Ancient DNA Genomics and the Renaissance of Herbaria

Hernán A. Burbano[1†]* and Rafal M. Gutaker[2†]

[1]Centre for Life's Origins and Evolution, Department of Genetic, Evolution and Environment, University College London, London, WC1E 6BT, UK.
[2]Royal Botanic Gardens, Kew, Kew Green, Richmond, Surrey, TW9 3AE, UK.
[†]These authors contributed equally to this work.
*Corresponding author: h.burbano@ucl.ac.uk (H.A.B).

Abstract

Herbaria are undergoing a renaissance as valuable sources of genomic data for exploring plant evolution, ecology, and diversity. Ancient DNA retrieved from herbarium specimens can provide unprecedented glimpses into past plant communities, their interactions with biotic and abiotic factors, and the genetic changes that have occurred over time. Here, we highlight recent advances in the field of herbarium genomics and discuss the challenges and opportunities of combining data from modern and time-stamped historical specimens. We also describe how integrating herbarium genomics data with other data types can yield substantial insights into the evolutionary and ecological processes that shape plant communities. Herbarium genomic analysis is a tool for understanding plant life and informing conservation efforts in the face of dire environmental challenges.

Teaser

A review on herbarium genomics to understand the history of plant evolution and ecological interactions.

Introduction

Herbarium specimens, meticulously collected and pressed plant samples, preserve a tangible record of botanical diversity and have long served as a foundation for botanical, taxonomical, and systematics studies (*1*). Hosted across 3,000 herbaria, these resources encompass close to 390 million specimens and their associated metadata (*2*). We owe this number to centuries of specimen custodianship and current curation that ensures continuous growth, preservation, and sustainable use of this rich collection. Alongside their uses for botanical monographs and systematics (*3*), herbarium specimens covering diverse taxa and all continents increase the power of macroevolutionary studies, enabling investigations into trait evolution (*4*) and plant family radiations (*5*).

Recent advances in high-throughput sequencing technologies have facilitated the retrieval of genome-scale data from herbarium specimens (*6*). Herbarium-derived DNA can be classified as ancient DNA (aDNA) by its biochemical characteristics; it is highly fragmented with distinct patterns of DNA misincorporations and breakage (*7*, *8*). Such patterns result from post-mortem DNA damage, which is accelerated during specimen preparation (*9*). DNA

damage does not reflect the evolutionary history of specimens, but allows authentication of the historical nature of the herbarium-derived DNA (*7*, *10*). Additionally, herbarium-derived DNA originates not only from the plant host but also its associated commensal or pathogenic microbiomes. While distinguishing between the genuine microbiome and post-mortem colonizers presents a methodological challenge (*11*), if resolved, herbarium metagenomics holds the potential to provide valuable insights into past biotic interactions.

Lack of cell and genome integrity makes functional and structural genomic tools such as chromatin conformation capture unsuitable for historical tissues, while the fragmented nature of DNA derived from herbarium specimens limits de novo assembly of full herbarium genomes with comparable quality to modern genomes. However, short-read sequencing approaches have been developed to generate genome-scale data from herbarium specimens (*12*, *13*). While whole-genome sequencing allows genome-wide scans for selection to address questions related to the genetic basis of adaptation, it can be substituted by reduced-representation techniques for addressing population history questions. By contrast, cost-efficient techniques such as enrichment capture of conserved genes and plastome sequencing are informative for macroevolutionary and systematics studies. Amplicon sequencing approaches have been reviewed exhaustively elsewhere (*14*) and are outside the scope of this article.

Despite the shortcomings associated with the degraded nature of herbarium-derived DNA, genomic studies of herbarium specimens offer a wide taxonomic and geographic sampling breadth to directly investigate the evolutionary responses of plants and their associated microbiomes over the last 300 years of anthropogenic global environmental change.

Contributions of herbarium genomics to taxonomic and evolutionary studies

Classification and macroevolutionary inference

The integration of sequencing data with morphological data has substantially increased our ability to group plants into taxonomic units. In botany, species definition comes from type specimens, which were used to describe their characteristics in the first place. Once the genetic composition of type specimens is established, a newly-collected plant can be identified to the species level. A prominent approach involves coupling genome skimming with chloroplast assembly or molecular barcode readout, allowing identification at the genus and species levels with high congruence with morphological identification (*15*). Such an approach was used to generate plastid and mitochondrial genomes from type specimens of red algae species (*Pyropia*) and allow molecular identification of unclassified material (*16*). Another challenge is to group species into higher taxonomic ranks, which often requires both morphological and genetic characterization. For example, genome-wide data helped resolve the challenging sweet potato genus (*Ipomea*) (*17*).

Herbarium collections harbor a wide range and diversity of species as well as geographies, which would not otherwise be accessible within the timeframe of contemporary research projects. This is particularly useful for large-scale phylogenomic projects such as the Plant and Fungal Tree of Life, in which over 28% of the taxa examined were sampled from herbarium specimens and were assessed using enrichment capture approach (*18*). The resulting phylogenetic trees inform our understanding of macroevolution and the history of

plant life on Earth. For example, this approach contributed to resolving the relationships of taxa in the mulberry family (*Moraceae*), and indicated repeated, parallel evolution of inflexed stamens that facilitate wind pollination (*19*). Other uses include facilitating predictions of plant-based compounds, with a special role for herbarium collections in predicting active compounds for medicinal use (*20*).

Preserving a permanent record of the source materials from which DNA was retrieved is also important. Consequently, large biogenome projects, such as the Darwin Tree of Life, which aim to build reference genome assemblies for vast quantities of species, have partnered with institutions that generate voucher herbarium specimens for the plant taxa sequenced (*21*).

Tracing the evolutionary history of wild plants

Herbarium collections offer valuable spatiotemporal data for investigating the genetic mechanisms underlying plant responses to global environmental change, including the influence of modern agricultural practices (*22*, *23*). With escalating global movement and trade, plants have become increasingly transported over long distances, heightening the likelihood of introduction of non-native species that might evolve into invasive species (*24*). Herbarium specimens have traditionally been employed as proxies for documenting first-plant colonizers (Fig. 1). Sequences from herbarium specimens over dense time series can test hypotheses in invasion biology and determine whether changes in phenotypic traits are genetically encoded. Genetic analysis of herbarium specimens of garlic mustard (*Alliaria petiolata*) spanning 50 years of North American invasion suggested that differential secretion of phytochemicals across this time period had an adaptive genetic basis (*25*). Herbarium genomics also reconstructed the North American population history of a prevalent colonizing semiclonal lineage of thale cress (*Arabidopsis thaliana*) (*26*). A phylogenetic approach determined the time of colonization and connected the maintenance of de novo mutations in this lineage to root traits and flowering time that are likely important for the ecology of invasions. Human-induced alterations of landscapes has led to genetic mixing in the native range of common ragweed (*Ambrosia artemisiifolia*) in North America, before its introduction to Europe, potentially playing a role in its subsequent success as an invasive species (*27*). Furthermore, sequencing of numerous present-day and historical herbarium genomes of ragweed revealed that European populations exhibit a significant turnover in genetic structure. This pattern is likely attributed to multiple introductions, founder effects, and genetic admixture with closely related European species (*28*).

Anthropogenic activities additionally generate habitats for wild plants. Agricultural fields face challenges from competition by weed species, with heavy use of herbicides to minimize interspecies competition and maintain crop yields. Excessive use of herbicides since their introduction in the 1950s has resulted in continuous evolution of weeds with resistance to these formerly deadly chemicals (Fig. 1) (*29*). Devising effective strategies for managing herbicide resistance necessitates determining whether resistance mutations originate *de novo* or were present in the existing genetic variation, and to what extent gene flow plays a role in spreading herbicide resistance. Genotyping herbarium specimens of blackgrass (*Alopecurus myosuroides*) revealed that a herbicide resistance mutation located in the direct target of herbicides (targeted site resistance) was segregating at a very low frequency approximately one century before the introduction of herbicides (*30*). A recent genomics study using

contemporary samples also suggested that standing genetic variation plays a crucial role in facilitating the rapid evolution of targeted site herbicide resistance (*31*). In addition to developing herbicide resistance, weeds must also adapt to the human-made environment of modern agricultural fields. A recent study of waterhemp (*Amaranthus tuberculatus*), an agricultural weed native to North America, employed geographical pairing of hundreds of present-day genomes from natural and agricultural populations and a time-series of hundreds of corresponding herbarium genomes (*32*). These allowed monitoring of changes in allele frequencies over time and estimation of selection coefficients of variants observed predominantly in agricultural settings.

A time machine for recent crop evolution

As with wild plants, we can use time-stamped herbarium specimens of crop plants to address many important questions about their movement, adaptation, and changes in genetic diversity. Crop herbaria span a period of over 500 years, with the oldest DNA sequences retrieved from a tomato specimen (*Solanum lycopersicum*) dated to circa 1558 (*33*). This time interval overlaps with the Columbian "exchange" of crops globally, a serendipitous test for the limits of environmental adaptation (Fig. 1). This period saw en masse movement of European staples to colonies in the Americas, the introduction of American crops to Eurasia, and the global movement of cash crops for colonial exploitation. All continents were now connected by maritime routes, which resulted in rapid introduction of new crops, livestock and diseases with unprecedented effect on people's livelihoods (e.g. illness) and the environment (e.g. agricultural pressures) (*34*). Herbarium specimens helped settle a long debate about the origins and adaptation of potato (*Solanum tuberosum*) when introduced from the Americas to Eurasia. Sequencing of herbarium specimens revealed the genetic turnover of introduced varieties over time (*35*), and showcased the increasing role of hybridization with geographically distinct species that came into direct contact on the European continent (*36*). Putative preadaptation through phytohormonal regulation preceded an important change in day length-mediated tuberization that allowed potatoes to be cultivated in the completely new European photoperiod regime (*36*).

The last 250 years have seen increases in agricultural productivity. From the 'British Agricultural Revolution' in the late 1700s to the Green Revolution in Latin America and South and Southeast Asia in the mid-1900s, new agrarian practices were accompanied by the adoption of new crop varieties (Fig. 1). Despite the general assumption that these processes had detrimental effects on the genetic diversity of crops, sequencing of seed bank materials has provided little evidence supporting massive genetic diversity loss (*37*). Genome sequencing of time-stamped crops from herbarium will yield a direct measure of genetic diversity change, together with other demographic parameters.

The postulated narrow gene pool available for crop breeding and the demand for genetic variants conferring stress tolerance, as well as recent developments in the field of plant 'neodomestication', have sparked interest in the genetic diversity of crop wild relatives (*38*). In addition to the detrimental effects of land use change and habitat fragmentation on genetic diversity, reports suggest that wild relatives of major crops are heavily influenced by gene flow from cultivated counterparts, even in self-pollinating species such as rice (*Oryza sativa*) (*39*). This raises concerns regarding genetically modified organisms where, for example, herbicide resistance variants could be passed to wild and weedy species.

Comparative genomics of time-stamped herbarium specimens of crop wild relatives will provide a direct measure of crop-to-wild gene flow risks over centuries.

Herbarium specimens as time capsules of ecological interactions

Shotgun sequencing of herbarium specimens enables characterization of not just plant genomes but also genomes belonging to associated organisms. These include plant pathogens and the diverse microbial communities residing in the rhizosphere and phyllosphere, habitats for microorganisms on the plant's aerial surface and at the root–soil interface, respectively. Sporadically, other organisms such as insects, molluscs, and even small reptiles can be found in leaf creases. Therefore, herbarium specimens serve as windows into past biotic interactions, revealing coevolutionary trajectories between plants and their associated organisms (Fig. 2A).

This field initially focused on detecting filamentous crop pathogens, such as fungi and water molds (oomycetes), stemming from the availability of infected plant herbarium specimens. These were collected by early plant pathologists and are curated alongside micro and macro fungi in specialized collections known as fungaria. The Irish potato famine pathogen *Phytophthora infestans* was detected by PCR amplification of a short fragment of its mitochondrial DNA genome (*40*). Sequencing whole genomes from multiple historical samples considerably expanded the study of late blight, with a pandemic clonal lineage of *P. infestans* identified as the causative agent triggering the Irish potato famine in the mid-1800s (*41*, *42*). The strain from the 19th century was completely replaced by other strains in the 20th century, although the 19th-century mitochondrial DNA is still in circulation (*43*). Combined analysis of historical and modern samples ascertained that polyploidization distinguishes 20th-century from diploid 19th-century samples (*41*). Moreover, changes in allele frequencies of pathogen genes modulating infection that are recognized by the plant immune system – effector genes – were identified following the introduction of resistant potato cultivars in the 20th century (*41*).

Bacteria can also be accurately identified through analysis of herbarium-derived shotgun reads. A proof-of-principle study employing a library preparation method that specifically enriches aDNA molecules successfully assembled *de novo* genomes for commensal and pathogenic bacteria belonging to the genus *Pseudomonas* from a 19th-century potato herbarium specimen (*44*). Herbarium-derived genomes of *Xanthomonas citri* pv. *citri* (*Xci*), the causal agent of Asiatic citrus canker, have been used to reconstruct the population history of this bacterial pathogen at both local (*45*) and global scales (*46*), revealing that the spread of Asian citrus canker disease are linked to climatic changes and the onset and global expansion of citriculture (*46*).

While agricultural disease outbreaks are often characterized by expansion of individual pathogenic lineages, wild plants are infected by genetically diverse groups of pathogens (*47*). Annotation of bacteriophages in present-day and herbarium-derived *Pseudomonas* genomes identified a putative bacteriophage-derived bacteriocin (tailocin) used by pathogenic bacteria for intralineage competition, which likely prevents expansion of single bacterial lineages (*48*). This study showcases the resolution attainable through analysis of herbarium-derived bacterial genomes: determining the genomic location of the tailocins within *Pseudomonas* genomes, characterizing the tailocin gene content, assessing conserved colinearity of tailocin genes, and elucidating the haplotypes of highly diverse proteins likely crucial for tailocin bactericidal activity (*48*). Outside of pathogenic microbes, a targeted genomics approach was

used to sequence and assemble symbiotic bacteria of wild yam (*Dioscorea sansibarensis*), revealing the horizontal transmissions of symbionts (*49*).

Studies of fungi, oomycetes, and bacteria have focused on analysis of pathogen-infected herbarium tissue or highly abundant microbial taxa expected to be found on plant tissues according to reference microbiomes retrieved from fresh plant samples. Pure metagenomic assignment of herbarium-derived sequencing reads is complicated by short read length and nucleotide misincorporation, which decrease the reliability of read mapping. Once reads are reliably assigned to taxa, subsequent challenges lie in distinguishing between the authentic microbiome and post-mortem colonizers. Investigating whether patterns of DNA misincorporation in reads mapped to specific microbial genomes resemble those commonly observed in aDNA (*50*), or employing library preparation methods that selectively incorporate aDNA reads (*44*) are insufficient to distinguish the authentic microbiome. It is crucial to assess the significance of identified taxa by comparing them to reference microbiomes from fresh plant tissues and herbarium contaminants. A promising alternative involves characterizing microbial taxa preferentially found in herbarium specimens across different plant taxa and multiple herbaria locations (*11*). This approach makes it possible to differentiate the authentic metagenome from potential contaminants within the herbarium metagenome. Shotgun metagenomics was employed on present-day and herbarium genomes of *A. artemisiifolia* to characterize the pathogen load of plants inhabiting invasive and native ranges. A lower prevalence of disease-inducing plant pathogens in the invasive range suggests that escape from naturally occurring pathogens contributed to the invasion of this species in Europe (*28*).

Herbarium specimens retain plant roots, a habitat rich in microorganisms, and in certain instances contain remnants of the original soil (Fig. 2A). Characterizing root microbiomes could potentially unveil the dynamics of microbial communities influenced by land management or agricultural practices, such as the application of agrochemicals.

Beyond herbarium genomics: integrating genomics with multiple data

Combining herbarium genomics with other molecular data

Herbaria have also proven valuable to scientists from other plant biology disciplines and for research on global environmental change (*22*, *23*). These alternative applications encompass the use of multiproxy molecular data. The biggest challenge in using molecules other than DNA from historic materials lies in the limited understanding of their degradation, and authentication procedures. While simple elements and their stable isotopes are not expected to degrade within herbarium timescales, care must be taken to avoid contamination. Ecologists as well as soil and plant scientists assess the state of ecosystems using nitrogen concentrations and isotope ratios, which reflect the cycling of key chemical elements. Herbarium specimens of multiple species allow monitoring of changes in nitrogen availability on a regional scale (Fig. 2B). Despite increased anthropogenic nitrogen deposition, foliar nitrogen concentrations have decreased (*51*). Combining analysis of nitrogen and carbon isotopes with phenology traits reveals heterogeneous physiological responses of *A. thaliana* that depend on growth habit, population structure, and changes in the environment (*52*). Combining such studies with genomic characterization might reveal the extent to which responses to anthropogenic pressures are genetic or plastic. This possibility is particularly

relevant for crops, whose genetic makeup changed contemporaneously with the onset of fertilization regimes during the Green Revolution.

Complex chemical molecules produced by plants can modulate biotic interactions such as plant defense, plant-plant competition and pollinator attraction. Thus, characterizing such molecules through time can reveal the evolution of multiple biotic interactions. However, isolating and quantifying complex chemical molecules is more challenging than isolating DNA and requires fine-tuning of multiple specialized approaches. It is important to understand the kinetics of each quantified molecule, i.e., its half-life and degradation products. A study comparing terpenoids, important metabolites for growth and defense, in different sages (*Salvia* sp.) noted a significant effect of species but not of herbarium collection date on the terpenoid composition, although concentrations decrease through time. Therefore, herbarium specimens are suitable for qualitative comparison of terpenoids, but care needs to be taken for their quantitative analysis. (*53*). Specimens of some plant species were annotated with their historical chemical composition at the time of collection, allowing direct comparison with modern measurements. For example, quinine alkaloid concentrations measured 150 years ago in fever tree (*Cinchona* sp.) barks are broadly congruent with recent measurements, suggesting that this group of compounds is stable (*54*). Degradation dynamics have been assessed for only a handful of biomolecules and should be evaluated individually for economically important plant products (medicinal, plant defense). Direct measurement coupled with genome sequencing should facilitate discovery of underlying genetic pathways (Fig. 2B).

Proteomes have been successfully isolated from ancient animal samples. Proteins degrade more slowly than DNA, and strict protocols for authenticating their historical or ancient nature have already been developed (*55*). However, these approaches have not yet been applied to herbarium specimens. Although RNA is much less stable than DNA and authentication procedures are not currently available, its sequencing can be authenticated indirectly via damage pattern signatures in complementary DNA, providing, for example, insights into plant–microbe interactions at the time of sampling.

Leveraging phenotypic and genomic data from herbarium specimens

Examining changes in the timing of phenological milestones such as leaf-out or flowering allows studying the effects of climate change. Herbarium specimens are valuable and reliable records for phenological research (Fig. 2B). Automation of phenological scoring through machine learning approaches will expand the scope of herbarium-based phenological studies, enabling assessment of more phenotypes (*56*). The collection date of flowering herbarium specimens provides a reliable estimation of past flowering time (*57*). For example, long-term phenology trends from herbarium specimens indicate that European forest wildflowers shifted their phenology in response to climate change (*58*). However, any genetic basis is yet to be determined. Integrating herbarium genomics with phenological research will reveal whether variation in phenotypic traits has a genetic basis or is solely attributable to phenotypic plasticity.

Another important phenotype in the context of global environmental change is stomatal density, referring to the pores plants use for gas exchange (Fig. 2B). Stomatal density is expected to decrease as atmospheric CO_2 levels rise with industrialization, as it has been revealed by herbarium specimens (*59*). A polygenic score-like metric that was devised by

integrating herbarium and present-day genomes with known effects of gene knock-out mutations on stomatal development showed a preliminary positive correlation with stomatal density in *A. thaliana* (*60*). This metric was validated using present-day genomes and, when applied to herbarium genomes, reproduces previously observed patterns of stomatal density decrease over time. This correlation was achieved without directly measuring stomatal density in herbarium specimens, indicating a potential genetic basis underlying the observed patterns (*60*).

Characterization of interactions between plants and insect herbivores is enabled by quantifying lesions on herbarium specimens (Fig. 1). Analysis of herbarium specimens spanning more than a century indicates that damage to plants by insect herbivores will likely continue to increase in the northeastern US as global temperatures rise (*61*). Herbarium genomics might enhance these studies by identifying herbivores through insect DNA left on plant tissue as well as plant genetic loci responsible for susceptibility to insect herbivory.

Conclusions

Herbarium genomic approaches reveal the profound influence of human activities such as agriculture on plant evolution. However, there is still much to explore, particularly regarding anthropogenic effects on plant interactions with the environment. Biotic interactions can be investigated through metagenomics of herbarium specimens, while isotope ratios, stomata numbers, and phenology provide insights into abiotic interactions. Although the influence of climate change on population diversity and structure is routinely established, its evolutionary consequences and adaptive responses are rarely investigated (*62*). Integration of present-day and herbarium genomics continues to be essential for studying rapid adaptation in natural populations.

With time, new ways to utilize herbarium collections emerge, and their potential grows with new analytical techniques. For example, recently, herbarium specimens have gained relevance for plant “de-extinction,” resurrecting plants considered extinct in the wild using preserved seeds or tissues from herbarium collections (*63*). Herbarium genomics holds potential for informing and validating such resurrection efforts.

Finally, herbarium collections have a temporal overlap with colonial history, reflecting the acquisition and preservation of botanical treasures from continents explored by Europeans since the 15th century. Most specimens were obtained without proper permits or consent from indigenous communities. Today, legal and ethical obligations regarding herbarium materials are guided by the Convention on Biological Diversity (CBD) and the benefit-sharing principles of the Nagoya Protocol. Although the CBD applies primarily to recent collections, ethical considerations extend to older materials, particularly plants cultivated on indigenous lands, including crops, medicinal plants, and sacred species.

References

1. V. A. Funk, 100 Uses for an Herbarium: well at least 72. *American Society of Plant Taxonomists Newsletter* (2003) (available at https://repository.si.edu/bitstream/handle/10088/11385/bot_2003_pr_Funk_100Uses.pdf).

2. B. M. Thiers, updated continuously. Index Herbariorum (2022).

3. O. M. Grace, O. A. Pérez-Escobar, E. J. Lucas, M. S. Vorontsova, G. P. Lewis, B. E. Walker, L. G. Lohmann, S. Knapp, P. Wilkie, T. Sarkinen, I. Darbyshire, E. N. Lughadha, A. Monro, Y. Woudstra, S. Demissew, A. M. Muasya, S. Díaz, W. J. Baker, A. Antonelli, Botanical Monography in the Anthropocene. *Trends Plant Sci.* **26**, 433–441 (2021).

4. J. M. Heberling, Herbaria as Big Data Sources of Plant Traits. *Int. J. Plant Sci.* **183**, 87–118 (2022).

5. S. Ramírez-Barahona, H. Sauquet, S. Magallón, The delayed and geographically heterogeneous diversification of flowering plant families. *Nat. Ecol. Evol.* **4**, 1232–1238 (2020).

6. L. Kistler, V. C. Bieker, M. D. Martin, M. W. Pedersen, J. Ramos Madrigal, N. Wales, Ancient Plant Genomics in Archaeology, Herbaria, and the Environment. *Annu. Rev. Plant Biol.* **71**, 605–629 (2020).

7. C. L. Weiß, V. J. Schuenemann, J. Devos, G. Shirsekar, E. Reiter, B. A. Gould, J. R. Stinchcombe, J. Krause, H. A. Burbano, Temporal patterns of damage and decay kinetics of DNA retrieved from plant herbarium specimens. *Open Science*. **3**, 160239 (2016).

8. B. Shapiro, M. Hofreiter, A paleogenomic perspective on evolution and gene function: new insights from ancient DNA. *Science*. **343**, 1236573 (2014).

9. M. Staats, A. Cuenca, J. E. Richardson, R. Vrielink-van Ginkel, G. Petersen, O. Seberg, F. T. Bakker, DNA Damage in Plant Herbarium Tissue. *PLoS One*. **6** (2011), doi:10.1371/journal.pone.0028448.

10. H. Jónsson, A. Ginolhac, M. Schubert, P. L. F. Johnson, L. Orlando, mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*. **29**, 1682–1684 (2013).

11. V. C. Bieker, F. Sánchez Barreiro, J. A. Rasmussen, M. Brunier, N. Wales, M. D. Martin, Metagenomic analysis of historical herbarium specimens reveals a postmortem microbial community. *Mol. Ecol. Resour.* **20**, 1206–1219 (2020).

12. S. Dodsworth, L. Pokorny, M. G. Johnson, J. T. Kim, O. Maurin, N. J. Wickett, F. Forest, W. J. Baker, Hyb-Seq for Flowering Plant Systematics. *Trends Plant Sci.* **24**, 887–891 (2019).

13. S. M. Latorre, P. L. M. Lang, H. A. Burbano, R. M. Gutaker, Isolation, Library Preparation, and Bioinformatic Analysis of Historical and Ancient Plant DNA. *Curr Protoc Plant Biol*. **5**, e20121 (2020).

14. V. C. Bieker, M. D. Martin, Implications and future prospects for evolutionary analyses of DNA in historical herbarium collections. *Botany Letters*. **165**, 409–418 (2018).

15. P. G. Nevill, X. Zhong, J. Tonti-Filippini, M. Byrne, M. Hislop, K. Thiele, S. van Leeuwen, L. M. Boykin, I. Small, Large scale genome skimming from herbarium material for accurate plant identification and phylogenomics. *Plant Methods*. **16**, 1 (2020).

16. J. R. Hughey, P. W. Gabrielson, L. Rohmer, J. Tortolani, M. Silva, K. A. Miller, J. D.

Young, C. Martell, E. Ruediger, Minimally destructive sampling of type specimens of Pyropia (Bangiales, Rhodophyta) recovers complete plastid and mitochondrial genomes. *Sci. Rep.* **4**, 5113 (2014).

17. P. Muñoz-Rodríguez, T. Carruthers, J. R. I. Wood, B. R. M. Williams, K. Weitemier, B. Kronmiller, Z. Goodwin, A. Sumadijaya, N. L. Anglin, D. Filer, D. Harris, M. D. Rausher, S. Kelly, A. Liston, R. W. Scotland, A taxonomic monograph of Ipomoea integrated across phylogenetic scales. *Nat Plants*. **5**, 1136–1144 (2019).

18. W. J. Baker, P. Bailey, V. Barber, A. Barker, S. Bellot, D. Bishop, L. R. Botigué, G. Brewer, T. Carruthers, J. J. Clarkson, J. Cook, R. S. Cowan, S. Dodsworth, N. Epitawalage, E. Françoso, B. Gallego, M. G. Johnson, J. T. Kim, K. Leempoel, O. Maurin, C. Mcginnie, L. Pokorny, S. Roy, M. Stone, E. Toledo, N. J. Wickett, A. R. Zuntini, W. L. Eiserhardt, P. J. Kersey, I. J. Leitch, F. Forest, A Comprehensive Phylogenomic Platform for Exploring the Angiosperm Tree of Life. *Syst. Biol.* **71**, 301–319 (2022).

19. E. M. Gardner, M. Garner, R. Cowan, S. Dodsworth, N. Epitawalage, O. Maurin, D. Arifiani, S. Sahromi, W. J. Baker, F. Forest, N. J. C. Zerega, A. K. Monro, A. L. Hipp, Repeated parallel losses of inflexed stamens in Moraceae: phylogenomics and generic revision of the tribe Moreae and the reinstatement of the tribe Olmedieae (Moraceae). *Taxon*. **70**, 946– 988 (2021).

20. O. A. Pérez-Escobar, J. E. Richardson, M.-J. R. Howes, E. Lucas, N. Álvarez de Róman, J. Collemare, I. A. Graham, J. Gratzfeld, P. J. Kersey, I. J. Leitch, A. Paton, P. M. Hollingsworth, A. Antonelli, Untapped resources for medical research. *Science*. **369**, 781–782 (2020).

21. M. K. N. Lawniczak, R. P. Davey, J. Rajan, L. L. Pereira-da-Conceicoa, E. Kilias, P. M. Hollingsworth, I. Barnes, H. Allen, M. Blaxter, J. Burgin, Others, Specimen and sample metadata standards for biodiversity genomics: a proposal from the darwin tree of life project. *Wellcome Open Research*. **7**, 187 (2022).

22. E. K. Meineke, C. C. Davis, T. J. Davies, The unrealized potential of herbaria for global change biology. *Ecol. Monogr.* **88**, 505–525 (2018).

23. P. L. M. Lang, F. M. Willems, J. F. Scheepens, H. A. Burbano, O. Bossdorf, Using herbaria to study global environmental change. *New phytologist*, ***221***(1), 110-122..

24. M. van Kleunen, W. Dawson, F. Essl, J. Pergl, M. Winter, E. Weber, H. Kreft, P. Weigelt, J. Kartesz, M. Nishino, L. A. Antonova, J. F. Barcelona, F. J. Cabezas, D. Cárdenas, J. Cárdenas-Toro, N. Castaño, E. Chacón, C. Chatelain, A. L. Ebel, E. Figueiredo, N. Fuentes, Q. J. Groom, L. Henderson, Inderjit, A. Kupriyanov, S. Masciadri, J. Meerman, O. Morozova, D. Moser, D. L. Nickrent, A. Patzelt, P. B. Pelser, M. P. Baptiste, M. Poopath, M. Schulze, H. Seebens, W.-S. Shu, J. Thomas, M. Velayos, J. J. Wieringa, P. Pyšek, Global exchange and accumulation of non-native plants. *Nature*. **525**, 100–103 (2015).

25. R. A. Lankau, V. Nuzzo, G. Spyreas, A. S. Davis, Evolutionary limits ameliorate the negative impact of an invasive plant. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 15362–15367 (2009).

26. M. Exposito-Alonso, C. Becker, V. J. Schuenemann, E. Reiter, C. Setzer, R. Slovak, B. Brachi, J. Hagmann, D. G. Grimm, J. Chen, W. Busch, J. Bergelson, R. W. Ness, J.

Krause, H. A. Burbano, D. Weigel, The rate and potential relevance of new mutations in a colonizing plant lineage. *PLoS Genet.* **14**, e1007155 (2018).

27. M. D. Martin, E. A. Zimmer, M. T. Olsen, A. D. Foote, M. T. P. Gilbert, G. S. Brush, Herbarium specimens reveal a historical shift in phylogeographic structure of common ragweed during native range disturbance. *Mol. Ecol.* **23**, 1701–1716 (2014).

28. V. C. Bieker, P. Battlay, B. Petersen, X. Sun, J. Wilson, J. C. Brealey, F. Bretagnolle, K. Nurkowski, C. Lee, F. S. Barreiro, G. L. Owens, J. Y. Lee, F. L. Kellner, L. van Boheeman, S. Gopalakrishnan, M. Gaudeul, H. Mueller-Schaerer, S. Lommen, G. Karrer, B. Chauvel, Y. Sun, B. Kostantinovic, L. Dalén, P. Poczai, L. H. Rieseberg, M. T. P. Gilbert, K. A. Hodgins, M. D. Martin, Uncovering the genomic basis of an extraordinary plant invasion. *Sci Adv*. **8**, eabo5115 (2022).

29. J. M. Kreiner, J. R. Stinchcombe, S. I. Wright, Population Genomics of Herbicide Resistance: Adaptation via Evolutionary Rescue. *Annu. Rev. Plant Biol.* **69**, 611–635 (2018).

30. C. Délye, C. Deulvot, B. Chauvel, DNA analysis of herbarium Specimens of the grass weed Alopecurus myosuroides reveals herbicide resistance pre-dated herbicides. *PLoS One*. **8**, e75117 (2013).

31. S. Kersten, J. Chang, C. D. Huber, Y. Voichek, C. Lanz, T. Hagmaier, P. Lang, U. Lutz, I. Hirschberg, J. Lerchl, A. Porri, Y. Van de Peer, K. Schmid, D. Weigel, F. A. Rabanal, Standing genetic variation fuels rapid evolution of herbicide resistance in blackgrass. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2206808120 (2023).

32. J. M. Kreiner, S. M. Latorre, H. A. Burbano, J. R. Stinchcombe, S. P. Otto, D. Weigel, S. I. Wright, Rapid weed adaptation and range expansion in response to agriculture over the past two centuries. *Science*. **378**, 1079–1085 (2022).

33. R. Vos, T. van Andel, B. Gravendeel, V. Kakakiou, E. Michels, E. Schijlen, A. Stefanaki, A Tomato Genome From The Italian Renaissance Provides Insights Into Columbian and Pre-Columbian Exchange Links And Domestication. *bioRxiv* (2022), p. 2022.06.07.495125.

34. A. W. Crosby, *The Columbian Exchange: Biological and Cultural Consequences of 1492* (Greenwood Publishing Group, 2003).

35. M. Ames, D. M. Spooner, DNA from herbarium specimens settles a controversy about origins of the European potato. *Am. J. Bot.* **95**, 252–257 (2008).

36. R. M. Gutaker, C. L. Weiß, D. Ellis, N. L. Anglin, S. Knapp, J. Luis Fernández-Alonso, S. Prat, H. A. Burbano, The origins and adaptation of European potatoes reconstructed from historical genomes. *Nat Ecol Evol*. **3**, 1093–1101 (2019).

37. M. van de Wouw, T. van Hintum, C. Kik, R. van Treuren, B. Visser, Genetic diversity trends in twentieth century crop cultivars: a meta analysis. *Theor. Appl. Genet.* **120**, 1241–1252 (2010).

38. A. Alsantely, R. Gutaker, M. E. Navarrete Rodríguez, G. Arrieta-Espinoza, E. J. Fuchs, A. Costa de Oliveira, J. Tohme, A. Zuccolo, R. A. Wing, A. Fornasiero, The International Oryza Map Alignment Project (IOMAP): the Americas-past achievements and future directions. *J. Exp. Bot.* **74**, 1331–1342 (2023).

39. H. Wang, F. G. Vieira, J. E. Crawford, C. Chu, R. Nielsen, Asian wild rice is a hybrid swarm with extensive gene flow and feralization from domesticated rice. *Genome Res.* **27**, 1029–1038 (2017).

40. J. B. Ristaino, C. T. Groves, G. R. Parra, PCR amplification of the Irish potato famine pathogen from historic specimens. *Nature*. **411**, 695–697 (2001).

41. K. Yoshida, V. J. Schuenemann, L. M. Cano, M. Pais, B. Mishra, R. Sharma, C. Lanz, F. N. Martin, S. Kamoun, J. Krause, M. Thines, D. Weigel, H. A. Burbano, The rise and fall of the Phytophthora infestans lineage that triggered the Irish potato famine (vol 2, e00731, 2013). *Elife*. **2** (2013), doi:10.7554/eLife.01108.

42. M. D. Martin, E. Cappellini, J. A. Samaniego, M. L. Zepeda, P. F. Campos, A. Seguin-Orlando, N. Wales, L. Orlando, S. Y. W. Ho, F. S. Dietrich, P. A. Mieczkowski, J. Heitman, E. Willerslev, A. Krogh, J. B. Ristaino, M. T. P. Gilbert, Reconstructing genome evolution in historic samples of the Irish potato famine pathogen. *Nat. Commun.* **4** (2013), doi:10.1038/ncomms3172.

43. M. D. Martin, S. Y. W. Ho, N. Wales, J. B. Ristaino, M. T. P. Gilbert, Persistence of the Mitochondrial Lineage Responsible for the Irish Potato Famine in Extant New World Phytophthora infestans. *Mol. Biol. Evol.* **31**, 1414–1420 (2014).

44. C. L. Weiß, M.-T. Gansauge, A. Aximu-Petri, M. Meyer, H. A. Burbano, Mining ancient microbiomes using selective enrichment of damaged DNA molecules. *BMC Genomics*. **21**, 432 (2020).

45. P. E. Campos, C. Groot Crego, K. Boyer, M. Gaudeul, C. Baider, D. Richard, O. Pruvost, P. Roumagnac, B. Szurek, N. Becker, L. Gagnevin, A. Rieux, First historical genome of a crop bacterial pathogen from herbarium specimen: Insights into citrus canker emergence. *PLoS Pathog.* **17**, e1009714 (2021).

46. P. E. Campos, O. Pruvost, K. Boyer, F. Chiroleu, T. T. Cao, M. Gaudeul, C. Baider, T. M. A. Utteridge, N. Becker, A. Rieux, L. Gagnevin, Herbarium specimen sequencing allows precise dating of Xanthomonas citri pv. citri diversification history. *Nat. Commun.* **14**, 4306 (2023).

47. T. L. Karasov, J. Almario, C. Friedemann, W. Ding, M. Giolai, D. Heavens, S. Kersten, D. S. Lundberg, M. Neumann, J. Regalado, R. A. Neher, E. Kemen, D. Weigel, Arabidopsis thaliana and Pseudomonas Pathogens Exhibit Stable Associations over Evolutionary Timescales. *Cell Host Microbe*. **24**, 168–179.e4 (2018).

48. T. Backman, S. M. Latorre, L. Eads, S. Som, D. Belnap, A. M. Manuel, H. A. Burbano, T. L. Karasov, A weaponized phage suppresses competitors in historical and modern metapopulations of pathogenic bacteria. *bioRxiv* (2023), p. 2023.04.17.536465.

49. B. Danneels, J. Viruel, K. Mcgrath, S. B. Janssens, N. Wales, P. Wilkin, A. Carlier, Patterns of transmission and horizontal gene transfer in the Dioscorea sansibarensis leaf symbiosis revealed by whole-genome sequencing. *Curr. Biol.* **31**, 2666–2673.e4 (2021).

50. F. M. Key, C. Posth, J. Krause, A. Herbig, K. I. Bos, Mining Metagenomic Data Sets for Ancient DNA: Recommended Protocols for Authentication. *Trends Genet.* **33**, 508–520 (2017).

51. K. K. McLauchlan, C. J. Ferguson, I. E. Wilson, T. W. Ocheltree, J. M. Craine, Thirteen

decades of foliar isotopes indicate declining nitrogen availability in central North American grasslands. *New Phytol.* **187**, 1135–1145 (2010).

52. V. L. DeLeo, D. N. L. Menge, E. M. Hanks, T. E. Juenger, J. R. Lasky, Effects of two centuries of global environmental variation on phenology and physiology of Arabidopsis thaliana. *Glob. Chang. Biol.* **26**, 523–538 (2020).

53. I. J. Foutami, T. Mariager, R. Rinnan, C. J. Barnes, N. Rønsted, Hundred Fifty Years of Herbarium Collections Provide a Reliable Resource of Volatile Terpenoid Profiles Showing Strong Species Effect in Four Medicinal Species of Salvia Across the Mediterranean. *Front. Plant Sci.* **9** (2018), doi:10.3389/fpls.2018.01877.

54. N. A. Canales, T. N. Gress Hansen, C. Cornett, K. Walker, F. Driver, A. Antonelli, C. Maldonado, M. Nesbitt, C. J. Barnes, N. Rønsted, Historical chemical annotations of Cinchona bark collections are comparable to results from current day high-pressure liquid chromatography technologies. *J. Ethnopharmacol.* **249**, 112375 (2020).

55. J. Hendy, F. Welker, B. Demarchi, C. Speller, C. Warinner, M. J. Collins, A guide to ancient protein studies. *Nat Ecol Evol*. **2**, 791–799 (2018).

56. K. D. Pearson, G. Nelson, M. F. J. Aronson, P. Bonnet, L. Brenskelle, C. C. Davis, E. G. Denny, E. R. Ellwood, H. Goëau, J. M. Heberling, A. Joly, T. Lorieul, S. J. Mazer, E. K. Meineke, B. J. Stucky, P. Sweeney, A. E. White, P. S. Soltis, Machine Learning Using Digitized Herbarium Specimens to Advance Phenological Research. *Bioscience*. **70**, 610–620 (2020).

57. C. C. Davis, C. G. Willis, B. Connolly, C. Kelly, A. M. Ellison, 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).

58. F. M. Willems, J. F. Scheepens, O. Bossdorf, Forest wildflowers bloom earlier as Europe warms: lessons from herbaria and spatial modelling. *New Phytol.* **235**, 52–65 (2022).

59. F. I. Woodward, Stomatal numbers are sensitive to increases in CO2 from pre-industrial levels. *Nature*. **327**, 617–618 (1987).

60. P. L. M. Lang, J. M. Erberich, L. Lopez, C. L. Weiß, G. Amador, H. F. Fung, S. M. Latorre, J. R. Lasky, H. A. Burbano, M. Expósito-Alonso, D. Bergmann, Century-long timelines of herbarium genomes predict plant stomatal response to climate change. *bioRxiv* (2022), p. 2022.10.23.513440.

61. E. K. Meineke, A. T. Classen, N. J. Sanders, T. Jonathan Davies, Herbarium specimens reveal increasing herbivory over the past century. *J. Ecol.* **107**, 105–117 (2019).

62. R. D. Clark, K. A. Catalano, K. S. Fitz, E. Garcia, K. E. Jaynes, B. N. Reid, A. Sawkins, A. A. Snead, J. C. Whalen, M. L. Pinsky, The practice and promise of temporal genomics for measuring evolutionary responses to global change. *Mol. Ecol. Resour.* (2023), doi:10.1111/1755-0998.13789.

63. G. Albani Rocchetti, A. Carta, A. Mondoni, S. Godefroid, C. C. Davis, G. Caneva, M. A. Albrecht, K. Alvarado, R. Bijmoer, R. Borosova, C. Bräuchler, E. Breman, M. Briggs, S. Buord, L. H. Cave, N. G. Da Silva, A. H. Davey, R. M. Davies, J. B. Dickie, M. Fabillo, A. Fleischmann, A. Franks, G. Hall, G. Kantvilas, C. Klak, U. Liu, L. Medina, L. G. Reinhammar, R. J. Sebola, I. Schönberger, P. Sweeney, H. Voglmayr, A. White, J. J.

Wieringa, E. Zippel, T. Abeli, Selecting the best candidates for resurrecting extinct-in-the-wild plants from herbaria. *Nat Plants*. **8**, 1385–1393 (2022).

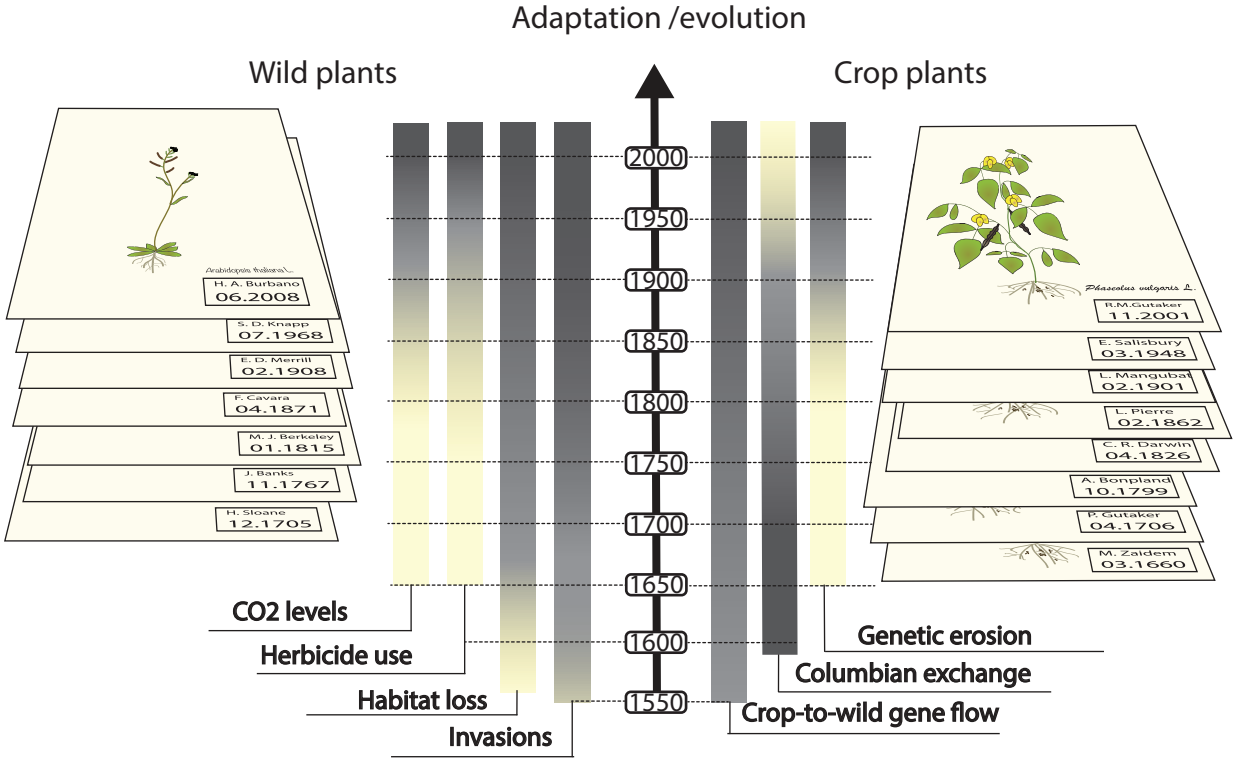
Acknowledgments

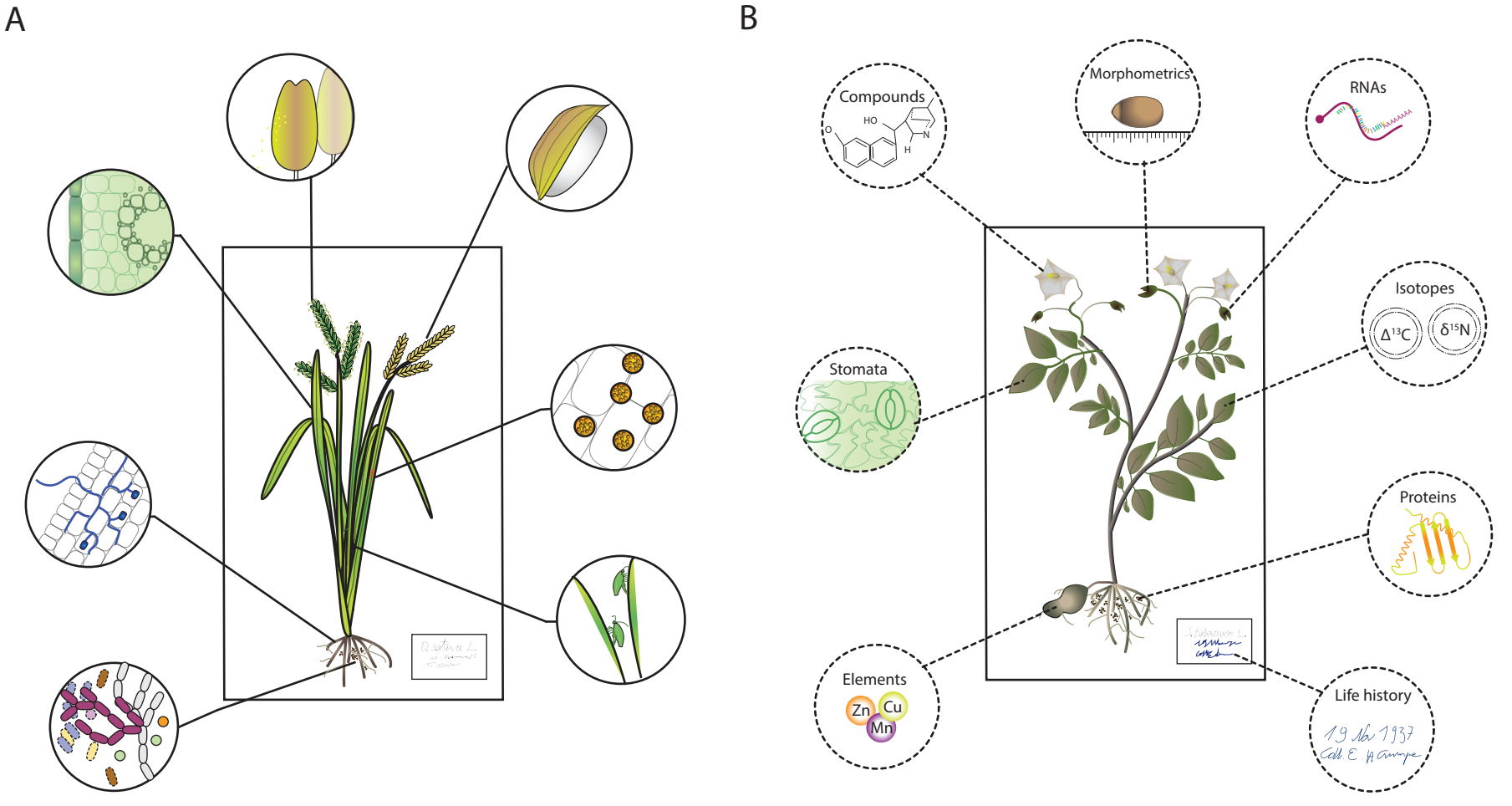
We thank P. Kersey and S. Latorre for their critical appraisal of the manuscript. **Funding:** H.A.B was supported by the Royal Society (RSWF\R1\191011) and the Leverhulme Trust (Philip Leverhulme Prize). R.M.G. was supported by UKRI Frontier Research Guarantee (EP/X022404/1). **Competing interests:** None declared.

Figure captions

Fig. 1: The sampling range of herbarium specimens covers the time of intensified human influence on the environment. Sequencing genomes of historical plants, as well as associated species, can address questions about their evolution and adaptation in response to anthropogenic pressures. Vertical bars indicate the intensity of human-mediated effects in the last five centuries (darker shades represent stronger effect) on wild (left) and crop (right) plants. Stacks of herbarium specimens illustrate temporal sampling.

Fig. 2: Environmental, molecular, and phenotypic wealth of herbarium specimens. A Herbarium sheets preserve not only the genomes of individual plants but also those of many associated species: microbes, such as bacteria and fungi, as well as macro-organisms, such as insects. **B** Herbarium specimens also harbor a plethora of other types of data. Phenology and, occasionally, environmental or ethnobotanical information is recorded on the label. Anatomical features allow direct morphometric measurements. Other molecular data are available, from simple micronutrients and atomic isotopes to complex molecules such as alkaloids, RNAs, and proteins.


Adaptation /evolution
Wild plants
Crop plants
H. A. Burbano
06.2008
S. D. Knapp
07.1968
E. D. Merrill
02.1908
F. Cavara
04.1871
M. J. Berkeley
01.1815
J. Banks
11.1767
H. Sloane
12.1705
2000
1950
1900
1850
1800
1750
1700
1650
1600
1550
R.M.Gutaker
11.2001
E. Salisbury
03.1948
L. Mangubat
02.1901
L. Pierre
02.1862
C. R. Darwin
04.1826
A. Bonpland
10.1799
P. Gutaker
04.1706
M. Zaidem
03.1660
CO2 levels
Herbicide use
Habitat loss
Invasions
Genetic erosion
Columbian exchange
Crop-to-wild gene flow


A
B
Compounds
Morphometrics
RNAs
Isotopes
Stomata
Proteins
Elements
Life history
Zn
Cu
Mn