



Politecnico di Bari

Repository Istituzionale dei Prodotti della Ricerca del Politecnico di Bari

Early-detection of plant stress using spectroscopy and chemometrics

This is a PhD Thesis

Original Citation:

Early-detection of plant stress using spectroscopy and chemometrics / Ahmed, Elhussein Mohamed Fouad Mourad. -
ELETTRONICO. - (2023). [10.60576/poliba/iris/ahmed-elhussein-mohamed-fouad-mourad_phd2023]

Availability:

This version is available at <http://hdl.handle.net/11589/252621> since: 2023-05-09

Published version

DOI:10.60576/poliba/iris/ahmed-elhussein-mohamed-fouad-mourad_phd2023

Publisher: Politecnico di Bari

Terms of use:

(Article begins on next page)



Al Magnifico Rettore
del Politecnico di Bari

Il sottoscritto Elhussein Mohamed Fouad Mourad Hussein AHMED nato ad Il Cairo (Egitto) il 06.12.1992 residente a Bari (BA) 70125 in Corso Alcide De Gasperi 405/A e-mail elhussein.ahmed@poliba.it iscritto al 3° anno di Corso di Dottorato di Ricerca in RISCHIO, SVILUPPO AMBIENTALE, TERRITORIALE ED EDILIZIO ciclo XXXV ed essendo stato ammesso a sostenere l'esame finale con la prevista discussione della tesi dal titolo:

Early-detection of plant stress using spectroscopy and chemometrics

DICHIARA

- 1) di essere consapevole che, ai sensi del D.P.R. n. 445 del 28.12.2000, le dichiarazioni mendaci, la falsità negli atti e l'uso di atti falsi sono puniti ai sensi del codice penale e delle Leggi speciali in materia, e che nel caso ricorressero dette ipotesi, decade fin dall'inizio e senza necessità di nessuna formalità dai benefici conseguenti al provvedimento emanato sulla base di tali dichiarazioni;
- 2) di essere iscritto al Corso di Dottorato di ricerca RISCHIO, SVILUPPO AMBIENTALE, TERRITORIALE ED EDILIZIO ciclo XXV, corso attivato ai sensi del "Regolamento dei Corsi di Dottorato di ricerca del Politecnico di Bari", emanato con D.R. n.286 del 01.07.2013;
- 3) di essere pienamente a conoscenza delle disposizioni contenute nel predetto Regolamento in merito alla procedura di deposito, pubblicazione e autoarchiviazione della tesi di dottorato nell'Archivio Istituzionale ad accesso aperto alla letteratura scientifica;
- 4) di essere consapevole che attraverso l'autoarchiviazione delle tesi nell'Archivio Istituzionale ad accesso aperto alla letteratura scientifica del Politecnico di Bari (IRIS-POLIBA), l'Ateneo archiverà e renderà consultabile in rete (nel rispetto della Policy di Ateneo di cui al D.R. 642 del 13.11.2015) il testo completo della tesi di dottorato, fatta salva la possibilità di sottoscrizione di apposite licenze per le relative condizioni di utilizzo (di cui al sito <http://www.creativecommons.it/Licenze>), e fatte salve, altresì, le eventuali esigenze di "embargo", legate a strette considerazioni sulla tutelabilità e sfruttamento industriale/commerciale dei contenuti della tesi, da rappresentarsi mediante compilazione e sottoscrizione del modulo in calce (Richiesta di embargo);
- 5) che la tesi da depositare in IRIS-POLIBA, in formato digitale (PDF/A) sarà del tutto identica a quelle **consegnate**/inviata/da inviarsi ai componenti della commissione per l'esame finale e a qualsiasi altra copia depositata presso gli Uffici del Politecnico di Bari in forma cartacea o digitale, ovvero a quella da discutere in sede di esame finale, a quella da depositare, a cura dell'Ateneo, presso le Biblioteche Nazionali Centrali di Roma e Firenze e presso tutti gli Uffici competenti per legge al momento del deposito stesso, e che di conseguenza va esclusa qualsiasi responsabilità del Politecnico di Bari per quanto riguarda eventuali errori, imprecisioni o omissioni nei contenuti della tesi;
- 6) che il contenuto e l'organizzazione della tesi è opera originale realizzata dal sottoscritto e non compromette in alcun modo i diritti di terzi, ivi compresi quelli relativi alla sicurezza dei dati personali; che pertanto il Politecnico di Bari ed i suoi funzionari sono in ogni caso esenti da responsabilità di qualsivoglia natura: civile, amministrativa e penale e saranno dal sottoscritto tenuti indenni da qualsiasi richiesta o rivendicazione da parte di terzi;
- 7) che il contenuto della tesi non infrange in alcun modo il diritto d'Autore né gli obblighi connessi alla salvaguardia di diritti morali od economici di altri autori o di altri aventi diritto, sia per testi, immagini, foto, tabelle, o altre parti di cui la tesi è composta.

Luogo e data Bari 24.04.2023

Firma 

Il/La sottoscritto, con l'autoarchiviazione della propria tesi di dottorato nell'Archivio Istituzionale ad accesso aperto del Politecnico di Bari (POLIBA-IRIS), pur mantenendo su di essa tutti i diritti d'autore, morali ed economici, ai sensi della normativa vigente (Legge 633/1941 e ss.mm.ii.),

CONCEDE

- al Politecnico di Bari il permesso di trasferire l'opera su qualsiasi supporto e di convertirla in qualsiasi formato al fine di una corretta conservazione nel tempo. Il Politecnico di Bari garantisce che non verrà effettuata alcuna modifica al contenuto e alla struttura dell'opera.
- al Politecnico di Bari la possibilità di riprodurre l'opera in più di una copia per fini di sicurezza, back-up e conservazione.

Luogo e data Bari 24.04.2023

Firma 



D.R.R.S

POLITECNICO DI BARI

01

Doctor of Philosophy in Environmental and Building Risk and Development

2023

Coordinator: Prof. Michele Mossa

XXXV CYCLE
Curriculum: CHIM/07

DICATECh

Department of Civil, Environmental, Building Engineering and Chemistry

Early-detection of plant stress using spectroscopy and chemometrics

Prof. Vito GALLO

Dr. Biagia MUSIO
DICATECh
Polytechnic University of Bari

Dr. Anna Maria D'ONGHIA
Dr. Stefania GUALANO
Eng. Franco SANTORO
CIHEAM Bari

Elhussein Mohamed Fouad Mourad Hussein AHMED



POLITECNICO DI BARI

D.R.R.S

01

Dottorato di Ricerca in Rischio e Sviluppo ambientale, territoriale ed edilizio

2023

Coordinatore: Prof. Michele Mossa

XXXV CICLO
Curriculum: CHIM/07

DICATECh

Dipartimento di Ingegneria Civile, Ambientale, del Territorio, Edile e di Chimica

Rilevamento precoce di condizioni stress nelle piante mediante l'applicazione di analisi spettroscopiche e chemiometriche

Prof. Vito GALLO
Dr.ssa Biagia MUSIO
DICATECh
Politecnico di Bari

Dr.ssa Anna Maria D'ONGHIA
Dr.ssa Stefania GUALANO
Ing. Franco SANTORO
CIHEAM Bari

Elhoussein Mohamed Fouad Mourad Hussein AHMED

EXTENDED ABSTRACT (eng)

Plant stress responses cause a biochemical cascade in which gene expression is altered. Such an event leads to further downstream changes that result in alterations in metabolism to reduce the effect of the stress so the plant can survive/resist it. Stress can induce phenotypic changes, i.e. changes observable to the human eye such as colour modification, inhibition of growth, and change in fruit size. Technological development leads to the ability to study organisms such as plants on different biological levels. Omics is a term that encompasses a wide range of domains in biology such as genomics, transcriptomics, proteomics, and metabolomics which are the science of studying genome, transcriptome, and metabolome of a sample, respectively. This PhD project focused on the chemometric (non-targeted) approach of spectroscopic data for studying both abiotic and biotic stress responses in plants. Three studies described in this PhD thesis demonstrated the feasibility of spectroscopic methods to decipher biological phenomena such as plant abiotic stress as well as biotic stress.

In the first study, hemp plant grown in soil contaminated with heavy metals was combined with spirulina for bioremediation purposes. The non-targeted nuclear magnetic resonance (NMR) analysis in addition to the quantitative inductively coupled plasma--atomic emission spectroscopy (ICP-AES) analysis showed specific changes at the metabolic and ionic levels for the hemp plant affected by the abiotic stress caused by heavy metals. ICP-AES provided an efficient way

for detecting residual heavy metals within plant tissues and soil. Importantly, non-targeted metabolomic analysis helped to reveal the relationships between metabolites' distribution in hemp tissues and the sequestered metals. Results reported in this work indicate that the hemp/spirulina system represents a suitable tool for remediation of metal-contaminated soils by modulating biomass production and metals uptake. Also, the combination of spirulina with the hemp plant aided the latter, mainly in terms of health condition, which allows for more survivability during bioremediation. It was demonstrated that hemp accumulates copper, chromium, nickel, and zinc preferentially in the leaves, while lead is distributed mainly in the stems of the plant. Moreover, it was found that, at higher concentrations, spirulina acts as a growth promoter, contributing to an increase in the final generated biomass.

In the second study, olive plants grown in a controlled environment were considered. The analysis focused on cultivated asymptomatic olive leaves. Young olive trees were artificially infected by *Xylella fastidiosa* subsp. *pauca* ST53 responsible of olive quick decline syndrome (OQDS). After 2 years of growth, samples from leaves were collected and analysed. By using ^1H NMR, HSR, and chemometrics, different OQDS-related diagnostic signals and wavelengths were identified for infected but asymptomatic leaves. These signals are necessary for the development of sensors capable of detecting the disease at the very early stages.

In the third study, olive plants grown in uncontrolled field conditions were considered. Though it is almost impossible to discriminate between healthy leaves and asymptomatic infected leaves at the early stage of infection using visual assessment, chemometrics applied to spectroscopic data provided another way to discriminate amongst both classes. Furthermore, the different supervised chemometric techniques applied to HSR and NMR data were able to discriminate between healthy and asymptomatic infected samples. Also, the Fisher Discriminant (FD) proved to be useful in decreasing the number of features to be used

while having good performance for the three tested supervised machine-learning techniques.

key words: plant stress; heavy metals; Xylella fastidiosa; hemp; spirulina; olive; ICP-AES; NMR; hyperspectral reflectance; metabolomics; spectranomics; chemometrics.

EXTENDED ABSTRACT (ita)

Le risposte delle piante allo stress causano una cascata biochimica in cui l'espressione genica è alterata. Un tale evento porta a ulteriori cambiamenti a valle che provocano alterazioni del metabolismo per ridurre l'effetto dello stress in modo che la pianta possa sopravvivere / resistere. Lo stress può portare a cambiamenti fenotipici, cioè cambiamenti osservabili all'occhio umano come la modifica del colore, l'inibizione della crescita e il cambiamento nelle dimensioni dei frutti. Lo sviluppo tecnologico porta alla capacità di studiare organismi come le piante a diversi livelli biologici. Omica è un termine che comprende una vasta gamma di domini in biologia come genomica, trascrittomica, proteomica e metabolomica che sono, rispettivamente, lo studio del genoma, del trascrittoma e del metaboloma. Questo progetto di dottorato si è concentrato sull'approccio chemiometrico (non-targeted) per analizzare i dati spettroscopici dalle tecnologie omiche mettendo in luce le reazioni nelle piante in seguito ad uno stress biotico o abiotico. Tre studi descritti in questa tesi di dottorato hanno dimostrato le potenzialità dei metodi spettroscopici al fine di individuare fenomeni biologici, come lo stress abiotico e biotico nelle piante.

Il primo studio, ha riguardato il biorisanamento di un terreno contaminato con metalli pesanti da parte della pianta di canapa in combinazione con la spirulina. L'analisi mediante Risonanza Magnetica Nucleare (NMR) non-targeted e l'analisi quantitativa mediante spettrometria di emissione atomica al plasma ad accoppiamento induttivo (ICP-AES) hanno permesso di rilevare cambiamenti specifici ai livelli metabolici e ionomici all'interno dei tessuti della pianta di

canapa affetta dallo stress abiotico dovuto alla presenza di metalli pesanti nel terreno. L'ICP-AES ha permesso di rilevare efficacemente i metalli pesanti residui all'interno dei tessuti vegetali e del suolo. È importante sottolineare che l'analisi metabolomica non-targeted ha contribuito a rivelare le correlazioni tra la distribuzione dei metaboliti nei tessuti della pianta e la tipologia dei metalli sequestrati. I risultati riportati in questo lavoro indicano che il sistema costituito da canapa/spirulina rappresenta uno strumento adatto per la bonifica dei suoli contaminati dai metalli modulando la produzione di biomassa e l'assorbimento dei metalli. Inoltre, l'utilizzo di acqua contenente spirulina per irrigare la pianta di canapa ha permesso a quest'ultima sopravvivere in modo più vigoroso durante il processo di biorisanamento. È stato dimostrato che la pianta di canapa accumula alcuni metalli pesanti (rame, cromo ecc.) preferibilmente nelle foglie, mentre il piombo viene localizzato principalmente nello stelo della pianta. È stato anche evidenziato come la spirulina, a maggiori concentrazioni agisca da promotore della crescita, contribuendo ad un aumento della biomassa finale.

Nel secondo studio sono state prese in considerazione piante di olivo coltivate in ambiente controllato. L'analisi si è concentrata sulle foglie asintomatiche. I giovani olivi sono stati infettati artificialmente da *Xylella fastidiosa* subsp. pauca ST53 responsabile del complesso del disseccamento rapido dell'olivo (OQDS). Dopo 2 anni di crescita, sono stati raccolti e analizzati campioni di foglie. Utilizzando ^1H NMR, HSR e chemiometria, sono stati identificati diversi segnali diagnostici e lunghezze d'onda correlate a OQDS per foglie infette ma ancora asintomatiche. La determinazione di metaboliti (mediante NMR) e di lunghezze d'onda (mediante HSR) specifici sono necessari per lo sviluppo di sensori in grado di rilevare la malattia nelle fasi iniziali.

Nel terzo studio sono state prese in considerazione piante di olivo coltivate in condizioni di campo non controllate. Sebbene sia quasi impossibile discriminare tra foglie sane e foglie infette asintomatiche nella fase iniziale delle infezioni utilizzando la valutazione visiva, la chemiometria applicata ai dati spettroscopici ha fornito un'alternativa per discriminare tra queste le classi.

Inoltre, le diverse tecniche chemiometriche supervisionate applicate ai dati HSR e NMR sono state in grado di discriminare tra campioni di foglie sane e infette ma asintomatiche. Inoltre, la discriminante di Fisher (FD) si è dimostrata utile per ridurre il numero di variabili da utilizzare mantenendo prestazioni soddisfacenti

key words: stress delle piante; metalli pesanti; Xylella fastidiosa; canapa, spirulina, ulivo, ICP-AES; NMR; riflettanza iperspettrale; metabolomica; spettrometrica; chemiometria

INDEX

<i>1.0 Introduction</i>	<i>11</i>
<i>2.1 Chapter 1</i>	<i>15</i>
<i>2.2 Chapter 2</i>	<i>27</i>
<i>2.3 Chapter 3</i>	<i>33</i>
<i>2.4 Chapter 4</i>	<i>43</i>
<i>2.5 Chapter 5</i>	<i>51</i>
<i>2.6 Chapter 6</i>	<i>75</i>
<i>2.7 Chapter 7</i>	<i>101</i>
<i>3.0 Conclusions</i>	<i>129</i>
<i>4.0 Acknowledgements</i>	<i>131</i>
<i>5.0 References</i>	<i>133</i>
<i>6.0 Curriculum</i>	<i>157</i>

INTRODUCTION

The general goal of this PhD thesis is to explore the application of spectroscopy in the field of plant stress and applying advanced chemometric methods for the analysis of omics data generated from spectroscopic techniques. This was carried out under the light of two independent projects. The first project is related to plant abiotic stress (heavy metals in hemp; results presented in chapter 5) and the other is related to biotic stress (*Xylella fastidiosa* in olive trees; results presented in chapter 6 and 7).

Plant stress causes a biochemical cascade in which gene expression is altered leading to further downstream changes that result in a change in metabolism to reduce the effect of the stress so the plant can survive/resist the stress. Different types of stress can induce phenotypic changes, of which are observable to the human eye such as colour modification, inhibition of growth, and change in fruit size. Plant stress factors are categorised as either abiotic or biotic.

As an abiotic stress factor, heavy metals are toxic to organisms when exceeding a certain threshold of concentration, usually parts per million (Edelstein & Ben-Hur, 2018). They generally do not get degraded and remain a risk to biological entities for several years after their introduction into an ecosystem (Zielonka et al., 2020; Yadav, 2010; Ahmad et al., 2016; Vardhan et al., 2019).

As a biotic stress factor, *Xylella fastidiosa* is the causal agent of various plant diseases, amongst which is the most recently disease affecting olive trees in southern Italy, the olive quick decline syndrome (OQDS). *Xylella fastidiosa* remain in an incubation period after infection until favourable conditions are available, e.g. warm temperature. Then, the bacteria start to form a biofilm that blocks the xylem vessels of the olive tree causing a hindrance in the water movement inside the plant and affecting vital processes such as photosynthesis and transpiration, which eventually leads to the death of the tree (Hornero et al.,

2020). Amongst the detection techniques, qPCR (real-time quantitative Polymerase Chain Reaction) is the most sensitive test. Nonetheless, the feasibility to adopt this technique in field conditions is limited by sampling frequency and the irregular distribution of *Xylella fastidiosa* in a canopy, especially at the early stage of the disease (Zarco-Tejada et al., 2018).

Over the past decades, technological development allowed for a more holistic approach to understanding plant stress in which a more comprehensive view is adapted for the identification of key components in the hidden patterns of the mechanisms involved behind (Crandall et al., 2020).

Technological development leads to the ability to study organisms, such as plants, on different biological levels. The comprehensive study of the complex interactions among molecules in different biological systems defines the “omics” sciences. Omics aims at the collective characterization and quantification of pools of biological molecules (Wheelock & Wheelock, 2013). Metabolomics aims at the systematic study of the chemical fingerprint correlated with specific cellular processes, genomics focuses on the structure and function of genomes, proteomics studies the entire set of proteins produced and modified by an organism. In contrast to genomics and proteomics, metabolomics is less developed in terms of hardware (instrumentation), software (algorithms and tools), and databases. This mainly due to the complexity of the metabolome and the period of development as compared to genomics and proteomics (Wishart, 2011; Wishart et al., 2022a).

Metabolomics is widely applied in studying various topics such as disease diagnosis, crop breeding, detection of adulteration of food, food authentication, tracing the origin of food production, metabolic pathway analysis, genetic engineering, monitoring of medical interventions effects as in organ transplantation, and effects of aging—nutrition—activities—and lifestyle—on the metabolic networks (Taha et al., 2022; Razaq et al., 2022). Metabolomics is less developed than genomics and proteomics in terms of databases, identification technologies, and algorithms (Ryan & Robards, 2006; Hall et al., 2022; van der Hooft et al., 2020). Nuclear Magnetic Resonance (NMR) spectroscopy is one of the main

techniques used in metabolomics. The potential of NMR spectroscopy is well known for the qualitative and quantitative analysis of chemical compounds of complex mixtures without the need to separate the chemical compounds. Furthermore, NMR spectroscopy is the most robust technique for the elucidation of unknown compounds. Since it is not selective, it allows for an unbiased view of the chemical composition of the sample. Moreover, NMR samples can be reused for subsequent analyses as it is not required for the sample to be pre-treated or destroyed (Wishart, 2011). Hyperspectral reflectance provides high-throughput data which can be related to plant biophysical and biochemical traits and is key-stone for the development of early-detection sensors for plant disease (Zheng et al., 2021a). Such spectroscopic techniques measure a large set of variables for every single sample analysed that result in high-dimensional datasets with thousands of variables. For the analysis of such high-dimensional data, chemometrics applies multivariate data analysis and machine learning with the aim of unravelling the hidden patterns of valuable information (Martens, 2015). Spectranomics is a new field in which the combination of spectroscopy with physiochemistry and taxonomy is used to link the plant spectral profile to its trait (Asner & Martin, 2009).

In this thesis, two research lines were followed under two different independent projects, with both having the common theme of the combination of spectroscopic and chemometric methods for the investigation of plant stress. On one hand, the first research line is considered with plant abiotic stress by means of hemp growing in heavy metals-contaminated soil with the addition of *Spirulina*. On the other hand, the second research line is considered with plant biotic stress by means of olive trees infected with *Xylella fastidiosa* under greenhouse and open field conditions.

CHAPTER 1

On plant stress

Plants are very important to humans as they provide us with many things to consider in a wide range of aspects including environmental, biological, medical, pharmaceutical, and industrial, as well as cultural and aesthetical. However, when plants are exposed to stress, these aspects get affected. Stress conditions are caused by biotic and/or abiotic factors. Biotic factors are biological entities such as insects, plants, as well as fungi, bacteria and viruses, the latter being often able to cause infections. On the other hand, abiotic factors are physical agents, including cold, heat, and light as well as chemical agents such as heavy metals, salinity, and oxidative stress. Biotic and abiotic factors can be found in combination leading to highly complex plant responses leading to particularly complex responses and, often, to particularly serious damages to the plant.

The result of biotic or abiotic stress factors is the production of reactive oxygen species (ROS) in the cell. If ROS are over-produced and accumulates in the cell environment, there is the onset of oxidative stress, a complex phenomenon that can lead to the damaging of cell component and to the dysfunction of the cell regulatory system. The effect of a certain stress factor can be very contrasting depending on the strength and time of exposure, in other words the effect of acute stress can be completely different from the effect of chronic stress, as reported in the study made by Choi and colleagues (2021) where was established that acute gamma irradiation of *Oryza sativa* causes instantaneous and significant damages to plant physiology, whereas chronic gamma irradiation leads mainly to reproductive failure causing long-term damage.

Plant stress responses cause a biochemical cascade in which gene expression is altered leading to further downstream changes that result in a change in metabolism to reduce the effect of the stress so the plant can survive/resist the stress. Stress can lead to a phenotypic change, i.e. changes observable, or made

observable by a technical procedure, such as in colour, in growth, or in sugar content. Technological development leads to the ability to study biological organisms such as plants on different biological levels (Fiorani & Schurr, 2013; Wahabzada et al., 2016; Mahlein et al., 2019).

Of interest is the stress caused by heavy metals in the application of using plants to remediate the environment from heavy metals in a process known as phytoremediation, as well as the use of eukaryotic algae and cyanobacteria.

Heavy metal(loid)s are elements with an atomic density greater than $4 - 6 \text{ g cm}^{-3}$, except for arsenic (As), boron (B), and selenium (Se). These elements are also classified as biologically essential such as cobalt (Co), chromium (Cr), copper (Cu), iron (Fe), manganese (Mn), nickel (Ni), and zinc (Zn), as well as non-essential such as As, cadmium (Cd), mercury (Hg), and lead (Pb), in which the formers are needed for the normal biological functions of the cell in specific ranges of concentrations. Nonetheless, they are toxic to organisms when exceeding a certain threshold of concentration, mainly in parts per million (Edelstein & Ben-Hur, 2018). They generally do not get degraded and remain a risk to biological entities for several years after their introduction into an ecosystem (Zielonka et al., 2020; Yadav, 2010; Ahmad et al., 2016; Vardhan et al., 2019).

Heavy metals cause toxicity in plants through four main mechanisms: i) analogy to nutrient cations, that consequently creates a state of competition in their absorption at the root surface e.g. Cadmium vs. Zinc; ii) disruption of enzymes function by changing their structure through interaction with sulfhydryl groups (-SH); iii) replacing enzymes cofactors through irreversible binding; iv) inducing oxidation stress through generation of reactive oxygen species (ROS), that leads to lipid peroxidation (Singh et al., 2016; Ahmad et al., 2016; Dalcorso et al., 2013; Sharma & Dietz, 2009).

Agricultural soils contaminated with heavy metals—either from fertilizers, residues from mining activities, or chemical plants—are of critical concern due to the risk they possess to the surrounding environment. Such risk can be the toxic effect on plants grown in that contaminated soil or the contamination of nearby

water supplies that in either way would end up in the whole food chain, (Edelstein & Ben-Hur, 2018).

The first response of plants upon exposure to a high dose of heavy metals is producing ROS either through Fenton reaction or accumulation of ROS, which will subsequently induce oxidative stress as well as lipid peroxidation (Yadav, 2010). Cd-affected plants have reduced water uptake, nutrient uptake, and photosynthesis. Subsequently, they show chlorosis, inhibition of growth, root tips browning, and eventually death (Wojcik & Tukiendorf, 2004; Mohanpuria et al., 2007; Pandey & Sharma, 2002). Cr-affected plants show chlorosis in young leaves, inhibition of growth, and injuries of the crown and root (Scoccianti et al., 2006). Although regarded as a micronutrient for plants, which is important in ATP synthesis and assimilation of CO₂, excess of Cu in soil induces phytotoxicity, which leads to poor growth and chlorosis (Lewis et al., 2001; Thomas et al., 1998). Ni-affected plants show chlorosis and necrosis, which is probably caused by the imbalance of nutrients as a consequence of the disruption of cell membrane functions, as well as the decrease in water uptake (Rahman et al., 2005; Pandey & Sharma, 2002). Pb-affected plants have water imbalance as a consequence of disruption of the membrane permeability and are also one of the most ubiquitous toxic elements in the soil (Sharma & Dubey, 2005). Even though nutritionally essential to the plant, Zn may also cause phytotoxicity when it is in higher concentrations than required by the plants. Zn toxicity appears as poor growth and decline. Symptoms first appear in younger leaves and extend to older leaves if the exposure to high levels of Zn is prolonged (Choi et al., 1996; Ebbs & Kochian, 1997; Fontes & Cox, 1998). Therefore, the remediation of heavy metals-contaminated soil is a critical prerequisite for the sustainable development of agriculture (Kos et al., 2003). In general, the phytoremediation process as well as the use of eukaryotic algae and cyanobacteria is of great interest for the remediation of metal contamination of soils. Also, the stress caused by heavy metals on the plants used for phytoremediation is an interesting parameter to investigate as well. According to Italian national law (*Decreto Legislativo*

152/2006; Decreto Legislativo 152/2006 - Supplemento ordinario; IP/A/ENVI/IC/2006-183), which is an amendment of the Italian law 99/1992 (Decreto Legislativo 99/1992) that follows the European Economic Council directive 86/278/EEC (86/278/EEC), the maximum levels of heavy metals in soils for the use in agriculture, residence, or green areas is 2, 150, 120, 100, 120, and 150 mg kg⁻¹ of dry weight for Cd, Cr, Cu, Pb, Ni, and Zn, respectively.

Metals are absorbed from the soil into the roots and shoots of the plant by the translocation/phytoextraction process. It is generally not feasible to remove plant roots, and therefore the translocation into the plant shoots is the main goal rather than the translocation into roots (Girdhar et al., 2014). A good candidate for bioremediation of heavy metals-contaminated soil is the hemp plant (Ahmad et al., 2016).

Hemp, the non-psychoactive variety of *Cannabis sativa* L., is an annual dioecious high yielding industrial crop, and it is mainly grown for its fibres and seeds and used for textiles, clothing, insulation, biodegradable plastics, food, animal feed, and biofuel production (Morin-Crini et al., 2019; Praspaliauskas et al., 2020). The bifurcation from the medicinal *C. sativa* L. is in the different ratios of cannabinoids they contain, with hemp having very low to no concentration, usually <0.3%, of the psychotropic compound tetrahydrocannabinol (THC), which in such amount is not adequate for psychological effect (Small & Beckstead, 1973).

Hemp has an industrial-favourable short life cycle of 180 days, a strong capability to eliminate heavy metals from contaminated soils, and even proved to be effective around the Chernobyl nuclear disaster area (Vandenhove & Van Hees, 2005). Hemp was strongly suggested for use in phytoremediation for its feasibility (Kos et al., 2003; Linger et al., 2005; Citterio et al., 2003; Ahmad et al., 2016). Depending on its genotype, hemp can tolerate, accumulate, and stabilize heavy metals. The ability to absorb heavy metals generally depends on the biomass produced as well as the ability of the plant to accumulate and translocate heavy metals in its biomass (Galić et al., 2019; Zielonka et al., 2020).

Spirulina [*Arthrospira platensis*; (Gomont, 1892)] is a filamentous cyanobacterium that is mostly known for its use as a nutritional supplement because of its high nutritional value and high content of protein (Ho et al., 2018; Ye et al., 2018). On one hand, spirulina was used as a plant growth promoter to enhance growth, increase yield, and speed up seed germination of different crops such as rice, cabbage, and other leafy vegetables (Watanabe et al., 1951; Tripathi et al., 2008; Wuang et al., 2016). On the other hand, spirulina was also used for the bioremediation of heavy metals contaminants (Zinicovscaia et al., 2019, 2015; Balaji et al., 2014; Rezaei, 2016; Nalimova et al., 2005; Şeker et al., 2008; Hernández & Olguín, 2002; König-péter et al., 2015; Cepoi et al., 2020).

Spirulina bioremediation of heavy metals depends mostly on the ability to absorb heavy metals inside their cells. This ability is affected by many physical and chemical factors such as initial metal concentration, dosage, adsorption time, temperature, and pH (Rezaei, 2016; König-péter et al., 2015).

Biotic stress in plants is caused by biological agents especially fungi, bacteria, and viruses as well as nematodes, insects, arachnids, and weeds (Dangi et al., 2018). Biotic stress agents directly compete against their hosts for their nutrients causing a diminishment of plant biological functions and in severe cases leading to the death of the plant. Biotic stress is a major cause of both pre- and post-harvest losses.

Plants respond to biotic stress through different defence mechanisms. After infection, plants can produce reactive oxygen species (ROS) to exert an oxidative condition to reduce the pathogen spread (Mithöfer et al., 2004). Also, plants increase cell lignification to block the spreading of parasites inside the plant. Defence mechanisms against biotic stress agents include physical and structural modification, as well as the production of certain metabolites, proteins, and enzymes such as β -aminobutyric acid (BABA).

Plant diseases are caused by pathogenic organisms such as viruses, bacteria, or fungi or by unfavourable environmental conditions such as the excess or deficiency of essential minerals or other substances, as well as by the presence

of toxic substances. Plant diseases cause drastic monetary harm to crops when exceeding the tolerance level. Early detection of plant diseases is essential for managing the diseases and is an integral part of the integrated pest management strategy of a disease. In fact, early detection of biotic stresses, in general, is a critical research area in epidemiology (Zarco-Tejada et al., 2018, 2021; Poblete et al., 2020; Camino et al., 2021; Rey et al., 2019; Galvan et al., 2022).

The severity of biotic stress occurrence depends on the environment, the host plant, and the biotic stress agent itself. This is known to phytopathologists as the plant disease triangle (Fig. 1). In general, hot and humid environments, susceptible cultivars, and lethal strains as well as poor agricultural practices result in a highly-vulnerable condition for the plant.

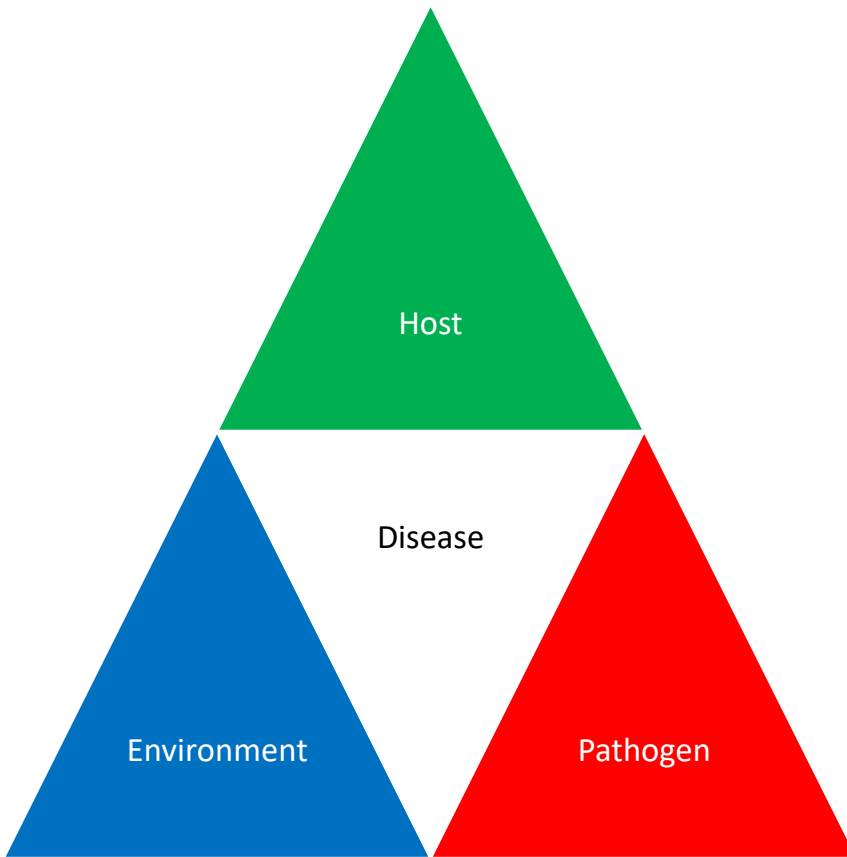


Fig. 1 - The plant disease triangle encompassing the three main axes for disease development. The host, the environment, and the pathogen.

Plant hormones such as ethylene, jasmonic acid, and salicylic acid play important roles in stress signalling. Plant defence mechanism against stress comprises several signal transduction pathways that also includes different transcription factor as mediators.

Xylella fastidiosa is the causal agent of various plant diseases, such as Pierce's disease of grapevine (PD), citrus variegated chlorosis (CVC), phony peach disease, plum leaf scald, almond leaf scorch, and most recently olive quick decline syndrome (OQDS).

Xylella fastidiosa is a gammaproteobacterium in the family *Xanthomonadaceae*. Taxonomically it is classified as follows:

Kingdom: *Bacteria*

Phylum: *Pseudomonadota* (synonym *Proteobacteria*)

Class: *Gammaproteobacteria*

Order: *Xanthomonadales*

Family: *Xanthomonadaceae*

Genus: *Xylella*

Species: *Xylella fastidiosa*

Initially, it was assumed to be a virus until the 1970s when it was proven to be a bacterium (Purcell, 2013). It was first named and described by Wells and colleagues (1987). Hitherto, the genus *Xylella* is composed of only single species, *Xylella fastidiosa*. Furthermore, *Xylella fastidiosa* itself has a considerable amount of genotypic and phenotypic diversity exerting a wide host range of 655 plants so far (Schuenzel et al., 2005; EFSA (European Food Safety Authority) et al., 2022).

Xylella fastidiosa has different subspecies (ssp.). Amongst those, are four widely reported viz. *Xylella fastidiosa* ssp. – *fastidiosa*, *multiplex*, *sandyi*, and *pauca* (Baldi & La Porta, 2017). Isolates within the *Xylella fastidiosa* spp. *pauca* are responsible for CVC in Brazil and are thoroughly characterized (Nunney et al., 2012b). The bacteria responsible for OQDS in Italy is a recombinant of alleles within the spp. *pauca* (Cariddi et al., 2014).

Several genotyping techniques were used to differentiate *Xylella fastidiosa* at different levels of genetic diversity (Sicard et al., 2018). In relationship to *Xylella fastidiosa* subspecies and host plant, multi locus sequence typing (MLST) a procedure for characterising isolates of bacterial species using the sequences of internal fragments of some house-keeping genes (typically) seven, proved to

be a sound method for the understanding of *Xylella fastidiosa* genetic diversity (Nunney et al., 2012a; Elbeaino et al., 2014; Jolley & Maiden, 2010).

Further genotypic analysis using multilocus sequence typing (MLST) found that the strain of *Xylella fastidiosa* spp. pauca infecting olive trees in the Apulia region had a novel sequence type profile (ST53). This strain is also known as the “De Donno” strain as well as the “CoDiRO” strain; an abbreviation of “Complesso del Disseccamento Rapido dell’Olivo” meaning olive quick decline syndrome (Elbeaino et al., 2014).

This plant pathogenic bacterium lives in the xylem tissue of the plant and is naturally transmitted by xylem-feeding insect vectors. It was originally present on the American continents but now spread to different continents. In Europe, it was first detected in the Apulia region in Southern Italy in October 2013 after the large damage in the olive groves (Almeida, 2016) and later in Southern France and Spain. In the case of Italy, the bacterium is suspected to be introduced earlier around 2009, and that is probably due transportation of an infected ornamental coffee plant from Latin America (Martelli, 2016). Furthermore, recent studies suggest that more than a million olive trees in the Salento Peninsula (south of the Apulia region) have been infected during the last couple of decades (A. M. D’Onghia et al.eds. , 2017).

Hitherto, the spread of *Xylella fastidiosa* has not been contained in the Apulia region, which has raised concerns about the risk *Xylella fastidiosa* might appoint to olive trees in the other areas of the Mediterranean region. Transmission of *Xylella fastidiosa* from an infected tree to a healthy one happens through insects that feed on the xylem-sap of the plants. These insect vectors include *Philaenus spumarius* known as the meadow spittlebug that pierces the plant with its stylet to feed on the sap of the tree.

The only available strategy is to uproot and cut infected trees. However, this strategy is not very effective since it is a destructive strategy and most importantly that trees could remain infected while showing slight to no symptoms at all and remaining asymptomatic for more than 5 months (Almeida, 2016)

Xylella fastidiosa remain in an incubation period after infection until favourable conditions are available, e.g. warm temperature. Then, the bacteria start to form a biofilm that blocks the xylem vessels of the olive tree causing a hindrance in the water movement inside the plant and affecting vital processes such as photosynthesis and transpiration. Moreover, the early symptoms of discoloration start to occur on the tree tops followed by defoliation. At later stages, olive trees show symptoms of dieback and then finally the death of the tree (Hornero et al., 2020).

Amongst the detection techniques, qPCR (real-time quantitative Polymerase Chain Reaction) is the most sensitive test. Nonetheless, the feasibility to adopt this technique in field conditions is limited by sampling frequency and the irregular distribution of *Xylella fastidiosa* in a canopy, especially at the early stage of the disease (Zarco-Tejada et al., 2018).

Since the combined agricultural, economical, and environmental adversity caused by *Xylella fastidiosa* in the Apulia region, the European Commission approved countermeasures in February 2014 to limit the spread and further introduction of *Xylella fastidiosa*. The measures included the elimination of infected trees as well as other hosts that are susceptible to *Xylella fastidiosa* within 100 m. Further measures are the treatment against the insect vectors to reduce the rate of transmission as well as the monitoring of areas around the spots where the infection is reported. Nonetheless, such measures were not followed by all member states of the EU (Almeida, 2016).

Xylella fastidiosa occupies the xylem vessels of plants and can move both up- and downstream. Growth of bacterial populations within the xylem inhibits water movement and it was found that disease symptoms development was linked to a high amount of blocked xylem vessels (Sabella et al., 2019, 2020).

Many host plants of *Xylella fastidiosa* remain asymptomatic after infection and act as a reservoir of *Xylella fastidiosa* for the insect vectors to spread (Hopkins & Purcell, 2002). The variability of disease severity from the persistence of asymptomatic infections to plant death within months depends on the host plant species as well as the genotype of the pathogen (Purcell et al., 1999; Purcell

& Saunders, 1999). By adding up the surface area of the infection foci that are individually scattered, around 10,000 hectares of olive were estimated to be affected by *Xylella fastidiosa*. Moreover, this was estimated to have around one million infected trees (Martinelli et al., 2019).

CHAPTER 2

On the omics technologies

Traditionally, managing plant stress consisted of a reductionist approach in research for single components concerning a certain disease instead of the complex interactive system amongst the components of the disease triangle, i.e. the pathogen, the host, and the environmental conditions. Over the past decades, technological development allowed for a more holistic approach to understanding plant stress in which a more comprehensive view is adapted for the identification of key components in the hidden patterns of the mechanisms involved behind (Crandall et al., 2020).

The ability to discern characteristics of such complex systems enabled advancement in management strategies, beginning from plant breeding programs to counter stress factors to agricultural practices to diminish the favourable conditions for plant stress especially plant disease (Moriana et al., 2002). The advances in omics-technologies had a fundamental impact on understanding plant stress, especially with the contemporary climate change and rapid change of environmental conditions (Brito et al., 2019). With the emergence of more biotic and abiotic plant stress factors that affects plant health, it is crucial for plants to adapt to such stress factors and overcome their consequences to be able to manage healthy agricultural systems and safe ecological areas.

Worldwide traveling and international trade can considerably change the ecological makeup of habitat by introducing invasive species that can be pathogenic to plants. This can be even more dangerous when such species are characterized by having a broad range of hosts, this can lead to the death of a significant number of trees within a landscape (Sicard et al., 2018; Scortichini, 2022; Abou Kubaa et al., 2019). Furthermore, climate change can affect the nature of interactions between a stress factor and the plant, as well as add further stress factors to the plant. This could be in the form of variation and fluctuation of

temperatures as well as changing the exposure frequency and duration of a certain meteorological state over a period. Consequently, such variations can lead to unfavourable conditions for the plant, making the plant's health at risk (Crandall et al., 2020)

The Central Dogma of Molecular Biology is a term first coined by Francis Crick in 1956 to introduce his proposal for the generalized process of gene expression. This process is described as the flow of information from the gene, in the form of a DNA sequence, through RNA, and then to protein, in the form of an amino acid sequence (Crick, 1970).

Omics is a term that encompasses a wide range of domains in biology, such as genomics, transcriptomics, proteomics, and metabolomics, which are the science of studying the genome, the transcriptome, and the metabolome of a sample, respectively. The suffix -ome is used to denote “whole”. Therefore, metabolomics can be defined as the study of the whole metabolic makeup of a sample (Wishart, 2011; Fiehn, 2001).

Omics is a suffix that is utilized to refer to research work of analysis of the qualitative and quantitative composition of the constitutive system (Karahalil, 2016; Pinu et al., 2019; Furbank & Tester, 2011). The related derivation suffix “-ome” is added to address the objects in consideration for measurements such as the genome (genetic material), metabolome (metabolites), ionome (ions i.e. elements with an electrical charge), or phenome (phenotypic traits). Therefore, the composition could be, but not restricted to, the genetic material (genomics and metagenomics), RNA transcripts (transcriptomics), proteins (proteomics), metabolites (metabolomics), elements (ionomics), or phenotypic traits (phenomics). Furthermore, the constitutive system could be but is not restricted to, cells, tissues, organisms, or environmental samples (Nielsen, 2017; Baxter et al., 2008; Tardieu et al., 2017).

Metabolomics refers to the branch of ‘omics’ that studies the metabolome. The metabolome refers to the whole set of metabolites found in a sample. Metabolites are defined as low atomic mass molecules of <1.5 kDa, that include

a wide range of endogenous and exogenous small molecules such as peptides, amino acids, nucleic acids, carbohydrates, organic acids, vitamins, drugs, drug metabolites, food additives, phytochemicals, toxins, and the vast array of chemicals that can be used, ingested or synthesized by a certain cell or organism (Wishart, 2008a; Fiehn, 2002).

The term “metabolomics” was first introduced by Oliver Fiehn (Fiehn, 2002). Nonetheless, Jeremy Nicholson and colleagues introduced the term “metabonomics” earlier in 1999 and defined it as “the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification” (Nicholson et al., 1999). Someone might argue that metabonomics would rather be concerned with endogenous metabolites (Ramsden, 2015). Nevertheless, both metabolomics and metabonomics terms are used interchangeably (Alseekh & Fernie, 2018; Ciampa et al., 2022).

The omics cascade is the flow of biological information from the genotype to the phenotype. This conforms to the central dogma of molecular biology that indicates the flow of information from DNA to RNA to proteins (Crick, 1970). The metabolome is downstream of the omics cascade and is the most sensitive to environmental or disease effects (Feussner & Polle, 2015). This is because metabolites are closer to the phenotype than genes and proteins. Changes at the genes and proteins level due to a pathogen include epigenetic and post-translational modifications, respectively. However, the effect of such changes is amplified at the metabolic level. Thus, the metabolomic profile is more consistent with the phenotype as opposed to the genomic, transcriptomic, and proteomic profiles in descending order (Fiehn, 2001; Kim & Lun, 2014).

Metabolomics is widely applied in studying various topics such as disease diagnosis, crop breeding, detection of adulteration of food, food authentication, tracing the origin of food production, metabolic pathway analysis, genetic engineering, monitoring of medical interventions effects as in organ transplantation, and effects of aging—nutrition—activities—and lifestyle—on the metabolic

networks (Taha et al., 2022; Razzaq et al., 2022). Metabolomics is less developed than genomics and proteomics in terms of databases, identification technologies, and algorithms (Ryan & Robards, 2006; Hall et al., 2022; van der Hooft et al., 2020).

Generally, there are two different schools of thought in studying metabolomics. On one hand, the quantitative approach, also mentioned as targeted approaches and quantitative or targeted metabolomics/profiling/methods, focuses on seeking to identify and/or quantify as many metabolites as possible in the sample (Wishart, 2008a). This is normally performed using a reference library of pure spectra of single compounds to compare against the sample's data. Moreover, after the metabolites in the sample are identified and quantified, the data are then further analysed using advanced statistical methods, such as principal component analysis or projection to latent structures-discriminant analysis to highlight significantly influential biomarkers. The quantitative approach may focus on targeting certain metabolites of interest, e.g. a class of compounds such as organic acids, amino acids, or lipids, or opt for a comprehensive quantification of all or almost all possibly detectable metabolites (Gika et al., 2014; Zheng et al., 2021b). On the other hand, the chemometric or non-targeted approach focuses on the statistical comparison of spectral patterns and intensities for identifying the relevant features that distinguish sample values or class membership. Then, after such important features are found, their corresponding metabolites are identified (Wishart, 2008b).

The most commonly used platforms in metabolomics studies are Nuclear Magnetic Resonance (NMR) and spectroscopy and Mass Spectrometry (MS) - based techniques (Labine & Simpson, 2020). In general, both platforms have their advantages and disadvantages, mainly correlated to reproducibility and sensitivity (Wishart, 2016; Pinu et al., 2019).

Regardless of the approach followed, a metabolomics study workflow adheres to typical steps. First, the hypothesis is formed in light of the biological question to be answered. Second, the experiment is designed accordingly to

reduce the source of error as much as possible and with careful consideration for the proper sample size. Third, the samples are collected and stored in a way that maintains their metabolomic status at the time of sampling. Fourth, the samples are prepared for analysis by the technique of interest. Fifth, the samples are analysed by the platform of interest, and the data are acquired. Sixth, the data are analysed following either a targeted or a non-targeted approach. Finally, results are interpreted taking into account the biological question (Wishart et al., 2022b; Bedia et al., 2018; Gorrochategui et al., 2016).

Ionomics is the process of studying the ionome as a whole, that is the elemental composition of a cell, tissue, organism, or other samples of interest (Baxter, 2009). Elemental analysis is commonly achieved by Inductively Coupled Plasma spectrometry, either Mass Spectrometry (ICP-MS) or Atomic Emission Spectrometry (ICP-AES) (Anguita-Maeso et al., 2021). This is very useful when studying the ability of plants to sequester heavy metals from soil (Musio et al., 2022).

Spectranomics combines the use of spectroscopy with physiochemistry and taxonomy. The recently introduced term was first mentioned by Greg Asner and colleagues (Asner & Martin, 2009). The premises on which spectranomics is built are that plants exert chemical fingerprints that are much more unique as an additional influential factor, such as genotype, phenological stage, and environmental conditions, is added and that spectroscopic signatures confer the chemical profile of the plants (Jacquemoud et al., 1995; Asner & Martin, 2011; Asner et al., 2014a; Feilhauer et al., 2015; Ustin et al., 2004). This concept proved to be useful in relationship plant metabolite (Fine et al., 2021; Gold et al., 2020a).

CHAPTER 3

On the spectroscopic methods for omics technology

As a spectroscopic method, Nuclear Magnetic Resonance (NMR) spectroscopy is based on measuring the interaction between electromagnetic waves and matter. However, what distinguishes it from other spectroscopic methods is that it requires a strong, static magnetic field.

NMR spectroscopy is one of the main techniques used in metabolomics. The potential of NMR spectroscopy is well known for the qualitative and quantitative analysis of chemical compounds in complex mixtures without the need to separate the chemical compounds as in mass spectrometry where they are coupled to a separation technique such as liquid or gas chromatography to be able to measure the compounds. Furthermore, NMR spectroscopy is the most robust technique for the elucidation of unknown compounds. Since it is not selective, it allows for an unbiased view of the chemical composition of the sample. Moreover, NMR samples can be reused for subsequent analyses as it is not required for the sample to be pre-treated or destroyed.

The nuclei of different atoms from the periodic table can be studied using NMR spectroscopy but, considering biological samples, the most valuable nuclei are ^1H , ^{13}C , ^{15}N , and ^{31}P . Amongst the mentioned nuclei, NMR is the most sensitive for the proton (^1H) as its natural abundance is very close to 100%. Although ^{31}P comes second for sensitivity, because many metabolites studied are phosphorylated. However, ^1H is the most studied nucleus in NMR spectroscopy-based metabolomics.

The NMR phenomenon was first discovered by Isidor Isaac Rabi and colleagues in 1938 (Rabi et al., 1938). He was then awarded the Nobile Prize in Physics in 1944 for his resonance method for recording the magnetic properties of atomic nuclei. The work of Rabi and his colleagues was successful in measuring the magnetic properties of different isolated nuclei. However, it was not until

1954 that NMR was demonstrated in condensed matter of water and paraffin by Felix Bloch and his colleagues at Stanford University and Mills Purcell and his colleagues at MIT, respectively (Bloch et al., 1946; Purcell et al., 1946). Both Bloch and Purcell were awarded the Noble Prize in Physics in 1952 for their development of new methods for nuclear magnetic precision measurements and discoveries in connection therewith. Furthermore, NMR spectroscopy applications in chemistry were not widespread until the discovery of the “chemical shift”; this is the phenomenon by which an atom is energetically affected by the interactions of its chemical surrounding.

Nuclear Magnetic Resonance is an intrinsic property of the nucleus inside the atom based on the nuclear magnetic moment that arises from the spin of protons and neutrons. Thus, we can consider the nucleus of the atom to have the ability to act as a special magnetic bar (Schmidt-Böcking et al., 2016). Furthermore, the spin angular momentum (I) is a conserved quantity carried by elementary particles and their composite. The spin of an atomic nucleus depends on the mass and the charge of the nuclei. The nuclei with even mass and charge numbers have no spin angular momentum. Such nuclei with no spin angular momentum are called NMR inactive or NMR silent nuclei because the nuclear spin is essential to allow NMR to occur. Proton (^1H), carbon (^{13}C), fluorine (^{19}F), and phosphorus (^{31}P) have $I = \frac{1}{2}$ and therefore NMR can be achieved and they can be analysed by NMR spectroscopy.

The nuclei spins are randomly oriented in every direction when they are not affected by any external magnetic field. However, if nuclei of $I \neq 0$ are kept in a magnetic field (B_0), they will undertake a possible number of different orientations, equal to $2I+1$, that will correspond to specific energy levels.

Applying this to the proton (^1H), the B_0 applied will make the protons undertake two possible orientations *viz.* α parallel of $I = \frac{1}{2}$ or β antiparallel of $I = -\frac{1}{2}$, with each corresponding to an energy level. The difference between the energy levels is equal to:

$$\Delta E = (h/2\pi) \cdot \gamma \cdot B_0$$

(1)

Where h is the Planck constant $6.62607015 \times 10^{-34}$ J·Hz, γ is the gyromagnetic ratio, which is a constant for a given nucleus. For the specific case of the proton its value is $\gamma_p = 2.6752 \times 10^8 \text{ s}^{-1} \cdot \text{T}^{-1}$, and B_0 is the applied magnetic field

Therefore, for an NMR active nucleus, the stronger the magnetic field the higher the gap in the energy levels observed, and, thus, higher the sensitivity in the NMR spectroscopy.

The protons are distributed according to the Boltzmann distribution between two energy states:

$$N_\alpha/N_\beta = e^{-\Delta E/(k T)}$$

(2)

where N_α and N_β are the number of protons in the lower and upper energy levels, respectively, k is the Boltzmann constant of 1.380649×10^{-23} J K⁻¹, and T is the temperature in kelvin (K).

The two states do not have an equal number of nuclei since the lower energy state (α) is energetically more favourable than the higher state (β). The net magnetization, M_0 , will be aligned with the applied magnetic field, B_0 . When a radio frequency pulse is applied, the nuclei will absorb energy and will induce a nuclear spin transition from lower to higher energy levels. The energy gap can be calculated as follows:

$$\Delta E = h \cdot \nu$$

(3)

Where h is Planck's constant = while ν is the frequency, of the excitation pulse, that induces the transitions between the energy levels. This frequency depends on the nucleus to be excited and the magnetic field applied and is known as the "Larmor frequency":

$$\nu = (\gamma \cdot B_0) / (2\pi)$$

(4)

When the sample is under a static magnetic field, NMR signals are not observable since the nuclei are at equilibrium and the net magnetization vector will have no component on the xy plane at which the signal is detected by the detector coil inside the machine. A 90° pulse brings the magnetization to the xy plane and a 180° pulse flips the net magnetization on the negative z -axis. The duration of the magnetization pulse is usually very short and measured in microseconds (μs). When the applied radiofrequency is turned off, the system will return to equilibrium, which is known as relaxation. The relaxation causes the NMR signal to decay over time and will produce the observed free induction decay (FID). The FID signals are in the time domain and are converted to the frequency domain using Fourier transformation (FT), which will produce the raw NMR spectra.

Since the proton is located in different positions in the molecule, they experience different magnetic fields because of the "shielding" phenomenon. This phenomenon is observed when the nucleus is surrounded by electrons that start a protective rotational motion when an external magnetic field (B_0) is applied. A local small magnetic field B_{loc} is generated, that may resist the external applied magnetic field. Thus, the shielded nucleus experiences a slightly reduced applied magnetic field. This reduced magnetic field is known as B_{eff} . Moreover, the density and the distribution of the cloud of electrons surrounding the nucleus are indicated by the shielding constant (σ). This constant ranges from 10^{-6} to 10^{-3} for lightly- and heavily shielded nuclei. This shielding effect is affected by the

presence of different functional chemical groups near the nucleus. For instance, an electronegative atom will withdraw electrons from the nucleus and, thus, reduce the density of the cloud of electrons creating what is known as the “deshielding” effect. Consequently, this deshielded nucleus will resonate at higher frequencies. Following Lenz’s law, B_{loc} is considered equal to $B_0 \sigma$ and B_{eff} can thus be calculated as follows:

$$B_{eff} = B_0 - B_{loc} = B_0 - B_0\sigma = B_0(1 - \sigma) \quad (5)$$

Therefore, the relationship between the degree of shielding and the resonance frequency can be given by:

$$\nu = (\gamma \cdot B_0) / (2 \cdot \pi) \cdot (1 - \sigma) \quad (6)$$

As a result, the protons located in different chemical environments will experience different magnetic fields. Therefore, these protons will resonate at different frequencies depending on their location in the molecule. Such effect is defined as the “chemical shift” (δ). The chemical shift is the basic metric for NMR experiments and it is rather fundamental to interpret the frequencies of different nuclei independently from the applied magnetic field. Therefore, the chemical shift is calculated as follows:

$$\delta = \nu \cdot \nu_{ref} / \nu_{ref} \cdot 10^6 \quad (7)$$

where ν is the frequency of the observed nucleus and ν_{ref} is the frequency of a reference compound. The main reference compounds used for ^1H NMR are tetramethylsilane (TMS) and the sodium salt of trimethylsilyl propionic acid (TSP)

for organic solvents and aqueous solutions, respectively. These compounds have the maximum shielding that shows a higher shielding constant (σ) than most nuclei analysed via NMR. Precisely, the frequency of the observed nucleus of TSP and TMS (ν) is equal to the Larmor frequency of the proton, due to their shielding constant of 1, thus their measured chemical shift is equal to 0 ($\delta_{\text{ref}} = 0$).

The last equation converts the chemical shift frequencies into parts per million from the Larmor frequency. This allows the depiction of NMR signals on an axis that is independent of the applied magnetic field. In this way the same spin will always have the same value regardless of the applied magnetic field. Information exerted from the NMR spectra does not confine to the chemical shift only, but other valuable features are present. Another feature is the J coupling. This phenomenon appears as a result of the influence wielded by the magnetic moments of the nuclei on each other through space or chemical bonds, known as dipolar and scalar coupling, respectively. Attention is given to the scalar coupling since it is visible from the NMR spectrum rather than the effect of the dipolar coupling that can be neglected due to the rapid molecular tumbling that results in interactions that average 0.

Moreover, the influence of one atom on another produces a split in the resonance signal acquired. This is visible for both nuclei, in which the perturbing nucleus will display a split in its signal when it is observed. This phenomenon is also known as spin-spin splitting. However, the observation of signal splitting by the interacting nuclei can occur when the interacting nuclei are bonded by proximity as in vicinal (neighbouring carbons) and geminal positions (same carbon), or being oriented in specific optimal configurations. The signal splitting is usually measured in Hz (Hertz, s^{-1}) and in ^1H NMR spectra can range from 1 to 18 Hz.

In order to enhance the sensitivity and resolution of the acquired data, different processing steps are performed before and after the Fourier transformation of the FID. Of interest are phase correction and baseline correction. Phase correction is applied to correct the phase errors in the spectra, as the name suggests. The errors in the phase come from two main sources: the delays between

the RF pulse and the receiver opening for FID acquisitions as well as the inability of the RF pulse to excite all the nuclei equally, which are referred to as off-resonance effects. Without being correctly phased, the NMR spectrum would have signals with dispersive line shapes and might as well have inverted signals. For such reasons, phase correction by both zero order and first order is applied. On one hand, the zero-order phase correction is independent of the chemical shift and affects all lines across the spectrum equally. On the other hand, the first-order phase correction is dependent on the frequency, and therefore on the chemical shift. The phase change increases linearly with the distance from the reference signal.

Baseline correction aims at solving issues usually caused by having the first few corrupted data points in the FID that add low frequency modulations during Fourier transformation, giving rise to a distorted baseline. This comes from two main sources: a too high signal amplification is applied or an incomplete recovery of information from the RF. Such a distorted baseline is commonly corrected using a polynomial fitting and then subtracting the offset from the spectrum.

For metabolomics research, one-dimensional (1D) ^1H NMR spectroscopy is the most widely applied, mainly due to the less required time for spectral acquisition as compared to 2D NMR. The most widely used pulse sequence for ^1H NMR is the 1D Nuclear Overhauser Effect Spectroscopy (NOESY). The 1D NOESY pulse sequence consists of a first increment of NOESY pulse sequence, with water suppression during the relaxation delay as well as during the mixing time. The pulse sequence form $-\text{RD}-90^\circ-\text{t}_m-90^\circ-\text{t}_m-90^\circ-\text{ACQ}$. Where RD, t, t_m , 90° , and the ACQ represent the relaxation delay, the short delay time of around $3\ \mu\text{s}$, the mixing time, the RF pulse degree, and the data acquisition period. Using this pulse sequence, all non-exchangeable protons can be detected in a deuterated buffer solution (Stryeck et al., 2018).

For ionomics studies, the samples are usually analysed by techniques such as the Inductively Coupled Plasma Atomic Emission Spectroscopy (ICP-AES)

following a standard procedure of sample drying and crushing to a fine powder, digestion for around 1h to 10 min at around 100°C to 180°C, respectively, using a microwave digestion system in mineral-free concentrated acid such as nitric acid. Then the samples are washed and subjected to simultaneous measurement of elements of such as Ca, Co, Cu, Fe, K, Mg, Mn, Mo, Na, P, S, and Zn. The concentrations of the measured elements are determined by comparing the emission intensities to a standard curve made from standard certified mineral solutions diluted serially with the same solvent used for the measured samples. The results are denoted as $\text{mg} \cdot \text{kg}^{-1}$ of dry matter (Musio et al., 2020).

As for spectranomics and phenomics studies, technologies based on spectral data acquisition have a broad application in both fields as they can be also for proximal and remote sensing applications in digital agriculture, which is also known as Agriculture 4.0. These technologies allow for the possibility of measuring features associated with plant physiology (Asner et al., 2014a). Moreover, these technologies are applied at different scales of studies, such as from the leaf-, to the canopy-, to the tree-, to the field- level.

In the last few decades, the leaf spectral properties have gained attention, probably due to the technological development that allowed for such studies. The leaf spectral properties follow the principle that any material subjected to a different wavelength of electromagnetic (EM) waves can reflect, absorb and transmit the EM waves differently, and thus different materials can be characterized based on their different spectral signature. The leaf spectral properties indicate the amount, in percentage, of incident radiation that is reflected, absorbed, and transmitted by the leaf surface. These properties are known as reflectance, absorbance, and transmittance. For homogeneous isotropic media without particles, transmission measurements respect the Beer–Lambert (Bouguer-Beer-Lambert) law:

$$A_{\lambda,c} = -\log(T_{\lambda,c}) = \epsilon_{\lambda} L c$$

(8)

Where $A_{\lambda,c}$ the absorbance, $T_{\lambda,c}$ the transmittance, ϵ_{λ} the extinction coefficient, L the light path length, c the concentration of the absorbing species, and λ the wavelength of the EM waves. In scattering media, two a modification of path length $L_{\lambda,c}$ and a loss of photons $f_{\lambda,c}$:

$$A_{\lambda,c} = -\log(T_{\lambda,c}) = \epsilon_{\lambda} L c + f_{\lambda,c} \quad (9)$$

The measured EM wavelengths for the leaf are usually in the range of 350 and 2500 nm, which consists of parts of ultraviolet (UV, 350-400 nm), visible (VIS, 400-780 nm), and near-infrared (NIR, 780-2500 nm) regions of the EM spectrum. The measurement can be carried out using a spectroradiometer that can measure EM wavelengths. Depending on the spectral resolution acquired, the spectral data are mainly classified as hyperspectral (narrow band) data with a final resolution of 1 nm wavelength bands or multispectral (wide-band) data with a final resolution around 10-20 nm wavelength bands. For healthy leaves, the low spectral reflectance in the portion of the blue (450 nm), green (550 nm), and red (670 nm) bands are related to the absorption by the leaf pigments such as chlorophyll. In the NIR region (700-1,300 nm), the high reflectance of the leaves can be related to cellular structure and the presence of water affects the region at 1450 nm (Carter & Knapp, 2001; Li et al., 2021; Couture et al., 2016; Zheng et al., 2021a).

Such spectral data can show different trends and patterns related to the plant species, cultivars, nutritional conditions, and more recently abiotic and biotic stress conditions (Galvan et al., 2022). Several studies showed a correlation between the spectral reflectance and leaf traits for different herbaceous, horticultural, and arboreal species in different agricultural and forest conditions (Asner et al., 2014a; Fine et al., 2021; Ustin et al., 2004; Asner & Martin, 2009,

2016, 2011; Feilhauer et al., 2015; Galvan et al., 2022; Mahlein et al., 2018; Rumpf et al., 2010).

Remote/proximal sensing offers a non-destructive way of measuring electromagnetic waves reflected by the plant surface, in which specific spectral bands/signatures are used to detect infections before the development of symptoms (Mahlein et al., 2018). Moreover, spectral signatures determine a gamut of chemicals found in plants, and the specificity of plant chemical fingerprints is reported to be strongly influenced by several factors (Jacquemoud et al., 1995; Curran, 1989). Furthermore, hyperspectral reflectance data has been proven to provide valuable information on plant biophysical and biochemical features. These features include non-pigment biochemical features such as nitrogen, phosphorus, cellulose/lignin, and water content as well as pigment biochemical features such as anthocyanins, carotenoids, and chlorophylls, in addition to biophysical features such as plant and leaf area, biomass, stem diameter, and estimated height (Pacumbaba & Beyl, 2011; Zheng et al., 2021a; Calderón et al., 2013; Nguyen et al., 2021; Asner et al., 2014b; Féret et al., 2021; Couture et al., 2016; Fine et al., 2021; McManus et al., 2016).

CHAPTER 4

On chemometrics for the analysis of spectroscopic data

Chemometrics is a branch of analytical chemistry that is concerned with the study of chemical and biochemical measurements. As the name suggests, “chemo-” implies chemical and “-metrics” implies measurement. Chemometrics addresses the application and development of mathematical and statistical methods to extract information from chemical data, usually hidden within the underlying patterns in the data. Chemometrics also addresses appropriate data collection, quality enhancement for the analytical signal through the identification of noise sources and reducing it, and models building that predict or classify future measurements (Sumner et al., 2007; Yi et al., 2016; Martens, 2015). Spectroscopic techniques record thousands of variables for every single sample analysed, which produces high-dimensional datasets that are not suitable for univariate data analysis. Chemometrics applies multivariate data analysis/machine learning taking into account causality to understand the hidden underlying pattern that contains valuable chemical information (Martens, 2021; Vitale et al., 2022).

Chemometrics techniques can also be categorized depending on the type of analysis, which includes exploratory, classification, and predictive analyses. Exploratory analysis unsupervised methods allow for an unbiased overview of the data to reveal general patterns and detect anomalies. Classification analysis allows for the distinction between groups, and in the case of biochemical data, it is used to find possible biomarkers that discriminate between groups. Predictive analysis allows for the quantitative comparison between blocks of data (Yi et al., 2016; Tortorella & Cinti, 2021; Powers & Riekeberg, 2017).

Unsupervised and supervised methods or machine learning are terms that are also used to categorize chemometric techniques. On one hand, unsupervised machine learning techniques look for the natural variation within the given data

without using *a priori*, which allows for exploratory analysis. On the other hand, supervised machine learning techniques look for variation between the samples in the given data based on prior knowledge of the sample class association or quantitative information, thus allowing for classification between groups of samples and predictive analysis of blocks of data (Granato et al., 2018; Behmann et al., 2015; Gromski et al., 2015).

Considering the chemometrics or non-targeted approach (mentioned in chapter 2), spectroscopic data, in most cases, are not naturally ready for downstream analysis by chemometric methods. For NMR spectra, the intensity of each data point of the spectrum is considered. For hyperspectral reflectance each wavelength band (1 nm) is considered (Kaddurah-Daouk et al., 2008; Gold et al., 2020b).

Bucketing or binning is the process of segmenting the spectra into equal or non-equal distant segments. Thus, aiming at removing residual noise, such as chemical shift drift for NMR spectra, and decreases the complexity of the data. For NMR data, small buckets of around 0.04 ppm are considered adequate to include variation in peaks of NMR signals. Each bucket is calculated using the area under the curve in the bucket, which gives the bucket intensity. Regular size bucketing, in which all buckets would have the same size e.g. 0.04 ppm, diminishes some chemical shift misalignments and decrease the noise in the spectra but could rather mask low-intensity peaks around huge signals as well as peak shapes. For hyperspectral data, bucketing of the spectra usually is done not to increase the resolution in the first place but mainly to decrease the amount of data points for computational analysis as well as for developing multispectral sensors from hyperspectral data. Spectral buckets of 10 nm bandwidth are the most common size for buckets and, usually, the average value within the bucket is taken (Parsons et al., 2009; Paulus & Mahlein, 2020; Augustijn et al., 2021; Abdulridha et al., 2016).

The normalization of the features present in each sample aims to make the observations comparable to each other. For NMR data, normalization can be

applied using an internal standard, a reference peak, Probabilistic Quotient Normalisation (PQN), an artificial signal, or the total intensity. Normalization is a row-wise operation since it is applied on the observation (sample) level (Gorrochategui et al., 2016; Wang et al., 2022; Weiss et al., 2020).

Scaling the dataset is also common in chemometrics, in which a mathematical operation is applied to the dataset to balance signal intensity variances. Unlike normalization methods, scaling is a column-wise operation that affects all observations together on the feature level. Scaling is applied to account for small, but maybe important, variations within features with low intensity. Scaling methods include mean centring, autoscaling (also known as standardization or unit variance scaling), and Pareto scaling (Wheelock & Wheelock, 2013).

Mean centring transforms the values of the features so they would represent variance around zero instead of the mean value. Thus, the values are now considered as distances from the mean instead of zero. This is applied by subtracting the column (feature) mean intensity from each value. Furthermore, autoscaling, or standardization or unit variance scaling, applies a unit variance to the features by dividing the feature by its standard deviation after mean centring, giving each feature a substantial amount of influence on the dataset. However, autoscaling may result in noise inflation if small variations in a distorted baseline are considered. Moreover, Pareto scaling follows the same principle of autoscaling but rather divides the mean-centred feature with the square root of its standard deviation, hence reducing the possibility of noise inflation accounted for by small variations (Femenias et al., 2021; Del Coco et al., 2021; Pane et al., 2022; Taraji et al., 2017).

Principal Component Analysis (PCA) is one of the most used unsupervised techniques in chemometrics. It was initially conceptualized and designed by Karl Pearson in 1901 in analogy to the principal axis theorem in mechanics (Pearson, 1901). Nevertheless, Harold Hotelling further developed PCA to its present form and gave it its current name (Hotelling, 1933). Furthermore, PCA started to gain attention in chemistry in the '60s (Wold et al., 1987).

PCA is defined as an unsupervised pattern recognition method that applies dimensionality reduction to provide a visual representation of the major variance of a dataset that consists of a large number of variables. PCA considers the dataset as a data matrix of N rows and K columns of observations and variables, respectively (Bro & Smilde, 2014).

On one hand, the observations are the samples, which can be biological samples from different individual organisms, technical/analytical samples, continuous process time point samples, batches from process, and so forth. On the other hand, variables, or features, are the values that can arise from measurements of different origins, including spectroscopic measurements such as NMR spectroscopy and hyperspectral reflectance (Rizzuti et al., 2018; Musio et al., 2022; Gupta et al., 2021).

Principally, the original set of variables is transformed into a new smaller set of variables that are orthogonal i.e. uncorrelated. These new orthogonal variables are called Principal Components (PCs) and are organized in descending order based on the amount/percentage of variance explained by each component. Thus, the first PC has the highest explained variance and the first few PCs, in most cases, contain most of the relevant information about the variance within the dataset. Each PC has two outcomes that are named scores and loadings that represent the observations and the variables, respectively (Tang & Hatzakis, 2020; Cozzolino et al., 2019; Jilil et al., 2021).

When plotted, the scores plot and the loadings plot will show points that represent the observations and the variables, respectively, with the new transformed values for each PC. These two plots are inspected together since the loadings plot shows the influence of variables on how the observation is represented in the scores plot. Furthermore, usually each two or three PCs are plotted in 2D or 3D plots, respectively, to visualize possible groupings or trends across the axes and the variables responsible for such groupings or trends (Luo et al., 2016; Hall et al., 2022).

The PCA mathematical model is represented as follows:

$$X = TP^T + E$$

(10)

where T is the scores matrix, P^T is the transposed loadings matrix, and E is the residual matrix.

The scores can also be considered as elements of the matrix T, in which each row is a single observation, and each column is the value of the observation along each PC. The loadings can also be considered as the weights of the variables for each PC. The residuals are not part of the PCA model and therefore should be minimized to reduce the amount of unmodeled information. Nevertheless, too many components would include irrelevant information that might come from noise in the measured data. Therefore, the data matrix is considered to be composed of two parts, a structure part and a noise part:

$$X = TP^T + E = \textit{Structure} + \textit{Noise}$$

(11)

$$\textit{Explained Variance} + \textit{Residual Variance} = 100\%$$

(12)

Partial Least Squares regression or discriminant analysis (PLS regression or PLS-DA, respectively), also known as a projection to latent structures, are supervised techniques applied also in chemometrics that use *a priori* knowledge about the observations, such as quantitative information or class membership, respectively (Brereton, 2015).

As a multivariate method, PLS evaluates the relationship between two blocks of data considering a descriptor matrix X that contains the measured data and a response matrix Y that contains the quantitative or class membership. When the Y matrix consists of quantitative data, the analysis is known as PLS

regression, while when the Y matrix consists of qualitative data, the analysis is called PLS-DA. Therefore, PLS-DA is different from PLS regression in the way the responses matrix Y is used:

$$X = TP^T + E$$

(13)

$$Y = UQ^T + F$$

(14)

where X is the descriptors matrix and Y is the responses matrix. T and U are the scores matrices, P^T and Q^T are the transposed loadings matrices, and E and F are the residuals matrices for X and Y, respectively.

The PLS model is built using the descriptor X and response Y matrices to highlight possible correlations. PLS performs a PCA, in which from the matrix X, a scores matrix T and loadings matrix P is obtained, while for the matrix Y, a scores matrix U and loadings matrix Q is obtained. Then, PLS uses the scores of the Y matrix to decompose the X matrix and calculate its loadings and the scores of Y scores is substituted with the X scores to decompose Y. Thus, the decomposition is influenced mutually and it is iterated until this process converges. This process seeks the intercorrelation between descriptor X and response Y matrices to reduce the variance of descriptor matrix X that is not correlated to response matrix Y. These decompositions of descriptor X and response Y matrices are performed to maximise the covariance between scores T and scores U matrices. In PLS-DA, the maximum separation amongst the classes is done using a dummy matrix Y that accounts for the variations amongst classes (Triba et al., 2015; Eriksson et al., 2001; Innamorato et al., 2020).

Correlation analysis and heatmaps representation allow for seeking a straightforward relationship between variables. It is regularly based on pairwise comparison, most commonly Pearson's correlation for linear correlation

between correlations as well as other correlation analyses such as Spearman correlation for non-linear correlation. Nevertheless, correlations do not entitle causations per se (Fine et al., 2021; Rizzuti et al., 2018; Rosato et al., 2018).

Furthermore, heatmaps can be considered as a tool to visualize correlation using a coloured scale to indicate the degree of positive and negative correlations. Nonetheless, other parameters besides the degree of correlation should be considered when analysing heatmaps, such as the significance of the correlation (p -values) as well as visualizing the correlation plots (Picone et al., 2016).

CHAPTER 5

A spectroscopic study to assess heavy metals absorption by a combined hemp/spirulina system from contaminated soil

In this chapter, the peer-reviewed published work in collaboration with ApuliaKundi S.r.L. in the project “BIO.SP.HE.RE” (BIO-integrated SPirulina and HEmp REmediation) is presented. This work describes the application of a combined system constituted of hemp (*Cannabis sativa*) of the industrial cultivar *Kompolti* and Spirulina algae (*Arthrospira platensis*). Hemp is very suitable for phytoremediation not only because it showed a strong ability to sequester heavy metals (such as Cd, Zn etc) in soil and in water, but also because it has high biomass, long roots and a short life cycle of 180 days. On the other hand, the use of eukaryotic algae and of cyanobacteria for the remediation is documented, although mainly for the treatment of wastewater. Among the cyanobacteria, *Arthrospira platensis* possesses excellent chelating properties towards heavy metals but is also effective as a plant growth promoter, enhancing the growth, increasing the yield, and speeding up the seed germination. In this study, hemp was chosen as the main agent for biological remediation, and spirulina was added as an enhancer of both the plant growth and the translocation of heavy metals in the hemp.

This system was tested for the bioremediation of soil contaminated with cadmium, chromium, nickel, lead, and zinc. The quantitative analysis of heavy metals inside parts of the hemp plant as well as in the contaminated soil provided valuable information on the accumulation of heavy metals in different parts of the hemp plant as well as the effect of the bioremediation system on the contaminated soil. The non-targeted NMR approach showed different metabolic responses to heavy metals by the hemp plant that was induced by the addition of spirulina that generally was associated with a positive health condition of hemp.

This published work is a prominent step towards the optimisation of the hemp-spirulina system in terms of spirulina dosage as well as for the localisation of heavy metals in specific tissues of hemp for the reuse of plant biomass of other useful purposes such as in the energy production. The article was first submitted to the journal “Chemosphere”, peer-reviewed, and transferred for publication in the journal “Environmental Advances” (Musio et al., 2022).



A spectroscopic study to assess heavy metals absorption by a combined hemp/spirulina system from contaminated soil

Biagia Musio^{a,b,*}, Elhussein Mohamed Fouad Mourad Hussein Ahmed^{a,c}, Marica Antonicelli^a, Danila Chiapperini^d, Onofrio Dursi^d, Flavia Grieco^d, Mario Latronico^{a,b}, Piero Mastrotrilli^{a,b}, Rosa Ragone^{a,b}, Raffaele Settanni^d, Maurizio Triggiani^{a,b}, Vito Gallo^{a,b,e}

^a Department of Civil, Environmental, Land, Building Engineering and Chemistry (DICATECh), Polytechnic University of Bari, Via Orabona, 4, Bari 70125, Italy

^b Innovative Solutions S.r.l. Spin-off company of Polytechnic University of Bari, Zona H 150/E, BA, NoCI 70015, Italy

^c International Centre for Advanced Mediterranean Agronomic Studies of Bari (CIHEAM Bari), Via Ceglie 9, BA, Valenzano, Apulia 70010, Italy

^d ApuliaKundi S.r.l., via N. Copernico snc, Zona PIP, BA, Gravina in Puglia 70024, Italy

ARTICLE INFO

Keywords:

Non-targeted nuclear magnetic resonance
Phytoremediation
Phycoremediation
Arthrospira platensis
Cannabis sativa L.
Metal quantification

ABSTRACT

The efficiency of hemp (*Cannabis sativa* L.) in remediating sites contaminated with heavy metals has received great attention in recent years. The main advantage of this technology relies on its inherent sustainability with a potential re-utilization of the significant amount of produced biomass which acts as a valuable flow resource. In this study, a combined system consisting of *Cannabis sativa* L. (hemp) and the blue-green alga *Arthrospira platensis* (spirulina) was tested to clean up soils contaminated with cadmium, chromium, copper, nickel, lead, and zinc. The application of non-targeted NMR methods combined with ICP-AES quantification provided an efficient strategy for detecting residual heavy metals within plant tissues and soil. Importantly, non-targeted metabolomic analysis helped to reveal the relationships between metabolites distribution in hemp tissues and the sequestered metals. It was demonstrated that hemp accumulates copper, chromium, nickel, and zinc preferentially in the leaves, while lead is distributed mainly in the stems of the plant. Moreover, it was found that, at higher concentrations, spirulina acts as a growth promoter, contributing to an increase in the final generated biomass. Results reported in this work indicate that the hemp/spirulina system represents a suitable tool for remediation of metal contaminated soils by modulating biomass production and metals uptake.

1. Introduction

Dispersion of heavy metals in soils is an age-old problem deriving from both natural and anthropic sources (Awa and Hadibarata, 2020). Among the anthropic contribution to soil contamination by metals, land application of treated wastewater, sewage sludge, fertilizers, and industrial activities are major concerns (Vareda et al., 2019). Unbalanced amounts of heavy metals may cause perturbation of soil parameters with consequent toxic effects on plants, in the nearby water supplies, and, ultimately, in the whole food chain (Arora et al., 2008; Kumar et al., 2019; Manzoor et al., 2018). Typically, elements, such as copper (Cu), nickel (Ni), zinc (Zn), and chromium (Cr) are biologically essential for plant growth but become toxic for animals and plants when their concentrations exceed certain threshold levels (Edelstein and Ben-Hur, 2018; Rizvi et al., 2020; Tiwari and Lata, 2018). Other heavy metals

often found in contaminated soils, such as cadmium (Cd) and lead (Pb) are not essential for plants growth, and many studies associated their presence with neurological and endocrinological toxicity for humans along with carcinogenic effects (Ali and Khan, 2019; Pratush et al., 2018; Rehman et al., 2018).

Since heavy metals are not biodegradable, they tend to accumulate in the environment becoming a high risk for biota over several years after their introduction in an ecosystem (Olsson et al., 1998; Tchounwou et al., 2012; Zwolak et al., 2019). The search for new solutions that can remediate heavy metals-contaminated soil is a critical prerequisite for the sustainable development of agriculture (Edelstein and Ben-Hur, 2018; Vardhan et al., 2019; Wuana and Okieimen, 2011), thus representing a topic of paramount importance. The most consolidated strategies to remediate heavy-metals contaminated soils include physical and chemical approaches, like isolation, through capping and subsurface

* Corresponding authors at: Department of Civil, Environmental, Land, Building Engineering and Chemistry (DICATECh), Polytechnic University of Bari, Via Orabona, 4, Bari 70125, Italy.

E-mail addresses: biagia.musio@poliba.it (B. Musio), vito.gallo@poliba.it (V. Gallo).

<https://doi.org/10.1016/j.envadv.2021.100144>

Received 4 October 2021; Received in revised form 17 November 2021; Accepted 19 November 2021

Available online 24 November 2021

2666-7657/© 2021 The Author(s).

Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license

<http://www.elsevier.com/locate/jenvadv>

barriers; immobilization, by solidification/stabilization, vitrification, and chemical treatment; physical separation; extraction, by soil washing, pyrometallurgical extraction, *in situ* soil flushing, and electrokinetic treatment (Dhalwal et al., 2020; Gong et al., 2018; Guslatin et al., 2020; Qin et al., 2020). Alternative approaches are gaining great attention as they combine cost-effectiveness, sustainability, low toxicity, and mobility decrease. They include bioaccumulation, phytoremediation (phytoextraction, phytostabilization, and rhizofiltration), bioleaching, and biochemical processes, in which living organisms such as plants or microbes are used to clean an area from contaminants.

In particular, phytoremediation is attracting the attention of the scientific community, since it has been demonstrated to be a cost-effective solution for the remediation of contaminated sites, and, in the meanwhile, to be a feasible method for bio-fixation of CO₂, resulting in highly sustainable technology (Awa and Hadibarata, 2020). The ability to absorb heavy metals generally depends on the biomass produced as well as on the ability of the plant to accumulate and translocate heavy metals in its biomass (Eid and Shaltout, 2016; Hernández-Allica et al., 2008; Pachura et al., 2016). According to recent scientific literature, a good candidate for phytoremediation of heavy metals-contaminated soil is the hemp plant (Ahmad et al., 2016; Morin-Crini et al., 2019; Zielonka et al., 2020). Kompoli, also known as hemp, the non-psychoactive variety of *Cannabis sativa* L., is an annual dioecious high yielding industrial crop, and it is mainly grown for its fibers and seeds and used for textiles, clothing, insulation, biodegradable plastics, food, animal feed, and biofuel production (Adesina et al., 2020; Crini et al., 2020; Schlutenhofer and Yuan, 2017; Vasantha Rupasinghe et al., 2020). Hemp possesses some characteristics that make it very suitable for phytoremediation, such as high biomass, long roots, and an industrial-favorable short life cycle of 180 days. Importantly, hemp demonstrated a strong capability to sequester heavy metals like cadmium, zinc, lead, nickel, copper, and chromium when are present in contaminated soil and water (Citterio et al., 2003; Galic et al., 2019; Piotrowska-Cyplik and Czarnecki, 2003; Zielonka et al., 2020).

Another attractive approach for the remediation of contaminated sites is the application of bioleaching technology which uses direct metabolism or by-products of microbial processes to uptake heavy metals adsorbed onto the soil surface and to transform them so that the elements can be extracted when water is filtered through. Bioleaching has several advantages over conventional physical and chemical strategies, such as low cost, environmental sustainability, low hazardous characteristics of waste/sludge, low energy demand, and absence of toxic chemicals (Bosecker, 1997; Drobiková et al., 2015; Mishra et al., 2005; Okoh et al., 2018; Rawlings, 2002; Sun et al., 2021). Besides, phytoremediation, which involves eukaryotic algae and cyanobacteria in remediation processes, has been extensively applied to the treatment of wastewater (Awa and Hadibarata, 2020). Nevertheless, its application to the remediation of heavy metals contaminated sediments is less documented. Among the cyanobacteria, *Arthrospira platensis* possesses excellent chelating properties both towards heavy metals present in humans and towards those present in the soil, water, and sludge (Balaji et al., 2014; Bhattacharya, 2020; König-péter et al., 2015; Nalimova et al., 2005; Zinicovscaia et al., 2019, 2016). The dried biomass of *Arthrospira platensis* is commonly known as spirulina and it finds many applications in agriculture as a plant growth promoter, enhancing the growth, increasing the yield, and speeding up the seed germination (Tripathi et al., 2008; Wuang et al., 2016). Recently, the employment of this blue-green alga to uptake heavy metals in contaminated sites has been explored (Cepoi et al., 2020; Wuang et al., 2016). The presence of a chloroplast-type ferredoxin in the active center has been reported as responsible for the chelating capability of Spirulina (Tsukihara et al., 1978), whereby its efficiency is affected by many physical and chemical factors such as initial metal concentration, dosage, adsorption time, temperature, and pH (Šeker et al., 2008).

The present study aimed at both exploring the ability of the unreported combined use of hemp and spirulina to uptake six selected heavy

metals (Cd, Ni, Cr, Pb, Cu, Zn) from artificially contaminated soil and investigating, under controlled plant growing conditions, their distribution into the plant tissues. Specifically, hemp was chosen as the main agent for biological remediation, and spirulina was added as an enhancer of both the plant growth and the translocation of heavy metals in the hemp. The application of a non-targeted Nuclear Magnetic Resonance (NMR) approach combined with an estimation of the residual metals by Inductively Coupled Plasma Atomic Emission Spectroscopy (ICP-AES) into the cultivation soil and within the different tissues of the plant was applied in view of gathering useful information on the efficiency of the integrated hemp/spirulina system. Obtaining this information is crucial for the potential re-utilization of the hemp plant or shoots of it, after the phytoremediation stage, for alternative usages, like production of bio-materials for textile industry, constructions, and bio-fuel.

2. Materials and methods

2.1. Materials

3-(Trimethylsilyl)-2,2,3,3-tetrahydropropionic acid sodium salt (TSP, CAS N. 24493-21-8, 99 %D, Armar Chemicals, Döttingen, Switzerland), hydrochloric acid (HCl, 37%, CAS N. 7647-01-0; >99.5%, Sigma-Aldrich, Milan, Italy), sodium oxalate (Na₂C₂O₄, CAS N. 62-76-0; >99.5%, Sigma-Aldrich, Milan, Italy) sodium azide (NaN₃, CAS N. 26628-22-8; >99.5%, Sigma-Aldrich, Milan, Italy) and deuterium oxide (D₂O, CAS N. 7789-20-0, 99.86 %D, Eurisotop, Saclay, France) were used for sample preparation. NMR tubes (Norell 509-UP 7) were provided by Norell, Landisville NJ, United States. The NMR samples were prepared using an automated system for liquid handling (SamplePro Tube, Bruker BioSpin).

The soil used during cultivation was Plagron Lightmix (Plagron, Ospel, The Netherlands), which had pH 6-7, electric conductivity (E.C. 1:5) 310-470 µS·cm⁻¹ and (E.C. 1:1.5) 0.7-1.1 µS·cm⁻¹, NPK (12-14-24) 1.5 kg·m⁻³, total N 180 g·m⁻³ (10S:75, NO₃:NH₄, P₂O₅: 210 g·m⁻³, K₂O 360 g·m⁻³, dry matter 37% (of which 82% organic matter), and water retention of 6 mL·g⁻¹ dry matter.

The soil was contaminated by Cd(NO₃)₂·4H₂O (CAS N. 10022-68-1, Carlo Erba Reagents, Milan, Italy), K₂Cr₂O₇ (CAS N. 7778-50-9, Carlo Erba Reagents, Milan, Italy), Cu(SO₄)₂·5H₂O (CAS N. 7758-99-8, ProLabo, Paris, France), Pb(CH₃COO)₂ (CAS N. 301-04-2, Carlo Erba Reagents, Milan, Italy), Ni(NO₃)₂·6H₂O (CAS N. 13478-00-7, Carlo Erba Reagents, Milan, Italy), Zn(CH₃COO)₂·2H₂O (CAS N. 5970-45-6, Carlo Erba Reagents, Milan, Italy).

Hydrochloric acid (HCl 37%, Merck, Darmstadt, Germany), nitric acid (HNO₃ 65%, Merck, Darmstadt, Germany), double-distilled water (DDW > 99.5%), standard multi-elemental reference solution (Ag, Al, Ba, Be, Bi, Ca, Cd, Co, Cr, Cu, Fe, K, Mg, Mn, Mo, Na, Ni, Pb, Sr, Ti, V, Zn [10 mg/l]) were used for heavy metals analysis by Inductively Coupled Plasma - Atomic Emission Spectrometry (ICP-AES).

2.2. Algal strain and growth conditions

Arthrospira platensis: strain was cultivated by ApuliaKundi S.r.l. in open ponds (3000 L per pond) under controlled greenhouse conditions with natural light at a temperature ranging from 22 to 28 °C.

The production cycle was monitored by turbidity measurements through a Secchi disk (200 mm diameter, Scubla, Remanzacco (UD), Italy). Once the disk was lowered into the algae suspension at 5-6 cm and could not be seen anymore, the algae were collected. Overall, the production cycle lasted approximately 3 days. The collected algal biomass was filtered (40 µm filter), extruded, and cold dried at a temperature value lower than 38 °C. The study was conducted using the dried biomass in powder form, commonly named spirulina.

of Cd, Cr, Cu, Ni, Pb, and Zn. The 6 elements indicated, taken from 1000 $\text{mg}\cdot\text{kg}^{-1}$ solutions, were brought in a single solution from $100 \text{ mg}\cdot\text{kg}^{-1}$. Dilutions of 1, 0.5, 0.1, 0.01, 0.005, and $0 \text{ mg}\cdot\text{kg}^{-1}$ of the concentrated standard reference solutions were made in final volumes of 100 mL using HNO_3 and HCl.

A calibration blank was used to obtain the analytical curve and was prepared by acidifying water with a mixture of 65% HNO_3 and 37% HCl in a way that the standard and the sample reached the same value of acidity. Furthermore, the method blank contained all reagents in the same volumes used in the sample preparation and was used to identify possible contamination resulting from the acidic reagents used or equipment used during the sample preparation, including the filtration step. The calibration standard analytical curve was made of 1.0, 0.50, 0.10, 0.010, 0.0050, and $0 \text{ mg}\cdot\text{kg}^{-1}$.

2.6. Statistical elaboration of NMR and ICP-AES data

The 48 processed 1D ^1H NOESY spectra (8 leaf samples \times 3 replicates, 8 stem samples \times 3 replicates) were reduced to a numerical matrix (bucket-table) manageable for multivariate statistical analysis. The bucket-table was obtained by dividing the entire spectrum in the range of [−0.514, 10.486] ppm into 249 rectangular intervals (buckets) of 0.04 ppm in width, excluding the region [5.126, 4.566] ppm corresponding to the residual water signal. The underlying area of each bucket was normalized to the total intensity. The bucket table was imported into SIMCA 13.0.3 software (Umetrics, Umea, Sweden) to perform multivariate statistical analyses (MVA). The coordinates of the observations in the new space are called scores, the weight of the original variables on each PC are called loadings. In the present study, the NMR spectra constituted the observations, the buckets constituted the x-variables.

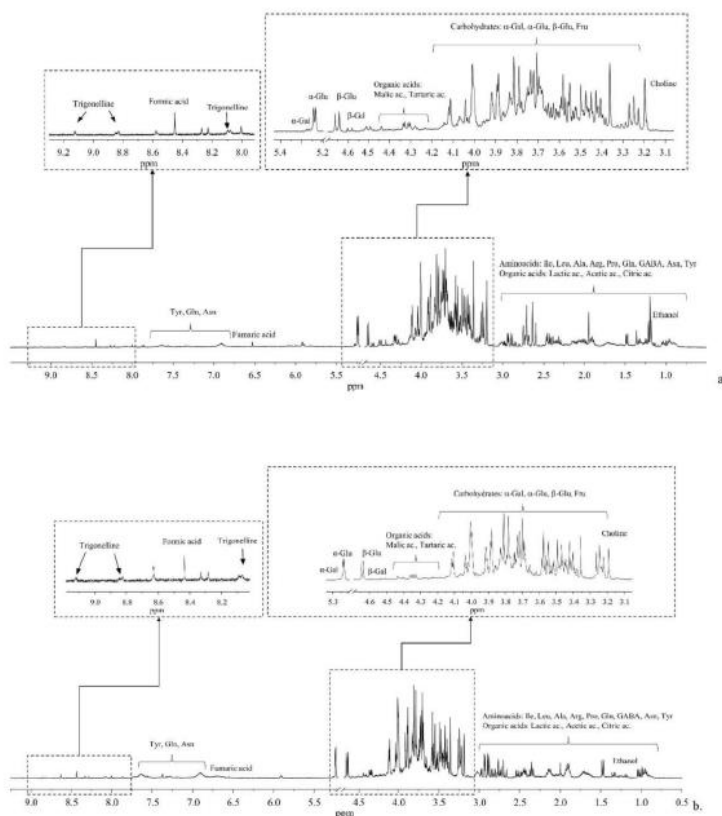


Fig. 1. Typical ^1H NOESY spectra (Bruker Avance 400 MHz, D_2O) of aqueous extracts of leaves (a) and stems (b) of *Cannabis sativa* L. Spectral region containing the residual water signal (4.78 ppm) was hid. The spectral regions included in rectangular were expanded for clarity.

Buckets were centered and subjected to Pareto scaling (each x_j -variable was scaled to $1/\sqrt{\text{sd}_j}$), where sd_j is the standard deviation of x_j -variable computed around the mean) to avoid noise inflation. The buckets in the range [-0.514, 0.486] ppm, including the TSP singlet, were excluded from statistical analysis. Data from ICP-AES analysis and morphometric data were also included in statistical analysis as y -variables and were scaled to Unit Variance (UV), as recommended for variables expressed in different units (for each y -variable the base weight is computed as $1/\text{sd}_j$, where sd_j is the standard deviation of y_j -variable computed around the mean). Principal Component Analysis (PCA) and Orthogonal Partial Least Squares analysis (OPLS) were performed. The quality of statistical models was assessed based on the parameters R^2 and Q^2 , which represent the descriptiveness (goodness-of-fit) and the predictivity in cross-validation (goodness-of-prediction in cross-validation), respectively. For OPLS models, permutation tests were carried out to verify the absence of model overfitting.

3. Results and discussion

NMR analysis (1D ^1H NOESY experiments) allowed for the identification of a pool of metabolites contained in typical aqueous extracts (pH = 4.2) of leaves (Fig. 1a) and stems (Fig. 1b) collected from a plant of hemp cultivated under controlled conditions in uncontaminated soil (see ESI for the full list of metabolites, Table S3). The main representative classes of metabolites included organic acids (lactic, citric, tartaric, fumaric, formic, and acetic acids), amino acids (alanine, asparagine, γ -aminobutyric acid, arginine, glutamine, isoleucine, proline, tyrosine, leucine), and carbohydrates (β -glucose, α -glucose, fructose, α -galactose, β -galactose). Moreover, 1D ^1H NOESY spectra contained also signals related to trigonelline, choline, and ethanol.

A holistic approach was applied during the present study by combining the information provided by the 1D ^1H NOESY experiments, the ICP-AES quantitative analysis of residual metal, and the data derived from the morphometric inspection.

As a first task, the relation between the variations in the metabolic composition and the calculated amounts of residual heavy metals in the plant was considered. Thus, the OPLS analysis was performed considering the spectral buckets as x -variables, and the metal content and plant biomass as y -variables (see ESI for OPLS model parameters, Table S4). Such a model was characterized by 5 predictive components P1-P5 that explained 59.5% of x -variance ($R^2X[\text{cum}] = 0.595$), 92.2% of y -variance ($R^2Y[\text{cum}] = 0.922$) and 87.6% of y -variance modelled by x -variables ($R^2 = 0.876$), and 8 orthogonal components O1-O8 that explained about 38% of the x -variance unrelated to y -variance ($R^2X[\text{cum}] = 0.383$). Two main separate clusters of observations were noticeable along P1. They

corresponded to the two parts of the plant, namely the leaves (LE) and the stems (ST) (Fig. 2a, LE vs ST). Inside each group, two evident subgroups were distinguishable along P2, which were identified as the vigorous plants (V) and the weak ones (W) (Fig. 2b, V vs W). The components P1 and P2 explained 41 and 9.8% of the information in x -space ($R^2X[P1] = 0.41$ and $R^2X[P2] = 0.098$), suggesting an important variability of the metabolic composition along with the different tissues of the plant and according to the wellness of the plant. Indeed, the same components P1 and P2 explained 48.5 and 15.2% of the information in y -space ($R^2Y[P1] = 0.485$ and $R^2Y[P2] = 0.152$), suggesting a differential distribution of the y -variables, namely the metals concentration and the plant biomass, according to the different tissue of the plant and the health status of the plant (Fig. 2c). The visual inspection of the original harvested plants (see ESI for further details, Fig. S1) combined with the plant biomass values suggested that the samples included in the subgroup with $t[2] < 0$ (Fig. 2b, red scores) corresponded to the less healthy plants (weak, W), whereas the observations included in the subgroup with $t[2] > 0$ corresponded to the more vigorous plants (V, Fig. 2b, yellow scores). The analysis of the loading plot (pq[1] vs pq[2]) allowed the identification of the most significant variables towards the distribution of the scores, including the metal concentration, biomass amount, and spectral regions (Fig. 2c).

The amounts of Cu, Cr, Zn, and Ni added during the present study were well tolerated by the plants since the cluster containing the vigorous plants was characterized by having higher levels of them. On the other side, a lower tolerance was observed towards Pb as the highest levels of this metal were found in the weakest plants (Fig. 2c). It is well documented that the tolerance of hemp towards heavy metals is highly selective (Zielonka et al., 2020), and it is related both to the ability of the plant to produce high levels of antioxidant agents and to accumulate metals in roots, hindering them to translocate to the upper parts of the plant (McPartland and McKernan, 2017; Petrová et al., 2012). A selective accumulation of the studied metals in the leaves was observed according to the following trend: $\text{Cu} > \text{Zn} > \text{Cr} > \text{Ni} > \text{Cd} \gg \text{Pb}$. Specifically, the observed accumulation of copper inside the leaves compared to the stems (Fig. 2c, pq[1] < -0.2), and particularly, in leaves collected from vigorous green plants (Fig. 2c, pq[2] > 0.1), can be justified by the fact that this metal is actively involved in cellular functions happening mostly in this part of the plant, such as the transport of electrons in mitochondria and chloroplasts, the control of the cellular redox state, and the remodeling of the cell wall (Printz et al., 2016).

The accumulation of zinc inside the leaves (Fig. 2c, pq[1] < -0.2) can be explained with both the increased production of organic acid and the vacuolar compartmentalization occurring in the leaves (Bradley et al.,

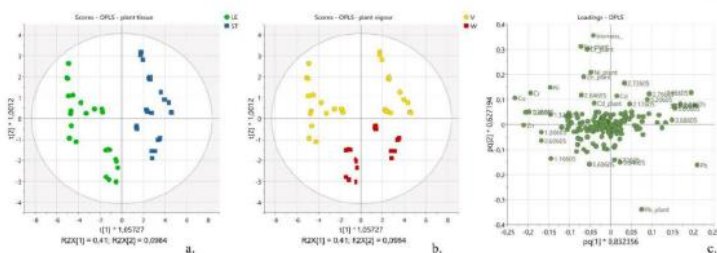


Fig. 2. OPLS-DA applied to spectral data (x -variables) and metal content and plant biomass data (y -variables). (a) Scores plot displaying $t[1]$ vs $t[2]$, the observations are colored according to the tissues of the plants: leaves (green circle, LE), stems (blue square, ST). (b) Scores plot $t[1]$ vs $t[2]$, the observations are colored according to the health status of the plants: vigorous (yellow circle, V), weak (red square, W). (c) Loading plot displaying the relationship between the x -variables (spectral regions) and the y -variables (metal content and plant biomass) for the first and second predictive components pq[1] vs pq[2] [For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article].

2007).

While cadmium did not show any preferential site accumulation (Fig. 2c, $pq[1] \approx 0$), lead presented a noticeable preference to distribute within the stems (Fig. 2c, $pq[1] > 0.2$). Interestingly, the tolerance of hemp towards high levels of lead was lower compared to the other selected heavy metals. Such evidence is not surprising considering the high toxicity of lead on living organisms (ATSDR, 2019). Indeed, lead impairs the physicochemical properties of soil and soil microbial community; it alters the water and nutrient uptake, generates reactive oxygen species (ROS), and reduces photosynthetic activity, hindering plant growth and development (Broadley et al., 2007). As a result of the present investigation, the residual amounts of lead were found higher in the less healthy plants (Fig. 2c, $pq[1] < -0.1$). Also, this evidence is confirmed by the inverse correlation observed in the loading plot

between the plant biomass (Fig. 2c, $pq[2] > 0.3$) and the level of lead in the plant (Fig. 2c, $pq[2] < -0.3$), suggesting a potential phytotoxic effect of lead.

The analysis of the correlation matrix (see ESI for further details, Table S5) related to the OPLS model helped to better understand the ability of hemp to uptake simultaneously and selectively the six metals under investigation. A positive correlation was found for the plant content of Cu with Zn ($r = 0.80$) and Ni ($r = 0.82$). Also, Zn well correlated with the plant content of Cd ($r = 0.84$) and Ni ($r = 0.81$). Conversely, a negative correlation was found between Cu plant content and Pb residual amount ($r = -0.82$). Such evidence may suggest a competition between lead and copper during the uptake stage. Importantly, it was found that the plant weight (biomass) positively correlated with the residual content of Cu ($r = 0.72$) and Cr ($r = 0.64$) in the plant.

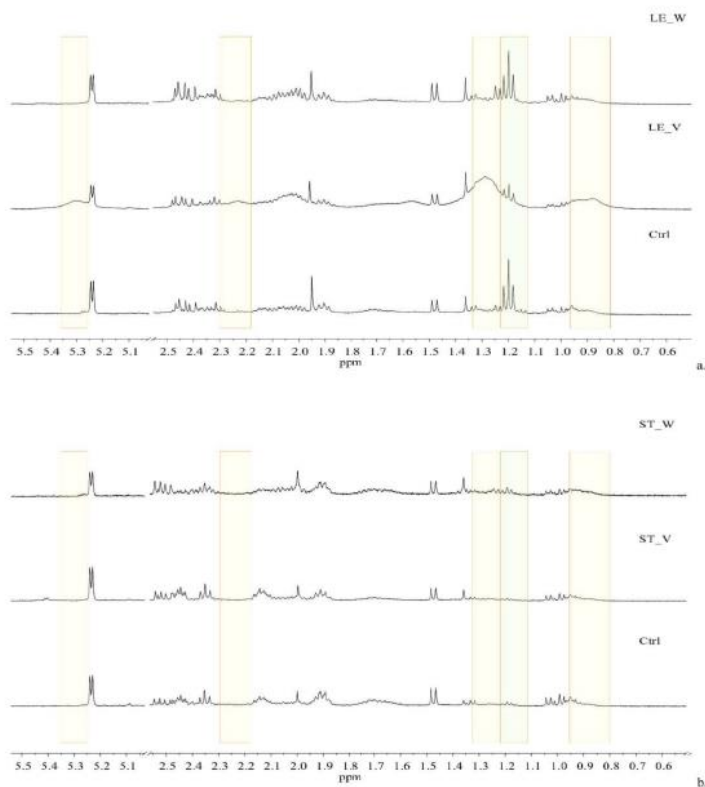


Fig. 3. ¹H NMR spectral regions [0.6–2.55 ppm; 5.03–5.54 ppm] of aqueous extracts of leaves (a) and stems (b) collected from the control plant (Ctrl), vigorous plant (LE_V; ST_V), and weak plant (LE_W; ST_W), respectively. The yellow rectangles delimit the broad signals (0.88, 1.29, 2.23, 5.31 ppm) observed predominantly in the aqueous extracts of leaves from vigorous plants (Fig. a, Ctrl vs LE_V vs LE_W) and not in the aqueous extracts of stems (Fig. b, Ctrl vs ST_V vs ST_W); the green rectangle highlights the ethanol signal (1.20 ppm) predominantly visible in the leaves (a) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Still, a negative correlation was found between the plant biomass and the Pb content ($r = -0.76$), supporting the hypothesis that hemp tolerance towards this metal is relatively low compared to the other metals investigated in the present work.

3.1. Metabolic analysis of leaves and stems

The variations in the metabolic composition of the aqueous extracts of leaves and stems were investigated by performing an OPLS-DA. A clear separation between the two classes of samples (leaves vs stems) along the first two components ($(R^2X[1] = 0.394, R^2Yo[1] = 0.27$; see ESI for the scores plot, Fig. S2a) was obtained and the analysis of the S-plot allowed for the detection of the most significant x -variables and, thus, the metabolites contributing to such clustering (see ESI for the S-plot, Fig. S2b). Specifically, the buckets having values of $VIP_{predictive} > 1$ and values of $p(\text{corr})1 > |0.5|$ were analyzed more deeply (see ESI for further details, Table S6).

The spectral regions containing signals related to glucose, fructose, and galactose moieties contributed significantly to the clustering of the stems. Also, the buckets containing the signals assigned to the organic acids were characterized by significant values of VIP and $p(\text{corr})1$. An analogous trend was observed for the buckets related to the signals assigned to γ -aminobutyric acid (GABA) at 1.89–1.93 ppm and 3.01–3.05 ppm, and glutamine at 2.13–2.17 ppm.

Interestingly, the clustering of the leaves was affected by some regions of the spectra containing very broad signals at 0.88, 1.29, 2.23, and 5.31 ppm (Fig. 3a, yellow rectangles), presumably attributable to the lipids of the cell membranes and/or walls (Cruciani et al., 2006). The comparison of the 1D 1H NOESY spectra of the leaves collected from control plants (Ctrl with no metal contamination, Fig. 3a), vigorous plants (LE_V, Fig. 3a), weak plants (LE_W, Fig. 3a) indicated that such broad signals were predominant in vigorous plants. Such broad signals were not observed in the aqueous extracts of stems (Fig. 3b). Also, the buckets containing the signals of ethanol at 1.17–1.25 ppm contributed relevantly to the grouping of the leaves towards the stems (Fig. 3a and b, green rectangular). These data could be explained by the fact that the leaves and, particularly the leaves sampled from vigorous green plants, have active photosynthetic processes and intact cell structures, whereas stems of weakest plants begin to enrich with lignin and have less active photosynthesis (Brazel and O'Maolfeidigh, 2019; Pfanz et al., 2002).

The spectral regions containing these broad signals correlated positively with the calculated amounts of residual copper in the leaves. Under physiological conditions, Cu is found in two common forms, Cu(I) (low oxidation state) preferentially binding sulfur-containing compounds having a thiol or a thioether group, and Cu(II) (high oxidation state), that coordinates mainly with oxygen or imidazole nitrogen groups. The main functions of Cu are the transport of electrons in mitochondria and chloroplasts, the control of the cellular redox state (being a major Cu-binding protein the Cu/Zn superoxide dismutase), and the remodeling of the cell wall, which is one of the major Cu-accumulation sites in hyperaccumulating plants (Printz et al., 2016). Moreover, Cu leaf content is inversely associated with the intensity of the buckets relating to glucose and fructose (in the range [3.26–4.08] ppm) and ethanol (centered at 1.18 and 3.62 ppm). In weakest plants, sucrose, the main soluble component of the phloem sap and translatable product of photosynthesis might be hydrolyzed into the two constituting monomers (glucose and fructose) and fermented by yeast and/or bacteria producing ethanol (Winter and Huber, 2000).

3.2. Effect of treatment with spirulina

In a second task, the effect of the treatment with spirulina during the cultivation of hemp towards the remediation of contaminated soil was investigated. Thus, a deeper look was taken at the higher components of the OPLS model described in paragraph 3.1. Indeed, a good clustering of samples according to the treatment was visible in the $t[2]$ vs $t[5]$ scores

plot (Fig. 4a). The second component and the fifth one explained together 11.75% of the x -variance predictive of the y -variance ($R^2XP[2] = 0.098$ and $R^2XP[5] = 0.019$), suggesting that the treatment with spirulina affected slightly the metabolic composition of the plant. Such evidence is relevant to make possible a re-utilization of the hemp, which does not change its metabolome during the phytoremediation stage.

Specifically, the samples collected from plants grown in contaminated soil (Fig. 4a, green circles, C) and the samples obtained from plants treated with a higher concentration of spirulina (Fig. 4a, red triangles, CS1) tended to cluster at values of $t[2] > 0$. Conversely, the samples derived from plants treated with a lower concentration of the blue-green alga (Fig. 4a, blue squares, CS0.5) and the control ones (Fig. 4a, yellow triangles, Ctrl) were distributed preferentially at values of $t[2] < 0$. A further piece of evidence was that samples C and the samples CS1 tended to separate into two clear groups along $t[5]$. Analysis of the loading plot (Fig. 4b) suggested that the uptake of all the metals, except for lead, contributed relevantly to the clustering of the plants treated with a higher dose of spirulina (CS1). Also, the average plant biomass ($pq[2] > 0.35$) was generally higher in contaminated plants compared to the uncontaminated ones. In addition, the plants that grew in contaminated soil and were irrigated with a higher dose of spirulina (CS1) presented a higher value of biomass, confirming the activity of these cyanobacteria to promote plant growth. The average plant contents of Ni ($pq[2] = 0.21$), Zn ($pq[2] = 0.19$), and Cd ($pq[2] = 0.09$) were greater in contaminated plants grown in soils added with spirulina than those grown without it.

To get deeper insights into the effects deriving from the dose of spirulina, OPLS was applied to CS1 and CS0.5 samples.

The observations, namely CS1 (Fig. 5a, red triangles) and CS0.5 (Fig. 5a, blue squares), separated along the second predictive component P2, explaining about 9.5% of the predictive x -variance ($R^2XP[2] = 0.095$). The distribution was strongly affected by the part of the plant the samples derived from, leaves or stems, with a noticeable clustering along the first predictive component P1 ($R^2XP[1] = 0.534$). The corresponding loading plot (Fig. 5b) showed important dose-dependent variations. Indeed, nickel, zinc, copper, and chromium were more abundant in samples treated with a higher dose of spirulina (CS1).

Generally, as the concentration of spirulina increased, the residual metal content increased both in the leaves and in the stems (for Cu and Zn, the increase was more noticeable in stems than in leaves), whereas the Cr content increased in leaves and decreased in stems. While no dose-effect was observed for Cd, noticeable variations were exerted by increased amounts of spirulina on the quantities of Pb incorporated into the plant. Specifically, the average content of Pb in the plant was higher in CS0.5 than in CS1, suggesting that a higher amount of the blue-green alga contained in the soil might compete with hemp to sequester this metal.

A piece of further evidence was that the average soil residual of all heavy metals resulted higher when it was irrigated with a higher concentration of spirulina (CS0.5 vs CS1), suggesting that the cyanobacteria stationing in the soil may exert a strong action of metal chelation, avoiding leaching of them during the watering stage and acting as a reservoir for hemp to absorb the chelated metals. Such results find fundamentals in the reported ability of *Arthrospira platensis* to adsorb (in form of dry biomass) and accumulate (in form of alive biomass) all the six metals here tested through various mechanisms (ion exchange, binding to functional groups, complexation, and microprecipitation) (Zincovacia et al., 2019).

Very small effects were observed on the metabolic profile when samples CS1 were compared with CS0.5. Interestingly, according to the loading plot (Fig. 5b), the buckets at 2.66 and 2.74 ppm, containing the signals assigned to citric acid, contributed importantly to the observed grouping between the samples treated with the two different doses of spirulina (see ESI for the variable trend, Fig. S3). Specifically, the signals of citric acid were more intense in samples treated with a higher dose of spirulina. Considering that citric acid is one of the main chelating agents

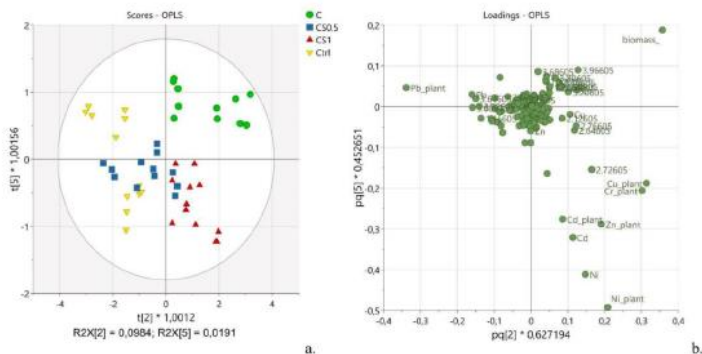


Fig. 4. (a) Scores plot displaying $t[2]$ vs $t[5]$ obtained from OPLS applied to spectral data (x-variables) and metal content and plant biomass data (y-variables). The observations are colored according to the treatment with spirulina: control plant (yellow triangles, Ctrl), plants grown in contaminated soil (green circles, C), plants treated with spirulina at a concentration of $1.0 \text{ g}\cdot\text{L}^{-1}$ (red triangles, CS1), plants treated with spirulina at a concentration of $0.50 \text{ g}\cdot\text{L}^{-1}$ (blue squares, CS0.5). (b) Loading plot displaying the second and fifth predictive components $pq[2]$ vs $pq[5]$ (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

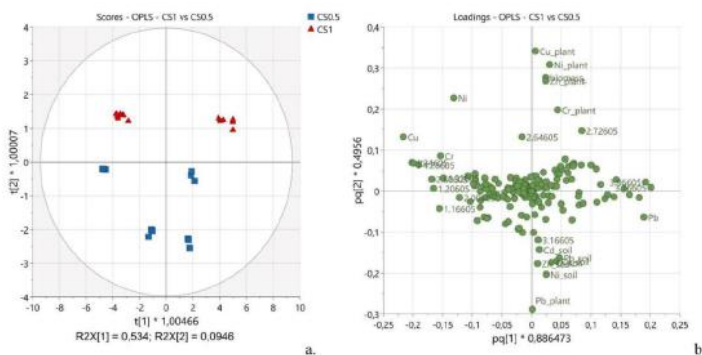


Fig. 5. OPLS was applied to spectral data of samples treated with the two different doses of spirulina (x-variables) and metal content and plant biomass data (y-variables). (a) Scores plot displaying $t[1]$ vs $t[2]$. The observations are colored according to the different doses of spirulina: plants treated with spirulina at a concentration of $1.0 \text{ g}\cdot\text{L}^{-1}$ (red triangles, CS1); plants treated with spirulina at a concentration of $0.50 \text{ g}\cdot\text{L}^{-1}$ (blue squares, CS0.5). (b) Loading plot displaying the first and the second predictive components $pq[1]$ vs $pq[2]$ (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

in plants, this evidence may support the hypothesis that, at a higher concentration, the cyanobacteria can chelate the heavy metals sequestering them into the soil. As a result, less quantity of heavy metal is available for the plant to uptake, and, thus, more chelating agents inside the plant are in the free form not coordinating any metal (Arru et al., 2004).

4. Conclusions

It was demonstrated that hemp accumulates copper, chromium, nickel, and zinc preferentially in the leaves, while lead is distributed

mainly in the stems of the plant. Such selective compartmentalization is enhanced when the plant is irrigated with water containing spirulina. It was found that, at higher concentrations, spirulina acts as a growth promoter, contributing to an increase in the final generated biomass. Also, it was demonstrated that the treatment with spirulina during the cultivation of the hemp induced an enhancement in the uptake of heavy metals, except for lead. Such a result may be explained by assuming a strong affinity of spirulina towards the lead, resulting in confinement of this metal inside the soil and hampering its uptake by the plant. The NMR analysis allowed to identify the crucial variations of the metabolic composition that were induced by the treatment with spirulina. Results

reported in this work pave the way to further studies aimed to understand the optimal dose of spirulina to enhance the efficiency of this innovative combined remediation bio-system. Furthermore, the results described may encourage the application of spectroscopic methods for the rapid detection of structural changes in the various environmental spheres, allowing prompt intervention through the adoption of remediation schemes.

Acknowledgments

This study was funded by Regione Puglia "Sezione Competitività delle filiere agroalimentari" under the call "Avviso pubblico per la presentazione di progetti di ricerca ed innovazione e interventi a carattere pilota con D.D. 189 del 17/10/2018 - Project "BIO.SP.HE.RE - Bio-integrated Spirulina and HEmp Remediation".

The authors acknowledge Dr. Vito Bruno, Dr. Francesca Portincasa, and Dr. Stefano Todisco for the helpful discussions and technical support.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.118758.

References

Adesina, I., Bhowmik, A., Sharma, H., Shahbazi, A., 2020. A review on the current state of knowledge of growing conditions, agronomic soil health practices and utilities of hemp in the United States. *Agriculture* 10, 129. <https://doi.org/10.3390/agriculture10040129>.

Ahmed, B., Telsin, Z., Malik, S.T., Asad, S.A., Shahzad, M., Bilal, M., Shah, M.M., Khan, S.A., 2016. Phytoremediation potential of hemp (Cannabis sativa L.): identification and characterization of heavy metals responsive genes. *Clean Soil Air Water* 44, 195–201. <https://doi.org/10.1002/clm.2011500117>.

Ali, H., Khan, E., 2019. Trophic transfer, bioaccumulation, and biomagnification of non-essential hazardous heavy metals and metalloids in food chains/web-s—concepts and implications for wildlife and human health. *Hum. Ecol. Risk Assess.* <https://doi.org/10.1080/10807039.2018.1466938>.

Arora, M., Kiran, B., Rani, S., Rani, A., Kaur, B., Mittal, N., 2008. Heavy metal accumulation in vegetables irrigated with water from different sources. *Food Chem.* 111, 811–815. <https://doi.org/10.1016/j.foodchem.2008.04.048>.

Arru, L., Roggioni, S., Barocchini, M., Bonatti, P.M., Perata, P., 2004. Copper localization in Cannabis sativa L. grown in a copper-rich solution. *Ecophytica* 140, 33–38. <https://doi.org/10.1007/s10681-004-4752-0>.

ATSDR, 2019. The ATSDR 2019 substance priority list. Agency Toxic Subst. Dis. Regist. <https://www.atsdr.cdc.gov/ptld/index.html>.

Awa, S.H., Hadjilovata, T., 2020. Removal of heavy metals in contaminated soil by phytoremediation mechanism: a review. *Water Air Soil Pollut.* 231, 1–15. <https://doi.org/10.1007/s11270-020-4426-0>.

Balaji, S., Kalavani, T., Rajasekaran, C., 2014. Biosorption of zinc and nickel and its effect on growth of different spirulina strains. *Clean Soil Air Water* 42, 507–512. <https://doi.org/10.1002/clm.201200340>.

Bhattacharya, S., 2020. The role of spirulina (Arthrospira) in the mitigation of heavy-metal toxicity: An appraisal. *J. Environ. Pathol. Toxicol. Oncol.* 39, 149–157. <https://doi.org/10.1615/JEnvironPatholToxicolOncol.2020034375>.

Bosecker, K., 1997. Bioleaching: metal solubilization by microorganisms. *FEMS Microbiol. Rev.* 20, 591–604. <https://doi.org/10.1111/j.1574-6976.1997.tb00340.x>.

Brazd, A.J., O'Maoléigh, D.S., 2019. Photosynthetic activity of reproductive organs. *J. Exp. Bot.* 70, 1737–1753. <https://doi.org/10.1093/jxb/erz033>.

Broadley, M.R., White, P.J., Hammond, J.P., Zelik, L., Lux, A., 2007. Zinc in plants. *New Phytol.* 173, 677–702. <https://doi.org/10.1111/j.1469-8137.2007.01996.x>.

Cepoi, L., Zinocroscalia, I., Rudi, L., Chiriac, T., Mucu, V., Djur, S., Strelkova, L., Veggel, K., Nekhodovskov, P., 2020. Growth and heavy metals accumulation by spirulina platensis biomass from multicomponent copper containing synthetic effluents during repeated cultivation cycles. *Ecol. Eng.* 142, 105637 <https://doi.org/10.1016/j.ecoeng.2019.105637>.

Citterio, S., Santagostino, A., Fumagalli, P., Prato, N., Ranalli, P., Sgorbati, S., 2003. Heavy metal tolerance and accumulation of Cd, Cr and Ni by Cannabis sativa L. *Plant Soil* 256, 243–252. <https://doi.org/10.1023/A:1026113905126>.

Cini, G., Lichtfouse, E., Charet, G., Morin-Crini, N., 2020. Applications of hemp in textiles, paper industry, insulation and building materials, horticulture, animal nutrition, food and beverages, nutraceuticals, cosmetics and hygiene, medicine, agrochemistry, energy production and environment: a review. *Environ. Chem. Lett.* 18, 1451–1476. <https://doi.org/10.1007/s10311-020-01629-2>.

Cruciani, O., Manzi, L., Sobolev, A.P., Casent, C., Segre, A.L., 2006. An improved NMR study of liposomes using 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylcholine as model. *Molecules* 11, 394–344. <https://doi.org/10.3390/11050344>.

De Backer, B., Maebe, K., Verstraete, A.G., Charlier, C., 2012. Evolution of the content of THC and other major cannabinoids in drug-type cannabis cuttings and seedlings during growth of plants. *J. Forensic Sci.* 57, 918–922. <https://doi.org/10.1111/j.1556-4029.2012.02068.x>.

Dhalwadi, S.S., Singh, J., Tanaji, P.K., Mandal, A., 2020. Remediation techniques for removal of heavy metals from the soil contaminated through different sources: a review. *Environ. Sci. Pollut. Res.* 27, 1319–1333. <https://doi.org/10.1007/s11356-019-06967-1>.

Drobňková, K., Rozumová, L., Otupalková, H., Seidlerová, J., 2015. Bioleaching of hazardous waste. *Chem. Pap.* 69, 1193–1201. <https://doi.org/10.1515/chempap-2015-0121>.

Edelstein, M., Ben-Hur, M., 2018. Heavy metals and metalloids: Sources, risks and strategies to reduce their accumulation in horticultural crops. *Sci. Hortic.* 234, 431–444. <https://doi.org/10.1016/j.scienta.2017.12.039>.

Eid, E.M., Shalhout, K.H., 2016. Bioaccumulation and translocation of heavy metals by nine native plant species grown at a sewage sludge dump site. *Int. J. Phytoremed.* 18, 1075–1085. <https://doi.org/10.1080/15220514.2016.1183578>.

Galić, M., Perić, A., Zgorelec, Z., Kisić, I., 2019. Evaluation of heavy metals accumulation potential of hemp (Cannabis sativa L.). *J. Cent. Eur. Agric.* 20, 700–711. <https://doi.org/10.5513/JCEA01/20.2.2201>.

Gong, Y., Zhao, D., Wang, Q., 2018. An overview of field-scale studies on remediation of soil contaminated with heavy metals and metalloids: Technical progress over the last decade. *Water Res.* 147, 440–460. <https://doi.org/10.1016/j.watres.2018.10.024>.

Gustafin, Z.M., Kulikowska, D., Kik, B., 2020. New-generation washing agents in remediation of metal-polluted soils and methods for washing effluent treatment: a review. *Int. J. Environ. Res. Public Health* 17, 1–19. <https://doi.org/10.3390/ijerph17172200>.

Hernández-Allier, J., Becerril, J.M., Garbón, C., 2008. Assessment of the phytoextraction potential of high biomass crop plants. *Environ. Pollut.* 152, 32–40. <https://doi.org/10.1016/j.envpol.2007.06.002>.

Konig-péter, A., Kilar, F., Felinger, A., Pernyeszi, T., 2015. Biosorption characteristics of spirulina and chlorella cells for the accumulation of heavy metals. *J. Serb. Chem. Soc.* 80, 407–419. <https://doi.org/10.2298/JSC140321060P>.

Kumar, S., Prasad, S., Yadav, K.K., Shivastava, M., Gupta, N., Nigir, S., Bae, O.V., Kanyah, H., Khan, S.A., Yadav, S., Matav, L.C., 2019. Hazardous heavy metals contamination of vegetables and food chain: role of sustainable remediation approaches—a review. *Environ. Res.* 179, 108792. <https://doi.org/10.1016/j.envres.2019.108792>.

Manzoor, J., Sharma, M., Wani, K.A., 2018. Heavy metals in vegetables and their impact on the nutrient quality of vegetables: a review. *J. Plant Nutr.* 41, 1744–1763. <https://doi.org/10.1080/01904002.2018.1462832>.

McPartland, J.M., McKernan, K.J., 2017. Contaminants of concern in cannabis: microbes, heavy metals and pesticides. In: Chandra, S., Lata, H., ElSohly, M.A. (Eds.), *Cannabis Sativa L. - Botany and Biotechnology*. Springer International Publishing, pp. 457–474. https://doi.org/10.1007/978-3-319-54584-6_22.

Mishra, D., Kim, D.J., Abu, J.G., Rhee, Y.H., 2005. Bioleaching: a microbial process of metal recovery; a review. *Met. Mater. Int.* 11, 249–256. <https://doi.org/10.1007/s10037-00302450>.

Morin-Crini, N., Loiacono, S., Placet, V., Tori, G., Bradu, C., Kostic, M., Cosentino, C., Chanet, G., Martel, B., Lichtfouse, E., Crini, G., 2019. Hemp-based adsorbents for sequestration of metals: a review. *Environ. Chem. Lett.* 17, 393–408. <https://doi.org/10.1007/s10311-018-0812-x>.

Nalimova, A.A., Popova, V.V., Troglin, L.N., Pronina, N.A., 2005. The effects of copper and zinc in spirulina platensis growth and heavy metal accumulation in its cells. *Russ. J. Plant Physiol.* 52, 229–234. <https://doi.org/10.1007/s11183-005-0035-4>.

Olokh, M.P., Olofayatan, I.W., Machunga, S.S., 2018. Bioleaching: a technology for metal extraction and remediation: mitigating health consequences for metal exposure. *International Journal of Development and Sustainability* 7, 2103–2118.

Olsson, P.E., Kling, P., Hogstrand, C., 1998. Mechanisms of heavy metal accumulation and toxicity in fish. *Metal Metabolism in Aquatic Environments*. Springer, US, pp. 321–350. https://doi.org/10.1007/978-1-4757-2761-6_10.

Pachura, P., Ociepa-Kubicka, A., Skowron-Grabowska, B., 2016. Assessment of the availability of heavy metals to plants based on the translocation index and the bioaccumulation factor. *Desalin. Water Treat.* 57, 1469–1477. <https://doi.org/10.1080/19443994.2015.1017330>.

Petrová, S., Benešová, D., Soudek, P., Vaněk, T., 2012. Enhancement of metal(loid)s phytoextraction by Cannabis sativa L. *J. Food Agric. Environ.* 10, 631–641.

Pfanz, H., Aschan, G., Lengenfeld-Heysler, R., Wittmann, C., Loose, M., 2002. Ecology and ecophysiology of tree stems: cuticular and wood photosynthesis. *Naturwissenschaften* 89, 147–162. <https://doi.org/10.1007/s00114-002-0209-z>.

Piotrowska-Cypk, A., Czarniecki, Z., 2003. Phytoextraction of heavy metals by hemp during anaerobic sewage sludge management in the non-industrial sites. *Pol. J. Environ. Stud.* 12, 779–784.

Pratush, A., Kumar, A., Hu, Z., 2018. Adverse effect of heavy metals (As, Pb, Hg, and Cr) on health and their bioremediation strategies: a review. *Int. Microbiol.* 21, 97–106. <https://doi.org/10.1007/s10123-019-012-3>.

Prinz, B., Letts, S., Haussman, J.F., Sergeant, K., 2016. Copper trafficking in plants and its implication on cell wall dynamics. *Front. Plant Sci.* 7, 1–16. <https://doi.org/10.3389/fpls.2016.00601>.

Qin, H., Hu, T., Zhai, Y., Lu, N., Aliyeva, J., 2020. The improved methods of heavy metals removal by biosorbents: a review. *Environ. Pollut.* 258, 113777. <https://doi.org/10.1016/j.envpol.2019.113777>.

Rawlings, D.E., 2002. Heavy metal mining using microbes. *Annu. Rev. Microbiol.* 56, 65–91. <https://doi.org/10.1146/annurev.micro.56.012302.161052>.

- Rehman, K., Fatima, F., Waheed, I., Akash, M.S.H., 2018. Prevalence of exposure of heavy metals and their impact on health consequences. *J. Cell. Biochem.* 119, 157–184. <https://doi.org/10.1002/jcb.26234>.
- Rivri, A., Zahid, A., Ameen, F., Ahmed, B., Alkahtani, M.D.F., Khan, M.S., 2020. Heavy metal induced stress on wheat: phytotoxicity and microbiological management. *RSC Adv.* 10, 38378–38403. <https://doi.org/10.1039/d0ra05610c>.
- Schlutenhofer, C., Yuan, L., 2017. Challenges towards revitalizing hemp: a multifaceted crop. *Trends Plant Sci.* 22, 917–920. <https://doi.org/10.1016/j.tplants.2017.08.004>.
- Şeker, A., Shahwan, T., Eroğlu, A.E., Yılmaz, S., Demirel, Z., Dalay, M.C., 2008. Equilibrium, thermodynamic and kinetic studies for the biosorption of aqueous lead (II), cadmium(II) and nickel(II) ions on spirulina platensis. *J. Hazard. Mater.* 154, 973–980. <https://doi.org/10.1016/j.jhazmat.2007.11.007>.
- Sun, W., Cheng, K., Sun, K.Y., Ma, X., 2021. Microbially mediated remediation of contaminated sediments by heavy metals: a critical review. *Curr. Pollut. Rep.* 7, 201–212. <https://doi.org/10.1007/s40726-021-00175-7>.
- Teboumou, P.R., Yedjou, C.G., Paolillo, A.F., Sutton, D.J., 2012. *Molecular, Clinical and Environmental Toxicology*, 101. Springer, Basel. https://doi.org/10.1007/978-3-7643-8340-4_6. Molecular, Clinical and Environmental Toxicology.
- Tiwari, S., Lata, C., 2018. Heavy metal stress, signaling, and tolerance due to plant-associated microbes: an overview. *Front. Plant Sci.* 9, 452. <https://doi.org/10.3389/fpls.2018.00452>.
- Tripathi, R.D., Dhotre, S., Shukla, M.K., Mishra, S., Srivastava, S., Singh, R., Rai, U.N., Gupta, D.K., 2008. Role of blue green algae biofertilizer in ameliorating the nitrogen demand and fly-ash stress to the growth and yield of rice (*Oryza sativa* L.) plants. *Chemosphere* 70, 1919–1929. <https://doi.org/10.1016/j.chemosphere.2007.07.038>.
- Tsukihara, T., Fukuyama, K., Tshara, H., Katsube, Y., Matsuura, Y., Tanaka, N., Kakudo, M., Wada, K., Matsubara, H., 1978. X-ray analysis of ferredoxin from spirulina platensis. *J. Biochem.* 84, 1645–1647. <https://doi.org/10.1093/oxfordjournals.jchem.a132293>.
- Vardhan, K.H., Kumar, P.S., Panda, R.C., 2019. A review on heavy metal pollution, toxicity and remedial measures: current trends and future perspectives. *J. Mol. Liq.* 10, 907. <https://doi.org/10.1016/j.molliq.2019.111197>.
- Varela, J.P., Valente, A.J.M., Duriães, L., 2019. Assessment of heavy metal pollution from anthropogenic activities and remediation strategies: a review. *J. Environ. Manage.* 246, 101–118. <https://doi.org/10.1016/j.jenvman.2019.05.126>.
- Vasanthi Rupasinghe, H.P., Davis, A., Kumar, S.K., Murray, B., Zhejazkov, V.D., 2020. Industrial hemp (*Cannabis sativa* subsp. *sativa*) as an emerging source for value-added functional food ingredients and nutraceuticals. *Molecules* 25, 4078. <https://doi.org/10.3390/molecules25184078>.
- Winter, H., Huber, S.C., 2000. Regulation of sucrose metabolism in higher plants: localization and regulation of activity of key enzymes. *CRC Crit. Rev. Plant Sci.* 19, 31–67. <https://doi.org/10.1080/07352680091139178>.
- Wuana, R.A., Okieimere, P.E., 2011. Heavy metals in contaminated soils: a review of sources, chemistry, risks and best available strategies for remediation. *ISRN Ecol.* 2011, 1–20. <https://doi.org/10.5402/2011/402647>.
- Wuang, S.C., Khin, M.C., Chua, P.Q.D., Luo, Y.D., 2016. Use of spirulina biomass produced from treatment of aquaculture wastewater as agricultural fertilizers. *Algal Res.* 15, 59–64. <https://doi.org/10.1016/j.algal.2016.02.009>.
- Zielonka, D., Szuk, W., Skowrońska, M., Rutkowska, B., Russel, S., 2020. Hemp-based phytocumulation of heavy metals from municipal sewage sludge and phosphogypsum under field conditions. *Agronomy* 10, 907. <https://doi.org/10.3390/agronomy1006907>.
- Zinicovscaia, L., Cepoi, L., Chiriac, T., Ana Culicov, O., Frontasyeva, M., Pavlov, S., Kifkenali, E., Akhmetiev, A., Rodionovskaya, E., 2016. Spirulina platensis as biosorbent of chromium and nickel from industrial effluents. *Desal. Water Treat.* 57, 11103–11110. <https://doi.org/10.1080/19443994.2015.1042061>.
- Zinicovscaia, L., Safonov, A., Ostalkevich, S., Gundorina, S., Nekhoroshkov, P., Groszov, D., 2019. Metal ions removal from different type of industrial effluents using spirulina platensis biomass. *Int. J. Phytoremed.* 21, 1442–1448. <https://doi.org/10.1080/15228314.2019.1633264>.
- Zwolak, A., Surzyńska, M., Szpyrka, E., Stawarczyk, K., 2019. Sources of soil pollution by heavy metals and their accumulation in vegetables: a review. *Water Air Soil Pollut.* 230, 164. <https://doi.org/10.1007/s11270-019-4221-y>.

Supporting Information

A spectroscopic study to assess heavy metals absorption by a combined hemp/spirulina system from contaminated soil

Biagia Musio,^{a,b,*} Elhoussein Mohamed Fouad Mourad Hussein Ahmed,^{a,c} Marica Antonicelli,^a Danila Chiapperini,^d Onofrio Dursi,^d Flavia Grieco,^d Mario Latronico,^{a,b} Piero Mastrorilli,^{a,b} Rosa Ragone,^{a,b} Raffaele Settanni,^d Maurizio Triggiani,^{a,b} Vito Gallo^{a,b,*}

^aDepartment of Civil, Environmental, Land, Building Engineering and Chemistry (DICATECh), Polytechnic University of Bari, Via Orabona, 4, I-70125, Bari, Italy

^bInnovative Solutions S.r.l. – Spin-off company of Polytechnic University of Bari, Zona H 150/B, 70015, Noci (BA), Italy.

^cInternational Centre for Advanced Mediterranean Agronomic Studies of Bari (CIHEAM Bari), Via Ceglie 9, 70010, Valenzano (BA), Italy.

^dApuliaKundi S.r.l., via N. Copernico snc, Zona PIP, 70024 Gravina in Puglia (BA), Italy.

INDEX

Table S1. Morphometric data recorded at three different developmental stages (t1, t2, and t3).	3
Results of ICP-AES quantitative analysis of residual metal (mg/kg) in all the matrices under study (leaves, stems, and soil), and values of biomass (kg).	4
Table S3. Metabolites identified in the 400 MHz ¹ H-NMR spectra of <i>Cannabis sativa</i> L. leaves and stems extracts in D ₂ O (pH 4.2).	7
Table S4. Parameters characterizing the OPLS model, built by using the data derived from 1D ¹ H NMR spectroscopy, ICP-AES quantitative analysis of residual metal, and morphometric inspection (biomass).	9
Figure S1. Visual inspection of the original harvested plants combined with the plant biomass values (pictures of less healthy plants have dashed borders).	10
Table S5. Extract from the correlation matrix, related to the OPLS model, highlighting the correlations among the six metals under investigation and the biomass (colour intensity is proportional to the absolute value of correlation, <i>r</i>).	10

Figure S2. Score plot (a) and S-plot (b) relating to the OPLS-DA model obtained by setting the plant tissue (LE vs ST) as class.	11
Table S6. The most significant buckets differentiating leaves (LE) from stems (ST) and their relative values of p1, p(corr)1 and VIPpredictive, obtained by applying OPLS-DA.....	12
Figure S3. Trend of the buckets (x-variables) at 2.646 - 2.686 ppm and 2.726 - 2.766 ppm, assigned to citric acid, in samples from plants treated with microalgae at different concentration (CS1 vs CS0.5).....	13

Table S1. Morphometric data recorded at three different developmental stages (t1, t2, and t3).

Parameter	Plant ID	t1	t2	t3
Plant height (m)	Ctrl_a	0.44	0.68	0.76
	Ctrl_b	0.50	0.88	0.75
	C_a	0.40	0.59	0.64
	C_b	0.51	0.87	0.80
	CS1_a	0.49	0.85	0.70
	CS1_b	0.36	0.69	0.80
	CS0.5_a	0.42	0.71	0.68
	CS0.5_b	0.50	0.94	0.80
Drum diameter (mm)	Ctrl_a	5	6	6
	Ctrl_b	5	7	7
	C_a	5	6	6
	C_b	5	8	9
	CS1_a	5	8	8
	CS1_b	5	8	7
	CS0.5_a	5	7	6
	CS0.5_b	4	6	6
Leaf stages (n.)	Ctrl_a	6	10	8
	Ctrl_b	7	10	6
	C_a	6	9	7
	C_b	7	10	5
	CS1_a	7	10	5
	CS1_b	8	12	10
	CS0.5_a	6	10	8
	CS0.5_b	6	10	5

Results of ICP-AES quantitative analysis of residual metal (mg/kg) in all the matrices under study (leaves, stems, and soil), and values of biomass (kg).

The determination of Cd, Cr, Cu, Ni, Pb, and Zn was performed by ICP - OES according to UNI EN 13657: 2004 and UNI EN ISO 11885: 2009. The sample under examination is treated with Aqua Regia (three parts concentrated hydrochloric acid (HCl) and one part concentrated nitric acid (HNO₃)) by means of microwave assisted digestion in a closed container and the solution obtained is analyzed by Inductively Coupled Plasma Emission Spectrometry (ICP - AES). The method is based on the measurement of Atomic Emission using Optical Spectroscopy technique. The samples are nebulized and the aerosol produced is transferred to a plasma torch, where the excitation takes place. The line spectra are produced by inductively coupled plasma by means of radio frequency (ICP). The spectra are dispersed by a grating spectrometer and the intensity of the lines is evaluated by the detectors. The signals coming from the detector are processed and controlled by a computer system. A background noise correction technique is used to compensate for variations in the contributions of the noise itself to the determination of trace elements. The concentration of analyt present in the sample is determined, by comparison, with a reference solution of known concentration. The microwave unit (CEM MARS system) provides a variable power that can be programmed within more or less 10W of the required power and provides a rated power from 250 to 1200W. Calibration of the power emitted by the microwave unit is performed by heating a known quantity of water for a "standard" period of 2 minutes at different microwave powers. Any temperature rise must be measured with an accuracy of ± 0.1 °C after each heating cycle. The absorbed power is calculated from the increase in temperature. The calibration is carried out for 3 points, measuring the power absorbed at the three different powers chosen. The calibration procedure is successful if the measured absorbed power corresponds to the chosen power (power difference higher than ± 10 W).

The reference standards used for the construction of the calibration curves are freshly prepared each time before carrying out the analysis of the samples. Multi-element reference standard concentrated solution at a concentration of 10 mg / l was used. All the elements present in the mixture have a concentration of 10 ppm. The sample under examination must be perfectly homogeneous and it is subjected to a pre-treatment as follows: the granulometry of the sample is reduced to particle size <250 μ m; the sample is dried in the air and in any case evaluating the percentage of humidity present since this data is inserted in the calculation formula; the sample is homogenized through a high-speed mixer. Weigh between 0.2 and 0.5 g of the sample, to the nearest 0.1 mg, is transferred into the vessel. 6 ml of HCl conc. and 2 ml of HNO₃ conc were added. The digestion vessels were transferred to the cavity of the microwave unit in accordance with the manufacturer's instructions and the following digestion procedure was started.

tempo in minuti	potenza in W
2	250
2	0
5	250
5	400
5	500

At the end of the program, the containers were allowed to cool to room temperature and then the contents of the vessels were transferred into a volumetric flask and bring to volume (10 ml). The sample was filtered through 0.45 μ m porosity filters and transferred to the measuring equipment.

To detect possible contamination from vessels and / or reagents, blank tests are carried out, in parallel, using the same digestion and filtration procedure, using the same quantities of all reagents but omitting the test portion.

The LOQ for all the analyzed metals is 0.01 mg/Kg.

For the evaluation of the calibration uncertainty, the following contributions were evaluated: uncertainty of repeatability determined through the execution of 10 independent tests; 5-point calibration uncertainty; uncertainty of control contribution (solution 0.500 mg / l) calibration curve; uncertainty of dilution of the reference materials (flask, pipette, temperature variation contribution). The summary table below shows the uncertainty values in percent (%) determined considering the contributions of type A and B. The last 2 columns show the uncertainty values determined at concentrations higher than 0.05 mg / kg and at the limits of the analytical quantification.

ANALYTE	UNC. pipette MAX		UNC. CALIBRATION 0.01		UNC. CALIBRATION 0.1		UNC. CONTROL		UNC. PIPETTE		UNC. ΔT PIPETTE		UNC. VOLUM. FLASK		UNC. ΔT VOLUM. FLASK		UNC. DILUTION		UNC. %	UNC. limite quant. %
	12	57	6	10	0,2	0,07	0,8	0,1	0,83	17	59									
Cd	2	33	3	10	0,2	0,07	0,8	0,1	0,83	11	35									
Cr	3	49	5	10	0,2	0,07	0,8	0,1	0,83	12	50									
Ni	3	52	7	10	0,2	0,07	0,8	0,1	0,83	13	53									
Pb	2	29	3	10	0,2	0,07	0,8	0,1	0,83	11	31									
Cu	2	65	3	10	0,2	0,07	0,8	0,1	0,83	11	66									
Zn																				

To calculate the results, the blank values of the method were subtracted from all the samples. The results are reported in milligrams of element per kilogram of sample (mg / kg) with a maximum of 2 significant figures. The following formula was applied:

$$\text{Concentration (mg/Kg)} = \frac{C * V * 10^5}{W * SS}$$

Where:

C = concentration read in mg/L

V = Final volume in liters after sample preparation

W = weight of the sample in g

SS = dry matter in g / 100 g

SS

Matrix	Plant_ID	Biomass (kg)	Cd (mg/kg)	Cr (mg/kg)	Cu (mg/kg)	Pb (mg/kg)	Ni (mg/kg)	Zn (mg/kg)
Leaves	Ctrl_a	0.0019	0.01±0.01	2.50±0.27	5.30±0.58	3.10±0.40	2.40±0.29	180.00±19.80
	Ctrl_b	0.0048	0.01±0.01	1.70±0.19	5.20±0.57	1.70±0.22	10.00±1.20	59.00±6.49
	C_a	0.0222	0.28±0.05	4.00±0.44	12.00±1.32	1.00±0.13	10.00±1.20	185.00±20.35
	C_b	0.0165	0.27±0.05	1.60±0.18	8.80±0.97	1.00±0.13	8.30±0.10	95.00±10.45
	CS1_a	0.0219	0.55±0.09	5.00±0.55	11.00±1.21	1.60±0.21	22.00±2.64	180.00±19.80
	CS1_b	0.0232	1.10±0.19	3.00±0.33	10.00±1.10	0.49±0.06	18.00±2.16	220.00±24.20
	CS0.5_a	0.0066	1.00±0.17	1.70±0.19	8.00±0.88	2.70±0.35	12.00±1.44	180.00±19.80
	CS0.5_b	0.0081	0.84±0.14	1.50±0.16	12.00±1.32	1.00±0.13	14.00±1.68	200.00±22.00
Stems	Ctrl_a	0.0049	0.01±0.01	0.79±0.087	2.10±0.23	2.70±0.35	2.00±0.24	38.00±4.18
	Ctrl_b	0.0040	0.01±0.01	0.85±0.09	1.60±0.18	2.80±0.36	4.50±0.54	28.00±3.08
	C_a	0.0181	0.34±0.06	2.80±0.31	5.10±0.56	1.10±0.14	7.50±0.90	63.00±6.93
	C_b	0.0160	0.29±0.05	1.80±0.20	4.00±0.44	3.60±0.47	6.10±0.73	25.00±2.75
	CS1_a	0.0049	1.20±0.20	0.90±0.10	6.10±0.67	2.40±0.31	13.00±1.56	120.00±13.20
	CS1_b	0.0106	1.40±0.24	0.50±0.05	5.60±0.62	3.10±0.40	13.00±1.56	120.00±13.20
	CS0.5_a	0.0041	0.99±0.17	1.70±0.19	2.60±0.29	1.50±0.19	5.80±0.70	54.00±5.94
	CS0.5_b	0.0073	1.10±0.19	1.30±0.14	5.20±0.57	2.60±0.34	11.00±1.32	78.00±8.58
Leaves + stems	Ctrl_a	0.0067	0.01±0.01	1.26±0.14	2.99±0.33	2.81±0.36	2.11±0.25	77.36±8.51
	Ctrl_b	0.0088	0.01±0.01	1.31±0.14	3.55±0.39	2.21±0.29	7.48±0.90	44.77±4.92
	C_a	0.0403	0.31±0.05	3.45±0.38	9.15±1.01	1.04±0.13	8.96±1.07	130.26±14.33
	C_b	0.0325	0.28±0.05	1.72±0.19	6.45±0.71	2.30±0.30	7.20±0.86	60.24±6.63
	CS1_a	0.0268	0.67±0.11	4.25±0.47	10.10±1.11	1.75±0.23	20.35±2.44	169.02±18.59
	CS1_b	0.0338	1.19±0.20	2.22±0.24	8.62±0.95	1.31±0.17	16.44±1.97	188.71±20.76
	CS0.5_a	0.0107	1.00±0.17	1.70±0.19	5.92±0.65	2.24±0.29	9.62±1.15	131.57±14.47
	CS0.5_b	0.0154	0.96±0.16	1.40±0.15	8.76±0.96	1.76±0.23	12.57±1.51	141.93±15.6
Soil	Ctrl_a	4	0.21±0.04	4.50±0.49	8.40±0.92	13.00±1.69	4.80±0.58	47.00±5.17
	Ctrl_b	4	0.33±0.06	8.10±0.89	9.40±1.03	14.00±1.82	5.00±0.60	42.00±4.62
	C_a	4	9.90±1.68	290.00±31.90	110.00±12.10	180.00±23.40	210.00±25.20	340.00±37.40
	C_b	4	9.90±1.68	330.00±36.30	150.00±16.50	270.00±35.10	220.00±26.40	320.00±35.20
	CS1_a	4	11.00±1.87	400.00±44.00	170.00±18.70	330.00±42.90	240.00±28.80	330.00±36.30
	CS1_b	4	7.40±1.26	240.00±26.40	120.00±13.20	250.00±32.50	160.00±19.20	260.00±28.60
	CS0.5_a	4	11.00±1.87	420.00±46.20	180.00±19.80	360.00±46.80	260.00±31.20	340.00±37.40
	CS0.5_b	4	9.70±1.65	270.00±29.70	110.00±12.10	180.00±23.40	190.00±22.80	310.00±34.10

Table S3. Metabolites identified in the 400 MHz ¹H-NMR spectra of *Cannabis sativa* L. leaves and stems extracts in D₂O (pH 4.2).

Compounds	LEAVES			STEMS		
	¹ H (ppm)	Multiplicity	[J (Hz)]	¹ H (ppm)	Multiplicity	[J (Hz)]
α-Glucose / β-Glucose	3.25	dd	9.15; 8.01	3.24	dd	9; 7.9
	3.43	m		3.44	m	
	3.51	d	8.86	3.49	d	8.8
	3.54	dd	9.23; 3.9	3.53	dd	9.8; 3.8
	3.74	m		3.73	m	
	3.83	m		3.83	m	
	3.89	dd	12.3; 2.3	3.9	dd	12.4; 2.3
	4.65	d	7.9	4.65	d	7.9
	5.24	d	3.7	5.23	d	3.7
	Fructose	3.55	s		3.55	s
	3.58	t	4.2	3.58	t	4.1
	3.69	m		3.69	m	
	3.73	m		3.73	m	
	3.89	d	3.5	3.89	d	3.6
	3.91	d	3.5	3.91	d	3.5
	4.01	m		4	m	
	4.04	d	1.4	4.04	d	1.4
	4.12	m		4.11	m	
α-Galactose / β-Galactose¹	5.27	d	3.79	5.26	d	3.79
	4.59	d	7.86	4.58	d	7.86
	4.10	m		4.09	m	
	3.99	m		3.98	m	
	3.93	m		3.92	m	
	3.86 overlapped	dd	10.21; 3.40	3.85 overlapped	dd	10.21; 3.40
	3.81 overlapped	dd	10.15; 3.65	3.80 overlapped	dd	10.15; 3.65
	3.75	m		3.74	m	
	3.66 overlapped	dd	9.90; 3.22	3.65 overlapped	dd	9.90; 3.22
	Ethanol	1.2	t	7.1	1.19	t
Acetic acid¹	1.95	s		2	s	
Citric acid	2.62	d	15.2	2.68	d	15.4
	2.73	d	15.3	2.78	d	15.3
Formic acid¹	8.45	s		8.43	s	
Fumaric acid¹	6.53	s		6.55	s	
Lactic acid	1.33	d	6.9	1.33	d	6.9
	4.15	q	6.9	4.14	q	6.9
Tartaric acid	4.44	s		4.43	s	
Choline¹	3.2	s		3.19	s	

¹ Assignment from literature (Ingallina et. Al, 2020)

Trigonelline¹	9.12	s		9.12	s	
	8.85	m		8.85	m	
	8.1	m		8.09	m	
Alanine	1.48	d	7.2	1.47	d	7.2
	3.79	q	7.2	3.78	q	7.2
Asparagine	2.91	dd	7.4; 4.4	2.9	dd	7.4; 4.4
	4.02	m		4.01	m	
	6.92	bs		6.92	bs	
	7.65	bs		7.65	bs	
γ-Aminobutyrate	1.9	q	7.4	1.91	q	7.4
	2.34	t	7.3	2.35	t	7.3
	3.02	t	7.6	3.02	t	7.6
Arginine	1.67	m		1.68	m	
	1.9	m		1.91	m	
	3.25	m		3.24	m	
	3.79	t	6.1	3.78	t	6.1
Glutamine	2.13	m		2.13	m	
	2.45	m		2.45	m	
	3.78	t	6.2	3.78	t	6.2
	6.91	bs		6.91	bs	
	7.65	bs		7.65	bs	
Isoleucine	0.92	t	7.4	0.93	t	7.4
	0.99	d	7	1	d	7
	1.23	m		1.24	m	
	1.45	m		1.46	m	
	1.96	m		1.97	m	
	3.6	d	3.9	3.68	d	3.9
Proline	2.04	m		2.03	m	
	2.35	m		2.34	m	
	3.36	m		3.37	m	
	4.12	m		4.11	m	
Tyrosine	3.06	dd	14.6; 7.7	3.05	dd	14.6; 7.7
	3.2	dd	14.6; 5.2	3.19	dd	14.6; 5.2
	3.96	dd	7.7; 5.2	3.95	dd	7.7; 5.2
	6.9	d	8.5	6.89	d	8.5
	7.2	d	8.5	7.19	d	8.5
Leucine	0.96	m		0.95	m	
	1.70	m		1.69	m	
	3.73	m		3.72	m	

Table S4. Parameters characterizing the OPLS model, built by using the data derived from 1D ¹H NMR spectroscopy, ICP-AES quantitative analysis of residual metal, and morphometric inspection (biomass).

Component	R ² X	R ² X(cum)	R ²	R ² (cum)	Q ²	Q ² (cum)	R ² Y	R ² Y(cum)
Model		0.978		0.929		0.820		0.922
Predictive		0.595		0.876		0.820		0.922
P1	0.410	0.410	0.472	0.472	0.161	0.161	0.485	0.485
P2	0.098	0.509	0.151	0.623	0.232	0.393	0.152	0.637
P3	0.043	0.552	0.169	0.792	0.153	0.546	0.182	0.819
P4	0.024	0.576	0.052	0.844	0.135	0.681	0.062	0.881
P5	0.019	0.595	0.032	0.876	0.139	0.820	0.041	0.922
Orthogonal in X (OPLS)		0.383		0.052				
O1	0.145	0.145	0.002	0.002				
O2	0.070	0.215	0.003	0.004				
O3	0.050	0.265	0.006	0.010				
O4	0.047	0.312	0.010	0.020				
O5	0.027	0.340	0.007	0.027				
O6	0.018	0.358	0.011	0.038				
O7	0.016	0.373	0.011	0.049				
O8	0.010	0.383	0.003	0.052				

Figure S1. Visual inspection of the original harvested plants combined with the plant biomass values (pictures of less healthy plants have dashed borders).

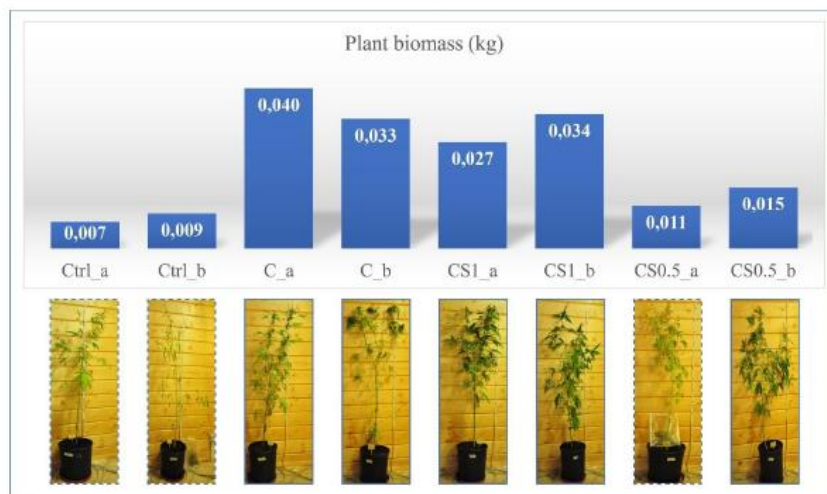
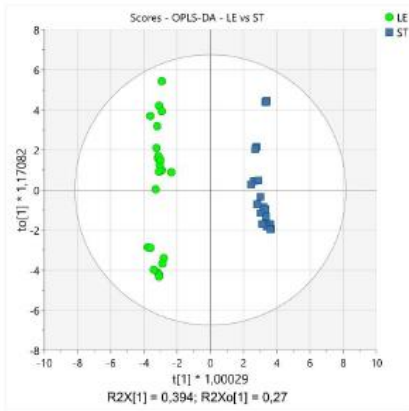


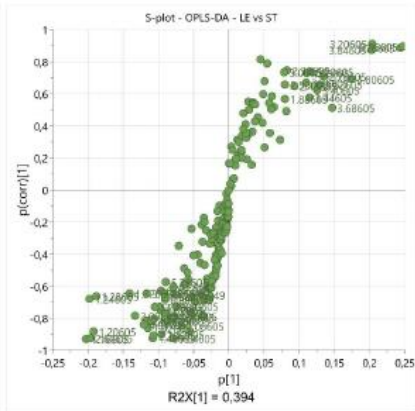
Table S5. Extract from the correlation matrix, related to the OPLS model, highlighting the correlations among the six metals under investigation and the biomass (colour intensity is proportional to the absolute value of correlation, r).

	Cd	Cr	Cu	Pb	Ni	Zn	Cd_plant	Cr_plant	Cu_plant	Pb_plant	Ni_plant	Zn_plant	sample biomass	plant biomass	
Cd	1,00	-0,18	0,24	-0,06	0,38	0,35	0,95	0,24	0,64	-0,46	0,74	0,84	0,05	0,22	
Cr		1,00	0,61	-0,49	0,46	0,51	-0,02	0,56	0,38	-0,37	0,22	0,25	0,71	0,40	
Cu			1,00	-0,62	0,72	0,84	0,36	0,41	0,61	-0,50	0,45	0,47	0,62	0,45	
Pb				1,00	-0,42	-0,40	-0,20	-0,30	-0,41	0,50	-0,24	-0,28	-0,54	-0,37	
Ni					1,00	0,63	0,64	0,54	0,74	-0,58	0,88	0,73	0,53	0,39	
Zn						1,00	0,46	0,29	0,43	-0,35	0,41	0,59	0,35	0,20	
Cd_plant							1,00	0,15	0,61	-0,46	0,69	0,84	0,17	0,20	
Cr_plant								1,00	0,74	-0,62	0,66	0,56	0,54	0,64	
Cu_plant									1,00	-0,82	0,82	0,80	0,61	0,72	
Pb_plant										1,00	-0,61	-0,68	-0,64	-0,76	
Ni_plant											1,00	0,81	0,35	0,42	
Zn_plant												1,00	0,35	0,42	
sample biomass													1,00	0,84	
plant biomass															1,00

Figure S2. Score plot (a) and S-plot (b) relating to the OPLS-DA model obtained by setting the plant tissue (LE vs ST) as class.



a.

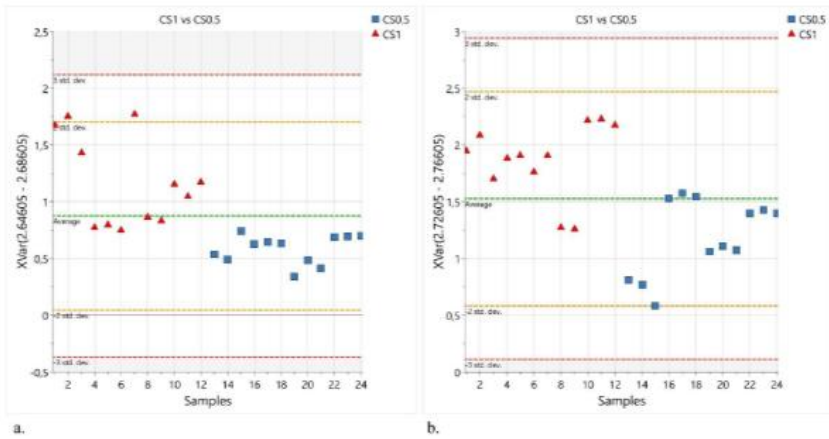


b.

Table S6. The most significant buckets differentiating leaves (LE) from stems (ST) and their relative values of p1, p(corr)1 and VIPpredictive, obtained by applying OPLS-DA.

Bucket	p[1]	p(corr)[1]	VIPpred
0.806049 - 0.846049	-0.075	-0.653	1,115
0.846049 - 0.886049	-0.097	-0.628	1,449
0.886049 - 0.926049	-0.091	-0.684	1,360
1.04605 - 1.08605	-0.079	-0.825	1,182
1.08605 - 1.12605	-0.067	-0.849	1,009
1.12605 - 1.16605	-0.095	-0.898	1,422
1.16605 - 1.20605	-0.203	-0.929	3,036
1.20605 - 1.24605	-0.192	-0.881	2,871
1.24605 - 1.28605	-0.199	-0.677	2,971
1.28605 - 1.32605	-0.188	-0.664	2,817
1.32605 - 1.36605	-0.141	-0.647	2,108
1.36605 - 1.40605	-0.092	-0.642	1,383
1.40605 - 1.44605	-0.075	-0.728	1,125
1.48605 - 1.52605	-0.108	-0.926	1,613
1.52605 - 1.56605	-0.106	-0.803	1,593
1.56605 - 1.60605	-0.096	-0.751	1,440
1.88605 - 1.92605	0.079	0.570	1,188
1.92605 - 1.96605	-0.115	-0.863	1,715
2.00605 - 2.04605	-0.133	-0.785	1,992
2.04605 - 2.08605	-0.117	-0.814	1,753
2.16605 - 2.20605	-0.078	-0.790	1,160
2.20605 - 2.24605	-0.109	-0.819	1,635
2.24605 - 2.28605	-0.088	-0.819	1,311
2.28605 - 2.32605	-0.107	-0.907	1,607
2.32605 - 2.36605	0.082	0.746	1,229
2.48605 - 2.52605	0.093	0.647	1,388
2.52605 - 2.56605	0.080	0.657	1,190
2.56605 - 2.60605	-0.091	-0.866	1,360
2.60605 - 2.64605	-0.198	-0.927	2,957
2.64605 - 2.68605	-0.071	-0.346	1,069
2.72605 - 2.76605	0.073	0.313	1,092
2.76605 - 2.80605	0.135	0.709	2,020
3.00605 - 3.04605	0.078	0.733	1,174
3.20605 - 3.24605	0.204	0.912	3,046
3.28605 - 3.32605	-0.098	-0.729	1,467
3.36605 - 3.40605	0.110	0.671	1,646
3.40605 - 3.44605	0.125	0.622	1,877
3.44605 - 3.48605	0.115	0.577	1,715
3.52605 - 3.56605	0.129	0.660	1,931
3.60605 - 3.64605	-0.116	-0.644	1,743
3.68605 - 3.72605	0.146	0.512	2,185
3.76605 - 3.80605	0.243	0.888	3,633
3.80605 - 3.84605	0.175	0.690	2,613
3.84605 - 3.88605	0.202	0.874	3,024
3.96605 - 4.00605	0.247	0.894	3,700
4.04605 - 4.08605	-0.121	-0.842	1,806
4.08605 - 4.12605	0.111	0.729	1,655
4.24605 - 4.28605	-0.081	-0.794	1,218
4.28605 - 4.32605	-0.079	-0.602	1,176
4.32605 - 4.36605	0.082	0.492	1,222
4.48605 - 4.52605	-0.089	-0.769	1,338
5.20605 - 5.24605	0.117	0.739	1,749
5.24605 - 5.28605	-0.106	-0.729	1,580
5.28605 - 5.32605	-0.090	-0.572	1,340

Figure S3. Trend of the buckets (x-variables) at 2.646 - 2.686 ppm and 2.726 - 2.766 ppm, assigned to citric acid, in samples from plants treated with microalgae at different concentration (CS1 vs CS0.5).



CHAPTER 6

Non-Targeted Spectranomics for Early-Detection of *Xylella fastidiosa*-Infected Asymptomatic Olive Leaves

In this chapter, we explored the possibility of performing an early-detection of the *Xylella fastidiosa*-infection in leaves collected exclusively from asymptomatic trees. ^1H NMR, hyperspectral reflectance (HSR), and chemometrics were exploited to select diagnostic signals in the NMR spectrum and specific wavelengths in the HSR spectrum, respectively, as a probe to assess the presence of *Xylella fastidiosa* -infection at an early stage.

INTRODUCTION

Olive quick decline syndrome (OQDS) is an olive disease that started to develop in the Salento peninsula (Apulia, southern Italy) ca. 2009 and subsequently heavily destroyed olive orchards in that region (Saponari et al., 2013).

The olive cultivar of “Cellina di Nardò” was the most susceptible to OQDS (Saponari et al., 2019; OEPP/EPPO, 2018). Tolerance of other cultivars to *X. fastidiosa* was related to the ability of the host plant to reduce *X. fastidiosa* colonisation as well as dysbiosis (Sabella et al., 2020). *X. fastidiosa* pathogenicity affects the universal function of water movement within the plant by hindering the transporting of nutrients and signals through the xylem vessels (Martelli et al., 2016). In the course of the disease, from infection to the decline of the plant, the olive tree undergoes phenotypic and metabolic fluctuations. However, in the early stages of infection, trees often remain asymptomatic for years, hindering disease detection in infected plants, which undergo rapid decline. So far, the main strategy adopted to control the spread of infection is to eradicate infected plants when they are symptomatic. (European Commission, 2020). Hence, efficient early detection of Xf infection would offer an advantage against the spread of the disease by allowing the early implementation of preventive actions to protect olive growing and the related agro-economic sector.

In the context of spectroscopy-based metabolomics, nuclear magnetic resonance (NMR) offers a non-destructive strategy to detect metabolic fluctuations due to the different stages of the disease (Hall, 2018; Bohnenkamp et al., 2019). In a parallel way, hyperspectral reflectance (HSR) remote/proximal sensing offers a non-destructive way of measuring electromagnetic waves reflected by the plant surface, in which specific spectral bands/signatures are used to detect infections before the development of symptoms (Mahlein et al., 2018). Moreover, spectral signatures determine a gamut of chemicals found in plants, and the specificity of plant chemical fingerprints is reported to be strongly influenced by several factors (Jacquemoud et al., 1995; Curran, 1989).

In this study, young trees of the Xf-susceptible variety, Cellina di Nardò, artificially infected with Xf, were cultivated for 2 years in a thermo-controlled environment. The spectral data obtained through the two analytical techniques were subjected to chemometric analysis, providing useful correlation matrices capable of linking the HSR spectral features with diagnostic NMR signals. Determining the diagnostic wavelength regions associated with specific metabolites is a keystone for the development of proximal and remote sensing devices capable of early detection of Xf-infection in olive trees.

MATERIALS AND METHODS

Cultivation of Xylella fastidiosa

A suspension was prepared from the “De Donno” strain (CFBP 8402) of *Xylella fastidiosa* subsp. *pauca* (Wells et al., 1987) taken from 8-10 days old Buffered Charcoal Yeast Extract Agar (BCYE) culture medium grown at 28°C, resuspended in Phosphate-Buffered Saline solution (PBS, 0.05 M, pH 7.2), and adjusted to 0.5-0.6 absorbance at an optical density of 600 (OD600), which corresponds to a concentration of 10⁹ colony forming unit (CFU)/mL (Koch, 1970). A sterile PBS solution was used as a control.

Cultivation and artificial infection of the olive plants

Olive plants (*Olea europaea* L. cv. Cellina di Nardò) were cultivated in a quarantine greenhouse under a controlled environment of 23-24°C in winter and 25-30°C in summer under >80% relative humidity at the CNR research area of Bari (Italy). After two years the plants were inoculated with Xf and according to a reported procedure (Saponari et al., 2017). Table 1 summarizes the adopted experimental design.

Artificial infection with Xf was performed via pinprick inoculation (Saponari et al., 2017, 2016; Hill & Purcell, 1995; OEPP/EPPO, 2019). Specifically, aliquots of 10 µL of the PBS suspension containing Xf, prepared according to the procedure described above, were punched with sterile entomological pins 5-6 times. For each plant, the 9-12 inoculation sites were carried out on three consecutive leaf nodes of 3 to 4 twigs placed at 40-50 cm from the ground.

Real-time Polymerase Chain Reaction (qPCR) for confirmation of Xf Presence in Inoculated Plants

DNA was extracted from leaves using a CTAB (cetyltrimethylammonium bromide)-based method (Murray & Thompson, 1980). Pieces of midveins and petioles (ca. 0.5 g) were hammer-smashed in sterilized plastic bags. Then, 5 mL of

CTAB buffer (2%, 0.1 M Tris-HCl pH 8, 20 mM EDTA, and 1.4 M NaCl) was added, followed by homogenization through Homex 6 homogenizer (Bioreba, Reinach Switzerland). Aliquots of the resulting homogenized suspension (1 mL) were transferred into 2 mL microcentrifuge tubes containing 1 mL of chloroform, followed by incubation at 65°C for 30 min using a water bath. Finally, the aqueous phase was separated and treated with 0.7 volumes of cold isopropanol inducing the DNA precipitation, which was, subsequently, used for qPCR analysis (Loconsole et al., 2014).

Quantitative real-time Polymerase Chain Reactions (qPCR) were performed using 20 µL reaction volumes containing 10 µL of 2 X qPCR Supermix-UDG (Invitrogen; Thermo Fisher Scientific Inc., Waltham, MA, USA), reaching a final concentration of 4 mM MgCl₂, 300 nM of *X. fastidiosa* forward (XF-F; 5`-CAC-GGCTGGTAACGGAAGA-3`) and reverse (XF-R; 5`-GGGTTGCGTGGTGAAATCAAG-3`) primers, 100 nM dual-labelled fluorescent probe (XF-P; 5`-TCG-CATCCCGTGGCTCAGTCC-3`) labelled by 5'-Fluorescein/Black Hole Quencher 1 (6-FAM/BHQ-1), bovine serum albumin (BSA) at 300 ng/µL (Sigma-Aldrich), and 2 µL of DNA template.

Thermocycling conditions were 50°C for 2 min, 94°C for 2 min, then 40 cycles of 94°C for 10 s and 62°C for 40 s (Harper et al., 2010). All samples were amplified in triplicates. Threshold values were applied automatically by the CFX Manager V1.6 software (Bio-Rad Laboratories).

Hyperspectral reflectance (HSR)

HSR was acquired by a hyperspectral acquisition system. This system consisted of a FieldSpec®3 spectroradiometer (Analytical Spectral Device [ASD], Boulder, CO, USA) linked by an optical fibre cable to a leaf probe (ASD) and a leaf clip holder (ASD) in addition to an instrument controller (laptop) to display and save the data. The plant probe was of 10 mm spot size with internal-illumination by a halogen bulb of 2901 K ±10% colour temperature. The HSR data were in the range of 350-1830 nm, with spectral sampling interval of 1.4 nm and 2 nm at spectral

ranges of 350–1050 nm (Full Width Half Maximum, FWHM: 3 nm) and 1000–1830 nm (FWHM: 10 nm), respectively.

The final resolution of 1 nm was obtained by subsampling and interpolation of the spectral channels and recorded as relative reflectance values by the RS³™ software (ASD). All the acquired raw HSR spectra were stored in a “.asd” format while the pre-processed raw data were stored in “.ref” format then exported in as “.txt” format (American Standard Code for Information Interchange, ASCII) and imported in a MATLAB R2021a routine (The MathWorks Inc., Natick, MA, USA) developed by the authors for further pre-processing.

The raw HSR spectra were pre-processed using a Savitzky-Golay filter (Savitzky & Golay, 1964; Nevius & Pardue, 1984) with frame size of 15 data points (2nd Degree polynomial) was applied. Then, the UV region from 350 nm to 400 nm was truncated and the interval ranging between 400 nm and 1,800 nm was considered for further analysis. The pre-processed spectra were exported in comma-separated values “.csv” format.

NMR Sample Preparation and Spectra Acquisition

After HSR acquisition, the collected olive leaves were lyophilized at –50°C under 0.180 mbar for 72 h in Christ Alpha 1-4 LSC lyophiliser (Martin-Christ Gefriertrocknungsanlagen GmbH, Osterode am Harz, Germany). The dried samples were then ground in a blender, sieved through a mesh of 0.5 mm pores, and stored at –20°C. For each NMR sample, an amount of 50 mg of olive leaf powder and 1.5 mL of oxalate buffer at pH 4.2 (pH value was reached after addition of 37% HCl to 100 mL an aqueous solution containing 0.25 M of Na₂C₂O₄ and 2.5·10^{–3} M of NaN₃) were mixed and then sonicated at 40 kHz for 10 min. After sonication, samples were vortexed at 2500 rpm for 5 min (Advanced Vortex Mixer ZX3, VELP Scientifica Srl, Italy), then centrifuged at 4700 g. for 15 min (ROTOFIX 32 A, Hettich, Italy). After centrifugation, an automated system for liquid handling (SamplePro Tube, Bruker BioSpin GmbH, Rheinstetten, Germany) transferred 630 µL of the

supernatant solutions into NMR tubes containing 70 μL of 0.20% of sodium salt of 3-trimethylsilyl-2,2,3,3-tetradeuteropropionic acid (TSP) solution in D_2O .

The one-dimensional ^1H NOESY pulse program was applied using a Bruker Avance I 400 MHz spectrometer equipped with an autosampler and a 5 mm inverse probe (Bruker BioSpin GmbH). The ^1H NOESY spectra were acquired using Topspin 3.0 software (Bruker BioSpin GmbH) under an automatic process that lasted around 22 min that included sample loading, 5 min temperature stabilization, tuning, matching, shimming. The free induction decays (FIDs) were Fourier transformed, the phase was manually corrected, the baseline was automatically corrected, and the spectra were shift referenced by setting the TSP singlet to 0 ppm.

Chemometric analysis

The FIDs of the 1D ^1H NOESY experiments, carried out on 55 aqueous extracts of olive leaves, were segmented into 237 buckets of 0.04 ppm intervals in the range of [10, 0.50] ppm using MestReNova 11.0 (Mestrelab Research SL, Santiago de Compostela, Spain). This is done with the aim of having a given peak in a bucket despite small spectral shifts between observations. The underlying area of each bucket was normalized to the total intensity. The 14 buckets in the region [5.18, 4.60] ppm, corresponding to the residual water signal, were set to 0. The data matrix was imported into SIMCA 16 software (Umetrics, Umea, Sweden) to carry out chemometric analyses. The NMR data matrix constituted 55 observations and 223 x-variables. Buckets were centred and subjected to unite variance scaling aka autoscaling. Initially, the unsupervised method of Principal Component Analysis (PCA) was performed to get an overview of the data. Then, the supervised Orthogonal Partial Least Square-Discriminant Analysis (OPLS-DA) was applied to identify variables that discriminate between observations, depending on a priori i.e. qPCR test, of the two classes of Xf-infected and non-infected, which count for 27 and 28 observations, respectively.

The relationship between NMR and HSR data was exploited by a correlation-based approach (Hall et al., 2022; Cavill et al., 2016). The pre-processed HSR spectra (see materials and methods) were bucketized to 141 buckets of 10 nm intervals. To obtain a homogenous sample size between both techniques, the 280 HSR samples were conformed to the 55 NMR samples by taking the median of each corresponding groups of HSR samples; around five HSR spectra for each NMR spectrum.

The calculation and visualisation of the correlation matrix were performed in RStudio IDE of R programming language using core packages as well as the tidyverse metapackage (R Core Team, 2021; RStudio Team, 2021; Wickham et al., 2019).

RESULTS AND DISCUSSION

Selection of asymptomatic leaves according to visual inspection

A number of 280 leaves were sampled that corresponded to 55 NMR samples (around five leaves per sample) belonging to the severity levels of 0 and 1 which were regarded as asymptomatic/having early symptoms of OQDS, respectively; By visual inspection, leaves were given an integer number scaled from 0 to 4 corresponding to disease severity of OQDS symptoms (Saponari et al., 2019).

For each of the Xf-infected leaves (146 out of 280), a qPCR test was done to molecularly check the presence of the bacterial cells in the plant. The qPCR test results were positive for 69 leaves. However, 77 leaves had a doubtful qPCR result, which might also be due to a false-negative qPCR test or very low bacterial cells in the leaf. Thus, each NMR sample from Xf-infected leaves were considered positive unless the majority of the leaves in that sample had a negative/doubtful qPCR result for Xf. Each Xf-infected sample had at least one qPCR positive leaf (Table 1).

Table 1 - Real-time qPCR results for Xf-infected samples:

	qPCR result	Frequency	Percentage (%)
Non-infected	Not-tested	27 (134)	49.1
Xf-infected (28 NMR samples; 146 leaves)	Positive	12 (69 leaves)	21.8
	Doubtful	16 (77 leaves)	29.1
	Total	55 (280 leaves)	100

Metabolic profile from NMR spectral analysis

The ¹H-NMR spectra considered, excluding the region containing the residual solvent signal, had signals related to oleuropein, oleuropein aglycone, malic acid,

mannitol, glucose, aromatic compounds, choline, alanine, tyrosol, and formic acid (Fig. 2). The identified metabolites were similar to those previously reported in the literature (Jililat et al., 2021; Girelli et al., 2017).

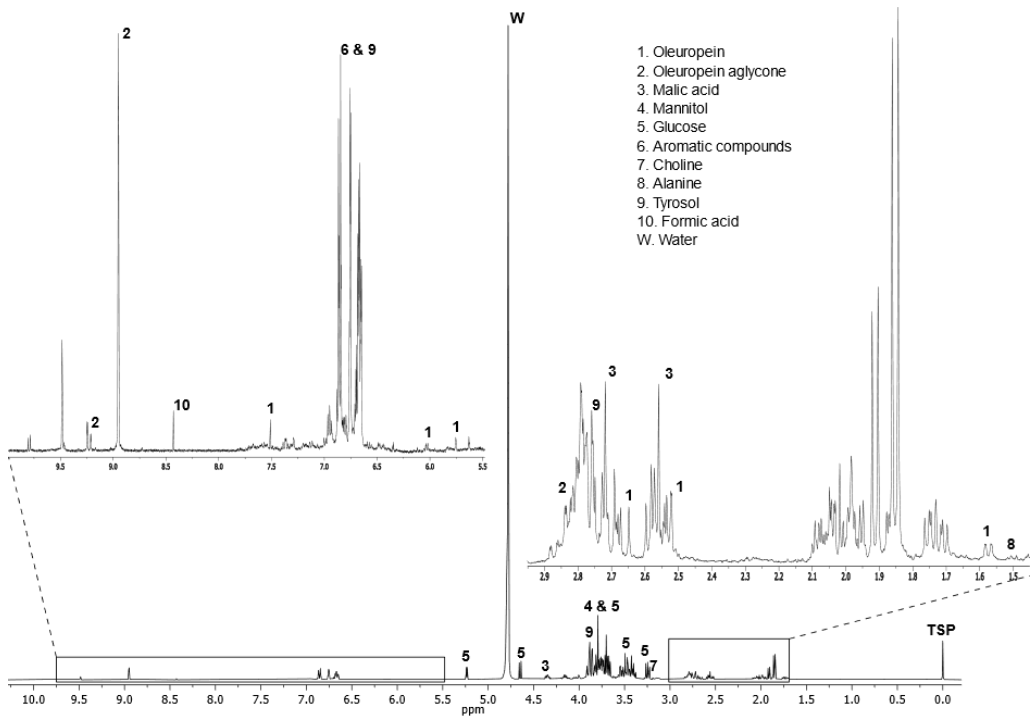


Fig. 2 - ^1H NOESY NMR spectrum of a water extract obtained from olive leaves sample. “W” refers to the residual water signal.

Chemometric analysis

To assess the quality and homogeneity of the data, Principal Component Analysis (PCA) was performed. PCA was carried out and indicated an explained variance (R^2) of 0.76 for 8 Principal Components [PCs] (Fig. 3). The first two components, PC1 and PC2, accounted for about 0.38 R^2 . Nonetheless, there was no clear trend related to Xf-infection and the same remark was also observed for the other PCs (Fig. 4).

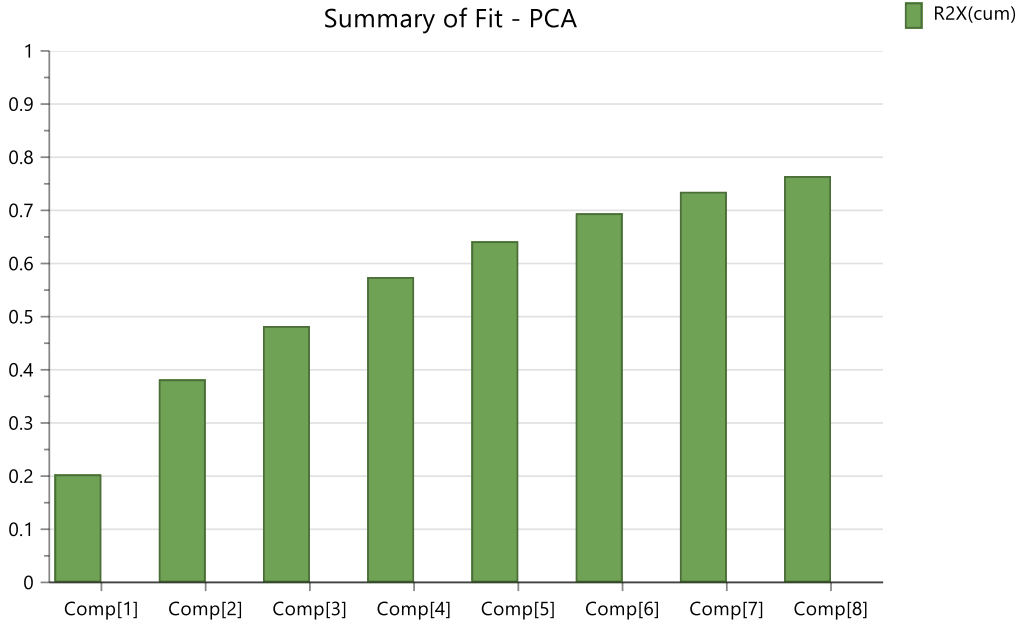


Fig. 3 - Summary of explained variance (R^2) cumulated for each principal component (Comp[n]) of the Principal Component Analysis (PCA) model generated.

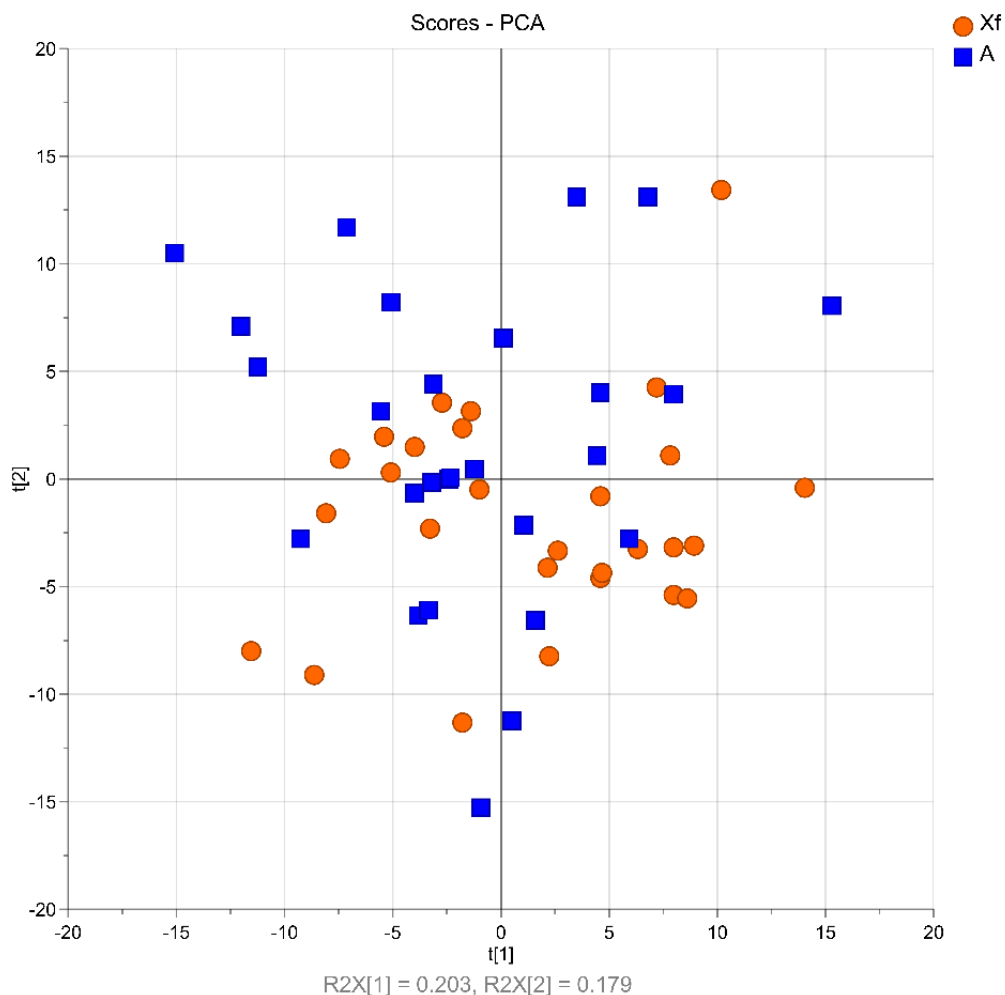


Fig. 4. Principal component analysis (PCA) of the NMR samples coloured according to *Xylella fastidiosa* infection; t[1]/t[2] scores plot relating to PC1/PC2. Xf, *Xylella*-infected; A, non-infected by *X. fastidiosa*.

To assess the differences in the NMR profile between related to Xf-infection, the orthogonal projection to latent structures discriminant analysis (OPLS-DA) was applied to the bucketed NMR spectra. The OPLS-DA model was created with a 1 predictive (P1) and 2 orthogonal (O1-O2) components (1+2+0). The O1

explained 17.7% of the systematic information in x-space orthogonal to y-space ($R^2X[O1] = 0.177$), O2 explained 11.7% ($R^2X[O2] = 0.117$, and together O1-O2 explained 28.9% having $R^2X[cum] = 0.289$. Furthermore, P1 explained 7.45% of x-variance ($R^2X = 0.0745$) and modelled 79.3% of y-variance ($R^2Y = 0.793$). In total, an x-variance of 28.9% was explained by O1–O2 ($R^2X[cum] = 0.289$).

Subsequently, the generated OPLS-DA model was validated by both permutation-test and Cross Validation ANalysis Of VAriance (CV-ANOVA). The 999-permutations showed a y-intercept of < 0.5 for R^2 and < -0.56 for Q^2 considering both classes (Fig. 5 and Fig. 6) and the CV-ANOVA test had a computed p-value of $8.91 \cdot 10^{-9}$.

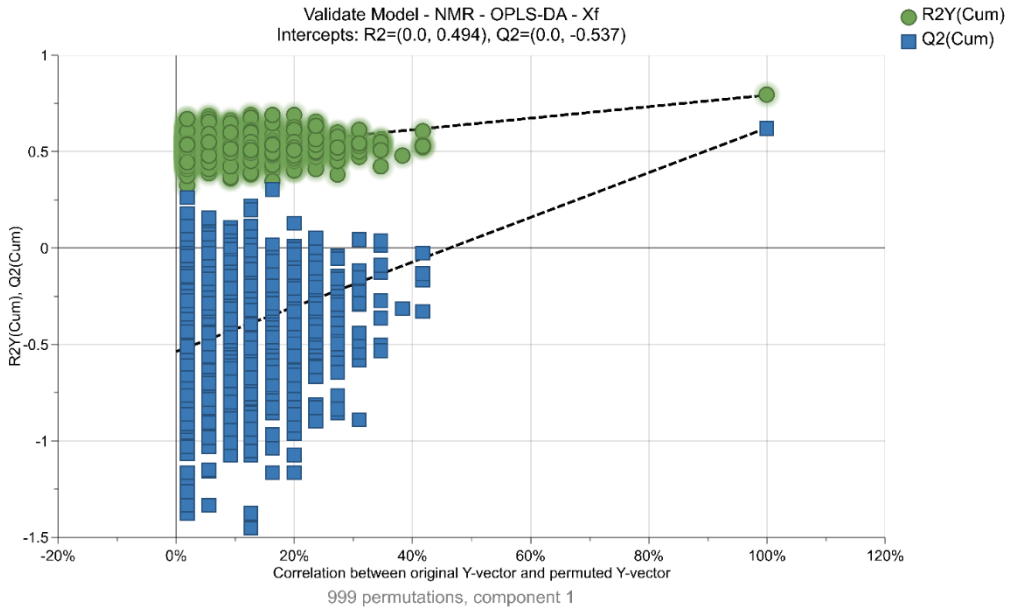


Fig. 5 - Permutation test for validation of the OPLS-DA model using 999 permutation for Xf-infected class (Xf).

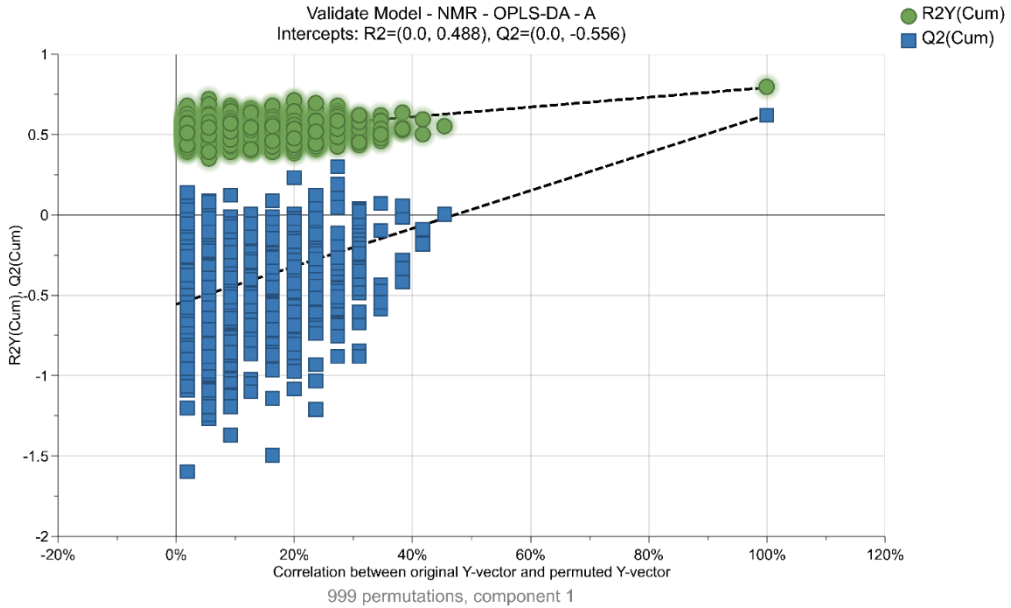


Fig. 6 - Permutation test for validation of the OPLS-DA model using 999 permutation for non Xf-infected class (A).

The P1 versus O1 ($t[1]$ vs. $t_0[1]$) scores plot showed a distribution of the observations in which they were separated along the P1 with one Xf-infected observation overlapping with a non-infected and another two lying in the area of non-infected class along P1 (Fig. 7). In fact, those three samples belonged to the doubtful positive Xf-infection according to the qPCR test (Fig. 8). Thus, they probably have a metabolic profile more similar to the non-infected samples.

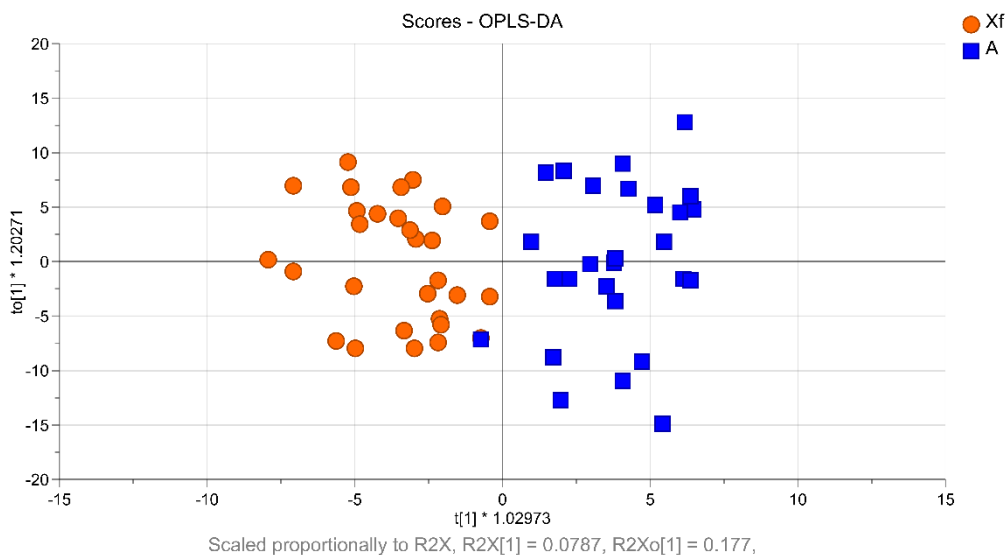


Fig. 7 - Orthogonal partial least squares discriminant analysis (OPLS-DA) score plot of the NMR data. Xf: samples from *X. fastidiosa* inoculated plants; A: samples from non-inoculated plants). The $t(1)/t_o(1)$ represent the first predictive and orthogonal components P1/O1, respectively.

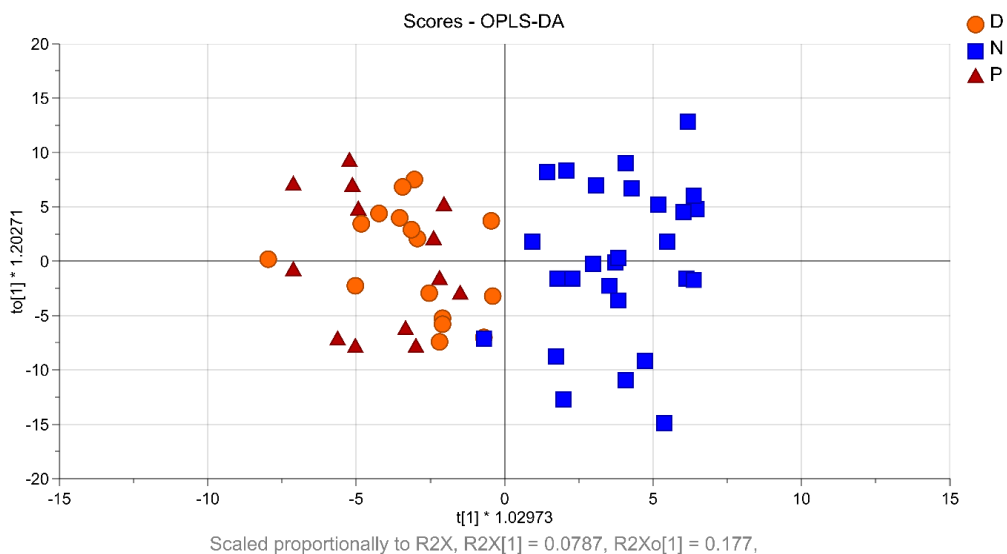


Fig. 8 - Orthogonal partial least squares discriminant analysis (OPLS-DA) scores plot of the NMR data. P: samples from *X. fastidiosa* inoculated plants with positive molecular test for Xf; N: samples from non-inoculated plants; and D, samples from *X. fastidiosa* inoculated plants with

doubtful positive molecular test for Xf). The $t(1)/t_o(1)$ represent the first predictive and orthogonal components P1/O1, respectively.

To identify the potential diagnostic NMR signals that were differently expressed between Xf-infected (Xf) and non-infected (A) samples from NMR bucketised data, both the S-plot and the predictive Variable Importance in the Projection (VIP) values and were inspected (Fig. 9).

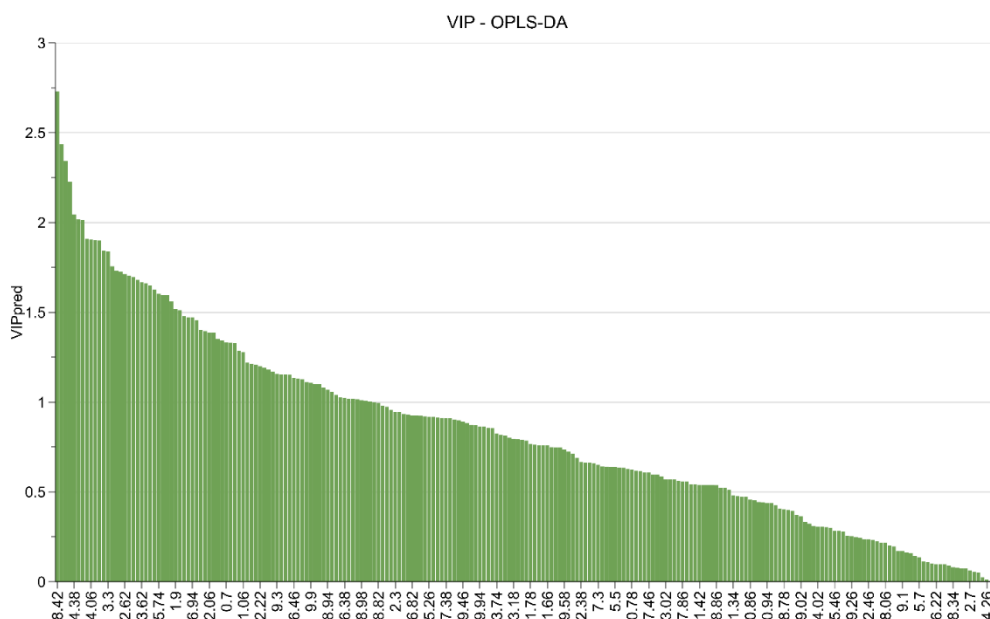


Fig. 9 - Variable Importance in the Projection (VIP) predictive values ordered from highest to lowest of the orthogonal partial least squares discriminant analysis (OPLS-DA) model presented in Fig. 7.

The S-plot displays the $p[1]$ versus $p_{(corr)}[1]$ vectors of the predictive component, where $p[1]$ is the loading vector that expresses the weight of each x-variable on the selected component P1, and $p_{(corr)}[1]$ is $p[1]$ scaled from -1 to 1 as a correlation coefficient between each x-variable and $t[1]$. In the S-plot, the x-

variables situated far out on the sides of the line do combine a high model influence with high reliability and are relevant in the search for up- or down-regulated signals (Fig. 10).

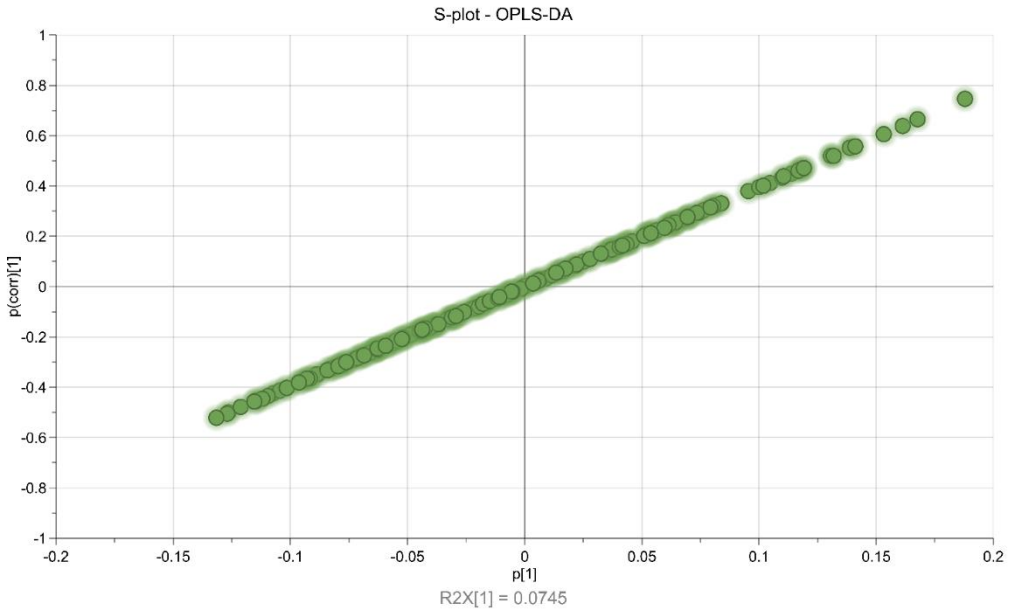


Fig. 10 - S-plot of NMR data from the OPLS-DA model (Fig. 1). The values shown are for $p(1)$ and $p_{corr}(1)$ of to the predictive component P1.

The VIP predictive values indicate the importance of each x-variable on the predictive part of the OPLS-DA model. Values larger than 1 are more relevant for explaining the y-response. The variables related to signals with the highest potential as diagnostic NMR signals were selected among those with $|p_{(corr)}| > 0.5$ and $VIP_{predictive} > 1$ (Fig. 11).

These signals were related formic acid, malic acid, lactic acid, fructose, glucose, choline, and tyrosol, as well as other aromatic compounds. On one hand, NMR signals related to fructose, glucose, and choline were positively correlated in relation to Xf-infection. On the other hand, signals of formic acid, malic acid, Lactic acid, Tyrosol/DHBA were negatively correlated to Xf-infection.

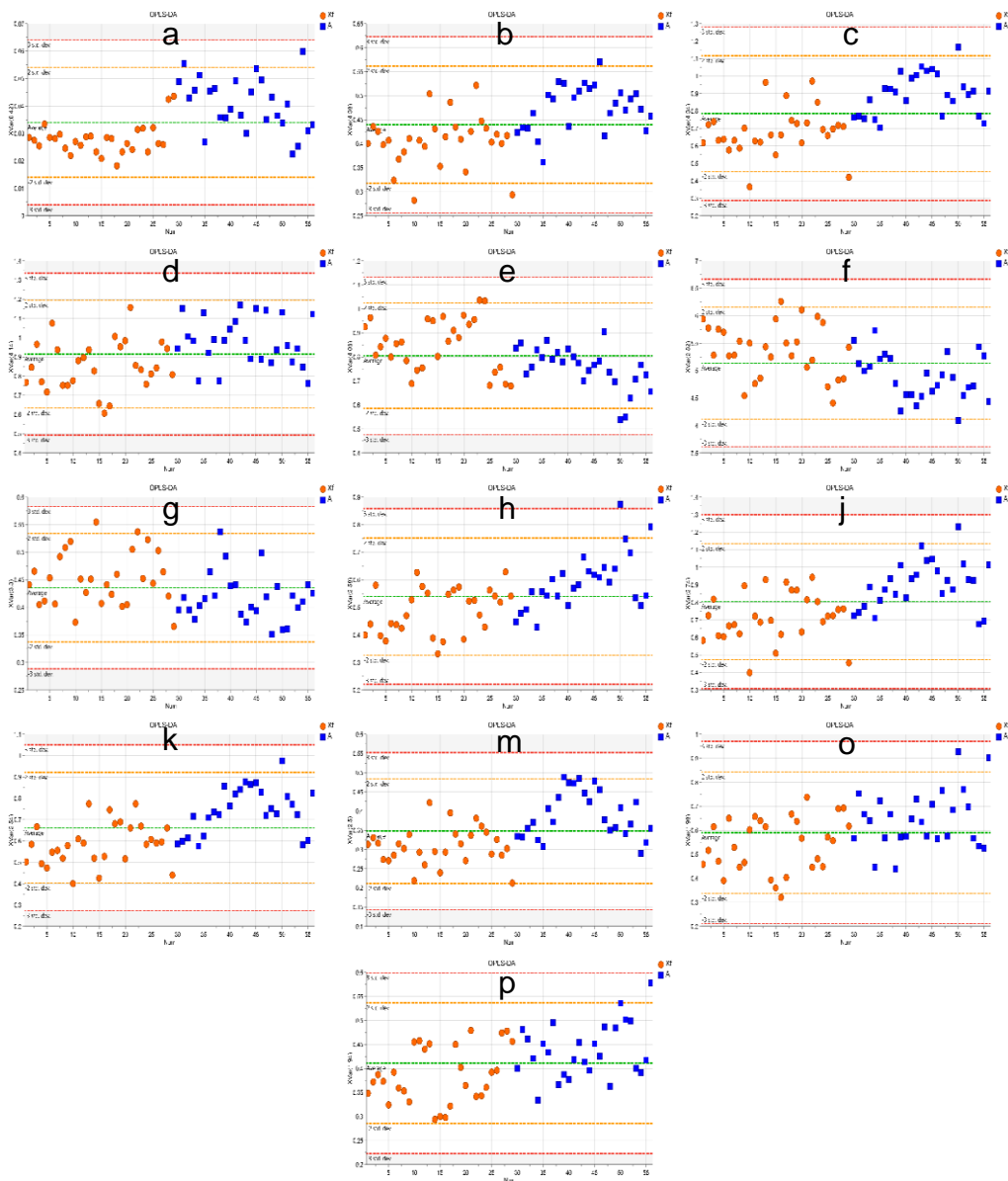


Fig. 11 - Variables trend plot representing buckets (of 0.04 ppm width) from the OPLS-DA model of $VIP_{\text{predictive}} > 1$ and $|p_{\text{corr}}| > 0.5$. The variables represent the buckets at: a, 8.42; b, 4.38; c, 4.34; d, 4.14; e, 4.06; f, 3.82; g, 3.30; h, 2.74; j, 2.58, k, 2.54, m, 2.50; o, 1.98; and p, 1.94 ppm. The letters i and l were not used to eliminate confusion. The compounds related to these signals are: a, Formic acid; b, c, j, k, and m, malic acid; d, lactic acid; e, fructose; f, glucose; f, galactose; g, choline; and j and k, citric acid.

Of interest, the relative decrease of malic acid is often associated salt and drought stress affected trees. Malic acid is a key metabolite in The Citric Acid (TCA)/Krebs cycle and was also associated to fatty acid synthesis (Skodra et al., 2021). Probably a plant reaction to the decrease of water supply as a result of Xf infection (Girelli et al., 2019). Fructose, glucose, and sucrose were also reported to play an important role in Xf infection. Coherently, higher amount of fructose was associated to Xf infection (De Pascali et al., 2022). Moreover, the decrease in sucrose and glucose was also reported in relation to Xf infection (De Pascali et al., 2022). Such decrease in sucrose and glucose, might be due the break-down of sucrose for the sake of consuming glucose but not fructose, hence the increase in fructose levels.

HSR analysis

The HSR spectra for both Xf-infected and non-infected leaves had an overall similar spectral shape with no particular pattern related to the Xf-infection overall. This is probably due to the asymptomatic nature of the studied leaves. In contrast to typical symptoms of late stage of OQDS, asymptomatic leaves show no apparent change neither in the colour nor the water content of the leaves, which appears in the visual (400-800) and NIR (800-1830) wavelengths regions, respectively (Fig. 12).

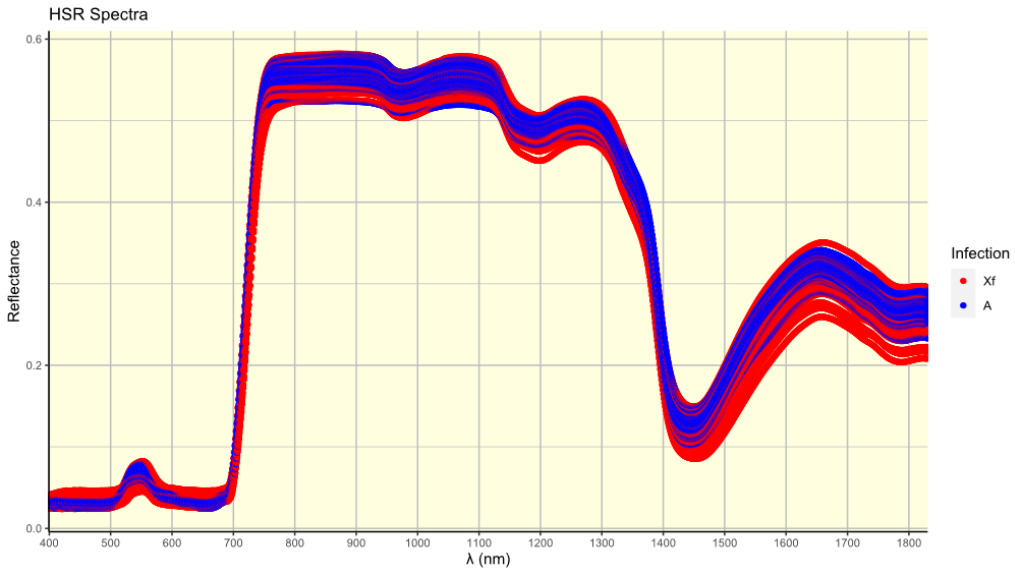


Fig. 12 - Hyperspectral reflectance (HSR) spectra of the analysed samples as bucketized at 10 nm intervals. The original measured wavelength region ranges from 400 to 1830 nm with a spectral resolution of 1 nm. Red lines represent HSR of non-infected samples (A), and blue lines represent Xf-infected samples (Xf).

Correlation of NMR diagnostics signals to HSR

To investigate the correlation between the metabolic profile (NMR) and phenotypic profile (HSR) of the olive leaves related to Xf-infection. HSR profile of the leaves together with the NMR profile were subjected to Pearson's correlation analysis using the *cor.test* function from the package *{stats}* in R programming language (R Core Team, 2021), and correlation matrices were obtained.

On one hand, if there was an association between NMR and HSR profiles, the correlation matrix would show a strong negative or positive correlation (slope of the line), by a maximum of $r = -1$ and $r = 1$, respectively. On the other hand, if there was no relationship, the correlation would exert values near zero (Fine et al., 2021). Importantly, the p-value determines how well this slope fits the data points where a p-value of <0.05 is considered to be significant.

Generally, diverse correlations between NMR and HSR profiles were observed for both Xf-infected and non-infected samples. On one hand, the

correlation for Xf-infected samples seemed to appear in six regions around i) 400-520, ii) 520-590, iii) 590-700, iv) 700-1150, v) 1150-1330, and vi) 1330-1830 nm (Fig. 13).

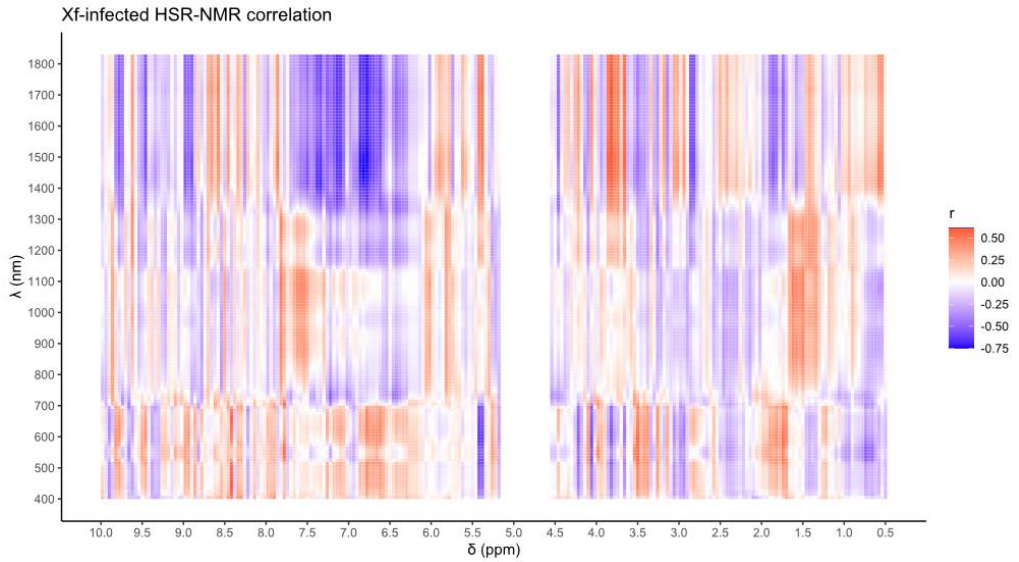


Fig. 13 - Correlation of hyperspectral reflectance (HSR) and NMR for Xf-infected samples.

On the other hand, the correlation for non-infected samples seemed to appear in six regions as well, viz i) 400-520, ii) 520-590, iii) 590-690, iv) 690-750, v) 750-1350, vi) 1350-1830 (Fig. 14).

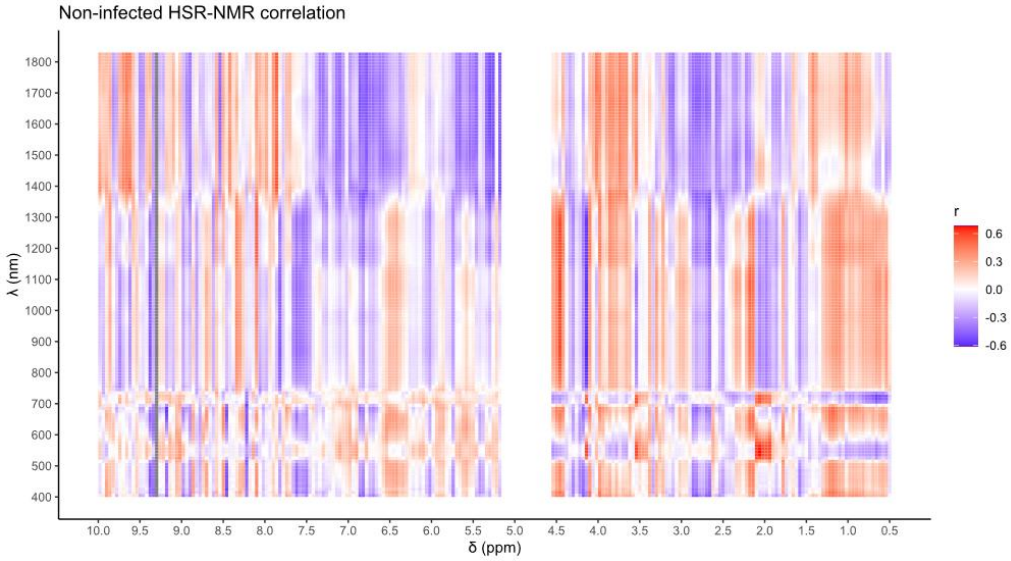


Fig. 14. Correlation of hyperspectral reflectance (HSR) and NMR for non-infected samples.

Furthermore, to identify the NMR signals correlating to HSR wavelengths in relation to Xf-infection, the Xf-infected correlation matrix was compared to the non-infected matrix using the Z-Fisher's method (Fisher, 1921, 1915). The Z-Fisher method first computes z scores from the correlation coefficients and then compares these z scores using the following equation: $z_1 = (1/2) \times \log((1+r_1) / (1-r_1))$; where n_1 and r_1 are the sample size and the correlation coefficient, respectively. Then, a test's statistic is computed from z_1 and z_2 : $Z = (z_1 - z_2) / \sqrt{(1 / (n_1 - 3)) + (1 / (n_2 - 3))}$; If Z is above the limit given by the alpha value, in our case 0.05, then the difference between r_1 and r_2 is significant. In principal, this function compares two correlation coefficients obtained from different sample sizes using Z-Fisher transformation (Savary et al., 2021). Then, non-significant correlations for both matrices, i.e. $P < 0.05$, were omitted. The resulted HSR-NMR significant correlations difference showed HSR bands of 10 nm at 405, 765-1165, 1215-1305, 1325, 1335, 1445, and 1455 nm related to NMR buckets of 0.04 ppm at 8.46, 7.82, 7.62, 7.58, 7.54, 6.82, 4.46, 4.1, and 3.54 ppm (Fig. 15).

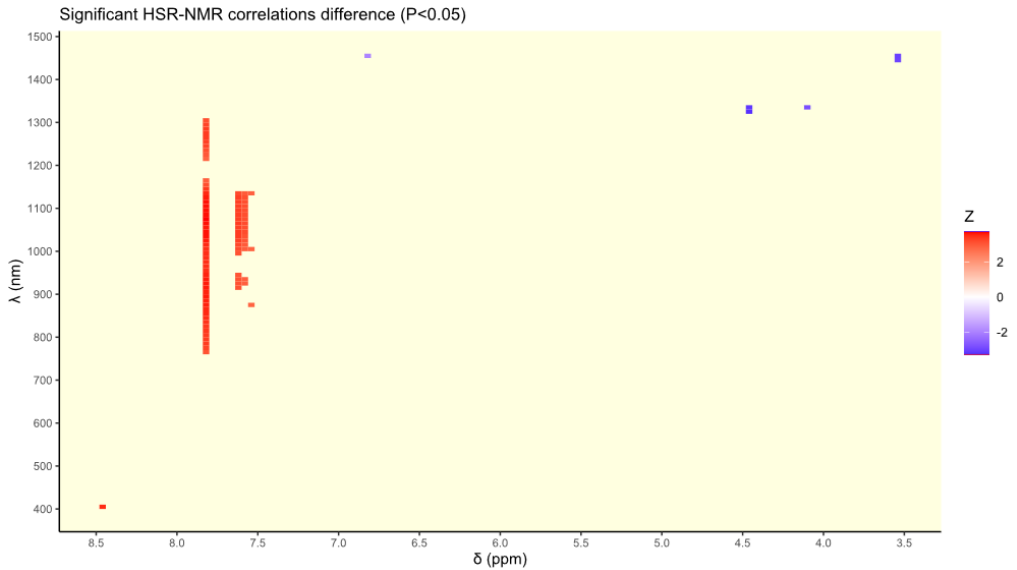


Fig. 15 - The difference between correlation of hyperspectral reflectance (HSR) and NMR profiles related to *X. fastidiosa* infection. Red indicates more positive correlation differences related to Xf-infection, while blue indicates a negative correlation differences related to Xf-infection.

The NMR buckets with significant correlation with the HSR bands were related to spectral noise, aromatic compounds at 6.82 ppm, fructose at 4.10 ppm, and glucose at 3.54 ppm (Fig. 16). Moreover, HSR bands with significant correlation with the NMR buckets were related to the bands of 1335, 1445, and 1455 nm (Fig. 17). These wavelength bands reside at the end of NIR and the beginning of SWIR. Importantly, the wavelength bands around 1400 nm are associated with water absorption which, interestingly, plays a major role in OQDS (Danson et al., 1992). In which, Xf grows and multiplies within the xylem by forming a biofilm that, over time, hinders the regular water movement through the xylem vessels (Scortichini, 2020).

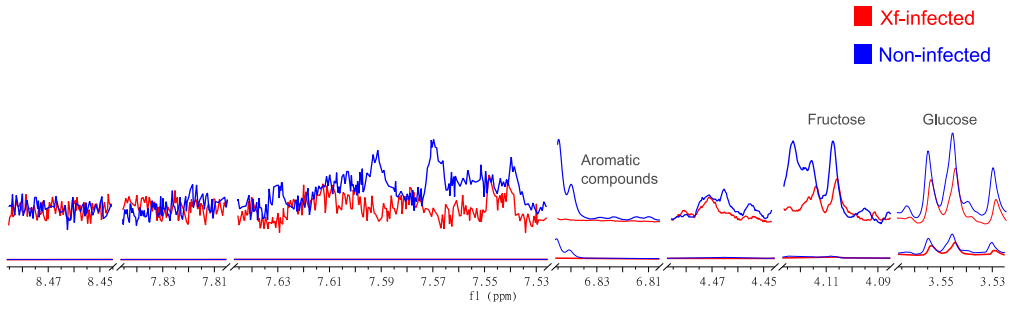


Fig. 16 - NMR Buckets with significant correlation to spectral reflectance.

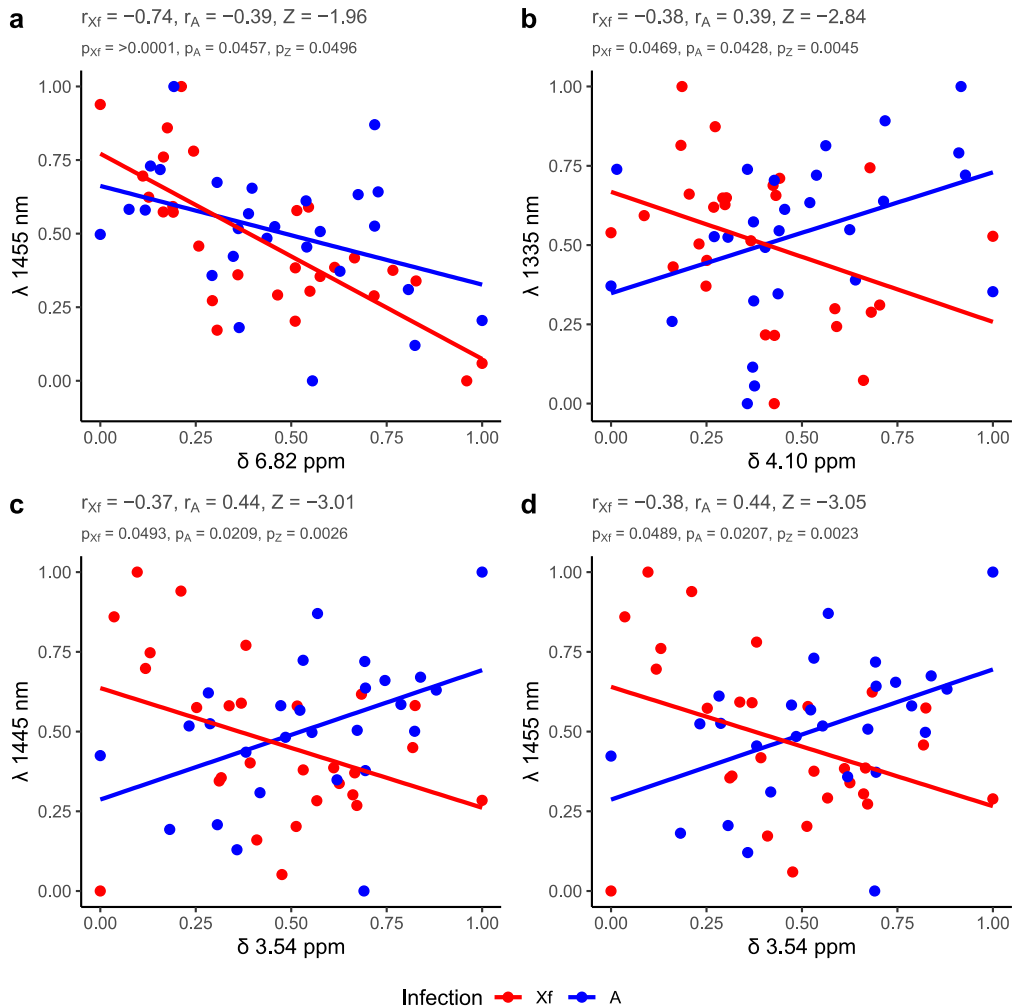


Fig. 17 - Scatter plots of significant correlations ($p < 0.05$) between NMR buckets (0.04 ppm, x-axis) and HSR buckets (10nm, y-axis) of significant differences ($p < 0.05$) between Xf-infected (Xf) and non-infected (A) samples; a, 4.10 ppm – 1335 nm; b, 3.54 ppm – 1445 nm; c, 3.54 ppm – 1455 nm; and d, 6.82 ppm – 1455 nm. r_{Xf} and r_A are the Pearson’s correlation coefficients for Xf-infected and non-infected samples respectively. Z is Fisher’s Z-score comparing both coefficients of correlation for Xf and A.

Conclusions

In conclusion, the analysis of the cultivated asymptomatic olive leaves, infected with *Xylella fastidiosa* subsp. *pauca* ST53 responsible for OQDS, using ^1H NMR, HSR, and chemometrics allowed for selecting diagnostic signals and wavelengths related to OQDS. In specific, the correlation matrices between NMR and HSR showed a significant relationship between diagnostic NMR signals to HSR features. This relationship revealed different wavelength regions with diverse associations with the corresponding metabolites. The determination of diagnostic wavelength bands associated with certain metabolites is an important step in the development of sensors capable of early-detection of Xf-infection in olive trees.

CHAPTER 7

Non-Targeted Spectranomics for Early-Detection of Xylella fastidiosa-Infected Asymptomatic Olive Leaves in Field Conditions

In this chapter, part of the work done in collaboration with the Big Data Cybernetics group at the Norwegian University of Science and Technology (NTNU) is presented. This collaboration was part of the training-abroad period of the PhD candidate in Trondheim, Norway. The period lasted less than 4 months from the beginning of September until half of December of the year 2022.

INTRODUCTION

Xylella fastidiosa subsp. *pauca* sequence type ST53 (Xf), also known as the “De Donno” strain, impacted severely the olive groves of the Apulia region in southern Italy by causing a disease known as olive quick decline syndrome [OQDS] (Saponari et al., 2013; Martelli et al., 2016; Saponari et al., 2019). A projection of such impact for the next 50 years is estimated to be more than 5 billion euros (Schneider et al., 2020; Godefroid et al., 2020). Since the main control strategy adopted by EU is to destroy the infected plants, early detection of infected trees remains a key advantage against the spread of the disease for the applying preventive measure in advance (Saponari et al., 2016; Blekos et al., 2021). As spectroscopic and non-destructive techniques. hyperspectral reflectance (HSR) and nuclear magnetic resonance spectroscopy (NMR) may offer means to detect the change at the phenotypic and metabolic level (Wishart et al., 2022a; Tanner et al., 2022). In this study, the samples of olive leaves from the susceptible variety Cellina di Nardò were collected from the field. These asymptomatic leaves were analysed using ^1H NMR, HSR, and chemometrics to look for the potential difference at the early stage of disease development.

MATERIALS AND METHODS

Sample collection from olive trees

Olive leaves (*Olea europaea* L) were collected from Olive trees in the open field of the Salento Peninsula. All olive trees belonged to the cultivar Cellina di Nardò which is known to be very susceptible to *Xylella fastidiosa* subsp. *pauca* ST53. More than 1000 leaves were collected over three time periods, of August, December, and May, to represent different seasonal variations as shown in Table. 2.

Table 2 - Frequency table of collected Olive leaf samples considering the different time points of sampling:

Time	Frequency	Percent	Valid Percent	Cumulative Percent
T0	398	39.641	39.641	39.641
T1	403	40.139	40.139	79.781
T2	203	20.219	20.219	100.000
Missing	0	0.000		
Total	1004	100.000		

According to the qPCR test, as detailed in the previous chapter, samples were characterized as negative (N) and positive (P), as well as doubtful (D) (Table 3).

Table 3 - Frequency table according to the qPCR test:

qPCR	Frequency	Percent	Valid Percent	Cumulative Percent
D	49	4.880	4.880	4.880
N	676	67.331	67.331	72.211
P	279	27.789	27.789	100.000
Missing	0	0.000		
Total	1004	100.000		

Leaves were first subjected to the acquisition of hyperspectral reflectance (HSR) by an HSR acquisition system that consisted of a FieldSpec®3 spectroradiometer (Analytical Spectral Device [ASD], Boulder, CO, USA) coupled to a leaf probe (ASD) and a leaf clip holder (ASD) by an optical fibre cable, in addition to a laptop to show and store the data acquired. The HSR data were in the range of 400-1830 nm with a final spectral resolution of 1 nm.

After HSR acquisition, the collected olive leaves were lyophilized at -45°C under 0.180 mbar for 48 h in Christ Alpha 1-4 LSC lyophiliser (Martin-Christ Gefriertrocknungsanlagen GmbH, Osterode am Harz, Germany). The dried samples were then ground by mortar and pestle, sieved through a mesh of 0.5 mm pores, and stored at -20°C . Around 10 leaves corresponded to each NMR sample.

For each NMR sample, an amount of 50 mg of olive leaf powder and 1.5 mL of oxalate buffer at pH 4.2 (pH value was reached after the addition of 37% HCl to 100 mL an aqueous solution containing 0.25 M of $\text{Na}_2\text{C}_2\text{O}_4$ and $2.5 \cdot 10^{-3}$ M of NaN_3) were mixed and then sonicated at 40 kHz for 10 min followed by vortexing at 2500 rpm for 5 min (Advanced Vortex Mixer ZX3, VELP Scientifica Srl, Italy), then centrifuged at 4700 g. for 15 min (ROTOFIX 32 A, Hettich, Italy). After centrifugation, an automated system for liquid handling (SamplePro Tube, Bruker BioSpin GmbH, Rheinstetten, Germany) transferred 630 μL of the supernatant solutions into NMR tubes containing 70 μL of a sodium salt of 3-trimethylsilyl-

2,2,3,3-tetradeuteropropionic acid (TSP- d_4) solution in D_2O (0.20 %_{wt}). The 1D 1H NOESY spectra, recorded on 120 aqueous extracts of olive leaves, were segmented into buckets of 0.04 ppm intervals in the range of [10, 0.50] ppm using MestReNova 11.0 (Mestrelab Research SL, Santiago de Compostela, Spain). The underlying area of each bucket was normalized to the total intensity. The buckets in the region [5.10, 4.60] ppm that corresponds to the residual water signal were set to 0.

For chemometrics, all calculations and visualizations were performed in RStudio IDE of R programming language using core packages as well as other packages that include ggfortify, ggpubr, reshape2, prospectr, kernlab, randomForest, e1071, caTools, MASS, klaR, forecast, caret, and the tidyverse metapackage (R Core Team, 2021; RStudio Team, 2021; Wickham et al., 2019).

Considering the HSR data, a particular splice around 1000 nm occurred due to the difference between the mode of the few different sensors in the spectroradiometer that cover different wavelength ranges when used in field conditions; this did not occur during HSR acquisition in chapter 6. Then, the hyperspectral reflection spectra were transformed to absorbance to obtain a linear unit and were further pre-processed. The raw absorbance spectra were pre-processed using Standard Normal Variate (SNV), mean-centring, and autoscaling. Pre-processing is a crucial step in chemometrics as it could improve and also cause noise-inflation. The pre-processing processes were used to reduce the effect of less-important variations to obtain more accurate models.

Considering NMR data, the data matrix was imported into R studio for chemometric analyses as well. Buckets were mean-centred and auto-scaled (unit variance), where each variable is scaled to the standard deviation: $[(value - mean)/st.dev]$, to give each variable a standard deviation of one, and therefore the data is analysed based on correlations instead of covariance.

Initially, the unsupervised method of Principal Component Analysis (PCA) was performed to get an overview of the data. The HSR spectra were acquired for 1004 leaves that accounted for 120 NMR samples. From the 1004 leaves, 676

were found negative for *Xylella fastidiosa* and 328 leaves were found infected including 49 that had doubtful results for the qPCR test. Then, supervised machine learning techniques were applied.

RESULTS AND DISCUSSION

In this chapter, the chemometric (non-targeted) approach to analyse the data was adopted (Wishart, 2008b; Nichani et al., 2021). The features are not initially identified and only their spectral patterns and intensities are recorded, statistically compared, and further used for the identification of the relevant spectral features that distinguish sample class membership.

Plant diseases combine physiological and visual alterations of the host plants that later appear as symptoms that become clearer as the disease progresses. Nevertheless, the visual alterations are rarely noticed at the very early stage of the disease, which is known to be asymptomatic to the disease. Therefore, the metabolic and spectral profiles of both asymptomatic and healthy trees would be very similar and very hard to distinguish between them. Nevertheless, such metabolic and spectral profiles could offer, to a certain degree, an advantage by retaining data that are untangled by the human eye. Thus, potentially providing a way to early detect plant diseases.

For Hyperspectral data, a splice at the wavelength of 1000 nm was detected using the 2nd derivative of the original spectra. The splice was corrected using the “spliceCorrection” function from the “prospectr” package. This function corrected the ASD FieldSpec spectroradiometer (Malvern Panalytical) which usually exhibits steps at the splice of the built-in detectors, positioned at 1000 nm [end of VNIR detector] (Fig. 18).

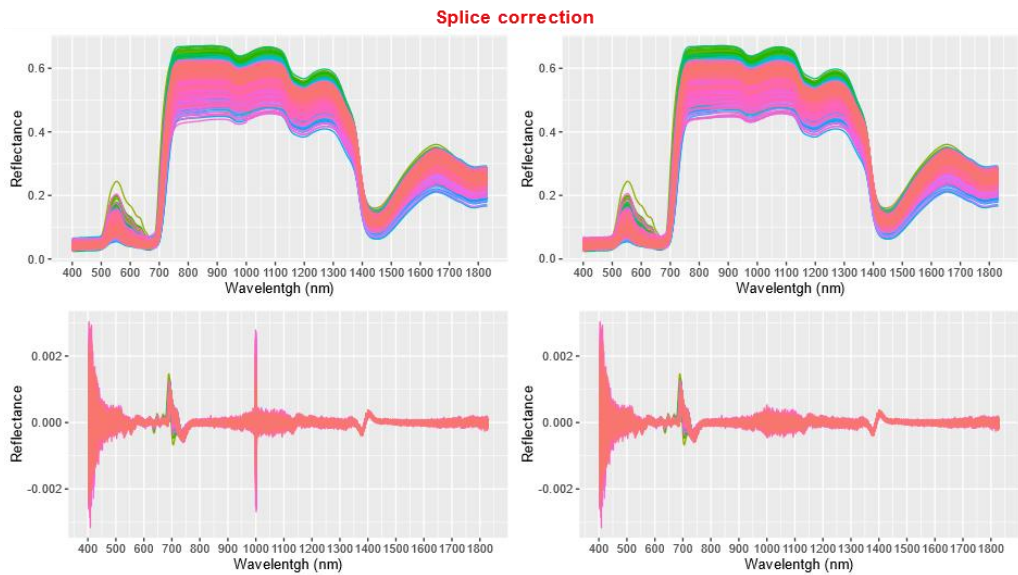


Fig. 18 - Splice detection and correction at 1000 nm. Top left, spectra with splice at 1000 nm; bottom left, 2nd derivative of top left; top right, splice-corrected spectra; bottom right, 2nd derivative of top right. Colours represents the different observations.

Furthermore, after correcting the splice in the spectral reflectance, the reflectance spectra were transformed to absorbance and then subjected to SNV pre-processing to decrease the additive and multiplicative baseline variation without altering the shape of the spectra as shown for the whole set of collected spectra at the general level as well as for the average spectrum of each class considering the qPCR test as a class discrimination criterion (Fig. 19 and Fig. 20).

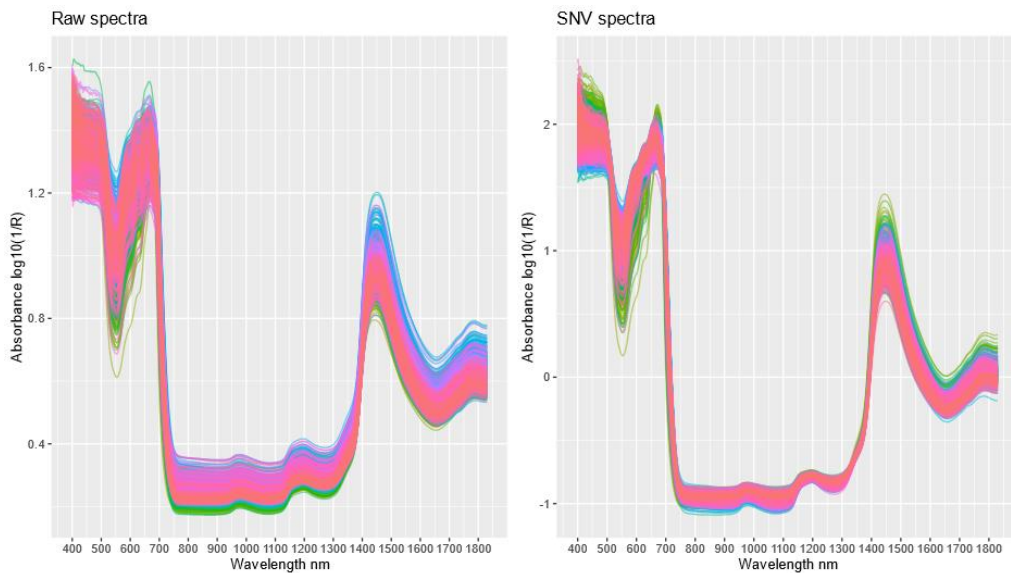


Fig. 19 - Absorbance spectra (right) and SNV pre-processed absorbance spectra (left)

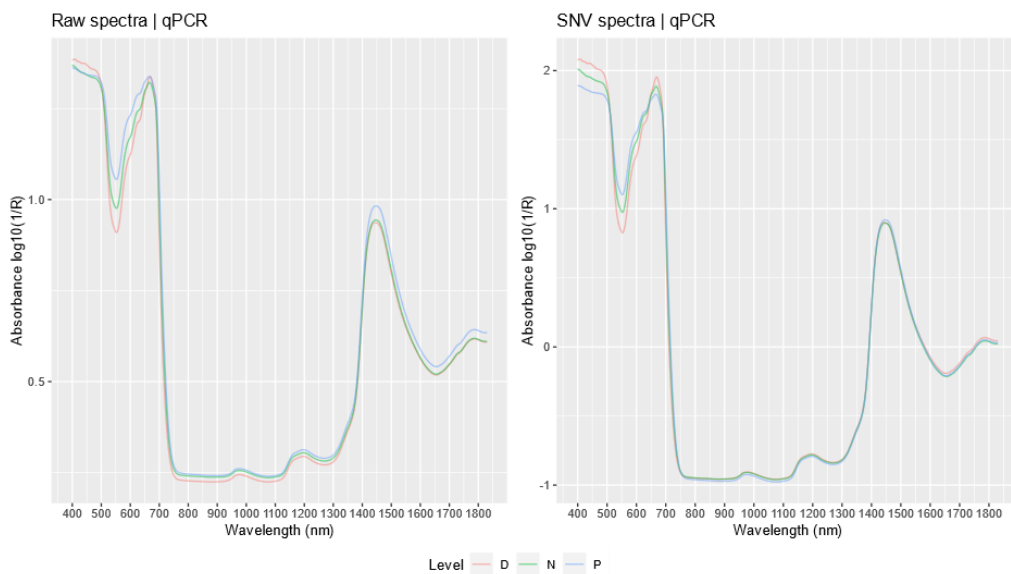


Fig. 20 - effect of SNV on the groups of belonging to different classes according to the qPCR test for *Xylella fastidiosa*.

After pre-processing with SNV, PCA was applied using mean-centring as well as autoscaling to gain an overview of the main variations amongst the samples. Considering the first five principal components, the cumulative variance explained reached 0.98 with first component explaining 69% of the variations, and the second principal component explained 25%, thus amounting for 94% of the variations amongst the samples (Fig. 21).

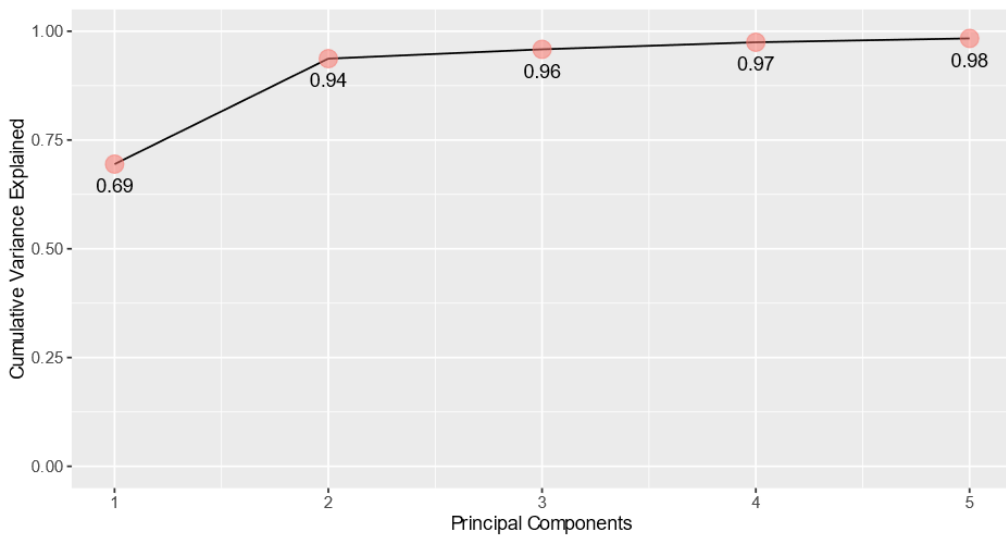


Fig. 21 - Cumulative variance plot showing the total variance explained by the new latent variables (Principal components; PCs) from PC1 to PC5 for HSR samples.

The different principal components were investigated for any particular pattern related to the qPCR test for *Xylella fastidiosa*. This was done by plotting each of the first five principal component against each other using their scores matrices to visualise the spreading of samples as well as using the loadings matrices to visualise the influence of the variables on the corresponding principal components. Nevertheless, there was no specific particular clustering amongst the samples was evident in relationship to the qPCR test results. This is probably due to the very little variations in the spectral profile of the collected healthy and asymptomatic leaf samples (Owomugisha et al., 2020; Martinelli et al., 2015). The

investigated combinations of principal components scores and loadings are shown in Fig 22-31.

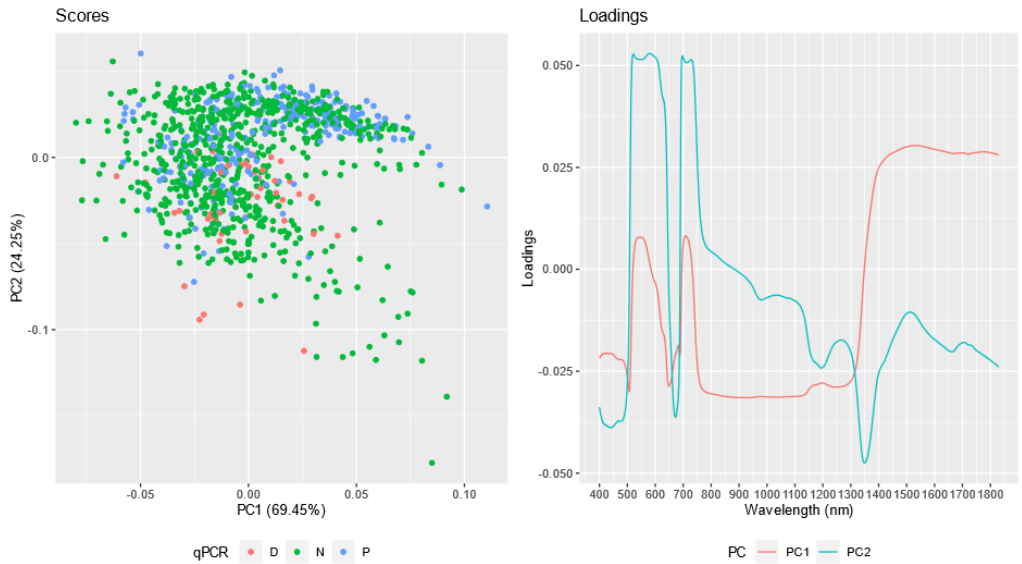


Fig. 22 - Principal component analysis (PCA) of the hyperspectral data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC1/PC2. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

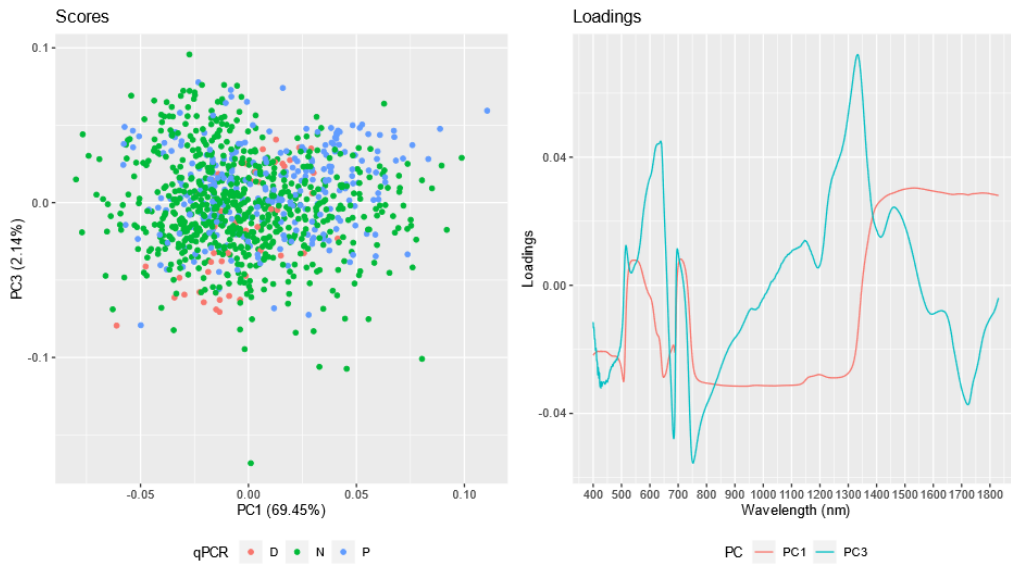


Fig. 23 - Principal component analysis (PCA) of the hyperspectral data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC1/PC3. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

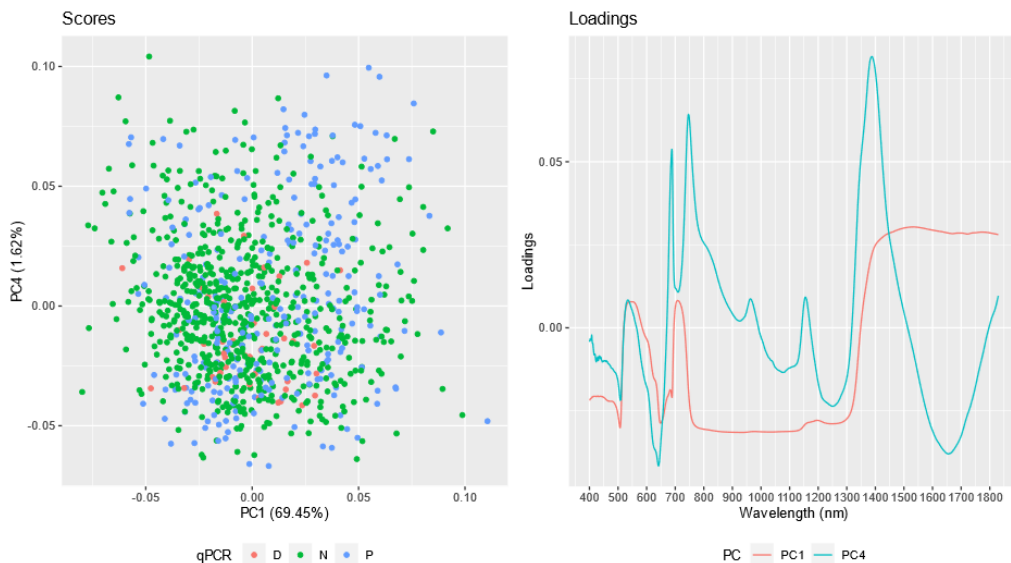


Fig. 24 - Principal component analysis (PCA) of the hyperspectral data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC1/PC4. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

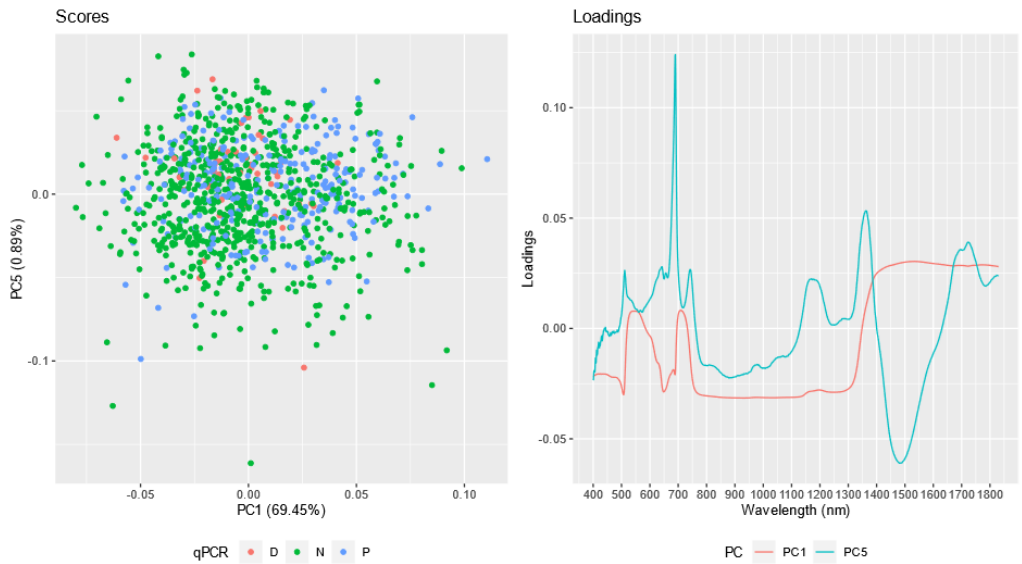


Fig. 25 - Principal component analysis (PCA) of the hyperspectral data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC1/PC5. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

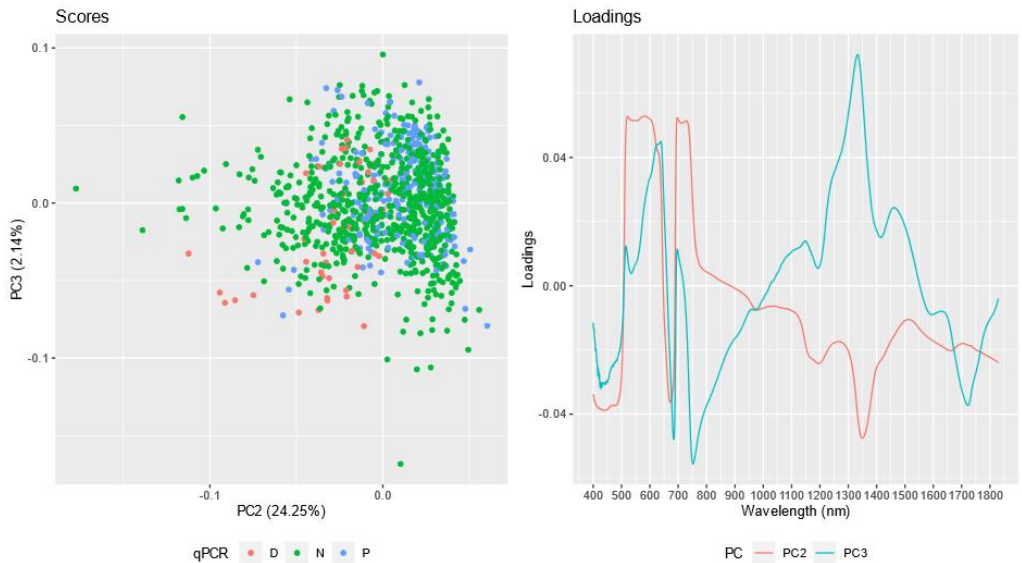


Fig. 26 - Principal component analysis (PCA) of the hyperspectral data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC2/PC3. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

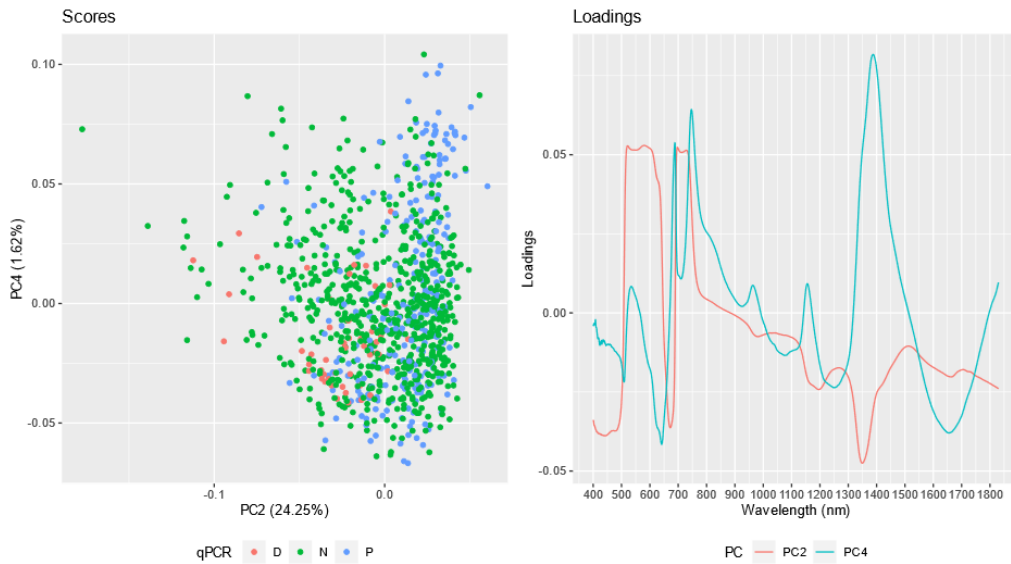


Fig. 27 - Principal component analysis (PCA) of the hyperspectral data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC2/PC4. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

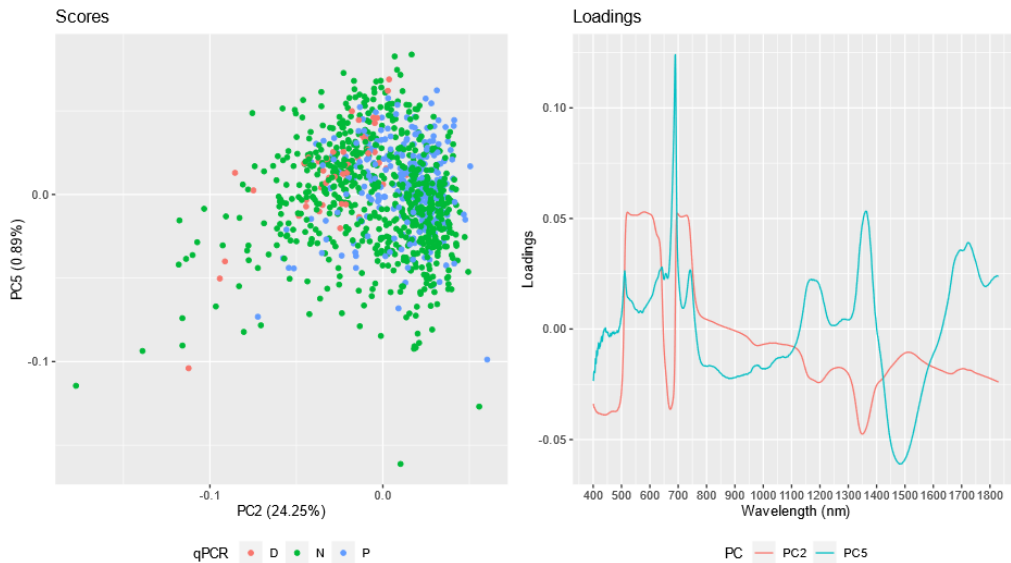


Fig. 28 - Principal component analysis (PCA) of the hyperspectral data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC2/PC5. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

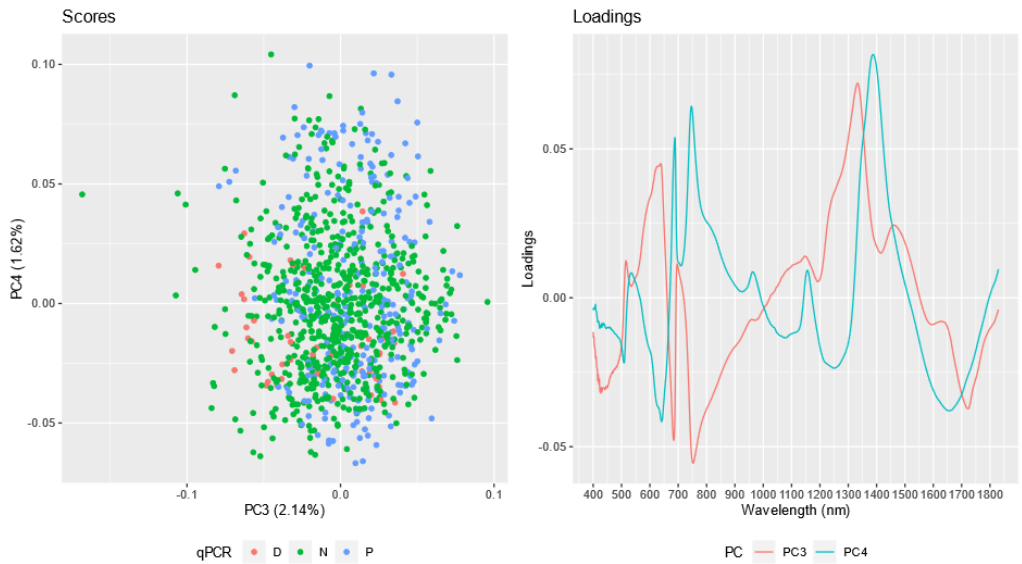


Fig. 29 - Principal component analysis (PCA) of the hyperspectral data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC3/PC4. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.



Fig. 30 - Principal component analysis (PCA) of the hyperspectral data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC3/PC5. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.



Fig. 31 - Principal component analysis (PCA) of the hyperspectral data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC4/PC5. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

For the NMR data, the bucketized spectra were subjected to PCA using both mean-centring and autoscaling. For the first five principal components (PCs), more than 85% of the variance was explained, with the first component accounted for almost 40% of the explained variance and the second principal component having 22%, both accounting for 62% of the explained variance. Furthermore, the third, fourth and fifth principal components had 14%, 6%, and 4% of explained variance respectively (Fig. 32).

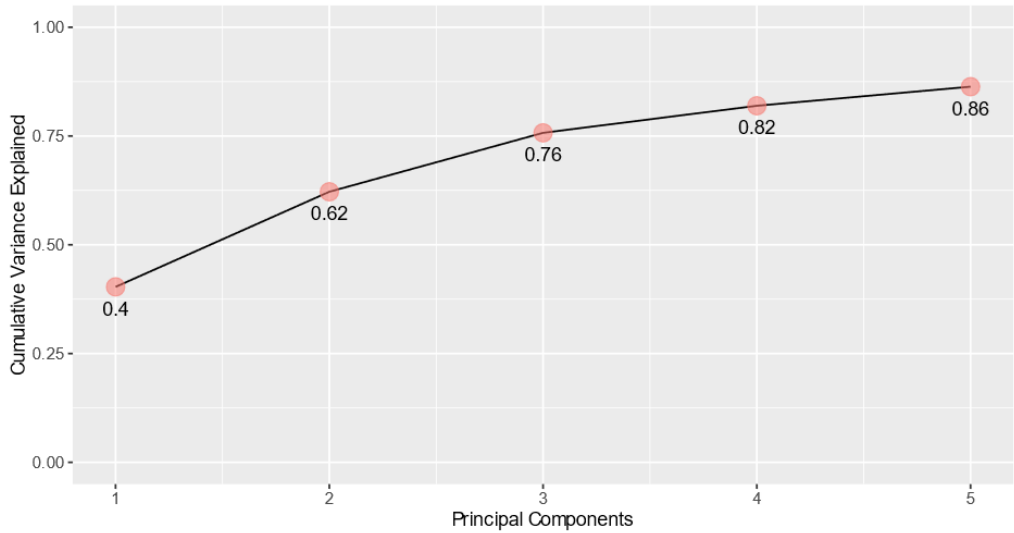


Fig. 32 - Cumulative variance plot showing the total variance explained by the new latent variables (Principal components; PCs) from PC1 to PC5 for NMR samples.

PCA did not show a clear clustering between samples according to qPCR test for *Xylella fastidiosa*. However, by observing the scores plots of the first five principal components, NMR samples showed more distinction than of HSR. This could be argued as if the change at the metabolic level is more sensitive to *Xylella fastidiosa* infection as well as the kind of information exerted from the HSR and NMR (Crandall et al., 2020; Ryan & Robards, 2006; Fine et al., 2021). The investigated combinations of principal components scores and loadings are shown in Fig. 33-42. Moreover, the clustering across PC1, as shown in Fig. 33-36, was rather related to the time of sampling i.e. environmental effect Fig. 43.

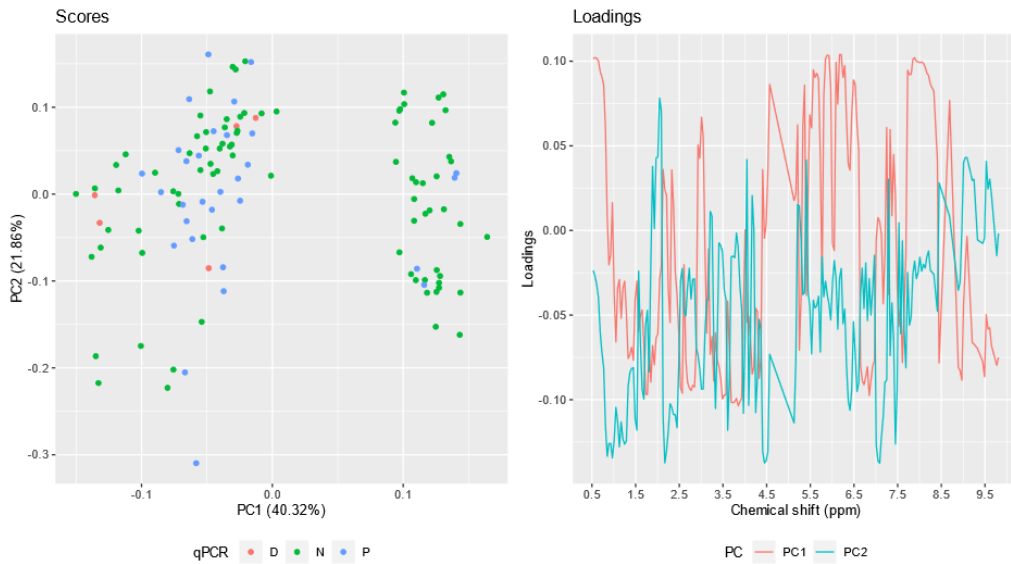


Fig. 33 - Principal component analysis (PCA) of the NMR data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC1/PC2. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

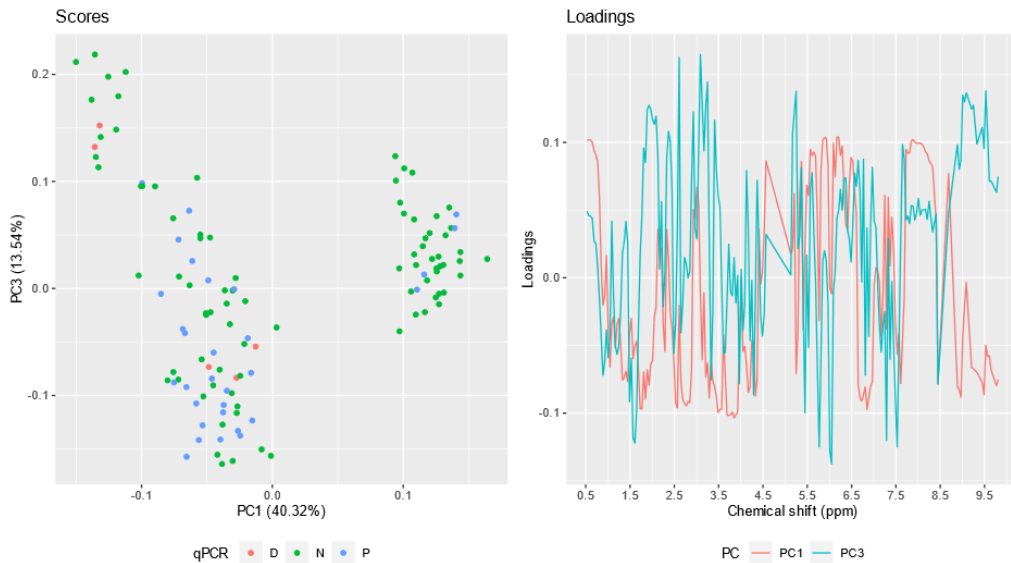


Fig. 34 - Principal component analysis (PCA) of the NMR data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC1/PC3. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

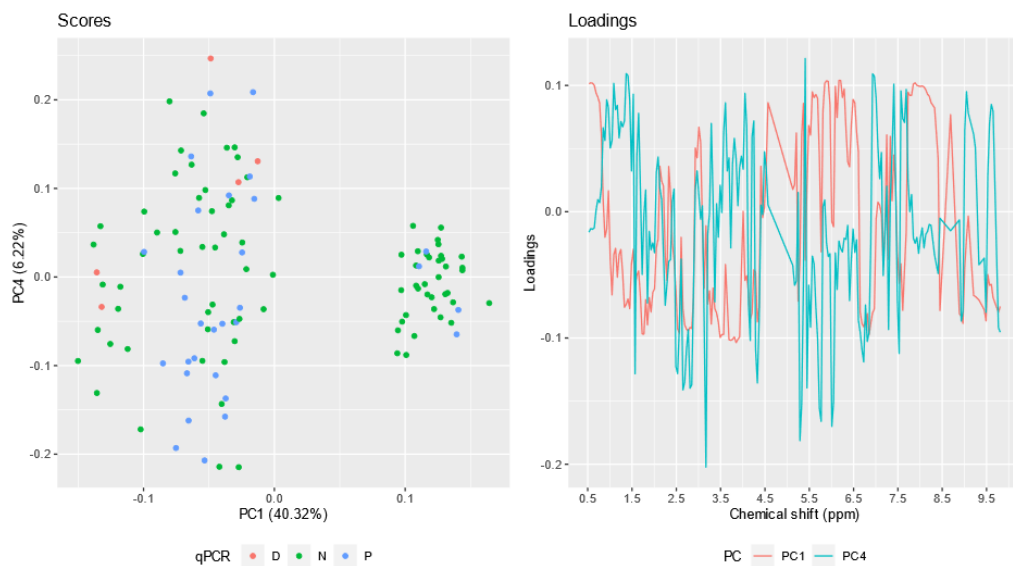


Fig. 35 - Principal component analysis (PCA) of the NMR data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC1/PC4. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

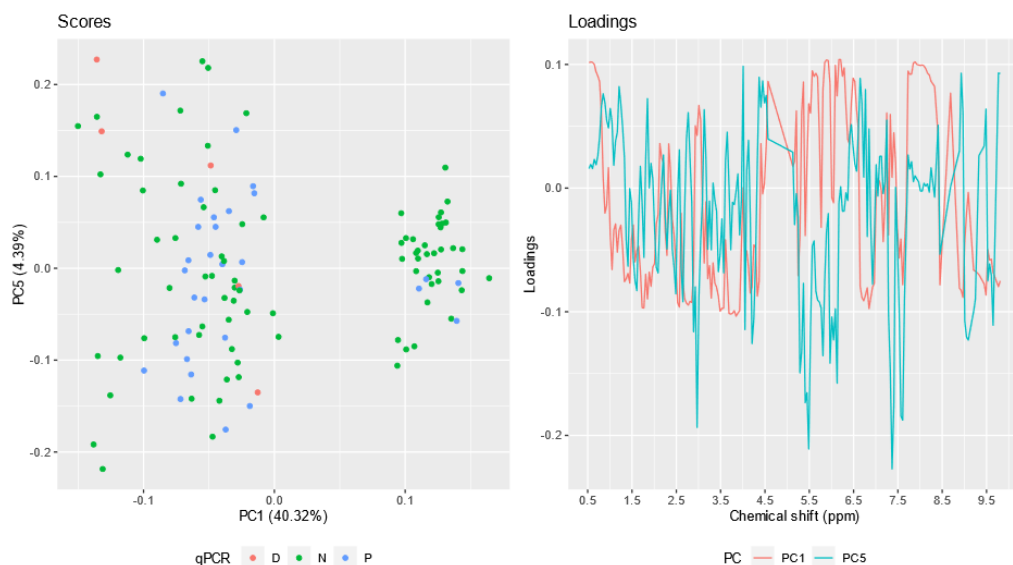


Fig. 36 - Principal component analysis (PCA) of the NMR data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC1/PC5. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

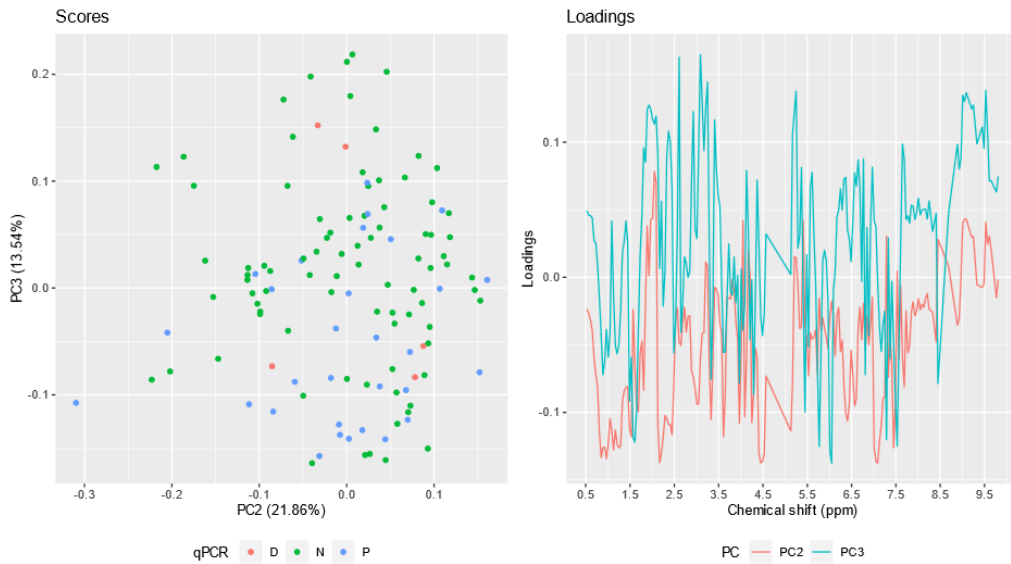


Fig. 37 - Principal component analysis (PCA) of the NMR data samples coloured according to *Xyloella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC2/PC3. P, positive; N, negative; and D, doubtful to *Xyloella fastidiosa* qPCR test.

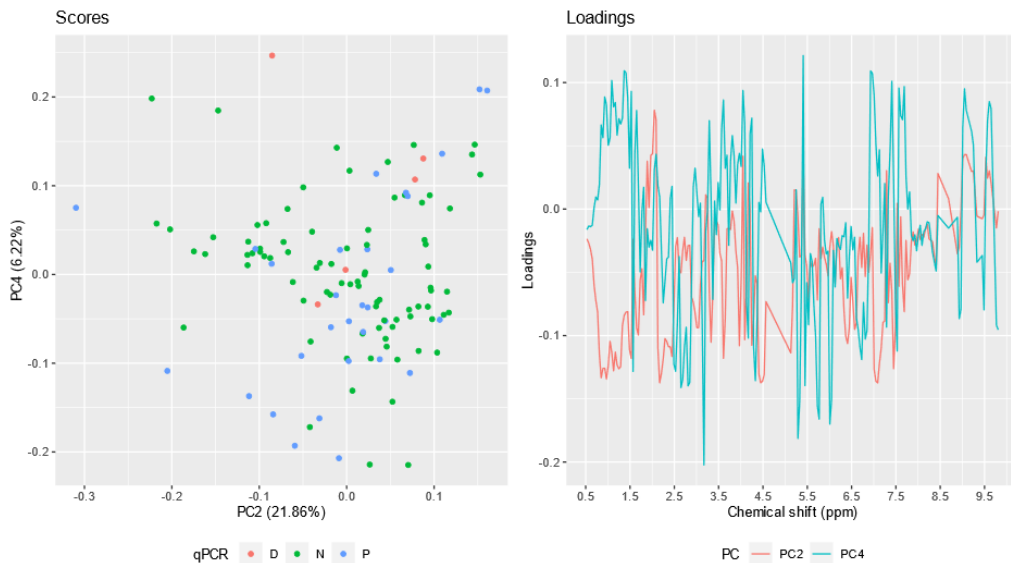


Fig. 38 - Principal component analysis (PCA) of the NMR data samples coloured according to *Xyloella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC2/PC4. P, positive; N, negative; and D, doubtful to *Xyloella fastidiosa* qPCR test.

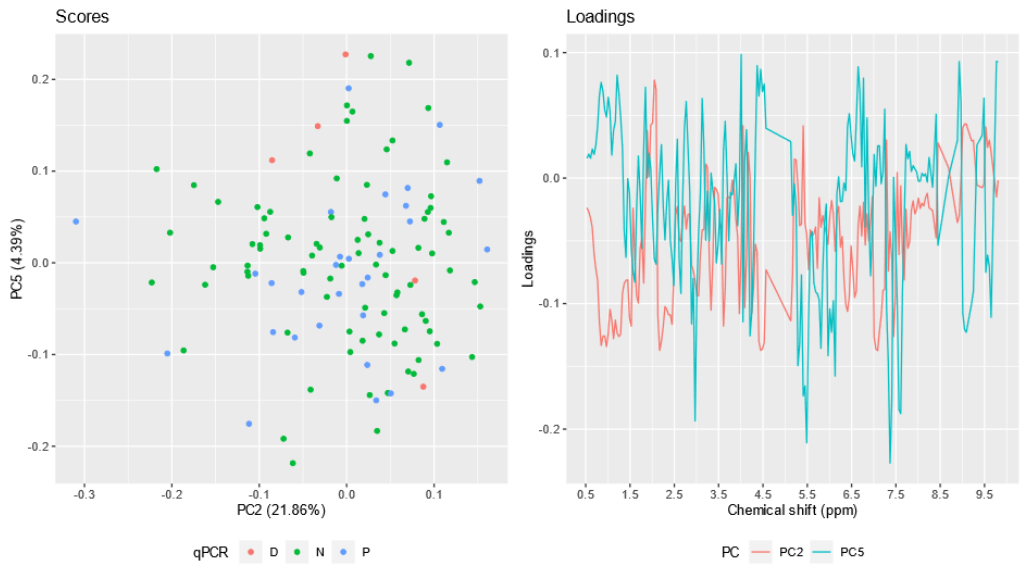


Fig. 39 - Principal component analysis (PCA) of the NMR data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC2/PC5. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* infection qPCR test.

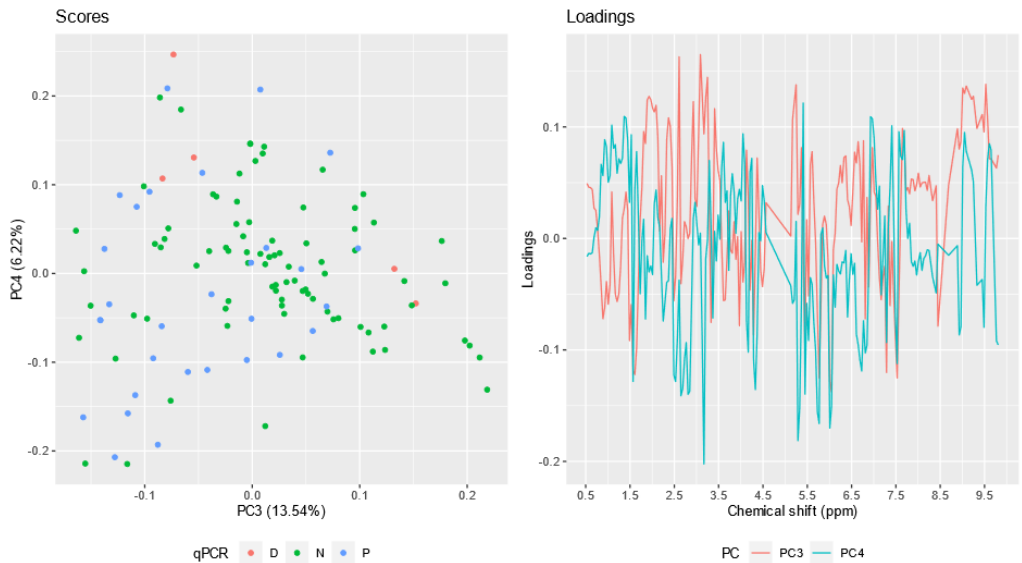


Fig. 40 - Principal component analysis (PCA) of the NMR data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC3/PC4. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

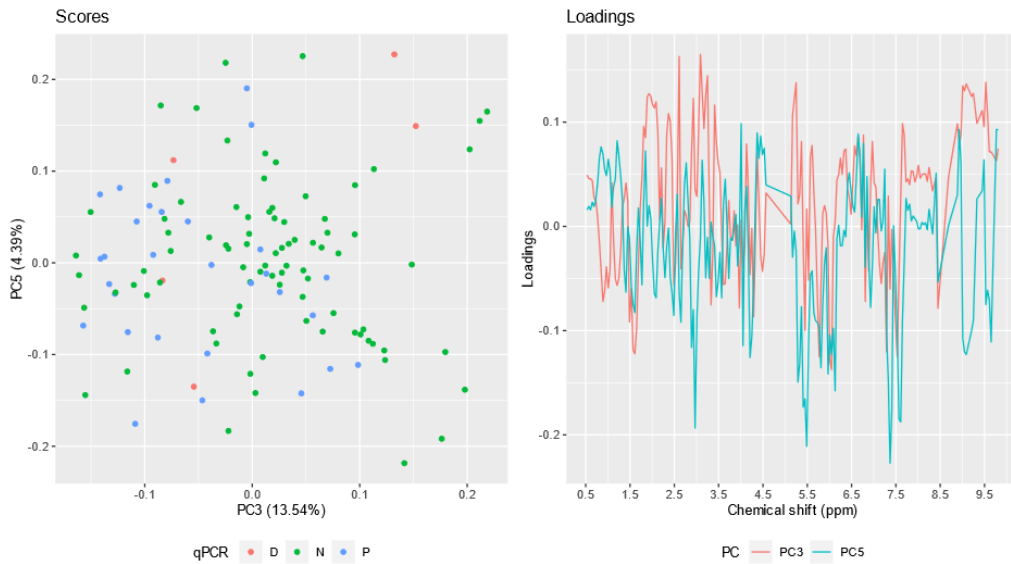


Fig. 41 - Principal component analysis (PCA) of the NMR data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC3/PC5. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

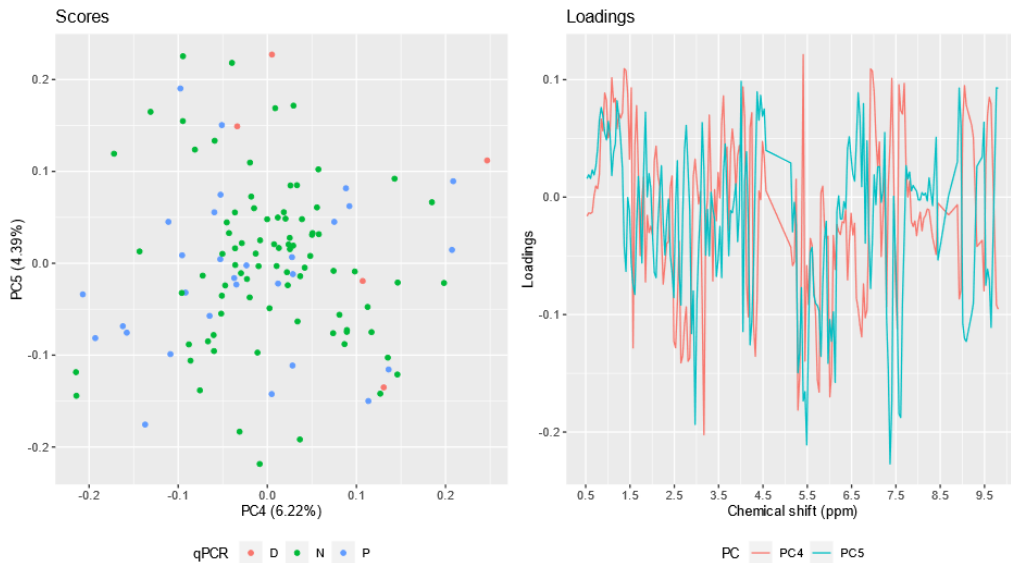


Fig. 42 - Principal component analysis (PCA) of the NMR data samples coloured according to *Xylella fastidiosa* qPCR test; scores plot as well as loadings plot relating to PC4/PC5. P, positive; N, negative; and D, doubtful to *Xylella fastidiosa* qPCR test.

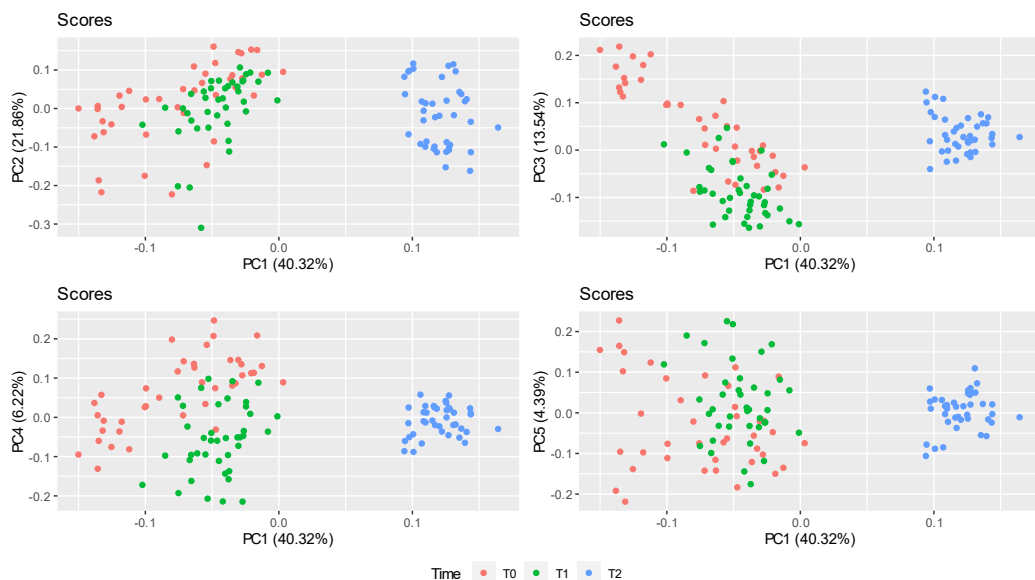


Fig. 43 - Principal component analysis (PCA) of the NMR data samples coloured according to the three sampling times (T0, T1 and T2); scores plot relating to PC1/PC2-5 corresponding to scores plots shown in Fig. 32-35.

Furthermore, to test the ability of hyperspectral data as well as NMR data to differentiate between infected and non-infected olive leaf samples, supervised machine learning methods were used. In particular, the support vector machines SVM, least squares support vector machines LS-SVM, and random forest RF were used. For each method, the data were split between training and testing set at a ratio of 0.7 and 0.3, respectively. The splitting of samples followed the EuroLab Guide which recommends a minimum of 20 samples per class in the test sets (Schönberger et al., 2016).

Interestingly, performing SV, LS-SVM, or RF using the full set of variables for the data considered would need a relatively large amount of computational power. In particular, the amount of memory needed for some the techniques to run was with terabytes (>3000 GBs), which was rather not possible and feasible to do.

Nonetheless, to reduce the amount of features to be used by such supervised machine learning techniques, Fisher Discriminant (FD) analysis was applied (Lovatti et al., 2020). Based on the FD value, features with higher loading in the discrimination were selected. Considering both the SVM and the LS-SVM, the kernel trick was used. In particular, the radial basis kernel was used.

For each of the three applied techniques, the training set was used to build the model. Then, the testing set was used as input for the model and the accuracy of the models was compared.

On one hand, the hyperspectral data showed a higher accuracy for the RF model which had an accuracy with 75%, followed by SVM and LS-SVM with both having an accuracy around 72% (Table 4). On the other hand, NMR data had both LS-SVM and RF models performing with the same exact accuracy of almost 73% as compared to SVM that had an accuracy of almost 68% (Table 5).

Table 4. Confusion Matrix and Statistics for Hyperspectral data:

	RF		SVM		LS-SVM	
	N	P	N	P	N	P
N	N 182	21	N 188	15	N 180	23
P	P 54	44	P 68	30	P 61	37
Accuracy	0.7508		0.7243		0.7209	
95% CI:	(0.698, 0.7987)		(0.6701, 0.774)		(0.6666, 0.7709)	
Sensitivity:	0.7712		0.7344		0.7469	
Specificity:	0.6769		0.6667		0.6167	
Pos. Pred. Value:	0.8966		0.9261		0.8867	
Neg. Pred. Value:	0.4490		0.3061		0.3776	
Prevalence:	0.7841		0.8505		0.8007	
Detection Rate	0.6047		0.6246		0.5980	
Detection Prevalence:	0.6744		0.6744		0.6744	
Balanced Accuracy:	0.7241		0.7005		0.6818	

Table 5. Confusion Matrix and Statistics for NMR data:

	RF	SVM	LS-SVM
	N P	N P	N P
N	N 23 3	N 24 2	N 22 4
P	P 7 4	P 10 1	P 6 5
Accuracy	0.7297	0.6757	0.7297
95% CI:	(0.5588, 0.8621)	(0.5021, 0.8199)	(0.5588, 0.8621)
Sensitivity:	0.7667	0.70588	0.7857
Specificity:	0.5714	0.33333	0.5556
Pos. Pred. Value:	0.8846	0.92308	0.8462
Neg. Pred. Value:	0.3636	0.09091	0.4545
Prevalence:	0.8108	0.91892	0.7568
Detection Rate:	0.6216	0.64865	0.5946
Detection Prevalence:	0.7027	0.70270	0.7027
Balanced Accuracy:	0.6690	0.51961	0.6706

CONCLUSIONS

The RF, SVM, LS-SVM were somehow able to discriminate infected samples. However, SVM seems to be the less effective method among the tested ones. FD proved to be useful in decreasing the number of features to be used while having good performance for the three tested supervised machine learning techniques. Indeed, it is challenging to discriminate between vegetative samples grown in field conditions as the environmental conditions generally account for the major variations amongst the sample and it is not easy to decompose it.

CONCLUSIONS

The work described in this PhD thesis demonstrate different examples of how spectroscopic methods can be of a great use to decipher biological phenomena such plant abiotic as well as biotic stress. In chapter 5, the analysis allowed for the identification of variations in the metabolic composition in hemp that were induced by the treatment with spirulina and the results described may encourage the application of spectroscopic methods for the rapid detection of structural changes in the various environmental spheres, allowing prompt intervention through the adoption of remediation schemes. In chapter 6, the case of young olive trees grown under controlled environmental conditions was considered. The artificially infected plants were studied after 2 years from infection considering only asymptomatic leaves. The analysis of the cultivated asymptomatic olive leaves, infected with *Xylella fastidiosa* subsp. *pauca* ST53 responsible of OQDS, using ^1H NMR, HSR, and chemometrics allowed for selecting diagnostic signals and wavelengths related to OQDS that are necessary for the development of sensors capable of early detection of the disease at the asymptomatic stage. In chapter 7, olive trees of the same cultivar as in chapter 6 were considered, but the samples were collected from the field, in a geographical area affected by OQDS. The different supervised chemometric techniques were somehow able to discriminate between healthy leaves and asymptomatic infected ones. Furthermore, the application of FD (Fisher's Discriminant) analysis was useful in decreasing the number of features to be used in the supervised chemometric techniques. Good performance was obtained for the three tested supervised machine learning techniques even though it is challenging to discriminate amongst vegetative samples at the early stage of asymptomatic infections in field conditions due to high environmental variations.

ACKNOWLEDGEMENTS

I would like to express my gratitude and deep appreciation to those whom enlightened me to get this work done and see the light, namely Dr. Biagia Musio and Prof. Vito Gallo from the Polytechnic University of Bari and Eng. Franco Santoro, Dr. Stefania Gualano, and Dr. Anna Maria D’Onghia from CIHEAM Bari. They kindly provided me with every support and limitless encouragement.

I would like also to extend my thanks to:

Prof. Damiano Varagnolo and Prof. Frank Ove Westad from the Big Data Cybernetics group of the Norwegian University of Science and Technology (NTNU) at which I spent my training abroad.

Prof. Harald Martens from NTNU

PhD colleagues at the Polytechnic University of Bari: Marica Antonicelli and Valentina Petrelli as well as Maria Stella Leone and Darya Nefedova, and Aras Izzaddin.

Polytechnic University of Bari research group: Prof. Piero Mastrorilli, Prof. Mario Latronico, Dr. Rosa Ragone, Dr. Alessandra Ciampa, Dr. Antonio Rizzuti, Dr. Ambra Fiore, Dr. Stefano Todisco, Dr. Maurizio Triggiani, Dr. Francesca Serena Abbatematteo, Prof. Maria Michela Dell’Anna, Dr. Matilda Mali, Dr. Maria Trisolini, Mrs. Stefania Pontrelli, and Mrs. Angelica Sicilia.

My sincere thanks to you.

REFERENCES

- Abdulridha, J., Ehsani, R. & De Castro, A. (2016). Detection and differentiation between laurel wilt disease, phytophthora disease, and salinity damage using a hyperspectral sensing technique. *Agriculture (Switzerland)*. 6 (4).
- Abou Kubaa, R., Giampetruzzi, A., Altamura, G., Saponari, M. & Saldarelli, P. (2019). Infections of the *Xylella fastidiosa* subsp. *Pauca* strain “De donno” in Alfalfa (*medicago sativa*) elicits an overactive immune response. *Plants*. 8 (9).
- Ahmad, R., Tehsin, Z., Malik, S.T., Asad, S.A., Shahzad, M., Bilal, M., Shah, M.M. & Khan, S.A. (2016). Phytoremediation Potential of Hemp (*Cannabis sativa* L.): Identification and Characterization of Heavy Metals Responsive Genes. *Clean - Soil, Air, Water*. 44 (2). p.pp. 195–201.
- Almeida, R.P.P. (2016). Can Apulia’s olive trees be saved? *Science*. 353 (6297). p.pp. 346–348.
- Alseekh, S. & Fernie, A.R. (2018). Metabolomics 20 years on: what have we learned and what hurdles remain? *Plant Journal*. 94 (6). p.pp. 933–942.
- Anguita-Maeso, M., Haro, C., Montes-Borrego, M., De La Fuente, L., Navas-Cortés, J.A. & Landa, B.B. (2021). Metabolomic, Ionic and Microbial Characterization of Olive Xylem Sap Reveals Differences According to Plant Age and Genotype. *Agronomy*. [Online]. 11 (6). p.p. 1179. Available from: <https://www.mdpi.com/2073-4395/11/6/1179>.
- Asner, G.P. & Martin, R.E. (2009). Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. *Frontiers in Ecology and the Environment*. [Online]. 7 (5). p.pp. 269–276. Available from: <http://doi.wiley.com/10.1890/070152>.
- Asner, G.P. & Martin, R.E. (2011). Canopy phylogenetic, chemical and spectral assembly in a lowland Amazonian forest. *New Phytologist*. 189 (4). p.pp. 999–1012.
- Asner, G.P. & Martin, R.E. (2016). Spectranomics: Emerging science and conservation opportunities at the interface of biodiversity and remote sensing. *Global Ecology and Conservation*. [Online]. 8. p.pp. 212–219. Available from: <http://dx.doi.org/10.1016/j.gecco.2016.09.010>.
- Asner, G.P., Martin, R.E., Carranza-Jiménez, L., Sinca, F., Tupayachi, R., Anderson, C.B. & Martinez, P. (2014a). Functional and biological diversity of foliar spectra in tree canopies throughout the Andes to Amazon region. *New Phytologist*. 204 (1). p.pp. 127–139.

- Asner, G.P., Martin, R.E., Tupayachi, R., Anderson, C.B., Sinca, F., Carranza-Jiménez, L. & Martinez, P. (2014b). Amazonian functional diversity from forest canopy chemical assembly. *Proceedings of the National Academy of Sciences of the United States of America*. 111 (15). p.pp. 5604–5609.
- Augustijn, D., de Groot, H.J.M. & Alia, A. (2021). HR-MAS NMR Applications in Plant Metabolomics. *Molecules*. 26 (4). p.p. 931.
- Balaji, S., Kalaivani, T. & Rajasekaran, C. (2014). Biosorption of Zinc and Nickel and Its Effect on Growth of Different Spirulina Strains. *Clean - Soil, Air, Water*. 42 (4). p.pp. 507–512.
- Baldi, P. & La Porta, N. (2017). Xylella fastidiosa: Host range and advance in molecular identification techniques. *Frontiers in Plant Science*. 8 (June).
- Baxter, I. (2009). *Ionomics : studying the social network of mineral nutrients*. p.pp. 381–386.
- Baxter, I.R., Vitek, O., Lahner, B., Muthukumar, B., Borghi, M., Morrissey, J., Guerinot, M. Lou & Salt, D.E. (2008). The leaf ionome as a multivariable system to detect a plant's physiological status. *Proceedings of the National Academy of Sciences of the United States of America*. 105 (33). p.pp. 12081–12086.
- Bedia, C., Cardoso, P., Dalmau, N., Garreta-Lara, E., Gómez-Canela, C., Gorrochategui, E., Navarro-Reig, M., Ortiz-Villanueva, E., Puig-Castellví, F. & Tauler, R. (2018). *Applications of Metabolomics Analysis in Environmental Research*. 1st Ed. [Online]. Elsevier B.V. Available from: <http://dx.doi.org/10.1016/bs.coac.2018.07.006>.
- Behmann, J., Mahlein, A.K., Rumpf, T., Römer, C. & Plümer, L. (2015). A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. *Precision Agriculture*. 16 (3). p.pp. 239–260.
- Blekos, K., Tsakas, A., Xouris, C., Evdokidis, I., Alexandropoulos, D., Alexakos, C., Katakis, S., Makedonas, A., Theoharatos, C. & Lalos, A. (2021). Analysis, Modeling and Multi-Spectral Sensing for the Predictive Management of Verticillium Wilt in Olive Groves. *Journal of Sensor and Actuator Networks*. [Online]. 10 (1). p.p. 15. Available from: <https://www.mdpi.com/2224-2708/10/1/15>.
- Bloch, F., Hansen, W.W. & Packard, M. (1946). Nuclear Induction. *Physical Review*. [Online]. 69 (3–4). p.pp. 127–127. Available from: <https://link.aps.org/doi/10.1103/PhysRev.69.127>.
- Bohnenkamp, D., Kuska, M.T., Mahlein, A.K. & Behmann, J. (2019). Hyperspectral signal decomposition and symptom detection of wheat rust disease at the leaf scale using pure fungal spore spectra as reference. *Plant Pathology*. 68 (6). p.pp. 1188–1195.

- Brereton, R.G. (2015). Pattern recognition in chemometrics. *Chemometrics and Intelligent Laboratory Systems*. [Online]. 149. p.pp. 90–96. Available from: <http://dx.doi.org/10.1016/j.chemolab.2015.06.012>.
- Brito, C., Dinis, L.T., Moutinho-Pereira, J. & Correia, C.M. (2019). Drought stress effects and olive tree acclimation under a changing climate. *Plants*. 8 (7). p.pp. 1–20.
- Bro, R. & Smilde, A.K. (2014). Principal component analysis. *Analytical Methods*. 6 (9). p.pp. 2812–2831.
- Calderón, R., Navas-Cortés, J.A., Lucena, C. & Zarco-Tejada, P.J. (2013). High-resolution airborne hyperspectral and thermal imagery for early detection of Verticillium wilt of olive using fluorescence, temperature and narrow-band spectral indices. *Remote Sensing of Environment*. [Online]. 139. p.pp. 231–245. Available from: <http://dx.doi.org/10.1016/j.rse.2013.07.031>.
- Camino, C., Calderón, R., Parnell, S., Dierkes, H., Chemin, Y., Román-Écija, M., Montes-Borrego, M., Landa, B.B., Navas-Cortés, J.A., Zarco-Tejada, P.J. & Beck, P.S.A. (2021). Detection of *Xylella fastidiosa* in almond orchards by synergic use of an epidemic spread model and remotely sensed plant traits. *Remote Sensing of Environment*. [Online]. 260 (October 2020). p.p. 112420. Available from: <https://www.sciencedirect.com/science/article/pii/S0034425721001383>.
- Cariddi, C., Saponari, M., Boscia, D., De Stradis, A., Loconsole, G., Nigro, F., Porcelli, F., Potere, O. & Martelli, G.P. (2014). Isolation of a *Xylella fastidiosa* strain infecting olive and oleander in Apulia, Italy. *Journal of Plant Pathology*. 96 (2). p.pp. 425–429.
- Carter, G.A. & Knapp, A.K. (2001). Leaf optical properties in higher plants: linking spectral characteristics to stress and chlorophyll concentration. *American Journal of Botany*. 88 (4). p.pp. 677–684.
- Cavill, R., Jennen, D., Kleinjans, J. & Briedé, J.J. (2016). Transcriptomic and metabolomic data integration. *Briefings in Bioinformatics*. 17 (5). p.pp. 891–901.
- Cepoi, L., Zinicovscaia, I., Rudi, L., Chiriac, T., Miscu, V., Djur, S., Strelkova, L., Vergel, K. & Nekhoroshkov, P. (2020). Growth and heavy metals accumulation by *Spirulina platensis* biomass from multicomponent copper containing synthetic effluents during repeated cultivation cycles. *Ecological Engineering*. [Online]. 142 (October 2019). p.p. 105637. Available from: <https://doi.org/10.1016/j.ecoleng.2019.105637>.
- Choi, J., Pak, C. & Lee, C.W. (1996). Micro nutrient toxicity in French marigold. *Journal of Plant Nutrition*. [Online]. 19 (6). p.pp. 901–916. Available from: <http://www.tandfonline.com/doi/abs/10.1080/01904169609365169>.

- Choi, J., Roy Choudhury, A., Park, S.Y., Oh, M.M. & Sa, T. (2021). Inoculation of acc deaminase-producing *Brevibacterium linens* rs16 enhances tolerance against combined uv-b radiation and heat stresses in rice (*Oryza sativa* L.). *Sustainability (Switzerland)*. 13 (18).
- Ciampa, A., Danesi, F. & Picone, G. (2022). NMR-Based Metabolomics for a More Holistic and Sustainable Research in Food Quality Assessment: A Narrative Review. *Applied Sciences*. [Online]. 13 (1). p.p. 372. Available from: <https://www.mdpi.com/2076-3417/13/1/372>.
- Citterio, S., Santagostino, A., Fumagalli, P., Prato, N., Ranalli, P. & Sgorbati, S. (2003). Heavy metal tolerance and accumulation of Cd, Cr and Ni by *Cannabis sativa* L. *Plant and Soil*. 256 (2). p.pp. 243–252.
- Del Coco, L., Girelli, C.R., Angilè, F., Mascio, I., Montemurro, C., Distaso, E., Tamburrano, P., Chiurlia, S., Clodoveo, M.L., Corbo, F., Amirante, R., Schena, F.P. & Fanizzi, F.P. (2021). NMR-based metabolomic study of Apulian Coratina extra virgin olive oil extracted with a combined ultrasound and thermal conditioning process in an industrial setting. *Food Chemistry*. 345 (December 2020).
- Couture, J.J., Singh, A., Rubert-Nason, K.F., Serbin, S.P., Lindroth, R.L. & Townsend, P.A. (2016). Spectroscopic determination of ecologically relevant plant secondary metabolites. *Methods in Ecology and Evolution*. 7 (11). p.pp. 1402–1412.
- Cozzolino, D., Power, A. & Chapman, J. (2019). Interpreting and Reporting Principal Component Analysis in Food Science Analysis and Beyond. *Food Analytical Methods*. 12 (11). p.pp. 2469–2473.
- Crandall, S.G., Gold, K.M., Jiménez-Gasco, M. del M., Camila, Filgueiras, C. & Willett, D.S. (2020). A multi-omics approach to solving problems in plant disease ecology. *PLoS ONE*. 15 (9 September). p.pp. 1–23.
- Crick, F. (1970). Central Dogma of Molecular Biology. *Nature*. [Online]. 227 (5258). p.pp. 561–563. Available from: <https://www.nature.com/articles/227561a0>.
- Curran, P.J. (1989). Remote sensing of foliar chemistry. *Remote Sensing of Environment*. 30 (3). p.pp. 271–278.
- A. M. D’Onghia, S. Brunel, & F. Valentini (eds.) (2017). *Xylella fastidiosa & the Olive Quick Decline Syndrome (OQDS) A serious worldwide challenge for the safeguard of olive trees*. In: *Options mediterraneenes Series A: Mediterranean Seminars*. [Online]. 2017, Bari: IAM Bari: CIHEAM (Centre International de Hautes Etudes Agronomiques Méditerranéennes), p. 172. Available from: www.ciheam.org/publications.

- Dalcorso, G., Manara, A. & Furini, A. (2013). An overview of heavy metal challenge in plants: From roots to shoots. *Metallomics*. 5 (9). p.pp. 1117–1132.
- Dangi, A.K., Sharma, B., Khangwal, I. & Shukla, P. (2018). Combinatorial Interactions of Biotic and Abiotic Stresses in Plants and Their Molecular Mechanisms: Systems Biology Approach. *Molecular Biotechnology*. [Online]. 60 (8). p.pp. 636–650. Available from: <http://dx.doi.org/10.1007/s12033-018-0100-9>.
- Danson, F.M., Steven, M. D., Malthus, T.J. & Clark, J.A. (1992). International Journal of Remote Sensing High-spectral resolution data for determining leaf water content. *International Journal of Remote Sensing*. 13 (3). p.pp. 461–470.
- Decreto Legislativo 152/2006*. (2006). [Online]. Available from: https://www.gazzettaufficiale.it/atto/serie_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=2006-04-14&atto.codiceRedazionale=006G0171.
- Decreto Legislativo 152/2006 - Supplemento ordinario*. (2006). [Online]. Available from: <http://www.gazzettaufficiale.it/eli/gu/2006/04/14/88/so/96/sg/pdf>.
- Decreto Legislativo 99/1992*. (1992). [Online]. Available from: <papers2://publication/uuid/8C3C2075-BF30-40B9-9FA0-5F736203E123>.
- Dg Internal Policies of the Union. *IP/A/ENVI/IC/2006-183. Status of Implementation of EU Environmental Laws in Italy*. (2006).
- Ebbs, S.D. & Kochian, L. V. (1997). Toxicity of Zinc and Copper to Brassica Species: Implications for Phytoremediation. *Journal of Environmental Quality*. 26 (3). p.pp. 776–781.
- Edelstein, M. & Ben-Hur, M. (2018). Heavy metals and metalloids: Sources, risks and strategies to reduce their accumulation in horticultural crops. *Scientia Horticulturae*. [Online]. 234 (June). p.pp. 431–444. Available from: <http://dx.doi.org/10.1016/j.scienta.2017.12.039>.
- EFSA (European Food Safety Authority), Delbianco, A., Gibin, D., Pasinato, L. & Morelli, M. (2022). Update of the *Xylella* spp. host plant database – systematic literature search up to 30 June 2021. *EFSA Journal*. 20 (1). p.pp. 1–67.
- Elbeaino, T., Valentini, F., Abou Kubaa, R., Moubarak, P., Yaseen, T. & Digiario, M. (2014). Multilocus sequence typing of *Xylella fastidiosa* isolated from olive affected by “olive quick decline syndrome” in Italy. *Phytopathologia Mediterranea*. [Online]. 53 (3 SE-Research Papers). p.pp. 241–252. Available from: <https://oajournals.fupress.net/index.php/pm/article/view/5592>.

- Eriksson, L., Hagberg, P., Johansson, E., Rännar, S., Whelehan, O., Åström, A. & Lindgren, T. (2001). Multivariate process monitoring of a newsprint mill. Application to modelling and predicting COD load resulting from de-inking of recycled paper. *Journal of Chemometrics*. 15 (4). p.pp. 337–352.
- European Commission. 86/278/EEC. *COUNCIL DIRECTIVE of 12 June 1986 on the protection of the environment, and in particular of the soil, when sewage sludge is used in agriculture*. (1986).
- European Commission (2020). Commission Implementing Regulation (EU) 2020/1201 of 14 August 2020 as regards measures to prevent the introduction into and the spread within the Union of *Xylella fastidiosa* (Wells et al.). *OJ*. [Online]. L 269. p.pp. 2–39. Available from: https://eur-lex.europa.eu/eli/reg_impl/2020/1201/2020-08-17.
- Feilhauer, H., Asner, G.P. & Martin, R.E. (2015). Multi-method ensemble selection of spectral bands related to leaf biochemistry. *Remote Sensing of Environment*. [Online]. 164. p.pp. 57–65. Available from: <http://dx.doi.org/10.1016/j.rse.2015.03.033>.
- Femenias, A., Gatus, F., Ramos, A.J., Sanchis, V. & Marín, S. (2021). Near-infrared hyperspectral imaging for deoxynivalenol and ergosterol estimation in wheat samples. *Food Chemistry*. [Online]. 341 (September 2020). p.p. 128206. Available from: <https://doi.org/10.1016/j.foodchem.2020.128206>.
- Féret, J.B., Berger, K., de Boissieu, F. & Malenovský, Z. (2021). PROSPECT-PRO for estimating content of nitrogen-containing leaf proteins and other carbon-based constituents. *Remote Sensing of Environment*. 252 (October 2020).
- Feussner, I. & Polle, A. (2015). What the transcriptome does not tell - proteomics and metabolomics are closer to the plants' patho-phenotype. *Current Opinion in Plant Biology*. [Online]. 26. p.pp. 26–31. Available from: <http://dx.doi.org/10.1016/j.pbi.2015.05.023>.
- Fiehn, O. (2001). Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comparative and Functional Genomics*. 2 (3). p.pp. 155–168.
- Fiehn, O. (2002). Metabolomics - The link between genotypes and phenotypes. *Plant Molecular Biology*. 48 (1–2). p.pp. 155–171.
- Fine, P.V.A., Salazar, D., Martin, R.E., Metz, M.R., Misiewicz, T.M. & Asner, G.P. (2021). Exploring the links between secondary metabolites and leaf spectral reflectance in a diverse genus of Amazonian trees. *Ecosphere*. 12 (2).
- Fiorani, F. & Schurr, U. (2013). Future scenarios for plant phenotyping. *Annual Review of Plant Biology*. 64. p.pp. 267–291.

- Fisher, R.A. (1915). Frequency distribution of the values of the correlation coefficient in samples from indefinitely large population. *Biometrika*. 10. p.pp. 507–521.
- Fisher, R.A. (1921). On the " Probable Error" of a Coefficient of Correlation Deduced from a Small Sample. *Biometrika*. 1. p.pp. 3–32.
- Fontes, R.L.F. & Cox, F.R. (1998). Zinc toxicity in soybean grown at high iron concentration in nutrient solution. *Journal of Plant Nutrition*. [Online]. 21 (8). p.pp. 1723–1730. Available from: <http://www.tandfonline.com/doi/abs/10.1080/01904169809365517>.
- Furbank, R.T. & Tester, M. (2011). Phenomics - technologies to relieve the phenotyping bottleneck. *Trends in Plant Science*. [Online]. 16 (12). p.pp. 635–644. Available from: <http://dx.doi.org/10.1016/j.tplants.2011.09.005>.
- Galić, M., Perčin, A., Zgorelec, Ž. & Kisić, I. (2019). Evaluation of heavy metals accumulation potential of hemp (*Cannabis sativa* L.). *Journal of Central European Agriculture*. [Online]. 20 (2). p.pp. 700–711. Available from: <https://jcea.agr.hr/en/issues/article/2201>.
- Galvan, F.E.R., Pavlick, R., Trolley, G., Ag-, S., Sousa, D., Starr, C., Forrestel, E., Bolton, S., Alsina, M., Dokoozlian, N. & Gold, K.M. (2022). *Scalable early detection of grapevine virus infection with airborne imaging spectroscopy*.
- Gika, H.G., Wilson, I.D. & Theodoridis, G.A. (2014). LC-MS-based holistic metabolic profiling. Problems, limitations, advantages, and future perspectives. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*. [Online]. 966. p.pp. 1–6. Available from: <http://dx.doi.org/10.1016/j.jchromb.2014.01.054>.
- Girdhar, M., Sharma, N.R., Rehman, H., Kumar, A. & Mohan, A. (2014). Comparative assessment for hyperaccumulatory and phytoremediation capability of three wild weeds. *3 Biotech*. 4 (6). p.pp. 579–589.
- Girelli, C.R., Angilè, F., Coco, L. Del, Migoni, D., Zampella, L., Marcelletti, S., Cristella, N., Marangi, P., Scortichini, M. & Fanizzi, F.P. (2019). 1H-NMR metabolite fingerprinting analysis reveals a disease biomarker and a field treatment response in xylella fastidiosa subsp. Pauca-infected olive trees. *Plants*. 8 (5).
- Girelli, C.R., Del Coco, L., Scortichini, M., Petriccione, M., Zampella, L., Mastrobuoni, F., Cesari, G., Bertaccini, A., D'Amico, G., Contaldo, N., Migoni, D. & Fanizzi, F.P. (2017). Xylella fastidiosa and olive quick decline syndrome (CoDiRO) in Salento (southern Italy): a chemometric 1H NMR-based preliminary study on Ogliarola salentina and Cellina di Nardò cultivars. *Chemical and Biological Technologies in Agriculture*. 4 (1). p.pp. 1–9.

- Godefroid, M., Morente, M., Schartel, T., Cornara, D., Purcell, A., Gallego, D., Moreno Lozano, A., Pereira, J. & Fereres, A. (2020). *The risk of Xylella fastidiosa outbreaks will decrease in the Mediterranean olive-producing regions*.
- Gold, K.M., Townsend, P.A., Chlus, A., Herrmann, I., Couture, J.J., Larson, E.R. & Gevens, A.J. (2020a). Hyperspectral measurements enable pre-symptomatic detection and differentiation of contrasting physiological effects of late blight and early blight in potato. *Remote Sensing*. 12 (2).
- Gold, K.M., Townsend, P.A., Larson, E.R., Herrmann, I. & Gevens, A.J. (2020b). Contact reflectance spectroscopy for rapid, accurate, and nondestructive phytophthora infestans clonal lineage discrimination. *Phytopathology*. 110 (4). p.pp. 851–862.
- Gomont, M. (1892). Monographie des Oscillariées (Nostocacées Homocystées). Deuxième partie. - Lyngbyées. *Annales Des Sciences Naturelles, Botanique*. 7 (16). p.pp. 91–264.
- Gorrochategui, E., Jaumot, J., Lacorte, S. & Tauler, R. (2016). Data analysis strategies for targeted and untargeted LC-MS metabolomic studies: Overview and workflow. *TrAC - Trends in Analytical Chemistry*. [Online]. 82. p.pp. 425–442. Available from: <http://dx.doi.org/10.1016/j.trac.2016.07.004>.
- Granato, D., Putnik, P., Kovačević, D.B., Santos, J.S., Calado, V., Rocha, R.S., Cruz, A.G. Da, Jarvis, B., Rodionova, O.Y. & Pomerantsev, A. (2018). Trends in Chemometrics: Food Authentication, Microbiology, and Effects of Processing. *Comprehensive Reviews in Food Science and Food Safety*. 17 (3). p.pp. 663–677.
- Gromski, P.S., Muhamadali, H., Ellis, D.I., Xu, Y., Correa, E., Turner, M.L. & Goodacre, R. (2015). A tutorial review: Metabolomics and partial least squares-discriminant analysis - a marriage of convenience or a shotgun wedding. *Analytica Chimica Acta*. [Online]. 879. p.pp. 10–23. Available from: <http://dx.doi.org/10.1016/j.aca.2015.02.012>.
- Gupta, A., Srivastava, P.K., Petropoulos, G.P. & Singh, P. (2021). Statistical unfolding approach to understand influencing factors for taxol content variation in high altitude himalayan region. *Forests*. 12 (12).
- Hall, R.D. (2018). Plant Metabolomics in a Nutshell: Potential and Future Challenges. *Annual Plant Reviews online*. 43. p.pp. 1–24.
- Hall, R.D., D’Auria, J.C., Silva Ferreira, A.C., Gibon, Y., Kruszka, D., Mishra, P. & van de Zedde, R. (2022). High-throughput plant phenotyping: a role for metabolomics? *Trends in Plant Science*. [Online]. xx (xx). p.pp. 1–15. Available from: <https://doi.org/10.1016/j.tplants.2022.02.001>.

- Harper, S.J., Ward, L.I. & Clover, G.R.G. (2010). Development of LAMP and Real-Time PCR Methods for the Rapid Detection of *Xylella fastidiosa* for Quarantine and Field Applications. *Phytopathology*®. [Online]. 100 (12). p.pp. 1282–1288. Available from: <https://apsjournals.apsnet.org/doi/10.1094/PHYTO-06-10-0168>.
- Hernández, E. & Olguín, E.J. (2002). Biosorption of heavy metals influenced by the chemical composition of spirulina sp. (arthrospira) biomass. *Environmental Technology (United Kingdom)*. 23 (12). p.pp. 1369–1377.
- Hill, B.L. & Purcell, A.H. (1995). Acquisition and retention of *Xylella fastidiosa* by an efficient vector, *Graphocephala atropunctata*. *Phytopathology*. 85 (2) p.pp. 209–212.
- Ho, S.H., Liao, J.F., Chen, C.Y. & Chang, J.S. (2018). Combining light strategies with recycled medium to enhance the economic feasibility of phycocyanin production with *Spirulina platensis*. *Bioresource Technology*. [Online]. 247. p.pp. 669–675. Available from: <https://doi.org/10.1016/j.biortech.2017.09.165>.
- van der Hooft, J.J.J., Mohimani, H., Bauermeister, A., Dorrestein, P.C., Duncan, K.R. & Medema, M.H. (2020). Linking genomics and metabolomics to chart specialized metabolic diversity. *Chemical Society reviews*. 49 (11). p.pp. 3297–3314.
- Hopkins, D.L. & Purcell, a H. (2002). *Xylella fastidiosa*: Cause of Pierce’s Disease of Grapevine and Other Emergent Diseases. *Plant Disease*. [Online]. 86 (10). p.pp. 1056–1066. Available from: <https://apsjournals.apsnet.org/doi/10.1094/PDIS.2002.86.10.1056>.
- Hornero, A., Hernández-Clemente, R., North, P.R.J., Beck, P.S.A., Boscia, D., Navas-Cortes, J.A. & Zarco-Tejada, P.J. (2020). Monitoring the incidence of *Xylella fastidiosa* infection in olive orchards using ground-based evaluations, airborne imaging spectroscopy and Sentinel-2 time series through 3-D radiative transfer modelling. *Remote Sensing of Environment*. [Online]. 236 (September 2019). p.p. 111480. Available from: <https://doi.org/10.1016/j.rse.2019.111480>.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*. 24 (6). p.pp. 417–441.
- Innamorato, V., Longobardi, F., Cervellieri, S., Cefola, M., Pace, B., Capotorto, I., Gallo, V., Rizzuti, A., Logrieco, A.F. & Lippolis, V. (2020). Quality evaluation of table grapes during storage by using 1H NMR, LC-HRMS, MS-eNose and multivariate statistical analysis. *Food Chemistry*. [Online]. 315 (August 2019). p.p. 126247. Available from: <https://doi.org/10.1016/j.foodchem.2020.126247>.

- Jacquemoud, S., Verdebout, J., Schmuck, G., Andreoli, G. & Hosgood, B. (1995). Investigation of leaf biochemistry by statistics. *Remote Sensing of Environment*. 54 (3). p.pp. 189–197.
- Jlilat, A., Ragone, R., Gualano, S., Santoro, F., Gallo, V., Varvaro, L., Mastrorilli, P., Saponari, M., Nigro, F. & D’Onghia, A.M. (2021). A non-targeted metabolomics study on *Xylella fastidiosa* infected olive plants grown under controlled conditions. *Scientific Reports*. [Online]. 11 (1). p.p. 1070. Available from: <https://doi.org/10.1038/s41598-020-80090-x>.
- Jolley, K.A. & Maiden, M.C.J. (2010). BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics*. [Online]. 11 (1). p.p. 595. Available from: <http://www.biomedcentral.com/1471-2105/11/595>.
- Kaddurah-Daouk, R., Kristal, B.S. & Weinshilboum, R.M. (2008). Metabolomics: A global biochemical approach to drug response and disease. *Annual Review of Pharmacology and Toxicology*. 48. p.pp. 653–683.
- Karahalil, B. (2016). Overview of Systems Biology and Omics Technologies. *Current Medicinal Chemistry*. 23 (37). p.pp. 4221–4230.
- Kim, M.K. & Lun, D.S. (2014). Methods for integration of transcriptomic data in genome-scale metabolic models. *Computational and Structural Biotechnology Journal*. [Online]. 11 (18). p.pp. 59–65. Available from: <http://dx.doi.org/10.1016/j.csbj.2014.08.009>.
- Koch, A.L. (1970). Turbidity measurements of bacterial cultures in some available commercial instruments. *Analytical Biochemistry*. 38 (1). p.pp. 252–259.
- Konig-péter, A., Kilar, F., Felinger, A. & Pernyeszi, T. (2015). Biosorption characteristics of *Spirulina* and *Chlorella* cells for the accumulation of heavy metals. *Journal of the Serbian Chemical Society*. 80 (3). p.pp. 407–419.
- Kos, B., Grčman, H. & Leštan, D. (2003). Phytoextraction of lead, zinc and cadmium from soil by selected plants. *Plant, Soil and Environment*. 49 (12). p.pp. 548–553.
- Labine, L.M. & Simpson, M.J. (2020). The use of nuclear magnetic resonance (NMR) and mass spectrometry (MS)–based metabolomics in environmental exposure assessment. *Current Opinion in Environmental Science and Health*. [Online]. 15. p.pp. 7–15. Available from: <https://doi.org/10.1016/j.coesh.2020.01.008>.
- Lewis, S., Donkin, M. & Depledge, M. (2001). Hsp70 expression in *Enteromorpha intestinalis* (Chlorophyta) exposed to environmental stressors. *Aquatic Toxicology*. [Online]. 51 (3). p.pp. 277–291. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0166445X00001193>.

- Li, H., Yang, W., Lei, J., She, J. & Zhou, X. (2021). Estimation of leaf water content from hyperspectral data of different plant species by using three new spectral absorption indices. *PLoS ONE*. [Online]. 16 (3 March). p.pp. 1–16. Available from: <http://dx.doi.org/10.1371/journal.pone.0249351>.
- Linger, P., Ostwald, A. & Haensler, J. (2005). Cannabis sativa L. growing on heavy metal contaminated soil: Growth, cadmium uptake and photosynthesis. *Biologia Plantarum*. 49 (4). p.pp. 567–576.
- Loconsole, G., Potere, O., Boscia, D., Altamura, G., Djelouah, K., Elbeaino, T., Frasherri, D., Lorusso, D., Palmisano, F., Pollastro, P., Silletti, M.R., Trisciuzzi, N., Valentini, F., Savino, V. & Saponari, M. (2014). Detection of *Xylella fastidiosa* in olive trees by molecular and serological methods. *Journal of Plant Pathology*. 96 (1). p.pp. 7–14.
- Lovatti, B.P.O., Nascimento, M.H.C., Rainha, K.P., Oliveira, E.C.S., Neto, Á.C., Castro, E.V.R. & Filgueiras, P.R. (2020). Different strategies for the use of random forest in NMR spectra. *Journal of Chemometrics*. 34 (12). p.pp. 1–10.
- Luo, P., Yin, P., Zhang, W., Zhou, L., Lu, X., Lin, X. & Xu, G. (2016). Optimization of large-scale pseudotargeted metabolomics method based on liquid chromatography-mass spectrometry. *Journal of Chromatography A*. [Online]. 1437. p.pp. 127–136. Available from: <http://dx.doi.org/10.1016/j.chroma.2016.01.078>.
- Mahlein, A.K., Kuska, M.T., Behmann, J., Polder, G. & Walter, A. (2018). Hyperspectral sensors and imaging technologies in phytopathology: State of the art. *Annual Review of Phytopathology*. 56. p.pp. 535–558.
- Mahlein, A.K., Kuska, M.T., Thomas, S., Wahabzada, M., Behmann, J., Rascher, U. & Kersting, K. (2019). Quantitative and qualitative phenotyping of disease resistance of crops by hyperspectral sensors: seamless interlocking of phytopathology, sensors, and machine learning is needed! *Current Opinion in Plant Biology*. [Online]. 50 (August). p.pp. 156–162. Available from: <https://doi.org/10.1016/j.pbi.2019.06.007>.
- Martelli, G.P. (2016). The current status of the quick decline syndrome of olive in southern Italy. *Phytoparasitica*. 44 (1). p.pp. 1–10.
- Martelli, G.P., Boscia, D., Porcelli, F. & Saponari, M. (2016). The olive quick decline syndrome in south-east Italy: a threatening phytosanitary emergency. *European Journal of Plant Pathology*. 144 (2). p.pp. 235–243.
- Martens, H. (2021). Interpretable machine learning with an eye for the physics: Hyperspectral Vis/NIR “video” of drying wood analyzed by hybrid subspace modeling. *NIR news*. 32 (7–8). p.pp. 24–32.
- Martens, H. (2015). Quantitative Big Data: Where chemometrics can contribute. *Journal of Chemometrics*. 29 (11). p.pp. 563–581.

- Martinelli, F., Marchese, A., Giovino, A., Marra, F.P., Noce, I. Della, Caruso, T. & Dandekar, A.M. (2019). In-field and early detection of xylella fastidiosa infections in olive using a portable instrument. *Frontiers in Plant Science*. 9 (January). p.pp. 1–5.
- Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., Villa, P., Stroppiana, D., Boschetti, M., Goulart, L.R., Davis, C.E. & Dandekar, A.M. (2015). Advanced methods of plant disease detection. A review. *Agronomy for Sustainable Development*. 35 (1). p.pp. 1–25.
- McManus, K.M., Asner, G.P., Martin, R.E., Dexter, K.G., Kress, W.J. & Field, C.B. (2016). Phylogenetic structure of foliar spectral traits in tropical forest canopies. *Remote Sensing*. 8 (3). p.pp. 1–16.
- Mithöfer, A., Schulze, B. & Boland, W. (2004). Biotic and heavy metal stress response in plants: Evidence for common signals. *FEBS Letters*. 566 (1–3). p.pp. 1–5.
- Mohanpuria, P., Rana, N.K. & Yadav, S.K. (2007). Cadmium induced oxidative stress influence on glutathione metabolic genes of *Camellia sinensis* (L.) O. Kuntze. *Environmental Toxicology*. [Online]. 22 (4). p.pp. 368–374. Available from: <http://doi.wiley.com/10.1002/tox.20273>.
- Moriana, A., Villalobos, F.J. & Fereres, E. (2002). Stomatal and photosynthetic responses of olive (*Olea europaea* L.) leaves to water deficits. *Plant, Cell and Environment*. 25 (3). p.pp. 395–405.
- Morin-Crini, N., Loiacono, S., Placet, V., Torri, G., Bradu, C., Kostić, M., Cosentino, C., Chanet, G., Martel, B., Lichtfouse, E. & Crini, G. (2019). Hemp-based adsorbents for sequestration of metals: a review. *Environmental Chemistry Letters*. [Online]. 17 (1). p.pp. 393–408. Available from: <https://doi.org/10.1007/s10311-018-0812-x>.
- Murray, M.G. & Thompson, W.F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research*. [Online]. 8 (19). p.pp. 4321–4326. Available from: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/8.19.4321>.
- Musio, B., Ahmed, E.M.F.M.H., Antonicelli, M., Chiapperini, D., Dursi, O., Grieco, F., Latronico, M., Mastroilli, P., Ragone, R., Settanni, R., Triggiani, M. & Gallo, V. (2022). A spectroscopic study to assess heavy metals absorption by a combined hemp/spirulina system from contaminated soil. *Environmental Advances*. [Online]. 7. p.p. 100144. Available from: <https://doi.org/10.1016/j.envadv.2021.100144>.

- Musio, B., Ragone, R., Todisco, S., Rizzuti, A., Latronico, M., Mastrorilli, P., Pontrelli, S., Intini, N., Scapicchio, P., Triggiani, M., Di Noia, T., Acquotti, D., Airoldi, C., Assfalg, M., Barge, A., Bateman, L., Benevelli, F., Bertelli, D., Bertocchi, F., Bieliauskas, A., Borioni, A., Caligiani, A., Callone, E., Čamra, A., Cesare Marincola, F., Chalasani, D., Consonni, R., Dambruoso, P., Davalli, S., David, T., Diehl, B., Donarski, J., Gil, A.M., Gobetto, R., Goldoni, L., Hamon, E., Harwood, J.S., Kobrlová, A., Longobardi, F., Luisi, R., Mallamace, D., Mammi, S., Martin-Biran, M., Mazzei, P., Mele, A., Milone, S., Molero Vilchez, D., Mulder, R.J., Napoli, C., Ragno, D., Randazzo, A., Rossi, M.C., Rotondo, A., Šačkus, A., Sáez Barajas, E., Schievano, E., Sitaram, B., Stevanato, L., Takis, P.G., Teipel, J., Thomas, F., Torregiani, E., Valensin, D., Veronesi, M., Warren, J., Wist, J., Zailer-Hafer, E., Zuccaccia, C. & Gallo, V. (2020). A community-built calibration system: The case study of quantification of metabolites in grape juice by qNMR spectroscopy. *Talanta*. [Online]. 214 (February). p.p. 120855. Available from: <https://doi.org/10.1016/j.talanta.2020.120855>.
- Nalimova, A.A., Popova, V. V., Tsoglin, L.N. & Pronina, N.A. (2005). The effects of copper and zinc on *Spirulina platensis* growth and heavy metal accumulation in its cells. *Russian Journal of Plant Physiology*. 52 (2). p.pp. 229–234.
- Nevius, T.A. & Pardue, H.L. (1984). Development and Preliminary Evaluation of Modified Savitzky-Golay Smoothing Functions. *Analytical Chemistry*. 56 (12). p.pp. 2249–2251.
- Nguyen, C., Sagan, V., Maimaitiyiming, M., Maimaitijiang, M., Bhadra, S. & Kwasniewski, M.T. (2021). Early detection of plant viral disease using hyperspectral imaging and deep learning. *Sensors (Switzerland)*. 21 (3). p.pp. 1–23.
- Nichani, K., Uhlig, S., Stoyke, M., Kemmlein, S., Ulberth, F., Haase, I., Döring, M., Walch, S.G. & Gowik, P. (2021). *Essential terminology and considerations for validation of non-targeted methods*. 17 (December). p.pp. 1–22.
- Nicholson, J.K., Lindon, J.C. & Holmes, E. (1999). ‘Metabonomics’: Understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*. 29 (11). p.pp. 1181–1189.
- Nielsen, J. (2017). Systems biology of metabolism. *Annual Review of Biochemistry*. 86. p.pp. 245–275.
- Nunney, L., Elfekih, S. & Stouthamer, R. (2012a). The importance of multilocus sequence typing: Cautionary tales from the bacterium *Xylella fastidiosa*. *Phytopathology*. 102 (5). p.pp. 456–462.

- Nunney, L., Yuan, X., Bromley, R.E. & Stouthamer, R. (2012b). Detecting genetic introgression: High levels of intersubspecific recombination found in *Xylella fastidiosa* in Brazil. *Applied and Environmental Microbiology*. 78 (13). p.pp. 4702–4714.
- OEPP/EPPO (2018). *EPPO A1 and A2 lists of pests recommended for regulation as quarantine pests*.
- OEPP/EPPO (2019). PM 7/24 (4) *Xylella fastidiosa*. *Bulletin OEPP/EPPO Bulletin*. 49 (2). p.pp. 175–227.
- Owomugisha, G., Nuwamanya, E., Quinn, J.A., Biehl, M. & Mwebaze, E. (2020). Early detection of plant diseases using spectral data. *ACM International Conference Proceeding Series*.
- Pacumbaba, R.O. & Beyl, C.A. (2011). Changes in hyperspectral reflectance signatures of lettuce leaves in response to macronutrient deficiencies. *Advances in Space Research*. [Online]. 48 (1). p.pp. 32–42. Available from: <http://dx.doi.org/10.1016/j.asr.2011.02.020>.
- Pandey, N. & Sharma, C.P. (2002). Effect of heavy metals Co^{2+} , Ni^{2+} and Cd^{2+} on growth and metabolism of cabbage. *Plant Science*. 163 (4). p.pp. 753–758.
- Pane, C., Manganiello, G., Nicastro, N. & Carotenuto, F. (2022). *Early Detection of Wild Rocket Tracheofusariosis Using Hyperspectral Image-Based Machine Learning*.
- Parsons, H.M., Ekman, D.R., Collette, T.W. & Viant, M.R. (2009). Spectral relative standard deviation: A practical benchmark in metabolomics. *Analyst*. 134 (3). p.pp. 478–485.
- De Pascali, M., Vergine, M., Negro, C., Greco, D., Vita, F., Sabella, E., De Bellis, L. & Luvisi, A. (2022). *Xylella fastidiosa* and Drought Stress in Olive Trees: A Complex Relationship Mediated by Soluble Sugars. *Biology*. 11 (1).
- Paulus, S. & Mahlein, A.-K. (2020). Technical workflows for hyperspectral plant image assessment and processing on the greenhouse and laboratory scale. *GigaScience*. 9 (8). p.pp. 1–10.
- Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*. [Online]. 2 (11). p.pp. 559–572. Available from: <https://www.tandfonline.com/doi/full/10.1080/14786440109462720>.
- Picone, G., Savorani, F., Trimigno, A., Mezzetti, B., Capozzi, F. & Engelsen, S.B. (2016). Metabolic changes of genetically engineered grapes (*Vitis vinifera* L.) studied by $^1\text{H-NMR}$, metabolite heatmaps and iPLS. *Metabolomics*. [Online]. 12 (10). p.p. 150. Available from: <http://link.springer.com/10.1007/s11306-016-1095-5>.

- Pinu, F.R., Beale, D.J., Paten, A.M., Kouremenos, K., Swarup, S., Schirra, H.J. & Wishart, D. (2019). Systems biology and multi-omics integration: Viewpoints from the metabolomics research community. *Metabolites*. 9 (4). p.pp. 1–31.
- Poblete, T., Camino, C., Beck, P.S.A., Hornero, A., Kattenborn, T., Saponari, M., Boscia, D., Navas-Cortes, J.A. & Zarco-Tejada, P.J. (2020). Detection of *Xylella fastidiosa* infection symptoms with airborne multispectral and thermal imagery: Assessing bandset reduction performance from hyperspectral analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*. [Online]. 162 (February). p.pp. 27–40. Available from: <https://doi.org/10.1016/j.isprsjprs.2020.02.010>.
- Powers, R. & Riekeberg, E. (2017). New frontiers in metabolomics: From measurement to insight. *F1000Research*. 6.
- Praspaliauskas, M., Žaltauskaitė, J., Pedišius, N. & Striūgas, N. (2020). Comprehensive evaluation of sewage sludge and sewage sludge char soil amendment impact on the industrial hemp growth performance and heavy metal accumulation. *Industrial Crops and Products*. [Online]. 150 (December 2019). p.p. 112396. Available from: <https://doi.org/10.1016/j.indcrop.2020.112396>.
- Purcell, A. (2013). Paradigms: Examples from the bacterium *xylella fastidiosa*. *Annual Review of Phytopathology*. 51. p.pp. 339–356.
- Purcell, A.H. & Saunders, S.R. (1999). Fate of Pierce’s disease strains of *Xylella fastidiosa* in common riparian plants in California. *Plant Disease*. 83 (9). p.pp. 825–830.
- Purcell, A.H., Saunders, S.R., Hendson, M., Grebus, M.E. & Henry, M.J. (1999). Causal role of *Xylella fastidiosa* in Oleander leaf scorch disease. *Phytopathology*. 89 (1). p.pp. 53–58.
- Purcell, E.M., Torrey, H.C. & Pound, R. V. (1946). Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Physical Review*. [Online]. 69 (1–2). p.pp. 37–38. Available from: <https://link.aps.org/doi/10.1103/PhysRev.69.37>.
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. [Online]. Available from: <https://www.r-project.org>.
- Rabi, I.I., Zacharias, J.R., Millman, S. & Kusch, P. (1938). A New Method of Measuring Nuclear Magnetic Moment. *Physical Review*. [Online]. 53 (4). p.pp. 318–318. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0031891439800040>.

- Rahman, H., Sabreen, S., Alam, S. & Kawai, S. (2005). Effects of Nickel on Growth and Composition of Metal Micronutrients in Barley Plants Grown in Nutrient Solution. *Journal of Plant Nutrition*. [Online]. 28 (3). p.pp. 393–404. Available from: <http://www.tandfonline.com/doi/abs/10.1081/PLN-200049149>.
- Ramsden, J. (2015). Metabolomics and Metabonomics. In: J. Ramsden (ed.). *Bioinformatics An Introduction*. [Online]. Springer London, pp. 265–270. Available from: http://link.springer.com/10.1007/978-1-4471-6702-0_18.
- Razzaq, A., Wishart, D.S., Wani, S.H., Hameed, M.K., Mubin, M. & Saleem, F. (2022). Advances in Metabolomics-Driven Diagnostic Breeding and Crop Improvement. *Metabolites*. 12 (6). p.pp. 1–22.
- Rey, B., Aleixos, N., Cubero, S. & Blasco, J. (2019). XF-ROVIM. A field robot to detect olive trees infected by *Xylella fastidiosa* using proximal sensing. *Remote Sensing*. 11 (3).
- Rezaei, H. (2016). Biosorption of chromium by using *Spirulina* sp. *Arabian Journal of Chemistry*. [Online]. 9 (6). p.pp. 846–853. Available from: <http://dx.doi.org/10.1016/j.arabjc.2013.11.008>.
- Rizzuti, A., Aguilera-Saez, L., Santoro, F., Valentini, F., Gualano, S., D’Onghia, A.M., Gallo, V., Mastroilli, P. & Latronico, M. (2018). Detection of *Erwinia amylovora* in pear leaves using a combined approach by hyperspectral reflectance and nuclear magnetic resonance spectroscopy. *Phytopathologia Mediterranea*. [Online]. 57 (2 SE-Research Papers). p.pp. 241–252. Available from: <https://oajournals.fupress.net/index.php/pm/article/view/5768>.
- Rosato, A., Tenori, L., Cascante, M., De Atauri Carulla, P.R., Martins dos Santos, V.A.P. & Saccenti, E. (2018). From correlation to causation: analysis of metabolomics data using systems biology approaches. *Metabolomics*. [Online]. 14 (4). p.p. 37. Available from: <http://dx.doi.org/10.1007/s11306-018-1335-y>.
- RStudio Team (2021). *RStudio: Integrated Development Environment for R*.
- Rumpf, T., Mahlein, A.K., Steiner, U., Oerke, E.C., Dehne, H.W. & Plümer, L. (2010). Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. *Computers and Electronics in Agriculture*. [Online]. 74 (1). p.pp. 91–99. Available from: <http://dx.doi.org/10.1016/j.compag.2010.06.009>.
- Ryan, D. & Robards, K. (2006). Metabolomics: The greatest omics of them all? *Analytical Chemistry*. 78 (23). p.pp. 7954–7958.

- Sabella, E., Aprile, A., Genga, A., Siciliano, T., Nutricati, E., Nicoli, F., Vergine, M., Negro, C., De Bellis, L. & Luvisi, A. (2019). Xylem cavitation susceptibility and refilling mechanisms in olive trees infected by *Xylella fastidiosa*. *Scientific Reports*. 9. p.pp. 1–11.
- Sabella, E., Moretti, S., Gärtner, H., Luvisi, A., De Bellis, L., Vergine, M., Saurer, M. & Cherubini, P. (2020). Increase in ring width, vessel number and $\delta^{18}\text{O}$ in olive trees infected with *Xylella fastidiosa*. *Tree Physiology*. [Online]. 40 (11). p.pp. 1583–1594. Available from: <https://academic.oup.com/treephys/article/40/11/1583/5875780>.
- Saponari, M., Boscia, D., Altamura, G., D’Attoma, G., Cavalieri, V., Zicca, S., Morelli, M., Tavano, D., Loconsole, G., Susca, L., Potere, O., Savino, V., Martelli, G.P., Palmisano, F., Dongiovanni, C., Saponari, A., Fumarola, G. & Carolo, M. Di (2016). Pilot project on *Xylella fastidiosa* to reduce risk assessment uncertainties. *EFSA Supporting Publications*. 13 (3).
- Saponari, M., Boscia, D., Altamura, G., Loconsole, G., Zicca, S., D’Attoma, G., Morelli, M., Palmisano, F., Saponari, A., Tavano, D., Savino, V.N., Dongiovanni, C. & Martelli, G.P. (2017). Isolation and pathogenicity of *Xylella fastidiosa* associated to the olive quick decline syndrome in southern Italy. *Scientific Reports*. 7 (1). p.pp. 1–13.
- Saponari, M., Boscia, D., Nigro, F. & Martelli, G.P. (2013). Identification of DNA sequences related to *Xylella fastidiosa* in Oleander, almond and olive trees exhibiting leaf scorch symptoms in Apulia (Southern Italy). *Journal of Plant Pathology*. 95 (3). p.p. 668.
- Saponari, M., Giampetruzzi, A., Loconsole, G., Boscia, D. & Saldarelli, P. (2019). *Xylella fastidiosa* in olive in apulia: Where we stand. *Phytopathology*. 109 (2). p.pp. 175–186.
- Savary, P., Foltête, J., Moal, H., Vuidel, G. & Garnier, S. (2021). graph4lg: A package for constructing and analysing graphs for landscape genetics in R O. Gaggiotti (ed.). *Methods in Ecology and Evolution*. [Online]. 12 (3). p.pp. 539–547. Available from: <https://onlinelibrary.wiley.com/doi/10.1111/2041-210X.13530>.
- Savitzky, A. & Golay, M.J.E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*. [Online]. 36 (8). p.pp. 1627–1639. Available from: <https://pubs.acs.org/doi/abs/10.1021/ac60214a047>.
- Schmidt-Böcking, H., Schmidt, L., Lüdde, H.J., Trageser, W., Templeton, A. & Sauer, T. (2016). The Stern-Gerlach experiment revisited. *The European Physical Journal H*. [Online]. 41 (4–5). p.pp. 327–364. Available from: <http://link.springer.com/10.1140/epjh/e2016-70053-2>.

- Schneider, K., van der Werf, W., Cendoya, M., Mourits, M., Navas-Cortés, J.A., Vicent, A. & Lansink, A.O. (2020). Impact of *Xylella fastidiosa* subspecies pauca in European olives. *Proceedings of the National Academy of Sciences of the United States of America*. 117 (17). p.pp. 9250–9259.
- Schönberger, T., Monakhova, Y.B., Lachenmeier, D.W., Walch, S., Kuballa, T. & NEXT-NMR-working, N.-P.E.T. (NEXT)-N. working group G. (2016). *Guide to NMR Method Development and Validation – Part II : Multivariate data analysis*.
- Schuenzel, E.L., Scally, M., Stouthamer, R. & Nunney, L. (2005). A multigene phylogenetic study of clonal diversity and divergence in North American strains of the plant pathogen *Xylella fastidiosa*. *Applied and Environmental Microbiology*. 71 (7). p.pp. 3832–3839.
- Scoccianti, V., Crinelli, R., Tirillini, B., Mancinelli, V. & Speranza, A. (2006). Uptake and toxicity of Cr(III) in celery seedlings. *Chemosphere*. [Online]. 64 (10). p.pp. 1695–1703. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0045653506000488>.
- Scortichini, M. (2022). The Epidemiology and Control of “Olive Quick Decline Syndrome” in Salento (Apulia, Italy). *Agronomy*. 12 (10).
- Scortichini, M. (2020). The multi-millennial olive agroecosystem of salento (Apulia, Italy) threatened by *xylella fastidiosa* subsp. Pauca: A working possibility of restoration. *Sustainability (Switzerland)*. 12 (17). p.pp. 1–15.
- Şeker, A., Shahwan, T., Eroğlu, A.E., Yılmaz, S., Demirel, Z. & Dalay, M.C. (2008). Equilibrium, thermodynamic and kinetic studies for the biosorption of aqueous lead(II), cadmium(II) and nickel(II) ions on *Spirulina platensis*. *Journal of Hazardous Materials*. [Online]. 154 (1–3). p.pp. 973–980. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0304389407016020>.
- Sharma, P. & Dubey, R.S. (2005). Lead toxicity in plants. *Brazilian Journal of Plant Physiology*. [Online]. 17 (1). p.pp. 35–52. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1677-04202005000100004&lng=en&tlng=en.
- Sharma, S.S. & Dietz, K.J. (2009). The relationship between metal toxicity and cellular redox imbalance. *Trends in Plant Science*. 14 (1). p.pp. 43–50.
- Sicard, A., Zeilinger, A.R., Vanhove, M., Schartel, T.E., Beal, D.J., Daugherty, M.P. & Almeida, R.P.P. (2018). *Xylella fastidiosa*: Insights into an emerging plant pathogen. *Annual Review of Phytopathology*. 56 (June). p.pp. 181–202.
- Singh, S., Parihar, P., Singh, R., Singh, V.P. & Prasad, S.M. (2016). Heavy metal tolerance in plants: Role of transcriptomics, proteomics, metabolomics, and ionomics. *Frontiers in Plant Science*. 6 (FEB2016). p.pp. 1–36.

- Skodra, C., Michailidis, M., Dasenaki, M., Ganopoulos, I., Thomaidis, N.S., Tanou, G. & Molassiotis, A. (2021). Unraveling salt-responsive tissue-specific metabolic pathways in olive tree. *Physiologia Plantarum*. [Online]. 173 (4). p.pp. 1643–1656. Available from: <https://doi.org/10.1111/ppl.13565>.
- Small, E. & Beckstead, H.D. (1973). Common cannabinoid phenotypes in 350 stocks of *Cannabis*. *Lloydia*.
- Stryeck, S., Horvath, A., Leber, B., Stadlbauer, V. & Madl, T. (2018). NMR spectroscopy enables simultaneous quantification of carbohydrates for diagnosis of intestinal and gastric permeability. *Scientific Reports*. 8 (1). p.pp. 1–8.
- Sumner, L.W., Amberg, A., Barrett, D., Beale, M.H., Beger, R., Daykin, C.A., Fan, T.W.M., Fiehn, O., Goodacre, R., Griffin, J.L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A.N., Lindon, J.C., Marriott, P., Nicholls, A.W., Reily, M.D., Thaden, J.J. & Viant, M.R. (2007). Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics*. 3 (3). p.pp. 211–221.
- Taha, K., Sharma, A., Kroeker, K., Ross, C., Carleton, B., Wishart, D., Medeiros, M. & Blydt-Hansen, T.D. (2022). Noninvasive testing for mycophenolate exposure in children with renal transplant using urinary metabolomics. *Pediatric Transplantation*. (November). p.pp. 1–11.
- Tang, F. & Hatzakis, E. (2020). NMR-Based Analysis of Pomegranate Juice Using Untargeted Metabolomics Coupled with Nested and Quantitative Approaches. *Analytical Chemistry*. 92 (16). p.pp. 11177–11185.
- Tanner, F., Tonn, S., de Wit, J., Van den Ackerveken, G., Berger, B. & Plett, D. (2022). Sensor-based phenotyping of above-ground plant-pathogen interactions. *Plant Methods*. [Online]. 18 (1). p.pp. 1–18. Available from: <https://doi.org/10.1186/s13007-022-00853-7>.
- Taraji, M., Haddad, P.R., Amos, R.I.J., Talebi, M., Szucs, R., Dolan, J.W. & Pohl, C.A. (2017). Prediction of retention in hydrophilic interaction liquid chromatography using solute molecular descriptors based on chemical structures. *Journal of Chromatography A*. [Online]. 1486. p.pp. 59–67. Available from: <http://dx.doi.org/10.1016/j.chroma.2016.12.025>.
- Tardieu, F., Cabrera-Bosquet, L., Pridmore, T. & Bennett, M. (2017). Plant Phenomics, From Sensors to Knowledge. *Current Biology*. 27 (15). p.pp. R770–R783.

- Thomas, J.C., Malick, F.K., Endreszl, C., Davies, E.C. & Murray, K.S. (1998). Distinct responses to copper stress in the halophyte *Mesembryanthemum crystallinum*. *Physiologia Plantarum*. [Online]. 102 (3). p.pp. 360–368. Available from: <http://doi.wiley.com/10.1034/j.1399-3054.1998.1020304.x>.
- Tortorella, S. & Cinti, S. (2021). How Can Chemometrics Support the Development of Point of Need Devices? *Analytical Chemistry*. 93 (5). p.pp. 2713–2722.
- Triba, M.N., Le Moyec, L., Amathieu, R., Goossens, C., Bouchemal, N., Nahon, P., Rutledge, D.N. & Savarin, P. (2015). PLS/OPLS models in metabolomics: The impact of permutation of dataset rows on the K-fold cross-validation quality parameters. *Molecular BioSystems*. [Online]. 11 (1). p.pp. 13–19. Available from: <http://dx.doi.org/10.1039/C4MB00414K>.
- Tripathi, R.D., Dwivedi, S., Shukla, M.K., Mishra, S., Srivastava, S., Singh, R., Rai, U.N. & Gupta, D.K. (2008). Role of blue green algae biofertilizer in ameliorating the nitrogen demand and fly-ash stress to the growth and yield of rice (*Oryza sativa* L.) plants. *Chemosphere*. 70 (10). p.pp. 1919–1929.
- Ustin, S.L., Roberts, D.A., Gamon, J.A., Asner, G.P. & Green, R.O. (2004). Using imaging spectroscopy to study ecosystem processes and properties. *BioScience*. 54 (6). p.pp. 523–534.
- Vandenhove, H. & Van Hees, M. (2005). Fibre crops as alternative land use for radioactively contaminated arable land. *Journal of Environmental Radioactivity*. 81 (2–3). p.pp. 131–141.
- Vardhan, K.H., Kumar, P.S. & Panda, R.C. (2019). A review on heavy metal pollution, toxicity and remedial measures: Current trends and future perspectives. *Journal of Molecular Liquids*. [Online]. 290. p.p. 111197. Available from: <https://doi.org/10.1016/j.molliq.2019.111197>.
- Vitale, R., Ruckebusch, C., Burud, I. & Martens, H. (2022). Hyperspectral Video Analysis by Motion and Intensity Preprocessing and Subspace Autoencoding. *Frontiers in Chemistry*. 10 (March). p.pp. 1–17.
- Wahabzada, M., Mahlein, A.K., Bauckhage, C., Steiner, U., Oerke, E.C. & Kersting, K. (2016). Plant Phenotyping using Probabilistic Topic Models: Uncovering the Hyperspectral Language of Plants. *Scientific Reports*. [Online]. 6 (September 2015). p.pp. 1–11. Available from: <http://dx.doi.org/10.1038/srep22482>.
- Wang, H.-P., Chen, P., Dai, J.-W., Liu, D., Li, J.-Y., Xu, Y.-P. & Chu, X.-L. (2022). Recent advances of chemometric calibration methods in modern spectroscopy: Algorithms, strategy, and related issues. *TrAC Trends in Analytical Chemistry*. [Online]. 153. p.p. 116648. Available from: <https://doi.org/10.1016/j.trac.2022.116648>.

- Watanabe, A., Nishigaki, S. & Konishi, C. (1951). Effect of nitrogen-fixing blue-green algae on the growth of rice plants. *Nature*. [Online]. 168 (4278) p.pp. 748–749. Available from: <http://www.nature.com/articles/168748b0>.
- Weiss, M., Jacob, F. & Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review. *Remote Sensing of Environment*. [Online]. 236 (December 2018). p.p. 111402. Available from: <https://doi.org/10.1016/j.rse.2019.111402>.
- Wells, J.M., Raju, B.C., Hung, H., Weisburg, W.G., Mandelco-paul, L. & Brenner, D.O.N.J. (1987). Limited , Fastidious Plant Bacteria Related to *Xanthomonas* spp . *International Journal of Systematic Bacteriology*. 37 (2). p.pp. 136–143.
- Wheelock, Å.M. & Wheelock, C.E. (2013). Trials and tribulations of 'omics data analysis: Assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine. *Molecular BioSystems*. 9 (11). p.pp. 2589–2596.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemond, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K. & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*. 4 (43). p.p. 1686.
- Wishart, D.S. (2011). Advances in metabolite identification. *Bioanalysis*. 3 (15). p.pp. 1769–1782.
- Wishart, D.S. (2008a). Applications of metabolomics in drug discovery and development. *Drugs in R and D*. 9 (5). p.pp. 307–322.
- Wishart, D.S. (2016). Emerging applications of metabolomics in drug discovery and precision medicine. *Nature Reviews Drug Discovery*. [Online]. 15 (7). p.pp. 473–484. Available from: <http://www.nature.com/articles/nrd.2016.32>.
- Wishart, D.S. (2008b). Quantitative metabolomics using NMR. *TrAC - Trends in Analytical Chemistry*. 27 (3). p.pp. 228–237.
- Wishart, D.S., Cheng, L.L., Copié, V., Edison, A.S., Eghbalnia, H.R., Hoch, J.C., Gouveia, G.J., Pathmasiri, W., Powers, R., Schock, T.B., Sumner, L.W. & Uchimiyama, M. (2022a). NMR and Metabolomics—A Roadmap for the Future. *Metabolites*. [Online]. 12 (8). p.p. 678. Available from: <https://www.mdpi.com/2218-1989/12/8/678>.
- Wishart, D.S., Rout, M., Lee, B.L., Berjanskii, M., LeVatte, M. & Lipfert, M. (2022b). Practical Aspects of NMR-Based Metabolomics. In: *Handbook of Experimental Pharmacology*. [Online]. Springer, Berlin, Heidelberg. Available from: https://link.springer.com/10.1007/164_2022_613.

- Wojcik, M. & Tukiendorf, A. (2004). Phytochelatin synthesis and cadmium localization in wild type of *Arabidopsis thaliana*. *Plant Growth Regulation*. [Online]. 44 (1). p.pp. 71–80. Available from: <http://link.springer.com/10.1007/s10725-004-1592-9>.
- Wold, S., Esbensen, K. & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*. [Online]. 2 (1–3). p.pp. 37–52. Available from: <https://linkinghub.elsevier.com/retrieve/pii/0169743987800849>.
- Wuang, S.C., Khin, M.C., Chua, P.Q.D. & Luo, Y.D. (2016). Use of *Spirulina* biomass produced from treatment of aquaculture wastewater as agricultural fertilizers. *Algal Research*. [Online]. 15. p.pp. 59–64. Available from: <http://dx.doi.org/10.1016/j.algal.2016.02.009>.
- Yadav, S.K. (2010). Heavy metals toxicity in plants: An overview on the role of glutathione and phytochelatins in heavy metal stress tolerance of plants. *South African Journal of Botany*. [Online]. 76 (2). p.pp. 167–179. Available from: <http://dx.doi.org/10.1016/j.sajb.2009.10.007>.
- Ye, C., Mu, D., Horowitz, N., Xue, Z., Chen, J., Xue, M., Zhou, Y., Klutts, M. & Zhou, W. (2018). Life cycle assessment of industrial scale production of spirulina tablets. *Algal Research*. [Online]. 34 (February). p.pp. 154–163. Available from: <https://doi.org/10.1016/j.algal.2018.07.013>.
- Yi, L., Dong, N., Yun, Y., Deng, B., Ren, D., Liu, S. & Liang, Y. (2016). Chemometric methods in data processing of mass spectrometry-based metabolomics: A review. *Analytica Chimica Acta*. [Online]. 914. p.pp. 17–34. Available from: <http://dx.doi.org/10.1016/j.aca.2016.02.001>.
- Zarco-Tejada, P.J., Camino, C., Beck, P.S.A., Calderon, R., Hornero, A., Hernández-Clemente, R., Kattenborn, T., Montes-Borrego, M., Susca, L., Morelli, M., Gonzalez-Dugo, V., North, P.R.J., Landa, B.B., Boscia, D., Saponari, M. & Navas-Cortes, J.A. (2018). Previsual symptoms of *Xylella fastidiosa* infection revealed in spectral plant-trait alterations. *Nature Plants*. [Online]. 4 (7). p.pp. 432–439. Available from: <http://dx.doi.org/10.1038/s41477-018-0189-7>.
- Zarco-Tejada, P.J., Poblete, T., Camino, C., Gonzalez-Dugo, V., Calderon, R., Hornero, A., Hernandez-Clemente, R., Román-Écija, M., Velasco-Amo, M.P., Landa, B.B., Beck, P.S.A., Saponari, M., Boscia, D. & Navas-Cortes, J.A. (2021). Divergent abiotic spectral pathways unravel pathogen stress signals across species. *Nature Communications*. [Online]. 12 (1). p.p. 6088. Available from: <https://www.nature.com/articles/s41467-021-26335-3>.
- Zheng, C., Abd-elrahman, A. & Whitaker, V. (2021a). Remote sensing and machine learning in crop phenotyping and management, with an emphasis on applications in strawberry farming. *Remote Sensing*. 13 (3). p.pp. 1–29.

- Zheng, J., Johnson, M., Mandal, R. & Wishart, D.S. (2021b). A comprehensive targeted metabolomics assay for crop plant sample analysis. *Metabolites*. 11 (5).
- Zielonka, D., Szulc, W., Skowrońska, M., Rutkowska, B. & Russel, S. (2020). Hemp-based phytoaccumulation of heavy metals from municipal sewage sludge and phosphogypsum under field conditions. *Agronomy*. 10 (6).
- Zinicovscaia, I., Duca, G., Cepoi, L., Chiriac, T., Rudi, L., Mitina, T., Frontasyeva, M. V., Pavlov, S. & Gundorina, S.F. (2015). Biotechnology of Metal Removal from Industrial Wastewater: Zinc Case Study. *Clean - Soil, Air, Water*. 43 (1). p.pp. 112–117.
- Zinicovscaia, I., Safonov, A., Ostalkevich, S., Gundorina, S., Nekhoroshkov, P. & Grozdov, D. (2019). Metal ions removal from different type of industrial effluents using *Spirulina platensis* biomass. *International Journal of Phytoremediation*. [Online]. 21 (14). p.pp. 1442–1448. Available from: <https://doi.org/10.1080/15226514.2019.1633264>.

CURRICULUM

Elhussein Mohamed Fouad Mourad Hussein AHMED

Work: Via E. Orabona 4, 70125, Bari, Italy

Email: elhussein.ahmed@poliba.it **Phone:** (+39) 3883986958

Gender: Male **Date of birth:** 06/12/1992 **Nationality:** Egyptian

ABOUT ME: Elhussein F. Mourad



WORK EXPERIENCE

Undergraduate trainee [07/2011 – 01/2014]

Environmental Studies and Research Unit (ESRU), Faculty of Agriculture, Cairo University. City: Giza. **Country:** Egypt

Research assistant [02/2014 – 10/2018]

Environmental Studies and Research Unit (ESRU), Faculty of Agriculture, Cairo University. City: Giza. **Country:** Egypt

Mandatory military service [02/2015 – 02/2016]

Egyptian Air Force. Country: Egypt.

Research internship [28/08/2022 – 15/12/2022]

NTNU - Norwegian University of Science and Technology City: Trondheim.
Country: Norway

EDUCATION AND TRAINING

BSc in Biotechnology (English Program) [10/2009 – 01/2014]

Faculty of Agriculture, Cairo University. Address: Giza, Egypt

Diploma in “Sustainable IPM technologies for Mediterranean fruit and vegetable crops” [10/2017 – 06/2018]

Libera Università Mediterranea / CIHEAM Bari. Address: Via Ceglie 9, 70010, Valenzano (BA), Italy

MSc in Agricultural Microbiology [03/2016 – 07/2019]

Faculty of Agriculture, Cairo University. Address: Giza, Egypt

MSc in “Sustainable IPM technologies for Mediterranean fruit and vegetable crops” [11/2018 – 10/2019]

CIHEAM BARI. Address: Via Ceglie 9, 70010, Valenzano (BA), Italy

PhD in “Risk and Environmental, Territorial and Building Development” (CHIM/07) [11/2019 – Current]

Polytechnic University of Bari. Address: Via E. Orabona 4, 70125, Bari, Italy

LANGUAGE SKILLS

Mother tongue(s): Arabic, **Other language(s):**

English: LISTENING C2 READING C2 WRITING C2 SPOKEN PRODUCTION C2 SPOKEN INTERACTION C2

French: LISTENING A2 READING A2 WRITING A1 SPOKEN PRODUCTION A1 SPOKEN INTERACTION A1

German: LISTENING A1 READING A1 WRITING A1 SPOKEN PRODUCTION A1 SPOKEN INTERACTION A1

Italian: LISTENING B2 READING B2 WRITING B1 SPOKEN PRODUCTION B1 SPOKEN INTERACTION B1

Norwegian: LISTENING A1 READING A1 WRITING A1 SPOKEN PRODUCTION A1 SPOKEN INTERACTION A1

DIGITAL SKILLS

R Language and R Studio | Good familiarity with MATLAB, Simulink | top spin (NMR spectroscopy) | Mestrelab - Mnova | SIMCA Umetrics | Mendeley software | Inkscape, Gimp | Internet, E-mail and Social Media | Microsoft Office (Microsoft) | Linux (basic) | Multivariate data analysis in R

PUBLICATIONS

[Google Scholar Profile](#)

[Website](#)