The colonial legacy of herbaria

Park, D.S., Feng, X., Akiyama, S, ... Bijmoer, R. ... et al.

DOI:
https://doi.org/10.1016/j.scitotenv.2023.165933

nature human behaviour

Article

https://doi.org/10.1038/s41562-023-01616-7

The colonial legacy of herbaria

Received: 19 August 2022

Accepted: 27 April 2023

Published online: 12 June 2023

Check for updates

Daniel S. Park[1,2,62] ✉, Xiao Feng[3,62], Shinobu Akiyama[4], Marlina Ardiyani[5], Neida Avendaño[6,7], Zoltan Barina[8], Blandine Bärtschi[9], Manuel Belgrano[10], Julio Betancur[11], Roxali Bijmoer[12], Ann Bogaerts[13], Asunción Cano[14], Jiří Danihelka[15,16], Arti Garg[17], David E. Giblin[18], Rajib Gogoi[19], Alessia Guggisberg[20], Marko Hyvärinen[21], Shelley A. James[22], Ramagwai J. Sebola[23,24], Tomoyuki Katagiri[25,26], Jonathan A. Kennedy[27], Tojibaev Sh. Komil[28], Byoungyoon Lee[29], Serena M. L. Lee[30], Donatella Magri[31], Rossella Marcucci[32], Siro Masinde[33], Denis Melnikov[34], Patrik Mráz[35], Wieslaw Mulenko[36], Paul Musili[33], Geoffrey Mwachala[33], Burrell E. Nelson[37], Christine Niezgoda[38], Carla Novoa Sepúlveda[39], Sylvia Orli[40], Alan Paton[41], Serge Payette[42], Kent D. Perkins[43], Maria Jimena Ponce[44], Heimo Rainer[45,46], L. Rasingam[47], Himmah Rustiami[5], Natalia M. Shiyan[48], Charlotte Sletten Bjorå[49], James Solomon[50], Fred Stauffer[51], Alex Sumadijaya[5,52], Mélanie Thiébaut[9], Barbara M. Thiers[53], Hiromi Tsubota[54,55], Alison Vaughan[56], Risto Virtanen[57,58], Timothy J. S. Whitfeld[59], Dianxiang Zhang[60], Fernando O. Zuloaga[10] & Charles C. Davis[61] ✉

Herbarium collections shape our understanding of Earth's flora and are crucial for addressing global change issues. Their formation, however, is not free from sociopolitical issues of immediate relevance. Despite increasing efforts addressing issues of representation and colonialism in natural history collections, herbaria have received comparatively less attention. While it has been noted that the majority of plant specimens are housed in the Global North, the extent and magnitude of this disparity have not been quantified. Here we examine the colonial legacy of botanical collections, analysing 85,621,930 specimen records and assessing survey responses from 92 herbarium collections across 39 countries. We find an inverse relationship between where plant diversity exists in nature and where it is housed in herbaria. Such disparities persist across physical and digital realms despite overt colonialism ending over half a century ago. We emphasize the need for acknowledging the colonial history of herbarium collections and implementing a more equitable global paradigm for their collection, curation and use.

The nearly 400 million specimens residing in the world's herbaria form the basis of the scientific understanding of our planet's flora and are a centrepiece of botanical research[1]. Since the sixteenth century, scientists including Linnaeus and Darwin have collected herbarium specimens principally to describe species and circumscribe taxonomic classifications. The past decade has seen a resurgence in herbarium collections research, which is driven in part by massive digitization efforts[2–4]. In particular, with advances in high-throughput methods

A full list of affiliations appears at the end of the paper. ✉e-mail: danielpark@purdue.edu; cdavis@oeb.harvard.edu

and image analyses, herbarium specimens are increasingly being used in innovative ways[5,6] beyond their original intended purpose, including research pertaining to global change[7–9]. For example, herbarium specimens have been used to uncover the effects of climate change on plant phenology[10], ecophysiology[11] and herbivory[7]; as barometers for pollution[12] and eutrophication trends[13]; and to reconstruct the origin and spread of invasive species[14,15].

However, these collections are not free from the many sociopolitical issues that define our modern era. Despite increased efforts by natural history museums and other cultural institutions to address their legacy of colonialism and representation, such efforts have largely been focused on human- and animal-related collections and public exhibits[16,17]. In contrast, herbaria have received comparatively less attention, sidelined by their lower visibility; few herbaria offer public displays, and plant awareness is generally lacking[18]. Nonetheless, botanists have contributed substantially to the colonial expansion of imperial powers through active participation in the overseas collection of plants and their scientific and economic development[19]. Much of the early exploration of colonized nations by colonialists (during which many botanical specimens were collected) was done with the assistance of Indigenous peoples who acted as guides during expeditions. The relationships between colonial collectors and Indigenous guides were complex. Sometimes Indigenous peoples participated in these endeavours of their own accord, driven by shared interests, potential rewards, political gain or goodwill, though remaining mostly unacknowledged. However, at other times, Indigenous peoples were often under duress or forced to disclose their scientific knowledge of plants and place[20–22]. Though not specifically quantified, it has thus been noted that herbaria in the Global North hold many of the voucher specimens and associated data from equatorial and southern hemisphere nations (that is, the Global South) owing to colonial-era explorations[23,24].

To address the appropriation of plant diversity and to open a dialogue to help move us towards a more expansive and inclusive herbarium of the future[25], we must first understand the extent of disparity in herbarium collections across the globe—specifically, a more robust quantification of where they were collected and where they currently reside. Here we, scientists and curators from herbaria across 31 countries from every continent, examine the colonial legacy of botanical collections by assessing the geopolitical distribution of herbarium collections and digitization efforts. Analysing over 85,621,930 plant specimen records from the Global Biodiversity Information Facility (GBIF) (23 April 2021; https://doi.org/10.15468/dl.nt5wkx), one of the largest biodiversity data infrastructures (N = 2,307,116,169; accessed on 23 April 2023), and assessing the state of selected herbarium collections across the world, we provide a view of the disparity present in herbarium collections and discuss the future of this colonial legacy and how its effects can be mitigated. Though here we primarily focus on the past territorial manifestations of colonialism, it should be noted that this is not the only facet of (ongoing) colonialism relevant to botanical collections, which have been shaped by the evolving distribution of demands, priorities and interests of imperial formations. We acknowledge that the structures of power in the colonial contexts are far more complex than our discourse here allows, and not all botanical collections are solely the product of simple asymmetrical power relationships. Along these lines, though we focus on characterizing the macro-level phenomenon of disparity in herbarium collections, the patterns we demonstrate have been influenced and mediated by micro-level individualistic motivations, structures and networks.

Results

The imprint of colonialism in online herbarium collections

Collection trends across the last four centuries strongly bear the imprint of colonialism. These trends can be readily observed in the plant specimen records hosted on GBIF, which represent a subset (~25%) of global herbarium collections (Fig. 1a). The majority of plant specimens collected across the globe are currently housed in European countries and the United States—the 10 largest herbaria in the world, which are estimated to collectively hold over 65 million specimens, are all from these regions[1]. Indeed, the currently widely adopted taxonomy of life originated from European scholars, most prominently Linnaeus and his disciples, who were associated with the relocation of massive numbers of plant collections from across the globe into European institutions and their associated systems of knowledge. Following the initial natural history expeditions by northern and western Europeans, the United States later mounted the US Exploring Expedition, collecting tens of thousands of living and preserved plant specimens from across the world forming the basis of the Smithsonian Institution's United States National Herbarium. This trend was further fuelled by the desire of imperial powers to exploit the biological resources of colonies abroad, a legacy of which is, for example, the pursuit of medicinal plants in tropical regions in search of profitable remedies for ailments that are of greater interest in developed countries, such as cancer or obesity[26].

The impact of this collecting legacy persists in the trends and patterns of more recent collecting activities. Despite the era of overt colonialism drawing to an end after the Second World War, the historical trend of specimen movement from Africa, Asia and South America to Europe and North America has largely remained constant (Fig. 1b,c), especially among countries that have historical connections[27]. In fact, the proportion of specimens collected from other continents has increased in Europe and North America over time. In particular, the United States emerged as the largest collector of overseas specimens after the Second World War, acquiring massive collections from countries such as Brazil and Madagascar. Notably, the proportion of specimens collected and housed in South America greatly increased during this period, while collection activity in Africa remained largely driven by European and North American countries, with the possible exception of South Africa. These patterns are largely consistent when limiting our sample to records with more complete information (for example, geographic coordinates; Supplementary Fig. 1) or to collection activity in the twenty-first century (Supplementary Fig. 2). However, we note that there are other factors, such as the degree of economic development, regional policies, political stability and scientific interest, that have likely influenced these patterns as well[28]. Also, though difficult to estimate, a portion of the specimens that have been dislocated likely have duplicates—separate physical specimens of the same taxon collected by the same people at the same place on the same date—deposited at local institutions. Among the specimen data we examined from GBIF, only 2.8% were of the same species collected at the exact same place and date and stored in different institutions.

Our analysis suggests that colonial exploitation has contributed to an inverse relationship between where plant biodiversity exists in nature and where it is housed in herbaria. In general, biodiversity is distributed along a latitudinal gradient, with most of the world's plant diversity located in the tropics[29]. However, when we examine the number of species collected in a given country—which reflects species richness—relative to the number of species with specimens housed in the same country, disparities emerge (Fig. 2a). Specifically, most of the world's flora is stored in temperate regions in a reverse-latitudinal gradient where absolute latitude is positively correlated with the number of species with specimens housed in a country ($r(243) = 0.26$; $P = 2.82 \times 10^{-5}$; 95% confidence interval (CI) = [0.14, 0.38]). In particular, herbaria in the United States and several nations in western and central Europe house over twice the number of species that occur in these nations, demonstrating the international appropriation of large numbers of specimens representing global plant diversity (Fig. 2a). In contrast, much of Africa and Asia house fewer species than are collected there, because North American and European herbaria currently house many of the specimens and associated data from these regions owing, in no small degree, to their colonial past. Indeed, nations from these two areas simultaneously house a disproportionate

https://doi.org/10.1038/s41562-023-01616-7

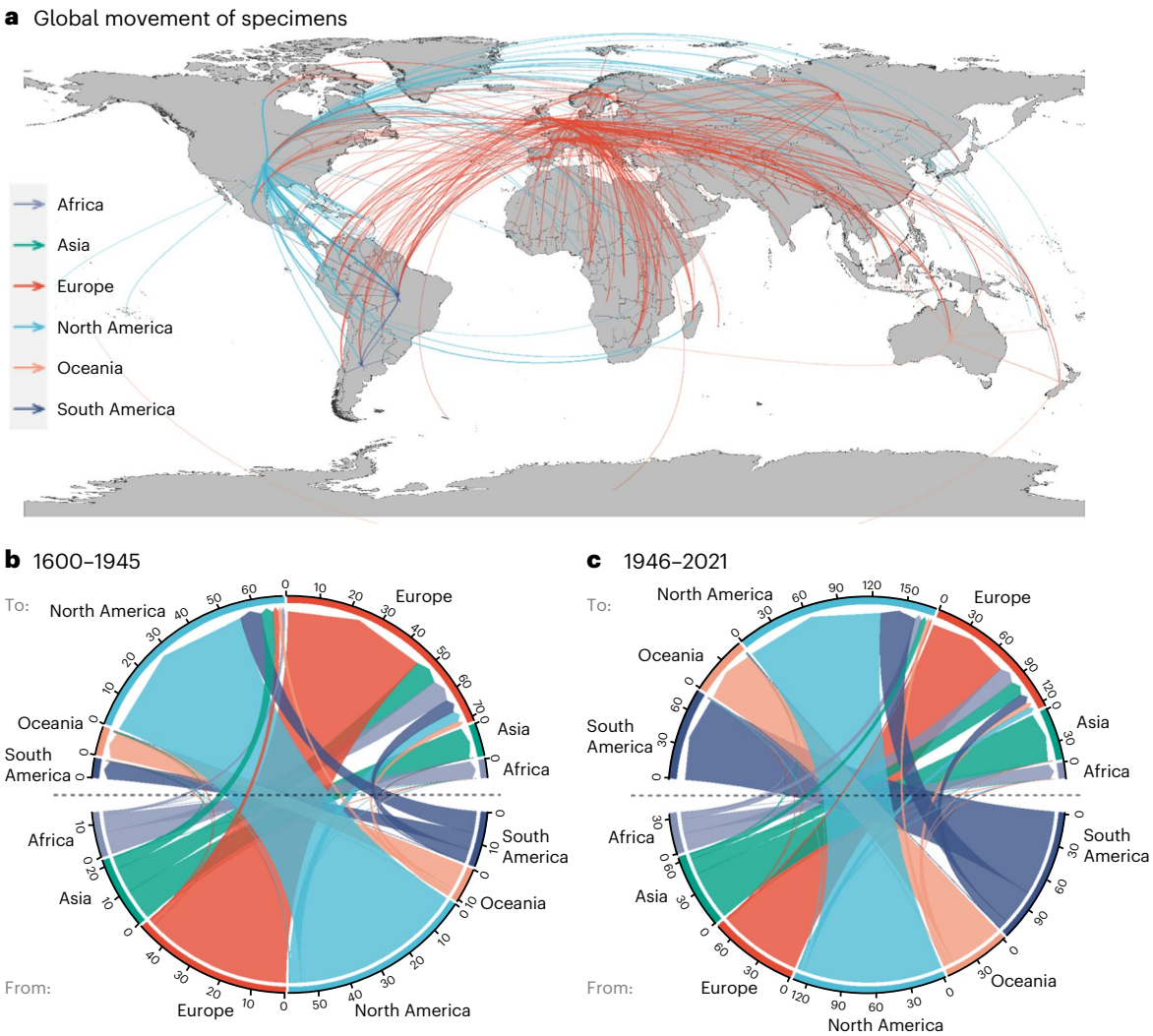

Fig. 1 | The past movement of plant specimens across the globe based on records from GBIF. a, The top tenth percentile of intercontinental connections between countries where specimens have been collected and where they are currently housed regardless of collection date. The widths of the arrows are proportionate to the number of specimens dislocated and are coloured by destination continent. Collections that remained in the country of collection are not depicted. **b**,**c**, The intercontinental dislocation of specimens before (**b**) and after (**c**) the end of overt colonialism post World War II (late 1945). The arrows are coloured by the continent of origin. The numbers on the outer ring indicate numbers of specimens collected from (lower half) or stored in (upper half) each continent and are in multiples of 100,000. The colours on the outer ring represent different continents. Political boundaries in panel **a** are based on data from GADM.

number of internationally collected specimens (Fig. 2b) and tend to have self-collected most of the specimens coming from their own countries (Fig. 2c). Furthermore, over 80% of the specimens with digital images are held by European and North American institutions, the majority of which were collected from Europe and North America, respectively (Supplementary Fig. 3). We note that not all countries in these two continents have actively participated in territorial colonialism, but some have nonetheless amassed sizable international plant collections (for example, Switzerland). There are also countries outside these regions that have sizeable international collections paired with more extensive self-collections (for example, Japan and New Zealand). Moreover, not all digitized specimen data online are available from GBIF—unique data can be found in smaller, regional repositories or institutional databases. Furthermore, such online databases harbour gaps and biases. The digital data assembled for this study are thus not entirely reflective of the complete distribution of specimens collected and stored across the world[30,31]. Nonetheless, our results are based on one of the largest biodiversity data infrastructures (N = 2,307,116,169; accessed on 23 April 2023) and represent our best estimates to date. To address these inherent limitations of our evaluation of digitized specimen content, we examined the distribution of specimens within physical herbaria across the world.

A glimpse inside the cabinet

Increasing digitization of specimen data and their online mobilization seem to have greatly decentralized and democratized access to herbarium data[32]. As demonstrated above, open-access biodiversity data infrastructures such as GBIF and iDigBio allow researchers from around the world to query aggregated specimen metadata and images, alleviating some of the need for extensive and prohibitive travels to consult materials and requests for loans. Institutional databases, although containing fewer specimens than global databases, efficiently contribute to make their own holdings available and encourage worldwide researchers to request free high-resolution images and better define loan requests. However, digitization requires substantial investments in infrastructure (that is, physical space, photographic devices and

https://doi.org/10.1038/s41562-023-01616-7

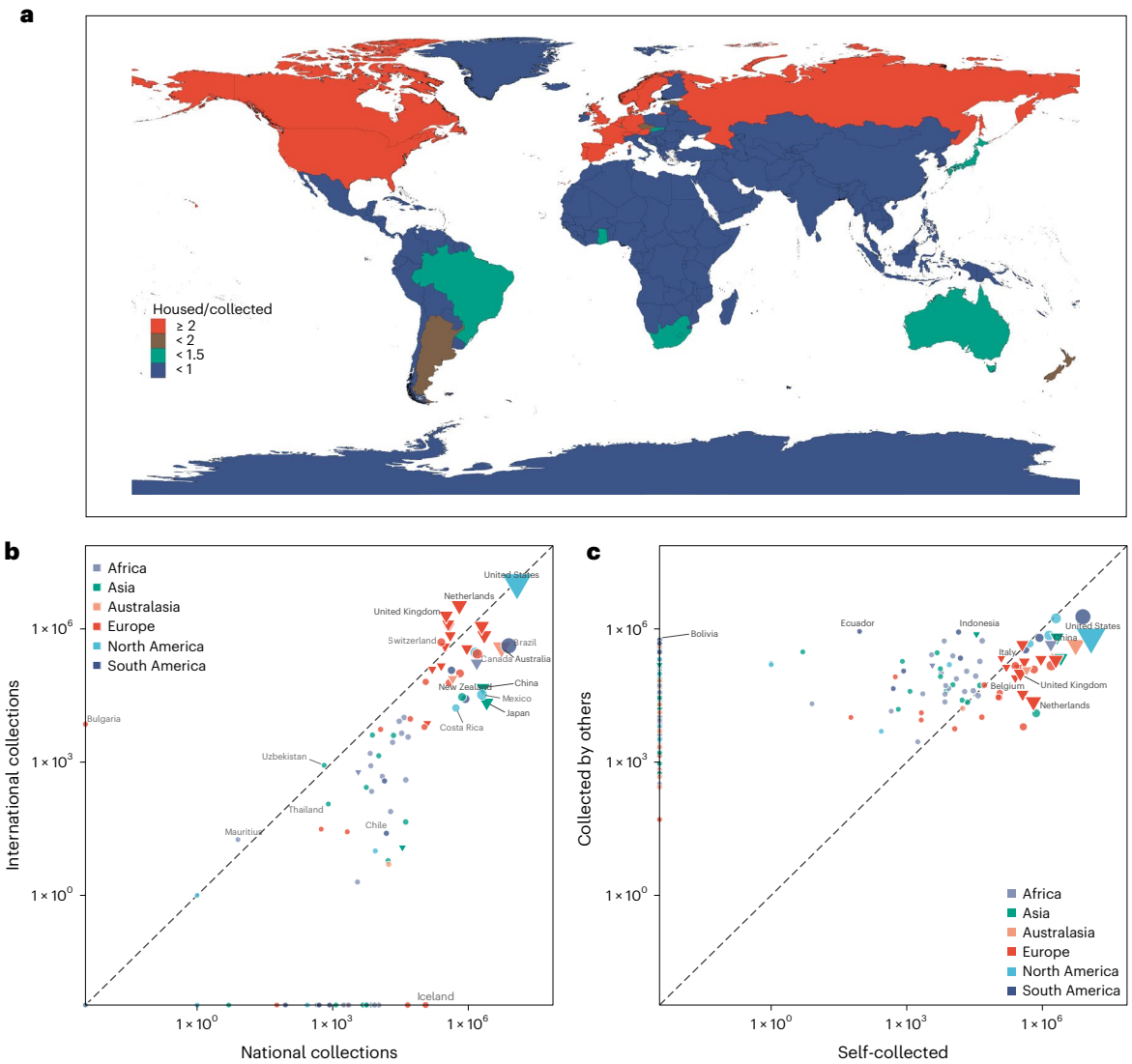

Fig. 2 | Disparity in the collection and housing of plant diversity. a, The ratio of the total number of species with specimens housed in a country to the total number of species collected in that country (species housed/collected). Ratios below one (blue) indicate areas where the number of species that have been collected from that country is higher than the number of species housed in that country. **b**, The ratio of nationally versus internationally collected specimens held by each country. **c**, The ratio of self-collected specimens in each country versus those collected by other countries. In **b** and **c**, the point sizes are log-scaled to the total number of specimens. The triangles represent countries that have overtly colonized other countries in the past following ref. 43. Political boundaries in **a** are derived from Esri, Global Mapping International and *The World Factbook*.

data storage) and personnel, which is often not financially feasible for small institutions and developing countries[4]. Along these lines, it has been argued that digitization could exacerbate the exploitation of intellectual property and biological resources by developed nations in a form of neo-imperialism[33]. Indeed, recent maps of biodiversity data gathered from mobile phone technologies appear to parallel many of the colonial-era trends we have identified here for herbaria[25]. Furthermore, only a small portion of specimen data are digitized and shared online at this time, and there are many studies that require access to physical specimens.

According to *Index Herbariorum*[1], there are at least 3,426 herbaria globally that together house approximately 400 million specimens. Over 60% of these herbaria and 70% of specimens are located in developed countries with colonial histories (Supplementary Fig. 4). To further understand the current state of the world's collections and their digitization, we conducted a collaborative assessment of major herbaria as listed by *Index Herbariorum* and targeted representative regional herbaria. A total of 92 herbaria across 39 countries and 6 continents submitted at least partial responses to our inquiries used in our analysis. Similar to the patterns observed using digitized data from GBIF, we identified that herbaria in developed nations with colonial histories in North America and Europe housed a higher proportion of internationally collected specimens on average ($t(59.10) = 3.58$; $P = 6.9 \times 10^{-4}$; $d = 0.82$; 95% CI = [9.50, 33.54]; Fig. 3a). This pattern generally held consistent across databased specimens with collection date and location information ($t(56.01) = 2.81$; $P = 6.8 \times 10^{-3}$; $d = 0.62$; 95% CI = [4.80, 28.62]; Fig. 3b) and specimens with digital images ($t(35.82) = 2.61$; $P = 0.01$; $d = 0.65$; 95% CI = [5.14, 41.24]; Fig. 3c) shared online. There were some notable exceptions; for instance, herbaria in Singapore hold a disproportionate number of international collections, possibly due to the country's small size, location, history as the main British colonial outpost in the area, and past and present association with Malaysia.

Our inquiries also revealed that the digitization of herbarium specimens remains in its infancy. We estimated that less than 30% of

https://doi.org/10.1038/s41562-023-01616-7

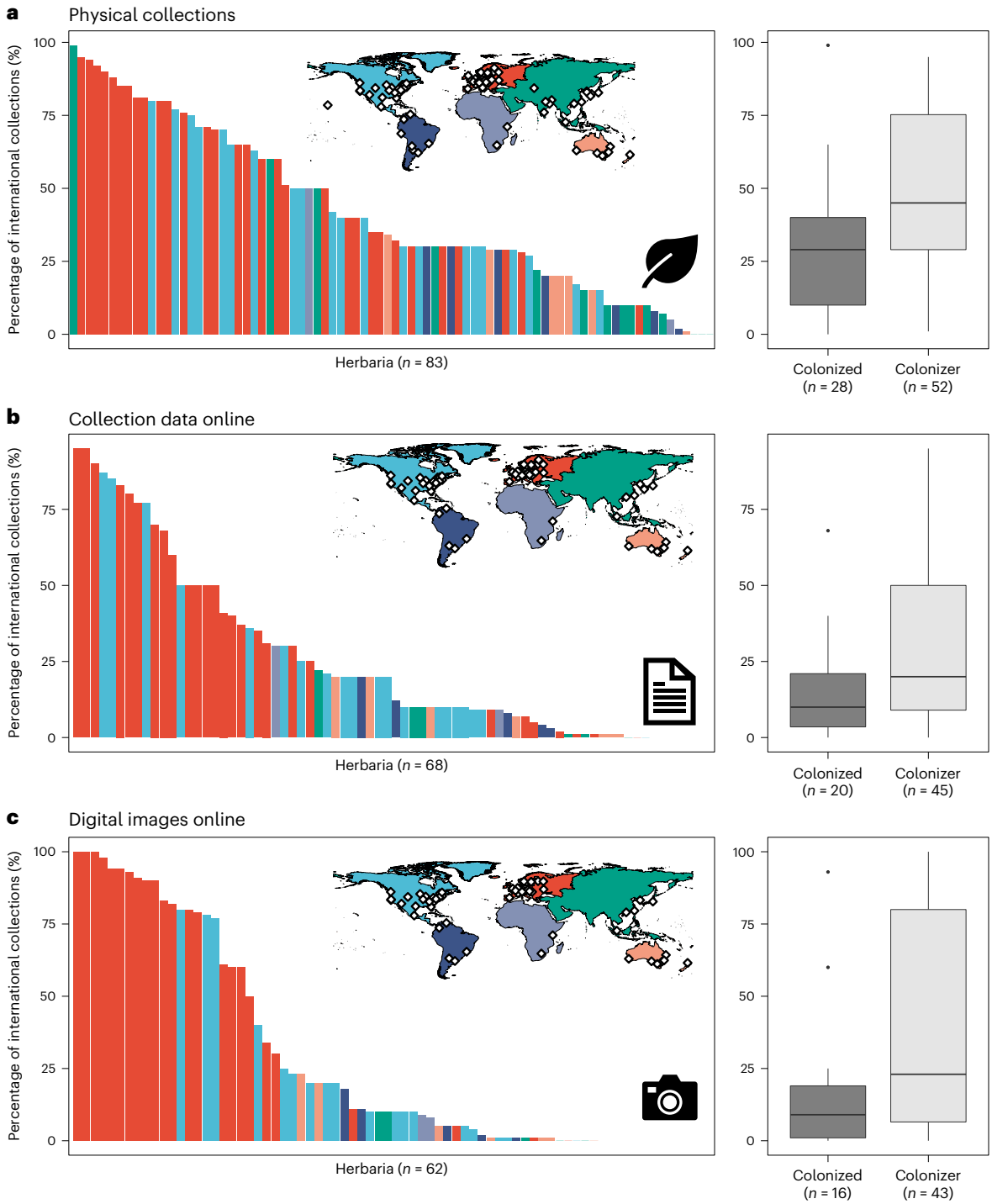

Fig. 3 | The percentage of internationally collected specimens in herbaria. **a**–**c**, Trends across physical specimens (**a**), specimens with at least a portion of their metadata available online (**b**) and specimens with digital images shared online (**c**); each bar represents a data-providing institution, and the colours indicate different continents. The box plots to the right summarize this information among countries that have been colonized (dark grey) versus those that have colonized others (light grey) following ref. 43. Countries that both experienced colonization and colonized others are depicted under their most recent category. The numbers represent the percentage of international collections in each category. The boxes in the box plots span the first and third quartiles around the median. The whiskers depict the 1.5 × interquartile range (IQR) unless minimum/maximum values fall within 1.5 × IQR of the quartiles. Points beyond the whiskers represent outliers. Political boundaries are derived from Esri, Global Mapping International and *The World Factbook*.

physical collections have at least collection location and date information online, and less than 10% have available digital images (Fig. 4a). Nearly all data-providing herbaria have ongoing digitization efforts with at least some specimen data provided online (Fig. 4b,c). However, these data are not always widely accessible, and they represent only the tip of the iceberg relative to the physical collections and are thus woefully insufficient to alleviate the reverse-latitudinal gradient of diversity inside herbarium cabinets. Our results suggest that the patterns we observe from GBIF data are likely representative of the larger reserves of specimen data yet to be digitized and mobilized online.

https://doi.org/10.1038/s41562-023-01616-7

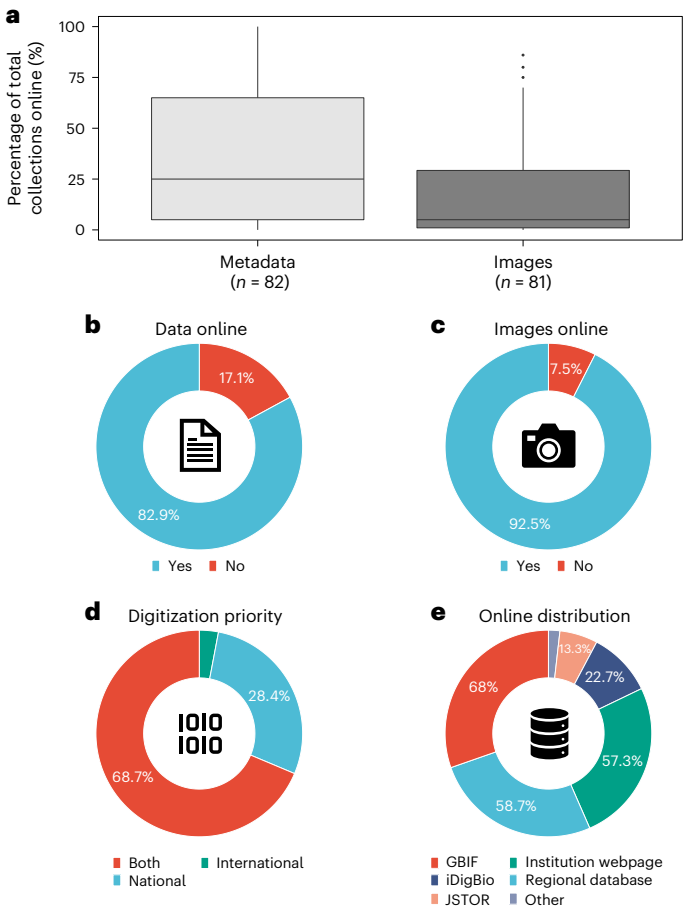

Fig. 4 | Trends in the digitization of herbarium specimens. a, The percentage of physical collections in data-providing herbaria that have at least location and date information (metadata; light grey) or digital images (dark grey) available online. The boxes span the first and third quartiles around the median. The whiskers depict the 1.5 × IQR unless minimum/maximum values fall within 1.5 × IQR of the quartiles. Points beyond the whiskers represent outliers. **b**,**c**, The percentage of herbaria that have at least some data (**b**) and images (**c**) of their collections shared online. **d**, The digitization priority of herbaria in terms of where specimens were collected. **e**, How the data-providing herbaria share and distribute digital specimen data (multiple answers were possible).

Indeed, most institutions gave equal priority to the digitization of national and international collections (Fig. 4d) and share their digitized specimen data with GBIF and/or regional databases that often also share data with GBIF (for example, the Consortium of California Herbaria, the Australasian Virtual Herbarium, eReColNat and Virtual Herbaria JACQ; Fig. 4e). Other aggregators such as JSTOR Global Plants and iDigBio share data with GBIF as well[34].

Our inquiries were sent to institutions that widely vary in their capacity and infrastructure. It is possible that institutions that supplied useable data were biased towards those with reliable internet connections, staff proficient in English and at least a portion of their collections already databased (that is, non-response bias), though our response rates were generally consistent across regions (Supplementary Fig. 4). The majority of herbaria that responded to our inquiries were located in developed countries. While this is reflective of the global distribution of herbarium collections (Supplementary Fig. 5), it also implies that we may be overestimating the state of specimen digitization and that the discrepancies in both physical and digital collections among previously colonized countries and their colonizers are much larger than our assessment suggests. Nonetheless, our collaborative assessments highlight the fact that we are still in the infancy of digitizing herbaria and thus have the opportunity to reassess how ongoing and future digitization and mobilization efforts can be organized to better address the colonial legacy of these collections and provide a deeper understanding of global plant diversity.

Discussion

Our study demonstrates a major disparity between where plant diversity naturally exists and where it is housed and catalogued. This renders much of the world reliant on botanical knowledge and resources housed and stewarded outside of their own borders. This disparity impacts not only the capacity for conservation and basic research but also commercial and government enterprises that seek to appropriate and monetize biological resources and their derivatives. In addressing this disparity, recent discussions regarding approaches to decolonizing cultural institutions, natural history museums and biogeographical practices in general could be applied to herbaria as well. We highlight some of these discourses below, though we stress that our study is only an initial step towards understanding how colonial activities have shaped herbaria and cannot adequately address the complexities of their colonial legacies.

First, as Das and Lowe note[35], it is important to acknowledge the colonial legacy of herbarium collections and to present the history and circumstance of these collections alongside existing interpretations about the specimens and their role in scientific research. They argue that such acknowledgment is a critical step towards bridging the gap between natural history collections and audiences in previously colonized nations and ensuring inclusiveness in the collection, curation and use of these collections. One way to openly share and communicate such narratives is via themed exhibitions and tours, such as the Black history tours of Hintze Hall and displays emphasizing unacknowledged collectors in the Darwin Centre of London's Natural History Museum or the First Nations-led and informed 'Unsettled' exhibition at the Australian Museum. These tours recognize and emphasize the (unrecognized) contributions of Indigenous peoples to the culture, science and natural history on display. Though most herbaria traditionally do not offer public exhibitions and herbarium specimens are rarely prominent in natural history museum displays (in part due to their fragility), increasing specimen digitization efforts have made it possible to curate digital exhibitions and virtual tours without competing for space and attention with others considered more charismatic (for example, large mammals and dinosaurs). Awareness and acknowledgement can also be facilitated by including positionality statements in grant proposals, research articles and other scientific communications that involve herbarium collections. Positionality statements describe the position of a researcher in relation to the social and political context of all phases of the research in question and are well established in the humanities but still rare in the natural sciences[36].

Second, we should continue to improve accessibility to the vast information held in herbaria worldwide, for both scientists and the public. Though digitizing and sharing specimen data is hardly a new idea, our collaborative assessment of herbarium collections suggests that the digitized data currently available online represent only a small portion of what resides inside herbarium cabinets. Indeed, several of our collaborators and authors noted that estimating the size and distribution of their collections is difficult—only a small portion of herbarium specimens have been databased by their respective institutions, and formal inventories seldom exist. Though massive digitization efforts have been funded, particularly at institutions in developed countries, even these institutions frequently lack funding for adequate curation and processing of specimens. Our analysis of available digital collection data also demonstrates that higher-level data products (that is, images) for many previously colonized areas are lacking (Supplementary Fig. 3). Digitization efforts focused on increasing representation from such areas could thus help bridge the reverse-latitudinal gradient of plant diversity knowledge. Furthermore, though much of the data that

have been digitized from herbarium collections are shared via open data infrastructures (for example, GBIF, iDigBio, BIEN, SpeciesLink, AVH and eReColNat), the portion that remains available only upon request (which can be denied), paywalled or inaccessible outside of specific groups is likely large (for example, limited access via Global Plants on JSTOR). Targeted initiatives and funding opportunities that prioritize the curation, digitization and sharing of collections from developing countries can be one way to address these discrepancies. There have been some promising efforts along these lines, such as the NSF-supported GLOBAL Bryophyte & Lichen Thematic Collections Network, GBIF's Biodiversity Information for Development programme, the United Kingdom's Darwin Initiative and the Mellon Foundation's African Plants Initiative[37]. We can also increase support for loan and exchange programmes across herbaria, facilitating access and repatriation of physical specimens as well. Such efforts must be mindful of the legacy of some herbarium collections. For instance, specimen returns in accordance with permits or agreements are traditionally referred to as 'gifts', but it may be preferable to use a different term, such as 'returns'[38]. We should also be mindful that specimens can contain biocultural information that is inappropriate for broader circulation and can risk further exploitation of Indigenous cultural knowledge. Thus, efforts to improve accessibility to botanical collections and share knowledge therein require careful discourse for all parties involved.

Third, in addition to recognizing the sovereignty of a nation's biological resources and that biodiversity can be best studied where it occurs[39], capacity-building in previously colonized countries through the sharing of tools and knowledge for contributing towards research is critical—if the science resulting from collections is globally relevant, the means of contributing should be distributed as such[36,40]. In particular, it is crucial to ensure that local contributions are sufficiently recognized and facilitate the development of local research priorities and agendas during this process. Acknowledging the providing country agency and personnel in all aspects from specimen labels to publication authorship to grant proposals is essential. Furthermore, the digital products of herbarium specimen data could be hosted and managed by researchers in the countries where they were originally collected as a form of repatriation, who could be trained and supported as necessary by institutions with greater capacity. Although the latter might not dispose of the necessary funding to support the local partner, they could play a major role when a grant request is addressed to an international agency, clearly stating their engagement in the transfer of technical and scientific knowledge. International collectors should be mindful to leave duplicate specimens in the host country—this practice has become increasingly the norm over recent decades, and at times enforced by local governments, especially since the Convention on Biological Diversity was signed in 1992 and the Nagoya Protocol on Access and Benefit Sharing was drafted in 2010 (https://www.cbd.int/abs/). Other forms of (inter)national regulations requiring the deposition of duplicate specimens have existed for longer and have largely been abided by major institutions even in non-signatory countries such as the United States. Still, many regions lack the facilities to properly store and curate collected specimens. In such cases, collectors could gather and treat duplicate specimens as loans until the necessary local infrastructure is established. This would in turn facilitate a more equitable, global view towards the collection, curation and use of herbarium specimens. We note that the term 'duplicate', though widely used and officiated by the International Association for Plant Taxonomy, implies that a separate original specimen exists elsewhere. Along these lines, we recommend that the country of origin should be the preferred place of deposition of type specimens (that is, specimens on which the description and name of a taxon is based) and those without duplicates (that is, unicates). Furthermore, developed nations could help establish the necessary infrastructure locally as part of collaborative endeavours with developing nations. To support such efforts, we strongly recommend that grant proposals involving the collection, curation and digitization of specimens associated with developing countries include requests for funding to support local colleagues and collaborators where possible. Institutions, scientists and funding agencies need to seek ways to expand opportunities for partners in providing countries to participate in research design and grant application, in addition to activities directly pertaining to the collection and curation of specimens. In turn, funding bodies should recognize the need to support local partners appropriately and guarantee access to the knowledge and benefits arising from plant collections sampled abroad. Importantly, these and other efforts to address the colonial history of herbaria should be guided by the needs and wishes of people who lived under colonial rule. One example of such a partnership can be found in a recent project to sequence and study the genome of the tuatara, a cultural treasure of the Māori people[41]. The Indigenous peoples provided access to the species and associated knowledge and were involved in all decision-making regarding the use of the genomic data generated by the study and any benefits that may accrue.

A profound set of challenges lie ahead if we are to address the still-persistent legacy of colonialism in our plant collections. Many of the examples and suggestions we discuss above require access to opportunities, resources and infrastructure (for example, internet, international travel and knowledge of a European language) that are not always widely available to people who lived under colonial rule. We emphasize that any such endeavours should centrally involve local and Indigenous peoples, and their knowledge, culture and interests should be respected and acknowledged. Ongoing digitization efforts have offered us new avenues of exchanging knowledge and infrastructure and sharing the benefits arising from the utilization of herbarium collections. These efforts can (and should) provide opportunities for the accommodation of Indigenous knowledge structures alongside the systems of classification, nomenclature and biodiversity data organization used in herbaria. Science is not exempt from sociopolitical realities, and we should not avert our gaze from the origins of these otherwise precious resources. To this end, we have endeavoured to provide a glimpse into the extent of the colonial legacy that provides the dominant and politically inflected shape for our herbarium collections. Embracing these realities represents the first step towards a more inclusive and expansive global herbarium.

Methods

We downloaded plant specimen data (kingdom, Plantae; basis of record, preserved specimen) from GBIF on 23 April 2021 (https://doi.org/10.15468/dl.nt5wkx). We only kept specimen records with accepted scientific names, valid country codes and publishing country names. With the remaining 50,303,354 records, we compiled a country-by-country matrix that summarized the number of specimens collected from one country and housed in another. The country where a specimen was collected was based on the field 'countryCode', and the country where a specimen was housed was based on the field 'publishingCountry'. We also grouped the country-by-country matrix into a continent-by-continent matrix. To examine the temporal trends of collection, we further examined the data after separating them into two subsets—before and after 1945, which marks the end of World War II and the era of overt colonialism. We finally verified our analyses on a subset of data that (1) had coordinates, (2) had the 'countryCode' field matching the location inferred from the coordinates and (3) were determined to be without geospatial issues by GBIF. Assessments of GBIF data at the species level were limited to specimens with full binomial species names, and the accepted scientific names following GBIF's taxonomic system were applied.

As records on GBIF represent a subset of the collections in herbaria across the world, we expanded our investigations to physical institutions. We sent out an informal request for information regarding the number and origin of plant collections in 2020 to major herbaria across the world as listed by *Index Herbariorum* (http://sweetgum.nybg.org/

science/ih/) and select representative regional herbaria. We distributed these queries to herbaria that fall within the top 100 size classes for collection size using Qualtrics (version 2020–2021). We also selected the largest herbaria from regions not well represented on this list (for example, Africa) and smaller herbaria from regions where the largest herbaria did not supply usable responses. These institutions do not represent a random sampling of global herbaria. Questions were focused on identifying the size of the collections, where they were collected and the proportion digitized (Supplementary Data 1). Links were sent out to curators, collection managers and directors listed as the point of contact for each institution on *Index Herbariorum* or institutional websites via email in August 2020. Monthly reminders were sent for a year, and the Qualtrics link was kept active until October 2021 to maximize participation. Those who provided relevant information were invited to collaborate on the study and offered authorship, but incentives were otherwise not used. A total of 172 institutions across 50 countries were contacted, and 92 herbaria across 39 countries and 6 continents submitted at least partial responses to our queries (response rate, 54%; Supplementary Fig. 5 and Supplementary Data 2). We did not weight the data collected by our informal assessment of herbaria, as we aimed to describe and present the state of collections as-is using as much data as possible.

We recognize that certain assumptions were made in our study. First, the Western scientific system is not the only way to understand and describe botanical knowledge, and though many of our discussions pertain to such as it is broadly adopted, we do not mean to devalue or reject other knowledge systems. The fact that most if not all (identified) specimens in databases such as GBIF are only assigned names following the Western scientific system (that is, Linnaean taxonomy) necessitated our use of this standardized nomenclature, but it is also stark proof of the imbalances wrought over centuries of colonialism. The idea that these are universally applicable can perpetuate an imperial mindset and undermine other ways of knowing. Though our work represents the views and values of scientists from over 30 countries spanning every continent, we recognize that these perspectives may not be shared universally and are based on the authors' collective knowledge and experiences. The examples we present do not represent the full extent of colonial activities and botanical collections by imperial powers. The practice of overt colonialism was not limited to European nations and their former colonies, and there are far more instances of colonialism by various entities than can be listed here. Moreover, the knowledge gathered through colonial activities comprises a complex history of "worlds and visions brought into contact" that cannot be characterized as a simple two-dimensional landscape[42]. Second, we use geopolitical constructs that are not free from the influence of colonialism. For instance, though we treat Australia as a single entity, it is home to over 500 Aboriginal nations. Our assessment of specimen movement thus does not encompass the appropriation of botanical collections and knowledge from Indigenous peoples within the confines of such entities. Finally, though we posit that the era of overt colonialism has ended, we realize that there was no single process of decolonization and that the idea that colonization is over can be problematic as its legacy persists to this day, even in botanical collections. Along these lines, here we use the term 'colonization' in a fairly general sense to describe a relationship between two countries, independent of their level of development, in which one has subjugated and governed the other over a period of time, contributing to the current state of its institutions (following ref. 43).

Positionality statement

The authors position themselves first and foremost as scientists who work with botanical collections. We acknowledge that our work reflects the biases of the Western scientific system that we rely on to understand and interpret the botanical world. Although our perspectives encompass those from a diverse array of backgrounds, cultures and orientations, we recognize that our science, and the institutions in which we conduct it, may have been shaped by imperialist thinking and colonial endeavours that privilege Western knowledge above other ways of knowing.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data discussed in the paper are either publicly available through GBIF (https://www.gbif.org/; https://doi.org/10.15468/dl.nt5wkx) or *Index Herbariorum* (https://sweetgum.nybg.org/science/wp-content/uploads/2021/01/The_World_Herbaria_2020_7_Jan_2021.pdf) or are in the Supplementary Information.

Code availability

The code used for data analysis is available at https://github.com/shandongfx/paper_specimen_2023.

References

1. Thiers, B. M. *The World's Herbaria 2020: A Summary Report Based on Data from Index Herbariorum* Issue 5.0 (NYBG Steere Herbarium, 2021).
2. Willis, C. G. et al. Old plants, new tricks: phenological research using herbarium specimens. *Trends Ecol. Evol.* **32**, 531–546 (2017).
3. Heberling, J. M., Miller, J. T., Noesgaard, D., Weingart, S. B. & Schigel, D. Data integration enables global biodiversity synthesis. *Proc. Natl Acad. Sci. USA* **118**, e2018093118 (2021).
4. Hedrick, B. et al. Digitization and the future of natural history collections. *Bioscience* **70**, 243–251 (2020).
5. Funk, V. A. 100 uses for an herbarium: well at least 72. *Am. Soc. Plant Taxon. Newsl.* **7**, 17–19 (2003).
6. Carine, M. A. et al. Examining the spectra of herbarium uses and users. *Bot. Lett.* **165**, 328–336 (2018).
7. Meineke, E. K., Classen, A. T., Sanders, N. J. & Davies, T. J. Herbarium specimens reveal increasing herbivory over the past century. *J. Ecol.* **107**, 105–117 (2019).
8. Lang, P. L. M., Willems, F. M., Scheepens, J. F., Burbano, H. A. & Bossdorf, O. Using herbaria to study global environmental change. *N. Phytol.* **221**, 110–122 (2019).
9. Lavoie, C. Biological collections in an ever changing world: herbaria as tools for biogeographical and environmental studies. *Perspect. Plant Ecol. Evol. Syst.* **15**, 68–76 (2013).
10. Davis, C. C., Willis, C. G., Connolly, B., Kelly, C. & Ellison, A. M. Herbarium records are reliable sources of phenological change driven by climate and provide novel insights into species' phenological cueing mechanisms. *Am. J. Bot.* **102**, 1599–1609 (2015).
11. Bonal, D. et al. Leaf functional response to increasing atmospheric CO_2 concentrations over the last century in two northern Amazonian tree species: a historical $\delta^{13}C$ and $\delta^{18}O$ approach using herbarium samples. *Plant. Cell Environ.* **34**, 1332–1344 (2011).
12. Rudin, S. M., Murray, D. W. & Whitfeld, T. J. S. Retrospective analysis of heavy metal contamination in Rhode Island based on old and new herbarium specimens. *Appl. Plant Sci.* **5**, 1600108 (2017).
13. Peñuelas, J. & Filella, I. Herbaria century record of increasing eutrophication in Spanish terrestrial ecosystems. *Glob. Change Biol.* **7**, 427–433 (2001).
14. Lees, D. C. et al. Tracking origins of invasive herbivores through herbaria and archival DNA: the case of the horse-chestnut leaf miner. *Front. Ecol. Environ.* **9**, 322–328 (2011).

15. Delisle, F., Lavoie, C., Jean, M. & Lachance, D. Reconstructing the spread of invasive plants: taking into account biases associated with herbarium specimens. *J. Biogeogr.* **30**, 1033–1042 (2003).
16. Vogel, G. Natural history museums face their own past. *Science* **363**, 1371–1372 (2019).
17. Wintle, C. Decolonizing the Smithsonian: museums as microcosms of political encounter. *Am. Hist. Rev.* **121**, 1492–1520 (2016).
18. Wandersee, J. H. & Schussler, E. E. Preventing plant blindness. *Am. Biol. Teach.* **61**, 82–86 (1999).
19. Brockway, L. H. Science and colonial expansion: the role of the British Royal Botanic Gardens. *Am. Ethnol.* **6**, 449–465 (1979).
20. Shellam, T., Nugent, M., Konishi, S. & Cadzow, A. *Brokers and Boundaries: Colonial Exploration in Indigenous Territory* (ANU Press, 2016).
21. Kennedy, D. *The Last Blank Spaces* (Harvard Univ. Press, 2013).
22. Reynolds, H. *With the White People* (Penguin, 1990).
23. Peterson, A. T., Soberón, J. & Krishtalka, L. A global perspective on decadal challenges and priorities in biodiversity informatics. *BMC Ecol.* **15**, 15 (2015).
24. Bakker, F. T. et al. The Global Museum: natural history collections and the future of evolutionary science and public education. *PeerJ* **8**, e8225 (2020).
25. Davis, C. C. The herbarium of the future. *Trends Ecol. Evol.* **38**, 412–423 (2022).
26. Voeks, R. A. *The Ethnobotany of Eden: Rethinking the Jungle Medicine Narrative* (Univ. of Chicago Press, 2018).
27. Webb, C. O., Slik, J. W. F. & Triono, T. Biodiversity inventory and informatics in Southeast Asia. *Biodivers. Conserv.* **19**, 955–972 (2010).
28. Hughes, A. C. et al. Sampling biases shape our view of the natural world. *Ecography* **44**, 1259–1269 (2021).
29. Brown, J. H. Why are there so many species in the tropics? *J. Biogeogr.* **41**, 8–22 (2014).
30. Paton, A. et al. Plant and fungal collections: current status, future perspectives. *Plants People Planet* **2**, 499–514 (2020).
31. Daru, B. H. et al. Widespread sampling biases in herbaria revealed from large-scale digitization. *N. Phytol.* **217**, 939–955 (2018).
32. Drew, J. A., Moreau, C. S. & Stiassny, M. L. J. Digitization of museum collections holds the potential to enhance researcher diversity. *Nat. Ecol. Evol.* **1**, 1789–1790 (2017).
33. Breckenridge, K. The politics of the parallel archive: digital imperialism and the future of record-keeping in the age of digital reproduction. *J. South. Afr. Stud.* **40**, 499–519 (2014).
34. Feng, X. et al. A review of the heterogeneous landscape of biodiversity databases: opportunities and challenges for a synthesized biodiversity knowledge base. *Glob. Ecol. Biogeogr.* **31**, 1242–1260 (2022).
35. Das, S. & Lowe, M. Nature read in black and white: decolonial approaches to interpreting natural history collections. *J. Nat. Sci. Collect.* **6**, 4–14 (2018).
36. Eichhorn, M. P., Baker, K. & Griffiths, M. Steps towards decolonising biogeography. *Front. Biogeogr.* **12**, e44795 (2019).
37. Smith, G. F. The African Plants Initiative: a big step for continental taxonomy. *Taxon* **53**, 1023–1025 (2004).
38. Rabeler, R. K. et al. Herbarium practices and ethics, III. *Syst. Bot.* **44**, 7–13 (2019).
39. Tydecks, L., Jeschke, J. M., Wolf, M., Singer, G. & Tockner, K. Spatial and topical imbalances in biodiversity research. *PLoS ONE* **13**, e0199327 (2018).
40. Asase, A., Mzumara-Gawa, T. I., Owino, J. O., Peterson, A. T. & Saupe, E. Replacing 'parachute science' with 'global science' in ecology and conservation biology. *Conserv. Sci. Pract.* **4**, e517 (2021).
41. Gemmell, N. J. et al. The tuatara genome reveals ancient features of amniote evolution. *Nature* **584**, 403–409 (2020).
42. Roque, R. & Wagner, K. A. in *Cambridge Imperial and Post-colonial Studies Series: Engaging Colonial Knowledge* (eds Roque, R. & Wagner, K. A.) 1–32 (Palgrave Macmillan, 2012).
43. Mayer, T. & Zignago, S. *Notes on CEPII's Distances Measures: The GeoDist Database* (SSRN, 2011); https://doi.org/10.2139/ssrn.1994531

Acknowledgements

We acknowledge the past and continuing contributions of colonized peoples to botanical science and knowledge. This work was largely conducted in institutions on the traditional territory of the Wampanoag, Massachusetts, Miami and Apalachee peoples. We thank the herbaria that contributed data to this work: AAU, AD, ASSAM, ASU, B, BISH, BM, BO, BP, BR, BRI, BRNU, BSA, BSHC, BSID, C, CANB, CAS, CHR, CL, CM, COL, CORD, DAO, DUKE, EA, F, FI, FLAS, FR, G, GB, GH, A, ECON, AMES, FH, NEBC, GJO, GOET, GZU, H, HAL, HIRO, IBSC, SI, K, KH, KW, L, LBL, LE, LIL, LY, M, MA, MEL, MEXU, MICH, MIN, MO, MT, MW, NCU, KB/NIBR, NICH, NSW, NY, O, OULU, P/PC, PAD, PE, PERTH, PRC, PRE/NBG/NH, QFA, RM/USFS, RO, S, SI, SING, SP, STR, TAIF, TASH, TENN, TEX/LL, TNS, TUR, UC, JEPS, UPS, US, USM, VEN, W, WTU and ZT. This work was supported by the Czech Academy of Sciences (grant no. RVO 67985939 to J.D.) and the Komarov Botanical Institute, RAS (grant no. AAAA-A19-119031290052-1 to D. Melnikov). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

D.S.P. and C.C.D. conceived the initial idea, which was refined through discussions with X.F. D.S.P and X.F. analysed the specimen data and designed the questionnaire with C.C.D. D.S.P. supervised the study and wrote the original draft with input from X.F. All other authors provided data from their respective institutions. All authors contributed to further revising the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-023-01616-7.

Correspondence and requests for materials should be addressed to Daniel S. Park or Charles C. Davis.

Peer review information *Nature Human Behaviour* thanks the anonymous reviewers for their contribution to the peer review of this work.

[1]Department of Biological Sciences, Purdue University, West Lafayette, IN, USA. [2]Purdue Center for Plant Biology, Purdue University, West Lafayette, IN, USA. [3]Department of Geography, Florida State University, Tallahassee, FL, USA. [4]Department of Botany, National Museum of Nature and Science, Tsukuba, Japan. [5]Herbarium Bogoriense, Research Center for Biosystematics and Evolution, National Research and Innovation Agency (BRIN), Cibinong, Indonesia. [6]Instituto Experimental Jardin Botánico 'Dr. Tobías Lasser', Avenida Salvador Allende, Ciudad Universitaria, Caracas, Venezuela. [7]Universidad Central de Venezuela, Caracas, Venezuela. [8]Hungarian Natural History Museum, Budapest, Hungary. [9]Herbier LY, FR-BioEEnVis, Université Claude Bernard Lyon 1, Lyon, France. [10]Instituto de Botánica Darwinion, San Isidro, Argentina. [11]Universidad Nacional de Colombia, Bogotá, Colombia. [12]Botany Section, Naturalis Biodiversity Center, Leiden, the Netherlands. [13]Meise Botanic Garden, Meise, Belgium. [14]Herbario San Marcos, Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, Lima, Peru. [15]Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic. [16]Institute of Botany of the Czech Academy of Sciences, Průhonice, Czech Republic. [17]Central Regional Centre, Botanical Survey of India, Allahabad, India. [18]University of Washington Herbarium, Burke Museum, Seattle, WA, USA. [19]Sikkim Himalayan Regional Centre, Botanical Survey of India, Gangtok, India. [20]Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland. [21]Botany Unit, Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland. [22]Department of Biodiversity, Conservation and Attractions, Western Australian Herbarium, Kensington, Western Australia, Australia. [23]South African National Biodiversity Institute, Pretoria, South Africa. [24]School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Johannesburg, South Africa. [25]Department of Biological Sciences, Faculty of Science and Technology, Kochi University, Kochi, Japan. [26]Hattori Botanical Laboratory, Nichinan, Japan. [27]Harvard University Herbaria, Cambridge, MA, USA. [28]Institute of Botany, Uzbekistan Academy of Sciences, Tashkent, Uzbekistan. [29]National Institute of Ecology, Seocheon, South Korea. [30]National Parks Board, Singapore Botanic Gardens, Singapore, Singapore. [31]Department of Environmental Biology, Sapienza University of Rome, Rome, Italy. [32]Herbarium Patavinum, University of Padua, Padua, Italy. [33]East African Herbarium, National Museums of Kenya, Nairobi, Kenya. [34]Komarov Botanical Institute, Russian Academy of Sciences, Saint Petersburg, Russian Federation. [35]Herbarium Collections & Department of Botany, Faculty of Science, Charles University, Prague, Czech Republic. [36]Institute of Biological Sciences, Maria Curie-Skłodowska University, Lublin, Poland. [37]Rocky Mountain Herbarium, University of Wyoming, Laramie, WY, USA. [38]Field Museum, Chicago, IL, USA. [39]Staatliche Naturwissenschaftliche Sammlungen Bayerns, Botanische Staatssammlung München, München, Germany. [40]Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA. [41]Royal Botanic Gardens, Kew, Surrey, UK. [42]Herbier Louis-Marie, Université Laval, Québec City, Québec, Canada. [43]University of Florida Herbarium, Florida Museum, Gainesville, FL, USA. [44]Instituto Multidisciplinario de Biología Vegetal (UNC-CONICET), Córdoba, Argentina. [45]Naturhistorisches Museum Wien, Vienna, Austria. [46]Department of Botany and Biodiversity Research, University of Vienna, Vienna, Austria. [47]Deccan Regional Centre, Botanical Survey of India, Hyderabad, India. [48]National Herbarium of Ukraine, M.G. Kholodny Institute of Botany, National Academy of Sciences of Ukraine, Kyiv, Ukraine. [49]Natural History Museum, University of Oslo, Oslo, Norway. [50]Missouri Botanical Garden, St. Louis, MO, USA. [51]Conservatory and Botanic Gardens of Geneva, Geneva, Switzerland. [52]Department of Plant Sciences, University of Oxford, Oxford, UK. [53]New York Botanical Garden, Bronx, NY, USA. [54]Graduate School of Integrated Sciences for Life, Hiroshima University, Hiroshima, Japan. [55]Miyajima Natural Botanical Garden, Hiroshima University, Hiroshima, Japan. [56]Royal Botanic Gardens Victoria, Melbourne, Victoria, Australia. [57]Ecology and Genetics Research Unit, University of Oulu, Oulu, Finland. [58]University of Oulu Botanical Museum, Oulu, Finland. [59]Bell Museum, University of Minnesota, St. Paul, MN, USA. [60]South China Botanical Garden Herbarium, Chinese Academy of Sciences, Guangzhou, China. [61]Department of Organismic and Evolutionary Biology, Harvard University Herbaria, Harvard University, Cambridge, MA, USA. [62]These authors contributed equally: Daniel S. Park, Xiao Feng. ✉e-mail: danielpark@purdue.edu; cdavis@oeb.harvard.edu

nature portfolio

Corresponding author(s): Daniel S. Park
Charles C. Davis

Last updated by author(s): Mar 27, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	We used the Qualtrics platform (version 2020–2021) to collect data from collaborators.
Data analysis	Data were analyzed with R version 4.0.3.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Data discussed in the paper are either publicly available through GBIF (https://www.gbif.org/; https://doi.org/10.15468/dl.nt5wkx), Index Herbariorum (https://sweetgum.nybg.org/science/wp-content/uploads/2021/01/The_World_Herbaria_2020_7_Jan_2021.pdf), or attached supplements.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We investigate the colonial legacy of botanical collections and discuss how we may move towards a more inclusive future by examining over 85 million online specimen records and assessing herbarium collections across the globe.
Research sample	Herbarium collections
Sampling strategy	We examined all plant specimen records publicly available on the Global Biodiversity Information Facility (GBIF), currently the largest biodiversity data infrastructure, and assessed the state of herbarium collections across the world with data provided by our collaborators/coauthors, who work in these institutions. We sent out an informal request for collaboration regarding assessing the number and origin of plant collections to herbaria that fall within the top 100 size classes across the world as listed by Index Herbariorum (http://sweetgum.nybg.org/science/ih/). We also selected the largest herbaria from regions not well represented on this list (e.g. Africa) and smaller herbaria from regions where the largest herbaria did not supply usable responses.
Data collection	Plant specimen data were directly downloaded from the GBIF website. Collaborators shared estimates of their institutions' collections through the Qualtrics platform (version 2020–2021).
Timing and spatial scale	The spatial breadth of the data used in the study is global. Data were downloaded from GBIF on April 23, 2021. Collaborators provided (updated) information on their collections throughout the duration of the project (August, 2020 onwards).
Data exclusions	We only examined specimen records from GBIF with accepted scientific names, valid country code and publishing country names - the minimum information necessary to address our questions.
Reproducibility	No experiments were conducted. We provide a DOI for the publicly available data used.
Randomization	N/A
Blinding	Blinding was not relevant to our study as the data involved were either publicly available or non-personal in nature (i.e. specimen counts).

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging