DIAGNOSTIC VOLATILES TO PREDICT RATOON STUNTING DISEASE

By

A.Z. BERNA[1], X. R. WANG[2], J. CASSELLS[1], B. CROFT[3], S. TROWELL[1]

[1]CSIRO Food and Nutrition Flagship. Clunies Ross, 2601 Acton, ACT, Australia
[2]CSIRO Digital Productivity Flagship. Corner of Vimiera and Pembroke Roads, Marsfield NSW 2122, Australia
[3]Sugar Research Australia, 90 Old Cove Rd, Woodford Q 4514 Australia

KEYWORDS: sugarcane, *Leifsonia xyli* subsp. x*yli,* E-nose

Abstract
Ratoon Stunting Disease (RSD), caused by *Leifsonia xyli* subsp. *xyli,* is one of the most significant diseases to affect sugarcane. Control relies on strict hygiene, including screening of seed cane, sterilisation of equipment and selective use of partially resistant germplasm. The current standard RSD diagnostic assay is based on an evaporative binding enzyme linked immunosorbent assay (EB-ELISA), which requires sap to be extracted and dried onto an immunoassay plate. Although high throughput is achievable with this method, it is a slow and complicated process.
Some species of pathogenic bacteria can be characterized by the volatile chemicals they produce. In this paper, we show that six volatiles are present at significantly higher levels in RSD-infected sugarcane, compared with uninfected plants. Gas chromatography-mass spectrometry analysis of the headspace of sugar sap from infected and uninfected plants identified specific alcohols, ketones and aldehyde as signatures of RSD infection. We were able to classify the infection status of plants with better than 95% accuracy using this information. We aim to use these volatile biomarkers of RSD to train electronic nose sensors to sniff out contamination with more convenience than current ELISA methods.

Introduction

There is an ongoing need for simple, sensitive, fast and accurate detection of microbes associated with plants and plant tissue. The requirement encompasses microbes that are serious plant pathogens as well as microbes that can cause illness in animals or humans. Examples of the latter include *Aspergillus flavus* and *Listeria monocytogenes*. *Leifsonia xyli* subsp. *xyli*, which causes Ratoon Stunting Disease (RSD) is a pathogen of economic significance. RSD is hard to detect and control (Saumtally *et al.*, 1996). It was discovered by BSES pathologists in 1944 and is probably the most economically important disease of sugarcane worldwide. RSD is found in all cane-growing districts of eastern Australia. Incidence of the disease depends on how strictly growers follow control measures aimed at excluding infected cane from the propagation cycle. RSD has no easily attributed symptoms but can cause yield losses of 5-60%, depending on the susceptibility of the variety and whether plants are stressed (Steindl, 1961). The current EB-ELISA diagnostic used by the industry has been effective in helping Australian growers to manage RSD but obtaining a diagnosis is slow and cumbersome and costs up to $0.5M per annum. Easier and faster

methods of detecting *L. xyli* subsp. *xyli* would be valuable because they would accelerate the decision-making process and cut costs.
Plant and animal pathogens emit, or trigger the plant to emit, characteristic volatile organic compounds (VOC) (Korpi *et al.*, 2009). The gold standard method for analysing VOCs is to collect a gas sample from the headspace above a plant or plant tissue and analyse it with a gas chromatograph coupled to a mass spectrometer (GC-MS). Although this equipment is very accurate and can usually identify VOCs that are diagnostic for disease, it is expensive and requires a well-equipped central laboratory with highly trained operators. Electronic nose (E-nose) technology, is simpler and cheaper and, albeit less accurate than GC-MS, offers the ability to make a diagnosis on the spot (Gardner *et al.*, 1998). Electronic nose technology has improved considerably over recent years and several papers have been published describing applications of E-nose technology to detecting plant pathogens or food contamination (Gardner *et al.*, 1998; Younts *et al.*, 2002; Younts *et al.*, 2003).

In this paper we report the volatiles associated with RSD infection and the development of a test based on the levels of RSD associated volatiles to classify healthy from infected samples. The knowledge of RSD volatiles will allow the development of sensors that can predict RSD plant infestation in a more convenient way than current methods.

Materials and methods

Plant samples and ELISA test

In total 145 samples of RSD-infected and disease-free cane was collected from the BSES Woodford Experiment Station. Samples came from three sites and 9 varieties (Table 1). The age of all plants tested in this study was 12 month old.

Table 1. Sugar cane variety, sample size, sample source, and average ELISA results of collected samples.

Variety	Number of samples	Source	Average ELISA
INFECTED SAMPLES			
KQ228	5	RSD resistance trial (A2)	1.98
Q110	8	RSD resistance trial (A2)	2.12
Q124	9	RSD resistance trial (A2)	2.11
Q138	10	RSD resistance trial (A2)	2.76
Q183	5	RSD resistance trial (A2)	1.41
Q232	8	RSD resistance trial (A2)	2.18
Q240	10	RSD resistance trial (A2)	0.72
Total 55			
HEALTHY SAMPLES			
KQ228	12	Woodford propagation (D1)	0.03
Q110	9	Woodford propagation (D1)	0.04
Q124	13	Woodford propagation (D1)	0.03
Q138	3	Woodford propagation (D1)	0.05

Q155	5	Woodford propagation (D1)	0.04
Q183	10	Woodford propagation (D1)	0.04
Q232	6	Woodford propagation (D1)	0.03
Q240	15	Woodford propagation (D1)	0.04
Q90	14	Woodford propagation (D1)	0.03
Q110	1	Kallangur propagation	0.02
Q155	3	Kallangur propagation	0.03
Total 91			

Stalk vascular extracts for ELISA and phase contrast microscopy were collected by applying positive air pressure to stalk pieces. The stalk pieces were taken from towards the base of the stalks (Figure 1). The extracts were collected in microcentrifuge tubes and frozen at –20 °C until required. In these experiments, an additional 1 ml of extract was collected for GC-MS analysis.

Figure 1: Xylem vascular extract from sugar cane using positive pressure

The ELISA procedure has been described by (Croft, 2002). Absorbance of the colour reaction in the ELISA was read on a Biorad microplate reader at 405 nm. Samples with an absorbance reading of >0.05 were considered positive.

GC-MS analysis

For GC-MS analysis of the sugar sap 1 mL of xylem extract was placed into 10 mL vials. Samples were incubated at 40° C with shaking (500 rpm) for 10 minutes. Headspace extraction was carried out for 40 minutes with a solid phase micro-extraction SPME fibre (Aldrich, Bellefonte, PA) composed of fused silica partially cross-linked with 65 µm polydimethylsiloxane/divinylbenzene (PDS/DVB). The SPME analysis was carried out using an autosampler (CombiPAL, Switzerland). After absorption, headspace volatiles were transferred to the GC injection port, which was equipped with a 0.8 mm i.d. splitless glass liner, at 240°C. Desorbed volatile compounds were separated in a Varian 3800 GC, equipped with a 30 m × 0.25 mm, 0.25 µm film thickness ZB-5MS fused silica capillary column. The oven temperature was programmed to rise from 30° C (held for 2 min) to 80° C at 20° C/min (held for 1 min) then to 100°C at 20°C/min (held for 1min) and finally to 230°C at 30°C/min (held for 2min). The GC column output was fed into a Varian 1200 mass selective detector (mass spectrometer). The GC-MS transfer line was heated at 250°C with the flow rate of the He carrier gas set to 1 mL min^{-1}. Mass spectrometry was performed in electron impact mode at 70 eV over the scan range 35-350 *m/z.*

GC-MS Data Processing

The GCMS data was processed in the following steps: (1) normalisation, (2) baseline correction, (3) peak detection, (4) peak alignment. As the GCMS analysis of the data was carried out over a period of several weeks, gradual changes in the GCMS instrument's sensitivity means the data analysed on different days are not consistent (i.e. a "day of analysis" effect). Therefore, we first normalised the data using the blank sample of the day of analysis (i.e. analysis of an empty sorbent tube). The baseline drift of each single ion chromatogram (SIC) is then corrected using Asymmetric Least Square (ALS)(Eilers and Boelens, 2005). We employed the peak detection algorithm as described in (Vivo-Truyols *et al.*, 2005) for each detected peak, we calculated the area under the peak, which is then used for alignment. Finally, alignment of the peaks is achieved through a combination of piecewise method (Johnson *et al.*, 2003) and dynamic time warping, where the target peaks were chosen based on dendrogram alignment scheme (Aberg *et al.*, 2004).

Feature Selection

We used the mutual information (MI) (Wang *et al.*, 2014) to find the optimal subset of features (peak areas) that give the best classification performance. The subset of features is selected by maximising the *mutual information* $I(\mathbf{Z}^v, C)$ between the selected features $\mathbf{Z}^v = \{Z^{i_1}, \ldots, Z^{i_n}\}$ and class C:

$$I(\mathbf{Z}^v, C) = \sum_{\mathbf{Z}^v, C} p(\mathbf{Z}^v, C) \log_2 \frac{p(C \mid \mathbf{Z}^v)}{p(C)}$$

We employed the Kraskov–Grassberger technique (Kraskov *et al.*, 2004) (estimator 2) for estimating the mutual information. The code used for the estimation is the publicly available Java Information Dynamics Toolkit (Lizier, 2014).

Classification

Two common classifiers, Support Vector Machine (SVM) (Cortes and Vapnik, 1995) and k nearest neighbours (kNN) (Russell and Norvig, 1995), were used to first train a model and then classify the test data. We performed the classification using the one-against-all (leave-one-out) cross-validation (Wang *et al.*, 2014).

Results and discussions

ELISA reading and volatiles levels in sugar sap

The average absorbance measured by the ELISA assay for all infected plots in the trial was between 0.72 and 2.76 (Table 1). The cultivar with the highest levels of infection was Q138 from plot A1.

Volatile analysis of all samples revelled that 81 compounds are present in the headspace of healthy and infected sugar sap. Alcohols, organic acids, aromatic compounds, aldehydes and ketones are among the chemical families found in the headspace. The majority of the compounds were identified using either NIST library or using standards.

Diagnostic RSD volatiles

Using GC-MS feature selection described above, seven m/z corresponding to seven volatiles showed to be diagnostic to RSD infection (Table 2). With the exception for 2-ethyl hexanoic

acid all other diagnostic volatiles were present in higher concentration in RSD infected samples compared to healthy samples.

Table 2- Diagnostic RSD volatiles and their characteristic mass to charge ratio (m/z)

Rt* (min)	Compound name	m/z	levels[ŧ]
1.957	Dimethyl sulfide	62	i>h
2.195	2-Methyl propanal	41	i>h
5.296	Ethylbenzene	106	i>h
5.681	Styrene	104	i>h
7.564	2-Ethyl-4-methyl 1-pentanol	57	i>h
7.634	2-Ethyl-1-hexanol	41	i>h
8.649	2-Ethyl hexanoic acid	73	h>i

*Rt=retention time, ŧ i=infected and h=healthy samples

Sulful compounds like dimethyl disulfide are the universal bacterial metabolites and it may be present in higher concentration in infected samples due to degradation of sulphur-containing amino acids such as methionine and cysteine, as has been reported previously for S. typhimurium contamination in vegetables and meat (Kai *et al.*, 2009; Xu *et al.*, 2010). Styrene, an aromatic hydrocarbon, is listed among volatiles emitted from plants (Tikunov *et al.*, 2005; Blasioli *et al.*, 2010). The sources of the potentially toxic compound styrene in the environment are of two types: anthropogenic and natural (Mooney *et al.*, 2006). Anthropogenic sources are mainly referable to petrochemical and polymer-processing industries and the automobile exhaust fumes. The natural styrene formation has been observed as a trace metabolite in foods, in particular cheeses, where it acts as an aroma defect (Pagot *et al.*, 2007). Styrene is also known to be naturally synthesized by select plant species, including several trees in the *Styracaceae* family (including several *Styrax* sp.) (Mckenna and Nielsen, 2011). Styrene is believed to be synthesized from excess L-phenylalanine, compatible with decarboxylation of cinnamic acid involved in secondary metabolism of plants (Mckenna and Nielsen, 2011). It is also known that microorganisms, such as *Penicillium* spp., *Aspergillus niger* and *Saccharomyces cerevisiae* do promote the decarboxylation of cinnamic acid (Blasioli *et al.*, 2010). On the basis of these studies, it is therefore plausible that styrene is naturally produced in healthy sugar sap and that *Leifsonia xyli* subsp. *Xyli* synthesises styrene and therefore its increase in levels is observed.

2-Ethyl-1-hexanol has been previously reported as a natural compound in grapes (Genoves *et al.*, 2005). In Table 1, 2-ethyl-1-hexanol and 2-ethyl hexanoic acid are associate since the acid is produced from the oxidation of the alcohol. Apparently *Leifsonia xyli* is able to reduce the organic acid and form the alcohol (Table 2). The enzyme that the bacteria uses for this chemical reaction is unknown.

With respect to the other volatile species, these have not been extensively reported in the literature however, it is known that VOCs are used by microorganisms to give the VOC producer an advantage in terms of colonisation by suppressing the growth of other bacteria. Also, when a plant is subjected to abiotic or biotic stress, a set of responses are activated to contrast the infection process culminating in the disease. A major discovery of the last

decade is that plants under stress emit much greater amounts and varieties of volatiles to serve mainly as indirect plant defence (Dudareva *et al.*, 2004).

Classification of samples using volatile profile

Of the two classifiers tested, kNN gave the best results for sugar sap. The first two volatiles selected in the mutual information as being the best compounds to discriminate healthy and infected samples were styrene and ethyl benzene. Ethyl benzene on its own can correctly classify 97.9% of samples while styrene can classify 98.6% of samples (Table 3).

Table 3: Confusion matrix using kNN classifier for ethyl benzene and styrene RSD diagnostic volatiles

Ethylbenzene

True class	Class	Predicted class	
		healthy	infected
	healthy	89	2
	infected	1	54

97.9% classification

Styrene

True class	Class	Predicted class	
		healthy	infected
	healthy	90	1
	infected	1	54

98.6% classification

If we use styrene only as a predictor of infection, two misclassified samples were found: one false negative and one false positive (Table 3). Future work will involve the analysis of the healthy sample using PCR to confirm if it was free of infection.

Correlation of ELISA vs diagnostic volatiles

Styrene and ethyl benzene peak areas were correlated with ELISA reading (Figure 2), the results showed that although the correlation is not strong, there is some relationship between the levels of infection and the levels of the volatiles: r^2= 0.42 and r^2= 0.59 for styrene and ethyl benzene respectively. The reason the correlation was not higher is because the measurement is linear in nature and seven healthy samples had high levels of ethyl benzene and styrene (Figure 2). Interesting to note is that of the seven 'outliers', only one or two healthy samples failed to be correctly classified as shown in the confusion matrix (Table 3). This is because both classifiers used employ a 'clustering' mechanism. As we can see from Figure 2, there are effectively three groups of samples: healthy samples with very low peak area, healthy samples with very large peak area, and infected samples with intermediate sized peak area. Therefore, since we use one-against-all cross-validation, for any test data in the group of the seven outlier healthy samples, it can be classified as 'healthy' due to the other outliers. The fact that the seven outliers had very large peak areas suggests that samples were either RSD infected not detected by ELISA or RSD free but infected with another plant disease. Further investigation of these outlier samples will be carried out. When the outliers are removed the correlations are 0.82 and 0.80 for ethylbezene and styrene respectively

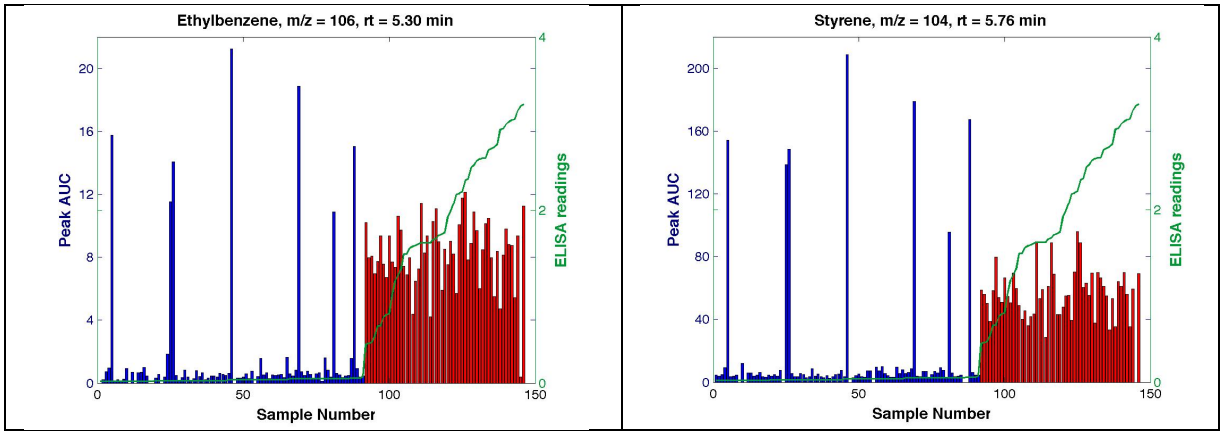

Figure 2: Peak intensity for ethyl benzene and styrene and ELISA readings for healthy (blue peaks) and infected (red peaks) samples.

Variety characterization

Another interesting finding, outside the scope of this research work, was that the volatile profile can be also used to discriminate cultivars. For example, cultivar KQ228 and Q90 coming from the same site and both healthy samples can be discriminated on the basis of their volatile profile (Figure 3). These results can potentially be used to test the origins of cultivars or characterize cultivars that are resistant to RSD.

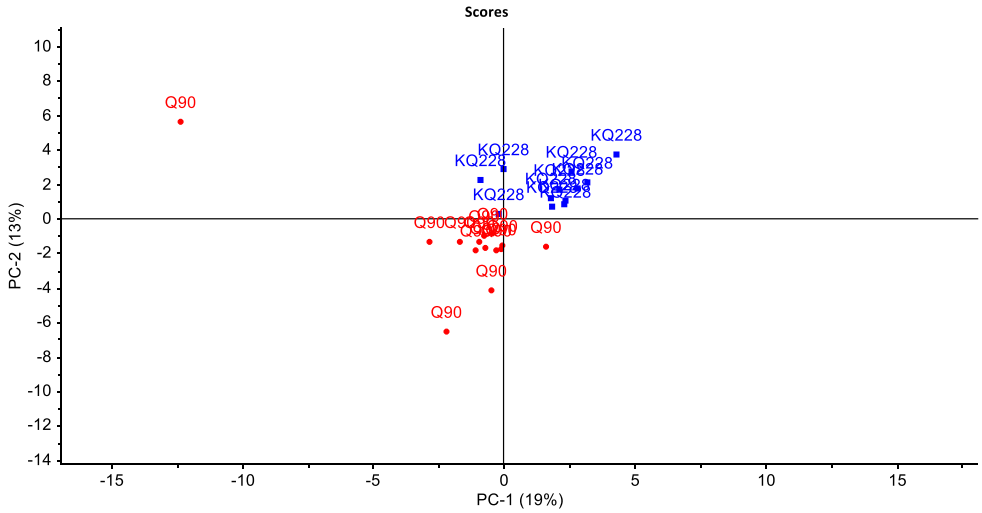

Fig. 3- Principal component analysis using GC-MS data of cultivars KQ228 and Q90. Both cultivars are healthy samples from a 12 month old plant and grown on same site.

Conclusions:

Gas chromatography-mass spectrometry analysis of the headspace of sugar sap from infected and uninfected plants identified specific alcohols, ketones and aldehydes as signatures of RSD infection. In particular styrene and ethyl benzene are the best signatures for RSD infection, using these volatiles we obtain more than 98% correct classification. There is also evidence of correlation between ELISA readings and peak area of the volatiles. Future work will involve validation of these results with other cultivars and testing the electronic nose for its ability to detect these volatiles.

Acknowledgements
We like to acknowledge Amanda Johnson who conducted the ELISA assays and microscopic examination of samples.

REFERENCES

Aberg KM, RJO Torgrip, SP Jacobsson (2004). Extensions to peak alignment using reduced set mapping: classification of LC/UV data from peptide mapping. *Journal of Chemometrics* **18**(10): 465-473.

Blasioli S, E Biondi, I Braschi, U Mazzucchi, C Bazzi, CE Gessa (2010). Electronic nose as an innovative tool for the diagnosis of grapevine crown gall. *Analytica Chimica Acta* **672**(1–2): 20-24.

Cortes C, V Vapnik (1995). Support vector networks. *Machine Learning* **20**(3): 273-297.

Croft BJ (2002). A method for rating sugarcane cultivars for resistance to ratoon stunting disease based on an enzyme-linked immunoassay. *Australasian Plant Pathology* **31**(1): 63-66.

Dudareva N, E Pichersky, J Gershenzon (2004). Biochemistry of plant volatiles. *Plant Physiology* **135**(4): 1893-1902.

Eilers PHC, HFM Boelens (2005). Baseline correction with asymmetric least squares smoothing. Leiden University Medical Centre Report

Gardner JW, M Craven, C Dow, EL Hines (1998). The prediction of bacteria type and culture growth phase by an electronic nose with a multi-layer perceptron network. *Measurement Science & Technology* **9**(1): 120-127.

Genoves S, JV Gil, S Valles, JA Casas, P Manzanares (2005). Assessment of the aromatic potential of Palomino fino grape must using glycosidases. *American Journal of Enology and Viticulture* **56**(2): 188-191.

Johnson KJ, BW Wright, KH Jarman, RE Synovec (2003). High-speed peak matching algorithm for retention time alignment of gas chromatographic data for chemometric analysis. *J Chromatogr A* **996**(1-2): 141-155.

Kai M, M Haustein, F Molina, A Petri, B Scholz, B Piechulla (2009). Bacterial volatiles and their action potential. *Applied Microbiology and Biotechnology* **81**(6): 1001-1012.

Korpi A, J Jarnberg, AL Pasanen (2009). Microbial Volatile Organic Compounds. *Critical Reviews in Toxicology* **39**(2): 139-193.

Kraskov A, H Stogbauer, P Grassberger (2004). Estimating mutual information. *Phys Rev E Stat Nonlin Soft Matter Phys* **69**(6 Pt 2): 066138.

Lizier JT (2014). JIDT: An information-theoretic toolkit for studying the dynamics of complex systems. *Frontiers in Robotics and AI* **1**.

McKenna R, DR Nielsen (2011). Styrene biosynthesis from glucose by engineered E. coli. *Metabolic Engineering* **13**(5): 544-554.

Mooney A, PG Ward, KE O'Connor (2006). Microbial degradation of styrene: biochemistry, molecular genetics, and perspectives for biotechnological applications. *Appl Microbiol Biotechnol* **72**(1): 1-10.

Pagot Y, JM Belin, F Husson, HE Spinnler (2007). Metabolism of phenylalanine and biosynthesis of styrene in Penicillium camemberti. *J Dairy Res* **74**(2): 180-185.

Russell S, P Norvig (1995). Articial Intelligence: A Modern Approach. E. C. Pentice Hall, N.J.

Saumtally S, A Dookun, LJC Autrey (1996). Ratoon stunting disease. *Sugarcane Germplasm Conservation and Exchange*(67): 116-117.

Steindl DRL (1961). Sugar cane diseases of the world, Elsevier: Amsterdam.

Tikunov Y, A Lommen, CHR de Vos, HA Verhoeven, RJ Bino, RD Hall, AG Bovy (2005). A novel approach for nontargeted data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles. *Plant Physiology* **139**(3): 1125-1137.

Vivo-Truyols G, JR Torres-Lapasio, AM van Nederkassel, Y Vander Heyden, DL Massart (2005). Automatic program for peak detection and deconvolution of multi-overlapped chromatographic signals part I: peak detection. *J Chromatogr A* **1096**(1-2): 133-145.

Wang XR, JT Lizier, AZ Berna, FG Bravo, SC Trowell (2014). Human breath-print identification by E-nose, using information-theoretic feature selection prior to classification. *Sensors and Actuators B: Chemical*(*in Press*).

Xu Y, W Cheung, CL Winder, R Goodacre (2010). VOC-based metabolic profiling for food spoilage detection with the application to detecting Salmonella typhimurium-contaminated pork. *Analytical and Bioanalytical Chemistry* **397**(6): 2439-2449.

Younts S, E Alocilja, W Osburn, S Marquie, J Gray, D Grooms (2003). Experimental use of a gas sensor-based instrument for differentiation of Escherichia coli O157 : H7 from non-O157 : H7 Escherichia coli field isolates. *Journal of Food Protection* **66**(8): 1455-1458.

Younts S, EC Alocilja, WN Osburn, S Marquie, DL Grooms (2002). Differentiation of Escherichia coli O157 : H7 from non-O157 : H7 E-coli serotypes using a gas sensor-based, computer-controlled detection system. *Transactions of the Asae* **45**(5): 1681-1685.