New Phytologist

Tansley review

Reversing extinction trends: new uses of (old) herbarium specimens to accelerate conservation action on threatened species

Giulia Albani Rocchetti[1] (iD)**, Chelsey Geralda Armstrong[2], Thomas Abeli[3]** (iD)**, Simone Orsenigo[4]** (iD)**, Caroline Jasper[5], Simon Joly[6,7]** (iD)**, Anne Bruneau[7]** (iD)**, Maria Zytaruk[8] and Jana C. Vamosi[9]** (iD)

[1]Department of Science, University Roma Tre, Viale G. Marconi 446, Roma 00154, Italy; [2]Indigenous Studies, Simon Fraser University, Burnaby, BC V5A 1S6, Canada; [3]Department of Science, University Roma Tre, Viale G. Marconi, Roma 00154, Italy; [4]Department of Earth and Environmental Sciences, University of Pavia, Pavia 27100, Italy; [5]Department of Biological Sciences, University of Calgary, Calgary, AB T2N 1N4, Canada; [6]Montreal Botanical Garden, Montréal, QC H1X 2B2, Canada; [7]Département de Sciences Biologiques et Institut de Recherche en Biologie Végétale, Université de Montréal, Montréal, QC H1X 2B2, Canada; [8]Department of English, University of Calgary, Calgary, AB T2N 1N4, Canada; [9]Department of Biological Sciences, University of Calgary, Calgary, AB T2N 1N4, Canada

Contents

Summary

Although often not collected specifically for the purposes of conservation, herbarium specimens offer sufficient information to reconstruct parameters that are needed to designate a species as 'at-risk' of extinction. While such designations should prompt quick and efficient legal action towards species recovery, such action often lags far behind and is mired in bureaucratic procedure. The increase in online digitization of natural history collections has now led to a surge in the number new studies on the uses of machine learning. These repositories of species occurrences are now equipped with advances that allow for the identification of rare species. The increase in attention devoted to estimating the scope and severity of the threats that lead to the decline of such species will increase our ability to mitigate these threats and reverse the declines, overcoming a current barrier to the recovery of many threatened plant species. Thus far, collected specimens have been used to fill gaps in systematics, range extent, and past genetic diversity. We find that they also offer material with which it is possible to foster species recovery, ecosystem restoration, and de-extinction, and these elements should be used in conjunction with

machine learning and citizen science initiatives to mobilize as large a force as possible to counter current extinction trends.

'...leaves laid up in a book of several plants. kept dry, which preserve colour however, and look very finely, better than any Herball.'

(Samuel Pepys' Diary, 5 November 1665)

I. Introduction

Current estimates that 39% of plant species are at risk of extinction (Antonelli *et al.*, 2020) require rapid responses to reverse these trends and optimize efforts towards species recovery. This estimate stems from the International Union for Conservation of Nature (IUCN), which provides what is widely recognized as the most comprehensive listing of species at risk of global extinction in its IUCN Red List of Threatened Species (IUCN, 2017). The process of correctly assessing a taxon's extinction risk requires data on range extent, population size, and population trends over time (IUCN, 2019) and plant natural history collections (NHCs) can be very useful sources of such information (Carrington *et al.*, 2017; Nic Lughadha *et al.*, 2018). Plant NHCs can include vascular plants and cryptogams (i.e. mosses, liverworts and hornworts) in herbaria and associated collections (fruits, seeds, wood, silica gel dried and frozen leaves, etc.). Recent articles have detailed the many uses of herbaria (Besnard *et al.*, 2018; Meineke *et al.*, 2018; Lang *et al.*, 2019) and how they are especially important for clarifying plant taxonomy and documenting changes in the range extent of rare or important species and their ecosystems through time (Graham *et al.*, 2004; Lavoie *et al.*, 2016; Nualart *et al.*, 2017). For instance, pairing the spatial and temporal features of data from herbaria with remote sensing imagery has greatly accelerated our ability to inform current and future conservation initiatives and identify putatively endangered plant species (Nic Lughadha *et al.*, 2018). While these studies have been extremely valuable for accelerating the rate at which we progress with Red Listing species, more work remains to be done. In order to reverse the alarming statistics on extinction, the academic and conservation communities will need to work together to bring about species recovery action. Much of the work of implementing legal protection for a species to reverse declines occurs within jurisdictions (e.g. nations, provinces, states, etc.) that operate within their respective communities. It is this important step, which requires the estimation of the magnitude of imminent and continuing threats acting upon a species (McCune *et al.*, 2013; Orsenigo *et al.*, 2018; Le Breton *et al.*, 2019), that is often incomplete and hinders species recovery (Akcakaya *et al.*, 2018). Although much of what follows could be applied to NHCs in general, we focus here on herbaria.

In order to reverse recently documented patterns of extinction, it has been argued that increased conservation action will be required to remove the threats to at-risk species (Williams *et al.*, 2020), and thus we devote this review to a discussion of how herbarium specimens are beginning to be used (and can be used further) for the specific purpose of contributing to

conservation action for plant species at a jurisdictional level (i.e. the governing body that would legislate and manage species protection). Conservation practitioners require isolation of the specific problems in need of solutions (e.g. the identification of conservation units) as well as transparent communication of the information that has led to decisions about which species are at risk of extinction (i.e. risk assessment or Red Listing (Roberts & McInerny, 2003; Loiselle *et al.*, 2008; Nic Lughadha *et al.*, 2018)), and we provide a general overview of how herbaria and advances in machine learning will accelerate these steps. We then further concentrate the review specifically on how herbarium specimens are currently being used to better understand the source and severity of human-influenced threats.

While we fully acknowledge that some of these questions can be addressed by combining geographical tools with occurrence data (e.g. species distribution modelling (SDM) approaches), some of the value of the herbarium is inherent to the physical specimens themselves. For example, chemical information linked to specimens sheds light on past and present threats, such as the accumulation of soil/water pollutants (Rodríguez Martín *et al.*, 2015; Rudin *et al.*, 2017a). A physical specimen is unique in that it holds morphological, anatomical, genetic and chemical features that are tied to spatial and temporal information recorded on the label (Meineke *et al.*, 2019; Bakker *et al.*, 2020) – and sometimes viable propagules (i.e. seeds and spores; Molnár *et al.*, 2015). We therefore explore the possibilities related to the use of herbaria as a source of genetic material for driving practical *in situ* and *ex situ* conservation actions (e.g. population genetics-driven seed banking (Nakahama *et al.*, 2015)), as a source of propagules for effective recovery of genetic variation lost from the wild (Godefroid *et al.*, 2011), and emerging fields of research dealing with de-extinction of extinct species from propagules contained in herbarium specimens (Abeli *et al.*, 2020). An exciting avenue of research that is emerging now is the use of machine learning to develop a number of facets of ecological research, and we examine its relevance to conservation and species recovery. Recent estimates indicate that there are > 390 million specimens in *c.* 3100 herbaria (Thiers, 2020). Coupled with technological resources such as spatial modeling, artificial intelligence (AI) and big data analysis, which maximize the information that can be gathered from such specimens, we have an enormous (unexpressed) potential for herbaria to contribute to plant conservation biology. This extent to which this potential can be exploited has increased further thanks to advances in digitization quality, even with only 21% of the world's specimens (representing 38% and 26% of vascular plant and fungi biodiversity, respectively) being available in a digital form through large repositories such as the Global Biodiversity Information Facility (GBIF; Paton *et al.*, 2020).

Categorizing a species as at risk of extinction does not have to be a death sentence for that species, yet the steps to recover a species that is in decline lag far behind the listing process in

many jurisdictions. If we take action, we can maintain that species within its respective ecosystem, and the probability of this process occurring will be greatly increased by engaging the interest of the public regarding at-risk species in their area. Digitization efforts are critical for species recovery in this sense, especially for providing the often forgotten historical context of which species previously comprised the surrounding ecosystems (Delisle *et al.*, 2003). Raising awareness of the loss of plant biodiversity among the general public through engagement and involvement has been aided by the advent of new user-friendly interfaces. Inevitably, documenting species on the brink of extinction will only get us so far. To improve conditions for future generations, we need to chart a path for the reversal of declines when they are observed, and we will need increased participation in order to achieve this goal. Here, we review recent, rapid advances in the use of information embedded in herbarium specimens and their relevance to the process of listing plant species at risk of extinction and identifying the reasons behind their endangerment. We further make the argument that the same exciting tools used for these purposes can also be directed towards planning and executing species recovery.

II. Conservation action – recognizing what to protect

The recovery of declining species requires active management, which in turn requires the provision of methods of recognizing species, delineating areas for protection, and parsing out the factors that contribute to the decline in a given unit of biodiversity. It is worth stating at the outset that effective conservation action can only be achieved when wildlife managers and other actors can focus efforts on a well-defined unit for conservation. These are variably referred to as a Conservation Unit (CU), an Evolutionary Significant Unit (ESU), or a Designatable Unit or DU (de Magalhães *et al.*, 2017). A DU (or a CU or an ESU) is often a species, but it can be a subspecies or variety (Fraser & Bernatchez, 2001; Mee *et al.*, 2015). These units are often identified and delimited by the examination of the material housed in herbaria (Shepherd *et al.*, 2015). In some jurisdictions (for example in Canada), DU can also designate populations that have a clear genetic distinctiveness, something herbaria can contribute to defining. Accurate species or DU delimitation, especially the examination of voucher specimens that allow for verification of species records, is critical for conservation planning and decision-making (Rojas, 1992; Agapow *et al.*, 2004) (Fig. 1).
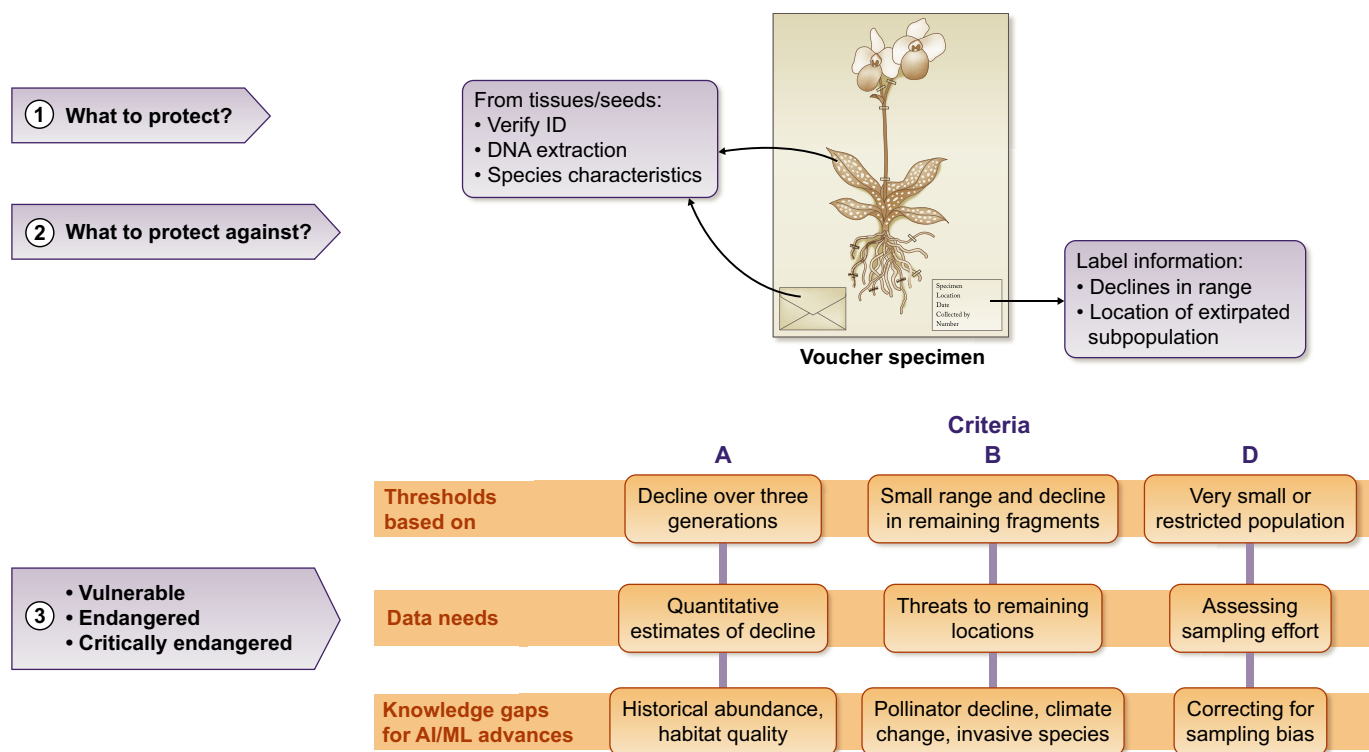

Fig. 1 Herbarium specimen components and their uses. Herbarium specimens historically provided information from the physical material of the plant itself. (1) The voucher specimen provides a verifiable report of a taxon's presence and is frequently reexamined to confirm the accuracy of a species, as well as to gather further information regarding the characteristics of the species and/or habitat (e.g. evidence of herbivory). Once there is convincing evidence of what the conservation community should aim to protect, there is further information that can be gleaned from the specimen regarding threats (2). For example, a variety of aspects of the health of the specimen can be traced back to its geographical coordinates on the label (see Table 1), which provide evidence of which locations have experienced declines or extirpations that pertain to whether the species meets IUCN criteria (3). Here, we have provided the IUCN criteria in a very simplified format to indicate the common avenues by which a plant species meets the criteria for designations of Critically Endangered, Endangered or Vulnerable (see Box 1 for Glossary), as well as how AI and machine learning (ML) advances might fill current knowledge gaps and increase the relevance of herbaria in future. Note that Criterion C is removed because it deals with the number of mature individuals, which can rarely be estimated with herbarium specimens. Further, vegetative tissue and seeds (often placed in an envelope) of a specimen can provide DNA which can be used to examine and potentially mitigate genetic diversity declines.

Clearly, the lumping or splitting of taxa can greatly affect the interpretation of species ranges and the number of individuals within taxa, which can have important and direct implications on the conservation status of taxa (Frankham *et al.*, 2012). As such, herbarium specimens continue to be used for the purpose of achieving our best hypothesis of the most appropriate unit for conservation. The study of the morphology of herbarium specimens to delimit species or subspecies has been ongoing for hundreds of years (Funk, 2018) and is set for a resurgence with machine learning (see Section VIII, below). The increase in the ease with which sequences from multiple independently evolving genetic markers can be obtained and the recent application of population genetic methods for understanding gene flow demonstrate some new avenues for the use of genetic data obtained from herbarium specimens of rare species (see Section V, below). The genetic data can be used alone or in combination with morphology, or complemented with material collected from the field (e.g. Bruneau *et al.*, 2011; Reeves & Richards, 2011; Lambert *et al.*, 2017; Bieker & Martin, 2018).

III. Herbaria and Red Listing

Further characterizing the DUs in need of protection (in terms of biology, habitat requirements, range size, and threats) is the next step in the conservation process and has become known as 'Red Listing'. To be categorized as 'at risk' (Vulnerable, Endangered or Critically Endangered, Fig. 1, Box 1), requires evidence to be gathered on where a species meets or exceeds a series of criteria. The details of the five criteria are well described in the IUCN procedures (IUCN, 2019), but, in brief, Criterion A focuses on evidence regarding rates of decline (which can be measured through abundance metrics or range contractions); Criterion B focuses on evidence regarding small range size, in addition to evidence of fragmentation, population fluctuations, rare and patchy distribution, and/or continuing decline; Criterion C draws on evidence of decline in the number of mature individuals; Criterion D relies on evidence that the species is susceptible to loss after stochastic events due to extremely low abundance or distribution (Rivers *et al.*, 2011); and Criterion E relies on quantitative modelling projections that the probability of extinction is above a certain threshold. Criteria C and E rely on population monitoring data that is often insufficiently documented in herbarium specimens (but see Case *et al.*, 2007; Kricsfalusy & Trevisan, 2014), and so we concentrate our attention on the potential for herbaria to contribute to Red Listing through gaining evidence for Criteria A, B and/or D (Fig. 1).

Many life-history characteristics have been shown to predict extinction risk through their roles in influencing susceptibility to anthropogenic threats, such as growth rates, phenology, habitat specialization and mating system (Fréville *et al.*, 2007). Red-listing procedures require that declines are adjusted such that they are estimated over a temporal window that is 10 yr or three generations, whichever is longer (up to a maximum of 100 yr; IUCN, 2019) and it is worth noting that for many Red-Listed plant species we have no estimate of the generation length (see Box 1). The default procedure is to then measure declines over 10 yr, which has been seen to

---

Box 1   Summary of key terms related to IUCN.

Conservation translocation: The intentional movement and release of a living organism where the primary objective is a conservation benefit (IUCN, 2013).
Critically endangered: A taxon considered to be facing an extremely high risk of extinction in the wild.
De-extinction: According to IUCN (2016), the creation of some proxy of an extinct species or subspecies. Here, the term is used in a more general sense, that is, to resurrect an extinct species.
Endangered: A taxon considered to be facing a very high risk of extinction in the wild.
Generation length: Commonly estimated as the average age of parents of the current cohort.
Haplotype: A group of alleles in an organism that are inherited together from a single parent; allows for inference of gene flow between separated sites
**Vulnerable:** A taxon considered to be facing a high risk of extinction in the wild.

---

underestimate the level of risk in long-lived species (Bird *et al.*, 2020 in press). Herbaria may be useful in this respect because specimens can provide an improved estimate of generation length before disturbance, without further disturbing the few remaining individuals of at-risk species, which can be accomplished, for example, by counting annual rhizome scars on specimens (McGraw, 2001) or by extracting information on tree height from the label (Greve *et al.*, 2016). As machine learning applications increase, these initial (and often sparse) estimates could be used to train machine learning models to develop more sophisticated estimates of age (and other life-history traits that influence extinction risk) that would allow for reconstruction of the trait before anthropogenetic disturbance (see Section VIII, below).

The historical perspective that herbaria can provide is essential for distinguishing between naturally rare species and species that are experiencing declines due to human activity, a step that is critical for placing a taxon on the IUCN Red List. As stated above, the process of correctly assigning risk to a taxon requires data on range extent, population size, and population trends over time (IUCN, 2019) (Fig. 1) and herbaria are common sources of such information (Willis *et al.*, 2003; Lister, 2011; Carrington *et al.*, 2017; Nic Lughadha *et al.*, 2018). The application of full IUCN criteria to all plant species represents an enormous task, and the majority of plant species have not been assessed (Bachman *et al.*, 2018). The current number of species that have been assessed using the IUCN Red List Guidelines represents *c.* 10% of the estimated 380 000 plant species (Nic Lughadha *et al.*, 2020). Efforts to automate the analysis of the large number of digitized records now available on the Global Biodiversity Information Facility (GBIF) website (https://www.gbif.org/), and thus accelerate the Red Listing of species, have been treated in other reviews (Nic Lughadha *et al.*, 2018).

Generally, these automated procedures are valuable for addressing range extent and can thus identify species that meet the criteria that rely heavily on occurrence data (i.e. Criterion B; Fig. 1). Occurrence data reported for herbarium specimens can be used to

Fig. 2 Herbarium specimens through time in conservation assessment. In aggregating specimens from a number of herbaria over time, information can be obtained on past and projected species declines of a given species. After comparison of each specimen allows for properly vouchered species identification, specimens can be divided into those coming from extant subpopulations (representing (1) the current distribution (in blue)) or extinct subpopulations according to the label information on the specimens, allowing for reconstruction of areas in the (2) historical distribution that have been lost (in gold, see Tables 1, 2). Examining what has occurred at the sites of extinct subpopulations can provide information on the scope and severity of anthropogenic threats. Species distribution modelling (SDM) can also use the present-day occurrences to identify areas of suitable habitat (in terms of predictor variables such as temperature and precipitation) as well as project the subset of extant sites (if any) that will remain or become suitable for a given taxon in the future ((3) projected future distribution, in red). This latter procedure can estimate the scope of threat of climate change within three generations.
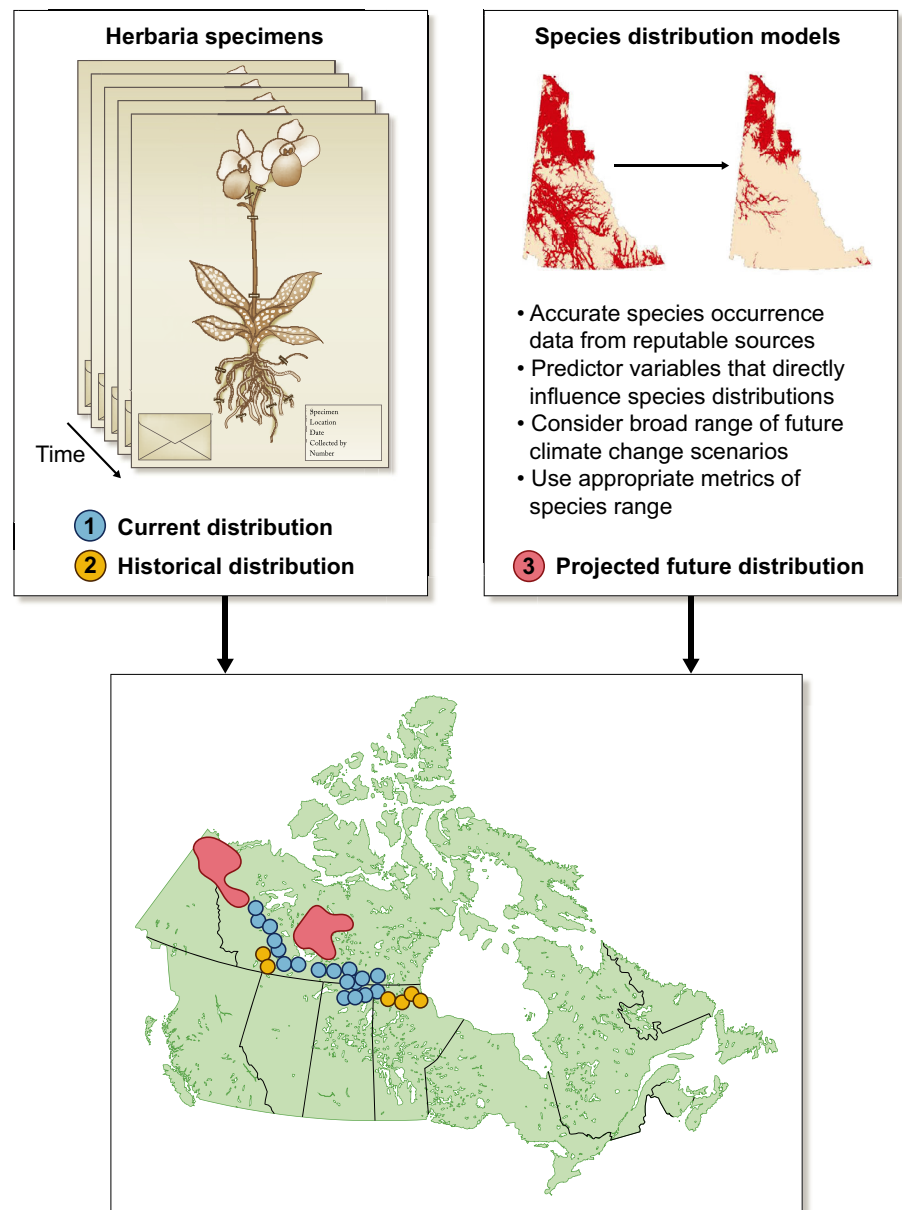
define a species range and assess past changes in range extent (Criterion A; Figs 1 and 2). However, this exercise is not exempt from biases (Tingley & Beissinger, 2009) because pseudo-absences (lack of observations that are inferred to be true absences) can exaggerate range contractions. More sophisticated algorithms based on the time since last collection have also been developed (Ungricht *et al.*, 2005; Farnsworth & Ogurcak, 2006; Colla *et al.*, 2012) and appear to show promise. In addition, occurrence records can be validated and additional data can be obtained from species checklists or targeted field surveys (Wolf *et al.*, 2016; Lewthwaite *et al.*, 2018). Analyses of past declines in range size using herbarium specimen data have been rare, with a few studies that have adopted this approach in the Lesser Antilles (Carrington *et al.*, 2017) and Hawaii (Krupnick *et al.*, 2009). Considering that the biodiversity research community has been working to digitize the vast stores of specimens (> 1 billion in the USA alone) by 2020 (Ellwood *et al.*,

2015), the lack of incorporation of Criterion A may be surprising. However, the procedure requires the collection of data from a comprehensive number of herbaria within the jurisdiction, which may still be prohibitive for larger countries or regions. Additionally, such analyses are subject to criticism that geographical and temporal biases exist in these datasets (Prather *et al.*, 2004; Daru *et al.*, 2018; James *et al.*, 2018) – most importantly that collections are concentrated near roads and the herbaria housing the specimens (Daru *et al.*, 2018), and that collecting may have declined recently in some regions and increased in others (Goodwin *et al.*, 2015). While there have been some useful advances in tackling these sampling biases by using range extent changes relative to associated species (e.g. Delisle *et al.*, 2003), these biases may hinder the confidence that jurisdictions have in applying Criterion A. For these reasons, herbaria may be more useful when examining the evidence that anthropogenic activities impact a species in a more

Table 1 Studies using herbarium-derived data contribute to the estimation of the threats responsible for the continuing decline of at-risk species. Threats are defined according to the IUCN Threats Classification Scheme (v.3.2).

Threat	Demonstration of proof-of-concept for using herbarium specimens in threat assessment	Reference(s)	Specimen data used for threat assessment
(1) Residential & commercial development	Assigned scope and severity of threat level based on historical proximity of ecosystem types to developed urban areas to assist with planning of protected areas. Found species near Red List thresholds due to past development but with inadequate data on continuing decline to highlight species needing further research.	Rebelo *et al*. (2011); Le Breton *et al*. (2019)	Occurrence
(2) Agriculture & aquaculture	Estimated the scope of eutrophication by examining diatom assemblages found on herbarium specimens.	van Dam & Mertens (1993)	Occurrence, assemblage of associated species.
(3) Energy production & mining	Estimated scope of the effects of heavy metals.	Reeves & Brooks (1983); Rudin *et al*. (2017b)	Heavy metal content.
(4) Transport & service corridors	Estimated extent to which transport corridors facilitate the spread of invasive plants by examining the rate of spread as a function of distance from corridors.	Lavoie & Delisle (2005)	Occurrence
(5) Biological resource use (gathering terrestrial plants)	Estimated the severity of decline in size of ginseng from preferential harvesting of larger plants as well as severity of declines in abundance.	McGraw (2001); Case *et al*. (2007)	Plant traits; occurrence.
(6) Human intrusions & disturbance (recreational activities)	Estimated severity (via C-values) of impact on species of urbanization activities such as gardening, landscaping.	Dolan *et al*. (2011)	Occurrence
(7) Natural system modifications (dams & water management/use; other ecosystem modifications)	Provide estimate of scope of historical impact of threat of water management.[1]	Abeli *et al*. (2012); Hutchings *et al*. (2018)	Occurrence; phenological traits.
(8) Invasive & other problematic species & genes	Predictions of changes to scope of (1) overlap with invasive species; and (2) introgression with other species; observations of severity of overabundance of native species/diseases.	Saltonstall (2002); Antonovics *et al*. (2003); Crawford & Hoagland (2009); Beauvais *et al*. (2017); Spriggs *et al*. (2019)	Occurrence; disease presence; collection date; plant traits; DNA sequences
(9) Pollution (agricultural & forestry effluents; airborne pollutants)	Scope of nitrogen pollution at endangered species sites; severity of air pollution estimated through declines/extirpations observed in several moss species.	Hallingbäck (1992); Abeli *et al*. (2014)	Occurrence
(10) Geological events (e.g. landslides)	None found.		
(11) Climate change & severe weather	Scope of unsuitable sites forecasted due to habitat shifting, temperature extremes and drought using (1) SDM approaches for numerous rare and endangered plant species or (2) severity forecasted by assuming a direct relationship between survival and suitability scores from SDM.[2]	Feeley (2012); Nualart *et al*. (2017); Tang *et al*. (2017); Kosanic *et al*. (2018); Yan & Tang (2019)	Occurrence

When the target is conservation action and informing jurisdictions of threats to an endangered species, past threats are considered of less importance compared to current or projected threats (if they happened in the past, there is little action that can be taken). Threats are scored as important (and the status of the species is more likely to be categorized as threatened or endangered) if they have high scope (what proportion of the current distribution is affected) and/or severity (the likelihood of loss of individuals within the affected scope). The information needed from herbarium specimens comes from the labels (i.e. occurrence data) and is also derived from the plant material itself (described in the rightmost column).

[1]Estimates of the scope and severity of the effect of loss of pollinators would be considered in this category and appear to be possible (Johnson *et al*., 2019). However, to our knowledge, there are very few examples where these methods have been applied to at-risk species (with the exception of Hutchings *et al*., 2018). Declines in reproduction are often noted in threat assessments but are not considered in depth due to their focus on the survivorship of mature individuals, representing a gap in current threat assessments.

[2]Many studies on common species have used this approach, but we have only summarized those few studies applying these methods to at-risk species.

direct fashion and when continuing declines are addressed, as is most commonly the case when assessing the applicability of Criterion B (Table 1). We note, however, that when species distribution modelling is used to predict declines in terms of Criterion B, these predictions may also be subject to biases (see Fig. 2).

The advances made in the use of AI and machine learning technologies to accelerate Red Listing cannot be overstated (see Section VIII and Table 3, below). Proper identification and enumeration of the elements that are at risk of extinction is a critical first step in conservation. The next steps involve discovery of the threats that are causing these species (or DUs) to be threatened

Table 2 Use of herbaria to assess the extent and provenance of genetic diversity declines.

Species	Scope	Genetic diversity parameter estimated	Citation
Ulmus rubra	USA	Estimate of past genetic bottlenecks	Brunet *et al.* (2016)
Anacamptis palustris	Italy	Distribution of haplotypes	Cozzolino *et al.* (2007)
Eligmocarpus cupuron	Madagascar	Distribution of haplotypes	Devey *et al.* (2013)
Dimorphandra exaltata	Brazil	Fragmentation; declines due to climate change	Muniz *et al.* (2019)
Vincetoxicum pycnostelma	Japan	Allelic richness; found unique alleles in seedlings grown from historical herbarium specimens	Nakahama *et al.* (2015)

and then to mitigate those threats and devise strategies for recovery of the species to a stable state. Herbaria, and to some extent the same machine learning technologies, have been employed for these subsequent steps, but to a lesser degree (Table 1).

IV. Historical baselines provided by herbaria and their use in threat assessment

With the frequent need for practitioners to analyze declines over the past century or longer, the relatively long history of herbaria is considered fortunate and is often the only window we have for the formulation of a historical baseline from which we can examine which species have been disappearing over time (Lang *et al.*, 2019). Early herbaria concentrating on the flora of natural environments are thought to have first been developed in the 1600s, but their primary function was knowledge and education rather than conservation (Soltis, 2017). The disappearance of certain plant species can be used to gather estimates of extinction rates in the recent past, and this can represent an important step for effective conservation if the calculation of a 'threat index' is used to determine the level of threat to extant rare species in the area (e.g. the Random Forest technique (a machine learning approach) in Nic Lughadha *et al.*, 2018).

Occurrence data can also be used to determine the environmental tolerances of a plant species (i.e. SDM or environmental niche modelling), which is of central importance for understanding the impact of climate change and land use change (see Soltis, 2017; Attorre *et al.*, 2018; Meineke *et al.*, 2018). Previous reviews have emphasized how temporal data are used to assess how climate change can cause range and/or phenological shifts (reviewed by Lavoie, 2013; Willis *et al.*, 2017a; Meineke *et al.*, 2018; Lang *et al.*, 2019) and possibly lead to mismatches in coevolved pollinators or mate limitation (Kharouba & Vellend, 2015; Bontrager & Angert, 2016; Hutchings *et al.*, 2018; Johnson *et al.*, 2019). While many studies have focused on phenology changes (Pearson *et al.*, 2020) and range shifts (Tingley & Beissinger, 2009), we know surprisingly little about whether these changes manifest as declines in survivorship and/or recruitment within the locations occupied by threatened plant species (Castro *et al.*, 2015). The tools to determine this are now available, yet the academic community has often shied away from translating their results into a form that conservation practitioners can use to protect and recover species. We argue that the time is right for better integration (see Table 2).

Table 3 References resulting from the literature search on big data and machine learning applied to herbaria, divided by main topics.

Topics	References
(1) Databases, data aggregators, and data management protocols	Wang *et al.* (2009); Matsunaga *et al.* (2013); Dugenie *et al.* (2017); Yagui *et al.* (2017); Dormontt *et al.* (2018); Heberling & Isaac (2018); Henning *et al.* (2018); Yost *et al.* (2018); Powell *et al.* (2019); Theeten *et al.* (2019)
(2) Evaluating plant extinction risks/conservation status	Tressou & Haevermans (2018); Nic Lughadha *et al.* (2018, 2019); Bachman *et al.* (2020); Hochkirch *et al.* (2020); Zizka *et al.* (2020)
(3) Gaps and biases in herbarium records and databases	Daru *et al.* (2018); Cornwell *et al.* (2019)
(4) Herbarium metadata extraction	Heidorn & Wei (2008); Steinke (2012); Silva (2016); Dagtekin *et al.* (2018); Kirchhoff *et al.* (2018); Meineke *et al.* (2020a); Triki *et al.* (2020); Walton *et al.* (2020)
(5) Overviews of big data and machine learning in plant science	Franklin *et al.* (2017); James *et al.* (2018); Soltis *et al.* (2018, 2020); Rønsted *et al.* (2020)
(6) Phenological features analyses	Corney *et al.* (2012); Henries & Tashakkori (2012); Václavík *et al.* (2017); Willis *et al.* (2017a); Reeb *et al.* (2018); Schneider *et al.* (2018a); Lorieul *et al.* (2019); McAllister *et al.* (2019); Borges *et al.* (2020); Davis *et al.* (2020b); Goëau *et al.* (2020); Kommineni *et al.* (2020); Ott *et al.* (2020); Pearson *et al.* (2020); Weaver *et al.* (2020); Yost *et al.* (2020); Younis *et al.* (2020)
(7) Species distribution analysis, modelling, protocols, and regional biodiversity prediction	Elith *et al.* (2006); Lorena *et al.* (2011); Biganzoli *et al.* (2013); Amici *et al.* (2014); Meyer *et al.* (2016); Glon *et al.* (2017); Huettmann & Ickert-Bond (2018); Park *et al.* (2020); Watkins *et al.* (2020); Zurell *et al.* (2020)
(8) Species identification and classification	Nordin *et al.* (2011); Clark *et al.* (2012); Wijesingha & Marikar (2012); Popescu *et al.* (2015); Grimm *et al.* (2016); Mata-Montero & Carranza-Rojas (2016); Unger *et al.* (2016); Carranza-Rojas *et al.* (2017, 2018); Kho *et al.* (2017); Schuettpelz *et al.* (2017); Wäldchen & Mäder (2018); Younis *et al.* (2018); Kamilov *et al.* (2019); Hussein *et al.* (2020); Joly *et al.* (2020); Little *et al.* (2020); Pryer *et al.* (2020)
(9) New technologies applied to herbarium analysis and management	Wu & Dietterich (2004); Delgado *et al.* (2005); Fan *et al.* (2012); Bhamra *et al.* (2014); Dillen *et al.* (2019); Villacis-Llobet *et al.* (2019); Davis *et al.* (2020a); Ledesma *et al.* (2020); White *et al.* (2020)

Information recorded in herbaria that is inherently linked to past occurrence sites of species has been used for identifying threats that have brought a species to extinction, local extirpation or endangerment in a number of different ways (Table 1). For instance, ecological conditions at current sites of occurrence for a given species may be compared to ecological conditions of sites at which the species was present in the past (from herbarium records) and from which it is now extirpated (current/historical growing site comparison approach (Lienert *et al.*, 2002; Abeli *et al.*, 2014)). Some of the ways in which researchers have used herbarium specimens to assess threats have been indirect but very clever – Van Dam & Mertens (1993) investigated changes in water characteristics in the Netherlands by examining the assemblages of diatoms attached to old and recent herbarium sheets of aquatic plants. These approaches have important implications for the current conservation management of species (e.g. threat identification, removal and mitigation, selection of suitable sites for reintroduction, habitat changes) and are not limited to examination of threats to plant species (e.g. one study on historical macroalgae collections was able to infer the effects of upwelling on sardines (Miller *et al.*, 2020)). In addition, increased herbivory and the presence of invasive pest species can be detected by looking at leaf damage on specimens (Fig. 1). Despite the tendency of botanists to collect attractive specimens (and those potentially free of pests; Kozlov *et al.*, 2020), recent evidence indicates that these collecting biases are not strong enough to offset the generation of meaningful conclusions (Meineke *et al.*, 2020b).

We summarize the various ways in which conservation biologists have generated estimates regarding the causes of population declines (known as a threat assessment) in Table 1. This threat assessment exercise is commonly conducted through the evaluation of 11 threats, according to the IUCN Threats Classification Scheme (v.3.2; see Table 1). When the objective is conservation action and informing jurisdictions of threats to an endangered species, past threats are considered less important than current or projected threats, yet herbarium specimens still provide critical baseline information for the quantification of ongoing threats. The impact of each threat depends on the scope (what proportion of the current distribution is affected) and severity (the likelihood of loss of individuals within the affected scope) of continuing threats. We find that herbaria have been employed in varied ways to assess the scope and severity of most of the measured threats, with their use in determining the threat from invasive species and climate change being the most common (see Table 1). Herbarium specimens are more often used to gain insight into the geographical scope of a threat, rather than its severity, as it appears that gaining an understanding of the severity of a threat may more often require examination of living individuals (but see Johnson *et al.*, 2019).

In estimating the scope of the threat of climate change (and to a lesser extent invasive species) species distribution modelling (SDM) approaches are often used to project the extent of the threat into the future and thus gain an estimate of the scope of an endangered species predicted to be subject to inhospitable conditions. Briefly, some SDM approaches (such as MAXENT) represent a type of machine learning (see Section VIII, below), where an algorithm is used to determine the influence of environmental variables (e.g. temperature, precipitation, soil types, etc.) on the presence of a species and then overlay these models with projections of climate change to detect where the species will find suitable conditions in the future. Modelling future responses to climate change requires baseline data of a species' occurrences (Table 1), and these data often can be obtained from herbaria (e.g. reviewed by Pyke & Ehrlich, 2010; Johnson *et al.*, 2011; Amici *et al.*, 2014; Nualart *et al.*, 2017; James *et al.*, 2018) (Fig. 2). Species distribution modelling has now reached a level of maturity such that IUCN guidelines instruct on their use (IUCN, 2019), and some regions now have estimates of the proportion of endemic plant species that will experience range declines with climate change projections (Loarie *et al.*, 2008; Yan & Tang, 2019). However, estimates in many tropical regions will still require more effort (Feeley & Silman, 2011; Feeley, 2012). The challenge of estimating the threat of climate change to endangered species in a way that can stimulate conservation action has long stymied conservation assessment efforts, leaving it as a threat that is acknowledged in a vague sense but not quantified, or even left as unknown (McCune *et al.*, 2013; Orsenigo *et al.*, 2018). The way in which plant species and communities have changed with temperature changes has begun to be studied in recent decades (i.e. over the past 40 years; Feeley *et al.*, 2020). This information can greatly strengthen the conclusions from projecting ongoing climate change threats and will hopefully lead to a much-needed acceleration of SDM for threatened plant species threat assessments.

V. Herbarium specimens as a source of conservation genetic data

Genetic data is not often incorporated into species assessments unless it provides evidence of fragmentation (See Table 2; Muniz *et al.*, 2019). In animal species, it has been estimated with NHCs that, on average, we have lost *c.* 6% of genetic diversity since the Industrial Revolution (Leigh *et al.*, 2019). It is therefore reasonable to think that declines in range and abundance are causing the genetic erosion of at least some endangered species. Knowledge of the level of inbreeding and historical loss of haplotypes can be very important for effective species recovery. Herbarium specimens are beginning to be used as the source of DNA to generate this information for plants.

The DNA extracted from herbarium specimens, while often degraded (Hart *et al.*, 2016), has been used successfully to shed light on the phylogenetics (Silva *et al.*, 2017), population genetics (Konrade *et al.*, 2019), biogeography (Zedane *et al.*, 2016), ethnopharmacology (Stepp & Thomas, 2005), and physiology (Besnard *et al.*, 2014) of various plant species. Genetic material from herbarium specimens can also be used for informing conservation practices, such as legal protection (e.g. newly defined species) and reintroduction (Sears, 2011; Muniz *et al.*, 2019). In particular, herbaria may have sufficient sampling over time to detect changes in genetic structure (Table 2), which allows for an understanding of how fragmentation increases due to declines in the number of subpopulations (Bieker & Martin, 2018). For instance, Cozzolino *et al.,* (2007) documented changes in genetic variation in the marsh orchid *Anacamptis palustris* (Jacq.)

R.M.Bateman, Pridgeon & M.W.Chase caused by human induced habitat changes. The analysis of historical herbarium specimens detected the extinction of private haplotypes and alleles, as well as a significant reduction in genetic diversity and an increase in the importance of random genetic drift in this species.

Herbarium specimens can allow for a more rapid investigation of population structure and fragmentation (and their causes) by using, for instance, landscape genetic approaches (Manel & Holderegger, 2013) that are particularly well suited to investigating conservation problems (Keller *et al.*, 2015). In contrast to population genetics, which focuses on well-delimited populations, landscape genetics benefits from samples collected across landscapes and time to identify points of rupture in gene flow. Because the use of herbarium specimens can reduce the need for expensive field trips to gather the necessary material, such strategies could be particularly useful for remote areas. Further, if high levels of gene flow appear to be maintained despite a distribution that indicates fragmentation, this result could indicate that further searching of suitable habitat (e.g. as indicated by SDM (McCune, 2016)) may reveal more occurrences of the species (Fig. 2).

Application of these techniques has previously been hindered because the DNA of aged herbarium specimens usually comes in small fragment sizes and holds various modifications, partially caused by spontaneous mutations (e.g. from depurination and hydrolysis of the DNA backbone (Weiß *et al.*, 2016)). DNA isolation and gene amplification of herbarium specimen materials can therefore be a challenge and can require the modification of frequently used protocols (Riahi *et al.*, 2010). DNA preservation is strongly affected by storage techniques; alcohol-dried specimens have more fragmented DNA sequences than air-dried specimens (Särkinen *et al.*, 2012), and these alterations also affect downstream applications (e.g. Brewer *et al.*, 2019; Forrest *et al.*, 2019). In addition, insects and other pests were historically killed using chemical treatments, which cause further DNA damage (e.g. DNA fragmentation in herbarium specimens treated with mercuric chloride (Lister *et al.*, 2008)), and fungi frequently present in NHCs could be a source of contamination (Bieker *et al.*, 2020). Nevertheless, next-generation sequencing methods are opening up new opportunities. The sequencing of traditional markers (plastid, nuclear ribosomal regions (ITS)) generally do not yield sufficient variation from independent loci to be optimal for conservation studies, but new developments that apply reduced-representation sequencing (Jordon-Thaden *et al.*, 2020), hybridization capture (Brewer *et al.*, 2019; Forrest *et al.*, 2019), shotgun sequencing (Staats *et al.*, 2013), or similar approaches (e.g. Lang *et al.*, 2020) can be used to obtain information from older herbarium specimens. Because many recent high-throughput sequencing approaches are designed to sequence small fragments of DNA (Lang *et al.*, 2020), which is a common characteristic of the usually degraded DNA that can be extracted from herbarium specimens (Neubig *et al.*, 2014), they have overcome many of the pitfalls of using fragmented archival DNA (Wandeler *et al.*, 2007) in numerous genera (Jordon-Thaden *et al.*, 2020). As an example, the complete nuclear genome of a 43-yr-old *Arabidopsis thaliana* (L.) Heynh. specimen has been recovered (Staats *et al.*, 2013), illustrating the potential of these techniques. Conservation biologists can thus employ these approaches and expand on the initial studies exploring where in a species' range we have lost genetic diversity (Table 2).

VI. Herbarium specimens as a source of diaspores and their implication for plant translocation and de-extinction

Although not always ideal, it has been proposed that herbarium specimens are valuable sources of viable diaspores (spores and seeds), which can provide propagules useful for recovering lost genetic variation, including species lost from the wild (e.g. for restoration of self-incompatible species or dioecious species, or for increasing genetic variation in inbred populations (Bowles *et al.*, 1993; Magrini *et al.*, 2010)). Diaspores may remain viable for centuries in herbarium specimens or similar collections (Telewski & Zeevaart, 2002; Molnár *et al.*, 2015) and may have the potential for recovering extinct species (de-extinction (Abeli *et al.*, 2020); Fig. 1). Although there are no documented examples of plant de-extinction using herbarium material, germination of very old seeds collected from herbaria demonstrate the potential feasibility of such actions (see Molnár *et al.*, 2015 and references therein). At present, the only reintroduction of a locally extirpated species using herbarium diaspores is reported by Sears (2011) – the locally extirpated stinking hawk's-beard (*Crepis foetida* L.) was successfully reintroduced in south-east England, UK.

De-extinction can also include approaches where 'extinct' genes are resurrected by genetic engineering, such as by cloning genes of organisms from extirpated sites (Table 2) or extinct species into present day species (Shapiro, 2017). As such, genetic information obtained from herbaria (see Section V, above) can provide the raw material for such projects. Seeds obtained from herbarium specimens could be used for recovering genetic material, even if this can sometimes be problematic because of the scarcity of material (seeds are not always collected), age of the specimens, preservation conditions, seed secondary-dormancy state, low germination rate, diaspore viability, and degraded DNA (see Abeli *et al.*, 2020 for a review of these issues). In this context, storage conditions of herbarium specimens are crucial to maintaining viable propagules, and in reducing DNA degradation (Godefroid *et al.*, 2011; Porteous *et al.*, 2019).

While herbaria may provide a source of seeds for species that we have failed to bank, future efforts should focus on storing seeds in seed banks designed for the task. Facilities such as the Millennium Seed Bank can clearly overcome some of the issues discussed here, but the diversity found in seed banks, while excellent, does not currently encompass a high proportion of rare and threatened species (Paton *et al.*, 2020) and in this regard still needs to be complemented with what is available in herbaria.

VII. Importance of herbaria in raising awareness and gaining momentum for conservation action

In some jurisdictions, the funding allocated to the conservation of animals outweighs that allocated to plants by an order of magnitude, in part due to people's tendency to gravitate toward

animal species and to experience what has been dubbed 'plant blindness' (Balding & Williams, 2016; Jose *et al.*, 2019) or 'plant awareness disparity' (Parsley, 2020). Herbaria can help mitigate some of the neglect for plant conservation, and they are increasingly used for teaching and educational activities in disciplines such as archaeology, paleobotany, ethnobotany and economic botany, as well as in history, art and literature (Heberling *et al.*, 2019), where the common goals center on dissemination and popularization through involvement (Willis *et al.*, 2017b; Park *et al.*, 2019). Species recovery can be greatly accelerated by including an engaged global community of participants, but many questions remain on how best to capture their interest and direct their actions where it is most impactful (Canteiro *et al.*, 2019). Here, we explore ways in which herbaria can best inform nonscientists on how plant diversity has been changed by human activity.

In the context of dissemination of knowledge regarding the current biodiversity crisis, herbaria and a number of different sources of historical records (e.g. ethnographic and ethnohistorical research, land use documents) can be used to infer environmental changes (Bonebrake *et al.*, 2010; Vellend *et al.*, 2013; Armstrong *et al.*, 2017) and show losses in biodiversity when they are integral to a museum's selection of exhibitions. Surprisingly, the cultural and/ or biological importance of a species is rarely extended towards conservation prioritization exercises (Dhar *et al.*, 2000), yet this can be what best captures the public's interest. Herbaria could be integrated into permanent or temporary exhibitions through glass walls as in other labs (typically fossil preparation labs), especially if they feature threatened species that are culturally important. In various prioritization exercises, a heightened risk of extinction is forefront in the discussion of high-priority species, yet the resolution of conflicts over which species are deemed 'more important' than others remains a challenge (Raymond *et al.*, 2018). Prioritization of conservation that includes the societal or cultural valuation of a species has been increasingly explored (Stepp & Thomas, 2005; Pfeiffer & Voeks, 2008; Martín *et al.*, 2010) and could engage a wider audience (Crisci *et al.*, 2020).

Digitization of herbaria is a further tool to make herbaria available to a broad audience and stimulate involvement of children and adults in citizen science projects (Hedrick *et al.*, 2020) such as the one offered by inaturalist.org (Barkworth & Murrell, 2012; Heberling & Isaac, 2018). These citizen science initiatives are a very rich and accessible way for nonscientists to learn more about extant biodiversity in their areas but can fail to depict biodiversity change without the baseline information provided by historical sources. Permanent, temporary or hi-tech thematic exhibitions of herbarium specimens can be effective in directing attention towards extinction (Heberling *et al.*, 2019) of, for example, flagship species, or extirpated taxa. Because many of these approaches link to online citizen science platforms that incorporate images taken with a smartphone, they have the added benefit of providing enriched information about the plant as it appeared before collection (e.g. with associated species, general habitat and population density features, and with unfaded colour).

An overview of the use of herbarium specimens throughout the ages to highlight the presence of important species in the surrounding habitat reveals a checkered history. On the one hand, many collections are rooted in colonial history; on the other hand, they reveal the diversity of people that have used and contributed to herbaria. For example, Victorian-age botany was one of the few scientific areas accessible to women (Shteir, 1997; Shteir & Cayouette, 2019) that was associated with the opportunity to explore different countries and continents, describing, painting and collecting plant specimens (e.g. Marianne North). The hope is that a close look at historical collections can allow for a stark examination of colonialism, while providing a way forward with an atmosphere of increased inclusivity for conservation action. For example, it is estimated that at least 28.1% of the world's land surface is owned or managed by Indigenous peoples, including some of the world's most ecologically intact landscapes (Garnett *et al.*, 2018). As such, ethnobotanical collections, in which cultural and ethnoecological knowledge of plant species (their uses as food and medicine, harvesting habitats, and phenological properties) are stored with vouchers, can be critical for verifying and co-managing important plant populations, as well as for identifying Cultural Keystone Species (CKS) (Przelomska *et al.*, 2020). For subsistence-based peoples especially, CKS play fundamental roles in diet, economy, and medicine, and feature prominently in language, education, governance, and ceremony (Garibaldi & Turner, 2004). Such multi-use and important species, if lost, could be fundamentally damaging to entire communities. Unfortunately, many CKS tend to be susceptible to biopiracy (Hamilton, 2004) (for example, Nestlé's attempt to copyright and genetically funnel the Rooibos plant (Golan *et al.*, 2019)), potentially putting them at greater risk (Robinson, 2010). Herbarium specimens and their attendant biocultural knowledge can be beneficial for Indigenous communities dealing with affronts to their intellectual property rights (Posey & Dutfield, 1996; Maffi & Woodley, 2012). Such information furthers conservation initiatives where Indigenous peoples continue to manage and govern ecologically important and diverse ecosystems (Stepp, 2002; Carlson & Maffi, 2004; Armstrong *et al.*, 2017).

VIII. The emerging role of machine learning in extracting information from herbaria

Increased participation in all of the above facets of conservation biology will become possible as herbarium specimens are rapidly digitized, mobilized, and easily accessible and searchable online (Hedrick *et al.*, 2020). With the additional increase in the use of machine learning tools (a type of AI), the herbaria of tomorrow are positioned to become more broadly relevant than they were in the past. Machine learning algorithms allow computers (machines) to learn patterns in a given dataset to make predictions, and they have played a part in some of the advancements reviewed in the sections above. In fact, the combination of machine learning (e.g. artificial neural networks, deep learning, optical character recognition; see Box 2 for terminology) and herbarium records is facilitating an assortment of noteworthy advances in the extraction of information from herbaria, and we review them here to highlight their breadth and utility. Using the queries 'herbaria + artificial intelligence', 'herbaria + big data', and 'herbaria + machine learning' we found 85 relevant papers, listed in Table 3. The results of the literature

> **Box 2** Summary of key terms related to artificial intelligence and big data.
>
> **Agent technology**: A methodology for the realization of an autonomous decentralized system with cooperative interactions among agents that model each element of the system.
> **Artificial neural network:** A type of machine learning algorithm whose computational model is (loosely) similar to biological neural networks.
> **Computer vision:** An interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos.
> **Deep learning:** A type of machine learning concerned with artificial neural networks, allowing advanced pattern recognition.
> **Machine learning:** A branch of artificial intelligence in which a computer generates rules underlying or based on raw data that has been fed into it.

review show that digitized specimen data are finding rapidly increasing applications in many fields pertinent to stemming the loss of species through extinction, for example, genomics, conservation assessment, ecology, phenology, and taxonomic revisions (Franklin *et al.*, 2017; Soltis *et al.*, 2018, 2020).

Thanks to online data aggregators such as GBIF, iDigBio (Integrated Digitized Biocollections; https://www.idigbio.org/), GPI (Global Plant Initiatives; http://gpi.myspecies.info/), and JACQ (https://www.jacq.org/), which allow for access to millions of digitized specimens together with a significant amount of metadata (annotations, label content, GPS coordinates etc.), herbaria have stepped into the field of big data, paving the way for innovative applications such as artificial intelligence and its branches (Delgado *et al.*, 2005) , agent technology (Bhamra *et al.*, 2014) and machine learning. Use of these aggregators is also often the first step in SDM analyses to estimate the threat of climate change (see Table 1, Fig. 2).

We are witnessing a new era of rapid development and improvement of methods and technologies for herbarium data gathering and management, and we now encourage the academic community to facilitate the conception of ways for the conservation community to learn about and utilize these advancements. Ongoing digitization, online publishing of specimens, collaborative networks for data management and cross-validation, and dataset building are the initial steps (Wang *et al.*, 2009; Matsunaga *et al.*, 2013; Dugenie *et al.*, 2017; Yagui *et al.*, 2017; Dormontt *et al.*, 2018; Theeten *et al.*, 2019), and these are followed by the development of scoring and data-sharing protocols (Yost *et al.*, 2018), the addition of non-traditional digitized data fields, user annotation capabilities, and born-digital field data collection, including citizen-science applications (Heberling & Isaac, 2018; James *et al.*, 2018; Powell *et al.*, 2019). These advancements will make machine learning increasingly user-friendly (e.g. see lobe.ai) and make herbaria accessible to a wider audience, thus allowing for the development of lesson plans on how to fill important knowledge gaps regarding at-risk species while still appealing to a generation raised on social media and simple graphics.

The accelerated development of big data management and machine learning techniques also expands their application to a wider spectrum of plant research fields (Soltis *et al.*, 2020). Machine learning in particular is an extremely useful tool that finds applications in a significant variety of naturalist fields (see Table 3). It is especially useful for analysis of phenological features, as it is possible to automatically identify and extract plant characteristics from digitized specimens (e.g. Corney *et al.*, 2012; Ott *et al.*, 2020; Weaver *et al.*, 2020), describe and annotate them (Reeb *et al.*, 2018; McAllister *et al.*, 2019; Younis *et al.*, 2020), count plants' elements (e.g. the number of flowers per inflorescence) (Davis *et al.*, 2020b), and evaluate intraspecific trait variability (Kommineni *et al.*, 2020), as well as to relate plant traits to environmental factors to analyze complex relationships and trends (Václavík *et al.*, 2017; Schneider *et al.*, 2018a; Park *et al.*, 2020; Yost *et al.*, 2020). Moreover, machine learning is used for species identification and classification, because of its contribution to the automatic collection of morphological characteristics and plant recognition at different taxonomic levels (e.g. Clark *et al.*, 2012; Carranza-Rojas *et al.*, 2018; Younis *et al.*, 2018). This will lead not only to improved species classification (Hussein *et al.*, 2020; Pryer *et al.*, 2020), but will also accelerate species discovery (Little *et al.*, 2020).

Furthermore, machine learning technologies (e.g. OCR, optical character recognition) can be applied to the extraction of metadata from herbarium specimens, for example, collectors names (Silva, 2016), plant traits from textual descriptions (Dagtekin *et al.*, 2018), and label content (Heidorn & Wei, 2008; Kirchhoff *et al.*, 2018; Walton *et al.*, 2020), enhancing integrative and cross-linked research on complex phenomena. Tiered data gathering and analysis from – to date – unconnected data expand the information that herbarium collections can provide (Soltis *et al.*, 2018; Lendemer *et al.*, 2019; Theeten *et al.*, 2019). Plant species distribution modelling of, for example, forest (Lorena *et al.*, 2011) and invasive species (Biganzoli *et al.*, 2013), ecological and climatic niche modelling (Amici *et al.*, 2014; Watkins *et al.*, 2020) and regional biodiversity pattern prediction (Park *et al.*, 2020) are some of the many ecological aspects for which the contribution of big data and the application of data mining techniques can be used with herbarium data and metadata is increasingly useful and promising (Huettmann & Ickert-Bond, 2018; Zurell *et al.*, 2020) (see Tables 1, 3).

In addition, the use of recent artificial intelligence technologies for evaluating and predicting plant conservation status is particularly relevant and represents a key application in conservation policy and biodiversity strategy (Hochkirch *et al.*, 2020). Tressou & Haevermans (2018) developed a method for plant species threat prediction (see Section IV, above) and Zizka *et al.* (2020) used deep learning for automated conservation assessment of at-risk families (see Section III, above). These innovative, automatically generated assessments show promising results for facilitating and speeding up Red List assessments (Bachman *et al.*, 2020) and species recovery, provided the scientific community accepts that the results are robust. The applications of artificial intelligence for herbarium collections analysis mentioned here are not free of bias or limitations. Herbarium collections have gaps and biases involving limited specimen availability, in both spatial and temporal terms,

and functional trait, phylogenetic, and collector characteristics that must be taken into account when using specimen data and datasets (Meyer *et al.*, 2016; Daru *et al.*, 2018; Cornwell *et al.*, 2019; Nic Lughadha *et al.*, 2019). Moreover, the accuracy of a herbarium's automated image analysis decreases as traits diminish in size, at lower resolution, and if organs overlap (Borges *et al.*, 2020). Phenological determination using machine learning is limited by factors such as damage obscuring certain traits, material being stored in opaque packets and the morphological characteristics of some species (e.g. seeds housed in fruits and indistinguishable male/female organs (Pearson *et al.*, 2020)). Finally, the application and development of novel methods for data extraction, management and analysis depends on the digitization of the collections, and although this is increasingly performed in many of the world's herbaria, it is still subject to funding and human effort limitations (James *et al.*, 2018). Nevertheless, these new data collection and management techniques suggest promising and more convenient solutions to the current limitations (see Fig. 1). The tools are available to determine which species are at risk with exceptional precision and accuracy. What remains to be developed is a way to mobilize and resource the coordinated community of people necessary to bring these species back from the brink.

IX. Conclusions and future perspectives

Conservation biology is a discipline that bridges science with policy to provide convincing arguments that the combination of features of isolation, migration, demography and human-induced disturbance are placing species at risk of extinction. Here we have specifically reviewed how information housed in herbaria provide information necessary for implementing conservation actions for plants. The vast quantity of digitized herbarium specimens has contributed to the successful acceleration of the designation of conservation units, conservation status assessments, and threat assessments. The long history of specimen collecting in a standardized way provides insight into biological, ecological, demographic, and cultural features of species that can have important impacts for their conservation. In addition to emerging as a tool for examining genetic diversity loss, the use of herbarium specimens can directly provide genetic material for translocation, including de-extinction. Finally, the increased accessibility of machine learning citizen science platforms can broaden the audience involved in conservation efforts and accelerate the understanding of the effects of global change among the general public.

We are left with an important question: how do we imagine herbaria in the next century? Despite the new uses of herbaria highlighted here and in other reviews, the word 'herbarium' often harkens images of dusty rooms with thousands of specimens stacked according to an obsolete taxonomic order. This 19th century vision is likely to change in the near future, and even more by the end of this and the next century. While in the past we could get information only from looking at the physical specimens, our technical and technological ability to extract data from the specimens (e.g. from images (Borges *et al.*, 2020)) makes this virtual information increasingly important. In other words, virtual data will enrich (rather than replace) the information provided by the physical objects themselves, with collections providing the essential first step to a new way of understanding plant diversity, ecology and physiology.

3-D laser scanning (Schneider *et al.*, 2018b), multispectral imaging (Schneider *et al.*, 2018b) (e.g. UV, infrared, X-ray; Durgante *et al.*, 2013; Schneider *et al.*, 2018b; Do *et al.*, 2020), multiscale imaging (e.g. scanning electron microscopy (SEM) and tomography; Vega *et al.*, 2019), fine-scale genetic analysis (Daru *et al.*, 2019), and environmental DNA (Lang *et al.*, 2019) coupled with AI and big data analysis will allow us to obtain an unprecedented amount of data from old and newly collected specimens. With the increasing resolution of cameras, herbaria may in some cases become detached from the physical specimens and may be associated with high-resolution images taken in the field. Future herbaria should therefore be regarded as 'extended specimens' (Lendemer *et al.*, 2019), available for different fields of research, and to a much broader type of end user (including citizens) in addition to botanists and conservationists. If we wish to stem the tide of plant extinction for future generations, we will need as large a crowd as we can gather to apply the tools developed in other disciplines to uncover the historical information herbaria provide.

ORCID

Thomas Abeli (iD) https://orcid.org/0000-0003-3096-2035
Giulia Albani Rocchetti (iD) https://orcid.org/0000-0001-7361-9179
Anne Bruneau (iD) https://orcid.org/0000-0001-5547-0796
Simon Joly (iD) https://orcid.org/0000-0002-0865-8267
Simone Orsenigo (iD) https://orcid.org/0000-0003-0348-9115
Jana C. Vamosi (iD) https://orcid.org/0000-0002-2376-0729

References

Abeli T, Barni E, Siniscalco C, Amosso C, Rossi G. 2012. A cost-effective model for preliminary site evaluation for the reintroduction of a threatened quillwort. *Aquatic Conservation: Marine and Freshwater Ecosystems* **22**: 66–73.

Abeli T, Rossi G, Smolders AJP, Orsenigo S. 2014. Nitrogen pollution negatively affects *Stratiotes aloides* in Central-Eastern Europe. Implications for translocation actions. *Aquatic Conservation: Marine and Freshwater Ecosystems* **24**: 724–729.

Abeli T, Dalrymple S, Godefroid S, Mondoni A, Müller JV, Rossi G, Orsenigo S. 2020. Ex situ collections and their potential for the restoration of extinct plants. *Conservation Biology* **34**: 303–313.

Agapow PM, Bininda-Emonds OR, Crandall KA, Gittleman JL, Mace GM, Marshall JC, Purvis A. 2004. The impact of species concept on biodiversity studies. *The Quarterly Review of Biology* 79: 161–179.

Akcakaya HR, Bennett E, Brooks T, Grace M, Heath A, Hedges S, Hilton-Taylor C, Hoffmann M, Keith D, Long B *et al.* 2018. Quantifying species recovery and conservation success to develop an IUCN Green List of Species. *Conservation Biology* 32: 1128–1138.

Amici V, Geri F, Bonini I, Rocchini D. 2014. Ecological niche modelling with herbarium data: a framework to improve Natura 2000 habitat monitoring. *Applied Ecology and Environmental Research* 12: 645–659.

Antonelli A, Fry C, Smith RJ, Simmonds MSJ, Kersey PJ, Pritchard HW. 2020. The state of the world's plants 2020. Royal Botanic Gardens, Kew. doi: 10.34885/ 172

Antonovics J, Hood ME, Thrall PH, Abrams JY, Duthie GM. 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.

Armstrong CG, Shoemaker AC, McKechnie I, Ekblom A, Szabo P, Lane PJ, McAlvay AC, Boles OJ, Walshaw S, Petek N *et al.* 2017. Anthropological contributions to historical ecology: 50 questions, infinite prospects. *PLoS ONE* 12: e0171883.

Attorre F, Abeli T, Bacchetta G, Farcomeni A, Giuseppe F, Sanctis M, Gargano D, Peruzzi L, Montagnani C, Rossi G *et al.* 2018. How to include the impact of climate change in the extinction risk assessment of policy plant species? *Journal for Nature Conservation* 44: 43–49.

Bachman SP, Nic Lughadha EM, Rivers MC. 2018. Quantifying progress toward a conservation assessment for all plants. *Conservation Biology* 32: 516–524.

Bachman S, Walker BE, Barrios S, Copeland A, Moat J. 2020. Rapid Least Concern: towards automating Red List assessments. *Biodiversity Data Journal* 8: e47018.

Bakker F, Antonelli A, Clarke J, Cook J, Edwards S, Ericson P, Faurby S, Ferrand N, Gelang M, Gillespie R *et al.* 2020. The Global Museum: natural history collections and the future of evolutionary science and public education. *PeerJ* 8: e8225.

Balding M, Williams KJH. 2016. Plant blindness and the implications for plant conservation. *Conservation Biology* 30: 1192–1199.

Barkworth M, Murrell Z. 2012. The US Virtual Herbarium: working with individual herbaria to build a national resource. *ZooKeys* 209: 55–73.

Beauvais M-P, Pellerin S, Dubé J, Lavoie C. 2017. Herbarium specimens as tools to assess the impact of large herbivores on plant species. *Botany-Botanique* 95: 153–162.

Besnard G, Christin P-A, Malé P-JG, Lhuillier E, Lauzeral C, Coissac E, Vorontsova MS. 2014. From museums to genomics: old herbarium specimens shed light on a C3 to C4 transition. *Journal of Experimental Botany* 65: 6711–6721.

Besnard G, Gaudeul M, Lavergne S, Muller S, Rouhan G, Sukhorukov AP, Vanderpoorten A, Jabbour F. 2018. Herbarium-based science in the twenty-first century. *Botany Letters* 165: 323–327.

Bhamra GS, Verma AK, Patel RB. 2014. Agent technology in bioinformatics: a review. *International Journal of Next-Generation Computing* 5: http://perpetua linnovation.net/ojs/index.php/ijngc/article/view/250.

Bieker VC, Sánchez Barreiro F, Rasmussen JA, Brunier M, Wales N, Martin MD. 2020. Metagenomic analysis of historical herbarium specimens reveals a postmortem microbial community. *Molecular Ecology Resources* 20: 1206–1219.

Bieker VC, Martin MD. 2018. Implications and future prospects for evolutionary analyses of DNA in historical herbarium collections. *Botany Letters* 165: 409–418.

Biganzoli F, Larsen C, Rolhauser A. 2013. Range expansion and potential distribution of the invasive grass *Bromus tectorum* in southern South America on the base of herbarium records. *Journal of Arid Environments* 97: 230–236.

Bird JP, Martin R, Akçakaya HR, Gilroy J, Burfield IJ, Garnett ST, Symes A, Taylor J, Şekercioğlu ÇH, Butchart SHM. 2020. Generation lengths of the world's birds and their implications for extinction risk. *Conservation Biology* 34: 1252–1261.

Bonebrake TC, Christensen J, Boggs CL, Ehrlich PR. 2010. Population decline assessment, historical baselines, and conservation. *Conservation Letters* 3: 371–378.

Bontrager M, Angert AL. 2016. Effects of range-wide variation in climate and isolation on floral traits and reproductive output of *Clarkia pulchella*. *American Journal of Botany* 103: 10–21.

Borges LM, Reis VC, Izbicki R. 2020. Schrödinger's phenotypes: Herbarium specimens show two-dimensional images are both good and (not so) bad sources of morphological data. *Methods in Ecology and Evolution* 11: 1296–1308.

Bowles ML, Betz RF, DeMauro MM. 1993. Propagation of rare plants from historic seed collections: implications for species restoration and herbarium management. *Restoration Ecology* 1: 101–106.

Brewer GE, Clarkson JJ, Maurin O, Zuntini AR, Barber V, Bellot S, Biggs N, Cowan RS, Davies NMJ, Dodsworth S *et al.* 2019. Factors affecting targeted sequencing of 353 nuclear genes from herbarium specimens spanning the diversity of angiosperms. *Frontiers in Plant Science* 10: 1102.

Bruneau A, Joly S, Starr J, Drouin J-N. 2011. Molecular markers indicate that the narrow Québec endemics *Rosa rousseauiorum* and *Rosa williamsii* are synonymous with the widespread *Rosa blanda*. *Canadian Journal of Botany* 83: 386–398.

Brunet J, Zalapa J, Guries R. 2016. Conservation of genetic diversity in slippery elm (*Ulmus rubra*) in Wisconsin despite the devastating impact of Dutch elm disease. *Conservation Genetics* 17: 1001–1010.

Canteiro C, Barcelos L, Filardi F, Forzza R, Green L, Lanna J, Leitman P, Milliken W, Pires Morim M, Patmore K *et al.* 2019. Enhancement of conservation knowledge through increased access to botanical information. *Conservation Biology* 33: 523–533.

Carlson TJ, Maffi L. 2004. *Ethnobotany and conservation of biocultural diversity*. Bronx, NY, USA: New York Botanical Garden.

Carranza-Rojas J, Goeau H, Bonnet P, Mata-Montero E, Joly A. 2017. Going deeper in the automated identification of herbarium specimens. *BMC Evolutionary Biology* 17: 181.

Carranza-Rojas J, Joly A, Goëau H, Mata-Montero E, Bonnet P. 2018. Automated identification of herbarium specimens at different taxonomic levels. In: Joly A, Vrochidis S, Karatzas K, Karppinen A, Bonnet P, eds. *Multimedia tools and applications for environmental & biodiversity informatics*. Cham, Switzerland: Springer, 151–167.

Carrington CMS, Krupnick GA, Acevedo-Rodríguez P. 2017. Herbarium-based preliminary conservation assessments of Lesser Antillean endemic seed plants reveal a flora at risk. *The Botanical Review* 83: 107–151.

Case M, Flinn K, Jancaitis J, Alley A, Paxton A. 2007. Declining abundance of American ginseng (*Panax quinquefolius* L.) documented by herbarium specimens. *Biological Conservation* 134: 22–30.

Castro S, Dostálek T, van der Meer S, Oostermeijer G, Münzbergová Z. 2015. Does pollen limitation affect population growth of the endangered *Dracocephalum austriacum* L.? *Population Ecology* 57: 105–116.

Clark JY, Corney DP, Tang HL. 2012. Automated plant identification using artificial neural networks. *IEEE Symposium on computational intelligence in bioinformatics and computational biology (CIBCB)*. San Diego, CA, USA: IEEE: 343–348.

Colla SR, Gadallah F, Richardson L, Wagner D, Gall L. 2012. Assessing declines of North American bumble bees (*Bombus* spp.) using museum specimens. *Biodiversity and Conservation* 21: 3585–3595.

Corney DPA, Clark JY, Tang HL, Wilkin P. 2012. Automatic extraction of leaf characters from herbarium specimens. *Taxon* 61: 231–244.

Cornwell WK, Pearse WD, Dalrymple RL, Zanne AE. 2019. What we (don't) know about global plant diversity. *Ecography* 42: 1819–1831.

Cozzolino S, Cafasso D, Pellegrino G, Musacchio A, Widmer A. 2007. Genetic variation in time and space: the use of herbarium specimens to reconstruct patterns of genetic variation in the endangered orchid *Anacamptis palustris*. *Conservation Genetics* 8: 629–639.

Crawford P, Hoagland B. 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.

Crisci JV, Katinas L, Apodaca MJ, Hoch PC. 2020. The End of Botany. *Trends in Plant Science* 25: 1173–1176.

Dagtekin O, Ulate W, Batista-Navarro R. 2018. Towards the automatic extraction of plant traits from textual descriptions. *10th International conference on ecological informatics – Translating ecological data into knowledge and decisions in a rapidly changing world*. Jena, Germany: ICEI.

van Dam H, Mertens A. 1993. Diatoms on herbarium macrophytes as indicators for water quality. *Hydrobiologia* 269: 437–445.

Daru BH, Park DS, Primack RB, Willis CG, Barrington DS, Whitfeld TJS, Seidler TG, Sweeney PW, Foster DR, Ellison AM *et al.* 2018. Widespread sampling biases in herbaria revealed from large-scale digitization. *New Phytologist* 217: 939–955.

Daru BH, Bowman E, Pfister D, Arnold A. 2019. A novel proof of concept for capturing the diversity of endophytic fungi preserved in herbarium specimens. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 374: 20170395.

Davis C, Kennedy J, Grassa C. 2020a. Back to the future: a refined single user photostation for massively scaling herbarium digitization. *bioRxiv*: doi: 10.1101/2020.09.09.288506.

Davis CC, Champ J, Park DS, Breckheimer I, Lyra GM, Xie J, Joly A, Tarapore D, Ellison AM, Bonnet P. 2020b. A new method for counting reproductive structures in digitized herbarium specimens using Mask R-CNN. *Frontiers in Plant Science* 11: 1129.

Delgado M, Fajardo W, Gibaja E, Pérez-Pérez R. 2005. BioMen: an information system to herbarium. *Expert Systems with Applications* 28: 507–518.

Delisle F, Lavoie C, Jean M, Lachance D. 2003. Reconstructing the spread of invasive plants: taking into account biases associated with herbarium specimens. *Journal of Biogeography* 30: 1033–1042.

Devey DS, Forest F, Rakotonasolo F, Ma P, Dentinger BTM, Buerki S. 2013. A snapshot of extinction in action: the decline and imminent demise of the endemic *Eligmocarpus Capuron* (Caesalpinioideae, Leguminosae) serves as an example of the fragility of Madagascan ecosystems. *South African Journal of Botany* 89: 273–280.

Dhar U, Rawal RS, Upreti J. 2000. Setting priorities for conservation of medicinal plants – a case study in the Indian Himalaya. *Biological Conservation* 95: 57–65.

Dillen M, Groom Q, Chagnoux S, Güntsch A, Hardisty A, Haston E, Livermore L, Runnel V, Schulman L, Willemse L *et al.* 2019. A benchmark dataset of herbarium specimen images with label data. *Biodiversity Data Journal* 7: e31817.

Do C, Abubakari F, Remigio AC, Brown GK, Casey LW, Burtet-Sarramegna V, Gei V, Erskine PD, van der Ent A. 2020. A preliminary survey of nickel, manganese and zinc (hyper)accumulation in the flora of Papua New Guinea from herbarium X-ray fluorescence scanning. *Chemoecology* 30: 1–13.

Dolan RW, Moore ME, Stephens JD. 2011. Documenting effects of urbanization on flora using herbarium records. *Journal of Ecology* 99: 1055–1062.

Dormontt EE, van Dijk K-j, Bell KL, Biffin E, Breed MF, Byrne M, Caddy-Retalic S, Encinas-Viso F, Nevill PG, Shapcott A *et al.* 2018. Advancing DNA barcoding and metabarcoding applications for plants requires systematic analysis of herbarium collections—an Australian perspective. *Frontiers in Ecology and Evolution* 6: 134.

Dugenie P, Freire N, Broeder D. 2017. Building new knowledge from distributed scientific corpus: HERBADROP & EUROPEANA: Two concrete case studies for exploring big archival data. *IEEE International Conference on Big Data (Big Data)*. Boston, MA, USA: IEEE, 2231–2239.

Durgante FM, Higuchi N, Almeida A, Vicentini A. 2013. Species spectral signature: discriminating closely related plant species in the amazon with near-infrared leaf-spectroscopy. *Forest Ecology and Management* 291: 240–248.

Elith J, H. Graham C, P. Anderson R, Dudík M, Ferrier S, Guisan A, J. Hijmans R, Huettmann F, R. Leathwick J, Lehmann A *et al.* 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129–151.

Ellwood E, Dunckel B, Flemons P, Robert G, Nelson G, Greg N, Newman S, Paul D, Riccardi G, Rios N *et al.* 2015. Accelerating the digitization of biodiversity research specimens through online public participation. *BioScience* 65: 383–396.

Fan YJ, Li FH, Zhao YL, Yan W. 2012. The preliminary design and implement of ectomycorrhizae digital information retrieval system. *International symposium on information technologies in medicine and education*. Hokodate, Hokkaido, Japan: IEEE, 495–498.

Farnsworth EJ, Ogurcak DE. 2006. Biogeography and decline of rare plants in New England: historical evidence and contemporary monitoring. *Ecological Applications* 16: 1327–1337.

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

Feeley KJ, Bravo-Avila C, Fadrique B, Perez TM, Zuleta D. 2020. Climate-driven changes in the composition of New World plant communities. *Nature Climate Change* 10: 965–970.

Feeley KJ, Silman MR. 2011. The data void in modeling current and future distributions of tropical species. *Global Change Biology*. 17: 626–630.

Forrest LL, Hart ML, Hughes M, Wilson HP, Chung K-F, Tseng Y-H, Kidner CA. 2019. The limits of Hyb-Seq for herbarium specimens: Impact of preservation techniques. *Frontiers in Ecology and Evolution* 7: 10-3389.

Frankham R, Ballou JD, Dudash MR, Eldridge MDB, Fenster CB, Lacy RC, Mendelson JR, Porton IJ, Ralls K, Ryder OA. 2012. Implications of different species concepts for conserving biodiversity. *Biological Conservation* 153: 25–31.

Franklin J, Serra-Diaz JM, Syphard AD, Regan HM. 2017. Big data for forecasting the impacts of global change on plant communities. *Global Ecology and Biogeography* 26: 6–17.

Fraser DJ, Bernatchez L. 2001. Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular Ecology* 10: 2741–2752.

Fréville H, McConway K, Dodd M, Silvertown J. 2007. Prediction of extinction in plants: interaction of extrinsic threats and life history traits. *Ecology* 88: 2662–2672.

Funk VA. 2018. Collections-based science in the 21st Century. *Journal of Systematics and Evolution* 56: 175–193.

Garibaldi A, Turner N. 2004. Cultural keystone species: implications for ecological conservation and restoration. *Ecology and Society* 9: 1.

Garnett ST, Burgess ND, Fa JE, Fernández-Llamazares Á, Molnár Z, Robinson CJ, Watson JEM, Zander KK, Austin B, Brondizio ES *et al.* 2018. A spatial overview of the global importance of Indigenous lands for conservation. *Nature Sustainability* 1: 369–374.

Glon H, Heumann B, Carter J, Bartek J, Monfils A. 2017. The contribution of small collections to species distribution modelling: a case study from Fuireneae (Cyperaceae). *Ecological Informatics* 42: 67–78.

Godefroid S, Van de Vyver A, Stoffelen P, Robbrecht E, Vanderborght T. 2011. Testing the viability of seeds from old herbarium specimens for conservation purposes. *Taxon* 60: 565–569.

Goëau H, Mora-Fallas A, Champ J, Love NLR, Mazer SJ, Mata-Montero E, Joly A, Bonnet P. 2020. A new fine-grained method for automated visual analysis of herbarium specimens: a case study for phenological data extraction. *Applications in Plant Sciences* 8: e11368.

Golan J, Athayde S, Olson EA, McAlvay A. 2019. Intellectual property rights and ethnobiology: an update on Posey's call to action. *Journal of Ethnobiology* 39: 90–109.

Goodwin Z, Harris D, Filer D, Wood J, Scotland R. 2015. Widespread mistaken identity in tropical plant collections. *Current Biology* 25: R1066–R1067.

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

Greve M, Lykke AM, Fagg CW, Gereau RE, Lewis GP, Marchant R, Marshall AR, Ndayishimiye J, Bogaert J, Svenning JC. 2016. Realising the potential of herbarium records for conservation biology. *South African Journal of Botany* 105: 317–323.

Grimm J, Hoffmann M, Stöver B, Müller K, Steinhage V. 2016. Image-based identification of plant species using a model-free approach and active learning. *Joint German/Austrian Conference on Artificial Intelligence* (Künstliche Intelligenz) Klagenfurt, Austria: Springer, 169–176.

Hallingbäck T. 1992. The effect of air pollution on mosses in southern Sweden. *Biological Conservation* 59: 163–170.

Hamilton AC. 2004. Medicinal plants, conservation and livelihoods. *Biodiversity & Conservation* 13: 1477–1517.

Hart ML, Forrest LL, Nicholls JA, Kidner CA. 2016. Retrieval of hundreds of nuclear loci from herbarium specimens. *Taxon* 65: 1081–1092.

Heberling JM, Prather LA, Tonsor SJ. 2019. The changing uses of herbarium data in an era of global change: an overview using automated content analysis. *BioScience* 69: 812–822.

Heberling JM, Isaac BL. 2018. iNaturalist as a tool to expand the research value of museum specimens. *Applications in Plant Sciences* 6: e01193.

Hedrick B, Heberling M, Meineke E, Turner K, Grassa C, Park D, Kennedy J, Clarke J, Cook J, Blackburn D *et al.* 2020. Digitization and the future of natural history collections. *BioScience* 70: 243–251.

Heidorn PB, Wei Q. 2008. Automatic metadata extraction from museum specimen labels. In: Greenberg J, Klas W eds. *International Conference on Dublin Core and Metadata Applications*. Universitätsverlag Göttingen, 57–68.

Henning T, Plitzner P, Güntsch A, Berendsohn WG, Müller A, Kilian N. 2018. Building compatible and dynamic character matrices – current and future use of specimen-based character data. *Botany Letters* 165: 352–360.

Henries DG, Tashakkori R. 2012. Extraction of leaves from herbarium images. *IEEE International Conference on Electro/Information Technology*. Indianapolis, IN, USA: IEEE, 1–6.

Hochkirch A, Samways MJ, Gerlach J, Böhm M, Williams P, Cardoso P, Cumberlidge N, Stephenson PJ, Seddon MB, Clausnitzer V *et al.* 2020. A strategy for the next decade to address data deficiency in neglected biodiversity. *Conservation Biology*. doi: 10.1111/cobi.13589.

Huettmann F, Ickert-Bond SM. 2018. On open access, data mining and plant conservation in the Circumpolar North with an online data example of the Herbarium, University of Alaska Museum of the North. *Arctic Science* 4: 433–470.

Hussein BR, Malik OA, Ong WH, Slik JWF. 2020. Automated classification of tropical plant species data based on machine learning techniques and leaf trait measurements. In:Alfred R, Lim Y, Haviluddin H, On C, eds. *Computational science and technology*. Singapore: Springer, 85–94.

Hutchings MJ, Robbirt KM, Roberts DL, Davy AJ. 2018. Vulnerability of a specialized pollination mechanism to climate change revealed by a 356-year analysis. *Botanical Journal of the Linnean Society* 186: 498–509.

IUCN. 2013. *Guidelines for reintroductions and other conservation translocations*, v.1.0. Gland, Switzerland: IUCN.

IUCN. 2016. *IUCN SSC Guiding principles on creating proxies of extinct species for conservation benefit*, v.1.0. Gland, Switzerland: IUCN.

IUCN. 2017. *The IUCN red list of threatened species*, v.2017-3. [WWW document] URL http://www.iucnredlist.org [accessed 05 December 2017].

IUCN. 2019. *Standards and Petitions Committee. Guidelines for using the IUCN red list categories and criteria*, v.14. Gland, Switzerland: IUCN.

James S, Soltis P, Belbin L, Chapman A, Nelson G, Paul D, Collins M. 2018. Herbarium data: global biodiversity and societal botanical needs for novel research. *Applications in Plant Sciences* 6: e01024.

Johnson AL, Rebolleda-Gómez M, Ashman T-L. 2019. Pollen on stigmas of herbarium specimens: a window into the impacts of a century of environmental disturbance on pollen transfer. *The American Naturalist* 194: 405–413.

Johnson KG, Brooks SJ, Fenberg PB, Glover AG, James KE, Lister AM, Michel E, Spencer M, Todd JA, Valsami-Jones E *et al.* 2011. Climate change and biosphere response: unlocking the collections vault. *BioScience* 61: 147–153.

Joly A, Goëau H, Kahl S, Botella C, De Castaneda RR, Glotin H, Cole E, Champ J, Deneu B, Servajean M *et al.* 2020. LifeCLEF 2020 teaser: biodiversity identification and prediction challenges. In: Jose J, ed. *Advances in Information Retrieval. ECIR 2020*. Lecture Notes in Computer Science. *vol. 12036*. Cham, Switzerland: Springer.

Jordon-Thaden IE, Beck JB, Rushworth CA, Windham MD, Diaz N, Cantley JT, Martine CT, Rothfels CJ. 2020. A basic ddRADseq two-enzyme protocol performs well with herbarium and silica-dried tissues across four genera. *Applications in Plant Sciences* 8: e11344.

Jose SB, Wu C-H, Kamoun S. 2019. Overcoming plant blindness in science, education, and society. *Plants, People, Planet* 1: 169–172.

Kamilov M, Hudayberdiev M, Khamroev A. 2019. Algorithm for the development of a training set that best describes the objects of recognition. *Procedia Computer Science* 150: 116–122.

Keller D, Holderegger R, van Strien MJ, Bolliger J. 2015. How to make landscape genetics beneficial for conservation management? *Conservation Genetics* 16: 503–512.

Kharouba HM, Vellend M. 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.

Kho SJ, Manickam S, Malek S, Mosleh M, Dhillon SK. 2017. Automated plant identification using artificial neural network and support vector machine. *Frontiers in Life Science* 10: 98–107.

Kirchhoff A, Bügel U, Santamaria E, Reimeier F, Röpert D, Tebbje A, Güntsch A, Chaves F, Steinke K-H, Berendsohn W. 2018. Toward a service-based workflow for automated information extraction from herbarium specimens. *Database* 2018: bay103.

Komminen VK, Kattge J, Gaikwad J, Baddam P, Tautenhahn S. 2020. Understanding intraspecific trait variability using digital herbarium specimen images. *Biodiversity Information Science and Standards* 4: e59061.

Konrade L, Shaw J, Beck J. 2019. A rangewide herbarium-derived dataset indicates high levels of gene flow in black cherry (*Prunus serotina*). *Ecology and Evolution* 9: 975–985.

Kosanic A, Anderson K, Harrison S, Turkington T, Bennie J. 2018. Changes in the geographical distribution of plant species and climatic variables on the West Cornwall peninsula (South West UK). *PLoS ONE* 13: e0191021.

Kozlov MV, Sokolova IV, Zverev V, Egorov AA, Goncharov MY, Zvereva EL. 2020. Biases in estimation of insect herbivory from herbarium specimens. *Scientific Reports* 10: 12298.

Kricsfalusy VV, Trevisan N. 2014. Prioritizing regionally rare plant species for conservation using herbarium data. *Biodiversity and Conservation* 23: 39–61.

Krupnick G, Kress W, Wagner W. 2009. Achieving Target 2 of the Global Strategy for Plant Conservation: building a preliminary assessment of vascular plant species using data from herbarium specimens. *Biodiversity and Conservation* 18: 1459–1474.

Lambert F, Clark JL, Joly S. 2017. Species delimitation in the Caribbean *Gesneria viridiflora* complex (Gesneriaceae) reveals unsuspected endemism. *Taxon* 66: 1171–1183.

Lang PLM, Willems FM, Scheepens JF, Burbano HA, Bossdorf O. 2019. Using herbaria to study global environmental change. *New Phytologist* 221: 110–122.

Lang PLM, Weiß CL, Kersten S, Latorre SM, Nagel S, Nickel B, Meyer M, Burbano HA. 2020. Hybridization ddRAD-sequencing for population genomics of nonmodel plants using highly degraded historical specimen DNA. *Molecular Ecology Resources* 20: 1228–1247.

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, Joly S, Bergeron A, Guay G, Groeneveld E. 2016. Explaining naturalization and invasiveness: new insights from historical ornamental plant catalogs. *Ecology and Evolution* 6: 7188–7198.

Lavoie C, Delisle F. 2005. The spread of reed canarygrass (*Phalaris arundinacea*) in Québec: a spatio-temporal perspective. *Ecoscience* 12: 366–375.

Le Breton TD, Zimmer HC, Gallagher RV, Cox M, Allen S, Auld TD. 2019. Using IUCN criteria to perform rapid assessments of at-risk taxa. *Biodiversity and Conservation* 28: 863–883.

Ledesma DA, Powell CA, Shaw J, Qin H. 2020. Enabling automated herbarium sheet image post-processing using neural network models for color reference chart detection. *Applications in Plant Sciences* 8: e11331.

Leigh DM, Hendry AP, Vázquez-Domínguez E, Friesen VL. 2019. Estimated six per cent loss of genetic variation in wild populations since the industrial revolution. *Evolutionary Applications* 12: 1505–1512.

Lendemer J, Thiers B, Monfils AK, Zaspel J, Ellwood ER, Bentley A, LeVan K, Bates J, Jennings D, Contreras D *et al.* 2019. The extended specimen network: a strategy to enhance US biodiversity collections, promote research and education. *BioScience* 70: 23–30.

Lewthwaite JMM, Angert AL, Kembel SW, Goring SJ, Davies TJ, Mooers AØ, Sperling FAH, Vamosi SM, Vamosi JC, Kerr JT. 2018. Canadian butterfly climate debt is significant and correlated with range size. *Ecography* 41: 2005–2015.

Lienert J, Fischer M, Diemer M. 2002. Local extinctions of the wetland specialist *Swertia perennis* L. (Gentianaceae) in Switzerland: A revisitation study based on herbarium records. *Biological Conservation* 103: 65–76.

Lister AM. 2011. Natural history collections as sources of long-term datasets. *Trends in Ecology & Evolution* 26: 153–154.

Lister DL, Bower MA, Howe CJ, Jones MK. 2008. Extraction and amplification of nuclear DNA from herbarium specimens of emmer wheat: a method for assessing DNA preservation by maximum amplicon length recovery. *Taxon* 57: 254–258.

Little DP, Tulig M, Tan KC, Liu Y, Belongie S, Kaeser-Chen C, Michelangeli FA, Panesar K, Guha RV, Ambrose BA. 2020. An algorithm competition for automatic species identification from herbarium specimens. *Applications in Plant Sciences* 8: e11365.

Loarie SR, Carter BE, Hayhoe K, McMahon S, Moe R, Knight CA, Ackerly DD. 2008. Climate change and the future of California's endemic flora. *PLoS ONE* 3: e2502.

Loiselle B, Jørgensen P, Consiglio T, Jiménez I, Blake J, Lohmann L, Montiel O. 2008. Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes? *Journal of Biogeography* 35: 105–116.

Lorena AC, Jacintho LFO, Siqueira MF, Giovanni RD, Lohmann LG, de Carvalho ACPLF, Yamamoto M. 2011. Comparing machine learning classifiers in potential distribution modelling. *Expert Systems with Applications* 38: 5268–5275.

Lorieul T, Pearson KD, Ellwood ER, Goëau H, Molino J-F, Sweeney PW, Yost JM, Sachs J, Mata-Montero E, Nelson G *et al.* 2019. Toward a large-scale and deep phenological stage annotation of herbarium specimens: Case studies from temperate, tropical, and equatorial floras. *Applications in Plant Sciences* 7: e01233.

Maffi L, Woodley E. 2012. *Biocultural diversity conservation: a global sourcebook.* Routledge.

de Magalhães RF, Lemes P, Camargo A, Oliveira U, Brandão RA, Thomassen H, Garcia PCdA, Leite FSF, Santos FR. 2017. Evolutionarily significant units of the critically endangered leaf frog *Pithecopus ayeaye* (Anura, Phyllomedusidae) are not effectively preserved by the Brazilian protected areas network. *Ecology and Evolution* 7: 8812–8828.

Magrini S, Olmati C, Onofri S, Scoppola A. 2010. Recovery of viable germplasm from herbarium specimens of *Osmunda regalis* L. *American Fern Journal* 100: 159–166.

Manel S, Holderegger R. 2013. Ten years of landscape genetics. *Trends in Ecology & Evolution* 28: 614–621.

Martín JL, Cardoso P, Arechavaleta M, Borges PAV, Faria BF, Abreu C, Aguiar AF, Carvalho JA, Costa AC, Cunha RT *et al.* 2010. Using taxonomically unbiased criteria to prioritize resource allocation for oceanic island species conservation. *Biodiversity and Conservation* 19: 1659–1682.

Mata-Montero E, Carranza-Rojas J. 2016. Automated plant species identification: challenges and opportunities. In: Mata F, Pont A, eds. *IFIP world information technology forum.* Cham, Switzerland: Springer, 26–36.

Matsunaga A, Thompson A, Figueiredo RJ, Germain-Aubrey CC, Collins M, Beaman RS, MacFadden BJ, Riccardi G, Soltis PS, Page LM *et al.* 2013. A computational-and storage-cloud for integration of biodiversity collections. *EEE 9th International Conference on e-Science.* Beijing, China: IEEE, 78–87.

McAllister CA, McKain MR, Li M, Bookout B, Kellogg EA. 2019. Specimen-based analysis of morphology and the environment in ecologically dominant grasses: the power of the herbarium. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 374: 20170403.

McCune JL, Harrower WL, Avery-Gomm S, Brogan JM, Csergő A-M, Davidson LNK, Garani A, Halpin LR, Lipsen LPJ, Lee C *et al.* 2013. Threats to Canadian species at risk: An analysis of finalized recovery strategies. *Biological Conservation* 166: 254–265.

McCune J. 2016. Species distribution models predict rare species occurrences despite significant effects of landscape context. *Journal of Applied Ecology* 53: 1871–1879.

McGraw JB. 2001. Evidence for decline in stature of American ginseng plants from herbarium specimens. *Biological Conservation* 98: 25–32.

Mee JA, Bernatchez L, Reist JD, Rogers SM, Taylor EB. 2015. Identifying designatable units for intraspecific conservation prioritization: a hierarchical approach applied to the lake whitefish species complex (*Coregonus* spp.). *Evolutionary Applications* 8: 423–441.

Meineke EK, Davis CC, Davies TJ. 2018. The unrealized potential of herbaria for global change biology. *Ecological Monographs* 88: 505–525.

Meineke EK, Davies TJ, Daru BH, Davis CC. 2019. Biological collections for understanding biodiversity in the Anthropocene. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 374: 20170386.

Meineke E, Tomasi C, Yuan S, Pryer K. 2020a. Applying machine learning to investigate long-term insect–plant interactions preserved on digitized herbarium specimens. *Applications in Plant Sciences* 8: e11369.

Meineke EK, Davis CC, Davies TJ. 2020b. Response to Kozlov *et al.*: Inaccurate estimation of biases in herbarium specimen data. *bioRxiv*: 2020.2009.2001.278606.

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

Miller EA, Lisin SE, Smith CM, Houtan KSV. 2020. Herbaria macroalgae as a proxy for historical upwelling trends in Central California. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 287: 20200732.

Molnár VA, Sonkoly J, Lovas-Kiss Á, Fekete R, Takacs A, Somlyay L, Toeroek P. 2015. Seed of the threatened annual legume, *Astragalus contortuplicatus*, can survive over 130 years of dry storage. *Preslia* 87: 319–328.

Muniz AC, Lemos-Filho JP, Buzatti RSdO, Ribeiro PCC, Fernandes FM, Lovato MB. 2019. Genetic data improve the assessment of the conservation status based only on herbarium records of a Neotropical tree. *Scientific Reports* 9: 5693.

Nakahama N, Hirasawa Y, Minato T, Hasegawa M, Isagi Y, Shiga T. 2015. Recovery of genetic diversity in threatened plants through use of germinated seeds from herbarium specimens. *Plant Ecology* 216: 1635–1647.

Neubig K, Whitten W, Abbott J, Elliott S, Soltis D, Soltis P. 2014. Variables affecting DNA preservation in archival DNA specimens. In Applequist W, Campbell L, eds. *DNA banking in the 21st Century: Proceedings of the U.S. workshop on DNA banking.* St Louis, MO, USA: The William L. Brown Center at the Missouri Botanical Garden.

Nic Lughadha E, Walker BE, Canteiro C, Chadburn H, Davis AP, Hargreaves S, Lucas EJ, Schuiteman A, Williams E, Bachman SP *et al.* 2018. The use and misuse of herbarium specimens in evaluating plant extinction risks. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 374: 20170402.

Nic Lughadha EM, Graziele Staggemeier V, Vasconcelos TNC, Walker BE, Canteiro C, Lucas EJ. 2019. Harnessing the potential of integrated systematics for conservation of taxonomically complex, megadiverse plant groups. *Conservation Biology* 33: 511–522.

Nic Lughadha E, Bachman SP, Leão TCC, Forest F, Halley JM, Moat J, Acedo C, Bacon KL, Brewer RFA, Gâteblé G *et al.* 2020. Extinction risk and threats to plants and fungi. *Plants, People, Planet* 2: 389–408.

Nordin S, Baba MS, Itam K. 2011. *Plant taxonomic keys repository.* International Conference on Research and Innovation in Information Systems: IEEE, 1–6.

Nualart N, Ibáñez N, Soriano I, López-Pujol J. 2017. Assessing the relevance of herbarium collections as tools for conservation biology. *The Botanical Review* 83: 303–325.

Orsenigo S, Montagnani C, Fenu G, Gargano D, Peruzzi L, Abeli T, Alessandrini A, Bacchetta G, Bartolucci F, Bovio M *et al.* 2018. Red Listing plants under full national responsibility: Extinction risk and threats in the vascular flora endemic to Italy. *Biological Conservation* 224: 213–222.

Ott T, Palm C, Vogt R, Oberprieler C. 2020. GinJinn: An object-detection pipeline for automated feature extraction from herbarium specimens. *Applications in Plant Sciences* 8: e11351.

Park D, Breckheimer I, Williams A, Law E, Ellison A, Davis C. 2019. Herbarium specimens reveal substantial and unexpected variation in phenological sensitivity across the eastern United States. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 374: 20170394.

Park DS, Willis CG, Xi Z, Kartesz JT, Davis CC, Worthington S. 2020. Machine learning predicts large scale declines in native plant phylogenetic diversity. *New Phytologist* 227: 1544–1556.

Parsley KM. 2020. Plant awareness disparity: a case for renaming plant blindness. *Plants, People, Planet* 2: 598–601.

Paton A, Antonelli A, Carine M, Forzza RC, Davies N, Demissew S, Dröge G, Fulcher T, Grall A, Holstein N *et al.* 2020. Plant and fungal collections: current status, future perspectives. *Plants, People, Planet* 2: 499–514.

Pearson KD, Nelson G, Aronson MFJ, Bonnet P, Brenskelle L, Davis CC, Denny EG, Ellwood ER, Goëau H, Heberling JM *et al.* 2020. Machine learning using digitized herbarium specimens to advance phenological research. *BioScience* 70: 610–620.

Pfeiffer JM, Voeks RA. 2008. Biological invasions and biocultural diversity: linking ecological and cultural systems. *Environmental Conservation* 35: 281–293.

Popescu MC, Tache IA, Sandu C. 2015. The classification of a plant family based on morpho-fractal dimension. *19th International Conference on System Theory, Control and Computing (ICSTCC).* Cheile Gradistei, Romania: IEEE, 780–784.

Porteous G, Nesbitt M, Kendon JP, Prychid CJ, Stuppy W, Conejero M, Ballesteros D. 2019. Assessing extreme seed longevity: the value of historic botanical collections to modern research. *Frontiers in Plant Science* 10: 1181.

Posey DA, Dutfield G. 1996. *Beyond intellectual property: toward traditional resource rights for indigenous peoples and local communities.* Ottawa, Canada: International Development Research Centre.

Powell C, Motley J, Qin H, Shaw J. 2019. A born-digital field-to-database solution for collections-based research using collNotes and collBook. *Applications in Plant Sciences* 7: e11284.

Prather L, Alvarez-Fuentes O, Mayfield M, Ferguson C. 2004. The decline of plant collecting in the United States: a threat to the infrastructure of biodiversity studies. *Systematic Botany* 29: 15–28.

Pryer KM, Tomasi C, Wang X, Meineke EK, Windham MD. 2020. Using computer vision on herbarium specimen images to discriminate among closely related horsetails (*Equisetum*). *Applications in Plant Sciences* 8: e11372.

Przelomska NAS, Armstrong CG, Kistler L. 2020. Ancient plant DNA as a window into the cultural heritage and biodiversity of our food system. *Frontiers in Ecology and Evolution* 8: 74.

Pyke GH, Ehrlich PR. 2010. Biological collections and ecological/environmental research: a review, some observations and a look to the future. *Biological Reviews of the Cambridge Philosophical Society* 85: 247–266.

Raymond CV, Wen L, Cooke SJ, Bennett JR. 2018. National attention to endangered wildlife is not affected by global endangerment: A case study of Canada's species at risk program. *Environmental Science & Policy* 84: 74–79.

Rebelo AG, Holmes PM, Dorse C, Wood J. 2011. Impacts of urbanization in a biodiversity hotspot: conservation challenges in Metropolitan Cape Town. *South African Journal of Botany* 77: 20–35.

Reeb C, Kaandorp J, Jansson F, Puillandre N, Dubuisson J-Y, Cornette R, Jabbour F, Coudert Y, Patiño J, Flot J-F et al. 2018. Quantification of complex modular architecture in plants. *New Phytologist* 218: 859–872.

Reeves PA, Richards CM. 2011. Species delimitation under the general lineage concept: an empirical example using wild North American hops (Cannabaceae: *Humulus lupulus*). *Systematic Biology* 60: 45–59.

Reeves R, Brooks R. 1983. Metallophytes from mining areas of Central Europe. *Environmental Pollution (Series A)* 31: 277–285.

Riahi M, Zarre S, Maassoumi AA, Attar F, Kazempour OS. 2010. An inexpensive and rapid method for extracting papilionoid genomic DNA from herbarium specimens. *Genetics and Molecular Research* 9: 1334–1342.

Rivers MC, Taylor L, Brummitt NA, Meagher TR, Roberts DL, Lughadha EN. 2011. How many herbarium specimens are needed to detect threatened species? *Biological Conservation* 144: 2541–2547.

Roberts D, McInerny G. 2003. When is a species extinct? Quantitative inference of threat and extinction from herbarium data. *Lankesteriana* 7: 17–20.

Robinson D. 2010. *Confronting biopiracy: challenges, cases and international debates*. Abingdon, UK: Taylor & Francis.

Rodríguez Martín JA, De Arana C, Ramos-Miras JJ, Gil C, Boluda R. 2015. Impact of 70 years urban growth associated with heavy metal pollution. *Environmental Pollution* 196: 156–163.

Rojas M. 1992. The species problem and conservation: what are we protecting? *Conservation Biology* 6: 170–178.

Rønsted N, Grace O, Carine M. 2020. Editorial: Integrative and translational uses of herbarium collections across time, space, and species. *Frontiers in Plant Science* 11: 1319.

Rudin S, Murray D, Whitfeld T. 2017a. Retrospective analysis of heavy metal contamination in Rhode Island based on old and new herbarium specimens. *Applications in Plant Sciences* 5: 1600108.

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, Staats M, Richardson JE, Cowan RS, Bakker FT. 2012. How to open the treasure chest? Optimising DNA extraction from herbarium specimens. *PLoS ONE* 7: e43808.

Schneider JV, Negraschis V, Habersetzer J, Rabenstein R, Wesenberg J, Wesche K, Zizka G. 2018a. Taxonomic diversity masks leaf vein–climate relationships: lessons from herbarium collections across a latitudinal rainfall gradient in West Africa. *Botany Letters* 165: 384–395.

Schneider JV, Rabenstein R, Wesenberg J, Wesche K, Zizka G, Habersetzer J. 2018b. Improved non-destructive 2D and 3D X-ray imaging of leaf venation. *Plant Methods* 14: 7.

Schuettpelz E, Frandsen PB, Dikow RB, Brown A, Orli S, Peters M, Metallo A, Funk VA, Dorr LJ. 2017. Applications of deep convolutional neural networks to digitized natural history collections. *Biodiversity Data Journal* 5: e21139.

Sears J. 2011. Re-introduction of stinking hawk's-beard into South-East England, UK. In: Soorae PS, ed. *Global reintroduction perspectives: additional case-studies from around the globe*. Abu Dhabi, United Arab Emirates: International Union for Conservation of Nature Species Survival Commission, Re-introduction Specialist Group, 234–238.

Shapiro B. 2017. Pathways to de-extinction: how close can we get to resurrection of an extinct species? *Functional Ecology* 31: 996–1002.

Shepherd KA, Thiele KR, Sampson J, Coates D, Byrne M. 2015. A rare, new species of *Atriplex* (Chenopodiaceae) comprising two genetically distinct but morphologically cryptic populations in arid Western Australia: implications for taxonomy and conservation. *Australian Systematic Botany* 28: 234–245.

Shteir AB. 1997. Gender and "modern" botany in Victorian England. *Osiris* 12: 29–38.

Shteir A, Cayouette J. 2019. Collecting with "botanical friends": Four Women in Colonial Quebec and Newfoundland. *Scientia Canadensis* 41: 1–30.

Silva C, Besnard G, Piot A, Razanatsoa J, Oliveira RP, Vorontsova MS. 2017. Museomics resolve the systematics of an endangered grass lineage endemic to north-western Madagascar. *Annals of Botany* 119: 339–351.

Silva LAE. 2016. A data mining approach for standardization of collectors names in herbarium Database. *IEEE Latin America Transactions* 14: 805–810.

Soltis PS. 2017. Digitization of herbaria enables novel research. *American Journal of Botany* 104: 1281–1284.

Soltis PS, Nelson G, James SA. 2018. Green digitization: online botanical collections data answering real-world questions. *Applications in Plant Sciences* 6: e1028.

Soltis PS, Nelson G, Zare A, Meineke EK. 2020. Plants meet machines: prospects in machine learning for plant biology. *Applications in Plant Sciences* 8: e11371.

Spriggs EL, Schlutius C, Eaton DA, Park B, Sweeney PW, Edwards EJ, Donoghue MJ. 2019. Differences in flowering time maintain species boundaries in a continental radiation of *Viburnum*. *American Journal of Botany* 106: 833–849.

Staats M, Erkens RH, van de Vossenberg B, Wieringa JJ, Kraaijeveld K, Stielow B, Geml J, Richardson JE, Bakker FT. 2013. Genomic treasure troves: complete genome sequencing of herbarium and insect museum specimens. *PLoS ONE* 8: e69189.

Steinke KH. 2012. Improvement of Omnipage 18's efficiency. *IEEE International Conference on Computer Science and Automation Engineering (CSAE): IEEE*, 434–438.

Stepp JR. 2002. *Ethnobiology and biocultural diversity: proceedings*. Athens, GA, USA: University of Georgia Press.

Stepp JR, Thomas M. 2005. Managing ethnopharmacological data: Herbaria, relational data bases, literature. In: Etkin N, Elisabetsky E, eds. *Ethnopharmacology: encyclopedia of life support systems*. Oxford, UK: UNESCO, EOLSS.

Tang CQ, Dong Y-F, Herrando-Moraira S, Matsui T, Ohashi H, He L-Y, Nakao K, Tanaka N, Tomita M, Li X-S et al. 2017. Potential effects of climate change on geographic distribution of the Tertiary relict tree species *Davidia involucrata* in China. *Scientific Reports* 7: 43822.

Telewski FW, Zeevaart JA. 2002. The 120-yr period for Dr. Beal's seed viability experiment. *American Journal of Botany* 89: 1285–1288.

Theeten F, Marielle A, Vandenberghe T, Dillen M, Semal P, Scory S, Herpers J-M, Spiegel D, Mergen P, Smirnova L et al. 2019. NaturalHeritage: Bridging Belgian natural history collections. *Biodiversity Information Science and Standards* 3: e37854.

Thiers B. 2020. *The Worlds Herbaria 2019. A summary report based on data from Index Herbariorum*. New York, NY, USA: Index Hebariorum. [WWW document] URL http://sweetgum.nybg.org/science/docs/The_Worlds_Herbaria_2019.pdf [accessed 18 November 2020].

Tingley MW, Beissinger SR. 2009. Detecting range shifts from historical species occurrences: new perspectives on old data. *Trends in Ecology & Evolution* 24: 625–633.

Tressou J, Haevermans T. 2020. Predicting plant endemicity based on herbarium data: application to French data. 4. *Conference for the International Society of Nonparametric Statistics, vol. 4*. Salerno, Italy.

Triki A, Bouaziz B, Mahdi W, Gaikwad J. 2020. Objects detection from digitized herbarium specimen based on improved YOLO V3. In: *Proceedings of the 15th international joint conference on computer vision, imaging and computer graphics theory and applications, vol. 4*: VISAPP. Valetta, Malta, : 523–529.

Unger J, Merhof D, Renner S. 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.

Ungricht S, Rasplus J-Y, Kjellberg F. 2005. Extinction threat evaluation of endemic fig trees of New Caledonia: priority assessment for taxonomy and conservation with herbarium collections. *Biodiversity & Conservation* **14**: 205–232.

Václavík T, Beckmann M, Cord AF, Bindewald AM. 2017. Effects of UV-B radiation on leaf hair traits of invasive plants – Combining historical herbarium records with novel remote sensing data. *PLoS ONE* **12**: e0175671.

Vega FE, Smith LT, Davies NMJ, Moat J, Góral T, O'Sullivan R, Davis AP. 2019. Elucidation of hosts, native distribution, and habitat of the coffee berry borer (*Hypothenemus hampei*) using herbaria and other museum collections. *Frontiers in Plant Science* **10**: 1188.

Vellend M, Brown CD, Kharouba HM, McCune JL, Myers-Smith IH. 2013. Historical ecology: using unconventional data sources to test for effects of global environmental change. *American Journal of Botany* **100**: 1294–1305.

Villacis-Llobet J, Lucio-Troya M, Calvo-Navarro M, Calderon-Ramirez S, Mata-Montero E. 2019. A first glance into reversing senescence on herbarium sample images through conditional generative adversarial networks. In Crespo-Mariño J, Meneses-Rojas E, eds. *Latin American High Performance Computing Conference.* Cham, Switzerland: Springer.

Wäldchen J, Mäder P. 2018. Machine learning for image based species identification. *Methods in Ecology and Evolution* **9**: 2216–2225.

Walton S, Livermore L, Bánki O, Cubey RWN, Drinkwater R, Englund M, Goble C, Groom Q, Kermorvant C, Rey I et al. 2020. Landscape analysis for the specimen data refinery. *Research Ideas and Outcomes* **6**: e57602.

Wandeler P, Hoeck PEA, Keller LF. 2007. Back to the future: museum specimens in population genetics. *Trends in Ecology & Evolution* **22**: 634–642.

Wang Z, Dong H, Kelly M, Macklin JA, Morris PJ, Morris RA. 2009. Filtered-Push: a map-reduce platform for collaborative taxonomic data management. *WRI World Congress on Computer Science and Information Engineering: IEEE.* 731–735.

Watkins H, Cameron R, Sjöman H, Hitchmough J. 2020. Using big data to improve ecotype matching for Magnolias in urban forestry. *Urban Forestry & Urban Greening* **48**: 126580.

Weaver WN, Ng J, Laport RG. 2020. LeafMachine: Using machine learning to automate leaf trait extraction from digitized herbarium specimens. *Applications in Plant Sciences* **8**: e11367.

Weiß CL, Schuenemann VJ, Devos J, Shirsekar G, Reiter E, Gould BA, Stinchcombe JR, Krause J, Burbano HA. 2016. Temporal patterns of damage and decay kinetics of DNA retrieved from plant herbarium specimens. *Royal Society Open Science* **3**: 160239.

White AE, Dikow RB, Baugh M, Jenkins A, Frandsen PB. 2020. Generating segmentation masks of herbarium specimens and a data set for training segmentation models using deep learning. *Applications in Plant Sciences* **8**: e11352.

Wijesingha D, Marikar F. 2012. Automatic detection system for the identification of plants using herbarium specimen images. *Tropical Agricultural Research* **23**: 42–50.

Williams DR, Balmford A, Wilcove DS. 2020. The past and future role of conservation science in saving biodiversity. *Conservation Letters* **13**: e12720.

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

Willis CG, Law E, Williams AC, Franzone BF, Bernardos R, Bruno L, Hopkins C, Schorn C, Weber E, Park DS et al. 2017b. CrowdCurio: an online crowdsourcing platform to facilitate climate change studies using herbarium specimens. *New Phytologist* **215**: 479–488.

Willis F, Moat J, Paton A. 2003. Defining a role for herbarium data in Red List assessments: a case study of *Plectranthus* from eastern and southern tropical Africa. *Biodiversity & Conservation* **12**: 1537–1552.

Wolf A, Zimmerman NB, Anderegg WR, Busby PE, Christensen J. 2016. Altitudinal shifts of the native and introduced flora of California in the context of 20th-century warming. *Global Ecology and Biogeography* **25**: 418–429.

Wu P, Dietterich TG. 2004. Improving SVM accuracy by training on auxiliary data sources. *Proceedings of the twenty-first international conference on machine learning* doi: 10.1145/1015330.1015436.

Yagui MMM, Maia LFMP, Oliveira J, Vivacqua AS. 2017. Curation of physical objects in botany: architecture and development of a linked open data-based application. *IEEE 15th international conference on dependable, autonomic and secure computing, 15th international conference on pervasive intelligence and computing, 3rd international conference on big data intelligence and computing and cyber science and technology congress* (DASC/PiCom/DataCom/CyberSciTech) IEEE, 889–892.

Yan Y, Tang Z. 2019. Protecting endemic seed plants on the Tibetan Plateau under future climate change: migration matters. *Journal of Plant Ecology* **12**: 962–971.

Yost JM, Sweeney PW, Gilbert E, Nelson G, Guralnick R, Gallinat AS, Ellwood ER, Rossington N, Willis CG, Blum SD et al. 2018. Digitization protocol for scoring reproductive phenology from herbarium specimens of seed plants. *Applications in Plant Sciences* **6**: e1022.

Yost JM, Pearson KD, Alexander J, Gilbert E, Hains LA, Barry T, Bencie R, Bowler P, Carter B, Crowe RE et al. 2020. The California phenology collections network: using digital images to investigate phenological change in a biodiversity hotspot. *Madrono* **66**: 130–141.

Younis S, Weiland C, Hoehndorf R, Dressler S, Hickler T, Seeger B, Schmidt M. 2018. Taxon and trait recognition from digitized herbarium specimens using deep convolutional neural networks. *Botany Letters* **165**: 377–383.

Younis S, Schmidt M, Weiland C, Dressler S, Seeger B, Hickler T. 2020. Detection and annotation of plant organs from digitized herbarium scans using deep learning. *Biodiversity Data Journal* **8**: e57090.

Zedane L, Hong-Wa C, Murienne J, Jeziorski C, Baldwin BG, Besnard G. 2016. Museomics illuminate the history of an extinct, paleoendemic plant lineage (*Hesperelaea*, Oleaceae) known from an 1875 collection from Guadalupe Island, Mexico. *Biological Journal of the Linnean Society* **117**: 44–57.

Zizka A, Silvestro D, Vitt P, Knight TM. 2020. Automated conservation assessment of the orchid family with deep learning. *Conservation Biology.* doi: 10.1111/cobi.13616.

Zurell D, Franklin J, König C, Bouchet PJ, Dormann CF, Elith J, Fandos G, Feng X, Guillera-Arroita G, Guisan A et al. 2020. A standard protocol for reporting species distribution models. *Ecography* **43**: 1261–1277.