

Laser-induced Fluorescence Spectroscopy (LIFS) for Discrimination of Genetically Close Sweet Orange Accessions (*Citrus sinensis* L. Osbeck)

Thiago Massaiti Kuboyama Kubota^{1,2},
Aida Bebeachibuli Magalhães¹, Marina Nery da Silva¹,
Paulino Ribeiro Villas Boas¹, Valdenice M. Novelli³,
Marinês Bastianel³, Cíntia H. D. Sagawa³,
Mariângela Cristofani-Yaly³, and
Débora Marcondes Bastos Pereira Milori¹

Applied Spectroscopy
2017, Vol. 71(2) 203–214
© The Author(s) 2016
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0003702816658668
journals.sagepub.com/home/asp


Abstract

Although there is substantial diversity among cultivated sweet oranges genotypes with respect to morphological, physiological, and agronomic traits, very little variation at DNA level has been observed. It is possible that this low DNA molecular variability is due to a narrow genetic basis commonly observed in this citrus group. The most different morphological characters observed were originated through mutations, which are maintained by vegetative propagation. Despite all molecular tools available for discrimination between these different accessions, in general, low polymorphism has been observed in all groups of sweet oranges and they may not be identified by molecular markers. In this context, this paper describes the results obtained by using laser-induced fluorescent spectroscopy (LIFS) as a tool to discriminate sweet orange accessions (*Citrus sinensis* L. Osbeck) including common, low acidity, pigmented, and navel orange groups, with very little variation at DNA level. The findings showed that LIFS combined with statistical methods is capable to discriminate different accessions. The basic idea is that citrus leaves have multiple fluorophores and concentration depends on their genetics and metabolism. Thus, we consider that the optical properties of citrus leaves may be different, depending on variety. The results have shown that the developed method, for the best classification rate, reaches an average sensitivity and specificity of 95% and 97.5%, respectively. An interesting application of this study is the development of an economically viable tool for early identification in seedling certification, in citrus breeding programs, in cultivar protection, or in germplasm core collection.

Keywords

Laser-induced fluorescence spectroscopy, LIFS, sweet orange accession, citrus leaf fluorophores, differentiation of very close orange varieties, instrumentation

Date received: 15 October 2015; accepted: 2 June 2016

Introduction

The sweet orange (*Citrus sinensis* L. Osbeck) is the most representative and recognizable species of the citrus group. Its production in 2012 was estimated at 51.8 million tons, totalling approximately 53% of citrus production for both fresh fruit and processed juice consumption. Brazil is the world's biggest producer of sweet oranges, which corresponds to about one-third of the world's production, that is, approximately 18 m/t.¹

¹Embrapa Instrumentation, São Carlos, SP, Brazil

²Sao Carlos Institute of Physics, University of Sao Paulo, São Carlos, SP, Brazil

³Centro de Citricultura Sylvio Moreira of Instituto Agrônomo (CCSM-IAC), Cordeirópolis, SP, Brazil

Corresponding author:

Aida Bebeachibuli Magalhães, Rua XV de Novembro, 1452, Centro, P.O. Box 741, São Carlos, SP, 13561-206, Brazil.
Email: habibe.aida@gmail.com

The sweet orange originates from Asia and several research results suggest the hybrid originates from crosses between the pomelo (*C. grandis* L. Osbeck) and the mandarin (*C. reticulata* Blanco).^{2–5} According to Hodgson,⁶ sweet oranges are classified into four groups: common, low acidity, pigmented, and navel oranges. In these groups, there is a wide diversity in phenotype traits such as size and shape of canopy, color, size, type, ripening season of the fruits, and the number of seeds per fruit. Although there is substantial diversity among cultivated genotypes in terms of morphological, physiological, and agronomic traits, very little variation at DNA level has been observed.^{7,8} It is possible that this low DNA molecular variability is due to a narrow genetic basis commonly observed in these citrus groups. Most of the different morphological characters observed originated through mutations, which are maintained by vegetative propagation.⁹

Several DNA markers have been utilized in genetic studies, cultivar characterization, and identification of sweet oranges, such as variable number of tandem repeats (VNTRs) loci,¹⁰ inter-simple sequence repeats (ISSRs) and restriction fragment length polymorphisms (RFLPs),^{8,11} random amplified polymorphic DNA (RAPD),^{12,13} isozymes,¹⁴ single nucleotide polymorphism (SNP),¹⁵ and simple sequence repeats (SSR).^{7,16} However, despite all these molecular tools, in general, low polymorphism has been observed in all groups of sweet oranges.

Sun et al.¹⁷ reported a new strategy cultivar identification diagram (CID) that was used with success to improve the efficiency of RAPD markers for the identification of 60 sweet orange cultivars. Recently, another technique called laser-induced fluorescence spectroscopy (LIFS) was employed as a tool to identify citrus varieties.^{18,19} The results showed that this technique was able to discriminate the different citrus varieties in a nursery. Milori et al.¹⁸ used an excitation laser of 561 nm and showed that LIFS combined with statistical methods is able to discriminate different citrus varieties. Several statistical tests, such as principal component regression (PCR), partial least squares regression (PLS), and factorial discriminant analysis (FDA), were assessed as classification methods for each sample class. The rate of success for this classification depends on the combination of canopy and rootstock varieties as well as the classification method used. The best results are given by the classic principal component analysis (PCA) and PLS-LDA classifiers (around 90%). However, FDA using Mahalanobis distance showed the best success rate for classification of the Natal variety (around 81%). In previous work described in Santana-Vieira et al.,¹⁹ two different excitation lasers, 405 nm and 561 nm, aimed to determine LIFS accuracy in the differentiation and grouping very closely varieties of four Sunki mandarin selections: Comum, Florida, Tropical, and Maravilha. The results obtained with LIFS was compared with the ones obtained with ISSR and SSR molecular markers for the same varieties.

Laser-induced fluorescence spectroscopy distinguished the four selections with an accuracy greater than 70% while molecular markers were able to distinguish clearly Tropical from Maravilha, but not Comum from Florida selections.

Laser-induced fluorescence spectroscopy is an analytical technique²⁰ used to monitor chemical and physical properties of plants which access not only color pigment such as chlorophyll and carotenoid,^{21,22} but also other secondary metabolites of the leaves, which comprise the fluorophore of the green leave. Thus, the fluorescence emission spectrums of the leaves are rich in information of various chemical compounds.

The fluorescence spectroscopy on the leaves of plants has emerged as a specific and important tool for remote sensing to access accurately the physiological state of plants and allow early diagnosis of biotic and abiotic stresses in vegetation.^{23,24} Their applications in agriculture, biology, and botanics has been published in a numerous of reviews in the last 20 years.^{18,23–29} Laser-induced fluorescence spectroscopy has an optical configuration which allows non-invasive interaction with the sample. Thus, the tool can be made very compact for the development of field devices and incorporated in agricultural vehicles. Such equipment allows large-scale and real-time monitoring as for citrus early identification in breeding programs and early biotic and abiotic stress diagnosis in crops. Beside, other features such as high sensitivity and reliability made LIFS an attractive investigation tool for the productive sector

In this context, this paper describes the results obtained by using LIFS associated with chemometric methods as a tool to discriminate sweet orange accessions (*C. sinensis* L. Osbeck) including common, low acidity, pigmented, and navel orange groups. The main features from these set of leaves is that they are indistinguishable by molecular markers. Even with little genetic variability among them, each accession is characterized by a specific combination of canopy and rootstock and presents a particular concentration of multiple fluorophores and chemical composition. These changes are successfully identified by optical techniques because it alters the characteristic optical signature of the leaves from each accession and discriminate them among its genetics.^{9,18,19} Two different LIFS systems, one with excitation wavelength at 405 nm and the other one with excitation wavelength at 561 nm, combined with chemometric tools, were used to discriminate these sweet oranges varieties. The evaluation was performed in two different sets of leaves, from ten different citrus varieties in each set collected in two different seasons: the first one in winter time and the second one in summer time. Both sample sets were received from Citrus Germplasm Bank of Centro de Citricultura Sylvio Moreira of Instituto Agronômico (CCSM-IAC), Cordeirópolis, SP, Brazil.

Table 1. First set of sweet orange (*C. sinensis*) accessions evaluated by spectroscopy.

ID	Groups		
	Navel	Common names	
BAG Citros 0006	Common	Baia Vale del Cauca	
BAG Citros 0008		Baianinha Piracicaba	
BAG Citros 0119		Pera Ipigua	
BAG Citros 0227		Ovale	
BAG Citros 0273		Pera Mahle	
BAG Citros 0295		Pera Olímpia	
BAG Citros 1546		Pera de Abril IAC 148	
BAG Citros 1557		Valencia B	
BAG Citros 1589		Westin IAC 115	
BAG Citros 1640		Seleta do Rio IAC 420	
BAG Citros 1667		Pera Roberto Gullo	
BAG Citros 1703		Pera Bianchi	
BAG Citros 0277		Pigmented	Red Pulp
BAG Citros 0278			Red Pulp de Mombuca
BAG Citros 0279			Sanguinelli- Faz. Reserva
BAG Citros 0280			Doppio Red Pulp
BAG Citros 1708			Valencia Puka
BAG Citros 1526	Low acidity	Piralima IAC 2	
BAG Citros 1538		Lima IAC 9	
BAG Citros 1701		Piralima IAC 11	

Materials and Methods

Plant Material

The sweet orange accession (*C. sinensis* L. Osbeck) belongs to the Citrus Germplasm Bank of Centro de Citricultura Sylvio Moreira of Instituto Agronômico (CCSM-IAC), Cordeirópolis, SP, Brazil. A total of 20 accessions of sweet orange were evaluated (Table 1) in two different seasons: the first set comprises leaves from ten different accessions collected in winter and the second set comprises leaves from another ten different citrus accessions collected in summer; they were separated into groups according to the measurements presented in Tables 2 and 3, respectively. All plants were grafted on *C. limonia* and cultivated in 60-L citrus pots containing commercial substrate on greenhouse conditions. All parameters such as soil management, irrigation, fertilization, temperature, and watering were adequate and favorable to growing the plants. These parameters were considered for each tree. The plants in a Germplasm Bank are not in the seedling phase or even adults. All plants were about 3 years old. Besides, no trees already showed any difference in phenotypic traits, which did not allow differentiation

Table 2. Separated groups of the first set of citrus accessions analyzed by LIFS.

Group	Accessions
Group 1	Pera Ipigua (BAG Citros 0119), Pera Mahle (BAG Citros 0273), and Pera Olímpia (BAG Citros 0295)
Group 2	Sa (BAG Citros 0277), Red Pulp of Mombuca (BAG Citros 0278), Sanguinelli – Faz. Reserva (BAG Citros 0279), and Doppio Red Pulp (BAG Citros 0280)
Group 3	Baianinha Piracicaba (BAG Citros 0006), Baia Vale del Cauca (BAG Citros 0008), and Ovale (BAG Citros 0227)

Table 3. Separated groups of the second set of citrus varieties analyzed by LIFS technique.

Group	Accessions
Group 1	Pera Abril IAC 148 Pera Bianchi Pera Roberto Gullo
Group 2	Seleta do Rio IAC 420 Valência B Valência Puka Westin IAC 115
Group 3	Lima IAC 9 Piralima IAC 11 Piralima IAC 2

between them by visual inspection. Figure 1 showed a protected BAG as maintained in the Germplasm Bank of CCSM-IAC.

Leaf samples

Three leaves of each accession were collected at CCSM-IAC and sent to Embrapa Instrumentation to be analyzed. First, these samples were cleaned with a piece of cotton made wet with distilled water and then dried with dry cotton wool in order to remove any soil or dirt that could affect and influence the LIFS spectra. After that, the leaves were kept in bags and refrigerated at 4 °C to prevent degradation. All measurements were performed with the leaves in natura²⁵ (without any kind of sample preparation) as shown in Figure 2. Measurements were performed within 24h after arrived at Embrapa Instrumentation. Accessions from the same variety were not measured in sequence to avoid interference in the reading of the equipment. The measurements were done on the abaxial leaf surface, on the right side of the midrib, next to the petiole.



Figure 1. Protected BAG as maintained in the Germplasm Bank of Centro de Citricultura Sylvio Moreira of Instituto Agronômico (IAC), Cordeirópolis, SP, Brazil.



Figure 2. All measurements were performed with the leaves *in natura* positioning the optical fiber probe perpendicularly on its ventral portion near the midrib, taking care that light does not focus directly on the veins of the leaf.

Laser-Induced Fluorescence Spectroscopy Apparatus

In this study, we used two portable LIF systems. Both systems were developed by the Optics and Photonics Laboratory from Embrapa Instrumentation (international patent WO2010069017A1).³⁰ The main difference between these two systems are excitation frequency and detection range of spectrometers. The first system has a diode laser CUBE Model produced by Coherent with excitation at 405 nm and maximum output power of 50 mW at room temperature, and a high sensitivity mini-spectrometer

manufactured by Ocean Optics (USB4000) with spectral range of 194–894 nm.^{18,19,25,30} It will be addressed as LIFS-405 from now on. Referred to as LIFS-561, the second system comprises a diode laser COMPASS model, also produced by Coherent, but with emission at 561 nm. LIFS-561 has a high sensitivity mini-spectrometer (USB4000) produced by Ocean Optics with detection range of 500–1200 nm.^{18,19} Both systems have an optical shutter, a bifurcated optical fiber (six illumination optical fibers around one reader fiber is also manufactured by Ocean Optics). There is an adjustable optical filter to minimize the scattered excitation light and a notebook with software developed to control, acquire, and evaluate the spectral signal.

The developed software, DIAGNÓSTICO, adjusts three parameters of the spectrometer in both pieces of LIFS equipment: integration time, average number, and boxcar. The integration time (in ms) is the time interval in which the system captures the light emission from each sample. The average number is the number of collected spectra to obtain the average spectrum which best represents the emission from the sample in order to minimize noise. The third parameter is the boxcar, which consists of a pre-treatment parameter used for smoothing the spectra. This smoothing is carried out by evaluating a medium at each point of the spectrum that involves a number of neighbors on either side of the respective point. In this work, the measurements performed with the LIFS-405 system adopted the following parameters: 60 ms for integration time, 20 for the spectral averages, and 2 for boxcar. For the measurements performed with the LIFS-560 nm, the spectrometer parameters adopted were: 2 ms for

integration time, 20 for the spectral averages, and 2 for boxcar.^{18,19}

The experimental procedure consists of positioning the optical fiber probe perpendicularly on the ventral portion of the leave *in natura* near the midrib, making sure that the light does not focus directly on the veins of the leaf. Before the measurement, each sample was exposed to excitation source for about 2 s, which is enough to reach fluorescence steady state.

Data Evaluation

In this work, the PCA^{31,32} was evaluated to study the potential to differentiate sweet orange accessions using the fluorescence spectra and create classifiers able to correctly identify each variety. Principal component analysis was evaluated on the entire fluorescence spectrum information only for the Pera sweet oranges accession as input variable and the PC scores delineated clearly the different accession. Once the discriminatory power of LIFS were shown, the spectra from all accession were analyzed and classifier induced as described below.

First, all spectra were pre-treated to remove electronic and optical offset, and normalized by the area under the curve, so only changes in the spectral profile will be considered. This pre-treatment enables the comparison among several spectra. For LIFS-405, the baseline was corrected by subtracting the average value of the intensities in the spectral range of 425–430 nm from the whole collected spectra, 400–892 nm. For LIFS-560, the baseline was corrected by subtracting the average value of intensities in the spectral range of 1090–1100 nm of the whole collected spectra, 600–1140 nm. All this analysis was performed with software developed by the group named DIAGNÓSTICO.

Second, these data were exported to an open source software Weka for the induction of a classifier via regression (CVR) based on the combination of two techniques: a model tree algorithm and a regression function, via PLSR.^{25,33–36} This induced classifier is based on chemical changes in citrus plants of different accessions. These changes are identified by leaf fluorescence, which offers information that is used to construct the classifier through a confusion matrix obtained via CVR. For our fluorescence data, 1593 points were used for LIFS-405 and 1680 points for LIFS-560 as input to the classifier for data training and data validation.

Partial least squares regression is a multivariate analysis technique that makes a linear transformation on the original data set, the input variables (fluorescence spectrum), and determines the best correlation with the response variables, namely the sweet oranges accessions. This technique has been widely used in the classification of diseases in citrus²⁵ and has shown to be very useful in the study of separating citrus varieties which are genetically very close. The classifiers induced by PLSR associate the sweet orange

accessions to numbers using a binarization process. The process starts with the selection of a training set, to calibrate the regression model, and a test set, to validate the generated model. Thus, value 1 is assigned for one accession of sweet orange and value 0 is assigned for the other sweet orange accessions. For each test set, the model shows values between 0 and 1. The closer the number is to 1 the stronger the similarity between the test spectrums to the reference class is. In order to validate the classification model generated by CVR, the cross-validation method was performed.³⁶ This method separates the data set in n folds or groups, and uses $(n-1)$ of them for training the classifiers and one for testing it. Then, the folder chosen for testing the classifier was returned to the data set. Iteratively, another fold was chosen for testing the classifier and the remaining folders composed the training group so the same procedure is adopted. This process was repeated until all groups are trained and tested.

The classification of the spectral data sets was carried out with ten randomly chosen cross-validation procedures of tenfold. For all tests, ten components were assigned. The choice of the number of components was made by analyzing the highest success rate and lowest root mean squared error (RMSE) as a function of the number of components with significant statistical differences. The success rate was obtained by the ratio of the number of correctly classified samples and the total number of samples.

With the free software Weka, the spectra treatment obtained with LIFS-405 and LIFS-560 were conducted and, as a result, it was possible to mount a confusion matrix. The columns of the confusion matrix show how the leaves were classified by the generated model and the rows show the actual varieties. Thus, the main diagonal shows the correctness of the model. The values represented in the confusion matrix correspond to the rights and wrongs obtained in the classification of the accessions in percentage.

Results and Discussion

Despite all phenotypic differences between the sweet orange groups, there is very little genotype variability between them, which cause little success in using genetic markers to characterize and identify each sweet orange accession.¹⁶ For this purpose, LIFS was successful, as this technique is based on the differences of the fluorescence spectra characteristic for each orange accession. This separability is due to chemical changes in citrus leaves identified by leaf fluorescence that is specific for each combination of canopy and rootstock.

Typical leaf fluorescence spectra obtained with LIFS-405 and LIFS-561 are shown in Figures 3 and 4. According to Figure 3, the fluorescence spectra with LIFS-405 show two distinct fluorescence bands: a blue–green emission between 400 and 630 nm, and a red to near-infrared (NIR) in the

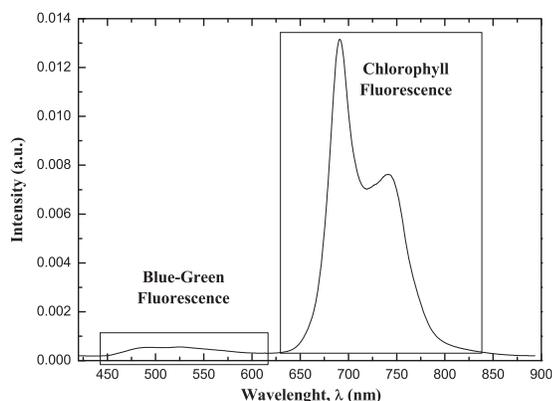


Figure 3. Typical fluorescence spectra of a leaf obtained with an excitation light of 405 nm. This spectrum displays two different bands that correspond to the blue–green fluorescence, in the range of 450–630 nm, and the characteristic chlorophyll fluorescence in the red to NIR region, in the range of 650–800 nm.

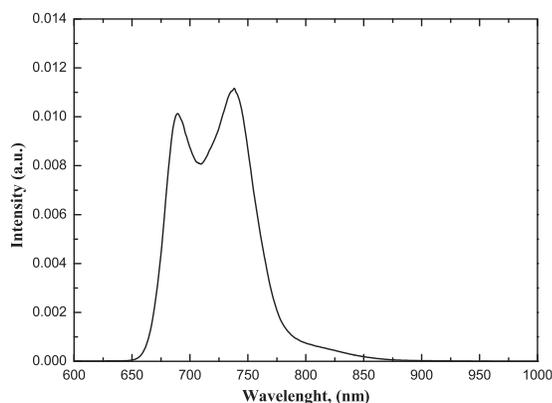


Figure 4. Typical fluorescence spectra of a leaf obtained with an excitation light of 561 nm. This spectrum displays only the characteristic chlorophyll fluorescence in the red to NIR region, in the range of 650–800 nm.

range of 650–800 nm.²⁵ Both play an important role in the construction of the induced classifier for citrus discrimination of different accession. The first band is strongly associated with the carotenoids, plant cell wall, and with the presence of secondary metabolites such as ferulic acid, flavonoids, coumarin, and quercetin.^{23–25,37–40} The second emission band possesses two peaks, one at 680 nm and another at 730 nm. They represent the characteristic emission of chlorophyll fluorescence (ChlF).^{23–25,37–40} Figure 4 shows the typical fluorescence spectra obtained with the LIFS-561 system, where the two emission peaks, at 680 nm and 730 nm, are characteristic of chlorophyll fluorescence emission (ChlF).^{26–29,40}

Figures 5 and 6 show the score plot for three PCs, $PC1 \times PC2 \times PC3$, of PCA analysis using the fluorescence spectra (LIFS-405 and LIFS-561, respectively) of leaves from

different Pera sweet orange accessions grafted in Rangpur lemon rootstock. One may see the existence of patterns in both graphics.

Although all sweet oranges are very close genetically, it is possible to observe three separate groups using both LIFS systems for Pera accession. Accordingly, an induced classifier will be able to be constructed with the aid of all the fluorescence spectra information to better characterize the leaves accession. In the next section, the induced classifier and the success rate are presented. The classifications of each group of leaves cited in Tables 2 and 3 are shown using the methods described above. A set of samples of leaves from different accessions, called set 1, was measured by both LIFS-405 and LIFS-560, and another set of leaves (set 2) was measured only with the LIFS-405 system.

First Set of Leaves Analyzed

After performing ten executions of tenfold stratified cross-validation runs to determine the number of components responsible for the best results of classification, a confusion matrix for all varieties was obtained by classification via regression using PLS with ten components. For the correct classification of ten accessions together, a global accuracy of 63% for LIFS-405 was obtained. Regarding LIFS-560, the result was 60.5%. This may be contemplated in the confusion matrix in Tables 4 and 5, respectively.

Analyzing the main diagonal of the confusion matrix, it is possible to observe the great potential of the technique once it was capable to separate all the accessions. From the confusion matrix of Tables 4 and 5, it is possible to infer that the classification error of the model is associated with an increase in the number of sweet orange accessions. This increase causes more confusion among them. Furthermore, as shown in both tables, the success rate for this classification depends on each combination of canopy and rootstock, and the best result obtained was 96.8% with LIFS-405 system for the Ipiqua orange. Considering the same variety, these results drops to 71% when the LIFS-560 technique is applied, as shown in Table 5. The best result obtained with the LIFS-560 technique concerns Pera Mahle, with a success rate of 94% in correctly classifying it from all other accession (Table 5).

In order to improve the discrimination among the sweet orange accessions, ten accessions from the first set were separated in three different groups according to their similarities, as shown in Table 2. The confusion matrix obtained for the first group (Pera Ipiqua, Pera Mahle, and Pera Olimpia) evaluated the fluorescence spectra obtained from LIFS-405 and LIFS-560. They are presented in Tables 6 and 7, respectively. According to these tables, it is possible to differentiate three sweet orange accessions, although they are genetically similar. Observing the confusion matrix of Table 6, the technique is able to differentiate three different accessions of Pera. Pera Olimpia was

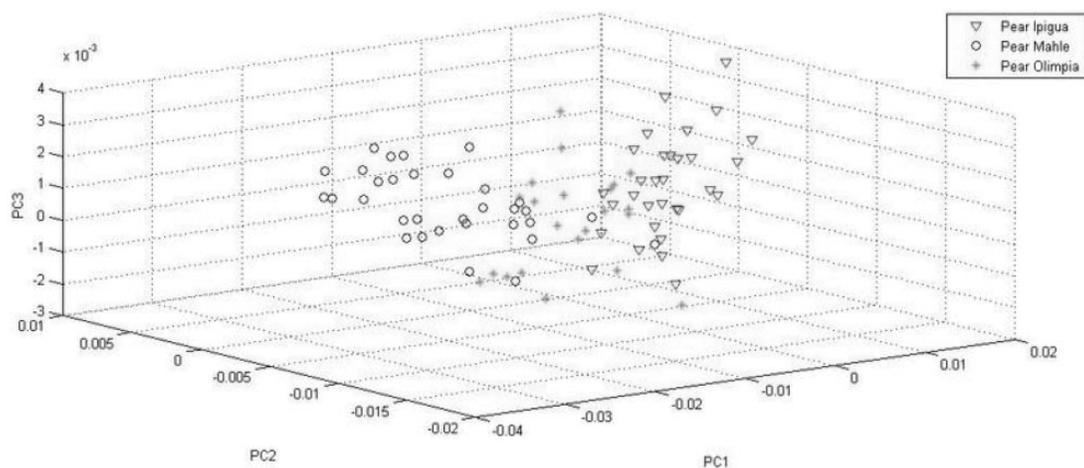


Figure 5. Principal component analysis using the fluorescence spectra (FIL-405) of leaves among Pera sweet orange accessions grafted in Rangpur lemon rootstock.

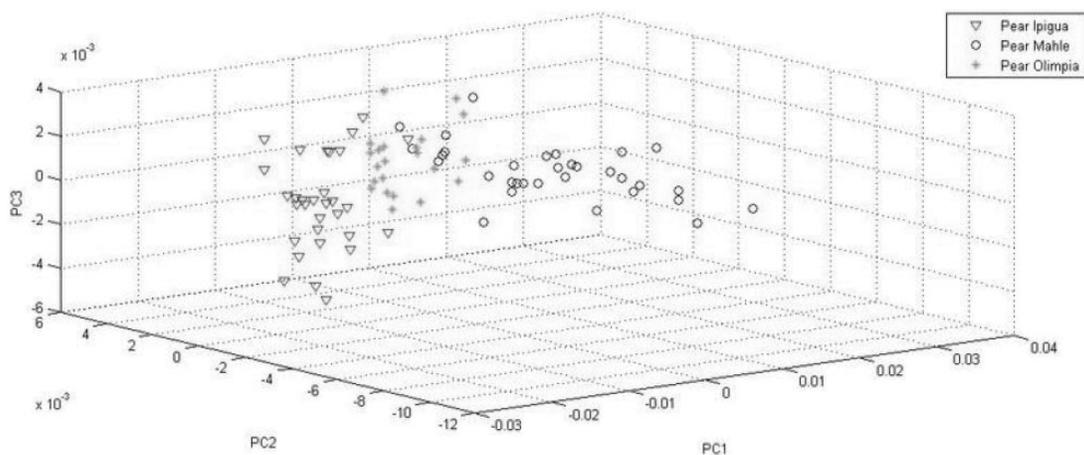


Figure 6. Principal component analysis using the fluorescence spectra (LIFS-561) of leaves from different Pera sweet orange accessions grafted in Rangpur lemon rootstock.

Table 4. Confusion matrix obtained when evaluating all ten accessions using the LIFS-405 technique.

Real variety \ Classified as	Classified as									
	Baianinha	Doppio	Ovale	Ipigua	Mahle	Olimpia	Sanguinea	Sanguinelli	Mombuca	Del Cauca
Baianinha	52.2%	8.8%	3.2%	0.0%	3.2%	0.0%	4.2%	0.0%	0.0%	16.6%
Doppio	4.3%	26.5%	0.0%	3.2%	0.0%	0.0%	0.0%	0.0%	0.0%	6.6%
Ovale	4.3%	0.0%	48.4%	0.0%	0.0%	4.2%	4.2%	3.2%	4.3%	3.3%
Ipigua	8.7%	0.0%	16.1%	96.8%	0.0%	0.0%	12.5%	3.2%	0.0%	6.6%
Mahle	0.0%	17.6%	3.2%	0.0%	90.3%	0.0%	4.2%	16.1%	4.3%	0.0%
Olimpia	0.0%	17.6%	6.5%	0.0%	0.0%	65.2%	0%	6.4%	4.3%	0.0%
Sanguinea	8.7%	11.7%	0.0%	0.0%	0.0%	13.0%	62.5%	0.0%	0.0%	10.0%
Sanguinelli	8.7%	11.7%	9.7%	0.0%	3.2%	17.4%	0.0%	61.3%	12.9%	0.0%
Mombuca	4.3%	2.9%	3.2%	0.0%	3.2%	0.0%	4.2%	6.4%	70.0%	0.0%
DelCauca	8.7%	2.9%	9.7%	0.0%	0.0%	0.0%	8.4%	3.2%	4.3%	56.6%

Table 5. Confusion matrix obtained when evaluating all ten accessions using the LIFS-560 technique.

Real variety \ Classified as	Classified as									
	Baianinha	Doppio	Ovale	Ipigua	Mahle	Olimpia	Sanguinea	Sanguinelli	Mombuca	Del Cauca
Baianinha	47.8%	0.0%	3.2%	3.2%	0.0%	0.0%	4.2%	8.7%	16.1%	16.7%
Doppio	13.0%	47.0%	0.0%	0.0%	0.0%	26.0%	16.7%	0.0%	9.7%	0.0%
Ovale	8.7%	8.8%	61.3%	6.4%	0.0%	4.3%	12.5%	0.0%	9.7%	3.3%
Ipigua	8.7%	0.0%	16.3%	71.0%	0.0%	0.0%	25.0%	0.0%	0.0%	13.3%
Mahle	0.0%	11.7%	3.2%	0.0%	94.0%	4.3%	0.0%	4.3%	29.0%	3.3%
Olimpia	0.0%	17.7%	0.0%	6.4%	3.2%	39.2%	0.0%	4.3%	0.0%	0.0%
Sanguinea	0.0%	0.0%	3.2%	9.7%	0.0%	8.7%	25.0%	0.0%	3.2%	0.0%
Sanguinelli	4.3%	5.9%	9.7%	0.0%	0.0%	0.0%	8.3%	61.0%	6.4%	3.3%
Mombuca	4.3%	0.0%	0.0%	0.0%	3.2%	17.4%	4.2%	14.4%	25.4%	16.6%
Del Cauca	13.0%	8.8%	3.2%	3.2%	0.0%	0.0%	4.2%	4.3%	0.0%	43.4%

Table 6. Confusion matrix obtained by analyzing the spectral information obtained with LIFS-405 technique.

real variety \ Classified as	Classified as		
	Ipigua	Mahle	Olimpia
Ipigua	94%	0	6%
Mahle	0	91%	9%
Olimpia	0	0	100%

Table 7. Confusion matrix obtained by analyzing the spectral information obtained with LIFS-560 technique.

real variety \ Classified as	Classified as		
	Ipigua	Mahle	Olimpia
Ipigua	94%	0	6%
Mahle	0	91%	9%
Olimpia	9%	4%	87%

differentiated from the other two accessions of Pera, Ipigua, and Mahle with a success rate of 100%. For the same group of accession, the model was able to separate and classify three accessions of sweet orange Pera using LIFS-560. It is shown in the main diagonal of the confusion matrix presented in Table 7. Higher confusion was observed in Table 7 than in Table 6. Consequently, with the LIFS-560 system, the accuracy decreases slightly and the highest success rate obtained is 94%, for Pera Ipigua.

The performance of methods based on the combination of two techniques—a model tree algorithm and a regression function, via PLSR—was measured through two parameters: sensitivity and specificity for discriminate the sweet oranges accession. The best confusion matrix (Table 6) was considered. For the evaluation of both parameters the classified samples were separated in terms of true positive (TP), true negative (TN), false negative (FN),

Table 8. The sensitivity and specificity obtained for the classification result among the pear variety presented in Table 6.

	Ipigua	Mahle	Olimpia
Sensitivity	94%	91%	100%
Specificity	97%	95.5%	100%

Table 9. Confusion matrix obtained by analyzing the spectral information obtained with LIFS-405 technique for pigmented sweet oranges Doppio, Sanguinea, Sanguinelli, and Mombuca.

real variety \ Classified as	Classified as			
	Doppio	Sanguinea	Sanguinelli	Mombuca
Doppio	65%	18%	5%	12%
Sanguinea	8%	92%	0%	0%
Sanguinelli	16%	10%	68%	6%
Mombuca	9%	4%	4%	83%

and false positive (FP). Table 8 shows the sensitivity and specificity for the classification rate obtained in the matrix shown in Table 6. The results have shown that the developed method, for the best classification rate, reaches an average sensitivity and specificity of 95% and 97.5%, respectively.

The second set of sweet oranges assembled is the pigmented accession (Doppio, Sanguinea, Sanguinelli, and Mombuca). Tables 9 and 10 show the confusion matrix obtained for these accessions. The global accuracies were 77% and 68% for the evaluation done with LIFS-405 and LIFS-560, respectively. Sanguinea got the best success rate, 92%, in differentiation among the other three classes, when the LIFS-405 system is used. With the aid of LIFS-560 system, the best success rate is 77% for Doppio.

For the last set of orange accession, consisting of Baianinha, Ovale, and Del Cauca, the global accuracy

Table 10. Confusion matrix obtained by analyzing the spectral information obtained with LIFS-560 technique for pigmented sweet oranges Doppio, Sanguinea, Sanguinelli, and Mombuca.

Real variety	Classified as			
	Doppio	Sanguinea	Sanguinelli	Mombuca
Doppio	77%	9%	6%	8%
Sanguinea	17%	71%	4%	8%
Sanguinelli	19%	16%	62%	10%
Mombuca	4%	17%	17%	62%

Table 11. Confusion matrix obtained by analyzing the spectral information obtained with LIFS-405 technique for Navel sweet orange Baianinha, Ovale, and Vale Del Cauca.

Real variety	Classified as		
	Baianinha	Ovale	Del Cauca
Baianinha	61%	26%	13%
Ovale	9.6%	84%	6.4%
Del Cauca	3.3%	10%	86.7%

Table 12. Confusion matrix obtained by analyzing the spectral information obtained with LIFS-561 technique for Baianinha, Ovale, and Vale Del Cauca.

Real variety	Classified as		
	Baianinha	Ovale	Del Cauca
Baianinha	74%	4.3%	21.7%
Ovale	3.2%	96.8%	0.0%
Del Cauca	13.3%	0.0%	86.7%

Table 13. Confusion matrix obtained when evaluating all ten sweet orange accessions together using the LIFS-405 technique.

Real variety	Classified as									
	Lima IAC9	Pera Abril	Pera Bianchi	Pera Roberto Gullo	Piralima IAC11	Piralima IAC2	Seleta do Rio	Valencia B	Valencia Puka	Westin
Lima IAC9	39.3%	10.0%	3.2%	3.3%	5.1%	0.0%	0.0%	13.8%	0.0%	0.0%
Pera Abril	21.4%	63.3%	0.0%	0.0%	5.1%	0.0%	3.0%	10.3%	0.0%	0.0%
Pera Bianchi	0.0%	0.0%	71.0%	23.3%	5.1%	6.7%	6.0%	0.0%	23.3%	0.0%
Gullo	3.6%	0.0%	3.2%	50%	0.0%	6.7%	0.0%	6.9%	10.0%	0.0%
Piralima IAC11	17.8%	6.7%	0.0%	0.0%	69.2%	23.3%	3.0%	13.8%	6.7%	0.0%
Piralima IAC2	3.6%	0.0%	6.5%	0.0%	5.1%	43.3%	0%	10.3%	6.7%	0.0%
Seleta do Rio	0.0%	3.3%	0.0%	0.0%	5.1%	3.3%	88%	0.0%	3.3%	0.0%
Valencia B	14.2%	13.3%	3.2%	0.0%	2.6%	6.7%	0.0%	27.6%	6.7%	9.7%
Valencia Puka	0.0%	0.0%	9.7%	13.3%	2.6%	6.7%	0.0%	6.9%	30.0%	0.0%
Westin	0.0%	3.3%	3.2%	10%	0.0%	3.3%	0.0%	10.3%	13.3%	90.3%

obtained among them is 77.23% with LIFS-405 and 85.83% with LIFS-561. The same model was evaluated to construct the confusion matrix for the fluorescence spectra data measured with both systems which are presented in Tables 11 and 12, respectively.

Apart from the difference between the success rates obtained with both systems for the same group of accession, the classification evaluations of a smaller number of accessions reduces confusion. This is due to the fact that the descriptor number used for classification, 1593 points for LIFS-405 and 1680 for LIFS-560, remains unchanged. This way, the success rate shown in Tables 6, 7, 9, 10, 11, and 12 increased when compared to the same accession in Tables 4 and 5.

Analyzing the main diagonal confusion matrix of Tables 4 and 5, which corresponds to the success rate in correctly classifying the orange accession, it is possible to observe an increase in differentiation accuracy when they are measured with LIFS-405. These results are due to higher excitation energy of LIFS-405, which may access multiple fluorophores, not only chlorophyll. Thus, the fluorescence spectra obtained with LIFS-405 get much more information which may be taken into account in the differentiation between the citrus varieties. Therefore, the second set of leaves accessions was measured only with LIFS-405.

Second Set of Leaves Analyzed

This set was measured only with LIFS-405 once the differentiation accuracy obtained in the previous set was better. Following the same procedure as before, several runs of cross-validation were performed to determine the number of components responsible for the best results of classification among ten different accessions. Again, a confusion matrix for all varieties together was evaluated using classification via regression associated with a PLSR classifier. Ten executions of tenfold stratified cross-validation with ten

Table 14. Confusion matrix obtained by analyzing the spectral information obtained with LIFS-405 technique.

Real variety	Classified as		
	Pera Abril	Pera Bianchi	Pera Roberto Gullo
Pera Abril	96.7%	0%	3.3%
Pera Bianchi	3.2%	90.4%	6.4%
Pera Roberto Gullo	13.3%	13.3%	73.4%

Table 15. Confusion matrix obtained by analyzing the spectral information obtained with LIFS-405 technique.

Real variety	Classified as			
	Seleta do Rio	Valencia B	Valencia Puka	Westin
Seleta do Rio	97.0%	3.0%	0.0%	0.0%
Valencia B	3.4%	75.9%	13.8%	6.9%
Valencia Puka	3.3%	3.3%	80%	13.3%
Westin	3.2%	3.2%	3.2%	90.4%

Table 16. Confusion matrix obtained by analyzing the spectral information obtained with LIFS-405 technique.

Real variety	Classified as		
	Lima	Piralima 11	Piralima 2
Ipigua	96.7%	0%	3.3%
Piralima IAC 11	3.2%	90.4%	6.4%
Piralima IAC 2	13.3%	13.3%	73.4%

components were also performed and it was possible to obtain a global success rate of 57.2% in correctly classifying each variety as shown in Table 13.

As with the first set, the second set of leaves was separated in three groups as presented in Table 3. For the first group, composed of the Pera accessions (Abril, Bianchi, and Roberto Gullo) the confusion matrix is depicted in Table 13. In this case, 86.83% of the instances were correctly classified. Observing the confusion matrix of Table 14, it is possible to see that the major confusion was with Pera Roberto Gullo and the other two accessions.

Using the same model to assemble the confusion matrix for the second group of accessions, it is possible to obtain the confusion matrix shown in Table 15. The accuracy in correctly classifying the orange accessions is 86.2%. The accession with highest accuracy rate was Seleta do Rio with a success rate of 97%.

Finally, Table 16 shows the confusion matrix obtained for the third group in this set of accessions, composed of sweet

orange Ipigua, Piralima IAC 11, and Piralima IAC12. The best success rate is 96.7% for Ipigua.

Conclusions

The results showed that LIFS, combined with statistical methods, is capable of discriminating different varieties of citrus based on the fact that the leaves of plants have multiple fluorophores whose concentration depends on its genetics and metabolism. Thus, it is possible to conclude that the optical properties of citrus leaves may be different, depending on their genetic variety. As demonstrated in this work, the success rate depends not only on the combination of canopy and rootstock varieties, but also of the laser excitation. The best results were obtained with LIFS-405 which has an excitation wavelength in the violet range and is able to access the information of several metabolites that compose the leaf. With the aid of LIFS-405, it was possible to achieve a 100% success rate in differentiating Pera Olimpia.

The application of this study is very interesting for the productive sector because it can result in economically viable tools for new seedlings' certification method and early identification, cultivar protection, and germplasm core collection. Besides, such equipment allows large-scale and real-time monitoring for citrus early identification in breeding programs. The spectroscopy may be used with success to discriminate citrus orange accessions, which is a quick and less expensive method when compared to others tools such as molecular techniques and visual inspection. When it comes to DNA markers, they are not as efficient as LIFS in differentiating accessions from the same species. Besides, LIFS is applicable in non-bearing nursery trees, which is an advantage. In most cases, the morphological characterization of citrus sweet orange cultivars may be only done post fruit bearing and has no practical value to be used in a certified program to produce and control the genetic background of commercialized citrus nursery plants.

Conflict of Interest

The authors report there are no conflicts of interest.

Funding

The authors are grateful to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) for financial support (Processo n. 2011/18605-0).

References

1. FAO (Food and Agriculture Organization). Faostat. Statistical database. <http://faostat.fao.org/site/567/default.aspx> [accessed Mar 29 2014].
2. G.A. Moore. "Oranges and Lemons: Clues to the Taxonomy of Citrus from Molecular Markers". *Trends Genet.* 2001. 17(9): 536–540.

3. E. Nicolosi, Z.N. Deng, A. Gentile, S. La Malfa, G. Continella, E. Tribulato. "Citrus Phylogeny and Genetic Origin of Important Species as Investigated by Molecular Markers". *Theor. Appl. Genet.* 2000. 100: 1155–1166.
4. Q. Xu, L.-L. Chen, X. Ruan, D. Chen, A. Zhu, C. Chen, D. Bertrand, W.-B. Jiao, B.-H. Hao, M.P. Lyon, J. Chen, S. Gao, F. Xing, H. Lan, J.-W. Chang, X. Ge, Y. Lei, Q. Hu, Y. Miao, L. Wang, S. Xiao, M.K. Biswas, W. Zeng, F. Guo, H. Cao, X. Yang, X.-W. Xu, Y.-J. Cheng, J. Xu, J.-H. Liu, O.J. Luo, Z. Tang, W.-W. Guo, H. Kuang, H.-Y. Zhang, M.L. Roose, N. Nagarajan, X.-X. Deng, Y. Ruan. "The Draft Genome of Sweet Orange (*Citrus sinensis*)". *Nat. Genet.* 2013. 45: 59–66.
5. G.A. Wu, S. Prochnik, J. Jemkins, J. Salse, U. Hellsten, F. Murat, X. Perrier, M. Ruiz, S. Scalabrin, J. Terol, M.A. Takita, K. Labadie, J. Poulain, K. Jabbari, F. Cattonaro, C. Del Fabbro, S. Pinosio, A. Zuccolo, J. Chapman, J. Grimwood, F.R. Tadeo, L.H. Estornell, J.V. Muñoz-Sanoz, V. Ibanez, A. Herrero-Ortega, P. Aleza, J. Pérez-Pérez, D. Ramón, D. Brunel, F. Luro, C. Chen, W.G. Farmerie, B. Desany, C. Kodira, M. Mohiuddin, T. Harkins, K. Fredrikson, P. Burns, A. Lomsadze, M. Borodovsky, G. Reforgiato, J. Freitas-Astúa, F. Quetier, L. Navarro, M. Roose, P. Wincker, J. Schmutz, M. Morgante, M.A. Machado, M. Talon, O. Jaillon, P. Ollitrault, F. Gmitter, D. Rokhsar. "Sequencing of Diverse Mandarin, Pummelo and Orange Genomes Reveals Complex History of Admixture During Citrus Domestication". *Nat. Biotechnol.* 2014. 32: 656–662.
6. R.W. Hodgson. "Horticultural Varieties of citrus". In: W. Reuther, H.J. Webber, L.D. Batchelor (eds) *The Citrus Industry*. Vol. 1, Berkeley, CA: University of California Press, 1967, pp.431–589.
7. V.M. Novelli, M. Cristofani, A.A. Souza, M.A. Machado. "Development and Characterization of Polymorphic Microsatellite Markers for the Sweet Orange (*Citrus sinensis* L. Osbeck)". *Genet. Mol. Biol.* 2006. 29: 90–96.
8. R.M. Moraes Filho, H.J. Jimenez, A.V.V. Montarroyos, R.S. Musser, M. M. Silva, E.F. Silva, L.S.S. Martins. "Variabilidade Genética em Genótipos da Coleção de Germoplasma de Citrus, do Instituto Agrônomo de Pernambuco Brejão-PE, Por Meio de Marcadores Moleculares ISSR". *Citrus Research & Technology Cordeirópolis*. 2011. 32(2): 67–76.
9. R. Herrero, M.J. Asíns, E.A. Carbonell, L. Navarro. "Genetic Diversity in the Orange Subfamily Aurantioideae: I—Intraspecies and Intra-genus Genetic Variability". *Theor. Appl. Genet.* 1996. 92: 599–609.
10. S.J. Orford, N.S. Scott, J.N. Timmis. "A Hypervariable Middle Repetitive DNA Sequence from Citrus". *Theor. Appl. Genet.* 1995. 91: 1248–1252.
11. D.Q. Fang, M.L. Roose. "Identification of Closely Related Citrus Cultivars with Inter-Simple Sequence Repeat Markers". *Theor. Appl. Genet.* 1997. 95: 408–417.
12. M.L.P.N. Targon, M.A. Machado, H.D. Coletta Filho, M. Cristofani. "Genetic Polymorphism of Sweet Orange (*Citrus sinensis* L. Osbeck) Varieties Evaluated by Random Amplified Polymorphic DNA". *Acta Horticulturae*. 2000. 535: 51–53.
13. T.G. Sankar, V. Gopi, B. Deepa, K. Gopal. "Genetic Diversity Analysis of Sweet Orange (*Citrus sinensis* Osbeck) Varieties/Clones Through RAPD Markers". *Int. J. Curr. Microbiol. App. Sci.* 2014. 3(4): 75–84.
14. V.M. Novelli, M.A. Machado, C.R. Lopes. "Isoenzymatic Polymorphism in *Citrus* spp and *P. trifoliata* (L.) Raf. (Rutaceae)". *Gen. Mol. Biol.* 2000. 23: 163–168.
15. V.M. Novelli, M.A. Takita, M.A. Machado. "Identification and Analysis of Single Nucleotide Polymorphisms (SNPs) in Citrus". *Euphytica*. 2004. 138: 227–237.
16. M.K. Biswas, Q. Xu, C. Mayer, X. Deng. "Genome Wide Characterization of Short Tandem Repeat Markers in Sweet Orange (*Citrus sinensis*)". *PLoS ONE*. 2014. 9(8): e104182.
17. X. Sun, Q. Mu, D. Jiang, C. Wang, X.C. Wang, J.G. Fang. "A New Strategy Employed for Identification of Sweet Orange Cultivars with RAPD Markers". *Genet. Mol. Res.* 2012. 11(3): 2071–2080.
18. D. M. B. P. Milori, M. Raunaud, P.R. Villas-Boas, A.L. Venâncio, S. Mounir, R.B. Bassaezi, R. Redon. "Identification of Citrus Varieties Using Laser-Induced Fluorescence Spectroscopy (LIFS)". *Computers and Electronics in Agriculture*. 2013. 95: 11–18.
19. D.D.S. Santana-Vieira, D.M.B.P. Milori, P.R. Villas Boas, M.F. Silva, M.G. Santos, F.A. Gaiotto, W.S. Soares Filho, A.S. Gesteira. "Rapid Differentiation of Closely Related Citrus Genotypes by Fluorescence Spectroscopy". *Advances in Bioscience and Biotechnology*. 2014. 5: 903–914.
20. J.R. Lakowicz. "Principles of Fluorescence Spectroscopy", Third Edition ed. Baltimore, MD: Springer Science+Business Media, LLC, 2006, p.954.
21. F.M.V. Pereira, D.M.B.P. Milori, E.R. Pereira-Filho, A.L. Venâncio, M.S.T. Russo, P.K. Martins, J. Freitas-Astúa. "Fluorescence Images Combined to Statistic Test for fingerprinting of Citrus Plants after Bacterial Infection." *Anal. Methods*. 2011. 3(3): 552–556.
22. F.M.V. Pereira, D.M.B.P. Milori, E.R. Pereira Filho, A.L. Venâncio, M.S.T. Russo, M.C.B. Cardinali, P.K. Martins, J. Freitas-Astúa, J. "Laser-Induced fluorescence Imaging Method to Monitor Citrus Greening Disease". *Computers and Electronics in Agriculture*. 2011. 79: 90–93.
23. Z.G. Cerovic, G. Samson, F. Morales, N. Trenblay, I. Moya. "Ultraviolet-Induced Fluorescence for Plant Monitoring: Present State and Prospect". *Agronomie*. 1999. 19: 543–578.
24. F. Stober, K. Lichtenthaler. "Characterization of the Laser-Induced Blue, Green and Red Fluorescence Signatures of Leaves of Wheat and Soybean Grown Under Different Irradiance". *Physiologia Plantarum*. 1993. 88: 696–704.
25. A.C. Ranulfi, M.C.B. Cardinali, T.M.K. Kubota, J. Frietas-Astua, E.J. Ferreira, B.S. Belleite, M.F.G.F. da Silva, P.R. Villas Boas, A.B. Magalhaes, D.M.B.P. Milori. "Laser-Induced Fluorescence Spectroscopy Applied to Early Diagnosis of Citrus Huanglongbing". *Biosystems Engineering*. 2016. 144: 133–144.
26. G.H. Krause, E. Weis. "Chlorophyll Fluorescence and Photosynthesis: The Basics". *Annu. Rev. Plant Physiol. Plant. Molec. Biol.* 1999. 42: 313–349.
27. A.N. Misra, M. Misra, R. Singh. "Chlorophyll Fluorescence in Plant Biology". In: A.N. Misra, editor. *Biophysics*. Croatia: InTech. 2012. Chapter 7, pp 171–192. <http://www.intechopen.com/books/biophysics/chlorophyll-fluorescence-in-plant-biology> [accessed Jun 27 2016].
28. J. Fernandes, W.F. Falco, S.L. Oliveira, A.R.L. Caires. "Changes in Chlorophyll α Fluorescence of Glyphosate-Tolerant Soybean Plants Induced by Glyphosate: *In Vivo* Analysis by Laser-Induced Fluorescence Spectroscopy". *Applied Optics*. 2013. 52(13): 3004–3011.
29. G. Agati. "Response of the *In Vivo* Chlorophyll Fluorescence Spectrum to Environmental Factors and Laser Excitation Wavelength". *Pure Appl. Opt.* 1998. 7: 797–807.
30. D.M.B.P. Milori, L.M. Neto, E.J. Ferreira, A.F. Zaghi. "Method, Apparatus and System for Diagnosis of Stresses and Diseases in Higher Plants". EP Patent. 2011. 2,376,897.
31. H. Cheng, M. Nikus, S.L. Jämsä-Jounela. "Evaluation of PCA Methods with Improved Fault Isolation Capabilities on a Paper Machine Simulator". *Chemometrics and Intelligent Laboratory Systems*. 2008. 92: 186–199.
32. S. Wold, K. Esbensen, P. Geladi. "Principal Component Analysis". *Chemometrics and Intelligent Laboratory Systems*. 1987. 2: 37–52.
33. I.E. Frank, J.H. Kalivas, B.R. Kowalski. "Partial Least Squares Solutions for Multicomponent Analysis". *Analytical Chemistry*. 1983. 55(11): 1800–1804.

34. E. Frank, Y. Wang, S. Inglis, G. Holmes, I.H. Witten. "Using Model Trees for Classification". *Machine Learning*. 1998. 32(1): 63–76.
35. M.C.D. Cardinali, P.R. Villas Boas, D.M.B.P. Milori, E.J. Ferreira, M.F.E. Silva, M.A. Machado, B.S. Belleste, M.F.D.F. da Silva. "Infrared Spectroscopy: a Potential Tool in Huanglongbing and Citrus Variegated Chlorosis Diagnosis". *Talanta*. 2012. 91: 1–6.
36. S. Wold, M. Sjöström, L. Eriksson. "PLS-Regression: a Basic Tool of Chemometrics". *Chemometrics and Intelligent Laboratory Systems*. 2001. 58: 109–130.
37. H.K. Lichtenthaler, U. Rinderle. "The Role of Chlorophyll Fluorescence in the Detection of Stress Conditions in Plants". *CRC Crit. Rev. Analyt. Chem.* 1988. 19(Suppl. 1): S29–S85.
38. H.K. Lichtenthaler. "Applications of Chlorophyll Fluorescence in Stress Physiology and Remote Sensing". In: M.D. Steven, J.A. Clark, editors. *Proceedings of the 48th Easter School in Agricultural Science*. London: Butterworths, 1990. Pp. 287–305.
39. H.K. Lichtenthaler, J.A. Miehe. "Fluorescence Imaging as a Diagnostic Tool for Plant Stress". *Trends Plant Sci.* 1997. 2(8): 316–320.
40. C. Buschmann. Variability and Application of the Chlorophyll Fluorescence Emission Ratio Red/Far-Red of Leaves. *Photosynth. Res.* 2007. 92(2): 261–271.