

1 Running head: Herbaria for understanding global change

2

3 The unrealized potential of herbaria for global change biology

4 Emily K. Meineke^{1,2*}, Charles C. Davis¹, T. Jonathan Davies²

5

6 ¹Department of Organismic and Evolutionary Biology, Harvard University Herbaria, 22 Divinity
7 Avenue, Cambridge, MA, USA 02138

8 ²Department of Biology, McGill University, 1205 Dr. Penfield Avenue, Montreal, QC H3A 1B1,
9 CA

10 *Emily K. Meineke, emily_meineke@fas.harvard.edu, phone: (+01) 919-308-3267, 22 Divinity
11 Avenue, Cambridge, MA, USA 02138

12

13 **Abstract**

14 Plant and fungal specimens in herbaria are becoming primary resources for investigating how
15 plant phenology and geographic distributions shift with climate change, greatly expanding
16 inferences across spatial, temporal, and phylogenetic dimensions. However, these specimens
17 contain a wealth of additional data—including nutrients, defensive compounds, herbivore
18 damage, disease lesions, and signatures of physiological processes—that capture ecological and
19 evolutionary responses to the Anthropocene but which are less frequently utilized. Here, we
20 outline the diversity of herbarium data, global change topics to which they have been applied,
21 and new hypotheses they could inform. We find that herbarium data have been used extensively
22 to study impacts of climate change and invasive species, but that such data are less commonly
23 used to address other drivers of biodiversity loss, including habitat conversion, pollution, and
24 overexploitation. In addition, we note that fungal specimens are under-explored relative to
25 vascular plants. To facilitate broader application of plant and fungal specimens in global change
26 research, we outline the limitations of these data and modern sampling and statistical tools that

27 may be applied to surmount challenges they present. Using a case study of insect herbivory, we
28 illustrate how novel herbarium data may be employed to test hypotheses for which few data
29 exist, despite potentially large biases. With the goal of positioning herbaria as hubs for global
30 change research, we suggest future research directions and curation priorities.

31

32 **Keywords:** climate change, invasive species, habitat conversion, extinction, museum specimens,
33 historical data, herbarium, global change

34

35 INTRODUCTION

36 A key challenge for biologists today is to determine how species are responding to the major
37 drivers of global change and biodiversity loss: habitat conversion and degradation, climate
38 change, invasive species, pollution, and overexploitation (Millennium Ecosystem Assessment
39 2005). Over the past decades, field observations and experiments have informed much of our
40 understanding of biological responses to these major drivers, particularly climate change.

41 However, like all scientific approaches, they have limitations. Experiments are almost always at
42 smaller spatial scales than inferences (e.g., Pelini *et al.*, 2011) and field observations are often
43 restricted to temperate biomes (Wolkovich *et al.*, 2012). Experiments and observations typically
44 only allow researchers to test hypotheses about a single driver of global change, while organisms
45 are generally exposed to many. Perhaps most critically, the majority of experiments and field
46 observations are short-term. Experiments addressing global change are commonly conducted for
47 a few years at most (e.g., Diamond *et al.*, 2012), limited by grant timelines and funding cycles.
48 Field observations are often designed to span latitude and altitude as a proxy for warming,
49 relying on the assumption that patterns across space will represent future patterns across time

50 (e.g., Kozlov *et al.*, 2013). While long-term field observations can span over a century in cases
51 where people across generations collect the same observations, these data are rare and only
52 available for a few phenomena (e.g., Keeling & Whorf 2005; Aono & Kazui 2008). Scientists
53 have increasingly turned to biological collections to expand data across time, space, and
54 taxonomy, thus better matching the scales at which recent global change is occurring (Pyke &
55 Ehrlich 2010). In particular, herbarium specimens, preserved (often pressed) plants and fungi,
56 have been the subject of a new wave of global change research.

57 The potential of herbarium data for global change biology stems in large part from its
58 temporal extent. Since the 1700s, scientists, including Linnaeus and Darwin, have collected
59 herbarium specimens to describe new species, aid taxonomic classifications, and as part of
60 regional floristic treatments (e.g., Moffett 2014). Although collecting has slowed in recent years
61 in many localities (Meyer *et al.*, 2016; e.g., Fig. 1a), scientists and amateurs continue to collect.
62 In many parts of the world, such as the northeastern US, the density of sampling extends far
63 beyond what is available from observations and experiments (Fig. 1b) and encompasses most
64 lineages of vascular plants (Fig. 1c), fungi, diatoms, and groups variously classified as algae.
65 Current estimates indicate that herbaria house over 350,000,000 specimens (Thiers 2016),
66 representing increasingly connected national (e.g., the Consortia of US Herbaria, Chinese Virtual
67 Herbarium [<http://www.cvh.ac.cn/news/8>]) and international networks of data (Fig. 1d; e.g.,
68 Australasian Virtual Herbarium [<https://avh.chah.org.au/>], iDigBio [<https://www.idigbio.org/>],
69 and GBIF [<https://www.gbif.org/>]).

70 Scientists today employ these specimens for purposes that could not have been imagined
71 by their collectors. Evolutionary biologists extract and sequence ancient DNA from herbarium
72 specimens to reconstruct phylogenetic relationships or infer population dynamics (Gugerli *et al.*,

73 2005; Wandeler *et al.*, 2007) and have documented selection on species traits through changes in
74 plant morphology (Kavanagh *et al.*, 2011; Burns *et al.*, 2012). More recently, ecologists
75 routinely use herbarium specimens as occurrence records for determining species distributions
76 (Graham *et al.*, 2004; Newbold 2010) and as records of leafing, flowering, and fruiting
77 phenology (Miller-Rushing *et al.*, 2006; Primack *et al.*, 2004; Bolmgren & Lonnberg 2005;
78 Everill *et al.*, 2014) to understand how these aspects of biology are influenced by rising global
79 temperatures. The value of such data is now well recognized (Lavoie & Lachance 2006; Pyke &
80 Ehrlich 2010; Johnson *et al.*, 2011; Lavoie 2013; Vellend *et al.*, 2013; Willis *et al.*, 2017a). The
81 unrealized potential of herbaria is perhaps most apparent in the opportunities presented by other
82 types of data that are only rarely extracted from specimens but which provide opportunities to
83 assess the effects of global change, including signatures of pollination efficiency, pollution
84 concentrations, physiological characteristics, nutrient concentrations, pathogen loads,
85 morphological and anatomical traits, genotypes, endophytes, and herbivory (Table 1). These
86 could allow researchers to address hypotheses about species and ecosystem responses to global
87 change, from shifts at the species level to changing environmental processes, including nutrient
88 cycles, changes in air quality, and biological control (Table 1).

89 Given the potential value of herbaria to global change research, it is perhaps surprising
90 that, with the exception of their use as records of species in occurrences and phenology, they
91 have not been more widely used. One explanation is that herbarium data present distinct
92 challenges not present in data gathered from experiments and field observations. Some of these
93 are associated with the vast and dispersed nature of herbarium collections. For example,
94 selecting appropriate focal taxa is critical in any ecological or evolutionary study. While
95 collections contain millions of specimens, they often lack easily accessible digital records. This

96 makes determining which taxa are well-represented within and across collections difficult. Other
97 challenges arise as a result of biases; plants are more frequently collected at certain times of year
98 to capture flowering or fruiting, and collection effort has been uneven across space and time
99 (Meyer *et al.*, 2016; Daru *et al.*, *In Press*). Challenges also arise because of preservation
100 artefacts; DNA degrades, specimens lose coloration over time, and insects often eat pressed
101 plants housed within museums. Extracting credible data from collections presents significant and
102 complex challenges, and, thus, much of their potential has remained untapped (Davis *et al.*,
103 2015; Meyer *et al.*, 2016).

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

118

119 **CONTRIBUTIONS OF HERBARIA TO GLOBAL CHANGE RESEARCH**

120 Interest in applying herbarium data for investigating global change has grown in recent decades.
121 While climate change has been the focus of most studies, use of herbarium data is slowly
122 permeating other areas of global change biology research. Here, we briefly review the ways in
123 which these data have been employed to examine biological responses to the five key drivers of
124 global change and biodiversity loss: climate change, habitat conversion and degradation,
125 invasive species, and—where studies are available—pollution, and overexploitation. We
126 categorize research addressing these drivers into three types of biotic responses: shifts in
127 distributions and population sizes, changes in physiology or morphology, and altered ecological
128 interactions. Within these categories, we consider each of the five key drivers of change ordered
129 by how thoroughly they have been investigated. We highlight particular studies that we consider
130 to be most innovative or that best represent a general approach and apologize in advance for not
131 being able to cite all the many excellent publications that have used herbarium data and which
132 have addressed a wide range of important and exciting questions.

133 **Shifts in distributions and population sizes**

134 Herbarium specimens typically include collection localities and therefore serve as occurrence
135 records that are now widely used to parameterize species distribution models and to understand
136 the effects of recent global change on species distributions (Feeley 2012; Calinger 2015;
137 D'Andrea *et al.*, 2009). For many species, herbarium specimens have revealed plant species
138 range-shifts both upwards in elevation and pole-wards in latitude in response to recent warming
139 (Feeley *et al.*, 2013), with some species' ranges contracting while others expand (Feeley 2012).
140 The use of herbarium specimens in this context is now widespread. Though most studies have
141 focused on terrestrial vascular plants, some have included algae (Riera *et al.*, 2015) and species

142 from aquatic systems (Wernberg *et al.*, 2011; Yaakub *et al.*, 2014), highlighting the taxonomic
143 and functional diversity represented in herbaria.

144 Herbaria occurrence records have also revealed the effects of habitat conversion on
145 species composition, ranges, and abundance. In developed countries, this research has focused on
146 urbanization, arguably the most profound form of global change in these regions (United Nations
147 2007). Herbarium specimens are among the only floristic records for these areas prior to
148 development and have revealed that urbanization drives native species declines across many
149 cities (e.g., Bertin 2002; DeCandido *et al.*, 2004; Dolan *et al.*, 2011; Gregor *et al.*, 2012; Celesti-
150 Grapow *et al.*, 2013). Herbaria can also provide insights into urban filters on plant communities
151 and have demonstrated that urbanization reduces the occurrence of species associated with
152 wetland habitats (Bertin 2002; DeCandido *et al.*, 2004; Dolan *et al.*, 2011; Gregor *et al.*, 2012;
153 Celesti-Grapow *et al.*, 2013). Unfortunately, while habitat loss through urbanization and
154 deforestation in remote, highly biodiverse regions, such as the wet tropics, is likely a major
155 driver of extinction (e.g., see Wearn *et al.*, 2012), herbarium specimens from many of these
156 regions are relatively sparse (Meyer *et al.*, 2016) and thus provide poor baselines from which to
157 derive estimates of biodiversity change (Feeley & Silman 2011). Nonetheless, herbarium data
158 can help identify species in decline and regions that may provide refuges (Farnsworth & Oğurcak
159 2006; Romeiras *et al.*, 2014), as well as capture the interactions between climate change and
160 habitat conversion, notably to demonstrate how human land use limits the extent to which
161 species can track their climatic niches (Feeley & Silman 2010).

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

165 Crawford & Hoagland 2009) and thus can help identify the mechanisms by which non-native
166 species have dispersed from continent-to-continent and expanded their geographic ranges.
167 Analyses of plant occurrence records from herbaria have revealed that human transportation
168 networks, notably paved roads and railroads, are important pathways for invasion (e.g., Barney
169 2006; Joly *et al.*, 2011). With recent advances in molecular techniques, it is now possible to map
170 the spread not just of species but also particular genotypes using herbarium data, as has been
171 illustrated for the common reed, *Phragmites australis* (Saltonstall 2002). These fine-scale
172 historical data can provide information on both the rate and direction of spread, revealing likely
173 centers of introduction and regions of high vulnerability.

174 **Physiological and morphological change**

175 Most species are unable to completely escape global change in space and thus must adapt or
176 acclimate *in situ* or risk extinction. Herbarium specimens can capture physiological and
177 morphological changes reflecting such responses, including shifts in morphology (Law & Salick
178 2005; Leger 2013), timing of life histories (Kharouba & Vellend 2015), and physiology (Miller-
179 Rushing *et al.*, 2009). In a small but growing number of studies, physiological and
180 morphological data, such as on plant leaf size (Guerin *et al.*, 2012), stomatal densities (Miller-
181 Rushing *et al.*, 2009), carbon and oxygen isotope measurements (Miller-Rushing *et al.*, 2009;
182 Bonal *et al.*, 2011), and specific leaf area (Reef & Lovelock 2014), have been extracted and
183 provide insights into changing photosynthetic rates and leaf palatability for herbivores, for
184 example.

185 The most studied of these shifts with herbarium data are changes in plant phenology.
186 Flower counts from herbarium specimens have revealed advances in peak flowering of
187 approximately 2.4 days for each degree Celsius rise in temperatures (Calinger *et al.*, 2013), and

188 similar advances due to the urban heat island effect (Primack *et al.*, 2004) that are particularly
189 pronounced in ephemeral species (Neil *et al.*, 2010). Researchers have also turned to herbaria to
190 identify potential cues driving plant phenology, revealing interactive effects of temperature,
191 precipitation, and latitude on the timing of flowering (Matthews & Mazer 2016). Such data have
192 proven extremely valuable as both a biotic index of climate change, and as a record of biotic
193 responses to climate warming. The extraction and application of phenological data from herbaria
194 has been reviewed extensively elsewhere (Miller-Rushing *et al.* 2006; Willis *et al.*, 2017a).

195 Plant phenology is just one response to changes in atmospheric chemistry. Concentrations
196 of pollutants in the atmosphere, including heavy metals, anthropogenic nitrogen, carbon dioxide
197 (CO₂), and other greenhouse gases have varied over time, with increases surprisingly early in
198 modern human history (Renberg *et al.*, 1994; Steffen *et al.*, 2007). However, because historical
199 data are sparse, and pollutant sources are often diffused, changes in pollution can be hard to track
200 using traditional ecological approaches. Herbarium specimens serve as records of pollutant
201 variation over space and time, helping link species' exposure to their responses. For example,
202 epiphytes, which accumulate atmospheric nitrogen, but do not uptake nitrogen from soil, can
203 serve as bioindicators of nitrogen pollution (Stewart *et al.*, 2002). Mosses similarly serve as
204 bioindicators of atmospheric metals (Weiss *et al.*, 1999), and lichens serve as bioindicators of
205 various pollutants, including a diversity of oxidants (Sigal & Nash 1983).

206 Data from herbaria can also be used to identify the traits associated with invasiveness. A
207 key question in invasion biology is why few species become invasive while most non-native
208 species remain at low abundances (Sakai *et al.*, 2001). By capturing physiological and
209 morphological data, herbarium studies have shown that adaptation to local conditions can
210 facilitate invasiveness (Vandepitte *et al.*, 2014) and, consistent with other types of data

211 (Wolkovich & Cleland 2011), that some invasive species are able to better track temperature and
212 thus take advantage of earlier springs resulting from warming (Calinger 2015).

213 **Shifts in Ecological Interactions**

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

223 The trace-record of interactions between plants and pollinators, herbivores, and
224 pathogens captured on herbarium specimens could be used to address a topic that has attracted
225 much interest in climate change research: potential phenological asynchronies and ecological
226 mismatches between associated species resulting from differential responses to warming climates
227 (Post *et al.*, 2008; Both *et al.*, 2009). Phenological asynchronies occur if the direction, rate, or
228 magnitude of change differs between associated species, if phenological responses differ in
229 space, or if responses within a single trophic level are varied, such that phenological tracking is
230 not possible (as discussed in detail elsewhere, e.g., Hegland *et al.*, 2009). Occurrence data from
231 herbaria and contemporaneous insect collections have suggested that climate change might lead
232 to asynchronies between some butterflies and their adult food plants (Kharouba & Vellend
233 2015). Other types of ecological mismatches can also disrupt mutualisms. Corolla tube length

234 data from herbarium specimens, along with corresponding bumble bee collections that provided
235 data on bee tongue length, showed that bees and flowers may be ecologically mismatched in
236 Colorado, USA; bee tongues lengths have decreased over time, while corolla tube lengths have
237 remained constant (Miller-Struttman *et al.*, 2015). However, there are few studies along these
238 lines, and we expand on the potential of herbarium data for studying phenological asynchronies
239 in the following section.

240 Similar data can serve as records of changing species interactions due to habitat
241 conversion, pollution, and invasions. Anthropogenic disturbance of a habitat can expose species
242 to novel conditions that disrupt co-evolved interactions. By rehydrating orchid flowers from
243 herbaria and counting the pollinaria (pollen clumps and associated tissues removed by bees),
244 Pauw & Hawkins (2011) demonstrated that the local decline of an orchid during urbanization
245 was driven by reduced pollination. Youngsteadt *et al.* (2015) counted scale insects on stems of
246 herbarium specimens and demonstrated that the urban heat island effect and natural warming
247 cycles in forests are associated with elevated abundance of a scale insect herbivore, suggesting
248 that ecological responses to warming in cities could predict the effects of climate change on
249 insect pests. In one of the few herbarium studies to explore how pollution affects species
250 interactions, Ziska *et al.* (2016) analyzed pollen preserved in herbarium specimens and
251 demonstrated that elevated CO₂ concentrations reduced pollen protein concentrations, altering
252 nutrient availability for pollinators, with possible fitness consequences for both pollinators and
253 plants. In a separate study, Zangerl & Berenbaum (2005) provided support for the natural enemy
254 release hypothesis of plant invasion (Keane & Crawley 2002; Mitchell & Power 2003) by
255 measuring exotic plant defensive compounds before and after introduction of their co-evolved
256 herbivores. They found evidence that invasive plants can escape their insect herbivores—and the

257 need to manufacture expensive defensive compounds in response to them—when first introduced
258 into new habitats.

259

260 **GAPS IN THE APPLICATION OF HERBARIUM DATA**

261 **Overview**

262 We have provided a brief overview of the diverse applications of herbarium data for
263 understanding biotic responses to global change. However, many of these data have only
264 recently become commonplace in ecological studies and we suggest their potential has not been
265 fully realized. Here, we identify significant gaps in current uses of herbarium specimens in
266 global change biology (Table 1). As we indicate above, data from herbaria have been used
267 extensively to explore plant distributional and phenological shifts in response to climate change.
268 However, herbarium specimens have been used less frequently to study plant responses to three
269 of the five key drivers of biodiversity loss: habitat conversion, pollution, and overexploitation.
270 By many estimates, habitat conversion currently drives most terrestrial species extinctions
271 (Millennium Ecosystem Assessment 2005), making this a particularly notable omission. With
272 these overarching themes in mind, we suggest future research directions we believe could benefit
273 from using herbarium data.

274 **Guidance for ecosystem management and restoration**

275 Herbarium specimens are rare records of historical biodiversity and thus could help guide
276 ecological restoration. A few studies have highlighted this potential, mostly within the context of
277 urban environments (DeCandido *et al.*, 2004; Atha *et al.*, 2016). However, less common are
278 examples of using herbarium data to explore other types of habitat conversion, such as

279 agricultural expansion and deforestation, much less integrate such findings into restoration
280 programs. As is the case for urbanization, these types of habitat transformations are likely to
281 have nuanced effects on local biodiversity that are not well understood but which may be
282 captured by collections. Herbarium specimens can additionally be used to determine areas and
283 species of conservation priority, as has been demonstrated for timber species in Angola
284 (Romeiras *et al.*, 2014). We suggest there may be many opportunities along these lines for
285 conservation of plants of economic and/or cultural importance, including both timber species and
286 wild crop relatives, which are often well-represented in herbaria, with many specimens collected
287 prior to intensification of global change.

288 **Markers of plant physiological change**

289 Specimens within herbaria provide more than just records of occurrence across space and time;
290 they also bear the imprint of past environments, including information on genotypic and
291 physiological shifts. For example, herbarium specimens could be employed to collect long-term
292 data on shifts in traits correlated with photosynthetic rates and nutrient concentrations across taxa
293 and habitats. These are ecosystem variables for which we have little historical data, but which
294 have important consequences. For example, leaf gas exchange rates can affect carbon balance
295 (Bonan 2008), and the net effects of global change, particularly climate change, on this and
296 similar ecosystem processes is a topic of great interest (Clark 2004; Wu *et al.*, 2011). Increasing
297 CO₂ levels are expected to increase photosynthetic rates in plants, increasing net carbon storage
298 in forests (Ainsworth & Long 2005). However, long-term experiments to evaluate the
299 relationship between CO₂ and photosynthetic rates have been in place for less than 20 years
300 (Norby *et al.*, 2016) and do not incorporate other recent global changes that may have interactive
301 effects on photosynthesis, most notably warming, changes in soil moisture, and nitrogen

302 deposition. While methods have been developed to extract such data from herbarium
303 specimens—for example, quantifying proxies for gas exchange preserved in leaves of pressed
304 plants (Miller-Rushing *et al.*, 2009; Bonal *et al.*, 2011)—and are not new (Woodward *et al.*
305 1987), they have much unrealized potential in global change research.

306 **Records of changing species interactions**

307 We suggest that herbarium specimens could provide unparalleled insights for understanding
308 shifting species interactions in the Anthropocene. Above, we outlined a few studies along these
309 lines, but we believe the data within herbarium specimens has not been fully exploited. For
310 example, presently, there is little consensus on whether climate change induced phenological
311 asynchronies are common or rare, in large part because data are sparse. It is, however, possible to
312 pair herbarium records with collections of species with which they are tightly associated, such as
313 their pollinators. This approach is potentially powerful, but requires substantial data. It is
314 possible, therefore, that for many species, collections will not yield data of sufficient spatial and
315 temporal resolution to quantify phenology for associated species. An alternative approach is to
316 draw indirect inferences on species interactions using data from only one partner; observations of
317 pollen, for example, can provide information on plant-pollinator interactions (Pauw & Hawkins
318 2011; Ziska *et al.*, 2016). Such methods can be more easily expanded to systems for which
319 corresponding collections on interactors are sparse or unavailable.

320 In a number of studies, researchers have used similar approaches to explore effects of
321 global change on antagonistic interactions between plants and their associates by extracting data
322 on herbivores, herbivory, and plant pathogens (Malmstrom *et al.*, 2007; Lees *et al.*, 2011;
323 Youngsteadt *et al.*, 2015; Syfert *et al.*, 2017). However, methods remain relatively
324 underdeveloped. Plant-insect interactions might be of particular interest, because insects are

325 ectothermic, and, thus, their abundances are expected to shift with global climate warming in
326 many cases (Kingsolver *et al.*, 2013). Herbivory is also ubiquitous among present day plants
327 (Turcotte *et al.*, 2014), has driven evolution of much of the biodiversity on earth (Futuyma &
328 Agrawal 2009), and has significant economic consequences in agriculture and forestry (Oerke &
329 Dehne 2004). Despite its importance as an ecological process, we are aware of few studies
330 quantifying plant-insect interactions on individual specimens (but see Morrow & Fox 1989).

331

332 **HERBARIA AS NOVEL DATA SOURCES: LIMITATIONS AND CHALLENGES**

333 We have shown that herbarium data can be applied to diverse topics and have suggested gaps
334 that warrant future exploration. Here, we outline challenges these data present and, when
335 available, approaches to reduce obstacles to their use, which we believe is a key hurdle to using
336 herbarium data in global change research. We consider some general methods that could be
337 applied widely to herbarium and biological collections data. Finally, we provide a table detailing
338 potential challenges posed by specific types of herbarium data, along with suggested ways to
339 approach these challenges (Table 2).

340 **Biases over space, time, and phylogeny**

341 Herbarium specimens are non-randomly collected across space and time, in part because their
342 historical purpose was to document species' ranges and capture morphological variation within
343 and between species, and not to address ecological questions. Using specimens for global change
344 research requires accounting for this unevenness in sampling over space, time, and taxonomy.
345 This topic has been reviewed recently by Meyer *et al.*, (2016), who analyze biases represented in
346 specimens and observational data, and Daru *et al.* (*In Press*), which includes biases represented

347 in herbarium specimens, specifically. Though the sampling biases described in these publications
348 should serve as a roadmap for those expected more generally in herbarium data, the gaps and
349 biases present across entire herbaria do not necessarily represent those in the subset of data
350 extracted for particular studies. For example, an herbarium may include few species from a
351 particular province in China, but have many specimens for each of these species.

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

370 Even when sampling procedures are designed carefully, however, it may not be possible to
371 select specimens evenly across all axes of variation. Numerous modern statistical tools are
372 available to account for uneven sampling and non-independence of data. These include, notably,
373 spatial regression techniques, such as spatial autoregressive models, spatial lag models, and
374 spatial error models (Plant 2012), that account for spatial non-independence in data. Various
375 phylogenetic comparative methods, including independent contrasts (Felsenstein 1985) and
376 phylogenetic generalized least squares regression (PGLS), are available for incorporating
377 phylogenetic non-independence among taxa. There is also a large literature on null models in
378 ecology (Gotelli & Graves 1996) that details methods for subsampling data to incorporate
379 potential biases into null models, thus holding bias constant when testing for relationships among
380 variables of interest. New machine learning tools and Bayesian techniques (notably in programs
381 such as Stan, WinBUGS, and JAGS) allow for the analyses of complex, hierarchically
382 structured, and incomplete datasets and are suited to analyzing large, sparsely sampled data, all
383 common features of collections data. Many of these methods are already well developed for
384 modelling species distributions—e.g., Maximum Entropy, Generalized Linear and Additive
385 Models, Boosted Regression Trees, and Random Forest (Elith & Leathwick 2009)— though they
386 are not yet integrated into ecology more generally (Thesen 2016).

387 **Data extraction and validation**

388 Once appropriate herbarium specimens are selected, depending on the analysis, it may then be
389 necessary to design sampling strategies to extract data from within individual herbarium sheets.
390 Like any other unit, such as a plot, agricultural field, or transect, individual herbarium specimens
391 can be subsampled to quantify features that may be hard to measure across entire specimens,
392 such as data on arthropod and pathogen associations, or other micro-traits including stomatal

393 density, cell shape etc. In some cases, it may be necessary to measure the surface area of
394 specimens to account for their size, a process that can be automated in programs such as ImageJ
395 (<https://imagej.nih.gov/ij/>), or to estimate specimen density, volume or other dimension to
396 standardize measures for differences in ‘sampling space’.

397 An important but frequently overlooked next step is to evaluate the reliability of derived
398 statistics. To characterize spatial and temporal biases, researchers have turned to data validation,
399 in which herbarium data are placed on common axes with trusted data. For example, several
400 recent studies have validated the utility of herbarium specimens for phenological research by
401 fitting common models to field observation and herbarium data (Robbirt *et al.*, 2011; Spellman
402 & Mulder 2016). By contrasting flowering data from herbaria to that from observations, Davis *et*
403 *al.* (2015) showed that herbarium specimens covered more climatic space than observations.
404 Observational data alone might thus provide a more limited estimate of future climate change
405 scenarios than herbarium specimens. This study highlights that traditional data sources also have
406 limitations and gaps in coverage, some of which can be alleviated by including data from
407 herbarium specimens. Similarly, pollution measures derived from herbarium specimens can be
408 calibrated by comparing pollutant concentrations to those from other historical data sources, such
409 as deposits in peat bogs and ice cores (Weiss *et al.*, 1999).

410 Cross validation with an independent dataset is perhaps the most robust approach for
411 detecting biases in data. However, in many cases, herbarium data cannot be directly compared to
412 independently derived data because no companion data exist, or the collection of such data
413 requires inordinate effort, e.g., the manual cleaning and standardization of multiple datasets.
414 When companion data are unavailable, we suggest another approach for assessing data
415 reliability: comparing summary statistics relevant to the hypothesis to be tested to theoretical

416 expectations from the literature. We demonstrate this approach focusing on insect herbivory, a
417 process that is likely to shift due to global change, but for which few historical data are available
418 (for detailed sampling and statistical methods, see Appendix S1). This is a particularly
419 challenging case study because collectors most likely select specimens that have little damage, so
420 absolute estimates of herbivory are probably strongly biased. Nonetheless, paleontologists have
421 used fossil herbivory, for which data are even more sparse, to assess changes in diversity and
422 abundance of herbivory with climatic changes across epochs (Wilf & Labandeira 1999).
423 Therefore, we suggest that the much higher resolution, more abundant data available from within
424 herbaria should provide at least as much information on contemporary herbivory patterns.

425 Despite potential biases, as a proof-in-concept example we demonstrate that herbarium
426 specimens harbor diverse types of herbivory damage (Fig. 2). We also provide evidence that they
427 could provide unique longitudinal data on plant-insect interactions (Fig. 3). Herbivory on
428 herbarium specimens has phylogenetic signal (Fig. 3a), with Blomberg's K value of 0.4 (above
429 zero indicates phylogenetic signal, with a value of one matching to Brownian motion), though
430 our dataset of 20 species precludes formal significance testing. Similarly, the composition of
431 herbivory, i.e., the relative amount of chewing, leaf galls, leaf mines, etc.—is significantly more
432 similar for individuals within than among genera and species (Fig. 3b). These taxonomic and
433 temporal patterns indicate that herbarium specimens may provide adequate data for addressing
434 questions related to how herbivore communities are changing over time and how herbivore
435 pressure varies across plant community members, questions that are of growing interest but for
436 which historical data are limited. In addition, herbivory measured on herbarium specimens is
437 highly skewed, with most specimens showing little or no damage, and a few showing heavy
438 damage (Fig. 3c), as also observed in field data (Turcotte *et al.*, 2014), and accumulates as the

439 growing season progressed (Fig. 3d), which suggests that these data are sensitive enough to
440 capture accumulating herbivory on individual plants through the growing season.

441 Collector bias toward intact specimens makes it difficult to infer the absolute magnitude
442 of herbivory from herbarium specimens. However, cross-validation with field data might allow
443 calibration of herbivory estimated from collections and allow researchers to quantify and correct
444 for under-estimates. Another limitation of these data is that we cannot infer herbivory rates, i.e.,
445 herbivory per day or month, without data on the timing of leaf-out, though degree-day data may
446 serve as a proxy and thus facilitate these estimates and associated inferences.

447 **Coarse, incomplete, or inaccurate metadata**

448 In some cases, specimens have been misidentified or taxonomic arrangements revised, but
449 herbarium records have not kept pace with changes. New databases and bioinformatics tools now
450 allow rapid taxonomic name synonymisation (e.g., the taxize R-package, [[https://cran.r-](https://cran.r-project.org/web/packages/taxize/index.html)
451 [project.org/web/packages/taxize/index.html](https://cran.r-project.org/web/packages/taxize/index.html)]). Judicious choices of exemplar species can also
452 help researchers avoid problematic taxa, as can sampling within herbaria with well-curated
453 metadata. However, in some cases, reviewing each specimen individually is necessary, and
454 herbarium specimens are themselves the key resource for resolving species identities. In contrast
455 to historical observations, for which it is usually impossible to verify species identifications,
456 herbarium specimens can always be revisited. However, numerous groups likely remain poorly
457 described in herbaria, and despite rapid efforts to mobilize collections online, the resolution of
458 digital images may not be sufficient to provide definitive identification, especially in taxonomic
459 groups with many closely related and morphological similar species, such as in the graminoids.

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

472 **Effects of preservation**

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

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

491

492 **NEXT GENERATION HERBARIA**

493 We have discussed how herbarium collections can contribute to global change research, and we
494 have described some simple approaches for sampling and analyzing data from herbaria, helping
495 overcome a key barrier to the use of herbarium specimens in ecology. We suggest consideration
496 of sampling strategies which are used in ecology more generally, the adoption of statistical
497 methods to help account for gaps in coverage, and new computational tools, such as those that
498 are appropriate for sparsely sampled data. We emphasize the importance of cross-validating data
499 extracted from herbarium specimens, and suggest how data reliability can be assessed even when
500 companion data are not available. We also identified gaps in the literature that point to future
501 research opportunities. Here, we detail new frontiers for using data derived from herbaria in
502 global change research.

503 Herbarium specimens could provide unprecedented amounts of data on fungal responses
504 to global change. While herbaria house millions of fungal specimens, these records are rarely
505 integrated into global change research. This presents opportunities for investigating fungal

506 responses to global change in ways that parallel ongoing efforts using plants. Though examples
507 are sparse, herbaria studies have shown increasing incidence of fungal plant diseases
508 (Antonovics *et al.*, 2003) and effects of climate on fungal phenology (Kausserud *et al.*, 2008;
509 Kausserud *et al.*, 2010; Diez *et al.*, 2013). Researchers could draw methods from these studies to
510 more deeply investigate fungal responses, which will be critical determinants of ecosystem
511 function in the future.

512 Plant and fungal data from herbaria, such as those described in Tables 1 and 2, could be
513 placed on common axes with data from traditional sources—field observations, experiments, and
514 fossils—to generate more robust predictions of how species will respond to drivers of global
515 change (Davis *et al.*, 2015; Youngsteadt *et al.*, 2015). Because all of these methods introduce
516 biases and have limitations, the most robust inferences would include data from multiple
517 approaches, and herbarium data could be uniquely suited in some cases to inform outstanding
518 debates about global change introduced by more traditional approaches. For example,
519 observations and experiments can yield different results about ecologically important
520 phenomena, such as phenological responses to warming (Wolkovich *et al.*, 2012). Herbarium
521 data may allow researchers to resolve such discrepancies, in part, because specimens can capture
522 both long-term processes, including evolution, and short-term processes, such as plastic
523 responses, that are generally not represented together in either observations or experiments. In
524 addition, herbarium records and experiments could be used together to tease apart mechanisms
525 that drive species responses, and herbarium data could be combined with data from fossils (as
526 other modern data have been compared to fossils, as described in Labandeira & Currano 2013) to
527 contrast effects of current and historical climate change by placing temperatures across epochs
528 and recent time on the same axes. While patterns and dynamics might differ due to radically

529 different timescales of these data, such differences might themselves reveal important insights
530 into universal drivers of and responses to global change (e.g., increased CO₂ levels and
531 warming).

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

551 Plant and fungal collections continue to grow, but the goals of herbaria frequently remain
552 unaligned with their ecological utility. There is a need to better integrate current research
553 demands into collecting methods and collection management. Therefore, we suggest a shift in
554 how herbaria operate to help maximize contributions of herbaria to global change research and to
555 center herbaria as key repositories of ecological data. We propose several key strategies to move
556 towards this goal, the last three of which are potentially the most resource intensive:

557 **1) Organize specimens by evolutionary history (phylogeny) and spatial location.**

558 Collections are most accessible when they are organized by taxonomy and specimen location
559 because researchers almost always sample within clades and areas. Following a more
560 phylogenetically oriented classification and ordering rather than outdated classifications that
561 do not align with current understanding of phylogeny might facilitate ease of usage (see:
562 [[http://www.bvaenviro.com/Public/Angiosperm%20Phylogeny%20Group%20III/APG%20II](http://www.bvaenviro.com/Public/Angiosperm%20Phylogeny%20Group%20III/APG%20II%20linear%20list.pdf)
563 [I%20linear%20list.pdf](http://www.bvaenviro.com/Public/Angiosperm%20Phylogeny%20Group%20III/APG%20II%20linear%20list.pdf)]).

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

- 572 3) **Retain or, at minimum, digitize damaged specimens.** Curators sometimes dispose of
573 specimens that were damaged by insects or pathogens before they were collected. They do
574 this for good reasons; herbaria have limited space and specimen storage comes at a cost.
575 However, damaged specimens contain valuable information on locations and about species
576 interactions. For suggestions on sterile ecological vouchers see below.
- 577 4) **Collect and preserve tissue for future molecular and chemical analyses.** Curators at many
578 herbaria now collect and preserve tissue samples, in addition to pressed and dried specimens,
579 explicitly for future molecular analyses. These samples can be dried and stored with silica gel
580 and preserved at room temperature or, more ideally, cryobanked for future extraction. Such
581 infrastructure is also invaluable for preserving RNA, which is essential for investigating gene
582 expression but is often degraded rapidly. These materials are best frozen as soon as possible
583 but field fixatives are now available to reduce the burden of collecting such samples. Recent
584 efforts in tissue banking (see, e.g., [<https://frozenark.org/>],
585 [http://www.ggbn.org/ggbn_portal/]) provide a useful guide on how such approaches could
586 be implemented.
- 587 5) **Add sampling information to specimen labels.** Specimens are most useful when labels
588 include metadata, which are recorded regularly as part of basic collecting protocol, i.e., date,
589 geolocation, species. We suggest that the type of sampling and habitat should also be
590 indicated as standard metadata. For example, one system would be to indicate sampling as
591 targeted or opportunistic, and if targeted, to indicate the intention of the collection, e.g., to
592 document galls on the specimen or to get a collection from a particular area. Indication of a
593 specimen's immediate and adjacent habitats, e.g., along a roadside, in an urban or

594 transformed environment, or in a forest, would also help ecologists to determine if specimens
595 were collected in appropriate ways for particular studies.

596 **6) Develop protocols for storing ecological vouchers.** Traditionally, herbaria have focused on
597 maximizing number and morphological diversity within species across their ranges.
598 However, because of increasing interest in historical ecology, there is a need to prioritize
599 specimens that represent ecological effects (Baker *et al.*, 2017). Space is an ongoing issue in
600 herbaria, and ecological vouchers are especially challenging to store, as specimens from a
601 single study can number in the thousands. We suggest that herbaria develop new protocols
602 that allow ecologists to deposit whole digital specimens and small physical samples from
603 plants in their studies. One example might be to reduce these collections in their physical size
604 and to minimally retain sufficient tissue for subsequent DNA extraction. In concert with
605 rapidly developing DNA barcoding methods (see also below), this approach would facilitate
606 the reconstruction of historical patterns of plant community assembly.

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

616 8) **Create barcode libraries.** As molecular sequencing costs have declined, it is now more
617 reasonable to create digital repositories of genetic data using species-specific markers,
618 allowing samples to be matched to collections using both molecules and morphology. The
619 former may be especially valuable when plant samples lack key identifying traits, such as
620 flowers or leaves. Although controversial, DNA barcoding efforts, such as the Barcode of
621 Life (<http://www.barcodeoflife.org/>), which now holds over 5 million barcode sequences,
622 illustrate the potential of such approaches.

623

624 **CONCLUSIONS**

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

638 the current use of collections while continuing high standards of preservation to benefit future
639 generations.

640

641 **ACKNOWLEDGEMENTS**

642 We thank Barbara Hanrahan, Anthony Brach, and Daniel Park for their knowledge of and
643 assistance within the Harvard University Herbaria. We thank Charlie Willis for providing expert
644 advice and data from Willis *et al.* (2008). We thank Matt Bertone, David Boufford, and Charley
645 Eiseman for their natural history expertise.

646 **Funding statement**

647 This project was supported by a Discovery Grant from The Natural Sciences and Engineering
648 Research Council of Canada (http://www.nserc-crsng.gc.ca/index_eng.asp) to TJD. This material
649 is based upon work supported by the National Science Foundation Postdoctoral Research
650 Fellowship in Biology under Grant No. (1611880) to EKM. Any opinions, findings, and
651 conclusions or recommendations expressed in this material are those of the authors and do not
652 necessarily reflect the views of the National Science Foundation. This work was also supported
653 by the Harvard University Herbaria and the Department of Organismal and Evolutionary Biology
654 at Harvard.

655

656

657

658

659

660

661

662

663

664 **LITERATURE CITED**

665

666 Agrawal, A.A. & Fishbein, M. (2008). Phylogenetic escalation and decline of plant defense
667 strategies. *Proceedings of the National Academy of Sciences of the United States of America*,
668 **105**, 10057–10060.

669

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

673

674 Ainsworth, E.A. & Long, S.P. (2005). What have we learned from 15 years of free-air CO₂
675 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy. *New*
676 *Phytologist*, **165**, 351–371.

677

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

681

682 Atha, D.E., Forrest, T., Naczi, R.F.C., Pace, M.C., Rubin, M., Schuler, J.A. *et al.* (2016). The
683 historic and extant spontaneous vascular flora of The New York Botanical Garden. *Brittonia*, **68**,
684 245–277.

685

686 Baker, T.R., Pennington, R. T., Dexter, K. G., Fine, P. V., Fortune-Hopkins, H., Honorio, E. N.,
687 Huamantupa-Chuquimaco, I., Klitgård, B. B., Lewis, G. P., de Lima, H. C. and Ashton, P.
688 (2017). Maximising Synergy among Tropical Plant Systematists, Ecologists, and Evolutionary
689 Biologists. *Trends in Ecology & Evolution*, **32**, 258–267.

690

691 Bale, J.S., Masters, G.J., Hodkinson, I.D., Awmack, C., Bezemer, T.M., Brown, V.K. *et al.*
692 (2002). Herbivory in global climate change research: direct effects of rising temperature on
693 insect herbivores. *Global Change Biology*, **8**, 1–16.

694

695 Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012). Selecting pseudo-absences for
696 species distribution models: how, where and how many? *Methods in Ecology and Evolution*, **3**,
697 327–338.

698

699 Barney, J.N. (2006). North American history of two invasive plant species: phylogeographic
700 distribution, dispersal vectors, and multiple introductions. *Biological Invasions*, **8**, 703–717.

701

702 Bertin, R.I. (2002). Losses of native plant species from Worcester, Massachusetts. *Rhodora*, **104**,
703 325–349.

704

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

707

708 Bonal, D., Ponton, S., Le Thiec, D., Richard, B., Ningre, N., Hérault, B. *et al.* (2011). Leaf
709 functional response to increasing atmospheric CO₂ concentrations over the last century in two

- 710 northern Amazonian tree species: A historical $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ approach using herbarium
711 samples. *Plant, Cell and Environment*, **34**, 1332–1344.
712
- 713 Bonan, G.B. (2008). Forests and climate change: Forcings, feedbacks, and the climate benefits of
714 forests. *Science*, **320**, 1444–1449.
715
- 716 Both, C., van Asch, M., Bijlsma, R.G., van den Burg, A.B. & Visser, M.E. (2009). Climate
717 change and unequal phenological changes across four trophic levels: constraints or adaptations?
718 *Journal of Animal Ecology*, **78**, 73–83.
719
- 720 Burns, K.C., Herold, N. & Wallace, B. (2012). Evolutionary size changes in plants of the south-
721 west Pacific. *Global Ecology and Biogeography*, **21**, 819–828.
722
- 723 Buswell, J.M., Moles, A.T. & Hartley, S. (2011). Is rapid evolution common in introduced plant
724 species? *Journal of Ecology*, **99**, 214–224.
725
- 726 Calinger, K.M. (2015). A functional group analysis of change in the abundance and distribution
727 of 207 plant species across 115 years in north-central North America. *Biodiversity and
728 Conservation*, **24**, 2439–2457.
729
- 730 Calinger, K.M., Queenborough, S. & Curtis, P.S. (2013). Herbarium specimens reveal the
731 footprint of climate change on flowering trends across north-central North America. *Ecology
732 Letters*, **16**, 1037–1044.
733
- 734 Celesti-Grapow, L., Capotorti, G., Del Vico, E., Lattanzi, E., Tilia, A. & Blasi, C. (2013). The
735 vascular flora of Rome. *Plant Biosystems - An International Journal Dealing with all Aspects of
736 Plant Biology*, **147**, 1059–1087.
737
- 738 Clark, D.A. (2004). Tropical forests and global warming: slowing it down or speeding it up?
739 *Frontiers in Ecology and the Environment*, **2**, 73–80.
740
- 741 Corney, D., Clark, J.Y., Tang, H.L. & Wilkin, P. (2012a). Automatic extraction of leaf characters
742 from herbarium specimens. *Taxon*, **61**, 231–244.
743
- 744 Corney, D.P., Tang, H.L., Clark, J.Y., Hu, Y. & Jin, J. (2012b). Automating digital leaf
745 measurement: the tooth, the whole tooth, and nothing but the tooth. *PLOS ONE*, **7**, e42112.
746
- 747 Crawford, P.H.C. & Hoagland, B.W. (2009). Can herbarium records be used to map alien species
748 invasion and native species expansion over the past 100 years? *Journal of Biogeography*, **36**,
749 651–661.
750
- 751 D'Andrea, L., Broennimann, O., Kozłowski, G., Guisan, A., Morin, X., Keller-Senften, J. *et al.*
752 (2009). Climate change, anthropogenic disturbance and the northward range expansion of
753 *Lactuca serriola* (Asteraceae). *Journal of Biogeography*, **36**, 1573–1587.
754

- 755 Daru, Barnabas H, Daniel S Park, Richard Primack, Charles G Willis, David S Barrington,
756 Timothy JS Whitfeld, Tristram G Seidler, Patrick W Sweeney, David R Foster, Aaron M Ellison,
757 and C. C. Davis. In press. Widespread sampling biases in herbaria revealed from large-scale
758 digitization. *New Phytologist*.
759
- 760 Davis, C.C., Willis, C.G., Connolly, B., Kelly, C. & Ellison, A.M. (2015). Herbarium records are
761 reliable sources of phenological change driven by climate and provide novel insights into
762 species' phenological cueing mechanisms. *American Journal of Botany*, **102**, 1599–1609.
763
- 764 DeCandido, R., Muir, A.A. & Gargiullo, M.B. (2004). A first approximation of the historical and
765 extant vascular flora of New York City: Implications for native plant species conservation.
766 *Journal of the Torrey Botanical Society*, **131**, 243–251.
767
- 768 Diamond, S.E., Nichols, L.M., McCoy, N., Hirsch, C., Pelini, S.L., Sanders, N.J. *et al.* (2012). A
769 physiological trait-based approach to predicting the responses of species to experimental climate
770 warming. *Ecology*, **93**, 2305–2312.
771
- 772 Diez, J.M., James, T.Y., McMunn, M. & Ibanez, I. (2013). Predicting species-specific responses
773 of fungi to climatic variation using historical records. *Global Change Biology*, **19**, 3145–3154.
774
- 775 Dolan, R.W., Moore, M.E. & Stephens, J.D. (2011). Documenting effects of urbanization on
776 flora using herbarium records. *Journal of Ecology*, **99**, 1055–1062.
777
- 778 Elith, J. & Leathwick, J.R. (2009). Species Distribution Models: Ecological Explanation and
779 Prediction Across Space and Time. In: *Annual Review of Ecology Evolution and Systematics*, pp.
780 677–697.
781
- 782 Everill, P.H., Primack, R.B., Ellwood, E.R. & Melaas, E.K. (2014). Determining past leaf-out
783 times of New England's deciduous forests from herbarium specimens. *American Journal of*
784 *Botany*, **101**, 1293–1300.
785
- 786 Farnsworth, E.J. & Ogurcak, D.E. (2006). Biogeography and decline of rare plants in New
787 England: Historical evidence and contemporary monitoring. *Ecological Applications*, **16**, 1327–
788 1337.
789
- 790 Feeley, K.J. (2012). Distributional migrations, expansions, and contractions of tropical plant
791 species as revealed in dated herbarium records. *Global Change Biology*, **18**, 1335–1341.
792
- 793 Feeley, K.J., Hurtado, J., Saatchi, S., Silman, M.R. & Clark, D.B. (2013). Compositional shifts in
794 Costa Rican forests due to climate-driven species migrations. *Global Change Biology*, **19**, 3472–
795 3480.
796
- 797 Feeley, K.J. & Silman, M.R. (2010). Land-use and climate change effects on population size and
798 extinction risk of Andean plants. *Global Change Biology*, **16**, 3215–3222.
799

- 800 Feeley, K.J. & Silman, M.R. (2011). Keep collecting: accurate species distribution modelling
801 requires more collections than previously thought. *Diversity and Distributions*, **17**, 1132–1140.
802
- 803 Felsenstein, J. (1985). Phylogenies and the comparative method. *The American Naturalist*, **125**,
804 1–15.
805
- 806 Fornoni, J. (2011). Ecological and evolutionary implications of plant tolerance to herbivory.
807 *Functional Ecology*, **25**, 399–407.
808
- 809 Futuyma, D.J. & Agrawal, A.A. (2009). Macroevolution and the biological diversity of plants
810 and herbivores. *Proceedings of the National Academy of Sciences of the United States of*
811 *America*, **106**, 18054–18061.
812
- 813 Gilman, S.E., Urban, M.C., Tewksbury, J., Gilchrist, G.W. & Holt, R.D. (2010). A framework
814 for community interactions under climate change. *Trends in Ecology & Evolution*, **25**, 325–331.
815
- 816 Gotelli, N.J. & Graves, G.R. (1996). Null models in ecology. Smithsonian Institution Press.
817
- 818 Graham, C.H., Ferrier, S., Huettman, F., Moritz, C. & Peterson, A.T. (2004). New developments
819 in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology &*
820 *Evolution*, **19**, 497–503.
821
- 822 Gregor, T., Bönsel, D., Starke-Ottich, I. & Zizka, G. (2012). Drivers of floristic change in large
823 cities - A case study of Frankfurt/Main (Germany). *Landscape and Urban Planning*, **104**, 230–
824 237.
825
- 826 Guerin, G.R., Wen, H. & Lowe, A.J. (2012). Leaf morphology shift linked to climate change.
827 *Biology Letters*, **8**, 882–886.
828
- 829 Gugerli, F., Parducci, L. & Petit, R.J. (2005). Ancient plant DNA: review and prospects. *New*
830 *Phytologist*, **166**, 409–418.
831
- 832 Harris, L.W. & Davies, T.J. (2016). A Complete Fossil-Calibrated Phylogeny of Seed Plant
833 Families as a Tool for Comparative Analyses: Testing the ‘Time for Speciation’ Hypothesis.
834 *PLOS ONE*, **11**, e0162907.
835
- 836 Hegland, S.J., Nielsen, A., Lazaro, A., Bjerknes, A.L. & Totland, O. (2009). How does climate
837 warming affect plant-pollinator interactions? *Ecology Letters*, **12**, 184–195.
838
- 839 Johnson, K.G., Brooks, S.J., Fenberg, P.B., Glover, A.G., James, K.E., Lister, A.M. *et al.* (2011).
840 Climate Change and Biosphere Response: Unlocking the Collections Vault. *Bioscience*, **61**, 147–
841 153.
842
- 843 Joly, M., Bertrand, P., Gbangou, R.Y., White, M.C., Dubé, J. & Lavoie, C. (2011). Paving the
844 way for invasive species: Road type and the spread of Common ragweed (*Ambrosia*
845 *artemisiifolia*). *Environmental Management*, **48**, 514–522.

- 846
847 Kauserud, H., Heegaard, E., Semenov, M.A., Boddy, L., Halvorsen, R., Stige, L.C. *et al.* (2010).
848 Climate change and spring-fruited fungi. *Proceedings of the Royal Society B-Biological*
849 *Sciences*, **277**, 1169–1177.
850
851 Kauserud, H., Stige, L.C., Vik, J.O., Okland, R.H., Hoiland, K. & Stenseth, N.C. (2008).
852 Mushroom fruiting and climate change. *Proceedings of the National Academy of Sciences of the*
853 *United States of America*, **105**, 3811–3814.
854
855 Kavanagh, P.H., Lehnebach, C.A., Shea, M.J. & Burns, K.C. (2011). Allometry of Sexual Size
856 Dimorphism in Dioecious Plants: Do Plants Obey Rensch's Rule? *American Naturalist*, **178**,
857 596–601.
858
859 Keane, R.M. & Crawley, M.J. (2002). Exotic plant invasions and the enemy release hypothesis.
860 *Trends in Ecology & Evolution*, **17**, 164–170.
861
862 Keeling, C. & Whorf, T. (2005). Atmospheric carbon dioxide record from Mauna Loa. *Carbon*
863 *Dioxide Research Group, Scripps Institution of Oceanography, University of California La Jolla,*
864 *California*, 92093–90444.
865
866 Kharouba, H.M. & Vellend, M. (2015). Flowering time of butterfly nectar food plants is more
867 sensitive to temperature than the timing of butterfly adult flight. *Journal of Animal Ecology*, **84**,
868 1311–1321.
869
870 Kingsolver, J.G., Diamond, S.E. & Buckley, L.B. (2013). Heat stress and the fitness
871 consequences of climate change for terrestrial ectotherms. *Functional Ecology*, **27**, 1415–1423.
872
873 Knepp, R.G., Hamilton, J.G., Mohan, J.E., Zangerl, A.R., Berenbaum, M.R. & DeLucia, E.H.
874 (2005). Elevated CO₂ reduces leaf damage by insect herbivores in a forest community. *New*
875 *Phytologist*, **167**, 207–218.
876
877 Kozlov, M.V., van Nieuwerkerken, E.J., Zverev, V. & Zvereva, E.L. (2013). Abundance and
878 diversity of birch-feeding leafminers along latitudinal gradients in northern Europe. *Ecography*,
879 **36**, 1138–1149.
880
881 Labandeira, C.C. & Currano, E.D. (2013). The Fossil Record of Plant-Insect Dynamics. In:
882 *Annual Review of Earth and Planetary Sciences, Vol 41* (ed. Jeanloz, R), pp. 287–+.
883
884 Lavoie, C. (2013). Biological collections in an ever changing world: Herbaria as tools for
885 biogeographical and environmental studies. *Perspectives in Plant Ecology Evolution and*
886 *Systematics*, **15**, 68–76.
887
888 Lavoie, C., Jodoin, Y. & de Merlis, A.G. (2007). How did common ragweed (*Ambrosia*
889 *artemisiifolia* L.) spread in Quebec? A historical analysis using herbarium records. *Journal of*
890 *Biogeography*, **34**, 1751–1761.
891

- 892 Lavoie, C. & Lachance, D. (2006). A new herbarium-based method for reconstructing the
893 phenology of plant species across large areas. *American Journal of Botany*, **93**, 512–516.
894
- 895 Law, W. & Salick, J. (2005). Human-induced dwarfing of Himalayan snow lotus, *Saussurea*
896 *laniceps* (Asteraceae). *Proceedings of the National Academy of Sciences of the United States of*
897 *America*, **102**, 10218–10220.
898
- 899 Lees, D.C., Lack, H.W., Rougerie, R., Hernandez-Lopez, A., Raus, T., Avtzis, N.D. *et al.* (2011).
900 Tracking origins of invasive herbivores through herbaria and archival DNA: The case of the
901 horse-chestnut leaf miner. *Frontiers in Ecology and the Environment*, **9**, 322–328.
902
- 903 Leger, E.A. (2013). Annual plants change in size over a century of observations. *Global Change*
904 *Biology*, **19**, 2229–2239.
905
- 906 Lloyd, J. & Farquhar, G.D. (2008). Effects of rising temperatures and CO₂ on the physiology of
907 tropical forest trees. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **363**,
908 1811–1817.
909
- 910 Mack, R.N., Simberloff, D., Lonsdale, W.M., Evans, H., Clout, M. & Bazzaz, F.A. (2000).
911 Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecological*
912 *Applications*, **10**, 689–710.
913
- 914 Malmstrom, C.M., Shu, R., Linton, E.W., Newton, L.A. & Cook, M.A. (2007). Barley yellow
915 dwarf viruses (BYDVs) preserved in herbarium specimens illuminate historical disease ecology
916 of invasive and native grasses. *Journal of Ecology*, **95**, 1153–1166.
917
- 918 Matthews, E.R. & Mazer, S.J. (2016). Historical changes in flowering phenology are governed
919 by temperature x precipitation interactions in a widespread perennial herb in western North
920 America. *New Phytologist*, **210**, 157–167.
921
- 922 Meineke, E.K., Dunn, R.R., Sexton, J.O. & Frank, S.D. (2013). Urban warming drives insect
923 pest abundance on street trees. *PLOS ONE*, **8**, e59687.
924
- 925 Meyer, C., Weigelt, P. & Kreft, H. (2016). Multidimensional biases, gaps and uncertainties in
926 global plant occurrence information. *Ecology Letters*, **19**, 992–1006.
927
- 928 Millennium Ecosystem Assessment, 2005. Ecosystems and Human Well-being: Synthesis. Island
929 Press, Washington, DC.
930
- 931 Miller-Rushing, A.J., Primack, R.B., Templer, P.H., Rathbone, S. & Mukunda, S. (2009). Long-
932 term relationships among atmospheric CO₂, stomata, and intrinsic water use efficiency in
933 individual trees. *American Journal of Botany*, **96**, 1779–1786.
934
- 935 Miller-Rushing, A.J., Primack, R.B., Primack, D. & Mukunda, S. (2006). Photographs and
936 herbarium specimens as tools to document phenological changes in response to global warming.
937 *American Journal of Botany*, **93**, 1667–1674.

- 938
939 Miller-Struttman, N.E., Geib, J.C., Franklin, J.D., Kevan, P.G., Holdo, R.M., Ebert-May, D. *et*
940 *al.* (2015). Functional mismatch in a bumble bee pollination mutualism under climate change.
941 *Science*, **349**, 1541-1544.
942
943 Mitchell, C.E. & Power, A.G. (2003). Release of invasive plants from fungal and viral
944 pathogens. *Nature*, **421**, 625–627.
945
946 Moffett, R. (2014). *A biographical dictionary of contributors to the natural history of the Free*
947 *State and Lesotho*. SUN MeDIA Bloemfontein, Bloemfontein.
948
949 Morrow, P.A. & Fox, L.R. (1989). Estimates of pre-settlement insect damage in Australian and
950 North American forests. *Ecology*, **70**, 1055–1060.
951
952 Neil, K.L., Landrum, L. & Wu, J. (2010). Effects of urbanization on flowering phenology in the
953 metropolitan phoenix region of USA: Findings from herbarium records. *Journal of Arid*
954 *Environments*, **74**, 440–444.
955
956 Newbold, T. (2010). Applications and limitations of museum data for conservation and ecology,
957 with particular attention to species distribution models. *Progress in Physical Geography*, **34**, 3–
958 22.
959
960 Norby, R.J., De Kauwe, M. G., Domingues, T. F., Duursma, R. A., Ellsworth, D. S., Goll, D. S.,
961 Lapola, D. M., Luus, K. A., MacKenzie, A. R., Medlyn, B. E., Pavlick, R. (2016). Model–data
962 synthesis for the next generation of forest free-air CO₂ enrichment (FACE) experiments. *New*
963 *Phytologist*, **209**, 17–28.
964
965 Oerke, E.-C. & Dehne, H.-W. (2004). Safeguarding production—losses in major crops and the
966 role of crop protection. *Crop protection*, **23**, 275–285.
967
968 Pachauri, R.K., Meyer, L., Plattner, G.K. and Stocker, T. (2015). *IPCC, 2014: Climate Change*
969 *2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment*
970 *Report of the Intergovernmental Panel on Climate Change*. IPCC.
971
972 Park, D.S. & Davis, C.C. (2017). Implications and alternatives of assigning climate data to
973 geographical centroids. *Journal of Biogeography*.
974
975 Park, D.S. & Potter, D. (2013). A test of Darwin's naturalization hypothesis in the thistle tribe
976 shows that close relatives make bad neighbors. *Proceedings of the National Academy of*
977 *Sciences*, **110**, 17915–17920.
978
979 Parmesan, C., Ryrholm, N., Stefanescu, C., Hill, J.K., Thomas, C.D., Descimon, H. *et al.* (1999).
980 Poleward shifts in geographical ranges of butterfly species associated with regional warming.
981 *Nature*, **399**, 579–583.
982

- 983 Pauw, A. & Hawkins, J.A. (2011). Reconstruction of historical pollination rates reveals linked
984 declines of pollinators and plants. *Oikos*, **120**, 344–349.
985
- 986 Pelini, S.L., Bowles, F.P., Ellison, A.M., Gotelli, N.J., Sanders, N.J. & Dunn, R.R. (2011).
987 Heating up the forest: open-top chamber warming manipulation of arthropod communities at
988 Harvard and Duke Forests. *Methods in Ecology and Evolution*, **2**, 534–540.
989
- 990 Pilson, D. (2000). The evolution of plant response to herbivory: simultaneously considering
991 resistance and tolerance in *Brassica rapa*. *Evolutionary Ecology*, **14**, 457–489.
992
- 993 Plant, R.E. (2012). *Spatial data analysis in ecology and agriculture using R*. cRc Press, Boca
994 Raton, FL, USA.
995
- 996 Pearse, W.D., Inouye, D.W., Primack, R.B., Davis, C.C., Davies, T.J. In Press. A statistical
997 estimator for determining the limits of contemporary and historic phenology. *Nature Ecology
998 and Evolution*.
999
- 1000 Post, E., Pedersen, C., Wilmers, C.C. & Forchhammer, M.C. (2008). Warming, plant phenology
1001 and the spatial dimension of trophic mismatch for large herbivores. *Proceedings of the Royal
1002 Society B-Biological Sciences*, **275**, 2005–2013.
1003
- 1004 Primack, D., Imbres, C., Primack, R.B., Miller-Rushing, A.J. & Del Tredici, P. (2004).
1005 Herbarium specimens demonstrate earlier flowering times in response to warming in Boston.
1006 *American Journal of Botany*, **91**, 1260–1264.
1007
- 1008 Pyke, G.H. & Ehrlich, P.R. (2010). Biological collections and ecological/environmental
1009 research: a review, some observations and a look to the future. *Biological Reviews*, **85**, 247–266.
1010
- 1011 Reef, R. & Lovelock, C.E. (2014). Historical analysis of mangrove leaf traits throughout the 19th
1012 and 20th centuries reveals differential responses to increases in atmospheric CO₂. *Global
1013 Ecology and Biogeography*, **23**, 1209–1214.
1014
- 1015 Renberg, I., Persson, M.W. & Emteryd, O. (1994). Preindustrial atmospheric lead contamination
1016 detected in Swedish lake sediments. *Nature*, **368**, 323–326.
1017
- 1018 Riera, R., Sangil, C. & Sanson, M. (2015). Long-term herbarium data reveal the decline of a
1019 temperate-water algae at its southern range. *Estuarine Coastal and Shelf Science*, **165**, 159–165.
1020
- 1021 Robbirt, K.M., Davy, A.J., Hutchings, M.J. & Roberts, D.L. (2011). Validation of biological
1022 collections as a source of phenological data for use in climate change studies: a case study with
1023 the orchid *Ophrys sphegodes*. *Journal of Ecology*, **99**, 235–241.
1024
- 1025 Romeiras, M.M., Figueira, R., Duarte, M.C., Beja, P. & Darbyshire, I. (2014). Documenting
1026 biogeographical patterns of African timber species using herbarium records: A conservation
1027 perspective based on native trees from Angola. *PLOS ONE*, **9**, e103403–e103403.
1028

- 1029 Sakai, A.K., Allendorf, F.W., Holt, J.S., Lodge, D.M., Molofsky, J., With, K.A. *et al.* (2001).
1030 The population biology of invasive species. *Annu. Rev. Ecol. Syst.*, **32**, 305–332.
1031
- 1032 Saltonstall, K. (2002). Cryptic invasion by a non-native genotype of the common reed,
1033 *Phragmites australis*, into North America. *Proceedings of the National Academy of Sciences of*
1034 *the United States of America*, **99**, 2445–2449.
1035
- 1036 Schaefer, H., Hardy, O.J., Silva, L., Barraclough, T.G. & Savolainen, V. (2011). Testing
1037 Darwin’s naturalization hypothesis in the Azores. *Ecology letters*, **14**, 389–396.
1038
- 1039 Särkinen, T., Staats, M., Richardson, J.E., Cowan, R.S. & Bakker, F.T. (2012). How to open the
1040 treasure chest? Optimising DNA extraction from herbarium specimens. *PLOS ONE*, **7**, e43808.
1041
- 1042 Sheppard, L., Bell, J.R., Harrington, R. & Reuman, D.C. (2016). Changes in large-scale climate
1043 alter spatial synchrony of aphid pests. *Nat. Clim. Chang.*, **6**, 610–+.
1044
- 1045 Sigal, L.L. & Nash, T.H. (1983). Lichen communities on conifers in southern California
1046 mountains: an ecological survey relative to oxidant air pollution. *Ecology*, **64**, 1343–1354.
1047
- 1048 Southwood, T.R.E., Henderson, P.A. (2009). *Ecological Methods*. John Wiley & Sons.
1049
- 1050 Spellman, K.V. & Mulder, C.P.H. (2016). Validating Herbarium-Based Phenology Models
1051 Using Citizen-Science Data. *Bioscience*, **66**, 897–906.
1052
- 1053 Steffen, W., Crutzen, P.J. & McNeill, J.R. (2007). The Anthropocene: Are humans now
1054 overwhelming the great forces of nature. *Ambio*, **36**, 614–621.
1055
- 1056 Stewart, G.R., Aidar, M.P.M., Joly, C.A. & Schmidt, S. (2002). Impact of point source pollution
1057 on nitrogen isotope signatures ($\delta N-15$) of vegetation in SE Brazil. *Oecologia*, **131**, 468–472.
1058
- 1059 Syfert, M.M., Serbina, L., Burckhardt, D., Knapp, S. & Percy, D.M. (2017). Emerging New Crop
1060 Pests: Ecological Modelling and Analysis of the South American Potato Psyllid *Russelliana*
1061 *solanicola* (Hemiptera: Psylloidea) and Its Wild Relatives. *PLOS ONE*, **12**, e0167764.
1062
- 1063 Thiers, B. (continuously updated). Index Herbariorum: A global directory of public herbaria and
1064 associated staff. New York Botanical Garden's Virtual Herbarium. Available at:
1065 <http://sweetgum.nybg.org/science/ih/> Last accessed December 2016.
1066
- 1067 Turcotte, M.M., Davies, T.J., Thomsen, C.J.M. & Johnson, M.T.J. (2014). Macroecological and
1068 macroevolutionary patterns of leaf herbivory across vascular plants. . *Proceedings of the Royal*
1069 *Society B-Biological Sciences*, **281**, 20140555.
1070
- 1071 United Nations, Department of Economic and Social Affairs, Population Division (2008). *World*
1072 *Urbanization Prospects: The 2007 Revision*.
1073

- 1074 Unger, J., Merhof, D. & Renner, S. (2016). Computer vision applied to herbarium specimens of
1075 German trees: testing the future utility of the millions of herbarium specimen images for
1076 automated identification. *BMC evolutionary biology*, **16**, 248.
1077
- 1078 Vandepitte, K., Meyer, T.D., Helsen, K., Acker, K.V., Roldan-Ruiz, I., Mergeay, J. *et al.* (2014).
1079 Rapid genetic adaptation precedes the spread of an exotic plant species. *Molecular ecology*, **23**,
1080 2157–2164.
1081
- 1082 Vellend, M., Brown, C.D., Kharouba, H.M., McCune, J.L. & Myers-Smith, I.H. (2013).
1083 Historical Ecology: using Unconventional Data Sources to Test for Effects of Global
1084 Environmental Change. *American Journal of Botany*, **100**, 1294–1305.
1085
- 1086 Vogel, G. (2017). Where have all the insects gone? *Science*, **356**, 576–579.
1087
- 1088 Wandeler, P., Hoeck, P.E.A. & Keller, L.F. (2007). Back to the future: museum specimens in
1089 population genetics. *Trends in Ecology & Evolution*, **22**, 634–642.
1090
- 1091 Warren, M.S., Hill, J.K., Thomas, J.A., Asher, J., Fox, R., Huntley, B. *et al.* (2001). Rapid
1092 responses of British butterflies to opposing forces of climate and habitat change. *Nature*, **414**,
1093 65–69.
1094
- 1095 Weiss, D., Shotyk, W., Kramers, J.D. & Gloor, M. (1999). Sphagnum mosses as archives of
1096 recent and past atmospheric lead deposition in Switzerland. *Atmospheric Environment*, **33**, 3751–
1097 3763.
1098
- 1099 Wernberg, T., Russell, B.D., Thomsen, M.S., Gurgel, C.F.D., Bradshaw, C.J.A., Poloczanska,
1100 E.S. *et al.* (2011). Seaweed communities in retreat from ocean warming. *Current Biology*, **21**,
1101 1828–1832.
1102
- 1103 Wilf, P. & Labandeira, C.C. (1999). Response of plant-insect associations to Paleocene-Eocene
1104 warming. *Science*, **284**, 2153–2156.
1105
- 1106 Wilf, P., Zhang, S., Chikkerur, S., Little, S.A., Wing, S.L. & Serre, T. (2016). Computer vision
1107 cracks the leaf code. *Proceedings of the National Academy of Sciences*, **113**, 3305–3310.
1108
- 1109 Williams, A., Willis, C. G., Davis, C. C., Goh, J. Ellison, A. M., Law, E. (In Review). Deja Vu:
1110 Characterizing worker quality using task consistency. *ACM CHI Conference on Human Factors*
1111 *in Computing Systems*.
1112
- 1113 Willis, C.G., Ellwood, E.R., Primack, R.B., Davis, C.C., Pearson, K.D., Gallinat, A.S. *et al.*
1114 (2017a). Old Plants, New Tricks: Phenological Research Using Herbarium Specimens. *Trends in*
1115 *ecology & evolution*, **32**, 531–546.
1116
- 1117 Willis, C.G., Law, E., Williams, A. C., Franzone, B. F., Bernardos, R., Bruno, L., Hopkins, C.,
1118 Schorn, C., Weber, E., Park, D. S., Davis, C. C. (2017b). CrowdCurio: an online crowdsourcing

- 1119 platform to facilitate climate change studies using herbarium specimens. *New Phytologist*, **215**,
1120 479–488.
- 1121
- 1122 Willis, C.G., Ruhfel, B., Primack, R.B., Miller-Rushing, A.J. & Davis, C.C. (2008). Phylogenetic
1123 patterns of species loss in Thoreau's woods are driven by climate change. *Proceedings of the*
1124 *National Academy of Sciences of the United States of America*, **105**, 17029–17033.
- 1125
- 1126 Wolkovich, E.M. & Cleland, E.E. (2011). The phenology of plant invasions: a community
1127 ecology perspective. *Frontiers in Ecology and the Environment*, **9**, 287–294.
- 1128
- 1129 Wolkovich, E.M., Cook, B.I., Allen, J.M., Crimmins, T.M., Betancourt, J.L., Travers, S.E. *et al.*
1130 (2012). Warming experiments underpredict plant phenological responses to climate change.
1131 *Nature*, **485**, 494–497.
- 1132
- 1133 Woodward, F.I. (1987). Stomatal numbers are sensitive to increases in CO₂ from pre-industrial
1134 levels. *Nature*, **327**, 617–618.
- 1135
- 1136 Wu, Z.T., Dijkstra, P., Koch, G.W., Penuelas, J. & Hungate, B.A. (2011). Responses of
1137 terrestrial ecosystems to temperature and precipitation change: a meta-analysis of experimental
1138 manipulation. *Global Change Biology*, **17**, 927–942.
- 1139
- 1140 Yaakub, S.M., McKenzie, L.J., Erftemeijer, P.L.A., Bouma, T. & Todd, P.A. (2014). Courage
1141 under fire: Seagrass persistence adjacent to a highly urbanised city-state. *Marine Pollution*
1142 *Bulletin*, **83**, 417–424.
- 1143
- 1144 Youngsteadt, E., Dale, A.G., Terando, A.J., Dunn, R.R. & Frank, S.D. (2015). Do cities simulate
1145 climate change? A comparison of herbivore response to urban and global warming. *Global*
1146 *Change Biology*, **21**, 97–105.
- 1147
- 1148 Zangerl, A.R. & Berenbaum, M.R. (2005). Increase in toxicity of an invasive weed after
1149 reassociation with its coevolved herbivore. *Proceedings of the National Academy of Sciences of*
1150 *the United States of America*, **102**, 15529–15532.
- 1151
- 1152 Zarnetske, P.L., Skelly, D.K. & Urban, M.C. (2012). Biotic Multipliers of Climate Change.
1153 *Science*, **336**, 1516–1518.
- 1154
- 1155 Zavala, J.A., Nability, P.D. & DeLucia, E.H. (2013). An Emerging Understanding of Mechanisms
1156 Governing Insect Herbivory Under Elevated CO₂. In: *Annual Review of Entomology, Vol 58* (ed.
1157 Berenbaum, MR). Annual Reviews Palo Alto, pp. 79–97.
- 1158
- 1159 Ziska, L.H., Pettis, J.S. & Edwards, J. (2016). Rising atmospheric CO₂ is reducing the protein
1160 concentration of a floral pollen source essential for North American bees. *Proceedings of the*
1161 *Royal Society B*. **283**.
- 1162

1163 **Table 1.** Key research questions and hypotheses from the global change literature for which herbaria may be well-suited as a data
 1164 source. Most of the research questions apply to multiple global change drivers and are grouped by the driver that has received the most
 1165 interest. These suggestions are biased toward plants, for which more diverse herbarium data extraction methods have been developed,
 1166 though we believe many could also be tested for fungi. We searched Web of Science to estimate the current interest in each research
 1167 topic. Searches were performed with “Herbari*” & the listed keywords. The “Citations” column contains the number of citations
 1168 retrieved and, when available, iconic, representative examples from the literature.

Global change driver	Research questions	Herbarium data	Hypotheses	Keywords	Citations
Climate change	Are phenological shifts increasing or decreasing in magnitude?	<ul style="list-style-type: none"> • Flowering • Plant, fungal fruiting • Leaf-out • Locality, date collected 	<p>Plant flowering and leaf-out are becoming less sensitive to climate as species reach their tolerance thresholds.</p> <p>Fungal fruiting patterns are changing across the season due to the redistribution of rainfall.</p>	“Climate change” & “Phenolog*” or “Flower*” or “Leaf-out” or “Fruit*” or “Seed*”	<p>95</p> <p>Diez <i>et al.</i> (2013) Kauserud <i>et al.</i> (2010) Matthews & Mazer (2016)</p>
	How does plant geographic location affect phenological responses to climate change?	<ul style="list-style-type: none"> • Flowering • Plant, fungal fruiting • Leaf-out • Locality, date collected (latitude, longitude) 	<p>Within and between species, plant phenological cueing mechanisms vary across latitude.</p>		

	<p>Are species' ranges shifting with climate change?</p>	<ul style="list-style-type: none"> • Locality, date collected (historical and current ranges) 	<p>Species are moving poleward and up elevations due to climate change.</p> <p>Human development limits plant species movement poleward.</p>	<p>“Climate change” & “Species distribution” or “Range” & “Latitude” or “Elevation” or “Urbanization”</p>	<p>20</p> <p>Feeley <i>et al.</i> (2013) Feeley & Silman (2010)</p>
	<p>Does dispersal syndrome influence plant range shifts due to climate change?</p>	<ul style="list-style-type: none"> • Fruit dispersal mode • Locality, date collected (historical and current ranges) 	<p>Bird-dispersed plants are able to migrate longer distances than those dispersed simply by gravity.</p>	<p>“Climate Change” & “Dispersal”</p>	<p>19</p>

	<p>How does climate change affect pest/pathogen abundance/diversity/community structure?</p>	<ul style="list-style-type: none"> • Herbivory • Insects and their damage with specimens that can be assigned species identity, e.g., leaf mines, galls (historical and current ranges, host shifts) • Pathogen lesions, DNA, RNA (historical and current ranges, host shifts) • Locality, date collected (community structure) 	<p>Herbivore abundance and damage increases with warming, consistent with fossil evidence.</p> <p>Herbivore distributions are expanding north and/or retracting at southern limits due to climate warming.</p> <p>Warmer climates increase the geographic spread of pathogens/herbivores, facilitating shifts to novel hosts.</p> <p>In areas that are warming, communities shift and become less structurally even due to increasing abundance of species with high thermal tolerances.</p>	<p>“Climate Change” & “Pest” or “Herbiv*” or “Pathogen” or “Disease”</p>	<p style="text-align: center;">12</p> <p style="text-align: center;">Youngsteadt <i>et al.</i> (2015)</p>
--	--	---	--	--	--

<p>Have pollination rates changed through time? If so, are declines driven by climate change? Does climate change interact with other global change drivers, such as urbanization?</p>	<ul style="list-style-type: none"> • Flower morphology, e.g., corolla length • Metrics of fitness, such as seed size, set • Pollen protein concentrations • Pollen removal 	<p>Pollination has decreased due to a combination of drivers, including climate change.</p>	<p>“Climate Change” & “Pollin*” or “Pollen”</p>	<p style="text-align: center;">25</p> <p style="text-align: center;">Miller-Struttman <i>et al.</i> (2015) Pauw & Hawkins (2011) Ziska <i>et al.</i> (2016)</p>
<p>How have climatic niches changed, and how well do current environmental niche models match past distributions?</p>	<ul style="list-style-type: none"> • Locality, date collected (realized historical and current niches) 	<p>Climate envelopes predict geographical distributions, such that species fill newly available niche space driven by climate change.</p>	<p>“Climate Change” & “Niche” or “Species distribution” or “SDM”</p>	<p style="text-align: center;">59</p>
<p>Are climatic changes and range shifts associated with shifts in gas exchange rates?</p>	<ul style="list-style-type: none"> • Gas exchange proxies, e.g., stomatal densities, isotope ratios, and guard cell lengths • Water use efficiency proxies, e.g., stomatal density, molecular hydrogen isotope composition δD 	<p>Climate change has increased photosynthetic rates, except when plants are water-stressed.</p>	<p>“Climate change” & “Photosynthe*” or “Gas exchange” or “Stomata” or “Stomatal conductance”</p>	<p style="text-align: center;">7</p> <p style="text-align: center;">Miller-Rushing <i>et al.</i> (2009)</p>

<p>Are plants evolving <i>in situ</i> to climate change?</p> <p>Are cooler adapted genotypes being displaced by warmer adapted genotypes via migration?</p>	<ul style="list-style-type: none"> • Plant DNA (allele frequencies) • Morphology, e.g., leaf size, shape, specific leaf area 	<p>Plant gene frequencies and phenotypes have shifted over time, and these shifts are consistent with changing climate, e.g., environments that have become drier select for more dry-tolerant phenotypes.</p>	<p>“Climate change” & “Evolution” or “Adaptation”</p>	<p>36</p>
<p>What roles do plant water relations play in responses to climate change across phylogeny and habitats?</p>	<ul style="list-style-type: none"> • Water use efficiency proxies, e.g., stomatal density, molecular hydrogen isotope composition δD • Gas exchange proxies, e.g., stomatal densities, isotope ratios, and guard cell lengths • Morphology, e.g., leaf size, shape, specific leaf area • Plant carbon content 	<p>Warming increases plant productivity at mid and upper latitudes, except when plants are water stressed.</p> <p>Plants worldwide are living close to their hydraulic limits.</p>	<p>“Climate change” & “Water-use efficiency” or “Water stress” or “Water potential”</p>	<p>4</p>

	<p>Does climate change promote phenological asynchrony and/or ecological mismatch between associated species?</p>	<ul style="list-style-type: none"> • Flowering • Flower morphology, e.g., corolla length • Plant, fungal fruiting • Leaf-out • Herbivory • Pollen removal 	<p>Plant and insect phenology shift at similar rates with temperature, such that pollination and herbivory rates are constant despite climate change.</p> <p>Plants and their insect/mycorrhizal associates respond to different cues and thus will become/are less synchronized due to climate change.</p> <p>Asynchronies driven by climate change will be reduced over time by rapid selection for insects to synchronize with plants and in the case of pollination, and vice versa.</p>	<p>“Climate change” & “Synchrony” or “Asynchrony” or “Ecological mismatch”</p>	<p style="text-align: center;">5</p> <p style="text-align: center;">Kharouba & Vellend (2015) Miller-Struttman <i>et al.</i> (2015)</p>
--	---	---	--	--	--

Invasive species	Is invasive species spread facilitated by genomic change?	<ul style="list-style-type: none"> • Plant DNA (allele frequencies) • Locality, date collected (time of introduction, spread) 	New mutations or gene combinations enable invasive species to overcome dispersal barriers, perhaps via gene surfing on expanding population fronts.	“Invasive” or “Non-native” & “Genome” & “Adaptation” or “Genomic change”	12 Buswell et al. (2011) Vandepitte <i>et al.</i> (2013)
	Have invasive species demonstrated greater phenological advancement with climate warming than native species?	<ul style="list-style-type: none"> • Flowering • Leaf-out 	Greater phenological advancement of non-native compared to native species facilitates invasions.	“Invasive” or “Non-native” & “Phenolog*” or “Flower*” or “Leaf-out” or “Fruit*” or “Seed*”	46 Calinger (2015)
	What are the physical pathways of invasive species spread?	<ul style="list-style-type: none"> • Plant, leaf miner DNA • Locality, date collected (time of introduction, spread) 	Natural pathways, such as waterways, were historically more important for invasive plant and insect species spread, but increasingly roads and railroads are key.	“Invasive” or “Non-native” & “Spread” or “Railroad” or “Road”	80 Barney (2006) Joly <i>et al.</i> (2011) Saltonstall (2002)

	<p>In novel habitats, does release from natural enemies promote invasive species spread?</p>	<ul style="list-style-type: none"> • Herbivory • Insects and their damage that can be assigned species identity, e.g., leaf mines, galls • Pathogen lesions, DNA, RNA • Plant defensive compounds 	<p>One mechanism by which species become invasive is escape from co-evolved natural enemies. (Enemy Release Hypothesis)</p>	<p>“Invasive” or “Non-native” & “Natural enem*” or “Natural enemy release”</p>	<p>2 Zangerl & Berenbaum (2005)</p>
	<p>What roles do diseases play in invasions?</p>	<ul style="list-style-type: none"> • Pathogen lesions, DNA, RNA • Locality, date collected (time of introduction, spread) 	<p>Diseases carried by non-native plants can facilitate their invasions via apparent competition.</p>	<p>“Invasive” or Non-native” & “Pathogen” or “Disease”</p>	<p>8 Malmstrom <i>et al.</i> (2007)</p>

	Does exotic species' relatedness to natives determine invasiveness?	<ul style="list-style-type: none"> • Herbivory • Insects and their damage that can be assigned species identity, e.g., leaf mines, galls (time of introduction, spread, host shifts) • Plant defensive compounds • Locality, date collected • (time of introduction, spread) 	<p>Exotic insect herbivores and pathogens are more likely to establish on novel host plants closely related to their co-evolved host plants.</p> <p>Exotic plant/pathogen/herbivore relatedness to native plants reduces the probability that they become invasive. (Darwin's Naturalization Hypothesis)</p>	<p>“Invasive” or “Non-native” & “Naturalization Hypothesis”</p>	<p>0</p> <p>Though this search returns no references, see Park & Potter (2013) and Schaefer <i>et al.</i> (2011).</p>
Habitat conversion	How do restored communities and their associates compare to pre-disturbance communities?	<ul style="list-style-type: none"> • Herbivory • Insects and their damage that can be assigned species identity, e.g., leaf mines, galls • Pathogen/endophyte/mycorrhizal DNA • Locality, date collected (community metrics) 	<p>Effects of restoration on diversity and community structure depend on land use legacies.</p>	<p>“Land use” or “Disturbance” & “Restoration”</p>	<p>5</p> <p>Bertin (2002) Celesti-Grapow <i>et al.</i> (2013) DeCandido <i>et al.</i> (2004) Dolan <i>et al.</i> (2011) Gregor <i>et al.</i> (2012)</p>

Can we use species occurrences prior to habitat change to guide restoration efforts?	<ul style="list-style-type: none"> • Locality, date collected (species composition prior to intensified anthropogenic change) 	Herbaria capture historical diversity and thus could serve as blueprints for restoration.	“Restoration”	37
Has global change led to no-analog communities?	<ul style="list-style-type: none"> • Locality, date collected (historical community structure) 	Habitat conversion, trade, climate change, among forms of global change, have led to novel plant and fungal communities.	“No-analog communit*” or “Novel communit*”	0
Are some community structures more robust to disturbance and/or invasion?	<ul style="list-style-type: none"> • Locality, date collected (historical community structure, diversity) • Plant DNA (phylogeny reconstructions) 	More diverse plant communities are more resilient to herbivore/invasive plant/pathogen pressure. (Biodiversity Insurance Hypothesis)	“Insurance Hypothesis” or “Resilience” & “Diversity” or “Richness”	2

	<p>How do different types of habitat change filter plant species and their microbial/arthropod associates?</p>	<ul style="list-style-type: none"> • Herbivory • Insects and their damage that can be assigned species identity, e.g., leaf mines, galls • Pathogen/endophyte/mycorrhizal DNA 	<p>Agriculture, urbanization, and other types of human development have signatures, such that biota in these habitats worldwide share common traits, and communities include certain species with global distributions. (Biotic Homogenization Hypothesis)</p>	<p>“Ecological filter” or “Habitat filter”</p>	<p>3</p>
	<p>What is the timescale of natural restoration?</p>	<ul style="list-style-type: none"> • Herbivory • Insects and their damage that can be assigned species identity, e.g., leaf mines, galls • Pathogen/endophyte/mycorrhizal DNA • Locality, date collected time series of (community structure, diversity) 	<p>Landscape connectivity increases the rate at which plants and their associates re-enter habitats.</p>	<p>“Re-establish*” or “Brownfield” or “Succession” or “Regeneration”</p>	<p>37</p>

	<p>How do plant associates respond to land use change?</p>	<ul style="list-style-type: none"> • Herbivory • Insects and their damage that can be assigned species identity, e.g., leaf mines, galls • Pathogen/endophyte/mycorrhizal DNA • Pollen protein concentrations • Pollen removal 	<p>Biodiversity of plant associates has decreased over time due to habitat conversion.</p> <p>Urbanization increases/decreases insect abundance/diversity/herbivory/pollination.</p> <p>Development releases some insects from their co-evolved natural enemies, allowing them to become pests (i.e., Enemy Release Hypothesis).</p> <p>Reduced endophyte loads and diversity in cities affects plant interactions with pollinators and herbivores.</p>	<p>“Land use” or “Disturbance” & “Herbiv*” or “Pathogen” or “Disease” or “Pollin*” or “Pollen” or “Endophyte” or “Mutualis*” or “Commensal*” or “Competit*” or “Predat*”</p>	<p style="text-align: center;">17</p> <p style="text-align: center;">Pauw & Hawkins (2011) Youngsteadt <i>et al.</i> (2015)</p>
--	--	---	---	--	--

	What selection pressures does land use change impose on plants?	<ul style="list-style-type: none"> • Morphology, e.g., leaf size, shape, specific leaf area • Plant DNA (allele frequencies) 	Cities worldwide, except deserts, select for species that can withstand relatively hot, dry, and open conditions.	“Land use” or “Disturbance” & “Adapt*” or “Selection” or “Evolution”	<p>14</p> <p>Dolan <i>et al.</i> (2011) Neil <i>et al.</i> (2010)</p>
	How has habitat conversion affected plants of cultural and economic importance, such as medicinal plants and wild crop relatives?	<ul style="list-style-type: none"> • Locality, date collected (abundance, historical and current ranges) • Morphology, e.g., leaf size, shape, specific leaf area • Metrics of fitness, such as seed size, set • Physiological characteristics, such as stomatal densities, isotope ratios, and guard cell lengths 	Habitat conversion has reduced abundance and range sizes of many species.	“Land use” or “Disturbance” & “Economic” or “Crop” or “Medicinal”	<p>14</p> <p>Farnsworth & Ogurcak (2006)</p>
Pollution	How has atmospheric pollution altered community composition of plants and fungi?	<ul style="list-style-type: none"> • Pollutant concentrations • Locality, date collected (community structure, diversity) 	Pollution selects for resistant species and lineages, driving phylogenetic underdispersion and reduced diversity.	“Pollut*” & “Communit*” or “Assembl*”	<p>8</p>

<p>Does biotic diversity enhance bioremediation, i.e., rates of pollutant removal, from a system?</p>		<p>More diverse communities remove pollutants more efficiently.</p>	<p>“Pollut*” & “Bioremediation” or “Recovery” & “Diversity” or “Richness”</p>	<p>0</p>
<p>How has exposure to atmospheric pollution varied through time and across regions?</p>		<p>Humans began polluting environments early in our history.</p>	<p>“Pollut*” & “Histor*” & “Human”</p>	<p>6</p>

	<p>How does atmospheric pollution affect plant associates?</p>	<ul style="list-style-type: none"> • Pollutant concentrations • Locality, date collected (community structure, phylogenetic diversity) • Pollutant concentrations • Herbivory • Insects and their damage that can be assigned species identity, e.g., leaf mines, galls • Leaf miner DNA (allele frequencies) • Pathogen/endophyte/mycorrhizal DNA • Pollen removal 	<p>Pollution reduces plant associate diversity by selecting resilient species and genotypes.</p>	<p>“Pollut*” & “Herbiv*” or “Pathogen” or “Disease” or “Pollin*” or “Pollen” or “Endophyte” or “Mutualis*” or “Commensal*” or “Competit*” or “Predat*”</p>	<p>3</p>
--	--	---	--	--	----------

	<p>Are pesticides, e.g., neonicotinoids/ chlorpyrifos, responsible for the sudden decline of many insects?</p>	<ul style="list-style-type: none"> • Insects and their damage that can be assigned species identity, e.g., leaf mines, galls • Pesticide residues in/on pollen and leaves • Pollen removal 	<p>Pesticide presence has increased in natural plant populations over time.</p> <p>Pesticides are present in a diversity of non-crop plants.</p> <p>Wild insect diversity/abundance/ herbivory has decreased due to pesticides.</p> <p>Pollen transfer has declined with pesticide use.</p>	<p>“Pollut*” & “Pesticide” & “Insect”</p>	<p>0</p>
--	--	---	---	---	----------

Exploitation	Do protected areas truly protect threatened and endangered species?	<ul style="list-style-type: none"> • Locality, date collected (historical and current occurrence within protected areas) 	<p>Because the locations of protected areas are driven by human concerns, and rarely informed by diversity data, many species are not protected.</p> <p>The amount of biodiversity preserved over time depends on management regimes, and resulting habitat patch size and connectivity.</p>	<p>“Protected area” and “Threatened species” or “Endangered species”</p>	<p>3</p> <p>Romeiras <i>et al.</i> (2014)</p>
	Can we predict extinctions and populations declines due to exploitation?	<ul style="list-style-type: none"> • Metrics of fitness, such as seed size, set • Morphology, e.g., leaf size, shape, specific leaf area • Physiological characteristics, such as stomatal densities, isotope ratios, and guard cell lengths • Plant DNA (allele frequencies) 	<p>Elevated rates of genotypic/phenotypic change precede population collapse.</p>	<p>“Harvest” and “Extinction” or “Extirpation” or “Decline”</p>	<p>4</p>

	What are the selection coefficients imposed by human harvests?	<ul style="list-style-type: none"> • Morphology, e.g., leaf size, shape, specific leaf area • Plant DNA (allele frequencies) 	Human harvests of wild plants reduce plant size by inducing selection pressure against larger individuals.	“Harvest” & “CITES” or “Exploitation” & “Selection” or “Adaptation” or “Morphology” or “Size” or “Height” or “Specific leaf area”	36 Law & Salick (2005)
--	--	--	--	---	----------------------------------

1169
 1170
 1171
 1172
 1173
 1174
 1175
 1176
 1177
 1178
 1179
 1180
 1181
 1182
 1183

1184
1185
1186
1187

Table 2. Herbarium data for global change research. Here, we highlight the different data types, challenges to their use, and potential solutions to overcome these challenges. We focus here on the data types and uses we have discussed in the main text; other applications may be subject to additional challenges (and potential solutions).

Herbarium data	Challenges	Potential solutions
All data	Biases over space, time, and phylogeny	<ul style="list-style-type: none"> • Careful focal herbaria/species selection aided by digitized specimens • Subsampling and techniques for inference with biased data, and statistical methods such as machine learning
Occurrence (locality, date collected)	Lack of recorded absences	<ul style="list-style-type: none"> • Statistical tools for simulating or otherwise analysing presence-only data, which are already well-developed for species distribution models and may be employed for other areas of research.
	Coarse-level geographical data.	<ul style="list-style-type: none"> • Careful focal herbaria/species selection aided by digitized specimens • In some cases, it may be possible to address hypotheses with environmental data at coarse scales, e.g., temperature data averaged at the county

		level in the U.S.
Phenology	Specimens are most likely to be collected at peak flowering times, thus missing early season phenological events and times of first event.	<ul style="list-style-type: none"> Use statistical estimators to infer timing of first events from a sampled distribution (Pearse <i>et al.</i> In Press).
	Collectors tend to retrieve specimens from near roadsides (Daru <i>et al.</i> In Press), which could affect phenology. For example, life events may be advanced relative to plants in more natural areas due to the urban heat island effect or delayed due to drought-like conditions.	<ul style="list-style-type: none"> When comparing across space, time, and/or phylogeny, relative measures may be sufficient to address hypotheses of interest. For studies that do require absolute measurements, use specimens with geolocations and model potential biases (e.g., urbanization). Additional studies are needed to assess the effects of local collection biases on global change data, especially for plant traits that are sensitive to temperature, such as phenology.
	For leaf-out, categorical criteria, such as pubescence, can be used to determine when specimens have newly flushed leaves (see Everill <i>et</i>	<ul style="list-style-type: none"> Continuous criteria can be developed, such as leaf size relative to fully flushed leaf size (also see Everill <i>et al.</i>

	<p><i>al.</i> (2014)). However, many species do not have unique qualities associated with new leaves.</p>	<p>2014), but such criteria introduce considerably more work than do categorical characteristics.</p>
<p>Other traits (morphology, physiology, fitness, and microbiomes)</p>	<p>Botanists may be more likely to collect specimens from more vigorous individuals or branches with greater seed set, seed size, leaf size, water use efficiency, etc.</p>	<ul style="list-style-type: none"> • When comparing across space, time, and/or phylogeny, relative measures may be sufficient to address hypotheses of interest. • Cross-validate recent collections with observational data to ascertain, and thus allow correction for, potential sampling biases.
	<p>Key tissues, morphological structures, or developmental stages may not be sampled or may be damaged. For example, roots are necessary for mycorrhizal studies but are not present on all specimens.</p>	<ul style="list-style-type: none"> • Check digital images and select intact specimens. • If alleles for a trait are known, it may be possible to genotype a sample without a visible phenotype and thus avoid the need for the actual structure possessing the trait of interest. • For certain species, roots can be easily collected and are thus present on many herbarium specimens. Overcoming this challenge should be a matter of

		<p>choosing appropriate focal taxa.</p>
Pollination	<p>Matching data on pollinators are often not available.</p>	<ul style="list-style-type: none"> Score pollination on herbarium specimens. This will only be possible for groups that have distinct pollen-holding structures (pollinaria), such as many orchids and milkweeds.
	<p>To build matching plant-pollinator collections, one needs to know which pollinators are associated with which plants.</p>	<ul style="list-style-type: none"> Identify plants via pollen morphology or DNA preserved on pollinator specimens.
Antagonistic interactions (herbivory, disease)	<p>Botanists may be likely to collect less damaged specimens.</p>	<ul style="list-style-type: none"> When comparing across space, time, and/or phylogeny, relative measures may be sufficient to address hypotheses of interest. Cross-validate recent collections with observational data to ascertain, and thus allow correction for, potential sampling biases.
	<p>Difficulty differentiating between taxa. For example, it may be difficult to identify taxa that created galls and leaf mines because of specimen degradation.</p>	<ul style="list-style-type: none"> Combined morphological and molecular methods may improve identifications. Focus on species with distinct,

		<p>well-preserved gall, mine, and pathogen damage morphologies.</p> <ul style="list-style-type: none"> • Direct sequencing of associates if preserved on plant specimens, for example, using DNA barcodes for species identification.
	<p>DNA, RNA degradation</p>	<ul style="list-style-type: none"> • For most species, RNA will be too degraded. However, Malmstrom <i>et al.</i> (2007) have shown it is possible to extract usable RNA in some (perhaps rare) cases. • Next generation sequencing techniques developed for amplifying degraded or ancient DNA • Collection of tissue samples preserved in silica or banked in ultra-low temperature storage.
	<p>Quantifying damage on individual specimens</p>	<ul style="list-style-type: none"> • We have demonstrated the potential of subsampling specimens to quantify herbivory damage or pathogen lesions. (See the “Limitations and Challenges” section),

		<p>analogous to the use of quadrats in vegetation sampling.</p> <ul style="list-style-type: none"> • In cases where visual signs of herbivory are difficult to quantify, defensive compounds might provide an alternative way to detect damage.
Pollution	Choosing appropriate indicator species for pollutants, regions, and time periods of interest	<ul style="list-style-type: none"> • Indicator species for many pollutants, such as nitrogen and heavy metals, have been established. Digitized records can then help select those with suitable temporal and geographic coverage.
	Pesticide degradation	<ul style="list-style-type: none"> • Some pesticides degrade faster than others, and recovering pesticide residues may be possible for chemicals that break down more slowly. This area of research has not been explored, and thus methods are not yet developed.
	Botanists may be unlikely to collect in heavily polluted environments or near areas where pesticides have been applied.	<ul style="list-style-type: none"> • Labels may indicate if specimens were collected near farms, roads, homes, and industry. These metadata can be included in subsequent

		analyses.
--	--	-----------

1188

1189

1190 **Figure captions**

1191

1192 **Fig. 1. The spatial, temporal, and phylogenetic extents of herbarium specimens.** Here we
1193 explore the richness of herbarium data, with emphasis on the major herbaria in and area within
1194 the northeastern USA referred to as New England (<http://neherbaria.org/>). (a) The herbaria of
1195 New England house more than half a million regional herbarium specimens collected as early as
1196 the 1800s, before industrialization and globalization, that can serve as baselines prior to
1197 anthropogenic change. The density plot represents the number of collections each year. (b) These
1198 specimens have spatial coverage across the region (shown here), and millions more specimens
1199 are available from other parts of the world. Though there is clear spatial bias, the spatial
1200 coverage exceeds that which is typically available in observations and experiments. (c) These
1201 specimens span most branches of the vascular plant tree of life. Here, bars represent log
1202 abundance of specimens in families, which are organized by evolutionary history (Harris &
1203 Davies 2016). (d) Herbaria are distributed worldwide. Many of the larger herbaria have amassed
1204 similar coverage to the New England herbaria across these axes. Here, symbols represent
1205 herbaria. Herbaria with more than one million specimens are indicated by white stars.

1206 **Fig. 2. The diversity of insect herbivory preserved on herbarium specimens.** We found
1207 several types of herbivory on herbarium specimens made by a diversity of arthropods, and
1208 quantified herbivory in five categories representing, (a) chewing, (b) skeletonization, (c)
1209 stippling, (d) leaf galls (here, with emerging gall wasp circled in red), and leaf mines: (e) typical
1210 leaf mine and (f) leaf mine with miner inside circled in red. We found other herbivores, such as
1211 aphids and caterpillars, pressed with plant specimens, but these were rarer. Chewing damage is
1212 typically made by caterpillars and beetles; leaf mines are made by flies, beetles, and mites;
1213 stippling is made by leafhoppers and other species that remove cell contents from leaves; leaf
1214 galls are typically made by gall wasps.

1215

1216 **Fig. 3. Herbivory recorded on herbarium specimens of New England.** We examined
1217 herbarium specimens from 20 species from the northeastern US for various types of herbivory
1218 (see Appendix S1) and contrasted observations to expectations from the literature. We found that
1219 (a) species that are closely related have similar herbivory (Blomberg's $K=0.4$), (b) composition
1220 of damage types on herbarium specimens is more similar within than between species and

1221 genera, illustrated here by two genera *Viola* and *Lespedeza* (PERMANOVA, $F_{19,507}=7.05$,
1222 $p<0.001$). Herbivory within genera clusters more tightly than between genera, as shown by the
1223 general separation of polygons. Similarly, herbivory composition is significantly different
1224 between species within genera, shown here as partial overlap of points colored by species. (c)
1225 Herbivory on herbarium specimens is highly skewed, and (d) herbivory increased as growing
1226 seasons progressed. Chewing damage is shown in (a) and (d) because it was the most prevalent
1227 type of herbivory (as shown in c). Detailed statistics are available in Appendix 1.

1228

1229

1230 **Appendix S1. Detailed sampling methods for scoring herbivory on herbarium specimens**

1231 Herbivory data collection

1232 We chose focal species by selecting from the list of native species included in Willis *et al.*, 2008.
1233 We filtered this list for species with over 15 specimens collected between 1900 and 1920, 1920
1234 and 1940, 1940 and 1960 and deposited at the Harvard University Herbaria. We randomly chose
1235 10 native rosids and 10 native asterids from this filtered database to ensure phylogenetic breadth
1236 in our survey. We chose only native species to avoid any effects of recent introduction on
1237 herbivory. We then randomly selected 40 specimens from each species collected between 1900
1238 and 1960 from Connecticut, Massachusetts, and Rhode Island and preserved in the Harvard
1239 University Herbaria collections. We chose these states to minimize spatial variation in the data
1240 and these years to exclude impacts of recent global warming, which became pronounced in the
1241 1970s (IPCC 2014).

1242 To quantify herbivory, we randomly subsampled five cells within a 40.64 by 25.40-cm grid
1243 overlaid on each specimen. If a cell did not include leaf material we selected another cell. We
1244 marked each grid cell for presence/absence of each type of herbivory, e.g., leaf mines, galls,
1245 chewing damage.

1246 We encountered several challenges when quantifying herbivory. Historically, many herbarium
1247 collections were previously kept in collectors' homes where they were not protected from
1248 insects. Insects can also infest tightly sealed cabinets where herbarium specimens are currently
1249 stored. We were only interested in herbivory that happened outdoors. Therefore, we developed
1250 methods for distinguishing indoor and outdoor herbivory. We determined that outdoor herbivory

1251 is generally distinct from indoor herbivory, in that live plants form toughened, necrotic tissue
1252 around the wounds made by herbivores. However, this tissue is more visible on some species
1253 than others. Therefore, certain species may not be appropriate for herbarium studies focused on
1254 herbivory. In this study, there were several cases in which individual specimens were damaged,
1255 and we could not distinguish indoor from outdoor herbivory. When we encountered this issue,
1256 we randomly selected another specimen to replace the specimen with ambiguous damage.

1257 Statistical analyses

1258 First, we examined the relationships between herbivore damage and host plants. We expected
1259 significant variation in herbivory between plant species, because plants have evolved to various
1260 extents to defend against different herbivores (Bale *et al.*, 2002). We also expected that closely
1261 related species would have similar levels of damage, because they are likely to share similar anti-
1262 herbivory traits, including defense and tolerance (Pilson 2000; Agrawal & Fishbein 2008;
1263 Futuyma & Agrawal 2009; Fornoni 2011), though we note that few studies explore plant
1264 tolerance in phylogenetically explicit frameworks. Consistent with these expectations, we
1265 demonstrated that there was (1) significant variation in herbivory between plant species
1266 (GLMER; $X^2_{1,19} = 855.50$, $P < 0.0001$), focusing on chewing damage, which was by far the most
1267 prevalent type of herbivory (Fig. 3c), and (2) evidence suggesting close relatives had similar
1268 herbivory (Fig. 3a; phylogenetic signal: Blomberg's $K = 0.40$), although low sample size ($n = 20$)
1269 limited our ability to test for statistical significance. We then evaluated the prediction that
1270 damage composition would cluster by plant species. Because herbivory is frequently host-plant
1271 specific, the types of herbivory on plant species, e.g., chewing, skeletonization, leaf mines, and
1272 the relative amounts of herbivory types, should be more similar among individuals within species
1273 than between species. We tested this expectation by building a dissimilarity matrix using the
1274 abundances of each damage type on each specimen and evaluating the distances between
1275 samples among individuals within vs between species. As expected, damage types cluster within
1276 species and genera (Fig. 3b), and centroids within the dissimilarity matrix were significantly
1277 different among species (PERMANOVA: $F_{19,507} = 7.05$, $p < 0.001$).

1278 Second, we examined the statistical distribution of herbivory and the accumulation of damage
1279 through time. Because plants vary in their defenses against herbivory, and insect attacks can be
1280 highly localized, herbivory data tend to have medians of zero, a pattern seen frequently in studies

1281 on living plants (Turcotte *et al.*, 2014). In addition, we would expect that as growing seasons
1282 progress, the diversity of damage types found on specimens should increase, as the leaves are
1283 exposed to herbivores for longer, and as galls and leaf mines grow and become detectable. We
1284 found support for both predictions in our data. Most observations showed no evidence of
1285 herbivory with a few observations showing high damage (Fig. 3c), but damage diversity,
1286 calculated as the number of types of damage per specimen, was higher on specimens collected
1287 later in the year (Fig. 3d; GLMER: Wald $Z=4.06$, $p<0.0001$).

1288 Last, we contrasted our estimates of herbivory from herbarium specimens to a database of
1289 herbivory collated by Turcotte *et al.*, (2014). We found five matching genera between datasets
1290 and demonstrated that, across these genera, herbivory was strongly correlated (Pearson
1291 correlation: $r = 0.71$).

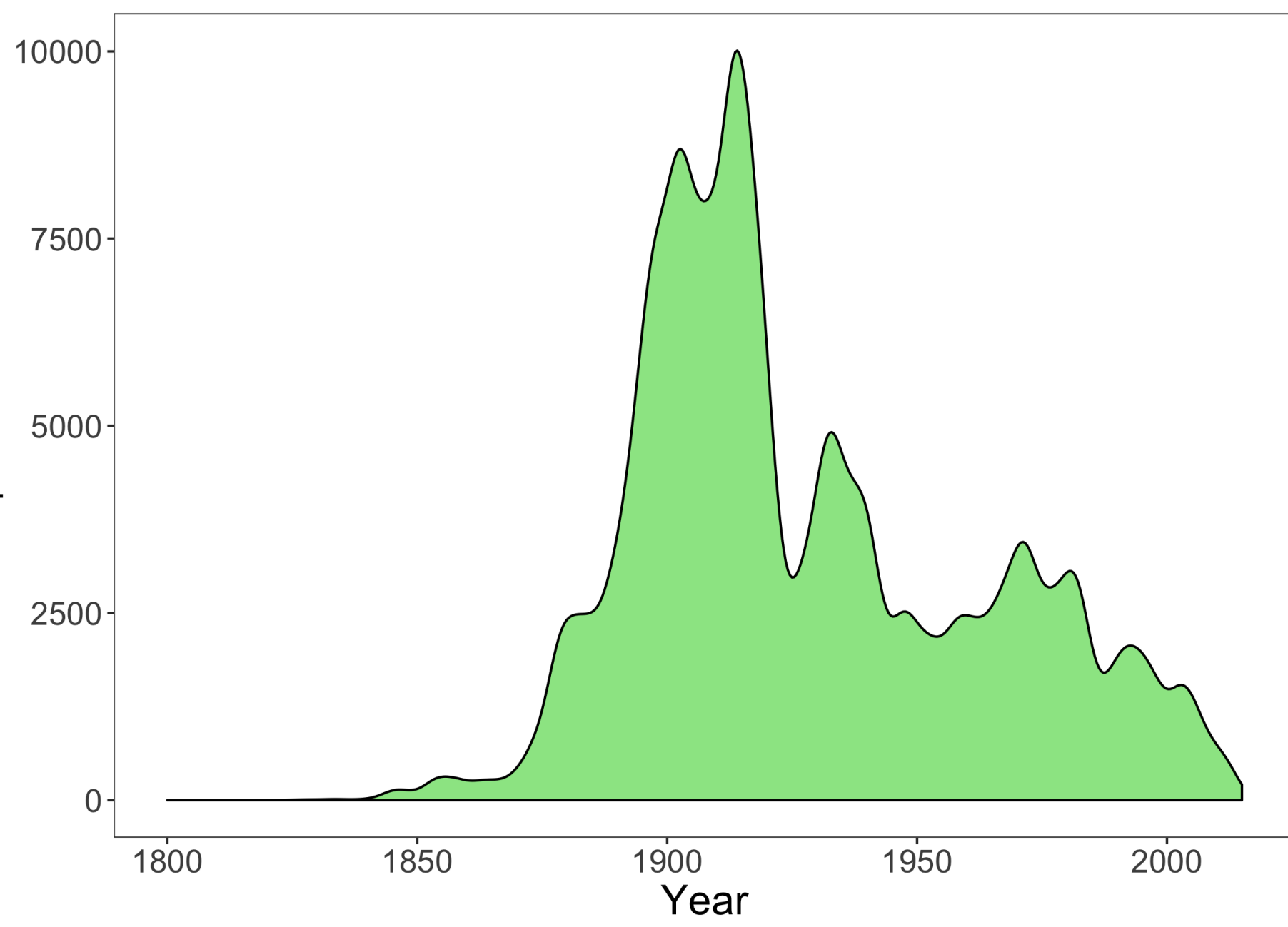
1292

1293

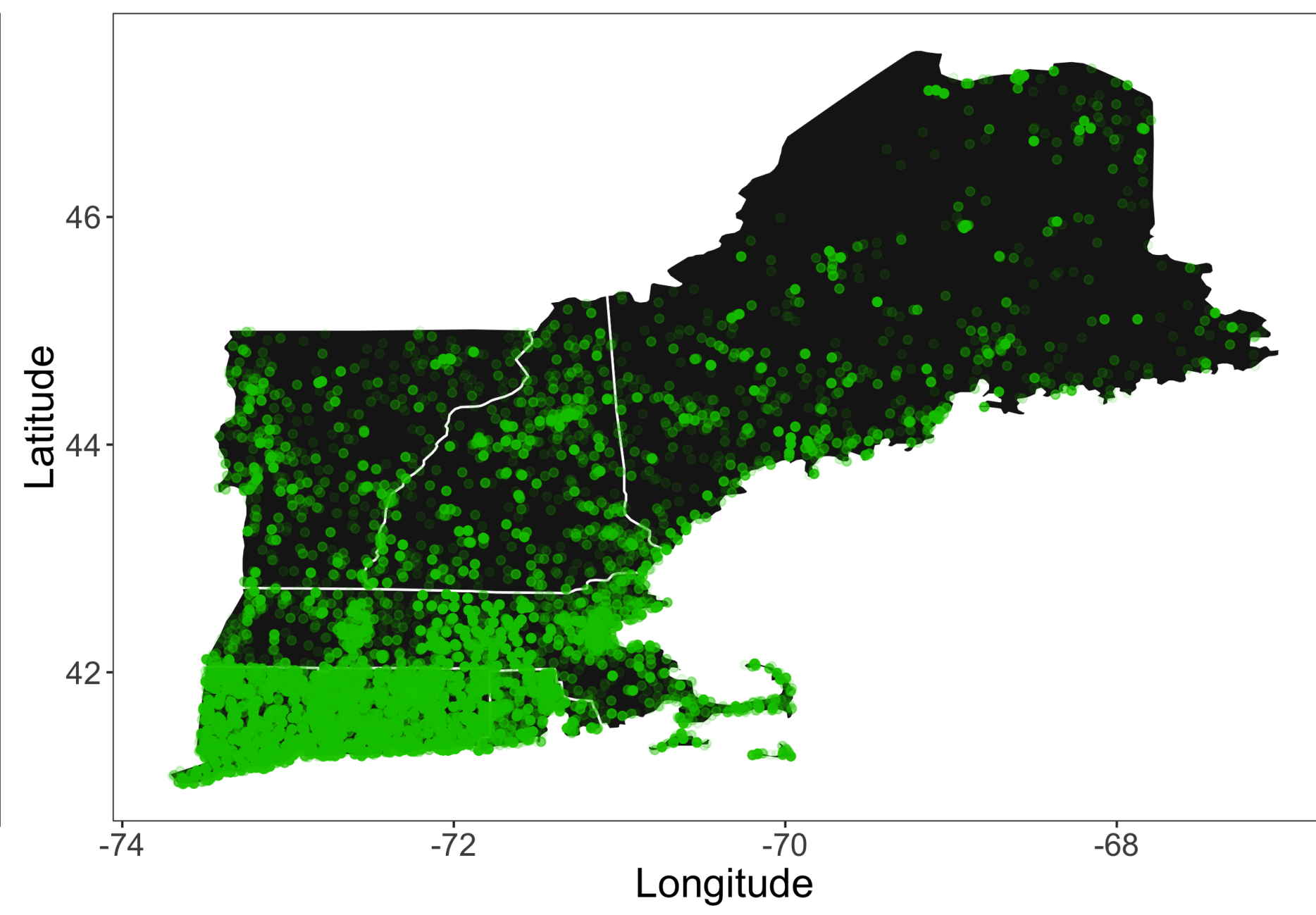
1294

1295

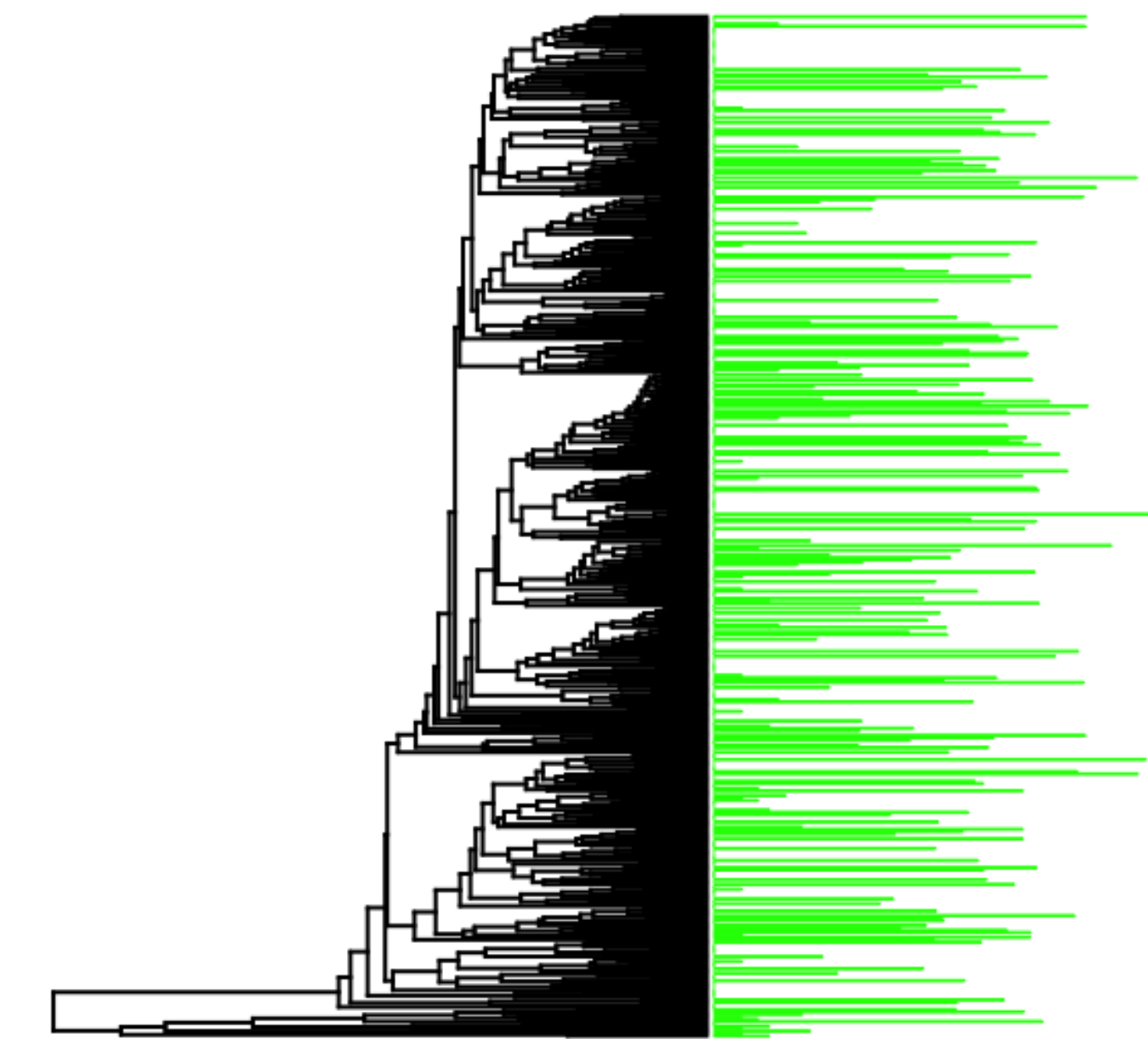
(a)



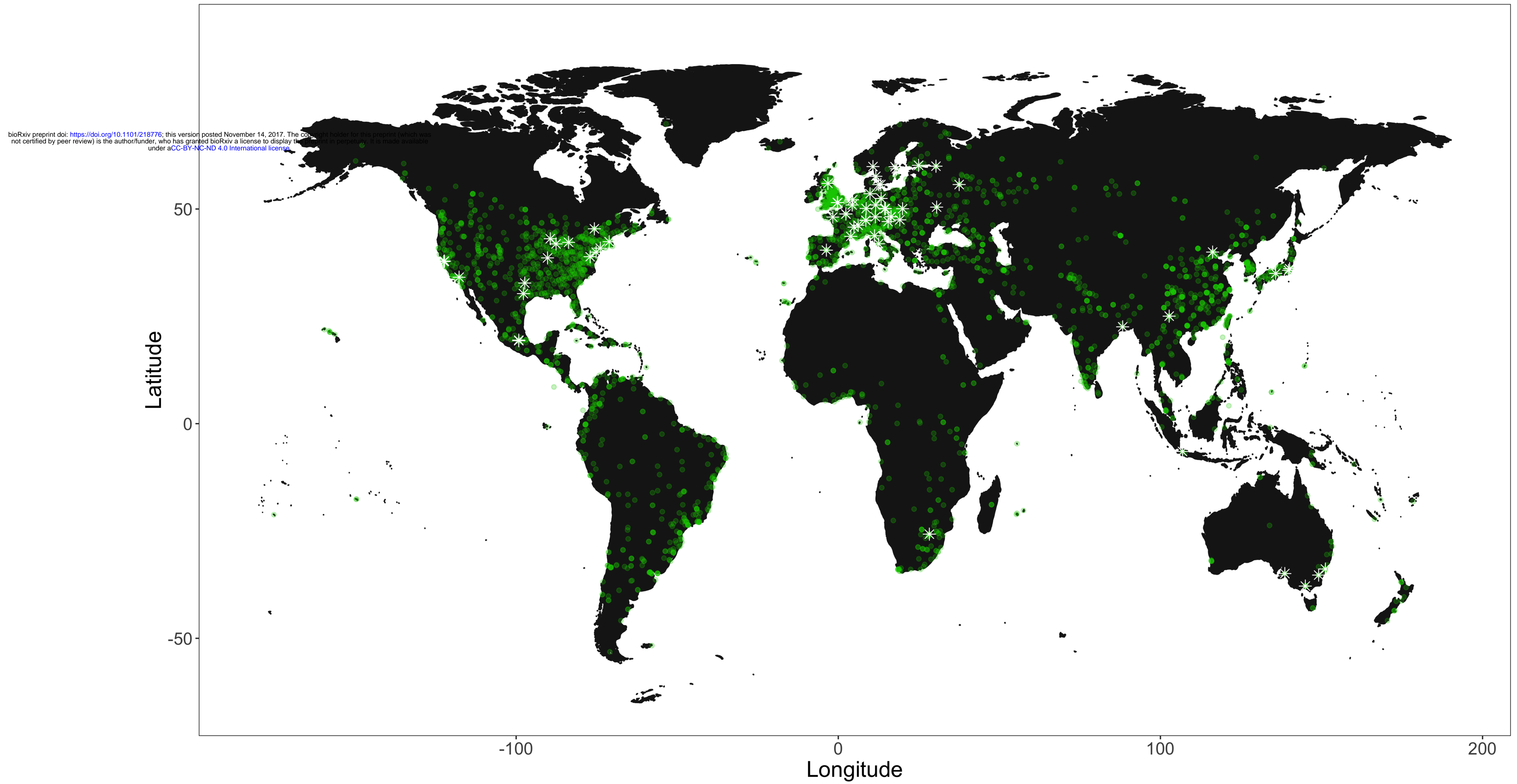
(b)



(c)



(d)



(a)



(b)



(c)



(d)



(e)



(f)



