

# IHerbSpec

International Herbarium Spectral Digitization Working Group

---

# IHerbSpec Protocol

**Version:** 1.0

## Authors

Dawson M. White<sup>1,\*</sup>, Natalie I. Ahlstrand<sup>2</sup>, Matthew W. Austin<sup>3</sup>, Denis Bastianelli<sup>4</sup>, Samantha Bazan<sup>4</sup>, Khalil Boughalmi<sup>5</sup>, Warren Cardinal-McTeague<sup>6</sup>, Thomas L.P. Couvreur<sup>5</sup>, Flavia Durgante<sup>7,8</sup>, Olwen M. Grace<sup>9</sup>, J. Antonio Guzmán Q.<sup>10</sup>, Kimberly Hansen<sup>11</sup>, Michael Hopkins<sup>8</sup>, Rykkar Jackson<sup>6</sup>, Jonathan A. Kennedy<sup>1</sup>, Shan Kothari<sup>12</sup>, Aaron K. Lee<sup>13</sup>, Étienne Léveillé-Bourret<sup>14</sup>, José Eduardo Meireles<sup>15</sup>, Barbara M. Neto-Bradley<sup>16</sup>, Cornelius Onyedikachi Nichodemus<sup>15</sup>, Natalia L. Quinteros Casaverde<sup>17,18</sup>, Carlos Rodrigues-Vaz<sup>5,19</sup>, Michaela Schmull<sup>1</sup>, Pamela S. Soltis<sup>20</sup>, Daniel Spalink<sup>21</sup>, Caroline C. Vasconcelos<sup>8</sup>, Jéssica Lira Viana<sup>22</sup>, Hanna Tuomisto<sup>22,23</sup>, Tom Wells<sup>24</sup>, Jeannine Cavender-Bares<sup>1,\*</sup>

## Affiliations

<sup>1</sup>Harvard University Herbaria and Department of Organismic and Evolutionary Biology, Harvard University, Cambridge MA 02138, USA

<sup>2</sup>Natural History Museum of Denmark, University of Copenhagen, Copenhagen, 1154, Denmark

<sup>3</sup>Herbarium, Missouri Botanical Garden, 4344 Shaw Blvd, St. Louis MO 63110

<sup>4</sup>CIRAD, Umr Selmet, Montpellier, France, UMR SELMET, F-34398 Montpellier, France.

<sup>5</sup>DIADE, University of Montpellier, CIRAD, IRD, Montpellier, France

<sup>6</sup>Department of Forest and Conservation Sciences, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada

<sup>7</sup>Department of Wetlands Ecology, Karlsruhe Institute of Technology, KIT, Germany.

<sup>8</sup>Herbário, National Institute of Amazonian Research, INPA, Brazil.

<sup>9</sup>Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, United Kingdom

<sup>10</sup>Department of Organismic and Evolutionary Biology, Harvard University, Cambridge MA 02138, USA

<sup>11</sup>Field Museum of Natural History, 1400 S. DuSable Lake Shore Dr., Chicago, IL 60625, USA

<sup>12</sup>Department of Renewable Resources, University of Alberta, Edmonton, AB T6G 2E3, Canada

<sup>13</sup>Department of Plant and Microbial Biology, University of Minnesota, 1479 Gortner Ave, Saint Paul, MN 55108, USA

<sup>14</sup>Université de Montréal and Marie-Victorin Herbarium (MT), Département de Sciences Biologiques, Institut de Recherche en Biologie Végétale (IRBV), Université de Montréal, Montréal, QC H1X 2B2, Canada

<sup>15</sup>School of Biology and Ecology, University of Maine, 5793 Hitchner Hall, Orono, ME 04469, USA

<sup>16</sup>Department of Plant Sciences and Conservation Research Institute, University of Cambridge, Cambridge, UK

<sup>17</sup>Biospheric Science Laboratory, NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771

<sup>18</sup>Southeastern Universities Research Association, Inc, 1201 New York Ave NW, Ste 430, Washington, DC 20005, US

<sup>19</sup>Institut de Systématique, Evolution, Biodiversité (ISYEB), Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, C.P. 39, 57 rue Cuvier, 75005, Paris, France

<sup>20</sup>Florida Museum of Natural History and University of Florida Biodiversity Institute, Gainesville, Florida, USA

<sup>21</sup>S.M. Tracy Herbarium (TAES), Texas A&M University, Texas, USA

<sup>22</sup>Department of Biology, Section of Ecoinformatics and Biodiversity, Aarhus University, Aarhus, Denmark

<sup>23</sup>Department of Biology, University of Turku, Turku, Finland

<sup>24</sup>Royal Botanic Gardens, Kew, UK

**\*Authors for correspondence:** Jeannine Cavender-Bares, jcavender@fas.harvard.edu; Dawson M. White, dawson.white@gmail.com

## Suggested Citation

White DM, NI Ahlstrand, MW Austin, D Bastianelli, S Bazan, K Boughalmi, W Cardinal-McTeague, TLP Couvreur, F Durgante, OM Grace, JA Guzmán Q, K Hansen, M Hopkins, R Jackson, JA Kennedy, S Kothari, AK Lee, É Léveillé-Bourret, JE Meireles, BM Neto-Bradley, CO Nichodemus, NL Quinteros Casaverde, C Rodrigues-Vaz, M Schmull, PS Soltis, D Spalink, CC Vasconcelos, JL Viana, H Tuomisto, T Wells, J Cavender-Bares. 2025. *IHerbSpec Protocol v1.0*. <https://doi.org/10.5281/zenodo.1584966>

## Acknowledgements

We gratefully acknowledge the herbaria, collection institutions, curators, collection managers, and staff who steward biodiversity collections and make them accessible for scientific discovery. Funding: US NSF Biology Integration Institute ASCEND, Advancing Spectral biology in Changing Environments to understand plant Diversity (DBI: 2021898), the Harvard University Herbaria, iDigBio (NSF DBI-2027654) to PSS and NSF grant DBI-1930007 to PSS and Douglas E. Soltis, Canadian NSERC Discovery Grant, support from the Missouri Botanical Garden's Revolutionizing Species Identification project. NSF Graduate Research Fellowship DGE-2237827 to AKL. 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). Digitisation at the Royal Botanic Garden Edinburgh is supported by the Botanics Foundation and American Friends of the Botanic Garden. NYBG Visiting Scientist Travel Award to CON. FD, MH, and CCV are funded by the Amazonas State Research Support Foundation (FAPEAM) 'Women in Science Program' No. 006/2024 (SPECTRA POP project, process number 01.02.016301.04984/2024-17) and the Federal Ministry of Research, Technology, and Space (BMBF) grant 01LK2101D. JLV is supported by Danish National Research Foundation grant DNRF179 to HT. NLQC is supported by Southeastern Universities Research Association award number 80NSSC22M0001. We thank Jorge M. Robles for his artistic contributions to Figs. 2.1 and 6.1.

## Author Contributions

All authors contributed intellectually to the development of the protocol, including reviewing, editing, and approving the final version. DMW organized and led protocol development and the writing process.

**Intellectual leadership:** Led by JCB with contributions from MWA, FD, JAK, JEM, JAGQ, MS, and DMW.

**Draft writing:** NIA, MWA, JCB, ELB, TLPC, FD, OMG, JAGQ, JAK, SK, DS, NLQC, DMW

**Figures and tables:** SB, FD, JAK, DMW

**Appendix II examples:** KB, SB, FD, AKL, CCV, CON, DMW

**Metadata standardization:** Led by JAK with contributions from MWA, JCB, ELB, FD, SK, JAGQ, PSS, DMW

**Validation:** MWA, CON, JLV, HT

## Contents

### Preface

- Abstract
- IHerbSpec and Protocol development

### Part 1: Overview

- 1.1 Purpose and structure of the IHerbSpec Protocol
- 1.2 Measurement and metadata overview

### Part 2: Measurement and Metadata Workflow

### Part 3: Filename Conventions and Formats

### Part 4: Metadata and Databasing

- 4.1 Metadata tables
- 4.2 Controlled vocabularies
- 4.2 Guidelines for Data Archiving and Sharing

### Part 5: Instrumentation and Materials Guidelines

- 5.1 Precautions
- 5.2 Spectra data files
- 5.3 Instrumentation and materials quality control
- 5.4 Specimen and tissue annotation

### Part 6: Selecting Tissues for Spectral Measurement

- 6.1 Sources of variation in herbarium plant tissues
- 6.2 General guidelines for tissue selection
- 6.3 Decision tree for tissue selection

### References

### Appendix I: Considerations Regarding the Number of Measurements per Tissue

### Appendix II: Representing Tissue Conditions in Metadata for Quality Control

## List of Tables

- Table 3.1: Recommended Filename Conventions
- Table 3.2: Filename Codes and Corresponding Metadata Fields
- Table 4.1: Session Metadata
- Table 4.2: Specimen Metadata
- Table 4.3: Tissue Metadata
- Table 4.4: Developmental Stage Class Codes
- Table 4.5: Target Class Codes
- Table 4.6: Tissue Descriptor Codes
- Table 4.7: Background and White Reference Class Codes
- Table A1: Select Metadata Fields and Scoring for Representative Specimens

## List of Figures

- Figure 1.1: Influence of Background on Leaf Spectral Measurements
- Figure 2.1: Diagram of the Measurement and Metadata Workflow
- Figure 2.2: Bench Setup and Optical Probe Measurements
- Figure 4.1: Schematic of Session Metadata Fields Related to Optical Setup
- Figure 4.2: Diagram of Grid Coordinate System for Scoring `tissueLocation`
- Figure 6.1: Decision Tree Diagram for Target Tissue Selection
- Figure A1: Variability in Spectral Measurements
- Figure A2: Accuracy of Taxonomic Discrimination as a Function of the Number of Averaged Measurements per Specimen
- Figure A3: Spectral Variability as a Function of the Number of Measurements Averaged per Specimen

## Preface

### Abstract

Reflectance spectroscopy is a powerful, broadly integrative tool for capturing plant phenotypes, but variation in instrumentation and measurement procedures introduces a high risk of data incompatibility and signal contamination. These challenges are amplified in herbarium specimens because they are subject to complex variation from age, preservation techniques, and mounting practices. To support global data harmonization and scientific reproducibility, this document presents a standardized, interoperable protocol for the spectral measurement of herbarium tissues. It provides data-driven justification and community-informed guidance for measurement procedures and metadata fields, with the goal of improving data quality and enabling confident data aggregation across institutions.

### IHerbSpec and Protocol development

The International Herbarium Spectral Digitization (IHerbSpec) Working Group was established in December 2024 as a global consortium focused on advancing the use of reflectance spectroscopy of herbarium specimens for ecological and evolutionary studies. The group actively collaborates to identify common challenges, facilitate collaborations, and develop best practices.

This protocol is the outcome of IHerbSpec's initial collaborative phase. From December 2024 to June 2025, members convened virtually to share project-specific workflows and align measurement strategies. The effort culminated in an in-person workshop held at the Harvard University Herbaria from May 1–3, 2025, where members reached consensus on key elements of the protocol, including:

1. The minimum and recommended number of measurements for leaf tissues (see [Section 1.2](#): Strategy and number of tissue measurements).
2. Tissue condition and contamination metadata (see [Section 1.2](#): Scoring metadata pertaining to tissue condition and contamination).
3. Tissue development stage metadata (see [Table 4.4](#)).

IHerbSpec will continue to refine and expand this protocol and welcomes inquiries from interested researchers or institutions wishing to ask questions, provide feedback, or explore opportunities for participation.

## Part 1: Overview

### 1.1 Purpose and structure of the IHerbSpec Protocol

This document outlines a “base” protocol for the spectroscopic measurement of herbarium plant tissues by defining clear minimum requirements and additional recommended practices that promote data quality, methodological consistency, reproducibility, and interoperability across projects and institutions. It is intended to support researchers collecting reflectance spectra for ecological and evolutionary analyses (Cavender-Bares et al. 2025).

The protocol specifies two categories of procedural elements:

- **Minimum requirements** are essential components that must be met for a project to be considered aligned with the IHerbSpec Protocol.
- **Recommended practices** are proposed additions that enhance data quality and support broader utility but are not mandatory for protocol alignment.

The IHerbSpec Protocol is organized into modular parts. Core procedures are presented in the overview ([Part 1](#)), the measurement and metadata entry workflow ([Part 2](#)), filename conventions and formats ([Part 3](#)), and metadata standards ([Part 4](#)). Additional guidance—especially for those new to spectroscopy or herbarium workflows—is provided in the instrumentation and materials guidelines ([Part 5](#)) and tissue selection guidance ([Part 6](#)). Supporting appendices provide data-driven justification for protocol decisions regarding the number of measurements per specimen ([Appendix I](#)) and guidelines and examples to assist with tissue metadata scoring ([Appendix II](#)).

While consistency is essential for integrating spectral datasets, the protocol is not intended to prescribe a rigid workflow. It is designed to accommodate a variety of project goals, instrumentation, and institutional capacities—balancing standardization with the flexibility needed for future refinement and innovation.

### 1.2 Measurement and metadata overview

This section outlines the required and recommended components (see [Section 1.1](#)) of the IHerbSpec Protocol, including filename conventions, reference and background measurements, tissue-level measurement strategies, and metadata documentation. Notes on adaptability are included where relevant.

#### Filename conventions

- To ensure key identifiability components for every unprocessed spectra file (white target, black target, calibrated reflectance standards, tissues), the IHerbSpec Protocol strongly **recommends** projects follow the filename formats for local file storage and data sharing that are described in [Part 3](#).

**Adaptability:** Projects can use the simplified filename provided in [Table 3.1](#) or similar during metadata scoring, then convert to the full filename format for permanent local storage and distribution.

## White reference and background measurements

- **White-reference** measurements are **required** at the start of the measurement session and at regular intervals during the session (e.g., 20 minutes; check with manufacturer).
  - This reference measurement is not recorded to a unique file (see [Section 5.3](#)).
  - A **measurement session** is defined as the duration between switching on the instrument to switching it off.
- **White, black, and/or gray calibrated reflectance standard** target measurements are **recommended** to be taken at the beginning of each measurement session or at regular intervals (e.g. once per week) and be locally archived (see [Section 5.3](#)).
- A single **white target** measurement is **required** at the beginning of each measurement session to confirm the performance of the instrument (see [Fig. 1.1](#)).
  - This measurement is linked to every tissue measurement in the session by the `sessionId` filename and metadata component ([Table 4.3](#)).
- The protocol **requires** projects to use a **<4% reflective black surface whenever possible** as a background behind tissues during measurement (see [Section 5.3](#)).
- A single **black-background** target measurement is **required** at the beginning of each measurement session to record its reflectance spectrum ([Fig. 1.1](#)).
  - This measurement is linked to every tissue measurement in the session by the `sessionId` metadata field ([Table 4.3](#)).
- If tissues are glued and must be measured with **herbarium sheet paper** as the background, then a target measurement of the paper is **required** and should be linked to the corresponding tissue spectra (see [Part 3](#)).

## Strategy and number of tissue measurements

- **Tissue selection strategy**
  - Considerations and guidelines for selecting specimens and tissues for measurement are provided in [Part 6](#), including a decision tree diagram ([Fig. 6.1](#)).
- **The number of leaf tissue measurements per specimen**
  - If both adaxial and abaxial leaf surfaces are accessible and suitable, a **minimum of three representative measurements per surface (six total) is required** per specimen. If only one surface is available or if the tissue class is **leaf** (see [Table 4.5](#)), then a minimum of three representative measurements are required.
    - **Representative measurements should be free of technical errors** and generally reproducible (see [Section 6.2](#); [Appendix I](#)).

- Five or more representative measurements per available and suitable leaf tissue class (`adaxial`, `abaxial`, or `leaf`) are **recommended** (see [Appendix I](#) for discussion).

**Adaptability:** Projects may choose to collect measurements from a single leaf or from multiple leaves to capture tissue- versus specimen-level variability. The different tissue units are **recommended** to be distinguished using the `targetTissueId` field and be identifiable via annotation labels and/or their location on the specimen (see [Table 4.3](#)).

- **Other tissue classes**

- No required minimum is prescribed.
- **Recommended:** At least three representative measurements per specimen per tissue class, if feasible.

### Measurement technique

- Multiple measurements per tissue unit may be taken by targeting multiple suitable areas on the tissue or—taking care not to heat-damage the tissue—by targeting a small area and slightly rotating or shifting the optical probe between measurements.
- Avoid overlapping tissue layers, debris, or partial coverage of the probe aperture ([Fig. A1](#)).
- If a measurement appears affected by technical error, delete and repeat it.

### Metadata requirements and flexibility

- All required and recommended metadata are available for use in the IHerbSpec Metadata Spreadsheet, which can be accessed at <https://iherbspec.github.io/protocol>.
- The protocol recommends recording tissue metadata for every individual measurement. This means that every tissue measurement will be recorded in a unique row in the metadata spreadsheet (e.g., 10 measurements have 10 rows).
- All fields in Session metadata ([Table 4.1](#)), Specimen Metadata ([Table 4.2](#)), and Tissue Metadata ([Table 4.3](#)) are marked as **required** or **recommended** in the field descriptions.

### Scoring metadata pertaining to tissue condition and contamination

- To record key elements of tissue condition and contamination, these fields are **required metadata** (see [6.1 Sources of Variation](#); see field descriptions in [Table 4.3](#)):
  - `tissueDevelopmentalStage`
  - `hasGlue`
  - `hasNonGlueContamination`
- To further describe tissue conditions, the `measurementFlags` and `tissueNotes` fields are **recommended**:
  - `measurementFlags` allows use of standardized tissue condition descriptors (see [Table 4.6](#)) for preservation status, damage, or contamination (e.g., ‘GoodPreservation|GluePresent’).

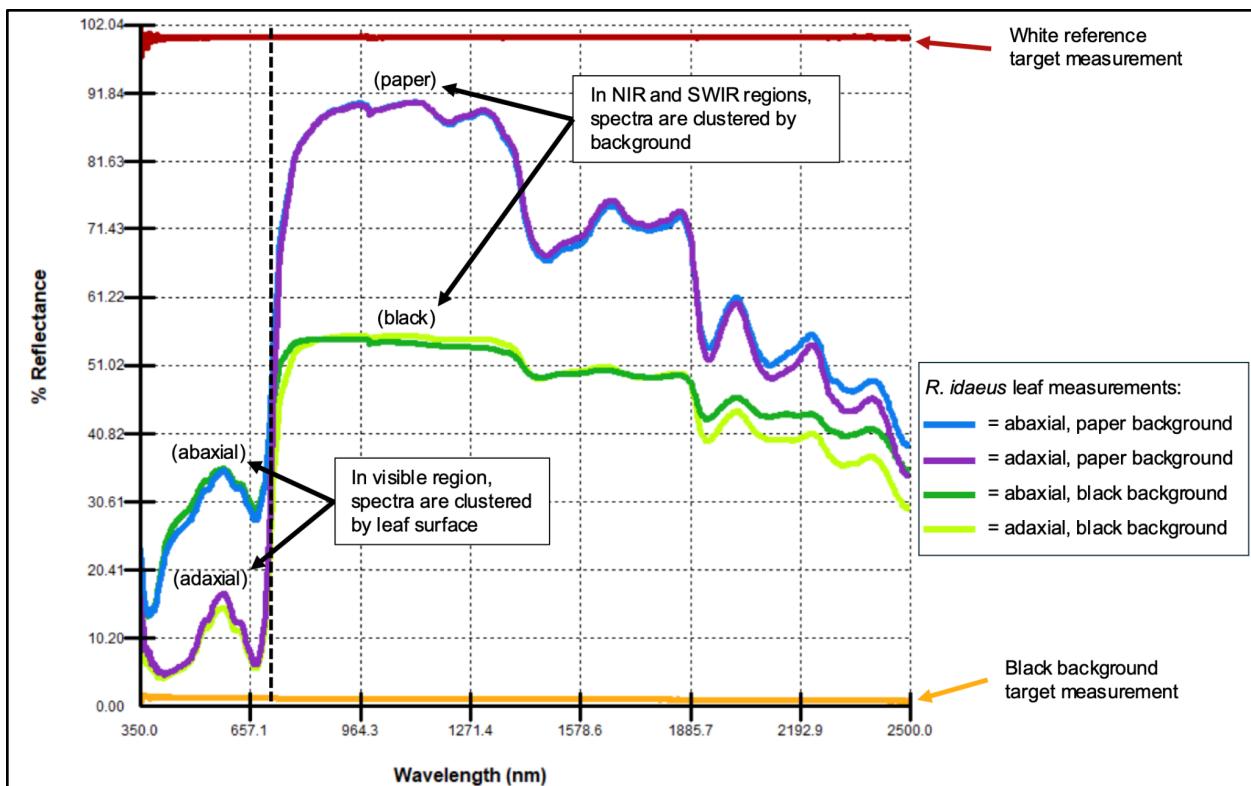
- `tissueNotes` enables free-text documentation of additional tissue conditions or context inside or outside the measurement area.

**Adaptability:**

- `tissueDevelopmentalStage` may be coded as ‘notScored’ if assessment is not possible (e.g., due to technician training or specimen condition). See [Table 4.4](#).
- `hasGlue` and `hasNonGlueContamination` may be marked ‘uncertain’ if presence is not confidently determined.

**Specimen and tissue annotation:**

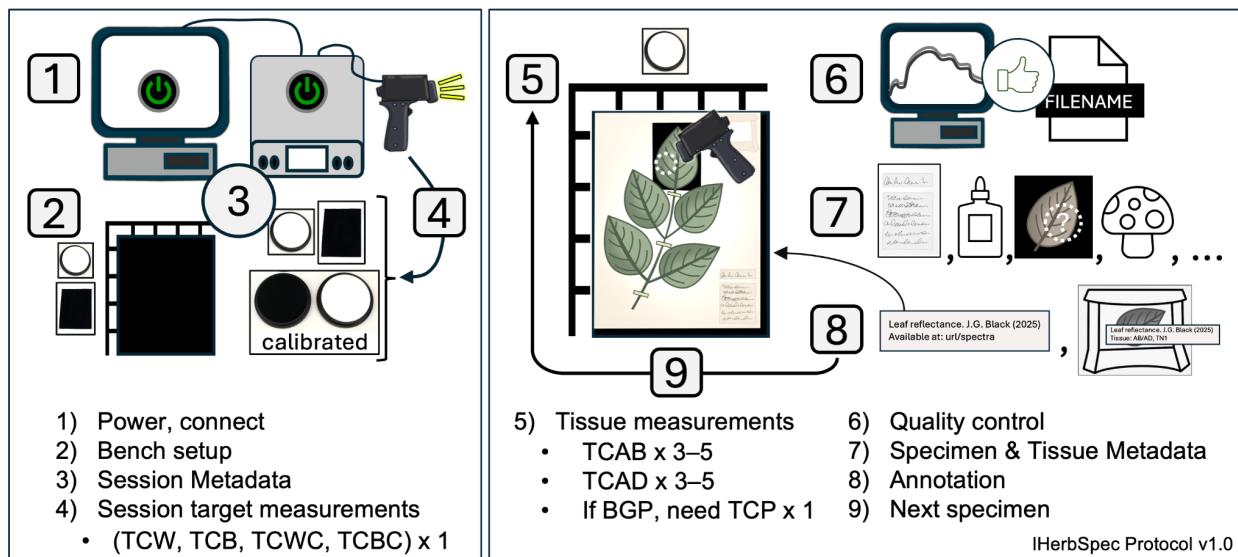
- The IHerbSpec Protocol **recommends** annotating the location of measured tissues on the sheet using the `targetTissueId` field. Additional annotations to the herbarium sheet or packet are encouraged in accordance with herbarium policies (see [Section 5.4](#)).



**Fig. 1.1:** Influence of background on leaf spectral measurements. For a single *Rubus idaeus* loose leaf (see [NEBC 02618198](#) in [Appendix II](#)), the green adaxial surface and strongly glaucous abaxial leaf surface influence strong effects on the spectrum in the visible region (400–700 nm) regardless of paper or black backgrounds. However, the spectra in the near-infrared (700–1,100 nm) and shortwave infrared (1,100–2,500 nm) regions are instead clustered by the background material, revealing both biological and background contamination effects on reflectance spectra. The plot shows target measurements of the white reference, black background, and adaxial and abaxial leaf surfaces over black and paper (herbarium sheet) backgrounds. All measurements were made with an SVC HR-1024i spectroradiometer with the LC-RP Pro leaf clip.

## Part 2: Measurement and Metadata Workflow

This part outlines the recommended step-by-step workflow for spectral data collection and metadata entry. Steps include references to protocol sections providing further details.



**Fig. 2.1:** Diagram of IHerbSpec Protocol measurement and metadata workflow. Steps 1–4 describe the setup procedure and steps 5–9 describe the measurement sequence. For decisions regarding specimen tissue selection and using black backgrounds, see [Part 6](#).

### Step 1) Instrument Setup and Connection

1. Plug in and turn on the instrument and light source.
  - Allow >15 minutes or follow manufacturer standard operating procedures (SOPs) for lamp warm-up and sensor cool-down.
2. Set up the computer and software, and connect the instrument following project SOPs.

### Step 2) Prepare bench and specimens

1. Select specimens and tissues for measurement (see [Part 6](#); [Fig. 6.1](#)).
2. Optionally, prepare filenames ([Part 3](#)) in a separate document to allow copy-paste entry during measurement.
3. Prepare bench with the following materials (see [Fig 2.2](#)):
  - **Required:** white reference, black background
  - **Recommended:** Benchtop black background, rulers for tissue coordinates, tweezers, and laboratory gloves (see [Part 5](#)). Annotation labels and archival envelopes (see [Section 5.4](#)).

### Step 3) Start session and score Session Metadata

1. Score Session Metadata (see [Table 4.1](#)):

- Record `sessionId` datetime in the format YYYYMMDDHHMM (e.g. 202507011351, see [Table 4.1](#)).
  - Score **required** fields: `projectId`, `sessionId`, `instrumentModel`, `opticalSetupDescription`, `measurementSettings`, `whiteReferenceDescription`
  - Score **recommended** fields: `operator`, `lightSourceType`, `distanceTargetToSensor`, `lensFieldOfView`, `angleLightToSensor`, `measurementAreaDiameter`
2. Create a storage folder for spectral data named with the `sessionId` (e.g., 202507011351), and set it as the destination folder for saving measurement files.

#### Step 4) Collect white and black session measurements

1. Take a white reference.
2. Set simple filename for white target (see [Table 3.1](#); `IDX` will be appended automatically):
  - `PI<projectId>_SN<sessionId>_TCW` (e.g. PIHUhCoca\_SN202507011351\_TCW)
3. Take one white target measurement.
4. Set simple filename for black target (see [Table 3.1](#)):
  - `PI<projectId>_SN<sessionId>_TCB` (e.g. PIHUhCoca\_SN202507011351\_TCB)
5. Take one black target measurement.
6. Optionally, take one target measurements of each calibrated reflectance standard (see [Section 5.3](#)), using filename conventions ([Table 3.1](#)):
  - `PI<projectId>_SN<sessionId>_TC<targetClass><serial number>` (e.g. PIHUhCoca\_SN202507011351\_TCCW379902)

#### Step 5) Tissue measurement sequence

1. Take white reference at regular intervals (e.g., 20 min.; see [Section 1.2](#)).
2. Determine if tissue is suitable for measurement (see [Part 6](#)). If not, skip to the next specimen.
3. Place specimen on benchtop black background with a ruler grid (**recommended**; see [Fig. 4.2](#)).
4. Place black background behind target tissue (**required** whenever possible; see [Fig. 2.2](#)).
5. Adaxial (AD) surface available:
  - a. Set filename: `SI<specimenId>_TC<targetClass>_TN<targetTissueId>`
    - i. e.g., SI02022418\_TCAD\_TN1
  - b. Take 3–5 adaxial measurements.
6. Abaxial (AB) surface available:
  - a. Set filename: `SI<specimenId>_TC<targetClass>_TN<targetTissueId>`
    - i. e.g., SI02022418\_TCAB\_TN1
  - b. Take 3–5 abaxial measurements.
7. If tissue was measured on herbarium paper:
  - a. set filename: `SI<specimenId>_TC<targetClass>_TN<targetTissueId>`

- i. e.g., `SI00746092_TCP_TN1`
- b. Take 1 paper measurement.
8. Additional tissues (optional)
  - a. Select other tissue units with new `targetTissueIds` (**recommended**) and/or additional target classes (see [Table 4.5](#)) and repeat tissue measurement steps.

## Step 6) Quality Control

1. Visually review each spectrum.
2. Delete and repeat any measurement that appears anomalous.
3. If unsure, take additional measurements.
4. Check for and fix any filename errors before proceeding.

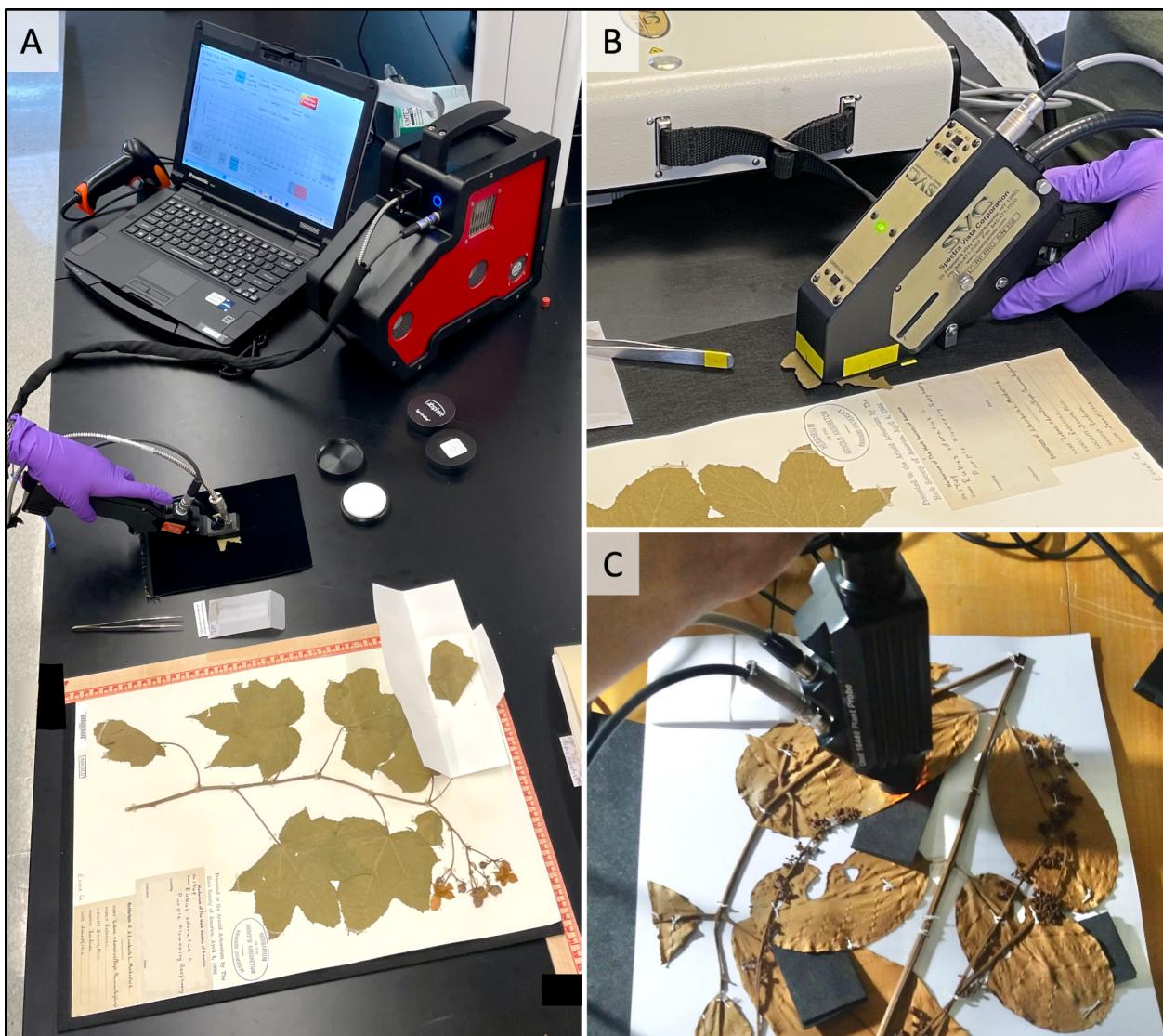
## Step 7) Score Specimen and Tissue Metadata

1. Score Specimen Metadata for every measurement (see [Table 4.2](#)):
  - o Score **required** fields: `herbariumCode`, `specimenId`
  - o Score **recommended** fields: `scientificName`, `identificationQualifier`, `identifiedBy`, `dateIdentified`, `isTempControlled`, `annualTempMin`, `annualTempMax`, `isHumidityControlled`, `annualHumidityMin`, `annualHumidityMax`
2. Score Tissue Metadata for every measurement (see [Table 4.3](#)):
  - o Score **required** fields: `backgroundClass`, `hasLowReflectanceBackground`, `targetClass`, `tissueDevelopmentalStage`, `hasBackgroundInMeasurement`, `hasGlue`, `hasNonGlueContamination`, `measurementIndex`
  - o Score **recommended** fields: `backgroundDescription`, `targetTissueId`, `percentBackgroundInMeasurement`, `measurementFlags`, `tissueNotes`, `tissueLocation`, `comment`

## Step 8) Specimen and tissue annotation (see [Section 5.4](#))

1. Add project annotation label to sheet (recommended).
2. Annotate attached target tissues (unless already recorded in `targetLocation`; **recommended**).
3. Annotate loose tissues (in packets/envelopes) with `targetTissueClass` and `targetTissueId` labels and store them in an envelope on the sheet ([Fig. 4.2](#); **recommended**).

## Step 9) Move to next specimen and return to Step 5



**Fig. 2.2:** Reflectance spectroscopy bench setup and optical probe measurements. (A) Benchtop setup with Spectral Evolution NaturaSpec spectroradiometer with herbarium probe measuring a loose leaf fragment on black background. The setup also includes a barcode scanner, a laptop with data acquisition software and display for quality control and recording metadata, 2-inch diameter white and black calibrated reflectance standards, a 2-inch diameter Spectralon® white reference (top removed), a benchtop black background underneath the herbarium sheet, and a tissue annotation label stored in a glassine envelope inside the herbarium sheet packet. (B) Spectral Vista Corporation HR-1024i spectroradiometer and LC-RP Pro leaf clip (with clip removed) measuring a loose leaf fragment over a black background. (C) ASD optical probe measuring mounted tissue with black background slid underneath. Photos A and B by D.M. White; specimen from A: Herbarium of the Arnold Arboretum of Harvard University. Photo C by F. Durgante.

## Part 3: Filename Conventions

The full and simple filename conventions specified here ([Table 3.1](#)) are strongly **recommended** but not mandatory. These conventions provide a consistent format for encoding key metadata into filenames, supporting traceability, metadata parsing, and long-term interoperability.

Each filename follows a defined format, built from metadata-coded segments with standardized prefixes (see [Table 3.2](#)). These **segments correspond to required metadata fields** and are designed for reliable parsing, regardless of segment order. The spectrometer software will automatically append a measurement index ([IDX](#)) to the end of each filename—no additional text should be added after the [IDX](#).

Filename formats differ slightly depending on the type of target material (i.e., [targetClass](#); [Table 4.5](#)). For white and black-background targets, the [projectId](#) ([PI](#)) and [sessionId](#) ([SI](#)) are key segments for traceability and the [SI](#) is also critical for linking to all associated tissue measurements via the tissue full filename. For tissues, the full filename convention also includes the minimum session-level, specimen-level, and tissue-level metadata needed for confident data aggregation.

A simplified filename convention may be used during measurement sessions to streamline data collection. However, projects should **convert all filenames to the full format** before archiving or sharing. When using the simplified format for local files, projects should maintain consistent file organization and take precautions to prevent ambiguity.

**Table 3.1: Recommended Full and Simple Filename Conventions for Target Types.**

Target type ( <a href="#">targetClass</a> )	Filename Convention Type	Filename format and example
White target (TCW)	Full	<p><a href="#">PI&lt;projectId&gt;_SN&lt;sessionId&gt;_TC&lt;targetClass&gt;_&lt;IDX&gt;</a></p> <ul style="list-style-type: none"> <li>• PIERYspec1_SN202406180932_TCW_0001</li> </ul>
Black- background target (TCB)	Full	<p><a href="#">PI&lt;projectId&gt;_SN&lt;sessionId&gt;_TC&lt;targetClass&gt;_&lt;IDX&gt;</a></p> <ul style="list-style-type: none"> <li>• PIERYspec1_SN202406180932_TCB_0001</li> </ul>
White calibrated reflectance standard (TCWC)	Full	<p><a href="#">PI&lt;project&gt;_SN&lt;session&gt;_TC&lt;target&gt;&lt;serialNumber&gt;_&lt;IDX&gt;</a></p> <ul style="list-style-type: none"> <li>• PIFagaceae_SN202506171532_TCWC7254_0000</li> </ul>
Black	Full	<a href="#">PI&lt;project&gt;_SN&lt;session&gt;_TC&lt;target&gt;&lt;serialNumber&gt;_&lt;IDX&gt;</a>

calibrated reflectance standard (TCBC)		<ul style="list-style-type: none"> <li>PIFagaceae_SN202506171532_TCBC7210_0000</li> </ul>
Tissue target on black background (BGB + TCAD/TCAB)	Full	<p>PI&lt;projectId&gt;_SN&lt;sessionId&gt;_BG&lt;backgroundClass&gt;_HC&lt;herbariumCode&gt;_SI&lt;specimenId&gt;_TC&lt;targetClass&gt;_TN&lt;targetTissueId&gt;_&lt;IDX&gt;</p> <ul style="list-style-type: none"> <li>PIERYspec1_SN202406180932_BGB_HCGH_SI02022418_TCAD_TN1_001</li> <li>PIERYspec1_SN202406180932_BGB_HCGH_SI02022418_TCAB_TN1_001</li> </ul>
Tissue target on black background (BGB + TCAD/TCAB)	Simple	<p>SI&lt;specimenId&gt;_TC&lt;targetClass&gt;_TN&lt;targetTissueId&gt;_&lt;IDX&gt;</p> <ul style="list-style-type: none"> <li>SI02022418_TCAD_TN1_0001</li> <li>SI02022418_TCAB_TN1_0001</li> </ul>
Tissue target on paper (BGP + TCAD/TCAB)	Full	<p>PI&lt;projectId&gt;_SN&lt;sessionId&gt;_BG&lt;backgroundClass&gt;_HC&lt;herbariumCode&gt;_SI&lt;specimenId&gt;_TC&lt;targetClass&gt;_TN&lt;targetTissueId&gt;_&lt;IDX&gt;</p> <ul style="list-style-type: none"> <li>PIERYspec1_SN202406180932_BGP_HCNEBC_SI00746092_TCAD_TN1_001</li> <li>PIERYspec1_SN202406180932_BGP_HCNEBC_SI00746092_TCAB_TN2_001</li> </ul>
Tissue target on paper (BGP + TCAD/TCAB)	Simple	<p>SI&lt;specimenId&gt;_TC&lt;targetClass&gt;_TN&lt;targetTissueId&gt;_&lt;IDX&gt;</p> <ul style="list-style-type: none"> <li>SI00746092_TCAD_TN1_0001</li> <li>SI00746092_TCAB_TN2_0001</li> </ul>
Paper target (BGB + TCP)	Full	<p>PI&lt;projectId&gt;_SN&lt;sessionId&gt;_BG&lt;backgroundClass&gt;_HC&lt;herbariumCode&gt;_SI&lt;specimenId&gt;_TC&lt;targetClass&gt;_TN&lt;targetTissueId&gt;_&lt;IDX&gt;</p> <ul style="list-style-type: none"> <li>PIERYspec1_SN202406180932_BGB_HCNEBC_SI00746092_TCP_TN1_001</li> </ul>
Paper target (BGB + TCP)	Simple	<p>SI&lt;specimenId&gt;_TC&lt;targetClass&gt;_TN&lt;targetTissueId&gt;_&lt;IDX&gt;</p> <ul style="list-style-type: none"> <li>SI00746092_TCP_TN1_0001</li> </ul>

**Table 3.2: Filename Components and Corresponding Metadata Fields.** Each segment of the filename follows a specific format, with a standardized prefix that links directly to a **required** metadata field (see [Part 4](#)). These components form the structured **conventions** shown in [Table 3.1](#) and enable automatic parsing and alignment between spectral files and metadata records.

Code	Metadata field (see <a href="#">Table 4.3</a> )	Description (see <a href="#">Table 4.3</a> )	Example
PI	<code>projectId</code>	Unique identifier for the spectral measurement project.	PIHUHcoca, PIFagales1
SN	<code>sessionId</code>	A unique identifier generated from the date and time (24-hour format) when the measurement session begins, typically at equipment startup or the first measurement. Format: <code>SNYYYYMMDDHHMM</code>	SN202406180932
BG	<code>backgroundClass</code>	Enumerated code from Background Class Codes ( <a href="#">Table 4.7</a> ) describing the background class under the target tissue. This field will be omitted for white and background target measurements (e.g., <code>TCW</code> , <code>TCB</code> , <code>TCP</code> ).	BGB, BGP, BGO
HC	<code>herbariumCode</code>	Herbarium acronym or collection identifier.	HCGH, HCINPA
SI	<code>specimenId</code>	Specimen identification code: GUID, barcode (with preceding zeros), accession number, or collector name and number.	SI03774853, SIThorne24070
TC	<code>targetClass</code>	Enumerated code from Target Class Codes ( <a href="#">Table 4.5</a> ) describing type of material or tissue being measured.	TCAD, TCAB, TCW, TCP, TCB
TN	<code>targetTissueId</code>	Number or character index tracking the measured tissue units (first unit = 1, second = 2, etc.). This field will be omitted for white or background target measurements (e.g., <code>TCW</code> , <code>TCB</code> , <code>TCP</code> ).	TN1, TN2
IDX	<code>measurementIndex</code>	Measurement index number, automatically appended by the spectrometer software after the base filename. Unlike other components (e.g., <code>HC</code> , <code>SI</code> , <code>PI</code> ), <code>IDX</code> is not included as a prefix, but is automatically appended during file generation.	0001, 0002

## Part 4: Metadata and Databasing

This section defines the metadata fields used to describe spectral reflectance measurements of herbarium specimens. This standardization supports interoperability, cross-project data aggregation, and future integration into biodiversity informatics platforms.

Projects can use the IHerbSpec Metadata Spreadsheet (available at <https://iherbspec.github.io/protocol>) as a foundation for organizing their metadata; this spreadsheet includes all **required** and **optional but recommended** fields defined the tables in [Section 4.1](#). The controlled vocabularies necessary for enumerating certain fields are provided in [Section 4.2](#). Finally, [Section 4.3](#) provides general guidelines for data and metadata storage and dissemination.

While metadata are expected to be disseminated in a 'flat file' format, such as the IHerbSpec Metadata Spreadsheet, fields are presented here in logical groups—project, specimen, and tissue—to support conceptual understanding and integration with local databases.

### 4.1: Metadata tables

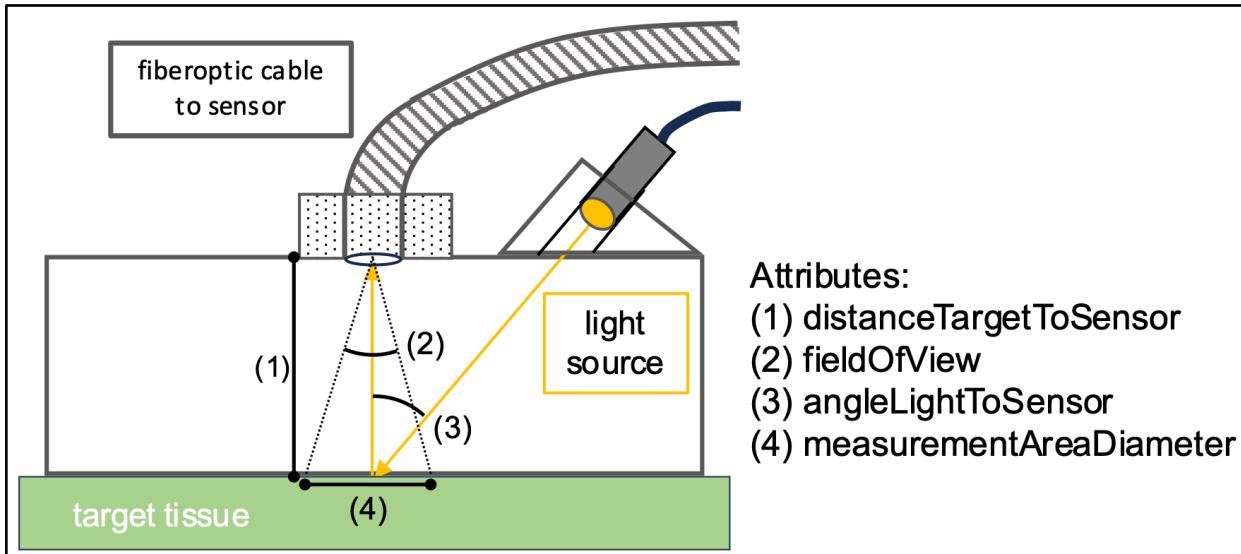
#### Session Metadata

Session metadata are those metadata usually associated globally with a continuous digitization project. These metadata can usually be captured once per project and automatically populated for each measurement instance.

**Table 4.1: Session Metadata**

Field Name	Status	Field Description	Data Type
projectId	Required	Unique identifier for the spectral measurement project. It is recommended that these identifiers be short and indicate something about location or institution.  Example: HUHERYspec1	TEXT
sessionId	Required	A unique identifier for a given measurement session, generated as the datetime when starting the equipment or conducting the first measurement. The session ends when the instrument is turned off. The datetime should use the format YYYYMMDDHHMM, with hour in 24-hour format.  Example: 20240617132251	TEXT
instrumentModel	Required	Spectroradiometer instrument model	TEXT

		name. Example: SVC HR-1024i	
opticalSetupDescription	Required	Description of optical probe including relevant instrument model.  Example: LC-RP contact probe with leaf clip removed	TEXT
measurementSettings	Required	Description of the instrument settings used for making spectral measurements, such as duration, light settings, number of measurements, etc.  Example: 2 seconds, high light setting	TEXT
whiteReferenceDescription	Required	Description of the material of the white reference.  Example: Spectralon SL Standard 99%	TEXT
operator	Optional	The name(s) of the person or people who conducted the spectral measurements during the session. Use full names or initials consistently.  Example: AB Cornejo Vargas	TEXT
lightSourceType	Optional	Light source for optical setup.  Example: tungsten halogen	TEXT
distanceTargetToSensor	Optional	Distance in millimeters between target tissue and the face of the fiber optic or sensor ( <a href="#">Fig. 4.1</a> ).  Example: 12	NUMERIC
lensFieldOfView	Optional	Angle in degrees of the field of view of the sensor lens or fiber optic cable ( <a href="#">Fig. 4.1</a> ).  Example: 22.5	NUMERIC
angleLightToSensor	Optional	Angle in degrees of the primary vector of the light source to the sensor ( <a href="#">Fig. 4.1</a> ).  Example: 10	NUMERIC
measurementAreaDiameter	Optional	The diameter in millimeters of the illuminated area of the target tissue that is within the field of view of the sensor or fiber optic and measured ( <a href="#">Fig. 4.1</a> ).  Example: 6	NUMERIC



**Fig. 4.1:** Schematic of Session Metadata fields related to optical setup (see [Table 4.1](#)):  
`distanceTargetToSensor`, `fieldOfView`, `angleLightToSensor`, `measurementAreaDiameter`.

## Specimen Metadata

Specimen metadata include identifiers and information about the physical specimen, with priority given to **required** fields needed to link spectral measurements to existing digital records (e.g., `specimenId`). Optional fields related to taxonomic determination and specimen storage environment are included to support integrative research, quality control, and downstream analysis.

Users should **avoid duplicating metadata that are already digitized, maintained, and available in herbarium or institutional platforms**, as these sources are better suited for future updates. Instead, users are encouraged to reference those records and supplement only missing required or recommended fields. Due to variation in the metadata recorded on institutional platforms, users should apply caution in the interpretation of presence or absence of determination information or any other recommended but optional fields.

**Table 4.2: Specimen Metadata**

Field Name	Status	Field Description	Data Type
<code>herbariumCode</code>	Required	The standardized acronym identifying the herbarium or collection where the specimen is deposited. This should correspond, whenever possible, to the <b>Index Herbariorum code</b> ( <a href="https://sweetgum.nybg.org/science/ih/">https://sweetgum.nybg.org/science/ih/</a> ), which is typically implemented as the institution code in GRSciColl. If referring to a collection within an institution code, collection identifiers may be appended using a dash “-”. Examples: 'GH', 'P', 'US-Botany'	TEXT
<code>specimenId</code>	Required	An identifier for the physical specimen or its digital record within the collection, such as a catalogue or accession number, barcode, or GUID. If no such number exists, enter the first collector name and number with no spaces between characters (for use as filename). Examples: '00238762', 'Thorne24070a'	TEXT
<code>scientificName</code>	Optional	The full scientific name of the specimen at the lowest taxonomic rank that can be confidently assigned (e.g., family, genus, species, subspecies, or variety), with or without author citation. It should not include qualifiers such as <i>cf.</i> or <i>aff.</i> , which should instead be recorded in the <code>identificationQualifier</code> field. Hybrid taxa should be indicated using the multiplication sign (×). If no taxonomic identification is available, the field should be left blank. Scientific names should conform to a recognized botanical authority or taxonomic reference to ensure consistency across	TEXT

		datasets (see <a href="https://worldfloraonline.org/">https://worldfloraonline.org/</a> ). Examples: 'Quercus bicolor', 'Erythroxylum coca ipadu', 'Agrostis stolonifera L.', 'Agavaceae'.	
identificationQualifier	Optional	A brief phrase or standard term indicating proximity or uncertainty of the taxonomic identification recorded in the <code>scientificName</code> field, following the Darwin core standard. If no qualifier is applicable, this field may be left blank.  Examples: 'cf.', 'confer', 'aff.', 'affinis', 'sp.'	TEXT
identifiedBy	Optional	The person, group, or organization that assigned the taxonomic information in the <code>scientificName</code> field. If no determination author is present, this field may be left blank.  Example: T. Plowman	TEXT
dateIdentified	Optional	The date on which the current taxonomic identification of the specimen was made . If no determination date is present, this field may be left blank.  Examples: '1999', '2004-12-30', '2010-06-09'	TEXT
isTempControlled	Optional	Room where the specimen is stored at home institution has active heating and air conditioning installations to maintain set temperatures.	BOOLEAN (true/false)
annualTempMin	Optional	The estimated minimum temperature in degrees celsius (°C) experienced in the specimen's storage environment over a typical year.  Examples: '18', '18°C'	TEXT
annualTempMax	Optional	The estimated maximum temperature in degrees celsius (°C) experienced in the specimen's storage environment over a typical year.  Examples: '26', '32°C'	TEXT
isHumidityControlled	Optional	Room where the specimen is stored at home institution has active humidity conditioning installations to maintain set humidity.	BOOLEAN (true/false)
annualHumidityMin	Optional	The estimated minimum relative humidity (%) experienced in the specimen's storage environment in a typical year.  Examples: '20', '50%'	TEXT
annualHumidityMax	Optional	The estimated maximum relative humidity (%)	TEXT

		experienced in the specimen's storage environment in a typical year. Examples: '60', '100%'	
--	--	--	--

## Tissue Metadata

Fields describing the type, condition, and position of the tissue measured. Includes **required** and **recommended** metadata for linking spectral measurements to individual tissue units. Note that timestamps for tissue measurement files are often captured within the file.

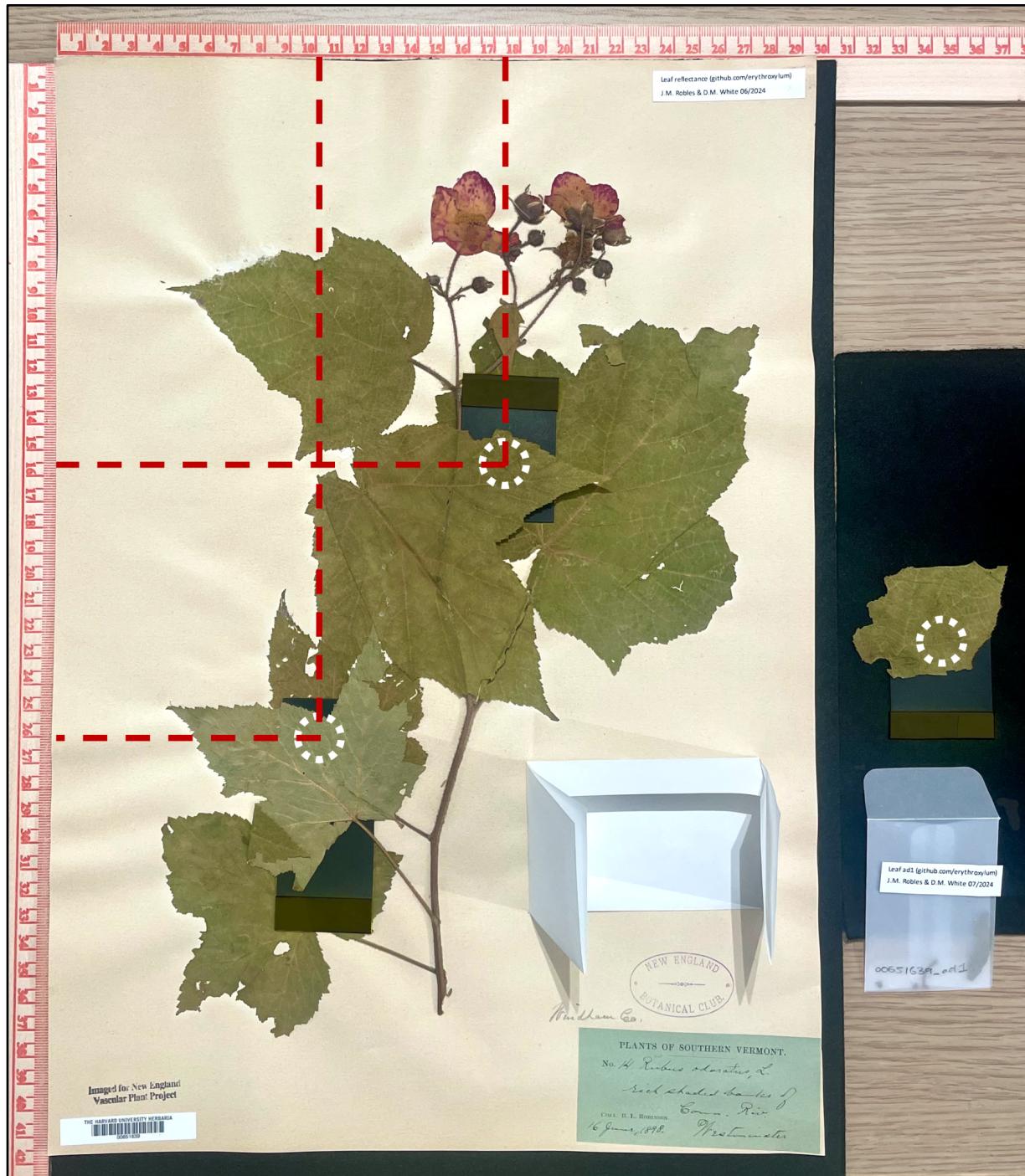
**Table 4.3: Tissue Metadata**

Field Name	Status	Field Description	Data type
backgroundClass	Required	Enumerated abbreviated code from Background Class Codes <a href="#">Table 4.7: Background Class Codes</a> describing the type of background used behind target tissue. Both abbreviated codes and descriptive codes are accepted.  Examples: "BGW", 'BGB', 'BGP'	ENUM ( <a href="#">Table 4.7: Background Class Codes</a> )
hasLowReflectanceBackground	Required	True or False statement that the background (black or paper) has low reflectance as defined as less than 4% reflectance across the spectral range of the instrument. For a paper background, this would be scored 'false'.	BOOLEAN(true/false)
backgroundDescription	Conditional Required	Description of the black or other background material, including manufacturer and product information when available. Not required for paper backgrounds. Required field when tissue has black or other type (not paper) of background.  Example: Musou IR Flock	TEXT
targetClass	Required	Free text or enumerated code from Target Class Codes ( <a href="#">Table 4.5: Target Class Codes</a> ) describing type of tissue or background being measured. Both abbreviated codes and full codes can be used.  Examples: 'AD', 'perigynium'	TEXT or ENUM ( <a href="#">Table 4.5: Target Class Codes</a> )
targetTissueId	Optional	Character index tracking the	TEXT

		<p>measured tissue units when multiple tissues are measured from a single specimen (e.g., 'loose1', '1', '2'). For compound or more complex structures, projects are encouraged to develop their own consistent naming convention.</p> <p>Examples: 'pinna1', 'leaflet1', 'petal1'</p>	
tissueDevelopmentalStage	Required	<p>Tissue developmental stage as coded in Developmental Stage Class Codes (<a href="#">Table 4.4</a>).</p> <p>Examples: 'Mature', 'Uncertain', 'NotScored'</p>	ENUM ( <a href="#">Table 4.4: Developmental Stage Class Codes</a> )
hasBackgroundInMeasurement	Required	<p>True or false values indicating that the target tissue does not cover the full measurement area and the background is part of the measurement.</p>	BOOLEAN (true/false)
percentBackgroundInMeasurement	Optional	<p>Numeric estimate for the percentage of the measurement area that is not covered by the target tissue and is background material (black background or herbarium paper). It is recommended to describe the estimation method in the <a href="#">comment</a> field.</p> <p>Example: 25</p>	INTEGER
hasGlue	Required	<p>True, false, or uncertain values indicating glue is present in the measurement area.</p>	ENUM (true/false/uncertain)
hasNonGlueContamination	Required	<p>True, false, or uncertain values indicating a contaminant other than glue is present in the measurement area. This includes foreign biotic or abiotic agents on the target tissue, such as fungus or preservatives.</p>	ENUM (true/false/uncertain)
measurementFlags	Optional	<p>Standardized categorical descriptors of the condition of the tissue <i>within the measurement area</i>. Values are selected from the predefined Tissue Descriptor Codes (<a href="#">Table 4.6</a>). Multiple descriptors should be separated with a pipe character.</p> <p>Example:</p>	ENUM ( <a href="#">Table 4.6: Tissue Descriptor Codes</a> )

		'GoodPreservation PathogenPresent'	
<code>tissueNotes</code>	Optional	<p>Free-text field used to record additional observations on the condition of the specimen that may aid interpretation of spectral data. It can be used to clarify or elaborate on descriptors already included in <code>measurementFlags</code>, such as the conditions evidencing the quality of preservation (e.g., a <code>MediumPreservation</code> flag could be explained with the note, 'measurement area discolored and wrinkled'). Notes should specify whether the information applies to the measurement area, the tissue unit, or the specimen as a whole.</p> <p>Examples: 'mold in measurement area, herbivory on leaf, specimen discolored and burnt', 'formaldehyde preserved specimen', 'very degraded specimen'.</p>	TEXT
<code>tissueLocation</code>	Optional	<p>The location of the target tissue on the herbarium sheet. For mounted tissues, record as an 'X,Y' coordinate in centimeters, with '0,0' at the top-left corner of the sheet (e.g., '17,29'; see <a href="#">Fig. 4.2</a>). If the sheet has non-square angles, align it flush with the left-side ruler. For unmounted tissues, provide a descriptive note indicating location.</p> <p>Examples: 'envelope TCAD_TN1', '17,29'</p>	<p>TEXT Coordinates preferred in the form of an integer tuple: 'X,Y'.</p>
<code>comment</code>	Optional	<p>Free-text field for recording any additional notes relevant to the measurement, including observations about the instrument, session, specimen, tissue, or data quality that are not captured elsewhere in the metadata.</p> <p>Example: 'Amazing specimen.'</p>	TEXT
<code>measurementIndex</code>	Required	<p>The measurement number index appended to the base filename (<a href="#">Part 3, Table 3.2: IDX</a>) to properly associate each row of metadata</p>	TEXT

		<p>with its single, corresponding measurement file.</p> <p>Example: 0001</p>	
--	--	--	--



**Fig. 4.2:** Diagram of coordinate system for scoring `tissueLocation` in [Tissue Metadata](#).

Herbarium sheet is placed on top of the benchtop black background with centimeter rulers at top and left sides for 'x,y' notation of measurement area (white dashed circles) in centimeters with '0,0' at the top left. Black background cards are placed under unglued portions of leaves. From left to right, tissue TCAB\_TN3 has the location '10,26', tissue TCAD\_TN2 has location '17,16' and tissue TCAD\_TN1 is stored with a label in a glassine envelope inside the packet with location 'envelope TCAD\_TN1'. For reference, specimen [NEBC\\_00651639](#) metadata fields are proposed in [Appendix II](#).

## 4.2: Controlled vocabularies

**Table 4.4: Developmental Stage Class Codes.** Available codes for enumerating the required `tissueDevelopmentStage` metadata ([Table 4.3](#)). Codes follow ‘CamelCase’ format with capitalized initial letters.

Code	Description
Young	Actively developing tissue that is not yet fully expanded; may appear thinner, lighter in color, or more pliable than mature tissue.
Mature	Fully developed and expanded tissue showing typical structural and color characteristics for the taxon; not visibly senescent.
Old	Senescent tissue showing visible signs of aging or decline, such as yellowing, darkening, curling, or drying.
Uncertain	The development stage has been assessed but cannot be confidently determined due to intermediate features, damage, or insufficient visual cues.
NotScored	Developmental stage was not assessed or recorded for this tissue.

**Table 4.5: Target Class Codes.** Available codes for enumerating the required `targetClass` metadata ([Table 4.3](#)). Either the abbreviated code or the full CamelCase-formatted code (with an initial capital letter) may be used.

Code	Full code	Description
W	WhiteReference	White reference.
WC	WhiteCalibratedReference	White calibrated reflectance standard (see <a href="#">Section 5.3</a> ).
B	BlackBackground	Black background material, recorded when used as background for other target tissue measurements.
BC	BlackCalibratedReference	Black calibrated reflectance standard (see <a href="#">Section 5.3</a> ).
P	Paper	Herbarium sheet paper, recorded when used as background for other target tissue measurements.
AB	LeafAbaxial	Abaxial leaf surface.
AD	LeafAdaxial	Adaxial leaf surface.
LF	Leaf	Leaf surface. Applied when the abaxial and adaxial side

		cannot be differentiated or when leaves are terete or otherwise not bifacial (e.g. <i>Senecio rowleyanus</i> ).
PT	Petal	Petal.
IF	Inflorescence	Inflorescence.
BR	Bract	Bract.
FR	Fruit	Fruit. The specific tissue (e.g. exocarp, mesocarp) can be described in <a href="#">tissueNotes</a> .
PSS	Photosynthetic-SucculentStem	Photosynthetic stem as in succulents like <i>Cactus</i> .
OB	OuterBark	Outer bark, rhytidome. Woody branch outer bark as in Hadlich <i>et al.</i> (2018).
IB	InnerBark	Phloem. Woody branch inner bark as in Hadlich <i>et al.</i> (2018).
HS	HerbaceousStem	Herbaceous stem; can be photosynthetic as in PSS.
WD	Wood	Wood.

### Tissue Descriptor Codes

The descriptors may be used individually or in combination to characterize contaminants or associated biological material present in the measurement area. These categories are not mutually exclusive—multiple flags may apply to the same tissue due to both the multiplicity of descriptors and overlap in their definitions (e.g., a specimen may have both *FungusPresent* and *PathogenPresent*). When ambiguity exists, further clarification should be recorded in the [tissueNotes](#) field.

**Table 4.6: Tissue Descriptor Codes.** Available codes for enumerating the optional `measurementFlags` tissue metadata ([Table 4.3](#)). See [Appendix II](#) for exemplary images of tissues and assigned descriptor codes

Code	Description
GoodPreservation	Tissue in the measurement area appears well preserved in color, structure, and texture (including original features from disease or herbivory) and shows minimal signs of degradation or breakage from pressing, drying, or storage. Tissues that are simply discolored may still be considered well preserved if other aspects of integrity are maintained.
MediumPreservation	Tissue in the measurement area shows moderate degradation, such

	as partial discoloration, wilting, or deformation. Some structural loss may be present, though not severe. Evidence of degradation from other parts of the specimen (e.g., mold elsewhere on the tissue) may support assigning this level of preservation, but such observations should be recorded in <a href="#">tissueNotes</a> if not present within the measurement area. This is expected to be the most common preservation condition for herbarium specimens used in spectral measurement.
PoorPreservation	Tissue in the measurement area shows clear signs of degradation, including severe discoloration, wrinkling, deformation, or breakage. Mold, insect damage, or other signs of poor preservation may also be present. As with other flags, evidence of degradation outside the measurement area (e.g., mold elsewhere on the tissue) may support the assigned flag but should be recorded in <a href="#">tissueNotes</a> if not directly observed in the measurement area. Note that natural discoloration tendencies of certain taxa should be considered when applying this flag (see <a href="#">Appendix II</a> ).
MidveinPresent	Target measurement area contains midvein or similarly prominent secondary venation.
OrganismPresent	Indicates that a visible organism (e.g., bryophyte, lichen, fungal structure) is present on or within the measurement area. This includes epiphyllous, endophytic, or other leaf-associated organisms, regardless of their ecological role (e.g., mutualistic, parasitic, or incidental). This flag serves as a general indicator and can be used in conjunction with more specific organism flags below.
BryophytePresent	Indicates that a visible bryophyte (e.g., moss, liverwort, or hornwort) is present on or within the measurement area.
LichenPresent	Indicates that a visible lichen thallus or fragment is present on or within the measurement area.
FungusPresent	Indicates that fungal structures are visible in the measurement area (e.g., hyphae, mycelium, fruiting body) and are presumed to be pre-mortem associates, such as endophytes or pathogens active while the plant was alive. This flag can overlap with <i>PathogenPresent</i> or <i>MoldPresent</i> .
PathogenPresent	Target measurement area contains necrotic tissue or other signs of pathogenic infection. Can be used with <i>FungusPresent</i> when fungal pathogens are suspected or known. This flag is based on visible tissue symptoms, not molecular confirmation.
MoldPresent	Indicates that the measurement area shows signs of post-mortem fungal growth (e.g., surface mold, fuzz, bloom), likely resulting from poor drying or storage conditions. In practice, mold may be difficult to distinguish from other fungal growth without microscopic or culture analysis. Use judgment and note uncertainty in <a href="#">tissueNotes</a> if

	needed.
HerbivoryPresent	Target measurement area contains herbivory.
AlcoholPresent	Target measurement area was preserved with ethanol or other alcohol.
PreservativePresent	Target measurement area contains chemical preservative contamination excluding alcohol preservatives (e.g., diatoms, formaldehyde).
BurnPresent	Target measurement area contains burned tissue.
DebrisPresent	Target measurement area contains non-specific material not described in other codes (e.g., dust, soil particles, insect parts, fibers, etc.) that may interfere with clean measurements. Can be elaborated in the tissueNotes field.

**Table 4.7: Background and White Reference Class Codes.** Available codes for enumerating the required tissueBackgroundClass metadata ([Table 4.3](#)). Either the abbreviated code or the full CamelCase-formatted code (with an initial capital letter) may be used.

Code	Full code	Description
W	WhiteReference	White reference.
B	BlackBackground	Black background. The protocol requires this material to be described in the <code>backgroundDescription</code> field ( <a href="#">Table 4.3</a> ).
P	PaperBackground	Herbarium sheet paper with or without glue present.
O	OtherBackground	Other background that is not black or herbarium sheet paper. Material to be described in the <code>backgroundDescription</code> field.

#### 4.3: Guidelines for Data Archiving and Sharing

- For each specimen, a copy of all **unprocessed spectral files and metadata should be archived** in the digital repository of the herbarium or institution that owns the specimen, or otherwise in accordance with their data storage practices. This ensures that the data remain co-located with the physical specimen and integrated with institutional capabilities for managing specimen metadata.
- In addition, projects should **upload unprocessed, original format data files and metadata to a persistent, open-access repository** that issues DOIs and supports versioning, such as Dryad, Zenodo, or Harvard Dataverse. These platforms are

preferred over tools like EcoSIS, which have shown reduced reliability and uncertain long-term support.

- To facilitate reuse, users may choose to also share tabular data files (e.g., .csv) with samples in rows and columns for metadata fields and spectral band values. If processed spectra are included (e.g., interpolated, splice/jump-corrected, or continuum-removed), the applied processing steps should be clearly documented.
- See example of data and metadata sharing here:  
<https://doi.org/10.7910/DVN/LXPHBC>

## Part 5: Instrumentation and Materials Guidelines

### 5.1 Precautions

1. Carefully handle backgrounds and white references (e.g., Spectralon®)
  - Use laboratory gloves to handle black backgrounds and white references and tweezers to handle herbarium tissues.
  - Avoid touching and dirtying the reflective surfaces.
  - Avoid blowing or cleaning backgrounds and reference standards with canned air, which can cause chemical contamination (other options, such as [O2 Hurricane](#), avoid this contamination).
2. Avoid burning target tissues.
  - High light intensities in optical probes can burn tissues, especially thin leaves! Use low light intensity and minimum measurement duration based on manufacturer's guidance to avoid this.
  - Avoid measuring tissues with glue on them, whenever possible.
  - Place a black background under the leaf or target tissue to avoid spectral contamination from the herbarium mounting paper whenever possible. Background surfaces, except for non-reflective black, will influence reflectance spectra (see [Fig. 1.1](#)).

### 5.2 Spectra data files

1. Output unprocessed percent reflectance data to files, if possible.
  - Percent reflectance should be present in the file at the wavelengths directly measured by the instrument.
  - If instruments and software output transformed data by default (e.g., Spectral Evolution PSR+ output is interpolated at 1 nm resolution), refer to user manuals to change settings to output percent reflectance values at measured wavelengths instead.
2. File formats differ among spectroradiometer manufacturers, models, and softwares. Make sure that files include percent reflectance values.

### 5.3 Instrumentation and materials quality control

1. Instrument maintenance:
  - Follow manufacturer SOP for spectroradiometer maintenance. If possible, send the instrument to the manufacturer at regular intervals for maintenance (e.g., every two years).
  - Have fiber optic cables checked by the manufacturer annually and replace damaged fiber optic cables.

2. For background and white reference measurement requirements, see [Section 1.2](#).
  - White reference scans are typically not saved as a separate file, but the white reference digital numbers (raw sensor values) should be included in each target spectrum file (e.g., `.sig` format). These values are used by the instrument software to calculate reflectance from raw radiance.
  - Black background materials are required to have less than 4% reflectance across all wavelengths ([Fig 1.1](#); e.g., IR Flock Sheet from Musou Black USA: <https://musoublackusa.com/products/ir-flock-sheet>).
  - If measuring leaves on herbarium paper on a benchtop, place a black mat under the herbarium sheets to prevent contaminating reflectance from the benchtop ([Fig. 4.2](#)). Black felt (3–5 mm thick) is a practical and cost-effective choice for this purpose.
3. Target measurements of background material (black or paper) must be linked to tissue measurements using filename conventions ([Part 3](#)); this supports quality control of background conditions and enables development of spectral unmixing approaches to isolate tissue reflectance.
4. Wear gloves and carefully handle all tissues, white reference, black backgrounds, and calibrated reflectance standards to keep them clean (see [5.1 Precautions](#)).
5. Calibrated reflectance standards:
  - Institutions should consider purchasing calibrated reflectance standards, such as Spectralon® Calibrated (SL Standard) Diffuse Reflectance Standards. These should be kept clean and separate from ‘working’ white references, and used for long-term assessment of instrument performance, optical drift, and the condition of reference materials (e.g., white references and black backgrounds).
  - These are supplied with a **calibration certificate and serial number** that verifies reflectance values across wavelengths.
  - Reflectance standards are white (e.g., Spectralon® 99%) and black (e.g., Spectralon® 2%) at a minimum, and can also include gradations of gray.
  - It is recommended that standards be measured at regular intervals depending on the level of activity (e.g., once per session; consult spectroradiometer manufacturer), and archived with the filename conventions provided in [Table 3.1](#).

## 5.4 Specimen and tissue annotation

1. Metadata recommendation:
  - IHerbSpec recommends recording the `tissueLocation` metadata field (see [Table 4.3](#)) to document the position of measured tissues on the specimen.
2. Consult with the herbarium collections manager(s) to define a suitable specimen annotation protocol that aligns with institutional policies.

3. Recommended specimen annotation practices:
  - Attach a project annotation label to the herbarium sheet (e.g., "*Leaf spectral reflectance; DM White (2025)*").
  - Use archival-quality materials for labels and annotations, such as acid-free paper.
  - Use pencil for direct annotation of measured tissues on sheets.
  - For loose tissues (from packets or envelopes):
    - i. place the measured tissue in a glassine envelope, which is archival and thin.
    - ii. Include a label inside the envelope indicating the `targetTissueClass` and `targetTissueId` ([Table 4.3](#); e.g., “*Tissue TN1*”).
    - iii. Store the envelope in the specimen packet or attach it to the herbarium sheet (see [Fig. 4.2](#)).

## Part 6: Selecting Tissues for Spectral Measurement

### 6.1 Sources of variation in herbarium plant tissues

Spectral data from herbarium specimens are influenced by two major sources of variation. The first is the natural **biological variation** of phenotypes constructed through genotype-by-environment interactions that biologists seek to understand. This includes differences in developmental stage (e.g., young, mature, senescent), within-individual morphological variation (e.g., sun vs. shade leaves), and biotic influences such as herbivory or pathogens.

The second source is **herborization-related variation**—herborization being defined as the process of preserving plant tissues in an herbarium, from the time of collection through long-term storage. This includes collection, pressing, drying, application of preservatives, and subsequent storage conditions. Each of these steps will affect the spectral data to some extent. For example, differences in technique (e.g., bagged vs. immediately pressed specimens; use of ethanol) and environmental factors during drying and storage can lead to discoloration, dehydration, tissue distortion, and contamination from biotic or abiotic agents such as fungus, glue, or insecticides. Among these, the drying protocol might be particularly influential: specimens not dried efficiently—especially in warm, humid conditions—are prone to rot and degradation, which significantly affects spectral quality. Over time, temperature and humidity fluctuations may

As a result, herbarium specimens exhibit a wide range of tissue conditions, reflecting both the biology of the living organism and the varied preservation environments they experience.

**Documenting these sources of variation as metadata is essential**, as they can be correlated with model accuracy (Kühn et al. 2024, White et al. 2025) and can be leveraged to improve the performance of predictive models

When available, contextual details from specimen labels—such as the use of ethanol or other treatments—should be captured in metadata. As it is rare to see such details, botanists should be encouraged to include preservation protocols on specimen labels. The metadata fields, `hasGlue`, `hasNonGlueContamination`, `measurementFlags`, and `tissueNotes` ([Table 4.3](#)), are designed to record the condition of the tissue **within the target measurement area at the time of measurement**. These fields are essential for downstream filtering and analysis of spectra. Examples of specimen tissue metadata records are provided in [Appendix II](#).

### 6.2 General guidelines for tissue selection

Note: Much of this text is tailored to measuring leaf tissues, but can generally be applied to other tissues.

#### 1. Quality over quantity.

- Under the goal of broad digitization of herbarium tissues, spectral data quality could be considered more important than specimen quantity. Targeting higher quality, less degraded specimens will yield more informative spectral data.

- If the majority of tissues on a specimen are degraded, damaged, or contaminated, that specimen should generally not be measured unless it holds particular research value.

**2. Select specimens based on their general tissue condition:**

- For digitization projects focused on broad taxonomic coverage, specimens should be evaluated as whole units when determining suitability for spectral measurement. Next, technicians should prioritize tissues that are both suitable for measurement and representative of the overall condition of the specimen, rather than selecting isolated tissues that are in unusually good condition. This approach helps ensure alignment between the general specimen condition and the resulting spectral data, and could avoid biases based on assumptions of exceptional tissue conditions. **If the majority of tissues on a specimen are degraded, damaged, or contaminated, that specimen should generally not be measured unless it holds particular research value.** Conversely, if the specimen consists mostly of clean, mature, fully expanded tissues, it is appropriate to proceed with measurement—while also considering additional measurements, when feasible, to capture variation such as developmental stage or damage from herbivory and scoring the appropriate metadata for these additional measurements.
- A caveat: while the overall condition of the specimen informs whether it should be measured, **metadata on developmental stage, contamination, and tissue condition should always record the characteristics of the specific measurement area of the tissue**—not the specimen as a whole. This ensures metadata accuracy and supports high-resolution filtering and analysis. Information about the specimen can be recorded in the [tissueNotes](#) field ([Table 4.3](#)).

**3. Score metadata with care.**

- Metadata collection is a critical component for understanding factors affecting spectral data quality and will provide confidence during data aggregation.
- Herbaria and research teams should train technicians in specimen selection strategies and scoring metadata to reduce subjectivity.

**4. Prioritize:**

- Tissue measurements on black backgrounds.
  - i. Detached tissues of appropriate size in fragment packets allow full assessment of the presence of glue and access to both leaf surfaces.
  - ii. For mounted specimens, taped or sewn tissues can have black backing inserted beneath (technicians should look for evidence of old glue).
- Specimens with mature tissues (e.g., fully expanded leaves) that are clean of biotic or abiotic contaminants.
- Tissues that fill the probe measurement area, even if midvein is measured.

- Leaves with flat, complete surfaces and intact laminae.

**5. Avoid:**

- Tissues with glue.
- Tissues attached to herbarium paper such that no black background can be inserted beneath the tissue.
- Tissues that are not pressed flat (e.g., deformed, folded, crumpled, or abnormal leaves).
- Tissues contaminated by biotic or abiotic agents (e.g., algae, lichen, fungus, chemical preservatives, glue).
- Damaged tissues (e.g., by herbivory, pathogens, herborization phenomena).
- Midribs and major venation in the leaves, or other types of vasculature or intrusions in target tissue.
- If larger tissues are available, avoid smaller tissues that don't cover the entire optical probe measurement area.

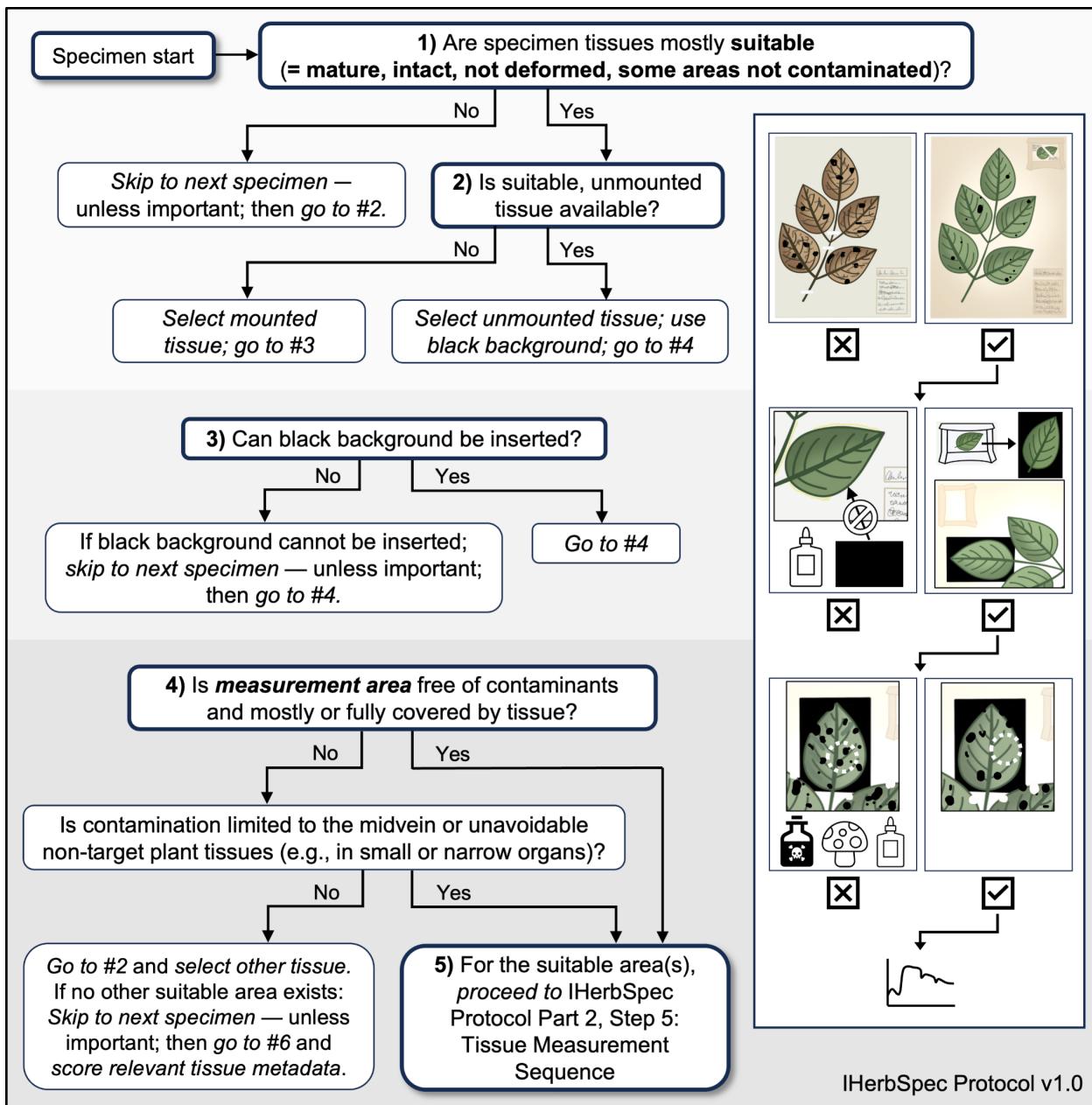
**6. Measuring small tissues:** Measuring taxa with naturally small tissues presents unique challenges, but several strategies can help obtain representative spectral measurements with minimal error.

- Center the probe on multiple areas of the small tissue and collect several measurements, then average the values to represent the tissue's spectral profile.
- Create a mosaic by arranging multiple small tissue units side by side to fill the measurement area. Avoid overlapping tissues, as stacking can alter light scattering and distort reflectance values ([Fig. A1](#)).
- Consider optical setup with small measurement area: For projects in the instrument selection phase, Spectral Evolution offers a custom optical probe with a 2-mm measurement area compatible with their NaturaSpec and PSR spectroradiometers.
- Avoid using narrow angle lenses as these might distort the spectrum ([Fig. A1](#)).

**7. Multiple individual plants within one herbarium sheet:**

- Each individual plant on a sheet should be demarcated with a unique **specimenId** (see [Table 3.3](#)).

### 6.3 Decision tree for tissue selection



**Fig. 6.1:** Decision tree diagram for target tissue selection.

## References

- Cavender-Bares J, White DM, Ahlstrand NI, Austin M, Bastianelli D, Bazan S, Boughalmi K, Cardinal-McTeague W, Chacón-Madrigal E, Couvreur TLP, Davis C, Durgante FM, Grace OM, Q JAG, Hansen K, Hernández-Leal MS, Hopkins MJG, Jackson R, Kothari S, Lee AK, Léveillé-Bourret É, Pinto-Ledezma J, Casaverde NLQ, Meireles JE, Nichodemus CO, Schmull M, Soltis DE, Soltis PS, Tuomisto H, Ustin S, Vasconcelos CC. 2025. NextGeneration specimen digitization: The international herbarium community goes spectral! *EcoEvoRxiv* [preprint]. <https://doi.org/10.32942/X2V927>.
- Durgante FM, Higuchi N, Almeida A, Vicentini A. 2013. Species Spectral Signature: Discriminating closely related plant species in the Amazon with Near-Infrared Leaf-Spectroscopy. *Forest Ecology and Management* 291:240–248. <https://doi.org/10.1016/j.foreco.2012.10.045>.
- Hadlich HL, Durgante FM, dos Santos J, Higuchi N, Chambers JQ, Vicentini A. 2018. Recognizing Amazonian tree species in the field using bark tissues spectra. *Forest Ecology and Management* 427:296–304. <https://doi.org/10.1016/j.foreco.2018.06.002>.
- Kühn P, Proß T, Römermann C, Wesche K, Brügelheide H. 2024. Using near-infrared spectroscopy to predict nitrogen and phosphorus concentrations of herbarium specimens under different storage conditions. *Plant Methods* 20:19. <https://doi.org/10.1186/s13007-024-01146-x>.
- Mersni J, Bazan S, Baby E, Bonnal V, Bastianelli D. 2025. Protocol for NIR spectra acquisition on herbarium specimen. *Agritrop*. <https://doi.org/10.18167/agritrop/00880>.
- Neto-Bradley BM, Bonnet P, Goëau H, Joly A, Cavender-Bares J, Coomes DA. 2025. Using reflectance spectra and PI@ntNet to identify herbarium specimens: a case study with Lithocarpus. *New Phytologist*. <https://doi.org/10.1111/nph.70258>.
- White, DM, Cavender-Bares, J, Davis, CC, Guzmán Q, J.A., Kothari, S, Robles, JM and Meireles, JE. 2025. Seeing herbaria in a new light: leaf reflectance spectroscopy unlocks trait and classification modeling in plant biodiversity collections. *New Phytologist*. <https://doi.org/10.1111/nph.70357>.

## Appendix I: Considerations regarding the number of measurements per specimen

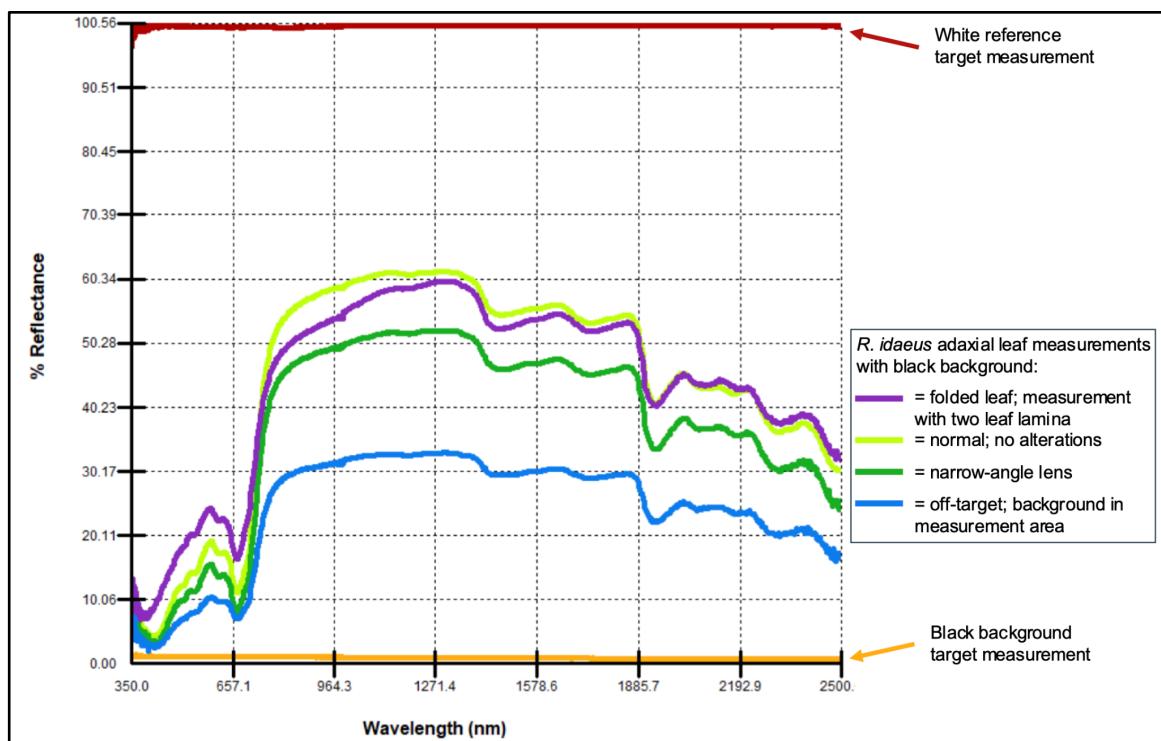
Spectral measurements of plant tissues are highly sensitive to both instrument configuration and small spatial differences across target tissue. Changes in the optical probe geometry in relation to the tissue—such as tilting, poor contact with the leaf surface, partial coverage of the measurement area ([Fig. A1](#)), or micro-contaminants like debris—can introduce technical errors. In addition, leaf surface microtopography, anatomical variation, and alterations from herborization (e.g., drying distortion, discoloration, or degradation) all contribute to variation in reflectance spectra (Cavender-Bares et al. 2025).

These considerations support the IHerbSpec Protocol requirement of collecting a minimum of three representative measurements per adaxial and abaxial leaf surface, when suitable tissues for both surfaces are available. Three measurements are necessary to calculate a mean spectrum and its variance. The protocol further recommends collecting five measurements per surface to achieve a more robust spectral characterization and to meet the threshold for improved performance in classification models. The protocol allows for users to make these minimum or higher measurements across one or many leaves. For other tissue types, at least three measurements are recommended, with more encouraged when feasible. Below, we elaborate justification for this recommendation.

Several studies have demonstrated improved model performance when multiple measurements per individual specimen are averaged, including those by Durgante et al. (2013; [Figs. A2, A3](#)), Neto-Bradley et al. (2025), and S. Bazan et al. (Unpubl.). The IHerbSpec Protocol's recommendation of five measurements is based on empirical evidence that model accuracy gains tend to stabilize beyond five to ten measurements per tissue ([Fig. A2](#)), although the optimal number may vary depending on tissue type, instrument sensitivity, and modeling goals. Beyond classification, IHerbSpec members are investigating how replicate measurements affect the accuracy and robustness of trait prediction models, evaluating both the benefits of averaging spectra and the variation in predicted trait values from unaveraged measurements of the same tissue.

Another important recommendation in the IHerbSpec Protocol is to measure both adaxial and abaxial leaf surfaces. These surfaces differ in anatomy and function due to evolutionary adaptations for light capture, gas exchange, and environmental stress. Such structural differences are consistently reflected in their spectral signatures (see main text [Fig. 1.1](#)), and we recommend measuring both surfaces—when suitable tissue is accessible—to capture this biologically meaningful variation.

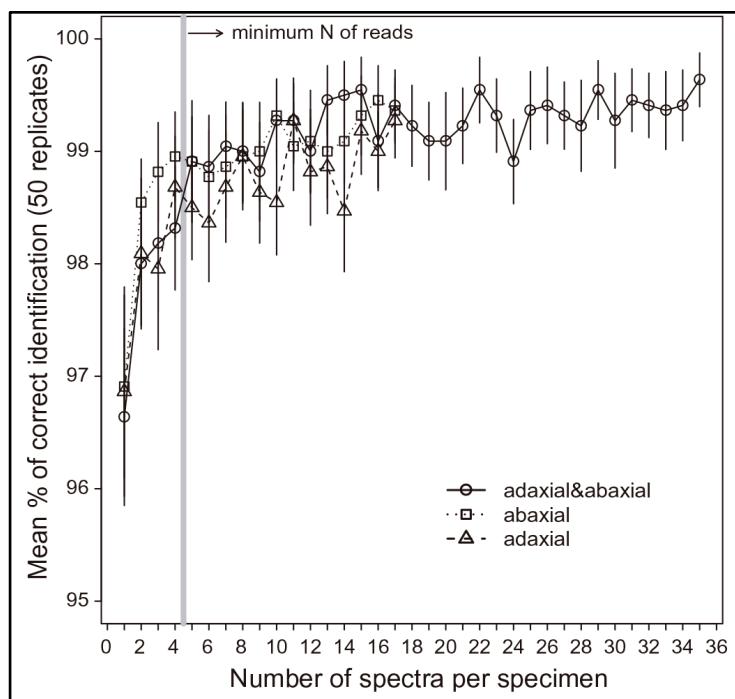
Within the context of a large-scale macroevolutionary project focusing on Annonaceae (ERC GLOBAL project), researchers aimed to measure spectra on one specimen per species with the intent to capture as much variation as possible for each specimen. In accordance with the protocol established at the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), researchers decided to make 20 spectral measurements per specimen to develop their spectral library for 1700 Annonaceae species preserved in over 30



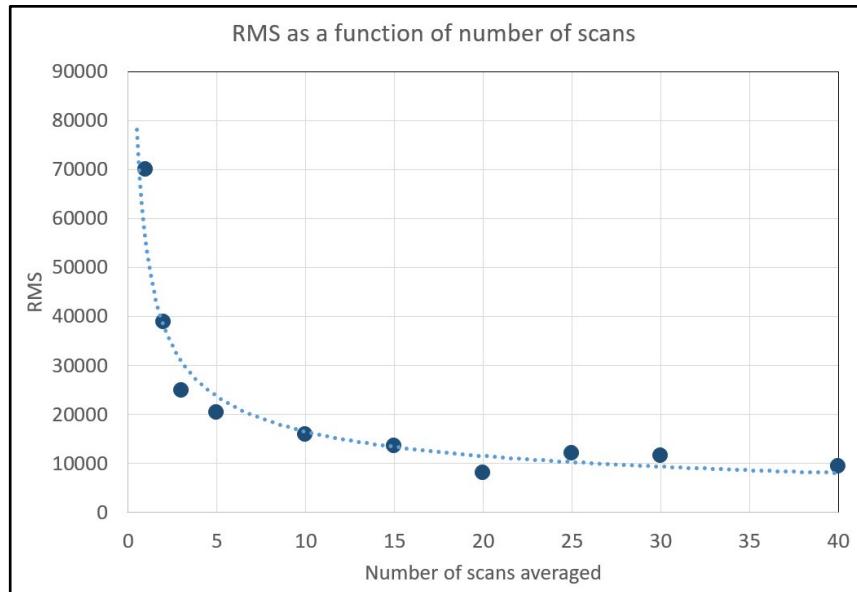
**Fig. A1:** Variability in spectral measurements due to instrument configuration and technical error. The following spectral measurements are shown: white calibration standard (red), black background (orange), standard adaxial measurement with black background (light green), standard adaxial measurement but with narrow-angle lens following recalibration (dark green), adaxial measurement where the tissue did not fully cover the optical field of view (i.e., `backgroundInMeasurement = TRUE`; blue), and a folded leaf (purple). All measurements were made of the same tissue using an SVC HR-1024i spectroradiometer.

different herbaria. This decision was informed by preliminary experiments demonstrating that spectral stability improves with larger numbers of measurements (Fig. A3). The choice of 20 spectra was also appropriate because of variation in the number and size of leaves, allowing for multiple measurements of multiple leaves or single large leaves that filled an entire herbarium sheet. It also allowed for poor-quality spectra to be discarded based on downstream filtering criteria while still retaining a large number of measurements per specimen.

Overall, given the sensitivity of spectral data, research projects are encouraged—when feasible—to collect more than five measurements per tissue, including across different tissue units within an individual. This enables further assessment of intra-individual spectral heterogeneity and its impact on downstream analyses. The marginal cost of additional measurements is relatively low across instruments, with a single measurement typically completed within 1 to 10 seconds. Nevertheless, determining the appropriate number of measurements remains a critical design consideration, particularly in large-scale digitization projects involving significant resource investment.



**Fig. A2:** Accuracy of taxonomic discrimination as a function of the number of averaged measurements per specimen. Three tissue class datasets are analyzed: combined adaxial and abaxial spectra (circles), abaxial only (squares), and adaxial only (triangles). Each point represents the average classification accuracy from 50 replicate linear discriminant analysis models, using randomized subsampling and test set validation. Error bars indicate 95% confidence intervals. The vertical gray line marks the minimum number of spectral measurements ( $n = 5$ ), below which accuracy significantly decreases (Tukey test,  $P \leq 0.05$ ). Data were collected using an FT-NIR spectrometer (1000–2500 nm) from 10 species in Lecythidaceae, including eight *Eschweilera* and two *Corythophora* species. Accuracy increases notably between one and five spectra per specimen, after which performance stabilizes, indicating diminishing returns beyond five measurements. This figure is reproduced from Durgante et al. (2013) with permission from the publisher.



**Fig. A3:** Spectral variability as a function of the number of measurements averaged per measurement. Here, a total of 100 spectral measurements were collected from a single *Annona* sp. leaf using an ASD LabSpec Pro spectrometer over a black background. Subsets of these measurements were then averaged in groups ranging from 2 to 40 to simulate different measurement strategies. The plot shows the root mean square (RMS) distance between replicate spectra as a function of the number of measurements averaged, illustrating that spectral variability decreases substantially as more measurements are combined. Notably, averaging fewer than five measurements did not adequately capture intra-tissue spectral variation. This analysis supported the required minimum of 10 measurements per specimen for the formal protocol currently implemented by teams at CIRAD and IRD in France (Mersni et al. 2025).

## Appendix II: Representing tissue conditions in metadata for quality control

This appendix provides a brief discussion and examples of the IHerbSpec approach for recording metadata pertaining to sources of biological and herborization variation that affect spectral data. As explained in [Section 6.1 Sources of Variation](#), we aim to capture biological variation that is scientifically meaningful, but the degraded nature of herbarium specimens requires a careful assessment and annotation of tissue quality and sources of contamination. The Tissue Metadata fields of `developmentalStage`, `hasGlue`, `hasNonGlueContamination`, `measurementFlags`, and `tissueNotes` ([Table 4.3](#)) are designed to support quality control of spectral data and enable downstream filtering.

The codes for enumerating the `developmentalStage` fields are provided in [Table 4.4](#), and the codes for `measurementFlags` are described in [Table 4.6](#).

Technicians should also be aware that some taxa (e.g. Myrtaceae) or tropical collections in general may naturally exhibit discoloration or non-flat, deformed surfaces (e.g., *Gasteranthus atratus*). Such characteristics do not necessarily indicate tissue degradation and poor-quality spectral data. In addition, some taxa exhibit more pronounced discoloration during herborization as a result of their biology, and such discoloration should not be interpreted as evidence of poor preservation condition. Some traits also covary; for example, young leaves wilt and become discolored faster and are more challenging to press and dry flat, so they may more often be scored as moderately or poorly preserved.

As a reminder, `measurementFlags` pertain only to the measurement area, not the whole tissue unit (e.g., leaf). Since suitable measurement areas of tissues (see decision tree in [Fig. 6.1](#)) will not contain any contaminants, pathogens, or other damage, technicians should not be enumerating any codes in the `measurementFlags` field. This could differ in projects collecting measurements pertaining to extended uses of the specimen, e.g., symbionts, degradation, etc., that might be measuring tissue areas with these other features.

The specimens and metadata examples in [Table A1](#) have been selected to guide consistent scoring of these fields during spectral measurement and promote metadata consistency across projects. Note that the images do not specify the exact measurement area corresponding to these metadata fields (as is done in the [IHerbSpec Metadata Spreadsheet](#) for specimen [NEBC\\_00651639](#) shown in main text [Fig. 2.2](#)).

**Table A1: Select Metadata Fields and Scoring for Representative Specimens.** This table provides example metadata records for herbarium specimens, illustrating how to score key fields related to tissue condition and contamination. The fields `developmentalStage`, `hasGlue`, `hasNonGlueContamination`, and `measurementFlags` refer specifically to the measurement area. Broader observations about the specimen or other tissues can be recorded in `tissueNotes` (see [Table 4.3](#)). Linked specimen images are included, though they do not indicate the precise measurement area corresponding to each metadata record. This table is accessible through the IHerbSpec Protocol documentation page: <https://iherbspec.github.io/protocol>.

Specimen	developmentalStage	hasGlue	hasNonGlueContamination	measurementFlags	tissueNotes
A_00631088	mature	true	false	MediumPreservation	discolored
A_00672503	mature	false	true	GoodPreservation PathogenPresent FungusPresent	Small necrotic spots in measurement area
A_00746288	mature	true	false	MediumPreservation	Discolored. Glue on loose leaf in measurement area.
A_00772865	young	false	false	PoorPreservation	Measured leaves uneven, mottled, discolored, breakage
A_2613405	mature	false	false	MediumPreservation	
A_2614928	mature	uncertain	true	PoorPreservation OrganismPresent FungusPresent MoldPresent	Discolored, herbivory on sheet. Glue used on sheet but unsure if in measurement area.
AAU_A.S.Barfod48062	mature	false	true	MediumPreservation MidveinPresent DebrisPresent Herbivory	Herbivory on specimen
AAU_I.A.Cacon1199	mature	false	true	MediumPreservation AlcoholPresent	discolored, grayish spots on specimen
AAU_R.C.Moran6023	mature	false	true	GoodPreservation AlcoholPresent	
ALF_035514	Mature, young	false	false	GoodPreservation	
ECON_00338371	young	false	false	MediumPreservation	Discolored, wrinkled. Specimen burnt but not in

					target area.
GH_00611866	mature	false	false	MediumPreservation	Discolored specimen
INPA_142198	mature	false	false	MediumPreservation	discolored, herbivory on specimen
INPA_179322	mature	false	false	GoodPreservation	
INPA_184434	mature	false	false	MediumPreservation	Herbivory on specimen, necrotic leaf
INPA_187932	mature	false	false	MediumPreservation	burnt leaves on specimen
INPA_203125	mature	false	false	MediumPreservation	discolored, wrinkled leaves
INPA_218399	mature	true	false	PoorPreservation	discolored, herbivory, wrinkled and burnt leaves
INPA_264413	mature	false	false	GoodPreservation	
INPA_267597	mature	false	false	GoodPreservation	
INPA_275373	mature	false	false	MediumPreservation	discolored, mold, grayish spots on older leaves specimen
INPA_278882	mature	false	true	MediumPreservation AlcoholPresent	discolored
INPA_4628	mature	false	true	PoorPreservation PreservativePresent	mold, whitish powder (diatoms?), poor storage conditions
INPA_59178	mature	true	true	PoorPreservation PathogenPresent OrganismPresent	discolored, herbivory, dark spots on leaves
INPA_72939	mature	true	true	PoorPreservation FungusPresent	discolored, herbivory on measured leaves
INPA_97794	mature	false	false	MediumPreservation	wrinkled leaves
INPA264208	mature	false	false	MediumPreservation	Measurement area not flat.

					Herbivory on sheet
M_BU681750	mature	false	uncertain	MediumPreservation	Discolored; Alcohol preservation suspected.
MIN_332436	mature	false	false	MediumPreservation MidveinPresent	wrinkled leaves
MIN_370477	mature	false	false	PoorPreservation	discolored, wrinkled leaves
MIN_588200	mature	false	false	GoodPreservation	
NEBC_00634726	young	false	false	PoorPreservation	Measurement area leaves uneven, discolored
NEBC_00636882	uncertain	false	true	PoorPreservation PathogenPresent	Senescent? Uneven, blight, necrosis in measurement. But no glue.
NEBC_00651639	mature	false	false	GoodPreservation	
NEBC_00677096	mature	false	true	MediumPreservation PathogenPresent FungusPresent	Leaves on sheet wrinkled and deformed
NEBC_00695035	mature	false	false	GoodPreservation	
NEBC_00746092	mature	false	true	MediumPreservation PathogenPresent	discolored
NEBC_00898559	mature	false	false	GoodPreservation	
NEBC_02618198	mature	false	false	MediumPreservation	Dark spots on measured leaf
NEBC_02618743	mature	false	false	GoodPreservation	
NY_00042623	mature	true	false	GoodPreservation	
NY_00043385	mature	false	true	MediumPreservation PathogenPresent	Discolored; dark spots on leaves
NY_00194824	mature	false	false	GoodPreservation	

## IHerbSpec Protocol v1.0

<a href="#">NY_00402473</a>	mature	true	false	MediumPreservation MoldPresent	Discolored; dark spots on leaves
<a href="#">NY_00709210</a>	mature	false	true	MediumPreservation PathogenPresent	Discolored; herbivory on specimen. Alcohol suspected.
<a href="#">NY_01163859</a>	mature	true	false	MediumPreservation	Discolored; herbivory on specimen
<a href="#">NY_02498899</a>	mature	true	false	MediumPreservation	Discolored
<a href="#">NY_02499746</a>	mature	false	true	PoorPreservation MoldPresent	Dark spots on leaves; Discolored
<a href="#">P_00391795</a>	mature	false	uncertain	MediumPreservation	wrinkled leaves; possible chemical preservative.
<a href="#">P_05198525</a>	mature	false	true	PoorPreservation FungusPresent	Herbivory on sheet.
<a href="#">P_05198660</a>	mature	false	true	MediumPreservation AlcoholPresent	discolored.