NextGeneration specimen digitization:
The international herbarium community goes spectral!

Jeannine Cavender-Bares[1*]

Dawson M. White[1]

Natalie Iwanycki Ahlstrand[2]

Matthew W. Austin[3]

Denis Bastianelli[4]

Samantha Bazan[4]

Khalil Boughalmi[5]

Warren Cardinal-McTeague[6]

Eduardo Chacón-Madrigal[7]

Thomas L.P. Couvreur[5]

Charles Davis[1]

Flávia M Durgante[8,9]

Olwen M. Grace[10]

J. Antonio Guzmán Q.[1]

Kimberly Hansen[11]

M.S. Hernández-Leal[1]

Mike John Gilbert Hopkins[9]

Rykkar Jackson[6]

Shan Kothari[12]

Aaron K. Lee[13]

Étienne Léveillé-Bourret[14]

Jesús Pinto-Ledezma[15]

Natalia L. Quinteros Casaverde[16]

Jose Eduardo Meireles[17]

Cornelius Onyedikachi Nichodemus[17]

Michaela Schmull[1]

Douglas E. Soltis[18]

Pamela S. Soltis[18]

Hanna Tuomisto[19,20]

Susan Ustin[21]

Caroline C. Vasconcelos[9]

*jcavender@fas.harvard.edu
[1]Harvard University Herbaria and Department of Organismic and Evolutionary Biology, Harvard University, Cambridge MA 02138, USA
[2] Natural History Museum of Denmark, University of Copenhagen, Copenhagen, 1154, Denmark
[3] Herbarium, Missouri Botanical Garden, 4344 Shaw Blvd, St. Louis MO 63110
[4] CIRAD, Umr Selmet, Montpellier, France, UMR SELMET, F-34398 Montpellier, France.
[5] DIADE, Univ Montpellier, CIRAD, IRD, Montpellier, France
[6] Department of Forest and Conservation Sciences, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada
[7] Herbario Nacional, Departamento de Historia Natural, Museo Nacional de Costa Rica & Herbario Luis Fournier Origgi, Centro de Investigación en Biodiversidad y Ecología Tropical (CIBET), Universidad de Costa Rica, Costa Rica
[8] Department of Wetlands Ecology, Karlsruhe Institute of Technology, KIT, Germany.
[9] Herbário, National Institute of Amazonian Research, INPA, Brazil.
[10] Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, United Kingdom
[11] Field Museum of Natural History, 1400 S. DuSable Lake Shore Dr., Chicago, IL 60625, USA
[12] Department of Renewable Resources, University of Alberta, Edmonton, AB T6G 2E3, Canada
[13] Department of Plant and Microbial Biology, University of Minnesota, 1479 Gortner Ave, Saint Paul, MN 55108, USA
[14] Département de sciences biologiques, Institut de recherche en biologie végétale (IRBV), Université de Montréal, Montréal, QC H1X 2B2, Canada
[15] Department of Ecology, Evolution and Behavior, University of Minnesota, 1479 Gortner Ave, Saint Paul, MN 55108, USA
[16] Biospheric Science Lab, NASA Goddard Space Flight Center, Greenbelt MD 20771, USA
[17] School of Biology and Ecology, University of Maine, Orono, ME 04469, USA
[17] School of Biology and Ecology, University of Maine, 5793 Hitchner Hall, Orono, Maine 04473, USA
[18] Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA
[19] Department of Biology, Section of Ecoinformatics and Biodiversity, Aarhus University, Aarhus, Denmark
[20] Department of Biology, University of Turku, Turku, Finland
[21] Institute of the Environment and Department of Land, Air and Water Resources, University of California Davis, Davis, CA 95616, USA

Abstract

1. Spectral reflectance measured from herbarium specimens represents a vast source of plant phenotypic and functional trait data.
2. The potential to capture data from specimens to enhance knowledge of plant function and taxon identification has inspired many laboratories worldwide to initiate next-generation spectral digitization from specimens.
3. Combining these datasets into a coordinated global database would enable prediction of traits from the world's plants and allow novel, impactful scientific questions to be addressed at global scale. These novel data streams will generate new capacity to model plant traits globally, enabling connection with remote sensing and ecological and biosphere models and to reconstruct their evolutionary history.
4. Coordination is needed to avoid downstream problems in data aggregation due to variation in data standards and technical specifications of the instruments, optical setups, or measurement protocols. The International Herbarium Spectral Digitization (IHerbSpec) working group has initiated a globally collaborative program, outlining the central issues to address in establishing protocols, standards, and best practices, and next steps. This collaborative effort will allow generation of replicable spectral reflectance data from plant specimens housed in herbaria around the world within ongoing digitization programs following community-defined standards and Findable, Accessible, Interoperable and Reusable (FAIR) principles.

Keywords: methodological standards, collection management, plant functional traits, spectral reflectance, next-generation digitization, global herbarium

Introduction

In an era of rapid global change and biodiversity loss, safeguarding our knowledge of plant diversity is essential. Herbaria serve as foundational repositories for this knowledge, both through their traditional applications to document species morphology, distribution, use, and phenology (National Academies of Sciences and Medicine 2020, Heberling et al. 2021, Davis 2023, Mandrioli 2023). Advancing technologies have now enabled investigations that apply whole-genome sequencing and metabolomic methods of herbarium specimens that are up to hundreds of years old (Burbano and Gutaker 2023, Davis 2023, Medeiros et al. 2024, Davis and Knapp 2025). These advances are now routinely used in macroecological and biogeographical studies providing a vast botanical record of species distributions in space and time relevant to monitoring biodiversity change in the Anthropocene (Willis et al. 2008, Meineke et al. 2018a, Meineke et al. 2018b). We broaden the extended specimen concept, which reenvisions the role of specimen data as a vast, connected repository of information about individual organisms (Webster 2017, National Academies of Sciences and Medicine 2020), to include spectral reflectance from dried plant specimens. A wealth of information about plant chemistry, function, structure can be inferred from plant reflectance spectra—the pattern of light's reflectance from plant tissues across wavelengths (Elvidge 1990, Gitelson and Merzlyak 1994, Sims and Gamon 2002, Ustin et al. 2004, Asner and Martin 2011, Serbin et al. 2014b, Cavender-Bares et al. 2017, Chlus and Townsend 2022, Wang et al. 2023). Spectra can provide information about taxonomic identity (Durgante et al. 2013), phylogenetic placement (Meireles et al. 2020a), defense chemistry in diverse genera (Fine et al. 2021), and phenotypic variation linked to genetic or phylogeographic variation (Cavender-Bares et al. 2016b, Deacon et al. 2017a, Stasinski et al. 2021, Hernandez-Leal et al. 2025). The contribution of spectral information to obtaining both functional and phylogenetic information in plants centers reflectance spectra as a critical data type in the plant sciences with high potential to integrate information about plant diversity across scales, from leaves to ecosystems and the biosphere (Cavender-Bares et al. 2017, Cavender-Bares et al. 2025, Wang et al. 2022, National Academies of Science and Medicine 2025).

Advances in spectroscopic technology and analytical approaches across scales that have enabled spectral data capture across scales (Jetz et al. 2016b, Jacquemoud and Ustin 2019, Serbin and Townsend 2020, Wang et al. 2023) have led to the development of the growing field of spectral biology. This field enables trait models and methods of taxonomic identification from spectral measurements of fresh or dried, pressed leaves (Durgante et al. 2013, Meireles et al. 2020a, Kothari et al. 2023), including actual herbarium collections (Kühn et al. 2025, Neto-Bradley et al. 2025, White et al. 2025), to address myriad questions in ecology, evolution, taxonomy, phylogeography, historical biogeography, biochemistry, and related realms of inquiry. When combined with even a fraction of the world's 400 million herbarium specimens (Thiers 2024), reflectance spectroscopy provides a new means to extend our inferences of plant phenotypic and functional variation across space and time – and across the entirety of plant taxonomic and phylogenetic diversity.

As with digitization standards for herbarium imaging over the past 25 years (Nelson et al. 2015, Hedrick et al. 2019, Davis et al. 2021), similar consensus on protocols and standards is now needed to incorporate spectral data in the 'global metaherbarium' (Davis 2023). Standardized workflows for spectral data and metadata collection will enable integration across herbarium collections, linking trait estimates to individual specimens and the time, place and environment in which they were collected (Davis et al. 2015, Willis et al. 2017, Meineke et al. 2018a, Meineke et al. 2018b, Pearson et al. 2020). ***This article documents the purpose and vision of the newly established international herbarium spectral digitization (IHerbSpec) working group, our progress to date in advancing spectral digitization within herbaria across all green plant groups as a scalable tool for biodiversity science. We describe the nature of spectra, the prospect of measuring them in the world's herbaria, highlight the benefits of this massive effort, and address key challenges and next steps.***

The expanding use of plant traits in ecology and evolution

Extensive work has been carried out by plant systematists to describe new species based on plant traits, often in conjunction with genetic and genomic data. Plant traits have become fundamental to understanding ecological processes such as interactions with the abiotic environment, interactions with herbivores and pollinators, growth responses to resource availability, community assembly processes, and the contributions to ecosystem and biosphere functions (Lavorel and Garnier 2002, Violle et al. 2007, Cavender-Bares et al. 2016a, Funk et al. 2017, Shipley et al. 2017, Dechant et al. 2024). Traits provide key insights into consistent patterns of resource acquisition and ecological functions across the green plant tree of life, termed the Leaf Economics Spectrum (Wright et al. 2004, Díaz et al. 2016). These patterns influence species distributions across environmental gradients and inform models of community assembly, vegetation dynamics, and biosphere function (Cavender-Bares et al. 2016a). Consequently,

significant effort has been invested in developing global plant trait databases (e.g., TRY (Kattge et al. 2020), BIEN (Enquist et al. 2016), AusTraits (Falster et al. 2021, Wenk et al. 2024)), and scaling plant traits to global maps to model biosphere dynamics (Dechant et al. 2024). Frequently collected functional traits for leaves include nitrogen and other nutrients, leaf mass per area, cellulose, lignin, pigments, stable carbon isotopies, and other bioechemical compounds, such as carbohydrates, proteins, lipids, and secondary metabolites (**Fig. 1**). Plant functional traits are well-studied and measured in temperate and arctic regions in well-resourced, easily accessible regions of North America, Europe, China, and Australia, but are undersampled in tropical regions (Bjorkman et al. 2018, Jetz et al. 2016). Common and widespread species are much better represented in plant functional trait databases than rare species, seasonally ephemeral species and species from understudied biogeographic regions. Given that herbaria provide access to plant specimens collected by generations of botanists, they offer a means to systematically obtain trait data from plants in taxa that are rare and/or occur in ecosystems or regions of the world that are difficult to access, filling in global gaps in plant functional traits of the known plant species and lineages (Heberling 2021). Although sampling biases also exist within herbaria (Daru et al. 2018), their coverage is better than that of trait databases.

What are reflectance spectra?

Reflectance spectra are observations of the interaction of electromagnetic radiation (light) with get reflected from surfaces resolved to narrow wavelength bands representing a few nanometers each (**Fig. 1**). Plant tissues from across the tree of life have various structural, anatomical, organellar and biomolecular compositions that influence their spectral properties. As a result, spectral data can be used to estimate a wide array of plant traits, including structural, chemical, and physiological properties (Ustin et al. 2009, Serbin et al. 2014a, Féret et al. 2017), and provide a holistic view of plant phenotypes (Kothari and Schweiger 2022). Reflectance spectra spanning visible, near-infrared and short-wave infrared light (400 - 2500 nm) from freshly-harvested leaves differ from those obtained from pressed and dried leaves because of changes in water content, pigments and structure during drying (**Fig. 1**). Spectra from fresh leaves can provide accurate estimates of plant pigments, traits related to water content, and mobile, degradable or small molecular weight carbon compounds (Wang et al. 2023). Dry leaves often provide more accurate prediction of macronutrients, micronutrients and carbon or nitrogren stable isotopes (Chlus and Townsend 2022, Kothari et al. 2023) but are unsuitable for predicting water-related traits and some pigments that degrade during drying (Ustin et al. 2009). Spectra also vary with changes in environmental conditions, including light, water availability, carbon dioxide, temperature and soil conditions (Cavender-Bares et al. 2016b, Jacquemoud and Ustin 2019, Kühn et al. 2025, Stefanski et al. 2025).

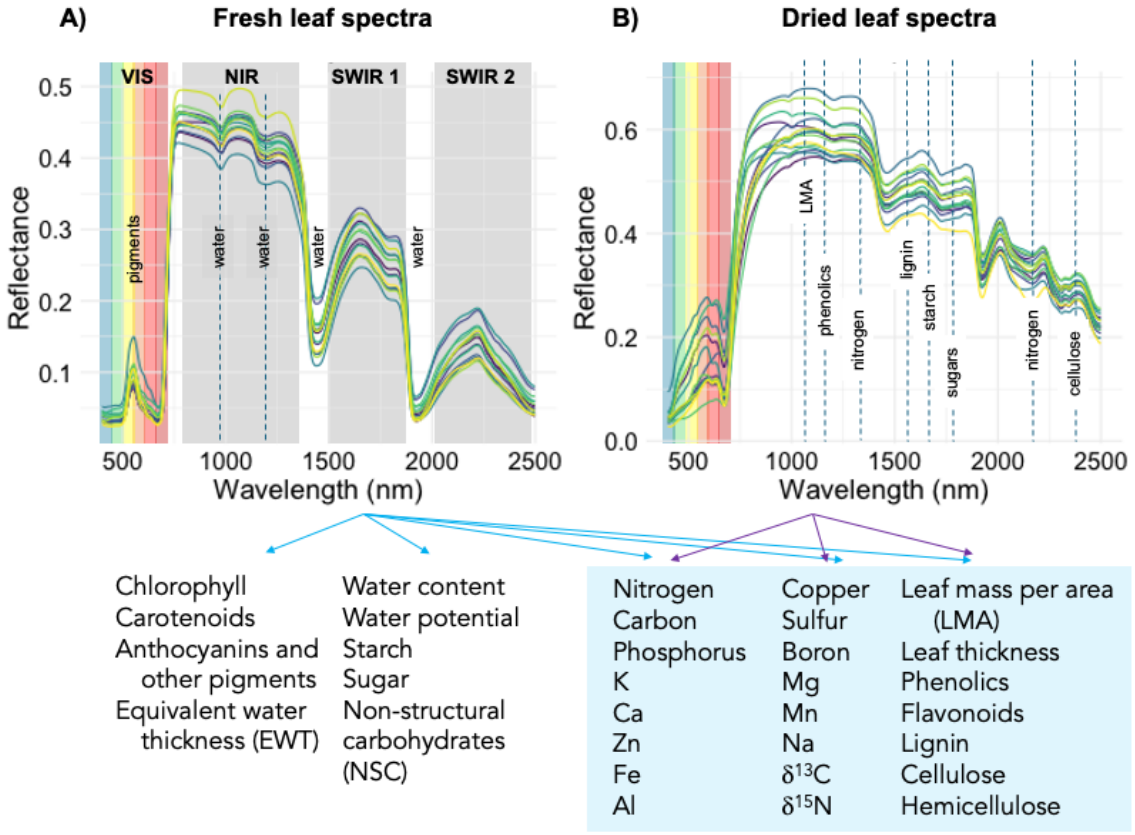
Given their utility in detecting variation in plants, reflectance spectra are widely used in applied contexts. High throughput methods of spectral detection are increasingly used in plant breeding programs to detect phenotype variation (e.g., Ge et al. 2016, Meacham-Hensold et al. 2019). Spectral signatures of plants are also used to detect heavy metals and other environmental contaminants in plant tissues (Zhang et al. 2022), to scale up detection of pathogen infections via remote sensing (e.g., Fallon et al. 2020, Sapes et al. 2024) and for detection of invasive species (e.g., Dao et al. 2021, Mallmann et al. 2023). Currently, major efforts are underway to scale up spectral data from plants for biodiversity monitoring (Jetz et al. 2016a, Cavender-Bares et al. 2020, Wang et al. 2020, Williams et al. 2021, Cavender-Bares et al. 2022, Blanchard et al. 2024, Dechant et al. 2024, Hadlich et al. 2025).

Estimating plant traits from spectra of herbarium specimens

Advancing technologies and a suite of studies over the past decade have set the stage for applying spectral biology to herborized specimens to predict plant function (**Fig. 1**). Some of the first spectral studies of dried leaves were conducted in the early 1990s and contributed to early vegetation index-based or radiative transfer model-based approaches to linking spectra with traits (Elvidge 1990; Fourty et al. 1996). More recently, Costa et al. (2018) showed the potential of using machine learning models trained on spectra on silica-dried leaves to predict traits of tropical tree species. Building on this work, Kothari et al. (2023) showed similar results for a wide range of traits in a large data set of unmounted pressed leaves from temperate forest species; indeed, for most chemical traits, spectra of pressed leaves performed better than the more conventional approach of using spectra of fresh leaves. Sampling the same species as Kothari et al. (2023), White et al. (2025) subsequently showed that spectral models trained using detached leaves from decades-old herbarium specimens accurately predicted key traits such as leaf mass per area. Kühn et al. (2025) applied a spectral model to 20th-century herbarium specimens to show temporal trends in leaf carbon, nitrogen, and phosphorus content associated with intensifying agricultural practices in Germany.

Figure 1. Plant traits derived from full-range spectral reflectance (400 - 2500 nm) from fresh leaf tissue (A) and dried, pressed leaves (B). Spectral regions can be characterized into the visible range (VIS, 400-700 nm, shown in rainbow colors in both A and B), the near infrared (NIR, ~700-1100), and the short-wave infrared (SWIR1 ~1100-2000, and SWIR2, ~2000-2500, shown by the gray shaded areas in (A). Plant pigments (chlorophyll a and b, carotenoids, anthocyanins), traits related to water content and water potential, including leaf dry mass concentration (LDMC), equivalent water thickness (EWT), and mobile, degradable or small molecular weight carbon compounds (sugars, nonstructural carbohydrates) can be derived from spectra in fresh leaves. In dry leaves, water-related and pigmentation traits may not be reliably derived due to degradation. In both fresh and dry leaves, it may be possible to estimate macronutrients (nitrogen (N), carbon (C), phosphorus (P), potassium (K)) and micronutrients (calcium (Ca), zinc (Zn), iron (Fe), copper (Cu), sulfur (S), boron (B), magnesium (Mg), manganese (Mn), aluminum (Al), sodium (Na)), as well as large stable carbon-based molecules (phenolics, lignin, cellulose or hemicellulose), defense compounds (phenolics, flavonoids) and stable isotopes ($\delta^{13}C$, $\delta^{15}N$). Indeed,

they may be more accurately detected from dried tissue due to the absence of water. Water absorbs energy throughout the spectral range shown, but particularly in the bands indicated in A. Some of the spectral features that are important for predicting traits are indicated in B, adapted from Cavender-Bares et al. 2025.



Box 1 Brief glossary of terminology applied to plant reflectance

Band - a range of wavelengths within a spectrum for which the radiation is detected.

Bandwidth - the range or span of the spectral band.

Dark current - the small electric current that flows through a spectroradiometer even in the absence of incident light.

Electromagnetic spectrum - Wavelengths of electromagnetic radiation that span short wavelength (high frequency) to long-wavelength (low frequency) wavebands.

Fiber optic - strands of glass or plastic (optical fibers) used to transmit light, for example as reflectance from a leaf to an instrument detector.

Full range spectrum - Spectroradiometers used for foliar spectroscopy commonly span the range of 400 - 2500 nm covering the visible (VIS), near-infrared (NIR) and shortwave infrared (SWIR), which is considered full range (VIS-NIR-SWIR) wavelengths. Alternately, some instruments only span 400-1000 nm, which covers only the visible and near-infrared.

Herbarium specimen - a specimen with associated taxonomic, geographic, temporal and other collection metadata stored in an herbarium, for which we might not know how it was dried or conserved

Hyperspectral - Spanning the visible to near-infrared or shortwave infrared measuring narrow wavelength bands every 3 to 5 nm.

InGaS sensor - photodiodes for near-infrared (NIR) and short wavelength infrared (SWIR) regions.

Near-infrared (NIR) - the range of electromagnetic spectrum from 700 - 1100 nm.

Partial Least Square Discriminant Analysis (PLS-DA) - a machine learning approach used for the classification of high-dimensional datasets such as spectral data. PLS-DA utilizes spectral phenotypic data to create a simplified classification method to distinguish among taxonomic or functional groups.

Pressed plant - a controlled dried specimen that was pressed flat to dry

Partial Least Square Regression (PLSR) - is a statistical method (multivariate regression) used to model the relationship between X (predictor) and Y (response) variables. In spectral biology, PLSR is used to estimate foliar traits due to its ability to handle multicollinearity and reduce the dimensionality of spectral data.

Radiative transfer models (RTM) - physical models using computer programs to simulate the reflectance, transmittance, and absorption of solar radiation in various media, including leaves. Governed by physical laws, RTMs can operate in forward mode to predict spectral responses of leaves based on material characteristics or in backward mode to infer material properties from leaf spectra. These models vary in complexity, balancing computational demand, accuracy, and scalability. The number of traits that can be simultaneously identified is limited. For remotely sensed imagery, inputs such as solar radiation and elevation angle, and parameters like leaf angle distribution and leaf area index help to solve the radiative transfer equation in optical or thermal domains.

Reflectance standards - A reflectance standard is a physical reference sample that includes ratio values between the total amount of radiation, as of light, reflected by a surface, and the total amount of radiation incident on the surface across the measured spectrum. Reflectance standards are used for the calibration and verification of spectrometers.

Spectral reflectance - Reflectance expressed as a function of wavelength (i.e., as a spectrum). Spectral reflectance is the fraction of the incident radiant flux that is reflected as a function of wavelength

Short wave infrared (SWIR) - The range of the electromagnetic spectrum between 1100 and 3000 nm.

Spectral biology -The study of biological processes and biodiversity using spectral resolved observations of light's interaction with biological systems to reveal plant chemistry, structure, and function across scales, from genomes to ecosystems.

Spectroradiometer -A device designed to measure electromagnetic radiation across a range of wavelengths; a radiometrically calibrated spectrometer (c.f. spectrometer).

Spectroscopy - The study of interactions between electromagnetic radiation and matter, used in imaging to measure reflected radiation from image pixels that is used to analyze the properties of leaves, canopies, ecosystems, and landscapes over time and space.

Spectral resolution - the measure of a sensor or spectrometer's ability to distinguish between closely spaced wavelengths within the electromagnetic spectrum. It is quantified by the width of the spectral bands, the

minimum resolvable wavelength interval, or, specifically for spectrometers, as the ratio of the measured wavelength to the full width at half maximum (FWHM) of a spectral peak ($R = \lambda/\Delta\lambda$). Higher spectral resolution, characterized by narrower bands or smaller FWHM values, allows for finer discrimination of spectral features and thus, more detailed identification of materials.

Spectral range - Span between the smallest and largest wavelengths that a sensor or system can detect.

Taxon discrimination methods - analytic methods aimed at classifying unidentified samples into taxonomic categories, or clustering samples according to taxonomic affinities.

Trait models - physical, statistical or machine learning models used to predict foliar traits from leaf spectra

Visible range (VIS) - the range of electromagnetic spectrum from 400 - 700 nm.

Applying spectral signatures to phenotype-genotype associations and evolutionary models

Spectroscopy provides a source of phenotypic data that can be measured on samples across the plant tree of life and coupled with genetic and genomic information (**Fig. 2**) or to scale up the study of phenotypic variation across large spatial, temporal, and phylogenetic scales from variation within species to the plant tree of life (Meireles et al. 2020, Cavender-Bares et al. 2025). Integrating spectra with genomic data allows researchers to determine factors and evolutionary processes involved in shaping underlying traits (Matsuda et al. 2012; Čepl et al. 2018). For example, spectral data can reveal variations in traits relevant to ecological niches, while genomic data provides insight into the genetic architecture underlying these traits (Blonder et al. 2020; Madritch et al. 2014). Together, these kinds of datasets can reveal how populations adapt to environmental pressures (Ge et al. 2019, Galan et al. 2020). Such integration revolutionizes our understanding of how adaptation unfolds at both microevolutionary and macroevolutionary scales. By comparing spectral data with genomic markers, it is possible to identify genetic loci associated with key traits, such as cold tolerance or drought resistance, and assess how these loci vary across populations (Madritch et al. 2014, Cavender-Bares et al. 2016b, Czyż et al. 2020). This integration illuminates the demographic history and selective pressures driving adaptation. For instance, closely related species of *Quercus* subsection *Virentes* exhibit various forms of evolutionary divergence, including sympatry, allopatry, and parapatry (Cavender-Bares et al. 2015). These species show spectral phenotypic divergence associated with specific microhabitats, shaped by local adaptation to ecological variation (Hernandez-Leal et al., 2025). This example highlights how demographic processes and selective pressures differ across spatial scales. Phenotype-genotype associations offer a robust framework for uncovering the interplay between natural selection, demographic processes, and adaptation (Kokaly et al. 2009, Cavender-Bares et al. 2016b, Deacon et al. 2017b, Blonder et al. 2020, Stasinski et al. 2021, Hernandez-Leal et al. 2025). Coupling spectral data with genomic and genetic information expands opportunities for understanding evolutionary processes across the plant tree of life (Ge et al. 2016).

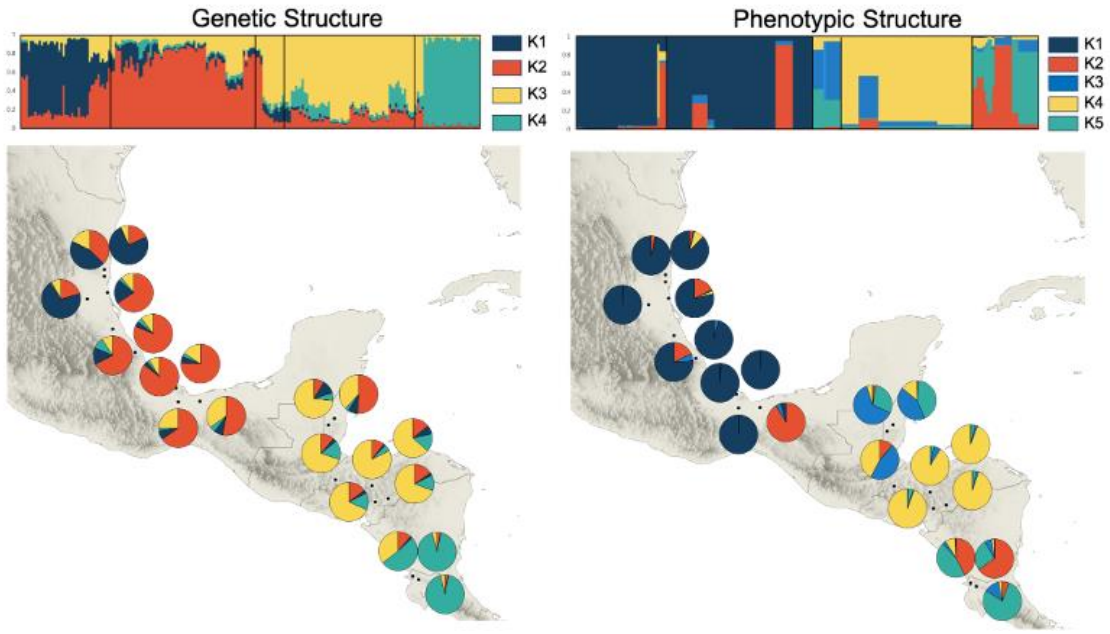
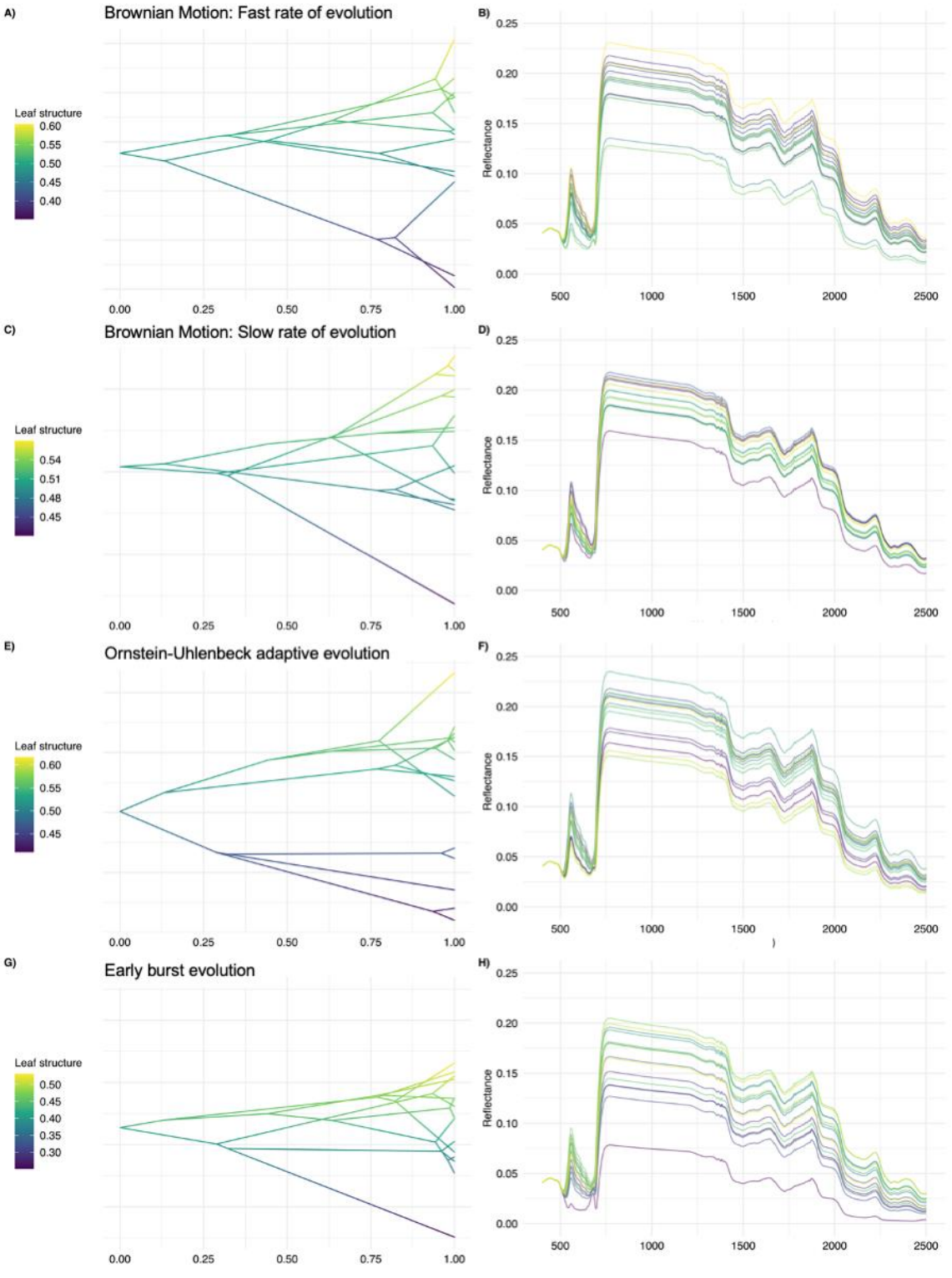


Figure 2. An example of combining genotypic and phenotypic data derived from pressed leaf spectra. The spatial distribution of genetic (left) and phenotypic (right) variation in *Quercus oleoides.* Four genetic groups were identified using STRUCTURE (Pritchard 2000) from 123 individuals and five phenotypic groups identified by GENELAND (Guillot et al., 2009) using six spectrally derived leaf traits (leaf mass area, thickness, solubles, hemicellulose, cellulose, lignin) derived from dried leaf spectra from all individuals. The percentage assignment to genetic or phenotypic groups is represented at both the individual tree level (upper bar plots) and subpopulation level (pie charts). Genetic data are from Cavender-Bares et al. 2015. Figure is adapted from Hernandez-Leal et al. 2025.

Spectral data can also be used to model species traits on phylogenetic trees using comparative methods (Harvey and Purvis 1991, Cavender-Bares 2019, Meireles et al. 2020a). Traits derived from spectra, or spectral indices and variables themselves, can be used to model evolutionary rates, constraints and selection when spectra are consistently measured across species. In a simulation exercise, we show how dried leaf spectra are expected to evolve according to four models of evolution, Brownian motion random walk model with fast and slow rates of evolution, a single-optimum Ornstein-Uhlenbeck model of adaptive evolution, and an Early Burst model with declining rates of evolution through time (**Fig. 3**). A major challenge, however, is the high dimensionality of full-range spectral data, which are comprised of hundreds of interdependent variables per measurement. Incorporating this high-dimensionality within a phylogenetic comparative framework is challenging both statistically and computationally because of the high number of coevolving traits in a lower number of taxa (Clavel et al. 2019). Novel comparative methods such as penalized likelihood approaches can now be applied to spectral data in phylogenetic frameworks (Clavel et al. 2015, Clavel et al. 2019).

340


A)
Brownian Motion: Fast rate of evolution
Leaf structure
0.60
0.55
0.50
0.45
0.40
0.00
0.25
0.50
0.75
1.00
B)
Reflectance
0.25
0.20
0.15
0.10
0.05
0.00
500
1000
1500
2000
2500
C)
Brownian Motion: Slow rate of evolution
Leaf structure
0.54
0.51
0.48
0.45
D)
E)
Ornstein-Uhlenbeck adaptive evolution
Leaf structure
0.60
0.55
0.50
0.45
F)
G)
Early burst evolution
Leaf structure
0.50
0.45
0.40
0.35
0.30
H)
Relative time
Wavelength (nm)

341

Figure 3. Simulated evolution of leaf structure (i.e., numbers of cell layers, N) under three models of evolution and the corresponding dried leaf spectra predicted from leaf traits using the PROSPECT D radiative transfer models (Féret et al. 2017) following Meireles et al. (2020a). Each row represents one model: Brownian motion with fast (A, B) or slow (C,D) evolutionary rates, a single-optimum Ornstein-Uhlenbeck adaptive evolution process (E,F), and an Early Burst process with declining rates over time (G,H). The graphs on the left show phylogenetic trees through relative time (x-axis) with y-values and branch colors indicating simulated leaf structure values. The graphs on the right show simulated reflectance spectra (400–2500 nm) for the final trait values of each lineage. (Starting values for the traits in the models are as follows: N=0.5, chlorophyll, Cab=10; carotenoids, Car=12.0; LMA/leaf dry matter content, Cm = 0.0005; water, Cw=0; leaf structure; sigma value for N= 0.1.)

Taxon discovery and discrimination

Spectral data have become a valuable tool for evaluating and refining taxonomic hypotheses, enabling rapid, non-destructive assessment of phenotypic cohesion and differentiation among taxa (2024). Across taxonomic ranks, spectra from dry leaves have been used to supplement DNA-based methods and support systematic studies, particularly in morphologically complex clades. In the Amazon, researchers have analyzed spectral absorption profiles with a Fourier-transform near-infrared (FT-NIR) spectrometer on unmounted leaf samples to distinguish closely related species within the genus *Eschweilera* (Lecythidaceae) (Durgante et al. 2013) as well as among different developmental stages of species in the Burseraceae (Lang et al. 2015). Paiva et al. (2021) applied spectroscopy to pressed fern fronds to classify species in the genus *Microgramma* with over 90% accuracy Both Prata et al. (2018) and Damasco et al. (2019) integrated DNA and spectral data to suggest taxonomic solutions in species complexes, including the reestablishment of *Protium cordatum* (Burseraceae) to species rank (Damasco et al. 2019). Spectroscopy is now being used to support new species hypotheses based on morphological and morphometric characteristics (Vasconcelos et al. 2020, da Cruz Vasconcelos et al. 2021, Gaem et al. 2022, Costa et al. 2025) or to differentiate hybrids from parental species (Deacon et al. 2017a).

Applications of these approaches to herbarium specimens are more recent. White et al. (2025) showed that spectral models from herbarium specimen leaves of diverse ages and sources (with a median age of 91 years) could classify specimens to species with up to 74% accuracy and genera up to 84% accuracy. In a broad study of stone oak (*Lithocarpus*) specimens from across Asia, Neto-Bradley et al. (2025) found that machine learning models trained on spectra work nearly as well as those from digitized images for identifying taxa. They concluded that spectra may be particularly important for identifying incomplete specimens of historical significance that may otherwise only be identified to family. In the future, it should be possible to combine digital images with spectral profiles to increase our power of discrimination at different taxonomic levels.

These studies highlight the feasibility of using spectra from herbarium specimens for species delimitation, taxonomic revision, and the detection of cryptic diversity. Discriminatory models offer an additional dimension of variation—rooted in chemistry and structure—that can complement morphology and genetics. Moreover, spectral classification can serve as a test of taxonomic validity, revealing whether phenotypic variation aligns with expected groupings. High within-taxon variation or convergent phenotypes can reduce accuracy, while informative spectral regions may shed light on the evolutionary processes shaping trait divergence.

For plant systematists, these advances represent a major opportunity to re-explore herbarium collections with new eyes. Spectral profiles can uncover consistent differences in chemical and structural traits in specimens that may not be evident through morphology. This offers an additional dimension of variation for species delimitation, taxonomic revisions, and the identification of cryptic diversity, especially in groups with limited diagnostic characters or incomplete material. As the number of spectral studies increases, the development of shared protocols and metadata standards will be essential to ensure reproducibility and cross-study comparability, enabling broader synthesis and integration into taxonomic workflows.

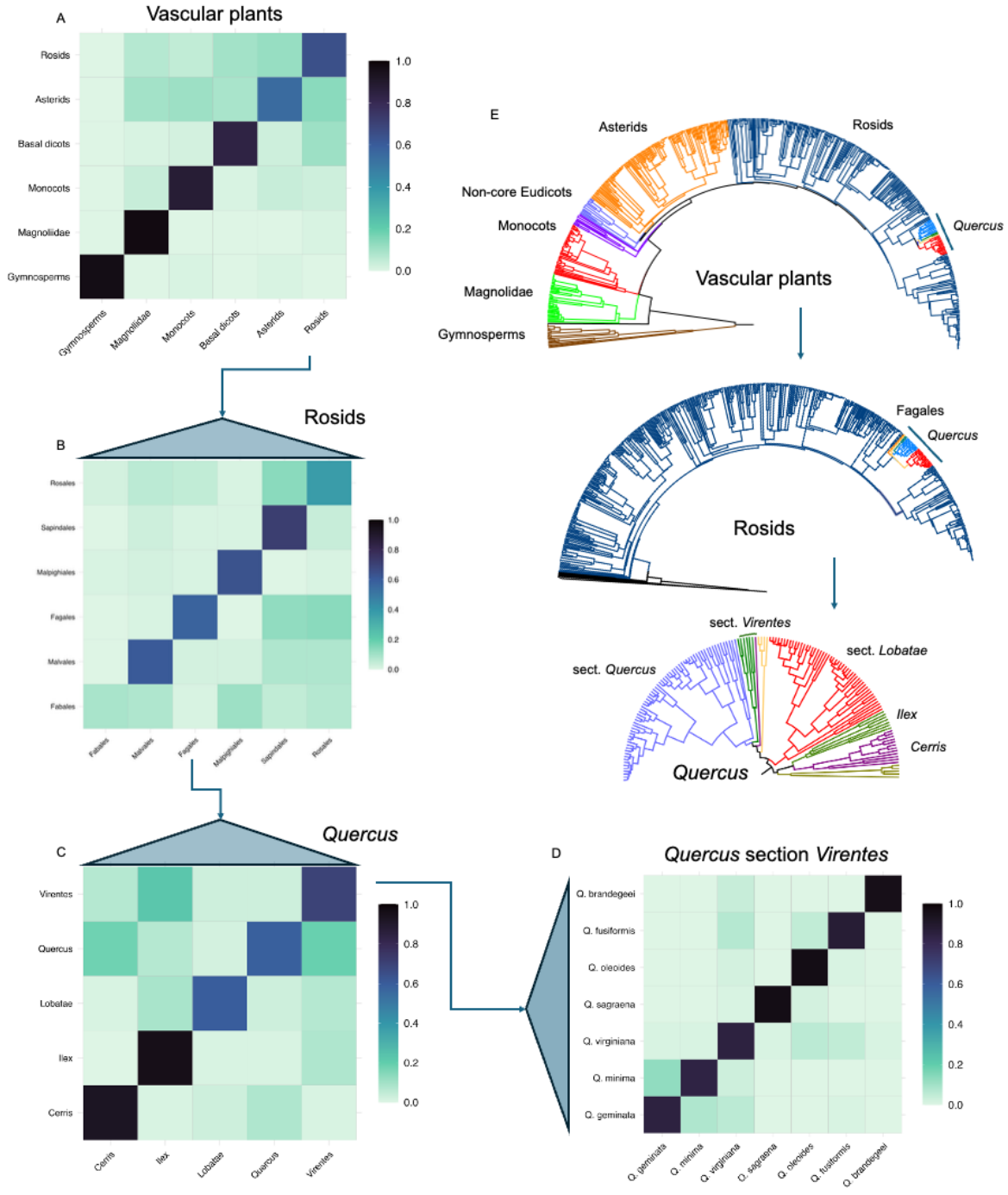
Operational workflows in botanical taxonomy are advancing rapidly, driven by the robustness and consistency of the results achieved across diverse taxonomic groups. This momentum will be further amplified by the development of compact handheld and benchtop spectrometers that have the potential to greatly expand the use of spectroscopy in herbarium collections worldwide. These devices enable individual herbaria to conduct spectral analysis of type specimens and other restricted materials, overcoming common barriers such as loan limitations and fragility of historical collections. Yet there are challenges to every aspect of this work, which if not addressed will make using the data collected difficult. Through an international collaborative working group, Meireles et al. (2020a) combined spectral data from fresh leaves collected from three different instruments models in different parts of the world by different investigators. By reconciling the instrument-specific variation and converting the data to a common format, they showed that reflectance signatures show phylogenetic signal and are tightly coupled to the tree of life with the capacity to reveal evolutionary history. This effort demonstrated both the difficulty and the benefits of aggregating data.

Generating spectral measurements from specimens of known identity will enable the development of reference datasets for identification purposes and may help to provide a preliminary taxonomic classification of the multitudes specimens in herbaria that have not been identified because they lack diagnostic structures, or have not yet been studied by sufficiently knowledgeable experts (Neto-Bradley et al. 2025). Fresh leaf spectra show promise for classifying species to all taxonomic ranks (Meireles et al. 2020b, Mallmann et al. 2023, Blanchard et al. 2024, Hadlich et al. 2025), and dried leaf

spectra may show even greater promise because the removal of the water absorption bands in the SWIR may reveal more unique absorption features (Kothari et al. 2023). Stepwise hierarchical approaches for classifying taxa first to broader and then narrower clades is a promising approach to accelerate the identification of undetermined (dried) herbarium material (**Fig. 4**). Such an approach would get around the problem of computational limits and reduced accuracy when large numbers of classes are discriminated among. To achieve high accuracies, highly populated spectral libraries at each phylogenetic scale would be necessary.

The success of taxonomic discrimination depends on the similarity of leaf and other plant part morphologies within each of the assigned taxa – and can be interpreted as a test of cohesion within groups and distinction between them. As such, an application of taxonomic discrimination models can be used to test the taxonomy itself. If there is nearly as much phenotypic variation within as between taxa, or if phenotypes evolve convergently, the success of taxonomic discrimination may be limited. Discriminatory models can potentially reveal the spectral regions that can best distinguish taxa, which may contribute to deciphering how taxa have diverged or converged and the evolutionary forces that have shaped their phenotypes.

Figure 4. Stepwise hierarchical approach to taxon classification to place specimens within the plant tree of life. Classification algorithms, such as partial least squares discriminant analysis (PLSDA), are limited by statistical power in the number of entities they can accurately discriminate between. One possible solution is a nested approach where spectral signatures are used to differentiate broad clades within the vascular plants, and stepwise within increasingly smaller clades, such as orders, families, genera—or increasingly narrow phylogenetic lineages. Shown are confusion matrices from broad to increasingly narrow taxonomic groups, starting with A) broad clades in the vascular plants, including the Rosid clade, B) orders within the Rosids, including the Fagales, C) sections of the genus *Quercus* within the Fagales, and D) species within *Quercus* section *Virentes*. Correct assignments are on the diagonal, incorrect assignments are in the off-diagonal cells. Data used in A and B and the vascular plant phylogeny are from Meireles et al. 2020; spectral data used C are from Cavender-Bares et al. 2016, spectral data in D are from Hernandez-Leal et al. 2025. The *Quercus* phylogeny is from Hipp et al. 2020.



Benefits and challenges to aggregating and scaling spectral data

The beauty of the 'global metaherbarium' lies in its ability to connect a wide array of data types that together comprise the extended specimen (Webster 2017, Lendemer et al. 2020, National Academies of Sciences and Medicine 2020, Davis 2023, Davis and Knapp 2025). These include not only the physical specimen itself, but also associated data such as digital images, species distribution models (SDMs), DNA sequences, and ecological trait measurements (Webster 2017, Lendemer et al. 2020) promising layers of

information that can significantly enhance the value of the specimen and the power of global-scale analyses if consistent or interoperable approaches can be established. While there is considerable excitement for the application of spectrometry in herbaria, a series of methodological, material, and data processing challenges must be addressed. Differences in instrumentation, specimen preservation, and data processing pipelines introduce variability that complicates large-scale data aggregation and cross-institutional compatibility (**Fig. 5**, (White et al. 2025). These issues were encountered and overcome by the NIMBioS (National Institute for Mathematical and Biological Synthesis) working group (Meireles et al. 2020c), which aggregated and harmonized spectral data collected from fresh leaves using different instruments and protocols around the world. Their work demonstrated both the promise and difficulty of harmonizing spectral data, given that spectra can vary significantly based on optical system design and measurement conditions. A further set of challenges arises from the varied nature of how herbarium specimens are treated upon intake, their age, their storage conditions, the impact of chemical additives such as glues, pest prevention treatments, and other specimen conservation factors (Kühn et al. 2025, White et al. 2025). To maximize the utility of spectral data in functional trait estimation, taxonomic classification, and evolutionary modeling, standardized measurement protocols across institutions, data quality controls, and robust correction methods are critical.

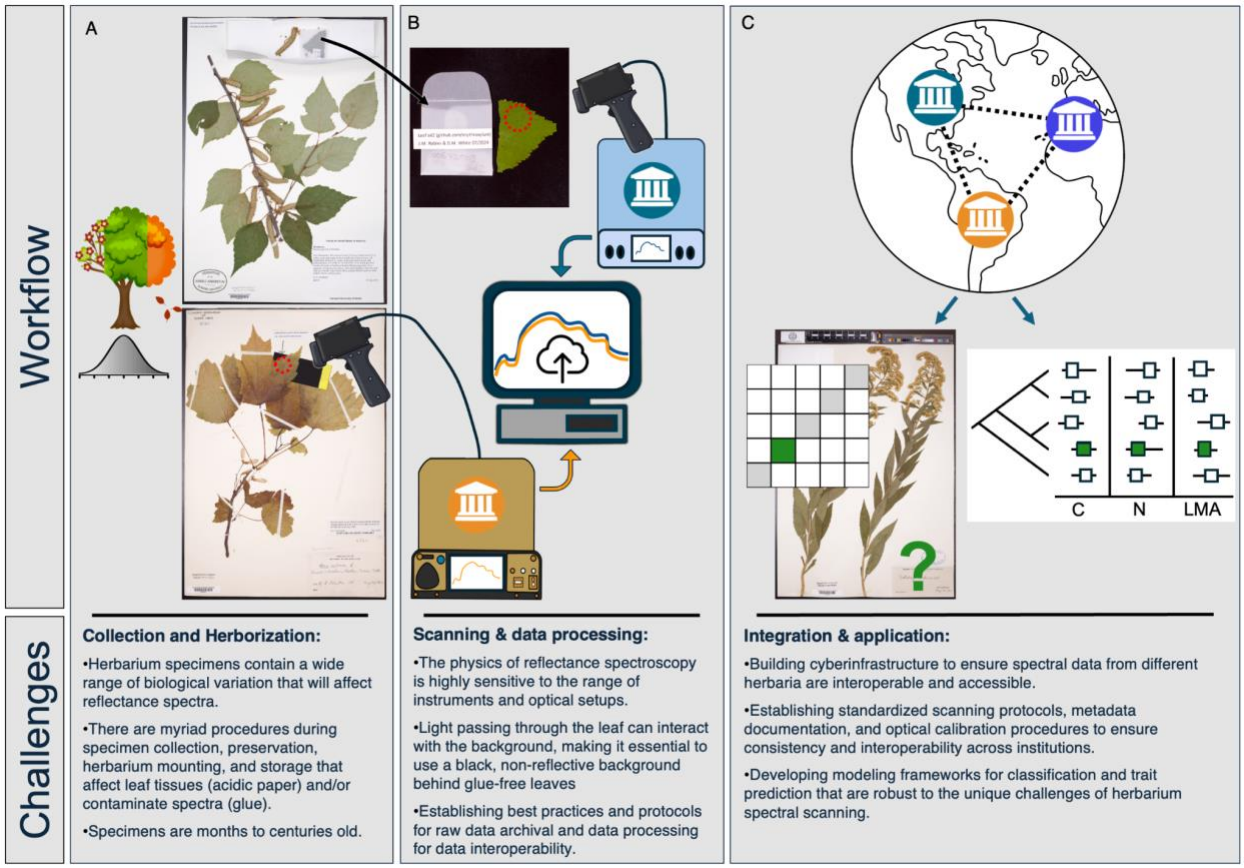


Figure 5. The herbarium spectral digitization workflow from specimen collection to global integration, and challenges. A) Variation in specimen condition and mounting methods, including changes in leaf preservation quality, use of adhesives or sewing, and presence of detached leaves in packets. B) Spectral digitization process using different spectroradiometers, with emphasis on black background placement and data upload; differences in instruments, optics, and data processing. C) Global integration of spectral data across herbaria, linked to specimen records and trait data such as carbon (C), nitrogen (N), and leaf mass per area (LMA).

Spectral regions to measure

Spectroscopy has been widely used to analyze fresh leaf tissue in the range of 400-2500 nmto obtain information about plant function and phenotype (Table S1). The visible range (400 - 700 nm) is strongly associated with chlorophyll, carotenoids, anthocyanins and other pigments (Gitelson et al. 1998, Gamon and Surfus 1999). Chlorophyll is partially degraded when exposed to light and may be reduced in herborized leaves, but it can still be detected. Spectral information related to carotenoids, flavonoids, lignin, and other chemical compounds remain readily detected and may be less susceptible to degradation. In the near-infrared-region (NIR, 1000-2500 nm), vibration frequencies associated with molecular functional groups (-CH, -NH, -OH) show features of primary metabolites (carbohydrates, lipids, and proteins) as well as of secondary metabolites (phenolic compounds, terpenoids, and alkaloids) (Türker-Kaya and Huck 2017, Jacquemoud and Ustin 2019, Ustin and Jacquemoud 2020). Spectral regions beyond 2500 nm reveal additional information. In the mid- and long-wave infrared regions (2500-25000 nm), for example, detailed chemical identification is possible, particularly for pure chemical compounds, (Türker-Kaya and Huck 2017), providing opportunities to detect ecologically and evolutionarily important differences among organisms, even when we do not have their taxonomic names.

Variation in foliar phenotypes. Spectral data capture variation in leaves due to a multitude of factors, including leaf developmental and ontogenetic stage, light exposure and environmental conditions during growth, and a range of other factors (Fajardo and Siefert 2016). Exposure of the foliar tissue to a range of environmental conditions within a single tree canopy, across environmental gradients or in response to temperature, water availability or CO_2 concentrations can modify tissue properties (Stefanski et al. 2025) and thus influence the reflection of light. Spectral properties and traits change over the lifecycle of the leaf (e.g., Fajardo and Siefert 2016, Chlus and Townsend 2022). Within a single leaf, tissue properties will change from the edges to the vascular tissue. With age, leaves add secondary cell wall material, cuticles become more structurally complex and thicker, and concentrations of biochemical compounds change, including compounds that are upregulated or produced in response to stress. Herbarium specimens representing species with asynchronous flowering or fruiting may disproportionately contain young, developing leaves or reproductive structures, leading to spectra that do not adequately represent

mature leaves (**Fig. 6**). Predictive models show promise in accommodating such biological variation as long as models are trained to "see" the full range of variation (Lang et al. 2015, Wang et al. 2023).

Adaxial and abaxial leaf surfaces. Adaxial and abaxial surfaces on vascular plants are often distinct. Given that leaf spectroscopy is often used to ground-truth remote sensing measurements, many protocols only involve measuring the adaxial surface of the leaf, which is the side most likely to be observed from above. Measurements of herbarium specimens have different goals, however, and the option is available to measure both leaf surfaces. The abaxial surface may contain important features for taxonomic identification, including pubescence or differing pigmentation. Whether both leaf surfaces should be measured as standard is an area for further investigation and discussion. In any case, it is important to document the leaf surface measured, given their spectral signatures are different (Ustin et al. 2001).

Leaf size. Small leaves present challenges to spectral measurement because they require adaptations to optical detection arrangements, such as fine leaf probes, small radius fiber optic cables, and/or small detection windows, which can reduce the signal-to-noise ratio and quality of the reflectance signal. Conifer needles, vascular plants with tiny leaves, and non-vascular plants such as mosses and lycopodia, can be difficult to measure due to their small size and the fact that many of them do not lie flat (**Fig. 6**). When leaves are not flat, the 'geometry' or angle of the light source relative to the surface of the leaf often varies from one measurement to the next, influencing the specular component of reflectance. When possible, it may help to measure the flattest parts of the leaves. Alongside these challenges, lichens also pose challenges due to their changing form and physiology when they dehydrate.

Specimen age. Over time specimens may undergo chemical or even structural degradation due to exposure to environmental factors. Although some aspects of this process have been documented, including the breakdown of chlorophyll and accumulation of brown pigments (Fourty et al. 1996), in general little is known about the suite of changes through time and the factors that slow or accelerate them. White et al. (2025) found that older specimens exhibit a slight decline in correct classification probability from spectra. The potential influences of degradation on spectra reinforces the need for careful specimen selection and possibly data filtering strategies that account for specimen age and preservation history when integrating spectral datasets (Durgante et al. 2013, Lang et al. 2015).

Figure 6. Variation in the challenges of measuring reflectance spectra on mounted specimens. A) An example of a relatively straightforward measurement of a specimen leaf against a black background using a one-sided leaf clip with a fiber optic probe that is smaller diameter than the leaf laminar surface area. Isotypes of (B) *Restio arcuatus* and (C) *Vulpia microstachys* are taxa with small leaves. D) *Abies*

balsamea, a conifer with typical narrow needles that is taped to the herbarium sheet but has also been glued, evidenced by the discolored glue indicated with the arrow. A fully taped *Salix arctica* (E) specimen from the Botanical Museum of Copenhagen with conspicuously gray discoloration representing variability in preservation techniques and quality. An herb (F) with delicate, thin leaves that are fully glued to the sheet with no packet containing extra material to measure. A *Sticta* lichen (G) with a non-flat thallus that will distort optical geometry. Circles on the top left demonstrate various aperture sizes (mm) typical of different optical instruments. For comparison, scale bars are shown in B, C and F, with a 2 cm bar in B.

Variation in preservation among specimens in herbarium collections.

Herbarium collections are generally curated to maximize the taxonomic and/or geographical diversity of specimens for given regions, and may represent centuries of collection efforts across biomes. The historical and biogeographic breadth has led to heterogeneity in specimen preservation techniques. Variability in storage conditions among herbarium facilities compounds the variation due to preservation methods. The most significant factors affecting specimen degradation and quality are the protocols implemented to initially press and dry the plant collection in the field. The best case scenario would be that a collected specimen is immediately pressed inside acid-free paper and dried with forced-air at ambient temperatures. However, this practice is rarely followed, and specimens, especially in the humid

554 tropics, are usually sealed in a bag containing 60-95% ethanol to kill fungus and prevent rotting. In all
555 latitudes, pressed specimens are routinely dried using industrial ovens.
556 Glue is a particularly significant contaminating source for herbarium spectra due to its potential
557 for direct contamination. For example, White et al. (2025) found that glue reduced the probability of
558 correct species identification from spectra. Sewing specimens to the herbarium sheets or using archival
559 grade mounting tape, both durable and secure mounting practices commonly used in European and South
560 American herbaria, can avoid this problem. However, it is highly labor-intensive and has not been a
561 standard practice in many herbaria. Consequently, glues are the most widely used adhesive in North
562 America. The leaves of some specimens may contain multiple layers of different adhesives. Other
563 common historical practices that may have resulted in the contamination of spectra include methyl
564 bromide fumigation (now banned) and the sprinkling of diatom powder and various poisons such as
565 arsenic or naphthalene. Efforts to 'unmix' or 'subtract' the glue or paper spectra from the leaf using
566 spectral libraries of these contaminants have not yet yielded solutions to isolating the leaf signal from a
567 spectral profile that contains these extra materials (A. Guzmán, B. Neto-Bradley, JE Meireles,
568 unpublished data). Although many large herbaria in well-funded institutions can maintain temperature
569 and humidity controls, humidity fluctuations remain a significant challenge worldwide, potentially
570 accelerating specimen degradation.
571
572 *Instruments and optics.* Variation in detectors and optical setups among instruments create differences in
573 spectral signatures, adding complexity to the use of herbaria for spectral data collection (**Fig. 5, Table 1**).
574 The signal-to-noise ratio, field-of-view, and spectral resolution and range are inherent instrument
575 variables that can influence optimal measurements on specimens. Standardization of spectral digitization
576 protocols and proper documentation is thus crucial for ensuring interoperability among spectral
577 digitization efforts and integration into broader analytical frameworks. The signal-to-noise ratio (SNR),
578 for instance, determines the quality and reliability of the measurements, where higher values mean clearer
579 and more precise spectral data. SNR has a major effect on the comparison of measurements between
580 instruments and, thus, the transferability of potential models derived from them. SNR is mostly
581 influenced by light detector sensitivity (particularly affecting the wavelengths near the edges of the
582 detector's spectral range) and light source intensity. Inappropriate selection of integration times, incorrect
583 calibration, or variation in viewing and illumination angles can impact the SNR. The removal of bands
584 with low SNR at the edge of the spectral range can help integrate data from different instruments into
585 models (White et al. 2025). However, removing bands should not be a default approach and should be
586 carefully considered in relation to project goals.

The field-of-view (FOV) – the angular extent of the observable area that can be seen through foreoptics or lenses – is another variable that often differs between instruments and influences how well specimens are captured in a single measurement. A wider FOV allows for the capture of a larger portion of the specimen, thereby increasing the representation of intraspecific variability. However, wider FOVs are usually associated with lower SNR measurements and are not ideal for small leaves and make it difficult to target specific leaf regions such as the blade rather than the central vein. Additionally, some instruments present a non-uniform FOV due to the optical fiber bundle and its integration with the light detector, which does not uniformly cover the viewing area (Lévesque et al. 2014). Using instruments with a narrow or non-uniform FOV might require several measurements of a leaf specimen to adequately capture the variability of optical information.

The spectral range and resolution are additional instrument variables inherent in instrument design. The spectral range of instruments is distinct from spectral resolution and is commonly used to differentiate between VIS-NIR spectrometers (e.g., 350 – 1000 nm), VIS-SWIR spectrometers (e.g., 350 – 2500 nm; also known as 'full range'), MIR (2.5-6.0 μm and and TIR spectrometers (e.g., 6.0 – 16 μm) because these detect photons in different regions of the electromagnetic spectrum. Although VIS-NIR spectrometers are less expensive than full-range spectrometers, some studies on dried leaves have begun to suggest that full-range spectra perform better for trait prediction (Kothari et al. 2024; White et al. 2025). Moreover, the spectral resolution of many full-range instruments varies within instruments because there are usually two or more light detectors, a silicon sensor (VIS and NIR) and an InGaAs (indium gallium arsenide) detector for the 1000–1700 nm wavelength region (e.g, Spectral Evolution) or an extended InGaAs covering the spectrum to 2500 nm (NIR and SWIR), each with distinct sensitivity and band sampling. Due to a trade-off between sensitivity and signal availability (i.e., light), silicon detectors designed for VIS-NIR wavelengths (e.g., 400 – 1000 nm)–where halogen bulbs generate peak irradiance–have high wavelength resolution. Consequently, they often achieve higher spectral resolution than detectors that span NIR-SWIR wavelengths (e.g., 1000 – 2500 nm). Instruments commonly report data output in 1-nm increments for spectral sampling of bandwidths 3 to 6 nm wide. Given the variation in design among instruments, the scanning interval (band center and band width) at different spectral ranges should be recorded in the metadata (see below).

Additional factors influencing the quality of spectral measurement data include the sensitivity of fiber optics, quality of standards, viewing and illumination angles, and the intensity of light sources (Grant 1987). Variations in fiber optic alignment can impact the signal-to-noise ratio, requiring careful handling and regular replacement (manufacturer maintenance should confirm whether all fibers are functional or, if not, determine the percentage of remaining fibers and their locations in the cluster). Use of a standardized leaf probe with integrated illumination avoids problems with the positioning of the

lights and the spectrometer. Regular replacement of calibration standards and routine instrument maintenance, such as cleaning and recalibrating sensors, are essential for maintaining instrument performance. The proper selection of light source intensity along with optimized integration times is critical to avoiding heat effects on specimens, which can alter both their optical properties and integrity. Lamps should be operated on direct current, not alternate current power supplies.

Table 1. Challenges to aggregating spectral data across specimens and institutions and potential solutions.

Challenge	Description	Mitigation Strategy
Variation in leaf phenotypes due to environment	Environmental variation (light, temperature, CO_2, nutrients) can influence spectra. Variation within a leaf	Collect spectra on specimens across a range of conditions to capture the full variation; avoid midrib and leaf edges, avoid leaves with herbivory and pathogens.
Variation in leaf phenotypes due to leaf age/development stage	Spectral properties and traits change over the lifecycle of the leaf	Capture information about leaf age in the metadata
Specimen age	Chemical degradation, pigmentation loss, and exposure to environmental factors over time can influence spectral signatures	Collect spectra on specimens across a range of ages to capture the full variation. When possible,conduct analyses on specimen age effects. Effects may not be large. Metadata should include date that specimen was acquired and location.
Preservation and storage variation	Drying, decontamination methods, adhesives, and storage environments can influence spectra	Standardize methods for incoming specimens. Conduct analyses on preservation and storage effects to determine if effects are large or small. Increase sample size to capture range of preservation methods.
Adhesives	Glues contaminate the reflectance spectrum	Prioritize leaves (or other tissues) without adhesives.
Leaf size	Small leaves are difficult to measure with standard set-ups	Fine leaf probes, narrow fiber optic cables, small detection windows
Leaf surface	Adaxial and abaxial surfaces differ	Measure both sides of the leaf. Capture metadata on which side(s) was/were measured
Measuring plant surfaces beyond leaves	Flowers, stems and other organs are also important	Measure on a project-by-project basis or using criteria at individual herbaria
Specimen geometry	Non-flat foliar surfaces generate specular influences on reflectance spectra (“noisy spectra”)	Prioritize flat tissues. Increase sample size. Use probe with a small optical aperture.
Instruments and optics	Different instruments give different spectral reflectance signatures. Optical setups also influence spectra.	Adhere to agreed upon protocols, instrument setup and metadata standards. Establish interconversion

		methods. Use the same reflectance standards across all labs
Reflectance background	White mounting paper introduces artefacts into the reflectance spectrum	Use a black non-reflective background on detached leaves or insert under a mounted leaf that is loose
Measurement settings and protocols	Sample count, leaf surface choice, integration times and other user-defined parameters influence SNR, data comparability and the time it takes to measure a specimen	Develop minimum sample counts and data standards across herbaria that can be surpassed for specific purposes and/or within individual herbaria
Destructive sampling for traits	Validation of models for many functional traits (N, P, cellulose, etc.) requires destroying leaf tissue to make direct observations.	Models may be generated from pressed leaves or tissues will need to be sampled ethically and in a coordinated manner.
Metadata collection and access	Data incompatibility due to the range of variables affecting herbarium spectral measurements	Establish metadata standards, similar to Darwin Core format, agreed upon, recorded and shared by all institutions.

Considerations for standardization of protocols and metadata

While detailed protocols and metadata standards remain to be established and agreed upon across institutions, it is important to provide some guidance at this stage, given the rapid increase in the use of spectroscopy for plant studies (Cavender-Bares et al. 2025). Here we make a series of recommendations and pose questions for consideration. Standardizing spectral digitization setups—including light source positioning, spectral calibration, and reference materials—will be essential to ensure compatibility between datasets collected using different instruments and workflows. Standardization of measurement protocols and metadata will be critical to maintaining consistency across instruments, collections, and institutions.

Sampling. We recommend sampling the variation among leaves on a specimen as well as variation within a single leaf, avoiding leaves with blemishes, contaminants or pathogens (unless these are relevant to the objectives of the study). When selecting among specimens for sampling, prioritizing those that have leaves where a black, non-reflective background can be readily inserted under the leaf will enable aggregation of data. Measuring sufficient leaf area to ensure that the spectra are representative of the variation within the specimen should be balanced against the care required to ensure non-destructive sampling. Taking 3-5 measurements across the leaf and 3-5 leaves per specimen can help capture relevant variation, but one fully expanded leaf per specimen may be a sufficient compromise to save time and avoid damage.

Reflectance standards. Given the variation in instrument and optical arrangements across institutions, regular measurement of standard cards of known reflectance using a set of certified reflectance standards ranging from black to gray to white provides a means to standardize data, making them comparable. A rare earth panel can be used to confirm wavelength calibration, especially if the instrument is returned for manufacturer recalibration infrequently. The black and gray standards would be measured regularly, such as once daily during measurement sessions. The white standard should be measured with each specimen, or at least every two to three specimens if spectral measurements are taken very rapidly.

Reflectance background. Background effects also play a critical role in spectral consistency across institutions. Light transmitted through leaves may reflect from the background (glue, paper, and even lab benches), which 'contaminates' the spectrum. The degree of contamination depends on the optical thickness of the leaves, which governs how much light is transmitted. Measuring leaves against a non-reflective black background to avoid contamination from other reflective surfaces is critical. This may create challenges for some specimens, which are often affixed to paper. If the specimen has been sewn or taped, it should be possible to slide a thin black sheet between the attached leaf and the paper. Herbarium specimens with loose leaves available in packets may be selected, and those leaves checked for glue before measuring them against a black background. A contact probe or modified leaf clip with the bottom portion removed will facilitate measurements. The importance of using a non-reflective black background will likely prevent automation through conveyor belts, as has been used for digitization of specimen images.

The selection of non-reflective black backgrounds is a critical component of standardization. A rule of thumb is that the background should have less than 4% reflection. EVA foam, black plastic, black card stock painted with Krylon® Camouflage Matte Black spray paint, and SpectralBlack® foil, or black backgrounds of manufacturer leaf clips for portable spectroradiometers (Malvern Panalytical [ASD], Spectra Vista Corporation, Spectral Evolution) are among the materials that have been used. It is worthwhile collecting a spectrum (or spectra) of the non-reflective black material to have as a reference if there are questions about the leaf measurements later. The identification and adoption of a universal background standard is an important objective of protocol development.

However, as noted above, vast portions of specimens already present in collections have been mounted with glue, such that using a non-reflective black background is not a feasible solution in all cases (Neto-Bradley et al., 2025). The extent to which measurements on mounted leaves should still be collected is an open area of discussion.

Metadata standards. Metadata standardization is critical for harmonizing datasets, as it facilitates the integration of phenomic data with associated specimen metadata, such as taxonomy, collection locality, and ecological context. Metadata will also provide users critical data on specimen preservation method, and potential artefacts due to glues or reflective backgrounds, that can inform aggregation and integration efforts. By adopting common metadata schemas and persistent identifiers (DOIs), researchers can link spectral data directly to digital databases, fostering seamless collaboration and data reuse. Experience gained from successful protocol standardization and data aggregation initiatives (e.g. Darwin Core and iDigBio; (Wieczorek et al. 2012, Soltis 2017) can be leveraged to implement a strategy for herbarium spectroscopic data. Recording metadata for instrument type, including the spectrometer brand and model, date of last maintenance and calibration by the manufacturer, whether the instrument is internally calibrated, as well as the spectral resolution and wavelength interval the instrument actually measures. The setup and standard measurements across a set of standard reference cards (white, gray, black) will be essential for aggregation. Also important is recording developmental, phenological, and ecological factors that influence leaf structure and physiology, as these are critical to untangling the effects of specimen processing and storage. Effective metadata curation and cyberinfrastructure development will be critical for integrating herbarium spectral data into global biodiversity platforms. Standardizing metadata fields for specimen age, preservation method, mounting medium, and measurement conditions will facilitate dataset comparison and allow researchers to apply appropriate quality control measures. Advances in spectral data repositories, such as linking reflectance data to GBIF, *species*Link and iDigBio records, will enhance accessibility and ensure that spectral datasets are fully interoperable with existing biodiversity databases (Heberling 2021, Davis 2023) and further enhance the extended specimen concept (Lendemer et al. 2020).

Data processing. Raw data should be stored as collected directly from the instrument to avoid artifacts in the data that cannot be undone, e.g., resampling or binning bands to achieve higher SNR. Processed data may also be included. There are various approaches to processing raw spectra through resampling and normalization or transformation using derivatives or continuous wavelet transforms (CWT) to standardize datasets from different instruments. Band resampling at a higher resolution than the true resolution could introduce artificial data to the spectrum but resampling to reduce the number of bands is common. A resampling interval of 5 nm is reasonable for accommodating the differences in spectral resolution between instruments (e.g., Spectral Evolution NaturaSpec, PSR+, or SVC HR-1024i). Resampling reduces the number of correlated bands for predictive models. Although small absorption features may be perceived as insignificant or as noise, they may be important for taxonomic classification, and binning or resampling to lower resolution could result in loss of information. If the raw spectral data is preserved,

any loss of information can be assessed by comparison to the raw data. In full-range instruments, the 350–400 nm and 2,400–2,500 nm regions are often trimmed due to the low signal to noise ratio in those regions. Reflectance values at the wavelengths where the silicon and InGaAs sensors meet need to be joined consistently. R packages like spectrolab (Meireles et al. 2017) have options for resampling, trimming, and joining data between the silicon and InGaAs sensors.

The success of aggregating herbarium reflectance spectra will depend on practices that ensure consistency, interoperability, and accessibility across institutions. Data quality controls must be ensured to include careful adherence to measurement protocols. It should also involve the preprocessing of spectral datasets to remove spectra of low quality. Routine validation processes will ensure that data meets the necessary standards for reproducibility and analysis. Finally, implementing analysis engines capable of handling high-dimensional datasets will be transformative. These engines should integrate spectral data with complementary datasets, such as genomic or spatial data, and provide tools for advanced modeling and visualization. Open-source analysis platforms with user-friendly interfaces will democratize access to these tools and foster collaboration across disciplines.

Fitting appropriate functional trait models to specimens. A key consideration is the transferability of spectral models between fresh, pressed, and herbarium specimens. While recent studies have confirmed that these specimen types are all informative for functional trait prediction and taxonomic classification (Wagner et al. 2019, Kothari et al. 2023), accounting for minor spectral shifts due to aging and preservation during model building will be important. Spectral models trained on recently pressed specimens may perform well on herbarium specimens, but in some cases where degradation will likely bias trait modeling (e.g., chlorophyll). Developing correction factors and harmonization techniques for cross-specimen spectral applications will be necessary to expand the utility of herbarium spectral databases. When seeking to apply a trait model, it is critical to collect some ground-truth data to ensure that it can yield accurate predictions for a particular class of specimens, based on taxa sampled, the herborization process, and instrumental set up. Without validation, there is no basis for ensuring that models developed for one set of specimens will accurately predict traits in another set of specimens.

Database sufficiency

The development of cyberinfrastructure has been pivotal in enabling large-scale aggregation of spectral data. Platforms like iDigBio, *species*Link and GBIF provide centralized repositories for biodiversity data, but dedicated cyberinfrastructure for spectral datasets, integrated with existing platforms, will be essential for advancing collections-based research. These systems should support real-time synchronization of available data from herbarium institutions, cross-referencing, and retrieval for

global accessibility. Any dedicated spectral cyberinfrastructure platforms will require Application Programming Interfaces (APIs) to enable researchers to query, retrieve, and contribute spectral datasets programmatically; facilitating the large-scale synthesis of data. While several cyberinfrastructure platforms have been developed specifically for spectral data and models, such as EcoSIS and EcoSML (Wagner et al. 2019), they are not appropriate or sufficient for capturing taxonomic information required in biodiversity collections. A more efficient solution will be to ensure that the databases from natural history museums can incorporate spectral data and the critical associated metadata.

Darwin Core. The Darwin Core (DC) standard is a compendium of terms and definitions that describe the observation of an organism at a particular time and space (Wieczorek et al. 2012). Broadening the Darwin Core to incorporate herbarium reflectance spectroscopy will enable data sharing across existing biodiversity informatics platforms (e.g., GBIF, *species*Link) and the use of spectral information in biodiversity research and conservation. In addition, extending DC for spectroscopy will supply new means for sharing protocols and practices for data aggregation that are critical for improving the quality of the herbarium reflectance data collection, storage, and distribution.

Effective, Equitable and Ethical Sampling

Herbarium users and stewards must balance innovation with preservation. Guidelines for the effectiveness and ethics of these global biodiversity heritage have been recently formalized to provide a path to better utilize and steward these collections to safeguard their continued use (Davis et al. 2025). Many of these guidelines have already been in place for decades and are often included in institutional policies that have been deliberated by stewards and shared with users (Richard et al. 2019, Shah 2023). The herbarium and Natural History Museum community has become increasingly aware of what we call the ‘destructive sampling conundrum’: how to foster innovative research that includes destructive sampling of specimens that are meant to be protected permanently (Davis et al. 2025). We additionally recognize that herbaria are differentially concentrated around the world, especially the global north (Heberling et al. 2019, Park et al. 2023), and are of different sizes and face a range of operating circumstances and challenges.

The effective implementation of spectroscopic sampling and measurement strategies requires meticulous consideration of both the specimen's characteristics and the preservation and repair methods that the world’s herbaria employ. Under ideal conditions, spectral digitization provides a non-destructive alternative to traditional trait measurements, allowing researchers to infer key plant traits without physically altering specimens (White et al. 2025). By non-destructive, we mean the specimen is left the way it was previously with no trace of tissue removal (Davis et al. 2025). Spectral data can be collected non-destructively when either detached leaves are available or when some leaves on a specimen are not

fully glued. Damage or removal of specimen tissue for spectroscopy efforts are likely to occur if non-destructive protocols are not enforced. While spectral measurements can be made non-destructively, if trait prediction is the ultimate goal, some amount of trait validation is necessary to improve the accuracy of trait prediction. Trait validation will require the destructive use of tissue from similar taxa ideally with similar herborization processes. The high performance of trait models from pressed leaves (Kothari et al. 2023) applied to herborized specimens (White et al. 2025) indicates that pressed leaves of the same taxa can probably be used for validation rather than herbarium specimens. If tissue removal from specimens is required, it should be absolutely minimal, avoiding key features for identification or possible future use, and documenting what was removed. It is crucial to seek permission for these efforts. To minimize the need for destructive sampling, unwanted duplicates could be retained for spectroscopic and destructive trait measurements for calibration of trait models. Alternatively, excess duplicate specimens could be set aside and offered to the network of herbarium spectroscopy researchers, advancing the development of spectroscopic libraries for trait estimation. Adhering to non-destructive measurement approaches should be strongly encouraged to maintain the structural and historical integrity of the specimens. Researchers should obtain explicit permission before conducting studies and ensure that their methodologies align with the conservation objectives of the hosting institutions.

Finally, it is important that any databased content or spectral signatures harvested be offered as shared data to the host institution as part of the global metaherbarium, making the data available immediately. Doing so will help to avoid duplication of efforts with the same collection at other institutions. Finally, it is essential to promote collaboration, including by recognizing herbaria scholars, staff, and scientists formally as contributors, collaborators, funded partners, and coauthors (Edwards et al. 2024, Davis et al. 2025) following practices that require extraction of tissue or other alterations to herbarium specimens (e.g., Burbano and Gutaker 2023, Davis 2023, Davis and Knapp 2025). Establishing clear guidelines for sampling ensures that scientific investigations do not compromise the preservation of these collections. For example, the American Society of Plant Taxonomists has outlined best practices that emphasize the importance of minimizing physical alterations to specimens. By adhering to such protocols, researchers can foster trust and facilitate ongoing access to these vital resources. All of these considerations are paramount in the application of leaf spectroscopy to herbarium specimens and for effective, equitable, and ethical sampling strategies (Davis et al. 2025). Sharing findings and derivative data with the broader scientific community will promote transparency and collective advancement of the plant sciences.

Next Steps

In closing, we declare that the international herbarium spectral digitization (IHerbSpec) working group is ready to go spectral! We acknowledge that we have considerable work to do as a community to ensure that data can be aggregated and to ensure fair and equitable practices. Our next step is to develop agreed upon protocols and metadata standards for the community. We pledge to be collaborative, inclusive and equitable in our work and to generate data from the world's herbaria to advance understanding and protection of the world's biodiversity.

Acknowledgements

We thank Jorge Robles for contributions to Figure 5. For funding, we thank the US NSF Biology Integration Institute ASCEND, Advancing Spectral biology in Changing Environments to understand plant Diversity (DBI: 2021898) to JCB, the Harvard University Herbaria, iDigBio (NSF DBI-2027654) to PSS and NSF grant DBI-1930007 to DES and PSS, Canadian NSERC Discovery Grant, support from the Missouri Botanical Garden's Revolutionizing Species Identification project. TLPC and KB are supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 865787).

Author contributions

All authors contributed intellectually to the manuscript, including the writing, editing and tables. JCB, JPL, DW, MHL, JEM and BNB prepared or contributed to the figures. JCB organized and led the writing process with help from DW.

References Cited

Asner, G., and R. Martin. 2011. Canopy phylogenetic, chemical and spectral assembly in a lowland Amazonian forest. . New Phytologist **189**:999-1012

Blanchard, F., A. Bruneau, and E. Laliberté. 2024. Foliar spectra accurately distinguish most temperate tree species and show strong phylogenetic signal. American Journal of Botany **111**:e16314.

Blonder, B., B. J. Graae, B. Greer, M. Haagsma, K. Helsen, R. E. Kapás, H. Pai, J. Rieksta, D. Sapena, C. J. Still, and R. Strimbeck. 2020. Remote sensing of ploidy level in quaking aspen (Populus tremuloides Michx.). Journal of Ecology **108**:175-188.

Burbano, H. A., and R. M. Gutaker. 2023. Ancient DNA genomics and the renaissance of herbaria. Science **382**:59-63.

Cavender-Bares, J. 2019. Diversification, adaptation, and community assembly of the American oaks (*Quercus*), a model clade for integrating ecology and evolution. New Phytologist **221**:669-692.

Cavender-Bares J, Gonzalez-Rodriguez A, Eaton DAR, Hipp AAL, Beulke A, Manos PS. 2015. Phylogeny and biogeography of the American live oaks (Quercus subsection Virentes): a genomic and population genetics approach. Molecular Ecology 24(14): 3668-3687.

Cavender-Bares, J., D. Ackerly, S. Hobbie, and P. Townsend. 2016a. Evolutionary legacy effects on ecosystems: Biogeographic origins, plant traits, and implications for management in the era of global change. Annual Review of Ecology, Evolution, and Systematics **47**:433-462.

Cavender-Bares, J., J. A. Gamon, S. E. Hobbie, M. D. Madritch, J. E. Meireles, A. K. Schweiger, and P. A. Townsend. 2017. Harnessing plant spectra to integrate the biodiversity sciences across biological and spatial scales. American Journal of Botany **104**:1-4 doi: 10.3732/ajb.1700061.

Cavender-Bares, J., A. Gonzalez-Rodriguez, D. A. R. Eaton, A. A. L. Hipp, A. Beulke, and P. S. Manos. 2015. Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): a genomic and population genetics approach. Molecular Ecology **24**:3668-3687.

Cavender-Bares, J., J. E. Meireles, J. J. Couture, M. A. Kaproth, C. C. Kingdon, A. Singh, S. P. Serbin, A. Center, E. Zuniga, G. Pilz, and P. A. Townsend. 2016b. Associations of leaf spectra with genetic and phylogenetic variation in oaks: prospects for remote detection of biodiversity. Remote Sensing **8**:221-238.

Cavender-Bares, J., J. E. Meireles, J. Pinto-Ledezma, P. B. Reich, M. C. Schuman, P. A. Townsend, and A. Trowbridge. 2025. Spectral biology across scales in changing environments. Ecology **in press**.

Cavender-Bares, J., F. Schneider, M. Santos, A. Armstrong, A. Carnaval, K. Dahlin, L. Fatoyinbo, G. Hurtt, D. Schimel, and P. Townsend. 2022. Integrating remote sensing with ecology and evolution to advance biodiversity conservation. Nature Ecology & Evolution1-14.

Cavender-Bares, J., P. A. Townsend, and J. A. Gamon. 2020. Remote Sensing of Plant Biodiversity. Springer International Publishing.

Chlus, A., and P. A. Townsend. 2022. Characterizing seasonal variation in foliar biochemistry with airborne imaging spectroscopy. Remote Sensing of Environment **275**:113023.

Clavel, J., L. Aristide, and H. Morlon. 2019. A Penalized Likelihood Framework for High-Dimensional Phylogenetic Comparative Methods and an Application to New-World Monkeys Brain Evolution. Systematic biology **68**:93-116.

Clavel, J., G. Escarguel, and G. Merceron. 2015. mvmorph: an r package for fitting multivariate evolutionary models to morphometric data. Methods in Ecology and Evolution **6**:1311-1319.

Costa, D. d. S., E. F. S. Rossetto, and L. L. Giacomin. 2025. Neea contracta (Nyctaginaceae), a New Species from Amazonia. Systematic Botany **49**:743-748.

Czyż, E. A., C. Guillén Escribà, H. Wulf, A. Tedder, M. C. Schuman, F. D. Schneider, and M. E. Schaepman. 2020. Intraspecific genetic variation of a Fagus sylvatica population in a temperate forest derived from airborne imaging spectroscopy time series. Ecology and Evolution **10**:7419-7430.

da Cruz Vasconcelos, C., I. D. K. Ferraz, M. U. Adrianzén, J. L. C. Camargo, and M. H. Terra-Araujo. 2021. Chromolucuma brevipedicellata (Sapotaceae, Chrysophylloideae), a new tree species from central Amazonia, Brazil. Brittonia **73**:211-219.

Damasco, G., D. C. Daly, A. Vicentini, and P. V. A. Fine. 2019. Reestablishment of Protium cordatum (Burseraceae) based on integrative taxonomy. Taxon **68**:34-46.

Dao, P. D., Y. He, and C. Proctor. 2021. Plant drought impact detection using ultra-high spatial resolution hyperspectral images and machine learning. International Journal of Applied Earth Observation and Geoinformation **102**:102364.

Daru, B. H., D. S. Park, R. B. Primack, C. G. Willis, D. S. Barrington, T. J. S. Whitfeld, T. G. Seidler, P. W. Sweeney, D. R. Foster, A. M. Ellison, and C. C. Davis. 2018. Widespread sampling biases in herbaria revealed from large-scale digitization. New Phytologist **217**:939-955.

Davis, C. C. 2023. The herbarium of the future. Trends in Ecology & Evolution **38**:412-423.

Davis, C. C., J. A. Kennedy, and C. J. Grassa. 2021. Back to the future: A refined single-user photostation for massively scaling herbarium digitization. Taxon **70**:635-643.

Davis, C. C., and S. Knapp. 2025. Exploring biodiversity through museomics. Nature Reviews Genetics **26**:149-150.

Davis, C. C., E. Sessa, A. Paton, A. Antonelli, and J. K. Teisher. 2025. Guidelines for the effective and ethical sampling of herbaria. Nature Ecology & Evolution **9**:196-203.

Davis, C. C., C. G. Willis, B. Connolly, C. Kelly, and A. M. Ellison. 2015. Herbarium records are reliable sources of phenological change driven by climate and provide novel insights into species' phenological cueing mechanisms. Am J Bot **102**:1599-1609.

Deacon, N. J., J. J. Grossman, A. K. Schweiger, I. Armour, and J. Cavender-Bares. 2017a. Genetic, morphological, and spectral characterization of relictual Niobrara River hybrid aspens (Populus × smithii). American Journal of Botany **104**:1878-1890.

Deacon, N. J., J. J. Grossman, A. K. Schweiger, I. Armour, and J. Cavender-Bares. 2017b. Genetic, morphological, and spectral characterization of relictual Niobrara River hybrid aspens (Populus x smithii). American Journal of Botany **104**:1878-1890.

Dechant, B., J. Kattge, R. Pavlick, F. D. Schneider, F. M. Sabatini, Á. Moreno-Martínez, E. E. Butler, P. M. van Bodegom, H. Vallicrosa, T. Kattenborn, C. C. F. Boonman, N. Madani, I. J. Wright, N. Dong, H. Feilhauer, J. Peñuelas, J. Sardans, J. Aguirre-Gutiérrez, P. B. Reich, P. J. Leitão, J. Cavender-Bares, I. H. Myers-Smith, S. M. Durán, H. Croft, I. C. Prentice, A. Huth, K. Rebel, S. Zaehle, I. Šímová, S. Díaz, M. Reichstein, C. Schiller, H. Bruelheide, M. Mahecha, C. Wirth, Y. Malhi, and P. A. Townsend. 2024. Intercomparison of global foliar trait maps reveals fundamental differences and limitations of upscaling approaches. Remote Sensing of Environment **311**:114276.

Díaz, S., J. Kattge, J. H. C. Cornelissen, I. J. Wright, S. Lavorel, S. Dray, B. Reu, M. Kleyer, C. Wirth, I. Colin Prentice, E. Garnier, G. Bönisch, M. Westoby, H. Poorter, P. B. Reich, A. T. Moles, J. Dickie, A. N. Gillison, A. E. Zanne, J. Chave, S. Joseph Wright, S. N. Sheremet'ev, H. Jactel, C. Baraloto, B. Cerabolini, S. Pierce, B. Shipley, D. Kirkup, F. Casanoves, J. S. Joswig, A. Günther, V. Falczuk, N. Rüger, M. D. Mahecha, and L. D. Gorné. 2016. The global spectrum of plant form and function. Nature **529**:167-171.

Durgante, F. M., N. Higuchi, A. Almeida, and A. Vicentini. 2013. Species Spectral Signature: Discriminating closely related plant species in the Amazon with Near-Infrared Leaf-Spectroscopy. Forest Ecology and Management **291**:240-248.

Edwards, E. J., B. D. Mishler, and C. D. Davis. 2024. University herbaria are uniquely important. Trends in Plant Science **29**:825-826.

Elvidge, C. D. 1990. Visible and near infrared reflectance characteristics of dry plant materials. International Journal of Remote Sensing **11**:1775-1795.

Enquist, B. J., R. Condit, R. K. Peet, M. Schildhauer, and B. M. Thiers. 2016. Cyberinfrastructure for an integrated botanical information network to investigate the ecological impacts of global climate change on plant biodiversity. PeerJ Preprints **4**:e2615v2612.

Fajardo, A., and A. Siefert. 2016. Phenological variation of leaf functional traits within species. Oecologia **180**:951-959.

Fallon, B., A. Yang, C. Lapadat, I. Armour, J. Juzwik, R. A. Montgomery, and J. Cavender-Bares. 2020. Spectral differentiation of oak wilt from foliar fungal disease and drought is correlated with physiological changes. Tree Physiology **40**:377-390.

Falster, D., R. Gallagher, E. H. Wenk, I. J. Wright, D. Indiarto, S. C. Andrew, C. Baxter, J. Lawson, S. Allen, A. Fuchs, A. Monro, F. Kar, M. A. Adams, C. W. Ahrens, M. Alfonzetti, T. Angevin, D. M. G. Apgaua, S. Arndt, O. K. Atkin, J. Atkinson, T. Auld, A. Baker, M. von Balthazar, A. Bean, C. J. Blackman, K. Bloomfield, D. M. J. S. Bowman, J. Bragg, T. J. Brodribb, G. Buckton, G. Burrows, E. Caldwell, J. Camac, R. Carpenter, J. A. Catford, G. R. Cawthray, L. A. Cernusak, G. Chandler, A. R. Chapman, D. Cheal, A. W. Cheesman, S.-C. Chen, B. Choat, B. Clinton, P. L. Clode, H. Coleman, W. K. Cornwell, M. Cosgrove, M. Crisp, E. Cross, K. Y. Crous, S. Cunningham, T. Curran, E. Curtis, M. I. Daws, J. L. DeGabriel, M. D. Denton, N. Dong, P. Du, H. Duan, D. H. Duncan, R. P. Duncan, M. Duretto, J. M. Dwyer, C. Edwards, M. Esperon-Rodriguez, J. R. Evans, S. E. Everingham, C. Farrell, J. Firn, C. R. Fonseca, B. J. French, D. Frood, J. L. Funk, S. R. Geange, O. Ghannoum, S. M. Gleason, C. R. Gosper, E. Gray, P. K. Groom, S. Grootemaat, C. Gross, G. Guerin, L. Guja, A. K. Hahs, M. T. Harrison, P. E. Hayes, M. Henery, D. Hochuli, J. Howell, G. Huang, L. Hughes, J. Huisman, J. Ilic, A. Jagdish, D. Jin, G. Jordan, E. Jurado, J. Kanowski, S. Kasel, J. Kellermann, B. Kenny, M. Kohout, R. M. Kooyman, M. M. Kotowska, H. R. Lai, E. Laliberté, H. Lambers, B. B. Lamont, R. Lanfear, F. van Langevelde, D. C. Laughlin, B.-A. Laugier-Kitchener, S. Laurance, C. E. R. Lehmann, A. Leigh, M. R. Leishman, T. Lenz, B. Lepschi, J. D. Lewis, F. Lim, U. Liu, J. Lord, C. H. Lusk, C. Macinnis-Ng, H. McPherson, S. Magallón, A. Manea, A. López-Martinez, M. Mayfield, J. K. McCarthy, T. Meers, M. van der Merwe, D. J. Metcalfe, P. Milberg, K. Mokany, A. T. Moles, B. D. Moore, N. Moore, J. W. Morgan, W. Morris, A. Muir, S. Munroe, Á. Nicholson, D. Nicolle, A. B. Nicotra, Ü. Niinemets, T. North, A. O'Reilly-Nugent, O. S. O'Sullivan, B. Oberle, Y. Onoda, M. K. J. Ooi, C. P. Osborne, G. Paczkowska, B. Pekin, C. Guilherme Pereira, C. Pickering, M. Pickup, L. J. Pollock, P. Poot, J. R. Powell, S. A. Power, I. C. Prentice, L. Prior, S. M. Prober, J. Read, V. Reynolds, A. E. Richards, B. Richardson, M. L. Roderick, J. A. Rosell, M. Rossetto, B. Rye, P. D. Rymer, M. A. Sams, G. Sanson, H. Sauquet, S. Schmidt, J. Schönenberger, E.-D. Schulze, K. Sendall, S. Sinclair, B. Smith, R. Smith, F. Soper, B. Sparrow, R. J. Standish, T. L. Staples, R. Stephens, C. Szota, G. Taseski, E. Tasker, F. Thomas, D. T. Tissue, M. G. Tjoelker, D. Y. P. Tng, F. de Tombeur, K. Tomlinson, N. C. Turner, E. J. Veneklaas, S. Venn, P. Vesk, C. Vlasveld, M. S. Vorontsova, C. A. Warren, N. Warwick, L. K. Weerasinghe, J. Wells, M. Westoby, M. White, N. S. G. Williams, J. Wills, P. G. Wilson, C. Yates, A. E. Zanne, G. Zemunik, and K. Ziemińska. 2021. AusTraits, a curated plant trait database for the Australian flora. Scientific Data **8**:254.

Féret, J. B., A. A. Gitelson, S. D. Noble, and S. Jacquemoud. 2017. PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. Remote Sens. Environ. **193**:204-215.

Fine, P. V. A., D. Salazar, R. E. Martin, M. R. Metz, T. M. Misiewicz, and G. P. Asner. 2021. Exploring the links between secondary metabolites and leaf spectral reflectance in a diverse genus of Amazonian trees. Ecosphere **12**:e03362.

Funk, J. L., J. E. Larson, G. M. Ames, B. J. Butterfield, J. Cavender-Bares, J. Firn, D. C. Laughlin, A. E. Sutton-Grier, L. Williams, and J. Wright. 2017. Revisiting the Holy Grail: using plant functional traits to understand ecological processes: Plant functional traits. Biological Reviews **92**:1156-1173.

Gaem, P. H., A. Andrade, F. F. Mazine, and A. Vicentini. 2022. Tree species delimitation in tropical forest inventories: Perspectives from a taxonomically challenging case study. Forest Ecology and Management **505**:119900.

Gamon, J., and J. Surfus. 1999. Assessing leaf pigment content and activity with a reflectometer. New Phytologist **143**:105-117.

Ge, Y., G. Bai, V. Stoerger, and J. c. Schnable. 2016. Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging. Comput. Electron. Agric. **127**:625-632.

Gitelson, A. A., C. Buschmann, and H. K. Lichtenthaler. 1998. Leaf chlorophyll fluorescence corrected for re-absorption by means of absorption and reflectance measurements. J. Plant Physiol. **152**:283-296.

Gitelson, A. A., and M. N. Merzlyak. 1994. Spectral reflectance changes associate with autumn senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. leaves. Spectral features and relation to chlorophyll estimation. . Journal of Plant Physiology **143**:286–292.

Grant, L. 1987. Diffuse and specular characteristics of leaf reflectance. Remote Sensing of Environment **22**:309-322.

Guillot G, Santos F. 2009. A computer program to simulate multilocus genotype data with spatially autocorrelated allele frequencies. Molecular Ecology Resources 9: 1112-1120.

Hadlich, H. L., J. Schöngart, F. Wittmann, C. C. Vasconcelos, C. L. Mallmann, M. L. G. Conde, P. Amaral de Sá, L. O. Demarchi, G. B. Mori, M. T. F. Piedade, and F. M. Durgante. 2025. Scaling up tree diversity inventories across Amazonian ecosystems using field spectroscopy. bioRxiv:2025.2003.2026.645444.

Harvey, P. H., and A. Purvis. 1991. Comparative methods for explaining adaptations. Nature **351**:619-624.

Heberling, J. M. 2021. Herbaria as Big Data Sources of Plant Traits. International Journal of Plant Sciences **183**:87-118.

Heberling, J. M., J. T. Miller, D. Noesgaard, S. B. Weingart, and D. Schigel. 2021. Data integration enables global biodiversity synthesis. Proceedings of the National Academy of Sciences **118**:e2018093118.

Heberling, J. M., L. A. Prather, and S. J. Tonsor. 2019. The Changing Uses of Herbarium Data in an Era of Global Change: An Overview Using Automated Content Analysis. BioScience **69**:812-822.

Hedrick, B., M. Heberling, E. Meineke, K. Turner, C. Grassa, D. Park, J. Kennedy, J. Clarke, J. Cook, D. Blackburn, S. Edwards, and C. Davis. 2019. Digitization and the future of natural history collections.

Hernández-Leal MS, Guzmán Q. JA, González-Rodríguez A, Cavender-Bares J. 2025. Foliar spectral signatures reveal adaptive divergence in live oaks (Quercus1section Virentes)across species and environmental niches. EcoEvoRxiv https://ecoevorxiv.org/repository/view/9054/: https://doi.org/10.32942/X32942BK32990.

Hipp AL, Manos PS, Hahn M, Avishai M, Bodénès C, Cavender-Bares J, Crowl AA, Deng M, Denk T, Fitz-Gibbon S, et al. 2020. Genomic landscape of the global oak phylogeny. New Phytologist 226(4): 1198-1212.

Jacquemoud, S., and S. Ustin. 2019. Leaf Optical Properties. Cambridge University Press, New York.

Jetz, W., J. Cavender-Bares, R. Pavlick, D. Schimel, F. W. Davis, G. P. Asner, R. Guralnick, J. Kattge, A. Latimer, P. Moorcroft, M. E. Schaepman, and S. L. Ustin. 2016a. Monitoring plant functional diversity from space. Nature Plants **2**.

Jetz, W., J. Cavender-Bares, R. Pavlick, D. Schimel, F. W. Davis, G. P. Asner, R. Guralnick, J. Kattge, A. M. Latimer, P. Moorcroft, M. E. Schaepman, M. P. Schildhauer, F. D. Schneider, F. Schrodt, U. Stahl, and S. L. Ustin. 2016b. Monitoring plant functional diversity from space. Nature Plants **2**:16024.

Kattge, J., G. Bonisch, S. Diaz, S. Lavorel, I. C. Prentice, P. Leadley, S. Tautenhahn, G. D. A. Werner, T. Aakala, M. Abedi, A. T. R. Acosta, G. C. Adamidis, K. Adamson, M. Aiba, C. H. Albert, J. M. Alcantara, C. C. Alcazar, I. Aleixo, H. Ali, B. Amiaud, C. Ammer, M. M. Amoroso, M. Anand, C. Anderson, N. Anten, J. Antos, D. M. G. Apgaua, T. L. Ashman, D. H. Asmara, G. P. Asner, M. Aspinwall, O. Atkin, I. Aubin, L. Baastrup-Spohr, K. Bahalkeh, M. Bahn, T. Baker, W. J. Baker, J. P. Bakker, D. Baldocchi, J. Baltzer, A. Banerjee, A. Baranger, J. Barlow, D. R. Barneche, Z. Baruch, D. Bastianelli, J. Battles, W. Bauerle, M. Bauters, E. Bazzato, M.

Beckmann, H. Beeckman, C. Beierkuhnlein, R. Bekker, G. Belfry, M. Belluau, M. Beloiu, R. Benavides, L. Benomar, M. L. Berdugo-Lattke, E. Berenguer, R. Bergamin, J. Bergmann, M. B. Carlucci, L. Berner, M. Bernhardt-Romermann, C. Bigler, A. D. Bjorkman, C. Blackman, C. Blanco, B. Blonder, D. Blumenthal, K. T. Bocanegra-Gonzalez, P. Boeckx, S. Bohlman, K. Bohning-Gaese, L. Boisvert-Marsh, W. Bond, B. Bond-Lamberty, A. Boom, C. C. F. Boonman, K. Bordin, E. H. Boughton, V. Boukili, D. Bowman, S. Bravo, M. R. Brendel, M. R. Broadley, K. A. Brown, H. Bruelheide, F. Brumnich, H. H. Bruun, D. Bruy, S. W. Buchanan, S. F. Bucher, N. Buchmann, R. Buitenwerf, D. E. Bunker, J. Burger, S. Burrascano, D. Burslem, B. J. Butterfield, C. Byun, M. Marques, M. C. Scalon, M. Caccianiga, M. Cadotte, M. Cailleret, J. Camac, J. J. Camarero, C. Campany, G. Campetella, J. A. Campos, L. Cano-Arboleda, R. Canullo, M. Carbognani, F. Carvalho, F. Casanoves, B. Castagneyrol, J. A. Catford, J. Cavender-Bares, B. E. L. Cerabolini, M. Cervellini, E. Chacon-Madrigal, K. Chapin, F. S. Chapin, S. Chelli, S. C. Chen, A. P. Chen, P. Cherubini, F. Chianucci, B. Choat, K. S. Chung, M. Chytry, D. Ciccarelli, L. Coll, C. G. Collins, L. Conti, D. Coomes, J. H. C. Cornelissen, W. K. Cornwell, P. Corona, M. Coyea, J. Craine, D. Craven, J. Cromsigt, A. Csecserits, K. Cufar, M. Cuntz, A. C. da Silva, K. M. Dahlin, M. Dainese, I. Dalke, M. Dalle Fratte, T. D. L. Anh, J. Danihelka, M. Dannoura, S. Dawson, A. J. de Beer, A. De Frutos, J. R. De Long, B. Dechant, S. Delagrange, N. Delpierre, G. Derroire, A. S. Dias, M. H. Diaz-Toribio, P. G. Dimitrakopoulos, M. Dobrowolski, D. Doktor, P. Drevojan, N. Dong, J. Dransfield, S. Dressler, L. Duarte, E. Ducouret, S. Dullinger, W. Durka, R. Duursma, O. Dymova, A. E-Vojtko, R. L. Eckstein, H. Ejtehadi, J. Elser, T. Emilio, K. Engemann, M. B. Erfanian, A. Erfmeier, A. Esquivel-Muelbert, G. Esser, M. Estiarte, T. F. Domingues, W. F. Fagan, J. Fagundez, D. S. Falster, Y. Fan, J. Y. Fang, E. Farris, F. Fazlioglu, Y. H. Feng, F. Fernandez-Mendez, C. Ferrara, J. Ferreira, A. Fidelis, B. Finegan, J. Firn, T. J. Flowers, D. F. B. Flynn, V. Fontana, E. Forey, C. Forgiarini, L. Francois, M. Frangipani, D. Frank, C. Frenette-Dussault, G. T. Freschet, E. L. Fry, N. M. Fyllas, G. G. Mazzochini, S. Gachet, R. Gallagher, G. Ganade, F. Ganga, P. Garcia-Palacios, V. Gargaglione, E. Garnier, J. L. Garrido, A. L. de Gasper, G. Gea-Izquierdo, D. Gibson, A. N. Gillison, A. Giroldo, M. C. Glasenhardt, S. Gleason, M. Gliesch, E. Goldberg, B. Goldel, E. Gonzalez-Akre, J. L. Gonzalez-Andujar, A. Gonzalez-Melo, A. Gonzalez-Robles, B. J. Graae, E. Granda, S. Graves, W. A. Green, T. Gregor, N. Gross, G. R. Guerin, A. Gunther, A. G. Gutierrez, L. Haddock, A. Haines, J. Hall, A. Hambuckers, W. X. Han, S. P. Harrison, W. Hattingh, J. E. Hawes, T. H. He, P. C. He, J. M. Heberling, A. Helm, S. Hempel, J. Hentschel, B. Herault, A. M. Heres, K. Herz, M. Heuertz, T. Hickler, P. Hietz, P. Higuchi, A. L. Hipp, A. Hirons, M. Hock, J. A. Hogan, K. Holl, O. Honnay, D. Hornstein, E. Q. Hou, N. Hough-Snee, K. A. Hovstad, T. Ichie, B. Igic, E. Illa, M. Isaac, M. Ishihara, L. Ivanov, L. Ivanova, C. M. Iversen, J. Izquierdo, R. B. Jackson, B. Jackson, H. Jactel, A. M. Jagodzinski, U. Jandt, S. Jansen, T. Jenkins, A. Jentsch, J. R. P. Jespersen, G. F. Jiang, J. L. Johansen, D. Johnson, E. J. Jokela, C. A. Joly, G. J. Jordan, G. S. Joseph, D. Junaedi, R. R. Junker, E. Justes, R. Kabzems, J. Kane, Z. Kaplan, T. Kattenborn, L. Kavelenova, E. Kearsley, A. Kempel, T. Kenzo, A. Kerkhoff, M. I. Khalil, N. L. Kinlock, W. D. Kissling, K. Kitajima, T. Kitzberger, R. Kjoller, T. Klein, M. Kleyer, J. Klimesova, J. Klipel, B. Kloeppel, S. Klotz, J. M. H. Knops, T. Kohyama, F. Koike, J. Kollmann, B. Komac, K. Komatsu, C. Konig, N. J. B. Kraft, K. Kramer, H. Kreft, I. Kuhn, D. Kumarathunge, J. Kuppler, H. Kurokawa, Y. Kurosawa, S. Kuyah, J. P. Laclau, B. Lafleur, E. Lallai, E. Lamb, A. Lamprecht, D. J. Larkin, D. Laughlin, Y. Le Bagousse-Pinguet, G. le Maire, P. C. le Roux, E. le Roux, T. Lee, F. Lens, S. L. Lewis, B. Lhotsky, Y. Z. Li, X. E. Li, J. W. Lichstein, M. Liebergesell, J. Y. Lim, Y. S. Lin, J. C. Linares, C. J. Liu, D. J. Liu, U. Liu, S. Livingstone, J. Llusia, M. Lohbeck, A. Lopez-Garcia, G. Lopez-Gonzalez, Z. Lososova, F. Louault, B. A. Lukacs, P. Lukes, Y. J. Luo, M. Lussu, S. Y. Ma, C. M. R. Pereira, M. Mack, V. Maire, A. Makela, H. Makinen, A. C. M. Malhado, A. Mallik, P. Manning, S. Manzoni, Z. Marchetti, L. Marchino, V. Marcilio-Silva, E. Marcon, M. Marignani, L. Markesteijn, A. Martin, C. Martinez-Garza, J. Martinez-Vilalta, T. Maskova, K. Mason, N. Mason, T. J. Massad, J. Masse, I. Mayrose, J. McCarthy, M. L. McCormack, K.

McCulloh, I. R. McFadden, B. J. McGill, M. Y. McPartland, J. S. Medeiros, B. Medlyn, P. Meerts, Z. Mehrabi, P. Meir, F. P. L. Melo, M. Mencuccini, C. Meredieu, J. Messier, I. Meszaros, J. Metsaranta, S. T. Michaletz, C. Michelaki, S. Migalina, R. Milla, J. E. D. Miller, V. Minden, R. Ming, K. Mokany, A. T. Moles, V. A. Molnar, J. Molofsky, M. Molz, R. A. Montgomery, A. Monty, L. Moravcova, A. Moreno-Martinez, M. Moretti, A. S. Mori, S. Mori, D. Morris, J. Morrison, L. Mucina, S. Mueller, C. D. Muir, S. C. Muller, F. Munoz, I. H. Myers-Smith, R. W. Myster, M. Nagano, S. Naidu, A. Narayanan, B. Natesan, L. Negoita, A. S. Nelson, E. L. Neuschulz, J. Ni, G. Niedrist, J. Nieto, U. Niinemets, R. Nolan, H. Nottebrock, Y. Nouvellon, A. Novakovskiy, K. O. Nystuen, A. O'Grady, K. O'Hara, A. O'Reilly-Nugent, S. Oakley, W. Oberhuber, T. Ohtsuka, R. Oliveira, K. Ollerer, M. E. Olson, V. Onipchenko, Y. Onoda, R. E. Onstein, J. C. Ordonez, N. Osada, I. Ostonen, G. Ottaviani, S. Otto, G. E. Overbeck, W. A. Ozinga, A. T. Pahl, C. E. T. Paine, R. J. Pakeman, A. C. Papageorgiou, E. Parfionova, M. Partel, M. Patacca, S. Paula, J. Paule, H. Pauli, J. G. Pausas, B. Peco, J. Penuelas, A. Perea, P. L. Peri, A. C. Petisco-Souza, A. Petraglia, A. M. Petritan, O. L. Phillips, S. Pierce, V. D. Pillar, J. Pisek, A. Pomogaybin, H. Poorter, A. Portsmuth, P. Poschlod, C. Potvin, D. Pounds, A. S. Powell, S. A. Power, A. Prinzing, G. Puglielli, P. Pysek, V. Raevel, A. Rammig, J. Ransijn, C. A. Ray, P. B. Reich, M. Reichstein, D. E. B. Reid, M. Rejou-Mechain, V. R. de Dios, S. Ribeiro, S. Richardson, K. Riibak, M. C. Rillig, F. Riviera, E. M. R. Robert, S. Roberts, B. Robroek, A. Roddy, A. V. Rodrigues, A. Rogers, E. Rollinson, V. Rolo, C. Romermann, D. Ronzhina, C. Roscher, J. A. Rosell, M. F. Rosenfield, C. Rossi, D. B. Roy, S. Royer-Tardif, N. Ruger, R. Ruiz-Peinado, S. B. Rumpf, G. M. Rusch, M. Ryo, L. Sack, A. Saldana, B. Salgado-Negret, R. Salguero-Gomez, I. Santa-Regina, A. C. Santacruz-Garcia, J. Santos, J. Sardans, B. Schamp, M. Scherer-Lorenzen, M. Schleuning, B. Schmid, M. Schmidt, S. Schmitt, J. V. Schneider, S. D. Schowanek, J. Schrader, F. Schrodt, B. Schuldt, F. Schurr, G. S. Garvizu, M. Semchenko, C. Seymour, J. C. Sfair, J. M. Sharpe, C. S. Sheppard, S. Sheremetiev, S. Shiodera, B. Shipley, T. A. Shovon, A. Siebenkas, C. Sierra, V. Silva, M. Silva, T. Sitzia, H. Sjoman, M. Slot, N. G. Smith, D. Sodhi, P. Soltis, D. Soltis, B. Somers, G. Sonnier, M. V. Sorensen, E. E. Sosinski, N. A. Soudzilovskaia, A. F. Souza, M. Spasojevic, M. G. Sperandii, A. B. Stan, J. Stegen, K. Steinbauer, J. G. Stephan, F. Sterck, D. B. Stojanovic, T. Strydom, M. L. Suarez, J. C. Svenning, I. Svitkova, M. Svitok, M. Svoboda, E. Swaine, N. Swenson, M. Tabarelli, K. Takagi, U. Tappeiner, R. Tarifa, S. Tauugourdeau, C. Tavsanoglu, M. te Beest, L. Tedersoo, N. Thiffault, D. Thom, E. Thomas, K. Thompson, P. E. Thornton, W. Thuiller, L. Tichy, D. Tissue, M. G. Tjoelker, D. Y. P. Tng, J. Tobias, P. Torok, T. Tarin, J. M. Torres-Ruiz, B. Tothmeresz, M. Treurnicht, V. Trivellone, F. Trolliet, V. Trotsiuk, J. L. Tsakalos, I. Tsiripidis, N. Tysklind, T. Umehara, V. Usoltsev, M. Vadeboncoeur, J. Vaezi, F. Valladares, J. Vamosi, P. M. van Bodegom, M. van Breugel, E. Van Cleemput, M. van de Weg, S. van der Merwe, F. van der Plas, M. T. van der Sande, M. van Kleunen, K. Van Meerbeek, M. Vanderwel, K. A. Vanselow, A. Varhammar, L. Varone, M. Y. Valderrama, K. Vassilev, M. Vellend, E. J. Veneklaas, H. Verbeeck, K. Verheyen, A. Vibrans, I. Vieira, J. Villacis, C. Violle, P. Vivek, K. Wagner, M. Waldram, A. Waldron, A. P. Walker, M. Waller, G. Walther, H. Wang, F. Wang, W. Q. Wang, H. Watkins, J. Watkins, U. Weber, J. T. Weedon, L. P. Wei, P. Weigelt, E. Weiher, A. W. Wells, C. Wellstein, E. Wenk, M. Westoby, A. Westwood, P. J. White, M. Whitten, M. Williams, D. E. Winkler, K. Winter, C. Womack, I. J. Wright, S. J. Wright, J. Wright, B. X. Pinho, F. Ximenes, T. Yamada, K. Yamaji, R. Yanai, N. Yankov, B. Yguel, K. J. Zanini, A. E. Zanne, D. Zeleny, Y. P. Zhao, J. M. Zheng, J. Zheng, K. Zieminska, C. R. Zirbel, G. Zizka, I. C. Zo-Bi, G. Zotz, C. Wirth, and N. Nutrient. 2020. TRY plant trait database - enhanced coverage and open access. Global Change Biology **26**:119-188.

Kokaly, R., G. Asner, S. Ollinger, M. Martin, and C. Wessman. 2009. Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. Remote Sensing of Environment **113**:S78-S91.

Kothari, S., R. Beauchamp-Rioux, E. Laliberté, and J. Cavender-Bares. 2023. Reflectance spectroscopy allows rapid, accurate and non-destructive estimates of functional traits from pressed leaves. Methods in Ecology and Evolution **14**:385-401.

Kothari, S., and A. K. Schweiger. 2022. Plant spectra as integrative measures of plant phenotypes. Journal of Ecology **110**:2536-2554.

Kühn, P., R. Umazekabiri, C. Römermann, H. Bruelheide, and K. Wesche. 2025. Nitrogen content of herbarium specimens from arable fields and mesic meadows reflect the intensifying agricultural management during the 20th century. Journal of Ecology **113**:555-569.

Lang, C., F. R. C. Costa, J. L. C. Camargo, F. M. Durgante, and A. Vicentini. 2015. Near Infrared Spectroscopy Facilitates Rapid Identification of Both Young and Mature Amazonian Tree Species. Plos One **10**:e0134521.

Lavorel, S., and E. Garnier. 2002. Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. Functional Ecology **16**:545-556.

Lendemer, J., B. Thiers, A. K. Monfils, J. Zaspel, E. R. Ellwood, A. Bentley, K. LeVan, J. Bates, D. Jennings, D. Contreras, L. Lagomarsino, P. Mabee, L. S. Ford, R. Guralnick, R. E. Gropp, M. Revelez, N. Cobb, K. Seltmann, and M. C. Aime. 2020. The Extended Specimen Network: A Strategy to Enhance US Biodiversity Collections, Promote Research and Education. BioScience **70**:23-30.

Madritch, M. D., C. C. Kingdon, A. Singh, K. E. Mock, R. L. Lindroth, and P. A. Townsend. 2014. Imaging spectroscopy links aspen genotype with below-ground processes at landscape scales. Philosophical Transactions of the Royal Society of London B: Biological Sciences **369**.

Mallmann, C. L., W. Pereira Filho, J. B. B. Dreyer, L. A. Tabaldi, and F. M. Durgante. 2023. Leaf-Level Field Spectroscopy to Discriminate Invasive Species (Psidium guajava L. and Hovenia dulcis Thunb.) from Native Tree Species in the Southern Brazilian Atlantic Forest. Remote Sensing.

Mandrioli, M. 2023. From Dormant Collections to Repositories for the Study of Habitat Changes: The Importance of Herbaria in Modern Life Sciences. Life (Basel) **13**.

Meacham-Hensold, K., C. M. Montes, J. Wu, K. Guan, P. Fu, E. A. Ainsworth, T. Pederson, C. E. Moore, K. L. Brown, C. Raines, and C. J. Bernacchi. 2019. High-throughput field phenotyping using hyperspectral reflectance and partial least squares regression (PLSR) reveals genetic modifications to photosynthetic capacity. Remote Sensing of Environment **231**:111176.

Medeiros, L. S., S. M. Q. Lima, M. d. Pinna, I. C. A. Souto-Santos, S. Pirro, and W. M. Berbel-Filho. 2024. Whole-Genome Sequencing of Two Listrura and Five Microcambeva Species (Trichomycteridae, Siluriformes), Rare and Threatened Catfishes from the Atlantic Forest. Biodiversity Genomes **December**.

Meineke, E. K., T. J. Davies, B. H. Daru, and C. C. Davis. 2018a. Biological collections for understanding biodiversity in the Anthropocene. Philosophical Transactions of the Royal Society B: Biological Sciences **374**:20170386.

Meineke, E. K., C. C. Davis, and T. J. Davies. 2018b. The unrealized potential of herbaria for global change biology. Ecological Monographs **88**:505-525.

Meireles, J. E., J. Cavender-Bares, P. A. Townsend, S. Ustin, J. A. Gamon, A. K. Schweiger, M. E. Schaepman, G. P. Asner, R. E. Martin, A. Singh, F. Schrodt, A. Chlus, and B. O'Meara. 2020a. Leaf reflectance spectra capture the evolutionary history of seed plants. New Phytologist **228**:485-493.

Meireles, J. E., B. O'Meara, and J. Cavender-Bares. 2020b. Linking leaf spectra to the plant tree of life. Pages 155-172 *in* J. Cavender-Bares, J. A. Gamon, and P. A. Townsend, editors. Remote Sensing of Plant Biodiversity. Springer.

Meireles, J. E., B. O'Meara, and J. Cavender-Bares. 2020c. Linking Leaf Spectra to the Plant Tree of Life. Pages 155-172 *in* J. Cavender-Bares, J. A. Gamon, and P. A. Townsend, editors. Remote Sensing of Plant Biodiversity. Springer International Publishing, Cham.

Meireles, J. E., A. K. Schweiger, and J. M. Cavender-Bares. 2017. spectrolab: Class and Methods for Hyperspectral Data. R package version 0.0.2.

National Academies of Sciences, E., and Medicine. 2020. Biological Collections: Ensuring Critical Research and Education for the 21st Century. The National Academies Press, Washington, DC.
Nelson, G., P. Sweeney, L. Wallace, R. Rabeler, D. Allard, H. Brown, R. Carter, M. Denslow, L. Ellwood, C. Germain-Aubrey, E. Gilbert, E. Gillespie, L. Goertzen, B. Legler, T. Marsico, A. Morris, Z. Murrell, M. Nazaire, and A. Mast. 2015. Digitization Workflows for Flat Sheets and Packets of Plants, Algae, and Fungi. Applications in Plant Sciences **3**.
Paiva, D. N. A., R. d. O. Perdiz, and T. E. Almeida. 2021. Using near-infrared spectroscopy to discriminate closely related species: a case study of neotropical ferns. Journal of Plant Research **134**:509-520.
Park, D. S., X. Feng, S. Akiyama, M. Ardiyani, N. Avendaño, Z. Barina, B. Bärtschi, M. Belgrano, J. Betancur, R. Bijmoer, A. Bogaerts, A. Cano, J. Danihelka, A. Garg, D. E. Giblin, R. Gogoi, A. Guggisberg, M. Hyvärinen, S. A. James, R. J. Sebola, T. Katagiri, J. A. Kennedy, T. S. Komil, B. Lee, S. M. L. Lee, D. Magri, R. Marcucci, S. Masinde, D. Melnikov, P. Mráz, W. Mulenko, P. Musili, G. Mwachala, B. E. Nelson, C. Niezgoda, C. Novoa Sepúlveda, S. Orli, A. Paton, S. Payette, K. D. Perkins, M. J. Ponce, H. Rainer, L. Rasingam, H. Rustiami, N. M. Shiyan, C. S. Bjorå, J. Solomon, F. Stauffer, A. Sumadijaya, M. Thiébaut, B. M. Thiers, H. Tsubota, A. Vaughan, R. Virtanen, T. J. S. Whitfeld, D. Zhang, F. O. Zuloaga, and C. C. Davis. 2023. The colonial legacy of herbaria. Nature Human Behaviour **7**:1059-1068.
Pearson, K. D., 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, and P. S. Soltis. 2020. Machine Learning Using Digitized Herbarium Specimens to Advance Phenological Research. BioScience **70**:610-620.
Prata, E. M. B., C. Sass, D. P. Rodrigues, F. M. C. B. Domingos, C. D. Specht, G. Damasco, C. C. Ribas, P. V. A. Fine, and A. Vicentini. 2018. Towards integrative taxonomy in Neotropical botany: disentangling the Pagamea guianensis species complex (Rubiaceae). Botanical Journal of the Linnean Society **188**:213-231.
Guillot G, Santos F. 2009. A computer program to simulate multilocus genotype data with spatially autocorrelated allele frequencies. Molecular Ecology Resources 9: 1112-1120.
Richard, K. R., T. S. Harlan, T. Barbara, L. A. Prather, A. M. James, P. L. Laura, C. M. Lucas, and J. F. Carolyn. 2019. Herbarium Practices and Ethics, III. Systematic Botany **44**:7-13.
Sapes, G., L. Schroeder, A. Scott, I. Clark, J. Juzwik, R. A. Montgomery, J. A. Guzmán Q, and J. Cavender-Bares. 2024. Mechanistic links between physiology and spectral reflectance enable previsual detection of oak wilt and drought stress. Proceedings of the National Academy of Sciences **121**:e2316164121.
Serbin, S. P., A. Singh, B. E. McNeil, C. C. Kingdon, and P. A. Townsend. 2014a. Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. Ecological Applications **24**:1651-1669.
Serbin, S. P., A. Singh, B. E. McNeil, C. C. Kingdon, and P. A. Townsend. 2014b. Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. Ecological Applications **7**:1651-1669.
Serbin, S. P., and P. A. Townsend. 2020. Scaling functional traits from leaves to canopies.*in* J. Cavender Bares, G. JA, and P. A. Townsend, editors. Remote Sensing of Plant Biodiversity. Springer, New York.
Shah, T. 2023. The Herbarium Handbook. Kew Publishing.
Shipley, B., M. Belluau, I. Kuhn, N. A. Soudzilovskaia, M. Bahn, J. Penuelas, J. Kattge, L. Sack, J. Cavender-Bares, W. A. Ozinga, B. Blonder, P. M. van Bodegom, P. Manning, T. Hickler, E. Sosinski, V. D. Pillar, V. Onipchenko, and P. Poschlod. 2017. Predicting habitat affinities of plant species using commonly measured functional traits. Journal of Vegetation Science **28**:1082-1095.
Sims, D. A., and J. A. Gamon. 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. Remote Sensing of Environment **81**:337-354.

Soltis, P. S. 2017. Digitization of herbaria enables novel research. Am J Bot **104**:1281-1284.
Sousa da Silva, G., and M. J. G. Hopkins. 2024. Utilização da espectroscopia do infravermelho próximo (NIRS) como ferramenta na discriminação de espécies do gênero Dimorphandra Schott (Leguminosae: Caesalpinioideae). Scientia Plena **20**.
Stasinski, L., D. M. White, P. R. Nelson, R. H. Ree, and J. E. Meireles. 2021. Reading light: leaf spectra capture fine-scale diversity of closely related, hybridizing arctic shrubs. New Phytologist **n/a**.
Stefanski, A., E. E. Butler, L. J. Williams, R. Bermudez, J. A. Guzmán Q., A. Larson, P. A. Townsend, R. A. Montgomery, J. Cavender-Bares, and P. Reich. 2025. All the light we cannot see: Climate manipulations leave short and long-term imprints in spectral reflectance of trees. Ecology **in press**.
Thiers, B. M. 2024. Strengthening Partnerships to Safeguard the Future of Herbaria. Diversity.
Türker-Kaya, S., and C. W. Huck. 2017. A Review of Mid-Infrared and Near-Infrared Imaging: Principles, Concepts and Applications in Plant Tissue Analysis. Molecules.
Ustin, S. L., A. A. Gitelson, S. Jacquemoud, M. Schaepman, G. P. Asner, J. A. Gamon, and P. Zarco-Tejada. 2009. Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. Remote Sensing of Environment **113**:S67-S77.
Ustin, S. L., and S. Jacquemoud. 2020. How the Optical Properties of Leaves Modify the Absorption and Scattering of Energy and Enhance Leaf Functionality. Pages 349-384 *in* J. Cavender-Bares, J. A. Gamon, and P. A. Townsend, editors. Remote Sensing of Plant Biodiversity. Springer International Publishing, Cham.
Ustin, S. L., S. Jacquemoud, and Y. Govaerts. 2001. Simulation of photon transport in a three-dimensional leaf: implications for photosynthesis. Plant, Cell & Environment **24**:1095-1103.
Ustin, S. L., D. A. Roberts, J. A. Gamon, G. P. Asner, and R. O. Green. 2004. Using Imaging Spectroscopy to Study Ecosystem Processes and Properties. BioScience **54**:523-534.
Vasconcelos, C. C., M. U. Adrianzén, J. L. C. Camargo, and M. H. Terra-Araujo. 2020. Pouteria kossmanniae (Sapotaceae): a new species from Central Amazonia, Brazil. Phytotaxa **447** 265–275.
Violle, C., M.-L. Navas, D. Vile, E. Kazakou, C. Fortunel, I. Hummel, and E. Garnier. 2007. Let the concept of trait be functional! Oikos **116**:882-892.
Wagner, E. P., J. Merz, and P. A. Townsend. 2019. EcoSIS: A Spectral Library and the Tools to Use It. American Geophysical Union, Fall Meeting 2019, abstract #B11F-2396 **2019AGUFM.B11F2396W**.
Wang, R., J. A. Gamon, and J. Cavender-Bares. 2022. Seasonal patterns of spectral diversity at leaf and canopy scales in the Cedar Creek prairie biodiversity experiment. Remote Sensing of Environment **280**:113169.
Wang, Z., A. Chlus, R. Geygan, Z. Ye, T. Zheng, A. Singh, J. J. Couture, J. Cavender-Bares, E. L. Kruger, and P. A. Townsend. 2020. Foliar functional traits from imaging spectroscopy across biomes in eastern North America. New Phytologist **228**:494-511.
Wang, Z., J.-B. Féret, N. Liu, Z. Sun, L. Yang, S. Geng, H. Zhang, A. Chlus, E. L. Kruger, and P. A. Townsend. 2023. Generality of leaf spectroscopic models for predicting key foliar functional traits across continents: A comparison between physically- and empirically-based approaches. Remote Sensing of Environment **293**:113614.
Webster, M. S. 2017. The Extended Specimen: Emerging Frontiers in Collections-Based Ornithological Research (1st ed.). CRC Press.
Wenk, E. H., H. Sauquet, R. V. Gallagher, R. Brownlee, C. Boettiger, D. Coleman, S. Yang, T. Auld, R. Barrett, T. Brodribb, B. Choat, L. Dun, D. Ellsworth, C. Gosper, L. Guja, G. J. Jordan, T. Le Breton, A. Leigh, P. Lu-Irving, B. Medlyn, R. Nolan, M. Ooi, K. D. Sommerville, P. Vesk, M. White, I. J. Wright, and D. S. Falster. 2024. The AusTraits plant dictionary. Scientific Data **11**:537.

White, D. M., J. Cavender-Bares, C. Davis, J. A. G. Q., S. Kothari, J. Robles, and J. E. Meireles. 2025. Seeing herbaria in a new light: leaf reflectance spectroscopy unlocks predictive trait and classification modeling in plant biodiversity collections. EcoEvoRxiv:1-41.

Wieczorek, J., D. Bloom, R. Guralnick, S. Blum, M. Döring, R. Giovanni, T. Robertson, and D. Vieglais. 2012. Darwin Core: an evolving community-developed biodiversity data standard. Plos One **7**:e29715.

Williams, L. J., J. Cavender-Bares, P. A. Townsend, J. J. Couture, Z. Wang, A. Stefanski, C. Messier, and P. B. Reich. 2021. Remote spectral detection of biodiversity effects on forest biomass. Nature Ecology & Evolution **5**:46-54.

Willis, C. G., E. R. Ellwood, R. B. Primack, C. C. Davis, K. D. Pearson, A. S. Gallinat, J. M. Yost, G. Nelson, S. J. Mazer, N. L. Rossington, T. H. Sparks, and P. S. Soltis. 2017. Old Plants, New Tricks: Phenological Research Using Herbarium Specimens. Trends in Ecology & Evolution **32**:531-546.

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

Wright, I. J., P. B. Reich, M. Westoby, D. D. Ackerly, Z. Baruch, F. Bongers, J. Cavender-Bares, T. Chapin, J. H. Cornelissen, and M. Diemer. 2004. The worldwide leaf economics spectrum. Nature **428**:821-827.

Zhang, J., M. Wang, K. Yang, Y. Li, Y. Li, B. Wu, and Q. Han. 2022. The New Hyperspectral Analysis Method for Distinguishing the Types of Heavy Metal Copper and Lead Pollution Elements. Int J Environ Res Public Health **19**.

Supplemental Tables

Table S1. Previous studies using dry-leaf spectra.

Application (dry leaves)	**Spectral Range**	**Author**
C, N, P, K, micronutrients, leaf pigments, d13C (see Fig. 1)	VIS-NIR: 400-2500 nm	Elvidge 1990, Fourty et al. 1996, Serbin et al. 2014, Prananto et al. 2020, Chlus and Townsend 2022, Chen et al. 2002, Kothari et al. 2023, 2024, Wang et al. 2023, Kühn et al. 2025
Nutrient contents and leaf traits	NIR: 1000-2500 nm	Prananto et al. 2021, Costa et al. 2018, Kothari et al. 2024
Species identification	VIS-NIR: 400-2,500 nm	Meireles et al. 2020; Vasconcelos et ali. 2025
Species identification	NIR: 1.000-2.500nm	Durgante et al. 2013; Lang et al. 2015;Paive et al. 2021; Vasconcelos et al. 2020, 2021
Species identification	IR: 1.250-25.000nm	Kim et al. 2004; Krajsek et al. 2008
Species characterization and chemical properties	Thermal IR: 2.5 - 20 μm	Elvidge 1990; Richardson et. al. 2000