



# Use of Near Infrared Spectroscopy (NIRS) as a tool to discriminate species of the genus *Dimorphandra* Schott (Leguminosae: Caesalpinioideae)

Utilização da espectroscopia do infravermelho próximo (NIRS) como ferramenta na discriminação de espécies do gênero *Dimorphandra* Schott (Leguminosae: Caesalpinioideae)

G. S. Silva<sup>1\*</sup>; M. J. G. Hopkins<sup>2</sup>

<sup>1</sup>*Institute of Biology, State University of Campinas, 13083-862, Campinas-São Paulo, Brazil*

<sup>2</sup>*Department of Botany, National Institute of Amazonian Research, 69060-001, Manaus-Amazonia, Brazil*

\**guilhermecx.cx@hotmail.com*

*(Recebido em 04 de maio de 2023; aceito em 01 de março de 2024)*

This work aimed to use near-infrared spectroscopy (NIRS) as a tool to discriminate species of the genus *Dimorphandra* Schott (Leguminosae, Caesalpinioideae). Spectra were collected from 315 individuals (six readings per individual) distributed in 20 species of *Dimorphandra* using a Thermo Nicolet spectrophotometer, FT-NIR Antaris II Method Development System (MDS) in the INPA (National Institute of Amazonian Research) Herbarium. Absorbance values comprise the wavenumbers from 4,000 to 10,000  $\text{cm}^{-1}$ , corresponding to the near infrared region, recorded for 16 scans at a resolution of 8  $\text{cm}^{-1}$ . Principal Component Analysis (PCA) was used to visualize the spectral distribution. Discriminant functions were generated in order to evaluate the potential of the data to correctly distinguish the species and the 70-30 cross-validation technique was used to validate the generated models, with selections randomized 1, 10, 50 and 100 times. Excellent results were obtained in the PCA, with prediction values of 95-92%, using the 70-30 validation test in the linear discriminant analyses (LDA), thus indicating high predictive power in the discrimination of species of the genus *Dimorphandra*. Thus, it is inferred that NIRS contributes to the discrimination of species of the genus and elucidation of taxonomic problems.

Keywords: spectra, species delimitation, Leguminosae.

O trabalho tem como objetivo utilizar a espectroscopia no infravermelho próximo (NIRS) como ferramenta na discriminação de espécies do gênero *Dimorphandra* Schott (Leguminosae, Caesalpinioideae). Coletou-se espectros de 315 indivíduos (seis leituras por indivíduo) distribuídos em 20 espécies de *Dimorphandra*, utilizando o espectrofotômetro Thermo Nicolet, sistema FT-NIR Antaris II Method Development System (MDS) do Herbário INPA - Instituto Nacional de Pesquisas da Amazônia, onde os valores de absorvância compreende os números de onda de 4.000 a 10.000 $\text{cm}^{-1}$ , faixa correspondente ao infravermelho próximo, sendo regulados com 16 scans e resolução de 8 $\text{cm}^{-1}$ . Foram realizadas Análises de Componentes Principais (PCAs), para visualização da distribuição espectral. Foram geradas funções discriminantes com o intuito de avaliar o potencial dos dados em distinguir corretamente as espécies, onde para validar os modelos gerados empregou-se a técnica de validação cruzada 70-30, onde seleção foi aleatorizada 1; 10; 50 e 100 vezes. Os resultados obtidos demonstram excelentes resultados utilizando a PCAs e com valores de predição de 95-92%, utilizando o teste de validação 70-30 nas análises lineares discriminantes (LDA), verificando taxas altas de predição na discriminação de espécies do gênero *Dimorphandra*. Dessa forma infere-se que a espectroscopia no infravermelho próximo (NIRs) contribui na discriminação das espécies do gênero, colaborando na elucidação de problemáticas taxonômicas.

Palavras-chave: espectros, delimitação de espécies, Leguminosae.

## 1. INTRODUCTION

The family Leguminosae comprises over 22.000 species in ca. 800 genera, what makes it the third largest angiosperm family in terms of species diversity [1-3]. Ranging in size from small annual herbs to long-lived giant trees, the Leguminosae are often ecologically dominant in tropical and temperate biomes [1, 3]. In 2017, a new subfamily classification was established for Leguminosae that dealt with the long-standing problem of paraphyly of the subfamily Caesalpinioideae DC. and formally divided the family into six subfamilies: Cercidoideae LPWG, Detarioideae Burmeist., Duparquetioideae LPWG, Dialioideae LPWG, Caesalpinioideae DC., and Papilionoideae DC. [3]. Since then, the idea that Leguminosae comprises six major lineages has been largely confirmed by phylogenomic analyses of

large datasets of nuclear gene and plastome sequences [4-6], providing robust support for the six subfamilies.

Caesalpinioideae sensu is the second largest subfamily of Leguminosae, with ca. 4.680 species and 163 genera. Within this subfamily, ca. 3.400 species and 90 genera are placed within the mimosoid clade, corresponding to the former subfamily Mimosoideae [3]. Caesalpinioideae has a pantropical distribution and many of its lineages form ecologically abundant or dominant elements in each of the main lowland tropical biomes – seasonally dry tropical forests, savannas, and humid tropical forests – thus occurring along the whole rainfall range of lowland tropical regions, from arid to hyperhumid, with only a small fraction of species found in the warm temperate zone and a subset of these species showing to be ice-tolerant. The main current problems to be solved in Caesalpinioideae are related to the delimitation of non-monophyletic genera [3]. [7] indicate that, although they included 420 taxa in their current phylogenomic analyses, additional sampling is still clearly necessary to solve all the possible problems of lack of monophyly within Caesalpinioideae. For example, a comprehensive sampling is still needed to address the long-standing lack of monophyly in *Dimorphandra* Schott.

*Dimorphandra* currently presents 26 species in the world and 23 species and 6 subspecies in Brazil [8-10]. The genus was described in 1827 [11], with *Dimorphandra exaltata* Schott as the type species [8]. It is restricted to Brazil and adjacent countries (Colombia, Venezuela, Peru and Bolivia) in South America, with some species endemic to some phylogeographic domains such as the Amazon (*D. campinaranum* Ducke and *D. urubuensis* Ducke) and Cerrado (*D. wilsonii* Rizzini) and other species presenting a wide distribution such as *D. gardneriana* Tul. and *D. mollis* Benth [8, 9, 12]. The first evidences that *Dimorphandra* is not a monophyletic group were pointed out by Manzanilla and Bruneau (2012) [13], who proposed a phylogeny for *Dimorphandra* through a Bayesian analysis, based on the sequencing of four cpDNA regions obtained from 45 individuals and 16 species.

In 2017, the phylogeny proposed by LWPG showed *Dimorphandra* as a polyphyletic group with the lineages corresponding to the sections proposed by Tulasne (1844) [14] and recognized as subgenera (i.e., *Dimorphandra* with 11 species, *Pocillum* with 10 species, and *Phaneropsia* with five species) in the latest taxonomic revisions of the group published by Silva (1986) [8], forming distinct clades. *Dimorphandra conjugata* (Splitg.) Sandwith (subgenus *Phaneropsia*) was resolved as the sister group of the genera *Mora*, *Stachyothyrsus* and *Burkea*; the species *D. parviflora* Spruce ex Benth., *D. jorgei* M.F. Silva and *D. gardneriana* Tul. (subgenus *Dimorphandra*) formed a monophyletic clade showing a sister relationship to the group *Phaneropsia*, *Mora*, *Stachyothyrsus*, and *Burkea*; and finally, *D. macrostachya* Benth. (*Pocillum*) formed a clade together with *Dinizia* [3].

Over time, the sampling for the *Dimorphandra* phylogeny was expanded (57 specimens and 17 species), with sequenced plastid (matK) and nuclear (ITS) markers, demonstrating again that the *Dimorphandra* species were grouped into three subclades, each one corresponding to the three subgenera, corroborating the lack of monophyly of the genus [15]. Thus, studies that contribute to solve this problem are fundamental, especially in the discrimination of species [16].

In *Dimorphandra*, the morphological overlap of some traits such as the indument, number of pinnae, leaflet shape, floral pedicels, the shape of flowers, stamens and staminodes among closely related taxa and the lack of description of some reproductive structures such as flowers and fruits is relatively high, making it difficult to delimitate the species. The main examples are the erroneous determinations of *D. gardneriana* Tul. and *D. mollis* Benth., especially when young branches are collected, because the diagnostic traits of the species, such as the number and arrangement of pinnae, the indument, the margin and shape of the base of the leaflets, and pubescence of the flower buds, are not yet evident [8, 17].

The delimitation of the subspecies of *D. macrostachya* Benth. and *D. cuprea* Sprague & Sandwith is also problematic, as they also show considerable morphological overlap of traits (leaflet shape and indument, number of pinnae, calyx shape, and pedicel size) besides the lack of description of some traits such as the staminodes. Another factor is the uncertain geographical distribution of the species, due to lack of geographical data and incorrect determinations that end up leading to mistaken reports of occurrences, further confusing the delimitation of the subspecies [8, 12].

New tools that contribute to taxonomy always bring new perspectives and better support to clarify the delimitation of species. One of these tools is Near Infrared Spectroscopy (NIRS), which emerged in this century as a clean, low-cost and rapid method that takes advantage of and improves conventional analyses [18-20]. When a sample of organic material is irradiated, the chemical bonds vibrate continuously, causing the elongation and flexion of the molecules and activating a type of wave

movement that is characteristic of the functional group that composes it. This information aligned with multivariate analyses serves as basis to discriminate species [21].

Infrared spectroscopy of leaves has surprising potential to aid in the taxonomy and systematics of plants [22-30]. This is due to the fact that the spectral behavior of a leaf is a function of its chemical composition, morphology and internal structure, data that help in the recognition of species [31-33]. Recent studies have shown the effectiveness of Fourier transform near-infrared (FT-NIR) spectroscopy in the identification of botanical species [20, 21, 30, 34, 35].

Thus, the works already published clearly show that spectroscopy using the near-infrared band – NIRS is an important tool to discriminate species and assist in the taxonomic resolution of plant species. Thus, this study aimed to use NIRS as a tool to discriminate species of the genus *Dimorphandra* Schott (Leguminosae, Caesalpinioideae).

## 2. MATERIAL AND METHODS

A Thermo Nicolet spectrophotometer, FT-NIR Antaris II Method Development System (MDS), in the INPA (National Institute of Amazonian Research) Herbarium, was used to collect leaf spectral data. The equipment was calibrated (white) every four hours of use, at each reading. For control, a black body was used on top of the point where the spectrum was collected, to avoid the dispersion of light.

Spectra were collected from whole and dried leaves of exsiccates deposited in the INPA Herbarium. Two spectral readings were obtained per leaf, one from the adaxial face (side or face of the leaf that is on or close to the plant axis) and another from the abaxial face (side or face of the leaf that is on the opposite side or away from the plant axis), using three leaves per individual, totaling six readings per individual. Undamaged leaves were prioritized, but leaves with fungi and that suffered herbivory were not discarded. For the reading of large leaflets (larger than the integrating sphere ca. of 1 cm in diameter), the spectra were directly collected, placing the leaves on the sphere and covering them with the black body. For leaflets smaller than the integrating sphere, fisheye lens – an accessory that calibrates the passage of light through the integration sphere to focus only on small organic material – was used (Figure 1).

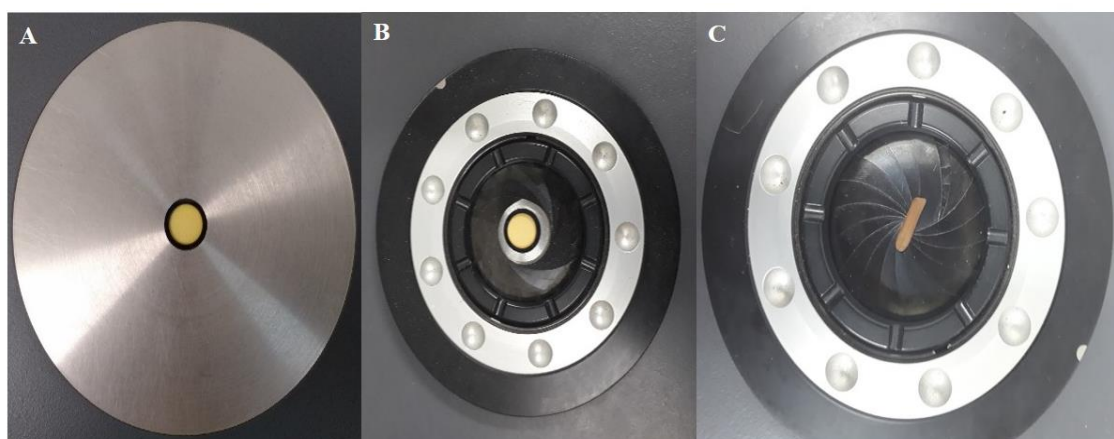


Figure 1. Demonstration of spectra collection with the use of fisheye lens. A) Integrating sphere without fisheye lens; B) Fisheye lens coupled to the integrating sphere; C) Leaflets on fisheye lens for spectra collection.

Spectra were collected from 315 individuals distributed in 20 *Dimorphandra* species (Figure 2). Absorbance values comprise the wavenumbers from  $4.000$  to  $10.000\text{ cm}^{-1}$ , recorded for 16 scans at a resolution of  $8\text{ cm}^{-1}$ . The spectra were not treated, only the Fourier transformation already integrated into the spectrophotometer was used. The collected data were organized in a table for visualization (Table 1).

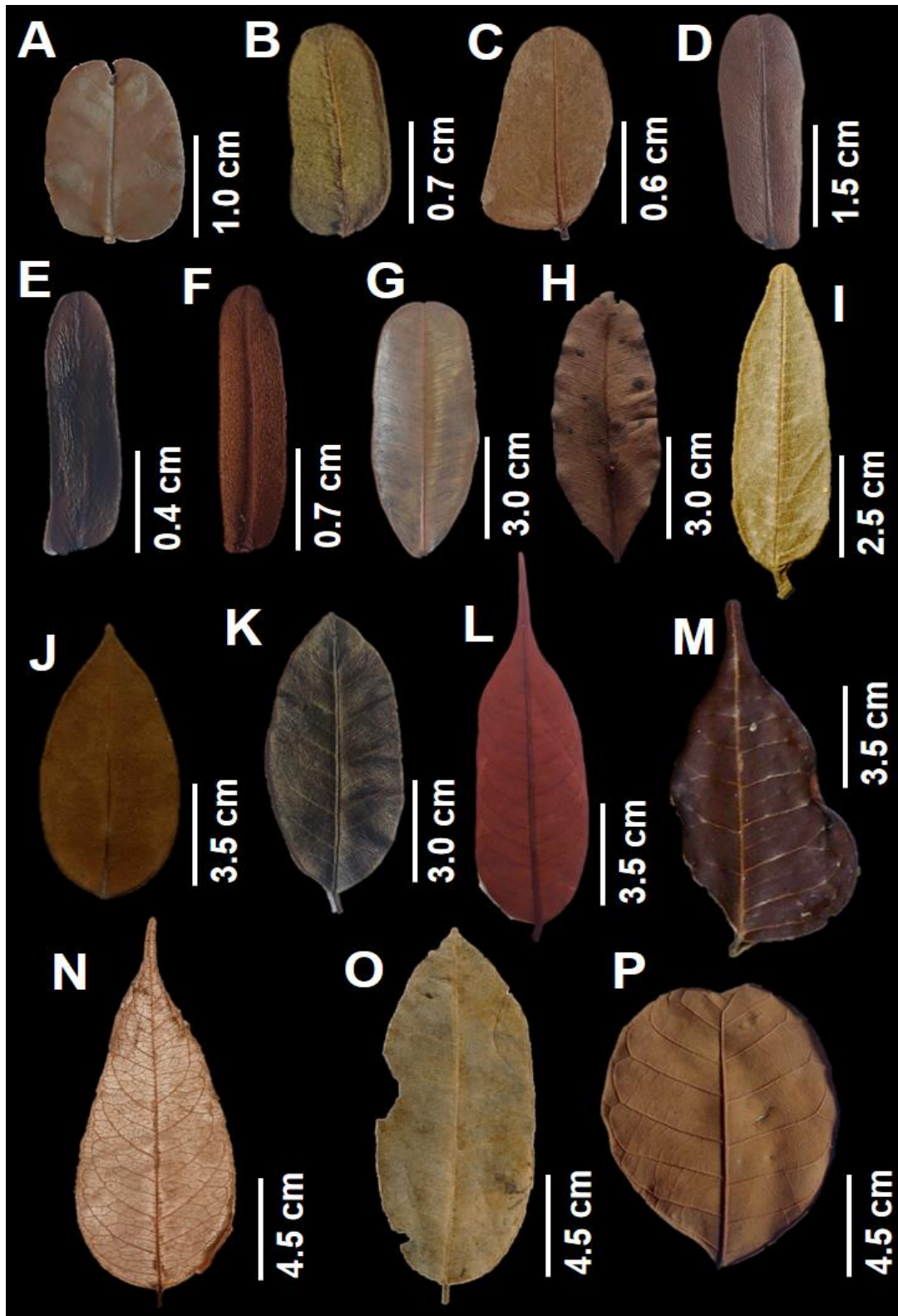


Figure 2. Size and shape variation of the *Dimorphandra* species analyzed. A) *D. gardneriana* Tul. (Prance, G.T. 18924-INPA) B) *D. mollis* Benth. (L.B. Pimentel 275-RB) C) *D. parviflora* Spruce ex Benth. (A. Ducke s.n-RB 16868) D) *D. polyandra* Benoist (B. Maguire 24365-RB) E) *D. penningera* Tul. (A. Ducke s.n-RB 20206) F) *D. cuprea* Sprague & Sandwith (J.G. de Carvalho-Sobrinho 1623-RB) G) *D. vernicosa* Spreng. ex Benth. (H.C. de Lima 2139-RB) H) *D. ignea* Ducke (A. Vicentini 985-RB) I) *D. exaltata* Schott (H.C. de Lima 2623-INPA) J) *D. multiflora* Ducke (R.L. Frões 24900-INPA) K) *D. jorgei* M.F.Silva (H.C. de Lima 6901-RB) L) *D. caudata* Ducke (N.T. Silva 3739-RB) M) *D. mediocres* Ducke (H.C. de Lima 3205-RB) N) *D. unijuga* Tul. (C.A.C Ferreira 9320-US) O) *D. davisse* Sprague & Sandwith (A.L. Bernardi 899-INPA) P) *D. conjugata* (Splitg.) Sandwith (Zanderij 186-RB).

Table 1. Number of specimens per species of *Dimorphandra* Schott used for spectral collection.

Species	N°	Species	N°
<i>D. caudata</i> Ducke	9	<i>D. cuprea</i> Sprague et Sandwith spp. <i>ferruginea</i> Ducke	5
<i>D. exaltata</i> Schott	7	<i>D. cuprea</i> Sprague et Sandwith spp. <i>velutina</i> Ducke	2
<i>D. gardneriana</i> Tul.	41	<i>D. gigantea</i> Ducke	4
<i>D. jorgei</i> M.F. da Silva	5	<i>D. ignea</i> Ducke	9
<i>D. mediocris</i> Ducke	2	<i>D. pennigera</i> Tul.	27
<i>D. mollis</i> Benth.	52	<i>D. polyandra</i> R.Benoist	2
<i>D. multiflora</i> Ducke	2	<i>D. urubuensis</i> Ducke	3
<i>D. parviflora</i> Spruce ex Benth.	28	<i>D. vernicosa</i> Spruce ex Benth.	73
<i>D. pullei</i> Amshoff	4	<i>D. unijuga</i> Tul.	6
<i>D. davisii</i> Sprague et Sandwith	8	<i>D. campinarum</i> Ducke	7
<i>D. cuprea</i> Sprague et Sandwith spp. <i>cuprea</i>	6	<i>D. coccinea</i> Ducke	13
		Total: 20 species	315 specimens

Principal Component Analysis (PCA) was used to determine the similarity between the spectra of the different species. The ordered attributes were all the absorbance values read at each wavelength and the readings per individual. The ordination was conducted using the average spectra per individual for all wavelengths, representing 1557 attributes. A PCA was performed with all species and two PCAs were performed for comparison of the species *D. gardneriana* and *D. mollis* and subspecies of *D. cuprea*: *D. cuprea* ssp. *cuprea*, *D. cuprea* ssp. *ferruginea* and *D. cuprea* ssp. *velutina*, whose determination is problematic.

Discriminant functions were generated according to Lang et al. (2015) [20] in order to evaluate the potential of the data to correctly distinguish the species. The independent variables in the discriminant functions generated were the 20 species analyzed and the dependent variables were the mean absorbance values at each wavelength of the readings per individual.

The cross-validation technique recommended by Durgante et al. (2013) [19] was used to test the efficiency of the generated model. In this technique, 70% of the individuals are used to generate the model and 30% are used to validate it; the selection was randomized 1, 10, 50 and 100 times. The procedure aims to detect the set of independent variables that best predict the species analyzed. For each of the tests described, the percentage of correctness in the identifications is to be obtained. The analyses were performed in the R software 2.10.0 [36].

### 3. RESULTS AND DISCUSSION

In the visual inspection of spectral behaviors, two patterns were observed, as shown in Figure 3. One of them was already expected for the spectra of leaves, with absorbances ranging from 0.55-0.20 (Behavior A), as described by [19], and another totally different from the one recorded in previous studies, with high absorbance values ranging from 1.0-0.7 (Behavior B).

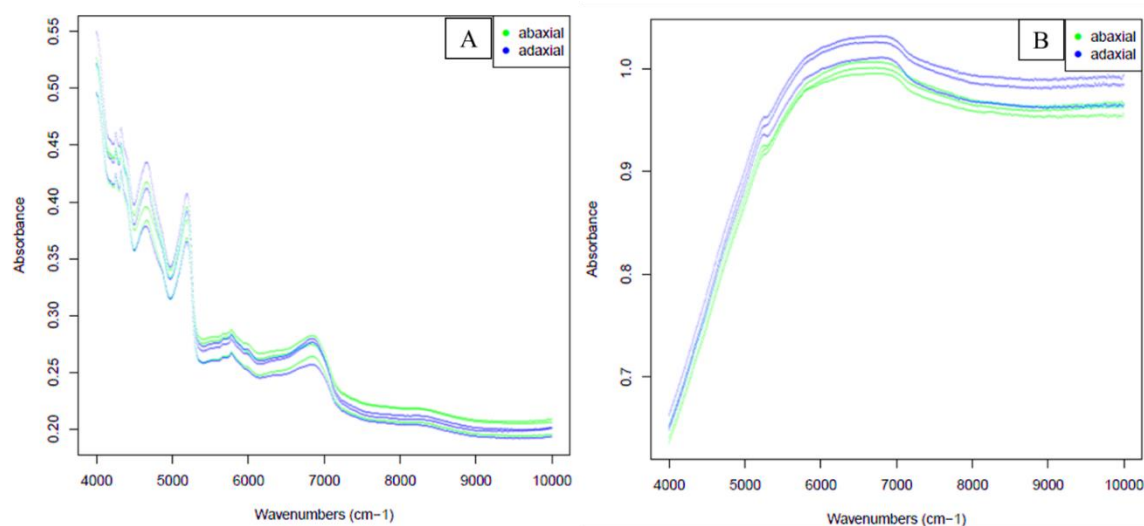


Figure 3. Spectral behavior of the genus *Dimorphandra* based on measurements in the abaxial and adaxial face of leaves. A) Behavior A; B) Behavior B.

It was found that of the 315 individuals who had spectra collected, 244 showed the behavior A and 71 the behavior B. The difference in spectral behaviors was related to the collection method: spectra collected directly without the use of fisheye lens resulted in behavior A, and those collected using fisheye lens resulted in behavior B. Thus, the use of fisheye lens modifies the visualized spectral behavior, producing different absorbance levels in relation to collections performed without this tool, as shown in previous works [19, 20, 29, 33-35]. However, the use of this spectrophotometer tool is recommended by the manufacturer for materials smaller than the integrating sphere [36].

The patterns informed by the spectra of each species are a reflection of the chemical and physical properties of the organic material that the technique is able to capture aligned with the form of spectrum collection, mainly the greater or lower absorbances that the tool can capture from the organic material. Differences in spectral behavior among genetically distinct groups are expected [31].

The ordering of the species in two dimensions captured 99% of the variation (PC1: 89% and PC2: 10%) using average values of the raw spectra in the entire wavelength (4000 to 10000  $\text{cm}^{-1}$ ) (Figure 4). Most species were clustered, such as *D. vernicosa*, *D. campinarum*, *D. giagantea* among others. However, some groups (the minority) were dispersed, as in the case of *D. gardneriana* and *D. exaltata*. Yet, even dispersed, some individuals formed distinct groups.

This is the first time that Fourier transform near-infrared spectroscopy (FT-NIR) has been tested to discriminate species of Legumes from the subfamily Caesalpinoideae, with bipinnate leaves and a wide geographic distribution. Our results show that FT-NIR is a powerful tool that can be easily applied to species identification using leaf spectral data [37, 38]. Considering that PCA values explain the similarity between the spectra of different species, the higher the value of the axes (99% in the present study), the better the explanation of this variation and the recognition of spectra-species groupings.

Comparing with other studies, the results are similar to works [25, 29, 32], where they obtained values greater than 95% capturing the variation of the spectra in the PCA method. Of the studies analyzed with a methodology similar to the present study, only studies [23, 27] obtained values lower than 70% capturing the variation of spectra in the PCA method, indicating that Near Infrared Spectroscopy (NIRS) as a tool to discriminate species is highly efficient.

These exploratory data demonstrate the importance of the tool in the discrimination of species, since many identification errors or the lack of curation of deposited samples lead to problems in the use of data. This problem is, in fact, common in *Dimorphandra* because of its wide morphological variation [39]. Spectral readings using dried plants from herbarium collections provide an alternative to solve identification problems using specimens already available, without the need for new collections [34].

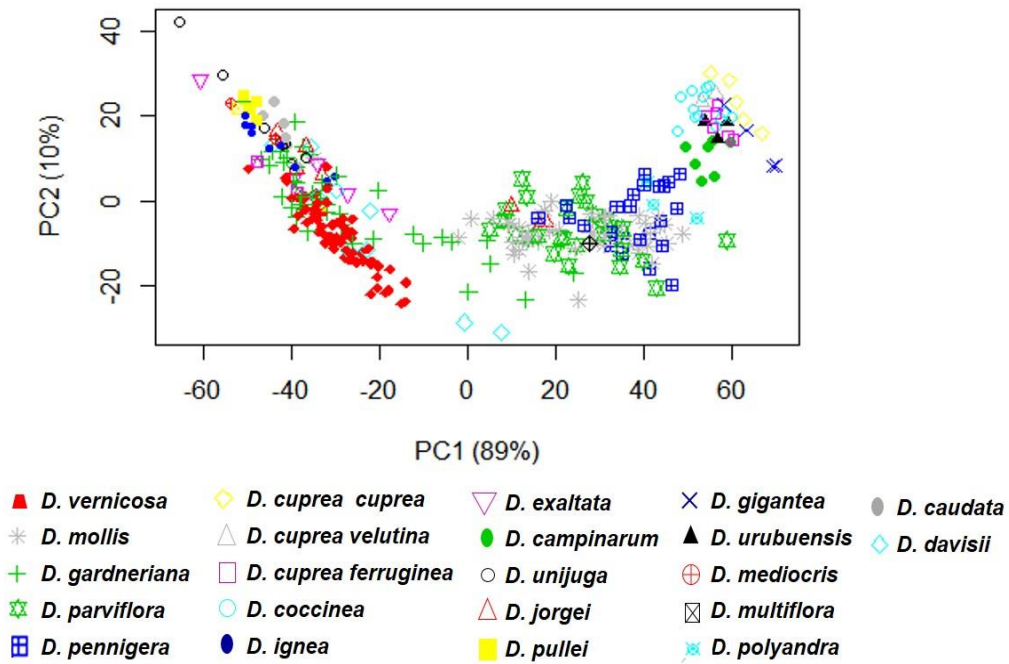


Figure 4. Principal component analysis (PCA) using average spectra of individuals of the 20 *Dimorphandra* species studied.

In the PCA, in general, the spectra formed two groups: one group that goes from 0 to 60 (left) and another that goes from 0 to 60 (right), based on the PC1, which explained most of the variation. The use of PCA to group samples according to differences and similarities of spectral data generated by infrared analysis, reducing the dimensionality of the data set while preserving the maximum information, proved to be positive [40].

The distinct spectral behaviors observed in the PCA was explained by the collection method: individuals of species with large leaves, in which the collection was performed directly without the use of fisheye lens, showed spectra from 0 to 60 in the PC1, while those in which collections were performed with the aid of fisheye lens showed a different behavior. This demonstrates that collection tool can influence the spectra behavior.

The use of an accessory introduces new variables in the spectrum, which may compromise the recognition of tissues in the sample. The materials of accessories have chemical components that add extra information unrelated to the spectral signatures of the species [40]. New tests and standardization in the use of accessories are of utmost importance in the definition of an effective protocol for spectral collections [41].

In the PCAs of *D. gardneriana* and *D. mollis* (Figure 5A) and of the subspecies of *D. cuprea* (Figure 5B), there was a clear separation of species. The ordering in two dimensions captured, in general, 99% to 97% of the variation. Problems in the delimitation of *D. gardneriana* and *D. mollis* are mainly associated to the collection of young or immature branches, in which the main distinguishing characters are not yet established. The determination of the two species solely based on vegetative structures of young branches is very difficult. Thus, the contribution of spectroscopy is fundamental to discriminate these species [8, 12].

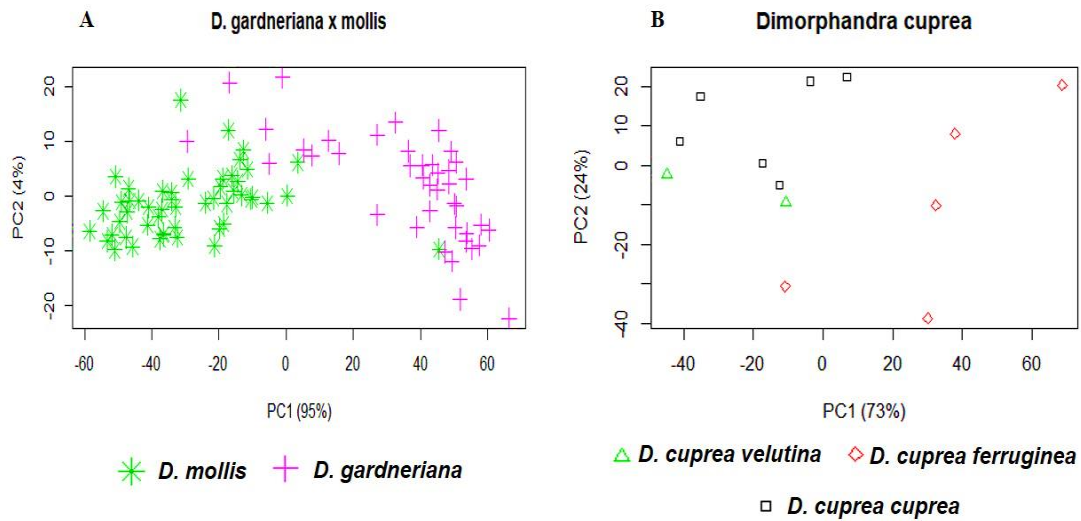


Figure 5. Principal component analysis (PCA) of the genus *Dimorphandra* using the average spectra of individuals of subspecies of *D. cuprea*, and of the species *D. gardneriana* and *D. mollis*. A) PCA of the species *D. gardneriana* and *D. mollis*; B) PCA of subspecies of *D. cuprea*.

The delimitation of the taxa included in *D. cuprea*, especially *D. cuprea* ssp. *cuprea* and *D. cuprea* ssp. *ferruginea*, which have some overlapping characteristics, is not clear. [8] uses a series of characteristics to separate the two subspecies, including the shape of the petals and size of the trees. These characteristics are essential, but they show some overlap. Thus, spectral data are important because they are able to separate these taxa. Exploratory data are satisfactory for closely related taxa such as those studied here, but it is noteworthy that for the efficient use of FT-NIR spectra in species identification, reference spectral data for calibration of models to be used in the future to classify new samples are necessary [42]. The spectral data managed here are pioneer for future calibrations.

The results for the 70-30 cross-validation test, with respective randomizations, are shown in Figure 6. An average prediction of approximately 93% was found among the randomizations performed, demonstrating a high predictive power for the discriminant functions generated. Thus, the model generated is efficient to discriminate *Dimorphandra* species. In a study with *Ephedra* species using FT-NIRS, Fan et al. (2010) [29] found results from 84.2 to 91.9% in the discriminant analysis.



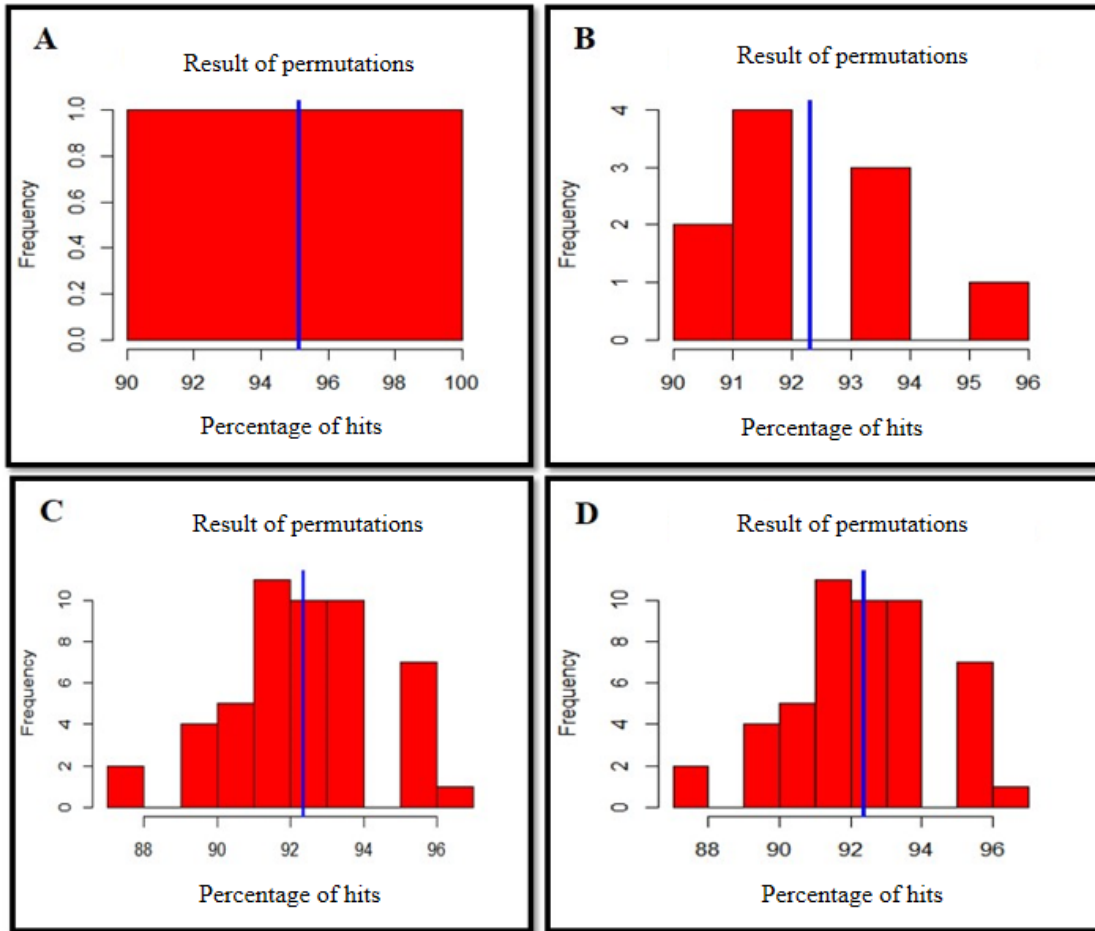


Figure 6. Percentage of accuracy by permutation obtained in the 70-30 validation test to validate discriminant functions. A) Randomized selection only 1 time; B) Randomized selection 10 times; C) Randomized selection 50 times; D) Randomized selection 100 times.

As for the categories with prediction error, a recurrent pattern was found for the species, as shown in Table 2. A prediction error consists of cases of non-matching species indicated by the generated model. It is noteworthy that, in these cases, the model indicated very close species with difficult morphological determination and that require a more accurate taxonomy. Each taxonomic approach accesses a different type of information about diverging lineages, and multiple lines of evidence are needed to separate the species, especially in closely related groups [43, 44].

Table 2. Demonstration of the categories with prediction error of species of the genus *Dimorphandra* demonstrated in the 70-30 validation test of the discriminant functions.

Categories	Prediction
<i>D. mollis</i>	<i>D. jorgei</i>
<i>D. gardneriana</i>	<i>D. mollis</i>
<i>D. exaltata</i>	<i>D. jorgei</i>
<i>D. ignea</i>	<i>D. vernicosa</i>
<i>D. cuprea ssp. ferr.</i>	<i>D. urubuensis</i>
<i>D. gigantea</i>	<i>D. urubuensis</i>

#### 4. CONCLUSION

The present results show that NIRS is a useful tool for the discrimination of species of the genus *Dimorphandra*. Excellent results, with prediction values of 92-95% in the 70-30 validation test in the linear discriminant analyses (LDA), were obtained. This is the first time that Fourier transform near-infrared spectroscopy (FT-NIR) has been tested to discriminate species of Legumes from the subfamily Caesalpinoideae, with bipinnate leaves and a with wide geographic distribution.

After the validation test, aligned with the change in spectral behavior using fisheye lens, it was found that despite the modification, the predictive values were high. However, the use of fisheye lens for spectral collection and the study of materials smaller than the integrating sphere need to be better investigated so as to determine the extent to which these changes influence the analyses performed.

PCA is an important tool for exploring and understanding the spectral behavior of species. This analysis provided important information on spectral data of *Dimorphandra* species, helping to discriminate species from a group with highly variable leaf morphology. Thus, it is concluded that NIRS is an auxiliary tool for the discrimination of species of the genus *Dimorphandra*, contributing to the identification and elucidation of taxonomic problems. Considering that spectroscopy provides a single fingerprint for each sample, the technique offers valuable data for taxonomic classification. The integration between spectroscopy and taxonomy is promising and should be consolidated in the near future for botanical studies.

#### 5. ACKNOWLEDGEMENTS

We thank Dr. Flávia Machado Durgante for her help in the analyses and collaboration in the writing of the text, Dr. Ricardo Perdiz for his help in organizing the data and performing the analyses, the CNPq for providing the master's scholarship of the first author, and the INPA herbarium for allowing the use of the spectrophotometer and providing the plant material for spectral collections.

#### 6. REFERENCES

- Lewis GP, Schrire B, Mackinder B, Lock M. Legumes of the world. London (UK): Royal Botanic Garden; 2005.
- Legume Phylogeny Working Group (LPWG). Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species-rich clades. *Taxon*. 2013;62:217-248. doi: 10.12705/622.8
- Legume Phylogeny Working Group (LPWG). A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny. *Taxon*. 2017;66(1): 44-77. doi: 10.12705/661.3
- Koenen EJM, Ojeda DI, Steeves R, Migliore J, Bakker FT, Wieringa JJ, et al. Large-scale genomic sequence data resolve the deepest divergences in the legume phylogeny and support a near-simultaneous evolutionary origin of all six subfamilies. *New Phytol*. 2020;225(3):1355-69. doi: 10.1111/nph.16290
- Zhang R, Wang YH, Jin JJ, Stull GW, Bruneau A, Cardoso D, et al. Exploration of plastid phylogenomic conflict yields new insights into the deep relationships of Leguminosae. *Syst Biol*. 2020;69(4):613-22. doi: 10.1093/sysbio/syaa013
- Zhao Y, Zhang R, Jiang K, Qi J, Hu Y, Guo J, et al. Nuclear phylotranscriptomics/phylogenomics support numerous polyploidization events and hypotheses for the evolution of rhizobial nitrogen-fixing symbiosis in Fabaceae. *Mol Plant*. 2021;14(5):748-73. doi: 10.1016/j.molp.2021.02.006
- Ringelberg JJ, Koenen EJM, Iganci JR, de Queiroz LP, Murphy DJ, Gaudeul M, et al. Phylogenomic analysis of 997 nuclear genes reveals the need for extensive generic redefinition in Caesalpinoideae (Leguminosae). In: Hughes CE, de Queiroz LP, Lewis GP, editors. *Advances in Legume Systematics 14. Classification of Caesalpinoideae Part 1: New generic delimitations*. *PhytoKeys*. 2022;205:3-58. doi: 10.3897/phytokeys.205.85866
- Silva MF. *Dimorphandra* (Caesalpiniaceae). New York (US): Flora Neotropica, New York Botanical Garden. 1986;44:127.
- Flora e Funga do Brasil [Internet]. Rio de Janeiro: Jardim Botânico do Rio de Janeiro; 01 jan 2020 [citado em 10 jan 2023]. Disponível em: <http://floradobrasil.jbrj.gov.br/>
- Silva GS, Hopkins MJG. New record of Fabaceae (Caesalpinoideae) for Brazil: *Dimorphandra davisii* Sprague & Sandwith. *Biota Amazôn*. 2018;8(2):63-4.
- Schott H. Fasciculus plantarum brasiliensium. In: Sprengel KPJ. *Systemtic vegetabilis*. 1827;4(2):403-10.

12. Souza VC, Lima AG. *Dimorphandra* in Flora e Funga do Brasil [Internet]. Rio de Janeiro: Jardim Botânico do Rio de Janeiro; 01 jan 2010 [citado em 25 abr 2023]. Disponível em: <https://floradobrasil.jbrj.gov.br/FB78675>
13. Manzanilla V, Bruneau A. Phylogeny reconstruction in the *Caesalpinieae* grade (Leguminosae) based on duplicated copies of the sucrose synthase gene and plastid markers. *Mol Phylogenet Evol.* 2012;65:149-62.
14. Tulasne MLS. Legumineuses arborescentes del Amerique du Sud. *Archs Mus Hist Nat Paris.* 1844;4:182-98.
15. Rocha VD, Dal'Sasso TCS, Williams CCV, Simon MF, Bueno ML, Oliveira LO. From forest to savanna and back to forest: evolutionary history of the genus *Dimorphandra* (leguminosae). *bioRxiv.* 2023;524261:34. doi: 10.1101/2023.01.16.524261
16. Hughes CE, de Queiroz LP, Lewis GP. Advances in Legume Systematics 14. Classification of Caesalpinioideae Part 1: New generic delimitations. *PhytoKeys.* 2022;205:1-470. doi: 10.3897/phytokeys.205.82821
17. Landim LP, Costa JGM. *Dimorphandra gardneriana* Tulasne (Fava d'anta) - Uma abordagem etnobotânica e riscos de extinção. *Rev Biol.* 2012;9(1):6-11. doi: 10.7594/revbio.09.01.02
18. Tralamazza SM, Bozza A, Destro JGR, Rodríguez JI, Dalzoto PR, Pimentel IC. Potential of Fourier Transform Infrared Spectroscopy (FT-IR) to differentiate environmental *Aspergillus* fungi species *A. niger*, *A. ochraceus* and *A. westerdijkiae* using two different methodologies. *Appl Spectrosc.* 2013;67(3):274-8. doi: 10.1366/12-06600
19. Durgante FM, Higuchi N, Almeida A, Vicentini A. Species Spectral Signature: Discriminating closely related plant species in the Amazon with Near-Infrared Leaf-Spectroscopy. *For Ecol Manag.* 2013;209:240-8. doi: 10.1016/j.foreco.2012.10.045
20. Lang C, Costa FRC, Camargo JLC, Durgante FM, Vicentini A. Near infrared spectroscopy facilitates rapid identification of both young and mature amazonian tree species. *PloS One.* 2015;10(8):e0134521. doi: 10.1371/journal.pone.0134521
21. Foley WJ, McIlwee A, Lawler I, Aragonés L, Woolnough AP, Berding N. Ecological applications of near infrared reflectance spectroscopy- a tool for rapid, cost-effective prediction of the composition of plant and animal tissues and aspects of animal performance. *Oecologia.* 1998;116:293-305.
22. Lu HF, Cheng CG, Tang X, Hu ZH. FTIR Spectrum of *Hypericum* and *Triadenum* with reference to their identification. *Acta Bot Sin.* 2004;46(4):401-6.
23. Kim SW, Ban SH, Chung H, Cho S, Chung HJ, Choi PS, et al. Taxonomic discrimination of flowering plants by multivariate analysis of Fourier transform infrared spectroscopy data. *Plant Cell Rep.* 2004;23:246-50. doi: 10.1007/s00299-004-0811-1
24. Vance CK, Tolleson DR, Kinoshita K, Rodriguez J, Foley WJ. Near infrared spectroscopy in wildlife and biodiversity. *J. Near Infrared Spectrosc.* 2016;24:1-25. doi: 10.1255/jnirs.1199
25. Dharmaraj S, Jamaludin AS, Razak HM, Valliappan R, Ahmad NA, Harn KGL, et al. The classification of *Phyllanthus niruri* Linn. according to location by infrared spectroscopy. *Vib. Spectrosc.* 2006;41:68-72. doi: 10.1016/j.vibspec.2005.12.012
26. Krajsenk SS, Buh P, Zega A, Kreft S. Identification of herbarium Whole-leaf Samples of *Epilobium* Species by ATR-IR Spectroscopy. *Chem Biodivers.* 2008;5:310-7. doi: 10.1002/cbdv.200890028
27. Castillo R, Contreras D, Freer J, Ruiz J, Valenzuela S. Supervised pattern recognition techniques for classification of *Eucalyptus* species from leaves NIR spectra. *J Chil Chem.* 2008;53(4):1709-13. doi: 10.4067/S0717-97072008000400016
28. Guo T, Guo L, Wang X, Li M. Application of NIR spectroscopy in classification of plant species. Wuhan (CH): First International Workshop on Education Technology and Computer Science. 2009:879-83. doi: 10.1109/ETCS.2009.734
29. Fan Q, Wang Y, Sun P, Liu S, Li Y. Discrimination of *Ephedra* plants with diffuse reflectance FT-NIRS and multivariate analysis. *Talanta.* 2010;80:1245-50. doi: 10.1016/j.talanta.2009.09.018
30. Mallmann CL, Pereira Filho W, Dreyer JBB, Tabaldi LA, Durgante FM. Leaf-level field spectroscopy to discriminate invasive species (*Psidium guajava* L. and *Hovenia dulcis* Thunb.) from native tree species in the Southern Brazilian Atlantic Forest. *Remote Sens Environ.* 2023;15(791):1-23. doi: 10.3390/rs15030791
31. Ponzoni FJ. Sensoriamento remoto no estudo da vegetação: Diagnosticando a mata atlântica. São José dos Campos (SP): Instituto Nacional De Pesquisas Espaciais. 2002;8984(62):1-28.
32. Damasco G, Daly DC, Vicentini A, Fine PVA. Reestablishment of *Protium cordatum* (Bursleraceae) based on integrative taxonomy. *Taxon.* 2019;68(1):34-46. doi: 10.1002/tax.12022
33. Lang C, Almeida DRA, Costa FRC. Discrimination of taxonomic identity at species, genus and family levels using Fourier Transformed Near-Infrared Spectroscopy (FT-NIR). *For Ecol Manag.* 2017;406:219-27. doi: 10.1016/j.foreco.2017.09.003
34. Hadlich HL, Durgante FM, dos Santos J, Higuchi N, Chambers JQ, Vicentini A. Recognizing amazonian tree species in the field using bark tissues spectra. *For Ecol Manag.* 2018;427:296-304. doi: 10.1016/j.foreco.2018.06.002

35. Prata EMB, Sass C, Rodrigues DP, Domingos FMBC, Specht CH, Damasco G, et al. Towards integrative taxonomy in neotropical botany: disentangling the *Pagamea guianensis* species complex (Rubiaceae). *Bot J Linn Soc.* 2018;188:213-31. doi: 10.1093/botlinnean/boy051
36. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 01 jan 2000 [citado em 10 Jan 2023]. Disponível em: <https://www.R-project.org/>
37. Rodriguez-Fernandez JI, de Carvalho CJB, Pasquini C, Lima KMG, Moura MO, Arízaga GGC. Barcoding without DNA? Species identification using near infrared spectroscopy. *Zootaxa.* 2011;2933:46-54. doi: 10.11646/zootaxa.2933.1.3
38. Paiva DNA, Perdiz RO, Almeida TE. Using near-infrared spectroscopy to discriminate closely related species: A case study of neotropical ferns. *J Plant Res.* 2021;134:509-20. doi: 10.1007/s10265-021-01265-9
39. Silva GS, Simon MFS. Tribe Dimorphandra. In: Bruneau A, Queiroz LP, Ringelberg J. *Advances in Legume Systematics 14. Classification of Caesalpinioideae Part 2: Higher-level classifications.* *PhytoKeys.* 2023;240:177-86. doi: 10.3897/phytokeys.240.101716
40. Barros Neto B, Scarminio IS, Bruns RE. 25 anos de quimiometria no Brasil. *Quím Nova.* 2006;29:1401-06. doi: 10.1590/S0100-40422006000600042
41. Torralvo K, Durgante F, Pasquini C, Magnusson WE. Near infrared spectroscopy for the identification of live anurans: Towards rapid and automated identification of species in the field. *J Near Infrared Spectrosc.* 2023;0(0):1-9. doi: 10.1177/09670335231156472
42. Pasquini C. Near Infrared Spectroscopy: Fundamentals, practical aspects and analytical applications. *J Braz Chem Soc.* 2003;14:198-219. doi: 10.1590/S0103-50532003000200006
43. de Queiroz K. The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations. In: Howard DJ, Berlocher SH. *Endless forms: Species and speciation.* New York (US): Oxford University Press; 1998. p. 57-75.
44. de Queiroz K. Species concepts and species delimitation. *Syst Biol.* 2007;56(6):879-86. doi: 10.1080/10635150701701083