University of São Paulo
"Luiz de Queiroz" College of Agriculture

Identification of resistant sugarcane genotypes to ratoon stunt disease by Fourier-transform infrared analysis

Marie Caroline Ferreira Laborde

Thesis presented to obtain the degree of Doctor in Science. Area: Plant Pathology

Piracicaba
2019

Marie Caroline Ferreira Laborde
Bachelor in Agronomy

Identification of resistant sugarcane genotypes to ratoon stunt disease by Fourier-transform infrared analysis

Advisor:
Prof. Dr. **LUIS EDUARDO ARANHA CAMARGO**

Thesis presented to obtain the degree of Doctor in Science. Area: Plant Pathology

Piracicaba
2019

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Laborde, Marie Caroline Ferreira

Identification of resistant sugarcane genotypes to ratoon stunt disease by Fourier-transform infrared analysis / Marie Caroline Ferreira Laborde. - - Piracicaba, 2019.

32 p.

Tese (Doutorado) - - USP / Escola Superior de Agricultura “Luiz de Queiroz”.

1. *Saccharum officinarum* 2. Infrared spectroscopy 3. *Leifsonia xyli* subsp. *xyli* 4. Melhoramento de plantas I. Título

CONTENTS

RESUMO

Identificação de genótipos de cana de açúcar resistentes ao raquitismo da soqueira por meio de técnica de espectroscopia de infravermelho com transformada de Fourier

O raquitismo da soqueira da cana (RS), causada pela bactéria *Leifsonia xyli* subsp. *xyl* (Lxx) pode produzir um impacto substancial negativo na produção de biomassa em todos os países produtores de cana-de-açúcar. Existem poucas opções de manejo para o controle da doença e todas são centradas em medidas de sanitização. Entretanto, este manejo não é eficiente, uma vez que não se elimina por completo a bactéria. Diante do exposto, uma alternativa interessante é a identificação e seleção de genótipos resistentes a doença. No entanto, por se tratar de um patossistema complexo, a seleção por métodos tradicionais não é viável em um programa de melhoramento. Uma nova abordagem, é a identificação de genótipos resistentes por meio de técnica de espectroscopia de infravermelho com transformada de Fourier (FT-IR). Este estudo demonstrou que por meio da técnica de reação em cadeia da polimerase em tempo real (qPCR) foi possível quantificar os títulos bacterianos e determinar os níveis de resistência de dezenove variedades comerciais inoculadas com Lxx. Além disso, por meio da técnica de FT-IR foi possível distinguir genótipos resistentes de suscetíveis baseados em grupos químicos específicos. Por meio de um algoritmo não linear, *support vector machine* (SVM) e uma regressão linear, *partial least square regression* (PLSR), foi possível distinguir genótipos de cana-de-açúcar resistentes e suscetíveis à doença além de predizer a concentração de Lxx nos genótipos por meio de dados espectrais e de título bacteriano. Nossos resultados demonstram que a resistência a esta doença está associada à variações bioquímicas e que os mecanismos de resistência envolvidos podem estar associados à regulação de poliaminas e ao acúmulo de compostos que tenha como principal grupo funcional anel aromático. Os resultados evidenciaram que ambas as abordagens, SVM e PLSR, foram capazes de fornecer estimativas precisas da resistência da cana-de-açúcar à RS e que a técnica FT-IR pode ser utilizada em programas de melhoramento na seleção de genótipos de cana-de-açúcar resistentes ao RS.

Keywords: *Saccharum officinarum*, Infrared spectroscopy, *Leifsonia xyli* subsp. *xyli*, Melhoramento da cana-de-açúcar

ABSTRACT

Identification of resistant sugarcane genotypes to ratoon stunt disease by Fourier-transform infrared analysis

Ratoon stunting disease (RSD), caused by *Leifsonia xyli* subsp. *xyli* (Lxx) can cause substantial impact on biomass production in all sugarcane producing countries. Few management options exist for controlling the disease, and they are all centered on measures of hygiene and sanitation. These measures are not efficient in eliminating the bacteria and can damage bud germination. Phenotypic selection would be the best management practice. However, selection based on Lxx-restrictive multiplication is costly, time-consuming, and destructive. One potentially useful approach for identifying resistant genotypes is Fourier-transform infrared (FT-IR) spectroscopy. In this study, we demonstrate that by using qPCR to quantify bacterial titers and to determine the resistance levels of nineteen commercial varieties inoculated with Lxx and FT-IR spectroscopy we could identify resistant and susceptible genotypes based on specific functional groups. Using a non-linear machine learning techniques (SVM) and a linear partial least square regression (PLSR) we distinguished between resistant and susceptible sugarcane genotypes and predict bacteria concentration using spectral data and Lxx titer quantification. Our Results suggest that resistance may be associated with variation in chemistry. The resistant mechanisms involved against RSD are probably associated to polyamines regulation and aromatic rings compounds accumulation. The results demonstrated that both approaches, SVM and PLSR, are capable of providing accurate estimates of sugarcane resistance to Lxx and can be applied to screen resistant sugarcane in breeding programs.

Keywords: *Saccharum officinarum*, Infrared spectroscopy, *Leifsonia xyli* subsp. *xyli*, Sugarcane breeding

1. GENERAL INTRODUCTION

1.1. Sugarcane crop

Climate change and depletion of fossil energy have stimulated the production of renewable energy (Carvalho-Netto et al 2014). For this reason, sugarcane (*Saccharum* spp.) is the largest source of biofuel and Brazil is the major producer followed by India, China, Pakistan, Thailand, and Mexico (FAO, 2010). In response to a growing demand for ethanol, Brazil has experienced a significant expansion of sugarcane cropping to the Center-South and Center-West drought-susceptible regions in the last years. The cultivated area in 2017/18 was 8.73 million ha and the production reached 633.26 million ton, resulting in 37.87 million ton of sugar and 27.76 billion liters of ethanol (CONAB 2018). The sugarcane agribusiness contributes US$ 43.4 billion to Brazil's gross domestic product per year, corresponding to almost 2% of the entire economy (Urashima et al. 2017).

Sugarcane is a semi-perennial crop and is vegetatively propagated through stalk cuttings (setts) containing lateral buds. When the cane is harvested, the subterranean buds remaining on the stubble give rise to a new crop, named the ratoon crop. In this way, crops can be harvested over multiple years until their productivity decline due to abiotic and biotic factors (Borba and Bazzo 2009), when new crops need to be established from new setts.

The commercial cultivars are derived from a hybridization process involving two species, *S. spontaneum* and *S. officinarum*, that began in the early 1900s (D'Hont et al. 1996). Thus, modern varieties have a high ploidy level and a cytogenetic complexity which make conventional breeding programs time-consuming and costly (Thirugnanasambandam, Hoang, and Henry 2018). In addition, sugarcane improvement based on high sugar content alone has reached a plateau with only slow increments over the years. Thus, in order to improve genetic gains in productivity, breeding programs must focus on traits other than sugar content, especially resistance to biotic and biotic stresses (Dal-Bianco et al. 2012).

Biotic disorders such as ratoon stunt (RS) caused by the vascular bacterium *Leifsonia xyli* subsp. *xyli* (Lxx) can severely affect the growth of sugarcane and reduce the lifetime of crops. Besides, plant genotype and water stress have demonstrated to be key determinant factors for disease severity (Davis, Dean, and Harrison 1988; Grisham 1991). Since cropping expansion to drought-prone areas in Brazil is a reality, it is expected that losses due to this disease will increase. In this sense, development of new chemical and genomic tools such as molecular/biochemical

markers that allow distinguishing between resistant and susceptible genotypes are essential to speed up breeding programs.

1.2. Ratoon stunt and *Leifsonia xyli* subsp. *xyli*

Ratoon stunt can cause substantial impact on biomass production in all sugarcane producing countries (Young 2016). Yield losses vary according to the geographical location and growing season: a 14% loss in the first-year crop increasing to 27% in the third year in Louisiana (Grisham 1991) and 18% yield in the first year followed by 28% in the next year in Ethiopia (Bailey 1997). Annual losses were estimated to be up to US$ 11 million in Australia, US$36 million in Florida and US$1 million in Brazil, in this case considering only 29% of the cultivated area (Urashima et al. 2017).

The major symptoms of the disease consist of a reduction in plant height and in the diameter of the stalks. Symptoms develop mainly on ratoon plants and can be mistaken by other effects that also affect plant growth. Since Lxx dissemination in a commercial field occurs during harvesting with machine or knives, disease management practices rely on preventive and eradication measures. In this context, the establishment of a health seed cane nursery from in vitro culture explants or from heat treatment cane cutting is the best management practices (Hoy et al. 2003; Damann and Benda 1983). The heat treatment consists of exposing the setts to water, air or steam in different temperature-time combinations. In Brazil, the most used treatment consists of immersing the plant material in hot water (50-52 °C) for 30-120 minutes. However, past reports demonstrated that heat treatment is not effective in eliminating Lxx from all plants and can affect negatively bud germination (Damann and Benda 1983; Carvalho et al. 2016). Carvalho et al. (2016) verified the efficiency of heat treatment by quantifying Lxx in two sugarcane varieties (SP80-3280 and SP70-3370). The authors demonstrated that heat treatment can be used to reduce the pathogen population but is not totally effective in eradicating it.

Previous studies related to bacterial density using different cultivars have already demonstrated that sugarcane genotypes differ in bacterial growth and consequently in level of resistance to RS. Thus, a promising control would be to identify and select resistant plants based ability of Lxx to multiply in the host tissues (Davis, Dean, and Harrison 1988). However, selection based on bacterial multiplication is time-consuming and not feasible. For example, screening resistance only for RS would involve inoculating thousands of seedlings and quantifying Lxx using expensive molecular techniques such as quantitative real-time polymerase

chain reaction (PCR). Therefore, it is necessary further studies to fill up the gaps of this pathosystem and also provide new tools for the disease control.

Leifsonia xyli subsp. *xyli* is a gram-positive and fastidious bacterium, member of the Actinomycetales that colonize the xylem vessels and the bundle sheath of sugarcane (Evtushenko et al. 2000; Quecine et al. 2016). Its genome has a high cytosine-guanine (CG) content. It has recently been proposed that the natural host of Lxx is *Saccharum spontaneum* and not the progenitor of the modern sugarcane cultivars, *S. officinarum* (Young 2016). It seems that that modern sugarcanes genotypes acquired the bacteria during the interspecific hybridization crosses performed at the beginning of the last century. This hypothesis also explains the genetic uniformity among Lxx strains (Young et al., 2006), since a single clone was worldwide disseminated by the adoption of a unique cultivar, POJ2878, in breeding programs in 1920 (Young 2016).

As Lxx is a peculiar pathogen that is capable of multiplying in the host tissues without causing apparent disease symptoms, the bacterium is considered and endophyte whose pathogenicity and consequent disease severity rely upon biotic and abiotic factors. Studies related to the genome of Lxx and the interaction of this bacterium with its host provided interesting clues about this dual endophytic/parasitic behavior. For instance, Lxx has a low number (105) of genes correlated to pathogenicity which explain the increase bacterial titer in plant tissue without causing significant symptom (Monteiro-Vitorello et al. 2004). Moreover, its genome carries 307 pseudogenes, a higher number than any other plant bacterial pathogen recorded so far. These non-functional genes, which include genes related to a free-living style, probably arose as a result of a genome decay process and limited the habitat of Lxx to its only known host, i.e., sugarcane (Monteiro-Vitorello et al. 2004).

1.3. Sugarcane-*Leifsonia xyli* subsp. *xyli* interactions

Despite the existence of many microbial species, only a few cause diseases on a given plant species. To prevent microbial infection, plant and animal hosts activate immune responses upon recognition of conserved microbe-associated molecular patterns (MAMPs). Nevertheless, some microbes are successful pathogens. To this end, plant pathogens secrete effectors to avoid or overcome MAMP-triggered immunity in their hosts (Zipfel 2014). Thus, in nature, coevolution of host microbe interaction is a dynamic process driven by selection pressures of diverse natures (Dangl 2006). Although knowledge of plant pathogen interactions has increased over the years, the mechanisms underlying endophyte-plant interaction are still poorly

understood. These organisms live within the plant tissues at least for part of their life cycle without producing any apparent disease symptoms in plants under normal conditions but turn pathogenic when the host plant is stressed or resource-limited (Wilson 1995). In this context, RS is an example of a disease caused by an endophyte described more than 70 years that remains an enigma in many aspects (Young 2016).

At the physiological level, infection by Lxx reduces the activities of phosphoenolpyruvate carboxylase (PEPC) and the photosynthetic and transpiration rates and increases the activities of compounds and enzymes related to oxidative stress, such as superoxide anion radical (O^{2-}), superoxide dismutase (SOD), and peroxidase (POD) (Zhang et al. 2017). In addition, in the complementary study, changes in the vascular tissue such as clogging of xylem vessels and levels of endogenous hormone abscisic acid (ABA), auxins (indoleacetic acid [IAA]), and gibberellic acids (GA3) were reported (Zhang et al. 2016). From both studies, the authors conclude that sugarcane growth is reduced as a consequence of the interference photosynthesis and hormone levels. Most of these changes were supported at the transcriptional level in a recent study that revealed a complex balance between synthesis and degradation of hormones. Genes related to the synthesis of ABA and to the degradation of IAA were up-regulated in plants with higher bacterial titers. Since the metabolism of ABA is correlated to drought-stress and plant growth, the authors suggest that this hormone plays a role in the development of RS symptoms. In addition, in plants with higher bacterial titers most genes involved in the control of the cell-cycle were downregulated, which could also explain the major symptom of the disease (Cia et al. 2018). At the biochemical level, recent study developed by our group defined some metabolites as candidate's resistance biomarkers to Lxx (Moretti 2017). Some phenolic compounds such as ferulic and chlorogenic clorogenic acids were more accumulated in a resistant genotype compared to a susceptible one.

1.4. Fourier transform infrared spectroscopy (FT-IR)

In recent years, infrared spectroscopy (e.g. Fourier-transform infrared, FT-IR) has been applied to plant sciences in an attempt to select genotypes, to monitor chemical changes in plants tissue and also study the pathogen host interactions (Conrad and Bonello 2016; Oh et al. 2018; Willick et al. 2018). FT-IR is a technique that provides information about the chemical structure of a given sample. This technique can be used to identify or discriminate between samples based on qualitative and/or quantitative characteristics associated with differences in molecular absorption (Diem 1993). As such, the spectra generated from FT-IR spectroscopy can be linked

to chemical functional groups and can be used to provide information on the metabolic composition of a sample at specific time point (Fiehn 2001b). When associated with chemometrics (i.e. multivariate analysis of chemical data), the technique can analyze chemical features in the spectrum and distinguish two or more groups of plants based on characteristics of interest in a study (e.g. level of resistance to a pathogen).

FT-IR has many advantages when compared to the traditional metabolic methods, like high-performance liquid chromatography-mass spectrometry (Fiehn 2001b), as it is less time-consuming, reproducible in the analyses of gaseous, liquid and solid samples and has a comparatively low cost. In agriculture, FT-IR has been used for detecting or differentiate microorganisms, health from infected plants and resistant from susceptible genotypes. For example, Chow, Rahman, and Ting (2019) determine changes in the composition of oil palm cell wall related to an endophyte and a host pathogen by quantifying plant cell wall degrading enzymes and also using the FT-IR technique. Results revealed changes in the carbohydrate, xylan and lignin constituents, suggesting that both endophytes and pathogen have elicited similar responses to plant cell walls. Skolik, Mcainsh, and Martin (2018) distinguished damaged and infected tomato fruit (*Solanum lycopersicum*) by the fungus *Geotrichum candidum* from health fruits through alterations in the spectral fingerprint. In a recent study, Oh et al. (2018) demonstrated the great utility of FT-IR spectroscopy in discriminating resistant and susceptible rice plant against to rice blast caused by the fungus, *Magnaporthe oryzae*. Furthermore, using the same technique Conrad et al. (2014) identified *Quercus agrifolia* (coast live oak) resistant to *Phytophthora ramorum* prior infection. Therefore, FT-IR spectroscopy is a useful approach for managing different diseases and a promising strategy to be directly applied to screen plants in breeding programs.

References

Aidan F., N. K. 2006. Learning to classify documents according to genre. Aidan F., Nicholas K. 57:1506–1518.

Bailey, R. A. 1997. The systemic distribution and relative occurrence of bacteria in sugarcane varieties affected by ratoon stunting disease. Proceeding South African Sugarcane Technol. 704:55–56.

Barker, M., and Rayens, W. 2003. Partial least squares for discrimination. J. Chemom. 17:73–166.

Behmann, J., Mahlein, A., Rumpf, T., Ro, C., and Plu, L. 2015. A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. :239–260.

Carvalho, G., da Silva, T. G. E. R., Munhoz, A. T., Monteiro-Vitorello, C. B., Azevedo, R. A., Melotto, M., et al. 2016. Development of a qPCR for Leifsonia xyli subsp. xyli and quantification of the effects of heat treatment of sugarcane cuttings on Lxx. Crop Prot.

Carvalho-Netto, O. V., Bressiani, J. A., Soriano, H. L., Fiori, C. S., Santos, J. M., Barbosa, G. V., ... & Pereira, G. A. (2014). The potential of the energy cane as the main biomass crop for the cellulosic industry. Chemical and

Biological Technologies in Agriculture 1(1), 20.

Cia, M. C., De Carvalho, G., Azevedo, R. A., Monteiro-Vitorello, C. B., Souza, G. M., Nishiyama-Junior, M. Y., et al. 2018. Novel insights into the early stages of ratoon stunting disease of sugarcane inferred from transcript and protein analysis. Phytopathology. 1–58.

Clemmensen, L., Hastie, T., Witten, D., and Ersbøll, B. 2011. Sparse discriminant analysis. Technometrics. 53:4.

Conrad, A. O., and Bonello, P. 2016. Application of Infrared and Raman Spectroscopy for the Identification of Disease Resistant Trees. Front. Plant Sci. 6:1–8 Available at: http://journal.frontiersin.org/Article/10.3389/fpls.2015.01152/abstract.

Conrad, A. O., Rodriguez-Saona, L. E., McPherson, B. A., Wood, D. L., and Bonello, P. 2014. Identification of Quercus agrifolia (coast live oak) resistant to the invasive pathogen Phytophthora ramorum in native stands using Fourier-transform infrared (FT-IR) spectroscopy. Front. Plant Sci. 5:1849.

Coutinho, I. D., Baker, J. M., Ward, J. L., Beale, H., Creste, S., and Cavalheiro, A. J. 2016. Metabolite profiling of sugarcane genotypes and identification of flavonoid glycosides and phenolic acids. J. Agric. Food Chem. 64:4198–4206.

Chow, Y. Y., Rahman, S., & Ting, A. S. Y. 2019. Evaluating the host defense responses in oil palm to complex biocontrol endophyte–pathogen–host plant interaction via Fluidigm® real-time polymerase chain reaction (RT-PCR). Biological Control, 129, 148-157.

Damann, A. J. K. E., and Benda, G. T. 1983. Evaluation of commercial heat treatment methods for control of ratoon stunting disease of sugarcane. Plant Dis. 67:966–967.

Davis, M. J., Dean, J. L., and Harrison, N. A. 1988. Quantitative variability of Clavibacter xyli subsp. xyli populations in sugarcane cultivars differing in resistance to ratoon stunting disease. Phytopathology. 78:462–468.

Devitt, G., Howard, K., Mudher, A., and Mahajan, S. 2018. Raman Spectroscopy: An Emerging Tool in Neurodegenerative Disease Research and Diagnosis. ACS Chem. Neurosci. 9:404–420.

Diem, M. 1993. *Introduction to Modern Vibrational Spectroscopy*.

DS Teakle, P. S. and D. S. 1973. Association of a small coryneform bacterium with the ratoon stunting disease of sugar-cane. Aust. J. Agric. Res. 24:869–874.

D'Hont A, Grivet L, Feldmann P, Rao S, Berding N, Glaszmann JC. 1996. Characterisation of the double structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. Mol Gen Genet. 250:405-413

Ehsani, M. R. et al. 1999. A NIR technique for rapid determination of soil mineral nitrogen. Precis. Agric. 1:217–234.

Evtushenko, L. I., Dorofeeva, L. V., Subbotin, S. A., and Cole, J. 2000. Leifsonia poae gen. nov., sp. nov., isolated from nematode galls on Poa annua, and reclassification of 'Corynebacterium aquaticum' Leifson 1962 as Leifsonia aquatic (ex Leifson 1962) gen. nov., nom. rev., comb.nov. and Clavibacter xyli Davis et al. 1984 w. Int. J. Syst. Evol. Microbiol. 50:371–380.

Fegan, M., Croft, B. J., Teakle, D. S., Hayward, A. C., and Smith, G. R. 1998. Sensitive and specific detection of Clavibacter xyli subsp. xyli, causal agent of ratoon stunting disease of sugarcane, with a polymerase chain reaction-based assay. Plant Pathol. 47:495–504 Available at: http://doi.wiley.com/10.1046/j.1365-3059.1998.00255.x.

Fiehn, O. 2001. Combining genomics, metabolome analysis, and biochemi- cal modelling to understand metabolic networks. Comp. Funct. Genomics. 2:155–168.

González-Camacho, J. M., Ornella, L., Pérez-Rodríguez, P., Gianola, D., Dreisigacker, S., and Crossa, J. 2018. Applications of Machine Learning Methods to Genomic Selection in Breeding Wheat for Rust Resistance. Plant Genome. 11 Available at: https://dl.sciencesocieties.org/publications/tpg/abstracts/11/2/170104.

Grisham, M. P. 1991. Effect of ratoon stunt on yield of sugarcane grown in multiple three-year plantings. Phytopathology 81:337-340.

Hoy, J. . W., Bischoff, K. P., Milligan, S. B., and Gravois, K. 2003. Effect of tissue culture explant source on sugarcane yield components. Euphytica. 129:237–240.

Huang, N., Ling, H., Su, Y., Liu, F., Xu, L., Su, W., et al. 2018. Transcriptional analysis identifies major pathways as response components to Sporisorium scitamineum stress in sugarcane. Gene. 678:207–218 Available at: https://doi.org/10.1016/j.gene.2018.08.043.

Hussein, H. M., Hameed, R. H., and Hameed, I. H. 2018. Screening of Bioactive Compounds of Ricinus communis Using GC-MS and FTIR and Evaluation of its Antibacterial and Antifungal Activity. Indian J. Public Heal. Res. Dev. 9:463–469.

Kucheryavskiy, S. 2018. Multivariate Data Analysis for Chemometrics. R Packag. Version 0.9.1.

Leme, G. M., Coutinho, I. D., Creste, S., Hojo, O., Carneiro, R. L., Bolzani, V. D. S., et al. 2014. HPLC-DAD method for metabolic fingerprinting of the phenotyping of sugarcane genotypes. Anal. Methods. 6:7781–7788.

Maríía Estrella Legaz, Roberto de Armas, D. P. and C. V. 1998. Relationships between phenolics-conjugated polyamines and sensitivity of sugarcane to smut (Ustilago scitaminea). J. Exp. Bot. 49:1723–1728.

Martínez-Martínez, V., Gomez-Gil, J., Machado, M. L., and Pinto, F. A. C. 2018. Leaf and canopy reflectance spectrometry applied to the estimation of angular leaf spot disease severity of common bean crops ed. Dawn L Arnold. PLoS One. 13:e0196072 Available at: https://dx.plos.org/10.1371/journal.pone.0196072.

Mathlouthi, M. K. and M. 1996. FTIR and laser-Raman spectra of oligosaccharides in water: characterization of the glycosidic bond. Carbohydr. Res. 284:145–157.

Meyer, L., Dimitriadou, E., Hornik, K., Weingessel, A., and Leisch, F. 2015. e1071: misc functions of the Department of Statistics, probability theory group (formerly: E1071), TU Wien. R package version 1.6-7.

Monteiro-Vitorello, C. B., Camargo, L. E. A., Sluys, M. A. Van, Kitajima, J. P., Truffi, D., Do Amaral, A. M., et al. 2004. The Genome Sequence of the Gram-Positive Sugarcane Pathogen Leifsonia xyli subsp. xyli. Mol. Plant. Microbe. Interact. 17:827–836.

Moretti, F. R. R. de C. 2017. Identification of candidate resistance metabolites to Leifsonia xyli subsp. xyli in sugarcane through metabolomic profiling.

Nagasubramanian, K., Jones, S., Sarkar, S., Singh, A. K., Singh, A., and Ganapathysubramanian, B. 2018. Hyperspectral band selection using genetic algorithm and support vector machines for early identification of charcoal rot disease in soybean stems. Plant Methods. 14:2–13.

Oh, M., Barnaby, J. Y., Lee, H., Jia, Y., and Kim, M. S. 2018. Infrared fingerprint profiles of compatible and incompatible interactions of the rice blast pathogen using FTIR spectroscopy. In *Sensing for Agriculture and Food Quality and Safety*,.

Pan, Y.-B., Grisham, M. P., Burner, D. M., Damann, K. E., and Wei, Q. 1998. A Polymerase Chain Reaction Protocol for the Detection of Clavibacter xyli subsp. xyli, the Causal Bacterium of Sugarcane Ratoon Stunting Disease. Plant Dis. 82:285–290 Available at: https://apsjournals.apsnet.org/doi/pdf/10.1094/PDIS.1998.82.3.285 [Accessed April 26, 2018].

Pospiech, A., and Neumann, B. 1995. A versatile quick-prep of genomic DNA from gram-positive bacteria. Trends Genet. 11:217–218.

Purcell, D. E., Oxley, J. P., Cox, M. C., Croft, B. J., and O'Shea, M. G. 2010. On-site rapid screening for sugarcane smut resistance using near-infrared (NIR) spectroscopy. In *Australian Society of Sugar Cane Technologists*, , p. 366–374.

Quecine, M. C. et al. 2015. A stable Leifsonia xyli subsp. xyli GFP-tagged strain reveals a new colonization niche in sugarcane. Plant Pathol. 65:154–162.

Quecine, M. C., Silva, T. M., Carvalho, G., Saito, S., Mondin, M., Teixeira-silva, N. S., et al. 2016. A stable Leifsonia xyli subsp . xyli GFP-tagged strain reveals a new colonization niche in sugarcane tissues. Plant Pathol. 65:154–162.

Saitta, L. 1995. Support-Vector Networks. In *Machine Leaming*, , p. 273–297.

Savitzky, A., and Golay, M. J. E. 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36:1627–1639.

Shi, Y., Huang, W., Ye, H., Ruan, C., Xing, N., Geng, Y., et al. 2018. Partial Least Square Discriminant Analysis Based on Normalized Two-Stage Vegetation Indices for Mapping Damage from Rice Diseases Using PlanetScope Datasets. Sensors. 18:1901 Available at: http://www.mdpi.com/1424-8220/18/6/1901.

Silva, L. J. 2015. Laercio: Duncan test, Tukey test and Scott-Knott test.

Skolik, P., McAinsh, M. R., & Martin, F. L. 2019. ATR-FTIR spectroscopy non-destructively detects damage-induced sour rot infection in whole tomato fruit. Planta, 249(3), 925-939.

Thirugnanasambandam, P. P., Hoang, N. V., & Henry, R. J. 2018. The challenge of analyzing the sugarcane genome. Frontiers in plant science, 9.

Thumanu, K., Wongchalee, D., Sompong, M., Phansak, P., Le Thanh, T., Namanusart, W., et al. 2017. Synchrotron-based FTIR microspectroscopy of chili resistance induced by *Bacillus subtilis* strain D604 against anthracnose disease. J. Plant Interact. 12:255–263 Available at: https://www.tandfonline.com/doi/full/10.1080/17429145.2017.1325523.

Urashima, A. S., Silva, M. F., Correa, J. J., Moraes, M. C., and Singh, A. V. 2017. Prevalence and Severity of Ratoon Stunt in Commercial Brazilian Sugarcane Fields. Plant Dis. 101:815–821 Available at: https://apsjournals.apsnet.org/doi/pdfplus/10.1094/PDIS-07-16-1030-RE [Accessed March 6, 2018].

Wilson D. 1995. Endophyte–the evolution of a term, and clarification of its use and definition. Oikos 73:274–276

Willick, I. R., Lahlali, R., Vijayan, P., Muir, D., and Karunakaran, C. 2018. Wheat flag leaf epicuticular wax morphology and composition in response to moderate drought stress are revealed by SEM , FTIR-ATR and synchrotron X-ray spectroscopy. :316–332.

Wold, S., Ruhe, A., Wold, H., and Dunn, III, W. J. 1984. The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses. SIAM J. Sci. Stat. Comput. 5:735–743 Available at: http://epubs.siam.org/doi/10.1137/0905052.

Young, A. J. 2016. Possible origin of ratoon stunting disease following interspecific hybridization of *Saccharum* species. Plant Pathol.

Zipfel, C. 2014. Plant pattern-recognition receptors. Trends Immunol. 35: 345–351

Zhan-yu, L., Jing-feng, H., Jing-jing, S., Rong-xiang, T., Wan, Z., and Li-li, Z. 2007. Characterizing and estimating rice brown spot disease severity using stepwise regression, principal component regression and partial least-square regression. J. zhejiang Univ. Sci. b. 8:738–744.

2. IDENTIFICATION OF RESISTANT SUGARCANE GENOTYPES TO RATOON STUNT DISEASE BY FOURIER-TRANSFORM INFRARED ANALYSIS

ABSTRACT

Ratoon stunting (RS), caused by *Leifsonia xyli* subsp. *xyli* (Lxx) is the most important biotic constraint on the production of sugarcane. The best management practice would be through the utilization of resistant germplasm. However, phenotypic selection based on inoculation trials is not feasible because the symptoms of the disease take a long time to develop and can be easily mistaken by general symptoms of poor plant growth caused by many stressing factors. One interesting approach to overcome this limitation is the identification of resistant genotypes by Fourier-transform infrared (FT-IR) spectroscopy. In this study, we demonstrated that by using qPCR to quantify bacterial titers and to determine the resistance levels of nineteen commercial varieties inoculated with Lxx and FT-IR spectroscopy to gather data on their metabotypes, we were able to distinguish resistant from susceptible genotypes based on specific chemical groups. Two trials were set up to assess the degree of resistance and to gather FT-IR spectra. Through the support vector machine (SVM) algorithm it was possible to distinguish between resistant and susceptible genotypes with an accuracy of 100% for model development of 80 and 83% for validation in the first and second trials, respectively. Interestingly, different spectral regions distinguished the genotypes in each trial, indicating the effect of experimental conditions on spectral data. For the first trial, resistant plants showed higher absorbance in peaks corresponding to polyamines vibrations and susceptible for amide and carbohydrates. In contrast, in the second trial resistant plants showed higher peaks in spectral regions associated to compounds with aromatic ring as a chemical functional group. Additionally, the amount of Lxx in resistant and susceptible genotypes was predicted using partial least squares regression (PLSR), demonstrating a strong correlation between predicted (based on FT-IR spectra) and measured bacterial titer levels.The results demonstrate that FT-IR spectroscopy can predict the degree of resistance of sugarcane to the multiplication of the pathogen in its tissues. Therefore, the technique can be applied to screen resistant sugarcane genotypes in breeding programs.

Keywords: *Saccharum officinarum*, Infrared spectroscopy, *Leifsonia xyli* subsp. *xyli*, Sugarcane breeding

2.1. INTRODUCTION

Global interest in sugarcane cropping has increased over the years due to its economic importance as an alternative to fossil fuels (Faostat, 2016). Brazil is the world's largest producer of sugarcane (*Saccharum* spp.), which contributes over 10.436 million US dollars in exports per year (Mapa, 2017). However, the cropping of sugarcane is hampered by several biotic and abiotic stresses. An important biotic constraint is ratoon stunt (RS), caused by *Leifsonia xyli* subsp. *xyli* (Lxx) (Urashima et al. 2017). Lxx is a gram-positive bacterium, member of the order Actinomycetales and only reported up to now as a pathogen of sugarcane. The bacteria systemically colonizes the xylem vessels, the mesophyll and also the bundle sheath cells (Bailey 1997; Quecine 2015). The symptoms of RS are characterized by a reduction in height and in stalk

diameter mainly on ratoon plants as a consequence of increased bacterial titers which leads to significant reduction in biomass over the years (Davis, Dean, and Harrison 1988). A few management options exist for controlling the disease and they are all centered on sanitation measures, such as the establishment of healthy seed cane nurseries from *in vitro* cultured explants or from heat-treated cane cuttings (setts) (Damann and Benda 1983; Hoy et al. 2003). However, these measures are not effective in eliminating the bacteria and can damage bud germination (Damann and Benda 1983; Carvalho et al. 2016). Therefore, the best management practice would be through and the use of resistant genotypes, since it is not feasible to control the bacterium after it becomes established in the field. However, screening trials to evaluate the level of sugarcane resistance are based on quantitative methods that require molecular or serological methods. As such, these approaches are labor-intensive, time-consuming and not practical as a breeding tool. Therefore, it is essential to develop alternative methods to to identify resistant genotypes.

Some studies support the hypothesis that specific plant metabolites, in particular phenolic compounds, are capable of producing comprehensive chemical fingerprints that may be used to distinguish susceptible genotypes from resistant ones (Conrad et al. 2014; Thumanu et al. 2017; Hussein, Hameed, and Hameed 2018). Recently, Moretti (2017) observed that susceptible (CB49-260) and resistant (SP80-3280) sugarcane varieties inoculated or not with Lxx accumulate distinct levels of compounds according to the resistance level of the genotype. In agreement, some studies revealed the chemical diversity of the main sugarcane genotypes used in Brazil and hypothesized that each genotype might respond to abiotic and biotic stress by accumulating different metabolites (Coutinho et al. 2016; Leme et al. 2014). Moreover, sugarcane genotypes with higher levels of luteolin-8-c-glucoside, a phenolic compound, are resistant to orange rust, caused by the fungus *Puccinia kuehnii* (Leme et al 2014).

One potentially useful approach for identifying resistant genotypes is Fourier-transform infrared (FT-IR) spectroscopy (Conrad et al., 2016). This technique can be used to identify or discriminate between samples based on qualitative and quantitative characteristics, such as resistance to pathogens, associated with differences in molecular absorption (Diem 1993). The spectra generated from FT-IR spectroscopy can be linked to chemical funtional groups and provide information on the metabolic composition of a sample at a given time (Fiehn 2001). A study developed with coast live oak (*Quercus agrifolia* Née) demonstrated the efficiency of this technique in identifying disease resistance genotypes (Conrad et al. 2014). The authors selected resistant and susceptible trees in a natural forest population prior to infection with *Phytophthora ramorum* and predicted the concentration of two phenolic compounds associated with resistance.

Further, a similar technique using near-infrared (NIR) spectroscopy was used as an on-site screening method to predict sugarcane smut resistance for plant breeding purposes (Purcell et al. 2010). Since infrared spectroscopy techniques are reproducible and require only basic simple preparation, it is a promising tool to identify resistant genotypes to complex diseases, such as RS.

Statistical models based on partial least square-regression (PLSR) and support vector machine (SVM) analysis presented good performance in predicting agronomic traits (Nagasubramanian et al. 2018; Shi et al. 2018). In general, these approaches can analyze many variables at the same time and identify distinct data patterns associated with specific phenotypes (Behmann et al. 2015; Wold et al. 1984). SVM is a powerful classification method capable of analyzing highly complex data with few statistical assumptions, while PLSR is a complementary tool for selecting resistant genotypes that can be used to estimate quantitative traits associated with plant disease severity. Successful applications using classification and regression analyses in crop protection include early pathogen detection, characterization, and disease estimation, as well as resistant genotype identification (Martínez-Martínez et al. 2018; González-Camacho et al. 2018; Zhan-yu et al. 2007).

Therefore, this study explored two objectives: (i) to develop a Fourier-transform infrared spectroscopy based-algorithm to identify and predict resistant sugarcane genotypes to Lxx using a non-linear support vector machine model, and (ii) to determine and predict the amount of bacteria in resistant and susceptible genotypes by partial least squares-regression.

2.2. MATERIALS AND METHODS

2.2.1. Resistance screening of commercial varieties – first trial

A first trial was conducted in the greenhouse from April to August 2017 to evaluate the level of resistance of commercial genotypes by quantifying bacterial titers in leaf tissue 85 days after inoculation. Single bud setts of approximately 5 cm in length were cut from stalks of *in vitro*-generated plants of 19 commercial sugarcane varieties (IAC91-1099, IACSP94-2094, IACSP94-2101, IAC91-2218, IACSP95-5000, IACSP96-7569, IACSP97-4039, IACSP93-3046, IACSP96-2042, RB83-5486, RB86-7515, RB-579, RB96-5917, RB96-6928, RB96-5902, SP78-4467, SP80-3280, SP80-1816, and CB49/260) provided by the Sugarcane Research Station of the Instituto Agronômico de Campinas (Ribeirão Preto, SP). The cane setts were germinated in trays containing commercial substrate (Basaplant, BASE) and the seedlings were transplanted to 3 L

plastic pots containing 3 L of substrate (Basaplant, BASE) amended with 5 g of 8-28-16 NPK fertilizer.

Inoculations were performed according to Carvalho et al. (2016) when the plants were 1-month-old by cutting off the leaf whorl just above the meristem and by placing 50 μL of a liquid culture (OD_{600}=0.7) of Lxx strain CTCB07 grown in M-SC medium (DS Teakle 1973; Monteiro-Vitorello et al. 2004) on top of the plant stub. The base of the first youngest expanded leaf (+1) was collected from each plant just before inoculation to assess if they were Lxx-free by conventional PCR using Lxx-specific primers(Pan et al. 1998). Leaf samples were kept at -80ºC until they were ground in liquid nitrogen using a mortar and a pestle for DNA extraction and FT-IR analysis. The presence of Lxx was evaluated in each sample using conventional PCR, which in this case can detect as low as 10^4 cells/mL of Lxx suspension (Fegan et al. 1998). The inoculated plants were distributed in a completely randomized design with five biological replicates per genotype and kept in the greenhouse.

At 85 days after inoculation (DAI) leaf from each genotype was collect (leaf +12) for the quantification of bacterial titers. Approximately 75g of ground tissue was used for DNA extraction according to Pospiech and Neumann (1995) following modifications by Carvalho et al. (2016). Quantification was carried out with the forward Lxx12950F1 (GCACATCGATCTGGAAAAAAGG) and the reverse Lxx12950R1 (CCGCAGTCTCACGCATACC) primers as described in Carvalho et al. (2016). Briefly, the quantitative PCR amplifications were performed in a 7500 FAST real-time thermocycler (Applied Bio-systems, U.S.A) using the Platinum SYBR® Green qPCR SuperMix UDG kit (Invitrogen, U.S.A.) in accordance with the manufacturer's instructions. The reactions were conducted using 12.5 μL of the SuperMix buffer, 0.5 μL a 10 mM of each primer, 0.5 μL of a 2.5 mM solution of ROX, 9 μL of nuclease-free water (Integrated DNA Technologies, U.S.A.), and 2 μL of the DNA solution (50 ng/μL). The amplification was carried out with an initial cycle of 50 °C for 2 min and 95 °C for 5 min followed by 40 cycles of 95 °C for 10 s and 60 °C for 30 s. All reactions were performed in duplicates using positive (Lxx DNA) and negative (nuclease-free water) controls. The threshold cycle (Ct) was used to estimate Lxx cell numbers in sugarcane samples based on Ct values of a dilution series of Lxx-DNA.

2.2.2. Confirmation of resistance levels of selected sugarcane genotypes – second trial

In order to confirm the results of the first trial, a second trial was conducted from January to May 2018 with five resistant (SP80-3280, SP80-1816, RB83-5486, IACSP96-7569, and IAC91-2218) and five susceptible (RB-579, SP78-4467, IACSP96-2042, RB86-7515, and IACSP97-4039) genotypes as determined in the first trial. Bud setts were taken from the same plants used in the first trial that were kept in the greenhouse for approximately 9 months. Sample collection and Lxx analyses prior and after inoculation were conducted and assessed as described for the first trial. The plants were distributed in a CRD with ten biological replicates per genotype and kept in the greenhouse.

2.2.3. FT-IR spectroscopy

For FT-IR spectroscopy, 200±1 mg fresh weight (FW) of finely ground leave tissue was extracted with 70% acetone, followed by purification with two volumes of chloroform (Villari et al. 2018). Extracts were then lyophilized for 48 hours. Lyophilized samples were placed inside sealed bags with silica gel and stored at room temperature. Prior to analysis, the samples were re-suspended in 1 mL Milli-Q water following the method of Villari et al. (2018), except that sample clean-up with C18 columns was not performed because the samples were colorless. Instead, the water was removed from the samples by lyophilization or evaporation using a SpeedVac at room temperature. After drying, samples were re-suspended in 200µl of methanol and 2µl of the concentrated extracts were transferred to the ATR accessory crystal and allowed to rest for 2 min to allow the evaporation of methanol. Two technical replicates were taken for each sample extract. Extracts from both trials were evaluated by FT-IR spectroscopy using a Cary 630 FT-IR spectrometer (Agilent Technologies Inc., Santa Clara, CA, USA) equipped with a one-bounce attenuated total reflectance (ATR) accessory. Spectra were collected over the range of 4000-700cm^{-1} wavenumbers at a resolution of 4cm^{-1}, with 64 scans for each sample. Spectral data were collected and displayed using MicroLab PC software (Agilent Technologies Inc., Santa Clara, CA, USA). Analyses were performed at the Laboratory of Molecular and Chemical Ecology of Trees of the Ohio State University.

2.2.4. Data analysis and model development

Data from bacterial titers (Lxx cells/100 ng of sugarcane DNA) were tested for normality by the Shapiro-Wilk test, transformed to log base 10 and compared with the Scott-Knott multiple range tests (P=0.05) using the R package Laercio (Silva 2015).

Spectral analyses were performed using the chemometric software Pirouette version 4.5. (Infometrix, Inc., Bothell, WA, USA) and by the R software version 3.5.1 (R Development Core Team, 2008). For Pirouette, spectral regions of interest were identified and included between 2150-700 and 800-1800 cm^{-1} wavenumbers for the first and second trials, respectively. Outliers were identified by Coomans plots and 3D class projection plots and excluded from the models.

Two different algorithms were used to correlate FT-IR spectra of leaf tissue with resistance to Lxx. The first consisted of developing a non-linear support vector machine (SVM) model with a radial base function kernel (Vapnik, 1998). The SVM is a classification method, which finds pattern in the data set with regard to pre-label classes (resistant or susceptible genotype). This algorithm transforms the original data (e.g. spectral data) from the low dimension space to the high dimension space and constructs a hyperplane to maximize the separation of the different sample classes (i.e. resistant or susceptible) (Saitta 1995). The second method was the partial least squares-regression (PLSR), which searches for the best relationship between spectral data (independent variables) and the resistance level of the genotypes expressed as bacterial titers (dependent variable) (Ehsani 1999; Barker and Rayens 2003).

The discriminant functions for the SVM (defining hyperplanes in a multidimensional space) were fitted following the model:

$$\mathbf{g(x)=w^{T}x + w_0} \quad (1)$$

where $\mathbf{w}$ is the weight vector (orthogonal to the decision hyperplane) and $\mathbf{w_0}$ is the threshold.

For the PLSR model, we fitted the following equations:

$$\mathbf{X = TP^{T}+E} \quad (2)$$

and

$$\mathbf{Y = UQ^{T}+F} \quad (3)$$

where $\mathbf{T}$ and $\mathbf{U}$ are n x p matrices of the p extracted score vectors (latent vectors), the $\mathbf{P}$ (N x p matrix) and $\mathbf{Q}$ (M x p matrix) represent matrices of loadings with number of columns equal to the number of PLS components, and E (n x N matrix) and F (n x M) are the residuals. The x-score in $\mathbf{T}$ are linear combinations of the x-variables (spectral data), and the y-score in $\mathbf{U}$ are linear combinations of the y-variables (bacterial titer).

For SVM analysis, a multivariate variable screening process based upon Gram–Schmidt QR-Decomposition was performed in order to remove data redundancy (Aidan F. 2006; Meyer et al. 2015). The best SVM model parameters were determined using 20-fold cross-validation. The datasets were not transformed prior analysis. For PLSR, data were transformed by taking the second derivative and smoothed, using the R packages sparse LDA (Clemmensen et al. 2011) and mdatools (Kucheryavskiy 2018) with 14 and 11 as width parameter for the first and second trial, respectively (Savitzky and Golay 1964). Two technical replicates of each genotype were used for both models.

The SVM and PLSR models were developed separately for each trial. The models developed from plant tissue collected in the first trial used the spectra from five resistant and five susceptible genotypes with nine biological replicates per genotype. The SVM model of the second trial used the plant tissue of all 19 genotypes with five biological replicates per genotype and the classification of genotypes (resistant/susceptible) as determined in the first trial. Whereas for the PLSR models we used the spectra and bacterial titer data only from genotypes which were previously assessed by qPCR.

2.2.5. Validation of PLSR and SVM models

In order to evaluate the predictive value of the SVM and PLSR models, a set of plants (testing data set) not used in the development of the models were selected from both trials and their spectral data were applied to the respective models. The testing set comprised 18 samples (extract of plants) randomly selected for the first trial and 19 samples for the second trial. Model performances and accuracy were assessed by the Root Mean Square Error (RMSEP) and the determination coefficient (R^2).

2.3. RESULTS

2.3.1. Resistance screening of commercial varieties – first trial

In the first trial, conventional PCR for Lxx resulted in no amplification (data not shown). Thus, the plants used in this assay were assumed to be Lxx-free. The quantification of bacterial titers at 85 DAI allowed to distinguish the genotypes in two groups, referred to as resistant and susceptible genotypes. Genotypes classified as susceptible (n=11) to RS had significantly (P=0.000) higher levels of Lxx titers in their tissues than the resistant (n=8) ones (Figure 1a).

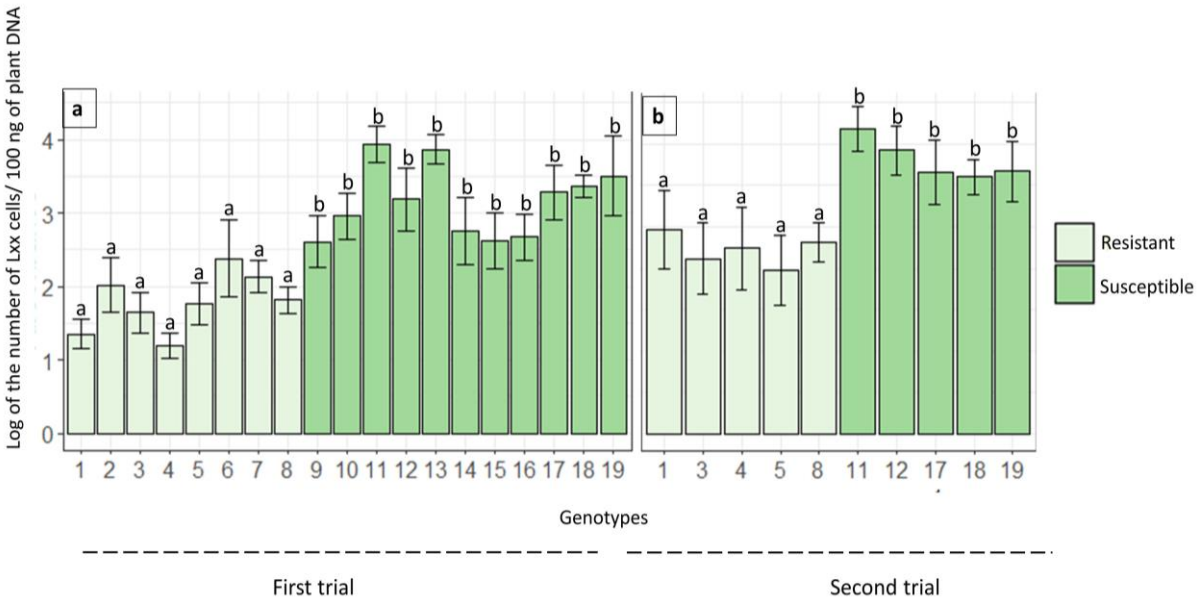

Figure 1. Log of Lxx cells in 100 ng of DNA of sugarcane genotypes inoculated with *Leifsonia xyli* subsp. *xyli* (Lxx) at 85 days after inoculation. The graphs represents data from two separate trials (A and B), with the genotypes: SP80-1816 (1), IAC91-1099 (2), RB83-5486 (3), SP80-3280 (4), IACSP96-7569 (5), RB96-6928 (6), RB96-5902 (7), IAC91-2218 (8), IACSP93-3046 (9), IACSP94-2094 (10), RB-579 (11), IACSP97-4039 (12), CB49/260 (13), IACSP94-2101 (14), IACSP95-5000 (15), RB96-5917(16), RB86-7515 (17), IACSP96-2042 (18), and SP78-4467 (19). Letters indicate significantly different means based on Scott-Knott test ($P < 0.05$) and vertical lines represent the standard errors

2.3.2. Confirmation of resistance levels of selected sugarcane genotypes – second trial

In the second trial, amplification products (positive for Lxx) were observed for some plants before inoculation (data not shown). Therefore, five replicates from ten genotypes were randomly collected for bacterial quantification using qPCR. Among the then genotypes evaluated, the variety SP80-1816 exhibited the lowest (P=0.0002) bacterial titer followed by the variety IACSP96-7569. The other genotypes did not differ statically (Figure 2). Regarding the bacterial titers after inoculation, the five resistant genotypes had significantly (P=0.003) lower titers compared to the susceptible ones, confirming the results obtained of the first trial (Figure 1b).

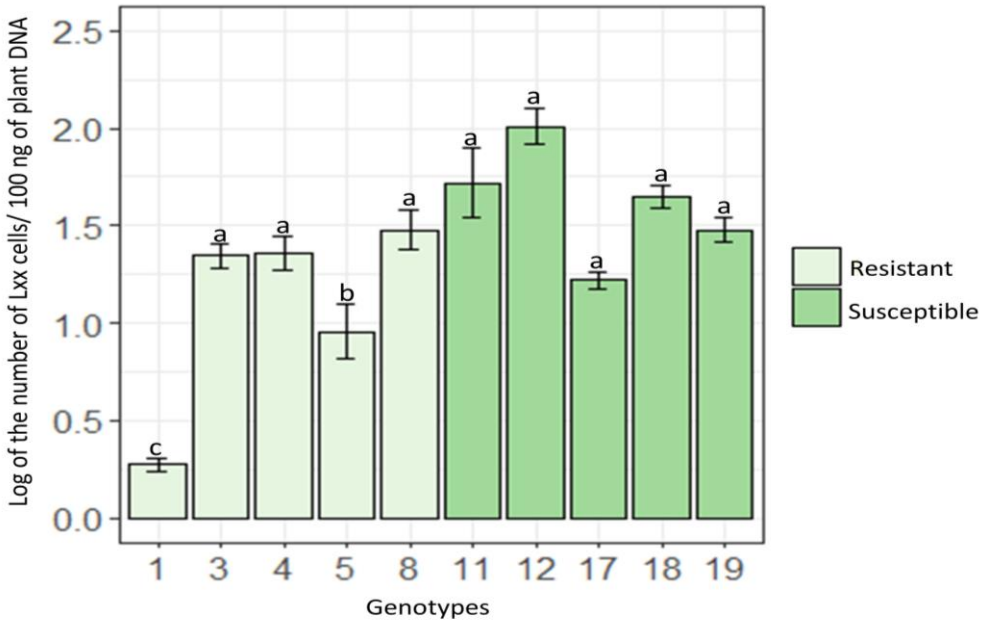

Figure 2. Log of Lxx cells in 100 ng of DNA in non-inoculated 1-month-old plants of SP80-1816 (1), RB83-5486 (3), SP80-3280 (4), IACSP96-7569 (5), IAC91-2218 (8), RB-579 (11), IACSP97-4039 (12), RB86-7515 (17), IACSP96-2042 (18), and SP78-4467 (19). Letters indicate significant different means based on Scott-Knott test ($P < 0.05$). Vertical bars represent standard errors.

2.3.3. Chemical fingerprint analysis

2.3.3.1. FT-IR spectra analysis – model development

The average absorption of the raw spectral data showed quantitative differences among sugarcane genotypes in the range of 700-1700 and 1400-1700 cm^{-1} wavenumbers for the first and second trials, respectively (Figure 3a, b). Using SVM multivariate analysis, it was possible to distinguish the resistant from the susceptible genotypes with an accuracy of 100%, for both trials (Table 1). After using pre-processing methods based upon Gram–Schmidt QR-Decomposition (Item 2.2.4) to improve model prediction, some regions of the spectra were removed, and we developed our model with 120 and 62 wavenumbers. Among the selected wavenumbers for model development only six and three spectral bands for the first and second trials, respectively, were significantly different between the resistant and susceptible genotypes (Figure 4).

Regarding the first trial, significant spectral differences were observed at 998, 1656, 1979, 1981, 2014 and 2035 cm^{-1}. In this trial, susceptible genotypes had greater absorbance at 998 and 1656 cm^{-1}, which are typically associated with carbohydrate and amide I α helix protein vibrations, respectively (Mathlouthi 1996; Devitt et al. 2018), whereas resistant plants had higher absorbance in spectra regions corresponding to 1979, 1981 to 2014 cm^{-1}, which are related to polyamines vibrations (Diem 1993). In the second trial, quantitative differences between resistant

and susceptible genotypes were observed around 1508, 1524, and 1541 cm^{-1}, which are associated with vibrations of the benzene ring group (C=C) (Diem, 1993). The absorbance was higher for resistant genotypes at the three wavenumbers when compared to the susceptible ones.

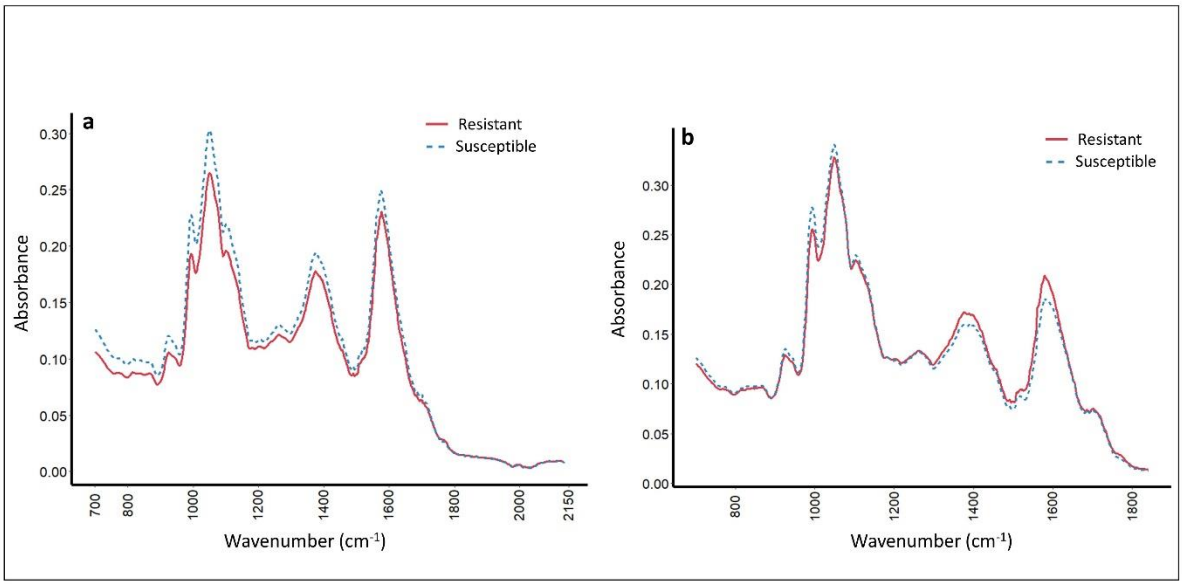

Figure 3. Average raw infrared spectra of leaf extracts of Lxx resistant (solid lines) and susceptible (dashed lines) sugarcane genotypes for the first (a) and second (b) trials.

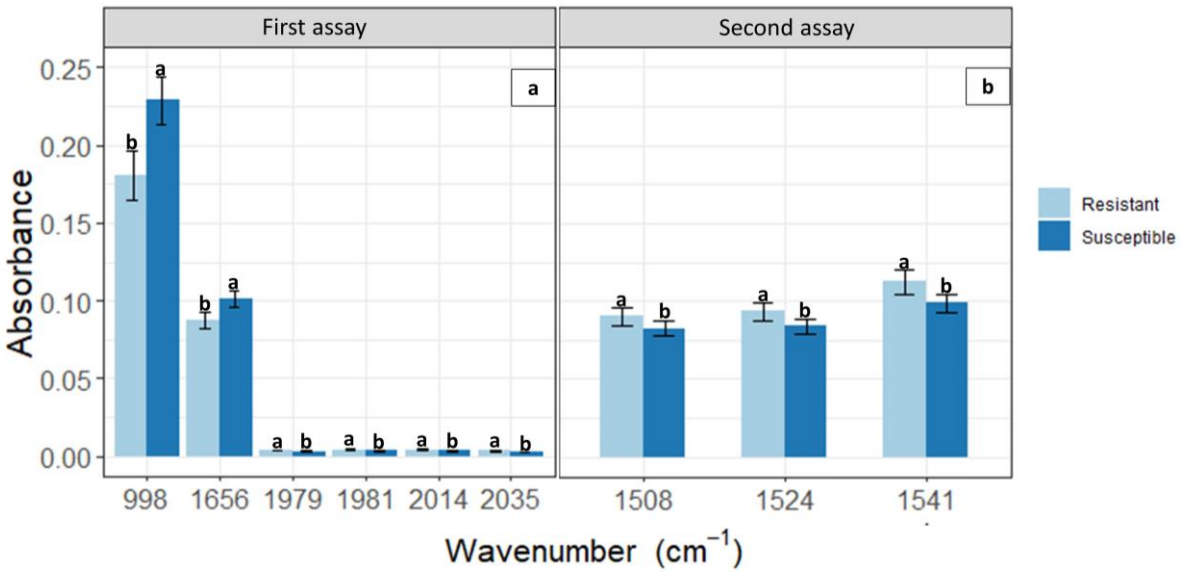

Figure 4. Medians and standard deviations (SD) of spectral bands with high discriminating power identified by SVM for the (a) first and (b) second trials. Different lowercase letters indicate significant differences in spectral band absorbance between resistant and susceptible genotypes according to the non-parametric Wilcoxon test ($p<0.05$).

Bacterial titers were predicted with PLSR model using the selected spectral data from SVM and the quantification data from each trial (Figure 5a, b). In the first trial, the cutoff between resistant and susceptible genotypes was approximately 10^2 (Log of the number of Lxx cells/100 ng of plant DNA). Whereas for the second trial the cutoff between resistant and susceptible genotypes was 10^3 (log of the number of Lxx cells/100 ng of plant DNA).

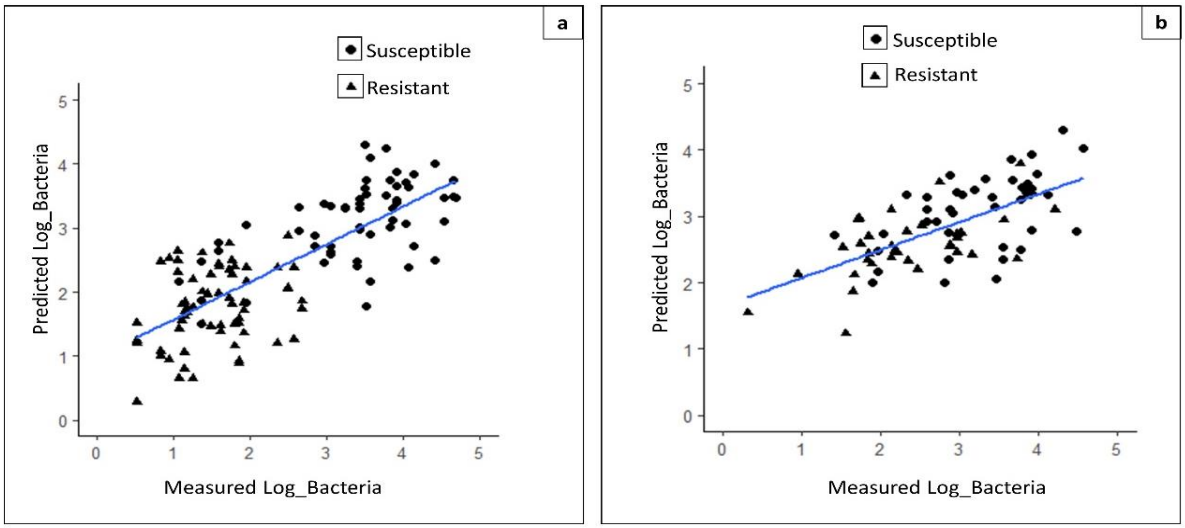

Figure 5. Partial least squares regression (PLSR) correlation plot for the (a) first and (b) second trials, showing the relationship between measured *Lefisonia xyli* subsp.*xyle* titer determined by qPCR at 85 DIA and the predicted bacteria titer on non-inoculated sugarcane plants based on FT-IR spectra.

2.3.3.2. Model validation SVM and PLSR

Even though only six and three spectral bands were significantly different between resistant and susceptible genotypes, for validate both models (SVM and PLSR) was necessary to include the 120 and 62 regions of the spectral for the first and second trial, respectively. The SVM accuracy for the first and second trials was 80 and 83% respectively (Table 1).

A 5-factor (spectral data) PLSR model explained the maximum variation in the concentration of bacterial cells, with a determination coefficient (R^2) of 0.77, whereas for the second trial, a 7-factor PLSR model accounted for 65% of the variation in bacterial titers. The root mean square errors (RMSEP) of the dataset was 1.064 and 1.017% for the first and second trials, respectively (Table 2).

Table 1. Support vector machine (SVM) classification results of discriminating sugarcane resistant genotypes and susceptible based on infrared spectrum

Data set	First trial				Second trial			
	Number of sample total	Number of correct classified samples	Number of misclassified samples	Accuracy (%)	Number of sample total	Number of correct classified samples	Number of misclassified samples	Accuracy (%)
Training set	123	123	0	100	148	148	0	100
Testing set	32	26	6	80	38	32	6	83

Table 2. Partial least square (PLSR) regression results for *Leifsonia xyli* subsp *xyli* titer in sugarcane plant

Data set	First trial				Second trial			
	Number of sample total	Factors	r value	RMSEP	Number of sample total	Factors	r value	RMSEP
Training set	123	5	0.77	1.064	148	7	0.65	1.017
Testing set	32	5	0.77	1.064	38	7	0.65	1.017

2.4. DISCUSSION

Several reports demonstrated a positive correlation between yield losses and Lxx titers in sugarcane (Grisham 1991; Urashima et al. 2017). Thus, breeding for resistance to ratoon stunt could be achieved by selecting genotypes that restrict the multiplication of the bacteria and our results confirmed that sugarcane genotypes differ with respect to this characteristic. However, breeding for this trait would be expensive and time-consuming if bacterial titers were determined by PCR-based techniques. Thus, the perspectives of using spectroscopic methods for this purpose are promising.

In our study, plants of the first trial were assumed to be free from infection whereas infected plants were detected in the second. A possible explanation would be that the plant matrices used as sources or setts for both trials were free from infection at the beginning but became infected after their prolonged cultivation in the greenhouse for 9 months. However, the mode of infection is unknown, since Lxx has no known insect vector and although it can be transmitted through wounding by cutting knives or other pruning tools, this was likely not the case since the plants were not subjected to any management practice that resulted in injuries. An alternative explanation is that the matrices were infected at the beginning with low bacterial titers that were not detected by conventional PCR and that bacterial concentration increased over 9 months to the point of detection in plants of the second trial. In any case, the classification of the genotypes as resistant or susceptible agreed in both trials regardless of the plants being infected before inoculation or not.

Besides being able to distinguish the genotypes in relation to the degree of resistance to bacterial growth, predictive models were developed by SVM and PLSR analyses for both trials. In the literature, SVM has been used to distinguish resistant from susceptible genotypes (Nagasubramanian et al. 2018; González-Camacho et al. 2018). For example, Kuska et al. (2018) used metabolic and phenomic profiles and SVM to identify resistant barley plants against powdery mildew. In a study on barley grain using near-infrared spectroscopy, Han et al. (2017) developed a predictive model for measuring phenolic compounds. The authors showed the efficiency of PLSR in determining important compounds for use in screening samples as part of a breeding process. However, few studies have used metabolic profiling and chemometric analyses for sugarcane genotype selection. These studies focused on metabolic profiling techniques such as high-performance liquid chromatography (HPLC) which is costly and requires expensive organic solvents and complex instruments. For instance, Leme et al. (2014) developed an HPLC-based protocol for phenotyping sugarcane genotypes. Later Coutinho et al. (2016) provided a

phytochemical catalogue of thirteen sugarcane genotypes by 1H nuclear magnetic resonance (NMR) and liquid chromatography coupled to a diode array detector (LC-DAD) methods. Even though these techniques are appropriate for metabolic analyses, in this study we proposed a faster and less expensive method for examining a broad spectrum of plant-derived chemicals. Regardless of whether the plants were not infected (first trial) or infected (second trial) with Lxx prior to inoculation, our results demonstrated that simple models with few wavenumbers distinguished the resistant from the susceptible genotypes. However, the fact that two different SVM and PLSR models were fitted to the data of the two trials indicates that environmental conditions and perhaps the prior colonization of the tissues by the pathogen in the second trial influence the spectra. In our study even though only six (first trial) and three (second trial) regions of the spectra differred between resistant and susceptible genotypes, it was necessary to use all the spectra (120 first trial and 62 second trial) regions for validating both models (SVM and PLSR). These results imply that there are more regions of the spectra responsible for the genotypes separation and other statistical analyses would be necessary to assess these regions.

In the first trial, we observed high absorbance of carbohydrates in susceptible genotypes. The Lxx genome has several genes involved in sugar intake and the bacterium is capable of utilizing a range of carbohydrates. These compounds play an important role in Lxx growth since it is an essential compound for its *in vitro* cultivation (Monteiro-Vitorello et al. 2004). In agreement, Cia et al. (2018) studied changes at the transcriptome and the proteome levels in the early stages of Lxx development in young plants and reported that infection results in the up-regulation of a SWEET14 gene related to the transport of sugar. Thus, as Lxx colonizes the relatively poor carbon environment of the xylem, carbohydrate allocation through the up-regulation of SWEET transporters would ensure its growth in host tissue. Additionally, since carbohydrates are essential substrates for fundamental processes of plant development, competition for sugar acquisition could explain the reduction in the rate of photosynthesis and biomass accumulation in infected plants (Zhang et al. 2017; Veillet et al. 2017).

Another intriguing finding relates to peaks previously correlated to alpha helix amide I protein and polyamines in studies related to plant cell wall compounds since the same wavenumbers were observed in our study (Sene et al. 1994; Kärkönen and Kuchitsu 2014; Romero et al. 2018). In the first trial, susceptible genotypes showed higher peaks for alpha helix amide I protein and lower for four different bands correlated to polyamines. Thus, we hypothesize that upon infection by Lxx resistant genotypes activate the defense mechanisms faster by regulating the catabolism and biosynthesis of free polyamines. The most common amide proteins related to plant disease responses is hydroxycinnamic acid amide (HCAA) which

is formed by the condensation of phenolic acids with polyamines (Zhou 2018). HCAA has been reported to contribute to resistance against pathogens when accumulated in the plant cell wall or hydrolyzed to release polyamines and phenolic compounds. As such, homeostasis of polyamines inside the cell is strictly regulated by HCAA and can be deregulated in response to some environment stress (e.g. drought or pathogen elicitation) or during the transition to a different developmental stage (e.g. flowering) (Martin-Tanguy 1997; Liang et al. 2017). It is known that polyamines play an intense role in regulatory and cellular processes, such as plant defense and abiotic stress tolerance (Takahashi 2016). In sugarcane, reports support a clear correlation between pathogen infection and polyamine accumulation (Marıía Estrella Legaz, Roberto de Armas 1998; Huang et al. 2018). At the transcriptome and proteome levels, Cia et al. (2018) suggested that Lxx modulates the regulation of genes involved in the synthesis and regeneration of a precursor of spermidine, an essential polyamine for prokaryotic cell growth and protection. As in our study, we could not identify which polyamine is accumulated in the plant tissue it would be interesting to develop studies focusing on the role of these compounds on the infection of Lxx in sugarcane.

In the second trial, resistant genotypes differed from the susceptible ones in wavenumbers related to benzene ring groups. Among the compounds of this functional group, phenolics seem to play an important role in resistance to lxx. Previous studies develop by our group on metabolic responses to RSD demonstrated that a resistant genotype had higher levels of phenolic compounds such as chlorogenic, hydroxybenzoic, cinnamic acid and luteolin (Moretti 2017). Another differentially represented peak identified in this trial was detected at 1508 cm^{-1}, which was previously associated to lignin (Stuart 2004). Even though we do not know exactly which compounds in our extracts are related to the spectral regions, we have the same absorbance in the same peak that is attributed to this compound. (Stuart 2004). Higher contents of this compound was observed in resistant genotypes when compared to susceptible ones. This is an interesting finding as well as resistance to Lxx is correlated to anatomical features of the cell wall, notably its rigidity (Teakle, D. S., Appleton, J. M., and Steindl 1978).

In order to complement our study we developed a prediction model using PLSR to estimate Lxx concentration for the first and second trials. Due to the complexity of our data, linear models such as PLSR could not provide optimal models. However, our results using this algorithm clearly demonstrated that there is a strong correlation between predicted (based on FT-IR spectra) and measured bacterial titer levels.

2.5. CONCLUSION

This work highlighted the potential of FT-IR spectroscopy to predict resistant genotypes to Lxx based on bacterial titers using two different algorithms. Our results suggest that resistance may be associated with variation in chemistry and would be probably associated with variations in the content of polyamines, carbohydrates and compounds with aromatic ring as a chemical functional group. Future studies should evaluate the performance of more high-throughput methods based on portable and handheld spectrometer devices to speed up the chemotyping of plants in a non-destructive way that could be directly used in the field.

References

Aidan F., N. K. 2006. Learning to classify documents according to genre. Aidan F., Nicholas K. 57:1506–1518.

Bailey, R. A. 1997. The systemic distribution and relative occurrence of bacteria in sugarcane varieties affected by ratoon stunting disease. Proceeding South African Sugarcane Technol. 704:55–56.

Barker, M., and Rayens, W. 2003. Partial least squares for discrimination. J. Chemom. 17:73–166.

Behmann, J., Mahlein, A., Rumpf, T., Ro, C., and Plu, L. 2015. A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. :239–260.

Carvalho, G., da Silva, T. G. E. R., Munhoz, A. T., Monteiro-Vitorello, C. B., Azevedo, R. A., Melotto, M., et al. 2016. Development of a qPCR for Leifsonia xyli subsp. xyli and quantification of the effects of heat treatment of sugarcane cuttings on Lxx. Crop Prot.

Cia, M. C., De Carvalho, G., Azevedo, R. A., Monteiro-Vitorello, C. B., Souza, G. M., Nishiyama-Junior, M. Y., et al. 2018. Novel insights into the early stages of ratoon stunting disease of sugarcane inferred from transcript and protein analysis. Phytopathology. :1–58.

Clemmensen, L., Hastie, T., Witten, D., and Ersbøll, B. 2011. Sparse discriminant analysis. Technometrics. 53:4.

Conrad, A. O., and Bonello, P. 2016. Application of Infrared and Raman Spectroscopy for the Identification of Disease Resistant Trees. Front. Plant Sci. 6:1–8 Available at: http://journal.frontiersin.org/Article/10.3389/fpls.2015.01152/abstract.

Conrad, A. O., Rodriguez-Saona, L. E., McPherson, B. A., Wood, D. L., and Bonello, P. 2014. Identification of Quercus agrifolia (coast live oak) resistant to the invasive pathogen Phytophthora ramorum in native stands using Fourier-transform infrared (FT-IR) spectroscopy. Front. Plant Sci. 5:1849.

Coutinho, I. D., Baker, J. M., Ward, J. L., Beale, H., Creste, S., and Cavalheiro, A. J. 2016. Metabolite profiling of sugarcane genotypes and identification of flavonoid glycosides and phenolic acids. J. Agric. Food Chem. 64:4198–4206.

Damann, A. J. K. E., and Benda, G. T. 1983. Evaluation of commercial heat treatment methods for control of ratoon stunting disease of sugarcane. Plant Dis. 67:966–967.

Davis, M. J., Dean, J. L., and Harrison, N. A. 1988. Quantitative variability of Clavibacter xyli subsp. xyli populations in sugarcane cultivars differing in resistance to ratoon stunting disease. Phytopathology. 78:462–468.

Devitt, G., Howard, K., Mudher, A., and Mahajan, S. 2018. Raman Spectroscopy: An Emerging Tool in Neurodegenerative Disease Research and Diagnosis. ACS Chem. Neurosci. 9:404–420.

Diem, M. 1993. *Introduction to Modern Vibrational Spectroscopy*.

DS Teakle, P. S. and D. S. 1973. Association of a small coryneform bacterium with the ratoon stunting disease of sugar-cane. Aust. J. Agric. Res. 24:869–874.

Ehsani, M. R. et al. 1999. A NIR technique for rapid determination of soil mineral nitrogen. Precis. Agric. 1:217–234.

Evtushenko, L. I., Dorofeeva, L. V., Subbotin, S. A., and Cole, J. 2000. Leifsonia poae gen. nov., sp. nov., isolated from nematode galls on Poa annua, and reclassification of 'Corynebacterium aquaticum' Leifson 1962 as Leifsonia aquatic (ex Leifson 1962) gen. nov., nom. rev., comb.nov. and Clavibacter xyli Davis et al. 1984 w. Int. J. Syst. Evol. Microbiol. 50:371–380.

Fegan, M., Croft, B. J., Teakle, D. S., Hayward, A. C., and Smith, G. R. 1998. Sensitive and specific detection of Clavibacter xyli subsp. xyli, causal agent of ratoon stunting disease of sugarcane, with a polymerase chain reaction-based assay. Plant Pathol. 47:495–504 Available at: http://doi.wiley.com/10.1046/j.1365-3059.1998.00255.x.

Fiehn, O. 2001. Combining genomics, metabolome analysis, and biochemi- cal modelling to understand metabolic networks. Comp. Funct. Genomics. 2:155–168.

González-Camacho, J. M., Ornella, L., Pérez-Rodríguez, P., Gianola, D., Dreisigacker, S., and Crossa, J. 2018. Applications of Machine Learning Methods to Genomic Selection in Breeding Wheat for Rust Resistance. Plant Genome. 11 Available at: https://dl.sciencesocieties.org/publications/tpg/abstracts/11/2/170104.

Hoy, J. . W., Bischoff, K. P., Milligan, S. B., and Gravois, K. 2003. Effect of tissue culture explant source on sugarcane yield components. Euphytica. 129:237–240.

Huang, N., Ling, H., Su, Y., Liu, F., Xu, L., Su, W., et al. 2018. Transcriptional analysis identifies major pathways as response components to Sporisorium scitamineum stress in sugarcane. Gene. 678:207–218 Available at: https://doi.org/10.1016/j.gene.2018.08.043.

Hussein, H. M., Hameed, R. H., and Hameed, I. H. 2018. Screening of Bioactive Compounds of Ricinus communis Using GC-MS and FTIR and Evaluation of its Antibacterial and Antifungal Activity. Indian J. Public Heal. Res. Dev. 9:463–469.

Kucheryavskiy, S. 2018. Multivariate Data Analysis for Chemometrics. R Packag. Version 0.9.1.

Leme, G. M., Coutinho, I. D., Creste, S., Hojo, O., Carneiro, R. L., Bolzani, V. D. S., et al. 2014. HPLC-DAD method for metabolic fingerprinting of the phenotyping of sugarcane genotypes. Anal. Methods. 6:7781–7788.

Marıía Estrella Legaz, Roberto de Armas, D. P. and C. V. 1998. Relationships between phenolics-conjugated polyamines and sensitivity of sugarcane to smut (Ustilago scitaminea). J. Exp. Bot. 49:1723–1728.

Martínez-Martínez, V., Gomez-Gil, J., Machado, M. L., and Pinto, F. A. C. 2018. Leaf and canopy reflectance spectrometry applied to the estimation of angular leaf spot disease severity of common bean crops ed. Dawn L Arnold. PLoS One. 13:e0196072 Available at: https://dx.plos.org/10.1371/journal.pone.0196072.

Mathlouthi, M. K. and M. 1996. FTIR and laser-Raman spectra of oligosaccharides in water: characterization of the glycosidic bond. Carbohydr. Res. 284:145–157.

Meyer, L., Dimitriadou, E., Hornik, K., Weingessel, A., and Leisch, F. 2015. e1071: misc functions of the Department of Statistics, probability theory group (formerly: E1071), TU Wien. R package version 1.6-7.

Monteiro-Vitorello, C. B., Camargo, L. E. A., Sluys, M. A. Van, Kitajima, J. P., Truffi, D., Do Amaral, A. M., et al. 2004. The Genome Sequence of the Gram-Positive Sugarcane Pathogen Leifsonia xyli subsp. xyli. Mol. Plant. Microbe. Interact. 17:827–836.

Moretti, F. R. R. de C. 2017. Identification of candidate resistance metabolites to Leifsonia xyli subsp. xyli in sugarcane through metabolomic profiling.

Nagasubramanian, K., Jones, S., Sarkar, S., Singh, A. K., Singh, A., and Ganapathysubramanian, B. 2018. Hyperspectral band selection using genetic algorithm and support vector machines for early identification of charcoal rot disease in soybean stems. Plant Methods. 14:2–13.

Oh, M., Barnaby, J. Y., Lee, H., Jia, Y., and Kim, M. S. 2018. Infrared fingerprint profiles of compatible and incompatible interactions of the rice blast pathogen using FTIR spectroscopy. In *Sensing for Agriculture and Food Quality and Safety*,.

Pan, Y.-B., Grisham, M. P., Burner, D. M., Damann, K. E., and Wei, Q. 1998. A Polymerase Chain Reaction Protocol for the Detection of Clavibacter xyli subsp. xyli, the Causal Bacterium of Sugarcane Ratoon Stunting Disease. Plant Dis. 82:285–290 Available at: https://apsjournals.apsnet.org/doi/pdf/10.1094/PDIS.1998.82.3.285 [Accessed April 26, 2018].

Pospiech, A., and Neumann, B. 1995. A versatile quick-prep of genomic DNA from gram-positive bacteria. Trends Genet. 11:217–218.

Purcell, D. E., Oxley, J. P., Cox, M. C., Croft, B. J., and O'Shea, M. G. 2010. On-site rapid screening for sugarcane smut resistance using near-infrared (NIR) spectroscopy. In *Australian Society of Sugar Cane Technologists*, , p. 366–374.

Quecine, M. C. et al. 2015. A stable Leifsonia xyli subsp. xyli GFP-tagged strain reveals a new colonization niche in sugarcane. Plant Pathol. 65:154–162.

Quecine, M. C., Silva, T. M., Carvalho, G., Saito, S., Mondin, M., Teixeira-silva, N. S., et al. 2016. A stable Leifsonia xyli subsp . xyli GFP-tagged strain reveals a new colonization niche in sugarcane tissues. Plant Pathol. 65:154–162.

Saitta, L. 1995. Support-Vector Networks. In *Machine Leaming*, , p. 273–297.

Savitzky, A., and Golay, M. J. E. 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36:1627–1639.

Shi, Y., Huang, W., Ye, H., Ruan, C., Xing, N., Geng, Y., et al. 2018. Partial Least Square Discriminant Analysis Based on Normalized Two-Stage Vegetation Indices for Mapping Damage from Rice Diseases Using PlanetScope Datasets. Sensors. 18:1901 Available at: http://www.mdpi.com/1424-8220/18/6/1901.

Silva, L. J. 2015. Laercio: Duncan test, Tukey test and Scott-Knott test.

Thumanu, K., Wongchalee, D., Sompong, M., Phansak, P., Le Thanh, T., Namanusart, W., et al. 2017. Synchrotron-based FTIR microspectroscopy of chili resistance induced by *Bacillus subtilis* strain D604 against anthracnose disease. J. Plant Interact. 12:255–263 Available at: https://www.tandfonline.com/doi/full/10.1080/17429145.2017.1325523.

Urashima, A. S., Silva, M. F., Correa, J. J., Moraes, M. C., and Singh, A. V. 2017. Prevalence and Severity of Ratoon Stunt in Commercial Brazilian Sugarcane Fields. Plant Dis. 101:815–821 Available at: https://apsjournals.apsnet.org/doi/pdfplus/10.1094/PDIS-07-16-1030-RE [Accessed March 6, 2018].

Willick, I. R., Lahlali, R., Vijayan, P., Muir, D., and Karunakaran, C. 2018. Wheat flag leaf epicuticular wax morphology and composition in response to moderate drought stress are revealed by SEM , FTIR-ATR and synchrotron X-ray spectroscopy. :316–332.

Wold, S., Ruhe, A., Wold, H., and Dunn, III, W. J. 1984. The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses. SIAM J. Sci. Stat. Comput. 5:735–743

Young, A. J. 2016. Possible origin of ratoon stunting disease following interspecific hybridization of *Saccharum* species. Plant Pathol.

Zhan-yu, L., Jing-feng, H., Jing-jing, S., Rong-xiang, T., Wan, Z., and Li-li, Z. 2007. Characterizing and estimating rice brown spot disease severity using stepwise regression, principal component regression and partial least-square regression. J. zhejiang Univ. Sci. b. 8:738–744.