Detection of Citrus Huanglongbing by Fourier Transform Infrared–Attenuated Total Reflection Spectroscopy

SAMANTHA HAWKINS,* BOSOON PARK, GAVIN POOLE, TIM GOTTWALD, BOB WINDHAM, and KURT LAWRENCE

US Department of Agriculture- Agricultural Research Service Russell Research Center, 950 College Station Rd, Athens, Georgia 30605 (S.H., B.P., B.W., K.L.); and US Department of Agriculture- Agricultural Research Service, US Horticultural Research Laboratory, Fort Pierce, Florida 34945 (G.P., T.G.)

Citrus Huanglongbing (HLB, also known as citrus greening disease) was discovered in Florida in 2005 and is spreading rapidly amongst the citrus growing regions of the state. Detection via visual symptoms of the disease is not a long-term viable option. New techniques are being developed to test for the disease in its earlier presymptomatic stages. Fourier transform infrared–attenuated total reflection (FT-IR-ATR) spectroscopy is a candidate for rapid, inexpensive, early detection of the disease. The mid-infrared region of the spectrum reveals dramatic changes that take place in the infected leaves when compared to healthy non-infected leaves. The carbohydrates that give rise to peaks in the 900-1180 cm^{-1} range are reliable in distinguishing leaves from infected plants versus non-infected plants. A model based on chemometrics was developed using the spectra from 179 plants of known disease status. This model then correctly predicted the status of >95% of the plants tested.

Index Headings: **Citrus greening disease; Fourier transform infrared spectroscopy; Attenuated total reflection; FT-IR-ATR; Chemometrics; HLB.**

INTRODUCTION

For several decades the occurrence of citrus greening disease, also known as Huanglongbing (HLB) or Yellow Dragon disease, has been of great concern in the citrus growing regions of the southern United States as well as in other major citrus regions of the world.[1–7] The citrus psyllid (*Diaphorina citri*) acts as a vector spreading the disease from infected trees to nearby and distant uninfected trees.[1–3,7] These psyllids were detected in Florida in 1998,[1] and HLB was confirmed to be in Florida citrus in 2005.[8] Since that time it has spread to 32 counties across the state, with the psyllid being found in two additional counties.[9] Unfortunately, once the tree is infected the disease does not always appear immediately. It may take multiple months to years for the infection to express visual signs depending on tree age and horticultural health. During this latent period the tree can, however, be contagious and infect numerous other trees. Thus, it is imperative that reliable techniques are developed for early, presymptomatic detection of the disease.

Currently, the technique used most frequently is polymerase chain reaction (PCR) DNA analysis.[4,5,7–13] The PCR detection method is very accurate and is able to detect the disease before the disease becomes symptomatic but not before the bacteria proliferate in the portion of the plant assayed. Unfortunately, this method is also somewhat time-consuming and is relatively expensive. Several studies have been reported using other techniques, including spectroscopic methods, to attempt to quickly identify the citrus greening disease's presence in plants.[14–17] Some of the reported methods are subjective or are not reproducible. Ideally, a technique needs to be developed that is highly reliable, quick, and inexpensive. Fourier transform infrared–attenuated total reflection (FT-IR-ATR) spectroscopy may be such a technique. In this study leaves from plants that were identified as infected and uninfected with the citrus greening disease using both conventional and real-time PCR methods were collected and studied by the ATR technique. The results of this study, which are presented below, are potentially very promising in providing a new method for the rapid detection of citrus greening.

EXPERIMENTAL DETAILS

Samples of leaves were collected in Florida from orange and grapefruit trees. There were several symptom types (subsets) identified during collection. These subsets include greenhouse quarantined negative (negative control), visual negative, and samples that showed visible symptoms of the HLB infection, including foliar blotchy mottle, yellow dragon (entire yellow shoots), foliar green islands, and various stages of chlorosis ranging from mild to severe. Several leaves were collected from each plant or area of a plant and combined as a single sample. HLB infected plants generally produce smaller leaves, which necessitates the use of several leaves per sample. The mid-rib veins of the leaves were separated for testing by PCR. The DNA was isolated using a modified sodium dodecyl sulfate/potassium acetate extraction method.[18] These samples were then normalized to contain 50 ng of DNA per μL of water. Analysis of 2 μL sample was performed in duplicate via a 7500 Fast Real-Time PCR system (Applied Biosystems, Foster City , CA) using Invitrogen Express Premix (Carlsbad, CA), along with primers and probe (labeled with NED-MGB™) developed by Li.[19] The thermal cycling protocol was 95 °C for 20 s, then forty cycles of 95 °C for 3 s and 60 °C for 30 s. PCR is a technique that amplifies and detects the amount of predetermined DNA strands.[20] Briefly, a probe to a specific piece of DNA, in this case a probe to a small section of the DNA sequence for HLB, is added to the DNA extracted from the sample. Through the cycling process, the DNA section to which the probe attaches is amplified. Under optimum conditions, and with perfect efficiency, the amount of DNA of interest is doubled with each cycle. With real-time PCR, the amplicon is "monitored" with each cycle. This can be done by tracking how much dye is turned on by the amplification process. In the technique used in this paper, a dye is attached to the DNA probe, along with a quencher. Once the probe attaches to the DNA of interest, the dye and quencher are cleaved from the probe (and each other), and therefore the dye fluoresces.[21] The amount of fluorescence is measured at the end

Received 29 July 2009; accepted 26 October 2009.
* Author to whom correspondence should be sent. E-mail: samantha.hawkins@ars.usda.gov.

0003-7028/10/6401-000000$2.00/0
© 2010 Society for Applied Spectroscopy

TABLE I. Summary of the status of the citrus plant samples analyzed in this study (standard deviations are in parenthesis). Predictions were obtained using the principal component regression model.

Description of leaves, # of samples in average	Real-time-PCR C_t value measured average	Predicted C_t average from FT-IR-ATR based chemometric model	Average difference of measured and predicted C_t values	Number of incorrect predictions in category
Quarantined HLB neg. orange trees, 20	39.28 (±1.40)	38.01 (±3.89)	1.27 (±2.16)	1 false positive
Quarantined HLB neg. Grapefruit trees, 10	39.03 (±1.74)	38.90 (±2.66)	0.13 (±2.16)	0
Visual negative, 11	38.34 (±1.49)	38.51 (±2.58)	0.17 (±2.02)	0
Visual negative but actually HLB positive, 4	20.33 (±0.38)	24.80 (±2.70)	4.47 (±2.71)	0
HLB negative, but with some visual symptoms, 14	38.31 (±2.08)	34.33 (±2.74)	3.98 (±2.34)	3 in grey area (30–32)
HLB positive with Green Islands, 17	19.14 (±1.31)	20.07 (±1.43)	0.93 (±1.18)	0
HLB positive with chlorosis, 32	20.99 (±1.10)	21.23 (±3.58)	0.24 (±2.14)	0
HLB positive Blotchy, 34	20.29 (±1.47)	21.78 (±3.74)	1.49 (±2.72)	1 false negative, 1 in grey area (30–32)
HLB positive Yellow Dragon, 24	19.63 (±1.18)	19.50 (±2.36)	0.13 (±1.38)	0
HLB positive, no visual symptoms, 2	19.70 (±0.28)	30.68 (±0.49)	10.98 (±0.21)	2 in grey area (30–32)
HLB positive, other symptoms, 11	21.25 (±1.27)	22.45 (±3.28)	1.20 (±1.77)	0

of each cycle and plotted on a graph with fluorescence on the *y*-axis and cycle number on the *x*-axis. Once the fluorescence exceeds a user defined threshold, the number of cycles needed to get that sample to that point is recorded as the C_t, or "Critical Threshold", for that sample.[20] The higher the C_t value, the lower the concentration of the DNA of interest in the sample.

For conventional PCR, 2 μL of each normalized sample was analyzed in duplicate using Las Long Primers with a modified thermo-cycling, Taq DNA polymerase (Sigma-Aldrich, St. Louis, MO), and FailSafe™ PCR PreMix D (EPICENTRE® Biotechnologies, Madison, WI) as per the instructions provided by the manufacturer.[22] The amplification protocol was 95 °C for 5 min, then thirty cycles of 94 °C for 15 s, 60 °C for 30 s, and 72 °C for 30 s. After amplification, the sample temperatures were raised to 72 °C for 10 min and then lowered to 4 °C until they were analyzed. The product was analyzed by electrophoresis on a 2% Agrose gel with ethidium bromide stain for 60–80 min at a constant voltage of 100 V. The gel was analyzed for the presence or absence of a band at 195 bp using a Kodak Gel Logic 200 Imaging System (Carestream Health-Molecular Imaging, New Haven, CT).

The remainder of the leaves (predominately the leaf lamina) were dried in a 1250 W microwave oven for 3 min at one-half power and ground into a powder in a Kleco 4200 Tissue Pulverizer (Garcia Manufacturing, Visalia, CA). The powder samples were kept in 2.5 mL screw-cap vials under refrigeration until ATR analysis to limit the amount of moisture in the samples and thus avoid interference in the FT-IR spectra due to moisture.

The spectra were collected using a Thermoelectron Nicolet (Madison, WI) Magna 850 FT-IR spectrometer with a DTGS detector. A single-bounce diamond crystal ATR (Durascope, Smiths Detection, Danbury, CT) accessory was used, negating the need for further sample preparation. The crystal and pressure arm were cleaned thoroughly with acetone and hexane between each sample and allowed to dry. The spectra were acquired with 2 cm^{-1} resolution and a co-addition of either 64 or 128 scans. The signal-to-noise difference between 64 and 128 scans was negligible. The sample spectra were background subtracted using a spectrum collected in the absence of any sample.

The spectra were analyzed using the Unscrambler® (version 9.8 from Camo Software, Woodbridge, NJ) software. The spectra were baseline corrected with a scalar baseline offset. The region from 1765 to 4000 cm^{-1} was not included in the analysis due to interference from the diamond ATR crystal and the large OH band due to remaining water in the samples. The remainder of the spectrum was normalized using area normalization. The 700 to 1765 cm^{-1} region of the mid-infrared was used to create a model for prediction of the presence of citrus greening disease. Real-time PCR testing identified the samples as either positive or negative. The value assigned during the real-time PCR testing was called the Critical Threshold (C_t) value, as discussed in the Experimental section. Based on previous research done in our lab, a sample with a C_t value above 32 was considered negative for citrus greening disease, whereas, a sample with a C_t below 30 was considered positive. However, C_t values between ~30 and 32 were considered to be in a "grey" area in which it is difficult to say with certainty that the sample is positive, negative, or in the early stages of developing the HLB. The conventional PCR technique is non-quantitative, and the primers used in this study were less susceptible to interference than the primers used in the real-time PCR.

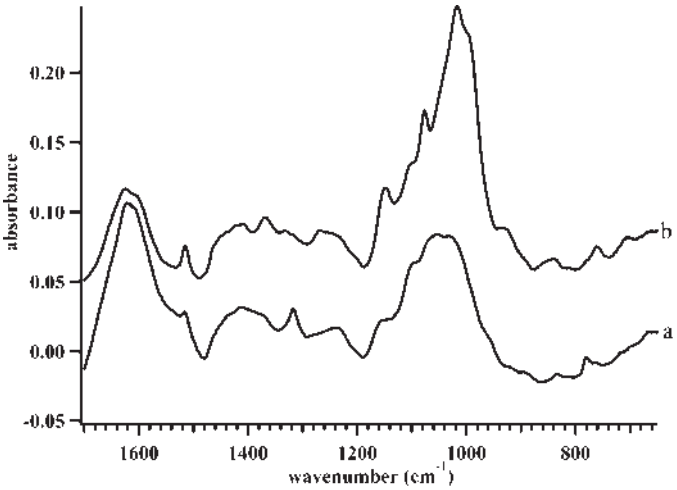

FIG. 1. (***a***) The ATR spectrum collected from an HLB negative citrus plant (C_t value 40); (***b***) ATR spectrum of an HLB positive citrus plant (C_t value 18.75). The region shown, 700-1765 cm^{-1}, is the region used in the chemometric models. Note the most obvious differences between the spectra between 900 and 1150 cm^{-1} , which correspond to the carbohydrate region. Spectra are offset for visual clarity.

TABLE II. Summary of the status of the citrus plant samples analyzed in this study (standard deviations are in parenthesis). Predictions were obtained using the partial least squares model.

Description of leaves, # of samples in average	Real-time PCR C_t value measured average	Predicted C_t average from FT-IR-ATR based chemometric model	Average difference of measured and predicted C_t values	Number of incorrect predictions in category
Quarantined HLB neg. orange trees, 20	39.28 (±1.40)	38.11 (±3.69)	1.17 (±3.00)	1 false positive
Quarantined HLB neg. Grapefruit trees, 10	39.03 (±1.74)	39.24 (±2.53)	0.21 (±2.29)	0
Visual negative, 11	38.34 (±1.49)	37.97 (±2.57)	0.37 (±2.09)	0
Visual negative but actually HLB positive, 4	20.33 (±0.38)	24.76 (±2.60)	4.43 (±2.61)	0
HLB negative, but with some visual symptoms, 14	38.31 (±2.08)	34.87 (±3.12)	3.44 (±2.40)	3 in grey area (30–32)
HLB positive with Green Islands, 17	19.14 (±1.31)	19.93 (±1.27)	0.79 (±1.00)	0
HLB positive with chlorosis, 32	20.99 (±1.10)	21.29 (±3.71)	0.30 (±2.27)	1 in grey area (30–32)
HLB positive Blotchy, 34	20.29 (±1.47)	21.62 (±3.48)	1.33 (±2.53)	1 false negative
HLB positive Yellow Dragon, 24	19.63 (±1.18)	19.40 (±2.17)	0.23 (±1.39)	0
HLB positive, no visual symptoms, 2	19.70 (±0.28)	30.89 (±0.27)	11.19 (±0.01)	2 in grey area (30–32)
HLB positive, other symptoms, 11	21.25 (±1.27)	22.79 (±3.27)	1.54 (±1.84)	1 in grey area (30–32)

RESULTS AND DISCUSSION

The samples were collected in two different sets, the first containing 85 samples (collected in July 2008) and the second containing 94 samples (collected in November 2008). Table I contains a summary of the sample sets. Figure 1 shows representative spectra from the first sample set. The most notable difference between HLB positive and negative samples is in the region 900-1150 cm^{-1}. This region is attributable to carbohydrate vibrational bands.[23–26] The carbohydrates represented by similar vibrational bands in this region include various sugars, starches, and cellulose. Both the positive and negative samples have peaks due to carbohydrates, but it is clear that a chemical change has occurred (in amount, type, or structure) to the carbohydrates present in the HLB positive samples. A key element to the destructive nature of the HLB disease is that the fruit becomes small, light, and very acidic.[7] The change in the fruit taste and appearance may be due, in large part, to the change in which carbohydrates are present. Future experiments may help determine the exact identity of the chemical changes. Weak shoulders at 1150 and 1077 cm^{-1} in the negative samples become sharp peaks in the positive samples. A broad, flat peak centered around 1052 cm^{-1} in the negative samples red-shifts to become a sharper asymmetrical peak centered at 1020 cm^{-1} in the positive samples. This entire region is potentially very useful in developing prediction models using the computer software.

The entire region of the ATR spectra from 700 to 1765 cm^{-1} was included in the modeling. Models were developed using principal component regression and partial least squares regression methods. The principal component regression results are summarized in Table I. The principal component regression method is a two-part method whereby principal component analysis of the spectra is followed by multiple linear regression. The method incorporates full cross-validation to arrive at an optimal model with which predictions can be made. The model developed during this study was optimized with three principal components. Outlier spectra were removed from the model in order to optimize the root mean square error (RMSE) of the model to a value below 3. The partial least squares method is also a two-part method incorporating partial least squares followed by multiple linear regression. Full cross-validation is also used in this method development, as is the removal of outliers for optimization. The prediction results for the partial least squares model are shown in Table II. The results are very similar whether the principal component regression method or the partial least squares method is used for the model calculations.

The area centered around 1020 cm^{-1}, attributed to carbohydrates, is the region that appeared to have the most striking effect on the modeling outcomes. The models were calculated using 162 of the 179 total samples. Seventeen (17) samples were left out of both models based on their outlier status. The models that were developed were then used to predict the positive (C_t value below 30) or negative (C_t value above 32) HLB status of all the samples, including the outliers. Samples were correctly predicted 171 out of 179 times using the above criteria for either a positive or negative prediction in the principle component regression model and 170 out of 179 times in the partial least squares regression model.

However, 6 and 7 samples (principal component regression and partial least squares, respectively) fall in the grey area for the value determining whether the plant is infected with HLB. These grey area predictions include almost all of the incorrect predictions in both models, and their status as a correct or

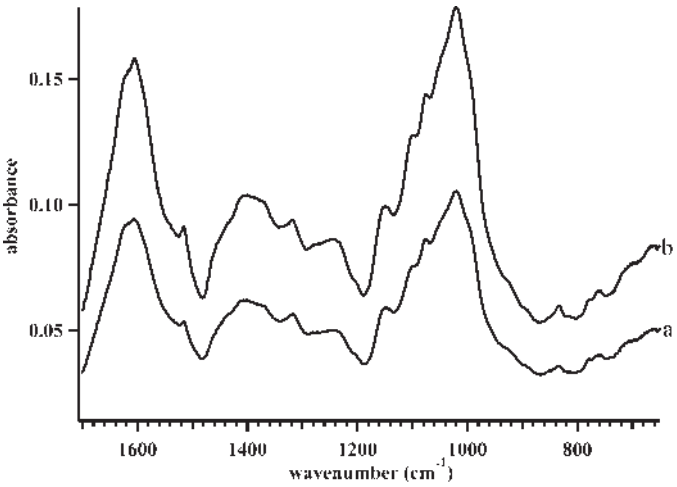

FIG. 2. FT-IR spectra from two plants that were identified by real-time PCR as HLB positive but had no visual symptoms. These samples were incorrectly identified by the prediction models as HLB negative but have FT-IR spectra that resemble those of positive samples in the carbohydrate region. (Spectra were offset for visual clarity.) The measured C_t values were (***a***) 19.89 and (***b***) 19.5, whereas the predicted values using the principle component regression model were 31.02 and 30.33, respectively. The predicted values are in the grey area and are very near the cutoff point between positive and negative, which would warrant further investigation of their disease status.

incorrect prediction may be in question due to the margin of error in the models. The principal component regression model had 1 false positive, 1 false negative, and 6 samples falling in the 30 to 32 grey area region (of which 3 were positive and 3 were negative). The partial least squares model had 1 false positive, 1 false negative, and 7 samples in the grey area (3 were negative and 4 were positive). The average difference (based on the C_t values) of the model predictions from the measured C_t values is approximately 2.6 for both methods. Thus, it would be highly recommended with the current modeling and spectroscopic techniques to do further testing on any samples predicted to be within the grey area of C_t values.

Comparisons of the predicted values and measured values are shown in Tables I and II. Of the samples that were correctly predicted using the current ATR spectral modeling, four samples were identified as visibly negative but tested positive by real-time PCR. Another 14 samples that were correctly predicted by the FT-IR based model had symptoms of chlorosis or other signs of greening disease but were actually HLB negative. Two samples were identified by real-time PCR as being positive for citrus greening disease but had no apparent visual symptoms. Both of these samples were incorrectly identified by the prediction model using their FT-IR spectra; however, they were both predicted very close to a C_t value of 30, falling in the grey area and warranting further testing. The spectra for these two incorrectly predicted samples are shown in Fig. 2. Note that the spectral features in the carbohydrate region more closely resemble those of the typical HLB positive samples than of the negative samples. Thus, relying solely on the prediction model would likely result in errors for samples with prediction values in the "grey" area. Closer inspection of the spectra and further refinement of the predictive model may help alleviate incorrectly diagnosed samples.

CONCLUSION

The highly contagious nature of the citrus greening disease amongst neighboring citrus plants, especially in orchards, is highly likely to cause drastic economic difficulties for producers if the problem is not kept in check. Thus far, PCR testing has proven to be the most accurate and widely used method to determine the status of citrus plants. However, it is very time consuming and costly. FT-IR-ATR spectroscopy is very quick and straightforward. If used in the field, a sample could be collected, dried, ground, and analyzed using portable versions of all equipment in a matter of minutes, versus the several hours needed for PCR analysis. In the current study of 179 samples, the ATR spectra produced a model that subsequently correctly predicted the status of being positive or negative with respect to having citrus greening disease of >95% of the samples. Clearly more work is needed to solve the dilemma of how to interpret the "grey" area predictions. Thus far the results indicate that FT-IR-ATR spectroscopy is a useful, rapid, and inexpensive tool in the early identification of plants containing citrus greening disease.

ACKNOWLEDGMENTS

The authors wish to express gratitude to Ashley Ford, Joanne Hodges, Kimberly Poole, Earl Taylor, and Leonard Therrien for their help and technical support in the production of the dataset used in this study.

1. S. E. Halbert and K. L. Manjunath, Florida Entomol. **87,** 330 (2004).
2. J. M. Bové, J. Plant Pathol. **88,** 7 (2006).
3. J. V. da Graça, Annu. Rev. Phytopathol. **29,** 109 (1991).
4. J. Cubero, J. H. Graham, and T. R. Gottwald, Appl. Environ. Microbiol. **67,** 2849 (2001).
5. D. D. C. Teixeira, C. Saillard, S. Eveillard, J. L. Danet, P. I. Da Costa, A. J. Ayres, and J. M. Bové, Int. J. Sys. Evol. Microbiol. **55,** 1857 (2005).
6. W. Li, Q. Song, R. H. Brlansky, and J. S. Hartung, Proc. Natl. Acad. Sci. U.S.A. **104,** 18427 (2007).
7. T. R. Gottwald, J. V. da Graça, and R. B. Bassanezi, Plant Health Progress, online doi:10.1094/PHP-2007-0906-01-RV 2007.
8. S. Halbert, Proc. HLB-Greening Int. Workshop: 50 (2005).
9. S. D. Halbert, K. L Manjunath, and M. Brodie, Proc. Int. Res. Conf. on Huanglongbing. (2009).
10. H. Li, H. Doddapaneni, X. Bai, J. Yao, X. Zhao, and E. L. Civerolo, Mol. Cell. Probes **22,** 30 (2008).
11. D. D. C. Teixeira, J. L. Danet, S. Eveillard, E. C. Martins, W. C. de Jesus, P. T. Yamamoto, S. A. Lopes, R. B. Bassanezi, A. J. Ayres, C. Saillard, and J. M. Bové, Mol. Cell. Probes **19,** 173 (2005).
12. W. Li, J. S. Hartung, and L. Levy, J. Microbiol. Methods **66,** 104 (2006).
13. K. L. Manjunath, S. E. Halbert, C. Ramadugu, S. Webb, and R. F. Lee, Phytopathology **98,** 387 (2008).
14. A. Mishra, R. Ehsani, G. Albrigo, and W. S. Lee, ASABE Annu. Int. Meeting paper 073056 (2007).
15. J. Qin, T. F. Burks, M. S. Kim, K. Chao, and M. A. Ritenour, Proc. SPIE-Int. Soc. Opt. Eng. **6983,** 698305-1 (2008).
16. J. Qin, T. F. Burks, D. G. Kim, and D. M. Bulanon, ASABE Food Processing Automation Conference paper 701P0508cd (2008).
17. J. Qin, T. F. Burks, M. A. Ritenour, and G. W. Bonn, ASABE Annu. Int. Meeting paper 084066 (2008).
18. J. J. De Paulo and C. A. Powell, Plant Dis. **79,** 246 (1995).
19. W. Li, J. Hartung, and L. J. Levy, Microbiol. Methods **66,** 104 (2006).
20. I. A. Mackay, K. E. Arden, and A. Nitsche, Nucleic Acids Res. **30,** 1292 (2002).
21. L. G. Lee, C. R. Connell, and W. Bloch, Nucleic Acids Res. **21,** 3761 (1993).
22. R. Shatters, USDA-ARS Ft. Pierce, FL, private communication (2009).
23. H. K. Khurana, S. Jun, I. K. Cho, and Q. X. Li, Appl. Eng. Agric. **24,** 631 (2008).
24. J. J. G. van Soest, H. Tournois, D. de Wit, and J. F. G. Vliegenthart, Carbohydrate Res. **279,** 205 (1995).
25. P. D. Vasko, J. Blackwell, and J. L. Koenig, Carbohydr. Res. **19,** 297 (1971).
26. J. C. Tewari and K. Malik, Int. J. Food Sci. Technol. **42,** 200 (2007).