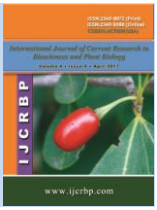
International Journal of Current Research in
Biosciences and Plant Biology
Volume 4 • Number 4 (April-2017) • ISSN: 2349-8080 (Online)
Journal homepage: www.ijcrbp.com



Original Research Article **doi:** https://doi.org/10.20546/ijcrbp.2017.404.001

The Identification Method for *Machilus* Nees Species Based on Visible-Near Infrared Spectral Analysis Technology

Yue Fei, Bo Xiao, Xianbin Liao, Sitao Wen, Biyu Zhou, Jiayao Pan and Die Hu*

Germplasm Resources Evaluation and Innovation Center of Phoebe Yangtze University, Hubei, Jingzhou, People's Republic of China

Corresponding author.

Abstract

A novel technology for rapid identification of *Machilus* Nees species using the visible-near infrared spectrum (300-1100 nm) is described in this study. The reflectivities of new leaves of seedlings from 9 species of the genus *Machilus* Nees were collected. Stepwise discriminant analysis was applied to the spectral information of the leaves, and 18 unique bands were selected from 126 bands total. After obtaining the spectral information for the unique bands, the Bayesian discriminant method was applied to establish the discriminant analysis model for *Machilus* Nees species. According to the discrimination model, combinations of 6, 12, and 18 unique bands were selected, and the discrimination accuracies of 180 training samples reached 76.111%, 83.889%, and 93.889%, respectively, while the accuracies of 90 testing samples were 77.778%, 84.444%, and 95.556%, respectively. These results validated the discrimination model for *Machilus* Nees species constructed from the spectral information of 18 selected unique bands. The application of visible-near infrared spectrum technology combined with discriminant analysis could provide a novel approach for the rapid and accurate identification of *Machilus* Nees species.

Article Info

Accepted: 13 March 2017
Available Online: 06 April 2017

Keywords

Discriminant analysis
Machilus Nees
Species identification
Visible-near infrared spectroscopy

Introduction

There are about 100 species in the genus *Machilus* Nees of the family Lauraceae with diverse varieties and worldwide distribution. There are 84 species and 4 varieties in China with 66 endemic species (Li et al., 2008), and most are phenetic species of tropical and subtropical forests (Lin, 2007). This genus of plants not only has great potential in gardens, buildings, and spice (Liu and Fei, 2011), but also plays important roles in the common economic forest in southern China. Therefore, they are of great interest to botanists.

However, due to their close genetic relationship and external morphology, especially the extremely similar leaf shape, plants in *Machilus* Nees are difficult to discriminate, and incorrect identifications occur easily. The traditional identification methods for *Machilus* Nees plants mainly include morphology (Zhong and Xia, 2010), leaf epidermis anatomy, palynology (Wang and Wei, 2003; Wei and Tang, 2006), and molecular systematics (Chen et al., 2009). Previously, traditional classification methods, the classic macro-morphology identification method based on flower and fruit characteristics, were the most utilized (Tang and Xiang,

1995). However, the most easily available organ of *Machilus* Nees plants, which are evergreen, is the leaf. In addition, due to the short flowering and fruiting seasons as well as significant biennial bearing, the flowers and fruits of *Machilus* Nees plants are difficult to collect increasing the difficulties of identification and classification for these plants.

Plants can reflect, absorb, and transmit electromagnetic waves. Because of the differences in nutrients, moisture (Li and Song, 2016), and chlorophyll (Song et al., 2008) in the leaves among plants, different plants have different absorption, reflection, and radiation characteristics of electromagnetic waves with different wavelengths, i.e., the differences in chemical components and organizational structure among plants can be represented through spectral data (Lin, 2011).

Recently, hyperspectral technology, which was developed from spectroscopy, having the characteristics of high distinguishability and large information content, may resolve the problem of plants with very similar morphological characteristics not being easy to recognize. For example, Pontius et al. (2005) used this hyperspectral technology to compare a mixed deciduous and coniferous forest as well as a pure forest near Harvard in America; classification processing was conducted, with the classification accuracy reaching 75%. Through the analysis of hyperspectral data during the rice growing period, Zhang et al. (2002) designed a hybrid decision tree classification algorithm for which the total classification accuracy of the test samples in the classification experiment reached 94.9%.

For wood identification, Zhuang et al. (2014) used the infrared spectrum to identify the four kinds of woods in the genera of *Machilus* and *Machilus* Nees; the two genera could be identified in the spectral regions of 1730–1740 cm^{-1} and 1640–1650 cm^{-1}. Through the comparison of extracted and treated wood flour, as well as the infrared spectral changes among the five sandalwood species, Zhang et al. (2014) demonstrated that the infrared spectroscopic technique could be used for the identification of wood species.

To sum up, using hyperspectral technology and the corresponding analytical methods to identify stable and effective spectral characteristics of the leaves of *Machilus* Nees plants could be powerful tools for *Machilus* Nees species identification.

Materials and methods

Instruments and software

The AvaField-1 Portable Hyperspectral Spectrometer (Dutch Avantes Company) with spectral region of 300–1100 nm and spectrum sampling interval of 0.6 nm was used in this experiment. Additional equipment included a blade clip, an optical fiber patch cord for field use, a reflective probe with built-in light resource, a reference white board, and computers. During the spectral determination measurements, correction with a standard white board was conducted after each measurement.

AvaField, AvaReader, and SPSS19 multivariate statistical software were used to process the data.

Sample treatment

Nine species of the genus *Machilus* Nees cultivated in the bonsai garden of College of Horticulture and Landscape Architecture on the west campus of Yangtze University, Hubei Province were selected, including *Machilus multinervia* Liou (A), *Machilus suaveolens* S.Lee (B), *Machilus kwangtungensis* Yang (C), *Machilus decursinervis* Chun (D), *Machilus oculodracontis* Chun (E), *Machilus litseifolia* S. Lee (F), *Machilus pyramidalis* H. W. Li (G), *Machilus microcarpa* Hemsl. (H), and *Machilus longipedicellata* Lec. (I). The above species were planted in fifteen centimeter plastic pots on March 18th, 2016, and placed on the open ground under shade. On September 9th 2016, the second new leaf from the stem tip was collected; a total of 30 leaves were sampled. Twenty training samples and 10 test samples were randomly selected. All samples were stored in valve bags at low temperature with moisture and kept in the dark. Samples were used for the spectrum data collection in the laboratory.

Spectrum data collection

Before data collection, blade surfaces were cleaned to prevent dust on the blades from impacting the measurements. Then, the instrument was connected and adjusted, leaves were placed face up on the reference white board, the probe was pointed at the blade for determination, and the data were read from AvaField on the computer and stored.

Due to interference from ambient illumination, granularity, density, and surface texture in the

environment during the measurements, the original spectral curve would have baseline drift (Chen et al., 2016) and noise. Therefore, the data from below 500 nm and over 1000 nm with high noise were removed, and the readings between 500 nm and 1000 nm with more recognizable spectral signatures were extracted.

Results and discussion

The analysis of the original spectral reflectivity

Fig.1 shows the curves for the average value of spectral reflectivities in the 9 species (in this figure, A, B, C, D, E, F, G, H, and I represent the 9 species of the genus *Machilus* Nees identified above; the abscissa represents the 126 bands between 500 nm and 1000 nm; hereafter, each band represents a corresponding characteristic). This figure reveals that the spectral reflectivities of the 9 *Machilus* Nees species have significant differences among the bands. Near the bands of 544 nm and 908 nm, the differences of spectral reflectivities among 9 species were more significant. When the spectral regions were below 500 nm and above 1000 nm, the reflectivities of the different species had very small differences, and they even crossed or overlapped. Owing to the jumbled information content of the original spectrum, it contained useful information for distinguishing the samples from the other species as well as some irrelevant information. The irrelevant information, including electrical noise, sample background, and stray light, could not be used as effective information for distinguishing species, which would directly negatively impact the modeling precision unless it was excluded.

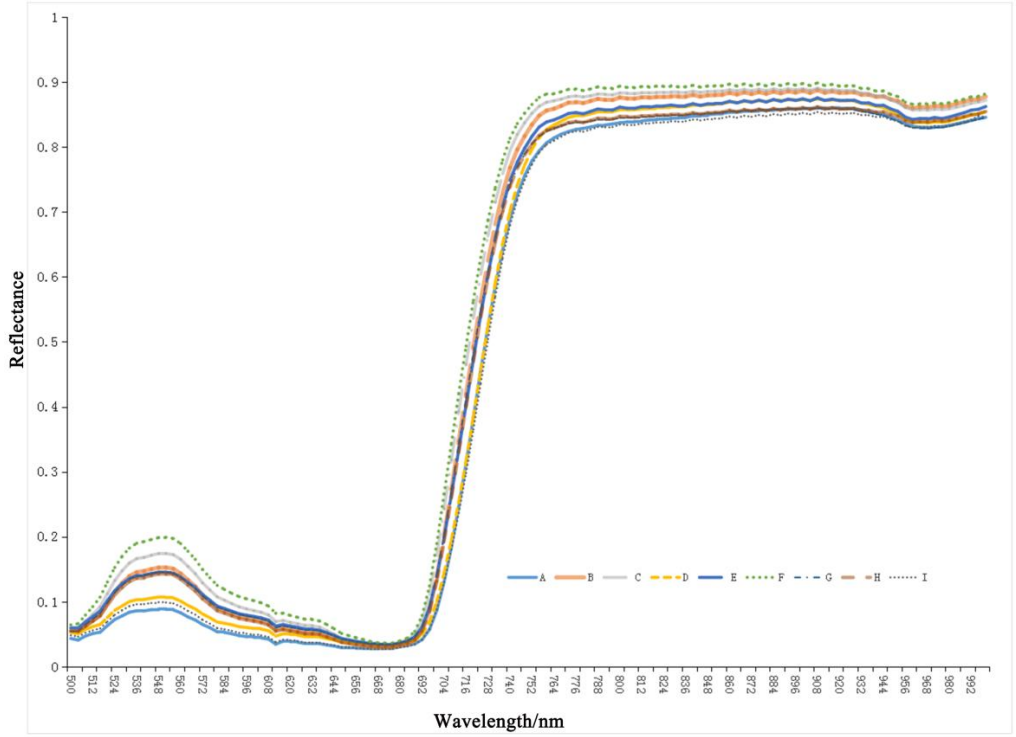


Fig.1: Mean spectral reflectance curve of nine species of *Machilus* Nees.

The selection of unique bands and the discriminant analysis

AvaReader software was utilized for the analysis of the obtained spectrum data. One band was selected every 4 nm for a total of 126 bands, which were equivalent to 126 variables. Each variable had a different function in the recognition of *Machilus* Nees species. If all the data were used for establishing the discrimination function, large computational resources are needed. In addition, because of the autocorrelation which might have existed between variables, the discrimination function lacked stability, and the model error increased. Therefore, it was necessary to conduct band selection.

The stepwise discriminant analysis method using multivariate statistical software was used; unique bands were screened and selected through introducing and

removing F limit. The Bayesian discriminant method was applied to the spectral information of the unique bands to establish the discrimination function, and then the different species of the genus *Machilus* Nees were distinguished. The spectral information of 18 selected unique bands was used to conduct the discriminant analysis for the 9 *Machilus* Nees species and establish the linear discriminant model presented in Table 1. Based on the established discrimination function, validation with 180 training samples involved in modeling and prediction of 90 test samples, which were not involved in modeling, were conducted.

Table 1. Stepwise discriminate function of the nine species.

Species	Functional equation
A	$Y(A) = -6687.52 + 32594.03R_{500} - 32337.24R_{508} - 14654.44R_{520} + 5648.01R_{544} + 4653.83R_{552} + 16060.08R_{556} - 58630.80R_{576} - 95047.69R_{580} + 158579.17R_{588} + 174433.18R_{648} - 283677.01R_{656} + 33146.11R_{664} + 49602.41R_{680} - 342347.46R_{908} + 74006.26R_{912} + 222414.94R_{928} + 32262.79R_{940} + 138836.33R_{944}$
B	$Y(B) = -6782.53 + 53001.07R_{500} - 37931.02R_{508} - 15879.96R_{520} + 23928.94R_{544} + 12648.88R_{552} - 14169.11R_{556} - 77800.24R_{576} - 54066.00R_{580} + 147264.76R_{588} + 154117.48R_{648} - 265715.70R_{656} + 17148.48R_{664} + 48307.32R_{680} - 292683.29R_{908} + 11730.38R_{912} + 218568.27R_{928} + 10840.23R_{940} + 197866.99R_{944}$
C	$Y(C) = -6745.82 + 44672.87R_{500} - 38459.40R_{508} - 14551.86R_{520} + 19894.14R_{544} + 3807.52R_{552} + 2849.52R_{556} - 86425.85R_{576} - 61098.94R_{580} + 159627.44R_{588} + 149902.56R_{648} - 269193.12R_{656} + 32187.66R_{664} + 48192.88R_{680} - 309772.40R_{908} + 26695.56R_{912} + 219035.18R_{928} + 29015.18R_{940} + 180230.83R_{944}$
D	$Y(D) = -6483.49 + 43002.79R_{500} - 32496.76R_{508} - 16313.36R_{520} + 18862.79R_{544} + 8856.61R_{552} - 5212.11R_{556} - 76832.06R_{576} - 52351.68R_{580} + 141299.81R_{588} + 157401.62R_{648} - 257763.16R_{656} + 12154.57R_{664} + 50160.38R_{680} - 307093.94R_{908} + 41500.22R_{912} + 203345.66R_{928} + 30078.56R_{940} + 169203.14R_{944}$
E	$Y(E) = -6785.23 + 56792.41R_{500} - 32315.57R_{508} - 19382.73R_{520} + 31900.29R_{544} + 20588.65R_{552} - 37392.76R_{556} - 63666.83R_{576} - 41350.60R_{580} + 128325.12R_{588} + 177243.00R_{648} - 302744.24R_{656} + 22359.75R_{664} + 50562.22R_{680} - 286677.93R_{908} + 9540.65R_{912} + 210069.49R_{928} + 9796.96R_{940} + 206115.12R_{944}$
F	$Y(F) = -6727.77 + 50356.11R_{500} - 32421.48R_{508} - 18158.88R_{520} + 28562.05R_{544} + 7722.63R_{552} - 14540.93R_{556} - 79307.51R_{576} - 44064.20R_{580} + 138545.41R_{588} + 162667.09R_{648} - 266855.88R_{656} + 4480.63R_{664} + 53897.87R_{680} - 289894.93R_{908} + 16462.82R_{912} + 206647.00R_{928} + 17394.98R_{940} + 194691.62R_{944}$
G	$Y(G) = -6510.77 + 38255.59R_{500} - 33519.06R_{508} - 16741.96R_{520} + 11819.82R_{544} - 2208.88R_{552} + 18093.18R_{556} - 60031.62R_{576} - 91440.55R_{580} + 155227.82R_{588} + 196204.14R_{648} - 315627.80R_{656} + 43084.08R_{664} + 47822.13R_{680} - 325495.66R_{908} + 58491.58R_{912} + 219356.33R_{928} + 33869.70R_{940} + 139324.73R_{944}$
H	$Y(H) = -6527.14 + 50023.26R_{500} - 38625.21R_{508} - 13608.40R_{520} + 22416.15R_{544} + 3581.61R_{552} - 3345.65R_{556} - 69998.61R_{576} - 69761.42R_{580} + 154125.19R_{588} + 147961.78R_{648} - 260225.51R_{656} + 18500.57R_{664} + 50347.11R_{680} - 294692.78R_{908} + 17788.77R_{912} + 210267.89R_{928} + 21320.66R_{940} + 190874.09R_{944}$
I	$Y(I) = -6621.34 + 45400.33R_{500} - 23309.30R_{508} - 21503.98R_{520} + 26865.31R_{544} + 19648.94R_{552} - 29639.75R_{556} - 52719.01R_{576} - 61973.74R_{580} + 133818.91R_{588} + 169717.70R_{648} - 272651.02R_{656} - 467.91R_{664} + 56658.50R_{680} - 299697.46R_{908} + 30218.01R_{912} + 209289.06R_{928} + 11920.20R_{940} + 187866.89R_{944}$

As the results shows in Table 2 below, when 18 band combinations were selected, the discrimination accuracies of the training samples and test samples were as high as 93.889% and 95.556%, respectively. For determining whether the unique bands selected by discriminant analysis were the best band combinations, 6 unique bands were successively added to reconstruct the Bayes discriminant model. Through analysis, for a single species, when 6 unique bands were selected, the discrimination accuracies of the training samples and test samples were as high as 76% and 77%, respectively; when the number of band combinations was increased to 12, the accuracies of the two samples were up to 83% and 84%, respectively. However, when the selected band combinations were less than 18, them both accuracies decreased, indicating that this linear discriminant model established by 18 unique bands in this experiment was satisfactory.

Table 2. Discrimination accuracy rate of model for training and prediction samples.

Band combination	Species	Training samples			Prediction samples		
		Number	Misjudgment	Accuracy%	Number	Misjudgment	Accuracy%
6 band combination	A	20	2	90	10	1	90
	B	20	8	60	10	0	100
	C	20	10	50	10	5	50
	D	20	5	75	10	4	60
	E	20	4	80	10	4	60
	F	20	4	80	10	1	90
	G	20	2	90	10	0	100
	H	20	4	80	10	4	60
	I	20	4	80	10	1	90
12 band combination	A	20	0	100	10	0	100
	B	20	2	80	10	0	100
	C	20	0	100	10	0	100
	D	20	0	100	10	1	90
	E	20	1	95	10	0	100
	F	20	3	85	10	2	80
	G	20	0	100	10	1	90
	H	20	11	45	10	3	70
	I	20	12	40	10	7	30
18 band combination	A	20	0	100	10	0	100
	B	20	1	95	10	0	100
	C	20	0	100	10	0	100
	D	20	0	100	10	1	90
	E	20	0	100	10	0	100
	F	20	0	100	10	0	100
	G	20	0	100	10	0	100
	H	20	8	60	10	2	80
	I	20	2	90	10	1	90

Conclusion

Spectral data for new leaves of one-year-old seedlings of 9 species of the genus *Machilus* Nees were measured with an instrument. The irrelevant information was removed, and stepwise discriminant analysis was utilized for screening the data. Eighteen unique bands were selected from 126 bands: 500 nm, 508 nm, 520 nm, 544 nm, 552 nm, 556 nm, 576 nm, 580 nm, 588 nm, 648 nm, 656 nm, 664 nm, 680 nm, 908 nm, 912 nm, 928 nm, 940 nm, 944 nm, and 956 nm. These 18 unique bands were used to establish the discriminant analysis model for species of the genus *Machilus* Nees through the Bayesian discriminant method. The correct recognition rate of this model achieved up to 93.889% on training samples and 95.556% on test samples in 9 *Machilus* Nees species. This model reached higher discrimination precision, indicating that this technology could be applied in the identification and classification of *Machilus* Nees species.

Because of the difficulties in obtaining flowers and fruits of *Machilus* Nees plants, traditional classification methods using characteristics of flower and fruits for discrimination were not chosen in this study; in contrast, leaves, which were more easily obtained, were chosen for analysis. The leaves of *Machilus* Nees plants could be readily collected and could be analyzed directly after cleaning without complicated treatments. This method increases the efficiency, decreases the experimental

error, causes very little damage to plants, and has small effect on the regular growth and development of plants; therefore, it is an ideal method for the identification of *Machilus* Nees plants.

Conflict of interest statement

Authors declare that they have no conflict of interest.

Acknowledgement

This work was supported by the Science and Technology Support Program in Hubei Province of P. R. China (2013BBB24), the Youth fund program in Yangtze University of P. R. China (2015cqn80), University Student Innovation Program in Yangtze, Hubei Province of P. R. China (20150118).

References

Chen, J.Q., Li, L., Li, J., 2009. Bayesian inference of nrDNA ITS sequences from *Machilus* (Lauraceae) and its systematic significance. Acta Bot. Yunnanica. 31(2), 117-126.

Chen, Z.G., Li, X., Fan, X.J., 2016. Method for the discrimination of the variety of potatoes with vis/nir spectroscopy. Spectrosc. Spect. Anal. 36(8), 2474-2478.

Li, J., Song, W.L., 2016. Water content model for strawberry leaves with spectral signature. J. Northeast Forestry Univ. 44(1), 72-74.

Li, X. W., Li, J., Huang, P. H., 2008. Lauraceae. In: Flora of China (Eds.: Wu, Z.Y., Raven, P.H., Hong, D.Y.). Vol. 7. Science Press, Beijing and Missouri Botanical Garden Press, St. Louis. pp.102-254.

Lin, H., Zang, Z., Liu, X.Y., 2011. Hyperspectral Remote Sensing of Forest Tree Species. China Forestry Press, Beijing. pp.110-119.

Lin, X.Z., 2007. Quantitative classification of plants of *Machilus* in Zhejiang Province. Sci. Silvae Sin. 43(11), 151-156.

Liu, Z.X., Fei, Y.J., 2011. Germplasm resources and conservation countermeasures on the *Phoebe* Nees and *Machilus* Nees species in China. J. Yangtze Univ. 8(5), 221-223.

Pontius, J., Hallett, R., Martin, M., 2005. Using AVIRIS to assess hemlock abundance and early decline in the Catskills, New York. Remote Sens. Environ. 97, 163-173.

Song, X.D., Jiang, H., Yu, S.Q., 2008. Relationship between chlorophyll concentrations and spectral reflectance feature of the typical evergreen hardwood species in subtropical region of China. Acta Ecol. Sin. 28(5), 5-10.

Tang, G.G., Xiang, Q.B., 1995. Pollen morphology of the family Lauraceae in China. Acta Phytotaxon. Sin. 33(2), 161-170.

Wang, Y. G., Wei, F. N., 2003. Comparative study of pollen morphology in *Machilus* and its related genus *Phoebe* under SEM. Guihaia. 22(1), 27-30.

Wei, F.N., Tang, S.C., 2006. A revision on several species of *Machilus* Nees (Lauraceae) from China and Vietnam. Guihaia. 26(4), 345-348.

Zhang, F., Xiong, Z., Kou, N., 2002. Airborne hyperspectral remote sensing image data is used for rice precise classification. J. Wuhan Univ. Technol. 24(10), 36-39.

Zhang, R., Xu, K. W., Zhang, L. S., 2014. Identification of five rosewood species by infrared spectrum characteristics. China Forestry Sci. Technol. 28(2), 95-99.

Zhong, Y., Xia, N. H., 2010. Leaf epidermal features of *Machilus* (Lauraceae) form China and their systematic significance. J. Trop. Subtrop. Bot. 18(2), 109-121.

Zhuang, L., Huang, Q., Xu, Y. H., 2014. Four kinds of *Phoebe* and *Machilus* wood infrared absorption spectrum identification. J. Fujian Forestry Sci. Technol. 41(4), 21-25.

How to cite this article:
Fei, Y., Xiao, B., Liao, X., Wen, S., Zhou, B., Pan, J., Hu, D., 2017. The identification method for *Machilus* Nees species based on visible-near infrared spectral analysis technology. Int. J. Curr. Res. Biosci. Plant Biol. 4(4), 1-6.
doi: https://doi.org/10.20546/ijcrbp.2017.404.001