FINAL REPORT 2014/082

A novel polyphasic framework to resolve Yellow Canopy Syndrome paradox.

Final report prepared by:	Brajesh Singh, Kelly Hamonts, Pankaj Trivedi, Galaxy Qui, Jasmine Grinyer
Chief Investigator(s):	Brajesh Singh
Research organisation(s):	Western Sydney University
Co-funder(s):	QLD Government Department of Agriculture and Fisheries
Date:	1st August 2018
Key Focus Area (KFA):	Pest, disease and weed management - Yellow Canopy Syndrome Impact Area

Sugar Research Australia

Queensland Government

WESTERN SYDNEY UNIVERSITY

Hawkesbury Institute for the Environment

Please cite as: Singh, BK, Hamonts, K, Trivedi, P, Qui, G, Grinyer, J. 2018. A novel polyphasic framework to resolve Yellow Canopy Syndrome paradox: Final Report 2014/082. Sugar Research Australia Limited, Brisbane.

ABSTRACT

Yellow Cane Syndrome (YCS) is a largely undiagnosed condition impacting sugarcane crops across Queensland, Australia, causing significant yield losses. Since the first observation of YCS in 2011, the condition has spread to all major cane-growing areas and is a critical issue for the industry. However, the cause of YCS remains unknown.

In this study, we collected asymptomatic and YCS-symptomatic samples from North Queensland regions (Tully, Herbert, Burdekin and Mackay) at multiple times. We used Next-generation sequencing including amplicon, metagenomics and metatranscriptomic sequencing along with several soil health measures to identify the cause of YCS. Additionally, we examined differences between microbial communities under geographical, physical, chemical and biological effects and identified the core and hub microbiota of sugarcane.

In the sequencing analyses, strong variety by YCS interaction effects on both the bacterial and fungal community compositions were observed, demonstrating variety-specific changes in the relative abundance of bacterial and fungal species between asymptomatic and symptomatic plant tissues. We also detected significant changes in certain microbial functional gene categories in the rhizosphere of YCS symptomatic vs. asymptomatic plants. Several potential pathogenic bacterial and fungal genera contributed to the differences observed between symptomatic and asymptomatic plants, as well as commensal and beneficial organisms. Further detailed analysis identified a limited number of bacterial and fungal operational taxonomic unit (OTUs) that were either significantly more or less abundant in YCS symptomatic plants and in the rhizosphere soil for each of the studied locations and cane varieties.

We identified the core and hub microbiota of sugarcane which can be harnessed for future microbial assisted plant breeding, an emerging discipline for sustainable farming. Our analyses suggested that in the core sugarcane microbiome shared by all studied varieties and growing regions, only few consistent bacterial and fungal signals associated with YCS emerged, including *Curtobacterium* in leaves and *Cladosporium* in leaves and stalks. Several of these signals have been reported in literature as potential pathogen facilitators or antagonists (e.g. *Cladosporium* in plant leaves).

Network analysis of the microbiome data showed that certain bacteria and fungi respond to elevated leaf sucrose content associated with YCS in the leaf lamina, midribs and in particular, the dewlaps. These bacteria and fungi, in turn, affected the relative abundance of numerous other bacteria and fungi through complex interactions, resulting in cascade effects of YCS through the entire sugarcane leaf microbiome.

Deep sequencing of small RNA from sugarcane leaves revealed an active antiviral response to a badnavirus with a genome similar, but not identical, to known Sugarcane bacilliform viruses and Banana streak CA virus. These viral signals were found in all leaf compartments (lamina, midribs and dewlaps) and in both asymptomatic and symptomatic plants of a Q240 block in Home Hill affected by YCS. Moreover, these viral signals were detected in leaf samples from Ayr showing yellow midribs rather than YCS symptoms, and were not detected in all YCS symptomatic leaf samples collected previously from different growing regions.

Both metagenomics sequencing and targeted PCR data suggested the presence of phytoplasma in some YCS sugarcane plants but results were inconsistent across all samples. Based on the above findings, we postulate that fungal and viral infection, and soil nutritional health are unlikely to cause YCS. Our data suggest some links between bacterial (e.g *Curtobacterium* spp) and phytoplasma presence and YCS, which needs further investigation to confirm whether either of these communities is directly involved in YCS incidence.

EXECUTIVE SUMMARY

Yellow Canopy Syndrome (YCS) is a critical issue for the sugarcane industry. Since the first observation of YCS in 2011 in the Central and Northern cane-growing regions in Australia, it has spread to all major cane-growing areas. YCS is still an undiagnosed condition and is not correlated to mineral nutrient deficiency or any of the well-known sugarcane pathogens. A 50% productivity loss has occurred when the severity of YCS expression is high and occurs several times during the growing season.

In this project, we have collected samples of asymptomatic and YCS-affected sugarcane as well as soil samples from across sugarcane farming regions in Queensland in 2014, 2015, and 2016. DNA samples were analysed by next generation sequencing including amplicon, shot-gun metagenomics and sRNA metatranscriptomics sequencing. We also analysed physical, chemical and biological factors of soil samples and detected potential pathogens in YCS samples.

We have identified the first core and hub microbiota of sugarcane from field samples at a national level and have found variety-specific changes in the relative abundance of bacterial and fungal species between asymptomatic and symptomatic plant tissues. A significant shift in certain microbial functional gene categories were observed in the rhizosphere of YCS symptomatic vs. asymptomatic plants. Several potential pathogenic bacterial and fungal genera contributed to the differences observed between symptomatic and asymptomatic plants, including *Curtobacterium* in leaves and *Cladosporium* in leaves and stalks. Shotgun sequencing of rhizosphere samples confirmed these findings but also identified the presence of a Phytoplasma signal. Targeted PCR-based assays suggested the presence of Phytoplasma in some symptomatic samples.

Soil health measurements performed on ~150 soil samples collected in Ingham in November 2014 revealed no significant differences in soil carbon and nitrogen content, soil microbial enzyme activity and substrate-induced soil respiration between soils from asymptomatic vs. YCS symptomatic blocks with the same soil type, sugarcane variety and similar management practices. Our analysis revealed that soil health parameters did not correlate with YCS incidence.

Deep sequencing of sRNA from sugarcane leaves resulted in the detection of an active sugarcane defence response against a pararetrovirus with a genome that is most similar to the badnaviruses Banana streak CA virus and Sugarcane bacilliform virus. However, this virus was present in both YCS symptomatic and asymptomatic plants.

Overall, our results suggest that fungal and viral infection, and soil health properties are unlikely to be the direct cause of YCS. However, some links between bacterial (e.g *Curtobacterium*) or Phytoplasma with symptomatic plants warrants further investigation to determine if either of these organisms can contribute to YCS in sugarcane.

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF FIGURES

1. BACKGROUND

Yellow Canopy Syndrome (YCS) was first observed in 2011 and is still an undiagnosed condition affecting the Central and Northern cane-growing regions in Australia and causing significant yield losses (Marquardt et al. 2016). YCS has affected different varieties of sugarcane and exhibited multiple symptoms like root rot, yellowing and mottling of leaves with yield reductions between 30-40% in severely affected blocks (SRA industry update Oct 21, 2013). There remains no conclusive evidence that YCS is caused by a biological agent and/or physiological and environmental stresses. The lack of knowledge on the cause of this condition hinders the development of specific diagnostic procedures and impedes efficient YCS control and management.

Initial analyses of YCS failed to find any correlation between the development of YCS and nitrogen, iron and other mineral nutrients, as well as known sugarcane pathogens. In addition, the development of YCS was distinctly different from that of Sugarcane Yellow Leaf Virus (SCYLV) which was detected initially in Hawaii and Brazil (Comstock et al. 1994; Goncalves et al. 2005) in the 1990's and is now recognised worldwide. The pathogen responsible, the sugarcane yellow leaf virus (ScYLV) was identified as a luteovirus (Moonan et al. 2000; Smith et al. 2000) and YCS affected sugarcane plants have consistently tested negative for the presence of ScYLV.

In recent breakthroughs, profiling the host's "microbiome" (set of genes encoded by host-associated microbes) by culture-independent methods (which requires no prior knowledge of host or pathogen)has played an important role in unravelling the causal agent(s) of certain mysterious diseases, for example Huanglongbing (Sagaram et al. 2008; Trivedi et al. 2010; Wang and Trivedi 2013) and other diseases which were considered to result from non-biological conditions, including citrus blight disease (Timmer et al. 1992). A comprehensive understanding of the sugarcane microbiome using next-generation sequencing (NGS) was therefore necessary to determine the role of biological agents in YCS development.

Among various factors, soil health and plant physiological changes were also postulated to contribute to YCS development (SRA industry update March 5, 2014). Identification of soil health conditions that facilitates or delays YCS through alterations to the microbial community structure is necessary to resolve the mystery behind YCS. A polyphasic approach to understand the complexity of the microbial community was undertaken to provide information on the cause/effect relationships between YCS and soil health/host response. This knowledge could further pave the way to exploit beneficial plant-microbe interactions to increase farm health and productivity through the rational utilization of probiotic microbes.

In this study, we aimed to use next-generation sequencing tools to (1) analyse microbial communities associated with asymptomatic and YCS symptomatic sugarcane to identify the causal agents of YCS, (2) compare two phenotypes on a small scale, to evaluate the effect of interactions between resident microbiome(s) and YCS on soil health, and manipulate this information for better YCS management. Further, we aimed to determine the impact of YCS on plant defence and stress responses to better understand the involvement of physiological disturbances in YCS development, and develop bio-resources for the management of YCS into the future. Our two working hypotheses were:

- **Hypothesis 1:**
  Unknown biological agents (bacteria, virus or phytoplasma) are responsible for YCS development either directly via pathogenicity or indirectly via inducing plant physiological stresses;

- **Hypothesis 2:**
  Soil health including rhizosphere functions are disrupted, leading to compromised nutrient uptake efficiency of YCS symptomatic sugarcane.

2. PROJECT OBJECTIVES

Yellow cane syndrome (YCS) is an enigmatic sugarcane condition recently observed in major sugarcane producing regions of Queensland (SRA industry update March 5, 2014). There is no conclusive evidence that YCS is caused by a biological agent and/or physiological and environmental stress. The lack of knowledge on the cause of this condition hinders development of specific diagnostic procedures and impedes efficient YCS control and management.

The overarching aim of this project was to apply a comprehensive 'microbiome' based approach, combining novel next-generation sequencing (Illumina) with conventional microbialculturing techniques to provide a broader view of the complexity of the micro-organisms present in YCS-affected sugarcane and determine the involvement of biological agent(s) in YCS development. Furthermore, we aimed to generate key knowledge on the impact of YCS on soil health and the plant hosts' response for a better understanding of the YCS facilitation process. The project worked closely with the two other SRA-funded projects on YCS.

The objectives of the project, prior to the project beginning were:

- Detailed metagenomic profiling of YCS expressing and healthy tissue of sugarcane varieties Q208 and Q200 in the Burdekin and Herbert regions.
- Identification of the changes in microbial structure and functions in the transition from 'healthy' to 'symptomatic' to reveal the involvement of biological agents in YCS development.
- Confirmation of potential causative agent(s) of YCS. If the metagenomic approach points towards the involvement of biological agent(s) then a combination of various approaches including qPCR, fluorescent in-situ hybridization will be employed to determine their role in YCS development. This will also include culturing and proving Koch postulates.
- Evaluation of the interactions between resident microbiome(s) and YCS on soil health. This analysis will reveal whether changes in microbial-mediated functions related to soil health cause poor root symptoms related to YCS.
- Identify the metabolic processes and host response related to stress and pathogenesis that are primarily influenced when the tissue goes from healthy to YCS symptomatic.
- Identify beneficial sugarcane associated microbes that promote plant growth and soil health thereby minimising the impact of YCS (on the advice of the SRP, this objective later became irrelevant)
- Identification of biological, biochemical and/or molecular indicators that can be targeted for early identification and suggestions for possible remedial actions.

3. OUTPUTS, OUTCOMES AND IMPLICATIONS

3.1. Outputs

The proposed research used an advanced 'microbiome approach' for generating key knowledge on plant-soil-microbe interactions to comprehensively determine the role of the sugarcane-associated microbiome, soil health and host responses in YCS development. The information can be used in multiple ways to provide informed management decisions regarding improved farming, YCS development and enhance the economic and sustainable production of the sugarcane industry.

The outputs of the projects were:

1. Identification of both the core and hub microbiota of sugarcane varieties in Queensland, Australia. This is the first time globally that the sugarcane microbiomes have been identified and could be harnessed for microbial-optimised plant breeding for sustainable and profitable farming in the future. The work is peer reviewed and published in the top ranked journal (Environmental Microbiology) in the discipline.

2. We found no relationship between soil nutritional health and YCS incidence. However, the functional potential of the rhizosphere (area surrounding the root) was compromised. This provides a tool to explore mechanisms to remove this constraint from the rhizosphere to minimise the impact of, or onset of YCS incidence. The work was peer reviewed and published in a top ranked journal (Soil Biology and Biochemistry).

3. Based on our data, we suggest that fungal and viral infections, or soil health is unlikely to be a direct cause of YCS but further research is needed to confirm / deny the role of bacteria and phytoplasma.

4. Results of this research were communicated to industrial stakeholders by multiple conference and symposia presentations, industrial presentations and to industry using other approaches.

3.2. Outcomes and Implications

The outcomes of this research provided the necessary building blocks for a resilient sugarcane industry, by providing knowledge on the core and hub microbiota which can be harnessed for multiple purposes (e.g. breeding for better performing varieties, harnessing the microbiome for sustainable and profitable farming). Our results suggest that fungal and viral infection, and soil health properties are unlikely to be the direct cause of YCS. Additionally, we highlight the need to prioritise research on the potential role of bacteria and phytoplasma and their potential role in the development of YCS. This can provide research direction to industry regarding YCS incidence for which management is important for the sustainability of the sugar industry.

4. INDUSTRY COMMUNICATION AND ENGAGEMENT

4.1. Industry engagement during course of project

Communication activities throughout the project:

- Panel members, SRA personnel and collaborators at the Yellow Canopy Syndrome Program Meeting, Lucinda, 30 June 2014.
- Dr Harjeet Khanna, Program Manager during her visit to Hawkesbury Institute of Environment and Research, UWS in October 2014.
- Sugarcane Industry through Industry Updates in YCS Handout, September 2014.
- Panel members, SRA personnel and collaborators at the Yellow Canopy Syndrome Program – Scientific Reference Panel Meeting, Townsville, 20 November 2014.
- Sugarcane Industry, through industry update meetings from 9-11 December 2014 in Mossman, Mulgrave, Ingham, Burdekin, Proserpine and Mackay.
- Prof. John Lovett, chair of the scientific reference panel, during his visit to the Hawkesbury Institute for the Environment, UWS on 12 February 2015.
- Sugar Research Australia YCS industry update, April 2015.
- Results obtained were communicated to panel members, SRA personnel and PEC officers at the Yellow Canopy Syndrome Integrated Research Program Scientific Reference Panel meeting in Brisbane, 5 June 2015.
- Preliminary results presented in Poster format at the Australian Society for Microbiology Annual Scientific Meeting, Canberra, 12-15 July 2015.
- Results communicated during Panel member visit to WSU in August 2015 with Harjeet Khanna (SRA Program Coordinator), Prof. Roger Hellens (YCS Scientific Reference Panel member), Belinda Bellings and Brad Pfeffer (SRA PEC).
- Results communicated at the YCS Integrated Research Program Expert Strategy Meeting, Brisbane, 1 September 2015.
- Sugar Research Australia YCS industry update, September 2015, featuring a section on the WSU work entitled: “Search for YCS answers continues”.
- Results communicated during a visit of Kelly Hamonts to Frikkie Botha’s team at SRA in Indooroopilly, Brisbane, 29-30 September 2015.
- Preliminary results presented in Poster format at the 9th Symposium for the International Society of Root Research, Canberra, 6-9 October 2015.
- Preliminary results communicated at the 1st International Symposium on Frontiers in Soil Microbiology, Beijing, China, 25-27 October 2015.
- Yellow Canopy Syndrome Integrated Research Program Scientific Reference Panel Meeting, Cairns, November 2015.

- During the YCS Integrated Research Program Expert Review Meeting in Cairns, 11-12 November 2015.
- Ongoing results presented at the YCS Integrated Research Program Scientific Reference Panel Meeting, Brisbane, 10 June 2016.
- Project results were presented at the 2016 American Phytopathology Society Annual Meeting in Tampa, Florida, USA (July 2016).
- Project results were presented at the 16th Symposium of the International Society for Microbial Ecology in Montreal, Canada (August 2016).
- Project progress presented at the YCS Integrated Research Program Scientific Reference Panel Meeting, Brisbane, 28-29 November 2016.
- Research presented in a poster format entitled "Unravelling the sugarcane microbiome to resolve the Yellow Canopy Syndrome in Australia" at the 2016 Theo Murphy Australian Frontiers of Science – Microbiome - Exploring the role of microorganisms in ecosystem processes and health symposium in Adelaide, 29 November-1 December 2016.
- Dr Hamonts presented research results at the first Australian Microbial Ecology conference in Melbourne, 13-15 February 2017. Dr Kelly Hamonts was awarded the best presentation by a mid-career scientist based on her quality of scientific research.
- 31 July – 1 August 2017: YCS Mid-year review. Brisbane.
- Research results presented at the SRA Panel meeting on the 12th December 2017 in Brisbane.
- Research results presented at the SRA Panel meeting: June 2018. Brisbane.
- Research publications in peer reviewed journals, to date:
    - Hamonts, K., Trivedi, P., Garg, A., Janitz, C, Grinyer, J., Holford, P., Botha, F.C., Anderson, I.C. and Singh, B.K. (2018), Field study reveals core plant microbiota and relative importance of their drivers. Environmental Microbiology, 20(1), 124-140
    - Hamonts, K., Trivedi, P., Grinyer, J., Holford, P., Drigo, B., Anderson, I.C. and Singh, B.K. (2018), Yellow Canopy Syndrome in sugarcane is associated with shifts in the rhizosphere soil metagenome but not with overall soil microbial function. Soil Biology and Biochemistry (accepted).

4.2. Industry communication messages

1. This project is the first to report the core and hub microbiota of sugarcane. Such knowledge has been identified as key information for sustainable and profitable farming as these are essential for developing better performing plant varieties for biotic and abiotic stresses. This information can also be exploited to develop emerging *in situ* microbiome engineering approaches for sustainable farming.
2. Based on our data, we suggest that fungal or viral infections and soil nutritional health are unlikely to cause YCS. We show some links between bacteria, phytoplasma and YCS incidence which need further investigation to confirm or deny their role.

5. METHODOLOGY

5.1. Sample collection

Asymptomatic and YCS symptomatic sugarcane samples were collected in North Queensland, Australia (Table 1). The first round of sampling was collected from Ayr and Ingham regions in coordination with Dr F Botha's team in May 2014. Due the lack of proper healthy controls from Ayr, additional YCS infected and healthy samples were requested in September 2014 and provided by Dr D Olsen.

In November 2014, we collected samples from asymptomatic vs. YCS-affected plants that were grown in the same soil type, under the same management regime and in close proximity from each other within the Ingham area. In addition, samples from healthy sugarcane were collected from particular blocks at the SRA station in Ayr, which are closely monitored by the SRA research team for the onset of YCS symptoms. The aim was to resample these blocks when YCS symptoms appeared to obtain a detailed picture of temporal changes in the sugarcane and soil microbiome associated with the onset of YCS symptoms. Additional symptomatic and asymptomatic leaves were kindly sent to us by John Agnew and colleagues from MAPS, Mackay in June 2015.

During sample collection in June-July 2015, we collected samples from 4 regions in North Queensland: Tully, Ingham (Herbert), Ayr (Burdekin) and Mackay. We collected samples in Tully because YCS has not yet been reported there. Samples were collected from the local seedbank, from a clean seed source that had received a hot water treatment. To date, these samples represent the cleanest YCS asymptomatic samples that we could collect in the field. In Ingham and Ayr, we resampled blocks sampled in November 2014. In Ayr, plants of all 6 blocks remained asymptomatic for YCS throughout the season. In Ingham, we resampled 6 of the 8 blocks sampled in November 2014. The blocks were affected by YCS multiple times throughout the season. At the time of sampling in June 2015, 2 of the 6 blocks showed similar YCS symptoms as in November 2014 (asymptomatic), whereas the remainder of the blocks switched from either YCS symptomatic in November 2014 to asymptomatic in June 2015, or vice versa. Comparison of the November 2014 and June 2015 Ingham data allowed temporal changes in the sugarcane microbiome to be studied, as well as shifts in the microbiome associated with YCS symptom switches and YCS severity based on the number of times a block was affected by YCS over the season.

Table 1: Sampling details in this study

Date of sampling	Location	Genotype	Material
May 2014 & Sep 2014	Ayr, Ingham	MQ239, Q240	leaf
Nov 2014 & June 2015	Ayr, Ingham, Tully, Mackay	KQ228, Q208, Q239, Q240	leaf 2, leaf 4, stalk, roots (+Rhz soil), bulk soil
June to July 2015	Ayr, Ingham, Tully, Mackay	KQ228, Q208, MQ239, Q240, Q242	leaf 2, leaf 1, stalk, roots, rhizophere soil, bulk soil
Feb 2016	Home Hill	Q240	leaf green, leaf yellow, midrib, dewlap, stalk, roots, bulk soil

All samples were snap-frozen in liquid nitrogen in the field, transported to the laboratory on dry ice and stored at -80 °C until analysis. Brief experimental procedures are available in Figure 1.

5.2. Sample processing

5.2.1.DNA and RNA extraction

For DNA extractions, sugarcane tissue samples were ground using mortars and pestles, and DNA was extracted from the resulting powder using the MoBIO PowerPlantVR Pro DNA isolation kit (Mo Bio Laboratories, Carlsbad, CA, USA), following the manufacturer's instructions. Each leaf and stalk sample comprised both epiphyte and endophyte DNA. Root samples were rinsed prior to grinding, to remove soil attached to the roots and comprised both rhizoplane (root surface) and endophyte DNA. The MoBIO PowerSoilVR DNA isolation kit (Mo Bio Laboratories, Carlsbad, CA, USA) was used to extract DNA from the soil samples, following the manufacturer's instructions.

For RNA extractions, three leaf samples were selected to test the small RNA sequencing approach: (1) a YCS symptomatic leaf sample collected in Mackay in June 2015, (2) a YCS symptomatic leaf sample collected in Ingham in June 2015, and (3) a leaf symptomatic for the Yellow Leaf Syndrome (yellow midrib) collected in Ayr in June 2015. The leaves were snap-frozen in liquid nitrogen immediately after sampling and ground to a fine powder using mortar and pestle (whilst keeping the samples frozen in liquid nitrogen). Total RNA was extracted from the resulting powder using Trizol (Invitrogen, CA, USA) following the manufacturer's instructions.

5.2.2.Sequencing

For amplicon sequencing, we targeted the V3-V4 region of the bacterial 16S rRNA gene (341F-805R, Herlemann et al., 2011) and the fungal ITS2 region (FITS7-ITS4R, Ihrmark et al., 2012) were sequenced at the Western Sydney University NGS facility (Sydney, Australia) using Illumina MiSeq 2x 300 bp paired end sequencing. For both 16S and ITS libraries, PNA clamps (PNA Bio Inc., Newbury Park, CA, USA) were added to the initial amplification PCR reaction at final concentrations of 3.8 mM each to minimise amplification of plant chloroplast and mitochondrial sequences (Lundberg et al., 2013).

For metagenomics sequencing, rhizosphere soil DNA sequencing libraries were prepared using the Illumina® Nextera DNA sample preparation kit, and the libraries were paired-end sequenced on three lanes of an Illumina HiSeq 2500 producing 125 bp read lengths. For RNA sequencing, RNA samples submitted to the NGS sequencing facility for small RNA sequencing using the Illumina MiSeq platform (1 x 50 bp single end reads).

5.2.3.Sequencing data processing

For amplicon sequencing, operational taxonomic units (OTUs) were built at 97% sequence similarity using UPARSE (Edgar, 2013) after joining the paired end reads using FLASH (Magoc and Salzberg, 2011) and several rounds of quality filtering. Singletons and chimeric sequences were discarded as recommend by Edgar et al. (2011). Taxonomy was assigned to OTUs in Mothur using the naïve Bayesian classifier (Wang et al., 2007) with a minimum bootstrap support of 60% and the Greengenes database version 13_8 (DeSantis et al., 2006; McDonald et al., 2012) for bacteria or the dynamic UNITE version 6 dataset (Koljalg et al., 2013) for fungi. OTUs representing sugarcane mitochondrial and chloroplast 16S sequences were removed, as well as OTUs matching other plant-derived sequences or Archaea. Good's coverage indices were calculated for the resulting OTU abundance tables, containing true bacterial and fungal OTUs, using mothur (Schloss et al.,

2009) and the tables were subsequently rarefied to an even number of sequences per sample, corresponding to the minimum number of sequences for a single sample obtained within a specific sample type or across all sample types. Alpha diversity metrics were calculated on the rarefied OTU abundance tables in mothur (Schloss et al., 2009).

For metagenomics sequencing, shotgun metagenomics was performed on 24 selected rhizosphere soil samples collected in Ingham in November 2014. The samples were prepared using Nextera NT and sequenced on 3 lanes of Illumina HiSeq 2500 v.4 using 2x 125 bp paired-end reads. Shotgun metagenomics produced between 14.6 and 23.8 million raw sequence reads per sample. These sequences were analysed through the MG-RAST metagenomics analysis server (https://metagenomics.anl.gov/), as well as through the HMP Unified Metabolic Analysis Network pipeline humann2 (http://huttenhower.sph.harvard.edu/humann2/manual) and a custom-built pipeline. Analysis of rhizosphere soil metagenomes was performed through the metagenomics RAST server using default parameters (Meyer et al., 2008). For each sample, the abundance of individual sequences matching SEED subsystem level 2 (groups of genes involved in a particular metabolic function; Overbeek et al., 2005) was obtained, normalised by sequencing effort and used to generate a functional profile of the rhizosphere soils.

For RNA sequencing, approximately 10,000,000 raw small RNA reads were obtained for each sample. After trimming the adapters and quality-filtering the reads, a *de novo* assembly was performed using the Trinity assembly pipeline. Assembled contigs were used to search the GenBank viral, reference RNA sequence, and Expressed Sequence Tags (EST) databases.

5.2.4. Phytoplasma Detection

The detection of Phytoplasma's from DNA extracted from sugarcane samples collected during the sampling campaign undertaken in June 2015 was prioritised after advice from the Panel in August 2017. Established detection methods were reviewed and nested PCR protocols optimised for 4 primer pairs reported to detect the DNA of Phytoplasma (Table 2).

Table 2: Primers used for Phytoplasma detection.

Primer Name	Primer Sequence 5'-3'	Reference
P1A	ACGCTGGCGGCGCGCCTAATAC	Lee *et al* 2004
P7A	CCTTCATCGGCTCTTAGTGC	
R16F2n	GAAACGACTGCTAAGACTGG	Gundersen and Lee, 1996
R16R2	TGACGGGCGGTGTGTACAAACCCCG	
P1	AAGAGTTTGATCCTGGCTCAGGATT	Smart *et al* 1996; Sharbatkhari *et al* 2008
P7	CGTCCTTCATCGGCTCTT	
fU5	CGGCAATGGAGGAAACT	Lorenz *et al* 1995; Sharbatkhari *et al* 2008
rU3	TTCAGCTACTCTTTGTAACA	

The initial PCR cycle was performed for 35 cycles using primers P1A / P7A prior to nested PCR using primers fU5 / rU3. Electrophoresis was conducted after PCR to visualise DNA products and Sanger sequencing of PCR products was undertaken to check for the presence of phytoplasma.

5.3. Data analyses

5.3.1. Extracting the core sugarcane microbiome

The core sugarcane microbiome was defined as bacterial and fungal OTUs detected in >75% of a particular sample type (leaf 2, leaf 4, stalk, roots, rhizosphere soil), that were in addition observed in at least one sample of each studied growing region and each variety, and in both young and mature sugarcane. For each sample type, rarefaction-normalized OTU abundance tables were used to extract the core OTUs. Given that YCS symptomatic plants might contain bacteria and fungi that were absent in asymptomatic plants, and vice versa, the core microbiome was extracted separately from YCS symptomatic plants and asymptomatic plants and subsequently merged to produce one core microbiome per sample type for analysis.

5.3.2. Statistical analyses and data visualization

To quantify the effects of sample type, growing region, crop age, sugarcane variety and YCS on bacterial and fungal assemblages of sugarcane leaves, stalks, roots, rhizosphere soil and bulk soil, rarefaction-normalised OTU abundance tables were analysed using PRIMER-E software version 6 (Plymouth, UK; Supporting Information Methods S3). Due to an unbalanced growing region (n54) 3 variety (n54) 3 crop age (n52) 3 YCS (n52) design, we used one-way ANOSIM (Clarke, 1993) to investigate the effects of sample type (leaf 2, leaf 4, stalk, roots, rhizosphere soil, bulk soil), and 2 way crossed ANOSIM to test the effects of growing region (Ayr, Ingham, Mackay, Tully), crop age (young plants in November 2014 vs. mature plants in June 2015), variety (KQ228A, Q240A MQ239, Q242A) and YCS across all sample types (Clarke and Warwick, 2001). Permutation-based multivariate analysis of variance (PERMANOVA; 999 permutations; Anderson, 2001) was applied on particular subsets of samples which permitted the investigation of interaction effects of certain drivers on bacterial and fungal community assemblages (e.g., variety 3 crop age effects for Ingham samples). Differences in the community assemblages were visualised using principal coordinates analysis (PCO; Gower, 1966).

Variation partitioning analysis was conducted to estimate the pure versus combined effects of the explanatory factors on the bacterial and fungal community assemblages as described in detail by Anderson and Cribble (1998). The significance of each pure effect was tested by 999 random permutations using the DISTLM procedure (McArdle and Anderson, 2001). The contribution of individual bacterial and fungal OTUs to differences in community assemblages among explanatory factors was determined by SIMPER analysis using PRIMER-E software version 6 (Plymouth, UK). To identify core OTUs that changed in relative abundance with YCS, multivariate linear models were applied for each sample type using MaAsLin (http://huttenhower.sph.harvard.edu/maaslin).

Cladograms containing bacterial and fungal genera contributing to sample type differences or core YCS signals, were constructed using GraPhlAn (https://huttenhower.sph.harvard.edu/graphlan). Venn diagrams showing bacterial and fungal OTUs shared between sample types were constructed based on rarefied OTU abundance tables using InteractiVenn (Heberle et al., 2015).

Statistical analyses on alpha diversity metrics were performed in R (https://www.R-project.org/), using the Kruskal-Wallis test to determine significant effects and Dunn tests for post-hoc pairwise comparisons.

5.3.3. Potential to use microbiome data as a YCS diagnostic tool

The involvement of biotic interactions in YCS development were studied by characterising the microbial community (bacteria and fungi) of 422 sugarcane tissue samples (leaves, stalks, roots) and 176 soil samples (rhizosphere and bulk soil) collected from YCS symptomatic and asymptomatic blocks. This included samples from 4 sugarcane varieties (KQ228, MQ239, Q240 and Q242) and 4 sugarcane-growing regions (Ingham, Ayr, Mackay, Tully).

Due to the collaborative nature of the FV10 sampling campaign, we obtained several metadata for the sugarcane leaf samples (determined in P2016/015), such as leaf chlorophyll content, sucrose content, water content, and metabolites that differed between control and YCS symptomatic leaves. Those metadata allowed the application of co-occurrence network analysis to interrogate the microbiome data together with the metadata. Briefly, co-occurrence network analysis (MINE; http://exploredata.net/) was applied to reveal both linear and non-linear associations between bacterial OTUs, fungal OTUs and metadata. This type of analysis is often used in medical science to identify associations between genes, transcripts and/or organisms, and has recently been postulated as a systems framework for identifying candidate microbial assemblages for plant disease management.

5.4. Physical, chemical and biological analysis of soil samples

Soils were sieved (<2 mm) prior to analysis. Soil moisture was measured gravimetrically by drying overnight at 105°C. Soil pH was determined in a 1:2.5 water suspension (Blakemore *et al*., 1987). Total carbon (TC) and nitrogen (TN) in soil were determined by combustion on a LECO macro-CN analyser (LECO, St Joseph, MI, USA). Available C (total dissolved organic (DOC) and inorganic (DIC) carbon) and N (dissolved nitrogen, DN) were measured on 1:5 soil:water extracts with a High TOC II analyser (Elementar, Germany).

Activities of extracellular enzymes were assessed according to Saiya-Cork *et al*. (2002) and Sinsabaugh *et al*. (2003). Six assays were performed to assess the activities of α-1,4-glucosidase, β-1,4-glucosidase, cellobiohydrolase, chitinase, β-xylosidase and acid phosphatase using the substrates 4-methylumbelliferyl (MUB)-α-D-glucopyranoside, 4-MUB-β-D-glucopyranoside, 4-MUB-β-D-cellobioside, 4-MUB-N-acetyl-β-D-glucosaminide, 4-MUB-N-acetyl-β-D-xylopyranoside and 4-MUB-phosphate, respectively. Soil respiration was assessed using the MicroResp^{TM} CO_2 detection system (Campbell *et al*., 2003). The respiratory response to addition of sterile, deionised water (basal respiration) and seven carbon sources was measured in triplicate for rhizosphere soil collected under each sugarcane stool.

5.5. Impact of YCS on plant defence and stress response.

Salicylic acid, which can be used as a biomarker for YCS prediction, was extracted from leaf samples of YCS symptomatic and healthy cane from the first sampling and these were submitted for HPLC analysis.

RNA was extracted from 48 leaf samples (24 leaf2 samples and 24 leaf4 samples) collected from 3 plants of each of the symptomatic and asymptomatic blocks of three sugarcane varieties that were sampled in Ingham in November 2014. Quantitative RT-PCR was performed on the RNA extracts for catalase, dehydrin and Pst2A genes involved in oxidative stress response, response to abiotic stress (in particular drought) and sugar transport, respectively. Catalase expression has previously been found to be upregulated in sugarcane plants resistant to sugarcane smut.

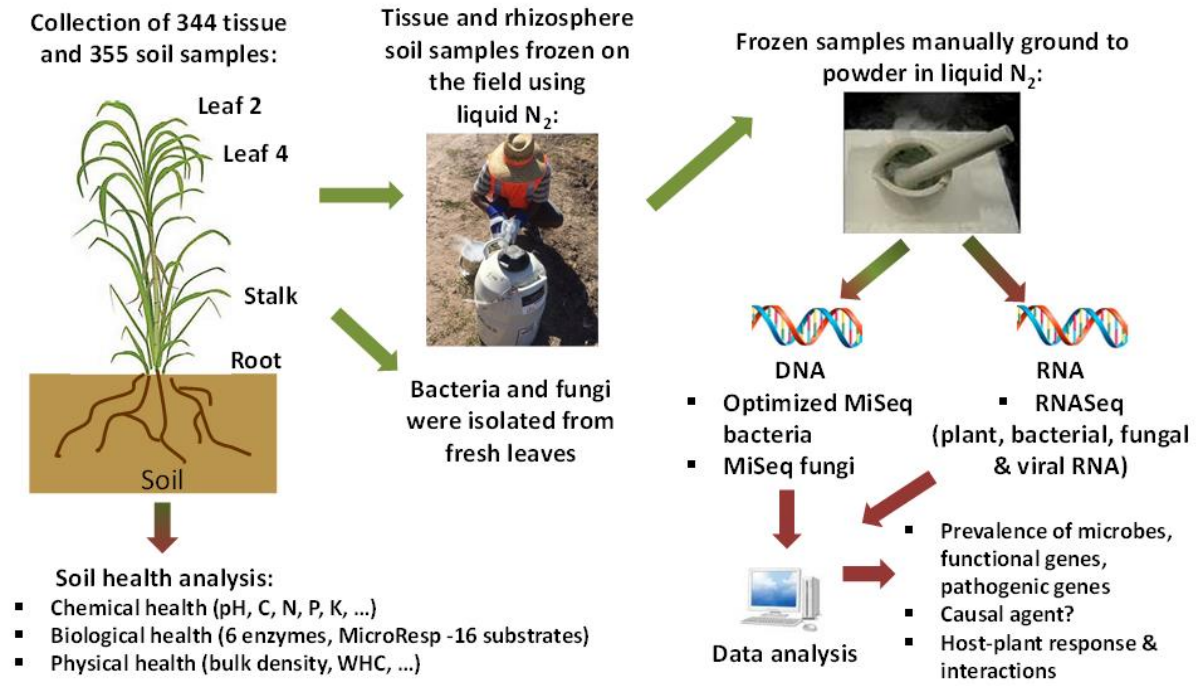

Figure 1: Schematic overview of samples and data analyses.

6. RESULTS AND DISCUSSION

6.1. Microbial communities and YCS incidence (this section is published in Environmental Microbiology)

6.1.1. Bacterial communities in plant samples

Clear differences were observed in the bacterial community composition of above-ground vs. below-ground samples, as demonstrated by the distinct clusters of leaf and stalk vs. root and soil samples in Figure 2 (PERMANOVA P<0.001). Within each plant, the bacterial communities detected in leaf2 and leaf4 samples did not significantly differ from each other (data not shown) but they were distinct from the bacterial community observed in the stalk, roots, rhizosphere and bulk soil.

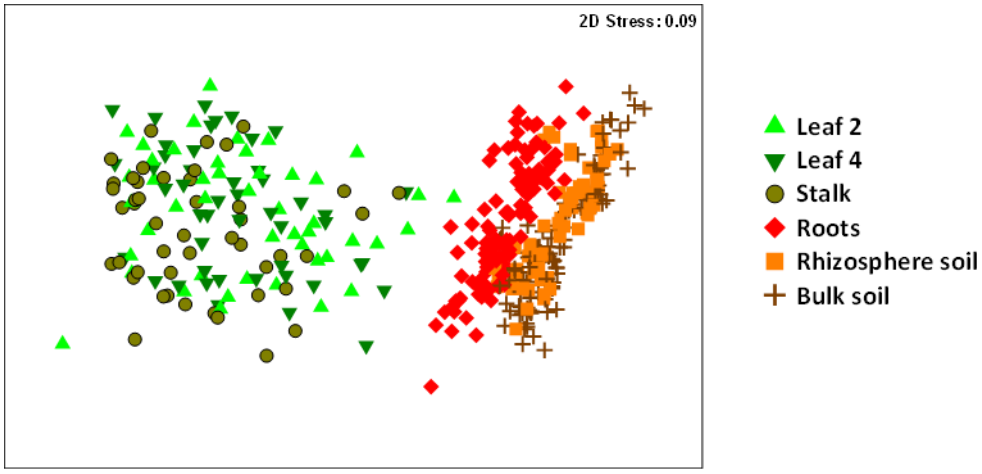

Figure 2: nMDS plot of Bray-Curtis similarities calculated based on the relative abundance of bacterial OTUs in sugarcane tissue and soil samples collected in Ingham in November 2014 and roots and soil samples collected in June 2015. Differences in the community composition of the different sample types are shown. Samples that plot in close proximity have similar bacterial community compositions, whereas samples that plot further away from each other have distinct bacterial communities.

In both the above-ground plant tissues and below-ground root and soil samples, Proteobacteria was the most predominant bacterial phylum, followed by Actinobacteria (Figure 3). Surprisingly, 31 to 41% of the bacterial sequences obtained from leaf and stalk samples were unclassified at phylum level (Figure 3). These sequences potentially belong to novel, uncharacterised bacteria that are currently not included in the curated bacterial 16S rRNA gene taxonomic databases (Greengenes, RDP). There was no obvious difference in relative abundance of these unclassified sequences between asymptomatic and YCS symptomatic plants, indicating that they might not be directly linked to YCS development. In contrast, almost all sequences obtained from root (on average 99.66%) and soil (on average 99.71%) samples were classified at phylum level (Figure 3). In roots and soil samples, the phyla Proteobacteria, Actinobacteria, Chloroflexi and Acidobacteria accounted for 75 to 79% of all sequences, respectively (Figure 3). In particular, Chloroflexi, Acidobacteria, Bacteriodetes and Verrucomicrobia were more abundant below- than above-ground.

A clear variety effect on the bacterial community composition in sugarcane leaves and stalks was detected (Figure 4, PERMANOVA P<0.001). Each sugarcane variety thus appears to have its own unique bacterial

microbiome. Apart from a Variety effect on the bacterial community composition of leaf and stalk samples, a strong Variety by YCS interaction effect was observed (P< 0.001) for leaf, stalk and root tissues (Figure 4 for leaves and stalks, data not shown for roots). Asymptomatic and YCS symptomatic plants of varieties Q240 and MQ239, but not KQ228, had distinct bacterial community compositions (Figure 4). Moreover, for each variety, the bacterial community of root samples of asymptomatic vs. symptomatic plants differed from each other (data not shown).

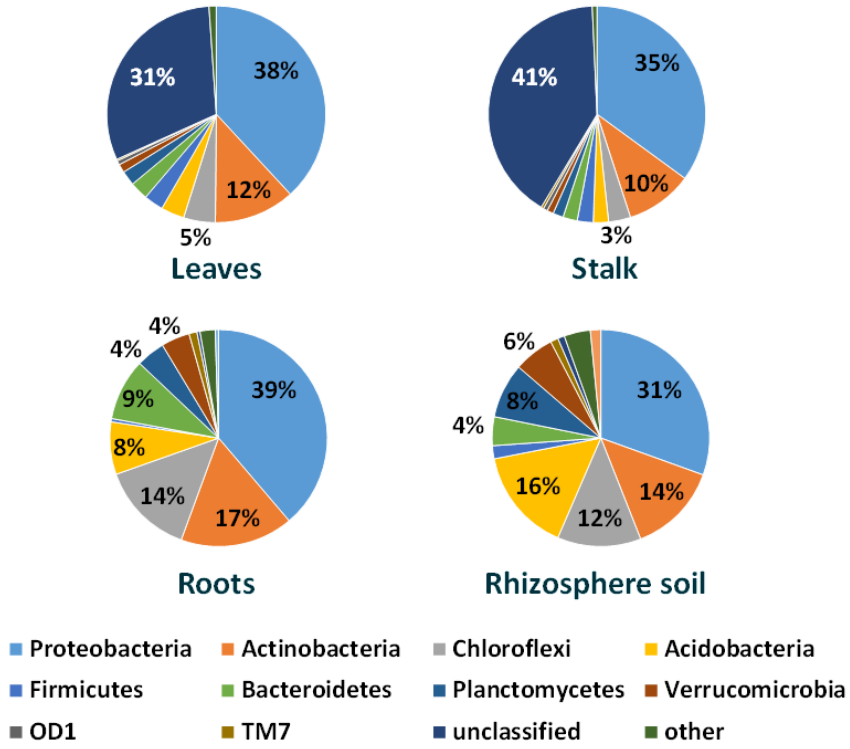

Figure 3: Pie charts showing the relative abundance (%) of the top 10 bacterial phyla detected in leaves, stalks, roots and rhizosphere soil samples. Values are averages calculated for all samples per sample type.

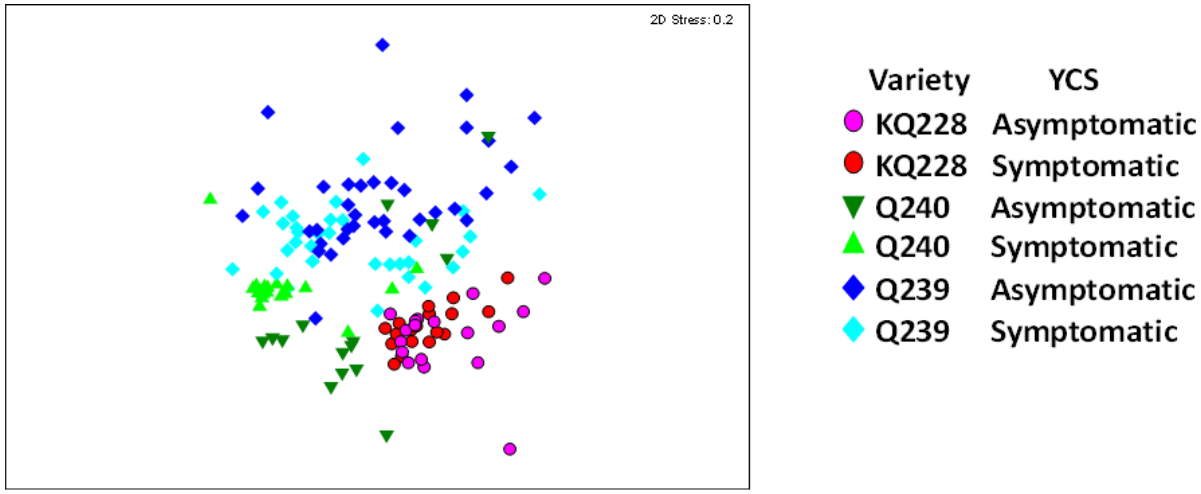

Figure 4: nMDS plot of Bray-Curtis similarities calculated based on the relative abundance of bacterial OTUs in leaf and stalk samples collected in Ingham in November 2014. Differences in the bacterial community composition of asymptomatic vs. YCS symptomatic leaves and stalks for the three sugarcane varieties (KQ228, MQ239 and Q240) are shown. Similar patterns were observed when leaf 2, leaf 4 and stalk samples were analysed separately.

Drivers of the differences between asymptomatic and YCS symptomatic leaf, stalk and root tissues included several bacterial genera containing known plant pathogens such as Sphingomonas, Curtobacterium, Agrobacterium, Streptomyces and Serratia (e.g. Figure 4A for leaf and stalk samples of Q240), along with many genera that are unclassified and might contain novel, uncharacterised bacteria (e.g. Figure 4B for root samples of KQ228). Variety-specific changes in the relative abundances of these genera between asymptomatic and symptomatic plants were observed, indicating that at first sight there was no clear signal associated with YCS incidence across the three studied varieties. Detailed analysis was thus required to identify a list of bacterial species whose relative abundances consistently changed with YCS incidence in all three varieties.

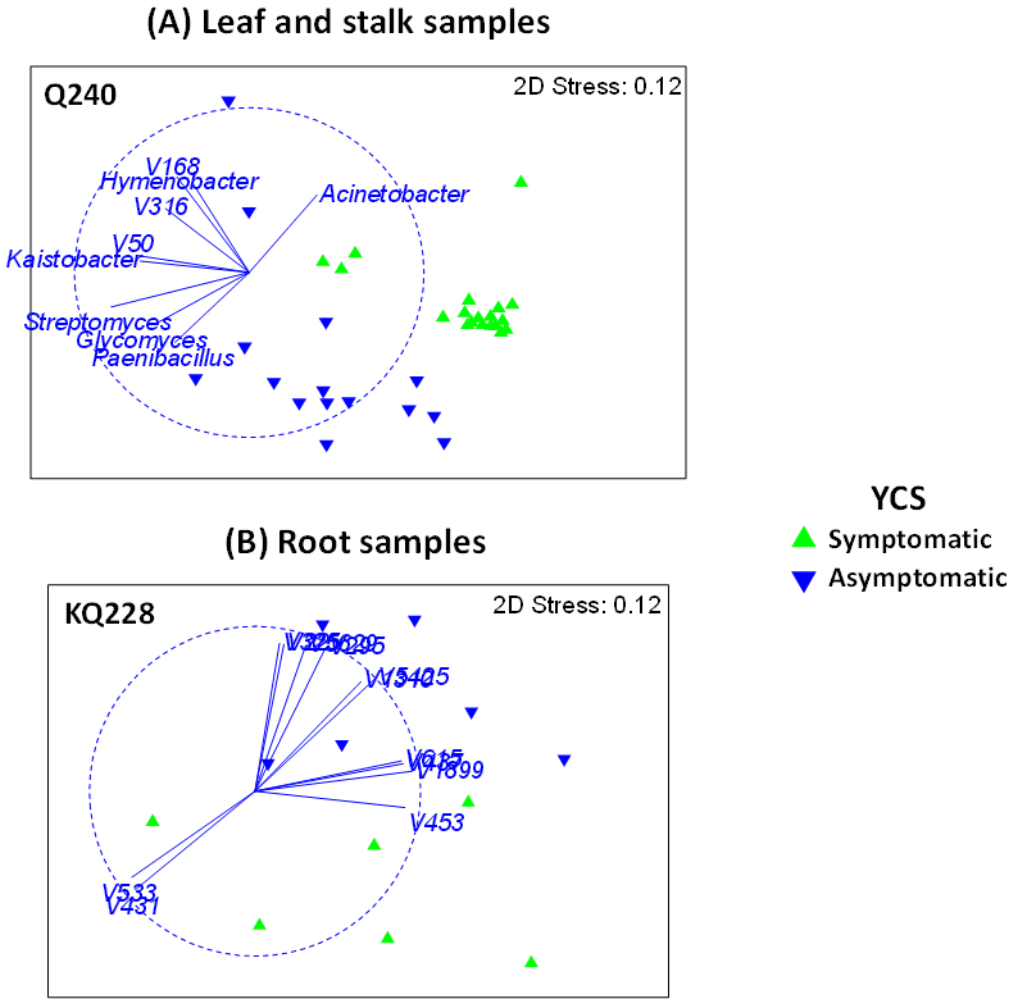

Figure 5: nMDS plot of Bray-Curtis similarities calculated based on the relative abundance of bacterial OTUs in (A) leaf and stalk samples of variety Q240 or (B) root samples of variety KQ228 collected in Ingham in November 2014. Differences in the bacterial community composition of asymptomatic vs. YCS symptomatic tissues are shown. Bacterial genera that correlated (Spearman ρ> 0.8) with the observed differences between symptomatic and asymptomatic tissues are depicted. Vx indicates unclassified genera.

Apart from the significant Variety and Variety x YCS interaction effects on the bacterial community composition, a strong Location effect (P=0.001) was observed for roots and soil samples (Figure 6A for root samples). For these samples, data from the June 2015 sampling campaign is already available. Interestingly, the strong Variety effect on the bacterial community composition observed for the Ingham samples (e.g.

Figure 3 for leaves and stalks) was less evident for all locations combined (Figure 6B). Instead, a Location x Variety interaction effect was detected (P=0.001), along with a Location x YCS effect (P=0.021) and a Variety x YCS effect (P=0.002; Figure 6C) on the bacterial community composition. As can be observed from Figure 6C, however, the strong Variety x YCS effect observed for the Ingham samples (e.g. Figure 4 for leaves and stalks) became less clear when samples collected at all locations were analysed together. Altogether, these data show that the effect of Location on the bacterial community composition of root samples is greater than the Variety and YCS effects, which is not surprising given the samples are collected over >600 km distance across different climatic regions. Moreover, these results demonstrate the difficulties associated with filtering out a clear YCS signal from the noise in the bacterial community created by shifts in the community composition with location, sugarcane variety, and other factors such as the crop cycle and chemicals applied to the field, which also significantly affected the bacterial community composition (data not shown).

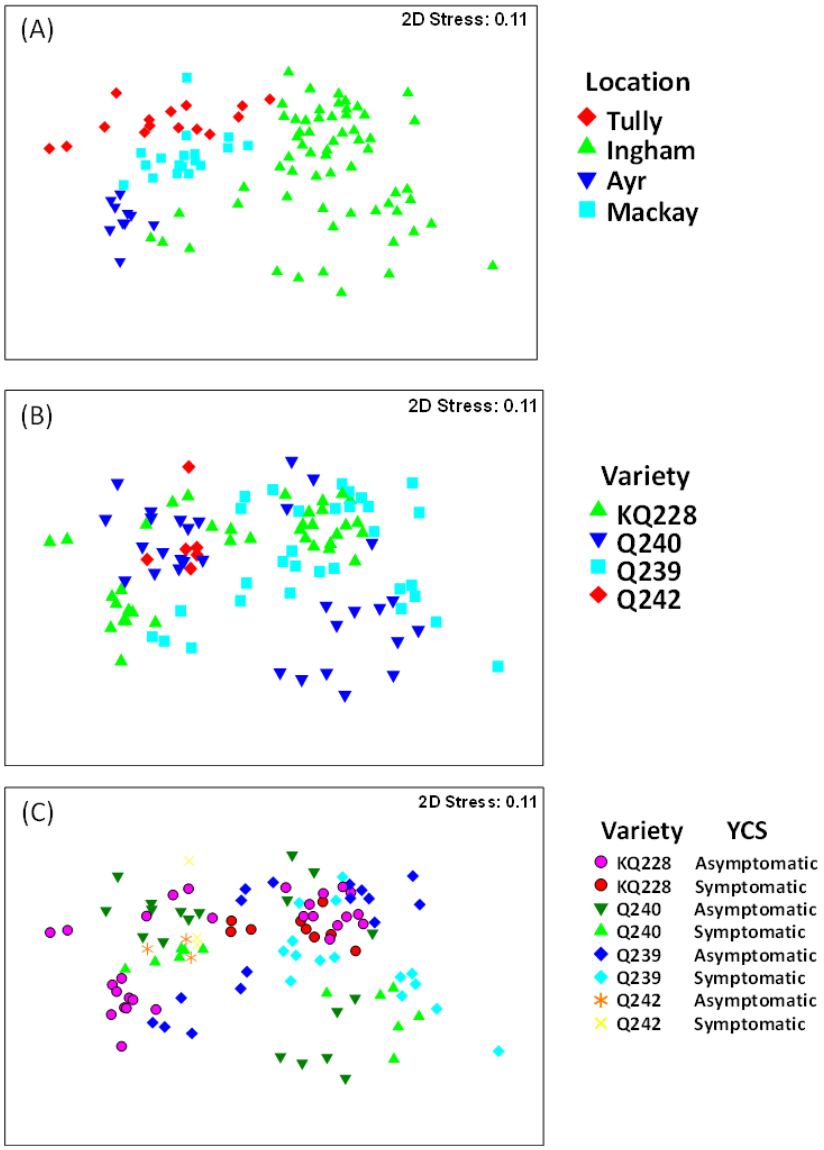

Figure 6: nMDS plot of Bray-Curtis similarities calculated based on the relative abundance of bacterial OTUs in root samples (collected November 2014 and June 2015). Differences in the bacterial community composition of root samples by (A) Location and (B) sugarcane Variety are depicted, as well as the observed (C) Variety x YCS effect.

Despite the challenge to identify significant changes in the relative abundance of 56,259 bacterial OTUs with YCS incidence that were consistent for all studied sugarcane varieties, a shortlist of bacterial OTUs was generated based on the sequence data processed (Table 3). The software STAMP (http://kiwi.cs.dal.ca/Software/STAMP) was used to visualise the relative abundance of all individual OTUs for the different groups (YCS symptomatic vs. asymptomatic samples) and OTUs that showed consistent trends in relative abundance with YCS incidence for each variety were selected based on visual inspection of bar plots. The OTUs with P<0.05 after Benjamini-Hochberg False Discovery Rate (FDR) correction while comparing differences in their relative abundance in YCS symptomatic vs. asymptomatic samples of all varieties combined were considered significant when they also showed consistent trends for each variety individually.

Table 3: Shortlist of bacterial OTUs showing significant and/or consistent changes in relative abundance (RA) between YCS symptomatic and asymptomatic samples.

Higher RA (%) in YCS symptomatic plants	Lower RA (%) in YCS symptomatic plants
CONSISTENT OTUs[1]:	
LEAVES: several unclassified bacteria (incl. OTU_710), *Brevundimonas*, *Acinetobacter*	**LEAVES:** several unclassified bacteria
STALK: Unclassified bacterium (OTU_710)	**STALK:** *Sphingomonas*
ROOTS: *Burkholderia*, *Ktedonobacter*, *Catenulispora*	**ROOTS:** *Streptomyces*
RHIZOSPHERE SOIL: several unclassified bacteria, *Bacillus*, *Acidobacteria* Gp3	**RHIZOSPHERE SOIL:** several unclassified bacteria, *Verrucomicrobia*, *Acidobacteria* Gp1
SIGNIFICANT OTUs[2]:	
LEAVES: Unclassified bacterium (OTU_710)	-

[1]OTUs showing consistent changes in relative abundance with YCS incidence for each variety after visual inspection
[2]Consistent OTUs with *P*<0.05 after Benjamini-Hochberg FDR correction

Among the bacterial OTUs that were consistently detected in higher abundance in YCS symptomatic samples (Table 3), several unclassified bacteria were found, as well as bacterial genera containing potential plant pathogens (e.g. Acinetobacter) or potential plant-growth promoting bacteria (e.g. *Brevundimonas*, *Burkholderia*, *Bacillus*). The single bacterial OTU that was present in significantly higher relative abundance in YCS symptomatic leaves was an unclassified bacterium. Among the bacterial OTUs that were consistently detected in lower abundance in YCS symptomatic samples, several genera containing potential plant-growth promoting bacteria were found (e.g. *Sphingomonas*, *Streptomyces*) but none of these OTUs were detected in significantly lower numbers after accounting for the false discovery rate (Table 3).

6.1.2. Fungal communities in plant samples

Fungal ITS2 amplicon sequence data for all samples collected in Ingham have been analysed in detail. In total, 29,807,612 fungal ITS sequences were obtained for 249 samples (183 plant tissue samples and 66 soil samples). After quality filtering of the sequences and removal of chimeras, OTUs were built at 97% sequence identity. Among the 249 samples, 9,260 fungal OTUs were identified.

As observed for bacterial beta diversity (Figure 7), the fungal community composition of above-ground plant tissues differed from root and soil samples (P<0.001; Figure 7). In contrast to the bacterial community

composition, the fungal community composition of leaf2, leaf4 and stalk samples significantly differed from each other, as well as from roots and soil samples (Figure 7).

The relative abundance of the 5 fungal phyla detected in the plant tissue and soil samples is depicted in Figure 8. The most predominant fungi in all sample types belonged to the phylum Ascomycota, followed by Basidiomycota. The relative abundance of Ascomycota was higher in above-ground compared to below-ground samples, whereas the opposite trend was observed for Basidiomycota (Figure 8). Leaves (3%) and in particular stalk samples (24%) contained a larger amount of unclassified fungi at the phylum level compared to roots (1.0%) and soil (3.6%) samples (Figure 8). This may indicate the presence of currently uncharacterized fungi in sugarcane leaf and stalk tissues.

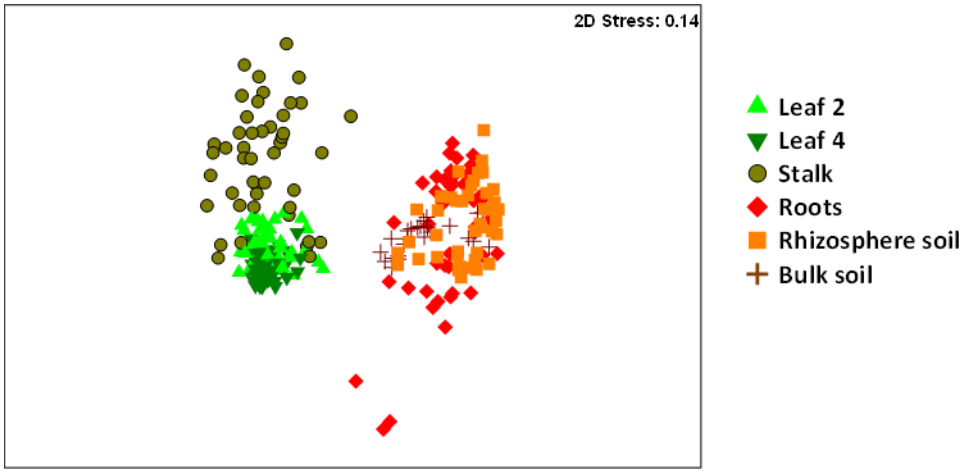

Figure 7: nMDS plot of Bray-Curtis similarities calculated based on the relative abundance of fungal OTUs in sugarcane tissue and soil samples collected in Ingham in November 2014. Differences in the community composition of the different sample types are shown.

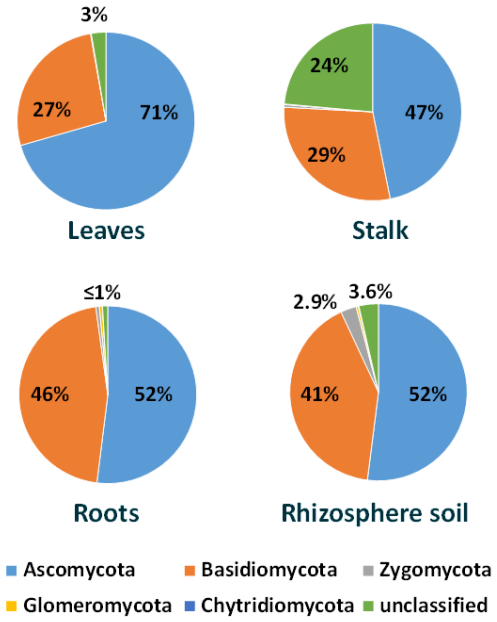

Figure 8: Pie charts showing the relative abundance (%) of fungal phyla detected in leaves, stalks, roots and soil samples. Values are averages calculated for all samples per sample type.

Interestingly, two fungal OTUs together accounted for at least 50% of the total fungal community detected in leaves and stalks of the three studied sugarcane varieties. The fungal ITS2 sequence of one of these OTUs is 100% identical to the ITS2 sequence of *Epicoccum nigrum*, a sugarcane endophyte known to produce antifungal compounds and induce root growth. The other OTU is a *Cladosporium* species. Members of the *Cladosporium g*enus are commonly found on plant leaf surfaces as saprobes or weak pathogens. No link was observed between the relative abundance of the *Cladosporium* OTU and YCS development, whereas *Epicoccum nigrum* was detected in higher abundance in YCS symptomatic leaves (see Table 4).

Similar to the bacterial community, strong Variety (P<0.001) and Variety x YCS (P<0.001) effects on the fungal community composition of sugarcane tissue samples were detected (Figure 9). Each sugarcane variety thus appears to harbour its own unique bacterial, as well as fungal community. For each variety, the fungal community composition of asymptomatic leaves, stalks and roots differed from the respective YCS symptomatic tissues (Figure 9).

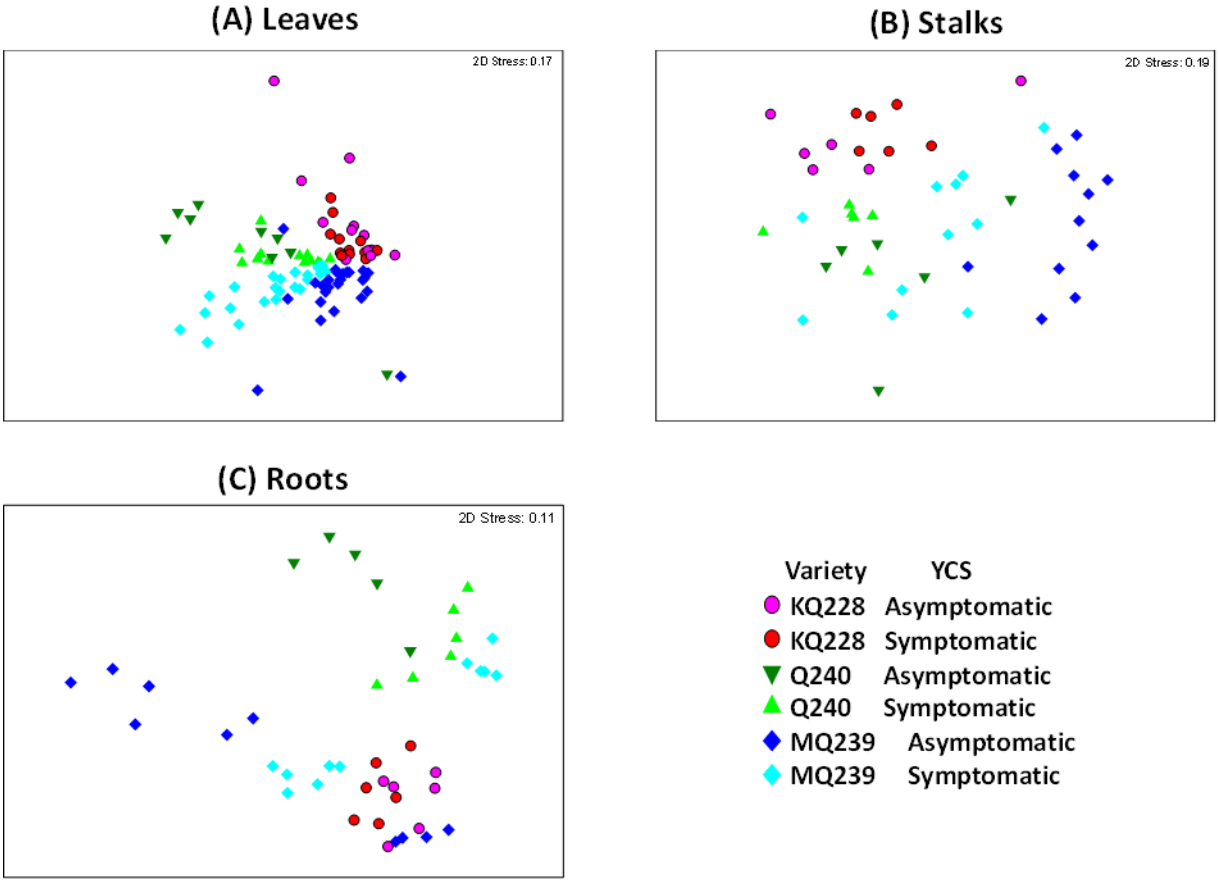

Figure 9: nMDS plot of Bray-Curtis similarities calculated based on the relative abundance of fungal OTUs in (A) leaves, (B) stalks and (C) roots collected in Ingham in November 2014. Differences in the fungal community composition of asymptomatic vs. YCS symptomatic tissues for the three sugarcane varieties (KQ228, MQ239 and Q240) are shown. When leaf 2 and leaf 4 samples were analysed separately, similar patterns as shown in (A) were observed.

Drivers of the differences between asymptomatic and YCS symptomatic tissues included several fungal genera containing known plant pathogens such as *Cladosporium*, *Fusarium*, *Ustilago*, *Mycosphaerella*, *Periconia*, *Libertella*, *Leptoxyphium*, *Acremonium*, *Aspergillus*, *Resinicium*, *Scytalidium*, *Alternaria* and *Hyphoderma* (e.g. Figure 10 for leaf samples). In addition, several unclassified fungal OTUs belonging to the

order *Pleosporales*, which contains a large number of highly destructive plant pathogens, contributed to the observed differences between asymptomatic and symptomatic samples. As for the bacterial genera, variety-specific changes in the relative abundances of these fungal genera between asymptomatic and symptomatic plants were observed. In addition, OTUs belonging to the same genus did not always show similar changes in relative abundance, indicating specific changes at the OTU level. As for the bacteria, complex shifts in the fungal community composition were observed between asymptomatic and YCS symptomatic plants and further detailed analysis of the obtained fungal dataset was thus required.

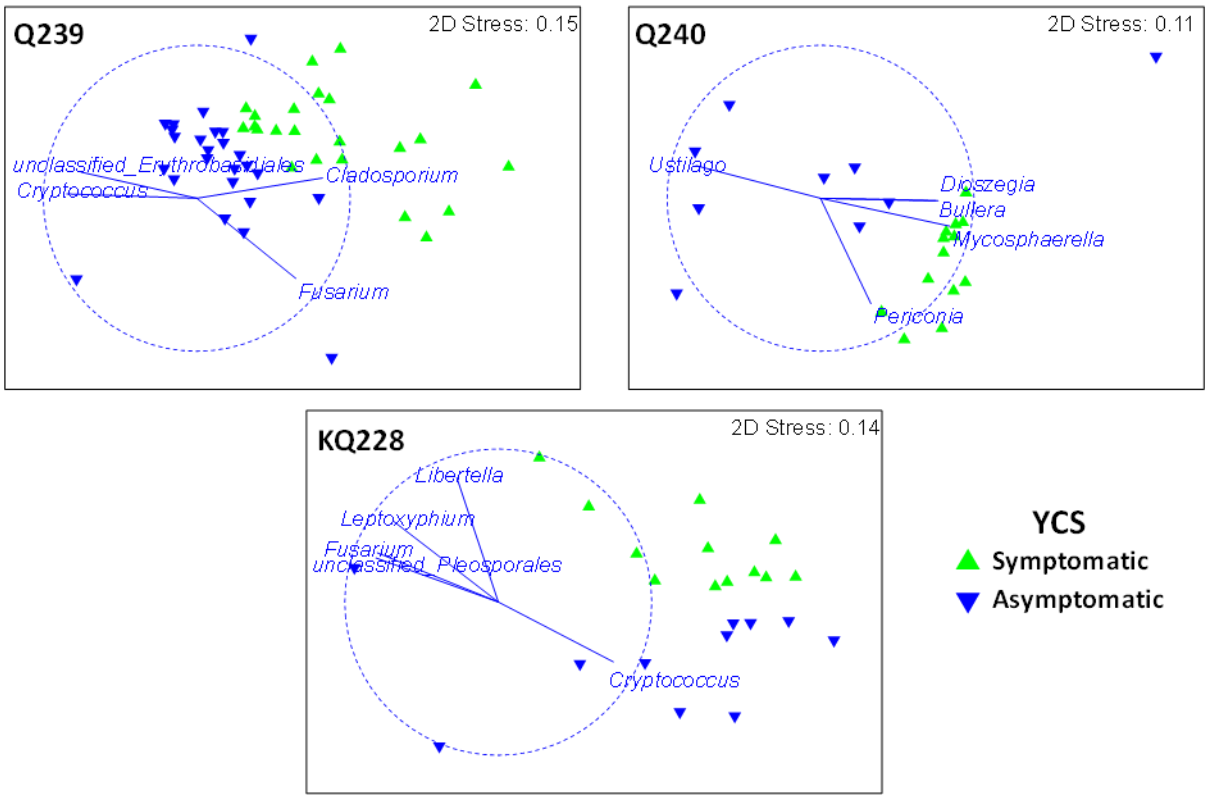

Figure 10: nMDS plot of Bray-Curtis similarities calculated based on the relative abundance of fungal OTUs in leaf samples of varieties MQ239, Q240 and KQ228 collected in Ingham in November 2014. Differences in the fungal community composition of asymptomatic vs. YCS symptomatic leaves are shown. Fungal genera that correlated (Spearman ρ> 0.8) with the observed differences between symptomatic and asymptomatic leaves are depicted.

Individual fungal OTUs that showed significant and/or consistent changes in relative abundance between YCS symptomatic and asymptomatic samples were identified (Table 4). Among the fungal OTUs that were consistently detected in higher abundance in YCS symptomatic samples (Table 4), several unclassified fungi were identified, as well as fungal genera containing potential plant pathogens (e.g. Periconia, Curvularia, Alternaria, Fusarium, Cochliobolus) or potential plant-growth promoting fungi (e.g. Epicoccum nigrum). Similarly, several unclassified fungi, fungal genera containing potential plant pathogens (e.g. Ustilago, Cryptococcus) and potential plant-growth promoting fungi (e.g. Ceratobasidium) were detected among the fungal OTUs that were consistently detected in lower abundance in YCS symptomatic samples (Table 4). In contrast to the bacteria, several of these fungal OTUs were present in significantly higher (or lower) relative abundance in YCS symptomatic leaves and rhizosphere soils collected from YCS symptomatic plants, as determined after taking the false discovery rate into account (Table 4).

Table 4: Shortlist of fungal OTUs showing significant and/or consistent changes in relative abundance (RA) between YCS symptomatic and asymptomatic samples.

Higher RA (%) in YCS symptomatic plants	Lower RA (%) in YCS symptomatic plants
CONSISTENT OTUs[1]:	
LEAVES: *Periconia* spp., *Curvularia intermedia*, *Alternaria*, *Epicoccum nigrum*, *Fusarium* cf *equiseti*	**LEAVES:** *Cryptococcus* spp., *Ustilago longissima* var. *macrospora*
STEM: unclassified fungi, *Hypocrea liscii*	**STEM:** several unclassified fungi, *Ustilago longissima*, *Cryptococcus* spp.
ROOTS: several unclassified fungi, *Sistotrema*, *Ceratobasidium* sp. JT092	**ROOTS:** *Chaetosphaeriales*, *Ceratobasidium* sp. AG-G
RHIZOSPHERE SOIL: several unclassified fungi, *Periconia*, *Curvularia intermedia*, *Pleosporales* sp. 14 MU_2012, *Cochliobolus lunatus, Fusarium*	**RHIZOSPHERE SOIL:** *Westerdykella* sp. WQ63
SIGNIFICANT OTUs[2]:	
LEAVES: *Periconia*, *Curvularia intermedia*, *Alternaria*, *Epicoccum nigrum*	**LEAVES**: *Ustilago longissima* var. *macrospora*
RHIZOSPHERE SOIL: *Fusarium*	

[1]OTUs showing consistent changes in relative abundance with YCS incidence for each variety after visual inspection
[2]Consistent OTUs with $P<0.05$ after Benjamini-Hochberg FDR correction

6.1.2.1. Improve sequencing yield of targeted sequences

Amplicon sequence analysis allows avoiding issues with obtaining enough microbial sequence reads from mixed plant-microbial DNA extracts since it in theory specifically amplifies microbial DNA. Unfortunately, applying this technique to sugarcane tissue samples collected in May and September 2014 demonstrated that using bacterial 16S rRNA gene primers resulted in non-specific amplification of sugarcane chloroplast and mitochondrial DNA. On average only 3% of the obtained sequences in leaf and stalk samples and 66% of sequences in root samples were of bacterial origin. This resulted in obtaining less than 100 bacterial sequences in 1/6 of the tissue samples analysed. We have applied Peptide Nucleic Acids (PNAs) designed for chloroplast and mitochondrial DNA to block amplification of plant-derived DNA during PCR for the bacterial and fungal amplicons. Optimisation steps included changing primer sets, changing PCR reagents, increasing DNA input concentrations and prevention of overclustering that resulted in a drop in sequence quality. By July 2015, we had successfully resolved the issue of non-specific amplification of plant DNA for sugarcane root samples (99-100% of the obtained sequences are bacterial) and significantly improved the amount of bacterial sequences obtained for leaf and stalk samples from 3% (without using PNAs) to on average 29 and 25%, respectively (Figure 11). Between 1000 and 20,000 bacterial sequences were obtained for leaf and stalk samples and between 5,000 and 140,000 bacterial sequences for root samples. We further optimised our protocol for leaf and stalk samples by testing the optimal ratio of sample DNA to PNA concentrations to further increase the percentage of bacterial sequences obtained and minimise the large sample to sample variation we were observing (Figure 11).

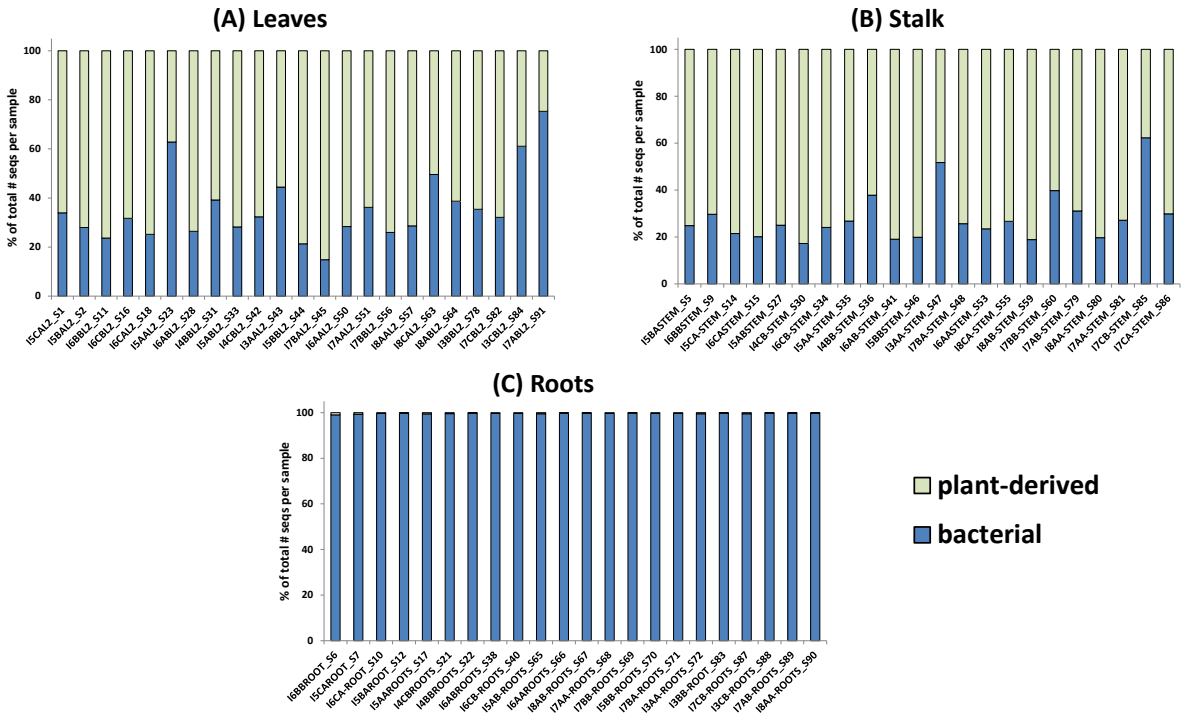

Figure 11: Typical percentages of plant-derived vs. bacterial quality-filtered sequences in leaf (A), stalk (B) or root (C) samples after optimisation of the Miseq protocol using PNAs. Bars represent the percentages in 22 randomly selected samples per sample type.

6.1.2.2. Potential to use microbiome data as a YCS diagnostic tool

We studied the involvement of biotic interactions in YCS development by characterising the microbial community (bacteria and fungi) of 422 sugarcane tissue samples (leaves, stalks, roots) and 176 soil samples (rhizosphere and bulk soil) collected from YCS symptomatic and asymptomatic blocks. This included samples from 4 sugarcane varieties (KQ228, MQ239, Q240 and Q242) and 4 sugarcane-growing regions (Ingham, Ayr, Mackay, Tully). A limited number of YCS-affected leaf samples were also studied for a viral signal (by electron microscopy and small RNA-Seq).

In total, over 34 million bacterial and 52 million fungal DNA sequences were obtained from 422 sugarcane plant tissue samples (leaves, stalks and roots) of four sugarcane varieties (KQ228, MQ239, Q240 and Q242) and 176 soil samples collected from asymptomatic and YCS symptomatic blocks near Ingham (November 2014, June 2015), Ayr (November 2014, June 2015), Mackay (June 2015) and Tully (June 2015). We detected over 25000 bacterial and over 11000 fungal species, phytoplasma and other mycoplasma-like organisms, (plant-parasitic) nematodes, bacteriophages, viruses, algae and insects.

Due to strong location, variety and crop age effects on the microbiome (Figure 12), few consistent YCS signals emerged. Finding YCS signals (significant changes in the relative abundance of bacterial or fungal OTUs with YCS incidence) that were consistent across varieties, locations and crop age, proved to be extremely challenging. Therefore, no promising biomarkers for YCS detection have yet been identified.

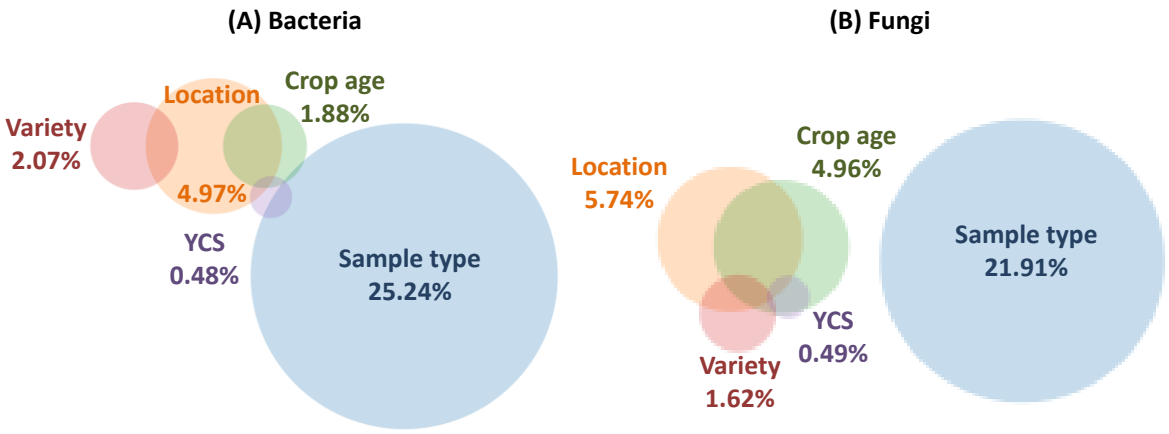

Figure 12: Venn diagrams showing the percentage variation in the bacterial (A) and fungal (B) community composition explained by sample type (leaves, stalk, roots, rhizosphere soil), location (Ingham, Ayr, Mackay, Tully), variety (KQ228, MQ239, Q240 and Q242), crop age, YCS, and their interactions, for all 598 samples collected prior to 2016. Pure effects of each factor, taking into account all other factors, are reported. All effects were statistically significant (P=0.001).

The core sugarcane microbiome has been defined as the bacterial and fungal OTUs detected in >75% of all samples of a particular sample type, that were in addition observed in at least 1 sample of each region and each variety. The core microbiome was extracted separately from YCS symptomatic plants and asymptomatic plants (as asymptomatic plants might not contain the biological agent responsible for YCS), and then merged to produce one core microbiome per sample type for analysis. Table 2 provides an overview of the number of core OTUs and their relative abundance in each sample type.

Table 5: Core OTU richness and relative abundance of core OTUs for each sample type based on the 598 sample set collected prior to 2016

Core microbiome	Bacteria		Fungi	
	OTU richness (%)[1]	Relative abundance (%)[2]	OTU richness (%)[1]	Relative abundance (%)[2]
Leaf 2	0.19	27 ± 18	1.5	68 ± 16
Leaf 4	0.22	35 ± 21	1.7	71 ± 19
Stalk	0.29	17 ± 13	0.4	25 ± 16
Roots	2.19	46 ± 8	0.9	35 ± 20
Rhizosphere soil	3.26	73 ± 15	1.2	45 ± 16

[1]Percentage of the number of core OTUs vs. the total number of OTUs observed in each sample type

[2]Average ± standard deviation

The core microbiome consisted of a very small number of bacteria and fungi (Table 5), comprising only 0.19-3.26% of all bacterial OTUs and 0.4-1.7% of all fungal OTUs observed in the different sample types. However, these core OTUs were dominant (relative abundant) members of the microbiome, accounting for on average 17-73% relative abundance of all bacteria and 25-71% relative abundance of all fungi detected in the samples. Core bacterial OTUs included *Methylobacterium*, *Bradyrhizobium*, *Sphingomonas*, *Erwinia* and *Curtobacterium* species that all have reported plant growth-promoting traits. Similarly, core fungal OTUs included several endophytes with reported plant-growth promoting traits such as *Trichoderma*, *Phoma*, *Cryptococcus* and *Fusarium spp*. However, some of the core OTUs included potential phytopathogens such as

Curtobacterium, which contains *pathovars* reported to cause bacterial wilt or leaf spot in several hosts, and potential phytopathogenic fungi belonging to the genera *Coniothyrium*, *Cladosporium*, *Fusarium*, *Periconia*, *Sporisorium*, and *Phoma*. Since several of these OTUs were not classified at species level, due to the amplified regions of the bacterial 16S rRNA gene or fungal ITS being identical for several different species of those particular genera, no information on their potential pathogenicity in sugarcane could be inferred.

Hub Microbiota: Co-occurrence network analysis of the core sugarcane leaf microbiome revealed 1061 significant associations (with MIC values ranging from 0.27 to 0.95) among bacterial and fungal core OTUs, of which most associations (96.5%) were non-linear. Fungal OTUs from the genera *Cladosporium*, *Periconia*, *Nigrospora* and *Bullera*, as well as four OTUs that were unclassified at the genus level were identified as hub microorganisms in the core leaf microbiome network. In the core stalk, roots and rhizosphere soil core microbiome networks, 2 (MIC values of 0.49 and 0.50), 338 (MIC ranging from 0.40 to 0.80) and 151 (MIC ranging from 0.48 to 0.77) significant associations were found among the core OTUs, respectively. Most of these significant associations (83 to 100%) were non-linear. Two bacterial OTUs belonging to the *Ktedonobacteraceae* and *Microsporaceae* families were identified as hub microorganisms in the sugarcane core root microbiome network, whereas two fungal OTUs (*Mortierella chlamydospora* and an unclassified Ascomycete) were identified as hub microorganisms in the core rhizosphere microbiome network (Hamonts et al., 2018).

Multivariate Association with Linear Models (Maaslin; https://huttenhower.sph.harvard.edu/maaslin) was used to study changes in the relative abundance of the core OTUs with YCS incidence. The bacterial and fungal OTUs that significantly increased or decreased with YCS incidence are depicted in Figure 13 for Leaf 4 and Root samples. Leaf 2 results were similar to Leaf 4 results and are therefore not shown. For stalks, only 1 core OTU (an unclassified species of *Cladosporium*) significantly increased in relative abundance with YCS.

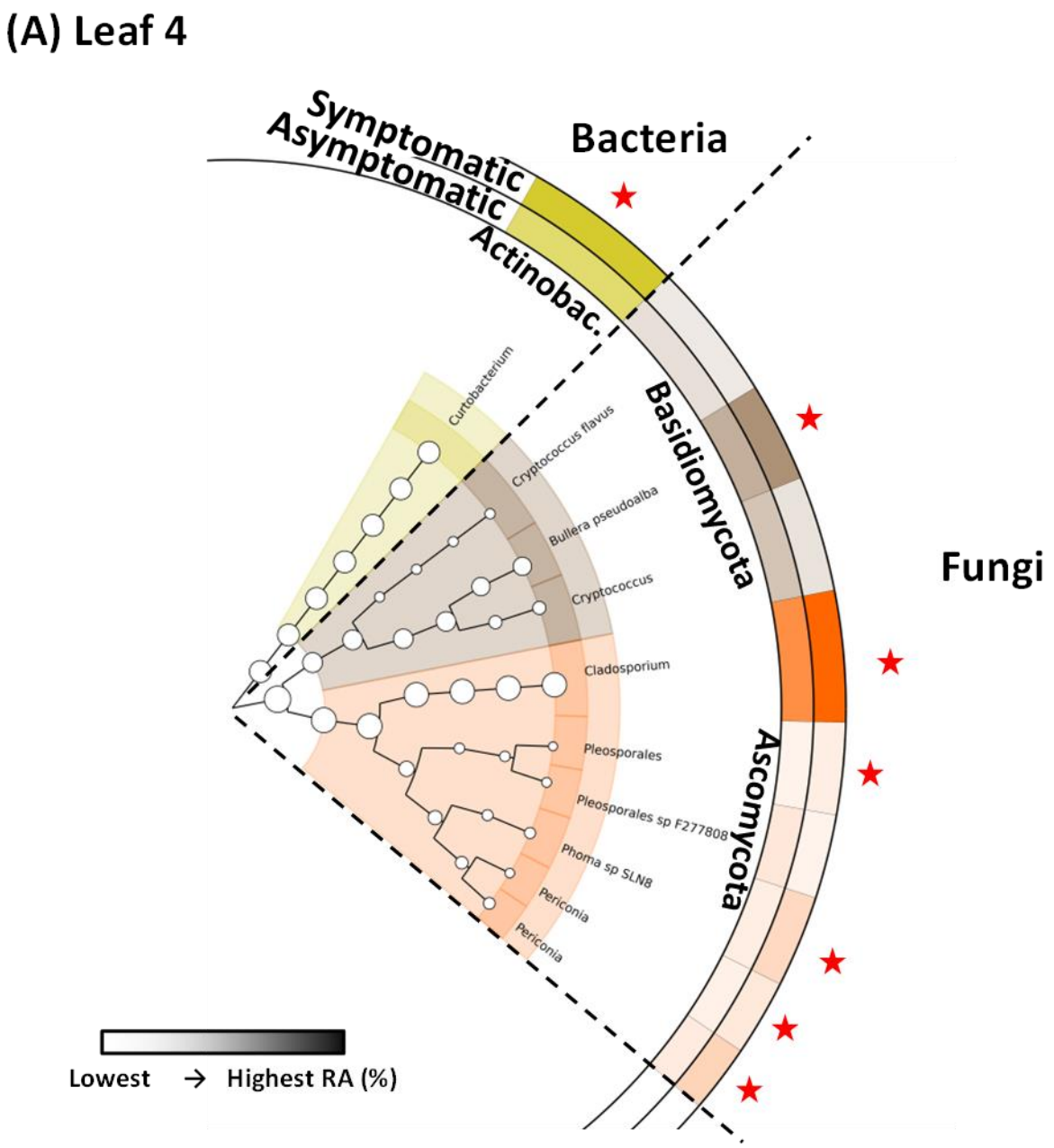

Figure 13: Cladograms showing the core bacterial and fungal OTUs whose relative abundance significantly increased or decreased with YCS incidence in (A) Leaf 4 and (B) Root samples collected prior to 2016. Size of the nodes indicates the average relative abundance of the bacterial or fungal OTUs across all samples. Colours indicate the phylogeny. Opaqueness of the outer rings reflects the average relative OTU abundance in asymptomatic and YCS symptomatic samples. Stars indicate OTUs that increased in relative abundance with YCS. Actinobac.: Actinobacteria.

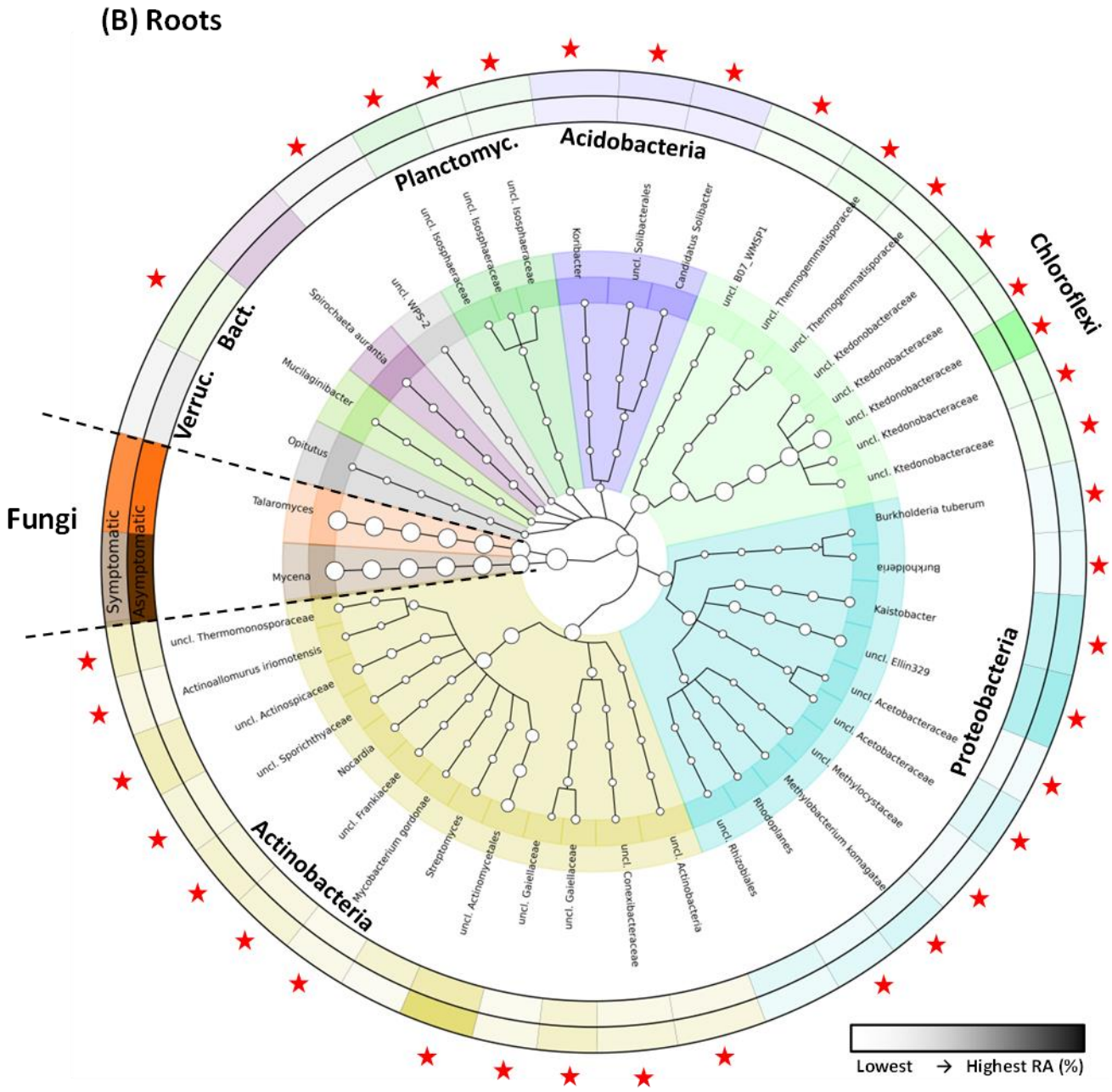

Figure 13 - ctd: Cladograms showing the core bacterial and fungal OTUs whose relative abundance significantly increased or decreased with YCS incidence in (A) Leaf 4 and (B) Root samples collected prior to 2016. Size of the nodes indicates the average relative abundance of the bacterial or fungal OTUs across all samples. Colours indicate the phylogeny. Opaqueness of the outer rings reflects the average relative OTU abundance in asymptomatic and YCS symptomatic samples. Stars indicate OTUs that increased in relative abundance with YCS. Verruc.: Verrucomocrobia; Bact.: Bacteroidetes; Planctomyc.: Planctomycetes.

Overall, more core OTUs were found to significantly change in relative abundance with YCS belowground compared to the above-ground leaf and stalk tissues (see Figure 13 for Leaf 4 and Roots). This likely reflects the overall higher bacterial diversity observed in below-ground samples compared to the aerial tissues.

The only bacterial core OTU that significantly increased in relative abundance in YCS symptomatic leaf tissues was an unclassified *Curtobacterium* species (Figure 13A, Figure 3). *Curtobacterium* is a common plant

endophyte, has been previously isolated from sugarcane stalk tissues and possesses plant-growth promoting properties (such as indole-3-acetic acid (IAA) production). Some *Curtobacterium* species have been postulated as biocontrol agents against phytopathogens, whereas others are pathogenic. Several *pathovars* of *Curtobacterium flacuumfaciens* have been reported to cause vascular diseases of beets, beans, soybeans, tulips and poinsettias. The *pathovars Curtobacterium flacuumfaciens pv. flacuumfaciens* and *Curtobacterium flacuumfaciens pv. beticola* have been identified as the causal agents of bacterial wilt and leaf spot in *Phaseolus spp.* and sugar beet plants, respectively. The wilting symptoms in *Phaseolus spp.* are caused by clogging of the xylem vessels by *Curtobacterium flacuumfaciens pv. flacuumfaciens*. Based on this information, the *Curtobacterium* core OTU might be of interest regarding YCS development, as it could fit with the "blockage hypothesis" potentially compromising water and sugar movement in YCS symptomatic leaves. However, further detailed analysis is required to reveal the identity of the *Curtobacterium* species present, and to delineate whether it's a beneficial/commensal endophyte residing in the sugarcane leaves, or whether it could be a pathogen clogging up vascular bundles.

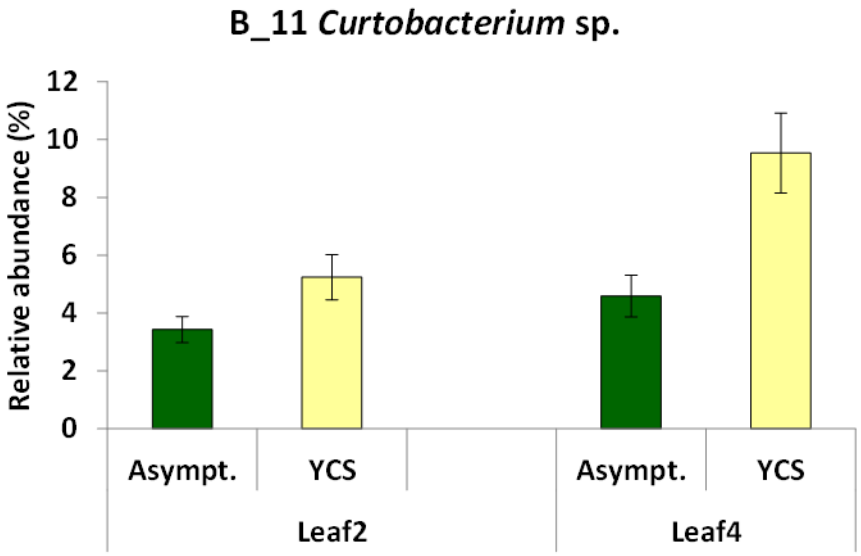

Figure 14: Relative abundance of core bacterial OTU B_11 in Leaf2 and Leaf4 samples collected before 2016. This OTU represents an unclassified *Curtobacterium* species that was significantly more abundant in YCS symptomatic leaf samples. Bars indicate averages and error bars represent the standard error of the mean.

Other core OTUs that significantly increased in abundance in YCS symptomatic leaves included several fungi belonging to the genera *Periconia*, *Cladosporium* and *Phoma* (Figure 13A). *Cladosporium* and *Phoma spp.* have reported plant-growth promoting traits and fungi belonging to the genus *Cladosporium* have been described as plant endophytes that can act as both pathogen facilitators and antagonists. Members of the genera *Periconia* and *Cladosporium* are also known plant pathogens, causing leaf and sheath spot, and the core OTUs increasing in abundance with YCS might therefore be involved in secondary effects on YCS symptomatic leaves. As discussed above for *Curtobacterium*, further research is required to identify the particular fungal species present, before their potential functions can be inferred.

In roots of YCS symptomatic plants, two fungal core OTUs (*Talaromyces* and *Mycena*) decreased in abundance with YCS and a phylogenetically diverse group of bacteria either significantly increased or decreased in relative abundance with YCS (Figure 13B). Among the bacterial core OTUs that were less

abundant with YCS, Streptomyces, a recognised biocontrol agent (pathogen antagonist), was detected. Its relative abundance decreased from an average 0.20% in roots from asymptomatic to 0.02% in YCS symptomatic plants.

Co-occurrence analysis revealed a highly connected network in the leaf dewlaps, with significantly more strong associations (Maximum Information Coefficient (MIC) $\geq$ 0.81) compared to the midribs and leaf lamina (Table 6). The higher connectivity observed in the dewlaps most likely resulted from the higher bacterial and fungal richness detected in the dewlap compared to the midrib and lamina samples. Numerous strong linear and non-linear associations were found between the YCS metadata and relative abundance of bacteria and fungi in all studied FV10 leaf sections, in particular in the dewlaps (Table 6). As expected, a strong response of the microbiome to elevated sugar levels detected in YCS symptomatic leaves was found. In the dewlaps for example, 106 positive linear associations (Pearson correlations coefficient $\rho > 0.6$) between sucrose content and the relative abundance of bacteria and fungi were detected, compared to 10 such associations in the midribs (Table 6). Bacterial and fungal OTUs that showed positive linear correlations with sucrose content did not necessarily belong to the core microbiome determined from all samples collected pre-2016. This indicates that particular bacteria and fungi can respond to the elevated sugar levels in the leaves associated with YCS, depending on the specific interactions between those micro-organisms and all other bacteria and fungi present in the sugarcane leaf of the particular sugarcane plant (dependant on the growing region, sugarcane variety and crop age). For example, in the studied FV10 field, the relative abundance of the fungus *Aureobasidium pullulans* responded very strongly to the elevated sucrose content in YCS symptomatic leaves (ρ=0.92 in dewlaps; ρ=0.96 in midribs; ρ=0.82 in lamina), whereas this particular fungus, who is a member of the core microbiome detected in all growing regions and all studied sugarcane varieties, did not respond to YCS incidence in all sugarcane fields studied in November 2016. We postulate that *A. pullulans* responded to the elevated sucrose levels in the FV10 sugarcane leaves because it could out-compete other bacteria and fungi in the particular FV10 microbiome, whereas it was in turn outcompeted by other bacteria and fungi in the microbial assemblages of leaves analysed previously. These contrasting results for A. pullulans between FV10 and the pre-2016 data are in agreement with the high number of inconsistent YCS signals observed to date, and demonstrate the complex associations between bacteria, fungi and YCS parameters such as sucrose content in sugarcane leaves.

Table 6: Summary of the number of strong (MIC ≥0.81) associations detected between the relative abundance of bacteria, fungi and several YCS variables in FV10 leaf sections

	Dewlaps (26618 associations[1])	Midribs (3866 associations[1])	Lamina (363 associations[1])
Chlorophyll content	154	19	2
Water content	155	19	2
Sucrose content	151 ▪ Non-linear: 22 ▪ Positive linear[2]: 106 ▪ Negative linear[2]: 23	28 ▪ Non-linear: 6 ▪ Positive linear: 10 ▪ Negative linear: 12	2 ▪ Non-linear: 1 ▪ Positive linear: 0 ▪ Negative linear: 1
Content of top 5 metabolites[3]	151 each	20 to 33	1 to 5

[1]Number of associations with MIC ≥0.81

[2]Positive linear: Pearson correlation ρ>0.6; negative linear: ρ <-0.6

[3]Top 5 metabolites showing the highest x-fold difference in content between controls and samples from YCS symptomatic leaves, excluding sucrose (as determined by P2015/016)

Networks containing the YCS metadata and their associations with bacteria and fungi are shown for the FV10 leaf dewlaps (Figure 15), midribs (Figure 16) and leaf lamina (Figure 17). In the dewlaps, the YCS variables and their first neighbour associations clustered centrally in the co-occurrence network (Figure 15), indicating that members of this "YCS module" that responded to shifts in water, chlorophyll, sucrose and metabolite content associated with YCS, were in turn highly connected to numerous other bacteria and fungi. The elevated sucrose content in the dewlaps due to YCS for example, thus results in cascade effects through the whole microbial community of the dewlap, affecting the relative abundance of numerous bacteria and fungi both directly and indirectly.

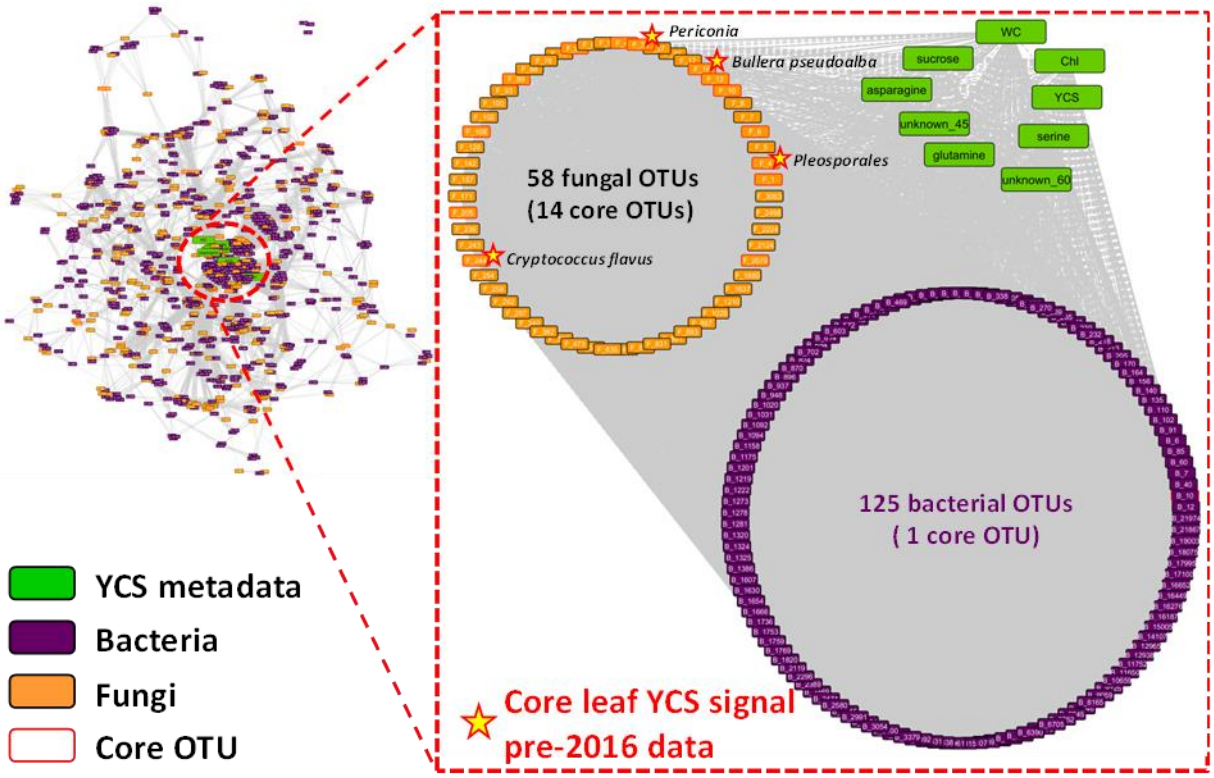

Figure 15: Co-occurrence network showing the top 18005 associations (MIC ≥ 0.99) in FV10 dewlaps between bacterial OTUs (purple), fungal OTUs (orange) and YCS metadata (water content (WC); chlorophyll (Chl); sucrose and 5 metabolites that showed the largest differences between controls and YCS samples). Nodes represent OTUs or metadata, and edges the correlations between the nodes. The central cluster in the co-occurrence network (enlarged on the right) contains the YCS metadata and their associations with bacterial and fungal OTUs. Essentially, the enlarged network on the right shows the high number of bacterial OTUs (purple) and fungal OTUs (orange) whose relative abundance is directly correlated to the leaf parameters sucrose, water content, chlorophyll and others (green), known to alter with YCS. This YCS module is in turn associated with numerous other bacteria and fungi (top left). OTUs representing core leaf YCS signals identified from pre-2016 data are highlighted by stars.

Overall, fewer strong associations between YCS variables and the relative abundance of bacteria and fungi were detected in the midribs and lamina compared to the dewlaps (Table 6, Figures 15-17). However, results from co-occurrence network analysis indicated potential YCS cascade effects through the entire microbial community of both midribs (Figure 16) and leaf lamina (Figure 17) via indirect effects of bacteria and fungi, whose relative abundance correlated with YCS variables, on the relative abundance of other co-occurring bacteria and fungi. Several members of the core microbiome were present in the “YCS modules” in both midribs (Figure 16) and lamina (Figure 17), including core leaf YCS signals identified in pre-2016 data along with numerous non-core members.

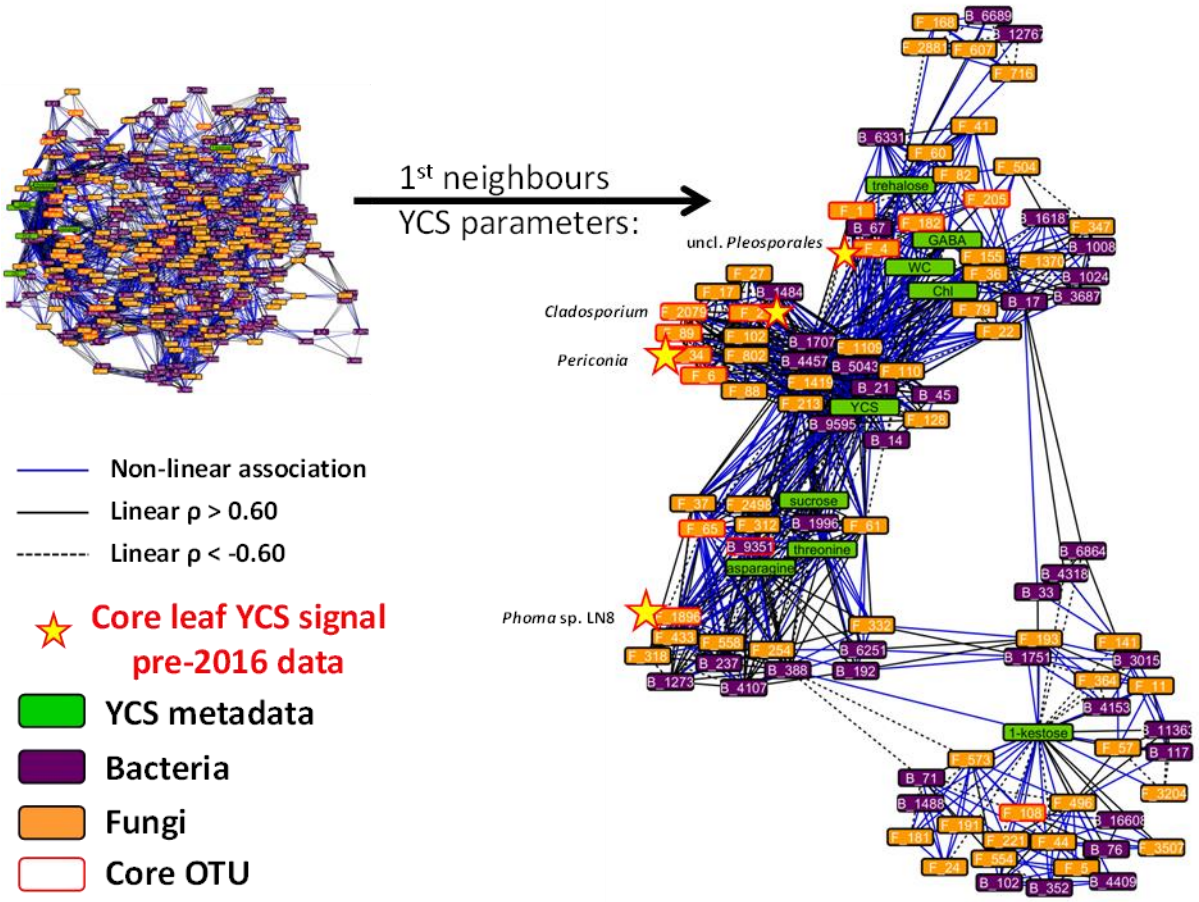

Figure 16: Co-occurrence network showing the top 3866 associations (MIC ≥ 0.81) in FV10 midribs between bacterial OTUs (purple), fungal OTUs (orange) and YCS metadata (water content (WC); chlorophyll (Chl); sucrose and 5 metabolites that showed the largest differences between controls and YCS samples). Nodes represent OTUs or metadata, and edges the correlations between the nodes. First neighbour associations between the YCS metadata and bacterial and fungal OTUs are shown on the right (YCS module). OTUs belonging to the core microbiome determined from pre-2016 samples have a red border, and core leaf YCS signals identified from pre-2016 data are highlighted by stars.

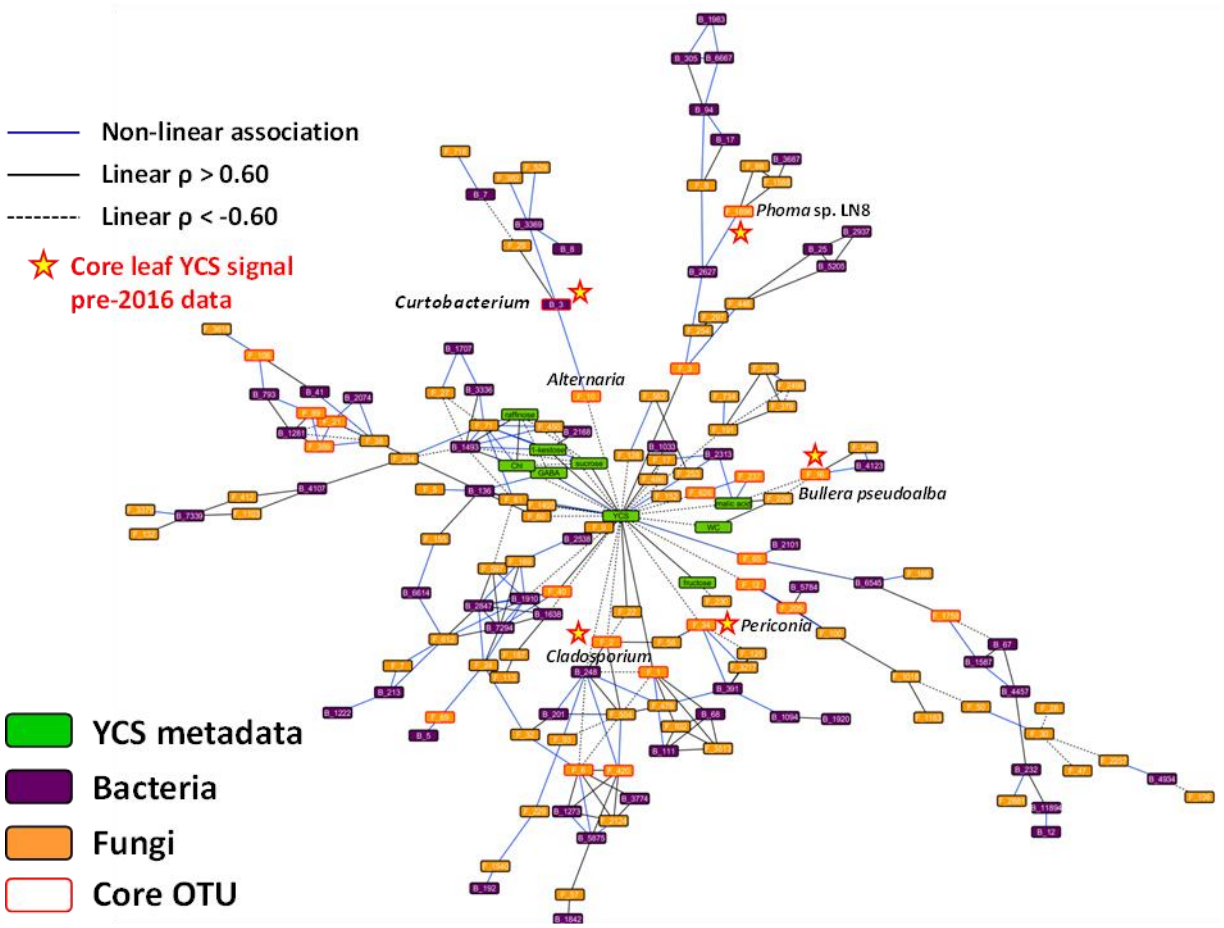

Figure 17: Co-occurrence network showing the top 363 associations (MIC ≥ 0.81) in FV10 leaf lamina samples between bacterial OTUs (purple), fungal OTUs (orange) and YCS metadata (water content (WC); chlorophyll (Chl); sucrose and 5 metabolites that showed the largest differences between controls and YCS samples). Nodes represent OTUs or metadata, and edges the correlations between the nodes. OTUs belonging to the core microbiome determined from pre-2016 samples have a red border, and core leaf YCS signals identified from pre-2016 data are highlighted by stars.

6.1.3. Untargeted shotgun metagenomics of rhizosphere soil (this work has been published in Soil Biology and Biochemistry)

The sugarcane rhizosphere soil harbours a complex community, including bacterial and fungal phytopathogens, plant-parasitic nematodes, phytoplasma- and other mycoplasma-like organisms, bacteriophages, viruses, algae and insects (data not shown). A significant Variety x YCS interaction effect (P=0.001) on the whole microbial community composition was detected (data not shown), confirming the results obtained for the rhizosphere soil samples analysed by amplicon sequencing for bacteria and fungi.

Interestingly, 10 Phytoplasma and 13 Spiroplasma species were detected in the rhizosphere soil (Table 7). No trend was found between the relative abundance of these species and YCS incidence in the studied soil samples (data not shown). However, the presence of these insect-transmitted plant pathogens in the rhizosphere raises the question whether they could be involved in YCS development. Since they normally reside in plant phloem, detection of these species in sugarcane plant tissues and/or phloem samples using specific primers targeting these organisms could identify their potential association with YCS incidence and should be further explored.

Table 7: Overview of the *Phytoplasma* and *Spiroplasma* species detected in rhizosphere soil samples.

Phytoplasma species	*Spiroplasma* species
Canadian peach X phytoplasma	*Spiroplasma gladiatoris*
Candidatus Phytoplasma mali	*Spiroplasma penaei*
Candidatus Phytoplasma solani	*Spiroplasma platyhelix*
Ipomoea phytoplasma	*Spiroplasma chrysopicola*
Aster yellows witches'-broom phytoplasma AYWB	*Spiroplasma* endosymbiont of Drosophila tenebrosa
Coconut lethal yellowing phytoplasma	*Spiroplasma insolitum*
Pepper stolbur phytoplasma	*Spiroplasma citri*
Catharanthus phyllody phytoplasma	*Spiroplasma eriocheiris*
Candidatus Phytoplasma australiense	*Spiroplasma* endosymbiont of Drosophila hydei
Loofah witches'-broom phytoplasma	*Spiroplasma litorale*
	Spiroplasma ixodetis

In addition to the phloem-restricted organisms (Table 7), xylem-restricted organisms such as *Leifsonia xyli subsp. xyli* were detected in the rhizosphere soil (data not shown). Although their relative abundance did not increase in the rhizosphere of YCS symptomatic plants, their potential involvement in YCS development should be investigated further.

In total, 11 plant viruses were detected in the rhizosphere soil samples, including 3 sugarcane viruses: sugarcane bacilliform virus, sugarcane streak Egypt virus and sugarcane streak Reunion virus. The sugarcane bacilliform virus was detected in most soil samples but did not show a pattern with YCS incidence, whereas the other two sugarcane viruses were only detected in 1 of the 24 studied samples.

Apart from gaining detailed information on the relative abundance of the whole microbial community in the sugarcane rhizosphere soil, the shotgun metagenomics approach allowed the functional potential of the rhizosphere soil microbial community to be investigated. Significant Variety (P=0.001) and YCS (P=0.013) effects on the functional gene categories obtained from MG-RAST (Subsystems level 2 categories) were observed (data not shown). Moreover, consistent and significant changes in the relative abundance of 33 functional gene categories were detected in the rhizosphere soil of YCS symptomatic vs. asymptomatic

plants (Table 6). The software STAMP (http://kiwi.cs.dal.ca/Software/STAMP) was used to determine which functional gene categories significantly (P<0.05) differed in relative abundance between YCS symptomatic and asymptomatic plants after Benjamini-Hochberg FDR correction.

Table 8: Overview of the functional gene categories (Subsystems level 2) showing consistent and significant changes in relative abundance (RA) between rhizosphere soils from YCS symptomatic and asymptomatic plants.

Higher RA (%) in rhizosphere of YCS symptomatic plants	Lower RA (%) in rhizosphere of YCS symptomatic plants
CRISPRs (prokaryotic immune system)	DNA recombination
CRISPs & associated hypotheticals	DNA replication
Bacterial cytostatics, differentiation factors and antibiotics	RNA processing & modification
	ATP synthases
Anaerobic degradation of aromatic compounds	Lipoic acid
Fatty acid metabolism	Plant alkaloids
Fermentation	Purines
Coenzyme F420 (involved in redox reactions in methanogens)	Pyrimidines
	Pyrimidine biosynthesis-related
Branched-chain amino acids	Protein biosynthesis
Biotin	Protein translocation across cytoplasmic membrane
Biosynthesis of galactoglycans and related lipopolysaccharides	Gene Transfer Agent
	Resistance to antibiotics and toxic compounds
Cell wall *Mycobacteria*	DNA polymerase III epsilon cluster
Cytochrome biogenesis	Acid stress
Putative associate of RNA polymerase sigma-54 factor rpoN	Hypothetical related to dihydroorate dehydrogenase
	Shiga toxin cluster
Peripheral pathways for catabolism of aromatic compounds	Signal transduction in eukaryotes
	TldD cluster (proteolytic complex)

We found a higher relative abundance of genes involved in prokaryotic immune system in the rhizosphere of YCS symptomatic plants (CRISPs and CRISPRs; Table 8), which might indicate that the prokaryotic community in the rhizosphere is defending itself against a potential pathogen (Skennerton et al. 2013; Tyson and Banfield, 2008). Moreover, we find a higher potential for microbial functions involved in anaerobic degradation of aromatic compounds, fatty acid metabolism, fermentation etc. (Table 8). In contrast, a lower abundance of genes involved in DNA, RNA and protein processing were detected in the rhizosphere of YCS symptomatic plants, as well as a lower resistance to antibiotics and toxic compounds (Table 8).

All rhizosphere functional gene categories significantly affected by YCS (Table 8) revealed complex associations with a range of bacteria, fungi and other eukaryotes that contributed most to the observed difference in community composition with YCS (Figure 18). This indicates that many organisms contribute to the changes in the abundance of the functional gene categories observed in the rhizosphere soil of YCS symptomatic plants, and confirms the complexity of shifts in the microbial community related to YCS and the associated challenges in detecting a clear YCS signal in the microbial community linked to YCS.

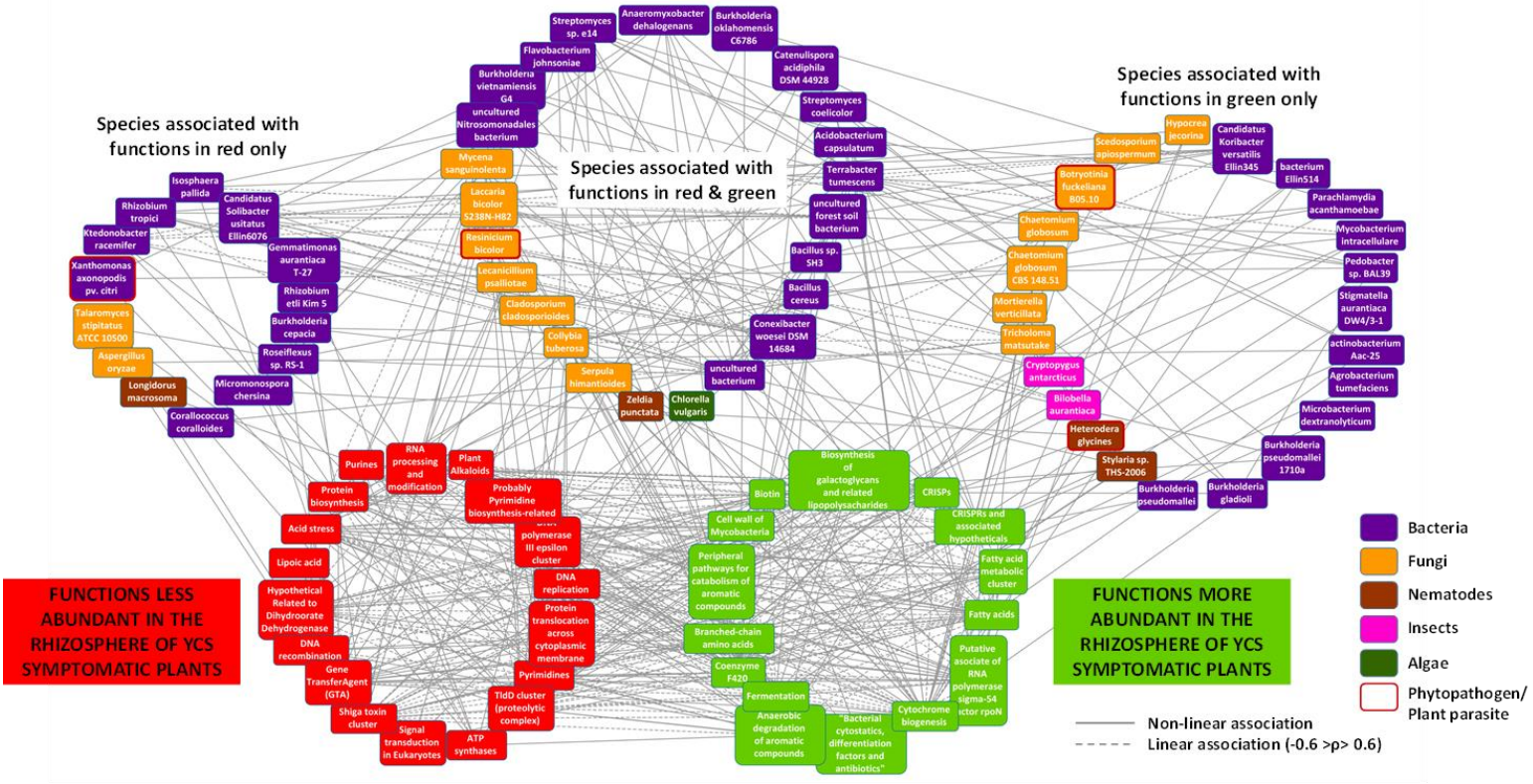

Figure 18: Co-occurrence network showing the 1st-neighbour associations between functional gene categories that were significantly less (red) or more (green) abundant in the rhizosphere of YCS symptomatic plants of three varieties, and the top 100 species contributing to the difference in the rhizosphere soil community composition observed between YCS symptomatic and asymptomatic plants (as indicated by SIMPER). The network was constructed using the maximal information coefficient and visualized using Cytoscape v 3.2.0.

6.1.4. Small RNA sequencing approach

Based on electron microscopy analysis of YCS symptomatic leaf and stalk samples by us and by Davey Olsen's team, no particular type of virus appeared associated with YCS. Therefore, enrichment of a particular type of virus potentially involved in YCS and subsequent sequence analysis to characterise this virus is not applicable. Instead, a small RNA sequencing approach was selected to study the potential involvement of a virus. As reported in literature (e.g. Kreuze et al., 2009; Wu et al., 2015), there are several reasons for enriching and sequencing small RNAs for the identification of plant viruses and viroids over other enrichment strategies. First, replication of RNA and DNA viruses and replication of subviral agents such as viroids and satellite RNAs in plants all induce accumulation of the pathogen-specific small interfering RNAs, which can represent up to 30% of total small RNAs sequenced from diseased plants. Second, unlike dsRNA and virus-like particle preparations, all replicating viruses and viroids in a diseased plant can be detected in principle from deep sequencing of a single library of small RNAs. Third, viral and subviral small interfering RNAs are the products of an active host immune response to infection. Finally, data mining of small RNA libraries might lead to the discovery of novel pathogens that exhibit no sequence similarity detectable by current available bioinformatics tools. In conclusion, sequencing small RNAs is a cost-effective approach for plant virus discovery. Moreover, such small RNA sequencing approaches have reportedly been successful in detecting novel plant viruses for a wide range of plant species, including sugarcane (see overview in Wu et al., 2015 and Candresse et al., 2014 for sugarcane).

Table 9: Overview of the leaf samples analysed by small RNA sequencing.

Sample_ID	Sample location	Symptoms	No. of raw reads	No. of assembled contigs
M2AAL6 (leaf 6)	Mackay, June 2015	YCS symptomatic	9.040.740	33 (225-1414 nt)
I10CL2 (leaf 2)	Ingham, June 2015	YCS symptomatic	10.065.145	39 (228-613 nt)
B10 EM (leaf 6)	Ayr, June 2015	Yellow Leaf Syndrome	9.799.257	32 (233-647 nt)

No significant homology to viral sequences was found for any of the contigs assembled by the Trinity pipeline. Since recognising viral or subviral pathogens in diseased tissue that share no significant homology with known pathogens is a common challenge in pathogen discovery by deep sequencing approaches, a homology-independent approach for discovering viroids was developed recently (Wu et al., 2012). Viroids are a distinct class of free circular RNA subviral pathogens that do not encode a protein and are known to infect plants only. A unique computational algorithm, progressive filtering of overlapping small RNAs (PFOR) was developed that retains viroid-specific small interfering RNAs for assembly (Wu et al., 2012). We applied the PFOR algorithm on our small RNA sequence data to investigate the potential presence of novel viroids but did not detect a viroid signal.

Around 20% of the assembled contigs showed significant homology to mRNA of proteins involved in photosynthesis (PS I and PS II chlorophyll apoproteins) and could therefore be involved in regulation of photosynthesis. Other contigs showed significant homology to expressed sequence tags (ESTs) expressed in a wide range of plants, including sugarcane and sorghum, in response to both biotic and abiotic stress. Biotic stressors included rust, red rot and anthracnose, while abiotic stressors triggering the expression of the homologous sequences included water stress (both drought and high rainfall or flooding), oxidative stress, heat stress, light stress, salt stress, and nitrogen or iron deficiencies. Matching the transcripts to the sugarcane transcriptome obtained by Frikkie Botha's team might enable to identify which transcripts were present in the small RNA libraries.

Among the assembled contigs, a few fungal signals were detected that need to be explored further. Fungal mRNAs homologous to mRNA of 3 fungal *Colletotrichum* species were detected in the YCS symptomatic leaves. The sequences were 90-96% identical over 86-97% of the length of the alignment, indicating that RNA from a closely related fungal species might be present. The *Colletotrichum* genus contains over 600 species that attack over 3200 plant species (both monocot and dicots). They are known to cause anthracnose, which has symptoms unlike YCS but could cause the brownish spots that become visible on YCS-affected leaves as a secondary effect. In the leaf sample from Ayr, symptomatic for Yellow Leaf Syndrome, another fungal signal was detected. In this sample, an assembled contig homologous to mRNA of the fungus *Leptoshaeria maculans* (98% coverage, 94% sequence identity) was observed.

We further conducted small RNA sequencing to study the involvement of viruses in YCS development. Small RNA-Seq was conducted on the 35 FV10 leaf samples including green lamina tissue, yellow lamina tissue, midribs and dewlaps. Per sample, between 8.7 and 14.8 million reads were obtained. The raw data were analysed as depicted in Figure 19.

Further analysis of the small RNA-Seq data was conducted as per pathway 3 in Figure 19. This resulted in the final assembly of 28 FV10 viral scaffolds (73-5322 nt length) showing the highest homology to Banana streak CA virus and Sugarcane bacilliform virus. Together, the viral scaffolds covered over 90% of the reference genomes. To obtain the full genome and identify the virus (or possible mixed badnavirus infection) in FV10

leaves, additional experiments are required such as e.g. performing rolling circle amplification reactions using degenerate primers or primers developed based on the assembled viral scaffolds.

Alignment of the FV10 small RNA reads to the assembled viral scaffolds (Figure 20) demonstrated that the active antiviral response detected in FV10 sugarcane leaves was present in all studied leaf sections and in both control and YCS symptomatic leaves. Only in dewlaps, the percentage of viral small RNA was significantly higher in YCS symptomatic leaves compared to the control leaves (Figure 20). Differences in the relative percentages of viral small RNAs between samples may reflect different proportions of infected cells in the analysed leaf tissues or different stages of virus replication (with different copy numbers of the viral DNA per cell).

Given that Sugarcane bacilliform virus and it's closely related variant Banana streak CA virus are commonly detected in commercial sugarcane varieties and the antiviral response to the (potential mixed) badnavirus infection was present in both asymptomatic and YCS symptomatic leaves, the detected virus(es) are likely not the single cause of YCS. However, information on the role of biotic stress caused by bacteria, fungi, nematodes, insects and other viruses in badnavirus disease development and severity is currently lacking. Therefore, without further research, we cannot rule out the possible involvement of the virus in YCS involvement in combination with other biological agent(s). However, if the detected badnaviral infection is involved in YCS development, we expected the virus to be present in YCS symptomatic leaf samples collected previously in June 2015 in Ingham and Mackay.

Alignment of the small RNA reads of 3 leaf samples collected in Ingham, Mackay and Ayr in June 2015 against the 28 final FV10 viral scaffolds indicated that a highly identical virus was present in the YCS symptomatic leaf collected in Ingham, but not in the YCS symptomatic leaf collected in Mackay (Table 5). Moreover, viral sequences identical to the FV10 viral scaffolds were detected in a leaf collected in Ayr that showed no YCS symptoms but instead had a yellow midrib (Table 10).

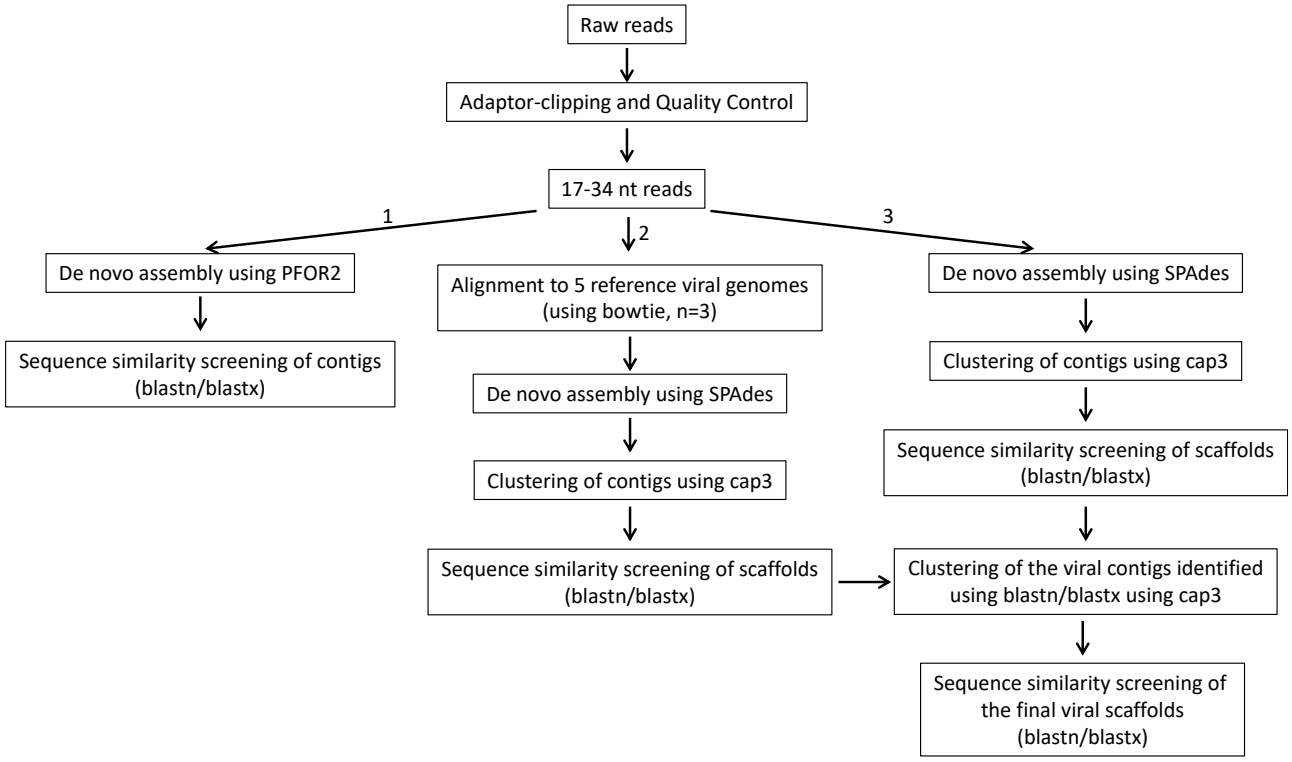

Figure 19: Simplified overview of the workflow applied for analysis of the small RNA reads.

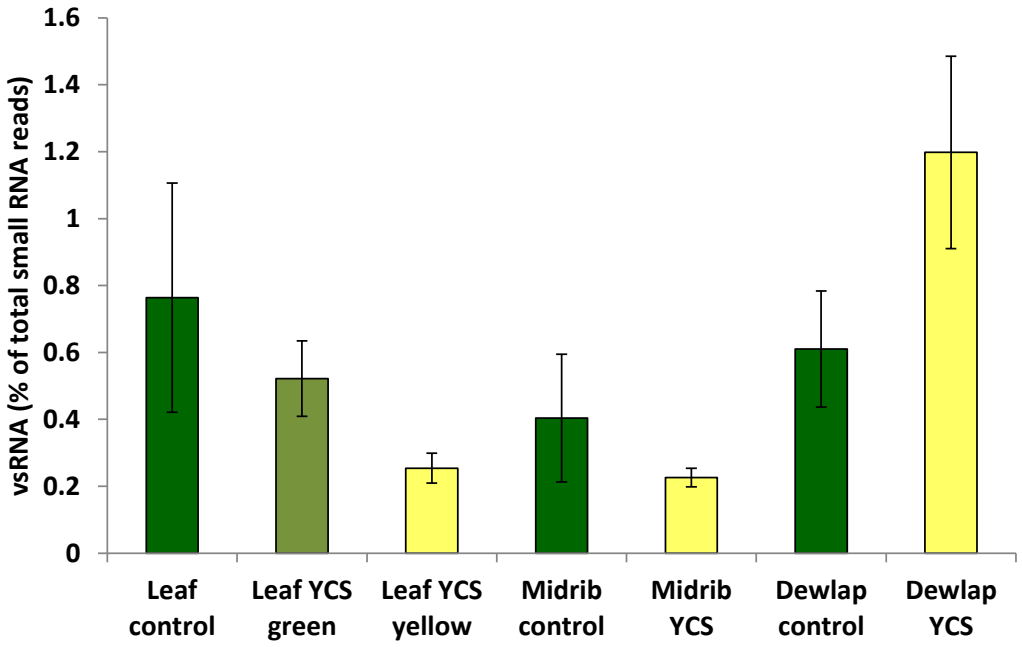

Figure 20: Percentage of viral small RNA reads (vsRNA) detected in the different FV10 sample types, calculated as the percentage of small RNA reads that aligned with zero mismatches to the final 28 FV10 viral scaffolds.

Table 10: Percentage of viral small RNA reads (vsRNA) detected in June 2015 leaf samples, calculated as the percentage of small RNA reads that aligned with zero mismatches to the final 28 FV10 viral scaffolds

Sample	Leaf symptoms	Sample location	%vsRNA
M2AAL6	YCS	Mackay	0.0003
I10CL2	YCS	Ingham	0.07
B10EM	YLS (yellow midrib)	Ayr	0.06

In conclusion, small RNA-Seq of FV10 samples resulted in the detection of an active sugarcane defence response against a pararetrovirus with a genome that is most similar to the badnaviruses Banana streak CA virus and Sugarcane bacilliform virus. Further research is required to reveal whether a novel virus or variants of the above badnaviruses were present. However, given (i) the presence of the antiviral response in both asymptomatic and YCS symptomatic FV10 leaves, (ii) the common detection of similar viruses in commercial sugarcane, (iii) the nature of the YCS symptoms being quite distinct from symptoms caused by such badnaviruses (chlorotic mottle or necrotic streaks, deformation of leaves, stunting of plants), and (iv) the absence of the virus in a YCS symptomatic leaf collected in Mackay, we postulate that the pararetrovirus(es) detected in FV10 sugarcane leaves are likely not involved in YCS development, unless YCS is the result of multiple pathogens acting simultaneously.

6.2. Physical, chemical and biological analysis of soil samples

Across the 8 cane blocks sampled in Ingham in November 2014, a trend of lower total carbon and nitrogen in YCS symptomatic vs. asymptomatic blocks was observed for the two of the three cane varieties analysed (Figure 21 A-B). No consistent differences were observed in available carbon or nitrogen between symptomatic and asymptomatic blocks (Figure 21 C-D). Soil pH was highly variable within some blocks, but didn't show a trend with presence of YCS symptoms (Figure 22).

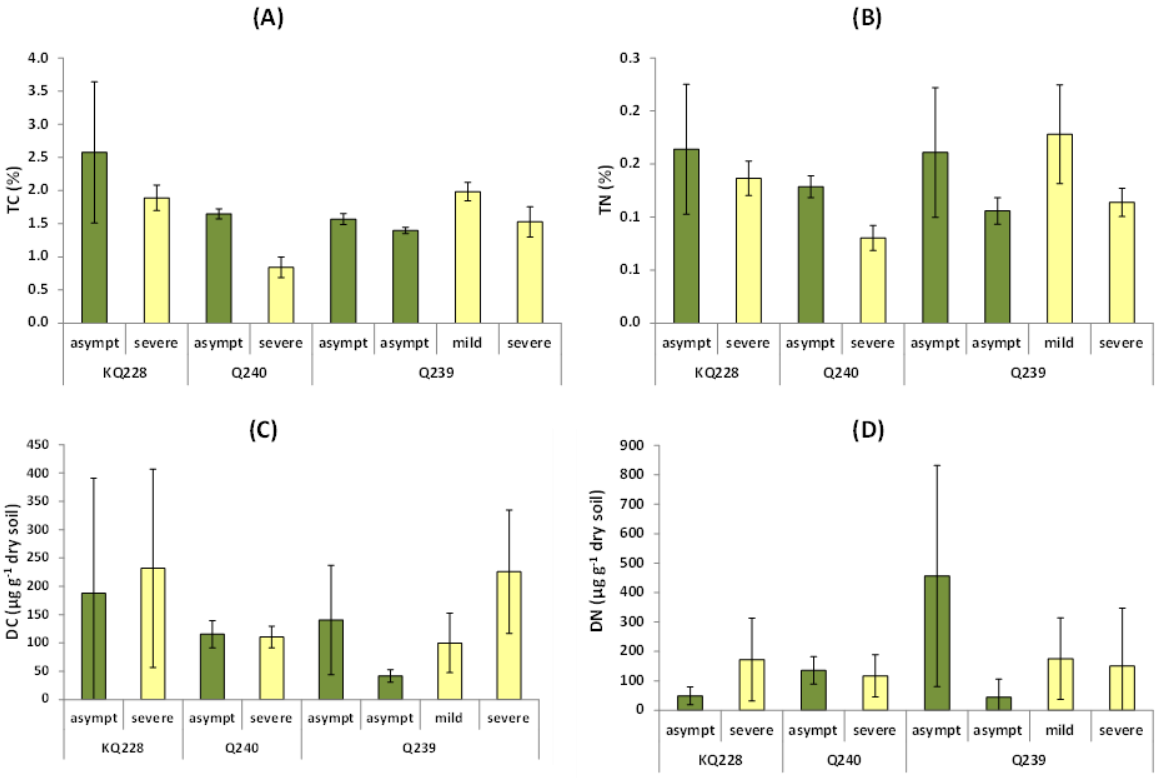

Figure 21: Total carbon (A), total nitrogen (B), available carbon (C) and available nitrogen (D) measured in soil samples collected from 8 cane fields in Ingham, November 2014. Bars represent the average, and error bars the standard deviation calculated for soil samples collected in each block (n=19). For each block, the sugarcane variety (KQ228, Q239 or Q240) and YCS symptoms (asymptomatic, mild YCS or severe YCS symptoms) are indicated.

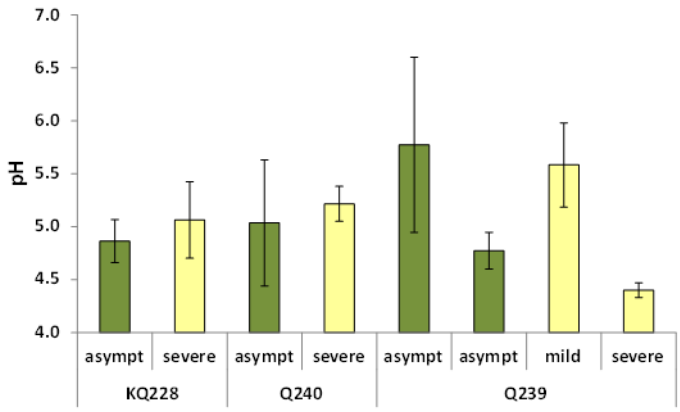

Figure 22: Soil pH measured in samples collected from 8 cane fields in Ingham, November 2014. Bars represent the average, and error bars the standard deviation calculated for soil samples collected in each block (n=19). For each block, the sugarcane variety (KQ228, Q239 or Q240) and YCS symptoms (asymptomatic, mild YCS or severe YCS symptoms) are indicated.

Soil microbial enzyme activity was analysed by performing assays for six extracellular enzymes involved in either C or P cycling (Figure 23). Enzyme activities differed between some of the blocks, but no consistent trend was observed between asymptomatic and symptomatic blocks for the three studied sugarcane varieties.

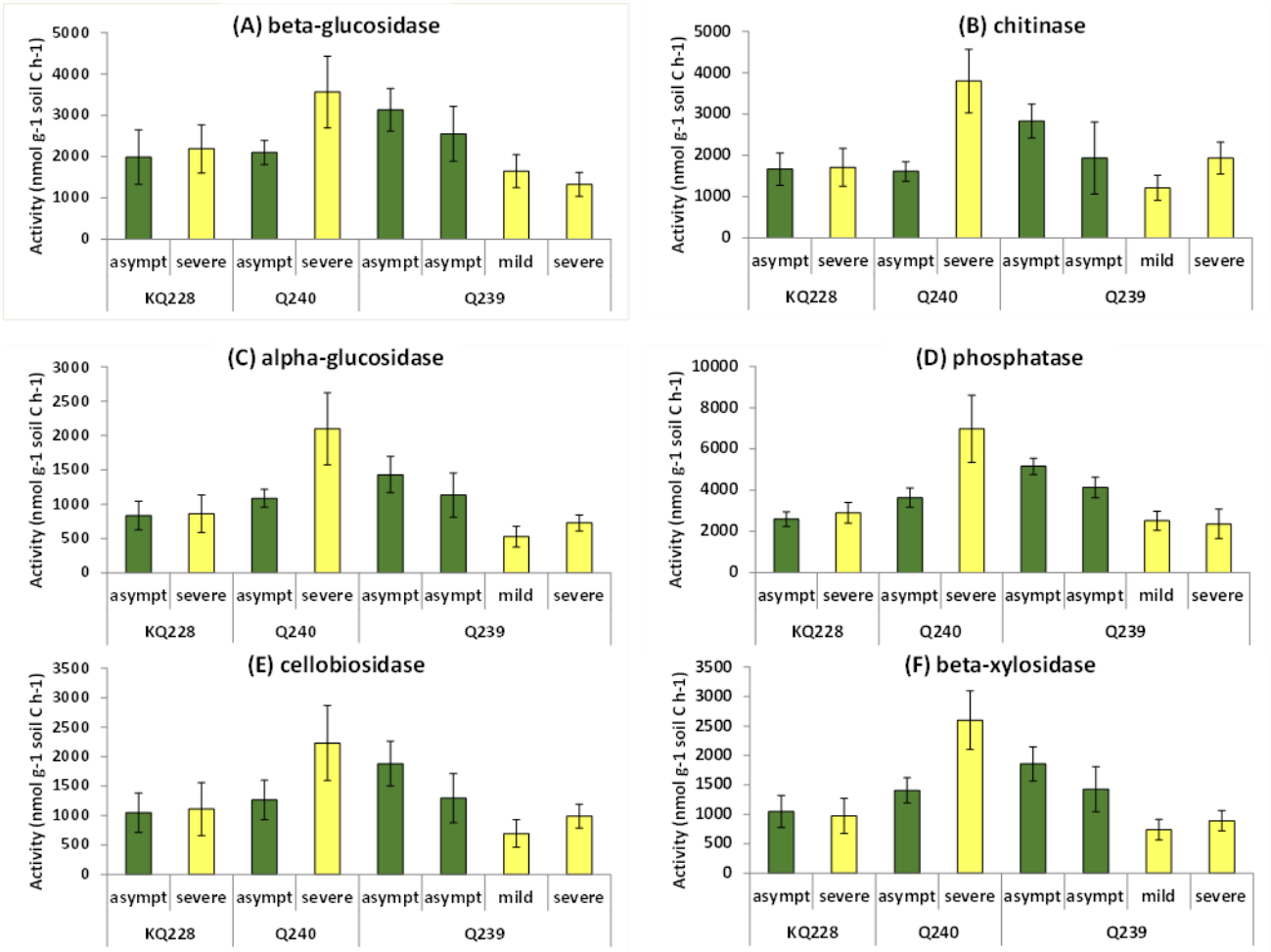

Figure 23: Microbial enzyme activity measured in soil samples collected from 8 cane fields in Ingham, November 2014. Results for beta-glucosidase (A), chitinase (B), alpha-glucosidase (C), phosphatase (D), cellobiosidase (E) and beta-xylosidase (F) are shown. Bars represent the average, and error bars the standard deviation calculated for soil samples collected in each block (n=19). Each soil sample was analysed in quadruplicate. For each block, the sugarcane variety (KQ228, Q239 or Q240) and YCS symptoms (asymptomatic, mild YCS or severe YCS symptoms) are indicated.

Soil respiration was measured after adding water or 7 substrates to the soils (Figure 24). Similar to the microbial enzyme activity results, no trend was observed for changes in substrate-induced soil respiration with YCS incidence.

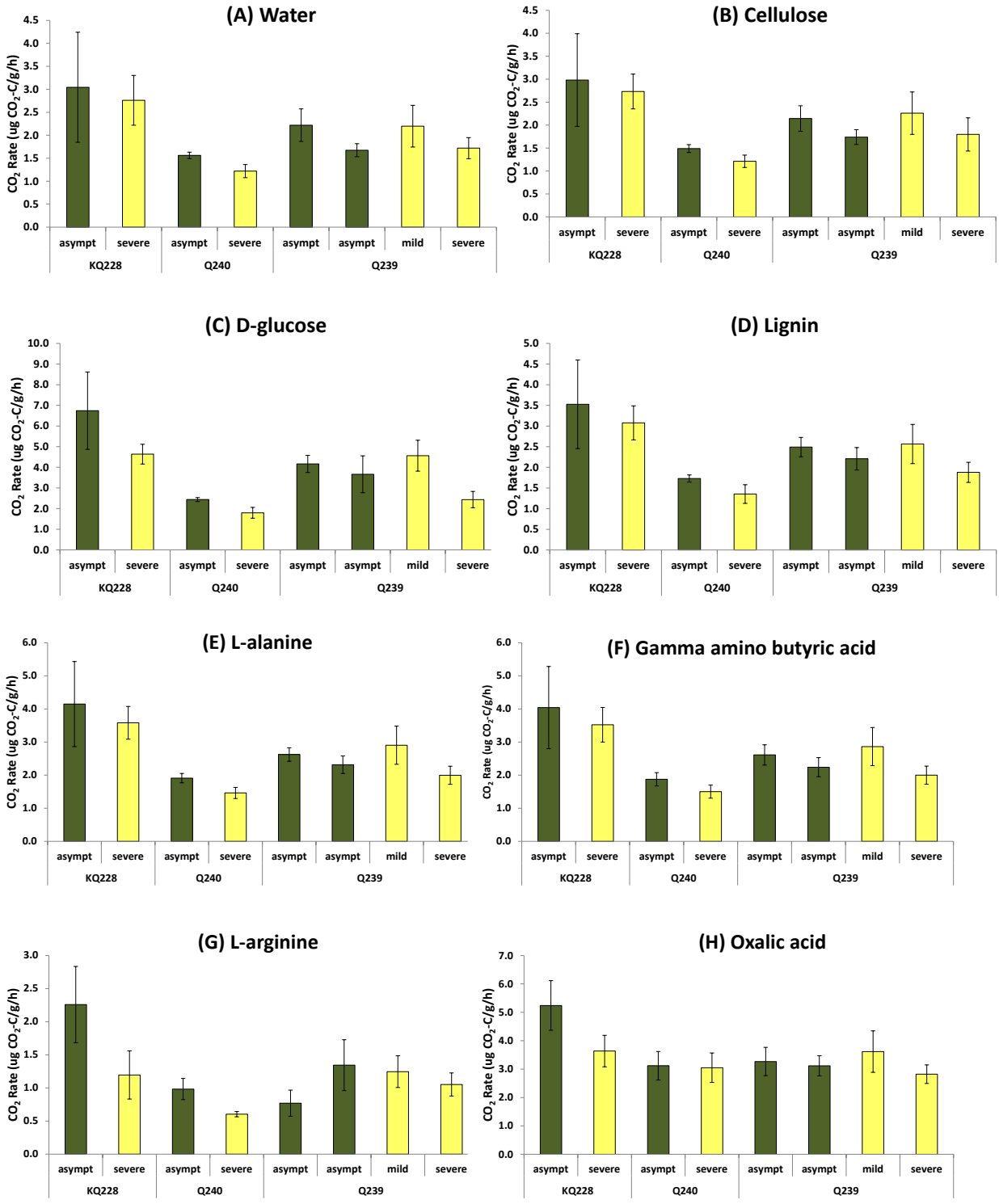

Figure 24: Soil respiration data for soil samples collected from 8 cane fields in Ingham, November 2014. Results after adding water (A), cellulose (B), D-glucose (C), lignin (D), L-alanine (E), Gamma amino butyric acid (F), L-arginine (G) and oxalic acid (H) are shown. Bars represent the average, and error bars the standard deviation calculated for soil samples collected in each block (n=19). Each soil sample was analysed in duplicate. For each block, the sugarcane variety (KQ228, Q239 or Q240) and YCS symptoms (asymptomatic, mild YCS or severe YCS symptoms) are indicated.

Based on these limited soil health measurements performed on soils collected from 8 cane fields in Ingham in November 2014, no clear correlation was found between YCS development and soil carbon and nitrogen

content, soil microbial enzyme activity or substrate-induced soil respiration. However, two of the four asymptomatic blocks have shown YCS symptoms since sample collection in November 2014.

6.3. Impact of YCS on plant defence and stress response

The abundance of both pathogenesis- (Figure 25a) and stress-related gene transcripts (Figure 25b) was higher in YCS symptomatic compared to asymptomatic leaves, demonstrating a host plant response to YCS. Moreover, symptomatic leaves contained more salicylic acid than asymptomatic leaves (Figure 26), demonstrating the potential use of this compound as a biomarker for YCS detection.

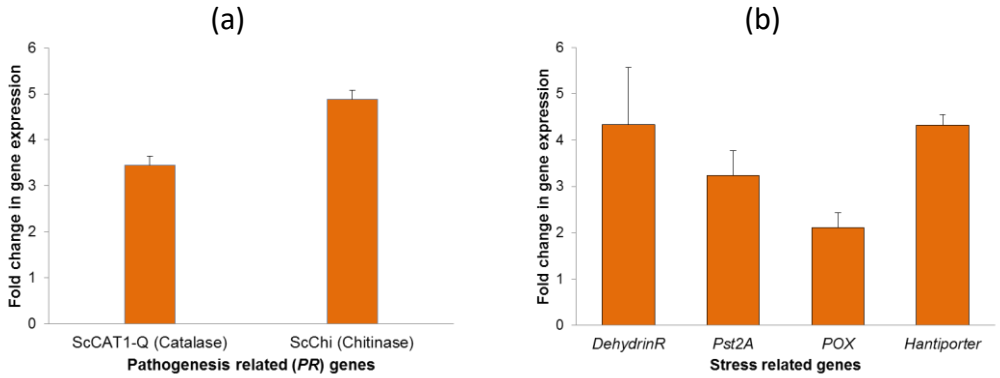

Figure 25: Upregulation of pathogenesis- and stress-related genes in symptomatic compared to asymptomatic leaf samples collected during the first two sampling periods, as determined by qRT-PCR.

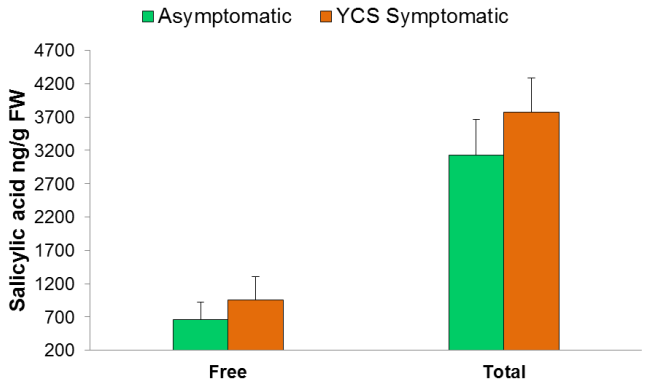

Figure 26: Salicylic acid content measured in asymptomatic and YCS symptomatic leaves.

Changes in the expression levels of these three genes in the tested sugarcane leaves were highly variable (Figure 27). Leaf 2 of variety Q240 overall showed the largest differences in gene expression between YCS symptomatic and asymptomatic plants. In Q240 leaves, catalase and dehydrin gene expression appeared upregulated in symptomatic plants (Figure 27 A-B), whereas expression of the Pst2A sugar transporter gene appeared downregulated in symptomatic plants (Figure 27 C). However, no strong or consistent effects were observed across the three sugarcane varieties.

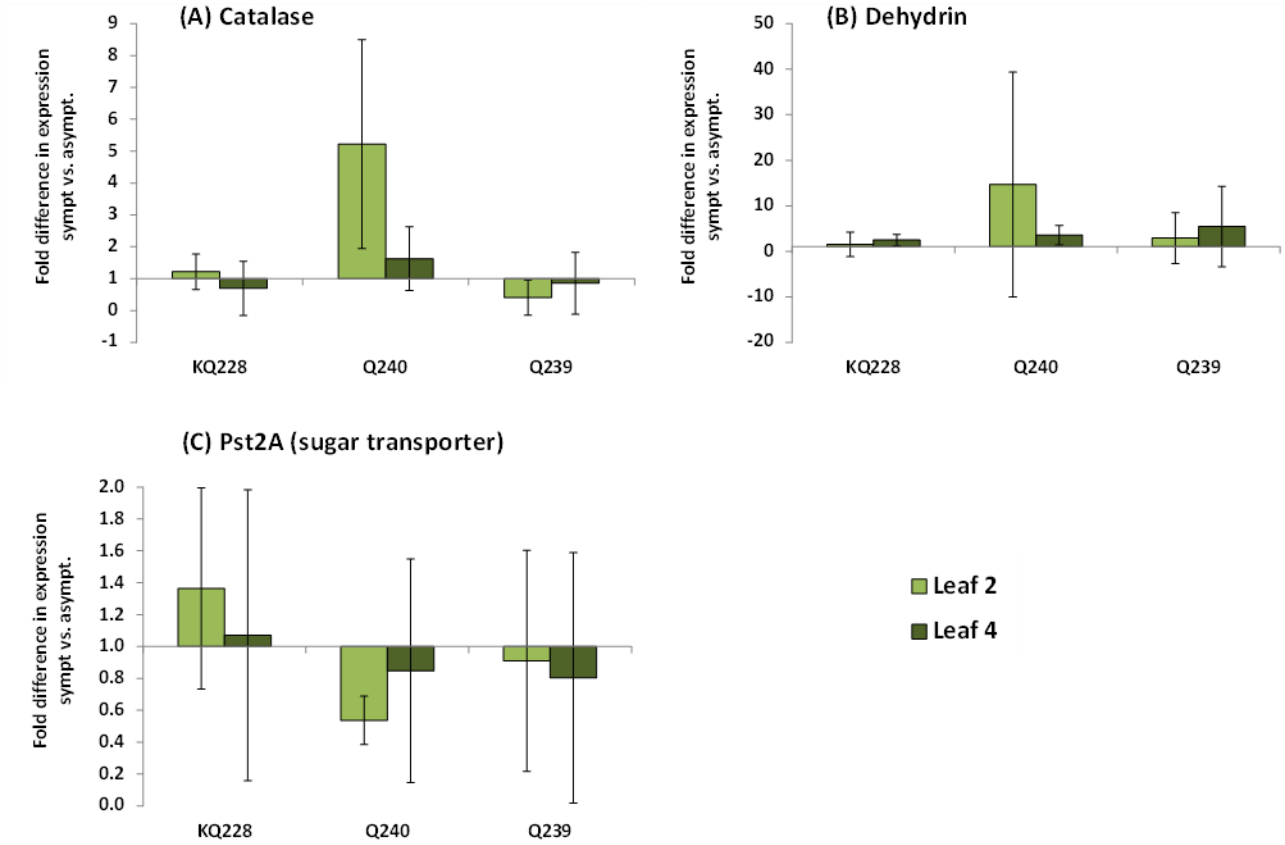

Figure 27: Difference in gene expression of catalase (A), dehydrin (B) and Pst2A (C) between symptomatic and asymptomatic leaves (leaf 2 and leaf 4), calculated after normalisation against GADPH expression levels, for three plants per block sampled in Ingham in November 2014. Bars represent averages and error bars standard deviations in the fold difference in gene expression in symptomatic vs. asymptomatic plants for each of the three sugarcane varieties (KQ228, Q239 and Q240). Positive values indicate upregulation of gene expression in YCS symptomatic compared to asymptomatic plants whereas negative values indicate downregulation.

Salicylic acid concentrations were higher in leaf 2 of symptomatic vs. asymptomatic plants of varieties KQ228 and Q240, but not Q239 (Figure 28 A). In leaf 4, no effect of YCS on salicylic acid content was detected (Figure 28 B).

Initially, a significant and consistent increase in salicylic acid content was detected in symptomatic compared to asymptomatic leaves (data not shown). This was based on results obtained for symptomatic and asymptomatic plants collected 40 km apart (due to widespread YCS incidence at the time of sample collection, during mid-2014). Additional results were based on samples collected in Ingham in November 2014 (asymptomatic and symptomatic blocks of three varieties sampled in close proximity, grown in the same soil type and under similar management), and thus show the importance of collecting adequate symptomatic and asymptomatic samples for comparison to obtain valuable results.

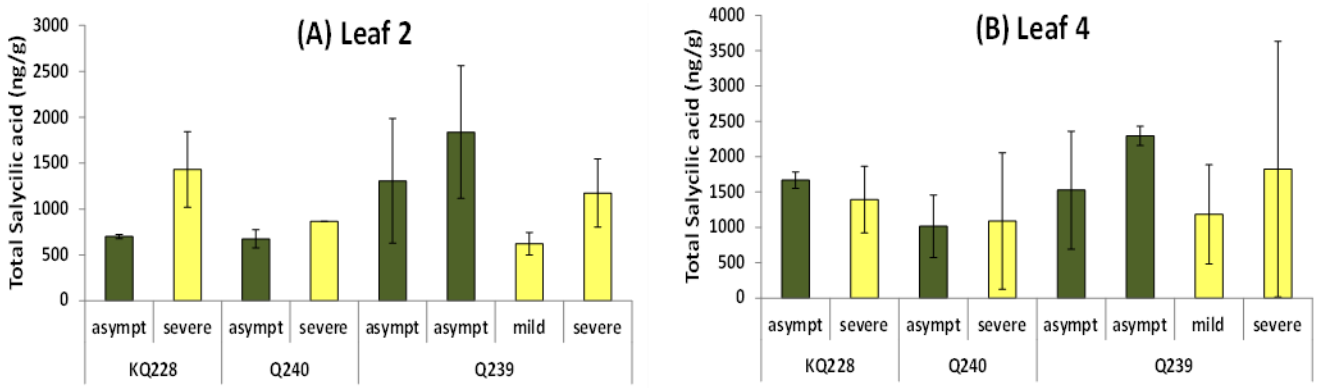

Figure 28: Total salicylic acid content in leaf 2 (A) and leaf 4 (B) for three plants per block sampled in Ingham in November 2014. Bars represent averages and error bars standard deviations for three sugarcane varieties (KQ228, Q239 and Q240). For each block, the sugarcane variety (KQ228, Q239 or Q240) and YCS symptoms (asymptomatic, mild YCS or severe YCS symptoms) are indicated.

6.4. Microscopic observations

6.4.1. Electron microscopy

Electron microscopy showed diverse assemblages of various microbes on YCS symptomatic leaves as compared to healthy leaves.

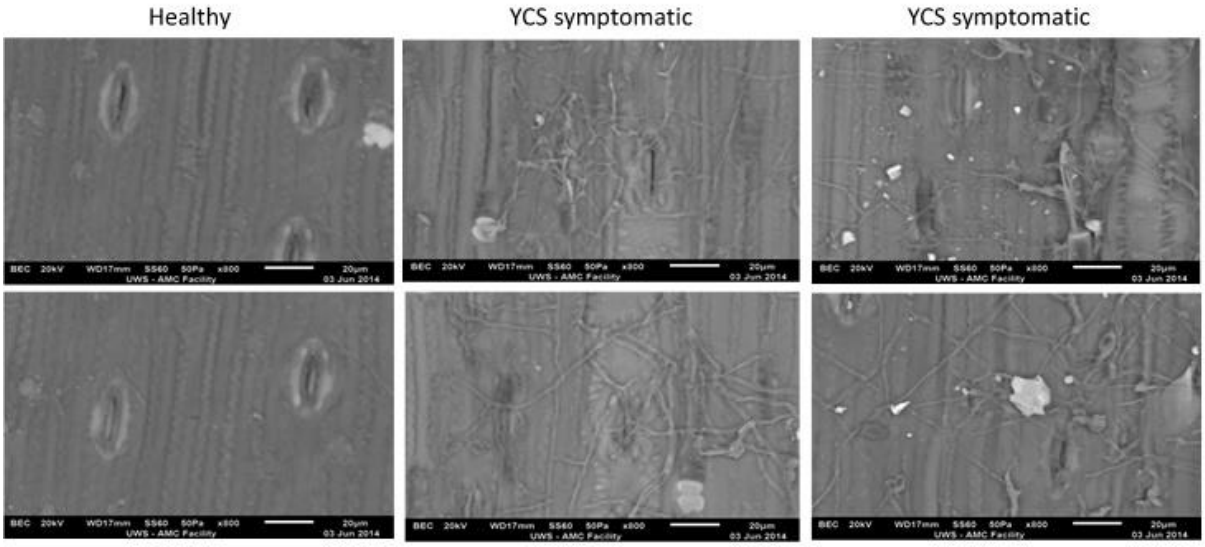

Figure 29: Electron microscopy images of healthy and YCS symptomatic sugarcane leaves; YCS symptomatic leaves show various microbial interactions.

6.4.2. Optical microscopy

Microscopic images revealed no abnormal symptoms nor any blockage or crushing of vessels in YCS symptomatic leaf and root samples collected during the initial sampling rounds (Figure 30). These samples consisted of Leaves 3 and 5 and roots from three asymptomatic plants and three symptomatic plants.

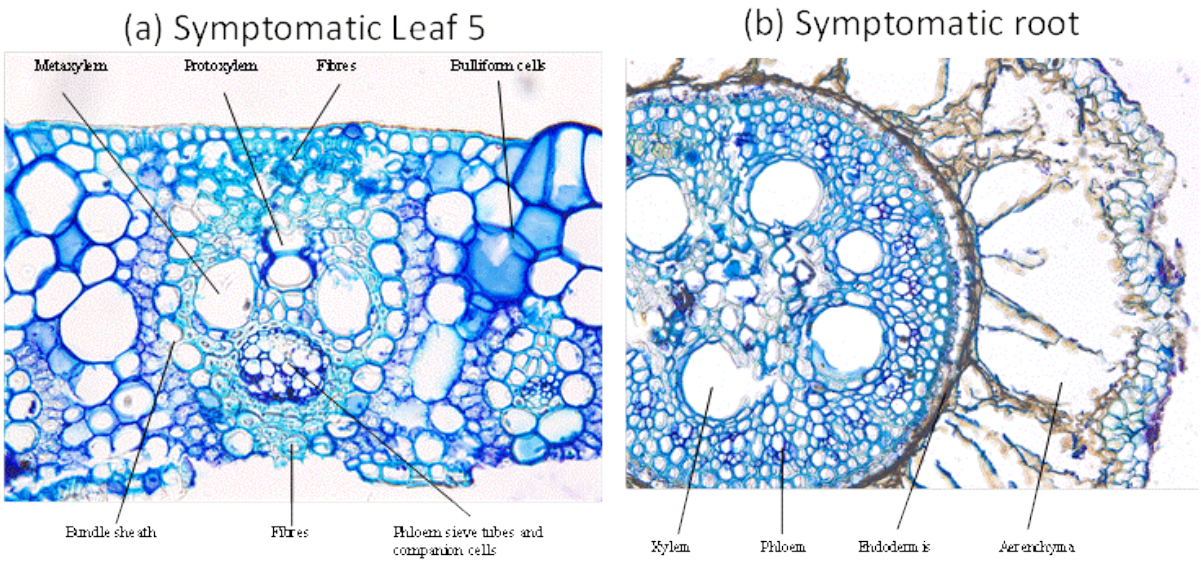

Figure 30: Typical microscopic images obtained for YCS symptomatic leaves (a) and roots (b).

6.5. Phytoplasma detection

The initial PCR reaction provided bands at the correct, estimated size of 1.8kb so the secondary, nested PCR reaction was performed providing bands at the expected size of 900bp. The darkest bands from the initial PCR and from the nested PCR were sequenced to confirm if DNA from Phytoplasma was detected. Here, only 1 out of 5 first-step PCR products were derived from DNA of Phytoplasma (Figure 31). From the nested PCR reaction, however, all sequenced PCR products were confirmed to be amplified from Phytoplasma DNA (Figure 32). From those Phytoplasma-positive sequences, a phylogenetic tree was formed (Figure 33). Further information regarding field site identification and sugarcane variety for samples are provided in Table 11.

Table 11: Sample identification used in Figures 31 and 32

Location	Field site ID	Sugarcane variety	YCS symptoms
Ingham	I9	MQ239	Asymptomatic
	I10	MQ239	Symptomatic
Mackay	M3	Q242	Symptomatic
	M6	Q242	Asymptomatic
	M4	Q240	Symptomatic
	M5	Q240	Asymptomatic

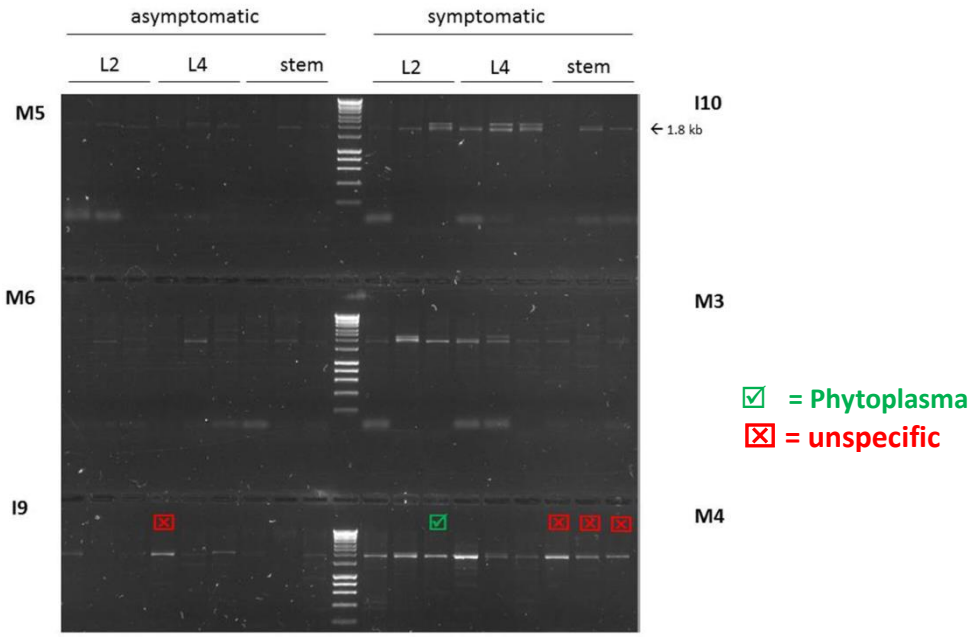

Figure 31: First-step of the nested PCR process to identify Phytoplasma from DNA extracted from sugarcane plant samples collected from Ingham and Mackay in 2015. L2 = Leaf2, L4 = Leaf4, M3-6 represent plants samples from Mackay, I9-I10 represent plants samples from Ingham. Samples from 3 plants were initially tested from each location, asymptomatic plants are represented on the left, symptomatic plants on the right side of the figure. Where indicated (by ticks or crosses), PCR bands were excised and sequenced to determine if DNA was derived from Phytoplasma.

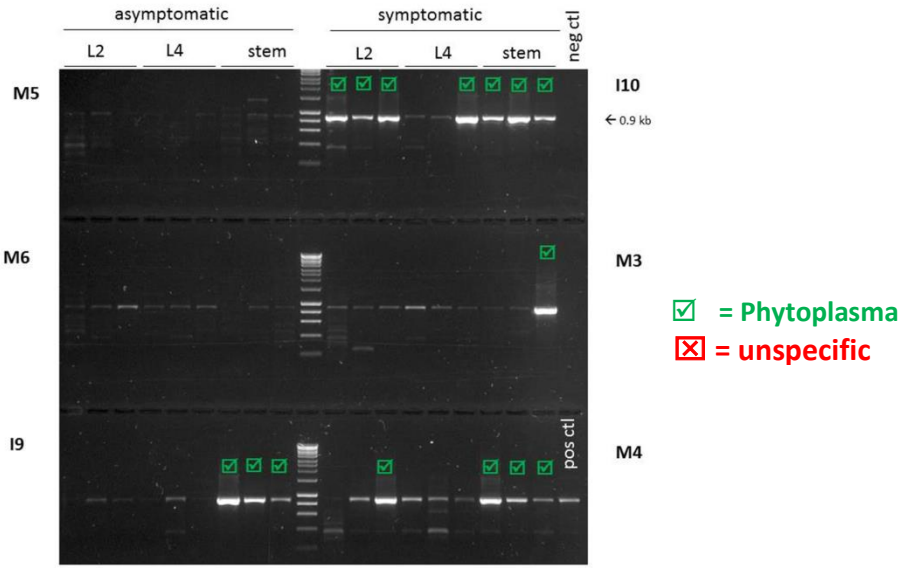

Figure 32: Second phase of the nested PCR process to identify Phytoplasma from DNA extracted from sugarcane plant samples collected from Ingham and Mackay in 2015. L2 = Leaf2, L4 = Leaf4, M3-6 represent plants samples from Mackay, I9-10 represent plants samples from Ingham. Samples from 3 plants were initially tested from each location, asymptomatic plants are represented on the left, symptomatic plants on the right side of the figure. PCR products from the initial PCR cycle were diluted 30-fold and used as template DNA for nested PCR. Where indicated (by ticks), PCR bands were excised and sequenced to determine if DNA was derived from Phytoplasma. Ticks indicated on Figure 2 indicate the DNA sequences positive for Phytoplasma.

Figure 33: : Phylogenetic tree of Phytoplasma samples. Molecular phylogenetic analysis by the maximum likelihood method (19 required sequences + 2 E. coli + 18 Phytoplasma sequences). The evolutionary history was inferred by the Maximum Likelihood method based on the Tamura-Nei model. The tree with the highest likelihood is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 39 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 194 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

However, samples from other time-points have returned inconsistent results. Samples collaboratively collected in February 2016 (FV10) from Burdekin (Home Hill) are an example of this inconsistency. Both YCS symptomatic and asymptomatic plants of variety Q240 were sampled. Leaf 6 showed clear YCS symptoms and was selected for sampling. In the laboratory, each leaf sample was photographed and the following sections were subsampled for grinding and subsequent analyses: leaf dewlap, leaf midrib, green leaf lamina, yellow leaf lamina. After grinding the leaf sections, samples were pooled to produce replicates to be shared by all collaborators. We used DNA extracted from these sub-samples to determine if Phytoplasma were present using the nested PCR approach described above (Figure 34 and 35). PCR results showed only a few samples detected as positive, and all others were negative.

As a result of these negative results, all other primer sets were trialled on the same samples, however none of the PCR products produced were specific for phytoplasma's (determined after DNA sequencing) apart from the positive controls. These negative results may due to (i) no phytoplasma present in these samples; or (ii) the primers were not specific enough to detect phytoplasma, despite published literature saying otherwise.

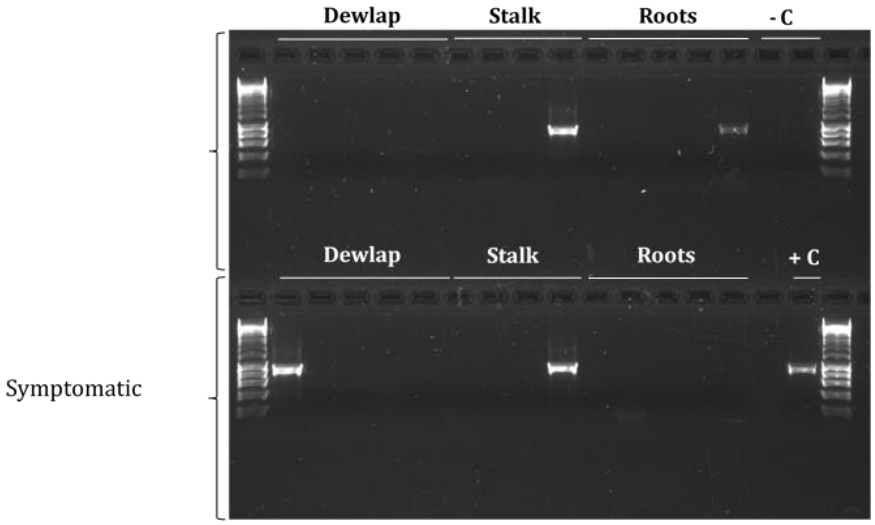

Figure 34: PCR from FV10 samples collected from Burdekin, February 2016. Sub-samples from leaves were dissected, ground in liquid nitrogen and DNA extracted from dewlaps, stalks and roots. Nested PCR was performed as described above.

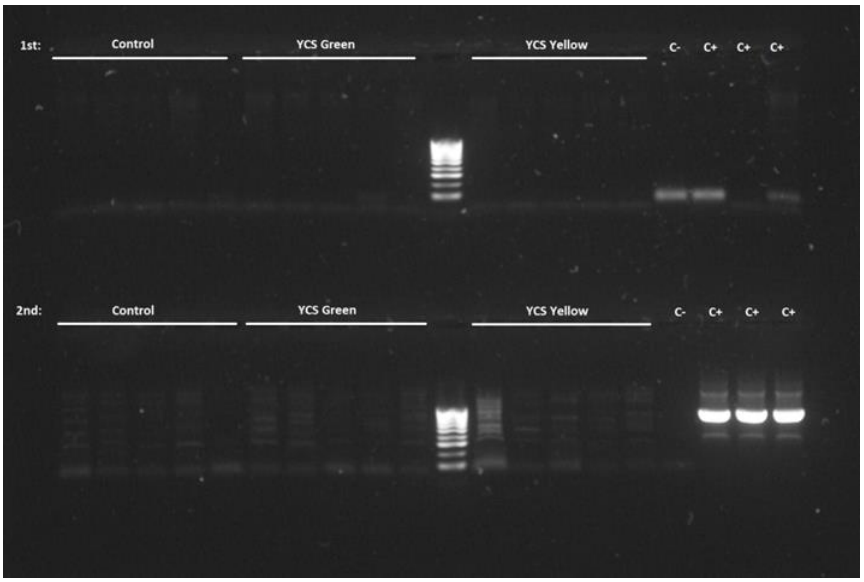

Figure 35: Nested PCR from FV10 samples collected from Burdekin, February 2016, with detection of leaf samples.

7. CONCLUSIONS

Our key findings include:

- The bacterial and fungal community composition of sugarcane leaves, stalks, roots and soil samples significantly differ from each other.
- The core and hub microbiota of sugarcane has been identified and reported at a national scale. This provides an important resource that can be harnessed in the future for sustainable farming using either microbial assisted breeding or *in situ* microbiome engineering.
- Within a growing region, each sugarcane variety harbours a unique bacterial and fungal community composition that changes over time during the growing season.
- Within each variety, the bacterial and fungal community composition of leaves, stalks, roots and rhizosphere soils differ between YCS symptomatic and asymptomatic plants.
- A limited number of individual bacterial and fungal OTUs significantly increased or decreased in relative abundance with YCS incidence in all studied varieties and at each studied location.
- Deep sequencing of rhizosphere soil samples revealed the presence of numerous phloem and xylem-restricted pathogens including *Curtobacterium* spp and Phytoplasma. The presence of Phytoplasma in some symptomatic plants were also found in PCR-based approaches but these results were inconsistent.
- The functional potential of the microbial rhizosphere soil community significantly differed for YCS symptomatic vs. asymptomatic plants.
- Genomes of two related viruses were detected in sRNA sequencing but were present in both YCS symptomatic and asymptomatic samples.
- All rhizosphere functional gene categories significantly affected by YCS revealed complex associations with a range of bacteria, fungi and other eukaryotes.
- *Curtobacterium* remains a candidate of interest because it has shown an increased relative abundance with YCS, predominantly in the leaf midribs and in dewlaps, which might correlate with the “blockage hypothesis”.
- Shifts in the microbiome (bacteria and fungi) between asymptomatic and YCS symptomatic leaves were more pronounced for dewlaps than for midribs and lamina samples.
- Bacteria have shown a dynamic response to YCS within sugarcane leaf compartments (lamina, midrib, dewlap).
- Within sugarcane leaf compartments, many linear and nonlinear associations exist between bacteria, fungi and parameters changing with YCS such as chlorophyll, water and sugar content, demonstrating the complex response of the sugarcane microbiome to YCS.

Based on the above findings, we suggest that fungal and viral infection or soil nutritional health are unlikely to be a direct cause of YCS. *Curtobacterium* and Phytoplasma remain candidates of interest and should be investigated further to confirm or deny their role in YCS incidence.

8. PEER REVIWED JOURNAL PUBLICATIONS

Hamonts, K., Trivedi, P., Garg, A., Janitz, C, Grinyer, J., Holford, P., Botha, F.C., Anderson, I.C. and Singh, B.K. (2018), Field study reveals core plant microbiota and relative importance of their drivers. Environmental Microbiology, 20(1), 124-140.

Hamonts, K., Trivedi, P., Grinyer, J., Holford, P., Drigo, B., Anderson, I.C. and Singh, B.K. (2018). Yellow Canopy Syndrome in sugarcane is associated with shifts in the rhizosphere soil metagenome but not with overall soil microbial function. Soil Biology and Biochemistry, accepted July 2018.

9. ACKNOWLEDGEMENTS

This study was funded by Sugar Research Australia (YCS-Integrated Research Program), the Queensland Government, Western Sydney University and the Global Centre for Land-based Innovation. We would like to thank YCS team members from each collaborative project including F Botha, G Scala, A Marquardt, K Wathen-Dunn, R Shafie.

We would also like to thank Anshu Garg, Melissa Martin, Krista Plett, Catarina Martins, Marcus Klein and Smriti Rayu for technical support during the project and Lawrence Di Bella, John Agnew, Mick Mackenzie, Steven Garrad, Jordan Villaruz, Greg Shannon, Davey Olsen, Leanne Hayes and Angela Zeilstra for assistance in identifying sampling sites. We acknowledge Susanna Tringe and colleagues (JGI, USA) for sharing Illumina MiSeq platform protocols using PNA clamps. BKS acknowledges his work on plant-microbial interactions is funded through the Australian Research Council (DP170104634).

10. REFERENCES

Anderson, M.J. (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecol 26: 32–46.

Anderson, M.J., and Cribble, N.A. (1998) Partitioning the variation among spatial, temporal and environmental components in a multivariate data set. Austral J Ecol 23: 158–167.

Blakemore, L.C., Searle, P.L., Daly, B.K. (1987). Methods for chemical analysis of soils. New Zealand Soil Bureau Scientific Report 80.

Campbell, C.D., Chapman, S.J., Cameron, C.M., Davidson, M.S., Potts, J.M. (2003). A rapid microtiter plate method to measure carbon dioxide evolved from carbon substrate amendments so as to determine the physiological profiles of soil microbial communities using whole soil. Applied and Environmental Microbiology, 69: 3593-3599.

Candresse, T., Filloux, D., Muhire, B., Julian, C., Galzi, S., Fort, G., Bernardo, P., Daugrois, J.-H., Fernandez, E., Martin, D.P., Varsani, A. and Roumagnac, P. (2014). Appearances can de deceptive: Revealing a hidden viral infection with deep sequencing in a plant quarantaine context. PloS ONE, 9(7): e102945. doi:10.1371/journal.pone.0102945.

Clarke, K.R. (1993) Non-parametric multivariate analyses of changes in community structure. Aust J Ecol 18: 117–143.

Clarke, K.R., and Warwick, R.M. (2001). Change in Marine Communities: An approach to statistical analysis and interpretation. Plymouth, UK: PRIMER-E.

Comstock, J. C. (1994). Yellow leaf syndrome appears on the United States mainland. Sugar J 56: 33-35.
DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., et al. (2006) Greengenes, a himerachecked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 72: 5069–5072.

Edgar, R.C. (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods 10: 996–998.

Gonçalves, Marcos C., et al. (2005). Sugarcane yellow leaf virus infection leads to alterations in photosynthetic efficiency and carbohydrate accumulation in sugarcane leaves. Fitopatologia brasileira 30: 10-16.

Gower, J.C. (1966) Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53: 325–338.

Gundersen, D.E., and Lee, I.M. (1996). Ultrasensitive detection of phytoplasmas by nested-PCR assays using two universal primer pairs. Phytopath Mediterranea, 35: 114-51

Hamonts, K., Trivedi, P., Garg, A., Janitz, C, Grinyer, J., Holford, P., Botha, F.C., Anderson, I.C. and Singh, B.K. (2018), Field study reveals core plant microbiota and relative importance of their drivers. Environmental Microbiology, 20(1), 124-140.

Heberle, H., Meirelles, G.V., da Silva, F.R., Telles, G.P., and Minghim, R. (2015) InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics 16: 169.

Herlemann, D.P., Labrenz, M., J€urgens, K., Bertilsson, S., Waniek, J.J., and Andersson, A.F. (2011) Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. ISME J 5: 1571–1579.

Ihrmark, K., Bodeker, I.T., Cruz-Martinez, K., Friberg, H., Kubartova, A., Schenck, J., et al. (2012) New primers to amplify the fungal ITS2 region - evaluation by 454- sequencing of artificial and natural communities. FEMS Microbiol Ecol 82: 666–677.

Koljalg, U., Nilsson, R.H., Abarenkov, K., Tedersoo, L., Taylor, A.F., Bahram, M., et al. (2013) Towards a unified paradigm for sequence-based identification of fungi. Mol Ecol 22: 5271–5277.

Kreuze, J.F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I. and Simon, R. 2009. Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: A generic method for diagnosis, discovery and sequencing of viruses. Virology, 388: 1-7.

Lee, I.M., Gundersen-Ridal, D.E., Davis, R.E., Bottner, K.D., Marcone, C., Seemuller, E. (2004) Candidatus Phytoplasma astris', a novel phytoplasm taxon associated with asters yellows and related diseases. Int J Sys Evol Microbiol, 54: 1037-48.

Lorenz, K.H., Schneider, B., Ahrens, U., Seemuller, E. (1995). Detection of the apple proliferation and pear decline phytoplasmas by PCR amplification of ribosomal and nonribosomal DNA. Amer Phytopath Soc, 85: 771-76.

Lundberg, D.S., Yourstone, S., Mieczkowski, P., Jones, C.D., and Dangl, J.L. (2013) Practical innovations for high-throughput amplicon sequencing. Nat Methods 10: 999–1002.

Magoc, T., and Salzberg, S.L. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27: 2957–2963.

McArdle, B.H., and Anderson, M.J. (2001) Fitting multivariate models to community data: a comment on distance-based redundancy analysis. Ecology 82: 290–297.

McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., et al. (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J 6: 610–618.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., Edwards, R.A. (2008). The metagenomics RAST server - a public resource for the automated phylogenetic and functional analysis of metagenomes. BMC Bioinformatics, 9: 386.

Moonan, F., Molina, J., Mirkov, T.E. Sugarcane yellow leaf virus: an emerging virus that has evolved by recombination between luteoviral and poleroviral ancestors. Virology 269: 156-171.

Overbeek R., Begley T., Butler R.M., Choudhuri J.V., Chuang H.Y., et al. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Research, 33: 5691–5702.

Sagaram, U.S., DeAngelis, K.M., Trivedi, P., Andersen, G.L., Lu, S.E., Wang, N. (2009). Bacterial diversity analysis of Huanglongbing pathogen-infected citrus, using PhyloChip arrays and 16S rRNA gene clone library sequencing. Applied and Environmental Microbiology, 75: 566-1574.

Saiya-Cork, K.R., Sinsabaugh, R.L., Zak, D.R. (2002). The effects of long term nitrogen deposition on extracellular enzyme activity in an *Acer saccharum* forest soil. Soil Biology and Biochemistry, 34: 1309-1315.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al. (2009) Introducing mothur: opensource, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75: 7537–7541.

Sharbatkhari, M., Bahar, M., Ahoonmanesh, A. 2008 Detection of the phytoplasmal agent of pear decline in Iran, Isfahan province, using nested PCR. Int J Plant Product 2; 167-174.

Sinsabaugh, R.L., Saiya-Cork, K.R., Long, T., Osgood, M.P., Neher, D.A., Zak, D.R., Norby, R.J. (2003). Soil microbial activity in a Liquidambar plantation unresponsive to CO_2-driven increases in primary production. Applied Soil Ecology, 24: 263-271.

Skennerton, C.T., Imelfort, M. and Tyson, G.W. 2013. Crass: identification and reconstruction of CRISPR from unassembled metagenomics data. Nucl Ac Res, 41(10): doi:10.1093/nar/gkt183.

Smart, C.D, Schneider, B., Blomquist, C.L., Guerra, L.J., Harrison, N.A., Ahrens, U., Lorenz, K.H., Seemuller, E., Kirkpatrick, B.C. (1996) Phytoplasma-specific PCR primers based on sequences of the 16S-23S rRNA spacer region. Appl Environ Micro, 62: 2988-2993.

Smith, Grant R., et al. (2000). Sugarcane yellow leaf virus: a novel member of the Luteoviridae that probably arose by inter-species recombination. J Gen Virology 81: 1865-1869.

Timmer, L.W., Lee, R.F., Brlansky, R.H., Graham, J.H., Albrigo, L.G., et al. (1992). The infectious nature of citrus blight. Proc. Fla. State Hortic. Soc. 105:21–26.

Trivedi, P., Duan, Y. and Wang, N., 2010. Huanglongbing, a systemic disease, restructures the bacterial community associated with citrus roots. Applied and Environmental Microbiology 76: 3427-3436.

Tyson, G.W. and Banfield, J.F. 2008. Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. Environ Microbiol, 10: 200-207.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naiıve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73: 5261–5267.

Wang, N., and Trivedi, P. (2013). Citrus huanglongbing: a newly relevant disease presents unprecedented challenges. Phytopathology 103: 652-65.

Wu, Q., Ding, S.-W., Zhang, Y. and Zhu, S. 2015. Identification of viruses and viroids by next-generation sequencing and homology-dependent and homology-independent algorithms. Annu. Rev. Phytopathol., 53: 425-444.

Wu, Q., Wang, Y., Cao, M., Pantaleo, V., Burgyan, J., Li, W.-X. and Ding, S.-W. 2012. Homology-independent discovery of replicating pathogenic circular RNAs by deep sequencing and a new computational algorithm. Proc. Nat. Ac. Sc., 109: 3938-3943.

Industry updates referenced:

SRA Industry Update, October 21, 2013. https://sugarresearch.com.au/growers-and-millers/pests-and-diseases/yellow-canopy-syndrome/

SRA Industry update, 5th March 2014. https://sugarresearch.com.au/growers-and-millers/pests-and-diseases/yellow-canopy-syndrome/

11. APPENDIX

11.1. Appendix 1 METADATA DISCLOSURE

Table 12 Metadata disclosure 1

Data	Sequencing data (including bacterial 16S rRNA sequences, fungal ITS2 sequences, shotgun metagenomics data, small RNA data)
Stored Location	HIE server
Access	Restricted; access to the data available upon request.
Contact	Gerard Devine, Data manager of the Hawkesbury Institute for the Environment E-mail: g.devine@westernsydney.edu.au Phone: 02 4570 1209

Table 13 Metadata disclosure 2

Data	Bacterial 16S rRNA amplicon sequencing data & fungal ITS2 amplicon sequencing data corresponding to November 2014 and June 2015 field sampling campaigns performed by Western Sydney University
Stored Location	NCBI Sequence Read Archive
Access	Publically accessible
Contact	NCBI Sequence Read Archive under Bioprojects PRJNA390435 (bacteria) and PRJNA390436 (fungi).