Advanced methods of plant disease detection. A review

To cite this version:

HAL Id: hal-01284270
https://hal.archives-ouvertes.fr/hal-01284270
Submitted on 7 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Advanced methods of plant disease detection. A review

Federico Martinelli · Riccardo Scalenghe · Salvatore Davino · Stefano Panno · Giuseppe Scuderi · Paolo Ruisi · Paolo Villa · Daniela Stroppiana · Mirco Boschetti · Luiz R. Goulart · Cristina E. Davis · Abhaya M. Dandekar

Abstract  Plant diseases are responsible for major economic losses in the agricultural industry worldwide. Monitoring plant health and detecting pathogen early are essential to reduce disease spread and facilitate effective management practices. DNA-based and serological methods now provide essential tools for accurate plant disease diagnosis, in addition to the traditional visual scouting for symptoms. Although DNA-based and serological methods have revolutionized plant disease detection, they are not very reliable at asymptomatic stage, especially in case of pathogen with systemic diffusion. They need at least 1–2 days for sample harvest, processing, and analysis. Here, we describe modern methods based on nucleic acid and protein analysis. Then, we review innovative approaches currently under development. Our main findings are the following: (1) novel sensors based on the analysis of host responses, e.g., differential mobility spectrometer and lateral flow devices, deliver instantaneous results and can effectively detect early infections directly in the field; (2) biosensors based on phage display and biophotonics can also detect instantaneously infections although they can be integrated with other systems; and (3) remote sensing techniques coupled with spectroscopy-based methods allow high spatialization of results, these techniques may be very useful as a rapid preliminary identification of primary infections. We explain how these tools will help plant disease management and complement serological and DNA-based methods. While serological and PCR-based methods are the most available and effective to confirm disease diagnosis, volatile and biophotonic sensors provide instantaneous results and may be used to identify infections at asymptomatic stages. Remote sensing technologies will be extremely helpful to greatly spatialize diagnostic results. These innovative techniques represent unprecedented tools to render agriculture more sustainable and safe, avoiding expensive use of pesticides in crop protection.

Keywords DNA-based methods · Immunological assays · Spectroscopy · Biophotonics · Plant disease · Remote sensing · Volatile organic compounds · Commercial kits

Abbreviations
ANN  Artificial neural networks
APAR  Absorbed photosynthetic active radiation
ARDRA Amplified 16S ribosomal DNA restriction analysis
AVIRIS Airborne visible/infrared imaging spectrometer
BAW Beet armyworms
BLAST Basic local alignment search tool
CDR Complementary determining regions
CMV Cucumber mosaic virus
Co-PCR Cooperative PCR
DMNT Dimethyonatriene
DMS Differential mobility spectrometry
dNTP Nucleoside triphosphates containing deoxyribose
dsDNA Double-stranded DNA
ELISA Enzyme-linked immunosorbent assay
EnMAP Environmental mapping and analysis program
EO Earth observation
EPPO European and Mediterranean Plant Protection Organization
FAIMS High field asymmetric waveform ion mobility spectrometry
FAO Food and Agriculture Organization
FISH Fluorescence in situ hybridization
GC-MS Gas chromatography mass spectrometry
ICA-PCA Independent-principal components analysis
LAMP Loop-mediated isothermal amplification
LAI Leaf Area Index
Landsat Earth Resources Technology Satellite
TM Thematic Mapper
LFM Lateral flow microarrays
M-PCR Multiplex PCR
MVA Multivariate data analysis
NASBA Nucleic acid sequence-based amplification
NIR Near-infrared wavelength
NMR Nuclear magnetic resonance
nPCR Nested PCR
PCA Principal component analysis
PFGE Pulsed-field gel electrophoresis
PCR Polymerase chain reaction
PDD Plant disease detection
PLRV Potato leafroll virus
PPV Plum pox potyvirus
PRISMA PRercursore IperSpettrale della Missione Applicativa
PTR-MS Proton-transfer-reaction mass spectrometry
RAPD Random amplified polymorphic DNA
rep-PCR Repetitive-sequence PCR
RFLP Restriction fragment length polymorphism
RS Remote sensing
RTM Radiative transfer modeling
RT-PCR Real-time PCR
SAIL Scattering by arbitrarily inclined leaves
SAM Spectral angle mapper classification
SBSE Stir bar sorptive extraction
scFv Single-chain variable fragment
SELEX Systematic evolution of ligands by exponential enrichment
SIFT-MS Selected ion flow tube mass spectrometry
SMA Spectral mixture analysis
SPME Solid-phase microextraction
ssDNA Single-stranded DNA
SNP Single nucleotide polymorphisms
SSEM Serologically specific electron microscopy
ssRNA Single-stranded RNA
STR Short tandem repeats
SVIs Spectral vegetation indices
SVM Support vector machine
SWIR Shortwave infrared wavelength
TIR Thermal infrared wavelength
TMTT trimethyltridecatetraene
TYLCD Tomato yellow leaf curl disease
UAV Unmanned aerial vehicle
VI Vegetation indices
VIS Visible wavelength
VOC Volatile organic compounds

Contents
1. Introduction ..................................
2. Traditional molecular methods for disease detection...
   2.1 Serological assays............................
   2.2 Nucleic acid-based methods......................
3. Innovative detection methods......................
   3.1 Lateral flow microarrays.........................
   3.2 Methods based on the analysis of volatile compounds as biomarkers..........................
      3.2.1 Instrumentation development for plant volatile organic compounds biomarker detection
      3.2.2 Advanced data analysis methods for plant volatile organic compounds analysis
3.3 Remote sensing of plant disease..................
   3.3.1 Definitions and theoretical basis ..............
   3.3.2 Overview of specific issues for remote sensing disease assessment
   3.3.2.1 Analysis techniques for disease detection using remote sensing
   3.3.3 Spectroscopy-based methods .............
      3.3.3.1 Non-imaging spectroscopy approaches
      3.3.3.2. Imaging spectroscopy approaches
   3.4 Potential technologies for biosensor development:
      phage display, electrochemistry, and biophotonics
      3.4.1. Phage display
      3.4.2. Electrochemistry and photonics
4. Conclusions ....................................
1 Introduction

Has agriculture driven the divergence of plant diseases or was it co-evolutionary processes in natural populations of the crops’ ancestors? Major plant taxa diverged millions of years ago, well before the time of plant domestication (Munkacsi et al. 2007). The ancient interaction between disease and future crop caused a never-ending cycle of adaptation, detection, and competition: pest control is, in fact, as old as agriculture. Early farmers competed against pathogen damage to their crops using supernatural or superstitious practices. However, it was only after the onset of modern plant pathology detection of a plant disease based on a scientific approach. The discipline was founded in 1767 by Giovanni Targioni Tozzetti (Fig. 1), who investigated the cause of a famine in central Italy and discovered *Puccinia graminis*, agent of the "rust" of wheat. At the Botanical Garden of Florence, identification of dozens of cryptogamic parasites on other plants started in 1765; at the same time, Felice Fontana, Mathieu de Tillet, and Joseph Banks studied other cryptogams at other European universities (Egerton 2012). Severe outbreaks of late blight on potato and powdery mildew on grapes inspired the new discipline. In the first half of the nineteenth century, Filippo Re and Carlo Berti Pichat started classifying plant diseases according to the symptoms they provoked. Infestations were fought with colloidal sulfur, copper salts, and lime.

There are diverse reasons why we need to estimate or measure disease on plants. Knowledge of the quantity of disease is particularly important for rapid management decisions, especially disease is closely related to yield loss. In plant breeding, crop need to be rated, in relation to resistance and susceptibility to diseases. In crop protection, it is very important to estimate plant resistance to apply pesticides in a cost-effective manner. The analysis of disease and symptom severity is extremely important for addressing fundamental questions in plant stress biology.

During the last 80 years, acceptable accuracy and precision of visual disease assessments have often been achieved using traditional disease scales. The recognition of visual symptoms is essential for diagnosis of plant diseases (Fig. 2). However, these methods are too subjective. New technologies offer opportunity to assess disease with greater objectivity (reliability, precision, and accuracy). One of these, visible light photography and digital image analysis has been increasingly used over the last 30 years, as software has become more sophisticated and user friendly (West et al. 2003). By contrast, hyperspectral imagery is relatively recent and has not been widely applied in plant pathology. Nonetheless, it offers interesting and potentially discerning opportunities to assess disease.

Vegetable crops represent an important economic segment of the agricultural production worldwide. Today, plant diseases still have both economic and social effects. The Food and Agriculture Organization (FAO) estimates that diseases,
insects, and weeds cause ~25% of crop failure. For example, enough rice to feed the entire population of Italy is destroyed by rice blast disease every year (Dean et al. 2005). A key element of crop diseases is how they spread. For instance, potato late blight by Phytophthora infestans, which caused the famous famine in Ireland, started at different locations in Europe during the 1845 epidemic, with an initial focus in Belgium (Mundt et al. 2009). The epidemic front advanced exponentially with time, and epidemic velocity increased linearly with distance (Mundt et al. 2009). This velocity has been demonstrated for other important diseases. Plant diseases threaten entire food crops worldwide, including citrus, banana, and grape. In the grape-growing regions of southwestern Europe, the disease Flavescence dorée (which is caused by phytoplasms) is widespread (Martinelli et al. 2014) and the economic consequences are immense. Infected plants must be removed and at least 5 years are required for new vineyards to reach full production. In this context, early detection of diseases is of key importance to prevent disease spread with the least damage to crop production (Yang et al. 2013). The traditional disease identification by visual assessment of plant symptoms (leaves become red or yellow and twigs stay soft) has been aided by advances in technology such as direct microscopic observation of pathogens and their manipulation in vitro. The introduction of polymerase chain reaction (PCR) by Nobel laureate Kary Mullis had a profound impact on plant disease diagnosis. While nucleic acid technology is the only choice for detecting pathogens that have not been cultured, DNA-based methods have not yet completely replaced classical microbiology and visual inspection; these three methods provide complementary information. The trend in the European Union for detecting plant pathogens, outlined in the European and Mediterranean Plant Protection Organization (EPPO) protocols, integrates phenotypic, serological, and molecular techniques. For newly discovered pathogens, it is easy to develop or adapt molecular assays within weeks of their discovery. Although nucleic acid-based techniques based on PCR and/or hybridization and biochemical assays are very sensitive, accurate, and effective for confirming visual scouting, they are unreliable as screening tests to monitor plant health status before the appearance of symptoms. They require detailed sampling procedures, expensive infrastructure, and may misrepresent the real status of infections.

In addition, these methods can be only effectively used for a restricted number of plants.

Current and future methods for plant disease detection (PDD) include proximate detection, immunological and DNA-based methods, approaches based on the analysis of volatile compounds and genes as biomarkers of disease, remote sensing (RS) technologies in combination with spectroscopy-based methodologies, and sensors based on phage display and biophotonics.

Several previous reviews address mostly biochemical and molecular methods (Hampton et al. 1990; Schaad and Frederick 2002) or applied RS techniques (West et al. 2003; Bock et al. 2010; Sankaran et al. 2010; Mahlein et al. 2012a). The present work provides information on new, alternative methods under development for effective, reliable, and early detection of pathogen infections.

2 Traditional molecular methods for disease detection

2.1 Serological assays

Because viruses cannot be cultivated ad hoc, serological assays were developed to detect them. More than a thousand other pathogens, bacteria, and fungi (Alarcon et al. 1990; Caruso et al. 2002) can now be detected using polyclonal and monoclonal antisera and techniques such as: enzyme-linked immunosorbent assay (ELISA), western blots, immunostrip assays, dot-blot immune-binding assays, and serologically specific electron microscopy (SSEM) (Van Vuurde et al. 1987; Hampton et al. 1990). Among them, ELISA, first employed in the 1970s, is by far the most widely used immunodiagnostic technique because of its high-throughput potential. The sensitivity of ELISA varies depending on the organism, sample freshness, and titre; for instance, bacteria can be detected at 100 cfu mL$^{-1}$ (Schaad et al. 2001, 2003). Antibodies need to be stored at lower than $-20^\circ$C and cannot be defrosted several times.

Polyclonal antisera for many viruses and bacteria have been developed for commercial use or research labs and have been used in numerous protocols (Nolasco et al. 2002), but their frequent cross-reactivity inspired the development of more effective monoclonal antisera using hybridoma technology (Holzloehner et al. 2013) with cell lines with specificity to single epitopes. To date, ELISA procedures using both poly- and monoclonal antibodies are available for numerous taxa of phytopathogenic bacteria; likewise, rapid detection kits are commercially available. Monoclonal antibodies recognize one epitome only and are generally more expensive. Polyclonal antibodies recognize multiple epitopes on an individual antigen (Robison 1995) and are less expensive. Unfortunately, their shelf life is short, and there is variation among batches.

2.2 Nucleic acid-based methods

Some pathogen detection methods are DNA based: fluorescence in situ hybridization (FISH) and the many PCR variants (PCR, nested PCR (nPCR), cooperative PCR (Co-PCR), multiplex PCR (M-PCR), real-time PCR (RT-PCR), and DNA fingerprinting). Others are RNA based: reverse transcriptase-PCR, nucleic acid sequence-based amplification (NASBA),
and AmpliDet RNA. All of these methods can overcome uncertain diagnosis or pathogen taxonomy, enabling a rapid and accurate detection and quantification of pathogens (López et al. 2009).

Sample preparation for molecular analysis is critical and requires reproducible and efficient protocols. Many published protocols for RNA and DNA isolation exist; all were developed to avoid the presence of inhibitory compounds that compromise detection (Louws et al. 1999). The primary compounds that inhibit DNA polymerase activity are polysaccharides, phenolic compounds, or humic substances from plants or other substrates (Minsavage et al. 1994; Mumford et al. 2006).

Several different protocols have been developed for PCR-based methods using different genomes: ssRNA, ssDNA, or dsDNA. Commercial kits specifically designed to extract nucleic acids from different types of plant material are widely used. The most noteworthy ones include the DNeasy and RNeasy Plant System from Qiagen, Ultra Clean Plant RNA and DNA isolation kits from MoBio, and the Easy-DNAPurification kit and Extract-N-Amp Plant PCR kit from Sigma. Lab-On-A-Chip devices have the potential to integrate treatment and purification using sonication (Taylor et al. 2001) or electrical devices (Gascoyne et al. 2004). The latest systems use miniaturized devices to extract DNA on a microchip using approaches such as laser irradiation or capillary electrophoresis (Lin et al. 2007). Although promising, these methods are not always effective with all types of plant material. They need to be evaluated for each combination of pathogen and plant or substrate before being adopted in routine detection (López et al. 2009).

All molecular detection methods for detecting plant pathogens are based on the accurate design of oligonucleotides and probes. Target sequences can be found using the GenBank® Nucleotide Sequence Search program provided by the National Center for Biotechnology Information (NCBI, Bethesda, MD, USA). Conserved regions for each target can be identified using the Basic Local Alignment Search Tool (BLAST), with the BLASTn program designed for analysis of nucleotides. Specific nucleotide regions are selected and primers specific for DNA or RNA targets can be easily designed. Different molecular methods are accessible to diagnose infections in open fields, orchards, or greenhouses. For example, tomato (Solanum lycopersicum L.) is one of the most common vegetable species in the world, being cultivated on about 4.6 million hectares worldwide (source FAO). This crop is nowadays cultivated in monoculture conditions, because of the continuous and uninterrupted demand made by the processing industry and consumers. This strongly favors the establishment and the increase in virulence of numerous pathogens. Today, the heightened incidence of many endemic diseases, as well as the emergence of new ones caused by fungi, bacteria, phytoplasmas, and viruses, worsens the crisis of the horticultural sector worldwide. One of the most important examples is the case study of tomato yellow leaf curl disease in Sicily (TYLCD): this disease is caused by different viral agent called begomovirus that apparently give the same symptoms in the plants. Depending on which different Begomovirus is present, the loss of production can vary from 20 to 100 %. The only way to distinguish the different begomovirus appears to be the molecular analysis (Davino et al. 2006).

PCR offers several advantages over immunoassays: the capability to detect a single target in complex mixtures, rapid and specific detection of multiple targets, and the potential to detect unculturable pathogens such as viruses or some bacteria and phytoplasma. In addition, costs of analysis are lower than 1 €/sample. In PCR-based diagnostics, primers are designed to pair with unique DNA regions from target organisms for DNA amplification and detection. Specific amplification of target nucleic acid sequences is widely used to detect and identify plant pathogens (Mumford et al. 2006). The presence of the amplification product confirms the presence of the organism in the tested sample. The original method of visualizing the amplified product through agarose gel electrophoresis with ethidium bromide (EtBr) staining has been replaced by less toxic and more sensitive SYBR GREEN detection under UV irradiation. PCR reliability and pathogen specificity has been improved by the use of dye-quenched probes (Morris et al. 1996; Thelwell et al. 2000). Generally, PCR can be performed in 2–3 h, but more advanced systems can deliver a result in minutes (Tomlinson et al. 2007). RT-PCR, with similar sensitivity to ELISA or hybridization techniques, is most practical for plant-virus detection.

Sensitivity and specificity can be increased with nPCR (Simmonds et al. 1990; Olmos et al. 1997; Pradhanang et al. 2000), performed in two steps, using one or two internal primers. This approach can lead to interference between primers of the two amplifications. Furthermore, contamination rates can increase with two rounds of amplification in different tubes, especially when the method is used routinely and on a large scale.

Co-PCR was developed for sensitive detection of plant viruses and bacteria (Caruso et al. 2003). Co-PCR, carried out in a single reaction, reduces contamination risks and has similar sensitivity to nPCR and RT-PCR. In addition, it can be coupled with dot-blot hybridization, allowing the characterization of the nucleotide sequence. However, the small volume of reagents can increase susceptibility to inhibitors, requiring higher RNA quality (Olmos et al. 2005). Coupled with colorimetric detection, the sensitivity for virus detection is at least 1,000-fold higher than that obtained with RT-PCR and is similar to that of nested RT-PCR.

M-PCR is a variant of PCR in which two or more target sequences are simultaneously amplified in the same reaction (López et al. 2009). M-PCR is useful because several
paths frequently infect a single plant (Duvino et al. 2012; Panno et al. 2012, 2014). The compatibility of the designed primers must be evaluated experimentally. Multiplex nested RT-PCR in a single tube reduces time and cost while allowing simultaneous detection of targets, melding the advantages of M-PCR and nPCR. Like M-PCR, it requires experimental verification of primer compatibility and an in silico analysis of the primers. Unlike conventional PCR, RT-PCR allows pathogen quantification and monitoring of the reaction while in progress. This high-throughput technique achieves high speed, specificity, and reliability, and many specific protocols have been developed. RT-PCR overcomes the risk of cross-contamination through modified primers and different labels of primers in combination with probes. Even RT-PCR must be adapted for variables such as primer design, reaction components, and conditions. It works well with small amplicons (50–200 bp) and at lower magnesium, primer, and dNTP concentrations than conventional PCR. Different chemistries are commercially available such as SYBRGreen, TaqMan, Scorpion, and Molecular Beacons. Primer design and probe type must be optimized for each assay. TaqMan is the most used real-time system, since it can discriminate sequences that differ by only 1 bp.

DNA fingerprinting is a molecular genetic method for identification of individuals based on unique patterns (polymorphisms) in their DNA, using hair, blood, semen, or other biological samples. When first described in 1984 by Alec Jeffreys, the technique focused on DNA sequences called mini-satellites that contained repeating patterns with no known function. These sequences are unique to each individual, with the exception of identical twins. This approach was adapted to detect plant disease.

Different DNA fingerprinting methods use either restriction fragment length polymorphism (RFLP), PCR, or both, to target areas of DNA with known variations in single nucleotides (single-nucleotide polymorphisms (SNP)), short tandem repeats (STR), or other repeating polymorphic regions. The odds of identifying an individual correctly depend on the number and size of repeating sequences tested.

Several fingerprinting methods are in use; most use PCR for fragment detection. They include pulsed-field gel electrophoresis (PFGE) (Grothues and Rudolph 1991), RFLP (Scholz et al. 1994), random amplified polymorphic DNA (RAPD) (Clerc et al. 1998), repetitive-sequence PCR (rep-PCR) (Little et al. 1998), amplified 16S ribosomal DNA restriction analysis (ARDRA) (Scortichini et al. 2001), and amplified fragment length polymorphism (AFLP) (Clerc et al. 1998; Manceau and Brin 2003). The choice of technique depends on the application, such as identification, DNA marker mapping, the organism under study, and the plant/pathogen system. Ideally, a fingerprinting technique should require no prior investment in sequence analysis, primer synthesis, or characterization of DNA probes. Presently, AFLP analysis is one of the most discriminating genomic methods to distinguish among phytopathogenic microorganisms (Vos et al. 1995). AFLP detects genomic restriction fragments by PCR amplification, and can be used for DNA of any origin or complexity. Fingerprints are produced without prior sequence knowledge using a limited set of generic primers. The number of fragments detected in a single reaction can be “tuned” by selection of specific primer sets. The AFLP technique is robust and reliable because stringent reaction conditions are used for primer annealing: the reliability of RFLP is combined with the power of PCR. To reduce the complexity of the original AFLP technique, the protocol was recently modified by the introduction of fluorescent dye-labeled primers and automated DNA sequencers for data capture. This improved method (fAFLP) has been used successfully to identify and/or type bacterial species (Manceau and Brin 2003; Cirvilleri et al. 2007a, b).

NASBA is commonly used to amplify RNA sequences. It was developed in the early 1990s for continuous amplification of nucleic acids in a single mixture at a single temperature (Compton 1991). NASBA does not require a thermal cycler, only a water bath. This technique has been used to detect viruses (Klerks et al. 2001; Olmos et al. 2005) and bacteria (Rodriguez-Lázaro et al. 2006; Scuderi et al. 2010). NASBA amplifies only RNA, so no DNase treatments are needed. This technique can selectively amplify mRNA sequences in a background of genomic DNA and can be used to specifically detect viable cells. The use of a hybridization device system (i.e., Hybrimax, Hybrio Limited) can simplify the time and reagents employed to develop the amplicons (Olmos et al. 2005). Target sequences were detected with good speed and specificity.

Loop-mediated isothermal amplification (LAMP) detects amplicons via photometry for solution turbidity (Mori et al. 2001). With SYBR Green, a color change can be seen without equipment. LAMP can be used easily as a simple, rugged screening assay and eliminates the need for expensive thermocyclers. It has been used to detect plant viruses such as PPV (Varga and James 2006).

In the early 1990s, DNA arrays caused a revolution in nucleic acid detection. In the last 2 years, improved image scanning allowed support miniaturization and increased spot density. The resulting “microarrays” or chips are only a few square centimeters and can include hundreds of thousands of probes representing an organism’s entire genome or transcriptome. Oligo DNA microarrays have been used extensively for PDD, using different oligo lengths and techniques to print the spots. Manufacturers like Affymetrix use a photolithography strategy to synthesize the oligo directly on the chip (Singh-Gasson et al. 2000). The length of the oligos varies from 25 to 70 bp. Generally, long oligos have higher sensitivity, but short oligos better recognize nonspecific hybridizations (Bates et al. 2005). Several international projects have developed diagnostic microarrays for plant pathogens. Microarrays can also multiplex diagnosis of multiple
Advanced methods of plant disease detection

pathogens. The technique allows different steps in molecular diagnosis such as nucleic acid extraction, PCR reactions, and detection to be performed directly on the array (Liu et al. 2007; van Doorn et al. 2007), with the potential for automation.

Many factors drive the choice of a molecular technique: the available budget, time of analysis, and the number of investigated species (López et al. 2009). Such methods are highly efficient and specific, but they have weaknesses: uneven pathogen distribution inside plants, particularly woody perennials, can render molecular tools unreliable, particularly at presymptomatic stages. In materials such as seeds, insect vectors, water, and soil, pathogen titers are often below the sensitivity limit of these methods that is usually 10–100 pg. False negatives can also occur when the DNA target sequence is degraded or reagents are of insufficient quality (Louws et al. 1999). Small sample sizes may misrepresent the real situation and sample cross-contamination can give false positives. The sensitivity of PCR-based methods is often lower than expected due to inhibitors of transcriptases and/or polymerases. PCR can amplify nonspecific products and artifacts due to mispriming or primer dimerizations and dead pathogens can give false-positive results; this last is particularly relevant when analyzing quarantine pathogens. Finally, the cost of equipment and reagents must be considered when selecting a molecular detection method.

3 Innovative detection methods

The potential to detect different infections in the same plant is desirable, since cultivated plants can be affected simultaneously by many pathogens, such as nematodes, fungi, bacteria, phytoplasmas, viruses, and viroids that traditional methods detect only at a late, symptomatic stage. Relatively novel approaches may offer pathogen detection that is easy, rapid, often inexpensive, and reliable at presymptomatic to early spread stages (Fig. 3). Conversely, novel methods allow detection of pathogen primary infections when symptoms are unclear and present in only few plants.

3.1 Lateral flow microarrays

Lateral flow microarrays (LFM) allow rapid, hybridization-based nucleic acid detection using an easily visualized colorimetric signal (Carter and Cary 2007). These arrays are built on miniaturized lateral flow chromatography nitrocellulose membrane, hybridize in minutes, have detection limits similar to microarrays, and can reduce the need for expensive laboratory instruments. The technology depends on the availability of strong and reliable host and pathogen biomarkers discovered through transcriptomic approaches (Martinelli et al. 2012a, 2013a). Metabolomics is widely used to identify key plant metabolites of primary and secondary metabolism usable as biomarkers for different environmental stresses or pathogen infections (Rizzini et al. 2010; Tosetti et al. 2012; Martinelli et al. 2012b, 2013b, 2014; Ibanez et al. 2014). An integrated omic approach can identify early pathogen infections such as Huanglongbing disease in citrus (Dandekar et al. 2010). Highly interactive proteins such as heat shock proteins or dehydrins, upregulated by different environmental factors (Natali et al. 2007), are possible indicators of plant health status.

3.2 Methods based on the analysis of volatile compounds as biomarkers

Plants emit many volatile organic compounds (VOCs) into their immediate surroundings that serve essential functions in growth, communication, defense, and survival (Baldwin et al. 2006). VOCs emitted from leaf surfaces are terminal metabolites of the host plant and can indicate its physiological health status. VOCs are low molecular weight biomolecules with a high vapor pressure and low boiling point. They readily exist in the gaseous phase under standard temperature and pressure and are generally present at ultra-low concentrations below the human olfactory threshold. VOC profiling is an emerging field with potential for immediate applications within the plant sciences; the ability to rapidly, frequently, and noninvasively monitor the health status of high-value commodity crops is highly desirable to growers. A new avenue of research is opened by VOC profiling, which may detect mechanisms for “plant-to-plant” and “plant-to-insect” communication, gaining new insights into host responses to pathogens and abiotic stressors. Novel analytical methods, instrumentation, and multivariate data analysis (MVA) methods are required to make and interpret these data sets; the development of such hardware and software tools is critical to bring these concepts into the field. The following section highlights several studies involving VOC profiling of plants using gas chromatography mass spectrometry (GC-MS).

In the late 1990s, VOCs profiling was applied to peanut plants (Arachis hypogaea). The emitted VOC profiles of healthy controls were significantly different than those infected with white mold (Sclerotium rolfsii); the major differences were in concentrations of methyl salicylate and 3-octanone (Cardoza et al. 2003). VOC markers such as hexenols, hexenals, hexenyl esters, and classes of terpenoids and indoles were detected, demonstrating the potential of VOC profiling to rapidly discriminate between fungal infection and insect vector feeding. The study also investigated the feeding behavior of beet armyworms (BAW) on healthy and white mold-infected leaves; the BAW preferentially consumed more infected leaves. Furthermore, the BAW naturally emitted trace amount of hexenyl acetate, linalool, and methyl salicylate, which retarded the growth of S. rolfsii. The emission of induced volatiles by infected plants, specifically methyl salicylate and 3-
octanone, thus attracting insect vectors, is a direct host response to fungal infections (Cardoza et al. 2002). This study highlights the complex interplay between host VOC responses to various stresses and pathogen infections, and the interaction of these signals with insects and herbivore predators.

VOCs profiling of potato tubers inoculated with late blight (Phytophthora infestans) and dry rot (Fusarium coeruleum) after harvest identified 52 volatiles. The most abundant compounds were benzothiazole, 2-ethyl-1-hexanol, hexanal, 2-methyl propanoic acid-2, 2-dimethyl-1-(2-hydroxyl-1-methylethyl)-propyl ester, 2-methylpropanoic acid-3-hydroxy-2,4,4-trimethyl-pentyl, various esters, and phenol. The differences in the emitted VOC profile of potato tubers inoculated with P. infestans and F. coeruleum were evident and might provide an early warning VOC system for postharvest disease in potato (de Lacy Costello et al. 2001), aiding postharvest cash crop management.

Cucumber mosaic virus (CMV) has also been investigated using VOC profiling. Cultivated squash (Cucurbita pepo cv. Dixie) were used as a model system to investigate the mode of action of the virus, the subsequent induced host response, and resulting downstream effects upon insect vectors. To account for variation in growing conditions and environment, both greenhouse and field-grown plants were used. CMV-infected plants showed an overall net increase in VOCs. (E)-2-hexenal, 6-methyl-5-hepten-2-one, (E)-β-ocimene, methyl benzoate, linalool, 4-ethyl-benzaldehyde, (Z)-3-hexen-1-yl butyrate, (Z)-3-hexen-1-yl 3-methylbutyrate, (E)-2-decenal, ethyl acetophenone, 3,5-dimethyl-1,2,4-trithiolane, citronellyl propionate, β-selinene, and (Z) jasmone were present in all plants, but no major qualitative difference in VOC profiles could be identified in infected plants (Mauck et al. 2010). Both insect vectors, Aphis gossypii and Myzus persicae, were preferentially attracted to CMV-infected plants, as with BAW to mouldy peanut plants, despite the smaller size and inferior quality of CMV-infected plants. The altered VOC profile emitted by CMV-infected plants may effectively mimic that of healthy plants. This demonstrates that the plant is inducing altered VOCs profile in response to viral infection, a mechanism known as “super normal stimulus” (Dawkins and Krebs 1979).

A significant body of work addresses emitted and induced VOC host response to plant pathogen infections. Several VOC markers are associated with host response, independent of plant species and disease type: (Z)-3-hexenol, methyl salicylate, (E)-β-ocimene, linalool, (E)-β-farnesene, (E)-4,8-dimethyl-1,3,7-triene (DMNT), and (E,E)-4,8,12-trimethyl-1,3,7,11-tridecatetraene (TMTT). While this list is not complete or exhaustive, no single VOC marker has been associated only with one specific plant disease (Jansen et al. 2009, 2011).

### 3.2.1 Instrumentation development for plant VOCs biomarker detection

One routinely measured plant VOCs marker is ethylene, a low molecular weight biomolecule that regulates a multitude of...
chemically in plants, from seed germination to organ senescence. Perhaps the most significant property of ethylene is its role as a trigger for fruit ripening (Bleecker and Kende 2000). Modern miniaturized ethylene sensors are limited in their accuracy and sensitivity; better detectors do exist, but are currently too large to be routinely used in crates or shipping containers and cost several thousand dollars per unit. Further development of ethylene sensors could greatly aid postharvest monitoring of agricultural commodities.

While single molecule targeted analysis is possible, it is not widely adopted because it provides a small fraction of the total available chemical information. Untargeted VOC profiling provides a more global perspective of the plant system under investigation; this can be very effective when many different classes of VOCs must be detected and quantified simultaneously. GC-MS is the analytical instrument of choice for VOC profiling (Lytvovenko et al. 2009), because it uses a gas phase separation and detection system to provide comprehensive structural and chemical information on individual components of VOC mixtures. Other spectrometric-based techniques for VOC profiling include proton-transfer-reaction mass spectrometry (PTR-MS) (Fehsenfeld et al. 1966; Ferguson et al. 1969; Blake et al. 2009) and selected ion flow tube mass spectrometry (SIFT-MS) (Spanel et al. 1996; Smith and Spang 2005); both of these provide near-real-time profiling and measurement of VOC samples.

Current MS-based systems are bulky and insufficiently robust for direct, in-field analysis but could be practical if significantly miniaturized and housed within suitable protective enclosures with adequate support infrastructure. VOCs are currently collected using various methods, and the samples are then sent to the laboratory for chemical analysis and characterization. VOCs can be collected through passive or active sampling. Passive sampling uses small, functionalized polymer traps like solid-phase microextraction (SPME) (Arthur and Pawliszyn 1990; Zhang et al. 1994; Pawliszyn 1997) or stir bar sorptive extraction (SBSE) (Baltussen et al. 1999, 2002). For SPME, different functionalized surfaces and film thickness are available for enhanced trapping of particular VOC classes such as volatile, semivolatile, polar, or nonpolar compounds. Alternatively, VOCs can be actively sampled by collecting the immediate headspace surrounding the plant several times and passing the VOCs through high-capacity sorbent traps such as Tenax® TA or Carbopack® D for sample preconcentration prior to chemical detection and analysis. Active sampling requires a specialized thermal desorption introduction system to transfer and focus the samples into the GC/MS system due to the large surface area of sorbent traps during desorption.

Alternative portable chemical detection systems are currently being explored for VOCs profiling. “High Field Asymmetric waveform Ion Mobility Spectrometry” (FAIMS) (Shnaydeman et al. 2005), or “Differential Mobility Spectrometry” (DMS), can be attached to a MS system (Krebs et al. 2006). They can also function as standalone separation and detection devices (Aksenov et al. 2012). DMS can be employed directly for in-field VOC analysis as it has low power requirements, less support infrastructure, and is portable. DMS systems offer sensitivity down to $10^{-9}$ (parts per billion) or $10^{-12}$ (parts per trillion) (Shnaydeman et al. 2005) and high-throughput near-real-time measurements but lower resolution than MS. The addition of a chromatographic separation system such as GC can vastly extend their analytical capabilities and has been used successfully to analyze complex bacterial samples (Shnaydeman et al. 2005; Cheung et al. 2009) and diagnose citrus disease (Aksenov et al. 2014).

The electronic nose (e-nose) is another platform for VOCs profiling. These systems use an array of specialized metal oxide sensors, each with selectivity toward particular classes of VOCs. When volatiles are introduced into the e-nose, an impedance response is generated from individual sensors, which are measured and displayed simultaneously (Gardner and Bartlett 1994; Doleman et al. 1997; Fend et al. 2006). E-nose can rapidly generate an overlay spectral response of a VOC mixture in near-real time, where the entire VOC profile is used without identifying individual components. E-noses have been used successfully to correlate specific physiological changes in plant health status (Laotawornkitkul et al. 2008). E-nose is a maturing technology, and issues with reproducibility, resolution and robustness remain to be addressed. The data generated by e-nose are extremely complex, requiring processing via multivariate statistical methods to be accurately and correctly interpreted. The future use of e-nose systems for in-field analysis holds promise, if these limitations are addressed.

Important technological development in sensors such as DMS, PTR-MS, SIFT-MS and e-nose highlights the need for specialized and sensitive instrumentation to analyze VOCs from biological systems. Continued development in this area will undoubtedly help and facilitate advances in the field.

3.2.2 Advanced data analysis methods for plant VOCs analysis

Data preprocessing and MVA are critical components of VOC profiling: essential tools for reliable and accurate identification and quantification of analytes. VOCs profiles are highly complex and present at ultra-low concentrations, often within a complex background. The data generated from GC/MS-based systems have a high dimensionality to them and require a suitable data analysis pipeline to accurately process the data for modeling and hypothesis testing such as identification of discriminating factors between healthy and infected plants. MVA is a four-step process: (1) Defining the number of classes and replicates required for initial experiments,
including an appropriate, balanced sample size and the number of replicates required for statistical significance (Hendriks et al. 2011). (2) Optimizing experimental parameters for maximum chemical information with the fewest required experiments (Trygg et al. 2007). (3) Developing a suitable data preprocessing pipeline and follow-up data analysis methodologies. The overall aim is to correct for any instrumental drift over time, identify and remove irrelevant information or artifacts from the data set, and detect outliers, all necessary to ensure that the data sets are of sufficient quality prior to MVA and trend visualization. Lastly, (4) applying different multiway regression algorithms to calibration- or classification-based problems (Marco and Gutiérrez-Gálvez 2012). Validation of results is required to determine the overall stability of the model (Szymańska et al. 2012) via independent training, test and validation data sets.

Of course other statistical and higher order numerical analysis approaches may be used to interpret VOC sensor outputs. The GC/DMS and other VOC sensors output complex data sets that are very rich in information, and it is appropriate to consider mining that data for latent trends and associations with biological phenomena. To reveal the rules that control complex systems, data mining and machine learning have been widely used in a variety of biofields such as gene sequence analysis, noninvasive medical instruments, and medical image analysis. Because of an exponential increase of data/information from novel sensor platforms and the complexity of the observed biological systems, it is clear that advanced and innovative data mining approaches are now poised to make significant and indispensable contributions to understanding many fundamental biological problems.

Data mining strategies provide powerful tools for us to concentrate information, prevent possible disturbance from noise contamination and unrelated variables, and establish a model to reveal complex biological system mechanisms. There have been reports of progress in this area for GC/DMS (Aksenov et al. 2012), but we still face challenges to efficiently apply data mining and machine learning approaches to large biological data systems. Unless a proper modeling strategy is employed, both sample number and sample dimension can significantly delay and hinder a knowledge discovery process. Particularly, an overly large sample dimension can result in a “curse of dimensionality” problem in the data mining and machine learning fields (Zhao et al. 2008; Zhao and Davis 2009). Also, increasing sample number does not necessarily increase model accuracy, as it is almost unavoidable to collect noise contaminated or unrelated samples. Therefore, using data mining strategies to select pertinent features and representative (core) samples becomes a key issue for revealing the rules that control complex biological processes.

Historically in the artificial intelligence field, we find that artificial neural networks have achieved great successes in many disciplines by mimicking the human brain’s learning processes. Many true biological systems (including the human brain) are exceptionally well-tuned for learning and self-adjusting, which allows them to survive in a competitive and harsh environment. It is this adaptive nature that is at the heart of all of these properties, and this strongly motivates us to develop and propose novel algorithms that mimic the mechanisms of biosystems as a promising way to reveal mechanisms of complex and large biological systems.

VOCs profiling is still an emerging field of analysis. Instrumentation and equipment are still being developed and refined, further advances in sampling allow greater coverage of VOCs to be collected, more sensitive instruments with lower detection limits and faster scan rates detect ultra-low abundance compounds, and suitable data analysis methods for data interpretation are all essential for further development of the field.

3.3 Remote sensing of plant disease

3.3.1 Definitions and theoretical basis

RS is a technique for obtaining information on an object without physical contact, by measuring the electromagnetic energy reflected/backscattered or emitted by the surface of the Earth (De Jong and Van der Meer 2006). As a noncontact technique, we include in the definition of RS also spectral measurements acquired by portable instruments such as handheld spectroradiometers (also called proximal sensing). These measurements are processed and analyzed to retrieve information on the object observed (i.e., plant health, in this case). RS is an indirect assessment technique, able to monitor vegetation conditions from distance, and evaluate the spatial extent and patterns of vegetation characteristics and plant health, in this application. Sensors can be distinguished into active or passive whether they emit artificial radiation and measure the energy reflected or backscattered (active sensors) or they measure the reflected solar radiation or the emitted thermal radiation (passive sensors). Radar and LIDAR are examples of active RS instruments. Here, we discuss the use of passive instruments which can measure the solar radiation reflected in the visible (VIS; wavelength range, 400–700 nm), near-infrared (NIR; wavelength range, 700–1,100 nm), and shortwave infrared (SWIR; 1,100–2,500 nm), and the energy emitted in the thermal infrared (TIR; 3 to 15 μm) wavelength regions of the electromagnetic spectrum. Passive instruments, for their specific characteristics, are employed in the vast majority of RS PDD applications.

The spectral signature of vegetation is influenced by variables describing canopy structure (leaf area and orientation, spatial arrangement, and roughness) and on the optical, dielectric, or thermal characteristics of the vegetation elements.
A plant which is in stressed conditions (induced by the disease) reacts with protection mechanisms that lead to suboptimal growth which show up as changes in variables such as leaf area index (LAI), chlorophyll content, or surface temperature; thus producing a spectral signature that is different from the signature of healthy, unstressed vegetation. Meroni et al. (2010) provided a clear and exhaustive description of the effects of stressors on a plant’s physiology and of how RS can detect early or chronic changes induced by them. Different sources of plant stress can be monitored by RS by analyzing changes in radiation harvesting and use by the plant. The total energy absorbed by the plant (absorbed photosynthetic active radiation (APAR)) is determined by the plant’s total leaf area and by the concentration of pigments (chlorophyll). A healthy plant uses APAR primarily for photochemical reactions (0–20 %) and dissipates the rest as heat (75–90 %) and fluorescence (2–5 %) (Meroni et al. 2010).

When plants are exposed to pathogens they activate defense responses whose molecular mechanisms are very complex. At the early stages, when visual symptoms such as lesions on the leaf surface are not present, plants react to the presence of a pathogen with physiological mechanism such as the reduction of the photosynthesis rate, which induces an increase of fluorescence and heat emission (West et al. 2003). For example, Polischuk et al. (1997) exploited spectral reflectance measurements for an early diagnosis of symptoms in *Nicotiana debneyi* plants at different stages of tomato mosaic tobamovirus infection and observed a decrease in leaf reflectance due to a reduction of chlorophyll content 10 days after inoculation; for the same experiment, visual symptoms of the presence of the pathogen were observed only 2 weeks later. Lorenzen and Jansen (1991) obtained similar results on barley leaves infected by cereal powdery mildew. Several studies investigated changes produced on leaf fluorescence by pathogens on different species (Daley 1995; Peterson and Aylor 1995). Generally, changes in fluorescence response do not provide unambiguously an indication of a specific stress, but it nevertheless can provide information able to anticipate nonnormality conditions, as disease symptoms are (West et al. 2003). Finally, the presence of stress factors changes the thermal properties of plants, which in turn influence the radiation emitted in the TIR domain of the spectrum, mainly produced by changes of the water content of leaves (Pinter et al. 1979; Mottram et al. 1983) which can also be detected at the early stage of the disease (Chaerle et al. 1999; Omasa 1990; Costa et al. 2013).

At a later stage, pathogens cause a reduction of leaf plant chlorophyll content due to necrotic or chlorotic lesions that increase reflectance in the VIS and cause a shift of the red-edge position in the spectrum. Browning effects by senescence of infected spots on the leaf influence the VIS and NIR regions due to dryness. At the canopy scale, infection can change canopy density and leaf area which can be observed in the NIR (Franke and Menz 2007). In summary, RS can provide information on both physiological processes ongoing under stress conditions, such as passive fluorescence effects or thermal energy dissipation, and plant parameters (leaf pigments, water content, and chlorophyll content) that indicate the actual condition. This section focuses on the second category, and in particular on RS assessment of plant disease due to biotic factors. During the last decades, scientific publications have described the capability and potential of RS approaches for PDD, from the early works of Nutter (1989) and the overview of capabilities shown by West et al. (2003) to the seminal review of methods performed by Bock et al. (2010), and more recently by the works of Sankaran et al. (2010), Mahlein et al. (2012b), and Prabhakar et al. (2012).

The RS scientific community defines plant disease monitoring as: detection (deviation from healthy), identification (diagnosis of specific symptoms among others and differentiation of various diseases), and quantification (measurement of disease severity, e.g., percent leaf area affected) (Mahlein et al. 2012a). Different sensors and techniques are required for detecting plant response to various diseases and disease severity. The ability of RS to diagnose plant disease and severity are shown in work describing the effect of disease severity on the spectral response of tomato plants (Zhang et al. 2003, Fig. 4).

### 3.3.2 Overview of specific issues for RS disease assessment

As noncontact techniques, RS allow observing the effect of the pathogen on the plants’ functionality and characteristics rather than the isolation of the pathogen molecules as with traditional techniques, which primarily aim at observing and identifying the pathogen and its direct effect on the plants. These features are determined by three factors: (a) observation conditions (RS measurements), (b) observation scale, and (c) the sensors and technology used. RS measurements of the spectral properties (optical and thermal) can be carried out on leaves and/or plants grown under controlled conditions in laboratory or greenhouses and climate chamber (Oerke et al. 2006; Costa et al. 2007; Purcell et al. 2009; Mahlein et al. 2012a; Bauriegel et al. 2011), as well as on leaves, plants, and/or canopies grown in the field where uncontrolled external factors (e.g., climate, nutrient deficiency) also influence vegetation conditions (Grisham et al. 2010; Malthus and Madeira 1993; Naidu et al. 2009; Bravo et al. 2003; Reynolds et al. 2012; Hadjiloucas et al. 2009; Lichtenhaler et al.1996; Camargo and Smith 2009; Agati et al. 2013; Bock et al. 2008), while some studies address both conditions. Measurements can be taken at a distance from which single leaves cannot be observed, often with imaging sensors mounted on elevated platforms, drones (Fig. 5; Torres-Sánchez et al. 2013), aircraft (Hillnhütter et al. 2011; Huang et al. 2007), or satellites (Seiffert and Schweizer 2005;
Johnson et al. 2003; Mewes et al. 2011; Mirik et al 2013). With the increasing altitude of acquisition of spectral measurements, issues such as atmospheric effect, pixel heterogeneity, and acquisition geometry have to be taken into account when analyzing the signal.

Observation conditions are strongly connected to the observation scale (Woodcock and Strahler 1987). While ecologists and environmental scientists are usually aware that spatial patterns (heterogeneity/homogeneity) of environmental phenomena are strongly scale dependent (upscaling), such issues can be overlooked in other branches of the natural sciences. Understanding upscaling mechanisms and spatial patterns of disease phenomena is crucial for operational disease assessment through spatialized techniques like RS (Hufkens et al. 2008). Knowledge of the dynamics that regulate the spatiotemporal heterogeneity of plant diseases is not as advanced as theoretical detection of diseases based on spectral features in the laboratory, though they are topics of equal importance (Anderson et al. 2004). Three scales of observation can be identified: (i) the single leaf scale, mainly used in theoretical studies and to test the methodology/algorithms, performed in laboratory or in the field, and relying on either destructive or nondestructive sampling (Smith and Dickson 1991; Delalieux et al. 2007; Stilwell et al. 2013; Chaerle et al. 1999; Chaerle and Van Der Straeten 2000; Lins et al. 2009; Graeff et al. 2006; Huang et al. 2012; Yuan et al. 2014; Zhang et al. 2014); (ii) the plant scale, which takes into account the structural characteristics of the plants, usually performed in situ for precision farming and pest management applications (Yang et al. 2007; Chen et al. 2008; Moshou et al. 2004, 2005; Mirik et al. 2006; Okamoto et al. 2007; Nutter 1989); and (iii) the spatilized scale, which can encompass an entire farm by exploiting airborne data (Seiffert and Schweizer 2005; Hillnhütter et al. 2011; Zhang et al. 2003) or drones (unmanned aerial vehicle (UAV)), or a region, by using satellite sensors (Mirik et al. 2011, 2013). The scale of observation regulates the outcomes of each approach, and upscaling theoretical methodologies for operational field- or regional-scale studies can introduce significant distortions and error sources that compromise the results in practice. Particular attention must be focused on this issue when selecting an appropriate method for a particular analysis. RS methods are closely linked to the instrument used for data collection (i.e., imaging or nonimaging, active or passive, ground-based, airborne, or satellite mounted, mono-, multi-, or hyperspectral sensor). A summary of RS methods to assess plant disease, including sensor technology used, disease origin, observation conditions
and scales, and methods and techniques used for data processing, is shown (Table 1). RS methods can be divided into nonimaging or imaging sensor-based applications. Among nonimaging sensors are radiometers-spectroradiometers (handheld or mounted on elevated platforms or tractors) and fluorescence radiometers. Imaging sensors include RGB cameras (visible or infrared), multispectral (broadband) sensors and hyperspectral (narrowband) sensors, thermal infrared sensors, and fluorescence imaging sensors. Less frequently used sensors include passive microwave radiometry in THz bands, nuclear magnetic resonance (NMR), and X-ray imaging.

RS has demonstrated its ability to monitor stress conditions in vegetation, and therefore the effects of plant disease. Nondestructive, spatialized techniques show promising flexibility and operational both in situ and at the farm to regional scale. RS offers the advantages of a large amount of data from the spectral response and the possibility of working at different spatial scales, with available sensor resolution from a single leaf level to an entire region.

Spectral datasets acquired by handheld, airborne, and satellite-borne sensors can be processed with data mining algorithms to detect and identify various diseases and the severity of the damage to plants as the disease advances. RS also had disadvantages: the high cost of in situ spectroradiometry and airborne platforms, the specialized experience needed to gather and process such data, and the nonuniversality of current protocols, which are concentrated on some diseases of cereals and other commercially valuable crops. RS techniques do offer good perspectives for operational implementation of disease monitoring. Moreover, the increased spatial resolution of recent satellite sensors and the decrease of the cost of data acquisition are making RS really competitive for the integration with traditional techniques. The increasing availability of small, inexpensive, high resolution spatial and spectral sensors has enhanced the operational capabilities of RS through UAV mounting spectral sensors for crop disease monitoring at the farm scale (Fig. 5). Methods are being improved through integrating different RS technologies, such as fusing spectroscopy and fluorescence data (Moshou et al. 2005; Blasco et al. 2007; Sighicelli et al. 2009). Spectral and fluorescence data were exploited to monitor winter wheat yellow rust, with greater accuracy than by using only fluorescence data (Moshou et al. 2012). In Berdugo et al. (2014), the joint monitoring of leaf temperature, chlorophyll fluorescence and hyperspectral vegetation indices (VIs) has provided good capabilities in identifying and distinguishing cucumber diseases (mosaic virus, green mottle mosaic virus, and powdery mildew).

The availability of UAV-mounted hyperspectral sensors and the integration of spectroscopy, fluorescence, and potentially thermal imaging techniques, used in combination with data from non-RS-based methods, could constitute an interesting and fruitful approach to studying plant diseases in the near future. As regards satellite platforms for earth observation (EO), a great progress will be made by the expected 2017 launch of the EnMAP (DLR, Germany) and PRISMA (ASI, Italy) hyperspectral imaging sensors. These sensors will boost hyperspectral RS data availability and usability (Yang et al. 2007).

Analysis techniques for disease detection using RS Similarly to VOC, the exploitation of RS data, which are intrinsically high dimensional, for plant-disease assessment, is relying on adequate and effective data processing techniques. Especially when dealing with hyperspectral data, in order to extract the largest amount of information data, processing and analysis techniques are a crucial asset. The data analysis techniques adopted in scientific literature of RS of plant disease can be categorized into four macro-groups: (1) correlation and regression analysis of disease presence and severity with spectral response in specific bands and/or intervals of electromagnetic spectrum (Nutter 1989; Chen et al. 2008; Huang et al. 2012); (2) assessment and derivation of spectral vegetation indices (SVIs), general or specifically introduced, which are sensitive to disease presence (Mirik et al. 2006; Huang et al. 2007; Naidu et al. 2009; Reynolds et al. 2012; Mahlein et al. 2013; Stilwell et al. 2013); (3) data mining algorithms applied to spectral data processing and feature extraction/selection for data dimensionality reduction (Malthus and Madeira 1993; Bravo et al. 2003; Zhang et al. 2003; Delalieux et al. 2007; Grisham et al. 2010; Bauriegel et al. 2011); and (4) machine learning and classification techniques, parametric and non-parametric, supervised and unsupervised, for producing results which are classified depending on disease presence/absence and possibly severity levels (Moshou et al. 2004, 2012; Rumpf et al. 2010; Hillnhütter et al. 2011; Mewes et al. 2011; Mirik et al. 2011; Mahlein et al. 2012a).

Although a relative abundance of data mining and statistical analysis software, which implement or can easily implement all the processing techniques cited above, are available (e.g., MATLAB, R, Weka, among the most used ones), such tools are currently used more by researchers specifically interested in RS general applications (among which one thread is about disease detection), than by stakeholders and generally target users of products covering plant disease assessment; the development of more specific and ready to use routines for producing information about plant disease from RS data could be a step forward towards operational use of this kind of data for vegetation health monitoring.

An alternative to empirically based method to quantify, from spectral measurements, changes in plant biochemical such as chlorophyll concentration and water content and structural parameters such as LAI, leaf angle distribution is the use of radiative transfer modeling (RTM). These mathematical approaches are used to model the interactions of photons and plant/soil components taking into account the
<table>
<thead>
<tr>
<th>Type</th>
<th>Sensor technology</th>
<th>Techniques used</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonimaging</td>
<td>Spectroscopy-spectroradiometry</td>
<td>Specific SVI, data mining and classification, features extraction, SMA, ANN, spectral decomposition (wavelets), and regression analysis</td>
<td>Nutter (1989), Malthus and Madeira (1993), Jacquemoud and Ustin (2001), Bravo et al. (2003), Graeff et al. (2006), Mirik et al. (2006), Costa et al. (2007), Delalieux et al. (2007), Yang et al. (2007), Chen et al. (2008), Naidu et al. (2009), Purcell et al. (2009), Grisham et al. (2010), Rumpf et al. (2010), Huang et al. (2012), Stilwell et al. (2013), Yuan et al. (2014), and Zhang et al. (2014)</td>
</tr>
<tr>
<td>Fluorescence spectroscopy</td>
<td></td>
<td>Features extraction, data mining, regression analysis, and passive fluorescence</td>
<td>Lins et al. (2009)</td>
</tr>
<tr>
<td>Multispectral-hyperspectral</td>
<td>SVI, ICA-PCA, image classification, feature extraction, SMA, ANN, and bio-optical vegetation parameters</td>
<td>Delwiche and Kim (2000), Bravo et al. (2003), Zhang et al. (2003), Moshou et al. (2004), Moshou et al. (2005), Okamoto et al. (2007), Blasco et al. (2007), Gowen et al. (2007), Huang et al. (2007), Sighicelli et al. (2009), Rumpf et al. (2010), Bauriegel et al. (2011); Hillnhüter et al. (2011), Mewes et al. (2011), Mirik et al. (2011), Mahlein et al. (2012a), Mahlein et al. (2012b), Reynolds et al. (2012), Wang et al. (2012), Mahlein et al. (2013), Mirik et al. (2013), Calderón et al. (2013), and Berdugo et al. (2014)</td>
<td></td>
</tr>
<tr>
<td>Thermal imaging</td>
<td>ICA, PCA, and image classification</td>
<td>Chaerle et al. (1999, 2003), Chaerle and Van der Straeten (2000), Oerke et al. (2006), Costa et al. (2013), and Berdugo et al. (2014)</td>
<td></td>
</tr>
<tr>
<td>Fluorescence imaging</td>
<td>Features extraction, data mining, and regression analysis</td>
<td>Lichtenhaller et al. (1996), Chaerle et al. (2003), Moshou et al. (2005), Blasco et al. (2007), Sighicelli et al. (2009), Agui et al. (2013), Calderón et al. (2013), and Berdugo et al. (2014)</td>
<td></td>
</tr>
<tr>
<td>Other (minor)</td>
<td>Passive microwave radiometry (THz)</td>
<td>Water emissivity assessment</td>
<td>Hadjiloucas et al. (2009)</td>
</tr>
<tr>
<td></td>
<td>Nuclear magnetic resonance</td>
<td>PCA</td>
<td>Choi et al. (2004)</td>
</tr>
</tbody>
</table>

parameters that control the electromagnetic interaction (e.g., soil properties, vegetation architecture, and leaf parameters). RTM are specifically developed for interpret leaf optical such as PROSPECT (Jacquemoud and Baret 1990) and canopy structural such as SAIL (Verhoef 1984) properties and are used in conjunction (leaf+canopy) in direct mode to reproduce vegetated land spectral behavior in different wavelengths as observed by RS sensors in relation to specific illumination condition. Model simulation is very important in designating vegetation index and performing sensitivity analyses to analyze the potentiality or limit in exporting field experimental results in other condition such as different plant characteristics, sensors and view geometry). When properly calibrated and using a priori constraint information, the inversion of RTM can become a powerful tool to predict target (i.e., vegetation compound) properties without depending on contemporary reference ground sample.

Among the different RTM proposed in literature the PROSAIL (PROSPECT+SAIL models; Jacquemoud et al. 2009) is one of the more diffuse and has been successfully applied in crop application with the purpose of identify anomalous plant condition related to chlorosis and vigour potential due to suboptimal growth, nutrient deficiency, and/or pathogen impact (Zarco-Tejada et al. 2004; Baret et al. 2007). Despite the potentiality of RTM applications in crop monitoring, some limitation and difficulties still reduce their use in comparison to local application based on the integration of EO data and field observation (smart scouting approach). Finally, when changes in plant/crop parameters are detected (e.g., LAI reduction of chlorophyll deficiency), the clear identification of stress factor is not often possible without additional information or taking into account field specific condition (Baret et al. 2007).

3.3.3 Spectroscopy-based methods

Spectroscopy is among the most used methods for RS of plant disease, including VIS, NIR, and/or SWIR, imaging or nonimaging sensors. These techniques hold particular promise for crop disease monitoring because of their potential as operational instruments, flexibility, efficacy, and cost-efficiency. The most relevant and recent advances in spectroscopy-based techniques are discussed below.

Nonimaging spectroscopy approaches Leaf scale RS of plant disease is based on inherent optical properties of leaf pigments, chemical components, properties, and structural characteristics (Jacquemoud and Ustin 2001). Leaf spectra collected in the laboratory or field were used to determine spectral regions (visible, NIR, and SWIR) and/or the VIs with which to detect diseases such as, among the most studied using RS: Fiji leaf gall sugarcane disease (Purcell et al. 2009), wheat powdery mildew and take-all disease (Graeff et al. 2006), curl mite (Stilwell et al. 2013), sugarcane yellow leaf virus (Grisham et al. 2010), grapevine viruses (Naidu et al. 2009), and winter wheat yellow rust (Zhang et al. 2014). A study on the differentiation of winter wheat disease due to pathogens (yellow rust, powdery mildew) and insects (wheat aphid) infestation was recently carried out by Yuan et al. (2014).

Of particular interest are methods for early stage disease detection (Malthus and Madeira 1993; Delalieux et al. 2007; Rumpf et al. 2010), although their actual application to crop management is inconsistent across crops. Available studies are crop specific and results cannot be generalized to other crops and/or places with comparable accuracy. Huang et al. (2012) compared leaf-scale to canopy-scale detection and severity assessment of rice leaf folder disease, finding good agreement of detection rates with VIS and NIR reflectance through linear regression modeling and high negative correlation between the NIR plateau (737–1,000 nm) and infestation severity.

Several authors have attempted to quantify the severity of crop damage using radiometry (Nutter 1989). Yang et al. (2007) studied rice brown planthopper and leaf folder diseases, Mirik et al. (2006) assessed winter wheat damage due to greenbugs, Chen et al. (2008) estimated the damage produced by the Verticillium wilt on cotton, and Prabhakar et al. (2011) studied leafhopper disease.

Spectroscopy has also been used for fruit quality assessment, sometimes in conjunction with supplementary information such as e-nose data, a successful integration of RS-based methods with VOC analysis (Costa et al. 2007).

Imaging spectroscopy approaches Recently, hyperspectral imaging instruments have been introduced for plant disease assessment and monitoring. Laboratory-based studies include Fusarium fungal infection and head blight disease in wheat (Bauriegel et al. 2011), early detection of sugar beet diseases (Mahlein et al. 2012a), and detection of Cercospora leaf spot, sugar beet rust, and powdery mildew on sugarbeet leaves (Mahlein et al. 2013). Of particular interest for an effective intervention is differentiating infections and their stage of development (Mahlein et al. 2012a). These studies applied a large range of statistical methods for image analysis such as linear regression, principal component analysis (PCA), spectral angle mapper (SAM) classification, and support vector machine (SVM) classification with very high accuracy of disease detection. However, these studies focused on one or few crops with little possibility for generalization. Field-based studies examined yellow rust disease on wheat plants (Bravo et al. 2003) and tried to distinguish among wheat diseases andabiotic stress conditions (Moshou et al. 2004). Reynolds et al. (2012) and Huang et al. (2007) used both field and airborne hyperspectral data to assess the severity of Rhizoctonia crown and root rot disease in sugarbeet and yellow rust in wheat, respectively. Airborne hyperspectral data are most suitable among RS applications at the farm and regional scales.
Zhang et al. (2003) used AVIRIS data to detect late blight disease in tomato plantations, while Hillnhütter et al. (2011) studied the soil pest-induced sugar beet disease using two different airborne sensors (AISA and HyMap). UAV technologies represent a potential solution for field or farm monitoring, reducing the cost of data acquisition.

Fewer works using a regional scale focus on plant disease assessment, but those are useful to determine the upscaling potential of RS based methods. Wheat streak mosaic was assessed using multispectral satellite data (Landsat TM) with the maximum likelihood-supervised classification technique (Mirik et al. 2011), or with spectral mixture analysis (SMA) techniques for subpixel classification (Mirik et al. 2013). Wang et al. (2012) discussed the multiscale capabilities (from leaf-to-satellite observation) offered by RS to detect winter wheat stripe rust.

There is underexploited potential in the integration of imaging spectroscopy in the VIS to SWIR ranges with fluorescence imaging. Application of this integrated method for crop monitoring have been exploited so far for enhancing detection of winter wheat yellow rust in field (Moshou et al. 2005), using UAV technologies and aerial platform (Panigada et al. 2014). This works demonstrate that integration of passive fluorescence measures with optical spectral measures is possible if spectral resolution of sensors is adequate to assess passive fluorescence features. Studies based on modeled data (Damm et al. 2011) have assessed the feasibility of fluorescence estimation as a function of the spectral sampling interval, spectral resolution and signal to noise ratio of the hyperspectral sensor used. However, extensive exploitation is limited to technology currently available and operational use of such integrated data is to be expected within the next years. Recently, the fusion of thermal, fluorescence and optical hyperspectral airborne data has been successfully used to assess olive plants infected by Verticillium wilt (Calderón et al. 2013). The success of integration is already evident in fruit safety and control studies (Gowen et al. 2007). Identification of defective fruits (some due to disease effects) through joint spectral and fluorescence data has been tested in Citrus (Blasco et al. 2007) and orange (Sighicelli et al. 2009).

In a related application, imaging spectroscopy was used to scan wheat kernels for head blight disease through machine vision techniques (Delwiche and Kim 2000); weed infestations have also been evaluated using image segmentation algorithms (Okamoto et al. 2007).

3.4 Potential technologies for biosensor development: phage display, electrochemistry, and biophotonics

Molecular detection of probe-target interactions, such as antigen-antibody or protein-protein, is based on specific peptide sequence recognition. The bound molecule is interrogated experimentally and the probe-target complex is identified through immunoassays such as ELISA. This protein complex interaction is usually optimized for pathogen detection, but with very low sensitivity of no fewer than $10^5$ molecules. This approach has the additional drawback of using entire proteins as probes, representing the exposure of multiple epitopes which may share common determinant regions with other targets and/or organisms, generating cross-reactions. To overcome such problems, probe size can be reduced to a specific epitope core to improve specificity and sensitivity, potentially aided by novel biosensors like electrochemistry and biophotonics. Different classes of portable biosensors have emerged from human diagnostics; among them, electrochemistry has become a major technology for fast field-based diagnostics. These sensors detect electron mobility and redox reactions. To better understand the novel biosensor technologies, we will specifically discuss probe discovery using Phage Display technology, and detection systems with focus on electrochemistry and biophotonics as emerging technologies.

3.4.1 Phage display

There are two main strategies to obtain short, highly specific ligands or markers, based on selecting nucleic acids, systematic evolution of ligands by exponential enrichment (SELEX) (Ellington and Szostak 1990), or proteins (peptides and antibody fragments by Phage Display) from combinatorial libraries.

Antigenicity is defined as the ability of a peptide or whole antigen to bind to a selector monoclonal antibody or target. Phage display technology is used to identify ligands that bind specific biological molecules through cycles of selection (Smith 1985). Phage display is also a subtractive proteomic technology for ligand discovery, which can be used as immunogens in vaccines, antigens in diagnostics, and as immune-modulating molecules in therapeutic systems. There are two types of combinatorial libraries for phage display selection of ligands: one relies on random peptides and the other, on combinatorial antibody fragments.

Antibody libraries can be constructed in the single-chain variable fragment (scFv) or Fab antibody fragment format. The scFv format has several advantages. First, library construction can be simplified by overlapping extension PCR to reduce the number of steps. Second, the ability of scFv to multimerize enhances antigen activity and facilitates selection against antigens such as cell-surface molecules. Third, the smaller scFv molecule tends to produce better yields than the Fab format (Barbas et al. 2001). Bacteriophage display of combinatorial antibody libraries allows monoclonal antibodies of a desired specificity to be selected without conventional hybridoma technology (Andris-Widhopf et al. 2000). Isolation of specific antibodies from a cloned immunological repertoire requires a large, diverse library and efficient selection. The keys to achieving this goal are generating a good
immune response and constructing the library from quality RNA and cDNA.

The first step in constructing an immune library is a successful course of vaccination with an immunogen of choice (Barbas et al. 2001). Evolution and diversity of antibodies in vivo is achieved by combining three complementary determining regions (CDR) provided by each heavy and light chain protein that act together by heterodimerization to form the antibody-binding site (Padlan 1994). The binding site, therefore, results from a convergence of six hypervariable regions, which produces antibodies of differing specificity. In vitro selection and evolution of antibodies derived from phage display has become a powerful tool for antibody engineering (Rader and Barbas 1997). The system for producing new antibodies can be improved using the properties of existing antibodies. When the affinity of a monovalent antibody fragment reaches a dissociation constant of ~0.1 nM, further improvements become difficult (Barbas et al. 2001). A panel of scFv against potato leafroll virus (PLRV) was isolated from a large nonimmunized human antibody library (Harper et al. 1997). The modified selected scFvs were suitable for use in assays to detect PLRV in infected potato leaves.

For peptide libraries, biopanning procedures may produce many different antigens, but with differential reaction due to its smaller sequence, which depends on linear or conformational structures that mimic the true antigenic epitope. This bacteriophage-fused peptide strategy is perhaps the most successful to date, resulting in many different markers for humans and animals (Goulart et al. 2010), but very few or none in plants. Therefore, this is a research field that deserves greater attention.

3.4.2 Electrochemistry and photonics

Biosensors are usually classified according to their signal transduction and biorecognition principles. Signal transduction can include electrochemical, optical, piezoelectric, and thermal sensors (Goulart et al. 2010). Optical biosensors measure light absorbed or emitted as the result of a biological or chemical reaction, while electrochemical biosensors are based on biochemical reactions that cause electron transfer between a functionalized electrode and an analyte in solution, and can make amperometric, voltammetric, or impedimetric measurements. Electrochemical biosensors are amenable to miniaturization, can function in turbid media, have compatible equipment, and do not experience environmental interference (Ahmed et al. 2008). They are usually used as immunoassays, affinity ligand-based biosensor solid-state devices, in which the immunoenzymatic reaction is coupled to a transducer. The fundamental basis of all immunoassays is the specificity of the molecular recognition of antigens by antibodies to form a stable complex, as with other immunoassays (Luppa et al. 2001).

4 Conclusions

Early detection of pathogen infections is pivotal to managing polycyclic diseases. Polycyclic epidemics are caused by pathogens capable of several infection cycles each season (Fig. 6). Several environmental parameters, including solar radiation, temperature, and runoff coefficient (D’Asaro and Grillone 2012; Grillone et al. 2012, 2014), showed significant variation among Mediterranean areas and might drastically effects pathology epidemiological dynamics. Usually, primary infections occur at the beginning of the growing season, while secondary infections, spread by vectors, lead to an irreversible, symptomatic disease stage with consequent production loss. Many pathogens remain in latent form inside plants during winter and then explode in the next season. Since volatile-based methods focus on the host, they may be able to detect early infections before primary infections are visible. When pathogen infects the host an induced stress response occurs before symptoms appear. Early induced volatiles may allow the detection of disease before that pathogen reaches detectable levels by molecular or serological methods. This is due to the activation of early pathogenetic mechanisms that although do produce any visible effects, they may be detected by the analysis of host induced biomarkers such as transcripts, proteins, and volatiles. Lateral flow microarrays that target both host- and pathogen-specific biomarkers might detect infections at primary sites. Biophotonic and phage display sensors could lead to early diagnosis before secondary infections occur. Other innovative methods like RS and spectroscopy-based technologies may speed diagnosis over traditional nucleic acid and serological assays and enable better analysis of the spatial and temporal variability of disease, especially when calibrated and coupled with reliable reference information. The practical application of each traditional or innovative method depends on the cost and availability of instruments, spatialization level (plant, field, farm, or agricultural district), rapidity of analysis, and the stage of disease at which detection is possible. An overview of the features of innovative methods is presented (Fig. 7). It is very difficult to compare costs/sample across methods, so this feature is not included. The issue of reliability is quite complicated. The co-authors of this work, coming from different scientific fields, have agreed that, today, a comparative scale of reliability between the different techniques would be misleading. For this reason, it was not included. Availability is pivotal to a successful, large-scale application. ELISA kits are widely available because of their low equipment costs compared with other detection methods. This method has good reliability, although false negatives are possible. However, ELISA is time consuming and has low potential for spatialization. Quantitative PCR is a highly available and effective method when symptoms are present; it is widely applied to confirm diagnosis after symptoms appear. The
decreasing costs of kits, reagents, and equipment render it available for use by large companies or consortia which can afford a molecular biology lab. However, like ELISA, this method is time consuming: several days are required to extract RNA, perform RNA qualitative analysis, synthesize cDNA, amplify it, and analyze the resulting data. This method is also unable to detect early infections: it is quite unreliable before symptoms appear. At the asymptomatic stage, both false negatives and false positives frequently occur in plants, due to the presence in plant extracts of amplification inhibitors, uneven pathogen distribution, and incorrect sampling in plant extracts. Spatialization is difficult although possible in the near future due to the essential simplicity of thermocyclers.

Biophotonic-based sensors have three strengths: (1) they can effectively detect early stages of infection, (2) results are rapid and presented directly in the orchard, and (3) they could...
be highly integrated with other systems, providing good spatialization. The major weakness is low availability. They are still under development, although they may be on the market soon. RS allows high spatialization of the results and represents a valid tool to detect early infections, since they can identify a few infected plants within orchards or fields. Thus, they may provide a tool to quickly analyze many plants and eradicate primary infections early, avoiding secondary spreads. As RS, volatile sensors are based on the analysis of host responses. Like biophotonics, they are not yet on the market. However, they deliver results rapidly and can effectively detect early infections using early induced volatiles as biomarkers. They have great potential for sustainable pest management, since they analyze host responses and can decipher how plants respond to treatments and their health status.

Technical improvements to traditional methods and novel approaches are needed in a world characterized by increasing air travel and free-trade agreements, where the borders of many countries have become open to unintentional or deliberate introduction of foreign plant pathogens. Once new protocols and methods have been developed, standardized, unbiased procedures are required to validate and certify protocols for diagnosis of crop diseases all around the world.

Acknowledgments We are grateful to Chiara Nepi for providing iconographic materials. We thank Minghua Zhang and colleagues for permission to use their data in our Fig. 4 and Jorge Torres-Sánchez and colleagues for permission to use their picture in our Fig. 5. CED was supported by the California Citrus Research Board (CRB), the Industry-University Cooperative Research Program (UC Discovery), the Florida Citrus Production Advisory Council (FCPRAC), and the National Science Foundation (no. 1255915).

Glossary

HyMap™ is a hyperspectral scanner that provides 128 bands across the reflective solar wavelength region of 0.45–2.5 μm with contiguous spectral coverage and bandwidths between 15 and 20 nm.

Omic refers to a field of study in biology aiming at the collective characterization of pools of biological molecules that translate into the function of organisms.

Microarrays integrate laboratory functions on a millimetric chip on a solid substrate (e.g., glass slide or silicon films) that assay large amounts of biological material using high-throughput screening miniaturized, multiplexed, and parallel processing and detection methods.

PROSPECT is a radiative transfer model based on the Allen’ plate model used by remote sensing techniques.

RGB is an additive color model in which red, green, and blue light is added together in various ways to reproduce a broad array of colors.

SYBR® Green is an asymmetrical cyanine dye used as a nucleic acid which absorbs blue light (λ_{max}=497 nm) and emits green light (λ_{max}=520 nm).

Trascriptome is used to address a specific object of a specific field of study in biology. It refers to the set of all RNA molecules produced in a population of cells. It differs from the exome, the sequences which when transcribed remain within the mature RNA after introns are removed by RNA splicing.

References


Fusarium coeruleum


