied over time, with increases surprisingly early in modern human history (Renberg et al. 1994, Steffen et al. 2007). However, because historical data are sparse, and pollutant sources are often diffused, changes in pollution can be hard to track using traditional ecological approaches. Herbarium specimens serve as records of pollutant variation over space and time, helping link species' exposure to their responses. For example, epiphytes, which accumulate atmospheric nitrogen, but do not uptake nitrogen from soil, can serve as bioindicators of nitrogen pollution (Stewart et al. 2002).

Mosses similarly serve as bioindicators of atmospheric metals (Weiss et al. 1999), and lichens serve as bioindicators of various pollutants, including a diversity of oxidants (Sigal and Nash 1983).

Data from herbaria can also be used to identify the traits associated with invasiveness. A key question in invasion biology is why few species become invasive while most nonnative species remain at low abundances (Sakai et al. 2001). By capturing physiological and morphological data, herbarium studies have shown that adaptation to local conditions can facilitate invasiveness (Vandepitte et al. 2014) and, consistent with inferences drawn from other types of data (Wolkovich and Cleland 2011), that some invasive species are able to better track temperature and thus take advantage of earlier springs resulting from warming (Calinger 2015).

### Shifts in ecological interactions

Ecological interactions are increasingly recognized as moderating species responses to global change (Gilman et al. 2010, Zarnetske et al. 2012), but empirical data are sparse. Herbarium specimens have served as records of the interactions between plants and their associates, revealing how these interactions have shifted over time. These plant associates, including insect pollinators, herbivores, and pathogens, are sometimes preserved on leaves (Lees et al. 2011) and branches (Youngsteadt et al. 2015) or are preserved as DNA or RNA (Malmstrom et al. 2007). In addition, flowers and leaves of herbarium specimens can contain signatures of interactions, such as the accumulation of defensive compounds induced by insect herbivores (Zangerl and Berenbaum 2005) and pollen (Ziska et al. 2016).

The trace-record of interactions between plants and pollinators, herbivores, and pathogens captured on herbarium specimens could be used to address a topic that has attracted much interest in climate change research: potential phenological asynchronies and ecological mismatches between associated species resulting from differential responses to warming climates (Post et al. 2008, Both et al. 2009). Phenological asynchronies occur if the direction, rate, or magnitude of change differs between associated species, if phenological responses differ in space, or if responses within a single trophic level are varied, such that phenological tracking is not possible (as discussed in detail elsewhere, e.g., Hegland et al. 2009). Occurrence data from herbaria and contemporaneous insect collections have suggested that climate change might lead to asynchronies between some butterflies and their adult food plants (Kharouba and Vellend 2015). Other types of ecological mismatches can also disrupt mutualisms. Corolla tube length data from herbarium specimens, along with corresponding bumble bee collections that provided data on bee tongue length, showed that bees and flowers may be ecologically mismatched in Colorado, USA; bee tongue lengths have decreased over time, while corolla tube lengths have remained constant (Miller-Struttmann et al. 2015). However, there are few studies along these lines; we expand on the potential of herbarium data for studying phenological asynchronies in the following section.

Similar data can serve as records of changing species interactions due to habitat conversion, pollution, and invasions. Anthropogenic disturbance of a habitat can expose

REVIEW

species to novel conditions that disrupt co-evolved interactions. By rehydrating orchid flowers from herbaria and counting the pollinaria (pollen clumps and associated tissues removed by bees), Pauw and Hawkins (2011) demonstrated that the local decline of an orchid during urbanization was driven by reduced pollination. Youngsteadt et al. (2015) counted scale insects on stems of herbarium specimens and demonstrated that the urban heat island effect and natural warming cycles in forests are associated with elevated abundance of a scale insect herbivore, suggesting that ecological responses to warming in cities might predict the effects of climate change on insect pests. In one of the few herbarium studies to explore how pollution affects species interactions, Ziska et al. (2016) analyzed pollen preserved in herbarium specimens and demonstrated that elevated $CO_2$ concentrations reduced pollen protein concentrations, altering nutrient availability for pollinators, with possible fitness consequences for both pollinators and plants. In a separate study, Zangerl and Berenbaum (2005) provided support for the natural enemy release hypothesis of plant invasion (Keane and Crawley 2002, Mitchell and Power 2003) by measuring exotic plant defensive compounds before and after introduction of their co-evolved herbivores. They found evidence that invasive plants can escape their insect herbivores, and the need to manufacture expensive defensive compounds in response to them, when first introduced into new habitats.

### GAPS IN THE APPLICATION OF HERBARIUM DATA

#### Overview

We have provided a brief overview of the diverse applications of herbarium data for understanding biotic responses to global change. However, many of these data have only recently become commonplace in ecological studies and we suggest their potential has not been fully realized. Here, we identify significant gaps in current uses of herbarium specimens in global change biology (Table 1). As we indicate above, data from herbaria have been used extensively to explore plant distributional and phenological shifts in response to climate change. However, herbarium specimens have been used less frequently to study plant responses to three of the five key drivers of biodiversity loss: habitat conversion, pollution, and overexploitation. By many estimates, habitat conversion currently drives most terrestrial species extinctions (Millennium Ecosystem Assessment 2005), making this a particularly notable omission. Pollution and overexploitation are difficult to track using modern observations, in part because they are highly heterogeneous across space and time. Thus, we argue that herbaria could provide important novel data on these global change drivers. With these overarching themes in mind, we suggest key future research directions we believe could benefit from using herbarium data. We acknowledge that some of the research areas described in this section and detailed in Table 1 may be more fruitful than others. We present these topics as research areas we believe deserve exploration based on the fact that data sampling and/or extraction methods have already been established, but have not yet been applied widely to address global change hypotheses.

#### Guidance for ecosystem management and restoration

Herbarium specimens are rare records of historical biodiversity and thus could help guide ecological restoration. A few studies have highlighted this potential, mostly within the context of urban environments (DeCandido et al. 2004, Atha et al. 2016). However, less common are examples of using plant occurrence records from herbaria to explore other types of habitat conversion, such as agricultural expansion and deforestation, much less integrate such findings into restoration programs. As is the case for urbanization, these types of habitat transformations are likely to have nuanced effects on local biodiversity that are not well understood but which may be captured by collections. Herbarium specimens can additionally be used to determine areas and species of conservation priority. For example, Romeiras et al. (2014) used occurrence records from herbarium specimens to determine timber species that were of high conservation priority in Angola based on their range sizes and the extent of their ranges that overlapped with protected areas in the region. We suggest there may be many opportunities along these lines for conservation of plants of economic and/or cultural importance, including both timber species and wild crop relatives, which are often well-represented in herbaria, with many specimens collected prior to intensification of global change.

#### Markers of plant physiological change

Specimens within herbaria provide more than just records of occurrence across space and time; they also bear the imprint of past environments, including information on genotypic and physiological shifts. For example, herbarium specimens could be employed to collect long-term data on shifts in traits correlated with photosynthetic rates and nutrient concentrations across taxa and habitats. These are ecosystem variables for which we have little historical data, but which have important consequences. For example, leaf gas exchange rates can affect carbon balance (Bonan 2008), and the net effects of global change, particularly climate change, on this and similar ecosystem processes is a topic of great interest (Clark 2004, Wu et al. 2011). Increasing $CO_2$ levels are expected to increase photosynthetic rates in plants, increasing net carbon storage in forests (Ainsworth and Long 2005). However, long-term experiments to evaluate the relationship between $CO_2$ and photosynthetic rates have been in place for <20 yr (Norby et al. 2016) and do not incorporate other recent global changes that may have interactive effects on photosynthesis, most notably warming, changes in soil moisture, and nitrogen deposition. Methods have been developed to extract such data from herbarium specimens, for example, quantifying proxies for gas exchange preserved in leaves of pressed plants (Miller-Rushing et al. 2009, Bonal et al. 2011), and are not new (Woodward 1987), but they have been underutilized for understanding the long-term effects of global change on plant physiology. In addition, it may be possible to use herbarium specimens to determine if nutrient limitation has changed over time due to the combined effects of $CO_2$ fertilization and nitrogen deposition. For example, McLauchlan et al. (2010) showed with carbon and nitrogen isotopes that

nitrogen availability is declining in grasslands in the midwestern United States.

### Records of changing species interactions

We suggest that herbarium specimens could provide unparalleled insights into shifting species interactions in the Anthropocene. Above, we outlined a few studies along these lines, but we believe the data within herbarium specimens has not been fully exploited. For example, presently, there is little consensus on whether climate change induced phenological asynchronies are common or rare, in large part because data are sparse. It is, however, possible to pair herbarium records with collections of species with which they are tightly associated, such as their pollinators. This approach is potentially powerful, but requires substantial data. It is possible, therefore, that for many species, collections will not yield data of sufficient spatial and temporal resolution to quantify phenology for associated species. An alternative approach is to draw indirect inferences on species interactions using data from only one partner; observations of pollen and pollinaria, for example, can provide information on plant–pollinator interactions (Pauw and Hawkins 2011, Ziska et al. 2016). Such methods can be more easily expanded to systems for which corresponding collections on interactors are sparse or unavailable. We outline potential hypotheses on pollination that herbaria might be used to test in detail in Table 1.

In a number of studies, researchers have used similar approaches to explore effects of global change on antagonistic interactions between plants and their associates by extracting data on herbivores, herbivory, and plant pathogens (Malmstrom et al. 2007, Lees et al. 2011, Youngsteadt et al. 2015, Syfert et al. 2017). However, methods remain relatively underdeveloped. Plant–insect interactions might be of particular interest, because insects are ectothermic, and, thus, their abundances are expected to shift with global climate warming in many cases (Kingsolver et al. 2013). Herbivory is also ubiquitous among present day plants (Turcotte et al. 2014), has driven evolution of much of the biodiversity on earth (Futuyma and Agrawal 2009), and has significant economic consequences in agriculture and forestry (Oerke and Dehne 2004). Despite its importance as an ecological process, we are aware of few studies quantifying plant–insect interactions on individual specimens (but see Morrow and Fox 1989). Herbarium specimens could be used to understand herbivore and pathogen responses to a number of aspects of global change, including global warming, urbanization, pollution, and plant range shifts. Below, we provide a case study detailing how one might approach using herbarium specimens for understanding species interactions.

### HERBARIA AS NOVEL DATA SOURCES: LIMITATIONS AND CHALLENGES

We have shown that herbarium data can be applied to diverse topics and have suggested gaps that warrant future exploration. Here, we outline challenges these data present and, when available, approaches to reduce obstacles to their use (see also Table 2), which we believe is a key hurdle to using herbarium data in global change research. We consider some general methods that could be applied widely to herbarium and biological collections data.

### Biases over space, time, and phylogeny

Herbarium specimens are non-randomly collected across space and time, in part because their historical purpose was to document species' occurrences and capture morphological variation within and between species, and not to address ecological questions. Using specimens for global change research requires accounting for this unevenness in sampling over space, time, and taxonomy. This topic has been reviewed recently by Meyer et al. (2016) and Daru et al. (2018), who analyze biases represented in herbarium specimens and observational data. Though the sampling biases described in these publications should serve as a roadmap for those expected more generally in herbarium data, the gaps and biases present across entire herbaria do not necessarily represent those in the subset of data extracted for particular studies. For example, an herbarium may include few species from a particular province in China, but have many specimens for each of these species.

A first step in any ecological analyses is to design appropriate sampling procedures to minimize biases. There is a rich literature on sampling techniques in ecology (see Southwood and Henderson 2009), but these have been largely overlooked when "sampling" herbarium specimens. Instead, when deriving data from specimens in ecological research, it is more common to conduct systematic sampling, analyzing specimens within a specified timeframe and/or spatial area rather than considering herbarium specimens as a sample drawn from a larger population. Now that centralized databases contain millions of specimens, traditional ecological tools, such as stratified random resampling or rarefaction, may allow researchers to minimize or quantify biases in their data. In addition, researchers can focus analyses on the richer parts of collections, thus reducing noise and bias introduced by taxa, time periods, or locations for which few specimens are available. For example, if we were to sample a particular plant species across the northeastern United States (Fig. 1b), we may find that more recent samples are available for Connecticut than for Massachusetts, a known pattern given the historical idiosyncrasies of institutions and collections in the region. Therefore, if we were interested in testing a global change hypothesis that requires a time series, we may decide to focus our sampling in Connecticut to take advantage of its time series and avoid the uneven temporal coverage across space in Massachusetts.

Even when sampling procedures are designed carefully, however, it may not be possible to select specimens evenly across all axes of variation. Numerous modern statistical tools are available to account for uneven sampling and non-independence of data. These include, notably, spatial regression techniques, such as spatial autoregressive models, spatial lag models, and spatial error models (Plant 2012), that account for spatial non-independence in data. Various phylogenetic comparative methods, including independent contrasts (Felsenstein 1985) and phylogenetic generalized least squares regression (PGLS), are available for incorporating phylogenetic non-independence among taxa. There is also a large literature on null models in ecology (Gotelli and

TABLE 2. Herbarium data for global change research.

| Challenges | Potential solutions |
| --- | --- |
| **All herbarium data** | |
| Biases over space, time, and phylogeny | Careful focal herbaria/species selection aided by digitized specimens; subsampling and techniques for inference with biased data, and statistical methods such as machine learning. |
| **Occurrence data (locality, date collected)** | |
| Lack of recorded absences | Statistical tools for simulating or otherwise analyzing presence-only data, which are already well developed for species distribution models and may be employed for other areas of research |
| Coarse-level geographical data | Careful focal herbaria/species selection aided by digitized specimens. In some cases, it may be possible to address hypotheses with environmental data at coarse scales, e.g., temperature data averaged at the county level in the United States. |
| **Phenology data** | |
| Specimens are most likely to be collected at peak flowering times, thus missing early season phenological events and times of first event. | Use statistical estimators to infer timing of first events from a sampled distribution (Pearse et al. 2017). |
| Collectors tend to retrieve specimens from near roadsides (Daru et al. 2018), which could affect phenology. For example, life events may be advanced relative to plants in more natural areas due to the urban heat island effect or delayed due to drought-like conditions. | When comparing across space, time, and/or phylogeny, relative measures may be sufficient to address hypotheses of interest. For studies that do require absolute measurements, use specimens with geolocations and model potential biases (e.g., urbanization). Post-hoc georeferencing of many historical specimens may greatly improve the data that are currently in online databases. Additional studies are needed to assess the effects of local collection biases on global change data, especially for plant traits that are sensitive to temperature, such as phenology |
| For leaf-out, categorical criteria, such as pubescence, can be used to determine when specimens have newly flushed leaves (see Everill et al. 2014). However, many species do not have unique qualities associated with new leaves. | Continuous criteria can be developed, such as leaf size relative to fully flushed leaf size (also see Everill et al. 2014), but such criteria introduce considerably more work than do categorical characteristics. |
| **Other trait data (morphology, physiology, fitness, and microbiomes)** | |
| Botanists may be more likely to collect specimens from more vigorous individuals or branches with greater seed set, seed size, leaf size, water use efficiency, etc. | When comparing across space, time, and/or phylogeny, relative measures may be sufficient to address hypotheses of interest. Cross-validate recent collections with observational data to ascertain, and thus allow correction for, potential sampling biases. |
| Key tissues, morphological structures, or developmental stages may not be sampled or may be damaged. For example, roots are necessary for mycorrhizal studies but are not present on all specimens. | Check digital images and select intact specimens. If alleles for a trait are known, it may be possible to genotype a sample without a visible phenotype and thus avoid the need for the actual structure possessing the trait of interest. For certain species, roots can be easily collected and are thus present on many herbarium specimens. Overcoming this challenge should be a matter of choosing appropriate focal taxa |
| **Pollination data** | |
| Matching data on pollinators are often not available. | Score pollination on herbarium specimens. This will only be possible for groups that have distinct pollen-holding structures (pollinaria), such as many orchids and milkweeds. It may also be possible to assess pollen/pollinator limitation through estimates of fruit or seed set per flower; or by estimating pollen deposition on stigmas. |
| To build matching plant–pollinator collections, one needs to know which pollinators are associated with which plants. | Identify plants via pollen morphology or DNA preserved on pollinator specimens. |
| **Antagonistic interaction data (herbivory, disease)** | |
| Botanists may be likely to collect less damaged specimens. | When comparing across space, time, and/or phylogeny, relative measures may be sufficient to address hypotheses of interest. Cross-validate recent collections with observational data to ascertain, and thus allow correction for, potential sampling biases. |
| Difficulty differentiating between taxa. For example, it may be difficult to identify taxa that created galls and leaf mines because of specimen degradation. | Combined morphological and molecular methods may improve identifications. Focus on species with distinct, well-preserved gall, mine, and pathogen damage morphologies. Direct sequencing of associates if preserved on plant specimens, for example, using DNA barcodes for species identification |
| DNA, RNA degradation | For most species, RNA will be too degraded. However, Malmstrom et al. (2007) have shown it is possible to extract usable RNA in some (perhaps rare) cases. Next-generation sequencing techniques developed for amplifying degraded or ancient DNA. Collection of tissue samples preserved in silica or banked in ultra-low temperature storage. |
| Quantifying damage on individual specimens. | We have demonstrated the potential of subsampling specimens to quantify herbivory damage or pathogen lesions (see *Limitations and Challenges*), analogous to the use of quadrats in vegetation sampling. In cases where visual signs of herbivory are difficult to quantify, defensive compounds might provide an alternative way to detect damage. |

REVIEW

TABLE 2. (Continued)

| Challenges | Potential solutions |
| --- | --- |
| **Pollution data** | |
| Choosing appropriate indicator species for pollutants, regions, and time periods of interest. | Indicator species for many pollutants, such as nitrogen and heavy metals, have been established. Digitized records can then help select those with suitable temporal and geographic coverage. |
| Pesticide degradation | Some pesticides degrade faster than others, and recovering pesticide residues may be possible for chemicals that break down slowly. This area of research has not been explored, and thus methods are not yet developed. |
| Botanists may be unlikely to collect in heavily polluted environments or near areas where pesticides have been applied. | Labels may indicate if specimens were collected near farms, roads, homes, and industry. This metadata can be included in subsequent analyses. |

*Notes:* Here, we highlight the different data types, challenges to their use, and potential solutions to overcome these challenges. We focus here on the data types and uses we have discussed in the text; other applications may be subject to additional challenges (and potential solutions).

Graves 1996) that details methods for subsampling data to incorporate potential biases into expectations, thus holding bias constant when testing for relationships among variables of interest. New machine learning tools and Bayesian techniques allow for the analyses of complex, hierarchically structured, and incomplete data sets and are suited to analyzing large, sparsely sampled data, all common features of collections data. Many of these methods are already well developed for modelling species distributions, e.g., maximum entropy, generalized linear and additive models, boosted regression trees, and random forest (Elith and Leathwick 2009), though they are not yet integrated into ecology more generally (Thessen 2016).

### Data extraction and validation

Once appropriate herbarium specimens are selected, depending on the analysis, it may then be necessary to design sampling strategies to extract data from within individual herbarium sheets. Like any other unit, such as a plot, field, or transect, individual herbarium specimens can be subsampled to quantify features that may be hard to measure across entire specimens, such as data on arthropod and pathogen associations, or other micro-traits including stomatal density, cell shape, etc. In some cases, it may be necessary to measure the surface area of specimens to account for their size, a process that can be automated in programs such as ImageJ, or to estimate specimen density, volume or other dimension to standardize measures for differences in sampling space (ImageJ *available online*).[9]

An important, but frequently overlooked, next step is to evaluate the reliability of derived statistics. To characterize spatial and temporal biases, researchers have turned to data validation, in which herbarium data are placed on common axes with trusted data. For example, several recent studies have validated the utility of herbarium specimens for phenological research by fitting common models to field observation and herbarium data (Robbirt et al. 2011, Spellman and Mulder 2016). By contrasting flowering data from herbaria to that from observations, Davis et al. (2015) showed that herbarium specimens covered more climatic space than observations. Observational data alone might thus provide a more limited estimate of future climate change scenarios than herbarium specimens. This study highlights that

[9] https://imagej.nih.gov/ij/

traditional data sources also have limitations and gaps in coverage, some of which can be alleviated by including data from herbarium specimens. Similarly, pollution measures derived from herbarium specimens can be calibrated by comparing pollutant concentrations to those from other historical data sources, such as deposits in peat bogs and ice cores (Weiss et al. 1999).

Cross validation with an independent data set is perhaps the most robust approach for detecting biases in data. However, in many cases, herbarium data cannot be directly compared to independently derived data because no companion data exist, or the collection of such data requires inordinate effort, e.g., the manual cleaning and standardization of multiple data sets, illustrating the uniqueness of collections data. When companion data are unavailable, we suggest another approach for assessing data reliability: comparing summary statistics relevant to the hypothesis to be tested to theoretical expectations from the literature. We demonstrate this approach focusing on insect herbivory, a process that is likely to shift due to global change, but for which few historical data are available. This is a particularly challenging case study because collectors most likely select specimens that have little damage, so absolute estimates of herbivory are likely biased downward. Nonetheless, paleontologists have used fossil herbivory, for which data are even more sparse, to assess changes in diversity and abundance of herbivory with climatic changes across epochs (Wilf and Labandeira 1999). Therefore, we suggest that the much higher resolution, more abundant data available from within herbaria should provide at least as much information on contemporary herbivory patterns.

We quantified insect damage (herbivory) preserved in herbarium specimens of 20 plant species native to the northeastern United States. For each species, we quantified the amount and diversity of damage by insect herbivores. Herbivory was quantified using a grid laid over each herbarium specimen and randomly sampled for damage. All sampling details are located in Appendix S1. Despite potential biases, as a proof-in-concept example we demonstrate that herbarium specimens harbor diverse types of herbivory damage (Fig. 2). We also provide evidence that they could provide unique longitudinal data on plant–insect interactions (Fig. 3). Herbivory on herbarium specimens follows several patterns that were predictable based on theory and empirical studies on contemporary herbivory. First, the extent of herbivory on herbarium specimens, measured as the proportion of grid cells scored for herbivory that had insect chewing
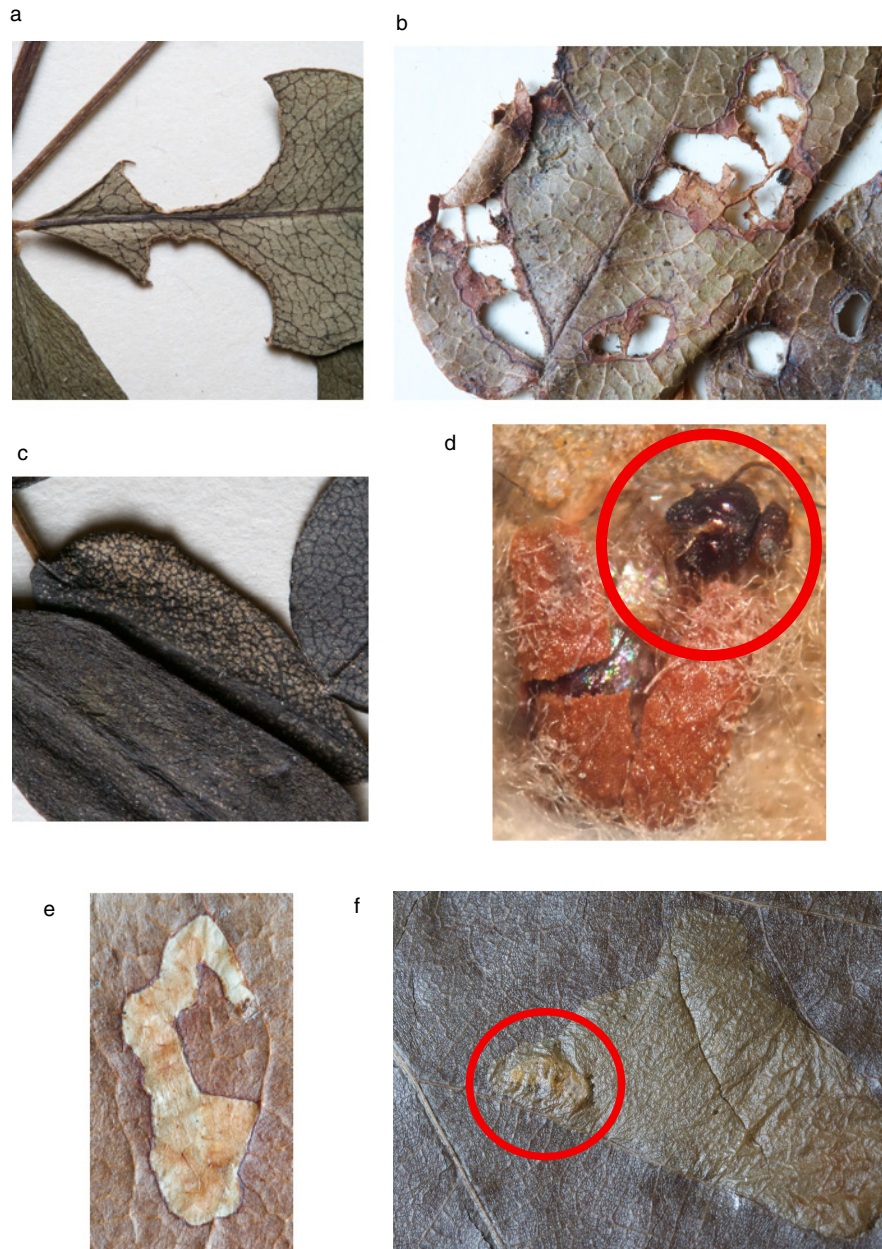
REVIEW

Fig. 2. The diversity of insect herbivory preserved on herbarium specimens. We found several types of herbivory on herbarium specimens made by a diversity of arthropods and quantified herbivory in five categories representing (a) chewing, (b) skeletonization, (c) stippling, (d) leaf galls (here, with emerging gall wasp circled in red), and leaf mines: (e) typical leaf mine and (f) leaf mine with miner inside circled in red. We found other herbivores, such as aphids and caterpillars, pressed with plant specimens, but these were rarer. Chewing damage is typically made by caterpillars and beetles; leaf mines are made by flies, beetles, and mites; stippling is made by leafhoppers and other species that remove cell contents from leaves; leaf galls are typically made by gall wasps.

damage, the most common type of herbivory, has phylogenetic signal, with a value of one matching to Brownian motion), though our data set of 20 species limits formal significance testing. Similarly, the composition of herbivory, i.e., the relative amount of chewing, leaf galls, leaf mines, etc., is significantly more similar for individuals within than among genera and species (Fig. 3b). These taxonomic and temporal patterns indicate that herbarium specimens may provide adequate data for addressing questions related to how herbivore communities are changing over time and how herbivore pressure varies across plant community members,

questions that are of growing interest but for which historical data are limited. In addition, herbivory measured on herbarium specimens is highly skewed, with most specimens showing little or no damage, and a few showing heavy damage (Fig. 3c), as also observed in field data (Turcotte et al. 2014), and accumulates as the growing season progressed (Fig. 3d), which suggests that these data are sensitive enough to capture accumulating herbivory on individual plants through the growing season.

Collector bias toward intact specimens makes it difficult to infer the absolute magnitude of herbivory from
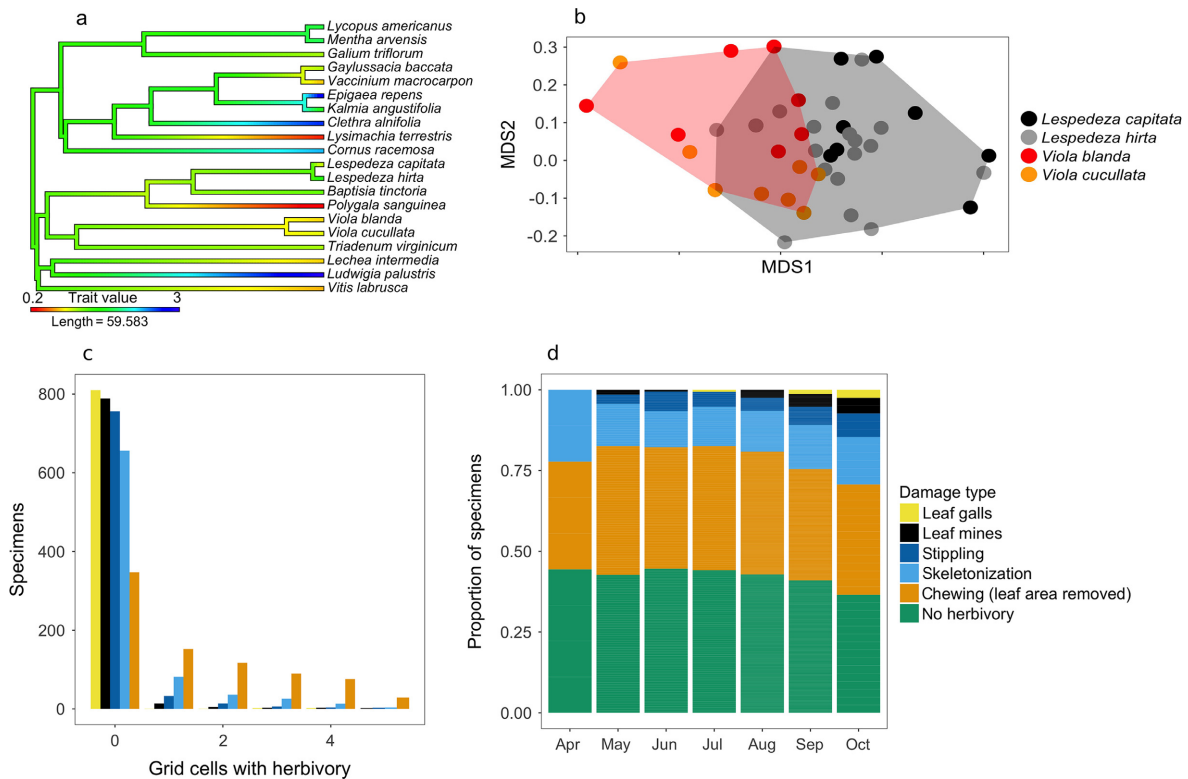
FIG. 3. Herbivory recorded on herbarium specimens of New England. We examined herbarium specimens from 20 species from the northeastern United States for various types of herbivory (see Appendix S1) and contrasted observations to expectations from the literature. We found that (a) species that are closely related have similar herbivory (Blomberg's $K = 0.4$) and (b) composition of damage types on herbarium specimens is more similar within than between species and genera, illustrated here by two genera *Viola* and *Lespedeza* (PERMANOVA, $F_{19,\ 507} = 7.05$, $P < 0.001$). Herbivory within genera clusters more tightly than between genera, as shown by the general separation of polygons. Similarly, herbivory composition is significantly different between species within genera, shown here as partial overlap of points colored by species. (c) Herbivory on herbarium specimens is highly skewed, and (d) herbivory increased as growing seasons progressed. Chewing damage is shown in (a) and (d) because it was the most prevalent type of herbivory (as shown in c). Detailed statistics are available in Appendix S1.

herbarium specimens. However, cross-validation with field data might allow calibration of herbivory estimated from collections and allow researchers to quantify and correct for underestimates. Another limitation of these data is that we cannot infer herbivory rates, i.e., herbivory per day or month, without data on the timing of leaf-out, though degree-day data may serve as a proxy and thus facilitate these estimates and associated inferences.

### Coarse, incomplete, or inaccurate metadata

In some cases, specimens have been misidentified or taxonomic arrangements revised, but herbarium records have not kept pace with changes. New databases and bioinformatics tools now allow rapid taxonomic name synonymization (e.g., the taxize R-package, available online).[10] Judicious choices of exemplar species can also help researchers avoid problematic taxa, as can sampling within herbaria with well-curated metadata. However, in some cases, reviewing each specimen individually is necessary, and herbarium specimens are themselves the key resource for resolving species identities. In contrast to historical observations, for which it is usually impossible to verify species identifications, herbarium specimens can always

be revisited. However, numerous groups likely remain poorly described in herbaria, and despite rapid efforts to mobilize collections online, the resolution of digital images may not be sufficient to provide definitive identification, especially in taxonomic groups with many closely related and morphological similar species, such as in the graminoids.

Another hurdle in using herbarium specimens, especially for ecological applications, is that locations and dates associated with specimens are sometimes absent or incomplete. Label information on specimen locations can be at coarse geographical scales (e.g., at the county level in the United States), or is not provided, especially for specimens collected long ago. This is further complicated by the fact that such specimens may be inaccurately georeferenced, which can propagate through subsequent analyses especially if climate data is linked to such data (Park and Davis 2017). Because we cannot go back in time and collect more fine-scale location data, its absence may prevent researchers from using specimens to address certain hypotheses. However, centralized databases, such as Integrated Digitized Biocollections,[7] now allow researchers to assess the availability and resolution of data across many herbaria, and thus evaluate whether there is sufficient information to address a hypothesis of interest before investing time in what could be wasted research effort.

[10] https://cran.r-project.org/web/packages/taxize/index.html

REVIEW

*Effects of preservation*

Preservation can degrade certain data sources, such as DNA, plant structure, and plant coloration. Technological advances have already helped overcome some of these limitations and may do so to an even greater extent in the future. For example, DNA quality varies considerably among herbarium specimens, but it is already possible to amplify very low concentration and fragmented DNA from specimens allowing population and phylogenetic studies using ancient DNA (Särkinen et al. 2012, Applequist and Campbell 2014). Genomic tools may eventually allow us to link genes to key plant traits, thus providing a means to characterize evolutionary responses to environmental stress, pathogens, or competitors by looking at changing gene frequencies, data that cannot be measured directly from preserved specimens or that have been lost as specimens have become degraded.

Other problems arise because of accidental damage to specimens, notably flooding and insect pests. For example, the global seedbank in Norway was almost flooded in the spring of 2017 due to melting permafrost. In addition, insect pests readily eat herbarium specimens that are not in sealed cabinets, and even herbaria with sealed cabinets can experience infestations, which may not be identified until hundreds of specimens are damaged. While many herbaria use integrated pest management to monitor and treat these occurrences today, historical specimens are often part of personal collections that were exposed to insects. Physical damage of specimens could affect data quality, notably for measurements of herbivory that occurred before collection.

### Next Generation Herbaria

Above, we identified gaps in the literature that point to future research opportunities in herbaria. Here, we detail new frontiers for using data derived from herbaria in global change research. First, herbarium specimens could provide unprecedented amounts of data on fungal responses to global change. While herbaria house millions of macrofungi, pathogenic fungi, and lichens, these records are rarely integrated into global change research. This presents opportunities for investigating macrofungal and lichen responses to global change in ways that parallel ongoing efforts using plants. Though examples are sparse, herbaria studies have shown increasing incidence of fungal plant diseases (Antonovics et al. 2003) and effects of climate on macrofungal phenology (Kauserud et al. 2008, 2010, Diez et al. 2013). Researchers could draw methods from these studies to more deeply investigate fungal responses, which might link to important components of ecosystem function in the future.

Plant and fungal data from herbaria, such as those described in Tables 1 and 2, could be placed on common axes with data from traditional sources—field observations, experiments, and fossils—to generate more robust predictions of how species will respond to drivers of global change (Davis et al. 2015, Youngsteadt et al. 2015). Because all of these methods introduce biases and have limitations, the most robust inferences would include data from multiple approaches, and herbarium data could be uniquely suited in some cases to inform outstanding debates about global change introduced by more traditional approaches. For example, observations and experiments can yield different results about ecologically important phenomena, such as phenological responses to warming (Wolkovich et al. 2012). Herbarium data may allow researchers to resolve such discrepancies, in part, because specimens can capture both long-term processes, including evolution, and short-term processes, such as plastic responses, that are generally not represented together in either observations or experiments. In addition, herbarium data could be combined with data from fossils (as other modern data have been compared to fossils, as described in Labandeira and Currano [2013]) to contrast effects of current and historical climate change by placing temperatures across epochs and recent time on the same axes. While patterns and dynamics might differ due to radically different timescales of these data, such differences might themselves reveal important insights into universal drivers of and responses to global change (e.g., increased $CO_2$ levels and warming).

Leveraging the potential of herbarium data will require advances that allow researchers access to "big data" that span the full range of spatial, temporal, and taxonomic information contained within herbaria and that will require new computational tools to explore. Museum specimen records and images are now more accessible than ever to a large diversity of researchers (Drew et al. 2017) thanks to large scale digitization efforts that have created centralized repositories of these data (e.g., Integrated Digitized Biocollections, Australia's Virtual Herbarium, Museum National d' Histoire Naturelle, Paris), although much data remain dark. Extracting ecologically meaningful data from digitized specimens presents additional challenges, notably in the time and resources required. The rapid growth of citizen science, in which the public aid in data collection efforts, has provided one way forward. Recent collaborations between biologists and computer scientists present new opportunities (Willis et al. 2017*b*) and have allowed for the development of crowdsourcing image annotation tools (e.g., CrowdCurio, *available online*) to extract phenological data from digitized herbarium specimens.[11] These tools are already being adopted and have enormous power for leveraging herbarium data for climate change research, and preliminary studies suggest they can generate highly usable data (Williams et al. *in Review*). Similarly, computer algorithms for analyzing digital images could provide an alternative approach for quantifying traits, such as leaf morphology (Corney et al. 2012*a,b*, Unger et al. 2016, Wilf et al. 2016) and offer the opportunity to rapidly collect data across large numbers of specimens.

Plant and fungal collections continue to grow, but the goals of herbaria frequently remain unaligned with their ecological utility. There is a need to better integrate current research demands into collecting methods and collection management. Therefore, we encourage a shift in how herbaria operate to help maximize their contributions to global change research and to center herbaria as key repositories of ecological data. We propose several key strategies to move toward this goal, the last three of which are potentially the most resource intensive. Of course, we realize there is a need to preserve specimens long-term while serving current

---

[11] https://www.crowdcurio.com/

research needs, and we believe that these recommendations offer the opportunity to better balance these objectives.

*Organize specimens by evolutionary history (phylogeny) and spatial location.*—Collections are most accessible when they are organized by taxonomy and specimen location because researchers almost always sample within clades and areas. Following a more phylogenetically oriented classification and ordering rather than outdated classifications that do not align with current understanding of phylogeny might facilitate ease of usage (Haston et al. 2009). However, we acknowledge that this may not be possible for some herbaria due to lack of resources. In these cases, we suggest that digital specimens include precise metadata on where specimens are located in physical collections.

*Digitize and generate publicly accessible online databases.*— While this is not a novel suggestion, we include it to emphasize the importance of publicly available databases and specimen images for supporting ecological research efforts and to acknowledge that the digital tools necessary for making such data available are still evolving. In many cases, research questions do not initially require physical specimens, and online specimen information facilitates assessment of project feasibility, research planning, and sampling. Online databases and images can also serve as sources of preliminary data that can be used in grant applications and proof-of-concept studies.

*Retain or, at minimum, digitize damaged specimens.*—Curators sometimes dispose of specimens that were damaged by insects or pathogens before they were collected. They do this for good reasons; herbaria have limited space and specimen storage comes at a cost. However, damaged specimens contain valuable information on locations and about species interactions. Suggestions on sterile ecological vouchers follow.

*Collect and preserve tissue for future molecular and chemical analyses.*—Curators at many herbaria now collect and preserve tissue samples, in addition to pressed and dried specimens, explicitly for future molecular analyses. These samples can be dried and stored with silica gel and preserved at room temperature or, more ideally, cryobanked for future extraction. Such infrastructure is also invaluable for preserving RNA, which is essential for investigating gene expression but is often degraded rapidly. These materials are best frozen as soon as possible but field fixatives are now available to reduce the burden of collecting such samples. Recent efforts in tissue banking provide a useful guide on how such approaches could be implemented (see examples *available online*).[12,13]

Besides storing tissues for DNA and RNA analysis, collecting and storing additional leaf and flower material in fragment envelopes attached to herbarium specimens would provide material for destructive sampling (nutrients, isotopes, pollutants), or morphological analyses, such as stomatal density, without damaging mounted specimens.

[12] https://frozenark.org/
[13] http://www.ggbn.org/ggbn_portal/

*Add sampling information to specimen labels.*—Specimens are most useful when labels include metadata, which are recorded regularly as part of basic collecting protocol, i.e., date, geolocation, species. We suggest that the type of sampling and habitat should also be indicated as standard metadata. For example, one system would be to indicate sampling as targeted or opportunistic, and if targeted, to indicate the intention of the collection, e.g., to document galls on the specimen or to get a collection from a particular area. Indication of a specimen's immediate and adjacent habitats, e.g., along a roadside, in an urban or transformed environment, or in a forest, would also help ecologists to determine if specimens were collected in appropriate ways for particular studies. If specimens are collected as part of an ecological or floristic study, referencing that study on the specimen label will help future researchers determine sampling methods.

*Develop protocols for storing ecological vouchers.*—Traditionally, herbaria have focused on maximizing number and morphological diversity within species across their ranges. However, because of increasing interest in historical ecology, there is a need to prioritize specimens that represent ecological effects (for a thorough discussion of this topic, see Baker et al. 2017). Space is a constant issue in herbaria, and ecological vouchers are especially challenging to store, as specimens from a single study can number in the thousands. We suggest that herbaria develop new protocols that allow ecologists to deposit whole digital specimens and small physical samples from plants in their studies. One example might be to reduce these collections in their physical footprint by minimally retaining sufficient tissue for subsequent DNA extraction and limited morphological/anatomical investigation. In concert with rapidly developing DNA barcoding methods, this approach could facilitate the reconstruction of historical patterns of plant community assembly. We suggest that ecologists take on some of the burden of processing and storing ecological vouchers by including storage in herbaria as part of grant budgets. We also suggest that ecologists discuss attribution with individual herbaria so that they are credited as data sources.

*Resample areas and clades for which there are strong historical collections.*—Many herbaria have collections that are temporally biased toward the 19th and 20th centuries, and few specimens are available from the last 30 yr. However, the intensification of global extinction drivers has largely occurred over the past few decades. We suggest that herbaria look to resample areas and clades with strong historical collections, particularly in cases where this objective can be incorporated into planned collection efforts and associated regional/geographic priorities. New collections would allow researchers to make contrasts before and after significant global change. Such efforts are common, and well organized among local amateur naturalist groups.

*Create barcode libraries.*—As molecular sequencing costs have declined, it is now more reasonable to create digital repositories of genetic data using species-specific markers, allowing samples to be matched to collections using both molecules and morphology. The former may be especially

REVIEW

valuable when plant samples lack key identifying traits, such as flowers or leaves. Although controversial, DNA barcoding efforts, such as the Barcode of Life, which now holds over 5 million barcode sequences, illustrate the potential of such approaches (*available online*).[14]

## Conclusions

Data from herbaria are increasingly being incorporated into global change research. Researchers are developing creative new methods to understand how diverse factors affect plants, fungi, and their associates. These methods include using herbarium specimens as occurrence and phenological records, sources of DNA, and information on physiology and morphology. Data derived from herbaria have wide breadth across space, time, and the tree of life. Several topics of broad interest in global change biology and to which herbaria can contribute remain underexplored but show great promise. Exciting applications include the use of herbaria as blueprints for restoration, signatures of physiological change, and records of changing species interactions. Collections of all kinds are threatened by declines in financial support. The future of herbaria will, in part, depend on their ability to adapt to current research demands and funding priorities. Here, we have emphasized applications to global change research, but a broader dialogue is needed to maximize collection utility across other disciplines. We should recognize that the value of such collections may only become apparent in the future. We must therefore maximize the current use of collections while continuing high standards of preservation to benefit future generations.

## Literature Cited

Ainsworth, E. A., and S. P. Long. 2005. What have we learned from 15 years of free-air $CO_2$ enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy. New Phytologist 165:351–371.

Antonovics, J., M. E. Hood, H. Thrall, J. Y. Abrams, and G. M. Duthie. 2003. Herbarium studies on the distribution of anther-smut fungus (*Microbotryum violaceum*) and Silene species (Caryophyllaceae) in the eastern United States. American Journal of Botany 90:1522–1531.

Aono, Y., and K. Kazui. 2008. Phenological data series of cherry tree flowering in Kyoto, Japan, and its application to reconstruction of springtime temperatures since the 9th century. International Journal of Climatology 28:905–914.

Applequist, W. L., and L. M. Campbell. 2014. DNA Banking for the 21st Century. in Proceedings of the US Workshop on DNA Banking, Vol. 194. The William L. Brown Center at the Missouri Botanical Garden, St. Louis, Missouri, USA.

Atha, D. E., T. Forrest, R. F. C. Naczi, M. C. Pace, M. Rubin, J. A. Schuler, and M. Nee. 2016. The historic and extant spontaneous vascular flora of The New York Botanical Garden. Brittonia 68:245–277.

Baker, T. R., et al. 2017. Maximising synergy among tropical plant systematists, ecologists, and evolutionary biologists. Trends in Ecology and Evolution 32:258–267.

Barney, J. N. 2006. North American history of two invasive plant species: phytogeographic distribution, dispersal vectors, and multiple introductions. Biological Invasions 8:703–717.

Bertin, R. I. 2002. Losses of native plant species from Worcester, Massachusetts. Rhodora 104:325–349.

Bolmgren, K., and K. Lonnberg. 2005. Herbarium data reveal an association between fleshy fruit type and earlier flowering time. International Journal of Plant Sciences 166:663–670.

Bonal, D., et al. 2011. Leaf functional response to increasing atmospheric $CO_2$ concentrations over the last century in two northern Amazonian tree species: A historical δ13C and δ18O approach using herbarium samples. Plant, Cell and Environment 34:1332–1344.

Bonan, G. B. 2008. Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. Science 320:1444–1449.

Both, C., M. van Asch, R. G. Bijlsma, A. B. van den Burg, and M. E. Visser. 2009. Climate change and unequal phenological changes across four trophic levels: Constraints or adaptations? Journal of Animal Ecology 78:73–83.

Burns, K. C., N. Herold, and B. Wallace. 2012. Evolutionary size changes in plants of the south-west Pacific. Global Ecology and Biogeography 21:819–828.

Buswell, J. M., A. T. Moles, and S. Hartley. 2011. Is rapid evolution common in introduced plant species? Journal of Ecology 99:214–224.

Calinger, K. M. 2015. A functional group analysis of change in the abundance and distribution of 207 plant species across 115 years in north-central North America. Biodiversity and Conservation 24:2439–2457.

Calinger, K. M., S. Queenborough, and P. S. Curtis. 2013. Herbarium specimens reveal the footprint of climate change on flowering trends across north-central North America. Ecology Letters 16:1037–1044.

Celesti-Grapow, L., G. Capotorti, E. Del Vico, E. Lattanzi, A. Tilia, and C. Blasi. 2013. The vascular flora of Rome. Plant Biosystems 147:1059–1087.

Clark, D. A. 2004. Tropical forests and global warming: Slowing it down or speeding it up? Frontiers in Ecology and the Environment 2:73–80.

Corney, D., J. Y. Clark, H. L. Tang, and P. Wilkin. 2012a. Automatic extraction of leaf characters from herbarium specimens. Taxon 61:231–244.

Corney, D. P., H. L. Tang, J. Y. Clark, Y. Hu, and J. Jin. 2012b. Automating digital leaf measurement: the tooth, the whole tooth, and nothing but the tooth. PLoS ONE 7:e42112.

Crawford, P. H. C., and B. W. Hoagland. 2009. Can herbarium records be used to map alien species invasion and native species expansion over the past 100 years? Journal of Biogeography 36:651–661.

D'Andrea, L., O. Broennimann, G. Kozlowski, A. Guisan, X. Morin, J. Keller-Senften, and F. Felber. 2009. Climate change, anthropogenic disturbance and the northward range expansion of *Lactuca serriola* (Asteraceae). Journal of Biogeography 36:1573–1587.

Daru, B. H., et al. 2018. Widespread sampling biases in herbaria revealed from large-scale digitization. New Phytologist 217:939–955.

[14] http://www.barcodeoflife.org/

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.

DeCandido, R., A. A. Muir, and M. B. Gargiullo. 2004. A first approximation of the historical and extant vascular flora of New York City: Implications for native plant species conservation. Journal of the Torrey Botanical Society 131:243–251.

Diamond, S. E., L. M. Nichols, N. McCoy, C. Hirsch, S. L. Pelini, N. J. Sanders, A. M. Ellison, N. J. Gotelli, and R. R. Dunn. 2012. A physiological trait-based approach to predicting the responses of species to experimental climate warming. Ecology 93: 2305–2312.

Diez, J. M., T. Y. James, M. McMunn, and I. Ibanez. 2013. Predicting species-specific responses of fungi to climatic variation using historical records. Global Change Biology 19:3145–3154.

Dolan, R. W., M. E. Moore, and J. D. Stephens. 2011. Documenting effects of urbanization on flora using herbarium records. Journal of Ecology 99:1055–1062.

Drew, J. A., C. S. Moreau, and M. L. J. Stiassny. 2017. Digitization of museum collections holds the potential to enhance researcher diversity. Nature Ecology and Evolution 1:1789–1790.

Elith, J., and J. R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology Evolution and Systematics 40: 677–697.

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.

Farnsworth, E. J., and D. E. Ogurcak. 2006. Biogeography and decline of rare plants in New England: Historical evidence and contemporary monitoring. Ecological Applications 16:1327–1337.

Feeley, K. J. 2012. Distributional migrations, expansions, and contractions of tropical plant species as revealed in dated herbarium records. Global Change Biology 18:1335–1341.

Feeley, K. J., and M. R. Silman. 2010. Land-use and climate change effects on population size and extinction risk of Andean plants. Global Change Biology 16:3215–3222.

Feeley, K. J., and M. R. Silman. 2011. Keep collecting: accurate species distribution modelling requires more collections than previously thought. Diversity and Distributions 17:1132–1140.

Feeley, K. J., J. Hurtado, S. Saatchi, M. R. Silman, and D. B. Clark. 2013. Compositional shifts in Costa Rican forests due to climate-driven species migrations. Global Change Biology 19:3472–3480.

Felsenstein, J. 1985. Phylogenies and the comparative method. American Naturalist 125:1–15.

Futuyma, D. J., and A. A. Agrawal. 2009. Macroevolution and the biological diversity of plants and herbivores. Proceedings of the National Academy of Sciences USA 106:18054–18061.

Gilman, S. E., M. C. Urban, J. Tewksbury, G. W. Gilchrist, and R. D. Holt. 2010. A framework for community interactions under climate change. Trends in Ecology and Evolution 25:325–331.

Gotelli, N. J., and G. R. Graves. 1996. Null models in ecology. Smithsonian Institution Press, Washington, DC, USA.

Graham, C. H., S. Ferrier, F. Huettman, C. Moritz, and A. T. Peterson. 2004. New developments in museum-based informatics and applications in biodiversity analysis. Trends in Ecology and Evolution 19:497–503.

Gregor, T., D. Bönsel, I. Starke-Ottich, and G. Zizka. 2012. Drivers of floristic change in large cities – A case study of Frankfurt/Main (Germany). Landscape and Urban Planning 104:230–237.

Guerin, G. R., H. Wen, and A. J. Lowe. 2012. Leaf morphology shift linked to climate change. Biology Letters 8:882–886.

Gugerli, F., L. Parducci, and R. J. Petit. 2005. Ancient plant DNA: review and prospects. New Phytologist 166:409–418.

Harris, L. W., and T. J. Davies. 2016. A complete fossil-calibrated phylogeny of seed plant families as a tool for comparative

analyses: testing the 'time for speciation' hypothesis. PLoS ONE 11:e0162907.

Haston, E., J. E. Richardson, P. F. Stevens, M. W. Chase, and D. J. Harris. 2009. The Linear Angiosperm Phylogeny Group (LAPG) III: a linear sequence of the families in APG III. Botanical Journal of the Linnean Society 161:128–131.

Hegland, S. J., A. Nielsen, A. Lazaro, A. L. Bjerknes, and O. Totland. 2009. How does climate warming affect plant-pollinator interactions? Ecology Letters 12:184–195.

Johnson, K. G., et al. 2011. Climate change and biosphere response: unlocking the collections vault. BioScience 61:147–153.

Joly, M., P. Bertrand, R. Y. Gbangou, M. C. White, J. Dubé, and C. Lavoie. 2011. Paving the way for invasive species: Road type and the spread of Common ragweed (*Ambrosia artemisiifolia*). Environmental Management 48:514–522.

Kauserud, H., L. C. Stige, J. O. Vik, R. H. Okland, K. Hoiland, and N. C. Stenseth. 2008. Mushroom fruiting and climate change. Proceedings of the National Academy of Sciences USA 105:3811–3814.

Kauserud, H., E. Heegaard, M. A. Semenov, L. Boddy, R. Halvorsen, L. C. Stige, T. H. Sparks, A. C. Gange, and N. Chr. Stenseth. 2010. Climate change and spring-fruiting fungi. Proceedings of the Royal Society B 277:1169–1177.

Kavanagh, P. H., C. A. Lehnebach, M. J. Shea, and K. C. Burns. 2011. Allometry of sexual size dimorphism in dioecious plants: Do plants obey Rensch's rule? American Naturalist 178:596–601.

Keane, R. M., and M. J. Crawley. 2002. Exotic plant invasions and the enemy release hypothesis. Trends in Ecology and Evolution 17:164–170.

Keeling, C., and T. Whorf. 2005. Atmospheric carbon dioxide record from Mauna Loa. Carbon Dioxide Research Group, Scripps Institution of Oceanography, University of California, La Jolla, California, USA.

Kharouba, H. M., and M. Vellend. 2015. Flowering time of butterfly nectar food plants is more sensitive to temperature than the timing of butterfly adult flight. Journal of Animal Ecology 84:1311–1321.

Kingsolver, J. G., S. E. Diamond, and L. B. Buckley. 2013. Heat stress and the fitness consequences of climate change for terrestrial ectotherms. Functional Ecology 27:1415–1423.

Kozlov, M. V., E. J. van Nieukerken, V. Zverev, and E. L. Zvereva. 2013. Abundance and diversity of birch-feeding leafminers along latitudinal gradients in northern Europe. Ecography 36:1138–1149.

Labandeira, C. C., and E. D. Currano. 2013. The fossil record of plant-insect dynamics. Annual Review of Earth and Planetary Sciences 41:287–311.

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.

Lavoie, C., and D. Lachance. 2006. A new herbarium-based method for reconstructing the phenology of plant species across large areas. American Journal of Botany 93:512–516.

Lavoie, C., Y. Jodoin, and A. G. de Merlis. 2007. How did common ragweed (*Ambrosia artemisiifolia* L.) spread in Quebec? A historical analysis using herbarium records. Journal of Biogeography 34:1751–1761.

Law, W., and J. Salick. 2005. Human-induced dwarfing of Himalayan snow lotus, *Saussurea laniceps* (Asteraceae). Proceedings of the National Academy of Sciences USA 102:10218–10220.

Lees, D. C., H. W. Lack, R. Rougerie, A. Hernandez-Lopez, T. Raus, N. D. Avtzis, S. Augustin, and C. Lopez-Vaamonde. 2011. Tracking origins of invasive herbivores through herbaria and archival DNA: The case of the horse-chestnut leaf miner. Frontiers in Ecology and the Environment 9:322–328.

Leger, E. A. 2013. Annual plants change in size over a century of observations. Global Change Biology 19:2229–2239.

Mack, R. N., D. Simberloff, W. M. Lonsdale, H. Evans, M. Clout, and F. A. Bazzaz. 2000. Biotic invasions: Causes, epidemiology,

REVIEW

global consequences, and control. Ecological Applications 10:689–710.

Malmstrom, C. M., R. Shu, E. W. Linton, L. A. Newton, and M. A. Cook. 2007. Barley yellow dwarf viruses (BYDVs) preserved in herbarium specimens illuminate historical disease ecology of invasive and native grasses. Journal of Ecology 95:1153–1166.

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

McLauchlan, K. K., C. J. Ferguson, I. E. Wilson, T. W. Ocheltree, and J. M. Craine. 2010. Thirteen decades of foliar isotopes indicate declining nitrogen availability in central North American grasslands. New Phytologist 187:1135–1145.

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

Millennium Ecosystem Assessment. 2005. Ecosystems and human well-being: synthesis. Island Press, Washington, D.C., USA.

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-Rushing, A. J., R. B. Primack, P. H. Templer, S. Rathbone, and S. Mukunda. 2009. Long-term relationships among atmospheric $CO_2$, stomata, and intrinsic water use efficiency in individual trees. American Journal of Botany 96:1779–1786.

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

Mitchell, C. E., and A. G. Power. 2003. Release of invasive plants from fungal and viral pathogens. Nature 421:625–627.

Moffett, R. 2014. A biographical dictionary of contributors to the natural history of the Free State and Lesotho. SUN MeDIA Bloemfontein, Bloemfontein, South Africa.

Morrow, P. A., and L. R. Fox. 1989. Estimates of pre-settlement insect damage in Australian and North American forests. Ecology 70:1055–1060.

Neil, K. L., L. Landrum, and J. Wu. 2010. Effects of urbanization on flowering phenology in the metropolitan phoenix region of USA: Findings from herbarium records. Journal of Arid Environments 74:440–444.

Newbold, T. 2010. Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. Progress in Physical Geography 34:3–22.

Norby, R. J., et al. 2016. Model–data synthesis for the next generation of forest free-air $CO_2$ enrichment (FACE) experiments. New Phytologist 209:17–28.

Oerke, E.-C., and H.-W. Dehne. 2004. Safeguarding production—losses in major crops and the role of crop protection. Crop Protection 23:275–285.

Park, D. S., and C. C. Davis. 2017. Implications and alternatives of assigning climate data to geographical centroids. Journal of Biogeography 44:2188–2198.

Park, D. S., and D. Potter. 2013. A test of Darwin's naturalization hypothesis in the thistle tribe shows that close relatives make bad neighbors. Proceedings of the National Academy of Sciences USA 110:17915–17920.

Parolo, G., and G. Rossi. 2008. Upward migration of vascular plants following a climate warming trend in the Alps. Basic and Applied Ecology 9:100–107.

Pauw, A., and J. A. Hawkins. 2011. Reconstruction of historical pollination rates reveals linked declines of pollinators and plants. Oikos 120:344–349.

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

Pelini, S. L., F. P. Bowles, A. M. Ellison, N. J. Gotelli, N. J. Sanders, and R. R. Dunn. 2011. Heating up the forest: open-top chamber warming manipulation of arthropod communities at Harvard and Duke Forests. Methods in Ecology and Evolution 2:534–540.

Plant, R. E. 2012. Spatial data analysis in ecology and agriculture using R. CRC Press, Boca Raton, Florida, USA.

Post, E., C. Pedersen, C. C. Wilmers, and M. C. Forchhammer. 2008. Warming, plant phenology and the spatial dimension of trophic mismatch for large herbivores. Proceedings of the Royal Society B 275:2005–2013.

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.

Pyke, G. H., and P. R. Ehrlich. 2010. Biological collections and ecological/environmental research: a review, some observations and a look to the future. Biological Reviews 85:247–266.

Reef, R., and C. E. Lovelock. 2014. Historical analysis of mangrove leaf traits throughout the 19th and 20th centuries reveals differential responses to increases in atmospheric $CO_2$. Global Ecology and Biogeography 23:1209–1214.

Renberg, I., M. W. Persson, and O. Emteryd. 1994. Preindustrial atmospheric lead contamination detected in Swedish lake sediments. Nature 368:323–326.

Riera, R., C. Sangil, and M. Sanson. 2015. Long-term herbarium data reveal the decline of a temperate-water algae at its southern range. Estuarine Coastal and Shelf Science 165:159–165.

Robbirt, 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.

Romeiras, M. M., R. Figueira, M. C. Duarte, P. Beja, and I. Darbyshire. 2014. Documenting biogeographical patterns of African timber species using herbarium records: A conservation perspective based on native trees from Angola. PLoS ONE 9:e103403.

Sakai, A. K., et al. 2001. The population biology of invasive species. Annual Review of Ecology and Systematics 32:305–332.

Saltonstall, K. 2002. Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. Proceedings of the National Academy of Sciences USA 99:2445–2449.

Särkinen, T., M. Staats, J. E. Richardson, R. S. Cowan, and F. T. Bakker. 2012. How to open the treasure chest? Optimising DNA extraction from herbarium specimens. PLoS ONE 7:e43808.

Schaefer, H., O. J. Hardy, L. Silva, T. G. Barraclough, and V. Savolainen. 2011. Testing Darwin's naturalization hypothesis in the Azores. Ecology Letters 14:389–396.

Sigal, L. L., and T. H. Nash. 1983. Lichen communities on conifers in southern California Mountains: an ecological survey relative to oxidant air pollution. Ecology 64:1343–1354.

Southwood, T. R. E., and P. A. Henderson. 2009. Ecological methods. Methuen and Co. Limited, London, UK.

Spellman, K. V., and C. P. H. Mulder. 2016. Validating herbarium-based phenology models using citizen-science data. BioScience 66:897–906.

Steffen, W., P. J. Crutzen, and J. R. McNeill. 2007. The Anthropocene: Are humans now overwhelming the great forces of nature. Ambio 36:614–621.

Stewart, G. R., M. P. M. Aidar, C. A. Joly, and S. Schmidt. 2002. Impact of point source pollution on nitrogen isotope signatures (delta N-15) of vegetation in SE Brazil. Oecologia 131:468–472.

Syfert, M. M., L. Serbina, D. Burckhardt, S. Knapp, and D. M. Percy. 2017. Emerging new crop pests: ecological modelling and analysis of the South American potato psyllid *Russelliana solanicola* (Hemiptera: Psylloidea) and its wild relatives. PLoS ONE 12: e0167764.

Thessen, A. 2016. Adoption of machine learning techniques in ecology and earth science. One Ecosystem 1:e8621.

Thiers, B. (continuously updated). 2016. Index Herbariorum: A global directory of public herbaria and associated staff. New York

Botanical Garden's Virtual Herbarium. http://sweetgum.nybg.org/science/ih/

Turcotte, M. M., T. J. Davies, C. J. M. Thomsen, and M. T. J. Johnson. 2014. Macroecological and macroevolutionary patterns of leaf herbivory across vascular plants. Proceedings of the Royal Society B 281:20140555.

Unger, J., D. Merhof, and S. Renner. 2016. Computer vision applied to herbarium specimens of German trees: testing the future utility of the millions of herbarium specimen images for automated identification. BMC Evolutionary Biology 16:248.

United Nations, Department of Economic and Social Affairs, Population Division. 2008. World Urbanization Prospects: The 2007 Revision. New York.

Vandepitte, K., T. D. Meyer, K. Helsen, K. V. Acker, I. Roldan-Ruiz, J. Mergeay, and O. Honnay. 2014. Rapid genetic adaptation precedes the spread of an exotic plant species. Molecular Ecology 23:2157–2164.

Vellend, M., C. D. Brown, H. M. Kharouba, J. L. McCune, and I. H. Myers-Smith. 2013. Historical ecology: using unconventional data sources to test for effects of global environmental change. American Journal of Botany 100:1294–1305.

Wandeler, P., P. E. A. Hoeck, and L. F. Keller. 2007. Back to the future: museum specimens in population genetics. Trends in Ecology and Evolution 22:634–642.

Warren, M. S., et al. 2001. Rapid responses of British butterflies to opposing forces of climate and habitat change. Nature 414:65–69.

Wearn, O. R., D. C. Reuman, and R. M. Ewers. 2012. Extinction debt and windows of conservation opportunity in the Brazilian Amazon. Science 337:228–232.

Weiss, D., W. Shotyk, J. D. Kramers, and M. Gloor. 1999. Sphagnum mosses as archives of recent and past atmospheric lead deposition in Switzerland. Atmospheric Environment 33:3751–3763.

Wernberg, T., B. D. Russell, M. S. Thomsen, C. F. D. Gurgel, C. J. A. Bradshaw, E. S. Poloczanska, and S. D. Connell. 2011. Seaweed communities in retreat from ocean warming. Current Biology 21:1828–1832.

Wilf, P., and C. C. Labandeira. 1999. Response of plant-insect associations to Paleocene-Eocene warming. Science 284:2153–2156.

Wilf, P., S. Zhang, S. Chikkerur, S. A. Little, S. L. Wing, and T. Serre. 2016. Computer vision cracks the leaf code. Proceedings of the National Academy of Sciences USA 113:3305–3310.

Williams, A. C., J. Goh, C. G. Willis, J. Goh, A. M. Ellison, J. H. Brusuelas, C. C. Davis, and E. Law. 2017. Deja Vu: characterizing work reliability using task consistency. Proceedings of the AAAI Conference on Human Computation (HCOMP 2017), Quebec City, Canada.

Willis, C. G., B. Ruhfel, R. B. Primack, A. J. Miller-Rushing, and C. C. Davis. 2008. Phylogenetic patterns of species loss in Thoreau's woods are driven by climate change. Proceedings of the National Academy of Sciences USA 105:17029–17033.

Willis, C. G., et al. 2017a. Old plants, new tricks: phenological research using herbarium specimens. Trends in Ecology and Evolution 32:531–546.

Willis, C. G., et al. 2017b. CrowdCurio: an online crowdsourcing platform to facilitate climate change studies using herbarium specimens. New Phytologist 215:479–488.

Wolkovich, E. M., and E. E. Cleland. 2011. The phenology of plant invasions: a community ecology perspective. Frontiers in Ecology and the Environment 9:287–294.

Wolkovich, E. M., et al. 2012. Warming experiments underpredict plant phenological responses to climate change. Nature 485:494–497.

Woodward, F. I. 1987. Stomatal numbers are sensitive to increases in $CO_2$ from pre-industrial levels. Nature 327:617–618.

Wu, Z. T., P. Dijkstra, G. W. Koch, J. Penuelas, and B. A. Hungate. 2011. Responses of terrestrial ecosystems to temperature and precipitation change: a meta-analysis of experimental manipulation. Global Change Biology 17:927–942.

Yaakub, S. M., L. J. McKenzie, P. L. A. Erftemeijer, T. Bouma, and P. A. Todd. 2014. Courage under fire: Seagrass persistence adjacent to a highly urbanised city-state. Marine Pollution Bulletin 83:417–424.

Youngsteadt, E., A. G. Dale, A. J. Terando, R. R. Dunn, and S. D. Frank. 2015. Do cities simulate climate change? A comparison of herbivore response to urban and global warming. Global Change Biology 21:97–105.

Zangerl, A. R., and M. R. Berenbaum. 2005. Increase in toxicity of an invasive weed after reassociation with its coevolved herbivore. Proceedings of the National Academy of Sciences USA 102:15529–15532.

Zarnetske, P. L., D. K. Skelly, and M. C. Urban. 2012. Biotic multipliers of climate change. Science 336:1516–1518.

Ziska, L. H., J. S. Pettis, and J. Edwards. 2016. Rising atmospheric $CO_2$ is reducing the protein concentration of a floral pollen source essential for North American bees. Proceedings of the Royal Society B 283:20160414.

REVIEW

## SUPPORTING INFORMATION

Additional supporting information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/ecm.1307/full

## DATA AVAILABILITY

Data associated with this study are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.903vs08.