ORIGINAL ARTICLE

CrossMark

ATR-FTIR spectroscopy non-destructively detects damage-induced sour rot infection in whole tomato fruit

Paul Skolik[1] · Martin R. McAinsh[1] · Francis L. Martin[2]

Abstract

Main conclusion ATR-FTIR spectroscopy with subsequent multivariate analysis non-destructively identifies plant–pathogen interactions during disease progression, both directly and indirectly, through alterations in the spectral fingerprint.

Plant–environment interactions are essential to understanding crop biology, optimizing crop use, and minimizing loss to ensure food security. Damage-induced pathogen infection of delicate fruit crops such as tomato (*Solanum lycopersicum*) are therefore important processes related to crop biology and modern horticulture. Fruit epidermis as a first barrier at the plant–environment interface, is specifically involved in environmental interactions and often shows substantial structural and functional changes in response to unfavourable conditions. Methods available to investigate such systems in their native form, however, are limited by often required and destructive sample preparation, or scarce amounts of molecular level information. To explore biochemical changes and evaluate diagnostic potential for damage-induced pathogen infection of cherry tomato (cv. Piccolo) both directly and indirectly, mid-infrared (MIR) spectroscopy was applied in combination with exploratory multivariate analysis. ATR-FTIR fingerprint spectra (1800–900 cm^{-1}) of healthy, damaged or sour rot-infected tomato fruit were acquired and distinguished using principal component analysis and linear discriminant analysis (PCA–LDA). Main biochemical constituents of healthy tomato fruit epidermis are characterized while multivariate analysis discriminated subtle biochemical changes distinguishing healthy tomato from damaged, early or late sour rot-infected tomato indirectly based solely on changes in the fruit epidermis. Sour rot causing agent *Geotrichum candidum* was detected directly in vivo and characterized based on spectral features distinct from tomato fruit. Diagnostic potential for indirect pathogen detection based on tomato fruit skin was evaluated using the linear discriminant classifier (PCA–LDC). Exploratory and diagnostic analysis of ATR-FTIR spectra offers biological insights and detection potential for intact plant–pathogen systems as they are found in horticultural industries.

Keywords MIR spectroscopy · Plant–pathogen interaction · Multivariate analysis · Crop biology · *Geotrichum candidum*

✉ Martin R. McAinsh
   m.mcainsh@lancaster.ac.uk

✉ Francis L. Martin
   flmartin@uclan.ac.uk

1 Lancaster Environment Centre, Lancaster University, Lancaster, UK

2 School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston, UK

Introduction

Providing food security for a rapidly growing global population of which a large fraction is malnourished is one of the greatest challenges in the modern era (IFPRI 2017). Conventional solutions such as increased land clearing and increasing usage of pesticides to produce sufficient food are unfavourable due to their environmental impacts and long-term unsustainability. Thus, novel alternatives are needed to efficiently produce the approximately 70% more food needed by 2050 (Beuchelt and Virchow 2012). Crop loss to pests and pathogens throughout food production/supply represent a major threat to this aim. Pests and pathogens may

reduce crop yield by 80%, thereby presenting a significant challenge to crop productivity (Oerke 2006). Detection of pests and pathogens within pre- and post-harvest settings is therefore essential to minimize the impacts on crop production. The post-harvest consumer stage can be viewed as one of the most critical points during food production/supply, because maximum resource allocation has occurred at this point, making plant–pathogen interactions especially relevant at this stage. Current methods for crop diagnosing crop health include remote sensing, molecular based methods and complex analytical techniques all of which have drawbacks. Remote sensing, including hyperspectral imaging and thermography, are highly responsive to environmental conditions and distance to the measured object making it difficult to determine disease specificity (Mahlein 2016). Molecular based techniques [for example, polymerase chain reaction (PCR), fluorescent in situ hybridization (FISH), and enzyme-linked immunosorbent assay (ELISA)] are time consuming and prone to contamination (Chitarra and Van Den Bulk 2003; Schaad and Frederick 2002; Wallner et al. 1993). Complex analytical methods, for example gas or liquid chromatography coupled to mass spectrometry (GC/LC–MS, require extensive sample preparation and are difficult to use in the field (Martinelli et al. 2015).

The development of flexible, non-destructive sensors capable of providing adequate detection sensitivity and pathogen specificity are keys goal for the detection of crop pathogens (Mahlein 2016; Skolik et al. 2018). This includes the ability to detect early effects of plant stress or disease, to differentiate between the effects of abiotic and biotic stresses and between different diseases, and to quantify the severity of the stress or disease (Mahlein 2016). While various molecular and imaging techniques to detect crop pathogens are under development, the limitations of many analytical methods combined with the above criteria for sensor technologies, has led to the development of label-free, non-destructive spectroscopic techniques that provide information about the chemical structure of analysed samples. The spectroscopic approach, originating in analytical chemistry, has been translated to the biological sciences mainly through advancements in computational analysis and the ability to measure live samples (Chan and Kazarian 2016). Application of these techniques to model plant and crop systems has the potential to both provide novel insight into plant–pathogen interactions, whilst generating a large number of variables for autonomous classification of disease states for detection of pests and pathogens.

Mid-infrared (MIR) spectroscopy has made substantial headway in the biological sciences as a non-destructive and rapid bioanalytical sensor technology (Martin et al. 2010). This is because MIR spectra have been effective at providing molecular insights into biological systems, while providing a large number of variables on which

to discriminate samples. More recently, spectrochemical techniques have made substantial progress in the plant and crop sciences, specifically with regard to the analysis of dynamic processes and plant biology-related to crop production (Butler et al. 2015, 2017; Ord et al. 2016; Skolik et al. 2018). This has been effective mainly through the development of new data analysis methods including multivariate analysis. For MIR biospectroscopy, data analysis can be split into two main types: exploratory and diagnostic. Exploratory analysis is aimed primarily at data visualization and pattern recognition (Trevisan et al. 2012). Diagnostic analysis employs the use of classifier algorithms to evaluate the potential for diagnosis of sample condition (healthy versus diseased, for example) (Trevisan et al. 2012). While the two frameworks typically have different objectives, they are closely linked and in general, the exploratory precedes the diagnostic framework. Among the many available data analysis options, unsupervised principal component analysis (PCA) and supervised linear discriminant analysis (LDA) have been used alone or in combination to successfully investigate a large number of biological phenomena based on MIR data (Li et al. 2015; Strong et al. 2017). Both PCA and LDA have formed core components of biospectroscopy data analysis. Related classifiers including support vector machine (SVM) and linear discriminant classifier (LDC) have also found ample application in the diagnostic framework. Such advancements have highlighted the potential for MIR biospectroscopy as an effective sensor technology for the plant and crop sciences (Skolik et al. 2018). Despite this progress, the number of investigations on intact samples has been limited, which is arguably and important prerequisite for the development of fully non-destructive horticultural sensors, and thus more research on intact samples is required. Recently, both attenuated total reflection Fourier-transform infrared (ATR-FTIR) and Raman spectroscopy have been favoured for studying intact samples of important crops (Butler et al. 2015; Ord et al. 2016; Fu et al. 2016; Trebolazabala et al. 2013). Raman spectroscopy is a complementary method to FTIR spectroscopy and the two are often combined for a more robust analysis, as each have specific drawbacks due to the distinct light–matter interactions they measure (Baker et al. 2014; Butler et al. 2016). Compared to macro-FTIR, Raman scattering as a low probability event can be highly variable, prone to interference from fluorescence, featuring a small measurement area typically between 20 and 30 μm, and uses more intense laser powers potentially leading to photobleaching (tissue decomposition) of delicate organic samples (Butler et al. 2016; Yeturu et al. 2016). Nevertheless, a strong case has been made in previous studies demonstrating the effectiveness of Raman spectroscopy for direct detection of microbial pathogens in intact crops

(Egging et al. 2018; Farber and Kurouski 2018; Yeturu et al. 2016). While direct detection of a plant–pathogen interaction generates spectral changes suitable for disease detection, indirect detection of plant infection through spectral changes in tissues because of pathogen attack remains difficult but offers a novel approach especially for early or pre-symptomatic stages of disease (Skolik et al. 2018). ATR-FTIR spectroscopy has the advantage of macro-measurements increasing the measurement area while also affording a very defined magnitude of light penetration into the sample (Kazarian and Chan 2013; Chan and Kazarian 2016). This may be more suitable for analysis at the whole-plant level as many experiments still rely on previous removal (cutting) of leaves and fruit, which is not truly non-destructive (Trebolazabala et al. 2017; Yeturu et al. 2016). It is therefore essential to evaluate Raman complementary methods such as FTIR, as a combination approach can overcome the limitations of a single technique, as well as the variable nature of crops and plant–pathogen systems covered by modern agriculture.

Herein we use ATR-FTIR spectroscopy to study the effects of damage and ambient infection by the sour rot causing agent *G. candidum*, both directly and indirectly, in commercially obtained, consumer-stage (red ripe) intact cherry tomatoes. Conventional spectral analysis for the characterization of main absorbance peaks of tomato and fungus *G. candidum* is followed by exploratory and diagnostic multivariate analysis to probe subtle biochemical changes induced indirectly by damage and infection. Changes to the surface of tomato fruit are characterized in response to damage and sour rot infection using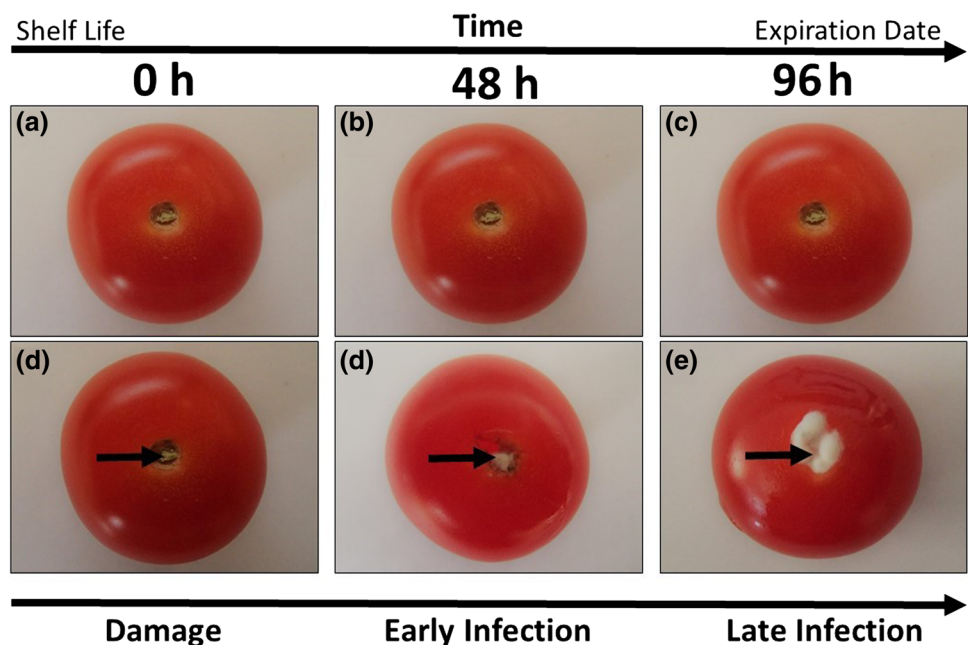 the tandem technique of PCA–LDA to maximize inter-class differences between damage, infected, and control fruit. The diagnostic potential of this approach is evaluated using the tandem classifier PCA–LDC, to distinguish damaged and infected tomato fruit from healthy controls indirectly and autonomously.

Materials and methods

Sample preparation and storage

Vine cherry tomatoes cv. Piccolo were obtained from a local supermarket (Sainsbury's Lancaster Main Store, UK). All analyses were performed prior to the advertised expiration date, which at the time of purchase presented a window of 8 days. Tomatoes were removed from their commercial packaging taken off the vine and adapted to room temperature (23 ± 1°C) and 35–40% relative humidity for 2 h prior to initial analysis. Loose tomatoes were split into two sets, including a control series accounting for changes occurring in naturally ripening tomatoes over the analysis timeframe (Fig. 1a–c), as well as a set of tomatoes punctured through the stem scar, at 0 h, to a depth of approximately 1 cm with a 21-gauge sterile syringe needle leaving the remainder of the skin intact (Fig. 1d). Damaged tomatoes were thus susceptible to ambient infection at the puncture site, which was visible starting at 48 h post-puncture, and full colonization of the punctured stem scar was observed at 96 h post-puncture (Fig. 1f). Control and damaged tomatoes were dark stored in cardboard boxes under identical conditions to compare pathogen development and allow stem scar infection with ambient fungal spores. Damaged tomato fruit (0 h) were

Fig. 1 Symptoms associated with tomato fruit damage (**d**), early (**e**), and late (**f**) infection of tomato fruit by *G. candidum* compared to their shelf-life-matched controls (**a–c**)

subsequently analysed opposite their shelf-life-matched controls, allowed to age naturally, at 48 and 96 h post-puncture, to assess pathogen infection caused by initial damage. Prior to analysis, tomato fruit were washed thoroughly with deionized water to remove dust and debris, as well as fungal growth and fluid exudate present on fruits at the early infection (48 h) and late infection (96 h) stages, prior to spectral acquisition, in order to characterize changes in fruit skin, without contribution from the fungus itself (analysed separately) or exudates from the site of infection. The fungal–fruit complex on fully colonized tomatoes at 96 h post-puncture (Fig. 1f) was analysed to obtain spectra from the fungus in its native state on tomato fruit.

ATR-FTIR spectroscopy

MIR spectra were acquired from intact tomato fruit, using a Bruker Tensor 27 IR spectrometer with Diamond ATR Helios attachment (Bruker Optics, Coventry, UK). Spectra were acquired over the range 4000–400 cm^{-1} with a spectral resolution of 8 cm^{-1}, 3.84 cm^{-1} data spacing, 32 co-additions and a mirror velocity of 2.2 kHz for optimum signal-to-noise ratio (Martin et al. 2010; Baker et al. 2014). Background spectra were taken prior to each sample to account for ambient atmospheric conditions. The diamond ATR crystal defined a spatial resolution (sampling area) of approximately 250 μm × 250 μm. The whole fruit was placed on the sample stage with no more than 0.1 kg of applied pressure. Between sample measurements, ATR cleaning wipes containing isopropyl-alcohol (Bruker Optics, Coventry, UK) were used to clean the ATR diamond crystal between measurements. Five points from around the fruit circumference were measured; two spectra from each of the five points for a total of 10 measurements per fruit. Ten spectra per fruit generally supply enough replicates for PCA–LDA, which provide intra-class differences (i.e., variance specific to 0 h control vs. 0 h damage, 48 h control vs. 48 h early infection, and 96 h control vs. 96 h late infection) while minimizing the effect of natural tissue heterogeneity potentially masking the subtle effects underpinning plant response to pathogen. Six fruits were measured for each treatment group. Measurements of *G. candidum* were taken in vivo without modification as part of the tomato–pathogen complex at late infection state (96 h post-puncture). The fungal mass completely covered the ATR crystal during measurements; six separate samples (10 spectra from each fungal sample) of *G. candidum* were measured to obtain a representative in vivo spectrum.

Pre-processing and computational analysis

All computational analysis was conducted using the open source IRootlab toolbox (https://github.com/trevisanj/iroot

lab) specialized for analysis of IR spectra (Trevisan et al. 2013), in conjunction with Matlab 2016 (The Math Works, MA, USA), unless otherwise stated. Raw spectra were truncated to the spectral fingerprint region between 1800 and 900 cm^{-1}, which is the primary region where biomolecules absorb IR radiation. Fingerprint spectra were pre-processed using the rubber band-like baseline correction algorithm and maximum normalized to account for differences in sample thickness and ATR diamond contact pressure. Class mean spectra were used for direct analysis. Exploratory PCA reduces the dataset down to factors that account for spectral variance; PCA was optimized using the IRootlab *pareto* function, where the first 10 PCs accounted for more than 99% of the variance in the dataset [see Supporting Information (SI) Figure S1]. These served as input variables for LDA forming the composite technique PCA–LDA (Trevisan et al. 2012). While PCA reduces the complexity of the spectral data, it is unsupervised, does not account for class labels, views all classes as one and therefore does not distinguish between control, damaged, or infected tomatoes for the purposes of extracting class-specific differences (Trevisan et al. 2012). Combined with a supervised approach, LDA following PCA (PCA–LDA) maximizes the spectral differences between classes (control vs. damage/infected), and thus allows the extraction of the class-specific biomarkers associated with damage and subsequent sour rot (Martin et al. 2010; Kelly et al. 2011; Trevisan et al. 2012). Pairwise comparisons between two classes generate one linear discriminant (LD), which summarizes the main class-specific differences between control and afflicted tomato fruit. This linear discriminant can be visualized as cluster plots, where each spectrum is defined as a point where overlap and separation of points indicate similar or dissimilar features, respectively (Trevisan et al. 2012). PCA–LDA loadings provide a 'spectrum-like' graph indicating the wavenumbers at which variance between classes is most pronounced, as indicated by the peak magnitude (variance). Peak maxima are used as 'spectral biomarkers' indicative of the biological process under investigation (Kelly et al. 2011).

Exploratory analysis by way of cluster separation along LD1 was explored, to determine whether significant alterations between control and damaged/infected groups were evident. In each case, a pairwise comparison conducted of damaged, early, and late-infected tomatoes at 0, 48, and 96 h, with their shelf-life-matched controls. For a characterization of the main spectral alterations, PCA–LDA loadings, in combination with a peak-pick algorithm (20 cm^{-1} minimum separation) identifying peak maxima, was used to tabulate the top six most prominent wavenumber alterations (spectral biomarkers). Identified spectral biomarkers were given chemical assignments matched to previously characterized spectral biomarkers, considering parameters including species, tissue type, instrumentation (method, measurement

area, interrogation depth, data analysis), and biological interaction (plant–pathogen).

Group classification was evaluated using PCA in combination with a linear discriminant classifier (PCA–LDC), which tests autonomous classification accuracy based on spectral differences (Butler et al. 2017; Gajjar et al. 2013; Friedman et al. 2001). PCA–LDA and PCA–LDC were cross-validated using 10 k-folds. Further information regarding analysis of biospectroscopy data can be found at https://github.com/trevisanj/irootlab and in the literature (Trevisan et al. 2012; Kelly et al. 2011). To test for statistically significant differences in PCA–LDA scores along the primary LD

between controls, damaged, and infected tomato, LD1 scores for each biological sample were averaged and tested for significance using unpaired two-tailed t tests (GraphPad Prism).

Results and discussion

Spectral characterization of surface structures of intact tomato fruit *S. lycopersicum*

Spectra from whole tomato fruit surface structures reflect prominent biochemical components present in the cuticle and cell wall (Serrano et al. 2014). There were no differences visible in the appearance of control, undamaged tomato fruit during the 96-h analysis window (Fig. 1a–c). In contrast, damaged tomato (0 h) had a small puncture wound from the syringe at which fungal infection developed (Fig. 1d–f). Figure 2 shows the primary absorbance intensities of intact tomato fruit corresponding to Fig. 1; for the control set at 0, 48, and 96 h (Fig. 2a) and damaged (0 h), early infected (48 h), and late infected (96 h) (Fig. 2b) over the baseline-corrected and normalized ATR-FTIR fingerprint spectrum over the region 1800–900 cm^{-1}. Comparison of spectra from both control and damage/infected classes shows that the top six main vibrational bands, and chemical assignments were identical as depicted in Fig. 2 and Table 1. Absorbance intensities shown in Fig. 2 and assigned in Table 1 reflect prominent biochemical components of plant surface structures including cutin, phenolic compounds, waxes, and potentially volatile organic chemicals (VOCs) (Baldassarre et al. 2015). Several of these compounds had been identified previously from the inner or outer face of isolated tomato cuticles (España et al. 2014; Heredia-Guerrero et al. 2014), despite differences in spectral resolution and equipment used to characterize isolated tomato cuticles (España et al. 2014). This is consistent with the thickness of the cuticle during the
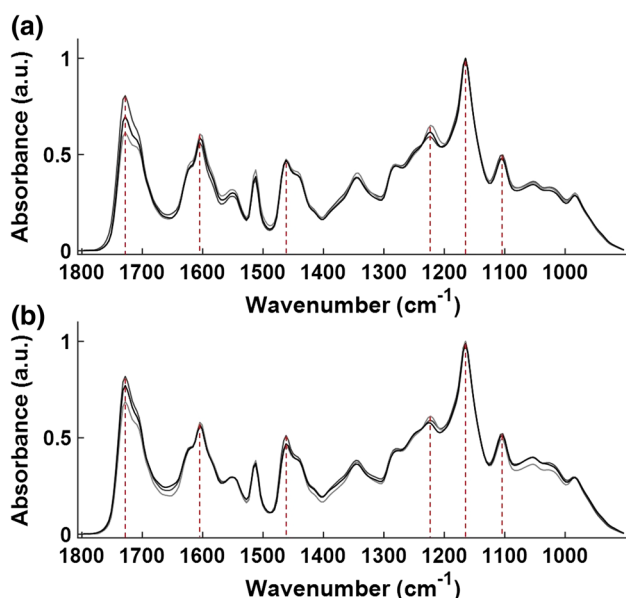

Fig. 2 ATR-FTIR spectrum of intact tomato fruit *S. lycopersicum* cv. Piccolo, over the fingerprint region (1800–900 cm^{-1}): **a** control series at 0 (light grey), 48 (grey), and 96 (black) h; **b** 0 h damaged (light grey), 48 h early infection (grey), and 96 h late infection (black)

Table 1 Primary absorbance peaks of intact tomato fruit *S. lycopersicum* cv. Piccolo

Wave-number (cm^{-1})	Vibrational mode	Biochemical assignment	References
1728	ν(C=O) ester	Cutin	Heredia-Guerrero et al. (2014), Movasaghi et al. (2008) and Ord et al. (2016)
1605	ν(C–C) aromatic	Phenolic compounds, pectin	Heredia-Guerrero et al. (2014), Movasaghi et al. (2008) and Schulz and Baranska (2007)
1462	δ(CH$_2$) scissoring	Cutin and other waxes, polysaccharides	Heredia-Guerrero et al. (2014), Movasaghi et al. (2008) and Schulz and Baranska (2007)
1223	ν_a(C–O–C) ester	Monoterpene	Movasaghi et al. (2008), Ord et al. (2016) and Schulz and Baranska (2007)
1165	ν_a(C–O–C) ester	Cutin, cellulose, pectin	Heredia-Guerrero et al. (2014), Movasaghi et al. (2008), Ord et al. (2016) and Schulz and Baranska (2007)
1103	ν_s(C–O–C) ester	Cutin, cellulose, pectin	Heredia-Guerrero et al. (2014), Movasaghi et al. (2008), Ord et al. (2016) and Schulz and Baranska (2007)

late red-ripe stage of tomato fruit (España et al. 2014; Heredia-Guerrero et al. 2014) and the shallow interrogation depth of the ATR-FTIR beam (~ 1–3 μm). Cuticle components readily identified include vibrational modes associated with the main polymer cutin at wavenumbers 1728, 1462, 1165, and 1103 cm⁻¹ (España et al. 2014; Heredia-Guerrero et al. 2014). Phenolic compounds are among other cuticle constituents that strongly absorb IR radiation, and were identified by absorption at 1605 cm⁻¹ (España et al. 2014; Heredia-Guerrero et al. 2014). We also identified an absorption peak at 1223 cm⁻¹ that is not present in isolated cuticle (Heredia-Guerrero et al. 2014) but which has been previously associated with monoterpenes, more specifically geranyl acetate, a structural component of many VOCs (Ord et al. 2016; Rodríguez et al. 2013; Schulz and Baranska 2007). Tomato fruit produce a characteristic profile of secondary metabolites including VOCs during ripening (Petro-Turza 1986; Buttery et al. 1988, 1990) and it likely that monoterpenes characteristic of VOCs present at the red-ripe stage, represent a unique contribution to the fingerprint spectrum of intact tomato fruit compared to isolated cuticle (Rodríguez et al. 2013). Alternatively, it is possible that the absorption at 1223 cm⁻¹ may simply be a broad absorption band related to the previously identified δ(OH) mode between 1246 and 1243 cm⁻¹ associated with both cutin and other polysaccharides (Heredia-Guerrero et al. 2014). Both cuticle and underlying plant layers including the cell wall have been well studied using MIR-based biospectroscopy (Heredia-Guerrero et al. 2014; Largo-Gosens et al. 2014). However, due to several caveats, the number of studies on intact, and hence physiologically competent, samples have been limited, limiting also the development of vibrational spectroscopy for applied horticulture (Skolik et al. 2018). Characterizing spectral features of tomato fruit in vivo, such as the cuticle, provides a first but important step in this endeavour and will contribute significantly to the sustainability of crop protection measures. Yet the role of the cuticle, and other epidermal structures as part of the tomato fruit skin, in post-harvest quality, shelf-life, and pathogen susceptibility remains debated especially at the molecular level, in part due to the difficult nature of this recalcitrant layer and the intimate relationship with the underlying cell wall (Domínguez et al. 2015; Lara et al. 2014). To shed light on this, analytical surface techniques such as ATR-FTIR spectroscopy are ideal, as demonstrated by the ability to measure delicate intact fruit truly non-destructively in vivo (Fig. 2). But before the molecular in vivo details can be uncovered, surface characterization of intact fruit is necessary to aid the interpretation of more subtle changes hidden in the spectral data, which can only be extracted through multivariate analysis, similar to how previous cuticle component characterization aids the interpretation of the tomato fruit skin in vivo shown in Fig. 2 (Heredia-Guerrero et al. 2014). Comparing between isolated
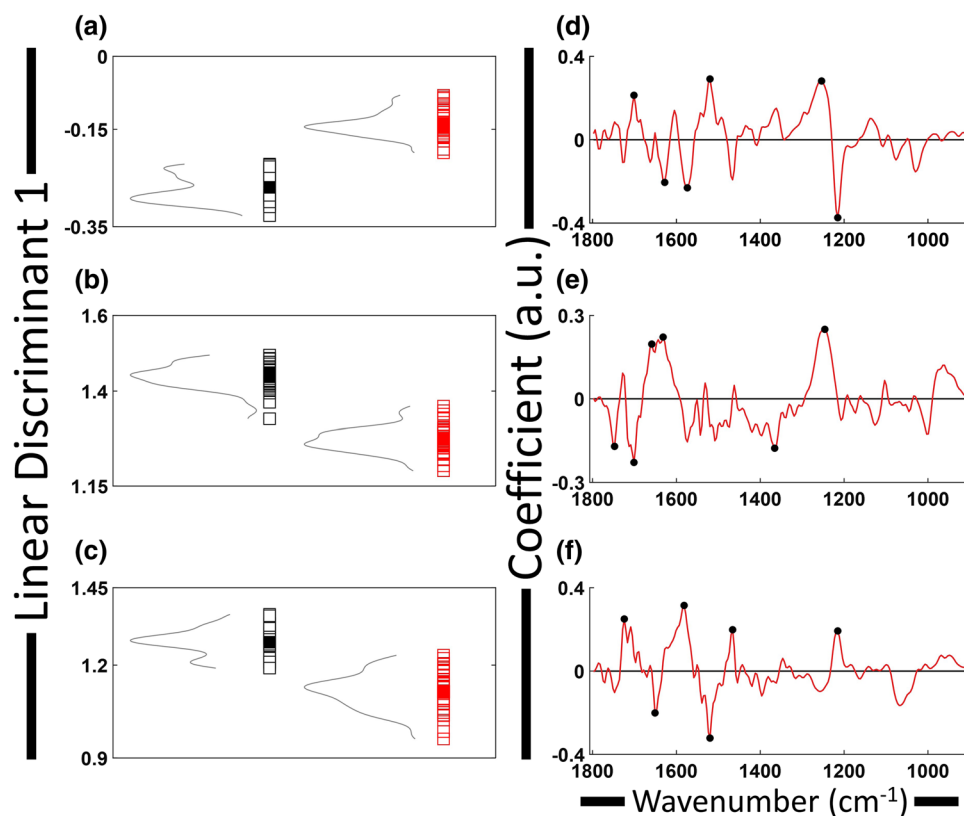
constituents and their native arrangements, in fruit or otherwise, will remain necessary to aid in the identification of candidate target compounds to serve as spectral biomarkers for varying conditions, especially dynamic physiologically driven ones, including plant–pathogen interactions. It is therefore important to characterize the candidate plant compounds being measured by MIR biospectroscopy techniques, for appropriate interpretation of spectral data from physiologically competent samples in vivo. Further, indirect detection of damage to the fruit surface (cuticle, cell wall, epidermis) and pathogens affecting crops, such as tomato which are easily compromised by damage leading to infection, would be of utmost interest for commercial development. Once characterized, changes in the MIR signature caused by abnormalities such as damage, pathogen infection, or stress, will prove useful for monitoring fruit condition as it pertains to shelf life through the post-harvest food system, thereby improving crop utilization.

Spectral alterations associated with tomato fruit damage and sour rot infection by *Geotrichum candidum*

The MIR spectrum of fruit surface structures is altered in response to damage through the stem scar and subsequent infection by *G. candidum*. Artificially damaged tomatoes (Fig. 1d) exposed to ambient conditions showed no initial signs of fungal infection after 24 h (data not shown), whilst at 48 h post-puncture (early infection) clear signs of infection were evident around the puncture site (black arrows) (Fig. 1e), and at 96 h post-puncture (late infection) substantial pathogen growth had covered the puncture site (Fig. 1f). Based on visible symptoms starting at 48 h post-puncture (Fig. 1d, f) the pathogen was determined to be *G. candidum*, a non-specific fungus known as a ubiquitous contaminant of tomato processing equipment (Thornton et al. 2010). Because the mean spectra of control and damage/infected tomatoes (Fig. 2a, b) were nearly identical, with respect to direct comparison of main vibrational bands (Fig. 2 and Table 1), PCA–LDA was employed to investigate if subtle class-specific effects were detectable between control and compromised tomatoes. This approach was intended to determine if any changes caused by damage and subsequent pathogen infection were observable indirectly without contributions from the fungus itself.

Class-specific differences for damage, early, and late infection were observed for tomato fruit compared to their healthy counterparts, as determined by multivariate analysis using PCA and LDA in tandem (PCA–LDA). Pairwise comparisons (class versus control) lead to the generation of a single LD, in this case generating three PCA–LDA score plots (Fig. 3a–c). Figure 3 shows a clear separation of clusters belonging to each paired class, indicating differences

Fig. 3 PCA–LDA 1-dimensional score plots showing class-specific cluster separation indicative of spectral differences between damaged, early, and late infection opposite their shelf-life-matched controls (**a**–**c**); corresponding loadings show specific wavenumbers responsible for clustering along LD1 (**d**–**f**)

in spectra acquired from controls and damaged, early, and late infected (Fig. 3a, b, and c, respectively). Separation along LD1 indicates significant differences within fingerprint spectra, which are specific to damage and infection. PCA–LDA score plots reveal significant data cluster separation, with statistical differences between damaged ($p = 0.003$), early infected ($p = 0.0001$), and late-infected ($p = 0.0003$) fruit, compared to their shelf-life-matched controls. This suggests that spectral changes are most pronounced for the early infected stage, showing the largest degree of separation along LD1 (Fig. 3b), followed by the late-infected stage (Fig. 3c) and damaged fruit (Fig. 3a), respectively. Loading plots (Fig. 3d–f) indicate the wavenumber regions responsible for the observed cluster separation within the PCA–LDA score plots (Fig. 3a–c) (Martin et al. 2007; Trevisan et al. 2012). Wavenumbers identified through peaks within loading pots, represent the areas with the highest degree of variance. Table 2 summarizes the top six discriminating wavenumbers identified from PCA–LDA loading plots. These top six wavenumbers, identified via the peak-picking algorithm described, are assigned as tentative spectral biomarkers responsible for the class-specific differences. Spectral biomarkers identified by PCA–LDA were considered a match if these were within ± 10 wavenumbers of those identified within the other classes. It is noteworthy that because PCA–LDA potentially extracts very subtle differences within complex tissue architectures, biomarkers

identified this way may not originate from the prominent cuticle components evident in the fingerprint spectra shown in Fig. 2, but may represent small fractions of molecules embedded in the epidermal matrix. For this reason and without extensive validation of their origin, spectral biomarkers are assigned tentatively.

Several wavenumbers identified for the various fruit conditions showed overlap, where biomarkers as discriminators for initial damage were also identified as discriminators for early and late infection. Vibrational modes at 1701, 1632–1628, 1254–1246 cm^{-1} were seen to be consistent between initial fruit damage and early *G. candidum* infection. Absorption at 1701 cm^{-1} was the only exact match between these two classes. These three wavenumber regions are assigned as carbonyl groups in fatty acid esters of cutin (1701 cm^{-1}) (España et al. 2014); carbon–carbon bonds in phenolic cuticular compounds (1632–1628 cm^{-1}) (Heredia-Guerrero et al. 2014); and hydroxyl group deformation in cutin or other polysaccharides (1254–1246 cm^{-1}) (Heredia-Guerrero et al. 2014), which are part of the epidermal surface. Alternatively, the region from 1254 to 1246 cm^{-1} has been associated with the amide III band of proteins or methylene functional groups of phospholipids (Movasaghi et al. 2008), which are also potential targets of ATR-FTIR as part of the epidermis. Consistency within spectral biomarkers was also observed between those indicative of damage and those identified within late-stage *G. candidum*

Table 2 Top six discriminating class-specific wavenumbers and tentative chemical assignments, from LD1 loading plots associated with tomato fruit damage, early infection or late infection versus control classes

Class	Wave-number (cm^{-1})	Vibrational mode	Biochemical assignment	References
Damaged	1701[a]	ν(C=O–H)	Fatty acid esters Cutin	Heredia-Guerrero et al. (2014), Ord et al. (2016)
	1628[a]	ν(C=C) ν_s(C–C) ring Amide I	Phenolic compounds Pectin Proteins	Butler et al. (2015, 2017), Heredia-Guerrero et al. (2014), Largo-Gosens et al. (2014), Ord et al. (2016)
	1574[b]	Amide II	Proteins	Movasaghi et al. (2008), Ord et al. (2016)
	1520[b]	ν(C–C) aromatic Amide II C=N or C=C	Phenolic compounds Proteins Nucleic acids?	Butler et al. (2017), Heredia-Guerrero et al. (2014), Largo-Gosens et al. (2014), Movasaghi et al. (2008), Ord et al. (2016)
	1254[a]	δ(OH) Amide III ν_a(CH$_2$)	Cutin/polysaccharides Proteins Phospholipids	Butler et al. (2017), Heredia-Guerrero et al. (2014), Movasaghi et al. (2008), Ord et al. (2016)
	1215[b]	ν_aPO$_2$ Amide III	Phosphate Proteins	Movasaghi et al. (2008), Ord et al. (2016)
Early infection	**1747**	ν(C=O) ν(C=C) Aliphatics	Polysaccharides Lipids (fatty acids) Suberin	Butler et al. (2017), Largo-Gosens et al. (2014), Movasaghi et al. (2008), Ord et al. (2016)
	1701[a]	ν(C=O–H)	Fatty acid esters Cutin	Heredia-Guerrero et al. (2014), Movasaghi et al. (2008), Ord et al. (2016)
	1659[c]	ν(C=C) Amide I	Pectin Proteins	Largo-Gosens et al. (2014), Movasaghi et al. (2008), Ord et al. (2016)
	1632[a]	ν(C=C) ν_s(C–C) ring Amide I	Phenolic compounds Proteins	Butler et al. (2017), Heredia-Guerrero et al. (2014), Largo-Gosens et al. (2014), Movasaghi et al. (2008), Ord et al. (2016)
	1366	δ(CH$_2$) ν_s(COO) δ_s(CH$_3$)	Cutin and waxes Cellulose Fatty acids Sesquiterpenes	Butler et al. (2017), Largo-Gosens et al. (2014), Movasaghi et al. (2008), Ord et al. (2016)
	1246[a]	δ(OH) Amide III ν_a(CH$_2$)	Cutin/polysaccharides Proteins Phospholipids	Butler et al. (2017), Heredia-Guerrero et al. (2014), Movasaghi et al. (2008), Ord et al. (2016)
Late infection	**1724**	ν(C=O)	Cutin and lipids Pectin/polysaccharide Phenolic ester	Butler et al. (2015, 2017), Heredia-Guerrero et al. (2014), Largo-Gosens et al. (2014), Ord et al. (2016)
	1651[c]	ν(C=C) Amide I	Pectin Proteins	Butler et al. (2015), Movasaghi et al. (2008), Ord et al. (2016)
	1582[b]	Amide II	Proteins	Movasaghi et al. (2008), Ord et al. (2016)
	1520[b]	ν(C–C) aromatic Amide II C=N or C=C	Phenolic compounds Proteins Nucleic acids	Butler et al. (2015, 2017), Heredia-Guerrero et al. (2014), Movasaghi et al. (2008), Ord et al. (2016)
	1466	δ(CH$_2$)	Cutin and waxes	Butler et al. (2015), Heredia-Guerrero et al. (2014), Largo-Gosens et al. (2014), Ord et al. (2016)
	1215[b]	ν_aPO$_2$ Amide III	Phosphate Proteins	Movasaghi et al. (2008), Ord et al. (2016)

Overlap between wavenumbers ± 10 was considered. **Bold** wavenumbers indicate class-specific wavenumbers

[a] Indicates overlapping wavenumbers between damaged and early infection

[b] Between damaged and late infection

[c] Between early and late infection

infection, specifically absorption bands at 1582–1574, 1520, and 1215 cm^{-1}. Interestingly, both absorption at 1520 and 1215 cm^{-1} were exact matches to wavenumbers related to

initial fruit damage and may therefore play a role in both damage response and response to pathogens (Table 2). Absorption bands between 1582 and 1574 cm^{-1} are strongly

associated with the amide II band of proteins (Movasaghi et al. 2008). The absorption band at 1520 cm^{-1} is potentially a shoulder region of the amide II peak but more likely associated with carbon–carbon bonds in phenolic compounds (Heredia-Guerrero et al. 2014), although this region has also been associated with alkene groups in aromatic compounds, or the imine group in nucleic acids (Movasaghi et al. 2008). Class-unique wavenumbers occur only in the early and late infection stages (upon appearance of visual symptoms). All absorption bands identified in damaged tomato occur also in either early or late infection and generate no unique absorbance peaks within the top six tentative biomarkers. Wavenumbers unique to early infection include absorbance at 1747 and 1366 cm^{-1} (Table 2). Vibrational modes at 1747 cm^{-1} are associated with double bonds in carbonyl and alkene functional groups of cutin, wax and suberin-like compounds, as well as lipids in general (España et al. 2014; Heredia-Guerrero et al. 2014). Besides compounds including cutin and waxes, also cellulose, pectin, polysaccharides, and sesquiterpenes are biomolecules, which have vibrational modes that absorb at 1366 cm^{-1} (Largo-Gosens et al. 2014; Heredia-Guerrero et al. 2014; Movasaghi et al. 2008). These spectral biomarkers appear to be unique to early infection of tomato fruit (Table 2). In contrast, late infection of tomato fruit shows specific absorbance at 1724 and 1466 cm^{-1} and are associated with carbonyl vibration of cutin, lipids, polysaccharides, or phenolic esters; and methylene vibration of cutin or other waxes, respectively (Heredia-Guerrero et al. 2014; Movasaghi et al. 2008). Taken together, these results indicate prominent changes occurring simultaneously across several compounds including lipids, proteins, and carbohydrates, many of which represent prominent components of the epidermal structure including cuticle and cell wall components.

Spectral alterations associated with tomato fruit damage are partially retained during subsequent early and late pathogen infection. Initially, tomato damage induces a wounding response, as colonization by *G. candidum* has not yet occurred (Figs. 1d, 3a, d), suggesting that the observed spectral alterations are specific to wounding. Both metabolic activity and VOC composition change in response to plant wounding; at the red-ripe stage, damage elicits changes in the VOC profile (Baldassarre et al. 2015). Wavenumbers identified as discriminators for fruit damage may therefore reflect prominent changes to the VOC profile, potentially combined with up-regulation of genes involved in defence reactions and the resulting changes in metabolism (Baldassarre et al. 2015). As VOCs diffuse through plant surface layers, their interaction with the cuticle, cell wall, or epidermis in general may produce alterations in these layers leading to the observed spectral changes (Peñuelas and Llusià 2001). Also, because damage has a direct effect on post-harvest deterioration and shelf life through various biochemical

and physiological events (Lara et al. 2014; Watada and Qi 1999), rapid damage detection of tomato fruit using spectrochemical analysis would help prevent subsequent infection and spoilage induced by spreading microorganisms such as *G. candidum.* Spectral biomarkers from the initial response to wounding are retained in part during subsequent early and late infection (Table 2). Although several spectral biomarkers are consistent between damaged and early as well as late-infected tomatoes, both early and late infection show unique spectral characteristics as well. It therefore seems plausible that changes in tomato fruit surfaces resulting from damage, share common biochemical alterations with early and late-stage infection, for example through a general stress response transitioning into a pathogen-specific response, explaining the overlap in biomarkers previously described (Table 2). As an increasing number of genetic and metabolic changes are induced by wounding and subsequent infection, the change in spectral profile likely reflects the move from damage response to plant–pathogen interaction, explaining the development of unique biomarkers at the early and late infection stages (Table 2).

Spectral alterations in plant surface structures of tomato, related to plant–pathogen interactions have been previously identified and may be related to conserved changes in epidermal surface structures in response to stress through the reactive oxygen species (ROS) network. Plant–pathogen interactions induce complex signalling networks leading to the induction of the hypersensitive response (HR) and/or systemic acquired resistance (SAR), both of which involve significant alterations to metabolism including lignification, suberization, callose deposition, changes in ion fluxes and lipid peroxidation (Camejo et al. 2016). The HR also involves the activation of programmed cell death (PCD). This is often accompanied by an oxidative burst generating ROS in the form of superoxide radical (O_2^-), hydrogen peroxide (H_2O_2) and hydroxyl radical ($\cdot OH$) accumulation (Apel and Hirt 2004; Hakmaoui et al. 2012; Suzuki et al. 2011). More generally, ROS signatures are altered in response to abiotic and biotic stresses alone and in combination (Camejo et al. 2016; Choudhury et al. 2017). Further, the oxidative burst, initiated during plant–pathogen interactions with fungi, generates essential ROS, which influence structural features of both cuticle and cell wall (AbuQamar et al. 2017). Part of the response to damage and pathogen attack, specifically at the late ripening stage is accelerated fruit softening caused by cutin depolymerization, which occurs naturally during the ripening program (Saladié et al. 2007; Brummell and Harpster 2001). Observed changes are therefore likely associated with a stress response initiated by fungal infection at a distance (in this case infection at the stem scar) and not caused by fungal released cutinases leading to cutin hydrolysis and depolymerization. To this end, the region 1750–1700 cm^{-1} has been implemented in the measurement

of cutin in tomato cuticles, with the potential to determine the degree of cutin esterification (España et al. 2014). This region was not only identified as a major cuticle component of intact tomato fruit (Table 1), but was also extracted by PCA–LDA for all classes (Table 2) making this spectral region a potentially robust biomarker indicative of spectral alterations associated with cuticle-dependent shelf-life and pathogen susceptibility. ROS signatures, or more specifically downstream targets of ROS present in epidermal surface structures such as the cuticle and cell wall, may therefore offer suitable targets for the detection and potential quantification of both abiotic/biotic stresses in various combinations (AbuQamar et al. 2017; Choudhury et al. 2017). Wavenumbers associated with the categories of damage, early, and late infection (Table 2), have also been identified as biomarkers related to abiotic and biotic stress in the epidermal surface structures of intact leaves of *Acer pseudoplatanus* (Sycamore) (Ord et al. 2016). Importantly, in this and the study by Ord et al. (2016), ATR-FTIR spectroscopy coupled with the composite technique PCA–LDA was employed emphasizing the effectiveness of this technique to extract biochemical information from dynamic biological processes. Spectral biomarkers identified in *A. pseudoplatanus* were associated with abiotic stresses caused by ozone and vehicle air pollution, as well as biotic stress caused by the tar spot leaf fungus *Rhytisma acerinum*. Changes in the cuticle and cell wall, as well as ROS signalling, are early events in the response of plants to environmental stress making it plausible that certain biochemical and biophysical changes occurring in plant surface structures in response to stress are conserved between species. Consequently, the observed alterations in the spectral signature of *A. pseudoplatanus* leaves (Ord et al. 2016) may be linked to the generation of ROS in response to stress, providing the connection between biomarkers observed here and those measured in *A. pseudoplatanus* stress response (Ord et al. 2016). This would explain the appearance of spectral biomarkers in tomato fruit related to damage and biotic stress, which have been previously associated with both abiotic and biotic stresses in surface structures of the distantly related *A. pseudoplatanus*. Although spectral biomarkers identified here in tomato match with stress biomarkers reported previously, the biomarkers occur in different combinations, which may be due to a combination of factors including inter-species differences, difference in tissue type, or differences conferred by disease (stress) specificity. Nevertheless, the identification of such a large number of spectral biomarkers point to strong commonalities between these two, different species, and suggests that spectral alterations relate to dynamic physiological changes pertaining to biotic and abiotic stress responses. While difficult to confirm through spectrochemical analysis alone, once additional data become available, the link between changes in the MIR signature, changes in epidermal structures, plant

stress, and specific signalling pathways such as ROS, will become increasingly clear.

In vivo spectral characterization of sour rot pathogen *Geotrichum candidum*

Interaction of *G. candidum* with tomato fruit in vivo appears to alter the MIR spectrum characteristic of typical fungi. The fungus *G. candidum* was measured on the tomato fruit as depicted in Fig. 1f at 96 h post-puncture. To date, MIR has been primarily used to study fungal pathogens from isolated and prepared samples (Salman et al. 2012, 2010). *G. candidum* is an economically important pathogen as it induces sour rot in many fruit and vegetable crops including tomato (Cantu et al. 2008). Its ubiquitous occurrence, as part of the human micro-biome, soil, as well as horticultural processing equipment, makes pathogenic strains of *G. candidum* a threat to crops (Thornton et al. 2010). Further, *G. candidum* can improve the conditions for infection by other pathogens thus contributing to further infection or synergistic pathogen interactions (Suzuki et al. 2014; Wade et al. 2003).

Figure 4 shows the ATR-FTIR fingerprint spectrum of *G. candidum* in vivo on tomato fruit. The main six vibrational bands of *G. candidum* in vivo are shown in Table 3. Identified vibrational bands are distinct from those of tomato fruit (Fig. 2 and Table 1) and contain several absorbance peaks consistent with those of other fungal pathogens (Salman et al. 2012, 2010). Absorbance peaks at 1639, 1547, 1404, and 1034 cm^{-1} could be assigned to a typical fungal MIR spectrum (Salman et al. 2012, 2010). In comparison, vibrational bands at 1342 and 1238, which are prominent peaks of *G. candidum* in vivo, appear to be much less pronounced or even absent depending on the fungal species under study (Salman et al. 2012). Main absorbance peaks of *G. candidum* in vivo, show vibrational modes associated with proteins between 1639 and 1342 cm^{-1} (Movasaghi et al. 2008; Salman et al. 2012) Specific absorbance peaks over this region include 1639, 1547, and 1404 cm^{-1}, corresponding to the fundamental protein vibrations amide I, amide II, and (C–N) vibration, respectively (Movasaghi et al. 2008;
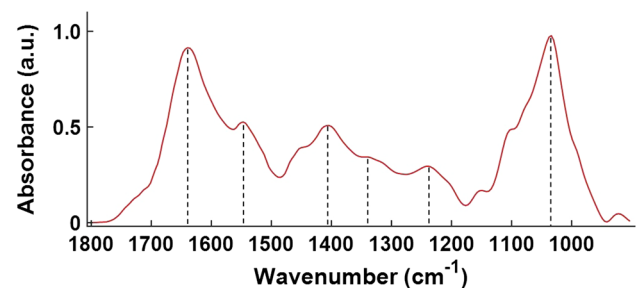

Fig. 4 ATR-FTIR fingerprint spectrum of in vivo sour rot pathogen *G. candidum* present on tomato fruit at the 96 h late infection stage

Table 3 Primary absorbance peaks of fungal pathogen *G. candidum* in vivo on tomato fruit

Wavenumber (cm^{-1})	Vibrational mode	Biochemical assignment	References
1639	Amide I	Proteins	Salman et al. (2010, 2012)
1547	Amide II	Proteins	Salman et al. (2012)
1404	ν(C–N)	Proteins	Salman et al. (2010)
1342	ν(CH$_2$)	Polysaccharides	Movasaghi et al. (2008)
1238	ν(PO$_4^{2-}$)	Nucleic acids	Movasaghi et al. (2008)
1034	ν(C–O)	Chitin	Salman et al. (2012)

Salman et al., 2010, 2012). Vibration at 1034 cm^{-1} is also readily identified as belonging to the chitin (C–O) bond (Salman et al. 2012, 2010). Absorbance at 1639, 1547, 1404, and 1034 cm^{-1} are thus all consistent with those previously characterized in fungal isolates of *Colletotrichum, Fusarium, Rhizoctonia* and *Verticillium* species (Salman et al. 2012, 2010). However, the vibrational bands identified here at 1238 and 1342 cm^{-1} do not appear to be a common constituent of other fungal pathogen isolates (Fig. 4). Phosphate (PO$_4^{2-}$) vibrational band at 1238 cm^{-1} is strongly associated with nucleic acids such as part of the DNA or RNA phosphate backbone (Movasaghi et al. 2008). Polysaccharide vibration (CH$_2$), atypical of fungi was also identified as a strong peak of *G. candidum* as part of the tomato fruit–pathogen complex (Fig. 4). It is likely that the discrepancy between the spectrum of *G. candidum* and those of other species is a result of in vivo analysis. The unique interaction between fungi and their host plants could influence the measured composition of the fungus, when compared to MIR spectra of fungal isolates, which are homogenized and taken out of their biological context. Although it cannot be ruled out that lack of sample preparation (dehydration and homogenization) prior to spectral acquisition led to a higher water content and more heterogeneous arrangement, which influenced the MIR spectrum (Fig. 4), the fundamentally different biochemical composition of fungal pathogens to that of tomato fruit is reflected in their respective MIR fingerprint spectra (Figs. 2 and 4). This fundamental difference in composition has led to the direct detection of fungal pathogens within plant tissues using differences in MIR spectral data. However, here we demonstrate that the typical fungal spectrum may have very unique features when measured intact. What remains to be seen is whether the differences in the MIR fungal spectrum arise due to simple fungal heterogeneity, or whether the interaction between plant and pathogen is the driving force for changes in its MIR fingerprint. Regardless, the characterization of pathogens in their native state, and as part of in vivo host–pathogen systems, is necessary to fully evaluate MIR for the non-destructive and rapid analysis of plant–pathogen interactions outside of the laboratory under the many variable conditions in which they occur.

Autonomous indirect detection of damage and infection based on alterations to tomato fruit surfaces

Diagnostic classifiers based on PCA–LDC are effective at detecting tomato fruit damage and infection indirectly and autonomously. PCA–LDC is one of many classifier algorithms used as training/validation datasets to evaluate the potential for autonomous classification based on MIR spectra (Butler et al. 2017; Strong et al. 2017). To evaluate the potential for autonomous detection of damage, early, and late infection, compared to healthy shelf-life-matched controls using MIR, spectra of intact tomato fruit were used as training/validation datasets for the PCA–LDC classifier (Fig. 5). Discrimination of classes using PCA–LDC has recently been applied to plant tissues with high accuracy (Butler et al. 2017). Classification of healthy controls, compared to their initially damaged but non-infected counterparts, showed the lowest observed accuracy at 78% for healthy controls, while freshly damaged tomatoes were identified correctly 83% of the time (Fig. 5a). In comparison, tomato fruit showing early signs of sour rot were accurately classified at 97%, and 92% for healthy controls at 48 h post-puncture (Fig. 5b). Late-stage *G. candidum*-infected tomatoes correctly classified 83% similar to freshly damaged tomatoes, was in contrast to the classification of control group at 96 h, which showed a classification accuracy of 96% (Fig. 5c). This was consistent with the separation observed along the primary LD for these classes (Fig. 3), as well as with the classification rates achieved by Butler et al. (2017) investigating calcium nutrient deficiency in tissues of *Commelina communis*. Interesting is the higher classification accuracy at early (97%) compared to late-stage infection (83%) (Fig. 5b, c). Late-stage infection leading to tissue breakdown and fruit softening, is likely more similar to the natural ripening process represented by control fruit at 96 h, when compared to younger control tomatoes at 48 h opposite their freshly colonized early infected counterparts. Also, the switch from damage to pathogen response may be more pronounced in comparison to shelf-life changes in younger tomatoes at 48 h. Such effects may lead to the higher classification accuracy in early infected compared to late-infected tomatoes. This is
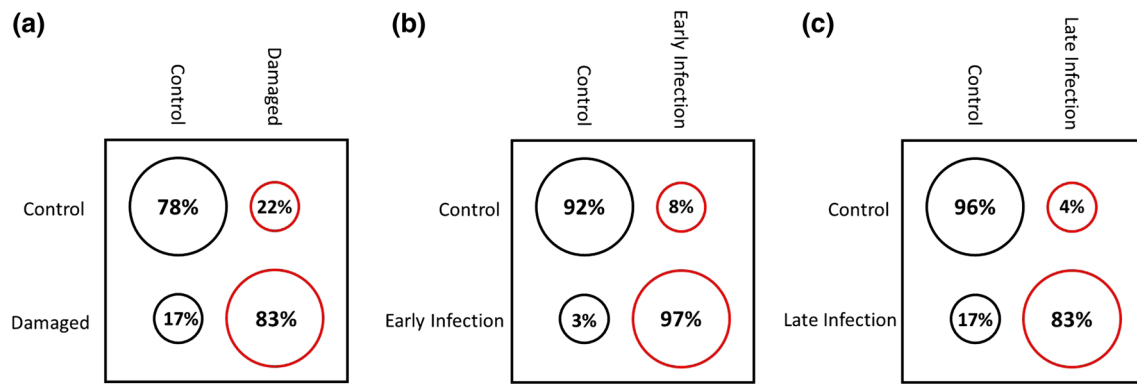
Fig. 5 Classification rates (%) of damage, early, and late infection, compared with shelf-life-matched controls, extracted from PCA input to linear discriminant classifier (PCA–LDC)

beneficial as detection of early infection is favourable over detection at later disease stages.

These collective data suggest that correct classification of infected tomato fruit improves with disease progression. Demonstrating that this is possible indirectly based only on changes in fruit epidermis not yet afflicted by pathogens, will be important to be able to detect damaged tomato fruit, prior to the development of the symptoms of fungal infection in order to reduce food waste through the repurposing of the affected crops. In addition, the early detection of fungal infection would help prevent the effects and spread of postharvest disease. It is well documented that damage to delicate fruits and vegetables leads to rapid spoilage (Tournas 2005) and thus early symptoms of damage may also serve as a pre-symptomatic indicator for imminent infection by ambient microorganisms. To this end, classifier performance may be further optimized by increasing the number of factors (PCs) fed into the LDC. For commercial development, appropriate training and test datasets would likely improve classification accuracy further. Nevertheless, preliminary classification accuracy of around 80% upwards is promising and certainly provides precedence for further development of spectrochemical analyses as a tool for crop protection.

Conclusions and future perspectives

Spectrochemical analysis combined with multivariate analysis offers a non-destructive sensor technology for the analysis of intact crops, active pathogens, and plant–pathogen interactions. Spectral characterization of intact tomato fruit showed prominent components from the cuticular layer of the plant epidermis including cutin, phenolic compounds, waxes, and VOCs. During healthy growth and plant–environment interactions, these compounds are notably modified with consequences for fruit quality and thereby provide unique groups of compounds serving as targets of dynamic

processes pertaining to crop biology. At the environmental interface, the cuticle is of specific importance due to its role, as part of the cell wall, in the determination of fruit qualities such as susceptibility to cracking and pathogen infection (Isaacson et al. 2009; Lara et al. 2014).

Multivariate analysis (PCA–LDA) can effectively discriminate healthy and compromised tomato fruit, based on damage and sour rot infection by *G. candidum*, effectively detecting pathogens indirectly. Spectral alterations in tomato fruit epidermis caused by damage and sour rot, induced changes in cuticle structure, which were assigned as tentative biomarkers. Damage, early and late-stage infected fruit thereby showed unique spectral profiles, while partial overlap of spectral markers between damage and early infection, as well as damage and late infection suggests a potential for disease specificity at these distinct stages. Disease specificity based on unique spectral markers is tentatively linked to complex and evolving stress responses. While the exact connection between spectral biomarkers of compromised tomato fruit and specific stress responses remains unclear, they are linked either directly or indirectly to plant responses such as ROS, SAR, and the HR. Clear alterations observed between healthy and damaged tomatoes further suggests the potential to identify damaged fruit prior to pathogen colonization. This may prevent disease spread, or to repurpose unmarketable specimens. Spectra of fungal pathogens and tomato fruit are fundamentally different offering direct detection of colonized pathogens within the intact fruit–pathogen complex.

Automatic detection of damage, early, and late infection through changes in fruit epidermal surface layers was evaluated based on the related classification model PCA–LDC. Indirect detection of damage and infection was shown to be effective with detection accuracy improving with disease development. Classification of tomato fruit damage and infection ranged between 83% and 97%, which may be improved through knowledge transfer, the use of more

sophisticated classification models, and trials with larger sample cohorts available to commercial growers.

Adapting spectrochemical analysis for fundamental plant science has been successful, yet more work is required to exploit the sensor potential of MIR spectrochemical analysis in complex crop systems. Herein, we demonstrate the ability to analyse individual parts of plant–pathogen complexes in vivo and show that effects of damage and infection generate unique spectral signatures reflecting common stress responses in fruits. These signatures are effective for the autonomous detection of compromised fruit crops non-destructively both directly and indirectly. This opens the door for future work, which may focus increasingly on intact or native plant systems. Portable spectrochemical analysis equipment including MIR and Raman probes are becoming increasingly available and just beginning to be explored for crop analysis (Egging et al. 2018; Farber and Kurouski 2018; Fu et al. 2016; Trebolazabala et al. 2013; Yeturu et al. 2016). Rapid developments in MIR spectrochemical analysis for plant and crop science will likely to lead to concrete large-scale applications for crop protection and production in the near future.

Author contribution statement FLM and MRM jointly initiated and led the research. FLM, MRM and PS all contributed to the research design. PS conducted experiments, analysed the data and wrote the manuscript. FLM and MRM provided feedback on data analysis and on the manuscript. All authors read and approved the manuscript.

References

AbuQamar S, Moustafa K, Tran LS (2017) Mechanisms and strategies of plant defense against Botrytis cinerea. Crit Rev Biotechnol 37(2):262–274

Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu Rev Plant Biol 55:373–399

Baker MJ, Trevisan J, Bassan P, Bhargava R, Butler HJ, Dorling KM, Fielden PR, Fogarty SW, Fullwood NJ, Heys KA, Hughes C, Lasch P, Martin-Hirsch PL, Obinaju B, Sockalingum GD, Sulé-Suso J, Strong RJ, Walsh MJ, Wood BR, Gardner P, Martin FL (2014) Using fourier transform IR spectroscopy to analyze biological materials. Nat Protoc 9(8):1771–1791

Baldassarre V, Cabassi G, Spadafora ND, Aprile A, Müller CT, Rogers HJ, Ferrante A (2015) Wounding tomato fruit elicits ripening-stage specific changes in gene expression and production of volatile compounds. J Exp Bot 66(5):1511–1526

Beuchelt TD, Virchow D (2012) Food sovereignty or the human right to adequate food: which concept serves better as international development policy for global hunger and poverty reduction? Agric Human Values 29(2):259–273

Brummell DA, Harpster MH (2001) Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. Plant Mol Biol 47(1–2):311–339

Butler HJ, McAinsh MR, Adams S, Martin FL (2015) Application of vibrational spectroscopy techniques to non-destructively monitor plant health and development. Anal Methods 7(10):4059–4070

Butler HJ, Ashton L, Bird B, Cinque G, Curtis K, Dorney J, Esmonde-White K, Fullwood NJ, Gardner B, Martin-Hirsch PL, Walsh MJ, McAinsh MR, Stone N, Martin FL (2016) Using Raman spectroscopy to characterize biological materials. Nat Protoc 11(4):664

Butler HJ, Adams S, McAinsh MR, Martin FL (2017) Detecting nutrient deficiency in plant systems using synchrotron Fourier-transform infrared microspectroscopy. Vib Spectrosc 90:46–55

Buttery RG, Teranishi R, Ling LC, Flath RA, Stern DJ (1988) Quantitative studies on origins of fresh tomato aroma volatiles. J Agric Food Chem 36(6):1247–1250

Buttery RG, Teranishi R, Flath RA, Ling LC (1990) Identification of additional tomato paste volatiles. J Agric Food Chem 38(3):792–795

Camejo D, Guzmán-Cedeño Á, Moreno A (2016) Reactive oxygen species, essential molecules, during plant–pathogen interactions. Plant Physiol Biochem 103:10–23

Cantu D, Vicente AR, Labavitch JM, Bennett AB, Powell AL (2008) Strangers in the matrix: plant cell walls and pathogen susceptibility. Trends Plant Sci 13(11):610–617

Chan KA, Kazarian SG (2016) Attenuated total reflection Fourier-transform infrared (ATR-FTIR) imaging of tissues and live cells. Chem Soc Rev 45(7):1850–1864

Chitarra LG, Van Den Bulk RW (2003) The application of flow cytometry and fluorescent probe technology for detection and assessment of viability of plant pathogenic bacteria. Eur J Plant Pathol 109(5):407–417

Choudhury FK, Rivero RM, Blumwald E, Mittler R (2017) Reactive oxygen species, abiotic stress and stress combination. Plant J 90(5):856–867

Domínguez E, Heredia-Guerrero JA, Heredia A (2015) Plant cutin genesis: unanswered questions. Trends Plant Sci 20(9):551–558

Egging V, Nguyen J, Kurouski D (2018) Detection and identification of fungal infections in intact wheat and sorghum grain using a hand-held raman spectrometer. Anal Chem 90(14):8616–8621. https://doi.org/10.1021/acs.analchem.8b01863

España L, Heredia-Guerrero JA, Segado P, Benítez JJ, Heredia A, Domínguez E (2014) Biomechanical properties of the tomato (*Solanum lycopersicum*) fruit cuticle during development are modulated by changes in the relative amounts of its components. New Phytol 202(3):790–802

Farber C, Kurouski D (2018) Detection and identification of plant pathogens on maize kernels with a hand-held raman spectrometer. Anal Chem 90(5):3009–3012

Friedman J, Hastie T, Tibshirani R (2001) The elements of statistical learning. Springer, New York

Fu X, He X, Xu H, Ying Y (2016) Non-destructive and rapid assessment of intact tomato freshness and lycopene content based on a miniaturized Raman spectroscopic system and colorimetry. Food Anal Methods 9(9):2501–2508

Gajjar K, Heppenstall LD, Pang W, Ashton KM, Trevisan J, Patel II, Llabjani V, Stringfellow HF, Martin-Hirsch PL, Dawson T, Martin FL (2013) Diagnostic segregation of human brain tumours using Fourier-transform infrared and/or Raman spectroscopy coupled with discriminant analysis. Anal Methods 5(1):89–102

Hakmaoui A, Pérez-Bueno ML, García-Fontana B, Camejo D, Jiménez A, Sevilla F, Barón M (2012) Analysis of the antioxidant response of *Nicotiana benthamiana* to infection with two strains of Pepper mild mottle virus. J Exp Bot 63(15):5487–5496

Heredia-Guerrero JA, Benítez JJ, Domínguez E, Bayer IS, Cingolani R, Athanassiou A, Heredia A (2014) Infrared and Raman spectroscopic features of plant cuticles: a review. Front Plant Sci 5:305

IFPRI International Food Policy Research Institute (2017) Global food policy report. IFPRI, Washington. https://doi.org/10.2499/9780896292529

Isaacson T, Kosma DK, Matas AJ, Buda GJ, He Y, Yu B, Pravitasari A, Batteas JD, Stark RE, Jenks MA, Rose JK (2009) Cutin deficiency in the tomato fruit cuticle consistently affects resistance to microbial infection and biomechanical properties, but not transpirational water loss. Plant J 60(2):363–377

Kazarian SG, Chan KA (2013) ATR-FTIR spectroscopic imaging: recent advances and applications to biological systems. Analyst 138(7):1940–1951

Kelly JG, Trevisan J, Scott AD, Carmichael PL, Pollock HM, Martin-Hirsch PL, Martin FL (2011) Biospectroscopy to metabolically profile biomolecular structure: a multistage approach linking computational analysis with biomarkers. J Proteome Res 10(4):1437–1448

Lara I, Belge B, Goulao LF (2014) The fruit cuticle as a modulator of postharvest quality. Postharvest Biol Tech 87:103–112

Largo-Gosens A, Hernández-Altamirano M, García-Calvo L, Alonso-Simón A, Álvarez J, Acebes JL (2014) Fourier transform mid infrared spectroscopy applications for monitoring the structural plasticity of plant cell walls. Front Plant Sci 5:303

Li J, Ying GG, Jones KC, Martin FL (2015) Real-world carbon nanoparticle exposures induce brain and gonadal alterations in zebrafish (*Danio rerio*) as determined by biospectroscopy techniques. Analyst 140(8):2687–2695

Mahlein AK (2016) Plant disease detection by imaging sensors–parallels and specific demands for precision agriculture and plant phenotyping. Plant Dis 100(2):241–251

Martin FL, German MJ, Wit E, Fearn T, Ragavan N, Pollock HM (2007) Identifying variables responsible for clustering in discriminant analysis of data from infrared microspectroscopy of a biological sample. J Comput Biol 14(9):1176–1184

Martin FL, Kelly JG, Llabjani V, Martin-Hirsch PL, Patel II, Trevisan J, Fullwood NJ, Walsh MJ (2010) Distinguishing cell types or populations based on the computational analysis of their infrared spectra. Nat Protoc 5(11):1748–1760

Martinelli F, Scalenghe R, Davino S, Panno S, Scuderi G, Ruisi P, Villa P, Stroppiana D, Boschetti M, Goulart LR, Davis CE (2015) Advanced methods of plant disease detection. A review. Agron Sustain Dev 35(1):1–25

Movasaghi Z, Rehman S, Ur Rehman DI (2008) Fourier transform infrared (FTIR) spectroscopy of biological tissues. Appl Spectrosc Rev 43(2):134–179

Oerke EC (2006) Crop losses to pests. J Agr Sci 144(01):31–43

Ord J, Butler HJ, McAinsh MR, Martin FL (2016) Spectrochemical analysis of sycamore (*Acer pseudoplatanus*) leaves for environmental health monitoring. Analyst 141(10):2896–2903

Peñuelas J, Llusià J (2001) The complexity of factors driving volatile organic compound emissions by plants. Biol Plantarum 44(4):481–487

Petro-Turza M (1986) Flavor of tomato and tomato products. Food Rev Int 2(3):309–351

Rodríguez A, Alquézar B, Pena L (2013) Fruit aromas in mature fleshy fruits as signals of readiness for predation and seed dispersal. New Phytol 197(1):36–48

Saladié M, Matas AJ, Isaacson T, Jenks MA, Goodwin SM, Niklas KJ, Xiaolin R, Labavitch JM, Shackel KA, Fernie AR, Lytovchenko A, O'Neill MA, Watkins CB, Rose JK (2007) A reevaluation of the key factors that influence tomato fruit softening and integrity. Plant Physiol 144(2):1012–1028

Salman A, Tsror L, Pomerantz A, Moreh R, Mordechai S, Huleihel M (2010) FTIR spectroscopy for detection and identification of fungal phytopathogenes. Spectrosc 24(3–4):261–267

Salman A, Shufan E, Lapidot I, Pomerantz A, Huleihel M, Tsror L, Moreh R, Mordechai S (2012) Identification of fungal phytopathogens using Fourier transform infrared-attenuated total reflection spectroscopy and advanced statistical methods. J Biomed Opt 17(1):017002

Schaad NW, Frederick RD (2002) Real-time PCR and its application for rapid plant disease diagnostics. Can J Plant Pathol 24(3):250–258

Schulz H, Baranska M (2007) Identification and quantification of valuable plant substances by IR and Raman spectroscopy. Vib Spectrosc 43(1):13–25

Serrano M, Coluccia F, Torres M, L'Haridon F, Métraux JP (2014) The cuticle and plant defence to pathogens. Front Plant Sci 5:274

Skolik P, McAinsh MR, Martin FL (2018) Biospectroscopy for Plant and Crop Science. Comprehensive analytical chemistry, vol 80. Elsevier, Amsterdam, pp 15–49

Strong R, Martin FL, Jones KC, Shore RF, Halsall CJ (2017) Subtle effects of environmental stress observed in the early life stages of the Common frog *Rana temporaria*. Sci Rep 7:44438

Suzuki N, Miller G, Morales J, Shulaev V, Torres MA, Mittler R (2011) Respiratory burst oxidases: the engines of ROS signaling. Curr Opin Plant Biol 14(6):691–699

Suzuki N, Rivero RM, Shulaev V, Blumwald E, Mittler R (2014) Abiotic and biotic stress combinations. New Phytol 203(1):32–43

Thornton CR, Slaughter DC, Davis RM (2010) Detection of the sour-rot pathogen Geotrichum candidum in tomato fruit and juice by using a highly specific monoclonal antibody-based ELISA. Int J Food Microbiol 143(3):166–172

Tournas VH (2005) Spoilage of vegetable crops by bacteria and fungi and related health hazards. Crit Rev Microbiol 31(1):33–44

Trebolazabala J, Maguregui M, Morillas H, de Diego A, Madariaga JM (2013) Use of portable devices and confocal Raman spectrometers at different wavelength to obtain the spectral information of the main organic components in tomato (*Solanum lycopersicum*) fruits. Spectrochim Acta A Mol Biomol Spectrosc 105:391–399

Trebolazabala J, Maguregui M, Morillas H, de Diego A, Madariaga JM (2017) Portable Raman spectroscopy for an in situ monitoring the ripening of tomato (*Solanum lycopersicum*) fruits. Spectrochim Acta A Mol Biomol Spectrosc 180:138–143

Trevisan J, Angelov PP, Carmichael PL, Scott AD, Martin FL (2012) Extracting biological information with computational analysis of Fourier-transform infrared (FTIR) biospectroscopy datasets: current practices to future perspectives. Analyst 137(14):3202–3215

Trevisan J, Angelov PP, Scott AD, Carmichael PL, Martin FL (2013) IRootLab: a free and open-source MATLAB toolbox for vibrational biospectroscopy data analysis. Bioinformatics 29(8):1095–1097

Wade WN, Vasdinnyei R, Deak T, Beuchat LR (2003) Proteolytic yeasts isolated from raw, ripe tomatoes and metabiotic association of Geotrichum candidum with Salmonella. Int J Food Microbiol 86(1–2):101–111

Wallner G, Amann R, Beisker W (1993) Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. Cytometry B Clin Cytom 14(2):136–143

Watada AE, Qi L (1999) Quality of fresh-cut produce. Postharvest Biol Tech 15(3):201–205

Yeturu S, Jentzsch PV, Ciobotă V, Guerrero R, Garrido P, Ramos LA (2016) Handheld Raman spectroscopy for the early detection of plant diseases: abutilon mosaic virus infecting Abutilon sp. Anal Methods 8(17):3450–3457