Discrimination of Annonaceae using herbarium leaf reflectance spectra under limited sample size conditions

Khalil Boughalmi[1], Paola G. Santacruz Endara[2], Lily A. Bennett[3] , Martin Ecarnot[4], Samantha Bazan[5], Denis Bastianelli[5], Laurent Bonnal[5], Thomas L.P. Couvreur[1]

1 : DIADE, Univ. Montpellier, CIRAD, IRD, Montpellier, France

2: Grupo de investigación Biodiversidad y Sociedad BioS, Jardín Botánico Padre Julio Marrero, PUCESD, Santo Domingo de los Tsáchilas, Ecuador

3: School of Natural Sciences, University of Tasmania, Hobart, Tasmania, Australia

4: AGAP, INRAE, Univ. Montpellier, CIRAD, Institut Agro, Montpellier, France

5: SELMET, Univ. Montpellier, CIRAD, INRAE, Institut Agro, F-34398 Montpellier, France

Corresponding author : Khalil Boughalmi; boughalmi.khalil@gmail.com

Number of words: 4987

Manuscript received _____; revision accepted _____.

ABSTRACT

Herbarium collections offer an unparalleled archive of plant biodiversity, but their use for species identification through spectral data remains constrained by uncertain effects of preservation histories. This study assesses whether a single herbarium specimen can reliably represent its species based on its leaf reflectance spectrum, despite variations in age, geographic origin, or conservation method. We scanned herbarium specimens of different ages and geographic distribution of 14 species of the pantropical Annonaceae. In addition, we used a second dataset of 9 species where some specimens were conserved in alcohol prior to drying and some not. We used five supervised classification models frequently used for high-dimensional data such as spectroscopy. All models achieved high accuracy (>80%) when trained on multiple specimens per species. However, when using only one specimen per species, accuracy varied substantially depending on the taxon. Alcohol preservation prior to drying had little impact on classification success. Our findings demonstrate that herbarium specimens often retain a strong taxonomic signal in their spectra, however, inter-individual variability affects accuracy in some taxa. These findings confirm the usefulness of herbarium spectroscopy as a non-destructive tool for species identification and offer a promising avenue for digitizing historical biodiversity data into high-dimensional trait space.

Keywords

Spectroscopy; Herbarium specimen; Support Vector Machine; Linear Discriminant Analysis; Identification; Preservation

INTRODUCTION

Herbaria hold hundreds of millions of preserved plant specimens, forming one of the largest and oldest repositories of biodiversity worldwide (Davis, 2023). These collections span more than four centuries of global plant exploration and constitute irreplaceable archives of taxonomic, geographic, and temporal information. Originally designed for taxonomy and floristics, herbaria have long been the foundation of plant classification and nomenclature. Yet, in recent decades, these collections are increasingly being reinterpreted as data-rich resources for ecology, evolution, and conservation biology (Besnard et al., 2018; Raxworthy and Smith, 2021). Herbarium specimens collected decades or even centuries ago are now used in trait-based ecology, functional biogeography, phylogenetics, and global change research (Nualart et al., 2017; Lang et al., 2019; Rosche et al., 2025), revealing new value in what was once considered passive documentation.

This shift has been catalyzed by the development of analytical tools capable of extracting meaningful data from historical plant material (Raxworthy and Smith, 2021; Davis, 2023). Among these, leaf reflectance spectroscopy, particularly in the near-infrared (NIR), has emerged as a powerful and non-destructive technique to characterize plant functional traits (Cavender-Bares et al., 2017). Spectra provides a rapid way to capture the chemical and structural composition of leaves, based on how they absorb light across specific wavelengths, ranging from 800 to 2500 nm, extended to the visible wavelengths from 400 to 800 nm. Reflectance is influenced by biochemical constituents such as cellulose, lignin, water, phenolics, and proteins, as well as by structural characteristics like leaf thickness or the presence of surface features such as hairs and stomata (Jacquemoud and Ustin, 2019).

Spectra therefore have the potential to serve as a rapid non-desctructive tool for plant species identification, as they may contain taxonomically informative features, sometimes referred to as the Signature Spectral Signal (SSS) (Durgante et al., 2013). This taxonomic signal has led to a growing number of studies using reflectance spectra to discriminate between plant genera, species, and even populations. For example, Durgante et al. (2013) demonstrated that leaf reflectance spectra could successfully classify tropical tree species, while Meireles et al. (2020) revealed a phylogenetic component embedded in spectral variation at the angiosperm level. Other studies have confirmed the potential of NIRS for identifying hybrids, assessing intraspecific variability, and monitoring trait shifts across environments (Cavender-Bares et al., 2017; Stasinski et al., 2021).

Based on these observations, leaf reflectance spectroscopy has recently emerged as a promising tool for botanists, ecologists, and taxonomists alike. It offers a fast, and scalable solution for identifying specimens, managing biodiversity datasets, and even predicting functional traits (Kothari and Schweiger, 2022; Kothari et al., 2023). Furthermore, its non-invasive nature makes it particularly attractive for use on valuable or rare specimens, including historical herbarium material (types) including extinct species.

However, there is a critical caveat to this optimism. Most studies to date have relied on freshly collected, purpose-pressed leaves, prepared under relatively standardized and controlled conditions (referred to here as “pressed specimens”) (Durgante et al., 2013; Meireles et al., 2020; Kothari et al., 2023). Pressed specimens are not equivalent to herbarium specimens as they have not undergone mounting, long-term storage, transportation, or historical treatment. Herbarium specimens have often been exposed to a wide variety of in-field drying methods, conservation practices, and chemical treatments, many of which are undocumented, and all of which can potentially affect the chemical integrity of the leaves. The effects of these processes

on leaf chemistry (and by extension, on spectral properties) are yet not well understood. This represents a fundamental limitation to the generalizability of existing findings.

For example, in humid tropical environments, alcohol is frequently used as a temporary preservative before heat drying, particularly during multi-day field expeditions (Hodge, 1947). While effective at preventing mold and tissue decay, alcohol is a strong solvent and oxidizer that can also impact DNA extraction (Forrest et al., 2019). Even today, practices such as applying glue, or drying methods vary considerably between herbaria and collectors. As a potential source of chemical modification, all these factors can introduce noise or bias into spectral data collected from herbarium sheets.

Moreover, the geographic origin of specimens introduces another layer of variation. Climate, elevation, soil chemistry, and other environmental parameters affect leaf chemistry (Maire et al., 2015; Asner and Martin, 2016; Cao et al., 2020), meaning that two specimens of the same species, preserved differently and collected in distinct locations, may display significant spectral divergence. If spectroscopy is to be used reliably for taxonomic or ecological inference on herbarium material, such sources of variation must be understood and ideally accounted for in the analytical framework.

Despite these challenges, recent evidence suggests that herbarium specimens still retain valuable chemical and even strong taxonomic information (Neto-Bradley et al., 2025; White et al., 2025b). For instance, Foutami et al. (2018) have shown that the age of a specimen does not drastically affect some part of its chemical profile, even after decades of storage. While certain surface properties or volatile compounds may degrade over time, the core chemical constituents that drive spectral variation often remain stable (Mendes Resende et al., 2020). The key challenge is not whether herbarium spectra contain useful information, but whether this information is consistent, reproducible, and interpretable enough to enable species-level

identification and classification, while remaining feasible at low sampling scales given the frequent scarcity of specimens per species and the logistical demands of projects involving hundreds or thousands of samples.

Here, we investigate several key questions regarding the application of spectroscopy to herbarium specimens. Specifically: (i) Can spectral data reliably support species-level identification? (ii) How many specimens per species are required to adequately represent intra-specific spectral variability? (iii) Are standard machine learning algorithms sufficiently robust to handle the heterogeneity typical of herbarium collections? and (iv) To what extent do pre-drying alcohol treatments affect the spectral signal used for classification? By addressing these questions, we aim to assess both the practical applicability and methodological limitations of herbarium spectroscopy under realistic working conditions.

To address the above questions, we focused on the pantropical plant family Annonaceae. With around 2500 species, it is one of the best-known tropical plant families in terms of taxonomy with global occurrence data and phylogenetic resources available (Couvreur et al., 2012, 2019; Nge et al., 2024). Despite this taxonomic maturity, species identification in Annonaceae can be challenging, particularly when reproductive structures (flowers or fruits) are absent or unavailable. This makes Annonaceae an ideal case study for evaluating the robustness of spectral approaches for species discrimination. Additionally, the family displays a degree of ecological and biogeographic amplitude within the tropics, spanning environments from Amazonia to African rain forests and Southeast Asian regions, thus offering a natural gradient of variation in both environmental conditions and specimen collection practices (Pennington et al., 2004; Couvreur et al., 2011; Chatrou et al., 2012).

METHODS

Taxon sampling

We generated two datasets. First, to address the objective of species identification using herbarium-generated spectra, we scanned leaves from 14 Annonaceae species (**Table 1**) all stored in the herbarium of the Muséum national d'Histoire naturelle in Paris (P) (Le Bras et al., 2017). For each species we scanned a minimum of 9 specimens. The selected species were: 1) chosen from recently revised genera, ensuring expert-based identification; 2) represent different growth habits (liana, understory and canopy trees); 3) have a broad geographical distribution, including species from South America, Africa, Madagascar, and Southeast Asia; 4) have diverse geographical representation within species, including specimens from various countries and ecotypes when possible; and 5) are phylogenetically diverse across Annonaceae (at least one species for each of the four subfamilies, Nge et al. 2024). We also included two species pairs belonging to the same genera in *Ambavia* and *Hexalobus* (**Table 1**). In general, we prioritized flowering or fruiting specimens as these are better identified. This dataset is referred to as ID_PARIS.

A second independent dataset was generated to examine the impact of alcohol preservation. For this we sampled nine species (**Table 2**) within the 50-ha Yasuni Forest Dynamics Plot (YFDP) located next to the Yasuní Research Station of the Pontificia Universidad Católica del Ecuador (Valencia et al., 2004). These species are the same as the study of Santacruz Endara et al. (2025) and cover the phylogenetic diversity of the family. First, specimens were sampled directly in the field at the Yasuni Research station of the Pontificia Universidad Católica del Ecuador (PUCE). Leaves were then pressed in newspaper and dried over an electric stove for 24 hours. For one or two specimens per species they were conserved in 70-80% alcohol for three days, and then dried in an electric stove for 24 hours. This dataset is referred to as ID_YASUNÍ.

To place the two datasets into context, **Figure 1** summarizes the overall workflow, from Specimen selection and scanning procedure acquisition to preprocessing and analytical steps, providing a framework for the methodological details that follow.

Spectral Measurements

Leaf specimens were scanned following a general protocol, with dataset-specific adaptations (**Table 3**). In both datasets (ID_PARIS and ID_YASUNI), spectra were collected in the 350–2500 nm wavelength range, on the adaxial side of mature leaves under controlled environmental conditions. Young or unfolding leaves were avoided, as well as the midrib region, and areas affected by fungi, bryophytes, or glue. Whenever possible, multiple leaves per specimen were scanned, and measurements were distributed across the cleanest and least damaged areas. Care was taken to ensure contact between the probe and the leaf without applying excessive pressure that could damage the specimen.

Data Processing

All data processing and analyses were conducted using R software (v4.4.1). The processing of spectral data ensures consistency, removes noise, and enhances relevant signals for further analysis and is composed of two steps. The first step involved applying a Standard Normal Variate (SNV) normalization (Barnes et al., 1989). For each spectrum, the mean and standard deviation of the 20 readings are calculated. Each reflectance intensity value is then normalized by subtracting the mean and dividing by the standard deviation, resulting in spectra with a mean of zero and unit variance. This standardization method removes a part of noise caused by external factors such as instrument settings, distance between the probe and the sample, or sample thickness. After normalization, the Savitzky-Golay filter (Savitzky and Golay, 1964) was used to compute the first derivative of the spectra. A polynomial order of 2 and a window size of 7 were used to smooth the data while capturing the rate of change at each wavelength. The first derivative accounts for variations in the slope of the spectra.

Analytical Approaches

Five classification models were implemented to evaluate taxonomic discrimination: Partial Least Squares Discriminant Analysis (PLS-DA), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Linear Discriminant Analysis (LDA), and Random Forest (RF). All models were trained and tested using stratified resampling with a strict separation by specimen identity to avoid data leakage. For PLS-DA, models were fitted using the plsda() function from the mixOmics package (Rohart et al., 2017). The number of components was optimized within each iteration by cross-validation on a dedicated tuning subset of the ID_PARIS dataset, testing values between 2 and 25. The optimal number of components was selected as the one minimizing cross-validated error rates. The SVM model was implemented using the caret interface to the e1071 and kernlab packages. Hyperparameter tuning was performed on the ID_PARIS dataset via cross-validation on a tuning subset, with a grid of sigma values (0.001, 0.01, 0.1) and cost values (0.1, 1, 10). The best combination of sigma and C was retained for final calibration and testing. All predictor variables were standardized based on calibration data. For KNN, the knn3() function from the caret package was used. The number of neighbors was tuned on the ID_PARIS dataset using cross-validation over k values from 1 to 10. Predictors were standardized according to calibration data, and the same transformation was applied to tuning and test sets. LDA was performed using the lda() function from the MASS package. As LDA assumes uncorrelated predictors, a principal component analysis (PCA) was applied to the calibration data. Components explaining at least 95% of the total variance were retained, and the same transformation was applied to tuning and test sets before fitting the LDA model. For Random Forest, the randomForest() function from the randomForest package was used with 500 trees (ntree = 500). All predictors with non-zero variance were included, and no additional

feature selection or scaling was performed. Classification was based on majority vote across the ensemble of trees.

The models were applied to two distinct datasets: ID_PARIS, which was used for model training and testing with herbarium data, and ID_YASUNI, which was used to assess the effect of alcohol preservation on species discrimination. For ID_PARIS, the dataset was split into tuning, calibration, and prediction subsets, enabling full hyperparameter optimization. In contrast, analyses on ID_YASUNI and additional analyses on ID_PARIS at varying calibration sizes were conducted in a robustness-testing framework, where fixed hyperparameters were reused across all conditions. This ensured comparability across datasets and specimen sampling levels.

To simulate real-world conditions where the preservation method is generally unknown, alcohol-preserved specimens were randomly included in either the training or test dataset. This allowed us to evaluate how alcohol preservation might affect the models' ability to distinguish between species. For each iteration, classification accuracy was calculated, and confusion matrices were pooled across resamples to evaluate model performance. To avoid bias, species were equally represented in all calibration datasets, irrespective of specimen availability. This balanced sampling ensured that species with more specimens did not dominate the models. Model predictions were compared to expert identifications to assess accuracy. Overall performance was computed for each model, and confusion matrices were used to highlight misclassification trends. We tested the models using both all available spectra per specimen (twenty spectra) and their average. To further explore the limitations of using a single herbarium specimen to represent a species, we conducted an analysis using the ID_PARIS dataset. We first evaluated model performance with only one randomly selected specimen per species, using its twenty spectra for training and the spectra of all remaining specimens for testing. We then

extended the analysis by progressively increasing the number of training specimens from one to eight, using all their spectra for training and the spectra of the remaining specimens for testing.

Finally, to ensure robustness and consistency, all measurements were repeated 100 times for each analysis. This repetitive testing helped to evaluate the stability of the models' performance across different sampling conditions and ensured reliable results.

RESULTS

Specimen spectrum sampling

A total of fourteen species were scanned in the P herbarium. Using the previously described workflow (Fig. 1), we were able to scan up to 18 specimens for some species. Preparing a specimen, entering all parameters, and collecting the 20 spectra takes about ten minutes. While preparing a specimen, entering all parameters, and beginning the scan takes nearly 8 minutes, an additional 10 scans require less than 2 minutes. This enabled us to scan 160 specimens over 5 days (see Table 1 & Table S1 for details). The scanned specimens ranged in age from 9 to 201 years, with an overall average of 69 years. Some species include older specimens (e.g., Annona senegalensis: mean 107 years, SD = 23; Uvaria grandiflora: mean 89 years, SD = 53; Anaxagorea dolichocarpa: mean 88 years, SD = 64), while others include younger ones (e.g., Xylopia humblotiana: mean 30 years, SD = 22; Ambavia spp.: mean 33 years, SD = 15) (Fig. S1). Several issues arose during sampling due to specific conditions encountered with some specimens, such as glue or highly curled leaves. Among the 160 specimens, less than five were affected by the scanning process, and the impacts were never severe. The two types of minor destructive effects observed were leaf tears on curled specimens and leaf detachment. In the case of the samples from Yasuní, the scanning workflow was much more efficient because the data had already been collected and the leaves were not fixed in an herbarium voucher. Despite

taking more scans per specimen (50 instead of 20), it took less than five minutes to measure a whole specimen.

Spectrum discrimination using herbarium specimen

For all five classification models tested, we observed an overall strong capacity to discriminate herbarium specimens at the species level. Using all spectra from six specimens per species for training and all spectra from the remaining specimens for testing, we obtained the following overall accuracies: 81.45% (SD = 4.50%) for SVM, 78.57% (SD = 4.77%) for PLS-DA, 75.63% (SD = 4.08%) for KNN, 84.74% (SD = 5.23%) for LDA, and 81.06% (SD = 4.47%) for RF (**Fig. 2A**). When each specimen in the test set was represented by a single averaged spectrum rather than the full set of 20 individual spectra, the discrimination capacity where even higher for all models: 86.35% (SD = 4.47%) for SVM, 83.03% (SD = 5.82%) for PLS-DA, 79.23% (SD = 4.86%) for KNN, 87.48% (SD = 5.57%) for LDA, and 84.96% (SD = 5.60%) for RF (**Fig. 2A**). This remained true independent of species diversity in terms of geographic origin or collection date (**Fig. 2B**). However, two species from the same genus, *Hexalobus crispiflorus* and *H. monopetalus*, appeared to be less accurately classified, with misclassifications from one to the other. In most models using individual spectra as the test, there was a tendency for spectra to be classified as the other species, with a slight bias toward misclassification in favor of *H. crispiflorus*. When using the average spectrum per specimen in the test set, this pattern persisted, where *H. monopetalus* was more frequently predicted as *H. crispiflorus* than the reverse (**Fig. 2B**). This trend was especially pronounced in LDA, where nearly all *H. crispiflorus* specimens were correctly classified (79.50%) but only 66.67% of *H. monopetalus,* the misclassified ones were all predicted as *H. crispiflorus*. The two *Ambavia* species showed a better discrimination rate without a distinct four-square misclassification pattern. However, there was a slight but consistent tendency to misclassify *A.capuronii* as *A. gerrardii*. *Uvaria*

grandiflora also appeared to be less well classified, where its correct classification rate was among the lowest observed. Overall, despite a general increase in model accuracy when using the average spectrum, the same species tended to be misclassified across both approaches and generally in the same predicted species (**Fig. 2B**).

Alcohol treatment effect

All models showed an accuracy above 90% when every spectrum was included in the test set, and no model accuracy was below 95% when the species average spectrum was used in the test set. Three models (SVM, PLS-DA and LDA) reached classification rates above 99%, with very low standard deviation (SVM and PLS-DA <0.005; LDA <0.02 (Fig. 2A)) , indicating nearly all specimens were correctly classified to their species, regardless of whether alcohol-treated specimens were included in the training or test datasets. We did not observe a strong genus-level misclassification pattern, similar to what was seen in Hexalobus, among the two Guatteria species included in this dataset especially with the average specimen in the test dataset (Fig. 2B). Additionally, we found that alcohol-preserved specimens do not appear to be distant in the 3D space defined by the first three PCA components of the average spectrum for each specimen (Fig. S2A). These specimens cluster closely with others from the same species and are never misclassified into the spectral space of a different species . The average raw spectra themselves also generally share the same overall shape as those of other non-alcohol-preserved specimens from the same species (Fig. S2B).

One specimen to represent them all

We observed substantial variability in classification accuracy depending on the species when training the models with only a single specimen but using its 20 spectra (**Fig. 3**). For some species, such as *Anaxagorea dolichocarpa*, *Duguettia calycina* or *Ambavia gerrardii*, accuracy remained high more than 80% across all models (Mean values *A. dolichocarpa* 85.22; *A.*

gerrardii 82.77; *D. calycina* 84.64 (**Fig. 3**)), while for others like *Uvaria grandiflora* or both *Hexalobus* species, it dropped below 50% (Mean values *H. monopetalus* 49.33; *H. crispiflorus* 43.47; *U. grandiflora* 37.64 (**Fig. 3**)). Despite this variability, model performance tended to follow similar patterns: species that were generally well classified remained so across models, and the same held for poorly classified species. Nonetheless, LDA generally outperformed the other methods, showing the highest classification accuracy in 12 out of 14 species. In certain cases, this performance gap was particularly marked (e.g. *Monanthotaxis enghiana*).

With an average accuracy of 65.4% (**Fig. 4A**), the models showed an important improvement, reaching 78.2% with only two specimens used for training, and stabilizing near 90% from five specimens onward, with only slight increases beyond that point (**Fig. 4B**). LDA consistently outperformed the other models at all specimen counts, achieving over 70% accuracy with just one specimen and reaching more than 90% from three specimens onward.

DISCUSSION

Our study demonstrates that herbarium leaf spectra contain sufficient information to enable accurate species classification, even decades after specimen collection and at small sampling sizes. Across most taxa tested, classification performance remained consistently high (**Fig. 2A**), supporting the idea that herbarium material can preserve relevant traits for classification (Cavender-Bares et al., 2017; Meireles et al., 2020; Kothari et al., 2023; Neto-Bradley et al., 2025; White et al., 2025b). Importantly, high accuracy (often >90%) was achieved using as few as five specimens per species for model training (**Fig. 4**). Even under single-specimen conditions, performance far exceeded random expectations (**Fig. 3**), indicating that the spectral signature (Durgante et al., 2013) of many species is both coherent and stable. In addition, we show that herbarium specimens conserved in alcohol prior to drying did not compromise classification accuracy. Even when included in training or test sets, models consistently

exceeded 95% accuracy, especially with SVM, LDA, and PLS-DA (**Fig. 2A**). Despite slight spectral shifts due to alcohol (**Fig. S1B**), the key taxonomic signal was preserved, with treated and untreated samples occupying overlapping species-specific space (**Fig. S1A**). This supports the broader inclusion of specimens preserved in alcohol prior to drying, a common practice especially in tropical collections. Our results thus highlight the untapped potential of historical collections in modern spectral applications and open promising avenues for species identification in data-limited contexts.

Classification success was, however, not universal. Some species, such as *Uvaria grandiflora* and the two *Hexalobus* species, presented consistent challenges across models. These difficulties may stem from greater intraspecific variability, specimen age, or unobserved artefacts rather than taxonomic misidentification, since all specimens were vetted by taxonomic experts. One striking example was the confusion between *Hexalobus crispiflorus* and *H. monopetalus* (**Fig. 2B**), two morphologically, ecologically and phylogenetically distinct African species that were difficult to separate spectrally (Botermans et al., 2011; Dagallier et al., 2023). While *H. crispiflorus* occurs in rain forest and *H. monopetalus* in savanna environments, their spectra did not reflect this divergence. Such cases suggest that while reflectance spectroscopy is generally effective, it may occasionally miss biologically meaningful differences, especially in species groups with broad geographic distributions or complex ecological variation. In contrast, *Greenwayodendron suaveolens*, a widespread generalist tree species in the Congo Basin (Lissambou et al., 2018), was robustly classified despite spanning multiple countries and time periods (**Fig. 3**), highlighting the potential of spectroscopy to detect stable traits even under wide natural variation.

We evaluated five supervised classification models (PLS-DA, LDA, RF, KNN and SVM) each chosen for its known capacity to handle high-dimensional spectral data (Kong et al., 2013;

Trainor et al., 2017; Jombo et al., 2020; Wang et al., 2020). Our results indicate that all models can achieve strong performance, but all not are equally robust or scalable (**Fig. 2-3-4**). LDA emerged as the most accurate method in this study, likely benefiting from the relatively clean structure of the dataset and the use of prior dimensionality reduction. Its strength lies in interpretability, making it a practical choice for small to mid-sized studies (Khondoker et al., 2016). SVM also performed very well leading to consistently high performance aligned with its mathematical strength in high-dimensional classification problems (Khondoker et al., 2016), especially where the decision boundaries between classes are not linearly separable. Using kernel functions such as the radial basis function (RBF), SVM effectively transforms the spectral space to identify optimal hyperplanes that maximize the margin between species classes. This ability to handle complex boundaries with relatively few tuning parameters makes SVM particularly suitable for herbarium-based datasets that may include dozens of species and thousands of spectral variables. In contrast, RF and KNN showed greater variability (**Fig 2-3**). While RF benefits from ensemble averaging, it may dilute fine-scale taxonomic signals, and KNN's reliance on local distance metrics can amplify noise in high-dimensional space (Zdunek et al., 2016). PLS-DA produced good results and remains useful, but it is mathematically less suited to classification when the number of classes increases, as it does not optimize for class separation in the same way as SVM (Trainor et al., 2017). These differences highlight the need to match model choice to dataset complexity and classification goals (e.g. in some cases PLS-DA might be the most suitable model (Richter et al., 2016)). We also explored how model performance evolved with the number of specimens used for training. As expected, accuracy improved with larger training sets, but gains diminished beyond five specimens (**Fig. 4**). This result is particularly encouraging for biodiversity-rich but specimen-poor groups, where limited material can still yield reliable classification.

379 While these results are highly encouraging, our study was conducted under controlled
380 conditions with specimens from a single herbarium (P). One limitation is that we did not isolate
381 or test the effects of specimen age or geographic origin independently. This choice was
382 intentional: in real herbarium settings, such metadata are sometimes missing, imprecise, or
383 impossible to control for. Our goal was to assess model performance under realistic conditions
384 that reflect the actual state of herbarium collections. In many cases, we do not know precisely
385 how or when a specimen was collected, and thus our approach tested models on a wide range
386 of real-life variability. This adds complexity; however, it ensures that our results reflect the
387 types of constraints commonly encountered when working with herbarium collections. In
388 practice, researchers rarely have full control over specimen metadata, and working with
389 whatever material is available is often necessary. By testing our models across realistic
390 conditions, we aimed to produce results that are not only scientifically valid, but also
391 operationally useful. This underlines a broader question: how generalizable are these findings
392 to other institutions, especially those with less controlled storage environments or differing
393 preparation protocols? Answering this will be essential for the integration of spectroscopy into
394 global workflows. This includes addressing variability not only in storage conditions but also
395 in specimen preparation protocols, which are rarely standardized across institutions. Pre-drying
396 preservation methods (like those involving alcohol) are highly variable and generally not well
397 documented. Although our targeted alcohol experiment showed minimal impact on
398 classification (**Fig. 2**), this may not hold for specimens exposed to prolonged or irregular
399 treatments. Reproducibility is also a key consideration. We found that averaging replicate
400 spectra improved model performance (**Fig. 2**), but this approach may obscure fine-scale
401 biological variation. Furthermore, artefacts introduced by glue, fungal contamination, or leaf
402 structure can affect measurements in unpredictable ways. While data preprocessing techniques
403 like SNV and Savitzky–Golay filtering help mitigate noise, they do not entirely eliminate it,

and the boundary between noise and true biological variation remains complex. In light of these challenges, harmonizing protocols across studies is essential for building interoperable datasets. The HerbSpec initiative (Cavender-Bares et al., 2025) provides a strong foundation, proposing a protocol for specimen preparation, scanning environment, and metadata standardization (White et al., 2025a). Adopting such protocols will ensure that spectral datasets are not only locally robust, but globally consistent and reusable, laying the groundwork for the future of large-scale, collaborative spectral biodiversity monitoring.

Conclusion and perspectives

The application of near infrared spectroscopy to herbarium specimens is still in its infancy. Our results show that even old specimens retain enough taxonomic signal in their leaf reflectance spectra to allow accurate species classification using standard machine learning approaches. This confirms spectral scanning as a valuable, non-destructive source of data for studies in taxonomy, ecology, and evolution (Meireles et al., 2020; Stasinski et al., 2021; Kothari and Schweiger, 2022).

The potential of spectroscopy applied to herbarium-preserved specimens opens exciting perspectives for plant biodiversity research, adding yet another important dimension to the use of herbaria (Davis 2023). In addition to complementing traditional taxonomic research, spectroscopy could also inform leaf functional traits and, in combination with advances in machine learning (Soltis et al., 2020), enable large-scale, non-destructive trait estimation, species identification, and even the detection of cryptic diversity across historical collections. Ensuring data quality and expanding spectral libraries across taxa and regions will be key to unlocking the full potential of this approach.

CRediT Author contributions:

KB: Conceptualization, Data curation, Formal analysis, Writing (original draft). **TLPC**: Conceptualization, Supervision, Methodology, Validation, Writing (review & editing), Investigation (field sampling), Funding acquisition. **PGSA**: Methodology, Investigation (field sampling), Writing (review) **LAB**: Methodology, Writing (review & editing). **ME**: Methodology, Writing (review & editing). **SB**: Methodology (instrumentation), Writing (review). **DB**: Methodology (instrumentation), Writing (review). **LB**: Methodology (instrumentation), Writing (review).

Acknowledgment

The authors thank the Muséum national d'Histoire naturelle (MNHN), Paris, for access to its herbarium collections, and the Yasuní National Park authorities for permitting field sampling and supporting research. We are grateful to Dr. Laura Holzmeyer for taxonomy training prior to specimen selection at P. This research was supported by the European Research Council (ERC) under the European Union's Horizon 2020 program (grant agreement No. 865787).

Data - Code availability

All code and data used in this study are available from the IRD Forge repository : https://forge.ird.fr/diade/global_erc/boughalmi_nirs_paris_yasuni_data_2025

REFERENCES

Asner, G. P., and R. E. Martin. 2016. Convergent elevation trends in canopy chemical traits of tropical forests. *Global Change Biology* 22: 2216–2227.

Barnes, R. J., M. S. Dhanoa, and S. J. Lister. 1989. Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy* 43: 772–777.

Besnard, G., M. Gaudeul , S. Lavergne, S. Muller, G. Rouhan , A. P. Sukhorukov, A. Vanderpoorten , and F. Jabbour. 2018. Herbarium-based science in the twenty-first century. *Botany Letters* 165: 323–327.

Botermans, M., M. S. M. Sosef, L. W. Chatrou, and T. L. P. Couvreur. 2011. Revision of the African Genus Hexalobus (Annonaceae). *Systematic Botany* 36: 33–48.

Cao, J., X. Wang, J. F. Adamowski, A. Biswas, C. Liu, Z. Chang, and Q. Feng. 2020. Response of leaf stoichiometry of *Oxytropis ochrocephala* to elevation and slope aspect. *CATENA* 194: 104772.

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: 966–969.

Cavender-Bares, J., D. M. White, N. I. Ahlstrand, M. Austin, D. Bastianelli, S. Bazan, K. Boughalmi, et al. 2025. NextGeneration specimen digitization: The international herbarium community goes spectral!

Chatrou, L. W., R. H. J. Erkens, J. E. Richardson, R. M. K. Saunders, and M. F. Fay. 2012. The natural history of Annonaceae: NATURAL HISTORY OF ANNONACEAE. *Botanical Journal of the Linnean Society* 169: 1–4.

Couvreur, T. L. P., A. J. Helmstetter, E. J. M. Koenen, K. Bethune, R. D. Brandão, S. A. Little, H. Sauquet, and R. H. J. Erkens. 2019. Phylogenomics of the Major Tropical Plant Family Annonaceae Using Targeted Enrichment of Nuclear Genes. *Frontiers in Plant Science* 9: 1941.

Couvreur, T. L. P., P. J. M. Maas, S. Meinke, D. M. Johnson, and P. J. A. KEßLER. 2012. Keys to the genera of Annonaceae: ANNONACEAE GENERIC KEYS. *Botanical Journal of the Linnean Society* 169: 74–83.

Couvreur, T. L. P., M. D. Pirie, L. W. Chatrou, R. M. K. Saunders, Y. C. F. Su, J. E. Richardson, and R. H. J. Erkens. 2011. Early evolutionary history of the flowering plant family Annonaceae: steady diversification and boreotropical geodispersal. *Journal of Biogeography* 38: 664–680.

Dagallier, L.-P. M. J., F. M. Mbago, M. Couderc, M. Gaudeul, A. Grall, C. Loup, J. J. Wieringa, et al. 2023. Phylogenomic inference of the African tribe Monodoreae (Annonaceae) and taxonomic revision of Dennettia, Uvariodendron and Uvariopsis. *PhytoKeys* 233: 1–200.

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

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.

Forrest, L. L., M. L. Hart, M. Hughes, H. P. Wilson, K.-F. Chung, Y.-H. Tseng, and C. A. Kidner. 2019. The Limits of Hyb-Seq for Herbarium Specimens: Impact of Preservation Techniques. *Frontiers in Ecology and Evolution* 7: 439.

Foutami, I. J., T. Mariager, R. Rinnan, C. J. Barnes, and N. Rønsted. 2018. Hundred Fifty Years of Herbarium Collections Provide a Reliable Resource of Volatile Terpenoid Profiles Showing Strong Species Effect in Four Medicinal Species of Salvia Across the Mediterranean. *Frontiers in Plant Science* 9.

Hodge, W. H. 1947. The Use of Alcohol in Plant Collecting. *Rhodora* 49: 207–210.

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

Jombo, S., E. Adam, M. J. Byrne, K. A. Ali, and S. W. Newete. 2020. Exploring the Potential of Feature Selection Methods in the Classification of Urban Trees Using Field Spectroscopy Data.

Khondoker, M., R. Dobson, C. Skirrow, A. Simmons, and D. Stahl. 2016. A comparison of machine learning methods for classification using simulation with multiple real data examples from mental health studies. *Statistical Methods in Medical Research* 25: 1804–1823.

Kong, W., C. Zhang, F. Liu, P. Nie, and Y. He. 2013. Rice Seed Cultivar Identification Using Near-Infrared Hyperspectral Imaging and Multivariate Data Analysis. *Sensors* 13: 8916–8927.

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.

Lang, P. L. M., F. M. Willems, J. F. Scheepens, H. A. Burbano, and O. Bossdorf. 2019. Using herbaria to study global environmental change. *New Phytologist* 221: 110–122.

Le Bras, G., M. Pignal, M. L. Jeanson, S. Muller, C. Aupic, B. Carré, G. Flament, et al. 2017. The French Muséum national d'histoire naturelle vascular plant herbarium collection dataset. *Scientific Data* 4: 170016.

Lissambou, B.-J., O. J. Hardy, C. Atteke, T. Stevart, G. Dauby, B. Mbatchi, B. Sonke, and T. L. P. Couvreur. 2018. Taxonomic revision of the African genus Greenwayodendron (Annonaceae). *PhytoKeys* 114: 55–93.

Maire, V., I. J. Wright, I. C. Prentice, N. H. Batjes, R. Bhaskar, P. M. Van Bodegom, W. K. Cornwell, et al. 2015. Global effects of soil and climate on leaf photosynthetic traits and rates. *Global Ecology and Biogeography* 24: 706–717.

Meireles, J. E., J. Cavender-Bares, P. A. Townsend, S. Ustin, J. A. Gamon, A. K. Schweiger, M. E. Schaepman, et al. 2020. Leaf reflectance spectra capture the evolutionary history of seed plants. *New Phytologist* 228: 485–493.

Mendes Resende, J. V., N. M. D. de Sá, M. T. L. de Oliveira, R. C. Lopes, R. Garrett, and R. Moreira Borges. 2020. Chemical profiling of herbarium samples of solanum (Solanaceae) using mass spectrometry. *Phytochemistry Letters* 36: 99–105.

Neto-Bradley, B. M., P. Bonnet, H. Goëau, A. Joly, J. Cavender-Bares, and D. A. Coomes. 2025. Using reflectance spectra and Pl@ntnet to identify herbarium specimens: a case study with Lithocarpus. *New Phytologist*.

Nge, F. J., T. Chaowasku, A. Damthongdee, C. Wiya, V. R. C. Soulé, C. Rodrigues-Vaz, D. Bruy, et al. 2024. Complete genus-level phylogenomics and new subtribal classification of the pantropical plant family Annonaceae. *TAXON* 73: 1341–1369.

Nualart, N., N. Ibáñez, I. Soriano, and J. López-Pujol. 2017. Assessing the Relevance of Herbarium Collections as Tools for Conservation Biology. *The Botanical Review* 83: 303–325.

Pennington, P. T., Q. C. B. Cronk, J. A. Richardson, J. E. Richardson, L. W. Chatrou, J. B. Mols, R. H. J. Erkens, and M. D. Pirie. 2004. Historical biogeography of two cosmopolitan families of flowering plants: Annonaceae and Rhamnaceae. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 359: 1495–1508.

Raxworthy, C. J., and B. T. Smith. 2021. Mining museums for historical DNA: advances and challenges in museomics. *Trends in Ecology & Evolution* 36: 1049–1060.

Richter, R., B. Reu, C. Wirth, D. Doktor, and M. Vohland. 2016. The use of airborne hyperspectral data for tree species classification in a species-rich Central European forest area. *International Journal of Applied Earth Observation and Geoinformation* 52: 464–474.

Rohart, F., B. Gautier, A. Singh, and K.-A. L. Cao. 2017. mixOmics: An R package for ‘omics feature selection and multiple data integration. *PLOS Computational Biology* 13: e1005752.

Rosche, C., O. Broennimann, A. Novikov, V. Mrázová, G. V. Boiko, J. Danihelka, M. T. Gastner, et al. 2025. Herbarium specimens reveal a cryptic invasion of polyploid Centaurea stoebe in Europe. *New Phytologist* 245: 392–405.

Santacruz Endara, P. G., A. Lozinguez, R. Valencia, S. A. Queenborough, M. Endara, B. Obando-Tello, and T. L. P. Couvreur. 2025. Discreet but diverse and specific: Determining plant-herbivore interactions across a species-rich plant family in a tropical rain forest. *Biotropica* 57: e13392.

Savitzky, Abraham., and M. J. E. Golay. 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry* 36: 1627–1639.

Soltis, P. S., G. Nelson, A. Zare, and E. K. Meineke. 2020. Plants meet machines: Prospects in machine learning for plant biology. *Applications in Plant Sciences* 8: e11371.

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* 232: 2283–2294.

Trainor, P. J., A. P. DeFilippis, and S. N. Rai. 2017. Evaluation of Classifier Performance for Multiclass Phenotype Discrimination in Untargeted Metabolomics. *Metabolites* 7: 30.

Valencia, R., R. B. Foster, G. Villa, R. Condit, J.-C. Svenning, C. Hernández, K. Romoleroux, et al. 2004. Tree species distributions and local habitat variation in the Amazon: large forest plot in eastern Ecuador. *Journal of Ecology* 92: 214–229.

Wang, Q., X. Cui, G. Teng, Y. Zhao, and K. Wei. 2020. Evaluation and improvement of model robustness for plastics samples classification by laser-induced breakdown spectroscopy. *Optics & Laser Technology* 125: 106035.

White, D., N. I. Ahlstrand, M. Austin, D. Bastianelli, S. Bazan, K. Boughalmi, W. Cardinal-McTeague, et al. 2025a. IHerbSpec/IHerbSpec-protocol: IHerbSpec Protocol v1.0.

White, D. M., J. Cavender-Bares, C. C. Davis, J. A. Guzmán Q., S. Kothari, J. M. Robles, and J. E. Meireles. 2025b. Seeing herbaria in a new light: leaf reflectance spectroscopy unlocks trait and classification modeling in plant biodiversity collections. *New Phytologist*.

Zdunek, R., M. Nowak, and E. Plínski. 2016. Statistical classification of soft solder alloys by laser-induced breakdown spectroscopy: review of methods. *Journal of the European Optical Society-Rapid Publications* 11: 16006i.

FIGURES/TABLES

TABLE 1. Paris herbarium specimens and species Information – ID_PARIS.

Species	Specimens scanned	Geographic origin	Specimen age span	Subfamily
Ambavia capuronii (Cavaco & Keraudren) Le Thomas	17	Madagascar	1953-2013	Ambavioideae
Ambavia gerrardi (Baill.) Le Thomas	18	Madagascar	1956-2010	Ambavioideae
Anaxagorea dolichocarpa Sprague & Sandwith	10	French Guiana	1845 - 2008	Anaxagoreoideae
Annona senegalensis Pers.	9	Zambia, Tanganyika, Mali, Sierra leone, Guinea	1893 - 1951	Annonoideae
Anonidium mannii (Oliv.) Engl. & Diels	11	Cameroon	1918 - 1995	Annonoideae
Duguetia calycina Benoist	12	French Guiana	1857 - 2003	Annonoideae
Greenwayodendron suaveolens (Engl. & Diels) Verdc.	11	Gabon, DR Congo, Central African Republic, Cameroon	1881 - 1991	Malmeoideae
Hexalobus crispiflorus A.Rich.	10	Congo republic, Guinea, Cameroon, Gabon, Senegal	1838 - 1988	Annonoideae
Hexalobus monopetalus (A.Rich.) Engl. & Diels	11	Mali, Cameroon, Central African Republic, Senegal, Niger, Ivory coast	1929 - 2007	Annonoideae
Monanthotaxis enghiana (Diels) P.H.Hoekstra	9	Cameroon, Gabon, Central African Republic	1948 - 1989	Annonoideae
Unonopsis guatterioides (A.DC.) R.E.Fr.	12	Brasil, Paraguay, Saint Vincent and the Grenadines, Surinam, Bolivia	1886 - 1979	Malmeoideae
Uvaria grandiflora Roxb. ex Hornem.	11	Vietnam, Malaysia, China, Indonesia	1823 - 2015	Annonoideae
Uvariastrum pierreanum Engl.	10	Gabon, Cameroon, Ivory c oast, Congo republic, Liberia, Nigeria	1948 - 2013	Annonoideae
Xylopia humblotiana Baill.	9	Madagascar	1924 - 2007	Annonoideae

TABLE 2: Yasuní National Park specimens count – ID_YASUNÍ.

Species	**Specimen count**	**Alcohol-treated specimen count**
Anaxagorea brevipes Benth.	7	2
Duguetia hadrantha (Diels) R.E.Fr.	6	1
Guatteria guianensis (Aubl.) R.E.Fr.	6	1
Guatteria scalarinervia D.R.Simpson	5	1
Oxandra riedeliana R.E.Fr.	5	1
Tetrameranthus globuliferus Westra	7	2
Trigynaea triplinervis D.M.Johnson & N.A.Murray	6	1
Unonopsis veneficiorum (Mart..) R.E.Fr.	7	2
Xylopia cuspidata Diels	6	1

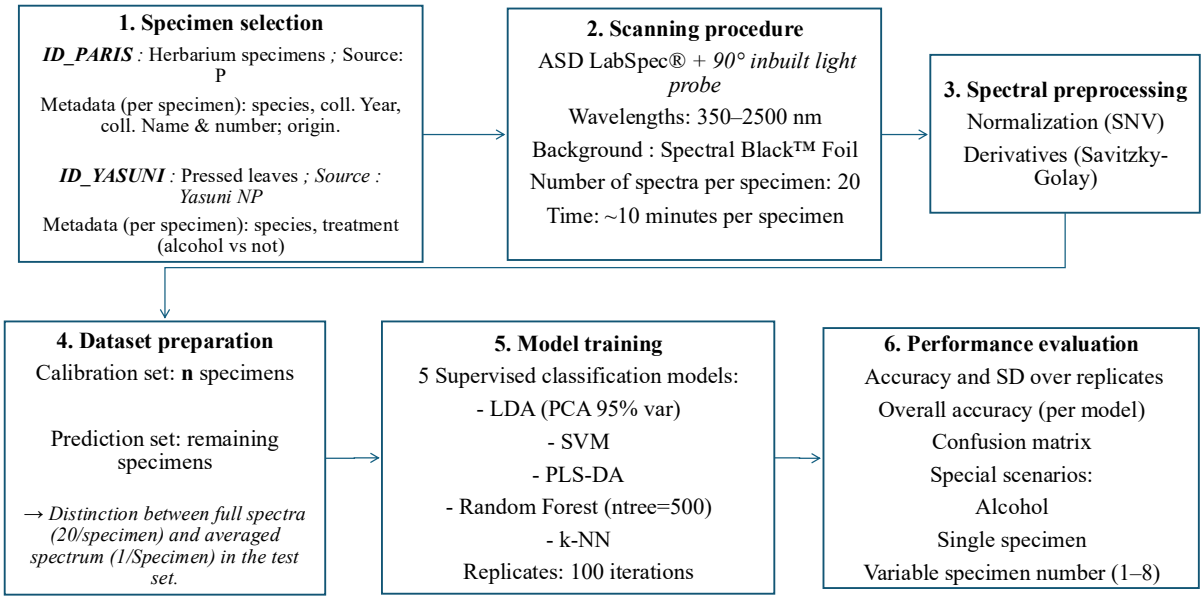


Figure 1. Workflow illustrating the classification pipeline used. Specimens from two datasets were scanned using spectrometers (ASD LabSpec), collecting 20 individual spectra per specimen. Spectral data was preprocessed, and databases were split into training and test sets across 100 replicates. Five supervised classification models were trained and then tested on either individual spectra or on spectra averaged per specimen. Model performance was assessed using accuracy and confusion matrices under varying conditions.

611 TABLE 3. Dataset-specific details of scanning protocols

Aspect	ID_PARIS	ID_YASUNI
Instrument	ASD LabSpec® Standard-Res Lab Analyzer	ASD LabSpec® Pro spectrophotometer
Probe type	Rapid Analysis Probe (90° angle) + Fiber Optic Illuminator	6 mm diameter pen probe
Software version	Indico Pro v6.5	Indico Pro v3.1
Reference procedure	White baseline references at start of session and every hour (Spectralon® white)	White reference every 15 minutes (Spectralon® Metal SpectralTM)
Black background	Spectral Black™ Foil beneath leaves	Spectral Black™ strip (gummed specimens) or larger black sheet (detached leaves)
Sample count/scan	20 readings averaged to generate one spectrum	Not specified
Integration time	68 ms^{-1}	Not specified
Number of scans	20 scans per specimen	50 scans per specimen
Operator/time	Single operator	Multiple operators

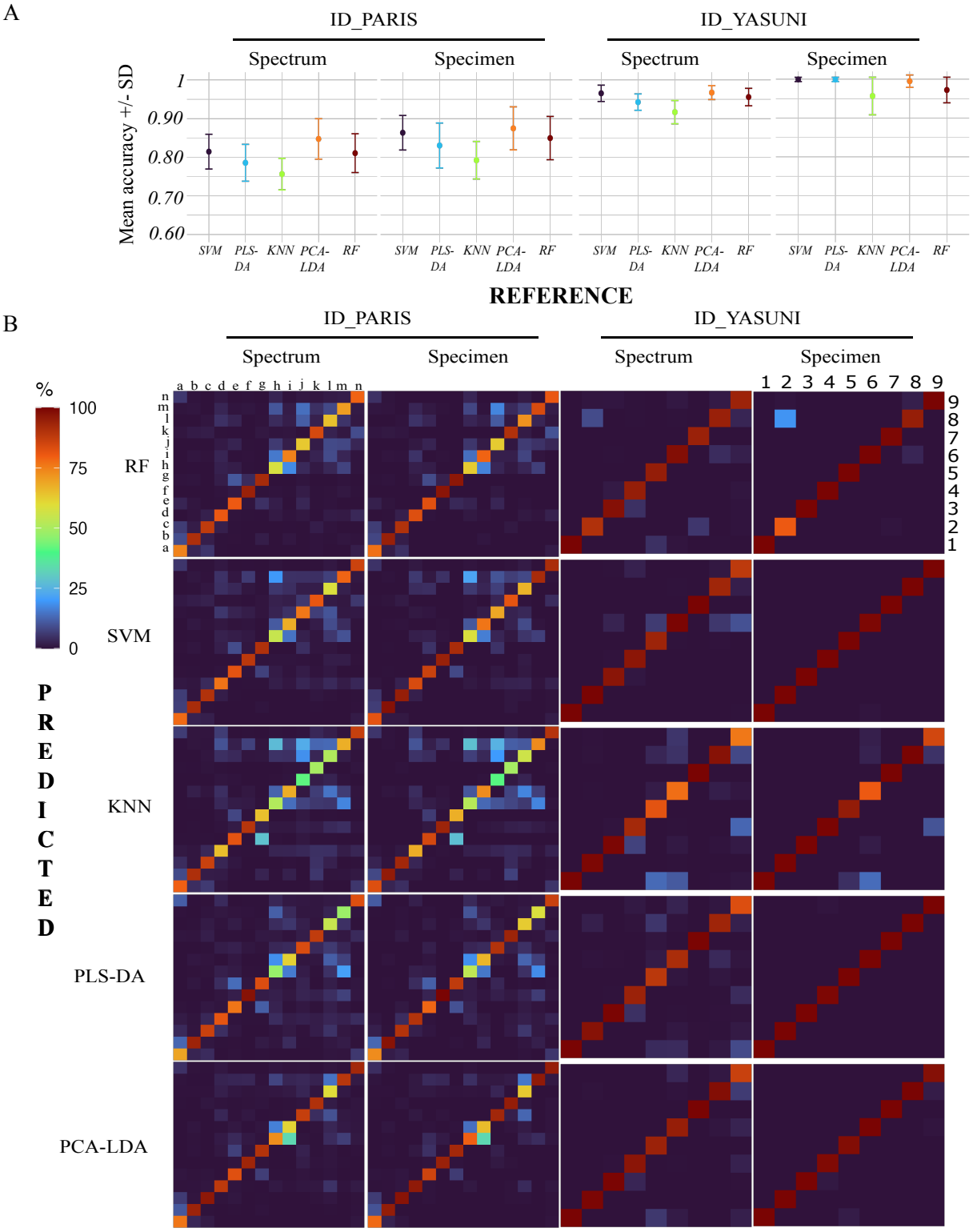


Figure 2. Species-level classification accuracy of leaf spectra. Classification performance for 14 species from ID_PARIS (a-*Ambavia capuronii*, b-*Ambavia gerrardi*, c-*Anaxagorea dolichocarpa*, d-*Annona senegalensis*, e-*Anonidium mannii*, f-*Duguetia calycina*, g-

616 *Greenwayodendron suaveolens*, h-*Hexalobus crispiflorus*, i-*Hexalobus monopetalus*, j-
617 *Monanthotaxis enghiana*, k-*Unonopsis guatterioides*, l-*Uvaria grandiflora*, m-*Uvariastrum*
618 *pierranum*, n-*Xylopia humblotiana*) and 9 species from ID_YASUNI (1-*Anaxagorea brevipes*,
619 2-*Duguetia hadrantha*, 3-*Guatteria guianensis*, 4-*Guatteria scalarinervia*, 5-*Oxandra*
620 *riedeliana*, 6-*Tetrameranthus globuliferus*, 7-*Trigynaeae triplinervis*, 8-*Unonopsis*
621 *veneficiorum*, 9-*Xylopia cuspidata*) using five models: Partial Least Squares Discriminant
622 Analysis (PLS-DA), k-Nearest Neighbors (KNN), Principal Component Analysis followed by
623 Linear Discriminant Analysis (PCA-LDA), Support Vector Machines (SVM), and Random
624 Forest (RF) from 100 independent replicates using 75 % randomly selected specimens in
625 training (6 for ID_PARIS, 3 for ID_YASUNI) all the others in the test, spectrum column mean
626 that all specimen spectra where classify and specimen column that only the average spectrum
627 is classified. **(A)** Overall model accuracy and standard deviation. **(B)** Confusion matrices
628 represent mean classification results. Values indicate the proportion of spectra predicted for
629 each species class.

630

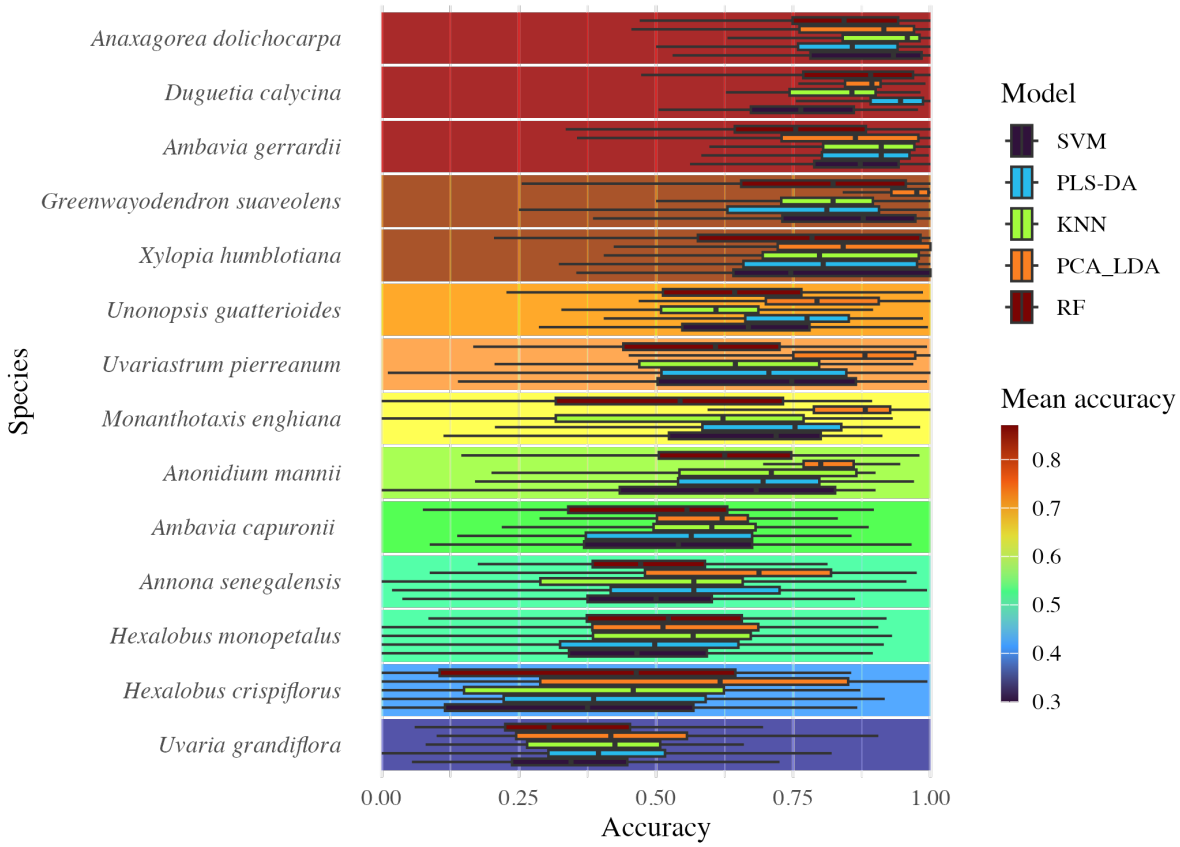


Figure 3. Performance of models trained on a single specimen per species. Classification accuracy (boxplots) across 100 replicates for each species and model, using one randomly selected specimen (20 spectra) as the training dataset. Background tiles indicate the mean classification accuracy across all models for each species. Results are based on the ID_PARIS dataset. Species are ordered from highest (top) to lowest (bottom) mean accuracy .

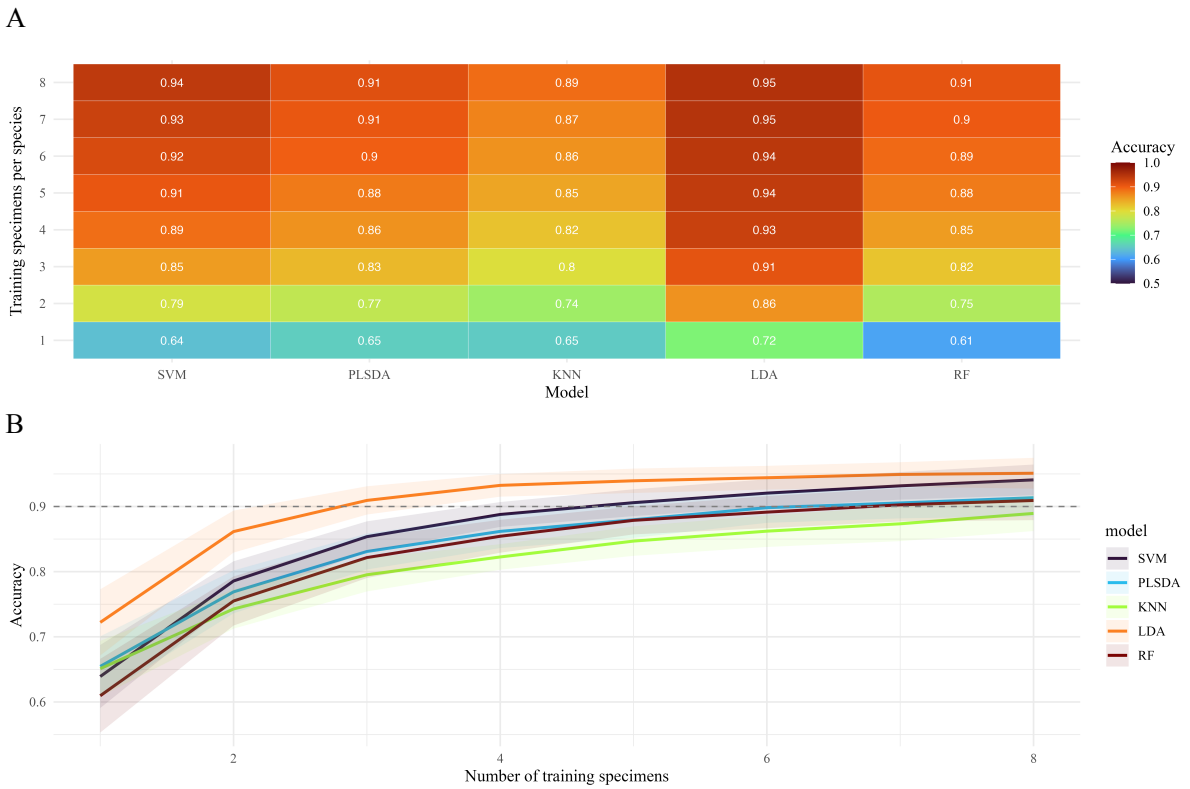


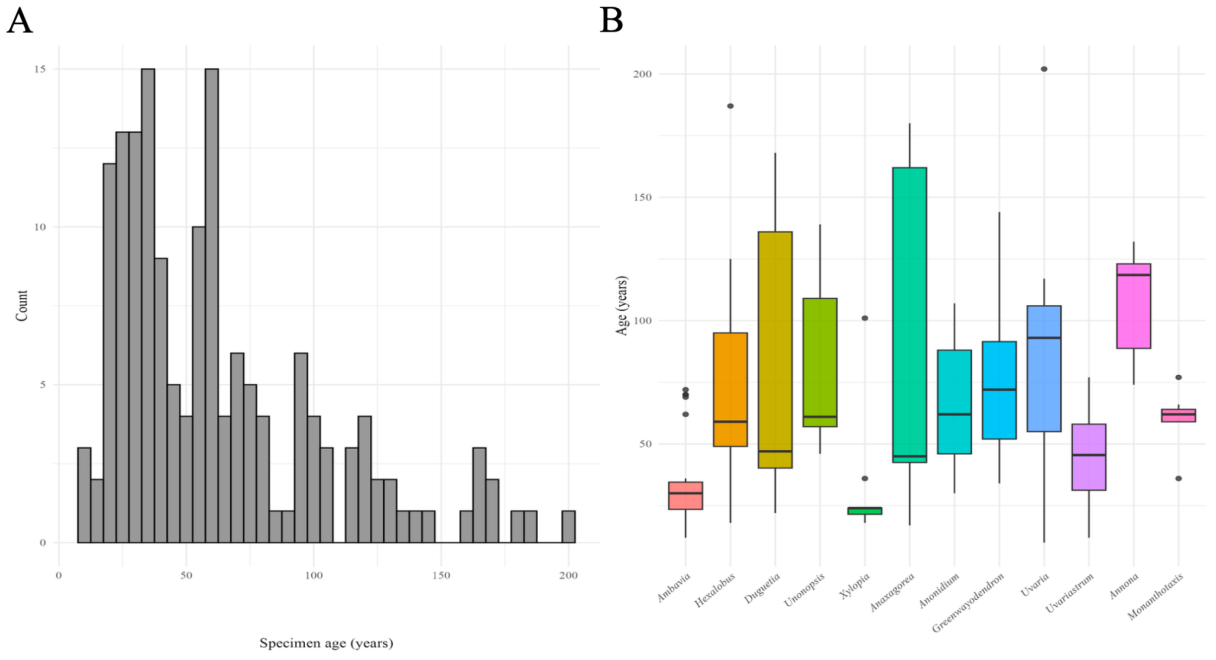
Figure 4. Effect of the number of training specimens per species on classification accuracy. **(A)** Heatmap of mean classification accuracy across 100 resampling iterations for each model and training size (1 to 8 specimens per species) using the ID_PARIS dataset. **(B)** Accuracy trajectories (± standard deviation) across the same iterations, showing model performance as a function of the number of training specimens per species.

644 # SUPPLEMENTARY

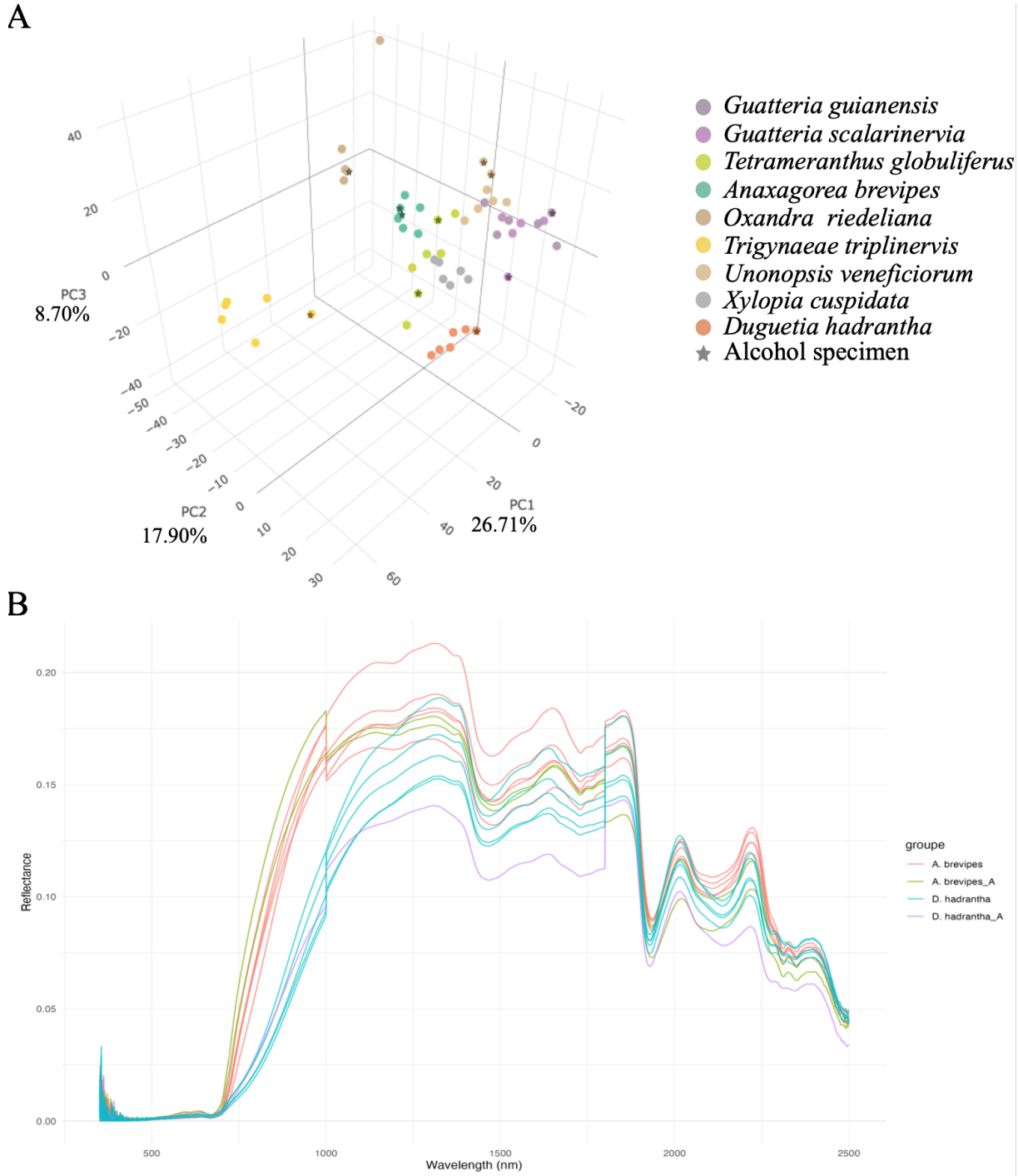
645 **Table S1.** List of herbarium specimens scanned from ID_Paris (barcode in brackets)

Species	Specimens
Ambavia capuronii	Schatz 2549 (P01966283); Razafimandimbison 247 (P01966290); Kotozafy 66 (P01966293); Rabe 6 (P01966295); Ravelonarivo 421 (P01966288); Ravelonarivo 617 (P01966289); Ravelonarivo 611 (P01966285); Turk 90 (P01966281); SF 8751 (P030238); SF 14256 (P02005959); SF 15301 (P01966279); Schatz 2536 (P01966282); Ranirison 812 (P06901328); Rakotondrajaono 107 (P01966280); Aridy 146 (P01966292); Malcomber 1028 (P01966294); Rakotonirina 128 (P01047823)
Ambavia gerrardi	Bosser 18703 (P02032521); Ranaivojaono 622 (P02032556); Rabenantoandro 1160 (P02032555), Randrianaivo 538 (P02032557); Randrianaivo 716 (P02032507); Rakotomalaza 441 (P02032523); SF 15383 (P02032525); Sauquet 94 (P00465816); Rahajasoa 608 (P02032505); Rabehevitra 3 (P06839299); Schatz 3096 (P02032508); Razanatsoa 522 (P06839300); Razakamalala 318 (P00858933); CNRE/PLARM 98 (P02032527); Evrard 11253 (P02032560); McPherson 18940A (P02032558); Rabehevitra 341 (P02133037); Rabevohitra 4893 (P06901451)
Anaxagorea dolichocarpa	Irwin 54545 (P01960236); Sagot 8 (P01960307); Lescure 875 (P01960277); Melinon 21 (P01960281); Melinon 31 (P01960283), Melinon 218 (P01960280); Granville 3706 (P01960643); Granville 5473 (P01960323); Granville 5950 (P01989415); Tostain 1669 (P01960545); Cremers 7680 (P01989410)
Annona senegalensis	s.c. 46 (P02032725); Nsama 1381 (P02032635); Mamboc 1411 (P02032637); Chevalier 2543 (P00517160); Chevalier 44019 (P00517162); Chevalier 13373bis (P02027965); Vuillet 410 (P00517154); Deigton 5442 (P02032682); Paroisse 107 (P02027962)
Anonidium mannii	Hedin 36bis (P02032735); Hedin 1337 (P02032738); Letouzey 194 (P02032741); Letouzey 11202 (P02032587); Kaji 98 (P02032588); de Wilde 7930 (P00353899); Cheek 7896 (P00956185); Cheek 6009 (P00956187); Raynal 9607 (P02032585); Annet 42 (P02032745); Breteler 2454 (P02032747)
Duguetia calycina	s.c. 7982 (P01984825); Melinon 93 (P01963746); Melinon 633 (P01963748); Deward 132 (P01984833); Mori 25703 (P06901455); Granville 5952 (P00124065); Granville 6253 (P01984844); Oldeman T-575 (P00124063); Villiers 3717 (P01748169); Sagot 1095 (P00083612); Poeppig 2772 (P01984856); Barrier 5112 (P01748171);
Greenwayodendron suaveolens	McPherson 15501 (P01965944); de Wilde 101 (P01965978); de Wilde 7940 (P01985243); Le Testu 7936 (P00363322); Gilbert 936 (P01965953); Tisserant 1172 (P01985835); Soyaux 218 (P00363356); Tisserant 2573 (P01084175); Zenker 1306 (P01985239); Leeuwenberg 7322 (P01985867); Letouzey 10456 (P01985871)

Hexalobus crispiflorus	Bouquet 895 (P00486232); Chevalier 13385 (P00486266); Pobéguin 176 (P00486263); Jaeger 9039 (P00486260); Félix 77 (P00486265); Fay 8299 (P00486244); Fleury 26627 (P00486257); Hallé 606 (P00486236); Harris 516 (P00486243); Heudelot 865 (P00486270); Birnbaum 1238 (P06901043)
Hexalobus monopetalus	Satabié 702 (P00486152); Megnikpa 2 (P00486153), Aubréville 415 (P00486158); Griaule N/A (P00486191); Chevalier 33487 (P00486200), Diallo 649 (P00486198), Boudouresque 5981 (P00486167); Baudet 3752 (P00486203); Peyre de Fabrègues 3408 (P00486160); Coussan 1158 (P00486176)
Monanthotaxis enghiana	Breteler 2137 (P01956022); Hallé 491 (P01985815); Hallé 367 (P01985814); Hallé 799 (P01985821); Hallé 3205 (P01985820); Hallé 2452 (P01985822); Tisserant 1257 (P01985786); Tisserant 2062 (P01985781); Harris 1721 (P01985779)
Unonopsis guatterioides	Labroy 107 (P00506137); Bernardy 19436 (P00506139); von Eggers 6882 (P00506141); Argent 6736 (P02132829); Westra 47623 (P02132839); Fróes 24503 (P02132835); Hostmann 1116 (P00506016); Rusby 1253 (P00506011); Fróes 1700 (P02132833); Harley 10597 (P02132825); Harley 10985 (P02132824); Santos 1660 (P02132826)
Uvaria grandiflora	Poilane 35624 (P00262277); Poilane 30328 (P00262274); Schaller KL3878 (P02132845); Poilane 40896 (P00262249); Hiêp 20 (P00262253); Couvreur 838 (P00938648); Robert 16 (P00262297); s.c. 836 (P00262239); Prawiroatmodjo 1387 (P02132755); Poilane 9789 (P00262263); Dilmy 12 (P02132754)
Uvariastrum pierreanum	McPherson 16128 (P01983191); Thomas 8094 (P01983291); Fogiel 1039 (P01983300); Letouzey 10225 (P01983295); Couvreur 590 (P02170712); Bakayoko 99 (P01983224); Koechlin 7758 (P01983223); Adam 26222 (P01983225); Akpabla 1121 (P01983228); Hallé 604 (P01983238)
Xylopia humblotiana	Razafindraibe 101 (P06901345); McPherson 98 (P06839280); Rabevohitra 3793 (P01031514); Rabevohitra 3792 (P01031513); Rogers 1007 (P02005890); Poncy 1597 (P00373125); McPherson 14371 (P01987117); Poncy 1594 (P00373123); Rakotonandrasana 676 (P01986879); Rakotovao 3823 (P01986878); Perrier de la bathie 15998 (P00524382); Lowry II 6030 (P01987062)



Supp. Figure 1. Age distribution of scanned Annonaceae specimens from the Paris herbarium (ID_Paris dataset)**. (A)** Histogram of specimen ages (calculated as 2025 minus collection year) showing the distribution of 160 scanned Annonaceae specimens from the Paris herbarium. **(B)** Boxplots of specimen ages grouped by genus. ID_PARIS.



Supp. Figure 2. Alcohol preservation effects on leaf spectral data. **(A)** three-dimensional PCA projection of the first three principal components, **(B)** mean raw reflectance spectra per specimen grouped by species and pre-drying treatment (*Anaxagorea brevipes* and *Duguetia hadrantha* alcohol vs. directly dried). An averaged spectrum is used to represent the specimen in both analyses. ID_YASUNI.