ATR-FTIR Spectroscopy-Linked Chemometrics: A Novel Approach to the Analysis and Control of the Invasive Species Japanese Knotweed

Claire Anne Holden

This dissertation is submitted for the degree of Doctor of Philosophy

January 2023

Lancaster Environment Centre

This thesis is dedicated in loving memory to Theresa Collins.

'An understanding of the natural world is a source of not only great curiosity, but great fulfilment' – Sir David Attenborough

Declaration

This thesis has not been submitted in support of an application for another degree at this or any other university. It is the result of my own work and includes nothing that is the outcome of work done in collaboration except where specifically indicated. Many of the ideas in this thesis were the product of discussion with my supervisors Professor Martin McAinsh and Professor Jane Taylor.

Excerpts of this thesis have been published in the following conference manuscripts and academic publications.

Emsley NEM, Holden CA, Guo S, Bevan RS, Rees C, McAinsh MR, Martin FL, Morais CLM. **2021**. Machine Learning Approach Using a Handheld Near-Infrared (NIR) Device to Predict the Effect of Storage Conditions on Tomato Biomarkers. *ACS Food Science & Technology* **2(1)**: 187-194.

Holden CA, Bailey JP, Taylor JE, Martin FL, Beckett P, McAinsh M. **2022**. Know your enemy: Application of ATR-FTIR spectroscopy to invasive species control (D Changwen, Ed.). *PLOS ONE* **17**: e0261742.

Holden CA, Morais CLM, Taylor JE, Martin FL, Beckett P, McAinsh M. **2021**. Regional differences in clonal Japanese knotweed revealed by chemometrics-linked attenuated total reflection Fourier-transform infrared spectroscopy. *BMC Plant Biology 2021 21:1* **21**: 1–20.

Heap, B, Holden, CA, Taylor, JE and McAinsh, M 2020. Crosstalk in Signalling Pathways. In *eLS*. Wiley. 1–9.

Abstract

Japanese knotweed (*Reynoutria japonica*), an invasive plant species, causes negative environmental and socio-economic impacts. A female clone in the United Kingdom, its extensive rhizome system enables rapid vegetative spread. Plasticity permits this species to occupy a broad geographic range and survive harsh abiotic conditions. It is notoriously difficult to control with traditional management strategies, which include repetitive herbicide application and costly carbon-intensive rhizome excavation. This problem is complicated by crossbreeding with the closely related species, Giant knotweed (*Reynoutria sachalinensis*), to give the more vigorous hybrid, Bohemian knotweed (*Fallopia x Bohemica*) which produces viable seed. These species, hybrids, and backcrosses form a morphologically similar complex known as Japanese knotweed *'sensu lato'* and are often misidentified. The research herein explores the opportunities offered by advances in the application of attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy-linked chemometrics within plant sciences, for the identification and control of knotweed, to enhance our understanding of knotweed biology, and the potential of this technique. ATR-FTIR spectral profiles of Japanese knotweed leaf material and xylem sap samples, which include important biological absorptions due to lipids, proteins, carbohydrates, and nucleic acids, were used to: identify plants from different growing regions highlighting the plasticity of this clonal species; differentiate between related species and hybrids; and predict key physiological characteristics such as hormone concentrations and root water potential. Technical advances were made for the application of ATR-FTIR spectroscopy to plant science, including definition of the environmental factors that exert the most significant influence on spectral profiles, evaluation of sample preparation techniques, and identification of key wavenumbers for prediction of hormone concentrations and abiotic stress. The presented results cement the position of concatenated mid-infrared spectroscopy and machine learning as a powerful approach for the study of plant biology, extending its reach beyond the field of crop science to demonstrate a potential for the discrimination between and control of invasive plant species.

Acknowledgements

Contents

List of Tables

List of Figures

List of Abbreviations and Acronyms

^{2}H	deuterium labelled hydrogen
ABA	abscisic acid
ACC	1-amino-cyclopropane-1-carboxylic acid
AFLP	amplified fragment length polymorphism
ANN	artificial neural network
ANN	artificial neural networks
ATR	attenuated total reflectance
B	boron (element or atom)
C	carbon (element or atom)
C:N	carbon to nitrogen ratio
Ca	calcium (element or atom)
Ca^{2+}	calcium ions
CABI	Centre for Agriculture and Bioscience International
CC	*cis* conformation
CE	capillary electrophoresis
CEBAS	Segura Center for Edaphology and Applied Biology
CKs	cytokinins
CO$_2$e	carbon dioxide equivalent
CSIC	Spanish National Research Council
CSR	competitive, stress, ruderal
Cu	copper (element or atom)
DNA	deoxyribonucleic acid
ECSt	thylakoid electrochromic shift
EDTA	ethylenediaminetetraacetic acid
EICA	evolution of increased competitive ability
ELISA	enzyme-linked immune sorbent assay
ERH	enemy release hypothesis
Fe	iron (element or atom)
FT	Fourier transformation
FTIR	Fourier transform infrared
GA$_1$	gibberellin A1
GA$_3$	gibberellin A3
GA$_4$	gibberellin A4
GAs	gibberellins
GC	gas chromatography
gH$^+$	proton conductivity of chloroplast ATP synthase
GUI	graphical user interface
H	hydrogen (element or atom)
HCA	hierarchical cluster analysis

HPLC	high pressure liquid chromatography
HRMS	high-resolution mass spectrometry
IAA	indole-3-acetic acid
IAS	invasive alien species
IMIDA	Institute for Agro-Environmental Research and Development of Murcia
iP	isopentyl adenine
IR	infrared
IRE	internal reflection element
JA	jasmonic acid
K	potassium (element or atom)
k-NN	k-nearest neighbours
LC	liquid chromatography
LDA	linear discriminant analysis
LEF	linear electron flow
LOI	loss on ignition
LSD	least significant difference
LTR	University of Leicester Herbarium
Mg	magnesium (element or atom)
MIR	mid infrared
Mo	molybdenum (element or atom)
MOA	mechanism of action
MPa	mega Pascal (unit of pressure)
MS	mass spectrometry
N	nitrogen (element or atom)
NA	not applicable
NCDPs	no-code development platforms
NIR	near infrared
NPQt	thylakoid non-photochemical quenching (light induced photoprotection through thermal dissipation of energy)
NRCS	Natural Resources Conservation Service
NWH	novel weapons hypothesis
O	oxygen (element or atom)
P	phosphorus (element or atom)
PAR	photosynthetic active radiation
PC	principal component
PCA	principal component analysis
PCA-LDA	principal component analysis-linear discriminant analysis
Phi2	quantum yield of Photosystem II
PhiNPQ	quantum yield of non-photochemical quenching
PLS-DA	partial least squares-discriminant analysis
PLSR	partial least squares regression
PS1	photosytem one
QDA	quadratic discriminant analysis

R:FR	red to far-red light ratio
RBF	radial basis function
RICS	Royal Institute of Chartered Surveyors
RNA	ribonucleic acid
ROS	reactive oxygen species
S	sulfur (element or atom)
SA	salicylic acid
SG	Savitzky–Golay
SNR	signal-to-noise ratio
SVM	support vector machine
***t*Z**	*trans*-zeatin
***t*ZR**	*trans*-zeatin riboside
UAST	universal adaptive strategy theory
UHPLC	ultraperformance liquid chromatography
UPLC	ultra-performance liquid chromatography
USDA	United States Department of Agriculture
UV	ultraviolet
vH$^+$	steady-state rate of photon flux
Zn	zinc (element or atom)
λ (Lambda)	wavelength
Ψ$_{leaf}$	leaf water potential
Ψ$_{root}$	root water potential

List of Appendices

1 Introduction

1.1 Invasive Species

Invasive alien species (IAS) pose a growing threat to global biodiversity, accelerated by anthropogenic activities (Pyšek *et al.*, 2020). Globalisation creates opportunity for new introductions through the movement of cargo or tourism between countries (Amano *et al.*, 2016). Extreme weather events, such as hurricanes and flooding, triggered by climatic change create opportunities for the spread of invasive species (Bellard *et al.*, 2018). Additionally, global warming increases the range for invasive species introduced from milder climates, allowing them to survive closer to the poles, for example the Northwards spread of invasive Japanese knotweed into Canada (Bourchier & Van Hezewijk, 2010; Clements & DiTommaso, 2012). In their introduced range, IAS have negative impacts on the environment, the economy, and people's wellbeing (Pyšek *et al.*, 2020). These problematic effects distinguish IAS from other non-native species. There are many hypotheses for the success of IAS (Gurevitch *et al.*, 2011; van Kleunen *et al.*, 2018). Table 1.1 summarises the main hypotheses for how IAS may achieve an advantage over native plants. Of these, plasticity (VanWallendael *et al.*, 2018), novel weapons (Tucker Serniak, 2016), hybridisation (Parepa *et al.*, 2014) and fluctuating resource availability (Parepa *et al.*, 2019) have been attributed to the success of Japanese knotweed, see Section 1.2.2.

Table 1.1: A summary of the hypotheses for the success of invasive species
(Gurevitch et al., 2011; van Kleunen et al., 2018).

Hypothesis	Explanation
Darwin's naturalization	Alien species distantly related to native species are more likely to naturalise
Diversity invasibility	More species-rich communities are frequently more resistant to invasions
Ecotype	genetic variations leading to local adaptations allow alien plants to thrive in varied habitats
Enemy Release (ERH)	Alien plants may be released from predation by herbivores and pathogens, particularly from specialists that have not been co-introduced, resulting in a competitive advantage over natives.
Evolution of increased competitive ability (EICA)	Alien plants may have reduced resource allocation to defences allowing an increased allocation to growth and reproduction
Fluctuating resource availability	Increases in resource availability make plant communities more susceptible to plant invasion
Hybridisation-invasion	Interspecific hybridisation promotes invasiveness by increasing the fecundity and size of the resultant offspring
Novel weapons (NWH)	Some chemical compounds of an alien species can negatively affect naïve native species
Plasticity	Phenotypic plasticity of traits allows populations of alien species with low genetic diversity to take advantage of a wider ecological niche
Shifting defence	Alien plants can shift in allocation of resources from specialist to generalist enemy defences

1.2 Japanese knotweed

1.2.1 History

Japanese knotweed is an herbaceous perennial plant which is amongst the one-hundred worst IAS both within Europe and globally (Lowe *et al.*, 2000; Nentwig *et al.*, 2018). It is known by three Latin names: *Polygonum cuspidatum, Reynoutria japonica* and *Fallopia japonica* (Centre for Agriculture and Bioscience International (CABI), 2018) and naturally takes the role of a pioneer species where it allows for succession on the exposed slopes of volcanoes in its native Asia (Bailey, 2003). Owing to its pleasing appearance and small white flowers, Japanese knotweed was introduced to Europe in 1849 as an ornamental plant by Philip Franz Von Siebold, a German physician and botanist (Bailey & Conolly, 2000). As Europe expanded colonially, it was fashionable to collect plants from around the world for botanical gardens, and the period between 1820-1900 focused on the introduction of tropical glasshouse plants and hardy plants from Japan and North America (Arber, 1986). Japanese knotweed has since become a successful invader, occupying a broad introduced range across North America, Europe, Australia and New Zealand (Fennell *et al.*, 2018) and colonising diverse habitats including riparian wetlands, urban transport courses, and coastal areas (Richards *et al.*, 2012b; Zhang *et al.*, 2016).

Historically the species has been used for numerous purposes such as cattle fodder (Kabat *et al.*, 2006), stabilisation of railway sidings (Macfarlane & Stuart, 2011), phytoremediation (Nguyen, 1997; Michalet *et al.*, 2017), biofuel (Seppälä *et al.*, 2013; Brunerová & Brožek, 2017), and Alternative Medicine (Nawrot-Hadzik *et al.*, 2019). In addition, owing to their rapid growth, two cultivars of the related species, Giant knotweed (*Reynoutria sachalinensis*), are also used as a novel biofuel crop by the German company IGNISCUM® in combined heat and power plants and for biogas production (Veste *et al.*, 2015).

1.2.2 Biological traits

Plants employ three basic approaches to survive their environment: C-competitive, S-stress tolerant, or R-ruderal (Grime & Pierce, 2012). The traits of each plant type are summarised in Table 1.1.2.

Table 1.2: C-S-R model (Universal Adaptive Strategy Theory).

		C- competitive	S- stress tolerant	R-ruderal
Environment occupied		High nutrient, low disturbance	Low nutrient, low disturbance e.g., desert	High nutrient, High disturbance e.g., following wildfires, construction, agriculture
Main energy allocation		Growth e.g., leaf construction	Maintenance e.g., anti-herbivory defences	Reproduction e.g., Seed production
Physiological trade off	*Growth rate*	High	Low	High
	Leaf longevity	Short	Long	Short
	Seed production	Low	Low	High
	Phenotypic plasticity	High	Low	High
Typical plant trait		Spatially dynamic foraging roots and shoots	Evergreen plants with small needles	Short-statured plants, annuals, and short-lived perennials

As a competitive 'C-strategist' species, Japanese knotweed grows vigorously into tall, dense monodominant clumps or 'stands' which possess a shared underground rhizome system extending several metres outwards and over a metre downwards from the parent plant (Aguilera *et al.*, 2010; Fennell *et al.*, 2018). In its native habitat, Japanese knotweed aids succession by improving soil properties (increased stability, decreased bulk density, increased nitrogen, organic matter and water; Hirose & Tateno, 1984) and proving a protective nursery for young plants via a reduction in shoot density of the central stand known as 'central die-back' (Adachi *et al.*, 1996). Japanese knotweed is known to exhibit 'allelopathy', the release of secondary compounds which influences the soil properties and development of neighbouring plants. Although the allelopathic nature of Japanese knotweed is useful in its native range for facilitating succession (Gant & Clebsch, 1975), in its introduced range this mechanism aids invasion and provides an advantage over native vegetation (Kato-Noguchi, 2022), see 'novel weapons hypothesis' in Table 1.1.1.

In addition to excreted allelopathic chemicals, phenolic phytochemicals sequestered in leaves (such as tannins, lignin and quercetin) form part of plant anti-herbivory and anti-pathogen defences (Singh *et al.*, 2021). Japanese knotweed and its related species are particularly high in phenolics and bioactive amides, such as feruloyltyramine and most notably the antioxidant resveratrol, as well as alpha-glucosidase inhibitory compounds (Zhang *et al.*, 2021). For this reason, knotweed species are traditionally used for the treatment of inflammatory diseases, infections, skin diseases, scald, and hyperlipidemia (Nawrot-Hadzik *et al.*, 2019). Importantly, invasive plants may be phytochemically unique in their new habitats, conferring advantages such as antiherbivore, antifungal, antimicrobial and allelopathic effects (Cappuccino & Arnason, 2006).

Japanese knotweed exhibits minimal genetic variation in Central Europe (Zhang *et al.*, 2016), Norway (Holm *et al.*, 2018) and the USA (Richards *et al.*, 2012b), and exists as a female clone in the United Kingdom from a single introduction (Bailey & Conolly, 2000; Hollingsworth & Bailey, 2000a). The ability of populations with low genetic diversity such as Japanese knotweed to take advantage of a wider ecological

niche has been attributed to phenotypic plasticity (Parker *et al.*, 2003; Bossdorf *et al.*, 2005; Richards *et al.*, 2006, 2008; Van Kleunen *et al.*, 2010; Pichancourt & van Klinken, 2012; Hagenblad *et al.*, 2015; Geng *et al.*, 2016), efficient resource partitioning (Price *et al.*, 2002) and vegetative regeneration (Bímová *et al.*, 2003). Despite its genetic uniformity, Japanese knotweed exhibits a high tolerance to abiotic stress, occupying extreme environments such as salt marshes (Rouifed *et al.*, 2012) and metal-polluted soil (Michalet *et al.*, 2017; Sołtysiak, 2020). This trait renders it particularly suitable for growth on contaminated land. British Japanese knotweed is not alone in its clonality, with a high proportion of successful invasive plants sharing this genetic trait; 70% of 468 studied species from the ICUN database and 81% of the one hundred worst invasive plants (Mounger *et al.*, 2021). Asexual plants may maintain genetic variation, and thereby the capacity to adapt to changing environmental conditions, through somatic mutation or epigenetics, which allows them to bypass the meiotic resetting of epigenetic modifications (Schoen & Schultz, 2019; Mounger *et al.*, 2021).

Japanese knotweed can hybridise and backcross with similar related species such as Giant Knotweed and Russian Vine, resulting in a mix of morphologically similar hybrids, backcrosses and related species that are referred to as Japanese knotweed *sensu lato (s.l.),* meaning Japanese knotweed 'in the broad sense' (Hollingsworth & Bailey, 2000b). Hybridisation is a strategy employed by IAS to overcome a genetic bottleneck (Ellstrand & Schierenbeck, 2000). It results in 'heterosis', the production of offspring with increased 'hybrid vigour' (Hollingsworth *et al.*, 1998). Hybrid descendants may have improved traits relative to their parents such as invasiveness (Ellstrand & Schierenbeck, 2000), growth rate, reproductive success and yield (Ben-Ari & Lavi, 2012), genetic variance (Lee, 2002), and stress tolerance e.g. to herbicides (Snow *et al.*, 1999) and cold (Milne & Abbott, 2000). The vigorous hybrid Bohemian Knotweed (*Reynoutria × bohemica)* has advantages over its maternal parent *Reynoutria japonica* var. *japonica*, including the ability to produce viable seed through genetic introgression without the need for cross-breeding (Mandák *et al.*, 2003). One of the three distinct introductions of Japanese knotweed to North America is also believed to have produced viable seeds (Del Tredici, 2017b). Different species, F1 hybrids and backcrosses can be identified using their

chromosome number and ploidy level, identifying features which are also positively related to invasiveness (Pandit *et al.*, 2014; Mounger *et al.*, 2021; Moura *et al.*, 2021). For example, *Reynoutria. japonica* var*. japonica* is exclusively octoploid (2n = 8x = 88), *Reynoutria japonica* var*. compacta* is exclusively tetraploid (2n = 4x = 44), while *Reynoutria sachalinensis* is predominantly tetraploid (2n = 4 x = 44) with occasional hexaploid and octoploid cytotypes, and *R. ×bohemica* is usually hexaploid (2n = 6x = 66) with occasional tetraploid and octoploid cytotypes (Mandák *et al.*, 2003).

1.2.3 Impacts

1.2.3.1 Impacts: Environmental

Japanese knotweed *s.l.* negatively impact ecosystem services (Fennell *et al.*, 2018), and native flora (Lavoie, 2017). Knotweeds alter soil chemistry for their own benefit, produce abundant polyphenol-rich leaf litter and possess a deep rhizome system (Tamura & Tharayil, 2014; Lavoie, 2017). The recalcitrant leaf litter produced can affect nutrient cycling and subsequent germination of other species (Tamura & Tharayil, 2014; Lavoie, 2017). Japanese knotweed's dense leaf canopy can act as umbrella, blocking light and diverting water away from native vegetation beneath. In combination with an expansive rhizome system this can have deleterious effects on water resources, with the capacity to lower the level of a small stream (Vanderklein *et al.*, 2014). The reduction in native plant species can have a knock-on effect on the abundance of native fauna, such as frogs (Lavoie, 2017). As discussed previously, Japanese knotweed is allelopathic which gives it an advantage over native flora (Murrell *et al.*, 2011; Parepa *et al.*, 2012; Parepa & Bossdorf, 2016). This has been demonstrated in previous studies, in which the growth of other plant species was altered by application of Japanese knotweed leachates (Dommanget *et al.*, 2014b) and isolated secondary metabolites resveratrol, emodin and (−)-epicatechin (Tucker Serniak, 2016).

1.2.3.2 Impacts: Socio-economic

Public opinion of Japanese knotweed has been negatively swayed by the sensationalist language used to describe its rapid growth. Click-bait headlines have drawn comparisons with the science-fiction horror creatures, 'Triffids', and have

described the species running 'rampant' and 'eating' homes, see (Brooks, 2022; Miller, 2022; Parry, 2022; Clare, 2022). This poor reputation of Japanese knotweed has acted as a self-fulfilling prophecy, imposing significant economic costs through both loss in property value and costly control measures. In 2010, the total cost to the UK economy was estimated to be £165,609,000 per annum (Williams *et al.*, 2010).

Whilst it is not illegal to have Japanese knotweed on your property, in the UK it is an offence to allow it to spread onto adjacent land, or to actively plant or otherwise cause the species to grow in the wild (United Kingdom Parliament, 1981). It is also classified as controlled waste and must be disposed of at a licenced landfill (United Kingdom Parliament, 1990). There have been two notable legal cases in recent years which mark a shifting precedent surrounding the responsibility for control of Japanese knotweed; Network Rail Infrastructure Ltd v Williams and Waistell, in which a company was sued for allowing the spread onto others property, and Ryb v Conway Chartered Surveyors in which a surveyor was sued for negligence after failure to identify knotweed during an inspection. In the Network Rail Infrastructure Ltd v Williams and Waistell case, the encroachment of Japanese knotweed was deemed to reduce the amenity value of the property by limiting the owners' enjoyment of their garden without the use of costly remediation methods, and thereby causing a nuisance (Warren, 2019).

To comply with the Consumer Protection Regulations, estate agents must also declare the presence of Japanese knotweed on a property. Historically, mortgage lenders have been following the Royal Institution of Chartered Surveyors (RICS) 'seven metre rule' which stated that if Japanese knotweed is within 7 metres of the property then it must be completely removed by a professional who can offer an insurance backed policy against its return (House of Commons Select Committee on Science and Technology, 2019). These strict rules have made it difficult and costly for homeowners to sell with some mortgage lenders refusing to lend while others require the homeowners to take on the expensive eradication costs. The knock-on social impacts of this policy on public mental health were highlighted in 2013, when the stress of Japanese knotweed invading their property was cited as the cause for

the murder of Mrs Jane McRae and the suicide of Mr Kenneth McRae (BBC News, 2015).

However, more recent studies have concluded that the 'seven-metre rule' is not a statistically robust tool for estimating likely rhizome extension because Japanese knotweed rhizome rarely extends more than four metres from above ground plants and is typically found within only 2-2.5 metres of stands (Fennell *et al.*, 2018). Additionally, this study concluded that Japanese knotweed does not cause significant damage to built-structures, and no more damage than other plant species which are not outlined in lending policies (Fennell *et al.*, 2018). These findings have been reflected in the updated guidance by the Royal Institution of Chartered Surveyors (RICS), which came into force in 2022. The 'seven-metre rule' has now been abolished in favour of a more lenient approach that allows surveyors to use their discretion when assessing the impact of knotweed infestations. This change mirrors the shifting mindset away from complete eradication to appropriate control.

1.2.4 Current Control Measures

1.2.4.1   Site Identification and Management Planning

The first step in Japanese knotweed control is the identification of inhabited sites, which can be problematic as rhizome buds may remain dormant underground (Bashtanova *et al.*, 2009a). Conservation dogs are effective at helping humans detect Japanese knotweed even below ground (Conservation Dogs Ireland, 2022). Working with detection dogs such as these could become increasingly useful following the previously described Ryb v Conway Chartered Surveyors legal case, which established a new model for penalty of surveyor negligence (Warren, 2019). Once identified, the affected land must be considered on a site-by-site basis as different situations may require different control measures. For example, if the affected site is in a conservation area near a water course, or in a public park where children play, then herbicide spraying may be inappropriate (Environment Agency, 2017). Treatment of a stand in a private garden or building site will differ from an area where the goal is to replant native vegetation (McManamen *et al.*, 2018). In some

cases safety may be a consideration, which could require the physical removal of knotweed, for example in order to maintain accessibility or visibility alongside transport infrastructure (Boyer & Brasier, 2019). To add further complexity, Japanese knotweed can hybridise and backcross with similar species, creating a complex known as Japanese knotweed *sensu lato* (in the broad sense) (Hollingsworth & Bailey, 2000b) (see Section 1.2.2). These hybrids respond differently to control measures, and some can produce viable seed (Mandák *et al.*, 2003). Therefore, it can be useful to identify the type of Japanese knotweed in an occupied site, although the leaves are morphologically very similar (Bailey *et al.*, 2009) suggesting a requirement for new methods for rapidly and reproducibly differentiating between knotweed hybrids.

1.2.4.2 Physical

Physical weed control methods such as tillage systems can be used to remove Japanese knotweed stands, examples include flex-tine harrow sweep cultivator and rotary hoe. These methods have a high energy requirement with an energy consumption of 4.2 to 5.5 MJ ha^{-1}, 13 to 14 MJ ha^{-1}, and 12 to 17 MJ ha^{-1}, respectively (Coleman *et al.*, 2019). These tilling methods may not be appropriate for conservation cropping, which aims to protect soil from erosion and degradation, therefore thermal based approaches may be considered. However, the energy requirements for thermal approaches are 100 to 1000-fold higher than tillage treatments: flaming (1,008 to 4,334 MJ ha^{-1}) and infrared (2,000 to 3,887 MJ ha^{-1}) (Bajwa *et al.*, 2015; Coleman *et al.*, 2019). Alternative technologies include electrocution and laser pyrolysis, which require 15-19, 15-249 MJ ha^{-1}, respectively (Coleman *et al.*, 2019). Other physical methods include burning/incineration, digging, root barrier methods to control spread, on-site burial/cell formation, soil and rhizome sifting, covering with geotextiles or wire meshes and rhizome crushing (Payne & Hoxley, 2012; Martin FM *et al.*, 2020a).

Contrasting opinions exist on the practice of mowing. Some authors argue strongly against mowing citing its limited efficacy, practicality, and environmental and economic sustainability (Jones D *et al.*, 2020). Others claim that the practice has some worth if applied with certain caveats, despite the risk of potentially increasing

spread in some cases; it must be complete mowing because partially mowed or cut knotweed stands are able to compensate for the loss of half of their aboveground organs, it is unlikely to lead to complete eradication because three mowing or cutting events per year was insufficient to kill young regenerating ramets; and its use is recommended alongside the restoration of native species (Martin FM *et al.*, 2020b,a).

1.2.4.3 Chemical

Several herbicides have been trialled for the treatment of Japanese knotweed, including imazapyr, glyphosate, and synthetic auxins (Bashtanova *et al.*, 2009b). Herbicides can be applied by foliar spray, stem injection, or foam application and generally require an energy consumption of 15 MJ ha^{-1} (Coleman *et al.*, 2019). Glyphosate treatment is still the most effective method, however eradication is still difficult after several years of application (Jones D *et al.*, 2018). When forming treatment plans for herbicide application it is important to take into account sink-source relationships in order to ensure sufficient glyphosate allocation to the extensive rhizome system, with late summer to early autumn being the most appropriate time for application (Bashtanova *et al.*, 2009b; Jones D *et al.*, 2018). However, the increased translocation to the rhizomes in autumn must always be balanced with the mechanism of action of glyphosate, which works by inhibiting protein synthesis and disrupting growth and should therefore theoretically have its greatest impact in summer when the plants are growing most profusely (Bashtanova *et al.*, 2009b).

Glyphosate has come into disrepute in recent years over concerns that it has been listed by the International Agency for Research on Cancer as a 'probable carcinogen' (Tarone, 2018), and due to the accumulation of both glyphosate and its degradation product, aminomethylphosphonic acid, in the environment (Van Bruggen *et al.*, 2018). There is a growing level of public discontent in both Europe and the United states, both concerning safety of glyphosate and adjuvants used and the involvement of large multinational companies (Van Straalen & Legler, 2018). Alternative herbicides which have been suggested as a replacement for glyphosate may be inappropriate for the treatment of recalcitrant species such as Japanese

knotweed due to its extensive rhizome system. For example, the manager of Italy's Ministry of Agricultural, Food and Forestry Policies, suggested pelargonic acid (also known as nonanoic acid) as an effective substitute for glyphosate (Marks & Paravicini, 2017). However, this bioherbicide causes plant death by desiccation through penetration of the waxy cuticle of the leaves and destroying the epidermal cell membranes, destroying foliage (Ciriminna *et al.*, 2019). Based on this mechanism of action, which primarily targets top growth, is unlikely that this herbicide would be suitable for the control of rhizomatous plants.

The allelopathic potential of the target plant could be an important consideration when looking for possible new herbicides. This would require further investigation of the impact of the intrinsic allelopathic capability of target species on their herbicide resistance, as previous research has predominantly focused on the use of allelopathic chemicals as herbicides for use on other species (Bertholdsson, 2012; Bajwa *et al.*, 2015). Plants which themselves produce phytotoxic compounds may be more resistant to allelochemicals as they possess detoxification mechanisms to prevent autotoxicity, or alternatively may have an increased responsiveness to allelopathic biochemicals due to auto-allelopathy (Cheng & Cheng, 2015; Thiébaut *et al.*, 2019). Another concern is insufficient herbicide efficacy because sub-lethal herbicide application can trigger plant's stress defences, arming them against other abiotic and biotic stresses and rendering them harder to eradicate, a process known as 'hormesis' (Belz & Duke, 2014).

1.2.4.4 Biocontrol

Biological control can be used as a method to mitigate the impact of a pestilential organism through a reduction in its prevalence (but not eradication) using a living agent to the benefit of a human stakeholder, for example the control of an IAS using natural enemies from a its original native environment (Stenberg *et al.*, 2021). These natural enemies put pressure on the limited resources of the target species, preventing it from competing as effectively with surrounding species. Although biological control agents take time to increase in number, and 'catch up' with their hosts, the eventual cost: benefit ratio can be very favourable, from both insect (McFadyen, 1998) and fungal (Evans, 2000) agents, particularly rust fungi (Tomley

& Evans, 2004). *Reynoutria japonica* was suggested as a highly appropriate target for classical biological control (Holden *et al.*, 1992), because it has been imported into its invasive range without any of its specialised natural enemies which are also native to Japan. Another trait of invasive *Fallopia japonica* which lends itself to biological control is its clonal growth (Hollingsworth & Bailey, 2000). A lack of sexual reproduction should mean that the plant is unable to show variation in resistance or tolerance to arthropod or microbial infection. Provided suitable agents can be found, plants which reproduce asexually are anticipated to be more susceptible to biological control (Burdon & Marshall, 1981; Crawley, 1990).

The development of biological control agents targeting Japanese knotweed has been led largely by the Centre for Agriculture and Biosciences International (CABI) in Wales, United Kingdom, and Agriculture and Agri-Food Canada (AAFC) at the Lethbridge Research and Development Centre in Canada (Shaw *et al.*, 2009; Grevstad *et al.*, 2018). To source the biocontrol agents, the closest genetic match to Japanese knotweed in Western Europe was sourced from current native populations in Japan (Pashley, 2003) and used to screen potential pests and pathogens which target Japanese knotweed in its native range as candidate biocontrol agents (Shaw *et al.*, 2009). Although several were ruled out due to problems overwintering in the UK and unspecific host plants, two potential control agents were selected: a psyllid, *Aphalara itadori*, and a rust fungus, *Mycosphaerella polygoni-cuspidati.* However, although promising, different species and hybrids within the Japanese knotweed *sensu lato* complex are known to have differing interactions with biocontrol agents (Mullerscharer *et al.*, 2004; Grevstad *et al.*, 2013; Gaskin *et al.*, 2014). For example, the northern biotype of *Aphalara itadori* favours Giant knotweed, whilst the southern biotype prefers Japanese knotweed and Bohemian knotweed (Grevstad *et al.*, 2013). This raises questions about the widespread use of biocontrol for Japanese knotweed.

1.2.4.5   Restoration

Due to its allelopathic nature, the soil environment following a knotweed invasion is not always hospitable to the growth of other plants (Dommanget *et al.*, 2014a). Allelopathic chemicals present in Japanese knotweed leachates can have a

phytotoxic effect on the growth of cuttings of native tree species and cause a reduction in final biomass compared to plants grown in bare soil (Dommanget *et al.*, 2014a). However, the impacts of leachates are complex and may not necessarily be adverse. In addition to the release of allelopathic chemicals, Japanese knotweed can also alter the microbiome beneath the stands. Japanese knotweed has been shown to stimulate the soil microbial community and may aid germination of native plant seedlings (Parepa *et al.*, 2012). The species selected for restoration of the area therefore needs to be carefully selected and resistant to the allelopathic chemicals produced by Japanese knotweed *s.l.,* with suggested candidates including *Salxi atrocinerea* (Dommanget *et al.*, 2014a). Planting native vegetation reduces the vigour and lateral expansion rate of knotweeds, increases the biotic resistance against re-invasion from knotweeds and other invasives by maintaining a dense vegetation cover, and stabilises riparian corridors thereby reducing downstream dispersal rate (Skinner *et al.*, 2012; Dommanget *et al.*, 2014b; Martin FM *et al.*, 2020a).

1.2.4.6   Future trends in management strategies

Misidentification within the Japanese Knotweed *s.l.* complex, has led to an underestimation of the prevalence of hybridisation (Zika & Jacobson, 2003; Gillies *et al.*, 2016), whilst similar morphology has complicated management strategies (Moody & Les, 2007). Despite its increased invasiveness, Bohemian Knotweed has not been recognised on the United States Department of Agriculture (USDA), Natural Resources Conservation Service (NRCS) Plants Database where it is still listed as "Absent/Unreported" in the United States of America (USDA Plants Database). Similarly overlooked due to morphological variation (Mandák *et al.*, 2003) is the dwarf variant *Reynoutria japonica* var. *compacta*. Misidentification is a particular concern in hybrids where viable seeds are produced because glyphosate, the main herbicide used to treat Japanese Knotweed, is applied post-flowering to increase herbicide allocation to rhizomes (Jones D *et al.*, 2018). However, an increasing prevalence of Bohemian Knotweed (Gillies *et al.*, 2016), and the occurrence of stands (clumps) of seeding Japanese Knotweed *s.l.* (Grimsby *et al.*, 2007), means late-season herbicide application may not be an appropriate 'cure-all' treatment program.

The discovery of seeding stands in North America (Bram *et al.*; Forman *et al.*, 2003; Grimsby *et al.*, 2007), putative introduction of a male plant in the United States (Del Tredici, 2017a), and viable seeds produced during hybridisation through genetic introgression (Hollingsworth *et al.*, 1998) has led to a shifting rhetoric in knotweed management towards addressing the issue of seed production (Soll, 2004; Grimsby *et al.*, 2007; Bashtanova *et al.*, 2009a). This issue has previously been neglected because rhizomes are considered the primary means of expansion and reproduction in the introduced ranges (Engler *et al.*, 2011). With climate change potentially improving conditions for seed germination and survival in its introduced ranges (Walther *et al.*, 2009; Diez *et al.*, 2012; Bellard *et al.*, 2013), it is a concern that any resultant increased propensity for Japanese knotweed *s.l.* to reproduce sexually could significantly worsen its impact (Groeneveld *et al.*, 2014).

The ultimate development goal for the control industry is a new, safe, effective and low-carbon treatment method to replace the current available treatment methods which are energy intensive (Coleman *et al.*, 2019) and potentially hazardous (Tarone, 2018). An improved ability to identify knotweed and its hybrids correctly is therefore essential for effective control and the future design of stand-specific treatment programs tailored to the plants present and the environmental conditions under which they are growing.

1.3 ATR-FTIR spectroscopy

The developing area of mid infrared vibrational spectroscopy (MIR; 4000–400 cm^{-1}; 2500–25,000 nm) allows the rapid, marker-free, non-destructive analysis of biological samples (Morais *et al.*, 2020). When infrared light contacts infrared-active bonds within a sample, light-matter interactions induce vibrations within the sample by absorption of light at specific defined wavelengths. This technique quantitatively measures the absorption of infrared light by a sample across various wavelengths to create a multivariate dataset. Although the spectroscopic approach originated in analytical chemistry, where it was typically used to identify functional groups in a relatively pure sample, advances in computational analysis have allowed the study of more complex molecular mixtures and expansion into biological science

applications (Chan & Kazarian, 2016). A branch of this spectroscopic method, attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy, followed by chemometric analysis has been successfully applied to investigate plant response to abiotic (McNear *et al.*, 2010; Euring *et al.*, 2012; Buitrago *et al.*, 2016; Butler *et al.*, 2017, 2020; Usman *et al.*, 2019) and biotic factors (Skolik *et al.*, 2019a), as well as for monitoring plant health and development (Butler *et al.*, 2015; Skolik *et al.*, 2019b). These studies focus on the wavelengths preferentially absorbed by biological materials, such as the fingerprint region (1800-900 cm^{-1}), amide I and II (1700–1500 cm^{-1}), stretching vibrations such as S-H, C-H, N-H and O-H (3500-2550 cm^{-1}), and the lower wavenumber regions which are typically due to bending and carbon skeleton fingerprint vibrations (Morais *et al.*, 2020). The fingerprint region includes important biological absorptions due to lipids, proteins, carbohydrates, nucleic acids and protein phosphorylation (Morais *et al.*, 2020).

Databases are available with catalogued definitions for characteristic peak frequencies (Movasaghi *et al.*, 2008; Talari *et al.*, 2017). For example, absorptions have been linked to biologically significant compounds including glucomannan (Kanter *et al.*, 2013), xyloglucan (Kacuráková *et al.*, 2000), succinate (Kang *et al.*, 2008), and pectin (Sharma & Uttam, 2018). However, the process from chemometric biomarker identification to physical biomolecular extraction is the subject of ongoing research focused on calibrating concentrations derived from biological spectra (Spalding *et al.*, 2018), consolidating the expanding database of key wavenumber changes and associated molecular definitions (Talari *et al.*, 2017), and trialling new biological applications (Skolik *et al.*, 2019a,b; Butler *et al.*, 2020; Holden *et al.*, 2021, 2022; Emsley *et al.*, 2021). Sample preparation, such as freeze-drying or grinding, may influence the spectra acquired from vegetative plant material and the resultant classification success of subsequent chemometric analysis. To ensure optimum spectral quality and molecular sensitivity, instrumental settings and sample preparation must be adjusted prior to spectral acquisition (Baker *et al.*, 2014; Chan & Kazarian, 2016; Pilling & Gardner, 2016; Morais *et al.*, 2019).

Other similar techniques include near infrared spectroscopy (NIR; 1000–2500 nm; 10,000–4000 cm^{-1} (Beć *et al.*, 2020), and Raman spectroscopy. Similar to MIR spectroscopy, Raman spectroscopy, has also been used for identification of abiotic and biotic stress markers through measurement of metabolites such as carotenoids and anthocyanins (Lew *et al.*, 2020). However, research is still ongoing to overcome the limitations caused by the interference between the strong fluorescent background of chlorophyll, carotene, lutein, and pathogens of dried plant samples which can obscure the Raman signal (Zhang *et al.*, 2010; Farber *et al.*, 2019).

1.4 Chemometrics

Although visual differences can be seen between infrared spectral profiles, mathematical analysis is beneficial to glean useful interpretations. Chemometrics involves the use of statistics to perform signal correction, pattern recognition and classification of measured data to gain chemically relevant information (Brereton, 2007). Machine learning algorithms can be used within programming platforms to analyse the multivariate datasets gained from the spectral absorbances (Houhou & Bocklitz, 2021). Specialised packages exist within programming platforms for analysis of spectral data (Trevisan *et al.*, 2013), and for performing these chemometric algorithms (Ballabio & Consonni, 2013; Ballabio, 2015; PLS_Toolbox with MIA_Toolbox 8.2.1, 2016; Valsecchi *et al.*, 2020). Figure 1.1 shows the chemometrics pipeline from data acquisition to data validation.
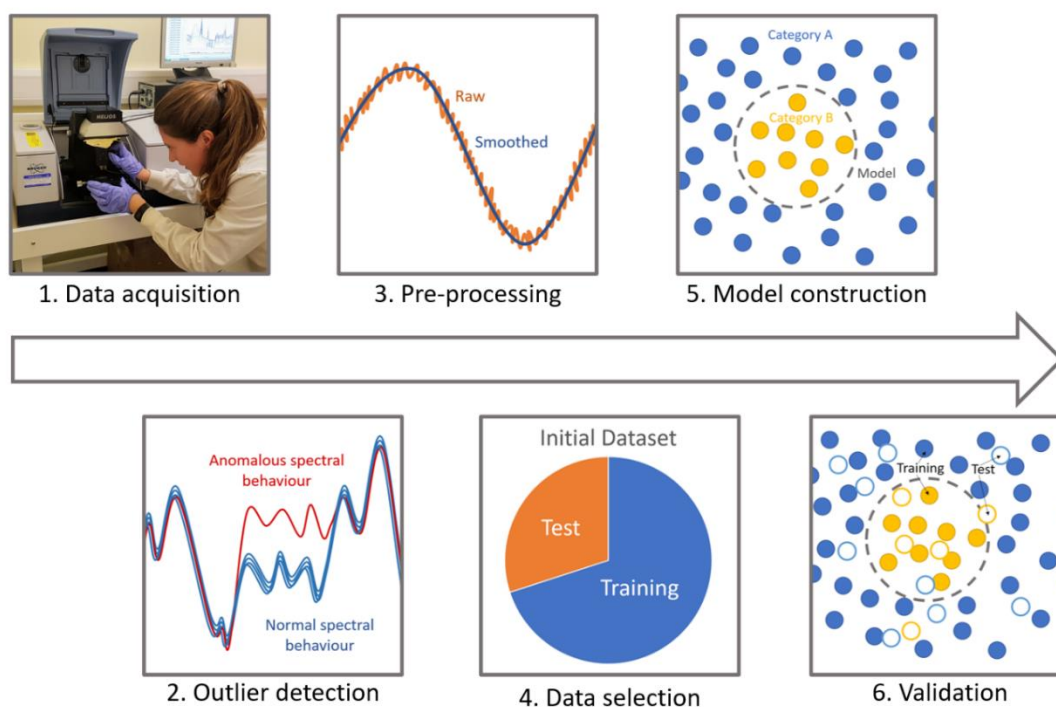
Figure 1.1: Chemometric data analysis process.

1.4.1.1 Pre-processing: Savitzky-Golay Second Differentiation and Vector Normalisation

Before any machine learning algorithms can be applied the acquired spectral data must be pre-processed. This is an essential step of all spectroscopic experiments and is used to improve the signal-to-noise ratio by correcting problems associated with random or systematic artefacts during spectral acquisition including different sample thickness (Butler *et al.*, 2018). There are numerous options for pre-processing, the choice of which is dependent on factors such as noise, light scattering, and baseline distortion, see decision tree in (Morais *et al.*, 2020). The experiments throughout this thesis have primarily used the techniques of Savitzky-Golay second differentiation and vector normalisation, and in some cases baseline correction.

1.4.2 Machine Learning

The large datasets generated by modern sensors has led to application of machine learning algorithms within many scientific disciplines, such as wildlife conservation (Tuia *et al.*, 2022), plant biology (Ma *et al.*, 2014), and soil science (Wadoux *et al.*, 2020). Machine learning computer algorithms can improve automatically, using historical data as an input to 'teach' the program to create a more accurate output (Jordan & Mitchell, 2015). There are many different available algorithms within machine learning (Figure 1.2). These can be divided into supervised and unsupervised learning (Ayodele, 2010), as well as more sophisticated techniques such as reinforcement learning (Mnih *et al.*, 2015). Which algorithm to choose can depend on the size of the data set, computing power, and the information desired from the dataset. Multivariate data such as spectral absorbance profiles for multiple samples can be simplified with dimensionality reducing algorithms such as principal component analysis (PCA; Oskolkov, 2022). These allow spectral absorbance values measured at multiple wavelengths to be plotted as a single point in a two-dimensional space and hence multi-dimensional patterns can be easily visualised.

Machine learning techniques known as 'classifiers' are supervised techniques which attempt to divide the input data into the correct classes (categories), using training data to recognise how the given input variables relate to each class (Mahesh, 2018). There are many models capable of partitioning and classifying data, such as those listed here in approximate order of increasing model complexity: Linear Discriminant Analysis (LDA) > Partial Least Squares-Discriminant Analysis (PLS-DA)> Quadratic Discriminant Analysis (QDA) > K-Nearest Neighbours (k-NN) > Support Vector Machines (SVM) > Artificial Neural Networks (ANN) > Random forests > Deep-Learning approaches (Brereton, 2015; Morais, 2020). The analysis reported in this thesis have used two unsupervised learning techniques, PCA and hierarchical cluster analysis (HCA), and three supervised techniques, principal component analysis-linear discriminant analysis (PCA-LDA), support vector machines, and partial least squares regression (PLSR).
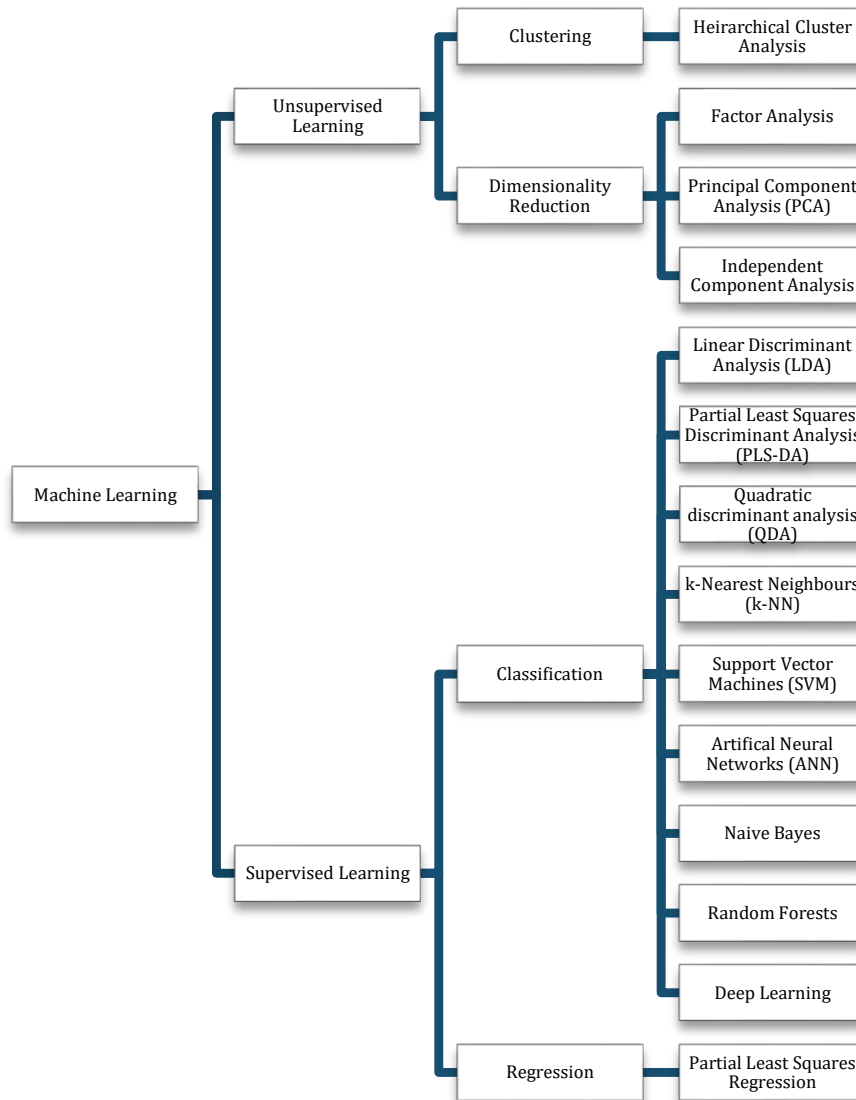
Figure 1.2: Branches of machine learning and some commonly used algorithms.

1.4.2.1   Unsupervised Learning

1.4.2.1.1   Dimensionality reduction: Principal Component Analysis

The traditional display of data on one variable per axis (x, y, z) becomes difficult if more than three variables are measured. PCA, an unsupervised learning technique, is a fast and reliable method which reduces the dimensionality of complex multivariate datasets, (Oskolkov, 2022). This simplicity allows spectral differences to be presented in an intuitive manner. PCA reduces the original data into a lower dimensional space based on a new set of axes called principal components (PCs), this decomposition is depicted in Figure 1.3 (see also Appendix A). These variables

are orthogonal to each other and account for most of the explained variance from the original data set. They are composed of scores and loadings that are used to identify similarities/dissimilarities among the samples and the weight that each variable contributes for the PCA model, respectively (Cirino de Carvalho *et al.*, 2016). For the example in Figure 1.3, hormone two is more important for explaining the variance in the data, has a higher loading score, and influences the placement of axis PC1 more than hormone 1. When PCA is used in biospectroscopy, for a given discriminant PC direction, the spectral bands with the highest loading coefficients, both positive and negative, can be selected as spectral biomarkers. Because PCA loadings represent the variance in the wavenumber direction, these selected loadings represent the key wavenumbers that distinguish between samples from different biological classes (Morais *et al.*, 2019). Advantages of PCA include prevention of collinearity problems and faster computational analysis. However, it is important to compare the effect of different number of PCs on the model's performance because as the data is projected into a lower dimensional space some high-dimensional subtleties are lost.
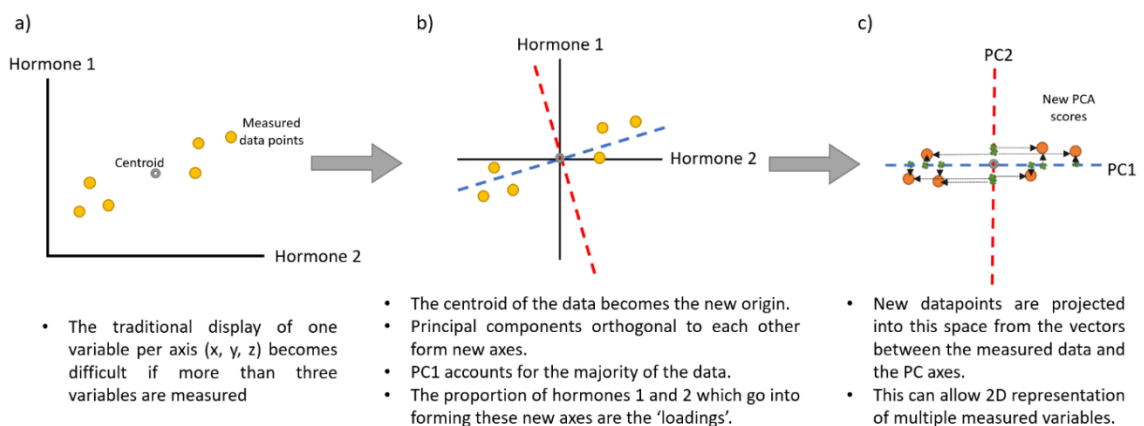
Figure 1.3: The process of PCA decomposition. Adapted from (Starmer, 2018).

As an unsupervised technique, the category variables are not used for this reduction. To perform a supervised classification model, the PCA scores can be employed as input variables for discriminant algorithms, for example in principal component analysis – linear discriminant analysis (PCA-LDA), see Section 1.4.2.2.1.

1.4.2.1.2 *Clustering: Hierarchical Cluster Analysis*

Clustering is the division of data points into groups using an algorithm, the choice of which is important because the way each data point is compared can alter the outcome of the group to which it is assigned. Hierarchical Cluster Analysis (HCA) is a clustering method which explores the organisation of samples within groups and can create nested categories, the hierarchy of which are depicted in a dendrogram; in HCA, objects are more related to nearby objects than to those farther away (Köhn & Hubert, 2015; Bridges, 2016). HCA determines, pairwise, which two points are most similar (Köhn & Hubert, 2015; Bridges, 2016). The samples distance is measured using the metric Euclidean distance. Ward's Method is a commonly used linkage criterion employed in this thesis. This is an agglomerative or 'bottom-up' approach, in which each individual item starts in its own cluster, and these are iteratively merged until all items are in one cluster. To find the merging combination with the smallest aggregate deviation, the centre points of each individual cluster are compared with the centroids of the combined clusters (Lavrenko, 2015; Figure 1.4). This is achieved by calculating the sum of the square deviations of the points from the new centroid (see Appendix B).
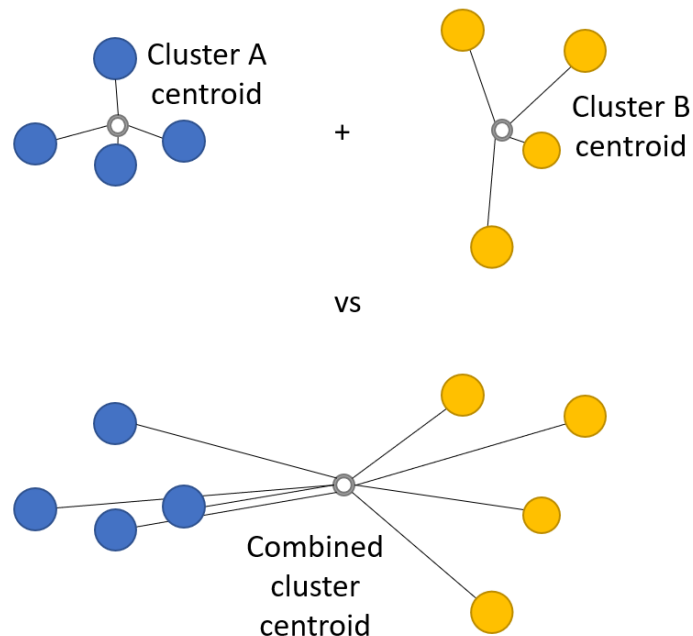
Figure 1.4: Ward's clustering method compares the centroids of individual clusters with their combined centre. Recreated from (Lavrenko, 2015).

1.4.2.2   Supervised Learning

1.4.2.2.1   *Classification: Principal Component Analysis – Linear Discriminant Analysis*

Linear discriminant analysis (LDA) is a discriminant algorithm that creates a linear classification rule between the classes based on a Mahalanobis distance. In PCA-LDA, the PCA scores are used as input variables for the LDA to generate the classification scores (see Appendix C).

1.4.2.2.2   *Classification: Support Vector Machines and kernel functions*

Real-world data is often hard to separate linearly, requiring more complex techniques for classification such as SVM (Boser *et al.*, 1992). For this, SVM classifiers work by finding a classification hyperplane that separates the data clusters providing the largest margin of separation. During model construction, the data are transformed into a different feature space by means of a kernel function that is responsible for the SVM classification ability (see Figure 1.5 and Appendix D). The most common kernel function is the radial basis function (RBF).
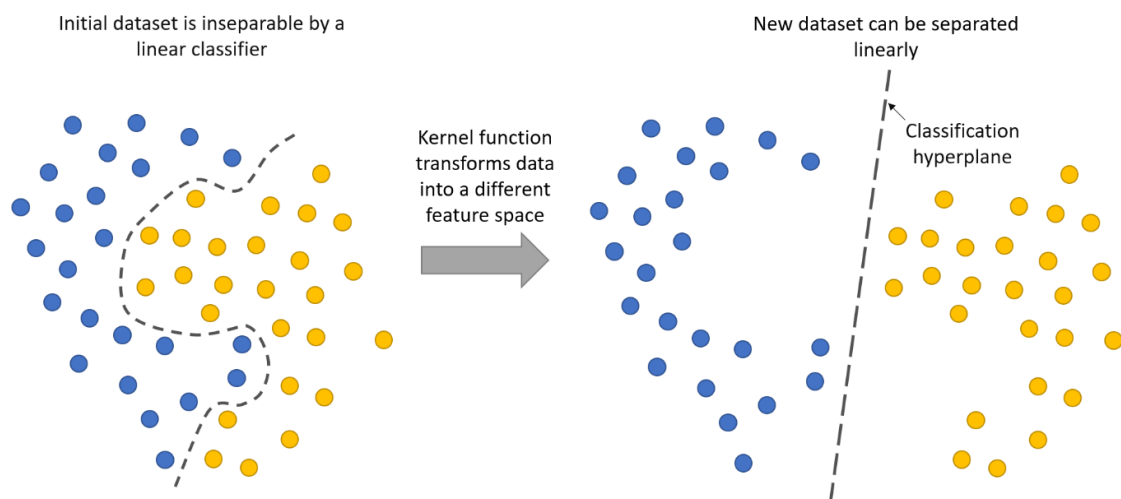
Figure 1.5: Transformation of data into a different feature space using a kernel function allows linear separation of data by a classifier such as SVM. Adapted from (Wang C *et al.*, 2007).

The number of components of PCA-LDA and all SVM parameters were optimized by venetian blinds (10 data splits) cross-validation. In the analyses reported in this

thesis, model validation was performed with 30% of the samples' spectra randomly selected as an external test set. Accuracy, sensitivity, and specificity are important factors in the evaluation of model success. The accuracy represents the total number of samples correctly classified considering true and false negatives; the sensitivity represents the proportion of positives that are correctly classified; and the specificity represents the proportion of negatives that are correctly identified (see Appendix E; Morais & Lima, 2018).

1.4.2.2.3 Regression: Partial least squares regression

Whilst classification is a useful tool to answer binary questions, such as the presence or absence of disease (Skolik *et al.*, 2019a), often further detail is desired. Partial least squares regression (PLSR), also known as 'projection on latent structures', is a supervised technique used to predict a set of dependent variables from a set of independent variables or predictors (Geladi & Kowalski, 1986). Its versatility lends it to a range of applications across social sciences, bioinformatics, machine learning and chemometrics (Mehmood *et al.*, 2012). To create the predicative model, PLSR extracts a set of orthogonal factors called latent variables, those which are inferred rather than directly observed. This method is often compared with the technique of PCA as described above, however rather than creating a separation hyperplane with maximum variance between independent and dependent variables, PLSR projects measured and predicted variables into a new space to create a linear regression model (Abdi, 2010). The quality of the prediction obtained from a PLS regression model is evaluated with cross-validation techniques such as the 'bootstrap' and 'jack-knife' (Abdi, 2010).

1.5 Aims and Objectives

The plasticity of Japanese knotweed (see Richards *et al.*, 2008; VanWallendael *et al.*, 2018) enables this species to survive in extreme abiotic conditions (Rouifed *et al.*, 2012; Richards *et al.*, 2012a; Michalet *et al.*, 2017) and occupy a broad geographic range (Bailey & Conolly, 2000; Barney *et al.*, 2006; Bzdega *et al.*, 2012; Clements *et al.*, 2016; Holm *et al.*, 2018). It is notoriously difficult to control (Jones D *et al.*, 2018), a matter complicated by the numerous morphologically similar hybrids (Bailey *et*

al., 2009). Available treatment methods are energy intensive (Coleman *et al.*, 2019) and potentially hazardous (Tarone, 2018). In addition, treatment is complicated difficulties with distinguish between morphologically similar species resulting in wasted efforts to control seeding hybrids (such as Bohemian knotweed) post-dispersal. Therefore, new approaches for the rapid and robust identification of knotweed and its hybrids and related species across the range of environmental conditions in which they are found are essential for the development of efficient control mechanism which have a reduced carbon footprint. The aim of the research reported in this thesis is to explore the opportunities offered by advances in the application of ATR-FTIR spectroscopy within plant sciences, for the identification and control of knotweed based on this technique and, in doing so, add to our understanding of knotweed biology and the potential of this technique.

1.5.1 Objectives

- Chapter 2: Sample Preparation of Leaf Material for Attenuated Total Reflection Fourier-Transform Infrared Spectroscopy

    o Determine the effect of sample preparation on the ATR-FTIR spectral signature.

- Chapter 3: Regional differences in clonal Japanese knotweed revealed by chemometrics-linked attenuated total reflection Fourier-transform infrared spectroscopy

    o Determine the effect of regional environmental conditions on the phenotypic plasticity of Japanese knotweed using ATR-FTIR spectroscopy.

- Chapter 4: Know your enemy: Application of ATR-FTIR spectroscopy to invasive species control

    o Develop a rapid screening method for hybrid identification in the Japanese knotweed complex

- Chapter 5: Environmental metabolomics using attenuated total reflection Fourier-transform infrared spectroscopy for the prediction of hormone concentrations in plants

    o Develop predictive models for plant hormone levels

    o Determine spectral biomarkers for plant hormones within the mid-infrared fingerprint region

- Chapter 6: Attenuated total reflection Fourier-transformed infrared spectroscopy reveals environment specific phenotypes in clonal Japanese knotweed

    o Determine which aspect of environmental conditions has the most pronounced effect on the spectral signature

    o Develop predictive models for physiological responses to environmental conditions such as water pressure

    o Determine biomarkers for abiotic stress in Japanese knotweed

2 Sample Preparation of Leaf Material for Attenuated Total Reflection Fourier-Transform Infrared Spectroscopy

Claire A Holden[1*], Jane E Taylor[1], Francis L Martin[2], Paul Beckett[4], Martin McAinsh[1]

1. Lancaster Environment Centre, Lancaster University, LA1 4YQ, UK

2. Biocel Ltd, Hull, HU10 7TS, UK

3. Phlorum Ltd, Brighton, BN2 6AH, UK

*Corresponding author: Claire Holden, c.holden6@lancaster.ac.uk

2.1 Abstract

The application of infrared spectroscopy to plant science is becoming increasingly widespread. Attenuated Total Reflection Fourier-Transform Infrared (ATR-FTIR) spectroscopy, in combination with chemometrics, has proven to be a rapid, low cost, and non-destructive method for analysing plant tissues, which has the potential for utilisation as a sensor to addresses emerging global environmental challenges. However, there are marked variations between studies in the sample preparation used and the treatment of plant vegetative material prior to spectral acquisition with ATR-FTIR spectroscopy. Five preparation methods of plant vegetative tissue prior to spectral acquisition were compared in order to assess how this affects the resultant spectra. Results show that sample preparation has a marked impact on the plant-derived spectra and that grinding of samples can allow spectral acquisition from different compounds of interest. Additionally, results provided evidence that historic herbarium specimen can prove a viable sample for ATR-FTIR analysis.

2.2 Introduction

In the age of the Anthropocene, there is a demand for new technology to addresses emerging global environmental challenges (Springer & Duchin, 2014; Fraser, 2020). Human interference is creating pressures on global food security in the face of climate change (Delabre *et al.*, 2021), increasing the spread of invasive plant species (Pyšek *et al.*, 2020), threatening biodiversity (Eriksson & Hillebrand, 2019; Kelly *et al.*, 2020), and producing pollution (MacLeod *et al.*, 2021; Vohra *et al.*, 2022). To feed an increasing population, expected to reach 9.3 billion by 2050 (Lee, 2011), efficient crop production is vital (Godfray *et al.*, 2010). An estimated 70% reduction in yield gaps is required in order to meet global food demands (Willett *et al.*, 2019). Increased agricultural productivity would allow reduction of the cropland area required, avoiding pressure on natural ecosystems. In order to meet these essential targets, food losses and waste throughout the supply chain must be at least halved (Willett *et al.*, 2019). Sensors which can rapidly detect early signs of crop pathogens

and plant stress are key goals (Jeger *et al.*, 2021). These would require high levels of sensitivity and pathogen specificity, be flexible and non-destructive enough to use in the field whilst retaining maximal yield, and ideally quantify the severity of the stress and differentiate between biotic and abiotic stresses (Mahlein, 2016; Skolik *et al.*, 2018, 2019a). One major cause of crop loss are invasive species, which threaten ecosystem stability and cause economic pressures (Bradshaw *et al.*, 2016; Shackleton *et al.*, 2019). Their entrance into countries in which they were previously absent is only set to be more prevalent due to increased globalisation and interconnection via world trade (Hulme, 2009a; Amano *et al.*, 2016). Movement of exotic species is influenced by changes in climate and land use (Paini *et al.*, 2016). Many countries, such as New Zealand, have strict border controls to monitor and predict transfer of species (Hulme, 2020). A key aspect of this strategy is the ability to accurately identify any intercepted specimen to the species-level despite the complications of morphologically indistinct immature life stages, cryptic species, and damaged specimens. However, there is disparity between the availability of trained taxonomic experts and the diagnostic need for a comprehensive worldwide identification strategy (Armstrong & Ball, 2005). Research now focuses on the development of portable technologies, such as deoxyribonucleic acid (DNA) barcoding, to aid point-of-entry detection of non-native species (Madden *et al.*, 2019; Trujillo-González *et al.*, 2022). However, DNA barcoding still struggles to identify specimens requiring urgent identification, where morphological assignment is the current best option (Madden *et al.*, 2019).

A potential alternative technology is attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy. This is an efficient method that has proven to be rapid, low cost, non-destructive (Morais *et al.*, 2020). In infrared (IR) spectroscopy, IR light passes through a sample causing bonds in the sample to absorb energy at specific wavelengths, changing their molecular vibrations to a different state (Smith B, 2018). These missing wavelengths can be quantitatively measured from the absorbance spectrum and are unique to the bond in which the vibration was triggered (Smith B, 2018). The fingerprint region of the spectrum 1800-900 cm^{-1} is particularly useful in biological applications of this technique, owing to its unique

nature for each sample (Morais *et al.*, 2019a). IR spectroscopy is increasingly used in plant sciences across a range of different applications, see Table 2. 1.

Table 2.1: Examples of current applications of infrared spectroscopy in plant science.

Application	Example References
Plant identification and classification	Ribeiro da Luz, 2006; Holden *et al.*, 2022
Phylogeny	Kim *et al.*, 2004; Demir *et al.*, 2015
Identification of plants from different growing regions	Rana *et al.*, 2008; Cirino de Carvalho *et al.*, 2016a; Traoré *et al.*, 2018; Holden *et al.*, 2021
Plant pharmaceutical properties	Gendrin *et al.*, 2008; Agatonovic-Kustrin *et al.*, 2020
Quality control: cereal grains	Williams *et al.*, 2009
Quality control: fruits	ElMasry *et al.*, 2007; Bureau *et al.*, 2019; Skolik *et al.*, 2019b; Emsley *et al.*, 2021
Monitoring plant health and development	Butler *et al.*, 2015; Skolik *et al.*, 2019b
Plant response to biotic factors: Detection of crop diseases and pathogens	Kos *et al.*, 2003; Berardo *et al.*, 2005; Sankaran *et al.*, 2013; Morais *et al.*, 2017; Skolik *et al.*, 2019a
Plant response to abiotic factors: soil fertility	Euring *et al.*, 2012
Plant response to abiotic factors: heavy metals	McNear *et al.*, 2010; Usman *et al.*, 2019
Plant response to abiotic factors: water and temperature stress	Buitrago *et al.*, 2016
Plant response to abiotic factors: nutrient deficiency and uptake	Butler *et al.*, 2017, 2020

Studies apply a range of different preparative methods to plant vegetative material prior to spectral acquisition with ATR-FTIR spectroscopy, these include; intact fresh leaf cuttings (Ribeiro da Luz, 2006; Ord *et al.*, 2016), *in planta* measurement of living leaves on intact plants (Butler *et al.*, 2015a), heat dried powdered leaves (Aouidi *et al.*, 2012), heat dried intact leaves (Gudi *et al.*, 2015; Holden *et al.*, 2021), freeze-dried powdered leaves (Khairudin *et al.*, 2014) and even herbarium intact leaf samples (Strgulc Krajšek *et al.*, 2008; Holden *et al.*, 2022). However, consistency of sample preparation between different research groups is essential to obtaining reliable analytical results and enables conclusions to be compared across studies.

Here, samples from the invasive plant species, Japanese knotweed, were used to investigate the effect of pre-spectral acquisition sample preparation on the resultant IR absorbance. Five preparation methods of plant vegetative tissue prior to spectral acquisition were compared: *in planta*, dried intact leaves, intact herbarium leaves, freeze-dried-ground leaves, heat-dried-ground leaves. The results clearly demonstrate that sample preparation affects resultant spectra. Additionally, grinding has been shown to allow spectral acquisition from different compounds of interest, and that historic herbarium specimens can prove a viable sample for analysis by ATR-FTIR spectroscopy despite their age.

2.3 Materials and Methods

2.3.1 Sample Preparation

2.3.1.1 Heat-dried intact

Leaf samples were collected from both wild grown plants from various locations in the United Kingdom as well as plants grown in controlled environments. Field specimens were collected late in the growing season from seven sites (Holden *et al.*, 2021). For controlled environment-grown plants, rhizomes were planted in fertilized organic loam (John Innes No. 1, J. Arthur Bowers, UK) and grown in a climate-controlled cabinet (Microclima 1750, Snijders Scientific BV, Netherlands). In both cases, leaves were dried at 37°C for 7 days and stored in a dry airtight container at room temperature before analysis using ATR-FTIR spectroscopy.

2.3.1.2 Heat-dried ground

Leaf samples were collected from glasshouse-grown plants at Lancaster University. Rhizomes were extracted from Scorton Picnic Site, GPS: SD 5044 5040, and planted into pots filled with a fertilized organic loam (John Innes No. 2, J. Arthur Bowers, UK). Plants were grown in a glasshouse under photosynthesis optimised light bulbs, Osram Plantastar 600W. Leaves were dried at 37°C for 7 days, ground in a ball mill, and stored in a dry airtight container at room temperature before analysis using ATR-FTIR spectroscopy.

2.3.1.3 Freeze-dried ground

Leaf samples were collected from plants grown in controlled environments at Lancaster University, grown as described above for heat-dried intact plants. The youngest leaf from the top of the plant was placed immediately into liquid nitrogen, freeze-dried, and finely ground in a ball mill for analysis by ATR-FTIR spectroscopy.

2.3.1.4 Herbarium

All herbarium samples were a kind gift from the University of Leicester herbarium (LTR) (Holden *et al.*, 2022). These samples range in collection date from 1935 to the 1990s. For large leaves a 2 cm² square was cut out of each leaf between the second and third veins from the bottom left corner of each sample, to preserve the herbarium samples for future users. In the case of small leaves where this was not possible, the whole leaf was taken.

2.3.1.5 In planta

Rhizomes were extracted from Scorton Picnic Site, GPS: SD 5044 5040, and planted in a fertilised organic loam (John Innes No. 2, J. Arthur Bowers, UK). Plants were glasshouse-grown at Lancaster University with the aid of photosynthesis optimised light bulbs, Osram Plantastar 600W.

2.3.2 ATR–FTIR spectroscopy

For *in planta* samples, ATR-FTIR measurements were taken between 10 am and 2 pm to control for differences induced by the plants' circadian rhythm, and all preserved samples were also harvested between 10 am and 2 pm. All sample types were analysed using a Tensor 27 FTIR spectrometer with a Helios ATR attachment (Bruker Optics Ltd, Coventry, UK). The sampling area, defined by the Internal Reflection Element (IRE), a diamond crystal, was 250 μm x 250 μm. Spectral resolution was 8 cm^{-1} with 2 times zero-filling, giving a data-spacing of 4 cm^{-1} over the range 4000 to 400 cm^{-1}; 32 co-additions and a mirror velocity of 2.2 kHz were used for optimum signal-to-noise ratio. To minimise bias, an even spread spectra were taken for each sample across the leaf surface. For herbarium samples and intact dried leaves, ten spectra were taken from each surface of the leaf, resulting in a total of twenty spectra per leaf. Each dried leaf was placed on a slide with the side to be analysed facing upwards, the slide is placed on a moving platform and moved upwards to ensure a good and consistent contact with the diamond crystal. For *in planta* leaves ten spectra taken from the upper leaf surface only as the plant was positioned carefully around the spectrometer.

2.3.3 Data Analysis

Pre-processing and computational analysis of the data were performed using a combination of in-house developed Matlab (Mathworks, Natick, USA) toolbox called IRootLab (Martin FL *et al.*, 2010; Trevisan *et al.*, 2013) and the PLS Toolbox version 7.9.3 (Eigenvector Research, Inc., Manson, USA). Firstly, all spectral information was converted from OPUS format to suitable files (.txt) via the 'mergetool' function of IRootLab, followed by pre-processing of acquired spectra. Pre-processing, an essential step of all spectroscopic experiments, improves the signal-to-noise ratio, corrects problems associated with random or systematic artefacts during spectral acquisition (Morais *et al.*, 2020). The pre-processing steps applied to all spectra were primarily the selection of wavelengths in the spectral biochemical fingerprint region (1800-900 cm^{-1}), followed by Savitzky–Golay (SG) second differentiation and vector normalisation. Differentiation was performed before vector normalisation to

prevent the possible transfer of additive effects into multiplicative effects (Stordrange *et al.*, 2002). All data were mean-centred before multivariate analysis. Multivariate analysis looks at multiple dependant variables simultaneously to determine a pattern.

The natural variation between samples was explored using the unsupervised technique, principal component analysis (PCA). PCA reduces the original data into a few sets of orthogonal variables called principal components (PCs) which account for much of the explained variance in the original data set. These PCs are composed of scores and loadings which can identify similarities and differences between samples and determine the weight contributed by each variable to the PCA model (Cirino de Carvalho *et al.*, 2016b). Two different chemometric techniques were used for the classification of groups; principal component analysis followed by linear discriminant analysis (PCA-LDA), in which the PCA scores are used as input variables for the LDA classifier, and support vector machines (SVM). These methods allow dimension reduction to visualise multivariate data more easily, grouping together correlated samples in a cluster away from those to which they are not correlated.

Hierarchical Cluster Analysis (HCA) is a clustering method which explores the organization of samples in groups and among groups depicting a hierarchy in the form of a dendrogram (Köhn & Hubert, 2015). In this study clustering was achieved by use of Euclidean distance as the metric of samples distance and Ward's Method as the linkage criterion. HCA was used to compare the relationship between spectra of different sample preparation methods. The spectral fingerprint regions were first averaged by sample type, then pre-processed by SG differentiation, vector normalisation, before analysis by HCA.

2.4 Results

2.4.1 Sample preparation affects the spectral profiles of plant material

Consistency of sample preparation is essential to obtaining reliable analytical results and enables conclusions to be compared across studies. This study shows that sample preparation has a marked effects on the spectral fingerprint of plant materials which can be seen visually in Figure 2.1a-c and with the use of chemometrics in Figures 2.2a-c, 2.3 and 2.4. This is consistent with observations in previous studies (Durak & Depciuch, 2020) suggesting that the choice of sample preparation method should be taken into consideration when planning experiments.

Figure 2.1 shows the class mean raw whole, raw, and pre-processed fingerprint spectra for the five different sample preparation methods. These spectral profiles show clear differences dependant on the preparatory technique used. The *in planta* spectral profile has a strong O-H peak centred at 3364 cm^{-1} due to the water in the leaves, as can be seen in Figure 2.1a. Absorbance amplitude in the 1050-1000 cm^{-1} region of the raw fingerprint region (Figure 2.1b) is markedly higher for the freeze-dried-ground samples. Visual differences between the fingerprint regions of spectra both of amplitude and horizontal shifts can be seen in the pre-processed fingerprint region before the use of chemometrics (Figure 2.1c).
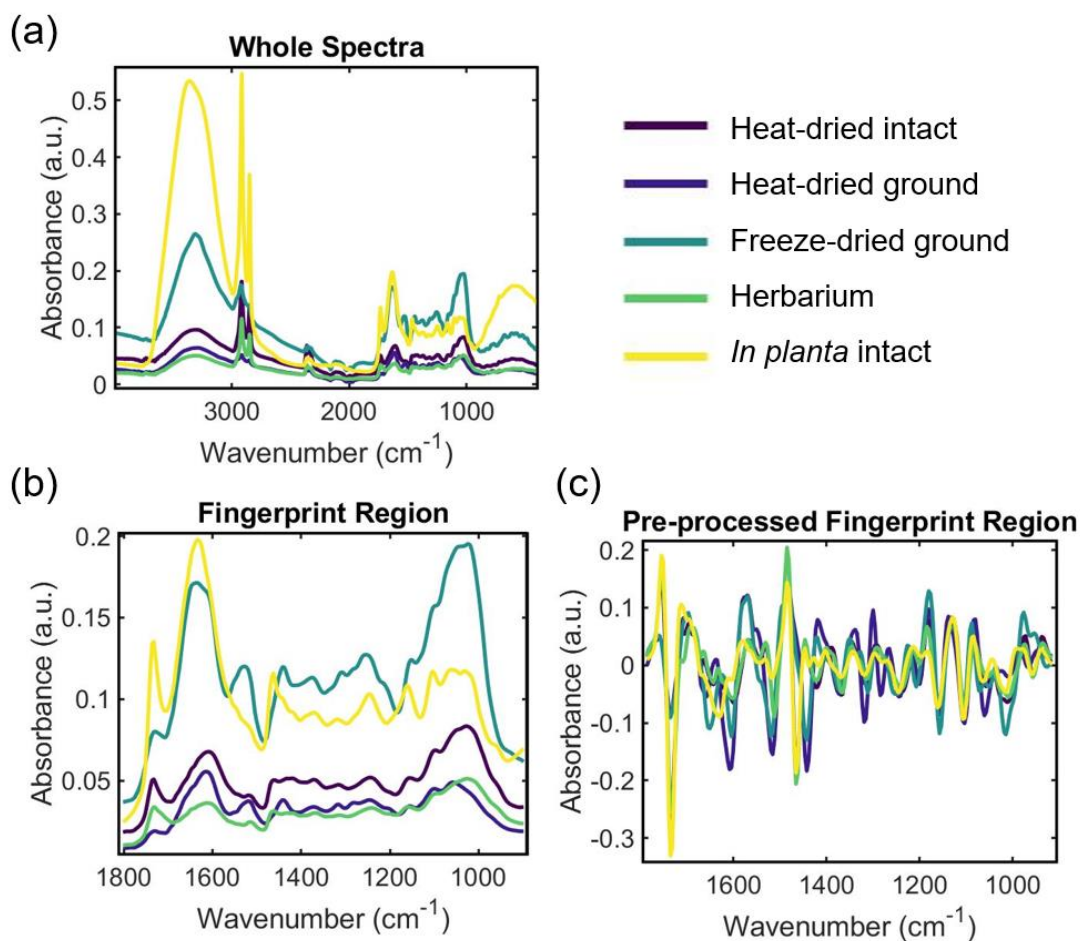
Figure 2.1: a) Whole spectrum, **b)** Fingerprint region and **c)** Pre-processed fingerprint region. These spectral profiles show differences in absorption dependent on the sample preparation technique used. Note the strong O-H hydration peak centred at 3364 cm^{-1} of the *in planta* spectral profile (Figure 2.1a), the high amplitude in the 1050-1000 cm^{-1} region of the raw fingerprint region of freeze-dried ground samples (Figure 2.1b), and the visual differences in the pre-processed spectral fingerprint region (Figure 2.1c).

PCA revealed the natural variation between samples. Figure 2.2a, which was constructed using six PCs, explains 79.03% of the variance. The accuracy for PCA classification was 86.73%; 83.23%, 100%, 87.88%, 69.19%, 93.33% for each sample preparation type respectively (heat-dried intact, heat-dried ground, freeze-dried ground, herbarium, and *in planta* intact). Ground samples, both heat-dried and freeze-dried are separated from the others by PC1. PC1 is the principal component most responsible for differences between samples, indicating that the freeze-drying and grinding caused the most pronounced spectral change. The supervised methods PCA-LDA and SVM were used for classification of samples. LDA effectively distinguishes freeze-dried-ground and *in planta* intact samples based on PCA factors. However, clusters were overlapping for heat-dried intact leaves, heart-dried ground leaves, and herbarium samples. PCA-LDA achieved a cross-validation accuracy of 87.78% in total; 83.94, 100, 87.88, 71.75, 95.33% for each sample preparation type respectively (heat-dried intact, heat-dried ground, freeze-dried ground, herbarium, and *in planta* intact). The predictive performance of PCA-LDA towards the external test set was relatively limited and left some groups overlapping (Figure 2.2b). On the other hand, SVM (cost = 10, γ = 3.1623, N_{SV}: 506) performed better in both training (100% accuracy) and test sets (99.7% accuracy in total; 99.6%, 100%, 99.8%, 99.1%, 99.9% accuracy for heat-dried intact, heat-dried ground, freeze-dried ground, herbarium, and *in planta* intact sample types respectively). The predictive performance of SVM towards the external test set was excellent (99.2%, 100%, 100%, 100%, 100% specificity; 99.9%, 100%, 99.7%, 98.2%, 99.7% sensitivity; 99.2%, 100%, 100%, 100%, 100% precision; for heat-dried intact, heat-dried ground, freeze-dried ground, herbarium, and *in planta* intact sample types respectively).
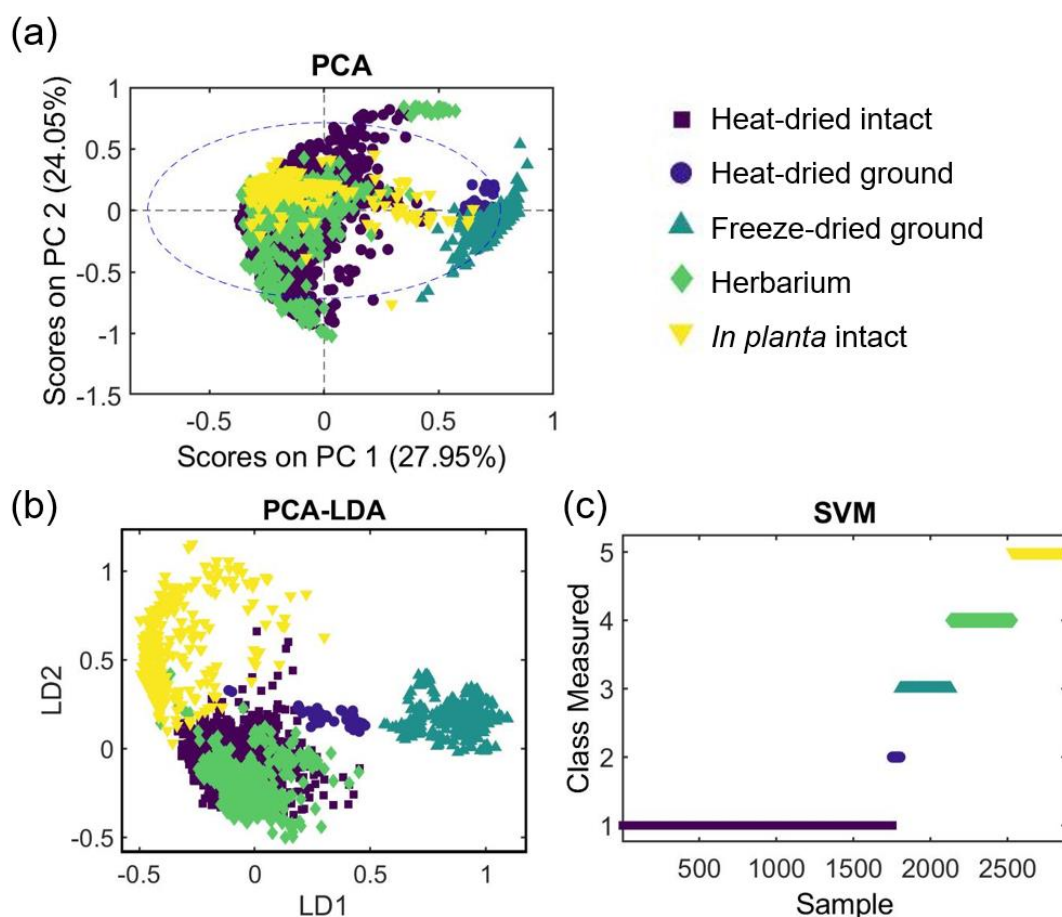
Figure 2.2: a) PCA, **b)** PCA-LDA and **c)** SVM. In Figure 2.2a, the unsupervised PCA analysis separates ground samples (both heat-dried and freeze-dried) from the others along axis PC1, suggesting that grinding is responsible for the most natural variation between samples. The linear discriminant analysis used in Figure 2.2b effectively distinguishes freeze-dried-ground and *in planta* intact samples based on PCA factors. However, clusters were overlapping for heat-dried intact leaves, heart-dried ground leaves, and herbarium samples. The classification performance of SVM, Figure 2.2c, towards the external test set was excellent.

2.4.2 Chemometrics highlights differential relationships between preparation methods

Some techniques are more closely related than others and will give more comparable results. For example, the confusion of herbarium and heat-dried intact leaves in the SVM confusion matrix in Figure 2.3, and their grouping in the HCA in Figure 2.4, is consistent with the similarity of sample preparation techniques between these two methods. Spectra from these two methods also overlap in the PCA (Figure 2.2a) and LDA (Figure 2.2b) two-dimensional space. As can be seen from the SVM confusion matrix in Figure 2.3, herbarium samples were the most often misclassified category and confused with dried intact leaves. The results of the HCA (Figure 2.4) concur with this and group Herbarium samples and Dried Intact samples as the most closely related. The next closest grouping is *in planta* leaves compared with the intact dried or intact herbarium leaves. Ground leaves both heat-dried and freeze-dried form a relationship branch of their own on the dendrogram, with freeze-dried ground being leaves the most distantly related from the others. Together, the results from Figures 2.3 and 2.4 suggest that comparisons between studies on herbarium samples and those on heat-dried intact leaves would give similar results. This is not the case for the other preparatory methods studied, and the biochemical changes caused by the sample preparation may cause differing results.
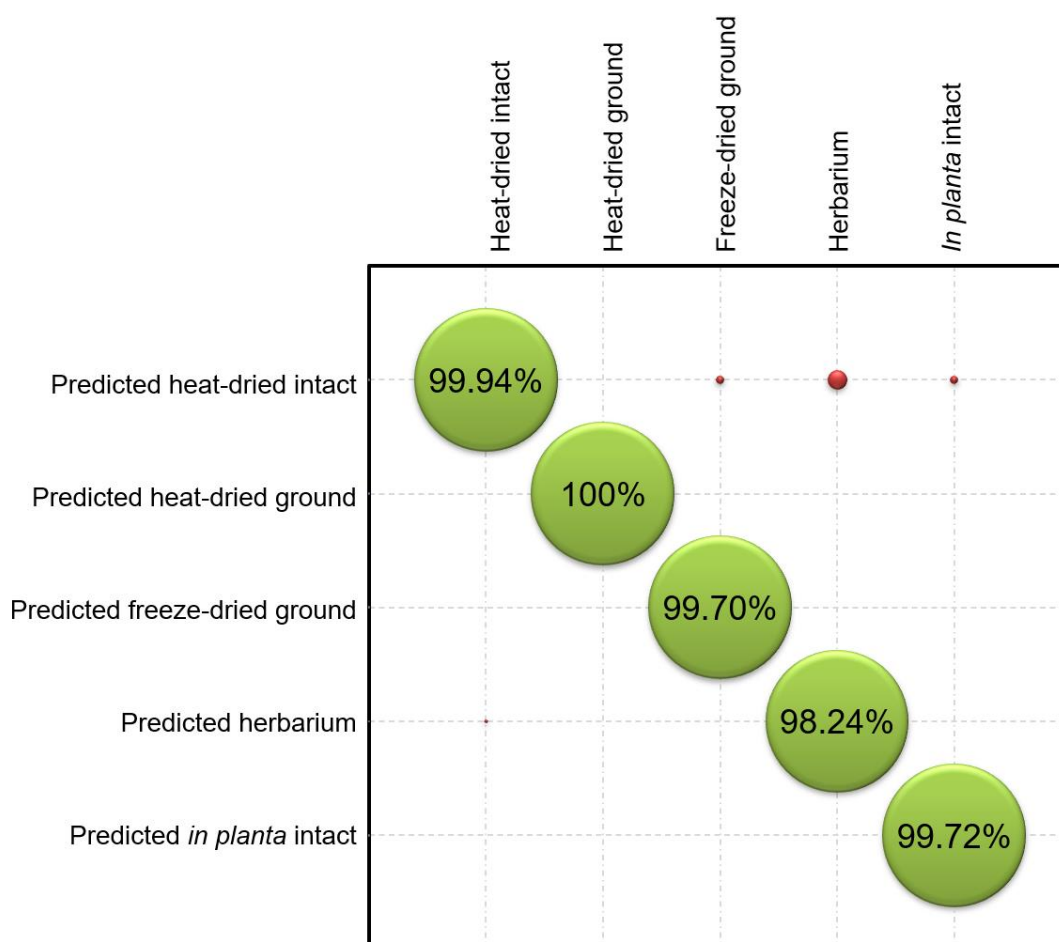
Figure 2.3: SVM confusion matrix. Predictive performance of SVM towards the external test set was overall excellent (99.9%, 100%, 99.7%, 98.2%, 99.7% sensitivity). The relative size of the circles indicates the number predicted as the horizontal category from each vertical sample type. Green indicates a correct prediction, whilst red indicates a misclassification. Herbarium samples were the most frequently misclassified category and were often confused with heat-dried intact leaves.

HCA Dendrogram of Fingerprint Region by Sample Preparation Method
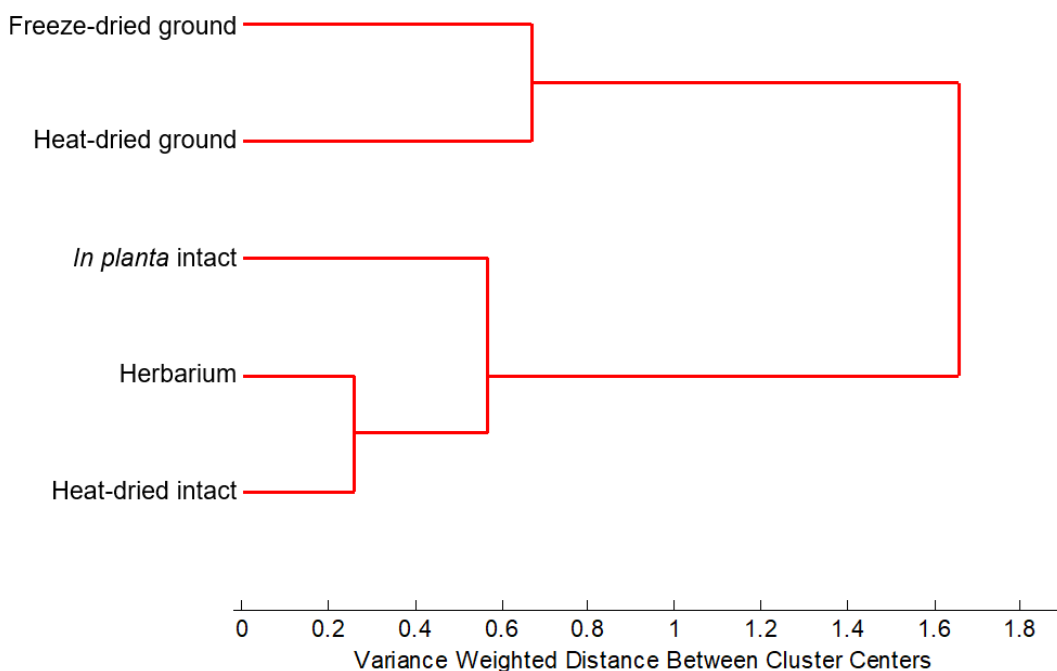
Figure 2.4: HCA dendrogram of fingerprint region. HCA analysis groups spectra from herbarium samples with heat-dried intact samples as the most closely related. The next closest grouping connects *in planta* leaves with the intact heat-dried and intact herbarium leaves. Ground leaves (both heat-dried and freeze-dried) appear on a separate branch within the dendrogram, with freeze-dried ground being leaves the most distantly related from the others.

2.4.3 Spectral biomarkers provide insight into the effect of sample preparation

To explore further the group clustering observed in the two-dimensional discriminant space shown in Figure 2.2b, PCA-LDA loadings were extracted to determine the specific spectral alterations associated with vegetative sample preparation. These markers provide a summary of the main biochemical changes occurring during the preparatory process. Figure 2.5 shows the selected top twelve markers from the LD1 loadings, which highlight the key wavenumbers discriminating samples of each category. The peak at 995 cm^{-1} is associated with "ring breathing" (Talari *et al.*, 2017). Although the curves are almost identical here

for intact heat-dried leaves and herbarium leaves, the spectral profiles greatly differ for the other three sample types. Dried ground spectra have a single peak shifted to lower wavelengths, whilst 995 cm^{-1} forms one half of the saddle-shaped double peak characteristic of in planta samples. Freeze-dried ground samples have a slight artefact on in the spectra here but do not have a peak.
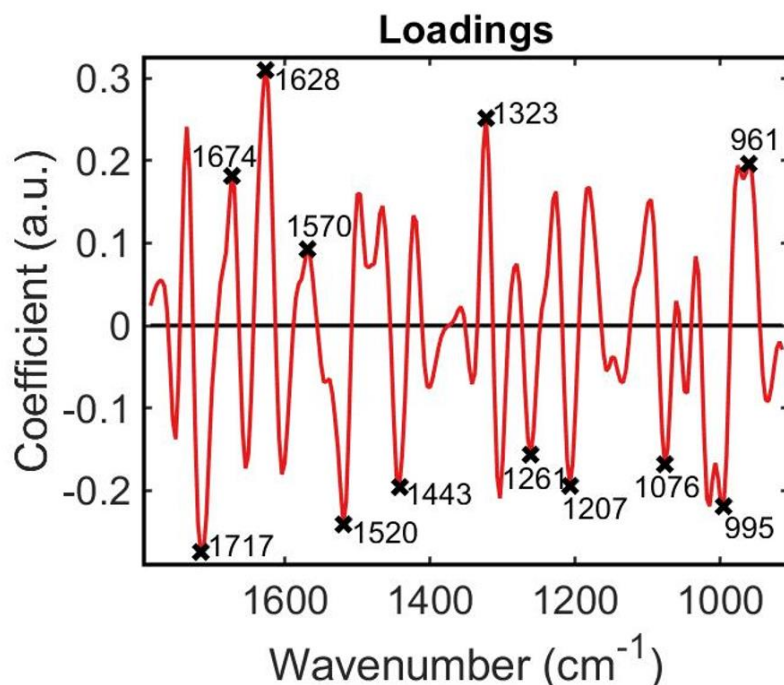
Figure 2.5: PCA-LDA loadings. The selected top twelve markers from the LD1 loadings highlight the key wavenumbers which discriminate between the spectral profiles of samples from each category. These are: 1717, 1674, 1628, 1570, 1520, 1443, 1323, 1261, 1207, 1076, 995, and 961 cm^{-1}.

For a more detailed summary of the main biochemical alterations occurring during the preparatory process PCA-LDA loadings (Figure 2.5) were extracted to determine molecular biomarkers (Table 2.2) for specific spectral variations. The key peaks which differentiate between spectra from different sample preparation methods are 1717 cm^{-1} (C=O thymine, Amide I (arises from C=O stretching vibration), C=O stretching vibration of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), C=O stretching vibration of purine base); 1674 cm^{-1} (Amide I); 1628 cm^{-1} (Amide I); 1570 cm^{-1} (Amide II); 1520 cm^{-1} (Amide II); 1443 cm^{-1} (δ(CH2), lipids, fatty acids); 1323 cm^{-1} (Amide III); 1261 cm^{-1} (PO$_2^-$ asymmetric (Phosphate I)); 1207 cm^{-1} (PO$_2^-$

asymmetric (Phosphate I)); 1076 cm^{-1} (Skeletal cis conformation (CC) of DNA, Symmetric phosphate [PO$_2^-$ (sym)] stretching); 995 cm^{-1} (Ring breathing); 961 cm^{-1} (C−O deoxyribose, C−C) (Qing *et al.*, 1996; Schmidt *et al.*, 2006; Morais *et al.*, 2017; Talari *et al.*, 2017).

Table 2.2: Tentative molecular assignments from PCA loadings. The PCA-LDA loadings were linked to biomarkers to provide a summary of the main biochemical changes occurring during the preparatory process.

Wavenumber/ cm^{-1}	Tentative Molecular Assignment	Reference
1717	C=O thymine, Amide I (arises from C=O stretching vibration), C=O stretching vibration of DNA and RNA, C=O stretching vibration of purine base	Talari *et al.*, 2017
1674	Amide I	Qing *et al.*, 1996
1628	Amide I	Schmidt *et al.*, 2006
1570	Amide II	Talari *et al.*, 2017
1520	Amide II	Talari *et al.*, 2017
1443	δ(CH$_2$), lipids, fatty acids	Talari *et al.*, 2017
1323	Amide III	Morais *et al.*, 2017
1261	PO^{2-} asymmetric (Phosphate I)	Talari *et al.*, 2017
1207	PO^{2-} asymmetric (Phosphate I)	Talari *et al.*, 2017
1076	Skeletal cis conformation (CC) of DNA, Symmetric phosphate [PO^{2-} (sym)] stretching	Talari *et al.*, 2017
995	Ring breathing	Talari *et al.*, 2017
961	C−O deoxyribose, C−C	Talari *et al.*, 2017

Our data show that the grinding process allows different molecules to be analysed, as indicated by the separation of ground samples along PC1 (Figure 2.2a), and the grouping of ground samples in the HCA analysis (Figure 2.4). Grinding allows the spectral acquisition of internal leaf compounds rather than those of the epidermis. For example, one of the key spectral differences between the different sample preparation methods as identified by the PCA loadings (Figure 2.5) was a peak at 1443 cm^{-1}, that was present only in ground samples (both heat and freeze-dried). This peak was assigned to lipids and fatty acids (Talari *et al.*, 2017; see Table 2.2). Conversely, the peak at 1717 cm^{-1} is only present in intact leaves (*in planta*, herbarium and heat dried intact ground), suggesting that it is from a compound found in the leaf epidermis. This peak has been assigned to vibrations from one of the carbonyl bonds of thymine or Amide I (arises from C=O stretching vibration) (Talari *et al.*, 2017). Two key peaks at 1261 and 1207 cm^{-1} are linked to the 'Phosphate I' vibrational mode, the asymmetric stretching of PO_2^- (Talari *et al.*, 2017). These peaks relate to compounds, the concentration of which are highest in the ground samples, indicating they are from the centre of the leaf structure. Therefore, the choice of sample preparation method should be informed by the plant physiology being studied and any molecules expected to be of interest. For example, previous studies have suggested that could freeze-dried ground samples could be useful in assessing viscosity and viscoelasticity (Lan *et al.*, 2020).

2.5 Discussion

2.5.1 Sample preparation has a marked effect on spectra

Preparation of plant vegetative material is often necessary because due to seasonal sample collection, rapid perishability, and time constraints due to access to spectral analysis equipment or number of samples for analysis (Vyankatrao *et al.*, 2014). The results from Figures 2.1-2.5 and Table 2.2 indicate that sample preparation influences the biochemical composition of vegetative tissues and in turn influences their ATR-FTIR spectral profiles. ATR-FTIR spectroscopy is a sensitive technique that is capable of detecting small changes to the fingerprint region induced by

alterations in biochemical composition (Morais *et al.*, 2020). The results of this study suggest that consistency in sample preparation is important in biological studies, as each preparation type is likely to affect different molecules within the sample. For example, excision of leaves can induce biochemical changes such as proline accumulation (Wang CY *et al.*, 1982) and chlorophyll loss (Mondal & Choudhuri, 1981), whilst some chemicals, such as terpenes, are more likely to be persistent after excision (Duff *et al.*, 2022). Drying can affect the chemical composition of leaves through oxidation changes and loss of low boiling chemicals, whilst nutrient levels in dried leaves can become three to four times more concentrated compared with fresh leaves (Babu *et al.*, 2018).

2.5.2 Grinding reveals different compounds of interest

The HCA dendrogram (Figure 2.4), suggests that grinding could allow capture of spectral information from different compounds compared with using intact leaves. All the intact leaves are grouped together (*in planta*, intact heat-dried and intact historic herbarium), whilst both ground sample types (heat-dried ground and freeze-dried ground) appear on a different branch of the HCA dendrogram (Figure 2.4). These results suggest that the spectral profile captures information from compounds in the outer leaf epidermis, such as epicuticular waxes (Dubis *et al.*, 2001), whilst grinding the leaves allows the inside of the leaf to be measured. This hypothesis is consistent with a shallow light penetration depth of 0.5 - 3 μm (Chan & Kazarian, 2006) in comparison with the leaf thickness of Japanese knotweed of approximately 300 μm (Kogami *et al.*, 2001). Although, the previously held assumption that only molecules in the epidermis are captured by ATR-FTIR spectroscopy has been recently challenged in a study reporting a penetration depth extending to over 60 μm (Götz *et al.*, 2020). Regardless of the mechanism, spectral acquisition of the outer surface of dried leaves has proven to be sufficient for the use of ATR-FTIR spectroscopy and chemometrics to classify closely related species and hybrids (Holden *et al.*, 2022). The successful application of this technique using intact leaves could be aided by the preservation of structural information, particularly as the trichome structure in plants is variable by species (Bailey *et al.*,

2009). Whilst grinding might be preferable when studying molecules within the central leaf structure, this preparative technique will result in loss of structural information and could be less suitable for certain uses, such as phylogenetic analysis. The additional step of grinding also adds more time and labour to the process of sample preparation. However, the grinding process reduces variation between samples thereby increasing the accuracy, specificity, and sensitivity of the results.

2.5.3 Advantages and disadvantages of freeze-drying

Freeze-drying, also known as lyophilization, theoretically better preserves less stable potential molecules of interest (Adams *et al.*, 2015). However, the process requires liquid nitrogen on hand at the time of harvest making it less suitable for field studies, where taking spectra directly from living leaves *in planta* on-site is likely to be the best choice for agricultural applications. Therefore, the question of which sample preparation method to use must also consider whether the time-consuming logistics justify the additional biochemical information gained. Two different forms of spectral pre-processing, rubber-band baseline correction and SG second derivative vector normalisation, can be used to provide insight into this issue and compare any concentration differences and bio-molecular structural changes between sample types. If the band intensity at a wavenumber of interest is higher or lower in the baseline corrected spectra (Supplementary Figure S2.1), then the differences can be attributed to a difference in concentration (Morais, 2020). However, if in the SG second derivative and vector normalised spectra (Figure 2.1b) there is a horizontal shift then the difference may be due to a molecular structural alteration (Morais, 2020). The (C–O deoxyribose, C–C) 961 cm^{-1} peak (Talari *et al.*, 2017) is only present in freeze-dried ground samples (shifted to 968 cm-1 in heat-dried samples).

Overall, the baseline corrected spectra show that the concentrations are highest across the full fingerprint region in the *in planta* samples and the freeze-dried ground samples, and the lowest concentrations are in the herbarium samples. This is likely to be because the herbarium samples are the oldest and time has allowed

for the degradation of molecules, whereas the freeze-drying process preserves these biomolecules through the prevention of oxidation and hydrolysis by endogenous enzymes (Lan *et al.*, 2020). At some wavelengths, the concentration of certain molecules in freeze-dried samples appears higher than the *in planta* samples which is likely due to the lack of interference by water in the sample (Lan *et al.*, 2020). However, care should be taken when forming conclusions about proteins from freeze-dried samples. Lyophilized proteins are prone to cross-linking reactions which form amide bonds between free amino groups of proteins with the reducing groups of sugars, and therefore freeze-drying is rarely used for proteomic studies (Carpentier *et al.*, 2007).

2.5.4 The importance of hydration: amide and phosphate vibrational modes

Several of the key peaks in the PCA loadings were associated with Amides I, II and III. These amide bands confer information on the secondary structure of proteins and are sensitive to the protein conformation (Kong & Yu, 2007). Amide I, II, and III bands differ in their sensitivities to changes in protein conformation (Van De Weert *et al.*, 2001). As they are affected by secondary protein structure (Kong & Yu, 2007), these bands are also affected by the hydration of the sample (Bridelli, 2017). Protein structure and activity is dependent on a monolayer of water molecules, known as the first hydration shell (Fogarty & Laage, 2014). The hydrogen bonds formed between protein moieties and water molecules are key to determining the protein's properties (Mallamace *et al.*, 2015).

The spectral region most sensitive to protein secondary structure and hydration is the Amide I band, originating from the C=O stretching vibration of the amide group coupled with the in-phase bending of the N-H bond and stretching of the C-N bond (Mallamace *et al.*, 2015; Yang H *et al.*, 2015). This sensitivity may explain the absence of the peak at 1674 cm^{-1} (Amide I; Qing *et al.*, 1996) in fully hydrated *in planta* samples and presence in dry intact leaves (herbarium and dried intact). New peaks in the Amide I region have been observed elsewhere in solid-state comparisons with solution state proteins, and are thought to be due to intermolecular beta-sheet

formation (Van De Weert *et al.*, 2001). This peak is also absent in the ground samples, suggesting a link to the leaf epidermis. The Amide I (Schmidt *et al.*, 2006) peak at 1628 cm^{-1} is present only in the *in planta* samples. The concentration in the baseline corrected spectra is higher in freeze-dried ground samples than heat-ground samples, suggesting a protein degradation in the heat-dried samples. Geometry, dipole-dipole interaction, and hydrogen bonding all affect the Amide I peak (Baronio & Barth, 2020). At high temperatures aggregation begins and alpha-helices begin to form beta-sheets and self-aggregate (Van De Weert *et al.*, 2001).

Unlike the amide I region, the amide II region is generally not used for determination of protein secondary structures, although it is conformationally sensitive (Yang H *et al.*, 2015). The amide II peak is a combination of N-H in plane bending and C-N stretching (Yang H *et al.*, 2015). The Amide II peak at 1570 cm^{-1} (Talari *et al.*, 2017) has a greater amplitude in ground samples (heat- and freeze-dried) than intact (herbarium and heat-dried) samples but interestingly is absent *in planta.* The Amide II peak at 1520 cm^{-1} (Talari *et al.*, 2017), present only in freeze-dried ground samples, is associated with alpha helix polypeptides (Venyaminov & Kalnin, 1990). This peak is horizontally shifted in the SG second derivative and vector normalised spectra to 1516 cm^{-1} for dried ground samples, suggesting a chemical change. The absence of this peak in the other samples suggests that it arises from compounds within the internal leaf structure.

The amide III associated region arises due to complex vibrational modes. The Amide III peak at 1323 cm^{-1} (Morais *et al.*, 2017), a key difference between sample preparation types, is higher in freeze-dried ground than heat-ground samples. Spectral acquisition of all samples in this study was performed following a background scan of the surrounding air. The differing affects in the Amide I and III regions of the *in planta* samples compared with their dried counterparts may be removed if a background scan of water is used instead (Chan & Kazarian, 2016), however this is outside the scope of this study.

Hydration not only affects protein structures but that of DNA (Berman, 1991). An increase in water content causes DNA to undergo a structural transition from A-

form to B-form (Banavali & Roux, 2005). Both symmetric and antisymmetric stretching vibrations of the phosphate linkage are sensitive to hydration, causing changes in intensity and frequency shifts (Pevsner & Diem, 2001). These hydration effects were observed in this study, with the peak at 1076 cm^{-1}, related to the symmetric PO_2^- stretching and the skeletal cis conformation (CC) of DNA (Talari *et al.*, 2017), present at a much lower concentration in the *in planta* samples. Additionally, the observed horizontal shift of the Phosphate I peak (PO_2^- asymmetric stretching) (Talari *et al.*, 2017) in the SG second differentiated vector normalised spectrum, suggests a biochemical structural change was induced by sample preparation. This peak shifted between 1207 cm^{-1} (freeze-dried-ground samples), 1200 cm^{-1} (dried intact, dried ground, and herbarium samples), and 1196 cm^{-1} (*in planta* samples). Therefore, spectral measurements focused on proteins, DNA or RNA should be *in planta* to maintain the secondary structure in its functioning state. However, this raises the additional challenge of maintaining the plant growth environment constant during spectral acquisition to prevent impacts to the resultant spectra from inducible plant responses to light (Legris *et al.*, 2017), temperature (Atkin & Tjoelker, 2003), or humidity (Osakabe *et al.*, 2014). Plants respond to abiotic stress through signalling cascades which affect a wide variety of biomolecules, including polyamines (Alcázar *et al.*, 2006), lipids (Hou Q *et al.*, 2016), sterols (Rogowska & Szakiel, 2020), RNA (Jha *et al.*, 2020) and secondary metabolites (Yang L *et al.*, 2018). Accurate spectral acquisition of *in planta* samples is therefore reliant on portable mid-infrared (MIR) spectrometers (Dhawale *et al.*, 2015; Ji *et al.*, 2016; Soriano-Disla *et al.*, 2018; Bureau *et al.*, 2019; Hutengs *et al.*, 2019) to allow spectral acquisition within the plant growth environment.

2.6 Conclusion

Our results clearly demonstrate that sample preparation of plant vegetative tissue prior to spectral acquisition affects resultant spectra. Additionally, grinding has been shown to allow spectral acquisition from different compounds of interest, and that historic herbarium specimen can prove a viable sample despite their age. Importantly, the sample preparation method used should be tailored to the

experimental design considering factors such as the expected location and stability of the compounds of interest, the type of plant material being studied and the physical location of plants from which spectra are to be obtained, for example, in a controlled growth environment or in the field.

2.7 References

Adams GDJ, Cook I, Ward KR. **2015**. The principles of freeze-drying. *Methods in Molecular Biology* **1257**: 121–143.

Agatonovic-Kustrin S, Doyle E, Gegechkori V, Morton DW. **2020**. High-performance thin-layer chromatography linked with (bio)assays and FTIR-ATR spectroscopy as a method for discovery and quantification of bioactive components in native Australian plants. *Journal of Pharmaceutical and Biomedical Analysis* **184**: 113208.

Alcázar R, Marco F, Cuevas JC, Patron M, Ferrando A, Carrasco P, Tiburcio AF, Altabella T. **2006**. Involvement of polyamines in plant response to abiotic stress. *Biotechnology Letters* **28**: 1867–1876.

Amano T, Coverdale R, Peh KSH. **2016**. The importance of globalisation in driving the introduction and establishment of alien species in Europe. *Ecography* **39**: 1118–1128.

Aouidi F, Dupuy N, Artaud J, Roussos S, Msallem M, Perraud-Gaime I, Hamdi M. **2012**. Discrimination of five Tunisian cultivars by Mid InfraRed spectroscopy combined with chemometric analyses of olive Olea europaea leaves. *Food Chemistry* **131**: 360–366.

Armstrong K., Ball S. 2005. DNA barcodes for biosecurity: invasive species identification. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**: 1813–1823.

Atkin OK, Tjoelker MG. **2003**. Thermal acclimation and the dynamic response of plant respiration to temperature. *Trends in Plant Science* **8**: 343–351.

Babu AK, Kumaresan G, Raj VAA, Velraj R. **2018**. Review of leaf drying: Mechanism and influencing parameters, drying methods, nutrient preservation, and mathematical models. *Renewable and Sustainable Energy Reviews* **90**: 536–556.

Bailey JP, Bímová K, Mandák B. **2009**. Asexual spread versus sexual reproduction and evolution in Japanese Knotweed *s.l.* sets the stage for the 'battle of the Clones'. *Biological Invasions*. ***11*(5)**: 1189-1203.

Banavali NK, Roux BB. **2005**. Free Energy Landscape of A-DNA to B-DNA Conversion in Aqueous Solution. **127(18)**: 6866-6876.

Baronio CM, Barth A. **2020**. The Amide i Spectrum of Proteins - Optimization of Transition Dipole Coupling Parameters Using Density Functional Theory Calculations. *Journal of Physical Chemistry B* **124**: 1703–1714.

Berardo N, Pisacane V, Battilani P, Scandolara A, Pietri A, Marocco A. **2005**. Rapid detection of kernel rots and mycotoxins in maize by near-infrared reflectance spectroscopy. *Journal of Agricultural and Food Chemistry* **53**: 8128–8134.

Berman HM. **1991**. Hydration of DNA. *Current Opinion in Structural Biology* **1**: 423–427.

Bradshaw CJA, Leroy B, Bellard C, Roiz D, Albert C, Fournier A, Barbet-Massin M, Salles JM, Simard F, Courchamp F. **2016**. Massive yet grossly underestimated global costs of invasive insects. *Nature Communications* **7**: 1–8.

Bridelli MG. **2017**. Fourier Transform Infrared Spectroscopy in the Study of Hydrated Biological Macromolecules. In: Fourier Transforms - High-tech Application and Current Trends. InTech. 191-213.

Buitrago MF, Groen TA, Hecker CA, Skidmore AK. **2016**. Changes in thermal infrared spectra of plants caused by temperature and water stress. *ISPRS Journal of Photogrammetry and Remote Sensing* **111**: 22–31.

Bureau S, Cozzolino D, Clark CJ. **2019**. Contributions of Fourier-transform mid infrared (FT-MIR) spectroscopy to the study of fruit and vegetables: A review. *Postharvest Biology and Technology* **148**: 1–14.

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

Butler HJ, Martin FL, Roberts MR, Adams S, McAinsh MR. **2020**. Observation of nutrient uptake at the adaxial surface of leaves of tomato (*Solanum lycopersicum*) using Raman spectroscopy. *Analytical Letters* **53**: 536–562.

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

Carpentier SC, Dens K, Van den houwe I, Swennen R, Panis B. **2007**. Lyophilization, a Practical Way to Store and Transport Tissues Prior to Protein Extraction for 2DE Analysis? *PROTEOMICS* **7**: 64–69.

Chan KLA, Kazarian SG. **2006**. ATR-FTIR spectroscopic imaging with expanded field of view to study formulations and dissolution. *Lab on a Chip* **6**: 864–870.

Chan KLA, Kazarian SG. **2016**. Attenuated total reflection Fourier-transform infrared (ATR-FTIR) imaging of tissues and live cells. *Chemical Society Reviews* **45**: 1850–1864.

Cirino de Carvalho L, de Lelis Medeiros de Morais C, Gomes de Lima KM, Cunha Júnior LC, Martins Nascimento PA, Bosco de Faria J, Henrique de Almeida Teixeira G. **2016a**. Determination of the geographical origin and ethanol content of Brazilian sugarcane spirit using near-infrared spectroscopy coupled with discriminant analysis. *Analytical Methods* **8**: 5658–5666.

Cirino de Carvalho L, de Lelis Medeiros de Morais C, Gomes de Lima KM, Cunha Júnior LC, Martins Nascimento PA, Bosco de Faria J, Henrique de Almeida Teixeira G. **2016b**. Determination of the geographical origin and ethanol content of Brazilian sugarcane spirit using near-infrared spectroscopy coupled with discriminant analysis. *Analytical Methods* **8**: 5658–5666.

Delabre I, Rodriguez LO, Smallwood JM, Scharlemann JPW, Alcamo J, Antonarakis AS, Rowhani P, Hazell RJ, Aksnes DL, Balvanera P, *et al.* 2021. Actions on sustainable food production and consumption for the post-2020 global biodiversity framework. *Science Advances* **7**: 8259.

Demir P, Onde S, Severcan F. **2015**. Phylogeny of cultivated and wild wheat species using ATR-FTIR spectroscopy. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy* **135**: 757–763.

Dhawale NM, Adamchuk VI, Prasher SO, Viscarra Rossel RA, Ismail AA, Kaur J. **2015**. Proximal soil sensing of soil texture and organic matter with a prototype portable mid-infrared spectrometer. *European Journal of Soil Science* **66**: 661–669.

Dubis EN, Dubis AT, Popławski J. **2001**. Determination of the aromatic compounds in plant cuticular waxes using FT-IR spectroscopy. *Journal of Molecular Structure* **596**: 83–88.

Duff T, Porteous MG, Goodger JQD. **2022**. Wildfires, Flammable Oils, and Eucalyptus Trees: The Persistence and Volatility of Terpenes in Excised Leaves. *SSRN Electronic Journal*.

Durak T, Depciuch J. **2020**. Effect of plant sample preparation and measuring methods on ATR-FTIR spectra results. *Environmental and Experimental Botany* **169**: 103915.

ElMasry G, Wang N, ElSayed A, Ngadi M. **2007**. Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry. *Journal of Food Engineering* **81**: 98–107.

Emsley NEM, Holden CA, Guo S, Bevan RS, Rees C, McAinsh MR, Martin FL, Morais CLM. **2021**. Machine Learning Approach Using a Handheld Near-Infrared (NIR) Device to Predict the Effect of Storage Conditions on Tomato Biomarkers. *ACS Food Science & Technology* **2(1)**: 187-194.

Eriksson BK, Hillebrand H. **2019**. Rapid reorganization of global biodiversity. *Science* **366**: 308–309.

Euring D, Löfke C, Teichmann T, Polle A. **2012**. Nitrogen fertilization has differential effects on N allocation and lignin in two Populus species with contrasting ecology. *Trees - Structure and Function* **26**: 1933–1942.

Fogarty AC, Laage D. **2014**. Water dynamics in protein hydration shells: The molecular origins of the dynamical perturbation. *Journal of Physical Chemistry B* **118**: 7715–7729.

Fraser EDG. **2020**. The challenge of feeding a diverse and growing population. *Physiology & Behavior* **221**: 112908.

Gendrin C, Roggo Y, Collet C. **2008**. Pharmaceutical applications of vibrational chemical imaging and chemometrics: A review. *Journal of Pharmaceutical and Biomedical Analysis* **48**: 533–553.

Godfray HCJ, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, Pretty J, Robinson S, Thomas SM, Toulmin C. **2010**. Food security: The challenge of feeding 9 billion people. *Science* **327**: 812–818.

Götz A, Nikzad-Langerodi R, Staedler Y, Bellaire A, Saukel J. **2020**. Apparent penetration depth in attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy of Allium cepa L. epidermis and cuticle. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **224**: 117460.

Gudi G, Krähmer A, Krüger H, Schulz H. **2015**. Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy on Intact Dried Leaves of Sage (Salvia officinalis L.): Accelerated Chemotaxonomic Discrimination and Analysis of Essential Oil Composition. *Journal of Agricultural and Food Chemistry* **63**: 8743–8750.

Holden CA, Bailey JP, Taylor JE, Martin FL, Beckett P, McAinsh M. **2022**. Know your enemy: Application of ATR-FTIR spectroscopy to invasive species control (D Changwen, Ed.). *PLOS ONE* **17**: e0261742.

Holden CA, Morais CLM, Taylor JE, Martin FL, Beckett P, McAinsh M. **2021**. Regional differences in clonal Japanese knotweed revealed by chemometrics-linked attenuated total reflection Fourier-transform infrared spectroscopy. *BMC Plant Biology 2021 21:1* **21**: 1–20.

Hou Q, Ufer G, Bartels D. **2016**. Lipid signalling in plant responses to abiotic stress. *Plant, Cell & Environment* **39**: 1029–1048.

Hulme PE. **2009**. Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology* **46**: 10–18.

Hulme PE. **2020**. Plant invasions in New Zealand: global lessons in prevention, eradication and control. *Biological Invasions* **22**: 1539–1562.

Hutengs C, Seidel M, Oertel F, Ludwig B, Vohland M. **2019**. In situ and laboratory

soil spectroscopy with portable visible-to-near-infrared and mid-infrared instruments for the assessment of organic carbon in soils. *Geoderma* **355**: 113900.

Jeger M, Beresford R, Bock C, Brown N, Fox A, Newton A, Vicent A, Xu X, Yuen J. **2021**. Global challenges facing plant pathology: multidisciplinary approaches to meet the food security and environmental challenges in the mid-twenty-first century. *CABI Agriculture and Bioscience 2021 2:1* **2**: 1–18.

Jha UC, Nayyar H, Jha R, Khurshid M, Zhou M, Mantri N, Siddique KHM. **2020**. Long non-coding RNAs: emerging players regulating plant abiotic stress response and adaptation. *BMC Plant Biology 2020 20:1* **20**: 1–20.

Ji W, Adamchuk VI, Biswas A, Dhawale NM, Sudarsan B, Zhang Y, Viscarra Rossel RA, Shi Z. **2016**. Assessment of soil properties in situ using a prototype portable MIR spectrometer in two agricultural fields. *Biosystems Engineering* **152**: 14–27.

Kelly LT, Giljohann KM, Duane A, Aquilué N, Archibald S, Batllori E, Bennett AF, Buckland ST, Canelles Q, Clarke MF, *et al.* 2020. Fire and biodiversity in the Anthropocene. *Science* **370**.

Khairudin K, Sukiran NA, Goh HH, Baharum SN, Noor NM. **2014**. Direct discrimination of different plant populations and study on temperature effects by Fourier transform infrared spectroscopy. *Metabolomics* **10**: 203–211.

Kim SW, Ban SH, Chung H, Cho S, Chung HJ, Choi PS, Yoo OJ, Liu JR. **2004**. Taxonomic discrimination of flowering plants by multivariate analysis of Fourier transform infrared spectroscopy data. *Plant Cell Reports* **23**: 246–250.

Kogami H, Hanba YT, Kibe T, Terashima I, Masuzawa T. **2001**. CO_2 transfer conductance, leaf structure and carbon isotope composition of *Polygonum cuspidatum* leaves from low and high altitudes. *Plant, Cell & Environment* **24**: 529–538.

Köhn H-F, Hubert LJ. **2015**. Hierarchical Cluster Analysis. In: Wiley StatsRef: Statistics Reference Online. John Wiley & Sons, Ltd, 1–13.

Kong J, Yu S. **2007**. Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures. *Acta Biochimica et Biophysica Sinica* **39**: 549–559.

Kos G, Lohninger H, Krska R. **2003**. Development of a method for the determination of Fusarium fungi on corn using mid-infrared spectroscopy with attenuated total reflection and chemometrics. *Analytical Chemistry* **75**: 1211–1217.

Lan W, Renard CMGC, Jaillais B, Leca A, Bureau S. **2020**. Fresh, freeze-dried or cell wall samples: Which is the most appropriate to determine chemical, structural and rheological variations during apple processing using ATR-FTIR spectroscopy? *Food Chemistry* **330**: 127357.

Lee R. **2011**. The outlook for population growth. *Science* **333**: 569–573.

Legris M, Nieto C, Sellaro R, Prat S, Casal JJ. **2017**. Perception and signalling of light and temperature cues in plants. *The Plant Journal* **90**: 683–697.

MacLeod M, Arp HPH, Tekman MB, Jahnke A. **2021**. The global threat from plastic pollution. *Science* **373**: 61–65.

Madden MJL, Young RG, Brown JW, Miller SE, Frewin AJ, Hanner RH. **2019**. Using DNA barcoding to improve invasive pest identification at U.S. ports-of-entry. *PLOS ONE* **14**: e0222291.

Mahlein AK. **2016**. Plant disease detection by imaging sensors – Parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease* **100**: 241–254.

Mallamace F, Corsaro C, Mallamace D, Vasi S, Vasi C, Dugo G. **2015**. The role of water in protein's behavior: The two dynamical crossovers studied by NMR and FTIR techniques. *Computational and Structural Biotechnology Journal* **13**: 33–37.

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

McNear DH, Chaney RL, Sparks DL. **2010**. The hyperaccumulator Alyssum murale uses complexation with nitrogen and oxygen donor ligands for Ni transport and storage. *Phytochemistry* **71**: 188–200.

Mondal R, Choudhuri MA. **1981**. Role of Hydrogen Peroxide in Senescence of Excised Leaves of Rice and Maize. *Biochemie und Physiologie der Pflanzen* **176**: 700–709.

Morais CL. **2020**. Novel chemometric approaches towards handling biospectroscopy datasets. *(PhD Thesis) University of Central Lancashire, Preston*.

Morais CLM, Costa FSL, Lima KMG. **2017**. Variable selection with a support vector machine for discriminating Cryptococcus fungal species based on ATR-FTIR spectroscopy. *Analytical Methods* **9**: 2964–2970.

Morais CLM, Lima KMG, Singh M, Martin FL. **2020**. Tutorial: multivariate classification for vibrational spectroscopy in biological samples. *Nature Protocols* **15**: 2143–2162.

Morais CLM, Paraskevaidi M, Cui L, Fullwood NJ, Isabelle M, Lima KMG, Martin-Hirsch PL, Sreedhar H, Trevisan J, Walsh MJ, *et al.* 2019. Standardization of complex biologically derived spectrochemical datasets. *Nature Protocols* **14**: 1546–1577.

Ord J, Butler HJ, McAinsh MR, Martin FL. **2016**. Spectrochemical analysis of

sycamore (Acer pseudoplatanus) leaves for environmental health monitoring. *The Analyst* **141**: 2896–2903.

Osakabe Y, Osakabe K, Shinozaki K, Tran LSP. **2014**. Response of plants to water stress. *Frontiers in Plant Science* **5**: 86.

Paini DR, Sheppard AW, Cook DC, De Barro PJ, Worner SP, Thomas MB. **2016**. Global threat to agriculture from invasive species. *Proceedings of the National Academy of Sciences of the United States of America* **113**: 7575–7579.

Pevsner A, Diem M. **2001**. Infrared spectroscopic studies of major cellular components. Part II: The effect of hydration on the spectra of nucleic acids. *Applied Spectroscopy* **55**: 1502–1505.

Pyšek P, Hulme PE, Simberloff D, Bacher S, Blackburn TM, Carlton JT, Dawson W, Essl F, Foxcroft LC, Genovesi P, *et al.* 2020. Scientists' warning on invasive alien species. *Biological Reviews* **95**: 1511–1534.

Qing H, Yanlin H, Fenlin S, Zuyi T. **1996**. Effects of pH and metal ions on the conformation of bovine serum albumin in aqueous solution: An Attenuated Total Reflection (ATR) FTIR spectroscopic study. *Spectrochimica Acta - Part A Molecular Spectroscopy* **52**: 1795–1800.

Rana R, Müller G, Naumann A, Polle A. **2008**. FTIR spectroscopy in combination with principal component analysis or cluster analysis as a tool to distinguish beech (Fagus sylvatica L.) trees grown at different sites. *Holzforschung* **62**: 530–538.

Ribeiro da Luz B. **2006**. Attenuated total reflectance spectroscopy of plant leaves: a tool for ecological and botanical studies. *New Phytologist* **172**: 305–318.

Rogowska A, Szakiel A. **2020**. The role of sterols in plant response to abiotic stress. *Phytochemistry Reviews* **19**: 1525–1538.

Sankaran S, Maja J, Buchanon S, Ehsani R. **2013**. Huanglongbing (Citrus Greening) Detection Using Visible, Near Infrared and Thermal Imaging Techniques. *Sensors* **13**: 2117–2130.

Schmidt P, Dybal J, Trchová M. **2006**. Investigations of the hydrophobic and hydrophilic interactions in polymer-water systems by ATR FTIR and Raman spectroscopy. *Vibrational Spectroscopy* **42**: 278–283.

Shackleton RT, Shackleton CM, Kull CA. **2019**. The role of invasive alien species in shaping local livelihoods and human well-being: A review. *Journal of Environmental Management* **229**: 145–157.

Skolik P, McAinsh MR, Martin FL. **2018**. Biospectroscopy for Plant and Crop Science. In: Comprehensive Analytical Chemistry. Elsevier B.V., 15–49.

Skolik P, McAinsh MR, Martin FL. **2019a**. ATR-FTIR spectroscopy non-

destructively detects damage-induced sour rot infection in whole tomato fruit. *Planta* **249**: 925–939.

Skolik P, Morais CLM, Martin FL, McAinsh MR. **2019b**. Determination of developmental and ripening stages of whole tomato fruit using portable infrared spectroscopy and Chemometrics. *BMC Plant Biology* **19**: 236.

Smith B. **2018**. Infrared spectral interpretation: A systematic approach. *CRC Press.* 1–304.

Soriano-Disla JM, Janik LJ, McLaughlin MJ. **2018**. Assessment of cyanide contamination in soils with a handheld mid-infrared spectrometer. *Talanta* **178**: 400–409.

Springer NP, Duchin F. **2014**. Feeding nine billion people sustainably: Conserving land and water through shifting diets and changes in technologies. *Environmental Science and Technology* **48**: 4444–4451.

Stordrange L, Libnau FO, Malthe-Sørenssen D, Kvalheim OM. **2002**. Feasibility study of NIR for surveillance of a pharmaceutical process, including a study of different preprocessing techniques. *Journal of Chemometrics* **16**: 529–541.

Strgulc Krajšek S, Buh P, Zega A, Kreft S. **2008**. Identification of Herbarium Whole-Leaf Samples ofEpilobium Species by ATR-IR Spectroscopy. *Chemistry & Biodiversity* **5**: 310–317.

Talari ACS, Martinez MAG, Movasaghi Z, Rehman S, Rehman IU. **2017**. Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Applied Spectroscopy Reviews* **52**: 456–506.

Traoré M, Kaal J, Martínez Cortizas A. **2018**. Differentiation between pine woods according to species and growing location using FTIR-ATR. *Wood Science and Technology* **52**: 487–504.

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

Trujillo-González A, Thuo DN, Divi U, Sparks K, Wallenius T, Gleeson D. **2022**. Detection of Khapra Beetle Environmental DNA Using Portable Technologies in Australian Biosecurity. *Frontiers in Insect Science* **2**. 795379.

Usman K, Al-Ghouti MA, Abu-Dieyeh MH. **2019**. The assessment of cadmium, chromium, copper, and nickel tolerance and bioaccumulation by shrub plant Tetraena qataranse. *Scientific Reports* **9**: 1–11.

Venyaminov SY, Kalnin NN. **1990**. Quantitative IR spectrophotometry of peptide compounds in water (H2O) solutions. II. Amide absorption bands of polypeptides and fibrous proteins in α-, β-, and random coil conformations. *Biopolymers* **30**:

1259–1271.

Vohra K, Marais EA, Bloss WJ, Schwartz J, Mickley LJ, Van Damme M, Clarisse L, Coheur PF. **2022**. Rapid rise in premature mortality due to anthropogenic air pollution in fast-growing tropical cities from 2005 to 2018. *Science Advances* **8**: 4435.

Vyankatrao N, Arts NT, Commerce VS. 2014. Effect of drying methods on nutritional value of some vegetables. *Proceeding of the National Conference on Conservation of Natural Resources & Biodiversity for Sustainable Development* **6**: 72-79.

Wang CY, Cheng SH, Kao CH. **1982**. Senescence of Rice Leaves VII. PROLINE ACCUMULATION IN SENESCING EXCISED LEAVES. *Plant Physiology* **69**: 1348–1349.

Van De Weert M, Haris PI, Hennink WE, Crommelin DJA. **2001**. Fourier transform infrared spectrometric analysis of protein conformation: Effect of sampling method and stress factors. *Analytical Biochemistry* **297**: 160–169.

Willett W, Rockström J, Loken B, Springmann M, Lang T, Vermeulen S, Garnett T, Tilman D, DeClerck F, Wood A, *et al.* 2019. Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems. *The Lancet* **393**: 447–492.

Williams P, Geladi P, Fox G, Manley M. **2009**. Maize kernel hardness classification by near infrared (NIR) hyperspectral imaging and multivariate data analysis. *Analytica Chimica Acta* **653**: 121–130.

Yang L, Wen KS, Ruan X, Zhao YX, Wei F, Wang Q. **2018**. Response of Plant Secondary Metabolites to Environmental Factors. *Molecules 2018, Vol. 23, Page 762* **23**: 762.

Yang H, Yang S, Kong J, Dong A, Yu S. **2015**. Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy. *Nature Protocols* **10**: 382–396.

2.8 Supporting Information
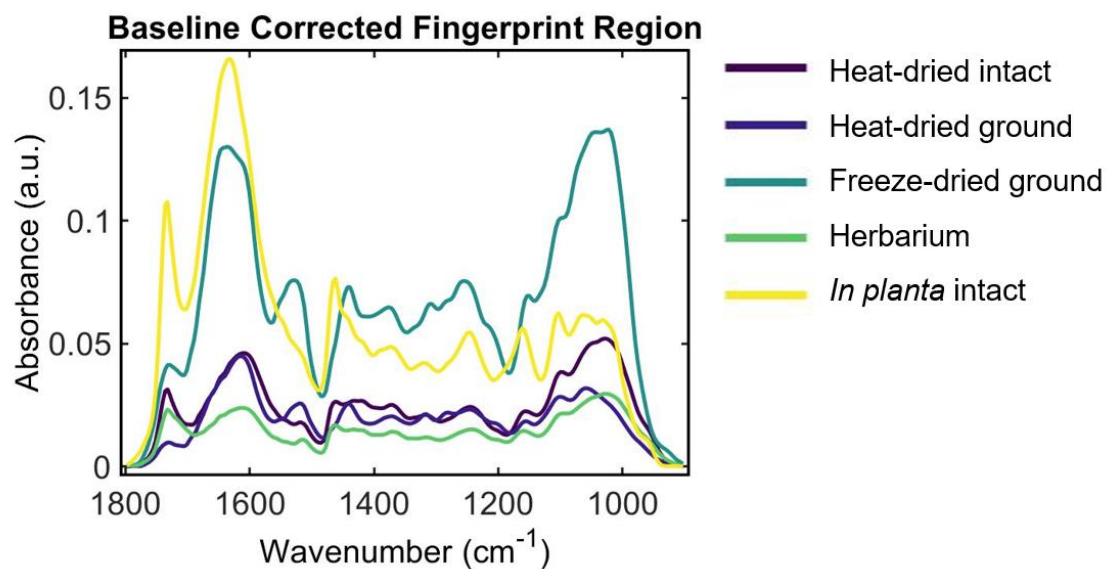
Figure S2.1: Rubber-band baseline corrected spectra fingerprint region grouped into class means for each sample preparation method: heat-dried intact, heat-dried ground, freeze-dried ground, herbarium and *in planta* intact.

3 Regional differences in clonal Japanese knotweed revealed by chemometrics-linked Attenuated Total Reflection Fourier-Transform Infrared Spectroscopy

Claire A Holden[1]*, Camilo LM Morais[2], Jane E Taylor[1], Francis L Martin[3], Paul Beckett[4], Martin McAinsh[5]

1. Lancaster Environment Centre, Lancaster University, LA1 4YQ, UK

2. School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston, PR1 2HE, UK

3. Biocel Ltd, Hull, HU107TS, UK

4. Phlorum Ltd, Brighton, BN2 6AH, UK

*Corresponding author: Claire Holden, c.holden6@lancaster.ac.uk

3.1 Abstract

Background: Japanese knotweed (*R. japonica* var. *japonica)* is one of the world's 100 worst invasive species, causing crop losses, damage to infrastructure, and erosion of ecosystem services. In the UK, this species is an all-female clone, which spreads by vegetative reproduction. Despite this genetic continuity, Japanese knotweed can colonise a wide variety of environmental habitats. However, little is known about the phenotypic plasticity responsible for the ability of Japanese knotweed to invade and thrive in such diverse habitats. Attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy, in which the spectral fingerprint generated allows subtle differences in composition to be clearly visualized, has been used to examine regional differences in clonal Japanese knotweed.

Results: Distinct differences in the spectral fingerprint region (1800–900 cm^{-1}) of Japanese knotweed from three different regions in the UK were sufficient to successfully identify plants from different geographical regions with high accuracy using support vector machine (SVM) chemometrics.

Conclusions: These differences were not correlated with environmental variations between regions, raising the possibility that epigenetic modifications may contribute to the phenotypic plasticity responsible for the ability of *R. japonica* to invade and thrive in such diverse habitats.

3.2 Keywords

3.3 Introduction

Invasive Alien Species (IAS) constitute a major threat to global biodiversity (Roy *et al.*, 2018). Japanese knotweed (*Reynoutria japonica* var. *japonica*) (Lowe *et al.*, 2000; Centre for Agriculture and Bioscience International (CABI), 2018) is particularly invasive throughout North America, Europe, Australia and New Zealand (Centre for Agriculture and Bioscience International (CABI), 2018; Nentwig *et al.*, 2018). It grows vigorously into tall, dense monodominant clumps or 'stands' which possess a shared underground rhizome system and spread vegetatively to expand up to several metres outwards and over a metre downwards from the parent plant (Aguilera *et al.*, 2010; Fennell *et al.*, 2018). These stands have marked negative effects on the environment (Lavoie, 2017) including: a reduction in ecosystem services in riparian zones (Kidd & Shaw, 2000; Urgenson, 2006; Gerber *et al.*, 2008; Environment Agency, 2013), the weakening of flood defences (Kidd & Shaw, 2000; Environment Agency, 2013; Booy *et al.*, 2015), changes in species abundance (Geng *et al.,* 2007, 2016; Estoup *et al.,* 2016; Jones D *et al.,* 2018) and a diminution in property values (Santo, 2017; Fennell *et al.*, 2018). The socio-economic cost also includes the expense of control measures, usually comprising repeated treatments with the controversial herbicide, glyphosate (Jones D *et al.*, 2018).

Many hypotheses have been proposed to explain invasive success, including: fluctuating-resource-availability, enemy-release, Evolution of Increased Competitive Ability (EICA), naturalization, diversity-invasibility, novel weapons (NWH), shifting-defence, hybridisation-invasion, ecotype, plasticity (Gurevitch *et al.*, 2011; Geng *et al.*, 2016; van Kleunen *et al.*, 2018; Banerjee *et al.*, 2019). The ecotype hypothesis suggests that the ability of invasive plant species to thrive across different habitats is underpinned by genetic variations leading to local adaptations (Geng *et al.*, 2007). Whilst in some cases a genetic bottleneck can reduce the fitness of the resultant population, paradoxically some invasive species appear to thrive despite a reduced genetic diversity in the founder population (Estoup *et al.*, 2016). Japanese knotweed exhibits minimal genetic variation in Central Europe (Zhang *et al.*, 2016), Norway (Holm *et al.*, 2018) and the USA (Richards *et al.*, 2012), and exists

as a female clone in the United Kingdom from a single introduction (Bailey & Conolly, 2000; Hollingsworth & Bailey, 2000a). The ability of populations with low genetic diversity such as Japanese knotweed to take advantage of a wider ecological niche has been attributed to phenotypic plasticity (Parker *et al.*, 2003; Bossdorf *et al.*, 2005; Richards *et al.*, 2006, 2008; Van Kleunen *et al.*, 2010; Pichancourt & van Klinken, 2012; Hagenblad *et al.*, 2015; Geng *et al.*, 2016), efficient resource partitioning (Price *et al.*, 2002) and vegetative regeneration (Bímová *et al.*, 2003), allowing *R. japonica* to colonise a broad geographic range across diverse habitats such as riparian wetlands, urban transport courses, and coastal areas (Richards *et al.*, 2012; Zhang *et al.*, 2016). However, there is also evidence that phenotypic plasticity is similar between invasive plants and native or non-invasive closely related species (Davidson *et al.*, 2011; Palacio-López & Gianoli, 2011). Previous studies have suggested that invasive plants are phytochemically unique in their new habitats, conferring advantages such as antiherbivore, antifungal, antimicrobial and allelopathic effects (Cappuccino & Arnason, 2006). Plant anti-herbivory and anti-pathogen defences are conferred in part by phenolic phytochemicals such as tannins, lignin and quercetin (Singh *et al.*, 2021).

Here, the technique of attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy combined with chemometrics has been employed to study the biomolecular adaptations of clonal *R. japonica* to growth in habitats with contrasting soil characteristics and climatic conditions. This technique provides rapid, marker-free, non-destructive analysis of biological samples (Martin FL *et al.*, 2010). Applications of ATR-FTIR studies in plants now include identification of plants from different growing regions (Rana *et al.*, 2008; Cirino de Carvalho *et al.*, 2016a; Traoré *et al.*, 2018); plant response to abiotic factors such as soil fertility (Euring *et al.*, 2012), heavy metals (McNear *et al.*, 2010; Usman *et al.*, 2019), water and temperature stress (Buitrago *et al.*, 2016), nutrient deficiency and uptake (Butler *et al.*, 2017, 2020); as well as monitoring plant health and development (Butler *et al.*, 2015a; Ord *et al.*, 2016; Skolik *et al.*, 2019b) and infection (Skolik *et al.*, 2019a).

ATR-FTIR works by using infrared light of wavenumbers 4000- 400 cm^{-1} (2.5–25 μm wavelengths) to induce atomic displacement and a change of dipole moment within the bonds of biomolecules (Baker *et al.*, 2014), which preferentially absorb light of wavenumbers 1800–900 cm^{-1}, a range known as the 'fingerprint region'. Spectral acquisition provides complex multivariate data and is therefore coupled with chemometrics. Subtle differences in sample composition can be analysed using mathematical techniques such as principal component analysis (PCA) and linear discriminant analysis (LDA), support vector machine (SVM), naïve bayes, and artificial neural networks (ANN) (Trevisan *et al.*, 2012; Li *et al.*, 2015; Chan & Kazarian, 2016; Strong *et al.*, 2017). This provides biochemical information about proteins, nucleic acids (DNA/RNA), lipids and carbohydrates (Luo *et al.*, 2021) because the absorption patterns are characteristic of the chemical composition, structure and function of the sample (Kazarian & Chan, 2013). Associated wavenumber shifts in the ATR-FTIR spectral fingerprint have been identified for biologically significant compounds of interest such as the herbicide glyphosate (Orcelli *et al.*, 2018), and the endogenous biological compounds; tannins (Falcão & Araújo, 2013), cutin (Heredia-Guerrero *et al.*, 2014), cutan (Heredia-Guerrero *et al.*, 2014), lignin (Butler *et al.*, 2015a), carotenoids (Butler *et al.*, 2015a), ellagic acid (Lopes & Sousa, 2018), and quercetin (Lopes & Sousa, 2018). Definitions for characteristic peak frequencies commonly seen in ATR-FTIR studies have been compiled in databases and are available in the literature from previous studies, for example see (Movasaghi *et al.*, 2008; Talari *et al.*, 2017).

The process from chemometric biomarker identification to physical biomolecular extraction is a developing area of spectroscopy with an ongoing research effort currently focused on optimising the quantification of biomolecular concentrations with the resultant spectra in biological extracts (Wagner *et al.*, 2010; Spalding *et al.*, 2018), consolidating the expanding database of key wavenumber changes and their associated molecular definitions (Talari *et al.*, 2017), and trialling new biological applications (Skolik *et al.*, 2019a,b; Butler *et al.*, 2020). Sample preparation such as freeze drying or grinding may influence the spectra acquired from vegetative plant material and the resultant classification success of subsequent chemometric

analysis. To ensure optimum spectral quality and molecular sensitivity, instrumental settings and sample preparation must be adjusted prior to spectral acquisition (Baker *et al.*, 2014; Chan & Kazarian, 2016; Pilling & Gardner, 2016; Morais *et al.*, 2019b).

To gain insight into how Japanese knotweed plants respond to and colonise varied environmental habitats, their spectral fingerprints were examined using the machine learning method, SVM. Variations in the obtained spectral fingerprint region were sufficiently distinct to differentiate between plants collected from sites in North-East England, North-West England, and Scotland with high accuracy. Key wavenumber changes indicated chemical differences between growing regions in several biomolecules: the cell wall component pectin, phenolic and antioxidant compounds (including carotenoids, tannins, ellagic acid, quercetin), lipids and fatty acids, the Amide I and II regions of proteins, and the nucleobases adenine and cytosine. To correlate spectral differences with environmental data, soil was collected from each site and climatic data collected by the United Kingdom Met Office was used (Met Office, 2018). Regional differences in the spectral fingerprint of *R. japonica* detected by ATR-FTIR spectroscopy and SVM could not be explained by pH, water content, organic matter content, plant available phosphorus, carbon to nitrogen ratio, maximum temperature, minimum temperature, air frost, days of rain, amount of rain, or the number of days of sunlight. Future studies will identify the mechanisms that underpin the regional differences in the spectral fingerprint of *R. japonica* and which contribute plasticity to *R. japonica* allowing it to thrive in such diverse habitats.

3.4 Methods

3.4.1 Field Sites

In late-summer 2018, plant and soil samples were collected from seven contrasting sites across the Northern United Kingdom where Japanese knotweed is known to be a problem (see Supplementary Table S3.1). Stands were identified according to their morphological features as described within the literature, see (Bailey *et al.*, 2009).

The data were then analysed by region (West Scotland [WS], North West England [NWE], North East England [NEE]) or site (Scotland [SOM, SAP, SLM, SRC], North East England [EDB], North West England [ESA, ESB]).

3.4.2 Sample Collection and Storage

Leaves were collected from three different canes per site of Japanese knotweed. On each cane, three leaves were collected from different positions on the plant, designated the labels 'New', 'Height', or 'Mature'. The relevant landowners were contacted for permissions to collect sample materials. The topmost newly unfurled leaf was collected, designated 'New'. 'Height' leaves were collected from 1 m above the soil surface from the main cane, to account for stands of different statures. 'Mature' leaves were the second leaf off the first stem branching off the main cane. Interestingly, the spectral profiles were affected by leaf position, data not shown. Therefore, to ensure that developmental stage was not a confounding factor when comparing sites all three leaf positions were included in the analysis. Leaves were dried at 37°C for one week and stored in a dry airtight container at room temperature before analysis using ATR-FTIR. Soil was collected from the base of each cane used in the leaf study, using a 25 cm long and 1 cm diameter bore (Professional Agricultural Analysis Group, 2013). The soil was passed through a 0.5 mm sieve and air dried before analysis (Emmett *et al.*, 2008).

3.4.3 ATR–FTIR spectroscopy

Dried leaves were analysed using a Tensor 27 FTIR spectrometer with a Helios ATR attachment (Bruker Optics Ltd, Coventry, UK). The sampling area, defined by the Internal Reflection Element (IRE), which was a diamond crystal, was 250 μm × 250 μm. Each leaf was placed on a slide with the section to be analysed facing upwards; the slide was then placed on a moving platform and moved upwards to ensure a good and consistent contact with the diamond crystal. Spectral resolution was 8 cm^{-1} with two-times zero-filling, giving a data-spacing of 4 cm^{-1} over the range 4000 to 400 cm^{-1}; 32 co-additions and a mirror velocity of 2.2 kHz were used for optimum signal to noise ratio. In total 1260 spectra were taken, ten spectra from each side of

sixty-three leaves (three leaves from each of three canes per seven sites). All spectra are available in the Supplementary Dataset. To minimise bias, an even spread of ten spectra were taken from each surface of the leaf, resulting in a total of twenty spectra per leaf. Approximately the same position on each leaf was located using a camera attachment, with five spectra taken either side of the central leaf vein.

3.4.4 Spectral data handling and analysis

All spectral information was converted from OPUS format to suitable files (.txt) before input to MATLAB (Mathworks, Natick, USA). Pre-processing of the acquired spectra is an essential step of all spectroscopic experiments and is used to improve the signal-to-noise ratio by correcting problems associated with random or systematic artefacts during spectral acquisition including different sample thickness (Butler *et al.*, 2018). Pre-processing and computational analysis of the data were performed using an in-house developed IRootLab toolbox (Martin FL *et al.*, 2010; Trevisan *et al.*, 2013) and the PLS Toolbox version 7.9.3 (Eigenvector Research, Inc., Manson, USA), according to standardised protocols for analysis of biochemical spectra (Morais *et al.*, 2019b, 2020). Spectra were cut at the biochemical fingerprint region (1800–900 cm^{-1}), Savitzky-Golay (SG) second differentiated and vector normalised. The number of points used in SG smoothing was nine. All data were mean-centred before multivariate analysis. To view the raw spectra see Supplementary Figure S3.1.

For the classification of groups, principal component analysis followed by linear discriminant analysis (PCA-LDA) and support vector machines (SVM) were used. PCA was used to reduce the original data into a few sets of variables called principal components (PCs). These variables, composed of 'scores' and 'loadings', are orthogonal to each other and account for most of the explained variance from the original data set. Scores were used to identify similarities and dissimilarities among the samples whilst loadings identify the weight contributed to the PCA model by each variable (Cirino de Carvalho *et al.*, 2016a). As PCA is an unsupervised technique, the category variables were not used for this dataset reduction. To perform a supervised classification model, the PCA scores were employed as input

variables for the discriminant algorithm, linear discriminant analysis (LDA; Morais & Lima, 2018). LDA created a linear classification rule between the classes based on a Mahalanobis distance. For exploratory data-analysis this study used the composite analysis PCA-LDA as well as the non-linear classifier, SVM (Morais *et al.*, 2017), which was additionally used for biomarker identification. The SVM classifier was used to find the classification hyperplane which provided the largest margin of separation between the data clusters. During model construction, the data were transformed into a different feature space by means of a kernel function that is responsible for the SVM classification ability. The most common kernel function, the radial basis function (RBF; (Cortes *et al.*, 1995)) was used here. Correlation between spectral differences and soil traits were assessed with PCA and partial least squares (PLS) regression. The relationship between spectra and climatic conditions (maximum temperature, minimum temperature, mean temperature, hours of sunshine, days of rainfall, days of rain ≥1 mm, and days of air frost) were also evaluated by PLS regression. Cross-validation is a model validation method used to evaluate the performance of the model when applied to an unknown sample. In this study, the number of components of PCA-LDA, PLS regression, and all SVM parameters were optimized by venetian blinds (10 data splits) cross-validation. The samples' spectra were randomly divided into a training set (70%, 882 spectra) and an external test set (30%, 378 spectra) to perform validation.

3.4.5 Soil Moisture Content and Organic Matter Level

For each biological replicate, two separate technical replicates were analysed. Approximately 6.5 g of air-dried soil was dried for 48 hours at 105 °C in an oven and the oven-dried mass was noted to calculate the percentage soil moisture content. The soil organic matter level was subsequently calculated by loss on ignition (LOI) (Hoogsteen *et al.*, 2015). The oven dried soil was placed in a furnace at 550 °C for six hours, and the final mass noted to calculate the percentage LOI.

3.4.6 C:N

In addition to calculating organic matter content by LOI, carbon and nitrogen levels were measured individually, and their values compared. Soil samples were dried overnight at 105 °C, before grinding for 2 min at 400 rpm. A microbalance was used to measure out 30 mg of dried-ground soil, which was then wrapped in tin foil boats for analysis in an Elemental Analyser (elementar vario EL III).

3.4.7 Plant available phosphate

Plant available phosphate was measured using the Olsen P method (Olsen *et al.*, 1954). This method uses bicarbonate as a chemical extractant to simulate the uptake of phosphorus by plants from the solution and exchange surfaces in soil in the form of phosphate. Three biological replicates were analysed per site. For each replicate, air-dried soil (2 g) was added to pH adjusted sodium bicarbonate ($NaHCO_3$, 0.5 M, 40 mL). This mixture was placed in an orbital shaker at 200 rpm for 30 mins, before filtration with Whatman 42 110 mm filter paper. Plant available phosphorus was measured using a SEAL AA3 AutoAnalyser with a SEAL XY-2 AutoSampler. The solutions from the first six samples were measured three times to check the consistency of the machine.

3.4.8 pH

Soil pH was measured based on the procedure created by Allen (Allen *et al.*, 1989). Soil (10 g fresh weight) was mixed with distilled mili-Q water (25 mL) for 30 min in an orbital shaker. The mixture was left in the fridge overnight to settle. The pH at the soil-water interface was then measured out using a Mettler Toledo SevenCompact™ pH meter.

3.4.9 Climatic Data

Met Office published climatic data were used for this study, for the regions West Scotland, North West England, and North East England, for the time period of the growing season 'Summer 2018' (Met Office, 2018). Maximum temperature,

minimum temperature, mean temperature, hours of sunshine (as a measure of photoperiod), days of rainfall, days of rain ≥1 mm, and days of air frost were considered.

3.4.10 Statistics

Statistical significance of measured soil parameters was calculated in R (R Core Team, 2014). A Shapiro–Wilk test indicated non-normal distribution therefore the data were analysed using the non-parametric Kruskal-Wallis test followed by a post hoc test using the criterium Fisher's least significant difference (LSD) within the package 'agricolae' (de Mendiburu, 2020) to determine where the difference lies between sites, signified by lowercase letters (Figure 3.6a-d). Alpha was set at 0.05. Within each graph, all bars which share letters are not significantly different from each other. Graphs were produced in RStudio using the package ggplot (Wickham, 2009).

3.5 Results and Discussion

3.5.1 Pre-processing of IR spectra in the fingerprint region reveals differences between regions

To capture and quantify the plant's response to its environment, ATR-FTIR spectroscopy was used. ATR-FTIR spectra were taken from both leaf surfaces of Japanese knotweed. Figure 3.1 shows the raw and pre-processed spectra, where the mean spectra at the fingerprint region are depicted by region (West Scotland [WS], North West England [NWE], North East England [NEE]; Figure 3.1a and b) and sites (Scotland [SOM, SAP, SLM, SRC], North East England [EDB], North West England [ESA, ESB]; Figure 3.1c and d) from which they were collected. There are subtle visual differences in the mean spectra from knotweeds collected in the WS, in particular at around 1648 cm^{-1} (Amide I band), 1586 cm^{-1} (Amide II band) and 1400 cm^{-1} (symmetric stretching of COO- of amino and fatty acids). Amide bands confer information on the secondary structure of proteins and are sensitive to protein conformation by differing degrees (Van De Weert *et al.*, 2001). The Amide I band is

the most sensitive, and originates from the C=O stretching vibration of the amide group coupled with the in-phase bending of the N-H bond and stretching of the C-N bond (Mallamace *et al.*, 2015; Yang H *et al.*, 2015). The Amide II peak is a combination of N-H in plane bending and C-N stretching. (Yang H *et al.*, 2015). The mean knotweed spectra collected from NWE also show subtle visual differences at around 1650 cm^{-1} and 1580 cm^{-1}.
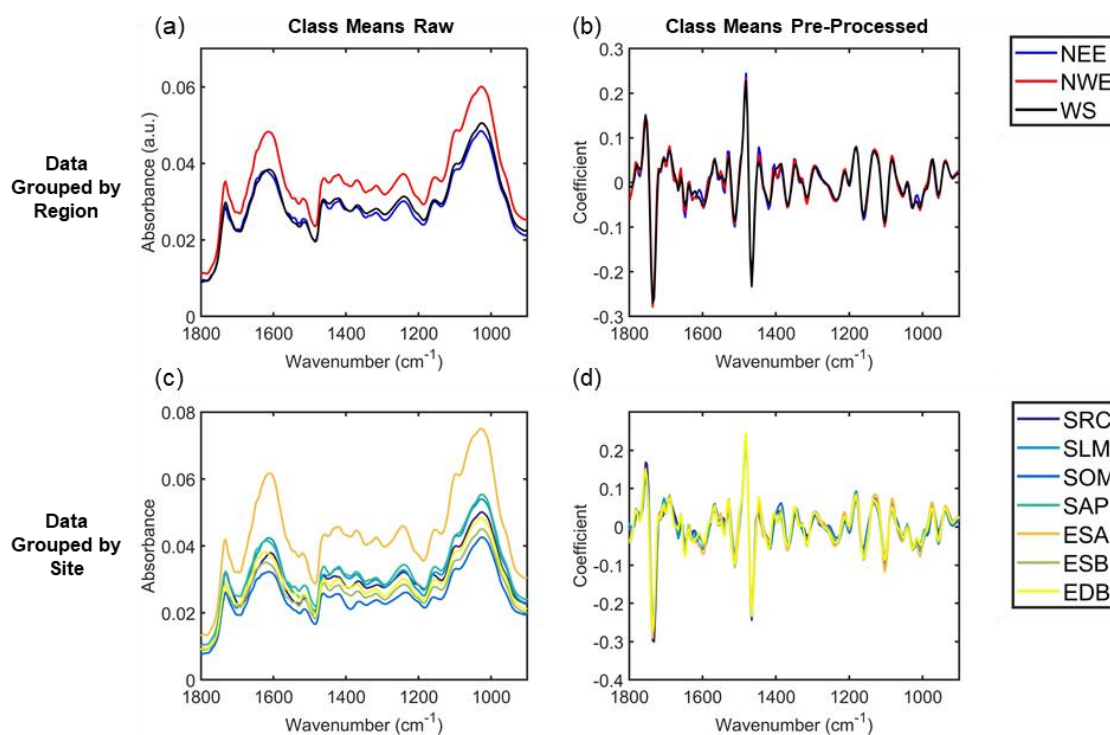
Figure 3.1: a) Class means raw and **b)** class means pre-processed (SG smoothed and vector normalised) IR spectra in the fingerprint region (1800-900 cm^{-1}) grouped by the different regions where knotweed samples were collected (NEE: North East England, NWE: North West England, WS: West Scotland); **c)** class means raw and **d)** class means pre-processed (SG smoothed and vector normalised) spectra in the fingerprint region (1800-900 cm^{-1}) grouped by the different sites where knotweed samples were collected (Scotland: SRC, SOM, SLM, SAP; North West England: ESA, ESB; North East England: EDB).

3.5.2 Knotweed from different regions were distinguishable on the basis of spectral profile

Changes in the spectral fingerprint region were sufficient to successfully identify sites from different geographical regions with high accuracy using SVM

chemometrics, indicating that Japanese knotweed from within each region share common properties that are distinct from those of plants from other regions. This resulted in them grouping together (Figure 3.2).
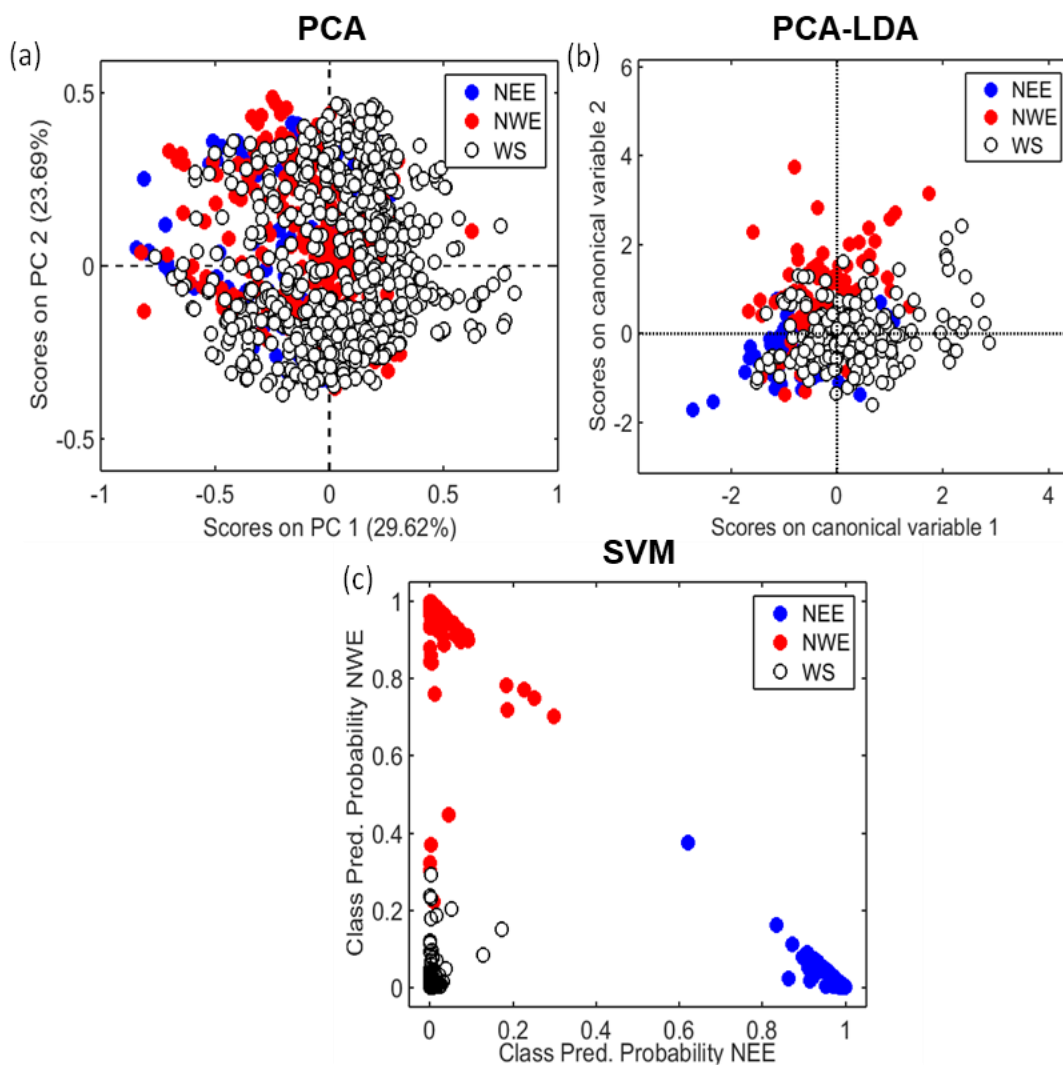
Figure 3.2: a) PCA scores, **b)** PCA-LDA canonical scores and **c)** SVM class predicted probability for the IR spectral dataset according to different regions where knotweed samples were collected (NEE: North East England, NWE: North West England, WS: West Scotland). Numbers inside parenthesis indicate the percentage of explained variance in each PC. Each spectral point in these scores plots represents a single spectral acquisition.

Unsupervised PCA was used to explore natural differences between knotweed samples collected from different regions (NEE, NWE and WS). No clear difference is observed in the scores on PC1 and PC2, indicating high similarities between the spectral profiles (Figure 3.2a). This is consistent with most Japanese knotweed in

the United Kingdom being a genetically uniform clone, described as a component of the "world's largest female" in biomass terms (Kidd & Shaw, 2000). Therefore, supervised methods of analysis, PCA-LDA and SVM, were applied to distinguish the samples based on their region. PCA-LDA was constructed using 10 PCs (93% explained variance) with a training performance of 68% accuracy (cross-validation accuracy of 67%). The predictive performance of PCA-LDA towards the external test set was relatively poor (96%, 90%, 50% specificity; 51%, 40%, 87% sensitivity; 69%, 62%, 70% precision for classes 1,2, and 3 respectively; Table 3.1), where the groups were found overlapping (Figure 3.2b). SVM (cost = 10, γ = 3.16, N_{SV}: 439) however, performed better in both training ($\sim$100 % accuracy) and test sets (95 %) (Table 3.1), indicating that knotweed plants can be differentiated by region based on their IR spectral profile. PCA-LDA assumes that the features of Japanese knotweed from each region are a realisation of a multivariate normal distribution, which is unlikely to be true. SVM on the other hand does not make this assumption, and given the right kernel function can uniformly approximate the true boundary between the classes of knotweed from each region by the universal approximation theorem (Hammer & Gersmann, 2003).

Table 3.1: Quality parameters for spectral classification based on different regions. The predictive performance of PCA-LDA towards the external test set was relatively poor. However, SVM performed well in training and test sets, indicating that knotweed leaf samples can be differentiated by region based on their IR spectral profile. The accuracy represents the total number of samples correctly classified considering true and false negatives; the sensitivity represents the proportion of positives that are correctly classified; and the specificity represents the proportion of negatives that are correctly identified.

Algorithm	Class	Accuracy	Sensitivity	Specificity
PCA-LDA	North East England	63%	30%	96%
	North West England	62%	37%	87%
	West Scotland	67%	88%	47%
SVM	North East England	100%	100%	100%
	North West England	98%	95%	100%

	West Scotland	98%	100%	97%
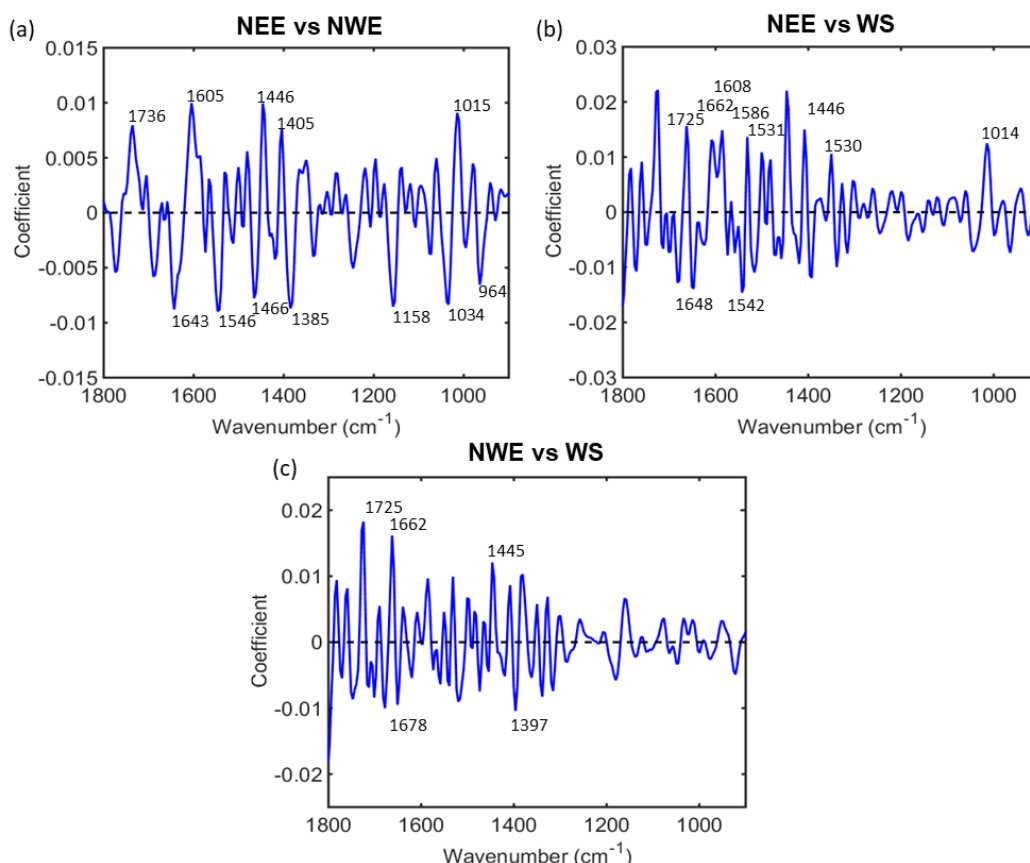
Figure 3.3: Difference between mean support vector spectra for **a)** North East England (+ coefficients) and North West England (- coefficients), **b)** North East England (+ coefficients) and West Scotland (- coefficients), **c)** North West England (+ coefficients) and West Scotland (- coefficients). The main wavenumbers responsible for class differentiation between the three regions are labelled, and have been used to identify spectral markers.

Examination of the difference-between-mean support vectors spectra found by SVM revealed spectral markers indicative of marked chemical differences between Japanese knotweed from different regions. Figure 3.3 shows the main wavenumbers responsible for class differentiation between the three regions which, through comparison with the literature, have been used to identify spectral biomarkers (Table 3.2). This indicates that there are marked chemical differences between

Japanese knotweed from Scotland, North West England and North East England. Differences between regions were identified at wavenumbers 1736, 1643, 1605, 1546, 1466, 1446, 1405, 1385, 1158, 1034, 1015, 964 cm^{-1} between NEE and NWE; 1725, 1662, 1648, 1608, 1586, 1542, 1531, 1446, 1530, 1014 cm^{-1} between NEE and WS; 1725, 1678, 1662, 1445, 1397 between NWE and WS cm^{-1} (Table 3.2). These are indicative of differences in the cell wall component pectin, phenolic and antioxidant compounds, lipids and fatty acids, the Amide I and II regions of proteins, and the nucleobases adenine and cytosine between knotweed from different regions.

Table 3.2: Biomolecular markers for class differentiation between the three regions (Scotland, North West England and North East England) and each site (SOM, SAP, SRC, SLM, ESA, ESB and EDB). Spectral markers were found by investigating the difference-between-mean support vectors spectra found by SVM, and linked to the biomolecules associated with each wavenumber from published literature.

Comparison	Wavenumber/ cm^{-1}	Tentative Molecular Assignment	Reference
NEE and NW	1736	C=O stretching [lipids]	(Movasaghi *et al.*, 2008)
	1643	C=O stretching [Amide I]	(Movasaghi *et al.*, 2008)
	1605	v_{as} (COO$^-$) [polysaccharides, pectin]	(Movasaghi *et al.*, 2008)
	1546	Amide II: [protein N–H bending, C–N stretching], α-helical structure	(Simsek Ozek *et al.*, 2010)
	1466	CH$_2$ bending in lipid	(Butler *et al.*, 2015a)
	1446	aromatic ring stretch vibrations, tannins	(Falcão & Araújo, 2013)
	1405	CH$_3$ asymmetric deformation	(Movasaghi *et al.*, 2008)
	1385	Ring stretching vibrations mixed strongly with CH in-plane bending	(Movasaghi *et al.*, 2008)

	1158	νC-O of proteins and carbohydrates	(Movasaghi *et al.*, 2008)
	1034	C-O stretch, tannins	(Falcão & Araújo, 2013)
	1015	ν(CO), ν(CC), δ(OCH), ring in pectin	(Butler *et al.*, 2015a)
	964	C-O deoxyribose, C-C	(Movasaghi *et al.*, 2008)
NEE and WS	1725	C=O stretching band mode of the fatty acid ester	(Movasaghi *et al.*, 2008)
	1662	Amide I, or fatty acid esters	(Movasaghi *et al.*, 2008)
	1648	Amide I	(Movasaghi *et al.*, 2008)
	1608	aromatic ring stretch vibrations, tannins	(Falcão & Araújo, 2013)
	1586	Amide II	(Movasaghi *et al.*, 2008)
	1542	Amide II	(Chen *et al.*, 2017)
	1531	Amide II	(Kaszowska *et al.*, 2013)
	1446	aromatic ring stretch vibrations, tannins	(Falcão & Araújo, 2013)
	1530	C=N adenine, cytosine	(Movasaghi *et al.*, 2008)
	1014	phosphodiester stretching bands [symmetrical and asymmetrical]	(Movasaghi *et al.*, 2008)
NWE and WS	1725	C=O stretching band mode of the fatty acid ester	(Movasaghi *et al.*, 2008)
	1678	Stretching C=O vibrations that are H-bonded [changes	(Movasaghi *et al.*, 2008)

		in the C=O stretching vibrations could be connected with destruction of old H-bonds and creation of the new ones]	
	1662	Amide I, or fatty acid esters	(Movasaghi *et al.*, 2008)
	1445	lipids	(Lopes & Sousa, 2018)
	1397	CH$_3$ symmetric deformation	(Movasaghi *et al.*, 2008)
SRC and others	1748	C=O stretching vibration of alkyl ester, pectin	(Heredia-Guerrero *et al.*, 2014)
	1728	ν (C=O) ester, cutin	(Heredia-Guerrero *et al.*, 2014)
	1678	Stretching C=O vibrations that are H-bonded [changes in the C55O stretching vibrations could be connected with destruction of old H-bonds and creation of the new ones]	(Movasaghi *et al.*, 2008)
	1651	phenolic compounds/ cutan [aromatic and C=C functional groups]	(Heredia-Guerrero *et al.*, 2014)
	1608	aromatic ring stretch vibrations, tannins	(Falcão & Araújo, 2013)
	1542	Amide II	(Chen *et al.*, 2017)
	1455	C-O-H	(Movasaghi *et al.*, 2008)
	1443	δ (CH$_2$) [lipids, fatty acids], or δ (CH) [polysaccharides, pectin]	(Movasaghi *et al.*, 2008)

SLM and others	1755	lipid	(Ord *et al.*, 2016)
	1735	C=O stretching, the phenolic compound ellagic acid/ the secondary metabolite quercetin	(Lopes & Sousa, 2018)
	1512	ν (C-C) aromatic (conjugated with C=C phenolic compounds	(Heredia-Guerrero *et al.*, 2014)
	1481	symmetric deformation NH_2^+, glyphosate[X]	(Orcelli *et al.*, 2018)
	1466	CH_2 bending in lipid	(Butler *et al.*, 2015a)
SOM and others	1755	lipid	(Ord *et al.*, 2016)
	1736	lipids	(Lopes & Sousa, 2018)
	1481	symmetric deformation NH_2^+, glyphosate[X]	(Orcelli *et al.*, 2018)
	1466	CH_2 bending in lipid	(Butler *et al.*, 2015a)
	1161	carbohydrate; stretching vibrations of hydrogen-bonding C–OH groups (found in serine, threonine and tyrosine residues of cellular proteins); cellulose	(Ord *et al.*, 2016)
	1103	ν(C–O–C) in ester	(Butler *et al.*, 2015a)
	1755	lipid	(Ord *et al.*, 2016)
	1736	lipid	(Lopes & Sousa, 2018)
	1481	symmetric deformation NH_2^+, glyphosate[X]	(Orcelli *et al.*, 2018)

SAP and others	1466	CH_2 bending in lipid	(Butler *et al.*, 2015a)
	1103	ν(C–O–C) in ester	(Butler *et al.*, 2015a)
ESA and others	1755	lipid	(Ord *et al.*, 2016)
	1732	lipid; fatty acid esters; hemicellulose	(Ord *et al.*, 2016)
	1647	amide I; pectin	(Ord *et al.*, 2016)
	1512	ν(C=C) in lignin, carotenoid or protein	(Butler *et al.*, 2015a)
	1481	symmetric deformation NH_2^+, glyphosate[X]	(Orcelli *et al.*, 2018)
	1466	aromatic ring stretch vibrations, tannins	(Falcão & Araújo, 2013)
ESB and others	1755	lipid	(Ord *et al.*, 2016)
	1736	C=O stretching [lipids]	(Movasaghi *et al.*, 2008)
	1481	symmetric deformation NH_2^+, glyphosate[X]	(Orcelli *et al.*, 2018)
	1466	CH_2 bending in lipid, or aromatic ring stretch vibrations, tannins	(Butler *et al.*, 2015a), (Falcão & Araújo, 2013)
EDB and others	1728	ν (C=O) ester, cutin	(Heredia-Guerrero *et al.*, 2014)
	1446	aromatic ring stretch vibrations, tannins	(Falcão & Araújo, 2013)
	1408	CH_3 deformation, ν_s(COO$^-$) in pectin	(Butler *et al.*, 2015a)

Chapter 3: Regional differences in clonal Japanese knotweed revealed by chemometrics-linked Attenuated Total Reflection Fourier-Transform Infrared Spectroscopy

Two of the key peaks for differentiation between NEE and NWE regions, 1015 and 1605 cm^{-1}, were linked to pectin. The rubber-band corrected spectra indicated that the pectin concentration was lowest for NEE at both peaks compared with other regions. A horizontal shift in the SG differentiated spectra for NEE at both peaks indicate an altered pectin structure for North East samples. This could be of interest as manipulation of pectin synthesis is often studied, due to the compound's importance in food products and biofuel production (Johnson *et al.*, 2018). Two peaks corresponding to tannins, 1034 cm^{-1} and 1608 cm^{-1}, were responsible for the differences between regions, NEE vs NWE and NEE vs WS respectively.

The concentration of tannins is higher in WS and NWE than in NEE, indicated by the higher absorbance levels in the rubber-band corrected spectra. A biochemical structural change in tannins at NEE is exhibited, indicated by a horizontal shift at both 1034 cm^{-1} and 1608 cm^{-1} in the SG differentiated spectra. This would suggest that the reduced tannin levels of Japanese knotweed at the site in North East England would make it a more favourable host for herbivorous biocontrol agents (Singh *et al.*, 2021).

3.5.3 ATR–FTIR spectral changes allowed discrimination on a site-by-site basis

In addition to investigating interregional variations between knotweed from different geographical areas, each site colonised by Japanese knotweed was further investigated individually to see how varying environmental conditions in each habitat affected the plant's spectral fingerprint. Due to the highly satisfactory classification performance of SVM, site differences were also investigated using this method. To view the PCA scores plots see Supplementary Figure S3.5. SVM was trained using cost = 10, γ = 3.16; N_{SV}: 675, generating accuracies at 100% for training and 94% for class validation. The predictive capability of SVM was tested in an external test set where accuracies were found ranging from 97% to 100%, sensitivities from 94% to 100%, and specificities from 99% to 100% (Table 3.3), indicating that the knotweed samples can be differentiated by the site at which they

were collected. Figure 3.4 shows the SVM class predicted probability for the IR spectral dataset on a site-by-site basis.

Using the same method described above, the difference between mean support vector spectra of a site was compared with the other six sites, for each site in turn. These comparisons were used to identify the key wavenumbers responsible for the differences between sites (Figure 3.5). The spectral markers associated with site difference were at wavenumbers: 1748, 1728, 1678, 1651, 1542, 1455, 1443 cm^{-1} for SRC and others; 1755, 1735, 1512, 1481, 1466 cm^{-1} for SLM and others; 1755, 1736, 1481, 1466, 1161, 1103 cm^{-1} for SOM and others; 1755, 1736, 1481, 1466, 1103 cm^{-1} for SAP and others; 1755, 1732, 1647, 1512, 1481, 1466 cm^{-1} for ESA and others; 1755, 1736, 1481, 1466 cm^{-1} for ESB and others; 1728, 1446, 1408 cm^{-1} for EDB and others, (Table 3.2). The changes in spectral profile were associated with the prominent cuticle components, cutan and cutin, and the cell wall component, pectin. While cuticle components were key for site-by-site classification, these were not used for regional distinction. Site-by-site classification required spectral biomarkers strongly associated with phenolic antioxidant compounds, including carotenoids, tannins, ellagic acid, quercetin. Other key identifiers included lipids, fatty acids, and the Amide I and II vibrational modes of proteins.

Table 3.3: Quality parameters for spectral classification based on different sites. Separation by PCA-LDA was relatively poor. However, SVM performed much better, indicating that knotweed leaf samples can be differentiated by the site at which they were collected using this method.

Algorithm	Class	Accuracy	Sensitivity	Specificity
PCA-LDA (10 PCs, 91% explained variance)	SRC	85%	80%	90%
	SLM	70%	57%	83%
	SOM	60%	24%	97%
	SAP	62%	35%	89%
	ESA	73%	52%	95%
	ESB	56%	20%	92%
	EDB	65%	43%	88%
SVM	SRC	100%	100%	100%
	SLM	100%	100%	100%
	SOM	99%	98%	100%
	SAP	100%	100%	100%
	ESA	100%	100%	100%
	ESB	97%	94%	100%
	EDB	99%	100%	99%
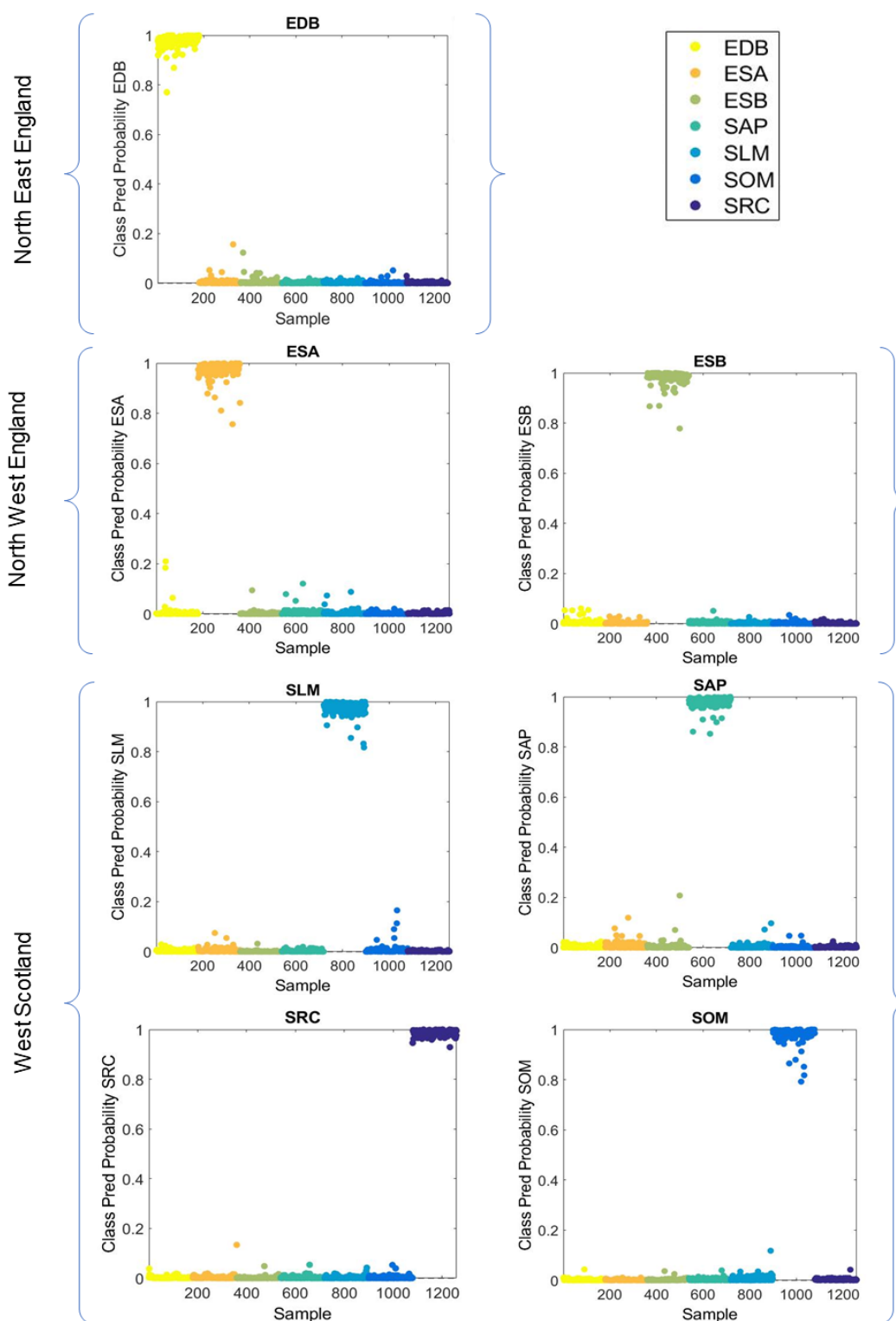
Figure 3.4: SVM class predicted probability for the IR spectral dataset according to different sites where knotweed samples were collected (Scotland: SRC, SOM, SLM, SAP; North West England: ESA, ESB; North East England: EDB). The clear separation indicates that the knotweed samples can be differentiated by the site at which they were collected.
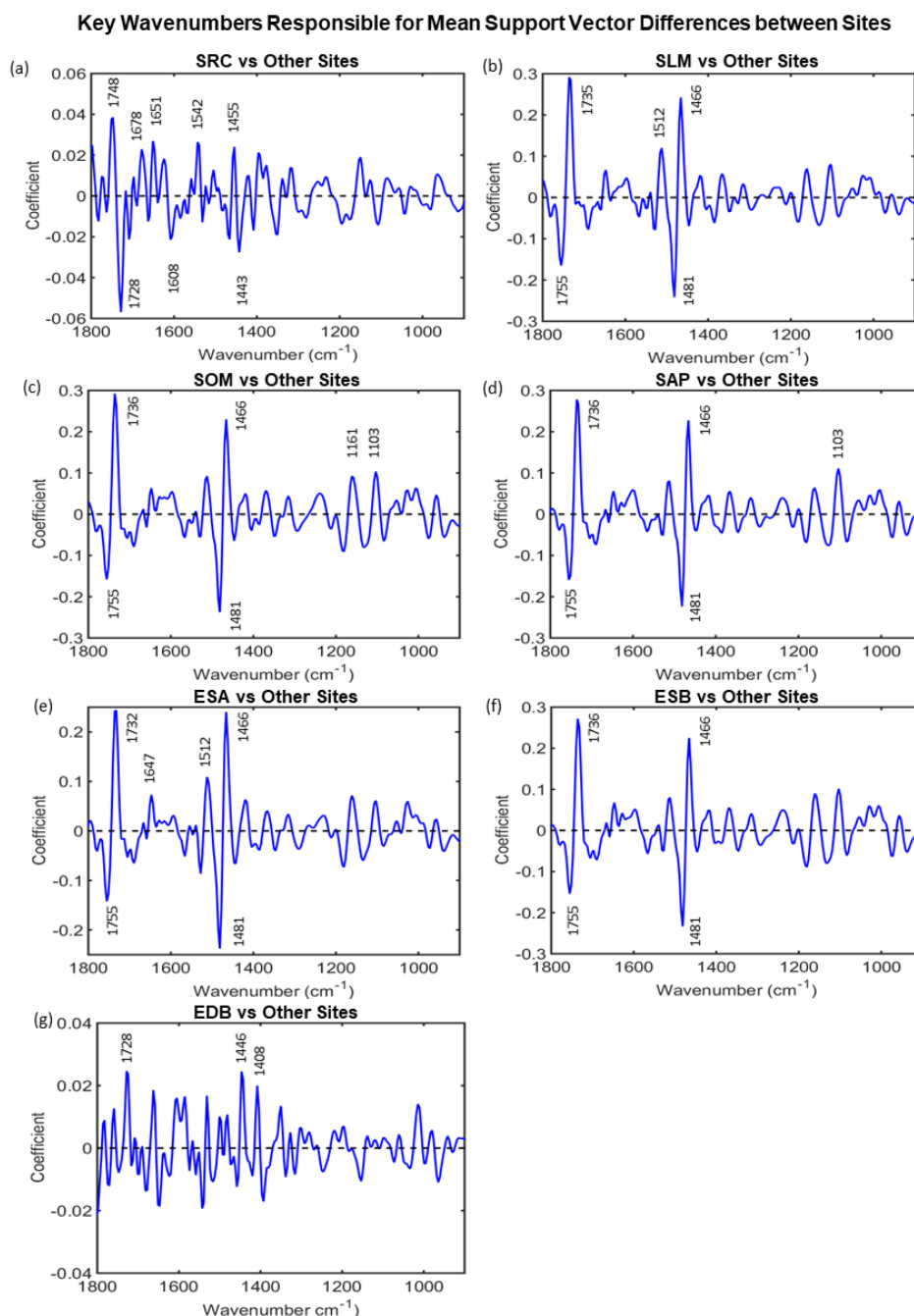
Figure 3.5: Difference between mean support vector spectra for **a)** SRC (+ coefficients) and others (- coefficients), **b)** SLM (+ coefficients) and others (- coefficients), **c)** SOM (+ coefficients) and others (- coefficients), **d)** SAP (+ coefficients) and others (- coefficients), **e)** ESA (+ coefficients) and others (- coefficients), **f)** ESB (+ coefficients) and others (- coefficients), and **g)** EDB (+ coefficients) and others (- coefficients). These comparisons can be used to identify the key wavenumbers responsible for the differences between sites, which have been labelled above, and can be used to find spectral biomarkers.

Five of the seven sites showed traces of treatment with glyphosate: EDB, ESA, SAP, SOM and SLM. Three of the sites, ESB, SOM, and SAP, appeared dead the following summer, indicating that they had received an effective dose of glyphosate the previous year. Upon collection at site SLM, the plants were noted to appear in good health, and despite showing traces of herbicide in ATR-FTIR spectroscopy studies, plants at the site were still alive and well the following year. The herbicide history of NEE and WS sites is unknown, however, ESA and ESB are known to have been treated the year prior to sample collection. Repetitive herbicide use is commonplace in Japanese knotweed removal, with a minimum of three years glyphosate treatment being the most common chemical treatment method. Jones D *et al.* (2018), found that repeated glyphosate use was the most effective control method. However, sub-lethal herbicide doses are known to have a hormetic effect on plants (Belz & Duke, 2014), with the potential to increase vigour of target plants and arm them against other stresses. Therefore, prior sub-lethal herbicide treatment may have resulted in altered gene expression being reflected in the spectral biomarkers in certain sites.

A family of membrane-stabilising plant pigments called carotenoids were flagged up as a biomarker for site ESA. Carotenoids participate in light-harvesting and are essential for photoprotection against excess light (Ruiz-Sola & Rodríguez-Concepción, 2012). It is therefore surprising that they would appear as a key identifier of a shaded woodland site like ESA. However, biomarkers for both glyphosate (at 1481 cm^{-1}), and carotenoid (at 1512 cm^{-1}), presented as differences between ESA and the other sites. Carotenoid levels are known to be affected by herbicide application, including a temporary increase in response to reactive oxygen species, followed by a decrease due to reduced biosynthesis (Stasiak *et al.*, 1992; Bashtanova *et al.*, 2009; Gomes *et al.*, 2017).

The presence of the biomarker for quercetin (at 1735 cm^{-1}) is consistent with current knowledge of this plant species (Fan *et al.*, 2009; Peng *et al.*, 2013). Quercetin is a flavonoid with antioxidant properties which also acts as a naturally occurring auxin polar transport inhibitor (Fischer *et al.*, 1997). The marker at 1608

cm^{-1} has been linked to tannins in the quebracho tree, and the marker at 1446 cm^{-1} is present in several tannins (Falcão & Araújo, 2013).

3.5.4 Soil analysis indicates the environmental diversity between sites

As a clonal species which spreads by physical disturbance of rhizome, crown and stem propagules, Japanese knotweed occupies habitats where disturbance occurs, such as near roads, railways and water courses (Hollingsworth & Bailey, 2000b). Although all the sites from which Japanese knotweed was collected share these characteristics, the environmental conditions and habitats colonised by this species were variable within each region. For example, within Scotland there were two riverside sites with adjacent forest: a brownfield site repurposed as a park which had historically been farmed, built on, mined, quarried, dug for clay and used as a landfill site and a railway siding; in addition to an urban site adjacent to a railway line, road, and public footpath (see Supplementary Table S3.1 for site descriptions, Google Maps coordinates, and photographs).

This environmental diversity between sites is reflected in differences in measured soil characteristics; considered individually in Figure 3.6, and in combination through the multivariate analysis shown in Figure 3.7. The two related graphs shown in Figure 3.7 are the PCA scores (7a) and PCA loadings (7b) plots for the relationship between the chemical soil parameters measured at each site. Overall, Figure 3.7a shows a segregation pattern which indicates site-by-site differences in measured soil traits. The three points of each category in the scores plot represent the three biological replicates collected from each site, their proximity to one another is an indication of variation within the site. The distance of scores away from the origin in Figure 3.7a show how different the sites were from one another, and the explanations for their separation can be found by looking at the loadings in Figure 3.7b. Scores which are close to a loading point have a higher value of this parameter. The trajectory of the samples can be thought of as a modulus, because a negative score or loading can still mean a higher value in terms of the soil trait. For example, the North East samples, EDB, which are closer to the C:N loading area had a higher C:N ratio than the others despite being in the negative region of the graph.
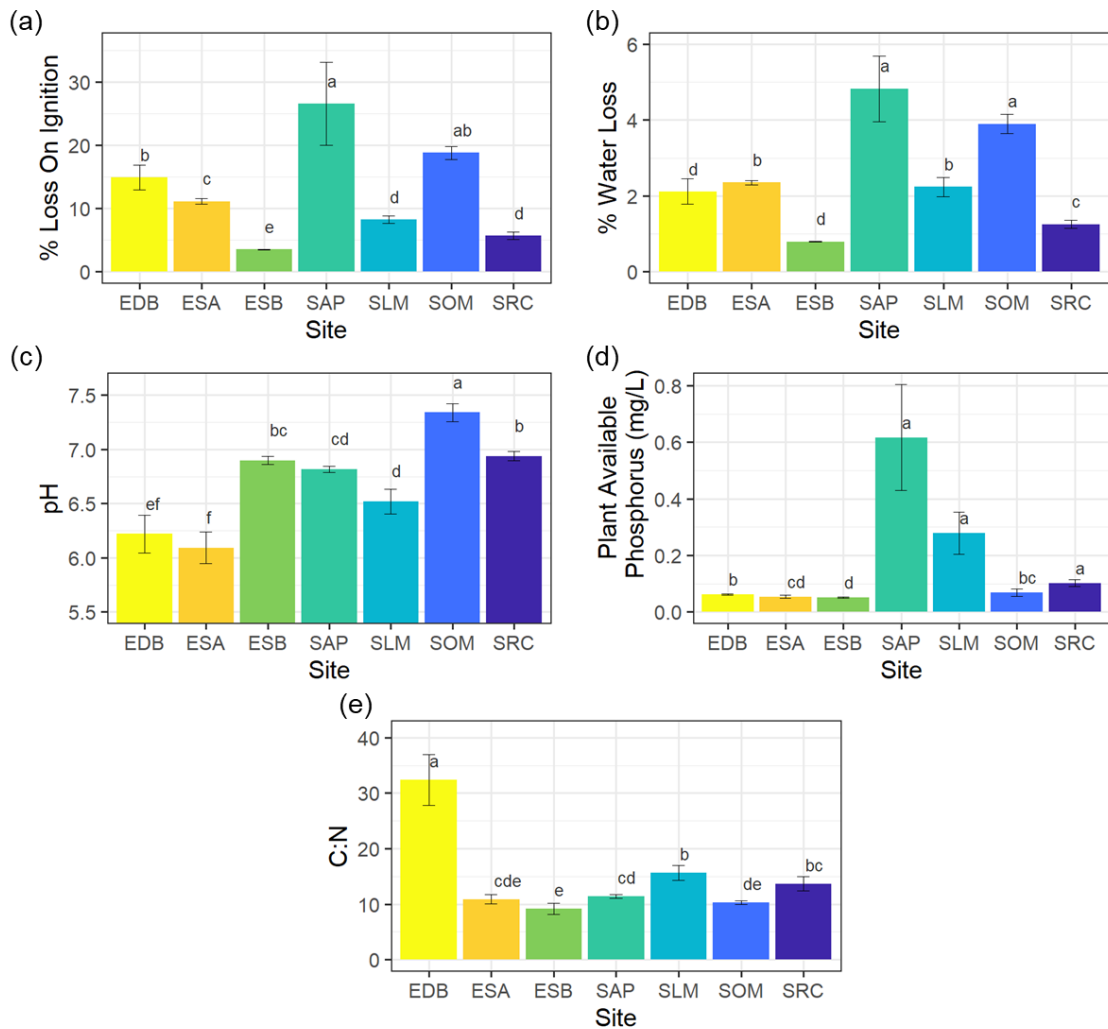
Figure 3.6: Soil parameters for each site, with error bars showing standard error; **a)** percentage mass lost on ignition (LOI), **b)** % water loss, **c)** pH, **d)** plant available phosphorus, **e)** carbon to nitrogen ratio (C:N). Statistical significance was calculated using a Kruskal-Wallis followed by a post hoc test using the criterium Fisher's least significant difference (LSD) to determine where the difference lies, signified by lowercase letters above the bars. Within each graph, all bars which share letters are not significantly different from each other. Data are mean +/- standard errors. pH, n=9; C:N n=9; % water loss mean, n=6; % loss on ignition mean, n=6; plant available phosphorus mean, n=3 except for EDB and ESB where n=9.
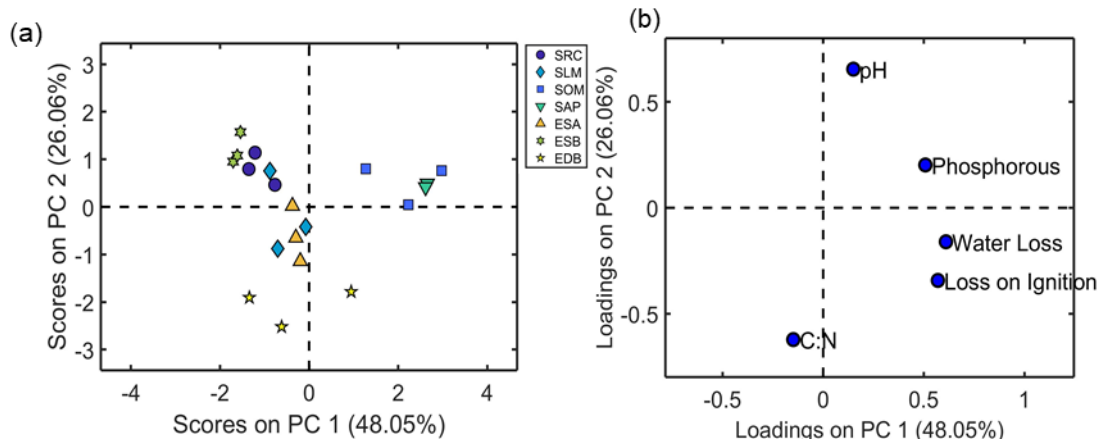
Figure 3.7: a) PCA scores and **b)** loadings for soil data (abbreviations define sites where samples were collected, Scotland: SRC, SOM, SLM, SAP; North West England: ESA, ESB; North East England: EDB). The North East England soil sample, EDB, has a high C:N ratio, and a lower pH than the other samples. EDB soil was found to be naturally different from the others on the PCA scores plot, with a greater separation in the Y-plane (PC2). ESB has a higher C:N than the other sites. SOM and SAP have high phosphorus, water loss, and LOI (organic carbon) compared with SRC, SLM, ESA, ESB, and EDB. SLM was a mixed sample, sharing similar soil traits with ESA and SRC. Note: An extreme sample was removed. SAP3 was a non-homologous urban environment and one of the three soil samples was an outlier. The bar graphs in Figure 3.6a and b show high standard deviation for loss on ignition and percentage water loss for SAP, due to this sample.

EDB soil was found to be naturally different from that of the other sites on the PCA scores plot, with a greater separation in the Y-plane (PC2). Although both Figures 3.6 and 3.7 agree that EDB had a higher C:N ratio compared with the other sites (nearly double), when considered individually in Figure 3.6e it is clear from the error bars that this high level varied between samples, whereas the lower ratios for SAP and SOM were more consistent between samples. Nitrogen is vital for the synthesis of chlorophyll, nucleic acids and proteins (Miller, 2014) and the C:N ratio is a good measure of decomposition rates, with higher C:N ratios generally leading to longer decomposition rates. Figure 3.7 highlights that ESB had a lower C:N than the other sites, suggesting that this riverside site in full sun had the fastest decomposition rate of all the sites.

In addition to a high C:N ratio, Figure 3.7 shows that EBD also had a lower pH than the other samples. This is consistent with Soilscapes' definition of this area as having 'slowly permeable seasonally wet acid loamy and clayey soils', with low fertility and

impeded drainage (UK Soil Observatory, 2020). Soil pH affects nutrient availability and in clay soils pH may alter the structure of the soil; pH in the range 5.5 to 6.5 is normally optimum for allowing plants access to most nutrients (Bolan *et al.*, 2004). The pH was higher in SOM compared with all other sites (Figure 3.6c). The pH of SRC and ESB were significantly higher than SLM, EDB and ESA, with SRC additionally higher than SAP (Figure 3.6c). SAP has a mid-range pH which is similar to SLM but higher than EDB and ESA. EDB and ESA have the lowest pHs, although there was variability between samples (Figure 3.6c). In general, phosphorus availability decreases with increasing pH, however the multivariate analysis placed these loadings within the same quadrant, connecting instead a high C:N ratio with a low pH. Low pH soil reduces the growth of the bacteria and fungi responsible for the breakdown of organic matter and nutrient cycling. As a species which produces abundant recalcitrant polyphenol-rich leaf litter (Tamura & Tharayil, 2014; Lavoie, 2017), Japanese knotweed may therefore be expected to have a more detrimental effect on the nutrient cycling in acidic soil sites such as EDB and ESA. However, only EDB exhibited this carbon accumulation.

Overall, SOM had the highest pH, organic matter, and water content, with little variability between samples. In Figure 3.7, PC1 separates SOM and SAP whilst PC2 separates the other samples. The loadings show that SOM and SAP have high phosphorus, water loss, and LOI (organic carbon) compared with SRC, SLM, ESA, ESB, and EDB. Although collected in a similar region to the mineral gleys of SLM and SRC, SOM and SAP were both urban sites with anthropic soils. Soil organic matter can improve the soil's water holding capacity, enhances aggregate stability, increases cation exchange capacity, acts as a nutrient source and alters the soil microbiome (Murphy, 2015). The site with the lowest organic matter content, ESB, is significantly different from the other sites (Figure 3.6a). Sites SLM and SRC have lower organic matter levels than all sites except ESB (Figure 3.6a). ESA had a mid-range organic matter content which is different from the other sites (Figure 3.6a). SOM has a similar organic matter content to both EDB and SAP. The site with the highest organic matter content, SAP, had a high level of variability between the samples (Figure 3.6a).

The water content of EDB and ESB was lower than the other sites. SRC soil contained the next highest level of water, followed by ESA and SLM, followed by SAP and SOM which have the highest (Figure 3.6b). Overall ESB, a clay-to-sandy loam soil (UK Soil Observatory, 2020) collected from a site on the River Wyre had both low organic matter and water contents with minimal variation between samples. The PCA in Figure 3.7 shows that SLM was a mixed sample, sharing similar soil traits with ESA and SRC. This grouping may be surprising as site ESB had a more comparable riverside location to SLM and SRC. Despite their differing proximity to water, the soils of ESA, SRC, and SLM were all a similar cohesive clay-like consistency whereas the soil of riverside site ESB was much sandier in texture. Additionally, the surrounding vegetation was similarly well-populated by trees at sites ESA, SRC and SLM, presumably providing similar levels of cover and nutrient deposition from leaf litter.

Phosphorus (P) is essential for both ATP and nucleic acids formation (Miller, 2014) and often limits plant productivity because of its low mobility in soil. ESA and ESB had similar levels of plant available phosphorus, whilst ESB was lower than the other sites, ESA was similar to SOM, which in turn is similar to EDB (Figure 3.6d). SAP, SLM and SRC had significantly higher plant available phosphorus values compared with the other sites (Figure 3.6d), although the variation between samples in SAP was much greater. One possible reason for the enhanced phosphorus content of soil from site SAP is its proximity to an urban footpath likely to contain dog excrement, which would support the high variability within the site of plant available phosphorus. When all five soil parameters are compared at once using PCA multivariate analysis to detect outlying data (Mejia *et al.*, 2017), one of the SAP samples appears as anomalous in terms of high values for organic matter and water content (Supplementary Figure S3.3). This is consistent with SAP being a non-homologous urban environment located between a footpath and a railway embankment with a nearby road, thereby being prone to contamination with organic matter, and partial shade. The spectral data for SAP3 concur with the anomalous nature of this sample, as does the variability shown in Figures 3.6a, 3.6b, and 3.6d. Consequently, this outlier was removed from subsequent multivariant

analyses (Figure 3.7a). This outlier effect supports that the bar graphs for loss on ignition and percentage water loss for SAP display a high standard deviation due to this sample.

3.5.5 Soil and climatic conditions do not explain the spectral differences between site

Chemometric analysis of ATR-FTIR spectral data shows that Japanese knotweed can be identified on a site-by site basis, but that, despite this intra-regional variation, enough commonality is present within a whole region to allow discrimination of samples on a regional-basis. This raises the question of what environmental stimulus causes the variation in functional groups detected by ATR-FTIR spectroscopy? In order to investigate this, spectral data and soil data were considered together (Supplementary Figure S3.2 for PCA scores of spectral and soil data combined). The main climatic difference between the three regions was higher rainfall in West Scotland, see Supplementary Figure S3.4. Partial least squares (PLS) regression of the soil characteristic data climatic data (maximum temperature, minimum temperature, mean temperature, hours of sunshine, days of rainfall, days of rain ≥1 mm, and days of air frost) from the Met Office records for Summer 2018 (Met Office, 2018) revealed only a minimal correlation between spectral differences and soil traits (Supplementary Figures S3.6a-e) or climatic conditions (Supplementary Figures S3.7a-e). Therefore, the chemical differences between Japanese knotweed from NWE, NEE and WS could not be explained by the soil or climatic parameters measured in this study.

Allelopathy, releasing chemicals into the soil to alter its characteristics (Murrell *et al.*, 2011) for example to alter nutrient availability in the rhizosphere, could provide an explanation for the absence of any correlation. Root-mediated localised acidification of the rhizosphere and soil microbes can have a marked effect on plant available phosphorus. Allelopathic plants, such as Japanese knotweed and other weed species, are particularly good at altering their soil environment (Murrell *et al.*, 2011; Parepa *et al.*, 2012; Parepa & Bossdorf, 2016). It is standard practice to collect soil samples at one of two depths: 7.5 cm for grassland, or 25 cm depth for

agricultural fields (Professional Agricultural Analysis Group, 2013). A 25 cm depth was chosen for this study, although the extensive rhizome system of Japanese knotweed spans much deeper than this (often up to 2 m below the surface (Smith JMD *et al.*, 2007). Therefore, the measured depth of topsoil may not have been representative of the soil environment experienced by a well-established site of Japanese knotweed, with an interconnected rhizome system. Furthermore, Japanese knotweed demonstrates stronger allelopathic effects in artificial soils which have greater aeration, water retention, permeability, and nitrogen content and lower bulk density (Parepa & Bossdorf, 2016), adding a greater level of complexity to the interaction of these plants with their habitats. Allelopathy in Japanese knotweed is thought to be subject to resource allocation and is increased when nutrient supply is high (Parepa & Bossdorf, 2016), inhibiting the growth rather than the germination of native species (Parepa *et al.*, 2012). This plant species is a known opportunist, which is able to take advantage of the fluctuation in resources associated with riparian areas (Parepa *et al.*, 2013) using its superior nitrogen-use efficiency compared with native species (Parepa *et al.*, 2019). Plasticity is also thought to be resource dependent, with invasive species showing greater phenotypic plasticity than native plants if resources are plentiful (Davidson *et al.*, 2011).

Additionally, rapidly growing plant species often promote nutrient cycling displaying exploitative traits such as high tissue nitrogen content and specific leaf area, because these plants input high-quality resources to the soil (Orwin *et al.*, 2010; de Vries *et al.*, 2012). Japanese knotweed can increase nutrient cycling, with the greatest impact on sites which if uninvaded would have low nutrient levels. Sites occupied by Japanese knotweed can increase topsoil concentrations of exchangeable nutrients when compared with nearby uninvaded sites; Cu: +45%, K: +34%, Mg: +49%, Mn: +61%, P: +44%, Zn: +75% (Dassonville *et al.*, 2007) possibly due to the deep extensive rhizome system allowing extraction of nutrients which are not easily accessible to other vascular plants (Groeneveld *et al.*, 2014). In addition to altering the nutritional value of its soil environment, Japanese knotweed can also capitalise on available resources when they arise. In fact it performs best when nutrients come in waves rather than at a consistent level (Parepa *et al.*, 2013). This

adaptability and the fluctuating nutrition of water-side sites may explain the minimal correlation between the soil measurements and the ATR-FTIR spectral data derived from vegetative tissues.

3.5.6 Differences between regions may result from phenotypic plasticity

ATR-FTIR spectroscopy with subsequent chemometric analysis has proven effective at differentiating Japanese knotweed from different geographical regions, despite this plant being considered clonal in the United Kingdom. Phenotypic plasticity, where one genotype can express different phenotypes, could be an explanation for the ability to identify plants from different regions, despite their supposable genetic consistency. This is particularly significant for clonal plants, such as Japanese knotweed (Mounger *et al.*, 2021). Phenotypic plasticity is a potentially important mechanism for introduced species in overcoming the genetic bottleneck, maintaining health components such as growth, survival, fertility and overall vigour (Parker *et al.*, 2003; Richards *et al.*, 2006; Geng *et al.*, 2007; Pichancourt & van Klinken, 2012). Populations of alien plants are known to have higher frequencies of clonality than native plant species (Silvertown, 2008), and clonality is thought to be an important characteristic of invasive alien plants (Pyšek, 1997; Mounger *et al.*, 2021). A high proportion of successful invasive plants are clonal; 70% of 468 studied species from the ICUN database and 81% of the one hundred worst invasive plants (Mounger *et al.*, 2021). Epigenetics could be an important mechanism for clonal plants as by reproducing asexually they are able to bypass the meiotic resetting of epigenetic modifications (Mounger *et al.*, 2021). Asexual species can additionally maintain genetic variation through somatic mutation, allowing adaptation to changing environmental conditions (Schoen & Schultz, 2019; Mounger *et al.*, 2021). Epigenetic modifications in gene expression and function have been recognized as key mechanisms behind phenotypic variation of plant traits in response to environmental cues (Herrera & Bazaga, 2010; Banerjee *et al.*, 2019). Although there are plastic responses which are not epigenetic such as provisioning and biochemical functioning (Herman & Sultan, 2011; Sultan, 2015; Richards *et al.*, 2017; Banta & Richards, 2018; Richards & Pigliucci, 2020; Mounger *et al.*, 2021), phenotypic

variation of plant traits in response to environmental cues could be a result of epigenetic modifications to gene expression and function (Herrera & Bazaga, 2010; Banerjee *et al.*, 2019). This raises the intriguing possibility that epigenetic modifications may contribute to the phenotypic plasticity allowing successful invasion of Japanese knotweed in a diverse range of habitats (Richards *et al.*, 2008, 2012; Zhang *et al.*, 2016). In Western Europe, very little genetic variation of Japanese knotweed has been found (Hollingsworth & Bailey, 2000a; Zhang *et al.*, 2016), which is consistent with the lack of dramatic variation shown in the PCA results in this study (Figure 3.2a). However, the ability to separate Japanese knotweed spectra with SVM (Figure 3.2c) indicates that there are differences common to each region. Despite its clonal nature, AFLP studies have shown an unprecedented level of epigenetic variation in *R. japonica*, particularly across Central Europe (Richards *et al.*, 2012; Zhang *et al.*, 2016). Additionally, in North America invasion of diverse habitats by few Japanese knotweed genotypes has been correlated with epigenetic differentiation, with the conclusion that some epigenetic loci may respond to local microhabitat conditions (Richards *et al.*, 2012). Japanese knotweed from different sources grown in the same greenhouse have been known to possess differing levels of physiological vigour, with the French variant growing more vigorously than its Japanese counterpart, suggesting either a rapid evolution or pre-adaptation (Rouifed *et al.*, 2018).

3.6 Conclusion and Future Work

Japanese knotweed can colonise a wide variety of environmental habitats despite its genetic continuity as the world's largest female clone. ATR-FTIR spectroscopy with subsequent chemometric analysis proved to be a successful tool for identifying Japanese knotweed grown in different environments, and even individual sites within the same geographical region. However, the chemical differences between Japanese knotweed from NWE, NEE and WS could not be explained by the soil or climatic parameters measured in this study. This lack of correlation raises important questions about the causes of these subtle variances because, as revealed by ATR-FTIR spectroscopy, subtle differences do exist between regions. These variations

may be due to phenotypic plasticity, a trait shared by other clonal invasive plants. Further studies will be necessary to elucidate the mechanistic basis for the effects of environmental conditions on Japanese knotweed, including the possible contribution of epigenetic modifications, and the connection with the robust growth habit of this species.

3.7 References

Aguilera AG, Alpert P, Dukes JS, Harrington R. **2010**. Impacts of the invasive plant *Fallopia japonica* (Houtt.) on plant communities and ecosystem processes. *Biological Invasions* **12**: 1243–1252.

Allen SE, Grimshaw HM, Parkinson JA, Quarmby C. **1989**. Chemical Analysis of Ecological Materials (Second Edition). *The Journal of Applied Ecology* **13**: 368.

Bailey JP, Bímová K, Mandák B. **2009**. Asexual spread versus sexual reproduction and evolution in Japanese Knotweed *s.l.* sets the stage for the 'battle of the Clones'. *Biological Invasions*. **11(5):** 1189-1203.

Bailey JP, Conolly AP. **2000**. Prize-winners to pariahs -A history of Japanese Knotweed *s.l.* (Polygonaceae) in the British Isles. *Watsonia* **23**: 93–110.

Baker MJ, Trevisan J, Bassan P, Bhargava R, Butler HJ, Dorling KM, Fielden PR, Fogarty SW, Fullwood NJ, Heys KA, *et al.* 2014. Using Fourier transform IR spectroscopy to analyze biological materials. *Nature Protocols* **9**: 1771–1791.

Banerjee AK, Guo W, Huang Y. **2019**. Genetic and epigenetic regulation of phenotypic variation in invasive plants - Linking research trends towards a unified framework. *NeoBiota* **49**: 77–103.

Banta JA, Richards CL. **2018**. Quantitative epigenetics and evolution. *Heredity* **121**: 210–224.

Bashtanova UB, Beckett KP, Flowers TJ. **2009**. Review: Physiological Approaches to the Improvement of Chemical Control of Japanese Knotweed (*Fallopia japonica*). *Weed Science* **57**: 584–592.

Beerling DJ, Bailey JP, Conolly AP. **1994**. *Fallopia Japonica* (Houtt.) Ronse Decraene. *The Journal of Ecology* **82**: 959.

Belz RG, Duke SO. **2014**. Herbicides and plant hormesis. *Pest Management Science* **70**: 698–707.

Bímová K, Mandák B, Pyšek P. **2003**. Experimental study of vegetative regeneration in four invasive Reynoutria taxa (Polygonaceae). *Plant Ecology* **166**: 1-11.

Bolan NS, Curtin D, Adriano DC. **2004**. Acidity. In: Encyclopedia of Soils in the Environment. Elsevier Inc., 11–17.

Booy O, Wade M, Roy H. **2015**. *Field guide to invasive plants and animals in Britain*. Bloomsbury Publishing.

Bossdorf O, Auge H, Lafuma L, Rogers WE, Siemann E, Prati D. **2005**. Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia* **144**: 1–11.

Buitrago MF, Groen TA, Hecker CA, Skidmore AK. **2016**. Changes in thermal infrared spectra of plants caused by temperature and water stress. *ISPRS Journal of Photogrammetry and Remote Sensing* **111**: 22–31.

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

Butler HJ, Martin FL, Roberts MR, Adams S, McAinsh MR. **2020**. Observation of nutrient uptake at the adaxial surface of leaves of tomato (*Solanum lycopersicum*) using Raman spectroscopy. *Analytical Letters* **53**: 536–562.

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

Butler HJ, Smith BR, Fritzsch R, Radhakrishnan P, Palmer DS, Baker MJ. **2018**. Optimised spectral pre-processing for discrimination of biofluids via ATR-FTIR spectroscopy. *Analyst* **143**: 6121–6134.

Cappuccino N, Arnason JT. **2006**. Novel chemistry of invasive exotic plants. *Biology Letters* **2**: 189–193.

Centre for Agriculture and Bioscience International (CABI). **2018**. Data Sheet: *Fallopia japonica* (Japanese knotweed).

Chan KLA, Kazarian SG. **2016**. Attenuated total reflection Fourier-transform infrared (ATR-FTIR) imaging of tissues and live cells. *Chemical Society Reviews* **45**: 1850–1864.

Chen J, Guo B, Yan R, Sun S, Zhou Q. **2017**. Rapid and automatic chemical identification of the medicinal flower buds of Lonicera plants by the benchtop and hand-held Fourier transform infrared spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **182**: 81–86.

Cirino de Carvalho L, de Lelis Medeiros de Morais C, Gomes de Lima KM, Cunha Júnior LC, Martins Nascimento PA, Bosco de Faria J, Henrique de Almeida Teixeira G. **2016**. Determination of the geographical origin and ethanol content of Brazilian sugarcane spirit using near-infrared spectroscopy coupled with discriminant analysis. *Analytical Methods* **8**: 5658–5666.

Cortes C, Vapnik V, Saitta L. **1995**. *Support-Vector Networks Editor*. Kluwer Academic Publishers.

Dassonville N, Vanderhoeven S, Gruber W, Meerts P. **2007**. Invasion by *Fallopia japonica* increases topsoil mineral nutrient concentrations. *Écoscience* **14**: 230–240.

Davidson AM, Jennions M, Nicotra AB. **2011**. Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. *Ecology Letters* **14**: 419–431.

Emmett B, Frogbrook Z, Chamberlain P, Giffiths R, Pickup R, Poskitt J, Reynolds B, Rowe E, Spurgeon D, Rowland P, *et al.* **2008**. CS Technical Report No 3/07 - Soils Manual Volume 1. *Centre for Ecology and Hydrology, Wallingford, UK.*

Environment Agency. **2013**. *Managing Japanese knotweed on development sites: the knotweed code of practice.*

Estoup A, Ravigné V, Hufbauer R, Vitalis R, Gautier M, Facon B. **2016**. Is There a Genetic Paradox of Biological Invasion? *Annual Review of Ecology, Evolution, and Systematics* **47**: 51–72.

Euring D, Löfke C, Teichmann T, Polle A. **2012**. Nitrogen fertilization has differential effects on N allocation and lignin in two Populus species with contrasting ecology. *Trees - Structure and Function* **26**: 1933–1942.

Falcão L, Araújo MEM. **2013**. Tannins characterization in historic leathers by complementary analytical techniques ATR-FTIR, UV-Vis and chemical tests. *Journal of Cultural Heritage* **14**: 499–508.

Fan P, Hay A, Marston A, Hostettmann K. **2009**. Allelopathic potential of phenolic constituents from *Polygonum cuspidatum* Sieb. & Zucc (Polygonaceae). *Planta Medica* **75**: 09.

Fennell M, Wade M, Bacon KL. **2018**. Japanese knotweed (*Fallopia japonica*): an analysis of capacity to cause structural damage (compared to other plants) and typical rhizome extension. *PeerJ* **6**: e5246.

Fischer C, Speth V, Fleig-Eberenz S, Neuhaus G. **1997**. Induction of Zygotic Polyembryos in Wheat: Influence of Auxin Polar Transport. *The Plant Cell* **9**: 1767–1780.

Geng Y, van Klinken RD, Sosa A, Li B, Chen J, Xu C-Y. **2016**. The Relative Importance of Genetic Diversity and Phenotypic Plasticity in Determining Invasion Success of a Clonal Weed in the USA and China. *Frontiers in Plant Science* **7**: 216.

Geng YP, Pan XY, Xu CY, Zhang WJ, Li B, Chen JK, Lu BR, Song ZP. **2007**. Phenotypic plasticity rather than locally adapted ecotypes allows the invasive alligator weed to colonize a wide range of habitats. *Biological Invasions* **9**: 245–256.

Gerber E, Krebs C, Murrell C, Moretti M, Rocklin R, Schaffner U. **2008**. Exotic invasive knotweeds (Fallopia spp.) negatively affect native plant and invertebrate assemblages in European riparian habitats. *Biological Conservation* **141**: 646–654.

Gomes MP, Le Manac'h SG, Hénault-Ethier L, Labrecque M, Lucotte M, Juneau P. **2017**. Glyphosate-Dependent Inhibition of Photosynthesis in Willow. *Frontiers in Plant Science* **8**: 207.

Groeneveld E, Belzile F, Lavoie C. **2014**. Sexual reproduction of Japanese knotweed (*Fallopia japonica s.l.*) at its northern distribution limit: new evidence of the effect of climate warming on an invasive species. *American journal of botany* **101**: 459–66.

Gurevitch J, Fox GA, Wardle GM, Inderjit, Taub D. **2011**. Emergent insights from the synthesis of conceptual frameworks for biological invasions. *Ecology Letters* **14**: 407–418.

Hagenblad J, Hülskötter J, Acharya KP, Brunet J, Chabrerie O, Cousins SAO, Dar PA, Diekmann M, De Frenne P, Hermy M, *et al.* 2015. Low genetic diversity despite multiple introductions of the invasive plant species Impatiens glandulifera in Europe. *BMC Genetics* **16**: 103.

Hammer B, Gersmann K. **2003**. A note on the universal approximation capability of support vector machines. *Neural Processing Letters* **17**: 43–53.

Hejda M, Pyšek P, Jarošík V. **2009**. Impact of invasive plants on the species richness, diversity and composition of invaded communities. *Journal of Ecology* **97**: 393–403.

Heredia-Guerrero JA, Benítez JJ, Domínguez E, Bayer IS, Cingolani R, Athanassiou A, Heredia A. **2014**. Infrared and Raman spectroscopic features of plant cuticles: a review. *Frontiers in plant science* **5**: 305.

Herman JJ, Sultan SE. **2011**. Adaptive transgenerational plasticity in plants: Case studies, mechanisms, and implications for natural populations. *Frontiers in Plant Science* **2**: 102.

Herrera CM, Bazaga P. **2010**. Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet Viola cazorlensis. *New Phytologist* **187**: 867–876.

Hollingsworth ML, Bailey JP. **2000a**. Evidence for massive clonal growth in the invasive weed *Fallopia japonica* (Japanese Knotweed). *Botanical Journal of the Linnean Society* **133**: 463–472.

Hollingsworth ML, Bailey JP. **2000b**. Hybridisation and clonal diversity in some introduced Fallopia species (Polygonaceae). *Watsonia* **23**: 111–121.

Holm AK, Elameen A, Oliver BW, Brandsæter LO, Fløistad IS, Brurberg MB. **2018**. Low genetic variation of invasive Fallopia spp. in their northernmost European distribution range. *Ecology and Evolution* **8**: 755–764.

Hoogsteen MJJ, Lantinga EA, Bakker EJ, Groot JCJ, Tittonell PA. **2015**. Estimating soil organic carbon through loss on ignition: effects of ignition conditions and structural water loss. *European Journal of Soil Science* **66**: 320–328.

Johnson KL, Gidley MJ, Bacic A, Doblin MS. **2018**. Cell wall biomechanics: a tractable challenge in manipulating plant cell walls 'fit for purpose'! *Current Opinion in Biotechnology* **49**: 163–171.

Jones D, Bruce G, Fowler MS, Law-Cooper R, Graham I, Abel A, Street-Perrott FA, Eastwood D. **2018**. Optimising physiochemical control of invasive Japanese knotweed. *Biological Invasions* **20**: 2091–2105.

Kaszowska Z, Malek K, Pańczyk M, Mikołajska A. **2013**. A joint application of ATR-FTIR and SEM imaging with high spatial resolution: Identification and distribution of painting materials and their degradation products in paint cross sections. *Vibrational Spectroscopy* **65**: 1–11.

Kazarian SG, Chan KLA. **2013**. ATR-FTIR spectroscopic imaging: recent advances and applications to biological systems. *The Analyst* **138**: 1940.

Kidd H, Shaw D. **2000**. Japanese knotweed - The world's largest female! *Pesticide Outlook* **11**: 99–100.

van Kleunen M, Bossdorf O, Dawson W. **2018**. The Ecology and Evolution of Alien Plants. *Annual Review of Ecology, Evolution, and Systematics* **49**: 25–47.

Van Kleunen M, Weber E, Fischer M. **2010**. A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecology Letters* **13**: 235–245.

Lavoie C. **2017**. The impact of invasive knotweed species (Reynoutria spp.) on the environment: review and research perspectives. *Biological Invasions* **19**: 2319–2337.

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

Lopes J, Sousa C. **2018**. Vibrational spectroscopy for plant varieties and cultivars characterization. In: Comprehensive Analytical Chemistry. 2–299.

Lowe S, Browne M, Boudjelas S, De Poorter M. **2000**. *100 of the world's worst invasive alien species: a selection from the global invasive species database*. Invasive Species Specialist Group Auckland, New Zealand **12**

Luo Y, Liu H, Wu C, Paraskevaidi M, Deng Y, Shi W, Yuan Y, Feng R, Martin FL, Pang W. **2021**. Diagnostic Segregation of Human Breast Tumours Using Fourier-transform Infrared Spectroscopy Coupled with Multivariate Analysis: Classifying Cancer Subtypes. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **255**: 119694.

Mallamace F, Corsaro C, Mallamace D, Vasi S, Vasi C, Dugo G. **2015**. The role of water in protein's behavior: The two dynamical crossovers studied by NMR and FTIR techniques. *Computational and Structural Biotechnology Journal* **13**: 33–37.

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

McNear DH, Chaney RL, Sparks DL. **2010**. The hyperaccumulator Alyssum murale uses complexation with nitrogen and oxygen donor ligands for Ni transport and storage. *Phytochemistry* **71**: 188–200.

Mejia AF, Nebel MB, Eloyan A, Caffo B, Lindquist MA. **2017**. PCA leverage: Outlier detection for high-dimensional functional magnetic resonance imaging data. *Biostatistics* **18**: 521–536.

de Mendiburu F. **2020**. agricolae. *Universidad Nacional Agraria La Molina*.

Met Office. **2018**. Regional values - Met Office. *https://www.metoffice.gov.uk/research/climate/maps-and-data/regional-values*.

Miller AJ. **2014**. Plant Mineral Nutrition. In: eLS. Chichester, UK: John Wiley & Sons, Ltd.

Morais CLM, Costa FSL, Lima KMG. **2017**. Variable selection with a support vector machine for discriminating Cryptococcus fungal species based on ATR-FTIR spectroscopy. *Analytical Methods* **9**: 2964–2970.

Morais CLM, Lima KMG, Singh M, Martin FL. **2020**. Tutorial: multivariate classification for vibrational spectroscopy in biological samples. *Nature Protocols* **15**: 2143–2162.

Morais CLM, Paraskevaidi M, Cui L, Fullwood NJ, Isabelle M, Lima KMG, Martin-Hirsch PL, Sreedhar H, Trevisan J, Walsh MJ, *et al.* 2019. Standardization of complex biologically derived spectrochemical datasets. *Nature Protocols* **14**: 1546–1577.

Mounger J, Ainouche M, Bossdorf O, Cavé-Radet A, Li B, Parepa M, Salmon A, Yang J, Richards C. **2021**. Epigenetics and the success of invasive plants. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376(1826)**: 20200117.

Movasaghi Z, Rehman S, ur Rehman I. **2008**. Fourier Transform Infrared (FTIR) Spectroscopy of Biological Tissues. *Applied Spectroscopy Reviews* **43**: 134–179.

Murphy BW. **2015**. Impact of soil organic matter on soil properties - A review with emphasis on Australian soils. *Soil Research* **53**: 605–635.

Murrell C, Gerber E, Krebs C, Parepa M, Schaffner U, Bossdorf O. **2011**. Invasive knotweed affects native plants through allelopathy. *American Journal of Botany* **98**: 38–43.

Nentwig W, Bacher S, Kumschick S, Pyšek P, Vilà M. **2018**. More than "100 worst" alien species in Europe. *Biological Invasions* **20**: 1611–1621.

Olsen SR, Cole C V, Watanabe FS, Dean LA, States. U, Agriculture. D of. **1954**. *Estimation of available phosphorus in soils by extraction with sodium bicarbonate*. Washington, D.C.: U.S. Dept. of Agriculture.

Orcelli T, di Mauro E, Urbano A, Valezi DF, da Costa ACS, Zaia CTB V., Zaia DAM. **2018**. Study of Interaction Between Glyphosate and Goethite Using Several Methodologies: an Environmental Perspective. *Water, Air, & Soil Pollution* **229**: 150.

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

Orwin KH, Buckland SM, Johnson D, Turner BL, Smart S, Oakley S, Bardgett RD. **2010**. Linkages of plant traits to soil properties and the functioning of temperate grassland. *Journal of Ecology* **98**: 1074–1083.

Palacio-López K, Gianoli E. **2011**. Invasive plants do not display greater phenotypic plasticity than their native or non-invasive counterparts: a meta-analysis. *Oikos* **120**: 1393–1401.

Parepa M, Bossdorf O. **2016**. Testing for allelopathy in invasive plants: it all depends on the substrate! *Biological Invasions* **18**: 2975–2982.

Parepa M, Fischer M, Bossdorf O. **2013**. Environmental variability promotes plant invasion. *Nature Communications* **4**: 1–4.

Parepa M, Kahmen A, Werner RA, Fischer M, Bossdorf O. **2019**. Invasive knotweed has greater nitrogen-use efficiency than native plants: evidence from a 15N pulse-chasing experiment. *Oecologia* **191**: 389–396.

Parepa M, Schaffner U, Bossdorf O. **2012**. Sources and modes of action of invasive knotweed allelopathy: the effects of leaf litter and trained soil on the germination and growth of native plants. *NeoBiota* **13**: 15–30.

Parker IM, Rodriguez J, Loik ME. **2003**. An Evolutionary Approach to Understanding the Biology of Invasions: Local Adaptation and General-Purpose Genotypes in the Weed Verbascum thapsus. *Conservation Biology* **17**: 59–72.

Peng W, Qin R, Li X, Zhou H. **2013**. Botany, phytochemistry, pharmacology, and potential application of *Polygonum cuspidatum* Sieb.et Zucc.: A review. *Journal of Ethnopharmacology* **148**: 729–745.

Pichancourt J-B, van Klinken RD. **2012**. Phenotypic Plasticity Influences the Size, Shape and Dynamics of the Geographic Distribution of an Invasive Plant (G Bonaventure, Ed.). *PLoS ONE* **7**: e32323.

Pilling M, Gardner P. **2016**. Fundamental developments in infrared spectroscopic imaging for biomedical applications. *Chemical Society Reviews* **45**: 1935–1957.

Price EAC, Gamble R, Williams GG, Marshall C. **2002**. Seasonal patterns of partitioning and remobilization of 14C in the invasive rhizomatous perennial Japanese knotweed (*Fallopia japonica* (Houtt.) Ronse Decraene). In: Ecology and Evolutionary Biology of Clonal Plants. Dordrecht: Springer Netherlands, 125–140.

Professional Agricultural Analysis Group. **2013**. Routine soil samples. *PAAG soil sampling guide*.

Pyšek P. **1997**. The ecology and evolution of clonal plants. *Backhuys Publishers, Leiden*.

R Core Team. **2014**. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.*

Rana R, Müller G, Naumann A, Polle A. **2008**. FTIR spectroscopy in combination with principal component analysis or cluster analysis as a tool to distinguish beech (Fagus sylvatica L.) trees grown at different sites. *Holzforschung* **62**: 530–538.

Richards CL, Alonso C, Becker C, Bossdorf O, Bucher E, Colomé-Tatché M, Durka W, Engelhardt J, Gaspar B, Gogol-Döring A, *et al.* 2017. Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecology Letters* **20**: 1576–1590.

Richards CL, Bossdorf O, Muth NZ, Gurevitch J, Pigliucci M. **2006**. Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecology Letters* **9**: 981–993.

Richards C, Pigliucci M. **2020**. Epigenetic Inheritance. A Decade into the Extended Evolutionary Synthesis. *Paradigmi* **XXXVIII**: 463–494.

Richards CL, Schrey AW, Pigliucci M. **2012**. Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation (M Velend, Ed.). *Ecology Letters* **15**: 1016–1025.

Richards CL, Walls RL, Bailey JP, Parameswaran R, George T, Pigliucci M. **2008**. Plasticity in salt tolerance traits allows for invasion of novel habitat by Japanese knotweed s. l. (*Fallopia japonica* and *F.*x *bohemica*, Polygonaceae). *American Journal of Botany* **95**: 931–942.

Rouifed S, Puijalon S, Bardon C, Meiffren G, Buonomo A, Sebei N, Poussineau S, Vallier F, Shimoda M, Piola F. **2018**. Comparison of defence and performance traits between one widespread clone and native populations in a major invasive plant species (I Ibáñez, Ed.). *Diversity and Distributions* **24**: 297–312.

Roy HE, Bacher S, Essl F, Adriaens T, Aldridge DC, Bishop JDD, Blackburn TM, Branquart E, Brodie J, Carboneras C, *et al.* **2018**. Developing a list of invasive alien species likely to threaten biodiversity and ecosystems in the European Union. *Global Change Biology* **25**: 1032–1048.

Ruiz-Sola MÁ, Rodríguez-Concepción M. **2012**. Carotenoid Biosynthesis in Arabidopsis: A Colorful Pathway. *The Arabidopsis Book* **10**: e0158.

Santo P. **2017**. Assessing diminution in value of residential properties affected by Japanese Knotweed. *Journal of Building Survey, Appraisal & Valuation* **6(11)**: 211-221.

Schoen DJ, Schultz ST. **2019**. Somatic Mutation and Evolution in Plants. *Annual Review of Ecology, Evolution, and Systematics* **50**: 49–73.

Silvertown J. **2008**. The evolutionary maintenance of sexual reproduction: Evidence from the ecological distribution of asexual reproduction in clonal plants. *International Journal of Plant Sciences* **169**: 157–168.

Simsek Ozek N, Tuna S, Erson-Bensan AE, Severcan F. **2010**. Characterization of microRNA-125b expression in MCF7 breast cancer cells by ATR-FTIR spectroscopy. *The Analyst* **135**: 3094.

Singh S, Kaur I, Kariyat R. **2021**. The Multifunctional Roles of Polyphenols in Plant-Herbivore Interactions. *International Journal of Molecular Sciences* **22**: 1442.

Skolik P, McAinsh MR, Martin FL. **2019a**. ATR-FTIR spectroscopy non-destructively detects damage-induced sour rot infection in whole tomato fruit. *Planta* **249**: 925–939.

Skolik P, Morais CLM, Martin FL, McAinsh MR. **2019b**. Determination of developmental and ripening stages of whole tomato fruit using portable infrared spectroscopy and Chemometrics. *BMC Plant Biology* **19**: 236.

Smith JMD, Ward JP, Child LE, Owen MR. **2007**. A simulation model of rhizome networks for *Fallopia japonica* (Japanese knotweed) in the United Kingdom. *Ecological Modelling* **200**: 421–432.

Spalding K, Bonnier F, Bruno C, Blasco H, Board R, Benz-de Bretagne I, Byrne HJ, Butler HJ, Chourpa I, Radhakrishnan P, *et al.* 2018. Enabling quantification of protein concentration in human serum biopsies using attenuated total reflectance – Fourier transform infrared (ATR-FTIR) spectroscopy. *Vibrational Spectroscopy* **99**: 50–58.

Stasiak MA, Hofstra G, Fletcher RA. **1992**. Physiological changes induced in birch seedlings by sublethal applications of glyphosate. *Canadian Journal of Forest Research* **22**: 812–817.

Strong R, Martin FL, Jones KC, Shore RF, Halsall CJ. **2017**. Subtle effects of environmental stress observed in the early life stages of the Common frog, Rana temporaria. *Scientific Reports* **7**: 1–13.

Sultan SE. **2015**. *Organism and environment: ecological development, niche construction, and adaption*. Oxford University Press, USA.

Talari ACS, Martinez MAG, Movasaghi Z, Rehman S, Rehman IU. **2017**. Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Applied Spectroscopy Reviews* **52**: 456–506.

Tamura M, Tharayil N. **2014**. Plant litter chemistry and microbial priming regulate the accrual, composition and stability of soil carbon in invaded ecosystems. *New Phytologist* **203**: 110–124.

Traoré M, Kaal J, Martínez Cortizas A. **2018**. Differentiation between pine woods according to species and growing location using FTIR-ATR. *Wood Science and Technology* **52**: 487–504.

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

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

UK Soil Observatory. **2020**. UKSO. *http://mapapps2.bgs.ac.uk/ukso/home.html*

Urgenson LS. **2006**. The Ecological Consequences of Knotweed Invasion into Riparian Forests. *(Master's Thesis) University of Washington*.

Usman K, Al-Ghouti MA, Abu-Dieyeh MH. **2019**. The assessment of cadmium, chromium, copper, and nickel tolerance and bioaccumulation by shrub plant Tetraena qataranse. *Scientific Reports* **9**: 1–11.

de Vries FT, Manning P, Tallowin JRB, Mortimer SR, Pilgrim ES, Harrison KA, Hobbs PJ, Quirk H, Shipley B, Cornelissen JHC. **2012**. Abiotic drivers and plant traits explain landscape-scale patterns in soil microbial communities. *Ecology letters* **15**: 1230–1239.

Wagner H, Liu Z, Langner U, Stehfest K, Wilhelm C. **2010**. The use of FTIR spectroscopy to assess quantitative changes in the biochemical composition of microalgae. *Journal of Biophotonics* **3**: 557–566.

Van De Weert M, Haris PI, Hennink WE, Crommelin DJA. **2001**. Fourier transform infrared spectrometric analysis of protein conformation: Effect of sampling method and stress factors. *Analytical Biochemistry* **297**: 160–169.

Wickham H. **2009**. ggplot2: elegant graphics for data analysis. *Springer-Verlag New York*.

Yang H, Yang S, Kong J, Dong A, Yu S. **2015**. Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy. *Nature Protocols* **10**: 382–396.

Zhang Y-Y, Parepa M, Fischer M, Bossdorf O. **2016**. Epigenetics of colonizing species? A study of Japanese knotweed in Central Europe. In: Barrett SCH, Colautti RI, Dlugosch KM, Rieseberg LH (Eds) Invasion Genetics. Chichester, UK: John Wiley & Sons, Ltd, 328–340.

3.8 Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: The datasets generated and analysed during the current study are available in a supplementary folder.

Competing Interests: The author, Martin McAinsh, is an editor for BMC Plant Biology. All other authors have no competing interests.

Author Contributions: CAH conceived, planned, and carried out the experiments and data analysis. The manuscript was written by CAH and MM with contribution from all the authors. CM aided CAH with the data analysis, designed the method to produce Figure 3.3, and ran the PLS correlations found in the supplementary information. FM provided equipment and expertise in the field of FTIR spectroscopy and chemometrics. PB edited and proofread the manuscript and provided expertise in Japanese knotweed. MM and JET supervised the project. All authors have read and approved the manuscript.

3.9 Supporting Information
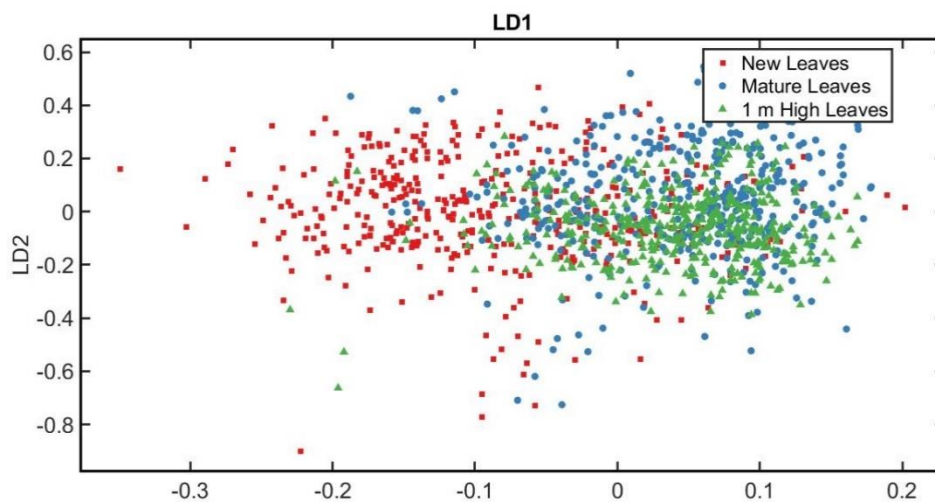
Figure S3.1: PCA-LDA Scatter by leaf type.
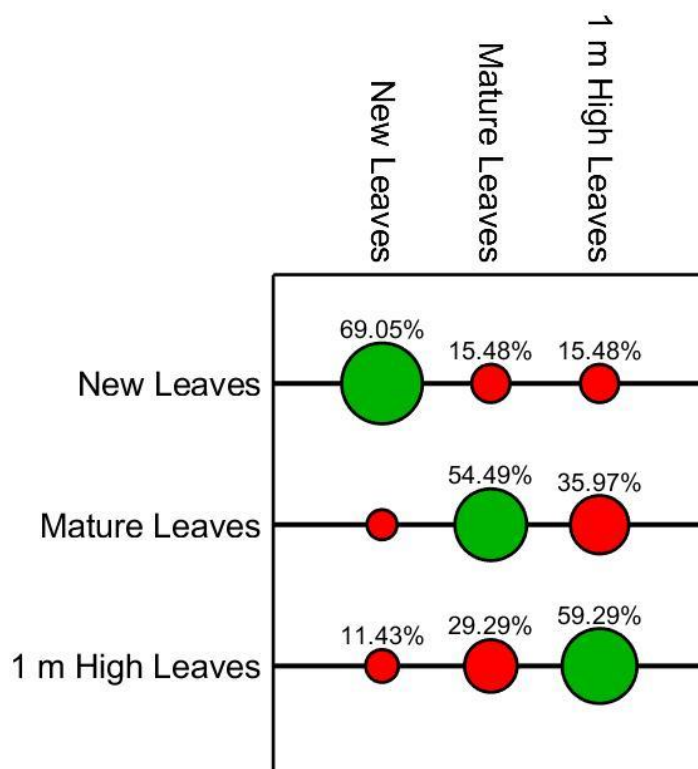
Figure S3.2: SVM classification by leaf type.
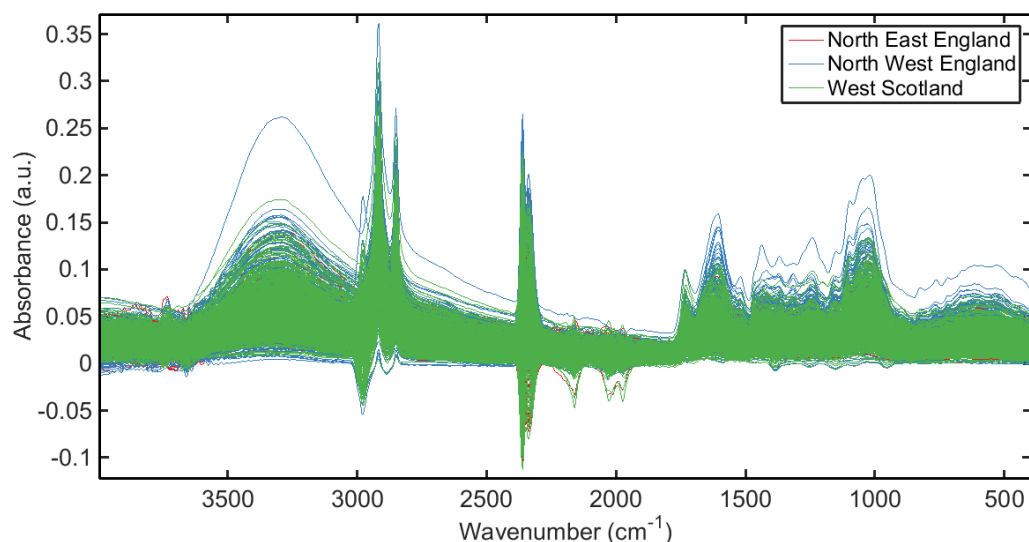
Figure S3.3: Raw IR data for different regions (North East England, North West England, West Scotland).
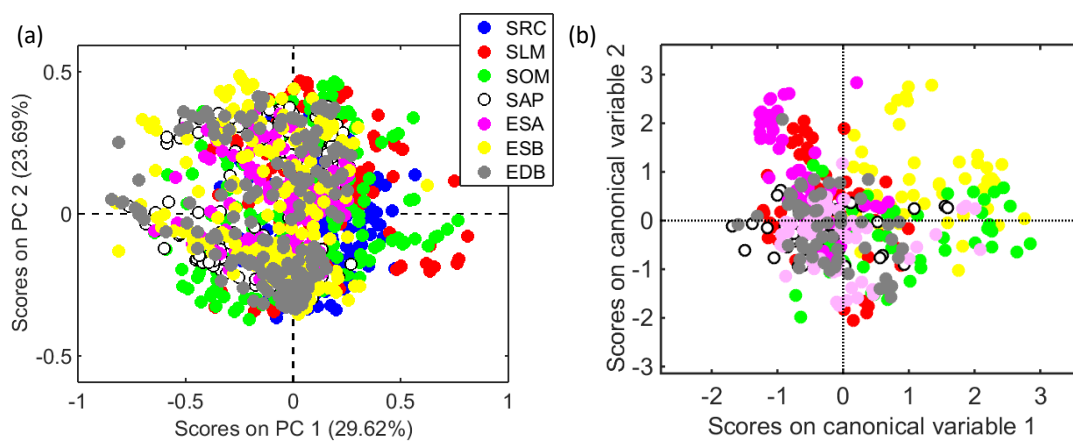
Figure S3.4: a) PCA scores plot and **b)** PCA-LDA canonical scores for the pre-processed spectral data in the fingerprint region for different sites where knotweed was collected (Scotland: SRC, SOM, SLM, SAP; North West England: ESA, ESB; North East England: EDB).
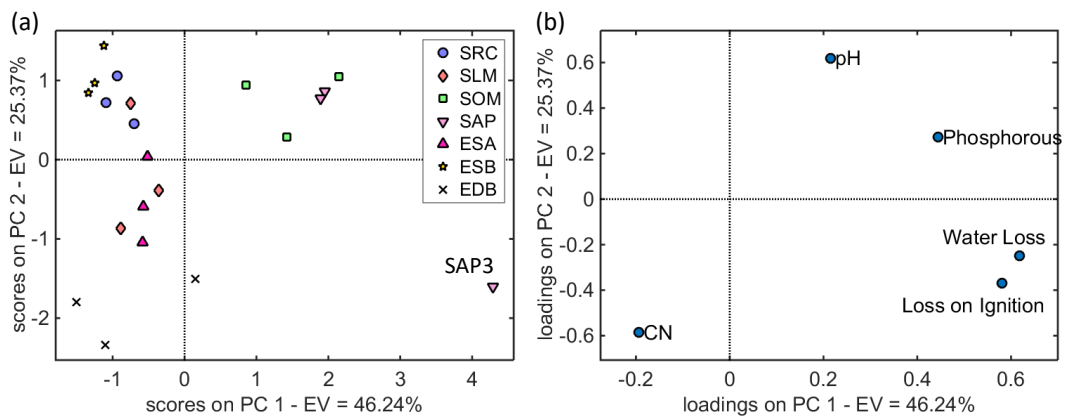
Figure S3.5: a) PCA scores and **b)** loadings for the soil data with extreme sample (SAP3).
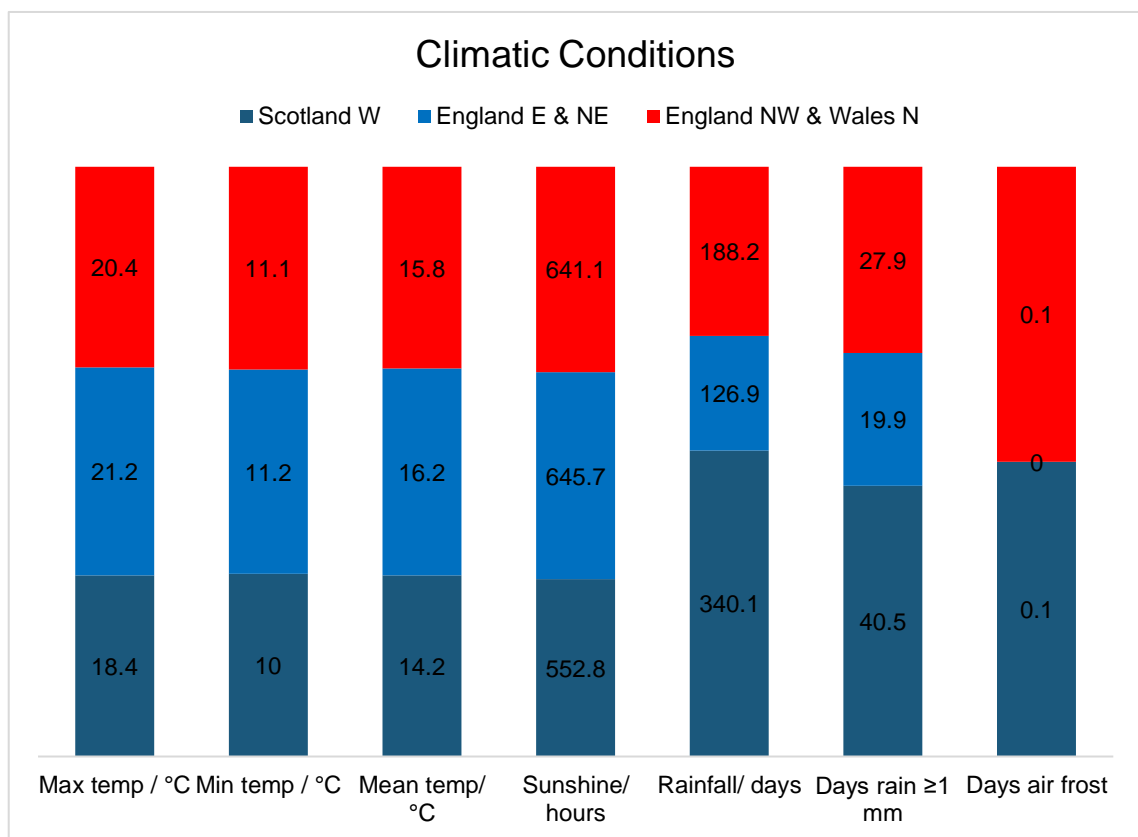
Figure S3.6: Climatic conditions for each region.
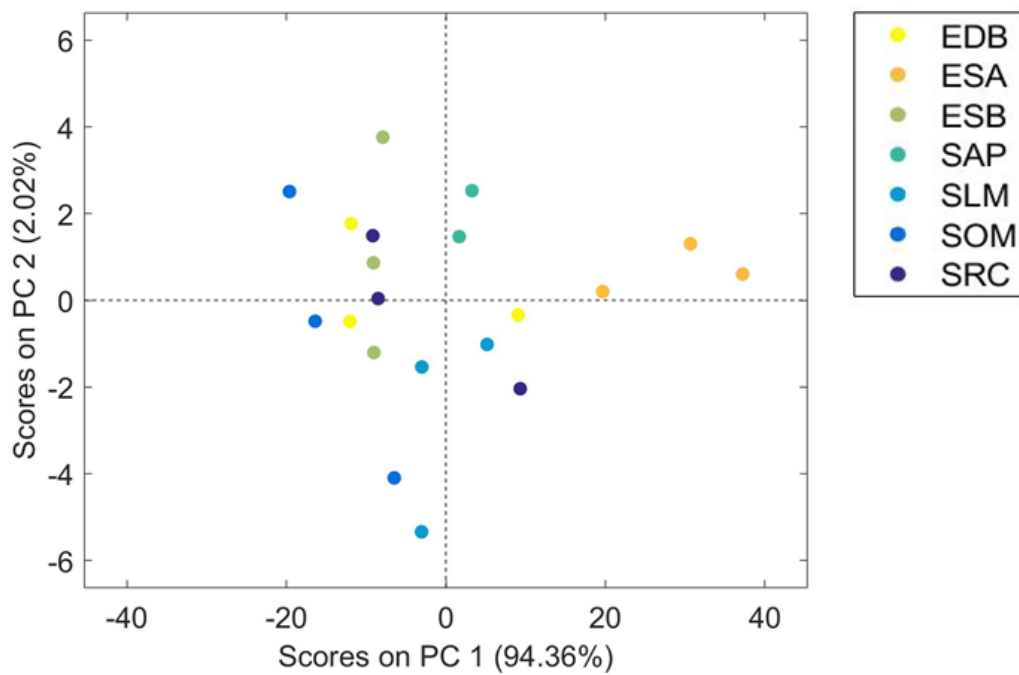
Figure S3.7: PCA scores for spectral and soil data combined (block-scaling applied before PCA).
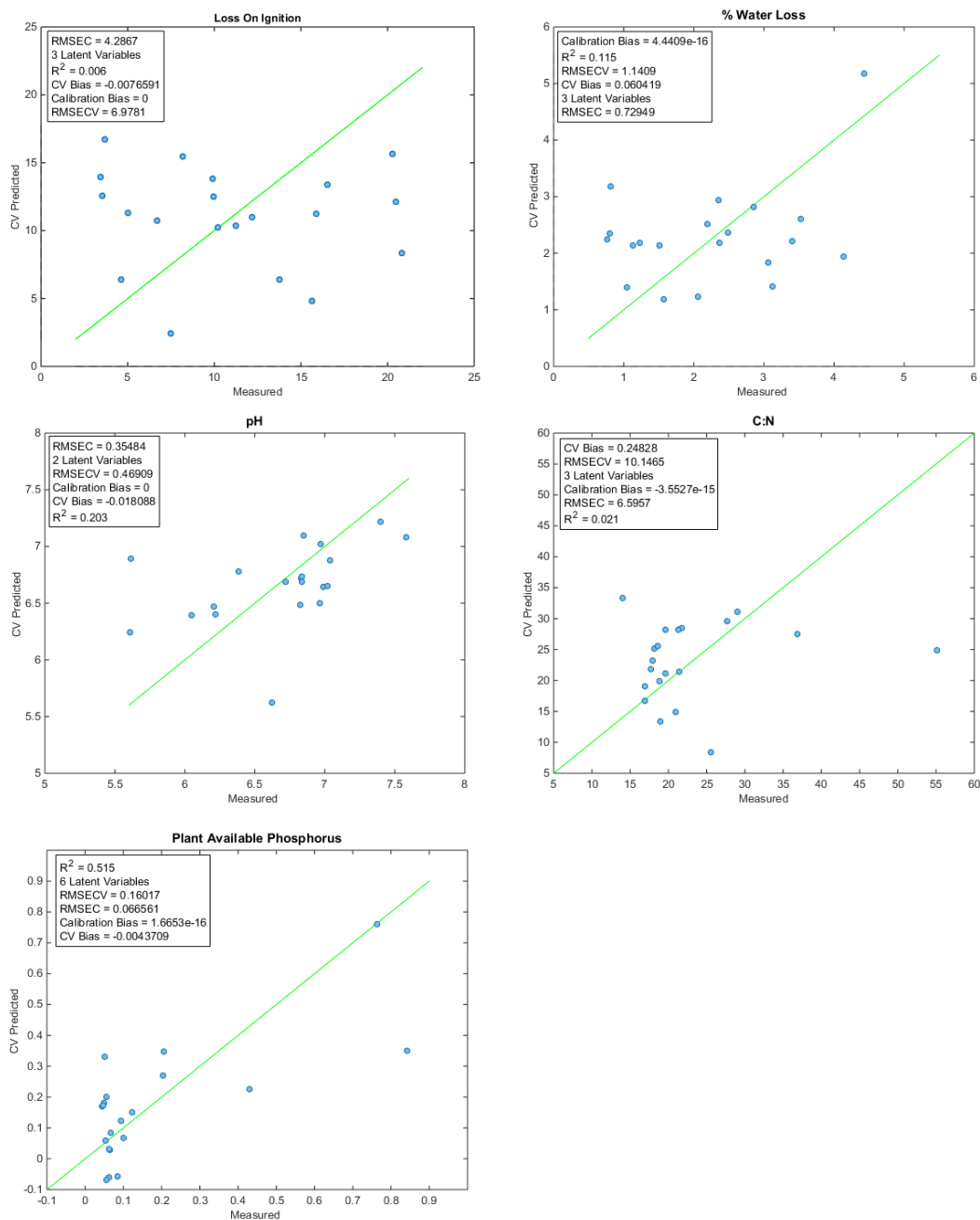
Figure S3.8: Measured versus predicted soil parameters based on the pre-processed spectral data using leave-one-out cross-validated partial least squares (PLS).
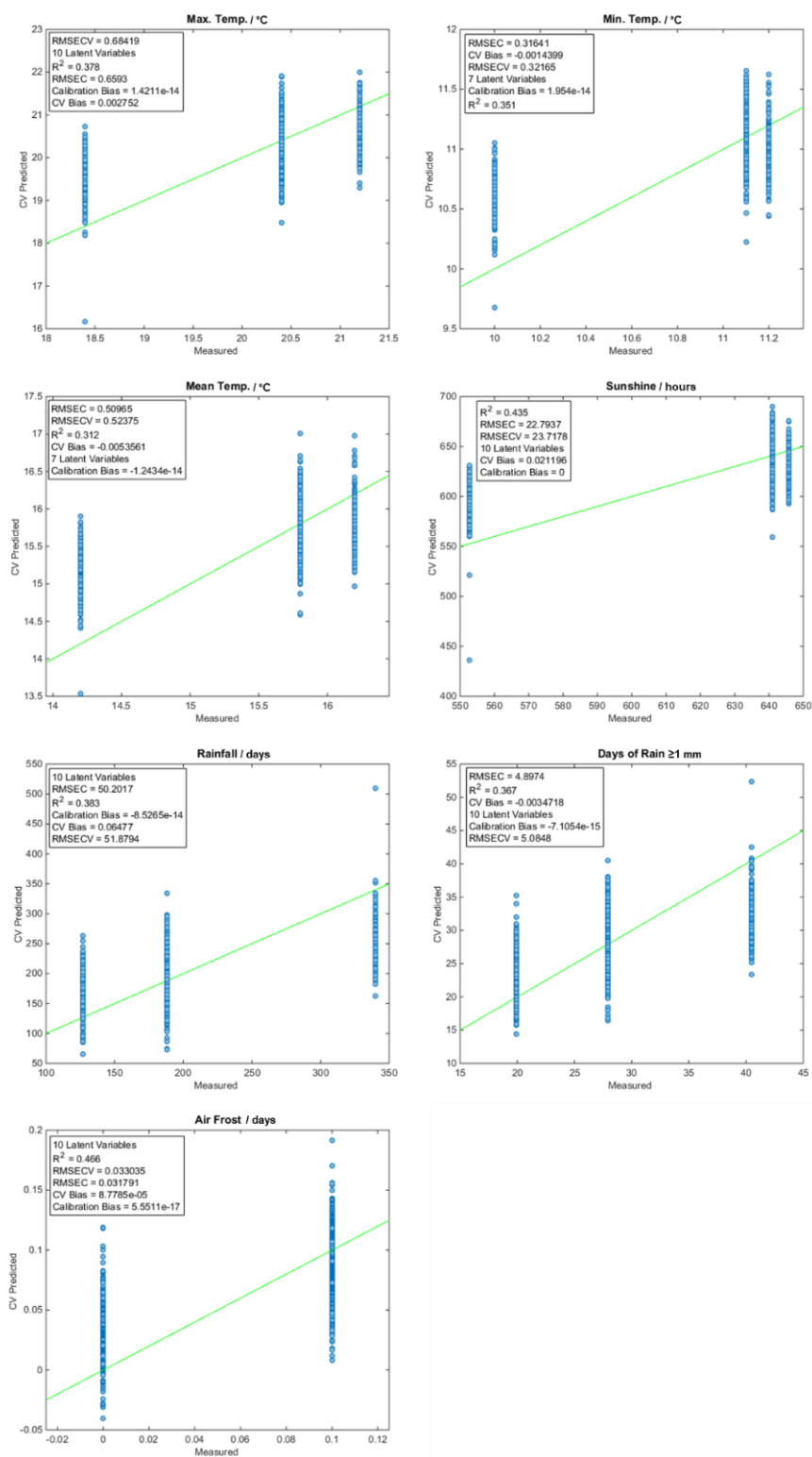
Figure S3.9: Measured versus predicted meteorological parameters based on the pre-processed spectral data using leave-one-out cross-validated partial least squares (PLS).
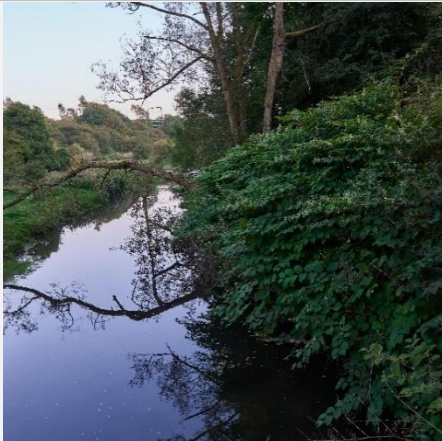
Table S3.1: Site Descriptions. Images taken by Claire A Holden.

Stand	Description	Soil Type	Google Map Co-ordinates	Photo
SAP	Between a public footpath and a railway line in an urban area, near a road. Dead the following year.	Urban	55.8236264, -4.0916304	
SRC	In full sun on Eastern bank of River Clyde	Mineral gleys	55.8186110, -4.0944440	
SLM	Western bank of River Clyde. Noticeably smaller leaves than those of SRC but appeared in overall better health	Mineral gleys	55.8026350, -4.0905780	

SOM	Brownfield site, previously used for numerous purposes including a railway siding and an old mine, but now transformed into a public park. Dead the following year.	Urban	55.8211260, -4.0554206
ESA	Wooded area further removed from river and path than ESB	Clay-to-sandy loam	53.9491670, -2.7550000
ESB	Edge of River Wyre. Main stand in full sun, however accessible leaves collected from the shaded side. Dead the following year.	Clay-to-sandy loam	53.94977780, -2.75541670

EDB	Adjacent to a wooded stream on the edge of one of the repurposed old railway lines now part of the Broompark cycle path network.	Slowly permeable seasonally wet acid loamy and clayey soils	54.764993, -1.609166	

4 Know your enemy: Application of ATR-FTIR spectroscopy to invasive species control

Claire Anne Holden[1*], John Paul Bailey[2], Jane Elizabeth Taylor[1], Francis L Martin[3], Paul Beckett[4], Martin McAinsh[1]

1. Lancaster Environment Centre, Lancaster University, LA1 4YQ, UK

2. Department of Genetics and Genome Biology, Leicester University, LE1 7RH, UK

3. Biocel Ltd, Hull, HU10 7TS, UK

4. Phlorum Ltd, Brighton, BN2 6AH, UK

*Corresponding author: Claire Anne Holden, c.holden6@lancaster.ac.uk

4.1 Abstract

1. Extreme weather and globalisation leave our climate vulnerable to invasion by alien species, which have negative impacts on the economy, biodiversity, and ecosystem services. Rapid and accurate identification is key to the control of invasive alien species. However, visually similar species hinder conservation efforts, for example hybrids within the Japanese Knotweed complex.

2. We applied the novel method of ATR-FTIR spectroscopy combined with chemometrics (mathematics applied to chemical data) to historic herbarium samples, taking 1580 spectra in total. Samples included five species from within the interbreeding Japanese Knotweed complex (including three varieties of Japanese Knotweed), six hybrids and five species from the wider Polygonaceae family. Spectral data from herbarium specimens were analysed with several chemometric techniques: support vector machines (SVM) for differentiation between plant types, supported by ploidy levels; principal component analysis loadings and spectral biomarkers to explore differences between the highly invasive *Reynoutria japonica* var. *japonica* and its non-invasive counterpart *Reynoutria japonica* var. *compacta*; hierarchical cluster analysis (HCA) to investigate the relationship between plants within the Polygonaceae family, of the *Fallopia, Reynoutria, Rumex* and *Fagopyrum* genera.

3. ATR-FTIR spectroscopy coupled with SVM successfully differentiated between plant type, leaf surface and geographical location, even in herbarium samples of varying age. Differences between *Reynoutria japonica* var. *japonica* and *Reynoutria japonica* var. *compacta* included the presence of two polysaccharides, glucomannan and xyloglucan, at higher concentrations in *Reynoutria japonica* var. *japonica* than *Reynoutria japonica* var. *compacta.* HCA analysis indicated that potential genetic linkages are sometimes masked by environmental factors; an effect that can either be reduced or encouraged by altering the input parameters. Entering the absorbance values for key wavenumbers, previously highlighted by principal component analysis loadings, favours linkages in the resultant HCA dendrogram corresponding

to expected genetic relationships, whilst environmental associations are encouraged using the spectral fingerprint region.

4. The ability to distinguish between closely related interbreeding species and hybrids, based on their spectral signature, raises the possibility of using this approach for determining the origin of Japanese knotweed infestations in legal cases where the clonal nature of plants currently makes this difficult and for the targeted control of species and hybrids. These techniques also provide a new method for supporting biogeographical studies.

4.2 Keywords

Cluster Analysis, Fourier Transform Infrared Spectroscopy, Introduced species, Japanese Knotweed, Plants, Polygonaceae, Principal Component Analysis, Support Vector Machine.

4.3 Introduction

Invasive alien species (IAS), such as Japanese Knotweed, detrimentally impact the economy (Williams *et al.*, 2010), ecosystem services (Fennell *et al.*, 2018), and native flora (Lavoie, 2017). The impacts of IAS are set to worsen as an increasing human population heightens the demand for healthy crops (Willett *et al.*, 2019), whilst globalisation (Hulme, 2009b) and extreme weather events (Parepa *et al.*, 2013) create further opportunities for introduction and spread of invasives. Accurate identification is the first step towards management of IAS. While many countries aim to intercept their introduction at border crossings, a lack of taxonomic experts or a world-wide comprehensive approach create barriers to identification (Armstrong & Ball, 2005).

Within the Japanese Knotweed complex, also known as *sensu lato (s.l.)*, misidentification has led to an underestimation of the prevalence of hybridisation (Zika & Jacobson, 2003; Gillies *et al.*, 2016), and similar morphology has complicated management strategies (Moody & Les, 2007). Hybridisation is a strategy employed by IAS to overcome a genetic bottleneck (Ellstrand & Schierenbeck, 2000). It results in 'heterosis', the production of offspring with increased 'hybrid vigour' (Hollingsworth *et al.*, 1998). Hybrid descendants may have improved traits relative to their parents such as invasiveness (Ellstrand & Schierenbeck, 2000), growth rate, reproductive success and yield (Ben-Ari & Lavi, 2012), genetic variance (Lee, 2002), and stress tolerance e.g. to herbicides (Snow *et al.*, 1999) and cold (Milne & Abbott, 2000).

The vigorous hybrid Bohemian Knotweed (*Reynoutria × bohemica)* has advantages over its maternal parent *Reynoutria japonica* var*. japonica*, including the ability to produce viable seed without the need for cross-breeding (Mandák *et al.*, 2003). Despite its increased invasiveness it has not been recognised on the United States Department of Agriculture (USDA), Natural Resources Conservation Service (NRCS) Plants Database where it is still listed as "Absent/Unreported" in the United States of America (USDA Plants Database). Similarly overlooked due to morphological variation (Mandák *et al.*, 2003) is the dwarf variant *Reynoutria japonica* var*. compacta*. Misidentification is a particular concern in hybrids where viable seeds

are produced because glyphosate, the main herbicide used to treat Japanese Knotweed, is applied post-flowering to increase herbicide allocation to rhizomes (Jones D *et al.*, 2018). However, an increasing prevalence of Bohemian Knotweed (Gillies *et al.*, 2016), and the occurrence of stands (clumps) of seeding Japanese Knotweed *s.l.* (Grimsby *et al.*, 2007), means late-season herbicide application may not be an appropriate 'cure-all' treatment program. Correct plant identification is therefore essential for the design of effective and stand-specific treatment programs.

Accurate identification is also important to the biocontrol of Japanese Knotweed. Two strains of psyllid currently under consideration as biocontrol agents exhibit differential development on different plant types; the northern Hokkaido biotype favours Giant Knotweed, whilst the southern Kyushu biotype prefers Japanese Knotweed and Bohemian Knotweed (Jones IM *et al.*, 2020). Consequently, accurate biogeographical information is required when sourcing biocontrol agents to ensure that introduced target species are matched with an agent from the same geographical origin. For Japanese Knotweed *s.l.*, the phylogenetic and biogeographic relationships between plants has been determined through comparison of genetic diversity levels both within, and between, introduced and native ranges (Hollingsworth & Bailey, 2000a; Inamura *et al.*, 2000; Pashley *et al.*, 2007). Phylogenetic studies using current genetic methods, can require specialist techniques and knowledge (Nadeem *et al.*, 2017). Whereas the development of rapid techniques which require minimal sample preparation that can accurately distinguish between morphologically similar species in the field, complementing traditional phylogenetic approaches, could improve the speed and accuracy of plant identification and support effective management strategies for this IAS.

Attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy allows the rapid, marker-free, non-destructive analysis of biological samples (Morais *et al.*, 2020). This technique is being increasingly used in plant science. Applications include differentiation of plants and pollen from different growing regions (Bağcıoğlu *et al.*, 2017; Traoré *et al.*, 2018; Holden *et al.*, 2021); phylogenetic studies (Kim *et al.*, 2004; Demir *et al.*, 2015), response to abiotic factors such as soil

fertility (Euring *et al.*, 2012), heavy metals (McNear *et al.*, 2010; Usman *et al.*, 2019), water and temperature stress (Buitrago *et al.*, 2016), nutrient deficiency and uptake (Butler *et al.*, 2017, 2020); as well as monitoring health and development (Butler *et al.*, 2015b; Skolik *et al.*, 2019b) and infection (Skolik *et al.*, 2019a). Therefore ATR-FTIR spectroscopy appears well suited to fulfil a role in the identification and management of IAS.

ATR-FTIR spectroscopy measures the absorption of infrared light by a sample at specific quantifiable wavenumbers. Energy from absorbed light ($4000–400$ cm^{-1} wavenumbers or $2.5–25\,\mu$m wavelengths) is transformed into vibrational energy through induction of atomic displacement and dipole moment changes (Morais *et al.*, 2020). Patterns of absorption are acquired as spectra comprising complex multivariate data that require chemometrics to derive subtle differences in sample composition. Available mathematical techniques include principal component analysis (PCA) and linear discriminant analysis (LDA), support vector machine (SVM), Naïve Bayes, and artificial neural networks (ANN) (Morais *et al.*, 2020). Biological molecules preferentially absorb light of wavenumbers $1800–900$ cm^{-1}, a range known as the 'fingerprint region', which includes important biological absorptions due to lipids, proteins, carbohydrates, nucleic acids and protein phosphorylation (see (Morais *et al.*, 2020)). Databases are available with catalogued definitions for characteristic peak frequencies (e.g. (Talari *et al.*, 2017)). For example, absorptions have been linked to biologically significant compounds including glucomannan (Kanter *et al.*, 2013), xyloglucan (Kacuráková *et al.*, 2000), succinate (Kang *et al.*, 2008), and pectin (Sharma & Uttam, 2018). However, the process from chemometric biomarker identification to physical biomolecular extraction is the subject of ongoing research focused on calibrating concentrations derived from biological spectra (Spalding *et al.*, 2018), consolidating the expanding database of key wavenumber changes and associated molecular definitions (Talari *et al.*, 2017), and trialling new biological applications (Skolik *et al.*, 2019a,b; Butler *et al.*, 2020).

This study aims to develop a tool to support management strategies through clarification of IAS species assignment, population dynamics, and biogeography. Spectral data were analysed using a combination of mathematical techniques: principal component analysis (PCA) to assess the natural variation; the classifier support vector machines (SVM), a supervised technique, to allow identification of closely related interbreeding species and hybrids, supported by ploidy levels; PCA-LDA and biomarkers to elucidate the biochemical differences between invasive and non-invasive varieties; and hierarchical cluster analysis (HCA) to explore species phylogeny.

4.4 Materials and Methods

4.4.1 Herbarium Samples

Samples were obtained from the University of Leicester herbarium (LTR), previously collected between 1935-2000, see Table S4.1 for detailed sample information. Leaves were air-dried at the time of collection and subsequently stored in cardboard folders within purpose-built cupboards to protect them from light exposure. Sample types included five interbreeding species, six hybrids, and five more distantly related 'out species' which were included in the analysis for the study of phylogeny, see Table 4.1. Species assignments were confirmed at the time of collection based on chromosome numbers (John Bailey, personal communication). For large leaves a four cm^2 square was cut out of each leaf between the second and third veins from the bottom left corner of each sample, to preserve the herbarium samples for future users. For small leaves where this was not possible, the whole leaf was taken.

Table 4.1: Species information for samples within the Polygonaceae family.

Latin Name	Contextual information
Reynoutria japonica var. *japonica*	• 'true' Japanese Knotweed • Western accessions are octoploid, 4n = 8x = 88 • Giant tetraploid plants are known in Japan, 2n = 4x = 44
Reynoutria japonica var. *compacta*	• 'dwarf' Japanese Knotweed • exclusively tetraploid, 2n = 4x = 44
Reynoutria japonica var. *uzenensis*	• 'hairy' Japanese Knotweed
Fallopia baldschuanica	• Russian vine
Reynoutria sachalinensis	• Giant Knotweed • predominantly tetraploid, 2n = 4 x = 44
Reynoutria x bohemica *OR* *Reynoutria japonica* var. *japonica x sachalinensis*	• Bohemian Knotweed (the most common hybrid) • 'True' Japanese Knotweed crossed with Giant Knotweed • - predominately hexaploid, 2n = 6x = 66
Reynoutria sachalinensis x Fallopia baldschuanica	• Giant Knotweed crossed with Russian Vine
Reynoutria japonica var. *japonica x Fallopia baldschuanica*	• 'True' Japanese Knotweed crossed with Russian Vine
Fallopia japonica var. *compacta x baldschuanica*	• dwarf variety of Japanese Knotweed crossed with Russian Vine
Reynoutria japonica var. *compacta x Reynoutria sachalinensis*	• dwarf variety of Japanese Knotweed crossed with Giant Knotweed
Fagopyrum esculentum	• Buckwheat
Rumex acetosella	• Sheep's Sorrel
Fallopia convolvulus	• Black-bindweed
Fallopia multiflora	• Tuber Fleece-flower
Fallopia cilinodis	• Fringed Bindweed

4.4.2 ATR-FTIR spectroscopy

Herbarium samples were analysed using a Tensor 27 FTIR spectrometer with a Helios ATR attachment (Bruker Optics Ltd, Coventry, UK). Ten spectra were taken from each leaf surface, resulting in twenty spectra per sample, 1580 spectra in total. A camera attachment was used to locate the area of interest and ensure an even spread of spectra across each surface for minimisation of bias. The ATR diamond crystal was cleaned between measurements of each leaf surface with wipes containing isopropyl alcohol (Bruker Optics, Coventry, UK), and background spectra were taken each time to account for ambient atmospheric conditions. Leaf material was placed on a slide with the analysed side facing upward, and then raised using a moving platform to make consistent contact with the Internal Reflection Element, a diamond crystal, defined the sampling area as 250 μm x 250 μm. Spectral resolution was 8 cm^{-1} with 2 times zero-filling, giving a data-spacing of 4 cm^{-1} over the range 4000 to 400 cm^{-1}; 32 co-additions and a mirror velocity of 2.2 kHz were used for optimum signal-to-noise ratio.

4.4.3 Spectral data handling and analysis

Acquired spectra were converted from OPUS format to .txt files before input to MATLAB (MathWorks, Natick, USA). Pre-processing of acquired spectra is essential for spectroscopic experiments to improve the signal-to-noise ratio, reduce spectral baseline distortions, and correct systematic variations in the absorbance intensity caused by different sample thickness (Butler *et al.*, 2018). Pre-processing and computational analysis of the data were performed using an in-house developed IRootLab toolbox (Martin FL *et al.*, 2010; Trevisan *et al.*, 2013) and the PLS Toolbox version 7.9.3 (Eigenvector Research, Inc., Manson, USA), according to standardised protocols for analysis of biochemical spectra (Morais *et al.*, 2019b, 2020). Spectra were cut at the biochemical fingerprint region (1800-900 cm^{-1}), Savitzky-Golay (SG) second differentiated, and vector normalised. All data were mean-centred before multivariate analysis.

The natural variation between samples was explored using the unsupervised technique, principal component analysis (PCA) (Cirino de Carvalho *et al.*, 2016a). For the classification of groups, PCA was followed by linear discriminant analysis (PCA-LDA; Morais & Lima, 2018) and the more complex technique, support vector machines (SVM)(Morais *et al.*, 2017). The classification hyperplane found by SVM provided the largest margin of separation between the data clusters. This was achieved using the most common kernel function, the radial basis function (RBF) (Cortes & Vapnik, 1995), that transformed the data into a different feature space during model construction. PCA-LDA was constructed using 10 principal components (PCs). The number of components of PCA-LDA and all SVM parameters were optimized by venetian blinds (10 data splits) cross-validation. Spectra were randomly divided into a training set (70%, 1106 spectra) and an external test set (30%, 474 spectra) to perform validation. For SVM parameters cost, gamma and number of support vectors see Supplementary Table S4.2. Increasing the cost and gamma values increases the complexity of the model. This makes the margin of separation between categories more specialised to the training data set, which results in fewer misclassifications but reduces the generalisation of the model.

The main spectral alterations were characterised with PCA loadings, for which peak maxima were identified with a peak-pick algorithm (20 cm^{-1} minimum separation). These spectral biomarkers were matched with previously characterised wavenumbers to give tentative chemical assignments. Further detail on biomarkers can be gained from comparisons of band intensity in the baseline corrected spectra (Supplementary Figure S4.1) and horizontal shifts in the vector normalised spectra, which indicate concentration and molecular structural alteration respectively.

Sample relationships were explored using the unsupervised pattern recognition method, hierarchical cluster analysis (HCA) and the resultant hierarchy was depicted in the form of a dendrogram (Figure 4.5a-c). Clustering was achieved using Euclidean distance as the metric of sample similarity and Ward's Method as the linkage criterion. HCA was used to analyse both the fingerprint region and the PCA loadings of the samples. Where genetic linkages were overridden by environmental factors, the wavenumbers highlighted by PCA loadings rather than the fingerprint

region were used for the HCA analysis. For the fingerprint region the spectra were first averaged by sample type, then pre-processed by SG differentiation, vector normalisation, before analysis by HCA. For the HCA based on the loading data, spectra were first pre-processed using second differentiation and vector normalisation. The absorbance values wavenumbers highlighted in the PCA loadings were selected, and spectra were averaged by sample type. HCA was then performed on the loadings and a dendrogram was produced.

4.5 Results

4.5.1 Sample types can be differentiated using ATR-FTIR spectroscopy and chemometrics

ATR-FTIR spectroscopy was used to explore the relationship between sample types (i.e. species, variety, hybrid, see Table 4.1), and to determine whether herbarium samples remain tractable to this type of analysis decades after collection. Figures 4.1a-b shows the raw and pre-processed spectra, where the mean spectra at the fingerprint region (1800-900 cm^{-1}) are grouped by sample type. Major change trends in intensity can be highlighted by visual comparison of averaged second-derivative spectra (Wang Q *et al.*, 2017). These spectra show clear differences; for example, some samples have peaks which others lack, or horizontal shifts visible at certain peaks.
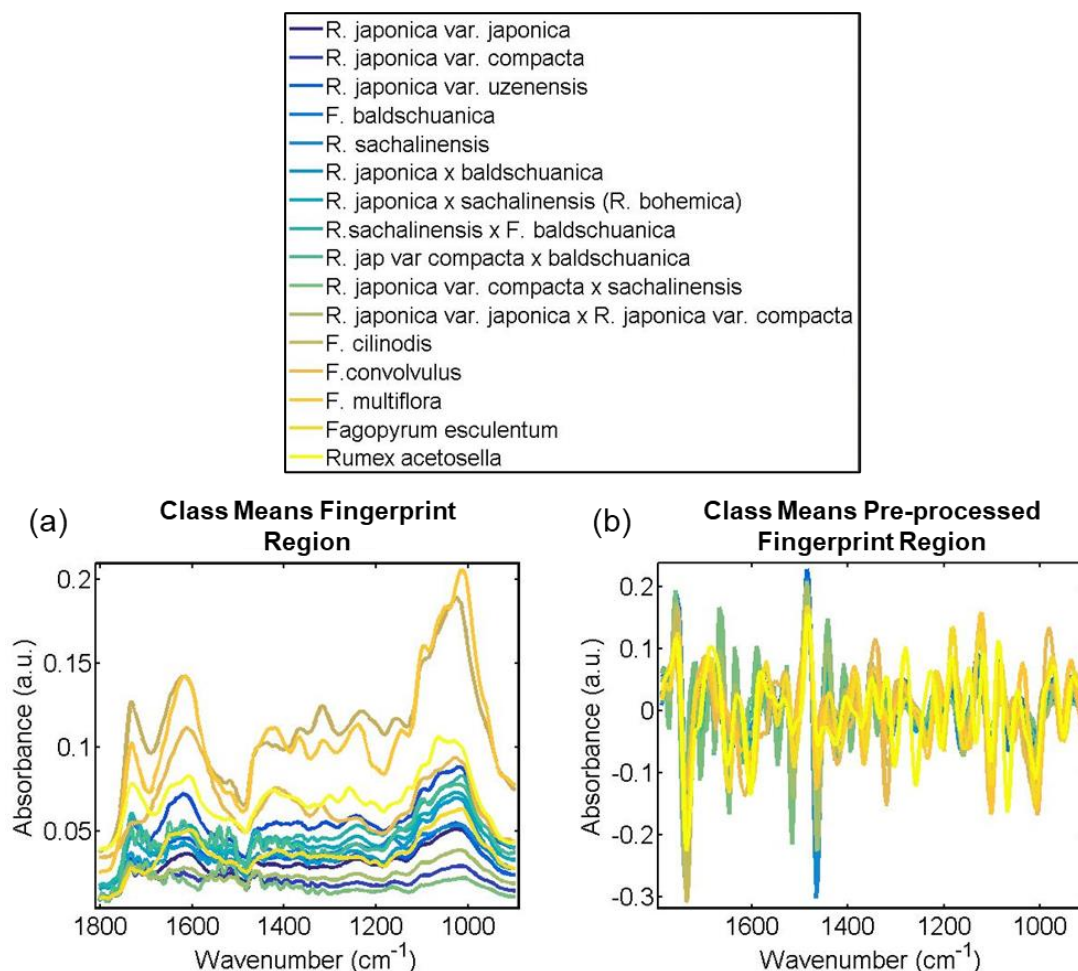
Figure 4.1: a) Raw and **b)** pre-processed class means IR-spectra for fingerprint region grouped by species. The pre-processing used for part b) was Savitzky-Golay (SG) second differentiation followed by vector normalisation.

Figure 4.2 shows that differences in the spectral fingerprint region (1800–900 cm^{-1}) of herbarium samples were sufficient to identify between sample types using PCA, PCA-LDA and SVM analyses. Importantly, the classification of samples by a combination of ATR-FTIR spectroscopy and chemometrics was consistent with chromosome counts performed at the time of collection (John Bailey, personal communication). Figure 4.2a shows the PCA scores which indicate the natural variation between samples. Although some clustering of spectra can be seen, a clear separation between samples was not observed in the scores on PC1 and PC2, which indicates high similarity between spectral profiles and the requirement for supervised analysis methods.

The predictive capability of PCA-LDA, as a supervised test was however also limited, with sample types overlapping and few distinct clusters (Figure 4.2b). Overall, PCA-LDA gave 86% accuracy, 52.30% sensitivity and 91.51% specificity (see Supplementary Table S4.3 for each type individually). Within this, the sensitivity scored 0% using PCA-LDA for two types: *Reynoutria japonica* var*. japonica x Reynoutria japonica* var*. compacta* and *Reynoutria japonica x baldschuanica.* In addition, of the *Reynoutria japonica* var*. japonica x Reynoutria japonica* var*. compacta* spectra, 75% were mistaken for *Reynoutria japonica x sachalinensis (Reynoutria x bohemica).*

For the plant type *Reynoutria japonica x baldschuanica*, spectra were acquired from only two samples (a plant artificially crossed at Leicester University by Dr. John Bailey and a plant collected from Cornwall) resulting in 40% of spectra being assigned incorrectly to *Reynoutria japonica* var*. compacta x sachalinensis*, another artificial hybrid. This suggests that environmental factors can outweigh the correct genetic assignment of samples when the sample numbers are too low to provide sufficient training data and highlights the need for caution if using PCA-LDA as a classification method for comparison of closely related hybrids and interbreeding species. In contrast, PCA-LDA achieved 100% specificity for the assignment of *Fallopia convolvulus, Fallopia multiflora,* and *Fagopyrum esculentum* using only one sample leaf each, which is likely a reflection of the distinct genetic nature, and therefore biochemical composition, of these species compared to the other plant types studied (Schuster *et al.*, 2015).
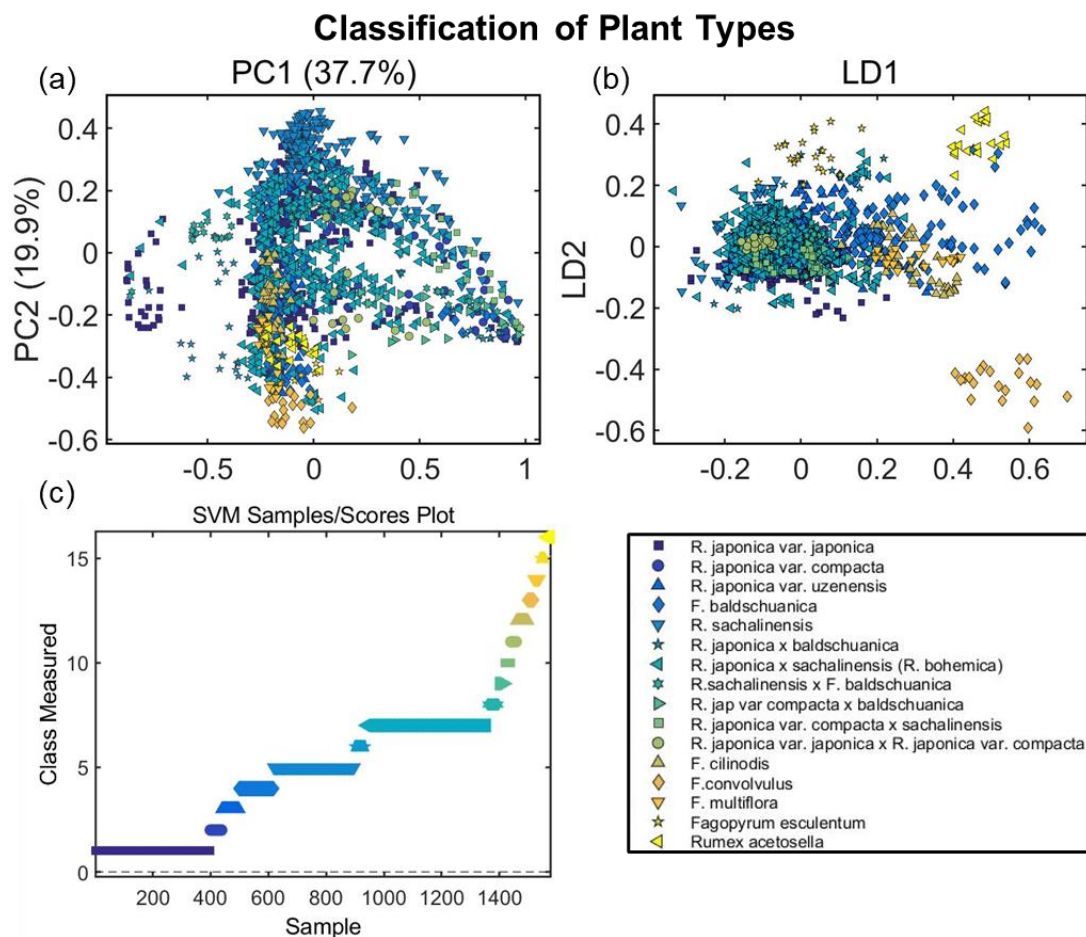
Figure 4.2: a) PCA**, b)** PCA-LDA and **c)** SVM of IR-spectra taken from both leaf surfaces for fingerprint region (1800-900 cm^{-1}) grouped by species, for all sixteen species with both sides of leaves included. Prior to multivariate analysis, the spectral fingerprint region was pre-processed using Savitzky-Golay (SG) second differentiation followed by vector normalisation and finally mean-centring.

SVM achieved excellent performance in both training (100 % accuracy) and test sets (99.04 % accuracy), successfully differentiating plant types based on their IR spectral profile (Figure 4.2c). SVM gave an average of 98.25% sensitivity and 98.32% specificity (see Table 4.2 for each type individually). However, focusing on the eleven most closely related species, then the average specificity increases to 100% for distinguishing between interbreeding species and hybrids of Japanese Knotweed, see Table S4.4. Unlike PCA-LDA, sensitivity remained high when using SVM despite a small available training set.

Table 4.2: Quality parameters (accuracy, sensitivity, and specificity) for spectral classification based on sample type of closely related species, hybrids, and varieties by SVM

SVM	% Accuracy	% Sensitivity	% Specificity
Reynoutria japonica var. *japonica*	98.61	98.49	98.50
Reynoutria japonica var. *compacta*	96.22	92.50	93.02
Reynoutria japonica var. *uzenensis*	100.00	100.00	100.00
Fallopia baldschuanica	99.13	98.32	98.35
Reynoutria sachalinensis	98.81	97.86	97.90
Reynoutria japonica x baldschuanica	97.40	94.87	95.12
Reynoutria japonica x sachalinensis (Reynoutria x bohemica)	98.32	97.15	97.21
Reynoutria sachalinensis x Fallopia baldschuanica	98.72	97.50	97.56
Reynoutria japonica var. *compacta x baldschuanica*	97.59	95.24	95.45
Reynoutria japonica var. *compacta x sachalinensis*	99.94	100.00	100.00
Reynoutria japonica var. *japonica x Reynoutria japonica* var. *compacta*	99.94	100.00	100.00
Fallopia cilinodis	100.00	100.00	100.00
Fallopia convolvulus	100.00	100.00	100.00
Fallopia multiflora	100.00	100.00	100.00
Fagopyrum esculentum	100.00	100.00	100.00
Rumex acetosella	100.00	100.00	100.00
Average	**99.04**	**98.25**	**98.32**

4.5.2 ATR-FTIR spectroscopy can distinguish between leaf surfaces in samples from the herbarium

PCA was used to explore natural differences between adaxial (upper) and abaxial (lower) leaf surfaces. Clustering of spectra from upper and lower leaf surfaces can be seen in the PCA scores (Supplementary Figure S4.2). PCA-LDA was constructed using 10 PCs, differentiating spectra from the upper and lower surfaces along axis LD1 of 2D scatterplot, see Supplementary Figure S4.2c. Classification performance of PCA-LDA was good with 74.1% accuracy, 72.3% sensitivity, and 75.9% specificity for the upper leaf surface, and 74.1% accuracy, 75.9% sensitivity, and 72.3% specificity for the lower leaf surface. SVM performed even better achieving a clear distinction between upper and lower surfaces. SVM test sets achieved an average of 98.4% accuracy, 98.4% sensitivity, and 98.4% specificity, see Supplementary Figure S4.2d.

4.5.3 Herbarium samples can be classified by their geographical origin using ATR-FTIR spectroscopy
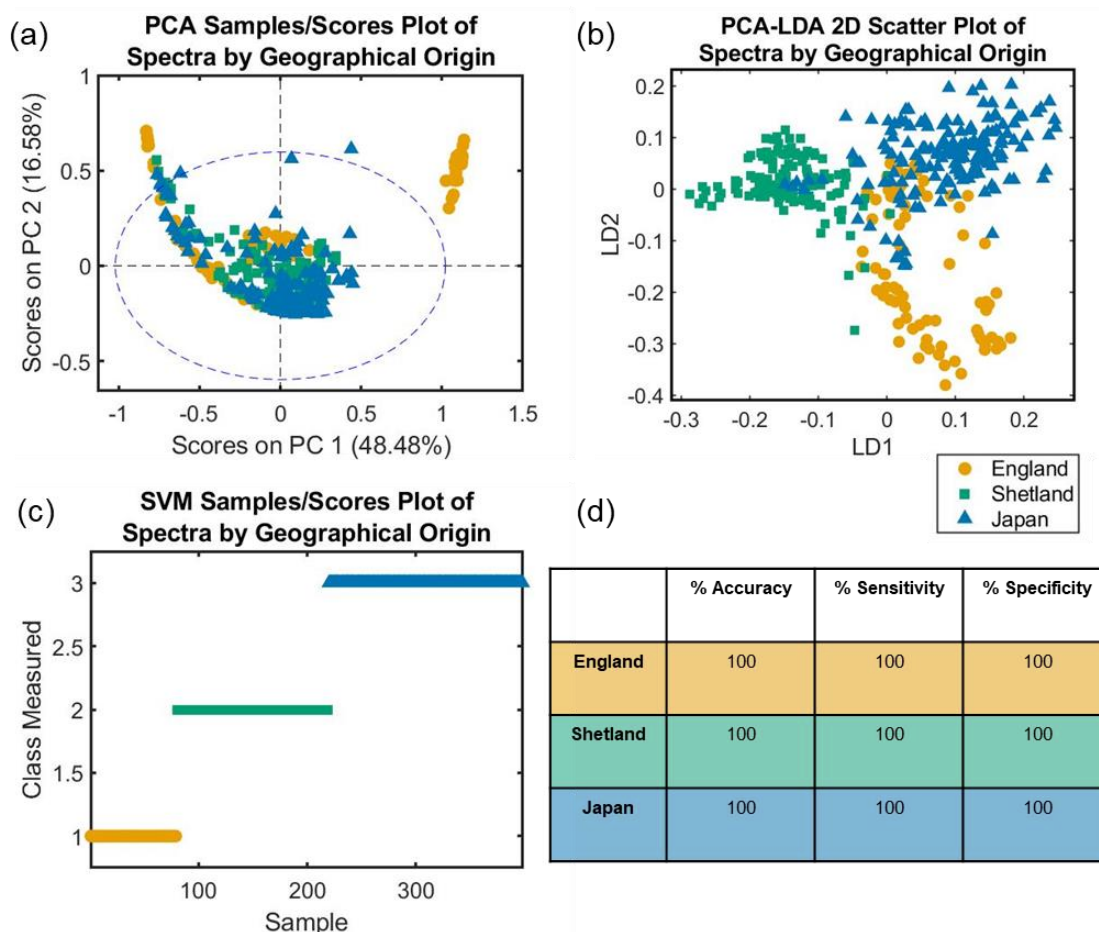
Figure 4.3: a) PCA scores plot, **b)** LDA 2D scatter plot, **c)** SVM scores plot and **d)** SVM classification table of fingerprint spectra (1800-900 cm^{-1}) grouped by geographical origin of *Reynoutira japonica* var. *japonica* samples: England (orange), Shetland (green) and Japan (blue). Prior to multivariate analysis, the spectal fingerprint region was pre-processed using Savitzky-Golay (SG) second differentiation followed by vector normalisation and finally mean-centring.

The material analysed in this study were of Japanese Knotweed (*Reynoutria japonica* var. *japonica*) collected from the Scottish island of Shetland, and various locations across England and Japan (see S5 Table). PCA was used to explore natural differences in plants between regions clustering of spectra into samples from England, Shetland, and Japan (Figure 4.3a). PCA-LDA resulted in an average 87.85% accuracy, 80.19% sensitivity, and 90.20% specificity, see S5 Table. Good separation was observed in the 2D scatterplot of spectra along the axes LD1 and LD2, revealing differences between spectra from different geographical locations (Figure 4.3b).

Spectra from English samples separated from the other two regions along the LD2 axis. Spectra from Shetlandic samples separated from the other two along the axis LD1. SVM performed the best, creating a clear distinction between the three locations (Fig3c), and achieving 100% in accuracy, sensitivity, and specificity (Figure 4.3d).

4.5.4 Chemometric analysis of spectral data highlights differences between dwarf and invasive knotweed varieties
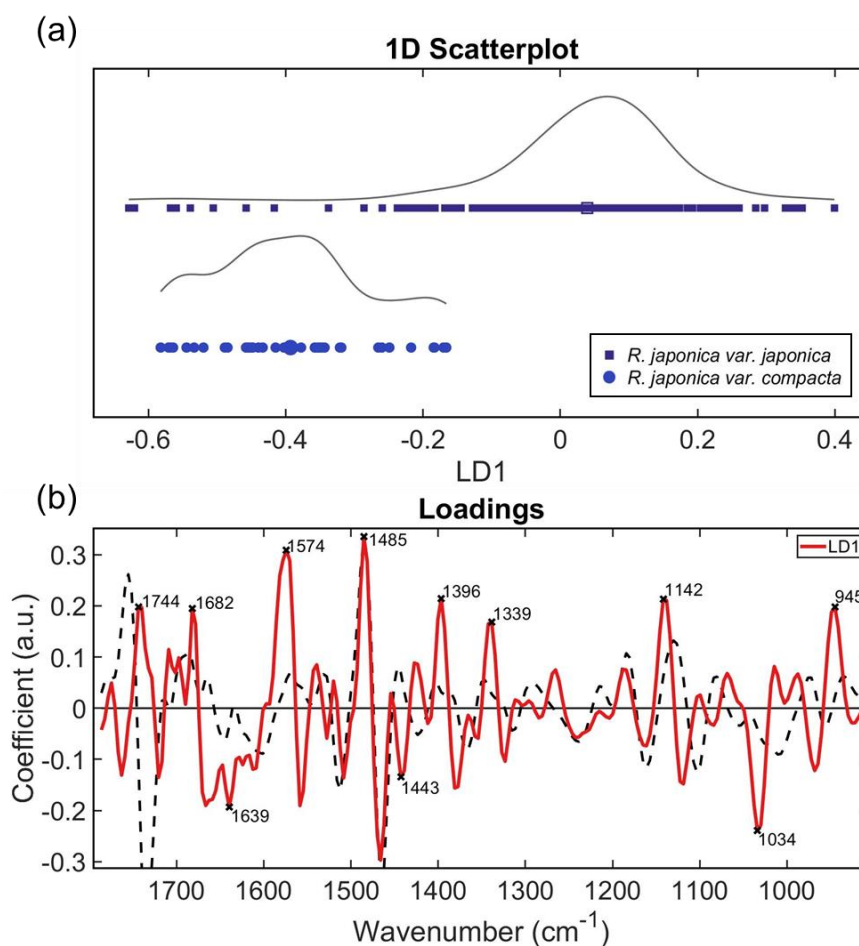
Figure 4.4: a) PCA-LDA **b)** loadings for *Reynoutria japonica* var. *japonica* vs *Reynoutria japonica* var. *compacta*. Figure 4.4b depicts the PCA loadings in red, and the total mean spectrum as the black dashed line, scaled to fit. Prior to multivariate analysis, the spectral fingerprint region was pre-processed using Savitzky-Golay (SG) second differentiation followed by vector normalisation and finally mean-centring.

Table 4.3: Main wavenumbers responsible for class differentiation between the highly invasive *Reynoutria japonica* var. *japonica* and its more easily controllable counterpart *Reynoutria japonica* var. *compacta*, and their assigned biomarkers.

Wavenumber/ cm	Assignment	Reference
1744	Ester carbonyl group C = O of triglycerides	(Raba *et al.*, 2015)
1682	Succinic acid (in pure solid form) Amide I, β-turns	(Kang *et al.*, 2008; Di Giambattista *et al.*, 2011)
1639	Amide I	(Talari *et al.*, 2017)
1574	C=N adenine	(Talari *et al.*, 2017)
1485	C_8-H coupled with a ring vibration of guanine	(Talari *et al.*, 2017)
1443	δ(CH) of pectin	(Sharma & Uttam, 2018)
1396	Symmetric CH_3 bending of the methyl groups of proteins	(Talari *et al.*, 2017)
1339	In-plane C-O stretching vibration combined with the ring stretch of phenyl	(Talari *et al.*, 2017)
1142	Phosphate and oligosaccharides; oligosaccharide C–O bond in hydroxyl group might interact with some other membrane components	(Talari *et al.*, 2017)
1034	Glucomannan	(Kanter *et al.*, 2013)
945	Xyloglucan	(Kacuráková *et al.*, 2000)

Molecular differences between two varieties, *Reynoutria japonica* var. *japonica* and *Reynoutria japonica* var. *compacta* were examined using PCA-LDA analysis (Figure 4.4a). The latter variety is intriguingly non-invasive and rarely naturalises (Bailey *et al.*, 2009). PCA loadings (Figure 4.4b) were subsequently used for the identification of biomarkers (Table 4.3). The PCA-LDA distributions of the two species differ in shape, with the mean clearly separated along the axis LD1 (Figure 4.4a). The following biomarkers were present at higher concentrations in *compacta* than *japonica*: 1744 cm^{-1} (ester carbonyl group of triglycerides; (Raba *et al.*, 2015), 1682 cm^{-1} (succinic acid; (Kang *et al.*, 2008) or the β-turns of Amide I; (Di Giambattista *et al.*, 2011), and 1485 cm^{-1} (C$_8$-H coupled with a ring vibration of guanine (Talari *et al.*, 2017). In contrast, the following biomarkers were present at higher concentrations in *japonica* than *compacta*: 1639 cm^{-1} (Amide I (Talari *et al.*, 2017), 1443 cm^{-1} (δ(CH) of pectin)(Sharma & Uttam, 2018), 1396 cm^{-1} (symmetric CH$_3$ bending of the methyl groups of proteins) (Talari *et al.*, 2017), 1339 cm^{-1} (in-plane C-O stretching vibration combined with the ring stretch of phenyl (Talari *et al.*, 2017), 1142 cm^{-1} (phosphate and oligosaccharides) (Talari *et al.*, 2017), 1034 cm^{-1} (glucomannan) (Kanter *et al.*, 2013), and 945 cm^{-1} (xyloglucan) (Kacuráková *et al.*, 2000). Particularly noticeable was the peak at 1034 cm^{-1} corresponding to the polysaccharide, glucomannan(Kanter *et al.*, 2013);which was two-fold higher in the invasive variety. In addition, horizontal shifts indicative of structural changes were present in the following biomarkers: 1744 cm^{-1} (ester carbonyl group of triglycerides) (Raba *et al.*, 2015), 1443 cm^{-1} (δ(CH) of pectin) (Sharma & Uttam, 2018), 1142 cm^{-1} (phosphate and oligosaccharides) (Talari *et al.*, 2017), 945 cm^{-1} (xyloglucan) (Kacuráková *et al.*, 2000).

4.5.5 ATR-FTIR spectroscopy offers a rapid alternative to genetic methods of phylogenetic analysis
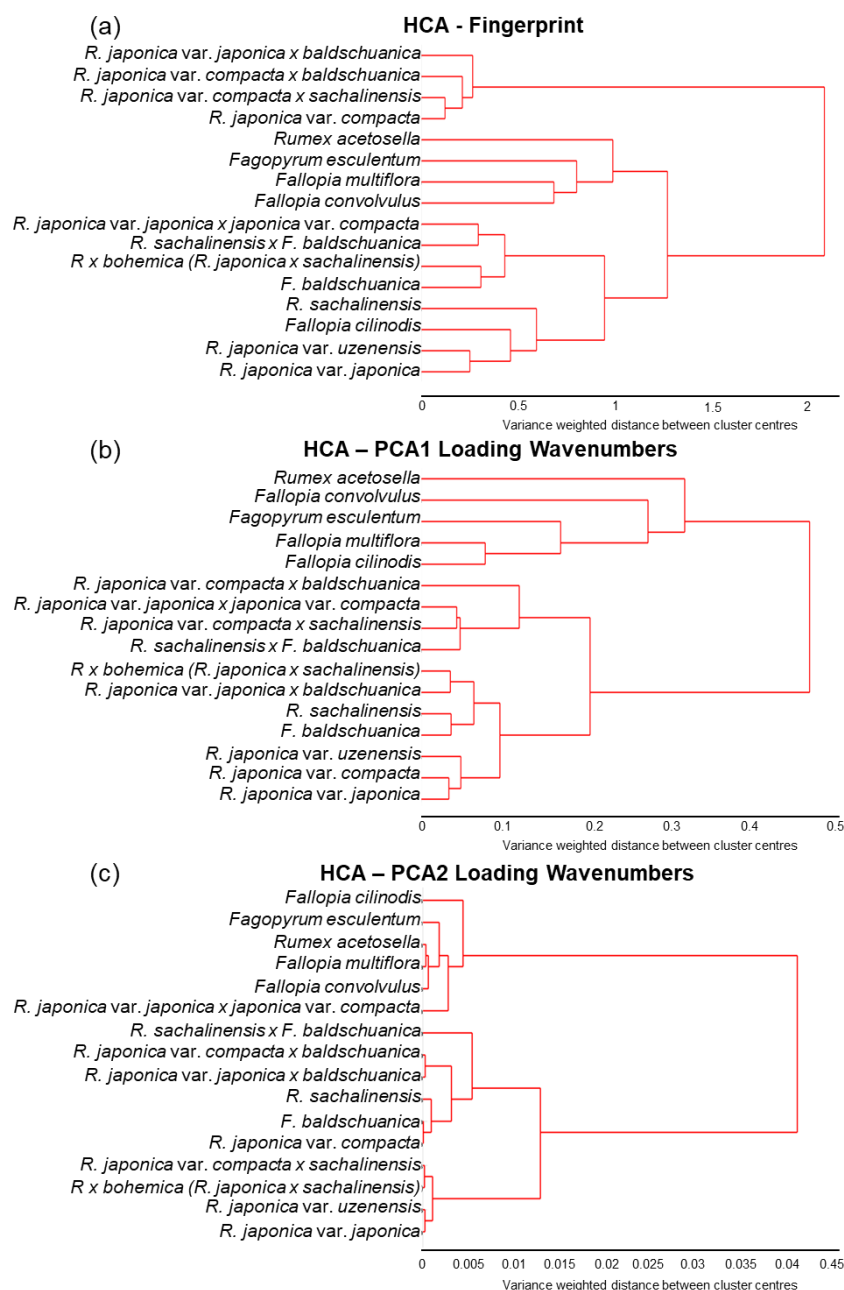
Figure 4.5: Hierarchical cluster analysis dendrogram results based on the Euclidean distance and Ward's method, comparison based on **a)** spectral fingerprint region (1800-900 cm^{-1}), or the wavenumbers from the **b)** PC1 and **c)** PC2 loadings. Prior to multivariate analysis, the spectal fingerprint region in part a) was pre-processed using Savitzky-Golay (SG) second differentiation followed by vector normalisation and finally mean-centring.

To determine the potential of ATR-FTIR spectroscopy and multivariate analysis for phylogenetic studies, and specifically within the Polygonaceae family, the

relationship between all sixteen plant types studied within this family was considered (Figures 4.5a-c). The resultant linkages derived from HCA analysis differed depending on the kinds of spectral information compared: the fingerprint region (Figure 4.5a); the wavenumbers from the PC1 and PC2 loadings (Figure 4.5b); or for loadings graphs (Supplementary Figure S4.3).

Using the fingerprint region, environmental growth conditions were the dominant influence over the relationship (Figure 4.5a). The most distantly related group in this linkage comprised four species grown in 'captivity': *Reynoutria japonica* var. compacta grown in Cambridge Botanic Gardens, and three artificial hybrids (cultivated by hand-pollination) grown at University of Leicester (UK), *Reynoutria japonica* var. *compacta x sachalinensis, Reynoutria japonica* var. *compacta x baldschuanica* and *Fallopia japonica* var. *japonica x baldschuanica.* Within this group, the parent dwarf variety and the artificial dwarf hybrids are genetically similar, however they group further away from *Reynoutria japonica* var. *japonica* than expected. The next cluster contained four out of the five 'out species', with the more distantly related *Rumex* and *Fagopyrum* genera placing further away than the genetically closer *Fallopia* species. Surprisingly, the fifth 'out species', *Fallopia cilinodis*, was placed within the *Reynoutria* species.

In contrast, when the wavenumbers from the PC1 loadings were selected for HCA (Figure 4.5b), the results more closely matched the genetic relationship. Two main clusters separated the 'out-species' from Japanese Knotweed *s.l.*. Within the knotweed group, artificial greenhouse grown hybrids group together. The parental species which can interbreed with Japanese Knotweed, *Reynoutria sachalinensis* and *Fallopia baldschuanica*, group together with their *Reynoutria japonica* hybrids, including the key naturally occurring hybrid *Reynoutria x bohemica*. All three varieties of Japanese Knotweed then grouped together in a close cluster.

When HCA was performed using the wavenumbers from the PC2 loadings (Figure 4.5c), the order of the out species was no longer reflective of genetic relationship with *Fallopia cilinodis* grouping furthest away despite its *Fallopia* genus and the *Reynoutria. japonica* var. *japonica x Reynoutria japonica* var. *compacta* unexpectedly grouping with the more genetically distant species. However, the artificial hybrids

grouped together with parent species *Fallopia baldschuanica*, *Reynoutria sachalinensis*, and *Reynoutria japonica* var. *compacta* despite being artificial and greenhouse grown. The *compacta* and *japonica* knotweed varieties crossed with *Reynoutria sachalinensis* grouped together as genetically similar and placed closer to the *Reynoutria japonica* parent than *Reynoutria sachalinensis*.

4.6 Discussion

4.6.1 ATR-FTIR spectroscopy combined with chemometrics as a tool for IAS identification

Our results show that ATR-FTIR spectroscopy of historic herbarium samples followed by analysis with SVM is able to effectively differentiate between plant type, even closely related hybrids, achieving 99% accuracy. This result opens the possibility of applying this method to historic samples to validate the species or hybrid assignment and conclusions drawn from previous studies, which have lacked cytological confirmation (Gillies *et al.*, 2016). This also suggests that this approach offers a solution to the misidentification and underestimation of hybridisation caused by morphological similarities within Japanese Knotweed *s.l.* (Gillies *et al.*, 2016) which has previously complicated management strategies (Moody & Les, 2007). Additionally, it raises the intriguing possibility of being able to distinguish between, or confirm the identity and sources of, phenotypically identical knotweed populations when apportioning responsibility for the damage caused by this species to property, which is a particular issue in the UK (Santo, 2017).

Interestingly, both PCA-LDA and SVM successfully differentiated between leaf surfaces. The observed differences (Supplementary Figure S4.2) are consistent with the presence of trichomes on the abaxial epidermis, differences which are also used to visually categorise hybrids and species within Japanese Knotweed *s.l.* (Bailey *et al.*, 2009), and the different functions performed by the two surfaces, with the upper (adaxial) surface primarily acting to conserve water and the lower (abaxial) surface more commonly involved in gas exchange (Taiz *et al.*, 2015). Water conservation is often enhanced by a thicker waxy cuticle on adaxial leaf surfaces, comprising different epicuticular waxes: adaxial containing primary alcohols and esters, and

waxes on the abaxial surface comprised of alkanes, aldehydes and secondary alcohols (Gniwotta *et al.*, 2005; Ringelmann *et al.*, 2009). Leaf surfaces have differing responses to the environment. For example epidermal wax composition alters susceptibility to fungal pathogens (Gniwotta *et al.*, 2005) and differing abaxial and adaxial epidermal Ca^{2+} concentrations affect stomatal guard cell sensitivity (De Silva *et al.*, 1986). Therefore, the choice of leaf surface should be a consideration for method design if applied to monitoring of disease by fungal pathogens, for example in the context of biocontrol agents. Variations in epidermal thickness may also play a role in species classification. The infrared light used for ATR-FTIR spectroscopy only penetrates ~0.5-2 μm deep into the sample, compared with the average plant cuticle thickness of ~1-10 μm; this may allow spectral acquisition of different compounds from deeper within the leaf in species with thinner cuticles. For example, *Reynoutria sachalinensis* leaves are thinner overall with a thicker cuticular layer than those of *Reynoutria japonica*, see Supplementary Figure S4.4 for scanning electron microscope images.

4.6.2 ATR-FTIR spectroscopy captures environmental information

Spectra of *Reynoutria japonica* var*. japonica* from England, Japan and the Scottish isle of Shetland were clearly distinguished by location using the SVM algorithm (Figure 4.3c), achieving 100% in accuracy, sensitivity, and specificity (Figure 4.3d). Japanese samples were likely genetically diverse as these plants can reproduce sexually within their native range and adapt to different environmental locations (Bailey, 2003). However, British specimens are believed to all originate from the female clone introduced in 1850 by Philipp von Siebold (Bailey & Conolly, 2000). Nevertheless, Shetlandic samples were distinct and effectively separated from both those from England and Japan, suggesting a strong environmental influence on spectra.

The differences detected using this approach between putatively genetically identical clonal plants from different geographical locations could be due, at least in part, to phenotypic plasticity, where one genotype can express different phenotypes (Pichancourt & van Klinken, 2012). Phenotypic plasticity is a result of environment-

genotype interactions. This is a particularly significant mechanism for invasive plants, a high percentage of which are clonal, such as Japanese Knotweed (Mounger *et al.*, 2021). Epigenetic modifications, somatic mutations, resource provisioning and biochemical functioning may contribute to the phenotypic plasticity allowing successful invasion of Japanese Knotweed in a diverse range of habitats (Richards *et al.*, 2008, 2012; Zhang *et al.*, 2016), particularly as clonal plants are able to bypass the meiotic resetting of epigenetic modifications through asexual reproduction (Mounger *et al.*, 2021).

ATR-FTIR spectroscopy examines the biochemical fingerprint of the leaf surface and environmental effects are reflected within the spectra, for example, the grouping of artificial greenhouse grown hybrids together in Figure 4.5a-c. Intriguingly, when spectra were grouped by individual sample rather than sample type and fingerprint regions were compared by HCA, a mixture of genetic and environmental factors appeared to influence the results. Samples group into clusters based on their geographical source or their genetic species class (Supplementary Figure S4.5). This observation has potential applications in the context of biocontrol of IAS. Biocontrol agents must be carefully selected for maximum efficacy, in the case of Japanese Knotweed taking into account the host preferences of different psyllid biotypes (Jones IM *et al.*, 2020). It is therefore important, when deciding from where to obtain biocontrol agents, to have access to combined genetic and geographical information about host plants in their native range (Pashley, 2003). In this context the weaving together of genetic and environmental information could prove an invaluable biogeographical tool.

In the present study, the fingerprint region (1800–900 cm^{-1}) of acquired spectra was selected for chemometric analysis because biological molecules preferentially absorb light of these wavenumbers, including important biological absorptions due to lipids, proteins, carbohydrates, nucleic acids and protein phosphorylation (Morais *et al.*, 2020). Isolation of this fingerprint region (1800–900 cm^{-1}) has achieved good results in other plant studies (Butler *et al.*, 2015c; Ord *et al.*, 2016; Skolik *et al.*, 2019a,b; Rahman *et al.*, 2019; Holden *et al.*, 2021), though the high region (3700-2800 cm^{-1}) has also yielded valuable information in a range of

applications (Du *et al.*, 2013; Minnes *et al.*, 2017; Festa *et al.*, 2019; Liu N *et al.*, 2020) since it contains additional biologically relevant absorbances such as those for water (~3275 cm^{-1}), protein (~3132 cm^{-1}), fatty acids and lipids (~3005, ~2970, ~2942 and ~2855 cm^{-1}) (Morais *et al.*, 2020).

4.6.3 Chemometric analysis of spectral data provides insights into why the dwarf variety of knotweed is less invasive than 'true' knotweed

Unexpectedly, two compound types involved with energy production, triglycerides and succinate, were higher in concentration in the dwarf variety of knotweed than the invasive variety (Figure 4.4). In vegetative tissues, triacylglycerol metabolism is used as an energy source for cell division and expansion, stomatal opening, and membrane lipid remodelling (Yang & Benning, 2018), whilst succinate is involved in the production of ATP and acts as a signalling hub (Tretter *et al.*, 2016). One explanation for this observation might be the differential effects of environmental conditions, time of day, or age of leaf of the sample at the time of collection (Bielczynski *et al.*, 2017). The structural differences in triglycerides, as indicated by a horizontal spectral shift, is consistent with the knowledge of their construction. Triglycerides are tri-esters comprised of a glycerol bound to three fatty acid molecules, the R-groups of which can vary in chirality and composition (Long *et al.*, 2021). Further work using samples from plants grown under controlled conditions is therefore required to investigate this observation before definitive conclusion can be drawn.

Conversely, pectin (a compound which strengthens the plant cell wall), oligosaccharides and two polysaccharides (glucomannan and xyloglucan) were present in higher concentrations in the invasive Japanese Knotweed variety compared with the dwarf variety (Figure 4.4). Interestingly, the C–O bond in the hydroxyl group responsible for the oligosaccharide wavenumber biomarker might interact with other membrane components (Talari *et al.*, 2017), leading to the indicated structural change. Plant oligosaccharides, which include fructans and raffinose family oligosaccharides (RFOs), act as important multifunctional compounds. They can act as reserve carbohydrates, membrane stabilizers and stress tolerance mediators, play a role in osmoregulation and source–sink

relationships, contribute to overall cellular reactive oxygen species (ROS) homeostasis by specific ROS scavenging, and act as phloem-mobile signalling compounds under stress (Van den Ende, 2013). In addition, the observed structural change in xyloglucan is consistent with the known species-dependent branching pattern of this molecule (Nishinari *et al.*, 2007). Xyloglucan is the most abundant hemicellulosic polysaccharide in the primary cell wall of most vascular plants, and together with cellulose gives the wall its strength. Glucomannan and xyloglucan act as storage polysaccharides in tubers and seeds. The presence of higher levels in invasive 'true' knotweed than in its dwarf counterpart raises the intriguing possibility that these polysaccharides could contribute to the comparatively rapid growth of *Reynoutria japonica* var*. japonica.*

4.6.4 HCA analysis of PCA loadings offers a rapid alternative to traditional phylogenetic analysis

The power of this technique has already been shown for phylogenetic studies in some flowering plants (Kim *et al.*, 2004) and agronomically important species such as wheat (Demir *et al.*, 2015). The results of the present study show potential for application to invasive species. ATR-FTIR spectroscopy combined with HCA analysis shows potential as a complementary technique alongside genetic methods to explore phylogeny and biogeographic relationships, without prior sequence knowledge. The HCA dendrogram shown in Figure 4.5b, where the PC1 loadings were used as the input, closely followed the expected phylogenetic relationship based on what is known of the genetics and phylogenies of this complex (Pashley, 2003; Desjardins, 2015). *Compacta* and *japonica*, varieties of the same species (Bailey *et al.*, 2009), were paired together. Both of the species *Reynoutria japonica*, these plants are morphologically similar and often confused (Mandák *et al.*, 2003), consistent with the close linkage. The interbreeding species and hybrids were present in a different cluster to the five more distantly related species from within the Polygonaceae family. Although of the *Fallopia* genus, *Fallopia cilinodis* places separately to other *Fallopia* and *Reynoutria* species in a previous likelihood tree produced using a molecular dataset (nrITS, matK, trnL-trnF) (Schuster *et al.*, 2015).

Fallopia cilinodis' placement in a cluster separate to the interbreeding species is also consistent with molecular studies (Schuster *et al.*, 2015).

The placement of *Rumex acetosella* as the furthest from *Reynoutria japonica* var*. japonica* is unexpected. Molecular studies place *Rumex* closer than *Fagopyrum* (Schuster *et al.*, 2015). Additionally, two potential biocontrol agents, the knotweed sawfly (*Allantus luctifer*) and a leaf beetle (*Gallerucida bifasciata* Motschulsky), were ruled out as candidates after host range testing confirmed it would feed on various native UK *Rumex* species (Grevstad *et al.*, 2018). This may be indicative of a similar composition of secondary metabolites to the target plant, Japanese Knotweed.

Importantly the hierarchical cluster analysis dendrograms in Figure 4.5 show that environmental factors can play a role in the determined linkages. Therefore, caution must be taken when using this method as a tool for phylogeny. Although the present study clearly demonstrates the power of this approach for the analysis of historic herbarium samples, fresh samples taken from plants grown together under controlled conditions prior to spectral acquisition would remove 'environment' as a variable to allow an unfettered comparison between plant types.

4.7 Conclusion

Outcomes from this study show ATR-FTIR spectroscopy coupled with SVM can accurately differentiate between leaf surfaces, plant types, and samples from different geographical locations even in herbarium samples of varying age of closely related species within the Polygonaceae family. This provides a rapid and robust method for hybrid identification, allowing informed decisions to be made regarding targeted control measures to tackle this invasive alien weed. This could be applied in the field using handheld mid and near-infra red devices (Tiwari *et al.*, 2013; Huang *et al.*, 2018). Additionally, spectra from the invasive *Reynoutria japonica* var. *japonica* knotweed variety indicated the presence of two polysaccharides, glucomannan and xyloglucan, in higher concentrations than in the dwarf variety. Results demonstrated that ATR-FTIR spectroscopy and hierarchical cluster analysis provides an additional methodology for investigating linkage between closely

related species. Adoption of this technology for the study of historic samples would increase the value of existing herbarium collections, which are currently threatened.

4.8 References

Armstrong K., Ball S. 2005. DNA barcodes for biosecurity: invasive species identification. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**: 1813–1823.

Bağcıoğlu M, Kohler A, Seifert S, Kneipp J, Zimmermann B. **2017**. Monitoring of plant–environment interactions by high-throughput FTIR spectroscopy of pollen. *Methods in Ecology and Evolution* **8**: 870–880.

Bailey J. **2003**. Japanese Knotweed s.l. at home and abroad. *Plant invasions: ecological threats and management solutions*: 183–196.

Bailey JP, Bímová K, Mandák B. **2009**. Asexual spread versus sexual reproduction and evolution in Japanese Knotweed s.l. sets the stage for the 'battle of the Clones'. *Biological Invasions*. **11(5)**: 1189-1203.

Bailey JP, Conolly AP. **2000**. Prize-winners to pariahs -A history of Japanese Knotweed s.l. (Polygonaceae) in the British Isles. *Watsonia* **23**: 93–110.

Ben-Ari G, Lavi U. **2012**. Marker-assisted selection in plant breeding. In: Plant Biotechnology and Agriculture. Elsevier Inc., 163–184.

Bielczynski LW, Łącki MK, Hoefnagels I, Gambin A, Croce R. **2017**. Leaf and Plant Age Affects Photosynthetic Performance and Photoprotective Capacity. *Plant Physiology* **175**: 1634.

Buitrago MF, Groen TA, Hecker CA, Skidmore AK. **2016**. Changes in thermal infrared spectra of plants caused by temperature and water stress. *ISPRS Journal of Photogrammetry and Remote Sensing* **111**: 22–31.

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

Butler HJ, Martin FL, Roberts MR, Adams S, McAinsh MR. **2020**. Observation of nutrient uptake at the adaxial surface of leaves of tomato (*Solanum lycopersicum*) using Raman spectroscopy. *Analytical Letters* **53**: 536–562.

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

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

Butler HJ, Smith BR, Fritzsch R, Radhakrishnan P, Palmer DS, Baker MJ. **2018**. Optimised spectral pre-processing for discrimination of biofluids via ATR-FTIR spectroscopy. *Analyst* **143**: 6121–6134.

Cirino de Carvalho L, de Lelis Medeiros de Morais C, Gomes de Lima KM, Cunha Júnior LC, Martins Nascimento PA, Bosco de Faria J, Henrique de Almeida Teixeira G. **2016**. Determination of the geographical origin and ethanol content of Brazilian sugarcane spirit using near-infrared spectroscopy coupled with discriminant analysis. *Analytical Methods* **8**: 5658–5666.

Cortes C, Vapnik V. **1995**. Support-vector networks. *Machine learning* **20**: 273–297.

Demir P, Onde S, Severcan F. **2015**. Phylogeny of cultivated and wild wheat species using ATR-FTIR spectroscopy. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy* **135**: 757–763.

Desjardins SD. **2015**. Evolutionary studies in subtribe Reynoutriineae (Polygonaceae). With contributions to the study of hybridisation in Helosciadium and Berula (Apiaceae) included as appendices. *Doctoral dissertation, University of Leicester*

Du C, Ma Z, Zhou J, Goyne KW. **2013**. Application of mid-infrared photoacoustic spectroscopy in monitoring carbonate content in soils. *Sensors and Actuators B: Chemical* **188**: 1167–1175.

Ellstrand NC, Schierenbeck KA. **2000**. Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences of the United States of America* **97**: 7043–7050.

Van den Ende W. **2013**. Multifunctional fructans and raffinose family oligosaccharides. *Frontiers in Plant Science* **4**: 247.

Euring D, Löfke C, Teichmann T, Polle A. **2012**. Nitrogen fertilization has differential effects on N allocation and lignin in two Populus species with contrasting ecology. *Trees - Structure and Function* **26**: 1933–1942.

Fennell M, Wade M, Bacon KL. **2018**. Japanese knotweed (*Fallopia japonica*): an analysis of capacity to cause structural damage (compared to other plants) and typical rhizome extension. *PeerJ* **6**: e5246.

Festa G, Andreani C, Baldoni M, Cipollari V, Martínez-Labarga C, Martini F, Rickards O, Rolfo MF, Sarti L, Volante N, *et al.* 2019. First analysis of ancient burned human skeletal remains probed by neutron and optical vibrational spectroscopy. *Science Advances* **5(6)**: 1292.

Di Giambattista L, Pozzi D, Grimaldi P, Gaudenzi S, Morrone S, Castellano AC. **2011**. New marker of tumor cell death revealed by ATR-FTIR spectroscopy. *Analytical and Bioanalytical Chemistry* **399**: 2771–2778.

Gillies S, Clements DR, Grenz J. **2016**. Knotweed (Fallopia spp.) Invasion of North America Utilizes Hybridization, Epigenetics, Seed Dispersal (Unexpectedly), and an Arsenal of Physiological Tactics. *Invasive Plant Science and Management* **9**: 71–80.

Gniwotta F, Vogg G, Gartmann V, Carver TLW, Riederer M, Jetter R. **2005**. What Do Microbes Encounter at the Plant Surface? Chemical Composition of Pea Leaf Cuticular Waxes 1. *Plant Physiology* **139(1)**: 519-530.

Grevstad FS, Andreas JE, Bourchier RS, Shaw R, Winston RL, Randall CB. **2018**. Biology and biological control of knotweeds. *United States Department of Agriculture, Forest Health Assessment and Applied Sciences Team.*

Grimsby JL, Tsirelson D, Gammon MA, Kesseli R. **2007**. Genetic diversity and clonal vs. sexual reproduction in Fallopia spp. (Polygonaceae). *American Journal of Botany* **94**: 957–964.

Holden CA, Morais CLM, Taylor JE, Martin FL, Beckett P, McAinsh M. **2021**. Regional differences in clonal Japanese knotweed revealed by chemometrics-linked attenuated total reflection Fourier-transform infrared spectroscopy. *BMC Plant Biology 2021 21:1* **21**: 1–20.

Hollingsworth ML, Bailey JP. **2000**. Evidence for massive clonal growth in the invasive weed *Fallopia japonica* (Japanese Knotweed). *Botanical Journal of the Linnean Society* **133**: 463–472.

Hollingsworth ML, Hollingsworth PM, Jenkins GI, Bailey JP, Ferris C. **1998**. The use of molecular markers to study patterns of genotypic diversity in some invasive alien Fallopia spp. (Polygonaceae). *Molecular Ecology* **7**: 1681–1691.

Huang Y, Lu R, Chen K. **2018**. Prediction of firmness parameters of tomatoes by portable visible and near-infrared spectroscopy. *Journal of Food Engineering* **222**: 185–198.

Hulme PE. **2009**. Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology* **46**: 10–18.

Inamura A, Ohashi Y, Sato E, Yoda Y, Masuzawa T, Ito M, Yoshinaga K. **2000**. lntraspecific Sequence Variation of Chloroplast DNA Reflecting Variety and Geographical Distribution of *Polygonurn cuspidatum* (Polygonaceae) in Japan. *Journal of Plant Research* **113(4)**: 419-426.

Jones D, Bruce G, Fowler MS, Law-Cooper R, Graham I, Abel A, Street-Perrott FA, Eastwood D. **2018**. Optimising physiochemical control of invasive Japanese knotweed. *Biological Invasions* **20**: 2091–2105.

Jones IM, Smith SM, Bourchier RS. **2020**. Establishment of the biological control agent Aphalara itadori is limited by native predators and foliage age. *Journal of Applied Entomology* **144**: 710–718.

Kacuráková M, Capek P, Sasinková V, Wellner N, Ebringerová A. **2000**. FT-IR study of plant cell wall model compounds: Pectic polysaccharides and hemicelluloses. *Carbohydrate Polymers* **43**: 195–203.

Kang S, Amarasiriwardena D, Xing B. **2008**. Effect of dehydration on dicarboxylic acid coordination at goethite/water interface. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **318**: 275–284.

Kanter U, Heller W, Durner J, Winkler JB, Engel M, Behrendt H, Holzinger A, Braun P, Hauser M, Ferreira F, *et al.* 2013. Molecular and Immunological Characterization of Ragweed (Ambrosia artemisiifolia L.) Pollen after Exposure of the Plants to Elevated Ozone over a Whole Growing Season. *PLoS ONE* **8(4)**: e61518.

Kim SW, Ban SH, Chung H, Cho S, Chung HJ, Choi PS, Yoo OJ, Liu JR. **2004**. Taxonomic discrimination of flowering plants by multivariate analysis of Fourier transform infrared spectroscopy data. *Plant Cell Reports* **23**: 246–250.

Lavoie C. **2017**. The impact of invasive knotweed species (Reynoutria spp.) on the environment: review and research perspectives. *Biological Invasions* **19**: 2319–2337.

Lee CE. **2002**. Evolutionary genetics of invasive species. *Trends in Ecology & Evolution* **17**: 386–391.

Liu N, Zhao L, Tang L, Stobbs J, Parkin I, Kunst L, Karunakaran C. **2020**. Mid-infrared spectroscopy is a fast screening method for selecting Arabidopsis genotypes with altered leaf cuticular wax. *Plant, Cell & Environment* **43**: 662–674.

Long F, Liu W, Jiang X, Zhai Q, Cao X, Jiang J, Xu J. **2021**. State-of-the-art technologies for biofuel production from triglycerides: A review. *Renewable and Sustainable Energy Reviews* **148**: 111269.

Mandák B, Pyšek P, Lysák M, Suda JAN, Krahulcová A, Bímová K. **2003**. Variation in DNA-ploidy Levels of Reynoutria Taxa in the Czech Republic. *Annals of Botany* **92**: 265–272.

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

McNear DH, Chaney RL, Sparks DL. **2010**. The hyperaccumulator Alyssum murale uses complexation with nitrogen and oxygen donor ligands for Ni transport and storage. *Phytochemistry* **71**: 188–200.

Milne RI, Abbott RJ. **2000**. Origin and evolution of invasive naturalized material of Rhododendron ponticum L. in the British Isles. *Molecular Ecology* **9**: 541–556.

Minnes R, Nissinmann M, Maizels Y, Gerlitz G, Katzir A, Raichlin Y. **2017**. Using Attenuated Total Reflection–Fourier Transform Infra-Red (ATR-FTIR) spectroscopy to distinguish between melanoma cells with a different metastatic potential. *Scientific Reports* **7**: 1–7.

Moody ML, Les DH. **2007**. Geographic distribution and genotypic composition of invasive hybrid watermilfoil (Myriophyllum spicatum × M. sibiricum) populations in North America. *Biological Invasions* **9**: 559–570.

Morais CLM, Costa FSL, Lima KMG. **2017**. Variable selection with a support vector machine for discriminating Cryptococcus fungal species based on ATR-FTIR spectroscopy. *Analytical Methods* **9**: 2964–2970.

Morais CLM, Lima KMG. **2018**. Principal Component Analysis with Linear and Quadratic Discriminant Analysis for Identification of Cancer Samples Based on Mass Spectrometry. *Article J. Braz. Chem. Soc* **29**: 472–481.

Morais CLM, Lima KMG, Singh M, Martin FL. **2020**. Tutorial: multivariate classification for vibrational spectroscopy in biological samples. *Nature Protocols* **15**: 2143–2162.

Morais CLM, Paraskevaidi M, Cui L, Fullwood NJ, Isabelle M, Lima KMG, Martin-Hirsch PL, Sreedhar H, Trevisan J, Walsh MJ, *et al.* 2019. Standardization of complex biologically derived spectrochemical datasets. *Nature Protocols* **14**: 1546–1577.

Mounger J, Ainouche M, Bossdorf O, Cavé-Radet A, Li B, Parepa M, Salmon A, Yang J, Richards C. **2021**. Epigenetics and the success of invasive plants. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376(1826)**: 20200117.

Nadeem MA, Amjad Nawaz M, Shahid MQ, Doğan Y, Comertpay G, Yıldız M, Hatipoğlu R, Ahmad F, Alsaleh A, Labhane N, *et al.* 2017. DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. *Biotechnology & Biotechnological Equipment* **32(2)**: 261-285.

Nishinari K, Takemasa M, Zhang H, Takahashi R. **2007**. 2.19 Storage Plant Polysaccharides: Xyloglucans, Galactomannans, Glucomannans. *In: Comprehensive glycoscience. Elsevier Oxford*, 613–652.

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

Parepa M, Fischer M, Bossdorf O. **2013**. Environmental variability promotes plant invasion. *Nature Communications* **4**: 1–4.

Pashley CH. **2003**. The use of molecular markers in the study of the origin and evolution of Japanese Knotweed sensu lato. *Doctoral dissertation, University of Leicester.*

Pashley CH, Bailey JP, Ferris C. **2007**. Clonal diversity in British populations of the alien invasive Giant Knotweed, Fallopia sachalinensis (F. Schmidt) Ronse Decraene, in the context of European and Japanese plants. *Watsonia* **26(3)**: 359-372.

Pichancourt J-B, van Klinken RD. **2012**. Phenotypic Plasticity Influences the Size, Shape and Dynamics of the Geographic Distribution of an Invasive Plant (G Bonaventure, Ed.). *PLoS ONE* **7**: e32323.

Raba DN, Poiana MA, Borozan AB, Stef M, Radu F, Popa MV. **2015**. Investigation on crude and high-temperature heated coffee oil by ATR-FTIR spectroscopy along with antioxidant and antimicrobial properties. *PLoS ONE* **10(9)**: e0138080.

Rahman MM, Feng X, Zhang H, Yan X, Peng Q, Yu P. **2019**. Using vibrational ATR-FTIR spectroscopy with chemometrics to reveal faba CHO molecular spectral profile and CHO nutritional features in ruminant systems. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **214**: 269–276.

Richards CL, Schrey AW, Pigliucci M. **2012**. Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation (M Vellend, Ed.). *Ecology Letters* **15**: 1016–1025.

Richards CL, Walls RL, Bailey JP, Parameswaran R, George T, Pigliucci M. **2008**. Plasticity in salt tolerance traits allows for invasion of novel habitat by Japanese knotweed s. l. (*Fallopia japonica* and *F*.x *bohemica*, Polygonaceae). *American Journal of Botany* **95**: 931–942.

Ringelmann A, Riedel M, Riederer M, Hildebrandt U. **2009**. Two sides of a leaf blade: Blumeria graminis needs chemical cues in cuticular waxes of Lolium perenne for germination and differentiation. *Planta* **230**: 95–105.

Santo P. **2017**. Assessing diminution in value of residential properties affected by Japanese Knotweed. *Journal of Building Survey, Appraisal & Valuation* **6(11)**: 211-221(11).

Schuster TM, Reveal JL, Bayly MJ, Kron KA. **2015**. An updated molecular phylogeny of Polygonoideae (Polygonaceae): Relationships of Oxygonum, Pteroxygonum, and Rumex, and a new circumscription of Koenigia. *Taxon* **64**: 1188–1208.

Sharma S, Uttam KN. **2018**. Early Stage Detection of Stress Due to Copper on Maize (Zea mays L.) by Laser-Induced Fluorescence and Infrared Spectroscopy. *Journal of Applied Spectroscopy* **85**: 771–780.

De Silva DLR, Cox RC, Hetherington AM, Mansfield TA. **1986**. The role of abscisic acid and calcium in determining the behaviour of adaxial and abaxial stomata. *New phytologist* **104**: 41–51.

Skolik P, McAinsh MR, Martin FL. **2019a**. ATR-FTIR spectroscopy non-destructively detects damage-induced sour rot infection in whole tomato fruit. *Planta* **249**: 925–939.

Skolik P, Morais CLM, Martin FL, McAinsh MR. **2019b**. Determination of developmental and ripening stages of whole tomato fruit using portable infrared spectroscopy and Chemometrics. *BMC Plant Biology* **19**: 236.

Snow AA, Andersen B, Jorgensen RB. **1999**. Costs of transgenic herbicide resistance introgressed from Brassica napus into weedy B. rapa. *Molecular Ecology* **8**: 605–615.

Spalding K, Bonnier F, Bruno C, Blasco H, Board R, Benz-de Bretagne I, Byrne HJ, Butler HJ, Chourpa I, Radhakrishnan P, *et al.* 2018. Enabling quantification of protein concentration in human serum biopsies using attenuated total reflectance – Fourier transform infrared (ATR-FTIR) spectroscopy. *Vibrational Spectroscopy* **99**: 50–58.

Taiz L, Zeiger E, Møller IM, Murphy A. **2015**. Plant physiology and development. *Plant physiology and development, Sinauer Associates Incorporated* **6**.

Talari ACS, Martinez MAG, Movasaghi Z, Rehman S, Rehman IU. **2017**. Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Applied Spectroscopy Reviews* **52**: 456–506.

Tiwari G, Slaughter DC, Cantwell M. **2013**. Nondestructive maturity determination in green tomatoes using a handheld visible and near infrared instrument. *Postharvest Biology and Technology* **86**: 221–229.

Traoré M, Kaal J, Martínez Cortizas A. **2018**. Differentiation between pine woods according to species and growing location using FTIR-ATR. *Wood Science and Technology* **52**: 487–504.

Tretter L, Patocs A, Chinopoulos C. **2016**. Succinate, an intermediate in metabolism, signal transduction, ROS, hypoxia, and tumorigenesis. *Biochimica et Biophysica Acta - Bioenergetics* **1857**: 1086–1101.

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

USDA Plants Database. Plant Profile: *Polygonum ×bohemicum* (J. Chrtek & Chrtková) Zika & Jacobson [cuspidatum × sachalinense]. *United States Department of Agriculture Plants Database:* https://plants.sc.egov.usda.gov/core/profile?symbol=POBO10.

Usman K, Al-Ghouti MA, Abu-Dieyeh MH. **2019**. The assessment of cadmium, chromium, copper, and nickel tolerance and bioaccumulation by shrub plant Tetraena qataranse. *Scientific Reports* **9**: 1–11.

Wang Q, He H, Li B, Lin H, Zhang Y, Zhang J, Wang Z. **2017**. UV–Vis and ATR–FTIR spectroscopic investigations of postmortem interval based on the changes in rabbit plasma. *PLoS ONE* **12(7):** e0182161.

Willett W, Rockström J, Loken B, Springmann M, Lang T, Vermeulen S, Garnett T, Tilman D, DeClerck F, Wood A, *et al.* **2019**. Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems. *The Lancet* **393**: 447–492.

Williams F, Eschen R, Harris A, Djeddour D, Pratt C, Shaw RS, Varia S, Lamontagne-Godwin J, Thomas SE, Murphy ST. **2010**. The economic cost of invasive non-native species on Great Britain. *CABI Proj No VM10066*: 1–99.

Yang Y, Benning C. **2018**. Functions of triacylglycerols during plant development and stress. *Current Opinion in Biotechnology* **49**: 191–198.

Zhang Y-Y, Parepa M, Fischer M, Bossdorf O. **2016**. Epigenetics of colonizing species? A study of Japanese knotweed in Central Europe. In: Barrett SCH, Colautti RI, Dlugosch KM, Rieseberg LH (Eds) Invasion Genetics. Chichester, UK: John Wiley & Sons, Ltd, 328–340.

Zika PF, Jacobson AL. **2003**. An overlooked hybrid Japanese knotweed (*Polygonum cuspidatum × sachalinense*; Polygonaceae) in North America. *Rhodora* 143-152.

4.9 Acknowledgements

4.10 Supporting Information

Table S4.1: Herbarium sample information.

P Number	F Number	Country	Location	Grid Reference	Species	Chromosome Number	Collected By	Collection Date	Additional Info
P942		Japan	Lowland		*R. sachalinensis*	2n=44			
P918		Japan	Lowland		*R. sachalinensis*	2n=44			
P920		Japan			*R. sachalinensis*	2n=44			
P994		Japan	Volcano Summit		*F. japonica* var. *japonica*	2n=44			
P995		Japan	Volcano Summit		*F. japonica* var. *japonica*	2n=cc44 (circa)			
P991		Japan	Volcano Summit		*F. japonica* var. *japonica*	2n=44			
P1051		Japan			*F. x bohemica*	2n=44			

P1052		Japan			*F. x bohemica*	2n=44		
P939		Japan			*F. japonica* var. *japonica*	2n=66		
P940		Japan			*F. japonica* var. *japonica*	2n=66		
P981		Japan	Mt. Yahiko		*F. japonica* var. *japonica*	2n=44		206 cm giant tetraploid
P977		Japan			*F. japonica* var. *uzenensis*	2n=88		hairy knotweed
P978		Japan			*F. japonica* var. *uzenensis*	2n=cc88		hairy knotweed
P980		Japan			*F. japonica* var. *uzenensis*	2n=88		hairy knotweed
P987		Japan			*F. japonica* var. *japonica*	2n=88		Hirokuta city
P986		Japan			*F. japonica* var. *japonica*	2n=44		160 cm tall
P976		Japan			*R. sachalinensis*	2n=44		Coastline

P975		Japan			*R. sachalinensis*	2n=44			Coastline
		England	Cambridge Uni Botanic Gardens		*F. baldschuanica*	2n=20			"labelled *baldschuanica*"
		England	Cambridge Uni Botanic Gardens		*F. baldschuanica*	2n=20		04/08/1996	"labelled *aubertii*"
		England	Hinckley Road Leicester		*F. baldschuanica*	2n=20		05/10/1985	
		UK	The Barn Aberech		*F. baldschuanica*	No chromosome data			
		UK	Tim Rich's Garden		*F. baldschuanica*	No chromosome data		1983	
	F461a	Wales	Mawddach, downstream from Dolgellau bridge	GR 23/720.183 vc 48					Group A Leaf 1

		Wales	Mawddach, downstream from Dolgellau bridge	GR 23/720.183 vc 49					Group A Leaf 2
		Wales	Mawddach, downstream from Dolgellau bridge	GR 23/720.183 vc 50					Group B Leaf 1
		Wales	Mawddach, downstream from Dolgellau bridge	GR 23/720.183 vc 51					Group B Leaf 2
		Japan	Hokkoido Ishikori		*R. sachalinensis*	2n=44	D. Ratcliffe + T. Rich		
	F86	Ireland	West end of L, East of Maam Cross		*F. x bohemica*		JP Bailey	12/07/1995	
	F586	Ireland	Booterstown Railway, Dulbin		*F. x bohemica*				

	F289	Ireland	West Galway, Coast road north of roundstone	GR 02/726.424	*F. x bohemica*	2n=66	Ann Conolly	Sep-79	
	F637	England	Haigh Country Park, Wigan. Site B	GR 600.078	*F. x bohemica*		Blackhall	1995	
	F638	England	Grove Park Timperly. Site A	GR/ 783.874	*F. x bohemica*		B. Blackhall	1995	
	F254	England	South Wylam, County Durham		*F. x bohemica*	2n=44	Ann Conolly		Between Railway and River Tyne
P1092		Japan	Hiroshima-Yamaguchi		*F. japonica* var. *japonica*	2n=44			3m tall, giant tetraploid
	F214		ex. Platt Bracken Hill		*F. japonica* var. *japonica x F. japonica* var. *compacta*		JP Bailey	09/10/1984	
	F563				*F. japonica* var. *compacta*		JP Bailey	26/09/1984	Artificial Hybrid: *F.*

								japonica var. *japonica* (race course) x *F. japonica* var. *compacta* (Comp 1)
	F282				*R. jap x R. sach*	2n=66		19/09/1984	Artificial hybrid: *R. jap* (race course) x *R. sachalinensis* (NF)
	F317	Wales	Howey, S of Llandrindod Wells	GR 32/0..5..	*R. sach*		Ann Conolly	Aug-62	*R. sachalinensis* "HOWEY"
	F512				*F. x bohemica*		JP Bailey	14/09/1987	RC x NF3 - leaf 1
	F350				*R. sachalinensis*		Ann Conolly	05/09/1977	NF - leaf 1
	F350				*R. sachalinensis*		Ann Conolly	05/09/1977	NF - leaf 2
	F512				*F. x bohemica*		JP Bailey	14/09/1987	RC x NF3 - leaf 2

	F512				*F. x bohemica*		JP Bailey	14/09/1987	RC x NF3 - leaf 3
	F186	England	Penwaltham Church, Preston		*F. japonica* var. *japonica*				
	F168	England	Mevagissey, East Cornwall		*F. japonica* var. *japonica*				
	F198	England	Highcross + Claybridge Magna. Leicester		*F. japonica* var. *japonica*				
	F558a				*R. sachalinensis x F. bald*	2n=32			Grown from seed collected from Cheshunt
	F2205				*R. jap x F.bald*	2n=54			
		England	Penwith, Cornwall	GR 10/4465.2645	*R. sachalinensis*				
	P140a	Wales	Buryas Bridge		*Sachalinensis x F. baldschuanica*				Open poll seed from P160 Llanishan (*sach*). Hybrid: Giant

									knotweed crossed with Russian Vine
	P1190	England	Cambridge Uni Botanic Gardens		*F. japonica* var. *compacta*				(spec F2053)
	F2148				*F. baldschuanica*	2n = 20			
	F2206				*F. japonica* var. *japonica x baldschuanica*				Artificially crossed at Leicester University by Dr. John Bailey, 2n= 54
	F2209				*F. japonica* var. *compacta x baldschuanica*	2n= 32			Artificial hybrid
	F87		Collected from race course cultivated at botanic gardens		*F. japonica* var. *japonica*				

	F236				*F. japonica* var. *compacta x male sachalinensis*				Compact variety of Japanese knotweed crossed with a male Giant knotweed from Nat Y Frith
	F518				*F. japonica* var. *japonica x sachalinensis*				*F. japonica* var. *japonica* from race course x male sachalinensis from Nat Y Frith, artificial hybrid P75D
	F493	Wales	Llanishan, Wales		*R. sachalinensis*				
					Fagopyrum esculentum			14/07/1930	
		England	Wisley Common, Surrey		*Rumex acetosolla*			10/05/1964	

	F401				*F. multiflora*		J. Bailey	Sep-85	Cultivated LJR
		England	Bloody Oaks Quarry, Rutland		*F. convolvulus*		EK Horwood	11/10/1951	
	F409	Canada	Champion Lookout, Quebec, Canada		*F. cilinodis*			05/08/1987	

Table S4.2: SVM parameters.

SVM Classification	Cost	Gamma (γ)	Number of support vectors (N_{SV})
Sample type of closely related species, hybrids, and varieties	100	1	613
Upper and lower surfaces	31.6228	3.1623	327
Geographical Areas	100	0.1	111

Baseline Corrected Spectra

Legend:
- R. japonica var. japonica
- R. japonica var. compacta
- R. japonica var. uzenensis
- F. baldschuanica
- R. sachalinensis
- R. japonica x baldschuanica
- R. japonica x sachalinensis (R. bohemica)
- R.sachalinensis x F. baldschuanica
- R. jap var compacta x baldschuanica
- R. japonica var. compacta x sachalinensis
- R. japonica var. japonica x R. japonica var. compacta
- F. cilinodis
- F.convolvulus
- F. multiflora
- Fagopyrum esculentum
- Rumex acetosella

Figure S4.1: Baseline corrected spectra for the fingerprint region (1800-900 cm^{-1}).

Table S4.3: Quality parameters (accuracy, sensitivity, and specificity) for spectral classification based on sample type of closely related species, hybrids, and varieties by PCA-LDA.

PCA-LDA	% Accuracy	% Sensitivity	% Specificity
R. japonica var. *japonica*	62.87	29.90	83.41
R. japonica var. *compacta*	86.59	12.50	90.74
R. japonica var. *uzenensis*	86.93	63.33	88.99
F. baldschuanica	89.07	46.22	97.39
R. sachalinensis	74.43	75.00	74.16
R. japonica x baldschuanica	92.35	0.00	97.75
R. japonica x sachalinensis (R. x bohemica)	60.65	19.71	87.00
R. sachalinensis x F. baldschuanica	83.59	47.50	85.54
R. japonica var. *compacta x baldschuanica*	97.02	47.62	98.62
R. japonica var. *compacta x sachalinensis*	86.70	30.00	88.25
R. japonica var. *japonica x R. japonica* var. *compacta*	80.10	0.00	82.12
F. cilinodis	95.04	70.00	96.59
F. convolvulus	100.00	100.00	100.00
F. multiflora	96.88	100.00	96.78
Fagopyrum esculentum	100.00	100.00	100.00
Rumex acetosella	96.74	95.00	96.79
Average	**86.81**	**52.30**	**91.51**

Table S4.4: Quality parameters for spectral classification based on sample type of closely related species, hybrids, and varieties by SVM.

Class	% Accuracy	% Sensitivity	% Specificity
R. japonica var. *japonica*	99	98	99
R. japonica var. *compacta*	96	93	100
R. japonica var. *uzenensis*	100	100	100
F. baldschuanica	99	98	100
R. sachalinensis	99	98	100
R. japonica x baldschuanica	97	95	100
R. japonica x sachalinensis (R. x bohemica)	98	97	99
R. sachalinensis x F. baldschuanica	99	98	100
R. japonica var. *compacta x baldschuanica*	98	95	100
R. japonica var. *compacta x sachalinensis*	100	100	100
R. japonica var. *japonica x R. japonica* var. *compacta*	100	100	100
Average	**99**	**97**	**100**
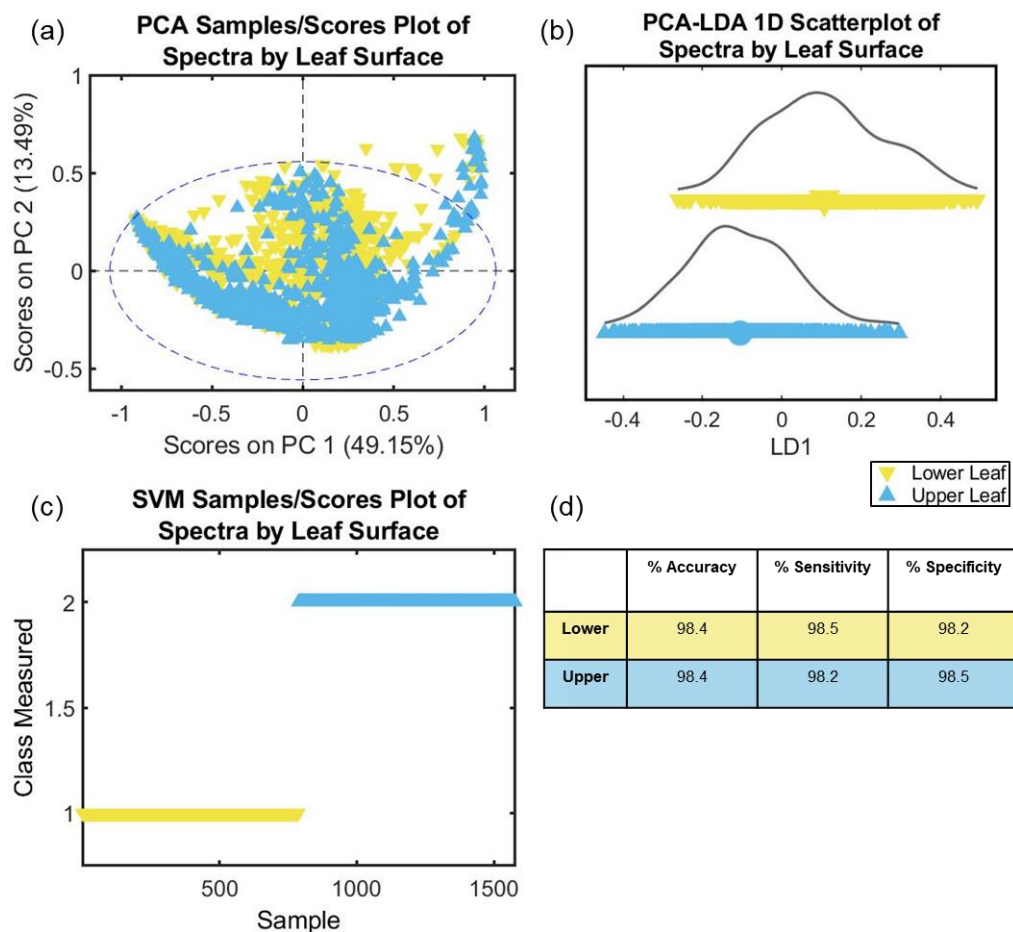
Figure S4.2: a) PCA scores plot, **b)** LDA 2D scatter plot, **c)** SVM scores plot and **d)** SVM classification table of fingerprint spectra grouped by leaf surface: upper (blue) and lower (yellow) leaf surfaces.

Table S4.5: Quality parameters (accuracy, sensitivity, and specificity) for spectral classification of *R. japonica* var. *japonica* based on geographical location by PCA-LDA.

PCA-LDA	% Accuracy	% Sensitivity	% Specificity
England	86.81	64.56	92.67
Scotland	92.16	95.00	90.32
Japan	84.58	81.01	87.62
Average	**87.85**	**80.19**	**90.20**
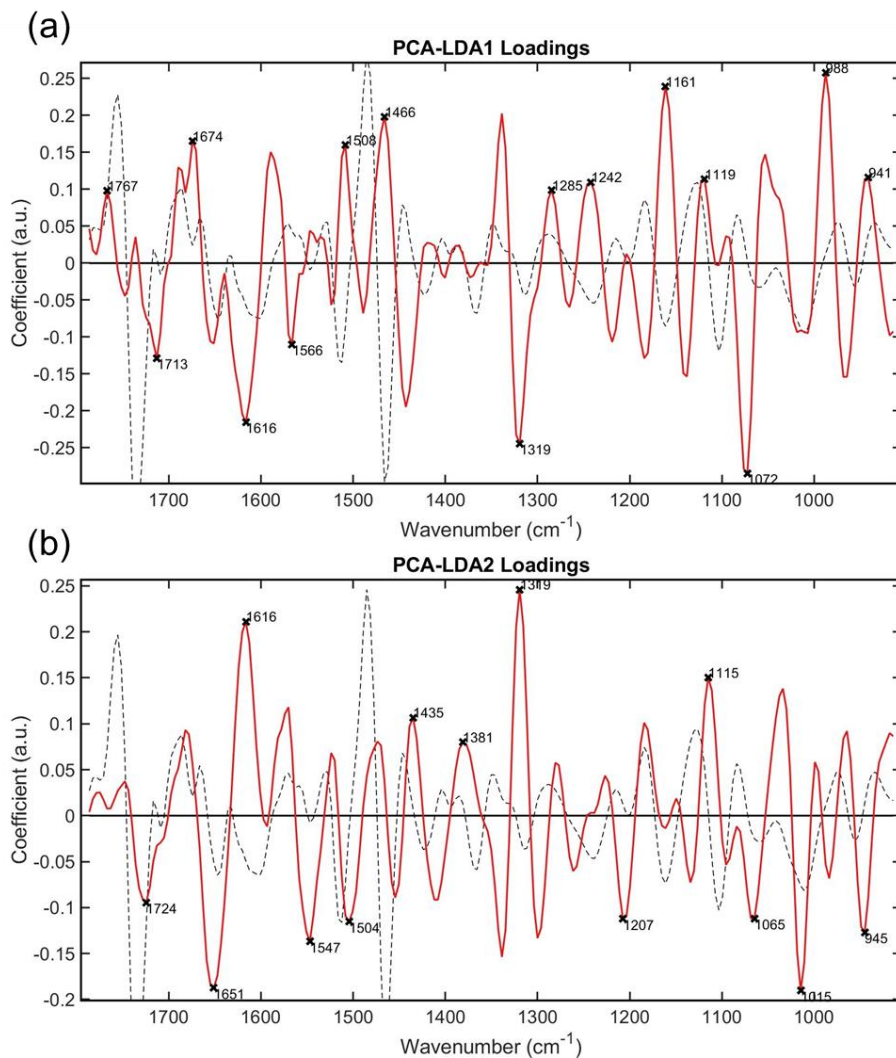
Figure S4.3: PCA-LDA loadings graphs and key wavenumbers used for HCA analysis.

(a) **Lower epidermis of *R. japonica***

(b) **Lower epidermis of *R. sachalinensis***
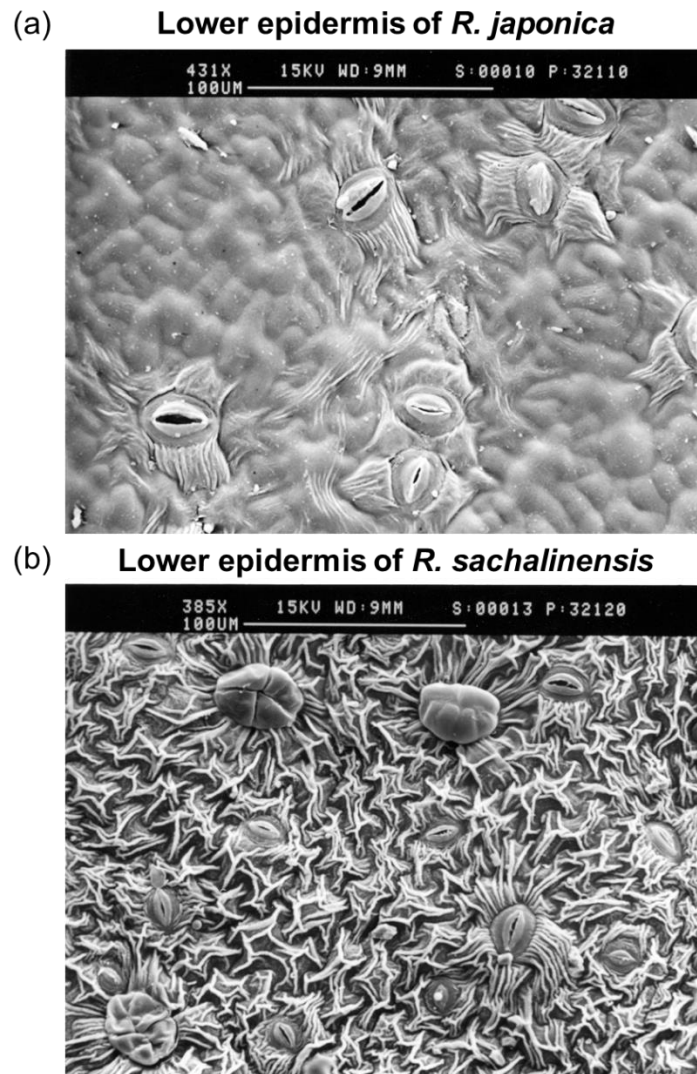
Figure S4.4: Scanning Electron Microscope images of the lower epidermis of **a)** *Reynoutria japonica* and **b)** *Reynoutria sachalinensis.*
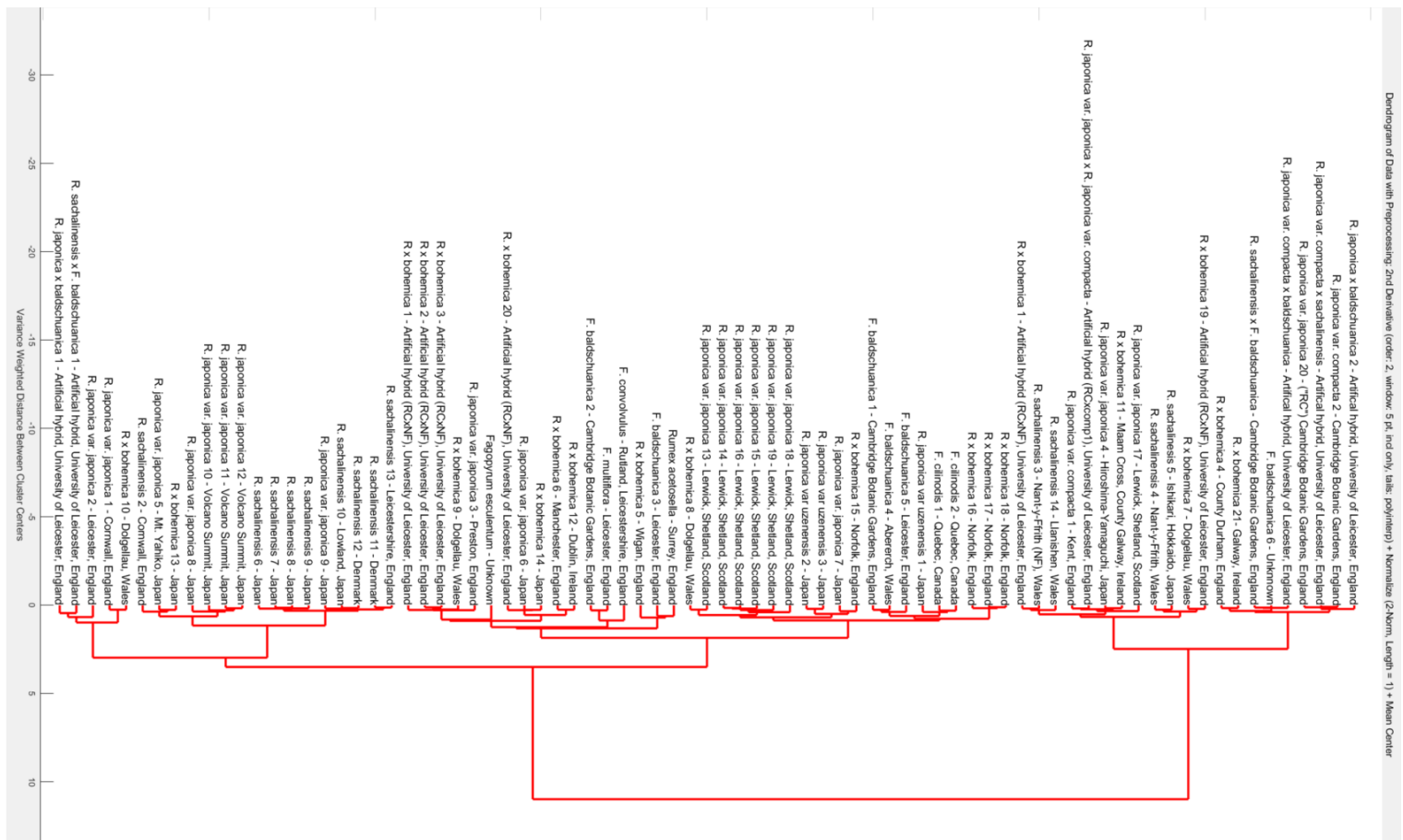
Figure S4.5: HCA analysis using fingerprint region (1800-900 cm⁻¹) for each sample with species and location information. A zoomable MATLAB file is available online at https://doi.org/10.1371/journal.pone.0261742.s006 (Holden *et al., 2022*).

5 Environmental metabolomics using attenuated total reflection Fourier-transform infrared spectroscopy for the prediction of hormone concentrations in plants

Claire A Holden[1]*, Jane E Taylor[1], Francis L Martin[2], Paul Beckett[3], Alfonso Albacete[4,5], Cristina Martínez-Andújar[5], Martin McAinsh[1]

1. Lancaster Environment Centre, Lancaster University, LA1 4YQ, UK
2. Biocel Ltd, Hull HU10 7TS, UK
3. Phlorum Ltd, Brighton BN2 6AH, UK
4. Institute for Agro-Environmental Research and Development of Murcia (IMIDA), Department of Plant Production and Agrotechnology, C/ Mayor s/n, E-30150 La Alberca, Murcia, Spain
5. CEBAS-CSIC. Department of Plant Nutrition. Campus Universitario de Espinardo, E-30100 Murcia, Spain

*Corresponding author: Claire A Holden, c.holden6@lancaster.ac.uk

5.1 Abstract

Background: Plant hormones are important in the control of physiological and developmental processes including seed germination, senescence, flowering, stomatal aperture, and ultimately the overall growth and yield of plants. Many currently available methods to quantify such growth regulators quickly and accurately require extensive sample purification using complex analytic techniques.

Methods: Herein ultra-performance liquid chromatography-high-resolution mass spectrometry (UHPLC-HRMS) was used to validate a prediction of hormone concentrations made using attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectral profiles of both freeze-dried ground leaf tissue and extracted xylem sap of Japanese knotweed (*Reynoutria japonica*) plants grown under different environmental conditions. In addition to these predictions made with partial least squares regression, further analysis of spectral data was performed using chemometric techniques, including principal component analysis, linear discriminant analysis, and support vector machines (SVM).

Results: Plants grown in different environments had sufficiently different biochemical profiles, including plant hormonal compounds, to allow successful differentiation by ATR-FTIR spectroscopy coupled with SVM. ATR-FTIR spectral biomarkers highlighted a range of biomolecules responsible for the differing spectral signatures between growth environments, such as triacylglycerol, proteins and amino acids, tannins, pectin, polysaccharides such as starch and cellulose, DNA and RNA. The potential for accurate prediction of plant hormone concentrations from ATR-FTIR spectral profiles was demonstrated using partial least squared regression, validated with hormonal data quantified by UHPLC-HRMS.

Conclusion: The application of ATR-FTIR spectroscopy and chemometrics offers accurate prediction of hormone concentrations in plant samples, with advantages over existing approaches.

5.2 Keywords

Fourier-transform infrared spectroscopy; Hormone analysis; Introduced species; Japanese knotweed; Principal component analysis; Support vector machine

5.3 Significance Statement

Plant hormones are crucial to plant environmental responses, controlling many processes including plant invasions. Current plant hormone measurement techniques are destructive and require lengthy sample preparation. This work presents a method to predict hormone concentrations using ATR-FTIR spectroscopic measurements and chemometrics, validated by UHPLC-HRMS. The spectral biomarkers detected could prove important indicators for rapid estimation of plant hormones for a variety of applications.

5.4 Introduction

As sessile organisms, plants rely on signalling molecules such as plant hormones to enable them to react appropriately to their environment; they contribute to a plastic adaptive response, regulating plant growth and stress tolerance (Anfang and Shani, 2021), and plants grown under different environmental conditions show significant differences in hormone profiles (Blázquez *et al.*, 2020; Davies, 2010). Plant hormones include: ethylene, auxin, gibberellins (GAs), cytokinins (CKs), abscisic acid (ABA), salicylic acid (SA), strigolactones (SLs), brassinosteroids (BRs) and jasmonic acid (JA) (Anfang and Shani, 2021; Davies, 2010). Plant hormone identification is challenging due to their low concentrations, ranging stabilities and similar core structures, including isomers with the same MS fragmentation patterns (e.g. cis- and trans-zeatin, topolin isomers, brassinolide and 24-epibrassinolide [24-epiBL], and castasterone and 24-epicastasterone; Šimura *et al.*, 2018). Current methods for plant hormone analysis include: gas chromatography-mass spectrometry (GC-MS), capillary electrophoresis (CE) (Porfírio *et al.*, 2016),

enzyme-linked immune sorbent assay (ELISA) (Pradko *et al.*, 2015), ultra-performance liquid chromatography-mass spectrometry (UPLC-MS) (Bosco *et al.*, 2014), high performance liquid chromatography-mass spectrometry (HPLC-MS) (Ge *et al.*, 2007) and liquid chromatography-ultraviolet detection (LC-UV) (Anagnostopoulos *et al.*, 2013). Liquid chromatography is a versatile method that allows the separation of compounds of a wide range of polarity, but these classical chromatographic techniques require destruction of the plant and lengthy sample preparation. More recently the research focus has shifted towards the development of non-destructive spectroscopic techniques for plant hormone detection, such as Raman spectroscopy (Naqvi *et al.*, 2022; Lew *et al.*, 2020).

Plant hormones control a range of complex physiological and developmental processes including seed germination, senescence, flowering, and stomatal control, and affect overall plant growth and crop yield (Anfang and Shani, 2021). Antagonistic hormonal crosstalk also regulates numerous factors influencing the success of invasive alien species (IAS), for example, the trade-off between growth and defence (Karasov *et al.*, 2017), adaptive transgenerational plasticity (Herman and Sultan, 2011), and the biosynthesis of allelopathic chemicals (Asif *et al.*, 2021). The importance of hormonal regulation in plant invasions has been demonstrated in the differential biomass allocation (Liu X *et al.*, 2021) and defence responses (Manoharan *et al.*, 2019) of invasive and native plants, and in locally adaptive chromosomal inversion in invasive plants (Lowry *et al.*, 2019)*. Additionally, many herbicides used for the control of IAS are plant hormone analogues or interfere with hormonal signalling and synthesis pathways (Grossmann, 2003). IAS have significant negative socio-economic (Fennell *et al.*, 2018; Santo, 2017) and environmental (Lavoie, 2017) impacts and therefore it is critical to gain an increased understanding of the factors, including the role of plant hormones, that enable the invasiveness and superior growth performance of these species (van Kleunen *et al.*, 2018; Parepa *et al.*, 2013; Urcelay and Austin, 2020; Liu Y *et al.*, 2021).

Japanese knotweed (*Reynoutria japonica*) is an IAS found across a broad geographic range, colonising diverse habitats including riparian wetlands, urban transport courses, and coastal areas (Zhang *et al.*, 2016; Richards *et al.*, 2012). It is very tolerant to abiotic stress, occupying extreme environments such as salt marshes (Rouifed *et al.*, 2012) and metal-polluted soil (Michalet *et al.*, 2017; Sołtysiak, 2020). Although its habitats are diverse, Japanese knotweed exhibits minimal genetic variation in Central Europe (Zhang *et al.*, 2016), Norway (Holm *et al.*, 2018) and the USA (Richards *et al.*, 2012), and exists as a female clone in the United Kingdom from a single introduction (Bailey and Conolly, 2000; Hollingsworth and Bailey, 2000). The ecological adaptability of Japanese knotweed as an invasive weed renders this species an ideal model for investigating the contribution of plant hormones to IAS invasiveness through a concatenated approach combining ultra-performance liquid chromatography-high resolution mass spectrometry (UHPLC-HRMS) and attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectral data.

This study used UHPLC-HRMS to quantitatively measure the concentrations of a set of plant hormones at nanogram per millilitre concentrations: the active CKs *trans*-Zeatin (t-Z), *trans*-zeatin riboside (tZR) and isopentyl-adenine (iP), the active GAs gibberellin A1 (GA_1), gibberellin A4 (GA_4), gibberellin A3 (GA_3), the active auxin indole-3-acetic acid (IAA), ABA, JA, SA, and the ethylene precursor 1-amino-cyclopropane-1-carboxylic acid (ACC); and compared these measured concentrations to those predicted from ATR-FTIR spectral profiles of both xylem sap and freeze-dried ground leaves. ATR-FTIR spectroscopy employs infrared (IR) light to alter the molecular vibrations of a sample, providing information on the compounds within. It is a rapid analytical technique well-suited to environmental monitoring with the advantages of a high degree of specificity and sensitivity, minimal sample preparation, and portable enough for use in the field. It can be used non-destructively on whole plant tissues, even *in planta* (Skolik *et al.*, 2019a,b). Chemometric algorithms were applied to gain further information from the absorbance profiles, such as molecular biomarkers associated with the plants'

environments. Chemometric techniques used included principal component analysis (PCA), PCA in combination with linear discriminant analysis (LDA), support vector machines (SVMs), and partial least squares regression (PLSR) (Morais & Lima, 2018; Morais *et al.*, 2017; Mehmood *et al.*, 2012). These highlighted a range of biomolecules responsible for the differing IR spectral signatures between growth environments, such as triacylglycerol, proteins and amino acids, tannins, pectin, polysaccharides such as starch and cellulose, DNA and RNA (Morais *et al.*, 2020). PLSR comparison of the ATR-FTIR spectral data with the quantitative data from UHPLC– HRMS analysis allowed the effect of each hormone on the spectral absorbances to be viewed in isolation. Key wavenumbers within the mid-infrared fingerprint region were identified for prediction of plant hormone concentrations using ATR-FTIR spectroscopy; predominantly in the region of 1200-1000 cm^{-1} for leaf samples and 1600-1500 cm^{-1} for xylem sap samples. In leaf samples these often related to polysaccharide molecules, whilst in xylem compounds these key wavenumbers were more commonly associated with nucleic acids and bases. Predictive models were built to consider the concentrations of each hormone in turn and also to detect concentrations of several different hormones at once.

5.5 Experimental Procedures

5.5.1 Plant growth

Ninety fragments of rhizome (10-50 g, volume 2-58 cm^3) were planted in fertilized organic loam (John Innes No. 1, J. Arthur Bowers, UK) in cylindrical pots designed to tightly fit in a Scholander-type pressure chamber (Soil Moisture Equipment Corp., Santa Barbara, CA, USA) measuring 6.5 cm in diameter and 23 cm in length with a volume of 763.2 cm^3, and featured a stainless-steel mesh (0.7 mm aperture) at the base to assist drainage. Pots were placed in one of two climate-controlled cabinets (Microclima 1750, Snijders Scientific BV, Netherlands) at 80% humidity, 16 h of photoperiod, and 19/11°C day/night temperature where the treatments were

applied and plants were grown for a total of fifty days before harvesting. The long photoperiod and temperature range were selected to simulate an average British Summer in the areas where Japanese knotweed usually colonises, using a comparison of temperature maps from the Met Office (Met Office, 2019) and a distribution map of Japanese knotweed in the British Isles (Bailey, 2013).

5.5.2 Treatments

Rhizome fragments were divided into eight treatment groups to give an even split of rhizome masses in each group. The treatments applied were: Light Control 'LC', Light Drought 'LD', Light Nitrogen 'LN', Light Low Nutrient 'LLN', Shade Control 'SC', Shade Drought 'SD', Shade Nitrogen 'SN' and Shade Low Nutrient 'SLN'. Four groups were placed in each of two growth cabinets. In both cabinets, the light emitted from the two high-pressure sodium lamps (SON-T 400 W, Philips Lighting, Eindhoven, The Netherlands) was reduced using a LEE 209 filter (LEE Filters Worldwide, Andover, Hampshire, UK). In one cabinet, a matrix of far-red LEDs (EPILEDS, 740-745 nm) distributed in five rows 30 cm apart was used to decrease the red: far-red ratio (R:FR) to simulate shading, see Supplementary Figure S5.1 for the spectrum produced by the LEDs alone. Wavelengths emitted were measured using an UPRtek (Taiwan) PG100N light spectrometer. The resultant combined light conditions (Supplementary Table S5.1) resulted in a 'light' treatment with a R:FR of 5.6 and a 'shade' treatment with a R:FR of 0.4 (see Figure S5.1 for the spectral profile). Plants were shuffled weekly within each cabinet to minimise positional effects from the LED matrix pattern. The R:FR of natural sunlight during the day is approximately 1.15 (Smith H, 1982) and the R:FR of 0.4 in the shade treatment was chosen to replicate that found within vegetative canopies such as sugar beet, deciduous woodland, coniferous woodland and tropical rainforest (Smith H, 1982). In both cases, the photosynthetic photon flux density (PPFD) was between 124.7 and 189.8 $\mu mol \cdot m^{-2} \cdot s^{-1}$ which is typical of growth cabinet studies (Larsen *et al.*, 2020; Pennisi *et al.*, 2020; Zou *et al.*, 2020; Park & Runkle, 2018).

Plants were provided with water (75 mL/pot / 48 h), apart from LD and SD in which water was withheld for 7 days prior to harvest. Once a week, four groups (LC, LD, SC, SD) were watered with 75 mL Hoagland solution to provide both nitrogen and micronutrients, see Supplementary Table S5.2 for details. LN and SN were fed with the commonly used agricultural dose of 50 kg ha^{-1} year^{-1} (Monaghan *et al.*, 2005); this was scaled down for a pot diameter of 6.2 cm and applied across a split-dose at 21 and 23 days to prevent leaching. Groups LLN and SLN were provided only with water and received no additional nitrogen or micronutrients.

5.5.3 Harvest

Two leaves were excised from each plant for the analysis 4-8 h into the photoperiod in order to fall within a stable period of the plants' circadian rhythm. The youngest leaf from the top of plants was placed in liquid nitrogen, freeze-dried, and finely ground for hormone analysis by U-HPLC-HRMS, and the second leaf down was treated similarly for analysis by ATR-FTIR spectroscopy. Following this, the plant was de-topped and the whole pot inserted into a Scholander-type pressure chamber (Soil Moisture Equipment Corp., Santa Barbara, CA, USA) with the stem protruding for xylem sap collection. The pressure was matched to the flow rate by increasing the pressure gradually above the balance pressure. For each trial pressure, the flow rate was calculated by weighing the sap collected for twenty seconds, until the flow rate matched that calculated by mass loss following the method previously described in (Dodd *et al.*, 2008). This was necessary as it has been shown that ABA concentration are influenced by sap flow rate (Dodd *et al.*, 2008). Sap was collected in Eppendorf vials, immediately frozen in liquid nitrogen and stored at −80°C for hormone determination, and ATR-FTIR spectral analysis.

5.5.4 Plant hormones

Plant hormones were quantified from frozen xylem sap and freeze-dried ground leaf material using UHPLC– HRMS as described previously with some modifications (Albacete *et al.*, 2008; Groãÿkinsky *et al.*, 2014). Freeze-dried ground leaf samples were prepared with several extraction steps and sonication before analysis, whilst only the filtration and centrifugation steps were necessary for the xylem sap samples. In the first extraction up to 250 mg of raw material was mixed with methanol (1.25 mL, 80%) and an internal-standards mix composed of deuterium labelled hormones ($[^2H_5]$tZ, $[_2H^5]$tZR, $[^2H_6]$iP, $[^2H_2]$GA$_1$, $[^2H_2]$GA$_3$, $[^2H_2]$GA$_4$, $[^2H_5]$IAA, $[^2H_6]$ABA, $[^2H_4]$SA, $[^2H_6]$JA, $[^2H_4]$ACC, Olchemim Ltd, Olomouc, Czech Republic) at a concentration of 5 µg mL^{-1} in 80% methanol. Samples were vortexed, incubated for 30 min at 4°C, and centrifuged (20000 g, 4°C, 15 min). Supernatants were passed through Chromafix C18 columns (MachereyNagel, Düren/Germany) previously pre-equilibrated with 80% methanol and filtrates were collected on ice. Extraction was repeated with 1.25 mL 80% methanol; second extracts were passed through the same columns. The combined extracts were collected and concentrated to complete dryness using the Integrated SpeedVac® Concentrator System AES1000 (Savant Instruments Inc., Holbrook/USA). The residues were resolved in 500 or 1000 µL 20% methanol, sonicated for 8 min using a ultrasonic bath, passed through 0.2-µm syringe filters (Chromafil PES-20/25) and placed in HPLC vials for analysis, and optionally stored at –80°C. Phytohormone analyses were performed using a UHPLC– HRMS system consisting of a Thermo ACCELA pump (Thermo Scientific, Waltham/USA) coupled to a tempered HTC-PAL autosampler (CTC Analytics, Zwingen/Switzerland), and connected to a Thermo Exactive Spectrometer (Thermo Scientific) with a heated electrospray ionization (HESI) interface. Due to the high resolution of the Orbitrap, the total ion chromatogram of the samples was recorded, and the molecules were not fragmented. A typical chromatogram for SA is shown in Supplementary Figure S5.2. The analysis was performed in the negative mode [M-H]$^-$ (Supplementary Table S5.3), and the instrument settings included: sheath gas

flow rate = 35 ml·min^{-1}, auxiliary gas flow rate = 10 ml·min^{-1}, spray voltage = 2.5 kV, capillary temperature = 275ºC, capillary voltage = -40 V, tube lens voltage = -110 V, skimmer voltage = -20 V. Mass spectra were obtained using the Xcalibur software version 2.2 (ThermoFisher Scientific, Waltham, MA, USA). For quantification of the plant hormones, calibration curves were constructed for each analysed component (1, 10, 50, and 100 µg l^{-1}) and corrected for 10 µg l^{-1} deuterated internal standards. Recovery percentages ranged between 92 and 95%.

5.5.5 ATR–FTIR spectral acquisition

Freeze-dried ground leaves and xylem sap were analysed using a Tensor 27 FTIR spectrometer with a Helios ATR attachment (Bruker Optics Ltd, Coventry, UK). The sampling area, defined by the Internal Reflection Element (IRE), which was a diamond crystal, was 250 µm x 250 µm. Spectral resolution was 8 cm^{-1} with 2 times zero-filling, giving a data-spacing of 4 cm^{-1} over the range 4000 to 400 cm^{-1}; 32 co-additions and a mirror velocity of 2.2 kHz were used for optimum signal to noise ratio. To minimise bias, ten spectra were taken for each sample. Each sample was placed on a slide with the side to be analysed facing upwards, placed on a moving platform, and then raised to ensure a consistent contact with the diamond crystal. For xylem sap samples, 30 µL of xylem sap was placed on a tin foil-covered slide and allowed to dry before analysis. For freeze-dried ground leaves a small amount of powder was transferred to each slide using a spatula. A total of 410 spectra were taken for xylem sap and 330 spectra were taken of freeze-dried ground leaf tissue.

5.5.6 Data analysis

The 'mergetool' function of an in-house developed MATLAB (Mathworks, Natick, USA) toolbox called IRootLab (Martin FL *et al.*, 2010; Trevisan *et al.*, 2013) was used to convert all spectral information from OPUS format to suitable files (.txt). Following this, it was necessary to pre-process the acquired spectra to improve the

signal-to-noise ratio. Pre-processing corrects problems associated with random or systematic artefacts during spectral acquisition and is an essential step of all spectroscopic experiments. Pre-processing and computational analysis of the data were performed using a combination of IRootLab toolbox (Martin FL *et al.*, 2010; Trevisan *et al.*, 2013) and the PLS Toolbox version 7.9.3 (Eigenvector Research, Inc., Manson, USA). The pre-processing steps applied to all spectra were firstly the selection of the spectral biochemical fingerprint region (1800-900 cm^{-1}), followed by Savitzky–Golay (SG) second differentiation (nine smoothing points) and vector normalisation. All data were mean centred before multivariate analysis, where multiple dependant variables are observed simultaneously to determine a pattern.

Four machine learning techniques were used in this study: an unsupervised dimensionality reduction method, two supervised classification methods and one regression. The unsupervised method PCA simplifies complex multivariate datasets, allowing them to be presented intuitively and enabling pattern recognition. Two supervised chemometric techniques, PCA-LDA and SVMs, were used for the classification of groups (Morais & Lima, 2018; Morais *et al.*, 2017). PCA-LDA was also used for the determination of biomarkers. Most importantly, hormone prediction was achieved using a multivariate analysis technique called PLSR of both ATR-FTIR spectral data and real hormone data as measured by UHPLC-HRMS (Mehmood *et al.*, 2012). Multivariate analysis techniques allow multiple variables to be compared at the same time enabling spectral absorbance values across a range of wavelengths to be simultaneously correlated against concentrations of multiple hormones for numerous samples. Observing all these data at once allows patterns to be seen and enables predictions to be made. To form these models, an X-block of ATR-FTIR spectral absorbance data for plants was analysed by PLSR against a Y-block of hormone concentrations for the corresponding plants as measured using UHPLC-HRMS. Environments were analysed separately, allowing a model to be created for each of them. Once made, these models can be applied to new ATR-FTIR

spectral data in the absence of UHPLC-HRMS data to predict plant hormone concentrations.

5.6 Results

5.6.1 ATR–FTIR spectral analysis classifies plants from different environments via spectral differences

The sensitive nature of IR spectroscopy allowed indications of plant responses to environment to be observed visually as differences between spectral profiles. The pre-processed fingerprint spectra exhibit distinguishable differences between spectra of different treatment groups, for both xylem sap and freeze-dried ground samples, at 950, 1050, 1150, 1250, 1325, 1400, 1525, 1575 and 1610 cm^{-1} (Supplementary Figure S5.3b) and 950, 1050, 1275, 1400, 1525 and 1610 cm^{-1} (Supplementary Figure S5.3d), respectively. Three chemometric techniques (PCA, PCA-LDA and SVM) were used to extract further information from the spectral absorbance profiles of xylem sap (Figures 5.1a-d) and freeze-dried ground leaves (Figures 5.2a-d). The unsupervised technique, PCA, showed poor separation between treatment groups in xylem sap samples (Figure 5.1a). However, addition of the supervised classifier LDA created biologically meaningful separation along the LD1 axis, with xylem sap relating to nutrient levels (Figure 5.1b) and leaf samples relating to light (Figure 5.2b). For the xylem sap samples, the left-hand side of the PCA-LDA scatter graph contains both control and drought plant samples (LC and LD) which were watered with Hoagland solution, the central portion contains clusters of nitrogen fed and low nutrient shaded plants (SN and SLN), and the right-hand side contains the light samples of the nitrogen and low nutrient categories (LN and LLN). The pattern observed in Figure 5.2a is distinctive due to the homogenisation introduced by the grinding process; PCA of freeze-dried ground leaves separated spectra from individual samples into clusters. PCA-LDA of freeze-dried leaf samples (Figure 5.2b) resulted in a separation along the axis LD1; LD to

the left, LC, LN and LLN in the central portion, and all shaded groups to the right (SC, SD, SN and SLN). The stronger chemometric technique, SVM, achieved the best classification results for both sample types. Analysis of spectra from xylem sap samples using SVM achieved 99.0% accuracy, 98.2% sensitivity, and 99.8% specificity (Figures 5.1c-d). However, application of SVM to spectra of freeze-dried ground leaves attained even better separation with 99.8% accuracy, 99.6% sensitivity and 100.0% specificity (Figures 5.2c-d). For SVM model parameters, cost, gamma and number of support vectors, see Supplementary Table S5.4.
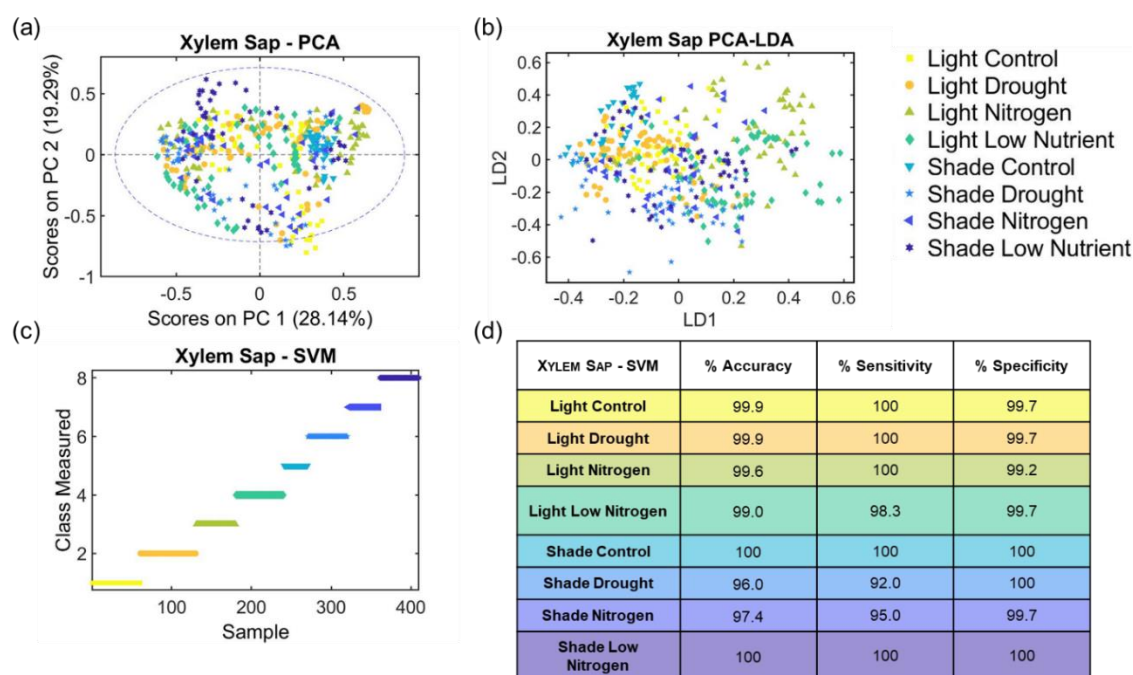
Figure 5.1: a) PCA scores plot showing poor separation between classes, b) PCA-LDA scatter plot showing some separation by nutrient levels, c) SVM sample/measured plot and d) SVM results for ATR-FTIR spectra taken of xylem sap samples showing excellent classification, grouped by treatments; Light Control (LC), Light Drought (LD), Light Nitrogen (LN), Light Low Nitrogen (LLN), Shade Control (SC), Shade Drought (SD), Shade Nitrogen (SN) and Shade Low Nitrogen (SLN).
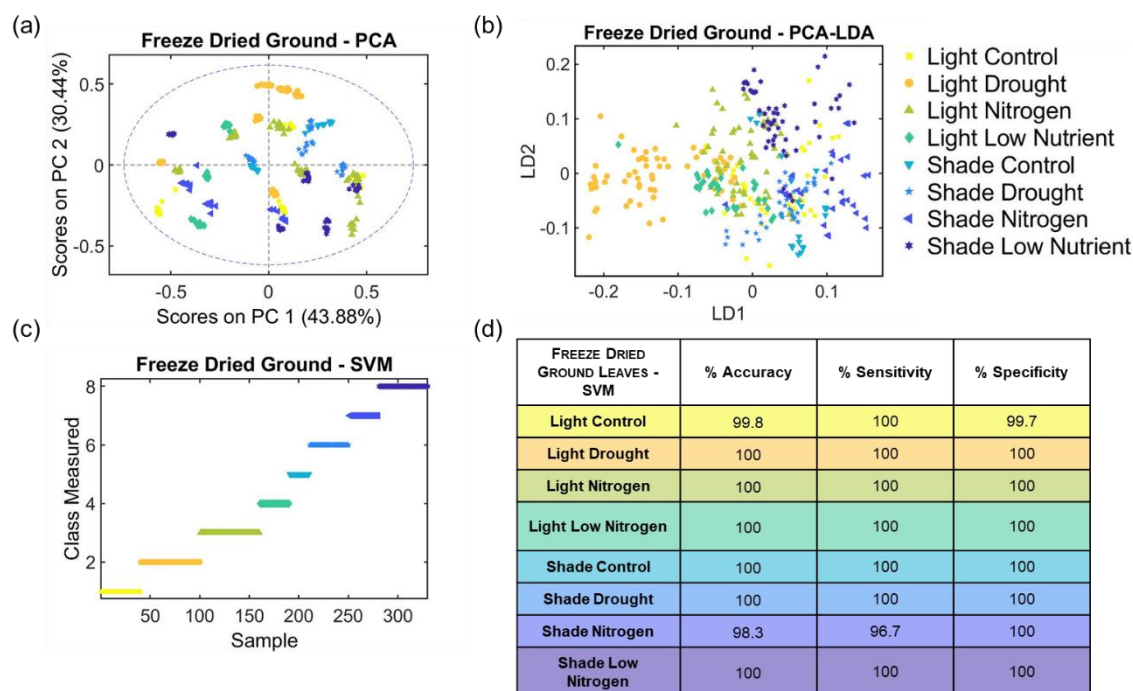
Figure 5.2: a) PCA scores plot in which each cluster is formed from separate samples due to the homogenisation introduced by the grinding process, **b)** PCA-LDA scatter plot showing some separation by light levels, **c)** SVM sample/measured plot and **d)** SVM results for ATR-FTIR spectra taken of freeze-dried ground leaves samples showing excellent classification, grouped by treatments; Light Control (LC), Light Drought (LD), Light Nitrogen (LN), Light Low Nitrogen (LLN), Shade Control (SC), Shade Drought (SD), Shade Nitrogen (SN) and Shade Low Nitrogen (SLN).

5.6.2 ATR-FTIR spectral analysis identifies biomolecular differences between treatments

ATR-FTIR spectroscopy can detect changes in concentration or molecular structure of compounds. Significant biomolecular differences can be deciphered by examination of the key wavenumbers, which differentiate spectral profiles of different treatment groups from one another. These wavenumbers are called loadings (Figure S5.4) and their tentative molecular assignments have been found through examination of the literature for both xylem sap and leaf sample types for biomarker information and references (Table S5.5). The peaks which differentiate treatment groups in xylem sap samples were related to a range of biomolecules such

as triacylglycerol, proteins, glutamate, cellulose, tannins, starch, and RNA (Nozahic & Amziane, 2012; Belfer *et al.*, 1998; Shivu *et al.*, 2013; Jin *et al.*, 2018; Moskal *et al.*, 2019; Talari *et al.*, 2017; Gorzsas, 2020; Falcão & Araújo, 2013; Morais *et al.*, 2017). For freeze-dried ground leaves, the differences were found in much the same compounds: triacylglycerol, proteins and amino acids, pectin, polysaccharides such as starch and cellulose, and DNA (Talari *et al.*, 2017; Belfer *et al.*, 1998; Shivu *et al.*, 2013; Rana *et al.*, 2018; Sharma and Uttam, 2018; Ajitha *et al.*, 2015).

5.6.3 UHPLC– HRMS hormone analysis indicates that hormone concentrations are impacted by applied treatments

Plants respond to their environment via signalling molecules such as hormones, to enable a plastic response. This is reflected in the concentrations of plant hormones measured by UHPLC-HRMS (ACC, tZ, iP, SA, ABA, JA, GA_1, GA_4, GA_3, tZR, and IAA) which were different between plants belonging to different treatment groups (see Figure 5.3a and c; Supplementary Figures S5.5 and S5.6. Figure 5.3a shows separation of LD and SD plants along PC1 based on xylem sap hormone concentrations accounting for 65.07% of the variance. This is primarily due to increased ABA and tZ (Figure 5.3b, PC1 loadings in blue). The separation along PC2 for xylem sap samples is due to the antagonistic relationship between JA and ABA (Figure 5.3b, PC2 loadings in green), which is variable within treatment categories (Figure 5.3a). Figure 5.3c also shows a separation along PC1 of droughted samples based on the hormone concentrations of freeze-dried ground leaves, accounting for 46.32% of the sample variance. High leaf ABA and low leaf ACC, JA and tZ concentrations were primary responsible for separation along axis PC1 (Figure 5.3d, PC1 loadings in blue). The PC2 axis of Figure 5.3c shows some separation by lighting treatment, however this separation was of lesser importance and only explained 38.23% of the variance. The green line in Figure 5.3d indicates that ABA, JA, tZ, and SA were all higher in LC and LD samples to create this separation along axis PC2,

whilst ACC was lower. JA concentrations in plants with a low red: far-red ratio were lower.
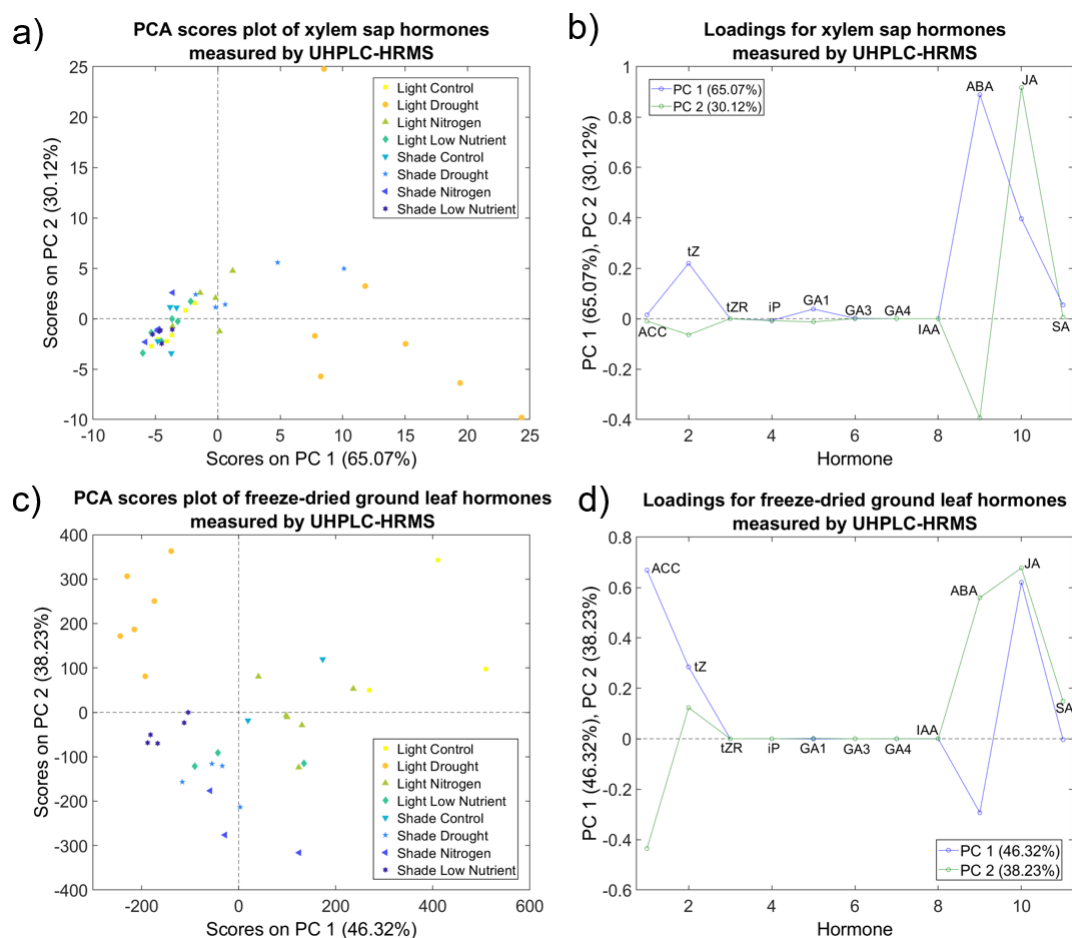
Figure 5.3: UHPLC-HRMS measurements of plant hormone concentrations analysed by PCA: **a)** xylem sap PCA scores showing separation of droughted plants along the PC1 axis, **b)** xylem sap loadings highlighting the importance of ABA in droughted samples, **c)** freeze-dried ground leaf scores showing separation by drought along PC1 and red: far-red ratio along PC2, **d)** freeze-dried ground leaf loadings indicating that droughted plants exhibited high ABA and low ACC, JA and tZ concentrations whilst plants with a high red: far-red ratio had high ABA, JA, tZ, and SA but low ACC concentrations.

In xylem sap samples (Supplementary Figure S5.4), ABA concentration was highest in the drought categories; LD and SD, at ~17 and ~7 ng·ml^{-1} of sap ABA respectively,

whilst the other categories ranged between ~1 and 3 ng·ml^{-1} sap. Leaf ABA concentrations (Supplementary Figure S5.5) were approximately quadruple in LD than those of the other categories. Shade plants had notably lower xylem SA concentrations, in the range of 0.7-1.1 ng·ml^{-1} sap compared with 1.6-4.5 ng·ml^{-1} sap for 'light' plants. Leaf tZ was 4.5-fold higher in LC plants than in those of SLN. Leaf JA concentration was significantly higher in the light control group LC (~710 ng·g^{-1} dry weight) compared to all other groups (ranging 170-420 ng·g^{-1} dry weight), except the shade control group SC (~460 ng·g^{-1} dry weight). LC had the highest iP concentrations at 0.25 ng·g^{-1} dry weight, significantly higher compared to groups LD, LN, SD, SN (ranging 0.03-0.6 ng·g^{-1} dry weight), with the other groups falling in between.

5.6.4 Combined ATR-FTIR UHPLC-HRMS analysis identifies key spectral wavenumber for hormone prediction via ATR-FTIR spectroscopy

Whilst the plant hormone concentrations quantified by using UHPLC-HRMS served to confirm that the applied treatments were effective at inducing a phenotypic response, importantly the UHPLC-HRMS data enabled the generation of predictive models for hormone concentrations using ATR-FTIR spectral data by means of a multivariate analysis technique called partial least squares regression. PLSR allows simultaneous comparison of multivariate datasets, in this case, the spectral absorbance values for either freeze-dried ground leaf tissue or from xylem sap compared with the plant hormone values obtained by HPLC-HRMS. Using PLSR, the extracted plant hormone concentrations measured by UHPLC-HRMS were accurately predicted from ATR-FTIR spectral profiles of the same sample material.
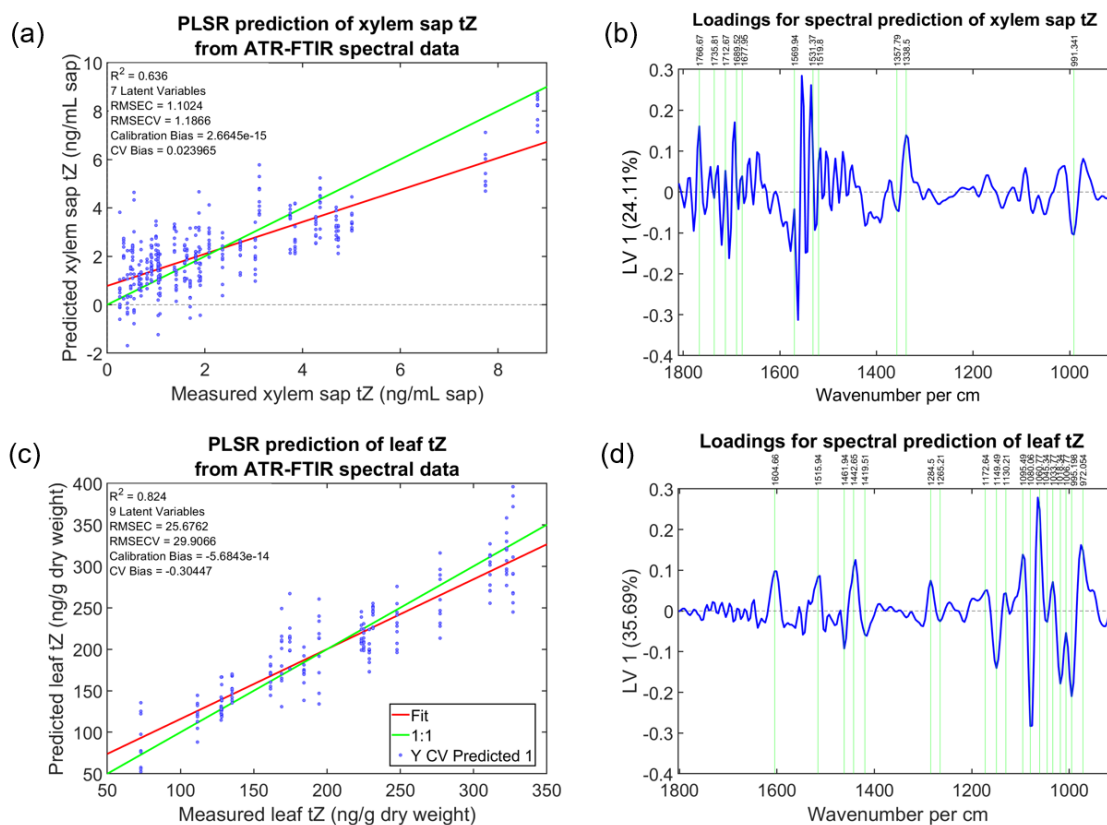
Figure 5.4: PLS regression and loadings of *trans*-Zeatin concentrations as measured using UHPLC-HRMS against predicted values in ng·g^{-1} dry weight using ATR-FTIR spectra of freeze-dried ground leaves grown under all treatment conditions. In panels **a)** and **c)**, the green line shows the ideal prediction gradient of one, which would be 100% accurate. The red line shows the achieved success of the model, which was a gradient of R^2 = 0.636 for xylem sap and 0.824 for freeze-dried ground leaves. These models were created using spectral data from all treatment categories for individual hormones. Panels **b)** and **d)** show the key wavenumbers involved in making this prediction.

The graphs in Figure 5.4 show the PLS regressions and loadings of tZ hormone concentrations as measured using UHPLC-HRMS against predicted concentrations using ATR-FTIR spectra of freeze-dried ground leaves from all treatment categories as an example of the predictive models generated using this approach (see Supplementary Figures S5.9 and S5.10 for of the predictive models for the other hormones). For the regressions in Figure 5.4a and Figure 5.4c, the green lines show the ideal prediction gradient of one, which would be 100% accurate, and the

gradient of the red lines indicate the achieved success of the model's prediction. Leaf samples achieved a more accurate prediction of R^2s= 0.807 to 0.839 compared with 0.477-0.681 for xylem sap samples (see Table S5.6). The PLSR models in Figures 5.4, S5.7 and S5.9 use hormonal data measured by UHPLC-HRMS to train them on the correlation between different hormone concentrations and the corresponding differences in ATR-FTIR spectral profiles. For each hormone, and each sample type, different spectral wavenumbers are important in making this prediction. These key wavenumbers are called loadings and are presented in Supplementary Figures S5.8 and S5.10 for each hormone and sample type. Key wavenumbers were detected mostly in the region of 1200-1000 cm^{-1} for prediction of leaf hormone concentration and between 1600-1500 cm^{-1} for xylem sap hormone concentration.

5.6.5 Combined ATR–FTIR UHPLC–HRMS analysis gives a high correlation between predicted and measured hormone concentrations

Analysis of data from each treatment separately allowed the generation of treatment-specific models. Table 5.1 shows the R^2 values for predicted against measured hormone concentrations, with each row being a separate treatment. The R^2 values for the predictions from xylem sap samples ranged between 0.48-0.9 (Table 5.1a). The R^2 values for the predictions from freeze-dried ground leaves ranged from 0.74-0.98 (Table 5.1b). LC and SC achieved the highest correlation between predicted and measured hormone concentrations. These models therefore provide a valuable resource that can be saved and applied to new spectral data obtained from plants grown under similar conditions thereby allowing the hormone concentrations to be accurately predicted without the requirement for exhaustive UHPLC– HRMS analysis.

Table 5.1: R^2 values for predicted against measured hormone concentrations from partial least squares regression between **a)** xylem sap and **b)** freeze-dried ground leaf ATR-FTIR spectral data against UHPLC-HRMS-measured hormone concentrations. Hormones with zero values for multiple plants were excluded from the model and are designated as NA. R^2 results in tables are colour coded from red (0.48) to green (0.98) and were higher for leaf samples, ranging from 0.74-0.98, whilst xylem sap predictions had R^2s of 0.48-0.93.

(a)

Xylem Sap R^2	tZ	iP	GA1	GA3	GA4	IAA	ABA	JA	SA
Light Control	0.59	0.74	0.76	NA	0.62	0.63	0.73	0.66	0.89
Light Drought	0.90	0.91	0.58	0.76	NA	NA	0.89	0.91	0.85
Light Nitrogen	0.66	0.69	0.74	0.70	NA	NA	0.76	0.81	0.82
Light Low Nutrient	0.8	0.91	0.83	NA	NA	NA	0.80	0.93	0.77
Shade Control	0.82	0.83	0.92	0.92	NA	NA	0.86	0.89	0.92
Shade Drought	0.83	NA	0.74	0.48	NA	NA	0.80	0.77	0.76
Shade Nitrogen	0.88	0.84	0.78	0.76	NA	0.72	0.75	0.82	0.78
Shade Low Nutrient	0.83	NA	0.79	NA	NA	NA	0.76	0.79	0.88

(b)

FDG Leaves R^2	ACC	tZ	ABA	JA	SA
Light Control	0.97	0.98	0.97	0.98	0.98
Light Drought	0.81	0.74	0.81	0.87	0.82
Light Nitrogen	0.83	0.83	0.83	0.83	0.83
Light Low Nutrient	0.92	0.95	0.92	0.95	0.90
Shade Control	0.98	0.98	0.98	0.98	0.98
Shade Drought	0.94	0.94	0.94	0.94	0.94
Shade Nitrogen	0.81	0.81	0.81	0.81	0.81
Shade Low Nutrient	0.89	0.88	0.84	0.88	0.85

Worst		Best
0.48		0.98

5.7 Discussion

5.7.1 Differences in ATR-FTIR spectral profiles are highlighted through chemometrics

Japanese knotweed and other invasive species with low genetic variation exhibit a plastic response to their environment which is thought to contribute to their invasion success (van Kleunen *et al.*, 2018; Geng *et al.*, 2016; Richards *et al.*, 2006). This phenotypic plasticity was reflected in the present study in the differences found between spectral profiles between treatment groups. This is consistent with the

results of studies in which ATR-FTIR spectroscopy has been successful in differentiating plants' nutrient status and plants from different growing environments (Butler *et al.*, 2015; Holden *et al.*, 2021; Traoré *et al.*, 2018; Holden *et al.*, 2022). The environmentally induced phenotypic changes were successfully captured by the ATR-FTIR spectral profiles, which were visibly different (see Figure S5.3). Figures 5.1 and 5.2 demonstrate the power of chemometrics to emphasise these differences. SVM was the most successful technique applied and had marginally more success in the freeze-dried ground samples, likely due to the homogenisation of the samples during the grinding process leading to more predictable results. The higher separation of spectra from freeze-dried ground leaves (Figure 5.2a) by PCA than that of xylem sap spectra (Figure 5.1a) could be due to the averaging effect of leaf growth over time, adapted to each environment, compared with the nature of the xylem-sap samples which capture a moment in time and could be influenced by compounds related to development stage. Leaf samples reflect a balance between synthesis and metabolism and the import and export of compounds, whilst xylem sap samples reflect instantaneous transport. The sample type more closely correlated to the physiological response therefore depends on the analyte of interest.

5.7.2 Hormone profiles reflect plant response to environment

It is well established that plant stresses such as drought, nutrient deficiency and shading can have a marked impact on the concentrations of plant hormones (Anfang & Shani, 2021; Davies, 2010). Results consistent with this were obtained through measurement of plant hormones with the highly specific technique, UHPLC-HRMS, from xylem sap (Figure S5.5) and leaves (Figure S5.6). This was explored using PCA which allowed all hormones to be considered together (Figure 5.3). The applied treatments (LC, LD, LN, LLN, SC, SD, SN and SLN) were sufficiently different to alter the hormone profiles in the plants, reflecting adaptations to each environment

(Wolters & Jürgens, 2009). Importantly, such a range of hormone concentrations was essential prerequisite to create good datasets for regression analysis.

5.7.3 Hormonal biomarkers identified for mid-infrared spectroscopy

The process from chemometric biomarker identification to physical biomolecular extraction is a developing area of spectroscopy with ongoing research to optimise concentration quantification (Spalding *et al.*, 2018; Wagner *et al.*, 2010), molecular definition databases (Talari *et al.*, 2017) and new applications (Skolik *et al.*, 2019a,b; Butler *et al.*, 2020; Holden *et al.*, 2022; Holden *et al.*, 2021). It was therefore crucial that predictions for expected hormone profiles from spectroscopic data were made and verified against actual hormone concentrations quantified by mass spectrometry. PLSR comparison of the ATR-FTIR spectral data with the quantitative data from UHPLC– HRMS analysis allowed the effect of each hormone on the spectral absorbances to be viewed in isolation. The loadings in Figure 5.4 show key spectral wavenumbers used in the model creation for tZ concentration prediction. The most important regions for prediction of hormone concentrations using ATR-FTIR spectral profiles were 1200-1000 cm^{-1} for leaf samples and 1600-1500 cm^{-1} for xylem sap samples.

Three wavenumbers used to predict ABA hormone concentration in leaf samples, 1612, 1566, and 1323 cm^{-1} are often attributed to the Amide I (Jin *et al.*, 2018), Amide II bands of proteins (N-H bending and C-N stretching; Rana *et al.*, 2018), and Amide III respectively (Morais *et al.*, 2017). As ABA does not contain nitrogen within its structure this suggests that ABA-associated biochemical changes in other compounds within the leaves are acting as proxy indicators for the estimation of ABA concentration. Similarly, 1516 cm^{-1} is also associated with Amide II vibrations of proteins and was one of the key indicators for prediction of tZ, JA and SA concentrations in leaves (Talari *et al.*, 2017). The amide III-associated (Morais *et al.*, 2017) peak at 1323 cm^{-1} was also used to predict leaf SA concentrations. Two

phosphorus-associated peaks were used for the prediction of leaf ABA concentration; 1211 cm^{-1} which is associated with PO_2^{-} asymmetric stretching (Phosphate I), and 1065 cm^{-1} linked to C–O stretching of the phosphodiester and the ribose of bases (Talari *et al.*, 2017). As ABA also does not contain phosphorus, this supports the hypothesis that compounds other than ABA contribute to a 'spectral signature' for ABA-associated biochemical change and suggest the use of associated compounds as a proxy, would be useful to gain an overall picture of plant health in agricultural and ecological settings.

In contrast, leaf SA concentrations were predicted using two peaks which could be directly related to the structure of SA; 1582 cm^{-1} which is linked to the ring C–C stretch of phenyl, and 1339 cm^{-1} is associated with in-plane C–O stretching vibration combined with the ring stretch of phenyl (Talari *et al.*, 2017). Peak 1339 cm^{-1} was used for prediction of leaf ABA and SA, as well as xylem ABA, tZ and SA. Other wavenumbers relating to Amides I and II (1663, 1547, 1570, 1555 cm-1) were also important for the prediction of hormone concentrations, see summary table in Supplementary Table S5.5 (Belfer *et al.*, 1998; Shivu *et al.*, 2013; Talari *et al.*, 2017; Strong *et al.*, 2017).

When plants are under stress, signalling cascades including hormones and reactive oxygen species (ROS) to induce biochemical changes (Heap *et al.*, 2020). As an important regulator in response to drought stress, ABA induces ROS accumulation to facilitate stomatal closure (Bharath *et al.*, 2021), whilst SA, which is part of the innate immune response (Maruri-López *et al.*, 2019), ameliorates oxidative damage through regulation of redox signalling and the antioxidant defence system (Saleem *et al.*, 2021). To prevent oxidative damage, excess ROS are absorbed and quenched by phenolic compounds, which have antioxidant properties (Zheng & Wang, 2001). This coordinated biochemical response perhaps explains why the biomarker at 1512 cm^{-1} which is associated with ν (C-C) aromatic (conjugated with C=C phenolic compounds (Heredia-Guerrero *et al.*, 2014) was a key wavenumber for the

prediction of xylem sap ABA and SA concentrations. Peak 1177 cm^{-1} is associated with the C–O stretch vibration of tannins (Falcão & Araújo, 2013), and was an important predictor of xylem JA concentrations.

These results also used wavenumbers associated with polysaccharides for the prediction of hormone concentrations, particularly within leaf samples. Peak 1038 cm^{-1}, associated with the polysaccharide galactan (Ord *et al.*, 2016), was important in the prediction of leaf SA concentrations. Leaf tZ and leaf JA concentrations were predicted using peak 1130 cm^{-1} which has previously been attributed to the stretching vibrations [ν(CO)] of the COC glycosidic linkage of polysaccharides (Liu X *et al.*, 2021). Pectin is associated with the biomarker at 1443 cm^{-1} (Sharma & Uttam, 2018), which was used for prediction of leaf tZ, SA, JA and ACC concentrations, and the peak at 972 cm^{-1} (specifically from the OCH$_3$ group of polysaccharides such as pectin) (Talari *et al.*, 2017) which was used in the prediction of leaf ABA, tZ, JA and ACC . This association with leaf JA could be linked to jasmonate-mediated accumulation of leaf soluble sugars in response to far-red light (Courbier *et al.*, 2020). The peak at 1636 cm^{-1} linked to C=O stretching of carbonyl group, typical of saccharide absorption (Talari *et al.*, 2017), was important in prediction of xylem JA and leaf SA concentrations. Leaf ABA was predicted using the peak at 1049 cm^{-1} which is associated with cellulose (Moskal *et al.*, 2019). Biomarker 1732 cm^{-1} which has been associated with both hemicellulose (Ord *et al.*, 2016), was a predictor of leaf ABA concentrations. As a key hormone in the drought response, it is fitting that ABA would be estimated using hemicellulose because the leaves of drought-treated plants are known to have a higher content of hemicellulosic polysaccharides (van der Weijde *et al.*, 2017). Biomarker 1732 cm^{-1} has also been associated with lipid fatty acid esters (Ord *et al.*, 2016), which is the more probable molecular assignment in its use for estimation of xylem JA concentrations because the fatty acid, linolenic acid, is an important precursor of JA synthesis (Gfeller *et al.*, 2010).

Whilst leaf hormone concentrations were strongly associated with sugar compounds, in xylem sap samples nucleic acids and bases generally proved more important indicators of hormone concentration. ABA, tZ and SA concentrations in xylem sap were predicted using peak 1690 cm^{-1} which is associated with nucleic acids due to the base carbonyl (C=O) stretching and ring breathing mode (Talari *et al.*, 2017). Like 1065 cm^{-1}, the peak 991 cm^{-1} is also associated with C–O stretching of the phosphodiester and the ribose of bases (Talari *et al.*, 2017). This peak was important in xylem sap samples for the prediction of ABA, tZ, SA, and GA1 concentrations. Peak 1713 cm^{-1}, associated with the C=O of the base thymine (Talari *et al.*, 2017), was important in prediction of tZ and SA concentrations in xylem sap samples. Peak 1690 cm^{-1}, linked to nucleic acids due to the base carbonyl (C=O) stretching and ring breathing mode (Talari *et al.*, 2017), was useful in prediction of xylem sap concentrations of ABA, tZ and SA. Peak 1574 cm^{-1} relating to the C=N of adenine (Talari *et al.*, 2017), was important in the prediction of xylem GA1 concentrations. Peak 1531 cm^{-1}, associated previously with modified guanine (Talari *et al.*, 2017), was used in the prediction of xylem tZ and SA.

5.7.4 ATR-FTIR spectral profiles allow prediction of hormone concentrations

The ATR-FTIR spectrum is information rich and provides an integrated holistic picture of the entire cellular biochemistry (Morais et al., 2020). In response to the growth environment, biomolecules unrelated, related and influenced by hormonal activity will be altered, presumably in a dose-related fashion. Chemometrics provides a method to extract this chemical information from spectral absorbances, considering the ratios of different biochemical entities and potentially allowing us to find the "needle in a haystack" of individual hormones in their natural state. PLSR models have previously been applied to the infrared and Raman spectroscopic absorbances of plant-derived samples to quantify individual components within

molecular mixtures (Naqvi *et al.*, 2022; Lew *et al.*, 2020; Zhu *et al.*, 2018; Romera-Fernández *et al.*, 2012; Bensemmane *et al.*, 2021).

This work provides a demonstration of PLSR for the accurate prediction of plant hormone concentrations from ATR-FTIR spectral profiles. The accuracy of PLSR prediction of tZ concentrations was lower for xylem sap (Figure 5.4a, R^2=0.636) compared with leaf samples (Figure 5.4c, R^2=0.824). This difference is due to the distribution of points for the regression, as the tZ concentrations present in the xylem sap were more similar between treatment categories than those present in the leaf. This highlights the importance of the applied treatment choice to produce a good calibration dataset for a chosen hormone. For example, to create an ABA specific model, application of a wide range of drought severities would be ideal, because ABA is the main regulator of the drought stress response (Bharath *et al.*, 2021) and appears as a key loading for separation of droughted plants in Figure 5.3, however this would not be the optimal calibration dataset for another hormone. The PCA loadings based on hormonal data alone (Figures 5.3b and 5.3d) show that in both leaf and xylem samples, tZ is a key loading for separation along the axis PC1 in Figures 5.3a and 3b. Whilst leaf samples in Figure 5.3b show a good distribution along PCA1, indicating a variety of leaf tZ levels, xylem samples Figure 5.3a show overlapping clusters. This overlap indicates similarity of xylem sap hormones concentrations across treatment categories, which explains why the xylem sap models have poorer predictive levels than those based on leaf samples.

This trend was also consistent when models were created by treatment categories, in which the hormone predictions based on xylem sap samples in Table 5.1a did not achieve as high a level of accuracy as those based on freeze-dried ground leaves (Table 5.1b); the high R^2 values achieved in Table 5.1b indicate an excellent level of prediction from leaf samples. This effect could also be attributed to the fact that these are liquid samples that were injected directly into the HPLC-MS system without any previous extraction, and the higher variability between xylem sap

samples (Supplementary Figure S5.5). Refinements to the technique used for collecting xylem sap (Netting *et al.*, 2012) and concentrating the samples prior to analysis with UHPLC-HRMS could improve the accuracy of xylem sap hormone quantification. Importantly, Figure 5.5 shows that it is possible to identify different hormones at the same time to a high accuracy, as these models predicted all hormones in a row simultaneously.

Once made, these models can be applied to new ATR-FTIR spectral data in the absence of UHPLC-HRMS data to predict plant hormone concentrations. As plant hormone concentrations are a key physiological interface for modulation of plant responses in relation to examined processes, the ability to predict them rapidly and non-destructively from spectral data makes it a valuable tool for efficient physiological phenotyping. This methodology has potential for application across a range of species as key plant hormones are conserved (Blázquez *et al.*, 2020; Wang C *et al.*, 2015). ATR-FTIR spectroscopy is a rapid and non-destructive tool, which although demonstrated here using sample preparation, can also be used *in planta* (Butler *et al.*, 2015)*.* Consequently, this method could be used in the field to monitor plant hormones and other key signalling molecules produced upon the perception of environmental stress. Biomolecular indications of stress can allow for intervention before the occurrence of phenotypic change, thereby reducing waste, increasing crop yield, and maintaining quality. As can be seen from the variation in R^2 values (Table 5.1) however the accuracy of prediction varies between leaf and xylem sap and between different hormones and environments, suggesting the choice of tissue and growth environment is important when creating models, and would be improved through calibration data.

5.8 References

Ajitha B, Ashok Kumar Reddy Y, Shameer S, Rajesh KM, Suneetha Y, Sreedhara Reddy P. 2015. Lantana camara leaf extract mediated silver nanoparticles: Antibacterial, green catalyst. *J. Photochem. Photobiol. B Biol.* **149**: 84–92.

Albacete A, Ghanem ME, Martínez-Andújar C, Acosta M, Sánchez-Bravo J, Martínez V, Lutts S, Dodd IC, Pérez-Alfocea F. 2008. Hormonal changes in relation to biomass partitioning and shoot growth impairment in salinized tomato (Solanum lycopersicum L.) plants. *J. Exp. Bot.* **59**: 4119–4131.

Anagnostopoulos CJ, Liapis K, Haroutounian S, Paspatis E. 2013. Simultaneous determination of different classes of plant growth regulator in high water content agricultural products by liquid chromatography tandem mass spectrometry and time of flight mass spectrometry. *J. Liq. Chromatogr. Relat. Technol.* **36**: 315–335.

Anfang M, Shani E. 2021. Transport mechanisms of plant hormones. *Curr. Opin. Plant Biol.* **63**: 102055.

Asif A, Baig MA, Siddiqui MB. 2021. Role of Jasmonates and Salicylates in Plant Allelopathy. *Jasmonates and Salicylates Signaling in Plants*. Springer, Cham. 115–127.

Bailey J. 2013. The Japanese knotweed invasion viewed as a vast unintentional hybridisation experiment. *Heredity* **110(2)**: 105-110.

Bailey JP, Conolly AP. 2000. Prize-winners to pariahs -A history of Japanese Knotweed s.l. (Polygonaceae) in the British Isles. *Watsonia* **23**: 93–110.

Belfer S, Purinson Y, Kedem O. 1998. Surface modification of commercial polyamide reverse osmosis membranes by radical grafting: An ATR-FTIR study. *Acta Polym.* **49**: 574–582.

Bensemmane N, Bouzidi N, Daghbouche Y, Garrigues S, de la Guardia M, El Hattab M. 2021. Quantification of phenolic acids by partial least squares Fourier-transform infrared (PLS-FTIR) in extracts of medicinal plants. *Phytochem. Anal.* **32**: 206–221.

Bharath P, Gahir S, Raghavendra AS. 2021. Abscisic Acid-Induced Stomatal Closure: An Important Component of Plant Defense Against Abiotic and Biotic Stress. *Front. Plant Sci.* **12**: 324.

Blázquez MA, Nelson DC, Weijers D. 2020. Evolution of Plant Hormone Response Pathways. *Annual Review of Plant Biology* **71**: 327–353.

Bosco R, Daeseleire E, Van Pamel E, Scariot V, Leus L. 2014. Development of an Ultrahigh-Performance Liquid Chromatography–Electrospray Ionization–Tandem Mass Spectrometry Method for the Simultaneous Determination of Salicylic Acid, Jasmonic Acid, and Abscisic Acid in Rose Leaves. *J. Agric. Food Chem.* **62**: 6278–6284.

Butler HJ, Martin FL, Roberts MR, Adams S, McAinsh MR. 2020. Observation of nutrient uptake at the adaxial surface of leaves of tomato (*Solanum lycopersicum*) using Raman spectroscopy. *Anal. Lett.* **53**: 536–562.

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

Courbier S, Grevink S, Sluijs E, Bonhomme P-O, Kajala K, Wees SCM Van, Pierik R. 2020. Far-red light promotes Botrytis cinerea disease development in tomato leaves via jasmonate-dependent modulation of soluble sugars. *Plant. Cell Environ.* **43**: 2769–2781.

Davies PJ. 2010. The Plant Hormones: Their Nature, Occurrence, and Functions. *Plant Hormones. Biosynthesis, Signal Transduction, Action!*, Springer, Dordrecht. 1–15.

Dodd IC, Egea G, Davies WJ. 2008. Abscisic acid signalling when soil moisture is heterogeneous: decreased photoperiod sap flow from drying roots limits abscisic acid export to the shoots. *Plant. Cell Environ.* **31**: 1263–1274.

Falcão L, Araújo MEM. 2013. Tannins characterization in historic leathers by complementary analytical techniques ATR-FTIR, UV-Vis and chemical tests. *J. Cult. Herit.*, **14**: 499–508.

Fennell M, Wade M, Bacon KL. 2018. Japanese knotweed (*Fallopia japonica*): an analysis of capacity to cause structural damage (compared to other plants) and typical rhizome extension. *PeerJ* **6**: e5246.

Ge L, Peh CYC, Yong JWH, Tan SN, Hua L, Ong ES. 2007. Analyses of gibberellins by capillary electrophoresis–mass spectrometry combined with solid-phase extraction. *J. Chromatogr. A* **1159**: 242–249.

Geng Y, van Klinken RD, Sosa A, Li B, Chen J, Xu C-Y. 2016. The Relative Importance of Genetic Diversity and Phenotypic Plasticity in Determining Invasion Success of a Clonal Weed in the USA and China. *Front. Plant Sci.* **7**: 216.

Gfeller A, Dubugnon L, Liechti R, Farmer EE. 2010. Jasmonate biochemical pathway. *Sci. Signal.* **3(109)**: cm3.

Gorzsas A. 2020. ATR-FTIR Microspectroscopy Brings a Novel Insight Into the Study of Cell Wall Chemistry at the Cellular Level. In *Proceedings of IPSC 2019-2nd International Plant Spectroscopy Conference*. Frontiers Media SA.

Groãÿkinsky DK, Albacete A, Jammer A, Krbez P, Van der Graaff E, Pfeifhofer H, Roitsch T. 2014. A Rapid Phytohormone and Phytoalexin Screening Method for Physiological Phenotyping. *Molecular Plant* **7**: 1053–1056.

Grossmann K. 2003. Mediation of Herbicide Effects by Hormone Interactions. *J. Plant Growth Regul.* **22**: 109–122.

Heap B, Holden CA, Taylor JE, McAinsh M 2020. Crosstalk in Signalling Pathways. In *eLS*. Wiley. 1–9.

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

Herman JJ, Sultan SE. **2011**. Adaptive transgenerational plasticity in plants: Case studies, mechanisms, and implications for natural populations. *Frontiers in Plant Science* **2**: 102.

Holden CA, Bailey JP, Taylor JE, Martin FL, Beckett P, McAinsh M. 2022. Know your enemy: Application of ATR-FTIR spectroscopy to invasive species control D. Changwen, ed. *PLoS One* **17**: e0261742.

Holden CA, Morais CLM, Taylor JE, Martin FL, Beckett P, McAinsh M. 2021. Regional differences in clonal Japanese knotweed revealed by chemometrics-linked attenuated total reflection Fourier-transform infrared spectroscopy. *BMC Plant Biol. 2021 211* **21**: 1–20.

Hollingsworth ML, Bailey, JP. 2000. Evidence for massive clonal growth in the invasive weed *Fallopia japonica* (Japanese Knotweed). *Bot. J. Linn. Soc.* **133**: 463–472.

Holm AK, Elameen A, Oliver BW, Brandsæter LO, Fløistad IS, Brurberg MB. 2018. Low genetic variation of invasive Fallopia spp. in their northernmost European distribution range. *Ecol. Evol.* **8**: 755–764.

Jin N, Semple KT, Jiang L, Luo C, Zhang D, Martin FL. 2018. Spectrochemical analyses of growth phase-related bacterial responses to low (environmentally-relevant) concentrations of tetracycline and nanoparticulate silver. *Analyst* **143**: 768–776.

Karasov TL, Chae E, Herman JJ, Bergelson J. 2017. Mechanisms to Mitigate the Trade-Off between Growth and Defense. *Plant Cell* **29**: 666–680.

van Kleunen M, Bossdorf O, Dawson W. 2018. The Ecology and Evolution of Alien Plants. *Annu. Rev. Ecol. Evol. Syst.* **49**: 25–47.

Larsen DH, Woltering EJ, Nicole CCS, Marcelis LFM. 2020. Response of Basil Growth and Morphology to Light Intensity and Spectrum in a Vertical Farm. *Front. Plant Sci.* **11**: 1893.

Lavoie C. 2017. The impact of invasive knotweed species (Reynoutria spp.) on the environment: review and research perspectives. *Biol. Invasions* **19**: 2319–2337.

Lew TTS, Sarojam R, Jang I-C, Park BS, Naqvi NI, Wong MH, Singh GP, Ram RJ, Shoseyov O, Saito K, Chua NH. 2020. Species-independent analytical tools for next-generation agriculture. *Nature Plants* **6(612)**: 1408–1417.

Liu Y, Oduor AMO, Dai ZC, Gao FL, Li J, Zhang X, Yu FH. 2021. Suppression of a plant hormone gibberellin reduces growth of invasive plants more than native plants. *Oikos* **130**: 781–789.

Liu X, Renard CMGC, Bureau S, Le Bourvellec C. 2021. Revisiting the contribution of ATR-FTIR spectroscopy to characterize plant cell wall polysaccharides. *Carbohydr. Polym.* **262**: 117935.

Lowry DB, Popovic D, Brennan DJ, Holeski LM. 2019. Mechanisms of a locally adaptive shift in allocation among growth, reproduction, and herbivore resistance in Mimulus guttatus*. *Evolution (N. Y).* **73**: 1168–1181.

Manoharan B, Qi SS, Dhandapani V, Chen Q, Rutherford S, Wan JSH, Jegadeesan S, Yang HY, Li Q, Li J, Dai ZC. 2019. Gene Expression Profiling Reveals Enhanced Defense Responses in an Invasive Weed Compared to Its Native Congener During Pathogenesis. *Int. J. Mol. Sci.* **20(19)**: 4916.

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

Maruri-López I, Aviles-Baltazar NY, Buchala A, Serrano M. 2019. Intra and Extracellular Journey of the Phytohormone Salicylic Acid. *Front. Plant Sci.* **0**: 423.

Mehmood T, Liland KH, Snipen L, Sæbø S. 2012. A review of variable selection methods in Partial Least Squares Regression. *Chemom. Intell. Lab. Syst.* **118**: 62–69.

Met Office. 2019. UK Regional Climates. *https://www.metoffice.gov.uk/research/climate/maps-and-data/regional-climates/index*.

Michalet S, Rouifed S, Pellassa-Simon T, Fusade-Boyer M, Meiffren G, Nazaret S, Piola F. 2017. Tolerance of Japanese knotweed s.l. to soil artificial polymetallic pollution: early metabolic responses and performance during vegetative multiplication. *Environ. Sci. Pollut. Res.* **24**: 20897–20907.

Monaghan RM, Paton RJ, Smith LC, Drewry JJ, Littlejohn RP. 2005. The impacts of nitrogen fertilisation and increased stocking rate on pasture yield, soil physical condition and nutrient losses in drainage from a cattle-grazed pasture. *New Zeal. J. Agric. Res.* **48**: 227–240.

Morais CLM, Costa FSL, Lima KMG. 2017. Variable selection with a support vector machine for discriminating: Cryptococcus fungal species based on ATR-FTIR spectroscopy. *Anal. Methods* **9**: 2964–2970.

Morais CLM, Lima KMG. 2018. Principal Component Analysis with Linear and Quadratic Discriminant Analysis for Identification of Cancer Samples Based on Mass Spectrometry. *Artic. J. Braz. Chem. Soc* **29**: 472–481.

Morais CLM, Lima KMG, Singh M, Martin FL. 2020. Tutorial: multivariate classification for vibrational spectroscopy in biological samples. *Nat. Protoc.* **15**: 2143–2162.

Moskal P, Wesełucha-Birczyńska A, Łabanowska M, Filek M. 2019. Adaxial and abaxial pattern of Urtica dioica leaves analyzed by 2DCOS ATR-FTIR as a function of their growth time and impact of environmental pollution. *Vib. Spectrosc.* **104**: 102948.

Naqvi SMZA, Zhang Y, Ahmed S, Abdulraheem MI, Hu J, Tahir MN, Raghavan V. 2022. Applied surface enhanced Raman Spectroscopy in plant hormones detection, annexation of advanced technologies: A review. *Talanta* **236**: 122823.

Netting AG, Theobald JC, Dodd IC. 2012. Xylem sap collection and extraction methodologies to determine in vivo concentrations of ABA and its bound forms by gas chromatography-mass spectrometry (GC-MS). *Plant Methods* **8**: 1–14.

Nozahic V, Amziane S. 2012. Influence of sunflower aggregates surface treatments on physical properties and adhesion with a mineral binder. *Compos. Part A Appl. Sci. Manuf.* **43**: 1837–1849.

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

Parepa M, Fischer M, Bossdorf O. 2013. Environmental variability promotes plant invasion. *Nat. Commun.* **4**: 1–4.

Park Y, Runkle ES. 2018. Spectral effects of light-emitting diodes on plant growth, visual color quality, and photosynthetic photon efficacy: White versus blue plus red radiation. *PLoS One* **13(8)**: e0202386.

Pennisi G, Pistillo A, Orsini F, Cellini A, Spinelli F, Nicola S, Fernandez JA, Crepaldi A, Gianquinto G, Marcelis LFM. 2020. Optimal light intensity for sustainable water and energy use in indoor cultivation of lettuce and basil under red and blue LEDs. *Sci. Hortic. (Amsterdam).* **272**: 109508.

Porfírio S, Sonon R, Gomes da Silva MDR, Peixe A, Cabrita MJ, Azadi P. 2016. Quantification of free auxins in semi-hardwood plant cuttings and microshoots by dispersive liquid–liquid microextraction/microwave derivatization and GC/MS analysis. *Anal. Methods* **8**: 6089–6098.

Pradko AG, Litvinovskaya RP, Sauchuk AL, Drach S V., Baranovsky A V., Zhabinskii VN, Mirantsova T V., Khripach VA. 2015. A new ELISA for quantification of brassinosteroids in plants. *Steroids* **97**: 78–86.

Rana R, Herz K, Bruelheide H, Dietz S, Haider S, Jandt U, Pena R. 2018. Leaf Attenuated Total Reflection Fourier Transform Infrared (ATR-FTIR) biochemical profile of grassland plant species related to land-use intensity. *Ecol. Indic.* **84**: 803–810.

Richards CL, Bossdorf O, Muth NZ, Gurevitch J, Pigliucci M. 2006. Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecol. Lett.* **9**: 981–993.

Richards CL, Schrey AW, Pigliucci M. 2012. Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation M. Vellend, ed. *Ecol. Lett.* **15**: 1016–1025.

Romera-Fernández M, Berrueta LA, Garmón-Lobato S, Gallo B, Vicente F, Moreda JM. 2012. Feasibility study of FT-MIR spectroscopy and PLS-R for the fast determination of anthocyanins in wine. *Talanta* **88**: 303–310.

Rouifed S, Byczek C, Laffray D, Piola F. 2012. Invasive Knotweeds are Highly Tolerant to Salt Stress. *Environ. Manage.* **50**: 1027–1034.

Saleem M, Fariduddin Q, Castroverde CDM. 2021. Salicylic acid: A key regulator of redox signalling and plant immunity. *Plant Physiol. Biochem.* **168**: 381–397.

Santo P. 2017. Assessing diminution in value of residential properties affected by Japanese Knotweed. *J. Build. Surv. Apprais. Valuat.* **6(11)**: 211-221.

Sharma S, Uttam KN. 2018. Early Stage Detection of Stress Due to Copper on Maize (Zea mays L.) by Laser-Induced Fluorescence and Infrared Spectroscopy. *J. Appl. Spectrosc.,* **85**: 771–780.

Shivu B, Seshadri S, Li J, Oberg KA, Uversky VN, Fink AL. 2013. Distinct β-Sheet Structure in Protein Aggregates Determined by ATR–FTIR Spectroscopy. *Biochemistry 52***(31)**: 5176-5183

Šimura J, Antoniadi I, Široká J, Tarkowská D, Strnad M, Ljung K, Novák O. 2018. Plant Hormonomics: Multiple Phytohormone Profiling by Targeted Metabolomics. *Plant Physiol.* **177**: 476.

Skolik P, McAinsh MR, Martin FL. 2019a. ATR-FTIR spectroscopy non-destructively detects damage-induced sour rot infection in whole tomato fruit. *Planta* **249**: 925–939.

Skolik P, Morais CLM, Martin FL, McAinsh MR. 2019b. Determination of developmental and ripening stages of whole tomato fruit using portable infrared spectroscopy and Chemometrics. *BMC Plant Biol.* **19**: 236.

Smith H. 1982. Light quality, photoperception, and plant strategy. *Annu. Rev. Plant Physiol.,* **33**: 481–518.

Sołtysiak J. 2020. Heavy metals tolerance in an invasive weed (*Fallopia japonica*) under different levels of soils contamination. *Journal of Ecological Engineering* **21**: 81–91.

Spalding K, Bonnier F, Bruno C, Blasco H, Board R, Benz-de Bretagne I, Byrne HJ, Butler HJ, Chourpa I, Radhakrishnan P, Baker MJ. 2018. Enabling quantification of protein concentration in human serum biopsies using attenuated total reflectance – Fourier transform infrared (ATR-FTIR) spectroscopy. *Vib. Spectrosc.* **99**: 50–58.

Strong R, Martin FL, Jones KC, Shore RF, Halsall CJ. 2017. Subtle effects of environmental stress observed in the early life stages of the Common frog, Rana temporaria. *Sci. Rep.* **7**: 1–13.

Talari ACS, Martinez MAG, Movasaghi Z, Rehman S, Rehman IU. 2017. Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Appl. Spectrosc. Rev.* **52**: 456–506.

Traoré M, Kaal J, Martínez Cortizas A. 2018. Differentiation between pine woods according to species and growing location using FTIR-ATR. *Wood Sci. Technol.* **52**: 487–504.

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

Urcelay C, Austin AT. 2020. Exotic plants get a little help from their friends. *Science* **368**: 934–936.

Wagner H, Liu Z, Langner U, Stehfest K, Wilhelm C. 2010. The use of FTIR spectroscopy to assess quantitative changes in the biochemical composition of microalgae. *J. Biophotonics* **3**: 557–566.

Wang C, Liu Y, Li S-S, Han G-Z. 2015. Insights into the Origin and Evolution of the Plant Hormone Signaling Machinery. *Plant Physiol.* **167**: 872–886.

van der Weijde T, Huxley LM, Hawkins S, Sembiring EH, Farrar K, Dolstra O, Visser RGF, Trindade LM. 2017. Impact of drought stress on growth and quality of miscanthus for biofuel production. *GCB Bioenergy* **9**: 770–782.

Wolters H, Jürgens G. 2009. Survival of the flexible: Hormonal growth control and adaptation in plant development. *Nat. Rev. Genet.* **10**: 305–317.

Zhang Y-Y, Parepa M, Fischer M, Bossdorf O. 2016. Epigenetics of colonizing species? A study of Japanese knotweed in Central Europe. In *Barrett SCH, Colautti RI, Dlugosch KM, Rieseberg LH (Eds) Invasion Genetics*. Chichester, UK: John Wiley & Sons, Ltd. 328–340.

Zheng W, Wang SY. 2001. Antioxidant Activity and Phenolic Compounds in Selected Herbs. *J. Agric. Food Chem.*, **49**, 5165–5170.

Zhu J, Agyekum AA, Kutsanedzie FYH, Li H, Chen Q, Ouyang Q, Jiang H. 2018. Qualitative and quantitative analysis of chlorpyrifos residues in tea by surface-enhanced Raman spectroscopy (SERS) combined with chemometric models. *LWT* **97**: 760–769.

Zou T, Huang C, Wu P, Ge L, Xu Y. 2020. Optimization of Artificial Light for Spinach Growth in Plant Factory Based on Orthogonal Test. *Plants* **9**: 490.

5.9 Supporting Information

- **Figure S5.1:** Spectra from a) 'Light' (LC, LD, LN, LLN) b) 'Shade' (SC, SD, SN, SLN) cabinets, providing red: far-red ratios of 5.6 and 0.4 respectively.
- **Table S5.1:** Lighting conditions within each Snijder cabinet.
- **Table S5.2:** Reagents used for Hoagland's solution.
- **Table S5.3:** Hormone descriptions and molecular ion masses.
- **Table S5.4:** SVM parameters for classification.
- **Figure S5.2:** Chromatogram and mass spectra for the hormone salicylic acid.
- **Table S5.3:** Hormone descriptions and molecular ion masses.
- **Figure S5.3:** a) Raw and b) pre-processed class means spectra in the fingerprint region from xylem sap, c) Raw and d) pre-processed (Savitzky–Golay 2nd differentiation, n=9, and vector normalisation) class means spectra in the fingerprint region from freeze-dried ground leaves.
- **Table S5.5:** PCA-loadings and biomarkers: key wavenumbers and compounds, which differentiate spectral profiles of plants from different growth conditions for both xylem sap and freeze-dried ground sample types.
- **Figure S5.4:** Loadings from spectra of a) xylem sap and b) freeze-dried ground leaf samples.
- **Figure S5.5:** Hormone profiles from xylem sap in ng·ml^{-1} sap for a) 1-amino-cyclopropanecarboxylic acid (ACC), b) trans-Zeatin (tZ), c) isopentyl-adenine (iP), d) salicylic acid (SA), e) abscisic acid (ABA), f) jasmonic acid (JA), g) gibberellin A1 (GA$_1$), gibberellin A4 (GA$_4$), gibberellic acid (GA$_3$), *trans*-zeatin riboside (tZR), and indole-3-acetic acid (IAA).
- **Figure S5.6:** Hormone profiles from freeze-dried ground leaves ng·g^{-1} dry weight for a) 1-amino-cyclopropanecarboxylic acid (ACC), b) trans-Zeatin (tZ), c) isopentyl-adenine (iP), d) salicylic acid (SA), e) abscisic acid (ABA), f) jasmonic acid (JA), g) gibberellin A1 (GA$_1$), gibberellin A4 (GA$_4$), gibberellic acid (GA$_3$), *trans*-zeatin riboside (tZR), and indole-3-acetic acid (IAA).
- **Figure S5.7:** PLS regression graphs for prediction of plant hormones from xylem sap.
- **Figure S5.8:** PLSR loadings graphs for prediction of plant hormones from xylem sap.
- **Figure S5.9:** PLS regression graphs for prediction of plant hormones from freeze-dried ground leaves.
- **Figure S5.10:** PLSR loadings graphs for prediction of plant hormones from freeze-dried ground leaves.
- **Table S5.6:** PLSR R^2 and loadings summary for predictions of plant hormone concentrations.
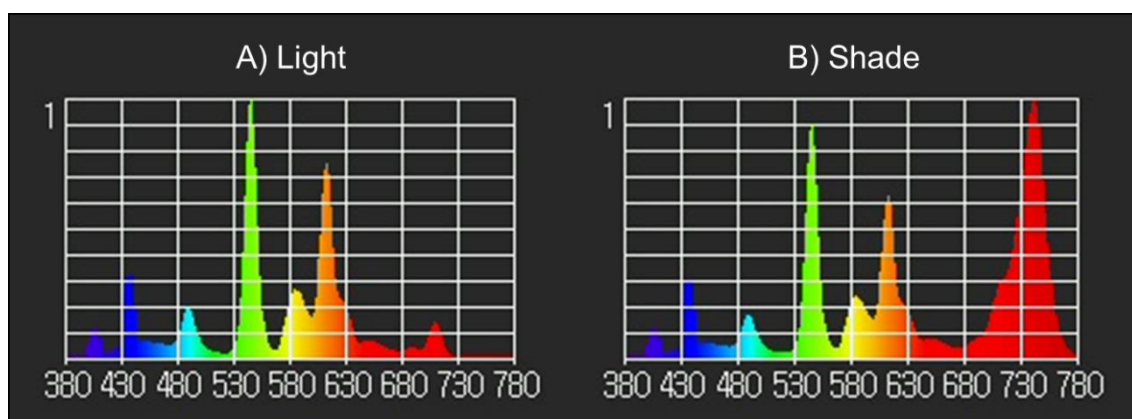
Figure S5.1: Spectra from a) 'Light' b) 'Shade' cabinets, providing red: far-red ratios of 5.6 and 0.4 respectively.

Table S5.1: Lighting conditions within each Snijder cabinet.

Light Quality	'Light' Groups:	'Shade' Groups:
PFD-R$_{(700-780\ nm)}$	72.51	49.28
PFD-FR$_{(600-700\ nm)}$	12.89	116.5
photosynthetic photon flux density	189.8	124.7
PFD-UV$_{(380-400\ nm)}$	0.5677	0.4402
PFD-B$_{(400-500\ nm)}$	33.93	21.58
PFD-G$_{(500-600\ nm)}$	83.40	53.87
peak wavelength λp / nm	545	741
peak wavelength value λpV / $mWm^{-2}nm^{-}$	827.7	576.0
Irradiance	43.2	45.8
Illuminance/ lux.	15128	9617

Table S5.2: Reagents used for Hoagland's solution. Full strength Hoagland's solution was made using 100 mL of solution A, 100 mL of solution B and 10 mL of solution C in 10 L of deionised water.

Solution	Reagent	Concentration/ gL^{-1}
A (100 mL)	NH_4NO_3	8.000
	$Ca(NO_3)_2.4H_2O$	82.600
	KNO_3	35.700
B (100 mL)	KNO_3	5.000
	KH_2PO_4	27.400
	$MgSO_4.7H_2O$ *added first	24.600
	$MnSO_4.5H_2O$	0.053
	H_3BO_3	0.140
	$CuSO_4.5H_2O$	0.015
	$(NH_4)_6Mo_7O_{24}.4H_2O$	0.008
	$ZnSO_4.7H_2O$	0.060
C (10 mL)	Fe-EDTA	36.71

Figure S5.2: Chromatogram and mass spectra for the hormone salicylic acid.

Table S5.3: Hormone descriptions and molecular ion masses.

Hormone class	Abbreviation	Hormone	Molecular formula	[M-H]⁻
Ethylene precursor	ACC	1-Aminocyclopropane-1-carboxylic acid	$C_4H_7NO_2$	100.04040
Cytokinins	t-Z	*trans*-Zeatin	$C_{10}H_{13}N_5O$	218.10473
	t-ZR	*trans*-Zeatin riboside	$C_{15}H_{21}N_5O_5$	350.14699
	iP	Isopentenyladenine	$C_{10}H_{13}N_5$	202.10982
Gibberellins	GA1	Gibberellin A1	$C_{19}H_{24}O_6$	347.15001
	GA3	Gibberellin A3	$C_{19}H_{22}O_6$	345.13436
	GA4	Gibberellin A4	$C_{19}H_{24}O_5$	331.15510
Auxins	IAA	Indole-3-acetic acid	$C_{10}H_9NO_2$	174.05605
Abscisic acid	ABA	Abscisic acid	$C_{15}H_{20}O_4$	263.12888
Salicylates	SA	Salicylic acid	$C_7H_6O_3$	137.02442
Jasmonates	JA	Jasmonic acid	$C_{12}H_{18}O_3$	209.11832
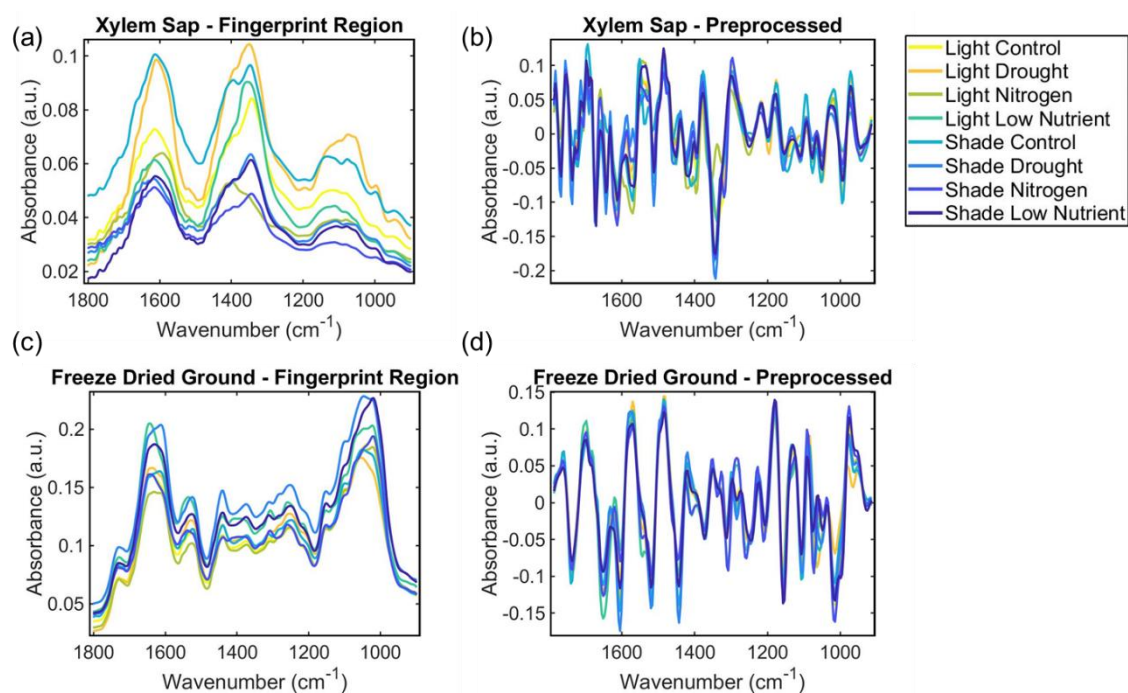
Figure S5.3: a) Raw and **b)** pre-processed class means spectra in the fingerprint region from xylem sap, **c)** Raw and **d)** pre-processed (Savitzky–Golay 2nd differentiation, *n*=9, and vector normalisation) class means spectra in the fingerprint region from freeze-dried ground leaves. Each class is grouped by treatment; Light Control (LC), Light Drought (LD), Light Nitrogen (LN), Light Low Nitrogen (LLN), Shade Control (SC), Shade Drought (SD), Shade Nitrogen (SN) and Shade Low Nitrogen (SLN).

Table S5.4: SVM parameters for classification.

	Cost	Gamma (γ)	Number of support vectors (N_{SV})
Xylem Sap	31.6228	3.1623	314
Freeze-dried ground leaves	100	3.1623	194

Table S5.5: PCA-loadings and biomarkers: key wavenumbers and compounds, which differentiate ATR-FTIR spectral profiles of plants from different growth conditions for both xylem sap and freeze-dried ground sample types.

Sample Type	Wavelength / cm	Tentative Molecular Assignment	Reference
Xylem sap	1771	v_1 symmetric stretching of C=O in the carboxylic acid of pectin or ester bond of triacylglycerol	(Nozahic and Amziane, 2012)
	1663	The N-C=O group of proteins. Amide I vibrations, specifically associated with disordered secondary structures or turns.	(Belfer *et al.*, 1998; Shivu *et al.*, 2013)
	1612	Amide I	(Jin *et al.*, 2018)
	1555	C-N stretching and N-H bending (Amide II vibration); C-O-O⁻ asymmetric stretching of proteins and glutamate	(Moskal *et al.*, 2019)
	1516	Amide II vibrations of proteins	(Talari *et al.*, 2017)
	1346	Cellulose	(Gorzsas, 2020)
	1312	Amide III vibrations of proteins	(Talari *et al.*, 2017)

	1177	C–O stretch vibration of tannins	(Falcão and Araújo, 2013)
	1053	Starch, νC–O and δC–O of carbohydrates	(Talari *et al.*, 2017; Jin *et al.*, 2018)
	991	C–O ribose	(Morais *et al.*, 2017)
Freeze-dried ground leaves	1744	Ester C=O stretch: triglycerides	(Talari *et al.*, 2017)
	1663	The N-C=O group of proteins. Amide I vibrations, specifically associated with disordered secondary structures or turns.	(Belfer *et al.*, 1998; Shivu *et al.*, 2013)
	1566	N-H bending; C-N stretching (Amide II band of proteins)	(Rana *et al.*, 2018)
	1443	Pectin	(Sharma and Uttam, 2018)
	1350	Phosphodiester stretching bands region (for absorbances due to starch)	(Talari *et al.*, 2017)
	1315	Cellulose	(Sharma and Uttam, 2018)
	1161	C–OH groups of serine, threonine and tyrosine of proteins, C-O stretching and hydrogen bonding	(Talari *et al.*, 2017)
	1057	Stretching C–O deoxyribose	(Talari *et al.*, 2017)
	1022	Starch	(Talari *et al.*, 2017)
	980	C-OH stretching of secondary alcohols and C-O-C vibrations of polysaccharides	(Ajitha *et al.*, 2015)
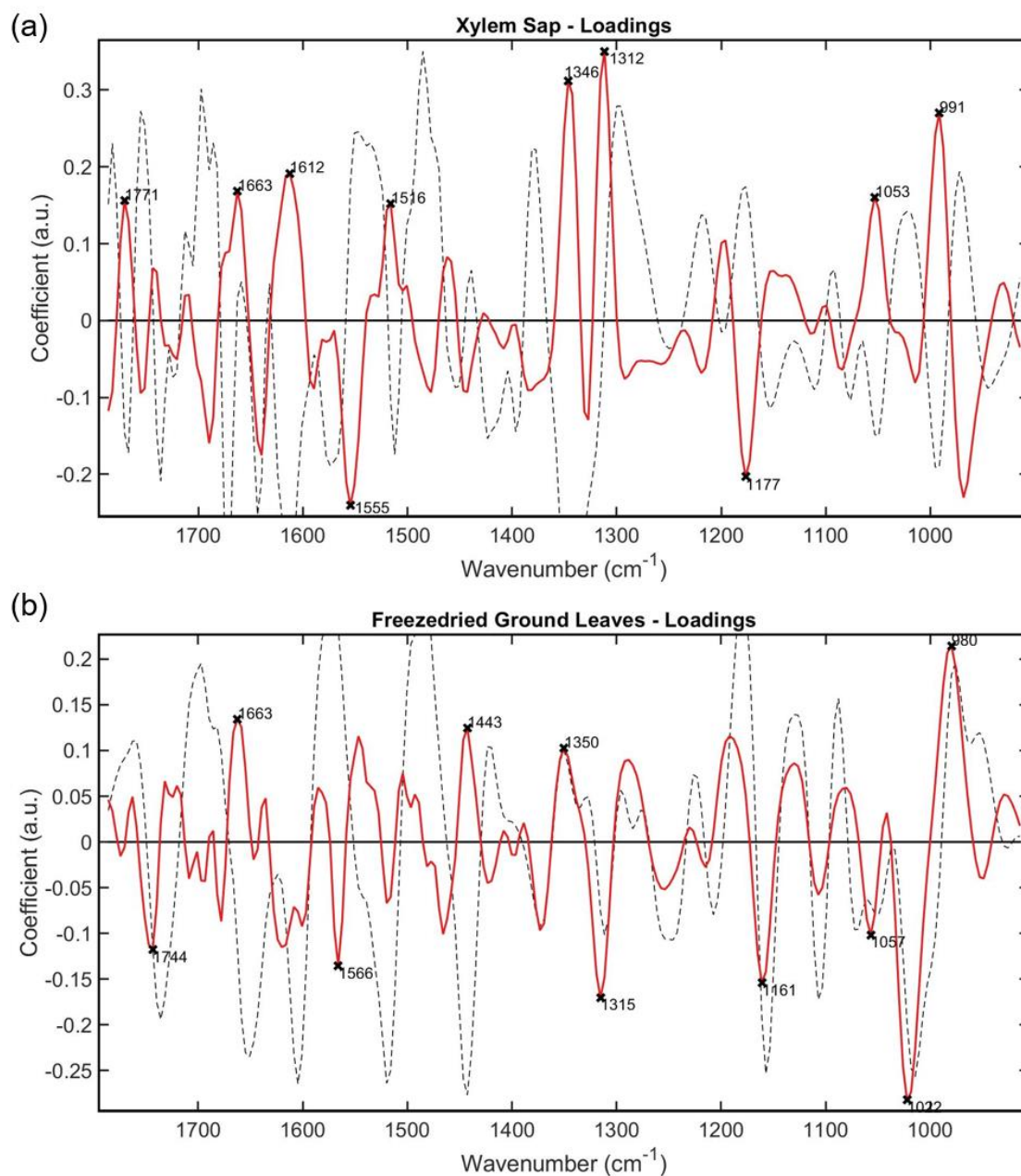
Figure S5.4: Loadings from spectra of **a)** xylem sap and **b)** freeze-dried ground leaf samples. These are the key wavenumbers which differentiate spectral profiles of different treatment groups from one another.
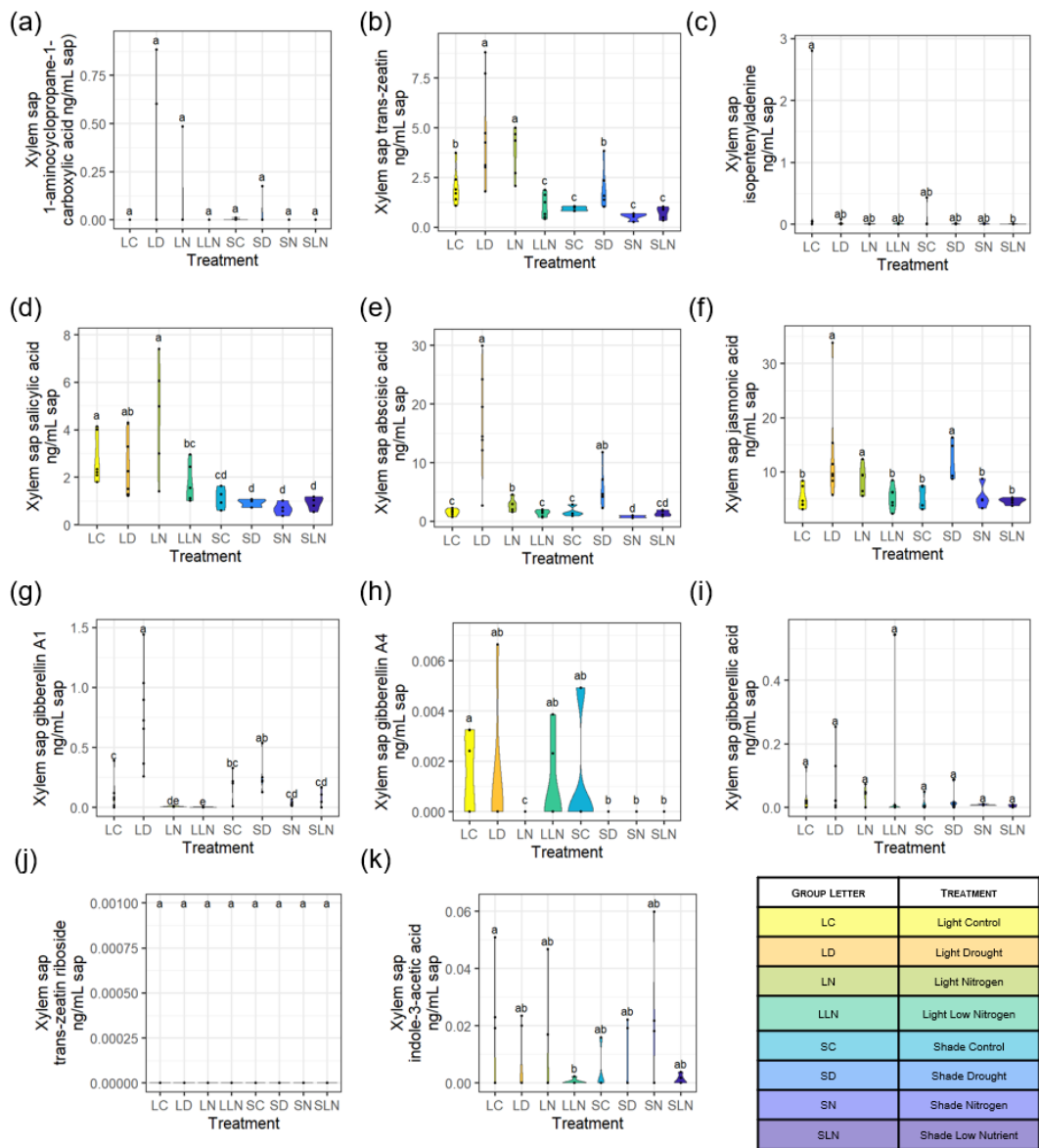
Figure S5.5: Hormone profiles from xylem sap measured using UHPLC– HRMS in ng·ml^{-1} sap for **a)** 1-amino-cyclopropanecarboxylic acid (ACC), **b)** *trans*-Zeatin (tZ), **c)** isopentyl-adenine (iP), **d)** salicylic acid (SA), **e)** abscisic acid (ABA), **f)** jasmonic acid (JA), **g)** gibberellin A1 (GA$_1$), **h)** gibberellin A4 (GA$_4$), **i)** gibberellic acid (GA$_3$), **j)** *trans*-zeatin riboside (tZR), and **k)** indole-3-acetic acid (IAA). ABA concentration was highest in the drought categories at approximately 17, 7, and a range of 1-3 ng·ml^{-1} sap for LD, SD, and the other categories respectively. Shade plants had lower xylem SA levels than light ones, in the range of 0.7-1.1 ng·ml^{-1} sap compared with 1.6-4.5 ng·ml^{-1} sap respectively. Xylem sap levels of GA$_1$ were approximately three times higher in LD than most other treatment groups, although this was not significantly different to the other drought category, SD, due to high variation.
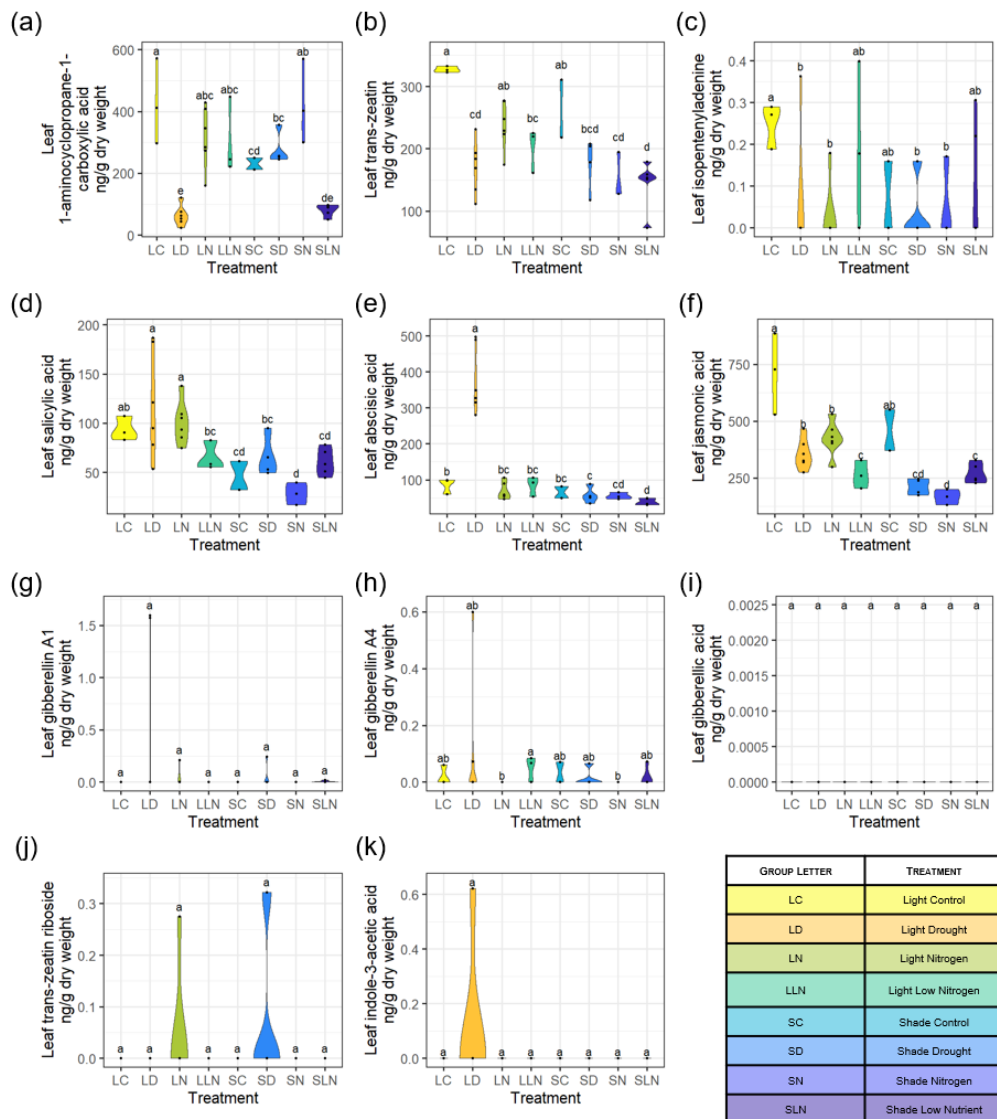
Figure S5.6: Hormone profiles from freeze-dried ground leaves measured using UHPLC– HRMS in ng·g^{-1} dry weight for **a)** 1-amino-cyclopropanecarboxylic acid (ACC), **b)** *trans*-Zeatin (tZ), **c)** isopentyl-adenine (iP), **d)** salicylic acid (SA), **e)** abscisic acid (ABA), **f)** jasmonic acid (JA), **g)** gibberellin A1 (GA$_1$), **h)** gibberellin A4 (GA$_4$), **i)** gibberellic acid (GA$_3$), **j)** *trans*-zeatin riboside (tZR), and **k)** indole-3-acetic acid (IAA). Leaf ABA levels (Figure S5.5) were approximately quadruple in LD than those of the other categories. Leaf tZ was approximately 4.5-fold higher in LC than SLN. Leaf JA concentration was significantly higher in the light control group LC (~710 ng·g^{-1} dry weight) compared to all other groups (ranging 170-420 ng·g^{-1} dry weight), except the shade control group SC (~460 ng·g^{-1} dry weight). The highest iP hormone concentration was found in leaves of category LC, at 0.25 ng·g^{-1} dry weight. This value was significantly higher compared to groups LD, LN, SD, SN (ranging 0.03-0.6 ng·g^{-1} dry weight), with the other groups falling in between.

Figure S5.7: PLS regression graphs for prediction of plant hormones from xylem sap.

Figure S5.8: PLSR loadings graphs for prediction of plant hormones from xylem sap.

Figure S5.9: PLS regression graphs for prediction of plant hormones from freeze-dried ground leaves.

Figure S5.10: PLSR loadings graphs for prediction of plant hormones from freeze-dried ground leaves.

Sample Type	Hormone	R2	Loadings
Leaf	ABA	0.839	1732, 1678, 1612, 1566, 1547, 1339,1323, 1304, 1285, 1261, 1230, 1211, 1196, 1173, 1149, 1134, 1115, 1099, 1084, 1065, 1049, 1034, 995, 972
	tZ	0.824	1605, 1516, 1462, 1443, 1420, 1285, 1265, 1173, 1150, 1130, 1095, 1080, 1061, 1045, 1034, 1018, 1007, 995, 972
	SA	0.842	1636, 1605, 1582, 1562, 1547, 1516, 1497, 1462, 1443, 1416, 1339, 1323, 1304, 1285, 1265, 1231, 1215, 1196, 1095, 1076, 1061, 1045, 1038, 1018, 1007, 995, 976, 914
	JA	0.807	1632, 1605, 1516, 1462, 1443, 1285, 1172, 1150, 1130, 1095, 1080, 1061, 1045, 1034, 1018, 1007, 995, 972
	ACC	0.830	1605, 1520, 1497, 1462, 1443, 1420, 1285, 1265, 1173, 1149, 1095, 1080, 1061, 1018, 1007, 995, 972
Xylem	ABA	0.674	1755, 1690, 1643, 1570, 1547, 1520, 1512, 1339, 991
	tZ	0.636	1767, 1736, 1713, 1690, 1678, 1570, 1531, 1520, 1358, 1339, 991
	SA	0.544	1782, 1767, 1751, 1740, 1713, 1690, 1578, 1570, 1547, 1531, 1520, 1512, 1373, 1358, 1339, 1092, 991
	JA	0.477	1798, 1778, 1763, 1747, 1732, 1678, 1663, 1647, 1636, 1555, 1543, 1524, 1462, 1439, 1389, 1369, 1327, 1177
	GA1	0.681	1720, 1705, 1693, 1643, 1574, 1566, 1555, 1543, 1535, 1524, 991

Best
0.842

0.477
Worst

Table S5.6: PLSR R^2 and loadings summary for predictions of plant hormone concentrations.

6 Attenuated total reflection Fourier-transformed infrared spectroscopy reveals environment specific phenotypes in clonal Japanese knotweed

Claire A Holden[1*], Jane E Taylor[1], Francis L. Martin[2], Paul Beckett[3], Martin McAinsh[1]

1. Lancaster Environment Centre, Lancaster University, LA1 4YQ, UK
2. Biocel Ltd, Hull, HU10 7TS, UK
3. Phlorum Ltd, Brighton, BN2 6AH, UK

* Corresponding author: Claire A Holden, c.holden6@lancaster.ac.uk

6.1 Abstract

Introduction: Biological invasions are a consequence of increasing globalisation and lead to negative environmental and socio-economic impacts. The problematic invasive species, Japanese knotweed (*R. japonica* var*. japonica*), has a wide geographical distribution and can occupy a variety of environments. The potential for attenuated total reflectance Fourier transform infrared (ATR-FTIR) spectroscopy and chemometrics to differentiate between Japanese knotweed plants grown in different regions based on their spectral profiles has been shown here previously (Chapter 3). However, the contribution of environment to the differences in the spectral signature of plants remains unclear. In this study, the influence of environment on the spectral profiles of clonal Japanese knotweed grown under different conditions has been examined to gain insight into how Japanese knotweed plants respond to and colonise varied environmental habitats.

Methods: Plants were grown in climate-controlled cabinets in which the red: far-red light ratio, and the availability of water, nitrogen, and micronutrients were manipulated creating eight different environments. The impacts of the growth environment on plant growth and photosynthetic parameters, as well as ATR-FTIR spectral profiles, was determined. Chemometric techniques were applied to explore the data, including principal component analysis (PCA), linear discriminant analysis, support vector machines (SVM) and partial least squares regression. The key wavenumbers responsible for spectral differences were identified with PCA loadings, and molecular biomarkers were assigned. Partial least squared regression (PLSR) of spectral absorbance and root water potential data was used to create a predictive model for root water potential.

Results: Spectra from plants grown in different environments were successfully differentiated using ATR-FTIR spectroscopy coupled with SVM. The biomarkers highlighted through PCA loadings corresponded to several molecules, most

commonly cell wall carbohydrates, suggesting that these wavenumbers could be consistent indicators of plant stress across species. Of the environmental variables tested, red: far-red light ratio (R:FR) most affected the ATR-FTIR spectra of intact dried leaf material, as demonstrated by the spectral PCA clustering pattern and PLSR. PLSR prediction of root water potential achieved an R^2 of 0.8, compared with the ideal of $R^2=1$, supporting the potential use of ATR-FTIR spectrometers as sensors for prediction of plant physiological parameters.

Conclusion: This study shows that Japanese knotweed exhibits specific environmentally induced phenotypes which are reflected in measurable differences in their ATR-FTIR spectra. This suggests that the high degree of plasticity exhibited by Japanese knotweed to environmental factors is reflected in key biomolecular changes that may contribute to its success as an invasive species. Light quality (R:FR) appears critical in defining the growth and spectral response of Japanese knotweed to environment. Many of the biomarkers identified, including those for cell wall carbohydrates, are conserved between species suggesting that they function as indicators of plant-environment interactions including abiotic stress responses and plant health.

6.2 Keywords

Cluster Analysis, Fourier Transform Infrared Spectroscopy, Introduced species, Japanese Knotweed, Plants, Phenotypic Plasticity, Principal Component Analysis, Support Vector Machine.

6.3 Introduction

In the current period of unprecedented global environmental change, the negative effects of invasive alien species (IAS) are compounded (Diez *et al.*, 2012; Parepa *et al.*, 2013; Head, 2017; Bellard *et al.*, 2018). They are a major driver of ecosystem degradation (Chinchio *et al.*, 2020; Pyšek *et al.*, 2020), and can trigger complex socio-

economic problems affecting human livelihoods and well-being (Pejchar & Mooney, 2009; Bradshaw *et al.*, 2016; Shackleton *et al.*, 2019). Effective establishment in a non-native environment by IAS is partly dependent on their ability to compete with native flora for key resources (light, nutrients, and water), the availability of which is limited (Davis *et al.*, 2000; Morris *et al.*, 2011). Human interference is creating new environments such as those with fluctuating water availability (Nicotra *et al.*, 2010) and high nitrogen ecosystems (Fowler *et al.*, 2013). Whilst native plants may be more able to withstand drought (Valliere *et al.*, 2019), the versatility of IAS frequently earns them the advantage. As opportunists, a non-uniform distribution of nutrients often increases the competitive edge of invasive plants (Valliere *et al.*, 2019). Japanese knotweed is an exemplar opportunistic invasive species, able to take advantage of fluctuating resources, perhaps because its nitrogen-use efficiency is superior to that of native plants (Parepa *et al.*, 2019). Additionally, increased nitrogen availability also exacerbates the impacts of invasive species on non-native species in aboveground competition (Wilson & Tilman, 1991), although nitrogen does not influence the consequences in belowground competition (Broadbent *et al.*, 2018). As an economically important weed, data on the growth response of Japanese knotweed under differing environmental conditions is of interest and was recently listed as a point of action in a Science and Technology Committee Report (Science and Technology Committee - House of Commons, 2019). Invasive species may also respond to abiotic stress differently than native species, for example, Japanese knotweed is known to be very tolerant to abiotic stress, occupying extreme environments such as salt marshes (Rouifed *et al.*, 2012) and metal-polluted soil (Michalet *et al.*, 2017; Sołtysiak, 2020). It has successfully achieved a wide geographical distribution and colonises diverse habitats such as riparian wetlands, urban transport courses, and coastal areas (Richards *et al.*, 2012; Zhang *et al.*, 2016). Despite this variation in habitat, it exhibits minimal genetic variation in Central Europe (Zhang *et al.*, 2016), Norway (Holm *et al.*, 2018) and the USA (Richards *et al.*, 2012), and exists as a female clone in the United Kingdom from a single introduction

(Bailey & Conolly, 2000; Hollingsworth & Bailey, 2000). Phenotypic plasticity (Parker *et al.*, 2003; Bossdorf *et al.*, 2005; Richards *et al.*, 2006, 2008; Van Kleunen *et al.*, 2010; Pichancourt & van Klinken, 2012; Hagenblad *et al.*, 2015; Geng *et al.*, 2016), efficient resource partitioning (Price *et al.*, 2002) and vegetative regeneration (Bímová *et al.*, 2003) allow clonal species such as Japanese knotweed to take advantage of a wide ecological niche regardless of their low genetic diversity.

A spectroscopic method, attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy, followed by chemometrics has been successfully applied to investigate plant response to abiotic (McNear *et al.*, 2010; Euring *et al.*, 2012; Buitrago *et al.*, 2016; Butler *et al.*, 2017, 2020; Usman *et al.*, 2019) and biotic factors (Skolik *et al.*, 2019a), as well as for monitoring plant health and development (Butler *et al.*, 2015a; Skolik *et al.*, 2019b). In combination with chemometric techniques, such as support vector machines, the multivariate spectral dataset produced through infrared spectroscopy allows the rapid, marker-free, non-destructive analysis of biological samples (Morais *et al.*, 2020). As previously shown in Chapter 3, the mid-infrared spectral profiles of plants from different growing regions can be distinguished with chemometric techniques despite the lack of genetic diversity in Japanese knotweed (Holden *et al.*, 2021). The relation between climate and differences in mid-infrared spectral profiles has allowed the differentiation of plants, pollen, and plant products from different growing areas in numerous species (Ruoff *et al.*, 2006; Bağcıoğlu *et al.*, 2017; Traoré *et al.*, 2018; Gordon *et al.*, 2019; Zeghoud *et al.*, 2021; Holden *et al.*, 2021). Mid-infrared spectroscopy uses the absorption pattern of mid-infrared light by a sample to gain biochemical information. Functional groups within a sample vibrate upon absorption of light energy at specific wavelengths (2.5–25 μm wavelengths or 4000–400 cm^{-1} wavenumbers), these wavelengths are then absent from the spectra (Morais *et al.*, 2020). Biological materials preferentially absorb light in the region of 1800–900 cm^{-1} wavenumbers, known as the fingerprint region, which includes important

information about key biomolecules such as lipids, proteins, nucleic acids, and carbohydrates (Morais *et al.*, 2020).

Although the biochemical differences induced by plant-environment interactions can be detected by spectral acquisition and allow discrimination between plants of different growing environments, it is not known which aspects of environmental conditions have the greatest impact on plant spectral profiles. Here, a controlled study was performed to unpick which environmental factors have the greatest influence over spectral profiles of Japanese knotweed grown in eight unique environments which differ in red: far-red light ratio (R:FR), and the availability of water, nitrogen, and micronutrients, whilst maintaining comparative temperature, humidity, and photoperiod. This study indicated that the light environment, in terms of the R:FR ratio, plays a key role in shaping the spectral profile of Japanese knotweed. Spectral differences between plants of different treatment groups allowed the identification of several biomarkers of environmental effects, highlighting changes in compounds such as cell wall carbohydrates. This raises the intriguing possibility that wavenumbers which are conserved between species could be used as indicators of plant health under abiotic stress.

6.4 Materials and methods

6.4.1 Plant growth

Ninety fragments of rhizome (10-50 g, volume 2-58 cm^3) were planted in fertilized organic loam (John Innes No. 1, J. Arthur Bowers, UK) in cylindrical pots designed to tightly fit in a pressure chamber (Soil Moisture Equipment Corp., Santa Barbara, CA, USA) measuring 6.5 cm in diameter and 23 cm in length, with a volume of 763.2 cm^3, and featured a stainless-steel mesh (0.7 mm aperture) at the base to assist drainage. Pots were placed in one of two climate-controlled cabinets (Microclima 1750, Snijders Scientific BV, Netherlands) at 80% humidity, 16 h of photoperiod, and

19/11 °C day/night temperature where treatments were applied and plants were grown for a total of fifty days before harvesting. The long photoperiod and temperature range were selected to simulate an average British summer in the areas Japanese knotweed usually colonises, using a comparison of temperature maps from the Met Office (Met Office, 2019) and a distribution map of Japanese knotweed in the British Isles (Bailey, 2013).

6.4.2 Treatments

Rhizome fragments were divided into eight treatment groups to give an even split of rhizome masses in each group. The treatments applied were: Light Control 'LC', Light Drought 'LD', Light Nitrogen 'LN', Light Low Nutrient 'LLN', Shade Control 'SC', Shade Drought 'SD', Shade Nitrogen 'SN', and Shade Low Nutrient 'SLN'. In both cabinets, the light emitted from the two high-pressure sodium lamps (SON-T 400 W, Philips Lighting, Eindhoven, The Netherlands) was reduced using a LEE 209 filter (LEE Filters Worldwide, Andover, Hampshire, UK). In one cabinet, a matrix of far-red LEDs (EPILEDS, 740-745 nm) distributed in five rows 30 cm apart was used to decrease the red: far-red ratio (R:FR) to simulate shading, see Supplementary Figure S6.1 for the spectrum produced by the LEDs alone. Wavelengths emitted were measured using an UPRtek (Taiwan) PG100N light spectrometer. The combined light conditions (see Supplementary Table S6.1) resulted in a 'light' treatment with a R:FR of 5.6 and a 'shade' treatment with a R:FR of 0.4 (see Supplementary Figure S6.1 for the spectral profile). Plants were redistributed weekly within each cabinet to minimise positional effects from the LED matrix pattern. The R:FR of natural sunlight during the day is approximately 1.15 (Smith H, 1982) and the R:FR of 0.4 in the shade treatment was chosen to replicate that found within vegetative canopies (Smith H, 1982). In both cases, the photosynthetic photon flux density (PPFD) was between 124.7 and 189.8 $\mu mol \cdot m^{-2} \cdot s^{-1}$ which is typical of growth cabinet studies (Park & Runkle, 2018; Zou $et\ al.$, 2020; Pennisi $et\ al.$, 2020; Larsen $et\ al.$, 2020).

Plants were regularly provided with water (75 cm^3 pot^{-1} 48 h^{-1}), apart from LD and SD in which water was withheld for 7 days prior to harvest. Once a week, four groups (LC, LD, SC, SD) were watered with 75 mL Hoagland solution to provide both nitrogen and micronutrients, see Supplementary Table S6.2 for details. LN and SN were fed with the commonly used agricultural dose of 50 kg ha^{-1} year^{-1} (Monaghan *et al.*, 2005); this was scaled down for a pot diameter of 6.2 cm and applied across a split-dose at 21 and 23 days to prevent leaching. Groups LLN and SLN were provided only with water and received no additional nitrogen or micronutrients.

6.4.3 Physiological and photosynthetic measurements

The following non-destructive measurements were taken the day prior to harvest: shoot height, number of leaves, stem diameter using a Vernier calliper, and chlorophyll concentration using Apogee chlorophyll concentration meter (Apogee Instruments Inc, Logan, Utah, USA). Preliminary observations of well-watered plants showed that stomatal conductance was maximal during the interval four to eight hours following the light switch on in the Snijder cabinets and dropped sharply after ten hours of light exposure, therefore, this interval was selected for measurements. Before measurements were taken, whole plant transpiration was calculated by comparison of initial and final pot weights. The pot surface was covered with duct tape to minimize soil evaporation and weighed twice on a precision balance to 0.01 g (Adventurer Pro AV4102, Ohaus, Thetford, UK), with the two readings a minimum of 30 minutes apart. Stomatal conductance was measured using an AP4 porometer (Delta-T Devices Ltd, Burwell, Cambridge, UK). A MultispeQ Beta (Kuhlgert *et al.*, 2016) was used to measure leaf thickness and the following photosynthesis data; Phi2, PhiNPQ, PhiNO, linear electron flow, leaf temp differential, NPQt, ECSt mAU, gH$^+$, vH$^+$, PS1 active centres, PS1 open centres, PS1 over reduced centres, PS1 oxidised centres (see Supplementary Figures S6.2 and S6.3).

A total of three leaves were excised from each plant for analysis 4-8 h into the photoperiod in order to fall within a stable period of the plants' circadian rhythm. The third leaf down was selected for bio-spectroscopy of dried whole leaves. Whole leaves were dried at 37 °C for one week and stored in a dry airtight container at room temperature before analysis using ATR-FTIR spectroscopy. The fourth leaf down was placed in a humidified plastic bag and taken immediately to the pressure chamber for determination of leaf water potential (Ψ_{leaf}). Following this, the plant was de-topped and the whole pot inserted into a pressure chamber (Model 3000F01 Plant Water Status Console, Soil Moisture Equipment Corp., Santa Barbara, CA, USA) with the stem protruding for determination of root water potential (Ψ_{root}). Leaf area was determined using a leaf area meter (model LI-3100C; Li-Cor, Lincoln, NE, USA) and finally aboveground fresh weight was measured. These were dried at 80 °C for four days to achieve constant mass before reweighing for aboveground dry weight.

6.4.4 ATR-FTIR spectral acquisition

Whole dried leaves were analysed using a Tensor 27 FTIR spectrometer with a Helios ATR attachment (Bruker Optics Ltd, Coventry, UK). The sampling area was 250 μm x 250 μm, as defined by the diamond crystal internal reflection element (IRE). Spectral resolution was 8 cm^{-1} with two times zero-filling, giving a data-spacing of 4 cm^{-1} over the range 4000 to 400 cm^{-1}; 32 co-additions and a mirror velocity of 2.2 kHz were used for optimum signal to noise ratio. To minimise bias, an even spread of ten spectra were taken from each surface of the leaf, twenty spectra per sample, resulting in 480 spectra in total. Each dried leaf was placed on a slide with the side to be analysed facing upwards, the slide was placed on an adjustable platform and moved upwards to ensure a good and consistent contact with the IRE.

6.4.5 Data analysis

The 'mergetool' function of an in-house developed Matlab (Mathworks, Natick, USA) toolbox called IRootLab (Martin FL et al., 2010; Trevisan et al., 2013) was used to convert all spectral information from OPUS format to suitable files (.txt). Following this, it was necessary to pre-process the acquired spectra to improve the signal-to-noise ratio. Pre-processing corrects problems associated with random or systematic artefacts introduced during spectral acquisition and is an essential step of all spectroscopic experiments. Pre-processing and computational analysis of the data were performed using a combination of IRootLab toolbox (Trevisan *et al.*, 2013) and the PLS Toolbox version 7.9.3 (Eigenvector Research, Inc., Manson, USA). The pre-processing steps applied to all spectra were firstly the selection of the spectral biochemical fingerprint region (1800-900 cm^{-1}), followed by Savitzky–Golay (SG) second differentiation (nine smoothing points) and vector normalisation. All data were mean-centred before multivariate analysis, where multiple dependant variables are observed simultaneously to determine a pattern.

Four machine learning techniques were used in this study: an unsupervised dimensionality reduction method, two supervised classification methods and one regression. The unsupervised method PCA simplifies complex multivariate datasets, allowing them to be presented intuitively and enabling pattern recognition. This process is known as dimensionality reduction. Two supervised chemometric techniques, PCA-LDA and SVM, were used for the classification of groups (Morais *et al.*, 2017b; Morais & Lima, 2018). PCA-LDA was also used for the determination of biomarkers.

To explore the relationship of controlled environmental variables with plant spectral profiles, each treatment category was scored either a zero (low) or a one (high) in each of the supplied variables of nitrogen, micronutrients, water, and R:FR, see Supplementary Table S6.3. This binary dataset was compared with ATR-FTIR

spectral absorbances using a multivariate analysis technique called PLSR. The resultant regressions gave an indication of which independent variable had the greatest impact on spectral absorbances. PLSR was performed using PLS Toolbox version 7.9.3 (Eigenvector Research, Inc., Manson, USA). Multivariate analysis techniques allow multiple variables to be compared at the same time, enabling spectral absorbance values across a range of wavelengths to be simultaneously correlated against environmental or physiological variables of numerous samples. Observing all these data at once allows patterns to be seen and enables predictions to be made.

PLSR was additionally used to predict plant physiological conditions, such as Ψ_{root}, by comparison of ATR-FTIR spectral data and measured water potential data. To form these models, an X-block of ATR-FTIR spectral absorbance data for plants was analysed by PLSR against a Y-block of physiological variables for the corresponding plants. Environments were analysed separately, allowing a model to be created for each of them. Once made, these models can be applied to new ATR-FTIR spectral data in the absence of measured physiological variables such as root water potential, to non-destructively predict these variables.

6.5 Results

6.5.1 Knotweed plants exhibit environmentally induced phenotypes

The response of Japanese knotweed to the environmental treatment conditions (LC, LD, LN, LLN, SC, SD, SN, SLN) was sufficient to induce measurable physiological differences. Leaf area was significantly lower for LD (510 cm^2) than LC (1028 cm^2), see Figure 6.1a. In general, shaded plants (SC, SD, SN, and SLN) were taller and etiolated, averaging 76.8 cm in height, compared with those of the light cabinet (LC, LD, LN and LLN), which averaged 61.3 cm (Figure 6.1b). However, plants substituted with nitrogen only (LN and SN) were different to the general trend for each light

treatment. SC and SLN plants had fewest leaves, ~15 and ~13 leaves respectively, whilst LN had the most, ~30, leaves (Figure 6.1c). Leaf thickness was higher SD and LLN than LD (Figure 6.1d). Of the plants with a high R:FR (LC, LD, LN and LLN) the aboveground dry weight was lowest under drought conditions (LD; Figure 6.1e). This differed from shaded plants (SC, SD, SN and SLN) in which SN had the lowest aboveground dry weight (Figure 6.1e). Aboveground fresh weight was lower in LD than SC, SN and SLN (Supplementary Figure S6.2a). Stomatal conductance in SN plants was three times that of LN (Figure 6.1f). Chlorophyll was highest in LD, LLN and LN (358, 343 and 325 μmol.m^{-2}), followed by SD, SC, LC and SLN (313, 305, 296, 290 μmol.m^{-2}), and the lowest levels were measured in SN (278 μmol.m^{-2}, Figure 6.1j).
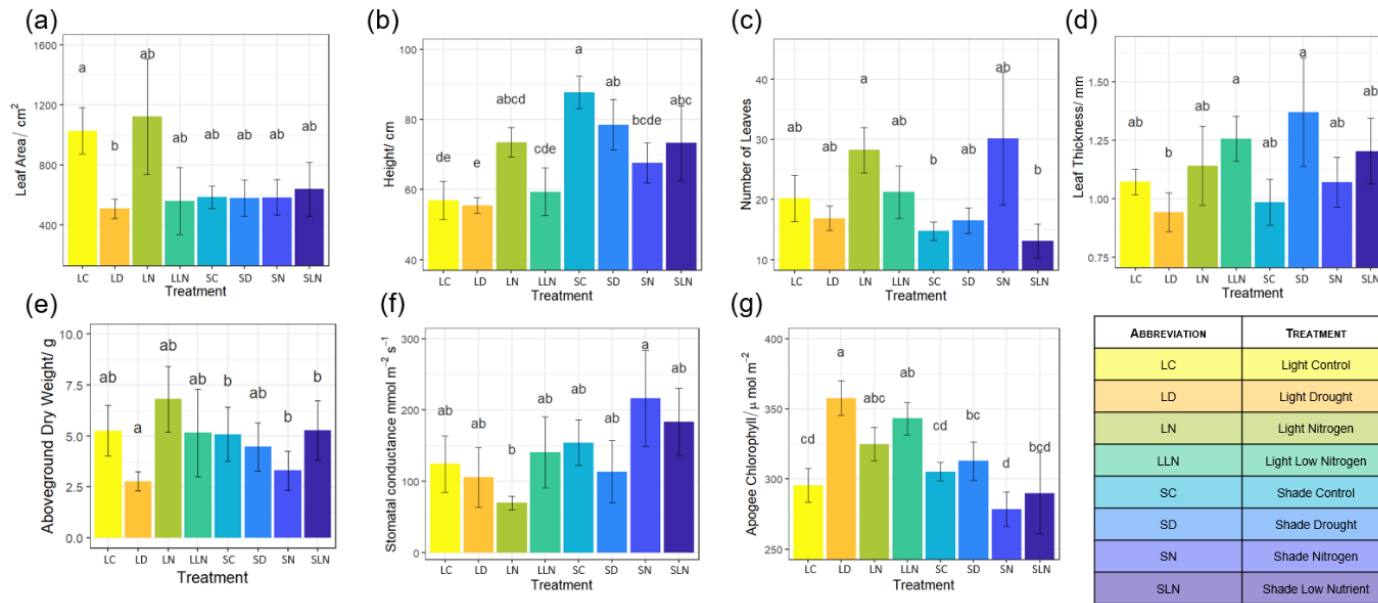
Figure 6.1: Effects of each treatment on the physiological parameters **a)** leaf area **b)** height **c)** number of leaves **d)** leaf thickness **e)** aboveground dry weight **f)** stomatal conductance **g)** Apogee chlorophyll, grouped by treatment. Statistical significance was calculated using a Kruskal-Wallis followed by a post hoc test using the criterium Fisher's least significant difference (LSD) to determine where the difference lies, signified by lowercase letters above the bars. Within each graph, all bars which share letters are not significantly different from each other. Data are mean +/- standard errors. N values for groups LC, LD, LN, LLN, SC, SD, SN, and SLN respectively: Leaf area 5, 5, 5, 4, 3, 6, 5, 6; height 10, 9, 8, 8, 8, 6, 7, 8, number of leaves 10, 9, 8, 8, 8, 6, 7, 8; leaf thickness 7, 5, 7, 7, 5, 3, 5, 6; aboveground dry weight 6, 8, 7, 5, 4, 6, 6, 7; stomatal conductance 5, 7, 4, 4, 6, 4, 4, 4; and chlorophyll 7, 7, 7, 8, 5, 6, 6, 7 (four technical replicates).

Plant water status, evidenced by Ψ_{leaf} and Ψ_{root}, was also affected markedly by treatments (Figure 6.2a-b). LD, SD, and SC plants had the most negative Ψ_{leaf}, (-6.8, -4.9 and -3.7 MPa respectively) indicating lowest hydration, followed by SN, LLN and LC (-3.6, -2.9, and -3.1 MPa respectively), whilst SLN and LN had the least negative leaf water pressures (-2.4 and -1.4 MPa respectively) indicating the highest levels of hydration (Figure 6.2a). LD, SD, and SN had the most negative root water potentials (-5.8, -3.4 and -2.5 MPa respectively) indicating lowest hydration, followed by SC (-1.8 MPa), then LN, LLN, SLN (-1.4, -1.0 and -1.2 MPa), and LC (-0.7 MPa) had the Ψ_{root} nearest to zero indicating the best levels of hydration (Figure 6.2b).
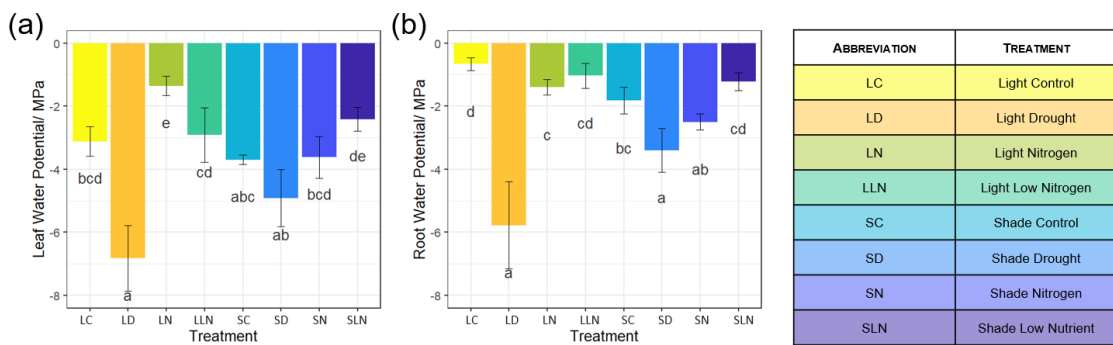
Figure 6.2: Effects of the growth environment on plant water relations: **a)** leaf and **b)** root water potential measured in MPa. Statistical significance was calculated using a Kruskal-Wallis followed by a post hoc test using the criterium Fisher's least significant difference (LSD) to determine where the difference lies, signified by lowercase letters above the bars. Within each graph, all bars that share letters are not significantly different from one another. Data are mean +/- standard errors. Leaf water potential n= 6, 8, 7, 6, 4, 6, 5, 7 and root water potential n= 6, 8, 5, 6, 4, 5, 4, 5 for groups LC, LD, LN, LLN, SC, SD, SN, and SLN respectively.

6.5.2 Physiological changes are reflected in the ATR-FTIR spectra of plants

The response of Japanese knotweed leaves to the environmental treatment conditions (LC, LD, LN, LLN, SC, SD, SN, SLN) was observed in marked differences in their spectral profiles. Figure 6.3a shows the class mean fingerprint spectra for the different environmental conditions, with the strongest absorbance peaks at approximately 1750, 1650 and 1030 cm^{-1}. Although further chemometric analysis is required to gain more information, Figure 6.3b demonstrated visual differences between pre-processed spectra taken from plants in different treatment categories. Where the spectral profiles do not overlap, such as the horizontal shifts at ~1400 and ~1550 cm^{-1}, and the vertical divergences at ~1600 and 1750 cm^{-1}, this is indicative of biomolecular differences in concentration and chemical structure respectively.
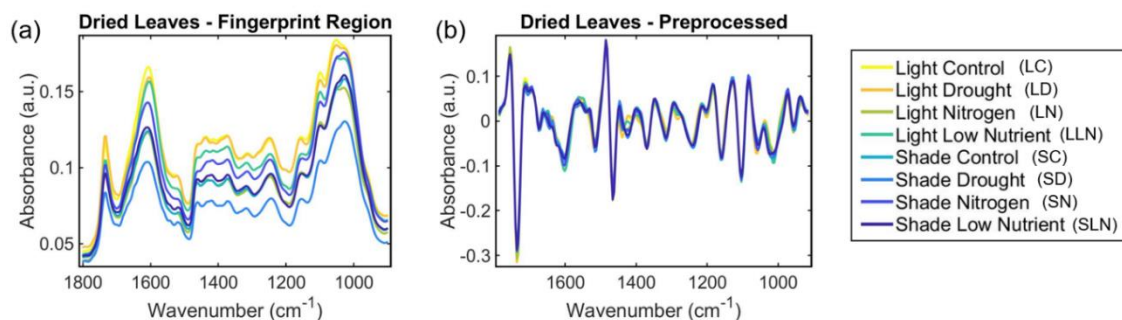
Figure 6.3: a) Class means fingerprint region (1800-900 cm^{-1}) and b) pre-processed (SG second differentiation n=9 and vector normalisation) class means spectra in the fingerprint region (1800-900 cm^{-1}) from dried leaves, grouped by treatment.

To further explore the response of Japanese knotweed to the environments (LC, LD, LN, LLN, SC, SD, SN, SLN), chemometric techniques were applied to the acquired spectra. PCA analysis alone did not achieve separation between treatment groups, with clusters overlapping (Figure 6.4a). The symmetrical pattern in Figure 6.4a was reflective of the two leaf surfaces, with separation of each in a different direction along the PC1 scores axis. This was confirmed by principal component analysis of

the spectral data from all treatment categories grouped by leaf surface; Supplementary Figure S6.2 shows lower surface spectra grouped on the left and upper surface spectra grouped on the right. Following the application of PCA-LDA to dried-leaf spectra, the different treatment types are separated into clusters along the axis LD1 (Figure 6.4b), in particular the separation between shaded and light grown samples. SVM of whole dried leaves (cost = 31.6228, γ = 3.1623, N_{SV}: 400) achieved excellent separation between categories (see Figures 6.4c-d), attaining the performance measures of 94.9% accuracy, 91.0% sensitivity, and 98.8% specificity.

DRIED LEAVES - SVM			
TREATMENT	% Accuracy	% Sensitivity	% Specificity
Light Control (LC)	94.4	90.0	98.8
Light Drought (LD)	97.4	96.3	98.5
Light Nitrogen (LN)	97.4	95.7	99.0
Light Low Nitrogen (LLN)	93.7	88.2	99.1
Shade Control (SC)	93.2	87.5	98.9
Shade Drought (SD)	96.3	95.0	97.6
Shade Nitrogen (SN)	91.9	84.0	99.8
Shade Low Nitrogen (SLN)	95.1	91.4	98.8
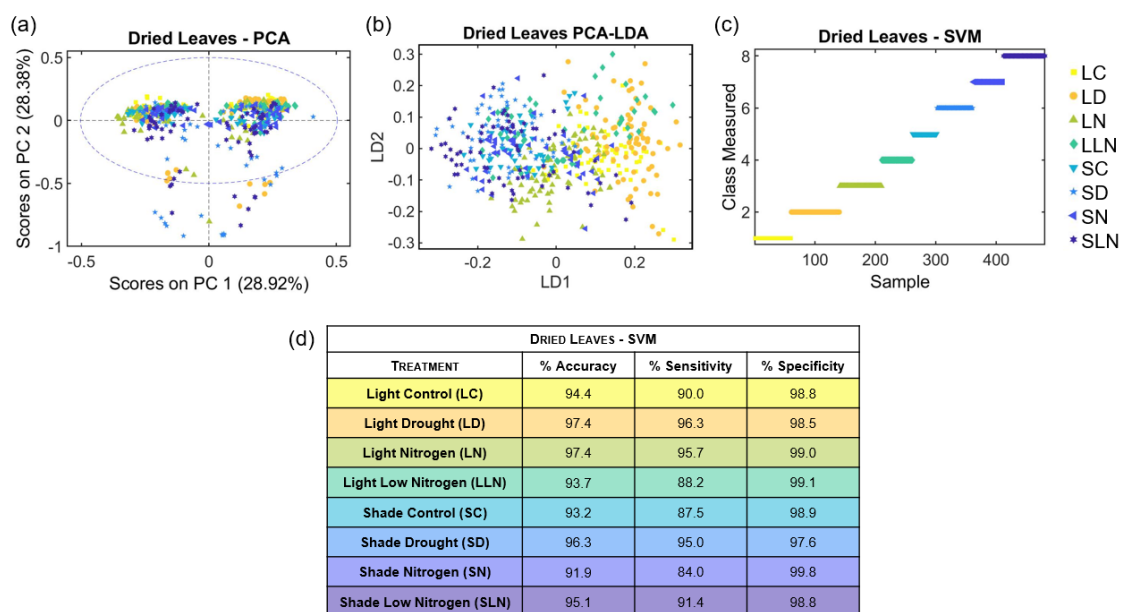
Figure 6.4: a) PCA scores plot, **b)** PCA-LDA scatter plot, **c)** SVM sample/measured plot and **d)** SVM results for spectra taken of dried leaves grouped by treatment. Note that the PC1 axis in a) separates the spectra taken from the lower leaf surface to the left of zero and those from the upper surface are on the right, see Supplementary Figure S6.2.

6.5.3 Plant responses to environment are associated with specific environmental response biomarkers

Molecular differences between plants of the eight treatments were examined using PCA-LDA analysis (Figure 6.4b). The top eight PCA loadings (Supplementary Figures S6.4-S6.15), which represent the main qualitative wavenumbers discriminating each treatment comparison, were subsequently selected for the identification of biomarkers to qualitatively characterise the biochemical compounds showing the greatest changes, see Table 6.1. Many of the assigned biomarkers are cell wall carbohydrates and the proteins which regulate them, as well as molecules which other studies have identified as indicators of abiotic stress.

Table 6.1: Loadings and biomarkers for discrimination between spectral profiles of plants from different treatment categories.

	Molecular Assignment	Wavenumbers cm⁻¹	Reference	Comparison											
				LC vs LD	LC vs LN	LC vs LLN	LC vs SC	LD vs SD	LN vs LLN	LN vs SN	LLN vs SLN	SC vs SD	SC vs SN	SC vs SLN	SN vs SLN
Carbohydrates	C=O stretching vibration of alkyl ester in pectin	1748, 1612	(Nogales-Bueno et al., 2017; Sharma & Uttam, 2017)			✓							✓	✓	
	Ring vibration of pectin	1319	(Sharma & Uttam, 2017)					✓	✓						
	OCH₃ (polysaccharides, pectin)	972	(Talari et al., 2017)							✓					
	Starch	1042, 1030, 1026, 1022	(Lei et al., 2018)	✓	✓		✓			✓		✓			
	β-glucans	1427	(Liu X et al., 2021)		✓		✓								
	Stretching vibrations [ν(CO)] of the COC	1130	(Liu X et al., 2021)			✓							✓		✓

	Assignment	Wavenumber	Reference												
	glycosidic linkage of polysaccharides														
	Bending vibration δ(CH) polysaccharides	1443	(Morais *et al.*, 2017a)								✓				
	C-O & C-C stretching; (C-O-H) bending of cellulose (abaxial surface)	1049	(Moskal *et al.*, 2019)											✓	
	Hemicellulose	1732	(Ord *et al.*, 2016)				✓	✓				✓			
	v(C–O) and *v*(C–C), disaccharides, sucrose	1126	(Talari *et al.*, 2017)		✓										
	C=O stretching of carbonyl group, typical saccharide absorption	1636	(Talari *et al.*, 2017)					✓							
	C-O-C bending vibration (fructo-oligosaccharides)	1134	(Krähmer *et al.*, 2021)					✓							
	C–C stretching ring (galactan and rhamnogalacturonan)	1072	(Liu X *et al.*, 2021)					✓							

240

Group	Assignment	Wavenumber	Reference												
	v(CC) skeletal cis conformation, v(CH$_2$OH); Galactan v(CO) stretching coupled with C–O bending;	1038	(Ord *et al.*, 2016)								✓				✓
Bases	C=O thymine	1713	(Talari *et al.*, 2017)			✓									
	C=N adenine	1609, 1574	(Talari *et al.*, 2017)	✓											✓
	Modified guanine	1531	(Talari *et al.*, 2017)	✓	✓										
	RNA stretching, ring bending of uracil	988	(Talari *et al.*, 2017)								✓				
	C–O stretching of the phosphodiester and the ribose	991, 1065	(Talari *et al.*, 2017)			✓								✓	
	C–N–C stretch of nucleic acids	968	(Talari *et al.*, 2017)						✓						
	nucleic acids due to the base carbonyl	1690	(Talari *et al.*, 2017)					✓							

Category	Feature	Wavenumber	Reference												
	(C=O) stretching and ring breathing mode														
	Phosphate I (stretching PO^{2-} symmetric vibration) in B-form DNA	1088	(Talari *et al.*, 2017)						✓		✓				
	Ring base	1566	(Talari *et al.*, 2017)												✓
Proteins	Protein	1709	(Butler *et al.*, 2015b)	✓	✓										
	Protein phosphorylation	953	(Ord *et al.*, 2016)	✓							✓		✓	✓	
	Amide I	1670, 1647, 1651, 1639, 1628, 1585	(Strong *et al.*, 2017; Talari *et al.*, 2017)	✓	✓	✓	✓		✓	✓	✓	✓	✓		✓
	Amide II: (protein N–H bending, C–N stretching), α-helical structure	1570, 1547	(Talari *et al.*, 2017)			✓	✓				✓	✓			
	Amide III	1323	(Morais *et al.*, 2017a)	✓						✓					

	Assignment	Wavenumber	Reference													
	Symmetric CH$_3$ bending of the methyl groups of proteins	1396	(Talari *et al.*, 2017)										✓			
	Tyrosine	1512	(Talari *et al.*, 2017)									✓				
Lipids	CH$_3$ bending of proteins and lipids	1393	(Strong *et al.*, 2017)		✓	✓	✓									
	Lipid; fatty acid esters	1732	(Ord *et al.*, 2016)				✓	✓					✓			
	CH$_2$ bending in lipids	1466, 1470	(Butler *et al.*, 2015b) (Talari *et al.*, 2017)				✓	✓					✓			✓	✓
	Ester C=O stretch: triglycerides, cholesterol esters	1744	(Talari *et al.*, 2017)							✓						
	Ester C=O stretching vibration (phospholipids)	1740	(Talari *et al.*, 2017)													✓	
Other	Pentanone	1385, 1157, 957	(Gandolfo *et al.*, 2016)	✓	✓		✓									
	Ring C–C stretch of phenyl	1589, 1582	(Talari *et al.*, 2017)			✓						✓				

244

Assignment	Wavenumber	Reference													
In-plane C–O stretching vibration combined with the ring stretch of phenyl	1339	(Talari et al., 2017)							✓					
Ring stretching (C–C)–d(C–H)	999	(Morais et al., 2017a)				✓								✓
C=O band	1728	(Talari et al., 2017)					✓							
Asymmetric phosphate; ν(C–O–C) of geranyl acetate (acyclic monoterpene)	1227	(Ord et al., 2016)							✓					
PO₂⁻ asymmetric stretching (Phosphate I)	1265, 1242, 1211, 1207	(Talari et al., 2017)					✓	✓		✓	✓			✓
CHα' rocking	1265	(Talari et al., 2017)								✓				
Symmetric phosphate stretching	1092	(Strong et al., 2017)										✓	✓	
Stretching vibrations of hydrogen-bonding C–OH groups	1161	(Talari et al., 2017)											✓	

6.5.4 PLS regression allows predictions of physiological responses

Plant environmental responses to different treatment groups were reflected in the differences in both physiological measurements (Figures 6.1 and 6.2) and in the acquired ATR-FTIR spectral profiles (Figure 6.3). Two such datasets can be combined to correlate spectral profiles with phenotypic responses, as shown previously (Chapter 5). Similarly, the phenotypic changes measured here can form training and test data for models which demonstrate the potential of mid-infrared (MIR) spectroscopy for the rapid, non-destructive prediction of plant physiology. Figure 6.2 shows that $\Psi_{leaf\ and}$ Ψ_{root} were affected markedly by treatments (Figure 6.2a-b). The largest Ψ_{root} was found in LD (-5.8 MPa) compared with LC (-0.7 MPa) which had the smallest. Therefore, these two categories were chosen to develop a model for the prediction of plant water status. Figure 6.5 is an exemplar PLS regression graph for the prediction of physiological parameters using ATR-FTIR spectral data. This shows the PLS regression of Ψ_{root}, as measured with a pressure chamber (Model 3000F01 Plant Water Status Console, Soil Moisture Equipment Corp., Santa Barbara, CA, USA), against predicted levels using ATR-FTIR spectral absorbances of intact dried leaves grown under LC and LD conditions as training data, showing an example of the predictive models generated using this approach. The green line in Figure 6.5 shows the ideal prediction gradient of one, which would be 100% accurate, and the red line shows the achieved success of the model, which was a gradient of $R^2 = 0.8$ indicating a successful prediction. The key wavenumbers for this PLSR prediction of Ψ_{root} are shown in a loadings graph (Supplementary Figure S6.16), which provide useful spectral indicators of drought. This further highlights the power of such models when investigating the responses of plants to environmental variables and the impacts these have on growth and development, particularly in species which are recalcitrant to the destructive and time-consuming physical techniques routinely used for such measurement.
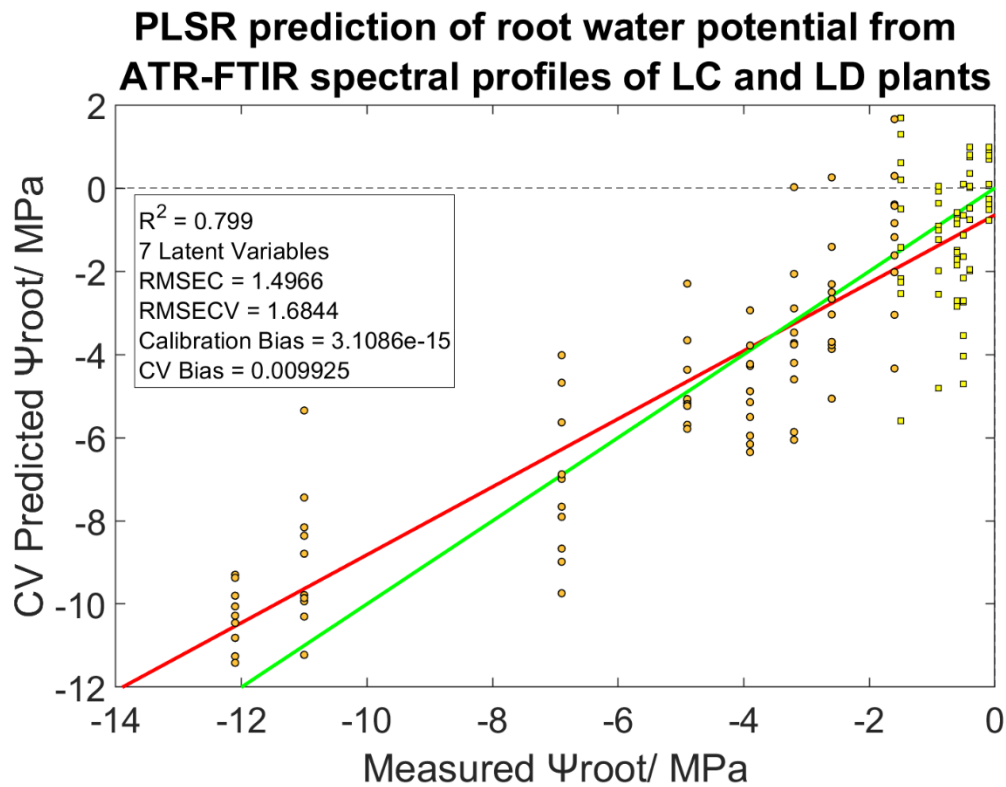
Figure 6.5: Prediction of root water potential (Ψ_{root}) from ATR-FTIR spectral data in the fingerprint region (1800-900 cm^{-1}) of plants from categories LC and LD using PLS regression. Model parameters are as follows: RMSEC: 1.4966, RMSECV: 1.6844, Calibration Bias: 3.1086e-15, CV Bias: -0.009925, R^2 Cal: 0.841332, R^2 CV: 0.799426. Yellow squares indicate data points from LC plants and orange circles represent data from LD plants.

6.5.5 Light environment R:FR ratio has the greatest impact on the plant spectral profiles

Insights into which independent variable had the greatest impact on spectral absorbances were gained through PLS regression of the binary dataset in Supplementary Table S6.3 against spectral absorbances. Table 6.2 shows the results

of this PLS regression, with higher R^2 values indicating a greater contribution from the corresponding abiotic variable. The red: far-red ratio had the greatest influence over the spectral profile of intact dried leaves, reaching the highest R^2 of 0.3, whilst water, micronutrients, and nitrogen all contributed R^2 values of 0.1. The RMSECV, representing the error in the predicted environmental parameter, was higher for prediction of micronutrient conditions at 0.5, than the other parameters at 0.4.

Table 6.2: Model parameters of PLSR correlation of intact leaf spectral profiles with controlled environments assigned binary values.

Model parameters	Nitrogen	Micronutrients	Red: far-red ratio	Water
R^2	0.1	0.1	0.3	0.1
RMSECV	0.4	0.5	0.4	0.4

6.6 Discussion

6.6.1 Environmental response can be detected through leaf spectral profiles

The measured physiological parameter in Figures 6.1 and 6.2 show clearly that the genetically identical plants responded to the applied treatment conditions with environmentally induced phenotypes, demonstrating the plasticity expected of this species (Parker *et al.*, 2003; Bossdorf *et al.*, 2005; Richards *et al.*, 2006, 2008; Van Kleunen *et al.*, 2010; Pichancourt & van Klinken, 2012; Hagenblad *et al.*, 2015; Geng *et al.*, 2016). The acquired ATR-FTIR spectral absorbances in Figure 6.3 demonstrate that the environmental conditions under which plants are grown have a marked impact on their spectral profiles suggesting subtle changes in plant composition. Although the overall absorbance pattern of the fingerprint region in Figure 6.3a is visually similar that of any other green vegetative plant tissue (Cao *et al.*, 2017;

Holden *et al.*, 2021; Macchioni *et al.*, 2022), chemometrics has the power to extract more information, including differences between treatment groups. The unique patterns produced by the ATR-FTIR spectral absorbance profiles of plants grown in different environments could be successfully differentiated through application of the discriminant algorithm, SVM (Figure 6.4c-d). This combined method achieved high accuracy, sensitivity, and specificity. This is consistent with observations from previous studies which have identified spectra from plants of different growing environments (Ruoff *et al.*, 2006; Bağcıoğlu *et al.*, 2017; Traoré *et al.*, 2018; Gordon *et al.*, 2019; Zeghoud *et al.*, 2021; Holden *et al.*, 2021) suggesting that the high degree of plasticity exhibited by Japanese knotweed to environmental factors is reflected in key biomolecular changes that may contribute to its success as an invasive species.

6.6.2 Detected biomarkers are associated with plant stress responses

A variety of biological compounds, such as proteins, ketones, terpenes, and carbohydrates, which differ under each condition were identified through linking molecular biomarkers with key wavenumbers affected by environment. The peak at 1709 cm^{-1} (Figures S6.4 and S6.7), which was assigned to protein absorbance (see Table 6.1), is a common biomarker across other plant species, and has been associated with plant development in tomato plants (Butler *et al.*, 2015b). This band indicates the relative concentrations of the significant proteins are highest in LD, followed by LC, and lowest in LN. The upregulation of these proteins under conditions of drought stress and increased micronutrient availability suggests that the peak at 1709 cm^{-1} may correspond to enzymes that break down polysaccharides in plant cell walls, such as pectin-methylesterase and β-glucosidase. These enzymes allow for leaf expansion during development and stress-induced alterations of cell-wall polymers (Wu *et al.*, 2018). Another response of Japanese knotweed plants to water deficit is highlighted by biomarker 1385 cm^{-1} which was assigned to the bioactive plant volatile, pentanone. Spectral comparison suggests that pentanone was higher in control plants (LC) than droughted plants (LD). This is consistent with

a similar spectroscopic study on citrus plants in which pentanone was retained in healthy leaves but released under biotic stress (Gandolfo *et al.*, 2016). It seems possible that this finding could be related to the hormone jasmonic acid, which contains a pentanone ring. This hypothesis is supported by the hormonal data in Chapter 5 Supplementary Figure S5.6f in which jasmonic acid levels were higher in LC than in plants grown under less ideal conditions. The jasmonic acid signalling pathway is a core component in plant response to both biotic and abiotic stress (Yang J *et al.*, 2019). The peak at 953 cm^{-1}, caused by protein phosphorylation, was a key difference in several comparisons; LC vs LD, LLN vs SLN, SC vs SN, and SC vs SLN. This indicates that wavenumber-953 is sensitive to changes induced by drought, and the combination of stresses from a low R:FR and deficiencies in nitrogen and micronutrients. Similarly, the peak at 1732 cm^{-1}, associated with hemicellulose (Ord *et al.*, 2016), was a key difference between LC vs SC, LD vs SD, and SC vs SN, whilst 1038 cm^{-1}, associated with the polysaccharide galactan (Ord *et al.*, 2016), allowed discrimination between LLN vs SLN and SN vs SLN. The association of these peaks with abiotic stress is consistent with other spectroscopic studies which have associated them with vehicular pollution in sycamore trees (Ord *et al.*, 2016). Wavenumber 1227 cm^{-1}, assigned to geranyl acetate, an acyclic monoterpene, was higher in LLN than SLN. This biomarker has previously been associated with response to ozone exposure in sycamore tree leaf tissue (Ord *et al.*, 2016). Amide I peaks at 1628 and 1585 cm^{-1}, have previously been associated with fungal infection in other studies (Ord *et al.*, 2016). In the present study, these two Amide I peaks were key differences between plants not provided with any nutrients and those provided with nitrogen only. It is common for plant responses to biotic and abiotic stresses to overlap because stress signalling pathways are known to share intermediates such as reactive oxygen species and calcium which allow for crosstalk (Heap *et al.*, 2020). Taken together these results suggest that wavenumbers 953, 1038, 1227, 1709, and 1732 cm^{-1} are key indicators of plant-

environment interactions, including abiotic stress responses, that are conserved between species.

Several cell wall carbohydrates, both structural and storage, were highlighted as significant changes between environments, detectable via the ATR-FTIR spectral profile. The storage molecule starch, reserves of which are known to be mobilised under conditions of abiotic stress (Thalmann & Santelia, 2017), was identified as a biomarker for five of the comparisons (Lei *et al.*, 2018). As outlined above, the peak for hemicellulose (1732 cm^{-1}) was a key difference between shaded and unshaded plants both under, and in the absence of, drought stress (Ord *et al.*, 2016). Shading stress is known to have a greater inhibitory effect on the biosynthesis of non-structural than that of structural carbohydrates (Hussain *et al.*, 2019). Some structural carbohydrates, such as xylose and mannose, decrease under shading stress (Hussain *et al.*, 2019). Beta- glucans were highlighted as a key differentiator between plants grown with differing red: far-red ratios, LC and SC again showing hemicellulose to be affected by shading. Also differentiated by this beta- glucan peak were LC and LN, indicating a consequence of environmental micronutrient levels. Wavenumber 1126 cm^{-1} identified sucrose, the major transport form of photo-assimilated carbon, as a key discriminator between LC and LN, but not their shaded equivalents SC and SN (Lemoine, 2000; Talari *et al.*, 2017). The peak at 1049 cm^{-1} for cellulose (Moskal *et al.*, 2019) was a key discriminator between SC and SLN, showing the impact of nutrient deficiency on plants also experiencing shading. Low levels of soluble sugar, sucrose, lignin and cellulose content can result in weak stem strength (Hussain *et al.*, 2019). However, this was not reflected in the measured stem diameter of plants, variations in which were insignificant across all eight treatments (see Figure 6.1k). Etiolation was observed in shaded plants from SC, SD and SLN which were significantly taller than those of the non-shaded groups LC, LD and LLN (see Figure 6.1b). One molecule which can alter elongation capacity of the cell wall is galactan (see peaks 1072 and 1038 cm^{-1}). It achieves this by controlling porosity and viscoelastic properties of the cell wall (Izquierdo *et al.*, 2018) and levels

increase during the cell expansion phase (Waldron & Faulds, 2007). Galactans also play a role as cell wall storage polysaccharide (Waldron & Faulds, 2007). The amplitude of peaks 1072 and 1038 in the rubber band normalised fingerprint spectra indicate that galactan was present in higher concentrations in LN than LLN, SN than SLN, and in LLN than SLN, but this peak was not identified as a key difference for other comparisons. These results suggest that Japanese knotweed plants which are plants experiencing shading and deficiency of both micronutrients and nitrogen have lower levels of galactan, because they have low requirements for carbohydrate storage and lack the excess resources for expansion and growth. 'Tarping', an herbicide-free control strategy where soil covered with a plastic tarp is heated by solar radiation and thought to reach a lethal temperature for knotweed growth (Dusz *et al.*, 2021), is more effective when black, light-blocking tarp is used (Miller *et al.*, 2013). The combined stresses of shade and low nutrients observed in this study suggests that the additional shading effect of tarping would be most effective in areas of poor soil quality or those prone to leaching. Artificial shading may also increase the efficacy of herbicides that function through interference with nutrient absorption and metabolism.

6.6.3 The red: far-red ratio had the greatest effect on leaf spectral profiles

Differences in spectral profiles indicate key biomolecular alterations occurring within the leaves under different growth environments, reflective of the high degree of environmental plasticity exhibited by Japanese knotweed that may contribute to its success as an invasive species (Liao *et al.*, 2016; Mounger *et al.*, 2021). Although IAS generally display greater plasticity, this is not always correlated with a fitness benefit (Davidson *et al.*, 2011). Physiological variation of Japanese knotweed grown in different habitats has been recorded in previous studies, including differences in height, number of leaves, leaf surface area and biomass allocation (Walls, 2010; Yuan *et al.*, 2022). All the rhizomes used for these controlled growth experiments

were extracted from the same source, but recent research indicates that the environmental adaptations observed here may be influenced by the original source of the rhizomes and could have differed if these were collected from another habitat type (Yuan *et al.*, 2022). The growth environment plays a significant role in phenotypic presentation, and this study has explored the influence of specific environmental variables.

Chlorophyll content under drought stress can both increase (da Costa & Tom Cothren, 2011) and decrease (Mafakheri *et al.*, 2010). Here, LD had significantly higher chlorophyll levels compared with LC (Figure 6.1g). This may be because LD had fewer (Figure 6.1c), smaller (Figure 6.1a), leaves compared with LC, leading to a necessity for increased chlorophyll levels per unit leaf area. Figure 6.1 shows that plants of the nitrogen supplemented category, SN, had a significantly greater number of leaves than SC or SLN, but this was not statistically different to SD. Although drought is usually associated with a reduced leaf number (Álvarez *et al.*, 2013; Kebbas *et al.*, 2015), this observation is consistent with water having only been withheld completely for seven days out of a total growth period of fifty days in the present study, representing a short-term drought rather than a long term water deficit. However, nitrogen availability, sugar demand, R:FR and auxin concentrations have all been linked to the control of apical branching (Mason *et al.*, 2014; Hou M *et al.*, 2021; Holalu *et al.*, 2021). These pathways interlink via common intermediates, for example a low R:FR promotes auxin signalling (Holalu *et al.*, 2021) and nitrogen fluctuations have a significant impact on auxin distribution (Hou M *et al.*, 2021). These complex interacting signals could explain the lack of significant differences in the number of leaves between categories. Growth was markedly affected in shaded plants (LC, LD, LN and LLN) which were generally taller in height (Figure 6.1). This is consistent with ethylene-mediated stem elongation which is a stress response known to be induced by low R:FR (Pierik *et al.*, 2004). Additionally, transpiration rate is elevated under high R:FR (Hoad & Leakey, 1994) thereby increasing the likelihood of plants experiencing drought under conditions of

reduced water availability, possibly leading to the larger root water potentials measured in category LD compared to category SD. This increased transpiration rate may also have altered the effect of drought on leaf quantity under different lightings; under a high R:FR LC plants had more leaves than LD, however under a low R:FR then SD had more leaves than SC, see Figure 6.1.

Of the altered environmental parameters, the R:FR of the growth environment had the greatest effect on the spectral profiles of intact dried leaf material. This was indicated by the highest PLS regression R^2 in Table 6.2. A possible explanation for this is that these spectra were taken from leaf vegetative tissue, which may be more prone to changes in light as leaves are photosynthetic organs. The primary function of leaves is to absorb sunlight for photosynthesis, and spectra of other plant organs may be affected differently by different environmental factors. The clustering pattern of the 2D PCA-LDA scatter graph (Figure 6.4b) provides an indication of why the R:FR has the greatest impact on the spectra. This displays a general separation along the axis LD1 of shaded samples on the left and non-shaded samples on the right. This LD1 axis is significant for identifying the wavenumbers most affected by the different environments, used for the classification of groups. Molecular biomarkers were found from the LD1 loadings and indicate which molecules differ most between leaves of plants grown in different environments: lipid (1732 cm^{-1}), CH$_2$ bending of the methylene chains in lipids (1470 cm^{-1}), ring breathing (995 cm^{-1}), C-H and O-H bending in hemicellulose (1423 cm^{-1}), C-O vibration in sucrose (1126 cm^{-1}), overlapping of the protein amide III and the nucleic acid asymmetric phosphate vibration (1231 cm^{-1}), cellulose (1319 cm^{-1}), peak of nucleic acids due to the base carbonyl stretching and ring breathing mode (1620 cm^{-1}), starch (1030 cm^{-1}), and CH$_3$ rocking (957 cm^{-1}) (Zhou et al., 2015; Talari et al., 2017; Rana et al., 2018; Jin et al., 2018; Lei et al., 2018; Kharrat et al., 2020).

The sensitivity of ATR-FTIR spectral profiles to the R:FR ratio of the growth environment could account for some of the spectral differences in plants grown in

different regions of the UK, observed previously by Holden *et al.* (2021). At higher latitudes, plants experience longer durations of sunlight from low solar angles (Chiang *et al.*, 2019). The R:FR ratio at low solar angles is lower (Kotilainen *et al.*, 2020) and more variable (Chiang *et al.*, 2019). Water vapour increases the R:FR photon ratio by preferentially absorbing the FR light, due to the water absorption band at 728 nm, meaning that plants in growing regions with overcast skies tend to experience a lower R:FR ratio than those under clear skies (Durand *et al.*, 2021). Additionally, modelled climate scenarios predict that increasing global temperatures will result in increased atmospheric water vapour, which will reduce the proportion of far-red photons in sunlight (Kotilainen *et al.*, 2020). Plants sense the R:FR ratio with phytochromes which allows them to trigger their shade-avoidance response and detect above-ground neighbours (Ballaré & Pierik, 2017). As a pioneer species, one could predict Japanese knotweed to be a competitive shade-avoider, likely to have a strong avoidance response compared with a shade-tolerant woodland floor species, however, leaves within the dense knotweed canopy are known to experience reduced light-penetration. Martin FM *et al.*, (2019) noted the lack of information on the significance of shading for Japanese knotweed, particularly in interaction with mechanical control, whilst observing its importance for ramet density (the space between independent members of a clone). Plants grown in shaded conditions (SC, SD, SN, SLN) tend to have lower aboveground fresh and dry weights compared with light (LC, LD, LN, LLN), see Figure 6.1 and Supplementary Figure S6.2, supporting the importance of light quality for Japanese knotweed. These results have been echoed in field studies which found that a reduction in soil fertility had no significant effect on knotweed biomass production, and concluded that light quality was the most important of the tested parameters (Dommanget *et al.*, 2013).

6.6.4 ATR–FTIR spectroscopy provides a novel tool for predicting physiological responses

The model in Figure 6.5 was created using training data from plants of LC and LD, which differed only in the amount of water supplied to them with the other controlled environmental variables remaining the same. The use of larger training sets would allow the generation of more robust models which take account of the breadth of the variables in the growth environment enabling this approach to be widely applied. Chapter 5 shows the power of ATR-FTIR spectroscopy for predicting plant physiological responses. Applying this approach to the analysis of plant water status in Figure 6.5 further highlights the importance of such predictive models for non-destructively studying the responses of plants to their environment in situ. Near-infrared (NIR) spectroscopy (Vohland *et al.*, 2022) using portable handheld NIR spectrometers, whilst less rich in the spectral information provided compared with ATR-FTIR spectroscopy, has been used for monitoring plant water (Diago *et al.*, 2017, 2018; Fernández-Novales *et al.*, 2018) and nutrient (Pandey *et al.*, 2017) status. Advances in technology mean that portable MIR spectrometers are now available (Dhawale *et al.*, 2015; Ji *et al.*, 2016; Soriano-Disla *et al.*, 2018; Bureau *et al.*, 2019; Hutengs *et al.*, 2019) highlighting the potential of this method for future applications of MIR spectroscopy to the prediction of physiological responses in the field, providing a more sensitive alternative to NIR spectroscopy. The use of MIR as an indicative tool to determine the efficacy of treatment approaches for invasive knotweeds could accelerate studies which normally span several years (Jones D *et al.*, 2018). ATR-FTIR spectral ground-readings could complement spatial dynamic data collected by remote sensing (Martin FM *et al.*, 2018), to create detailed predictive maps that enhance our ability to monitor invasive alien species, providing further information on the 'what' in addition to the 'where'.

6.7 Conclusions

This study indicated that the R:FR ratio of the light environment, plays a key role in shaping the spectral profile of Japanese knotweed. Spectral differences between plants of different treatment groups allowed the identification of several biomarkers for environmental effects, highlighting changes in compounds such as cell wall carbohydrates. Conservation of specific wavenumbers across plant species raises the potential for their use as indicators of plant health under abiotic stress.

6.8 Authors' Contributions

CAH conceived, planned, and carried out the experiments and data analysis. The manuscript was written by CAH and MM with contribution from all the authors. FM provided equipment and expertise in the field of FTIR spectroscopy and chemometrics. PB provided funding for CAH's studentship and expertise in Japanese Knotweed. MM and JET supervised the project.

6.9 Acknowledgements

CAH is a member of the Centre for Global Eco-Innovation that is funded by the European Union Regional Development Fund and mediates the collaboration between Lancaster University and Phlorum Ltd. The authors declare that there is no conflict of interest.

6.10 Conflict of Interest

The authors declare that there is no conflict of interest.

6.11 Data Availability Statement

The datasets generated and analysed during the current study are available in a supplementary folder.

6.12 References

Álvarez S, Bañón S, Sánchez-Blanco MJ. **2013**. Regulated deficit irrigation in different phenological stages of potted geranium plants: Water consumption, water relations and ornamental quality. *Acta Physiologiae Plantarum* **35**: 1257–1267.

Bağcıoğlu M, Kohler A, Seifert S, Kneipp J, Zimmermann B. **2017**. Monitoring of plant–environment interactions by high-throughput FTIR spectroscopy of pollen. *Methods in Ecology and Evolution* **8**: 870–880.

Bailey J. **2013**. The Japanese knotweed invasion viewed as a vast unintentional hybridisation experiment. *Heredity* **110(2)**: 105-110.

Bailey JP, Conolly AP. **2000**. Prize-winners to pariahs -A history of Japanese Knotweed s.l. (Polygonaceae) in the British Isles. *Watsonia* **23**: 93–110.

Ballaré CL, Pierik R. **2017**. The shade-avoidance syndrome: multiple signals and ecological consequences. *Plant, Cell & Environment* **40**: 2530–2543.

Bellard C, Jeschke JM, Leroy B, Mace GM. **2018**. Insights from modeling studies on how climate change affects invasive alien species geography. *Ecology and Evolution* **8**: 5688–5700.

Bímová K, Mandák B, Pyšek P. **2003**. Experimental study of vegetative regeneration in four invasive Reynoutria taxa (Polygonaceae). *Plant Ecology* **166**: 1–11.

Bossdorf O, Auge H, Lafuma L, Rogers WE, Siemann E, Prati D. **2005**. Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia* **144**: 1–11.

Bradshaw CJA, Leroy B, Bellard C, Roiz D, Albert C, Fournier A, Barbet-Massin M, Salles JM, Simard F, Courchamp F. **2016**. Massive yet grossly underestimated global costs of invasive insects. *Nature Communications* **7**: 1–8.

Broadbent A, Stevens CJ, Peltzer DA, Ostle NJ, Orwin KH. **2018**. Belowground competition drives invasive plant impact on native species regardless of nitrogen availability. *Oecologia* **186**: 577–587.

Buitrago MF, Groen TA, Hecker CA, Skidmore AK. **2016**. Changes in thermal infrared spectra of plants caused by temperature and water stress. *ISPRS Journal of Photogrammetry and Remote Sensing* **111**: 22–31.

Bureau S, Cozzolino D, Clark CJ. **2019**. Contributions of Fourier-transform mid infrared (FT-MIR) spectroscopy to the study of fruit and vegetables: A review. *Postharvest Biology and Technology* **148**: 1–14.

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

Butler HJ, Martin FL, Roberts MR, Adams S, McAinsh MR. **2020**. Observation of nutrient uptake at the adaxial surface of leaves of tomato (*Solanum lycopersicum*) using Raman spectroscopy. *Analytical Letters* **53**: 536–562.

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

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

Cao Z, Wang Z, Shang Z, Zhao J. **2017**. Classification and identification of Rhodobryum roseum limpr. And its adulterants based on fourier-transform infrared spectroscopy (FTIR) and chemometrics. *PLoS ONE* **12(2):** e0172359.

Chiang C, Olsen JE, Basler D, Bånkestad D, Hoch G. **2019**. Latitude and weather influences on sun light quality and the relationship to tree growth. *Forests* **10(8):** 610.

Chinchio E, Crotta M, Romeo C, Drewe JA, Guitian J, Ferrari N. **2020**. Invasive alien species and disease risk: An open challenge in public and animal health (LJ Knoll, Ed.). *PLOS Pathogens* **16**: e1008922.

da Costa VA, Tom Cothren J. **2011**. Drought Effects on Gas Exchange, Chlorophyll, and Plant Growth of 1-Methylcyclopropene Treated Cotton. *Agronomy Journal* **103**: 1230–1241.

Davidson AM, Jennions M, Nicotra AB. **2011**. Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. *Ecology Letters* **14**: 419–431.

Davis MA, Grime JP, Thompson K. **2000**. Fluctuating resources in plant communities: A general theory of invasibility. *Journal of Ecology* **88**: 528–534.

Dhawale NM, Adamchuk VI, Prasher SO, Viscarra Rossel RA, Ismail AA, Kaur J.
2015. Proximal soil sensing of soil texture and organic matter with a prototype
portable mid-infrared spectrometer. *European Journal of Soil Science* **66**: 661–669.

Diago MP, Bellincontro A, Scheidweiler M, Tardaguila J, Tittmann S, Stoll M.
2017. Future opportunities of proximal near infrared spectroscopy approaches to
determine the variability of vineyard water status. *Australian Journal of Grape and
Wine Research* **23**: 409–414.

Diago MP, Fernández-Novales J, Gutiérrez S, Marañón M, Tardaguila J. 2018.
Development and validation of a new methodology to assess the vineyard water
status by on-the-go near infrared spectroscopy. *Frontiers in Plant Science* **9**: 59.

**Diez JM, D'Antonio CM, Dukes JS, Grosholz ED, Olden JD, Sorte CJB, Blumenthal
DM, Bradley BA, Early R, Ibáñez I, *et al.* 2012**. Will extreme climatic events
facilitate biological invasions? *Frontiers in Ecology and the Environment* **10**: 249–
257.

Dommanget F, Spiegelberger T, Cavaillé P, Evette A. **2013**. Light availability
prevails over soil fertility and structure in the performance of Asian knotweeds on
riverbanks: New management perspectives. *Environmental Management* **52**: 1453–
1462.

Durand M, Murchie EH, Lindfors A V., Urban O, Aphalo PJ, Robson TM. 2021.
Diffuse solar radiation and canopy photosynthesis in a changing environment.
Agricultural and Forest Meteorology **311**: 108684.

**Dusz MA, Martin FM, Dommanget F, Petit A, Dechaume-Moncharmont C, Evette
A**. **2021**. Review of Existing Knowledge and Practices of Tarping for the Control of
Invasive Knotweeds. *Plants 2021, Vol. 10, Page 2152* **10**: 2152.

Euring D, Löfke C, Teichmann T, Polle A. **2012**. Nitrogen fertilization has
differential effects on N allocation and lignin in two Populus species with contrasting
ecology. *Trees - Structure and Function* **26**: 1933–1942.

Fernández-Novales J, Tardaguila J, Gutiérrez S, Marañón M, Diago MP. **2018**. In
field quantification and discrimination of different vineyard water regimes by on-
the-go NIR spectroscopy. *Biosystems Engineering* **165**: 47–58.

**Fowler D, Coyle M, Skiba U, Sutton MA, Cape JN, Reis S, Sheppard LJ, Jenkins A,
Grizzetti B, Galloway JN, *et al.* 2013**. The global nitrogen cycle in the twenty-first
century. *Philosophical Transactions of the Royal Society B: Biological Sciences*
368(1621): 20130164.

Gandolfo DS, Mortimer H, Woodhall JW, Boonham N. **2016**. Fourier transform infra-red spectroscopy using an attenuated total reflection probe to distinguish between Japanese larch, pine and citrus plants in healthy and diseased states. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **163**: 181–188.

Geng Y, van Klinken RD, Sosa A, Li B, Chen J, Xu C-Y. **2016**. The Relative Importance of Genetic Diversity and Phenotypic Plasticity in Determining Invasion Success of a Clonal Weed in the USA and China. *Frontiers in Plant Science* **7**: 216.

Gordon R, Chapman J, Power A, Chandra S, Roberts J, Cozzolino D. **2019**. Mid-infrared spectroscopy coupled with chemometrics to identify spectral variability in Australian barley samples from different production regions. *Journal of Cereal Science* **85**: 41–47.

Hagenblad J, Hülskötter J, Acharya KP, Brunet J, Chabrerie O, Cousins SAO, Dar PA, Diekmann M, De Frenne P, Hermy M, *et al.* 2015. Low genetic diversity despite multiple introductions of the invasive plant species Impatiens glandulifera in Europe. *BMC Genetics* **16**: 103.

Head L. **2017**. The social dimensions of invasive plants. *Nature Plants* **3**: 1–7.

Heap, B, Holden, CA, Taylor, JE and McAinsh, M 2020. Crosstalk in Signalling Pathways. In *eLS*. Wiley. 1–9.

Hoad SP, Leakey RRB. **1994**. Effects of light quality on gas exchange and dry matter partitioning in Eucalyptus grandis W. Hill ex Maiden. *Forest Ecology and Management* **70**: 265–273.

Holalu S V., Reddy SK, Finlayson SA. **2021**. Low Red Light:Far Red Light Inhibits Branching by Promoting Auxin Signaling. *Journal of Plant Growth Regulation* **40**: 2028–2036.

Holden CA, Morais CLM, Taylor JE, Martin FL, Beckett P, McAinsh M. **2021**. Regional differences in clonal Japanese knotweed revealed by chemometrics-linked attenuated total reflection Fourier-transform infrared spectroscopy. *BMC Plant Biology 2021 21:1* **21**: 1–20.

Hollingsworth ML, Bailey JP. **2000**. Evidence for massive clonal growth in the invasive weed *Fallopia japonica* (Japanese Knotweed). *BokanicalJouml ofrhc Linnean So&&* **133**: 463–472.

Holm AK, Elameen A, Oliver BW, Brandsæter LO, Fløistad IS, Brurberg MB. **2018**. Low genetic variation of invasive Fallopia spp. in their northernmost European distribution range. *Ecology and Evolution* **8**: 755–764.

Hou M, Wu D, Li Y, Tao W, Chao L, Zhang Y. **2021**. The role of auxin in nitrogen-modulated shoot branching. *https://doi.org/10.1080/15592324.2021.1885888* **16**.

Hussain S, Iqbal N, Rahman T, Liu T, Brestic M, Safdar ME, Asghar MA, Farooq MU, Shafiq I, Ali A, *et al.* 2019. Shade effect on carbohydrates dynamics and stem strength of soybean genotypes. *Environmental and Experimental Botany* **162**: 374–382.

Hutengs C, Seidel M, Oertel F, Ludwig B, Vohland M. **2019**. In situ and laboratory soil spectroscopy with portable visible-to-near-infrared and mid-infrared instruments for the assessment of organic carbon in soils. *Geoderma* **355**: 113900.

Izquierdo L, Martín I, Albornos L, Hernández-Nistal J, Hueso P, Dopico B, Labrador E. **2018**. Overexpression of Cicer arietinum βIII-Gal but not βIV-Gal in arabidopsis causes a reduction of cell wall β-(1,4)-galactan compensated by an increase in homogalacturonan. *Journal of Plant Physiology* **231**: 135–146.

Ji W, Adamchuk VI, Biswas A, Dhawale NM, Sudarsan B, Zhang Y, Viscarra Rossel RA, Shi Z. **2016**. Assessment of soil properties in situ using a prototype portable MIR spectrometer in two agricultural fields. *Biosystems Engineering* **152**: 14–27.

Jin N, Semple KT, Jiang L, Luo C, Zhang D, Martin FL. **2018**. Spectrochemical analyses of growth phase-related bacterial responses to low (environmentally-relevant) concentrations of tetracycline and nanoparticulate silver. *Analyst* **143**: 768–776.

Jones D, Bruce G, Fowler MS, Law-Cooper R, Graham I, Abel A, Street-Perrott FA, Eastwood D. **2018**. Optimising physiochemical control of invasive Japanese knotweed. *Biological Invasions* **20**: 2091–2105.

Kebbas S, Lutts S, Aid F. **2015**. Effect of drought stress on the photosynthesis of Acacia tortilis subsp. raddiana at the young seedling stage. *Photosynthetica 2015 53:2* **53**: 288–298.

Kharrat F, Khlif M, Hilliou L, Haboussi M, Covas JA, Nouri H, Bradai C. **2020**. Minimally processed date palm (Phoenix dactylifera L.) leaves as natural fillers and processing aids in poly(lactic acid) composites designed for the extrusion film blowing of thin packages. *Industrial Crops and Products* **154**: 112637.

Van Kleunen M, Weber E, Fischer M. **2010**. A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecology Letters* **13**: 235–245.

Kotilainen T, Aphalo PJ, Brelsford CC, Böök H, Devraj S, Heikkilä A, Hernández R, Kylling A, Lindfors A V., Robson TM. **2020**. Patterns in the spectral composition of sunlight and biologically meaningful spectral photon ratios as affected by atmospheric factors. *Agricultural and Forest Meteorology* **291**: 108041.

Krähmer A, Böttcher C, Gudi G, Stürtz M, Schulz H. **2021**. Application of ATR-FTIR spectroscopy for profiling of non-structural carbohydrates in onion (Allium cepa L.) bulbs. *Food Chemistry* **360**: 129978.

Kuhlgert S, Austic G, Zegarac R, Osei-Bonsu I, Hoh D, Chilvers MI, Roth MG, Bi K, TerAvest D, Weebadde P, *et al.* 2016. MultispeQ Beta: A tool for large-scale plant phenotyping connected to the open photosynQ network. *Royal Society Open Science* **3(10):** 160592.

Larsen DH, Woltering EJ, Nicole CCS, Marcelis LFM. **2020**. Response of Basil Growth and Morphology to Light Intensity and Spectrum in a Vertical Farm. *Frontiers in Plant Science* **11**: 1893.

Lei Y, Hannoufa A, Christensen D, Shi H, Prates LL, Yu P. **2018**. Molecular Structural Changes in Alfalfa Detected by ATR-FTIR Spectroscopy in Response to Silencing of TT8 and HB12 Genes. *International Journal of Molecular Sciences 2018, Vol. 19, Page 1046* **19**: 1046.

Lemoine R. **2000**. Sucrose transporters in plants: update on function and structure. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1465**: 246–262.

Liao H, D'Antonio CM, Chen B, Huang Q, Peng S. **2016**. How much do phenotypic plasticity and local genetic variation contribute to phenotypic divergences along environmental gradients in widespread invasive plants? A meta-analysis. *Oikos* **125**: 905–917.

Liu X, Renard CMGC, Bureau S, Le Bourvellec C. **2021**. Revisiting the contribution of ATR-FTIR spectroscopy to characterize plant cell wall polysaccharides. *Carbohydrate Polymers* **262**: 117935.

Macchioni V, Picchi V, Carbone K. **2022**. Hop leaves as an alternative source of health-active compounds: Effect of genotype and drying conditions. *Plants* **11**: 99.

Mafakheri A, Siosemardeh A, Bahramnejad B, Struik PC, Sohrabi Y. **2010**. Effect of Drought Stress on Yield, Proline and Chlorophyll Contents in Three Chickpea Cultivars. *Australian Journal of Crop Science* **4**: 580–585.

Martin FM, Dommanget F, Janssen P, Spiegelberger T, Clément Viguier ·, Evette A. **2019**. Could knotweeds invade mountains in their introduced range? An analysis of patches dynamics along an elevational gradient. *Alpine Botany* **129**: 33–42.

Martin FM, Müllerová J, Borgniet L, Dommanget F, Breton V, Evette A. **2018**. Using Single- and Multi-Date UAV and Satellite Imagery to Accurately Monitor Invasive Knotweed Species. *Remote Sensing 2018, Vol. 10, Page 1662* **10**: 1662.

Mason MG, Ross JJ, Babst BA, Wienclaw BN, Beveridge CA. **2014**. Sugar demand, not auxin, is the initial regulator of apical dominance. *Proceedings of the National Academy of Sciences of the United States of America* **111**: 6092–6097.

McNear DH, Chaney RL, Sparks DL. **2010**. The hyperaccumulator Alyssum murale uses complexation with nitrogen and oxygen donor ligands for Ni transport and storage. *Phytochemistry* **71**: 188–200.

Met Office. **2019**. UK Regional Climates. *https://www.metoffice.gov.uk/research/climate/maps-and-data/regional-climates/index*.

Michalet S, Rouifed S, Pellassa-Simon T, Fusade-Boyer M, Meiffren G, Nazaret S, Piola F. **2017**. Tolerance of Japanese knotweed s.l. to soil artificial polymetallic pollution: early metabolic responses and performance during vegetative multiplication. *Environmental Science and Pollution Research* **24**: 20897–20907.

Miller JH, Manning ST, Enloe SF. **2013**. A management guide for invasive plants in southern forests. *Gen. Tech. Rep. SRS–131. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station* **131**: 1–120.

Monaghan RM, Paton RJ, Smith LC, Drewry JJ, Littlejohn RP. **2005**. The impacts of nitrogen fertilisation and increased stocking rate on pasture yield, soil physical condition and nutrient losses in drainage from a cattle-grazed pasture. *New Zealand Journal of Agricultural Research* **48**: 227–240.

Morais CLM, Costa FSL, Lima KMG. **2017a**. Variable selection with a support vector machine for discriminating Cryptococcus fungal species based on ATR-FTIR spectroscopy. *Analytical Methods* **9**: 2964–2970.

Morais CLM, Costa FSL, Lima KMG. **2017b**. Variable selection with a support vector machine for discriminating: Cryptococcus fungal species based on ATR-FTIR spectroscopy. *Analytical Methods* **9**: 2964–2970.

Morais CLM, Lima KMG. **2018**. Principal Component Analysis with Linear and Quadratic Discriminant Analysis for Identification of Cancer Samples Based on Mass Spectrometry. *Article J. Braz. Chem. Soc* **29**: 472–481.

Morais CLM, Lima KMG, Singh M, Martin FL. **2020**. Tutorial: multivariate classification for vibrational spectroscopy in biological samples. *Nature Protocols* **15**: 2143–2162.

Morris TL, Esler KJ, Barger NN, Jacobs SM, Cramer MD. **2011**. Ecophysiological traits associated with the competitive ability of invasive Australian acacias. *Diversity and Distributions* **17**: 898–910.

Moskal P, Wesełucha-Birczyńska A, Łabanowska M, Filek M. **2019**. Adaxial and abaxial pattern of Urtica dioica leaves analyzed by 2DCOS ATR-FTIR as a function of their growth time and impact of environmental pollution. *Vibrational Spectroscopy* **104**: 102948.

Mounger J, Ainouche ML, Bossdorf O, Cavé-Radet A, Li B, Parepa M, Salmon A, Yang J, Richards CL. **2021**. Epigenetics and the success of invasive plants. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376 (1826):** 20200117.

Nicotra AB, Atkin OK, Bonser SP, Davidson AM, Finnegan EJ, Mathesius U, Poot P, Purugganan MD, Richards CL, Valladares F, *et al.* **2010**. Plant phenotypic plasticity in a changing climate. *Trends in Plant Science* **15**: 684–692.

Nogales-Bueno J, Baca-Bocanegra B, Rooney A, Miguel Hernández-Hierro J, José Heredia F, Byrne HJ. **2017**. Linking ATR-FTIR and Raman features to phenolic extractability and other attributes in grape skin. *Talanta* **167**: 44–50.

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

Pandey P, Ge Y, Stoerger V, Schnable JC. **2017**. High throughput in vivo analysis of plant leaf chemical properties using hyperspectral imaging. *Frontiers in Plant Science* **8**: 1348.

Parepa M, Fischer M, Bossdorf O. **2013**. Environmental variability promotes plant invasion. *Nature Communications* **4**: 1–4.

Parepa M, Kahmen A, Werner RA, Fischer M, Bossdorf O. **2019**. Invasive knotweed has greater nitrogen-use efficiency than native plants: evidence from a 15N pulse-chasing experiment. *Oecologia* **191**: 389–396.

Park Y, Runkle ES. **2018**. Spectral effects of light-emitting diodes on plant growth, visual color quality, and photosynthetic photon efficacy: White versus blue plus red radiation. *PLoS ONE* **13(8)**: e0202386.

Parker IM, Rodriguez J, Loik ME. **2003**. An Evolutionary Approach to Understanding the Biology of Invasions: Local Adaptation and General-Purpose Genotypes in the Weed Verbascum thapsus. *Conservation Biology* **17**: 59–72.

Pejchar L, Mooney HA. **2009**. Invasive species, ecosystem services and human well-being. *Trends in Ecology and Evolution* **24**: 497–504.

Pennisi G, Pistillo A, Orsini F, Cellini A, Spinelli F, Nicola S, Fernandez JA, Crepaldi A, Gianquinto G, Marcelis LFM. **2020**. Optimal light intensity for sustainable water and energy use in indoor cultivation of lettuce and basil under red and blue LEDs. *Scientia Horticulturae* **272**: 109508.

Pichancourt J-B, van Klinken RD. **2012**. Phenotypic Plasticity Influences the Size, Shape and Dynamics of the Geographic Distribution of an Invasive Plant (G Bonaventure, Ed.). *PLoS ONE* **7**: e32323.

Pierik R, Cuppens MLC, Voesenek LACJ, Visser EJW. **2004**. Interactions between Ethylene and Gibberellins in Phytochrome-Mediated Shade Avoidance Responses in Tobacco. *Plant Physiology* **136**: 2928–2936.

Price EAC, Gamble R, Williams GG, Marshall C. **2002**. Seasonal patterns of partitioning and remobilization of 14C in the invasive rhizomatous perennial Japanese knotweed (*Fallopia japonica* (Houtt.) Ronse Decraene). In: Ecology and Evolutionary Biology of Clonal Plants. Dordrecht: Springer Netherlands, 125–140.

Pyšek P, Hulme PE, Simberloff D, Bacher S, Blackburn TM, Carlton JT, Dawson W, Essl F, Foxcroft LC, Genovesi P, *et al.* 2020. Scientists' warning on invasive alien species. *Biological Reviews* **95**: 1511–1534.

Rana R, Herz K, Bruelheide H, Dietz S, Haider S, Jandt U, Pena R. **2018**. Leaf Attenuated Total Reflection Fourier Transform Infrared (ATR-FTIR) biochemical profile of grassland plant species related to land-use intensity. *Ecological Indicators* **84**: 803–810.

Richards CL, Bossdorf O, Muth NZ, Gurevitch J, Pigliucci M. **2006**. Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecology Letters* **9**: 981–993.

Richards CL, Schrey AW, Pigliucci M. **2012**. Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation (M Vellend, Ed.). *Ecology Letters* **15**: 1016–1025.

Richards CL, Walls RL, Bailey JP, Parameswaran R, George T, Pigliucci M. **2008**. Plasticity in salt tolerance traits allows for invasion of novel habitat by Japanese knotweed s. l. (*Fallopia japonica* and *F.*x *bohemica*, Polygonaceae). *American Journal of Botany* **95**: 931–942.

Rouifed S, Byczek C, Laffray D, Piola F. **2012**. Invasive Knotweeds are Highly Tolerant to Salt Stress. *Environmental Management* **50**: 1027–1034.

Ruoff K, Luginbu̇hl W, Luginbu̇hl L, Ku̇nzli R, Ku̇nzli K, Äa M, Iglesias T, Bogdanov S, Bosset JO, Von Der Ohe K, von der Ohe, W. and Amadò, R. 2006. Authentication of the Botanical and Geographical Origin of Honey by Mid-Infrared Spectroscopy. *Journal of agricultural and food chemistry* **54(18)**: 6873-6880.

Science and Technology Committee - House of Commons. **2019**. Japanese knotweed and the built environment. *Science and Technology Committee Report*: 13.

Shackleton RT, Shackleton CM, Kull CA. **2019**. The role of invasive alien species in shaping local livelihoods and human well-being: A review. *Journal of Environmental Management* **229**: 145–157.

Sharma S, Uttam KN. **2017**. Rapid analyses of stress of copper oxide nanoparticles on wheat plants at an early stage by laser induced fluorescence and attenuated total reflectance Fourier transform infrared spectroscopy. *Vibrational Spectroscopy* **92**: 135–150.

Skolik P, McAinsh MR, Martin FL. **2019a**. ATR-FTIR spectroscopy non-destructively detects damage-induced sour rot infection in whole tomato fruit. *Planta* **249**: 925–939.

Skolik P, Morais CLM, Martin FL, McAinsh MR. **2019b**. Determination of developmental and ripening stages of whole tomato fruit using portable infrared spectroscopy and Chemometrics. *BMC Plant Biology* **19**: 236.

Smith H. **1982**. Light quality, photoperception, and plant strategy. *Annual review of plant physiology* **33**: 481–518.

Sołtysiak J. **2020**. Heavy metals tolerance in an invasive weed (*Fallopia japonica*) under different levels of soils contamination. *Journal of Ecological Engineering* **21**: 81–91.

Soriano-Disla JM, Janik LJ, McLaughlin MJ. **2018**. Assessment of cyanide contamination in soils with a handheld mid-infrared spectrometer. *Talanta* **178**: 400–409.

Strong R, Martin FL, Jones KC, Shore RF, Halsall CJ. **2017**. Subtle effects of environmental stress observed in the early life stages of the Common frog, Rana temporaria. *Scientific Reports* **7**: 1–13.

Talari ACS, Martinez MAG, Movasaghi Z, Rehman S, Rehman IU. **2017**. Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Applied Spectroscopy Reviews* **52**: 456–506.

Thalmann M, Santelia D. **2017**. Starch as a determinant of plant fitness under abiotic stress. *New Phytologist* **214**: 943–951.

Traoré M, Kaal J, Martínez Cortizas A. **2018**. Differentiation between pine woods according to species and growing location using FTIR-ATR. *Wood Science and Technology* **52**: 487–504.

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

Usman K, Al-Ghouti MA, Abu-Dieyeh MH. **2019**. The assessment of cadmium, chromium, copper, and nickel tolerance and bioaccumulation by shrub plant Tetraena qataranse. *Scientific Reports* **9**: 1–11.

Valliere JM, Escobedo EB, Bucciarelli GM, Sharifi MR, Rundel PW. **2019**. Invasive annuals respond more negatively to drought than native species. *New Phytologist* **223**: 1647–1656.

Vohland M, Ludwig B, Seidel M, Hutengs C. **2022**. Quantification of soil organic carbon at regional scale: Benefits of fusing vis-NIR and MIR diffuse reflectance data are greater for in situ than for laboratory-based modelling approaches. *Geoderma* **405**: 115426.

Waldron KW, Faulds CB. **2007**. Cell Wall Polysaccharides: Composition and Structure. *Comprehensive Glycoscience: From Chemistry to Systems Biology* **1–4**: 181–201.

Walls RL. **2010**. Hybridization and Plasticity Contribute to Divergence Among Coastal and Wetland Populations of Invasive Hybrid Japanese Knotweed s.l. {Fallopia spp.). *Estuaries and Coasts* **33**: 902–918.

Wilson SD, Tilman D. **1991**. Components of plant competition along an experimental gradient of nitrogen availability. *Ecology* **72**: 1050–1065.

Wu H-C, Bulgakov VP, Jinn T-L. **2018**. Pectin Methylesterases: Cell Wall Remodeling Proteins Are Required for Plant Response to Heat Stress. *Frontiers in Plant Science* **9**: 1612.

Yang J, Duan G, Li C, Liu L, Han G, Zhang Y, Wang C. **2019**. The Crosstalks Between Jasmonic Acid and Other Plant Hormone Signaling Highlight the Involvement of Jasmonic Acid as a Core Component in Plant Response to Biotic and Abiotic Stresses. *Frontiers in Plant Science* **10**: 1349.

Yuan W, Pigliucci M, Richards CL. **2022**. Rapid phenotypic differentiation and local adaptation in Japanese knotweed s.l. (*Reynoutria japonica and R. × bohemica*, Polygonaceae) invading novel habitats. *bioRxiv*: 2022.03.07.483296.

Zeghoud S, Rebiai A, Hemmami H, Seghir B Ben, Elboughdiri N, Ghareba S, Ghernaout D, Abbas N. **2021**. ATR–FTIR Spectroscopy, HPLC Chromatography, and Multivariate Analysis for Controlling Bee Pollen Quality in Some Algerian Regions. *ACS Omega* **6**: 4878–4887.

Zhang Y-Y, Parepa M, Fischer M, Bossdorf O. **2016**. Epigenetics of colonizing species? A study of Japanese knotweed in Central Europe. In: Barrett SCH, Colautti RI, Dlugosch KM, Rieseberg LH (Eds) Invasion Genetics. Chichester, UK: John Wiley & Sons, Ltd, 328–340.

Zhou C, Jiang W, Via BK, Fasina O, Han G. **2015**. Prediction of mixed hardwood lignin and carbohydrate content using ATR-FTIR and FT-NIR. *Carbohydrate Polymers* **121**: 336–341.

Zou T, Huang C, Wu P, Ge L, Xu Y. **2020**. Optimization of Artificial Light for Spinach Growth in Plant Factory Based on Orthogonal Test. *Plants* **9**: 490.

6.13 Supporting Information
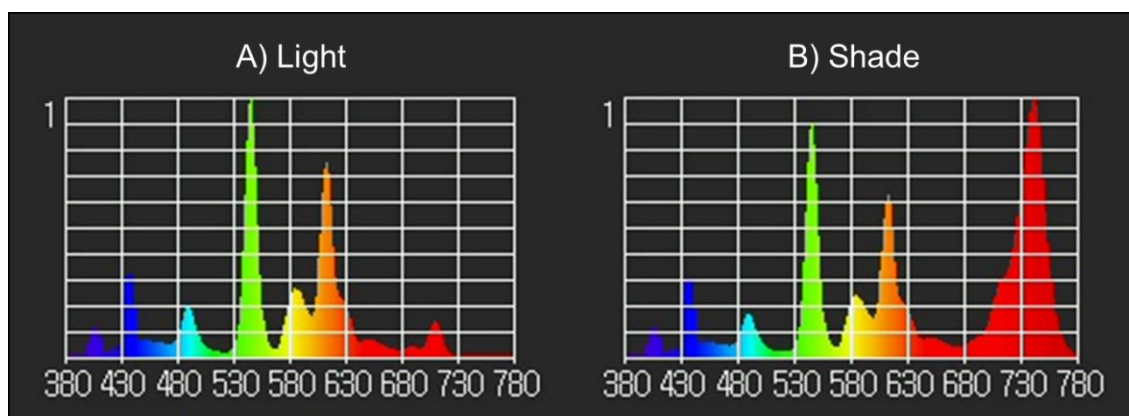
Figure S6.1: Spectra from **a)** 'Light' **b)** 'Shade' cabinets, providing red: far-red ratios of 5.6 and 0.4 respectively.

Table S6.1: Lighting conditions within each Snijder cabinet.

	'Light' Groups: LC, LD, LN and LLN	'Shade' Groups: SC, SD, SN and SLN
PFD-R$_{(700-780\ nm)}$	72.51	49.28
PFD-FR$_{(600-700\ nm)}$	12.89	116.5
photosynthetic photon flux density PPFD$_{(400-700\ nm)}$	189.8	124.7
PFD-UV$_{(380-400\ nm)}$	0.5677	0.4402
PFD-B$_{(400-500\ nm)}$	33.93	21.58
PFD-G$_{(500-600\ nm)}$	83.40	53.87
peak wavelength λp / nm	545	741
peak wavelength value λpV / mWm^{-2}nm^{-1}	827.7	576.0
Irradiance	43.2	45.8
Illuminance/ lux.	15128	9617

Table S6.2: Reagents used for Hoagland's solution. Full strength Hoagland's solution was made using 100 mL of solution A, 100 mL of solution B and 10 mL of solution C in 10 L of deionised water.

Solution	Reagent	Concentration/ gL^{-1}
A (100 mL)	NH_4NO_3	8.000
	$Ca(NO_3)_2.4H_2O$	82.600
	KNO_3	35.700
B (100 mL)	KNO_3	5.000
	KH_2PO_4	27.400
	$MgSO_4.7H_2O$ *added first	24.600
	$MnSO_4.5H_2O$	0.053
	H_3BO_3	0.140
	$CuSO_4.5H_2O$	0.015
	$(NH_4)_6Mo_7O_{24}.4H_2O$	0.008
	$ZnSO_4.7H_2O$	0.060
C (10 mL)	Fe-EDTA	36.71

Table S6.3: The eight different treatment environments described in binary format
with each controlled variable (nitrogen, micronutrients, water, and R:FR) scored a
value of one (high) or zero (low).

Characterising the cabinet environments in binary format				
Environment letter	Nitrogen	Micronutrients	Red: Far-red Ratio	Water
A	1	1	1	1
B	1	1	1	0
C	1	0	1	1
D	0	0	1	1
E	1	1	0	1
F	1	1	0	0
G	1	0	0	1
H	0	0	0	1

Figure S6.2: Environmental effects on **a)** aboveground fresh weight, **b)** gH+, **c)** vH+, **d)** PhiNO and **e)** PS1 active centres, grouped by treatment (LC, LD, LN, LLN, SC, SD, SN, and SLN). Statistical significance was calculated using a Kruskal-Wallis followed by a post hoc test using the criterium Fisher's least significant difference (LSD) to determine where the difference lies, signified by lowercase letters above the bars. Within each graph, all bars which share letters are not significantly different from each other. Data are mean +/- standard errors. N values for each group respectively: 6, 8, 7, 6, 4, 6, 6, 6 for aboveground fresh weight and 7, 5, 7, 7, 5, 3, 5, 6 for all other categories (gH+ vH+, PhiNO and PS1 active centres). Thylakoid proton conductance, gH+, was 11% lower in LC than SLN (Figure S6.2b). Steady state proton flux from the stroma to the lumen, vH+, was ~24% higher in SD, SN, and SLN than in LD (Figure S6.2c). PhiNO, the amount of incoming light that is neither used for photosynthesis nor is dissipated that has the potential to damage the leaf, is lowest in LN, and up to ~8% higher in SC and SN (Figure S6.2d). PS1 active centres, which represents the amount of photosystem I which was operational and could pass or receive electrons, in LD were approximately three times those of SC, SN and SLN (Figure S6.2e).
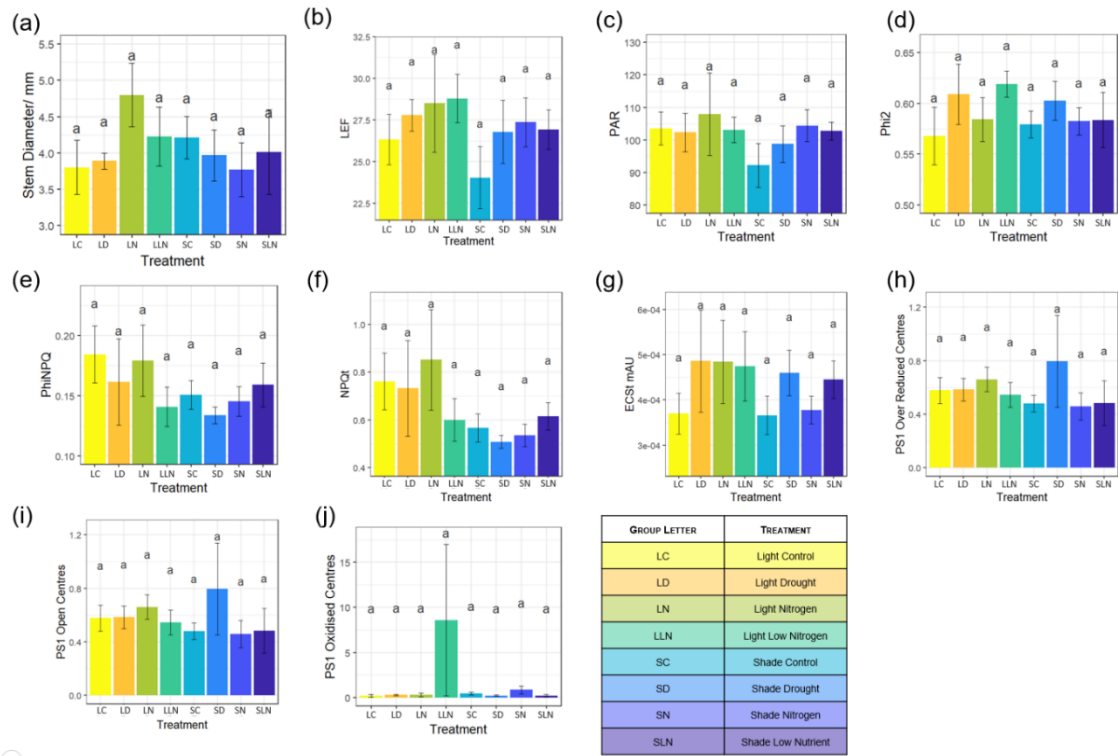
Figure S6.3: Environmental effects on **a)** stem diameter (which averaged 4.08 mm), **b)** linear electron flow (LEF), **c)** PAR, **d)** Phi2, **e)** PhiNPQ, **f)** NPQt, **g)** ECSt, **h)** PS1 over reduced centres, **i)** PS1 open centres, and **j)** PS1 oxidised centres, grouped by treatment (LC, LD, LN, LLN, SC, SD, SN, and SLN). Statistical significance was calculated using a Kruskal-Wallis followed by a post hoc test using the criterium Fisher's least significant difference (LSD) to determine where the difference lies, signified by lowercase letters above the bars. Within each graph, all bars which share letters are not significantly different from each other. Data are mean +/- standard errors. N values for each group respectively: 7, 7, 7, 8, 5, 6, 6, 7 for stem diameter and 6, 8, 7, 6, 4, 6, 6, 6 for all other categories. There was no significant difference between treatments in measured stem diameter (which averaged 4.08 mm), linear electron flow (LEF), PAR, Phi2, PhiNPQ, NPQt, ECSt, PS1 over reduced centres, PS1 open centres, or PS1 oxidised centres.
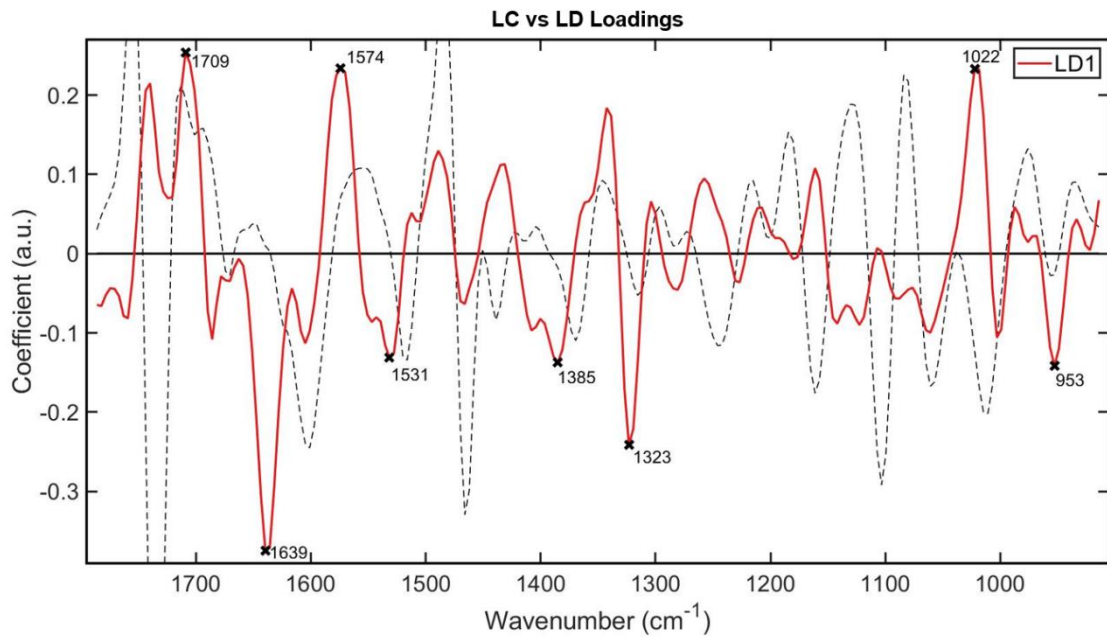
Figure S6.4: LC vs LD PCA loadings, the key wavenumbers which differentiate spectral profiles of intact dried leaf samples grown under the different treatment conditions LC and LD from one another.
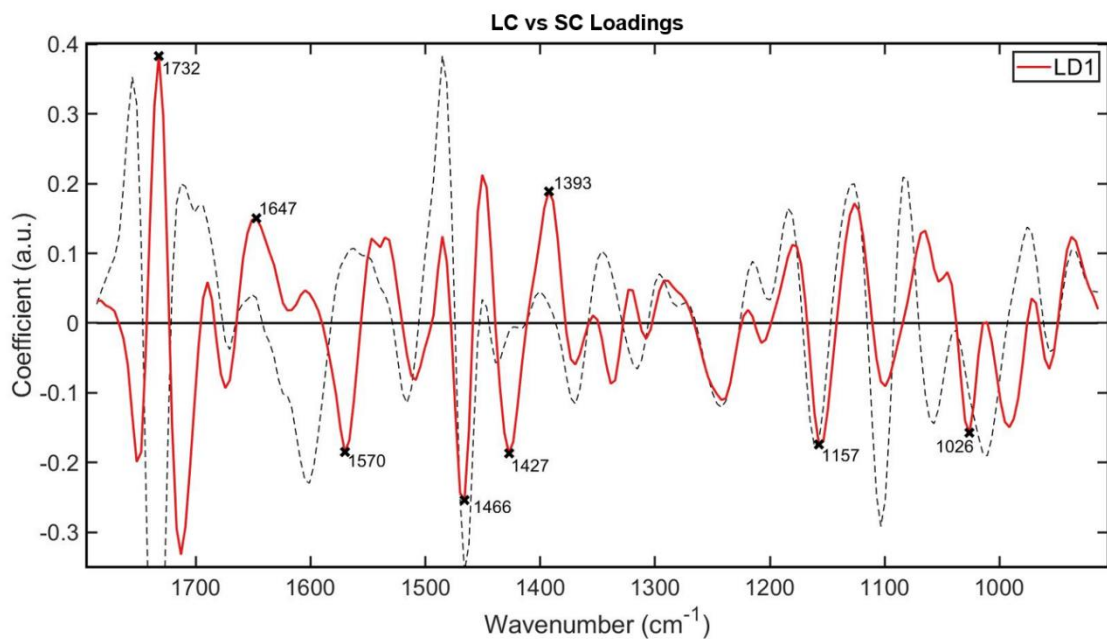
Figure S6.5: LC vs SC PCA loadings, the key wavenumbers which differentiate spectral profiles of intact dried leaf samples grown under the different treatment conditions LC and SC from one another.
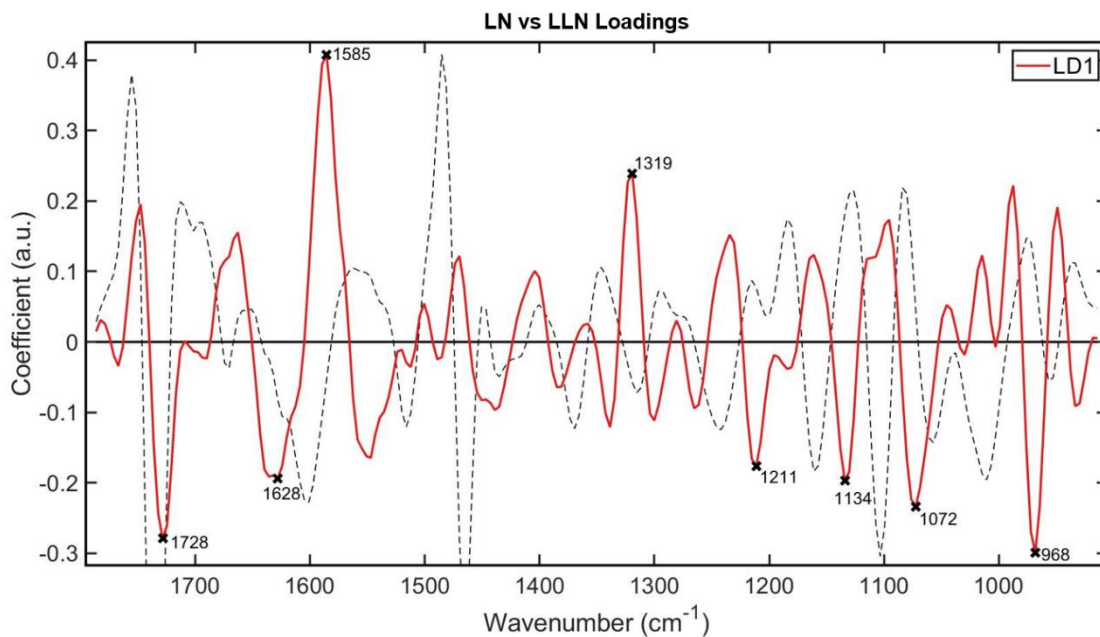
Figure S6.6: LN vs LLN PCA loadings, the key wavenumbers which differentiate spectral profiles of intact dried leaf samples grown under the different treatment conditions LN and LLN from one another.
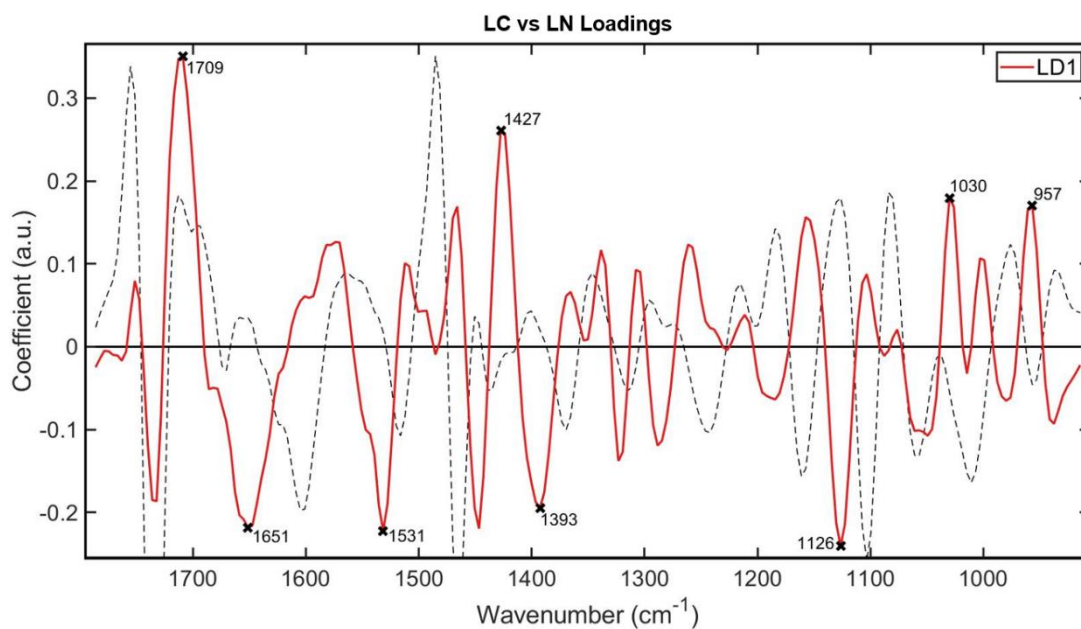
Figure S6.7: LC vs LN PCA loadings, the key wavenumbers which differentiate spectral profiles of intact dried leaf samples grown under the different treatment conditions LC vs LN from one another.
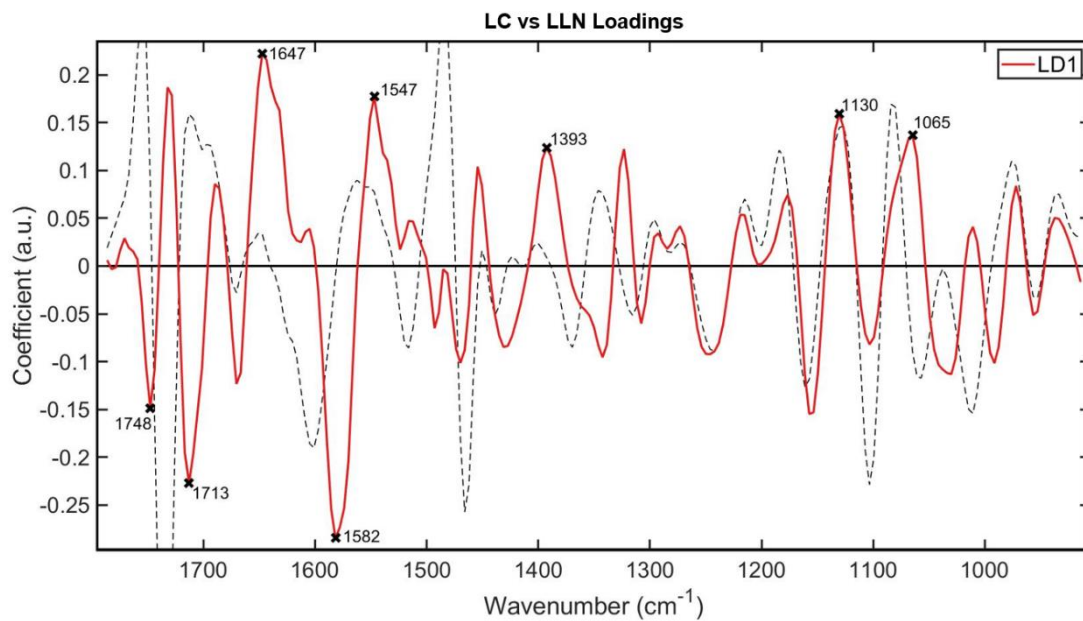
Figure S6.8: LC vs LLN PCA loadings, the key wavenumbers which differentiate spectral profiles of intact dried leaf samples grown under the different treatment conditions LC vs LLN from one another.
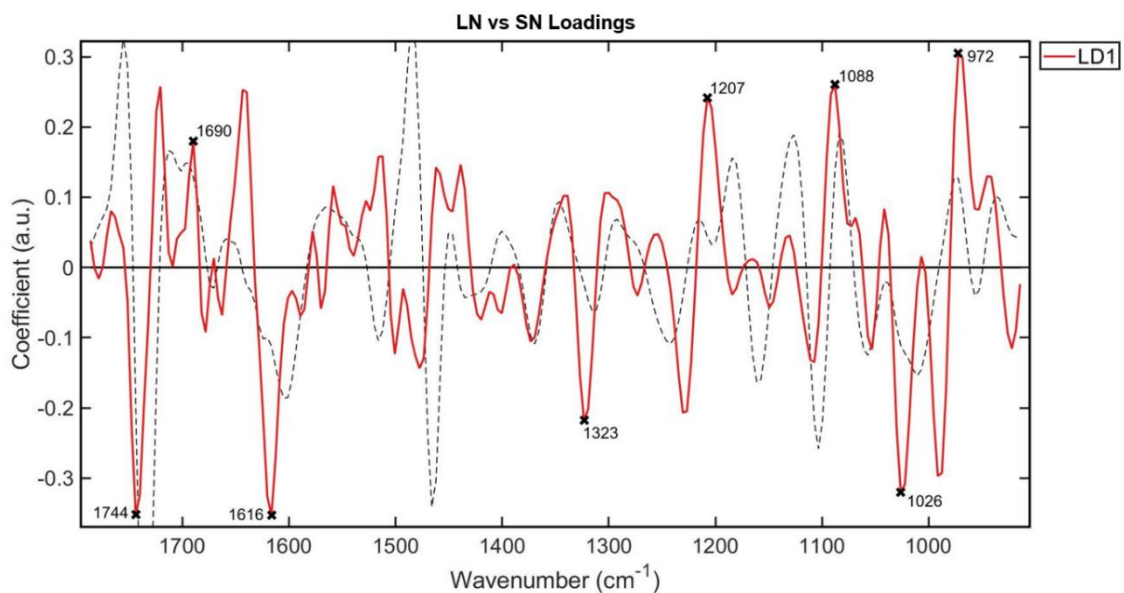
Figure S6.9: LN vs SN PCA loadings, the key wavenumbers which differentiate spectral profiles of intact dried leaf samples grown under the different treatment conditions LN vs SN from one another.
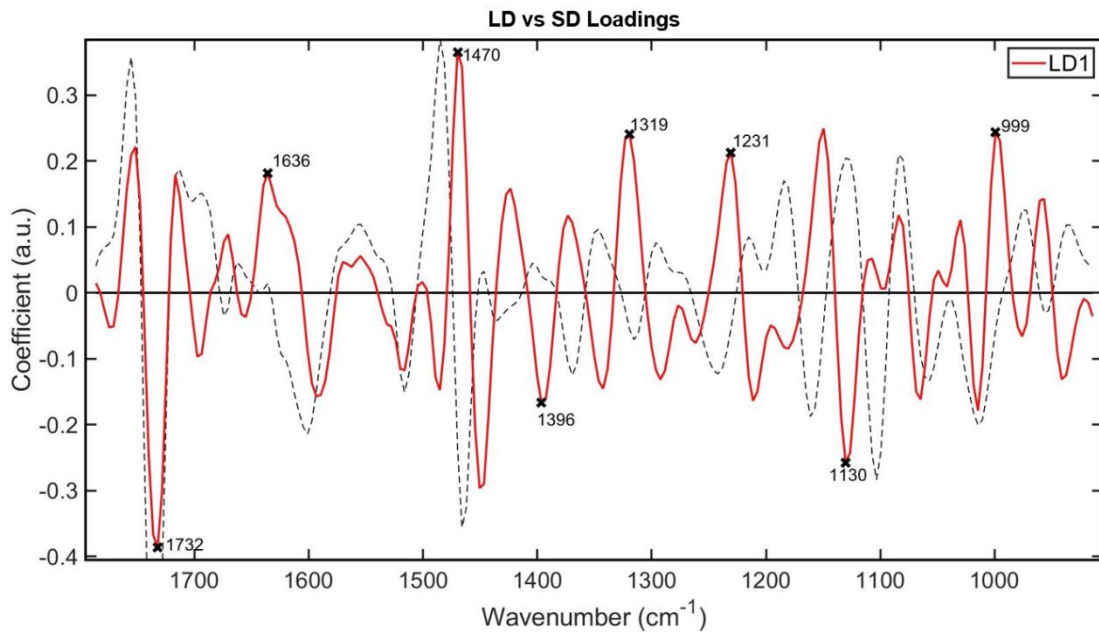
Figure S6.10: LD vs SD PCA loadings, the key wavenumbers which differentiate spectral profiles of intact dried leaf samples grown under the different treatment conditions LD vs SD from one another.
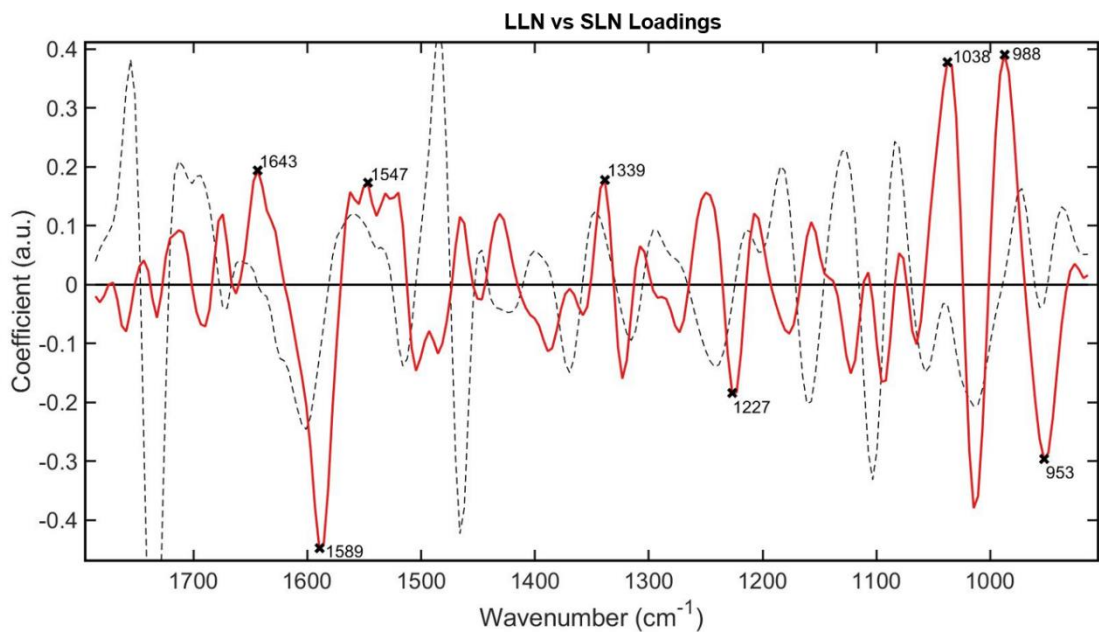
Figure S6.11: LLN vs SLN PCA loadings, the key wavenumbers which differentiate spectral profiles of intact dried leaf samples grown under the different treatment conditions LLN vs SLN from one another.
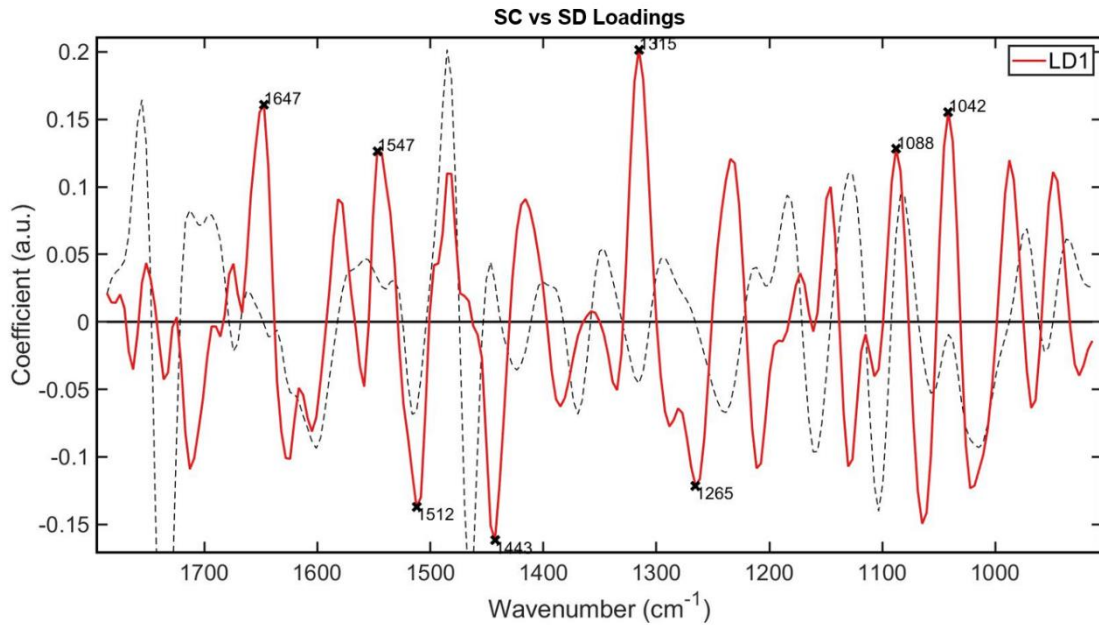
Figure S6.12: SC vs SD PCA loadings, the key wavenumbers which differentiate spectral profiles of intact dried leaf samples grown under the different treatment conditions SC vs SD from one another.
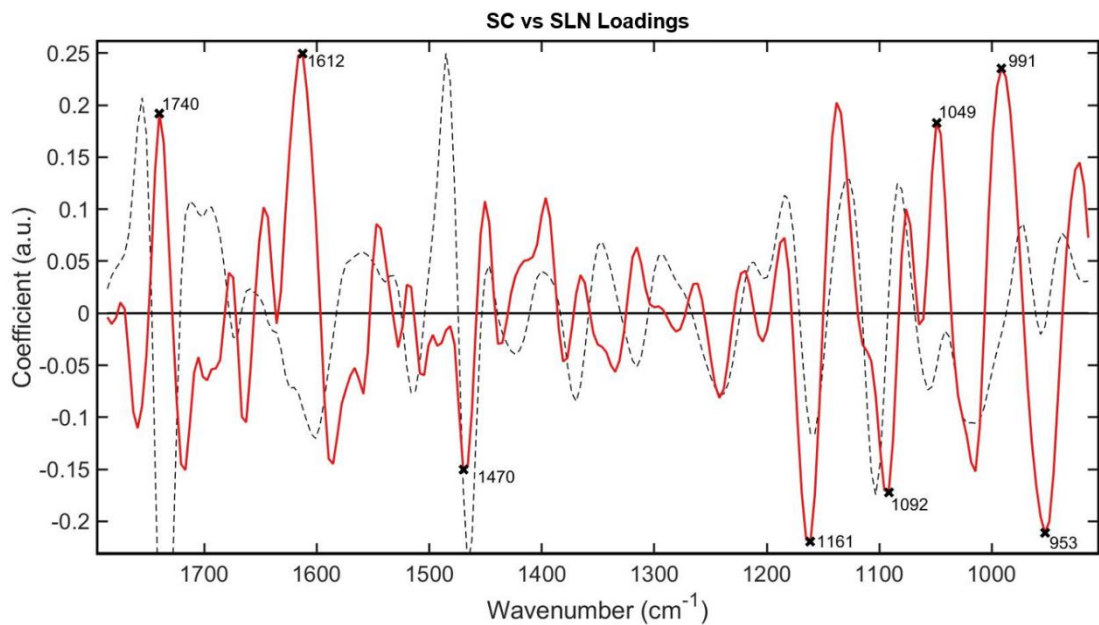
Figure S6.13: SC vs SLN PCA loadings, the key wavenumbers which differentiate spectral profiles of intact dried leaf samples grown under the different treatment conditions SC vs SLN from one another.
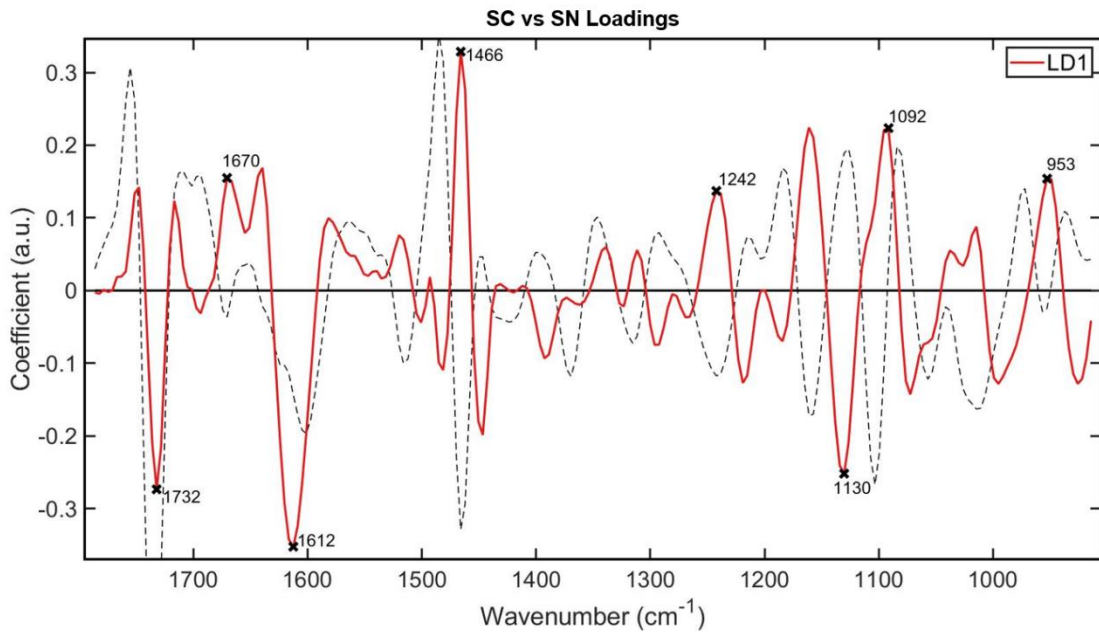
Figure S6.14: SC vs SN PCA loadings, the key wavenumbers which differentiate spectral profiles of intact dried leaf samples grown under the different treatment conditions SC vs SN from one another.
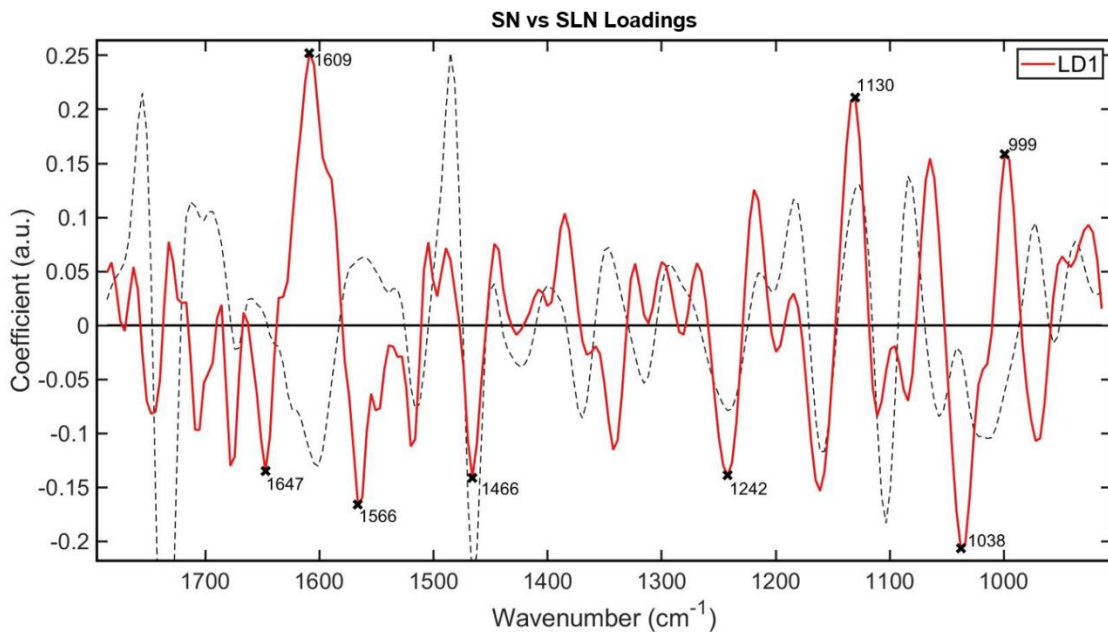
Figure S6.15: SN vs SLN PCA loadings, the key wavenumbers which differentiate spectral profiles of intact dried leaf samples grown under the different treatment conditions SN vs SLN from one another.
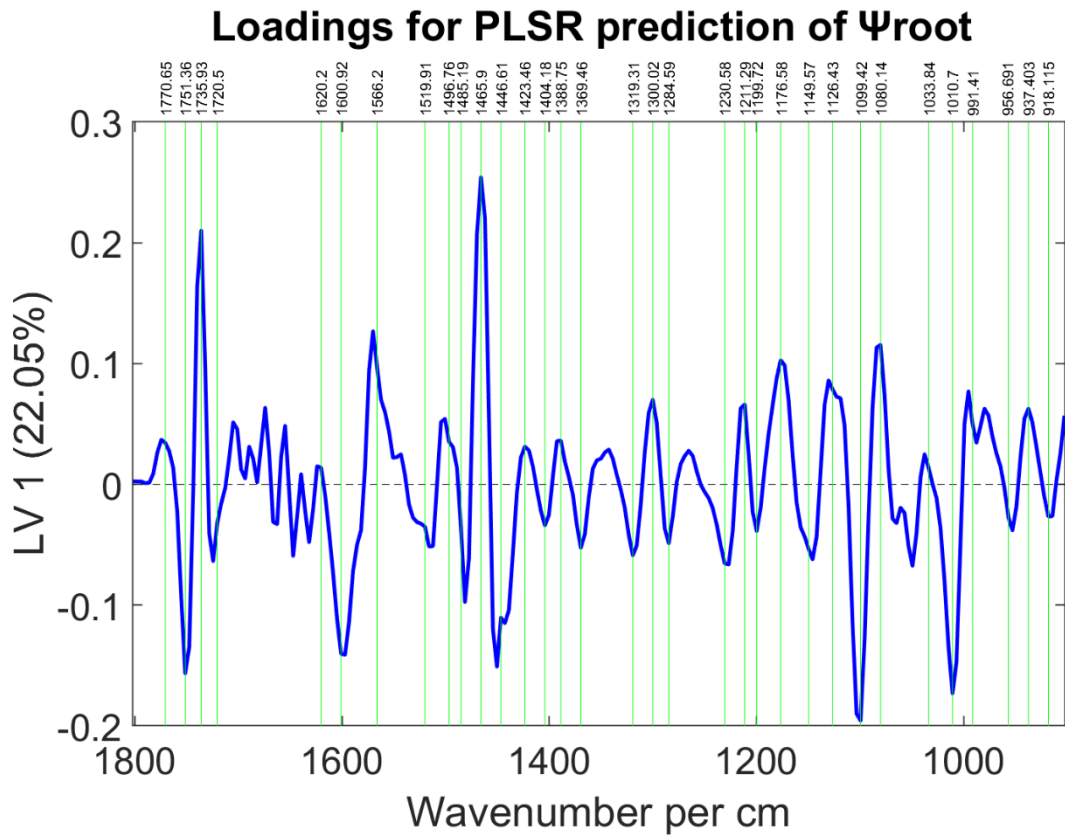
Figure S6.16: PLSR loadings, the key wavenumbers which form the prediction of root water potential (Ψ_{root}) based on the spectral profiles of intact dried leaf samples grown under the treatment conditions LC and LD.

7 General Discussion

7.1 Adaptive plasticity may contribute to the success of Japanese knotweed

The overarching aim of this series of studies was to mitigate the negative impact of the invasive plant species, Japanese knotweed, and related species, with consideration of the capability of vibrational spectroscopy as a method to extract valuable chemical information from samples. This plant remains a challenging problem and inflicts negative environmental (Lavoie, 2017) and socio-economic (Williams *et al.*, 2010) impacts in its introduced range. It successfully occupies a variety of environments (Rouifed *et al.*, 2012; Richards *et al.*, 2012; Michalet *et al.*, 2017) and a broad geographic range (Bailey & Conolly, 2000; Barney *et al.*, 2006; Bzdega *et al.*, 2012; Clements *et al.*, 2016; Holm *et al.*, 2018). Plasticity has been attributed as the key to the success of invasive species (Geng *et al.*, 2016), and reports of this trait in Japanese knotweed have been echoed in many studies (Richards *et al.*, 2008; VanWallendael *et al.*, 2018). These reports of environmentally induced phenotypes were supported by the findings of current research, in which plant environmental responses were detected though chemometric investigation of differences in their ATR-FTIR spectral profiles; see Chapter 3 for regional variation in Japanese knotweed (Holden *et al.*, 2021), Chapter 4 for geographical differentiation in historic herbarium samples (Holden *et al.*, 2022), and see Chapters 5 and 6 for phenotypic responses to controlled environmental conditions. Distinct differences in the spectral fingerprint region (1800-900 cm^{-1}) of Japanese knotweed samples were sufficient to successfully identify plants grown in different geographical regions and controlled environments to a high accuracy using support vector machine chemometrics. These results support the hypothesis that epigenetic

modifications or somatic mutation may contribute to the phenotypic plasticity responsible for the ability of Japanese knotweed to invade and thrive in such diverse habitats. This hypothesis is all the more likely due to the clonal nature of the plant, as asexual plants are thought to bypass the meiotic resetting of epigenetic modifications, thereby maintaining their genetic variation (Schoen & Schultz, 2019; Mounger *et al.*, 2021).

7.2 Outlook for future IAS control measures

7.2.1 Phenotypic plasticity and herbicide interaction

High levels of plasticity in this species highlight a need for closer examination of its interaction with herbicides because: i) adaptive plasticity has proven to be an important mechanism underlying herbicide resistance in other weed species (Josephs *et al.*, 2021); and ii) a bi-phasic dose response to herbicides under stressful conditions has been suggested as a method to quantitatively measure biological plasticity (Agathokleous *et al.*, 2019). This bi-phasic response to herbicides is a phenomenon known as 'hormesis' in which a sub-lethal dose provides a beneficial outcome, for example the nitrogen and phosphorus provided by the glyphosate can act as a fertiliser in some weeds (Cedergreen *et al.*, 2016). The possibility of hormesis in Japanese knotweed is a concern expressed elsewhere (see Peters, 2019), but has yet to be quantified in this species. Japanese knotweed is notoriously difficult to control, a trait demonstrated by field trials in which a range of physiochemical treatment methods failed to eradicate it within a three year period (Jones D *et al.*, 2018). The study by Jones D *et al.* (2018) suggests that this species frequently receives sub-lethal doses of herbicides, which was supported here in Chapter 3 by the observation of a biomarker for glyphosate in seemingly healthy-looking plants (Holden *et al.*, 2021).

7.2.2 Reducing the carbon footprint of invasive species control

Management of Japanese knotweed on private land is largely reliant on application of the herbicide, glyphosate, or alternatively by the physical excavation of plants and their rhizomes (Jones D *et al.*, 2018). In ecological contexts, remediation also includes supporting the growth of competitive native species (Brown *et al.*, 2008; Martin FM *et al.*, 2020). Treatment methods for weed-control are energy intensive (Table 7.1), see detailed comparisons in (Coleman *et al.*, 2019). The carbon footprint of the control process is augmented by the need for repetitive glyphosate applications over a minimum of three years. For example, using the company Phlorum Ltd as a case study, annually only 15.6 CO_2e tonnes were released to produce the glyphosate itself compared with 2505 CO_2e tonnes released from the transportation of staff and equipment between offices and treatment sites (data correct as of May 2018). These figures demonstrate a great potential for reduction of the carbon footprint created by Japanese knotweed management, by way of a more effective single-application herbicide. In addition to these energy requirements, the demand for a new herbicide effective against Japanese knotweed has been amplified by the safety concerns surrounding glyphosate after recently being named as a putative carcinogen (Tarone, 2018). An innovative herbicide must use new chemistry with novel mechanisms of action (MOA). Herbicide target sites include photosynthetic electron transport, hormones, microtubules, protein regulation, and synthesis of cellulose, lipids, nucleic acids, carotenoid, plastoquinone, amino acids, and chlorophyll (Dayan, 2019).

Table 7.1: Energy consumption for weed-control strategies (Tehrani *et al.*, 2014; Coleman *et al.*, 2019).*Excavation values were collected from a different source so the carbon-scopes may not be comparable.

Treatment approach	Treatment type	Energy consumption MJ ha^{-1}
Tillage	Flex-tine harrow	4.2-5.5
	Sweep cultivator	13-14
	Rotary hoe	12-17
Thermal	Flaming	1,008-4,334
	Infrared	2,000-3,887
Alternative	Electrocution	15-19
	Laser pyrolysis	15-249
	Hoeing	17
	Herbicides	15
*Excavation**	Hydraulic	12
	Backhoe	15
	Steer loader	14

7.2.3 ATR–FTIR spectroscopy for herbicide screening

Current herbicide screening techniques can be target-based, which works in 'reverse' from the effector-molecule to discover the whole-plant phenotypic effect, or phenotype-based, which works 'forward' from an observed phenotype to elucidate the mechanistic pathway (Dejonghe & Russinova, 2017). In ligand or target-based drug discovery, a putative target site is immobilised on a column and potential new drugs are passed through, in the search for a 'hit' (Lein *et al.*, 2004; Schenone *et al.*, 2013). This screening process can also be done virtually if the three-dimensional structure of the target protein is available (López-Ramos & Perruccio, 2010). However, these target-based approaches require prior

knowledge of an experimentally identified target compound. In the phenotype-based method, induced changes in a phenotype are detected first, and then the molecular process and target site is discovered afterwards (Schenone *et al.*, 2013).

ATR-FTIR spectroscopy is a sensitive technique which can pick up small changes in molecules of interest. For example, when traces of glyphosate were detected using ATR-FTIR spectroscopy in plants which appeared visually healthy, differences in carotenoids were also noticed, one of the target sites targeted by some herbicides, see Chapter 3 (Holden *et al.*, 2021). This sensitivity to minute change makes ATR-FTIR spectroscopy an ideal technique for screening new herbicides, showing potential to reduce the cost of the screening process by detecting changes caused by even sub-lethal doses of expensive compounds. For recalcitrant invasive species such as Japanese knotweed, this species-specific approach would have a targeted advantage over current high throughput methods such as the use of green algae suspensions of *Chlorella* as a model organism to detect a phenotypic response (Ma *et al.*, 2002). Plant responses to novel chemicals could be monitored using ATR-FTIR spectroscopy and scanned for changes in key stress biomarkers such as plant hormones and other herbicide target sites. This would allow selection of herbicides with the most potential before refining their chemical structure. Potential biomarkers would include changes to cell signalling molecules. Cell signalling, a process in which a cascade of biomolecules transfer specific information, is the first essential step to allow a response to perceived threats (Heap *et al.*, 2020). Here results have shown that biochemical stress signalling molecules such as hormones can be quantitatively predicted using ATR-FTIR spectrometry, see Chapter 5. The biomarkers detected in response to abiotic stress were common to other studies examining biotic stress, for example those in (Ord *et al.*, 2016). Physiological plant stress responses to abiotic stress, such as root water pressure, were also predicted using mid-infrared (MIR) spectroscopy and chemometrics, see Chapter 6.

7.2.4 ATR-FTIR for hybrid identification

One of the challenges presented by invasive plants is that they can sometimes hybridise with natives, and whilst they may look alike, their response to treatment strategies can be very different. This is certainly the case for the Japanese knotweed hybrid complex (Bailey, 2013), in which the Bohemian hybrid may be more invasive than the traditional variety (Parepa *et al.*, 2014). Indeed, Japanese, Giant and their hybrid Bohemian knotweed can all appear similar (see Figure 7.1).
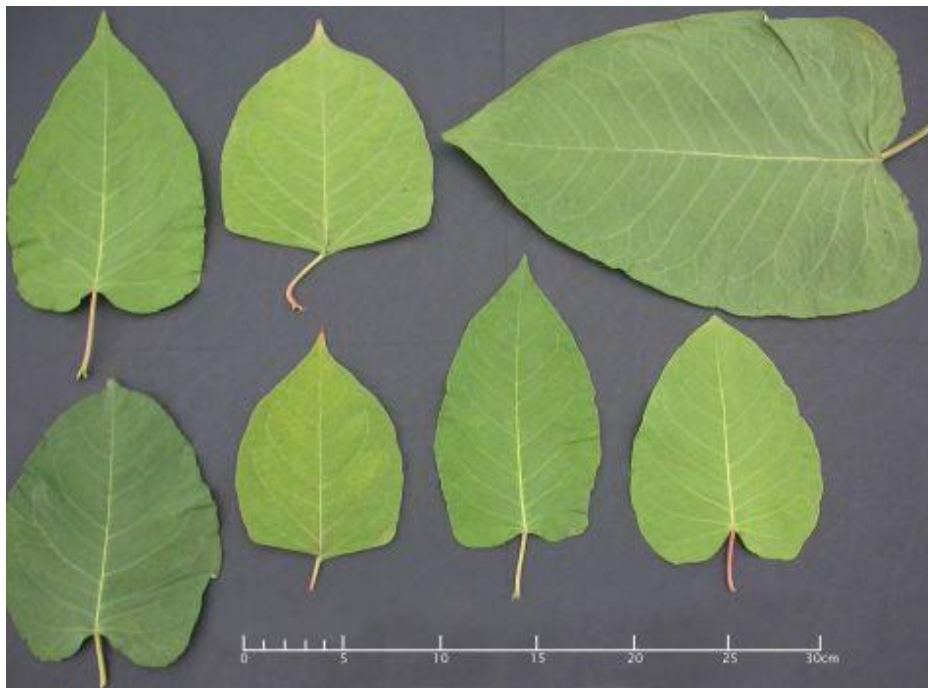
Figure 7.1: Spectrum of morphologically similar leaf shapes of Japanese, Giant, and hybrid Bohemian knotweed varieties.(From Lincoln Soil and Water Conservation District, http://www.lincolnswcd.org/bohemian-knotweed-page.html).

Misidentification impedes management of invasive species (Islam *et al.*, 2022), with 61% of 1285 plant species identified as invasive in the US remain available through the plant trade, including 50% of state-regulated species and 20% of federal noxious weeds (Beaury *et al.*, 2021). This lack of plant-awareness is exacerbated by a decline in botanical education as students are presented with diminutive plant content, compared with animal content particularly with regards to identification (Stroud *et al.*, 2022).

This project has shown that ATR-FTIR spectral absorbance profiles of hybrid plants and different species can be differentiated using chemometrics (Holden *et al.*, 2022). This presents an opportunity for the development of new identification methodologies, particularly important as we move away from traditional botany skills and rely more heavily on apps and technology for problem solving. The availability of hand-held spectrometers (Dhawale *et al.*, 2015; Ji *et al.*, 2016; Soriano-Disla *et al.*, 2018; Bureau *et al.*, 2019; Hutengs *et al.*, 2019) make this a possibility for use in the field. The potential for this methodology extends beyond the Japanese knotweed complex to the control of other invasive species. There are many other examples of morphologically similar invasive hybrids, and look-alike non-invasives which could benefit from an identification program. For example, the invasive aquatic plant, hybrid watermilfoil, could be more auxin-tolerant and faster-spreading than the traditionally invasive Eurasian variety (Thum *et al.*, 2017). The two parent milfoil plants and their hybrid look virtually identical to an untrained eye (see Figure 7.2).
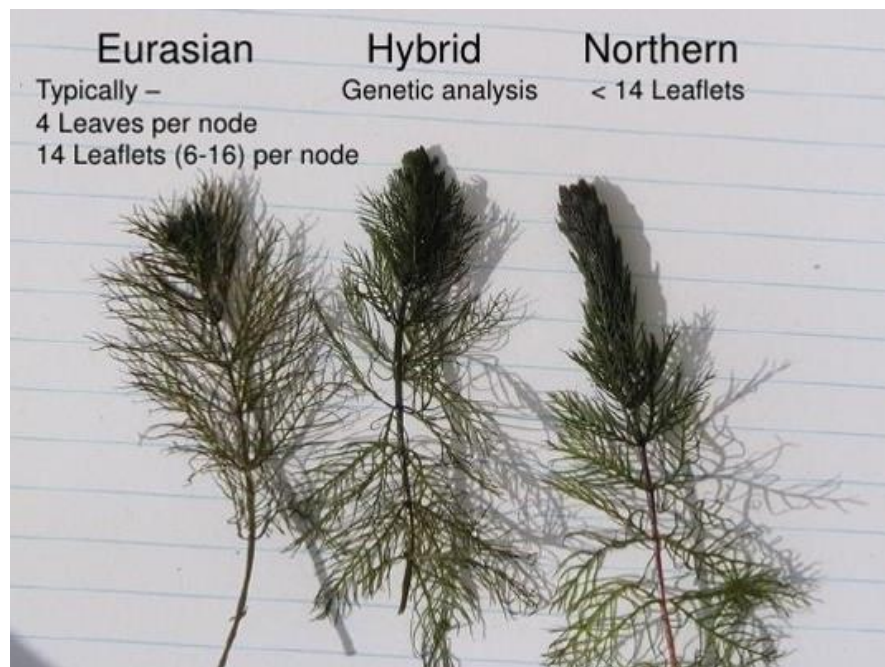
Figure 7.2: Morphologically similar Eurasian, hybrid, and Northern milfoil plants. (From Minnehaha Creek Watershed District, 2013).

This problem also transcends plant biology across other invasive species. Morphologically similar Zebra mussel (*Dreissena polymorpha*) and Quagga mussel (*Dreissena rostriformis bugensis*) populations have a co-occurring invasive range within Canadian freshwater ecosystems and have been hybridised under laboratory conditions (Therriault *et al.*, 2013; see Figure 7.3).
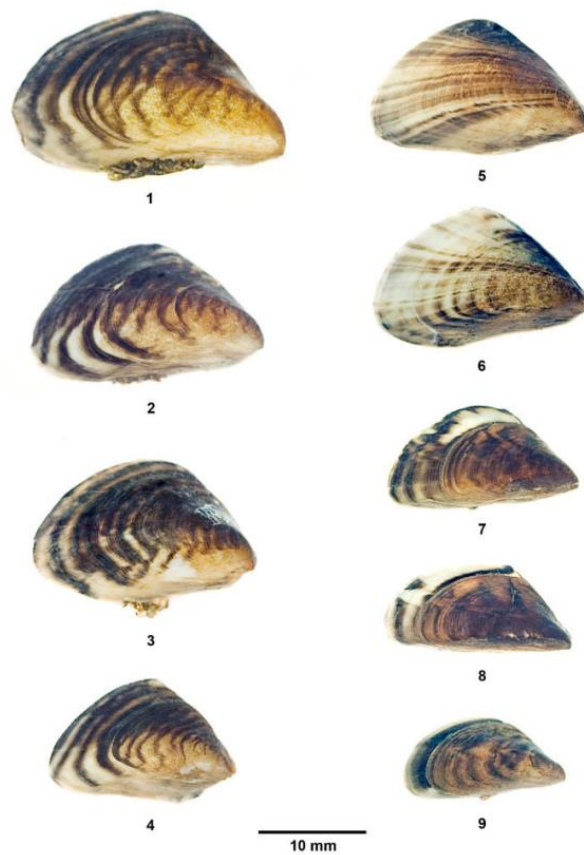
Figure 7.3: Quagga mussels (*Dreissena rostriformis bugensis*) (panels 1-6) and Zebra mussels (*Dreissena polymorpha*) (panels 7-9) collected at the same locality in the River Main, Germany. (From Van Der Velde & Platvoet, 2007).

7.3 Potential for ATR–FTIR spectroscopy in other applications

The large amount of data provided in a rapid and non-destructive manner by spectroscopic analysis lends this method to agricultural applications. This series of studies have successfully predicted plant hormone concentrations (see Chapter 5) and physiological parameters such as root water pressure (see Chapter 7) using MIR spectroscopic absorbances. Previous studies using similar portable near-infrared (NIR) spectrometers have shown the potential to use the predictive capacity of this technology for high-throughput monitoring of plant nutrient status (Pandey *et al.*, 2017) and even in-field, effectively mounting NIR spectrometers onto an all-terrain-vehicle for application in viticulture to predict plant water status and plan irrigation regimes (Diago *et al.*, 2017, 2018; Fernández-Novales *et al.*, 2018). However, research into the applications of mid-infrared spectroscopy have been previously limited due to the increased expense and lack of transportability of MIR spectrometers compared with NIR spectrometers (Soriano-Disla *et al.*, 2018). MIR used in the current research has the advantage over NIR of being more biologically informative, and often more accurate as NIR can suffer from co-linearity problems because the broad absorption bands due to overtones and combination modes overlap, reducing its sensitivity and making molecular assignment more difficult (Shiroma & Rodriguez-Saona, 2009; Bellon-Maurel & McBratney, 2011; de Oliveira et al., 2014). Going forward, this trade-off between cost and accuracy is set to reduce as advances in technology mean that portable MIR spectrometers have become more readily available (Dhawale *et al.*, 2015; Ji *et al.*, 2016; Soriano-Disla *et al.*, 2018; Bureau *et al.*, 2019; Hutengs *et al.*, 2019).

7.4 Practical application of ATR–FTIR spectroscopy

To successfully apply ATR-FTIR spectroscopy to real world problems, the influence of sample preparation and growing environment on the drawn-conclusions must be considered (Morais *et al.*, 2019; Durak & Depciuch, 2020). As shown in Chapter 2, the way in which vegetative tissue sample are prepared has an influence over the

spectral profiles. Caution must therefore be taken to compare like-with-like. This requirement was also recognised when using ATR-FTIR spectrometry to quantitatively predicted concentrations of hormones in leaves and xylem sap, see Chapter 5. Additionally, localised effects of the growing environment such as nearby tree canopies must be considered when interpreting the data. Particularly because the red: far-red ratio affects the spectral fingerprint of vegetative tissue more than levels of water or nutrients. For example, whilst there is a legal demand to prove the source of Japanese knotweed which has crossed a boundary and spread onto property, a nearby tree could introduce a spectral change which overrides other information about the origin of the plant material. Further work is needed to decipher how best to use this technology. Chemometric analysis could be adapted to gain the most information from the acquired spectra, for example Chapter 3 shows that the results of HCA can be steered towards organising sample relationships by either geographical origin or species, by inputting the fingerprint region or the LD1 loadings respectively (Holden *et al.*, 2022).

The process from chemometric biomarker identification to physical biomolecular extraction is a developing area of spectroscopy with an ongoing research effort currently focused on optimising the quantification of biomolecular concentrations with the resultant spectra in biological extracts (Wagner *et al.*, 2010; Spalding *et al.*, 2018), consolidating the expanding database of key wavenumber changes and their associated molecular definitions (Talari *et al.*, 2017), and trialling new biological applications (Skolik *et al.*, 2019a,b; Butler *et al.*, 2020). The continued development of biomarker databases calibrated against empirical data from biomolecule extraction is an important focus. Further method development is required to optimise spectral acquisition, such as the number of technical replicates required for reproducible classification and biomarker identification.

The requirement for prior knowledge of scripting, coding and machine learning can present a barrier to the widespread use of this technology in other fields. This can be overcome using no-code development platforms (NCDPs) with integrated

graphical user interfaces (GUIs) to create custom-built software packages such as IRootLab (Trevisan *et al.*, 2013) and FRUITNIR-GUI (Mishra *et al.*, 2021). These can make spectroscopic analysis more accessible to biological scientists, clinicians, and non-coders whose expertise lie elsewhere, thereby opening this method up to more applications. As new spectrometers are developed, it would be helpful to accompany them with code which can format the generated spectral data into user-friendly formats ready for input into one of these data analysis packages. One caveat to this is that the more of a 'black-box' process the analysis becomes, the more potential for misinterpretation of the biological meaning of the data. As an interface between disciplines, this field is best progressed in collaboration between researchers of different specialties, with maintained communication throughout from initial planning stages to data interpretation. Plant science still lacks a domain-specific repository for sharing infrared spectral data between research groups, see for example 'CLIRSPEC' which exists for clinical applications (Henderson, 2015).

Another area of required development is the practicalities of high-throughput and in-field use. For example, a standard for the best sample preparation so that studies can be compared. Whilst Chapter 2 shows that the homogenisation of freeze-dried ground leaves provides the most consistent results, in planta spectral acquisition remains most practical for in-field use but would require portable spectrometers. Although portable MIR spectrometers do exist (Dhawale *et al.*, 2015; Ji *et al.*, 2016; Soriano-Disla *et al.*, 2018; Bureau *et al.*, 2019; Hutengs *et al.*, 2019) these have mainly been used in soil sciences. Further work is needed to compare the quality of spectral data taken with portable spectrometers to bench-top or lab-based instruments. The proficiency to cost of mid-infrared spectrometry should be compared with similar techniques such as Raman spectroscopy and near-infrared spectroscopy, and this balance may vary dependant on the application.

7.5 Conclusions

The main conclusions from this series of studies were:

- Sample preparation of vegetative tissue affects the ATR-FTIR spectral signature.

- Regional environmental conditions induce differences in Japanese knotweed vegetative tissue which can be detected using ATR-FTIR spectroscopy.

- ATR-FTIR spectroscopy provides a rapid screening method for hybrid identification in the Japanese knotweed complex.

- Of the environmental conditions studied (water, red: far-red light, nitrogen, and micronutrients) the red: far-red light ratio has the most pronounced effect on the ATR-FTIR spectral signature of dried leaf tissue.

- Predictive models can be created using PLSR to non-destructively estimate physiological responses to environmental conditions such as water potential and hormone concentrations from the spectral signatures of plant samples.

- Several key biomarkers for abiotic stress in Japanese knotweed were conserved in other species and could be indicator wavenumbers (953, 1038, 1227, 1709, and 1732 cm^{-1}).

- Within the mid-infrared fingerprint region, spectral biomarkers for prediction of plant hormone concentrations differ for samples from xylem sap (1600-1500 cm^{-1}) and leaf tissue (1200-1000 cm^{-1}).

7.6 Future perspectives

Although mortgage options for home-buyers with Japanese knotweed on their properties are likely to become more lenient in proportion to the risk of structural damage (Fennell *et al.*, 2018; Baird, 2022), it will take time to repair the poor reputation of Japanese knotweed with the public. Global change and human interference favour invasive species, driving poleward expansion and providing increased opportunities for their introduction and spread. Therefore, whilst Japanese knotweed may be perceived less as a property issue in future, it remains

an environmental challenge. As public awareness of the climate crisis increases so does the pressure to find an ecologically friendly, low-carbon solution, with a view to move away from the traditional herbicide, glyphosate. This calls for the use of novel technologies to approach this problem. ATR-FTIR spectroscopy has the potential to aid invasive species control via hybrid identification and targeted treatment plans to conserve resources, as a screening tool for novel herbicides, for the monitoring of target-plant health and prevention of sub-lethal herbicide application. There is ever-expanding evidence for the possible applications of ATR-FTIR spectroscopy to plant science, including monitoring health and development, disease detection, and invasive species identification. Implementation of these in real-world scenarios will require cooperation between physicists, engineers, entrepreneurs, and plant biologists. Through integration with technical advances from other fields, it is expected that detection sensitivity and acquisition speed will increase, and a reduction in both instrument size and the overall cost of the technology can be achieved (Cheng & Xie, 2015). Although still in its developmental stage, the application of bio-spectroscopy to plant and crop science is a rapidly progressing area with largely unexploited opportunities.

8 Bibliography

Abdi H. **2010**. Partial least squares regression and projection on latent structure regression (PLS Regression). *Wiley Interdisciplinary Reviews: Computational Statistics* **2**: 97–106.

Adachi N, Terashima I, Takahashi M. **1996**. Central Die-back of Monoclonal Stands of *Reynoutria japonica* in an Early Stage of Primary Succession on Mount Fuji. *Annals of Botany* **77**: 477–486.

Adams GDJ, Cook I, Ward KR. **2015**. The principles of freeze-drying. *Methods in Molecular Biology* **1257**: 121–143.

Agathokleous E, Kitao M, Calabrese EJ. **2019**. Hormesis: A Compelling Platform for Sophisticated Plant Science. *Trends in Plant Science* **24**: 318–327.

Agatonovic-Kustrin S, Doyle E, Gegechkori V, Morton DW. **2020**. High-performance thin-layer chromatography linked with (bio)assays and FTIR-ATR spectroscopy as a method for discovery and quantification of bioactive components in native Australian plants. *Journal of Pharmaceutical and Biomedical Analysis* **184**: 113208.

Aguilera AG, Alpert P, Dukes JS, Harrington R. **2010**. Impacts of the invasive plant *Fallopia japonica* (Houtt.) on plant communities and ecosystem processes. *Biological Invasions* **12**: 1243–1252.

Ajitha B, Ashok Kumar Reddy Y, Shameer S, Rajesh KM, Suneetha Y, Sreedhara Reddy P. **2015**. Lantana camara leaf extract mediated silver nanoparticles: Antibacterial, green catalyst. *Journal of Photochemistry and Photobiology B: Biology* **149**: 84–92.

Albacete A, Ghanem ME, Martínez-Andújar C, Acosta M, Sánchez-Bravo J, Martínez V, Lutts S, Dodd IC, Pérez-Alfocea F. **2008**. Hormonal changes in relation to biomass partitioning and shoot growth impairment in salinized tomato (Solanum lycopersicum L.) plants. *Journal of experimental botany* **59**: 4119–4131.

Alcázar R, Marco F, Cuevas JC, Patron M, Ferrando A, Carrasco P, Tiburcio AF, Altabella T. **2006**. Involvement of polyamines in plant response to abiotic stress. *Biotechnology Letters* **28**: 1867–1876.

Allen SE, Grimshaw HM, Parkinson JA, Quarmby C. **1989**. Chemical Analysis of Ecological Materials (Second Edition). *The Journal of Applied Ecology* **13**: 368.

Álvarez S, Bañón S, Sánchez-Blanco MJ. **2013**. Regulated deficit irrigation in different phenological stages of potted geranium plants: Water consumption, water relations and ornamental quality. *Acta Physiologiae Plantarum* **35**: 1257–1267.

Amano T, Coverdale R, Peh KSH. **2016**. The importance of globalisation in driving the introduction and establishment of alien species in Europe. *Ecography* **39**: 1118–1128.

Anagnostopoulos CJ, Liapis K, Haroutounian S, Paspatis E. **2013**. Simultaneous determination of different classes of plant growth regulator in high water content agricultural products by liquid chromatography tandem mass spectrometry and time of flight mass spectrometry. *Journal of Liquid Chromatography and Related Technologies* **36**: 315–335.

Anfang M, Shani E. **2021**. Transport mechanisms of plant hormones. *Current Opinion in Plant Biology* **63**: 102055.

Aouidi F, Dupuy N, Artaud J, Roussos S, Msallem M, Perraud-Gaime I, Hamdi M. **2012**. Discrimination of five Tunisian cultivars by Mid InfraRed spectroscopy combined with chemometric analyses of olive Olea europaea leaves. *Food Chemistry* **131**: 360–366.

Arber A. **1986**. Herbals : their origin and evolution : a chapter in the history of botany, 1470-1670. : 358.

Armstrong K., Ball S. 2005. DNA barcodes for biosecurity: invasive species identification. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**: 1813–1823.

Asif A, Baig MA, Siddiqui MB. **2021**. Role of Jasmonates and Salicylates in Plant Allelopathy. In *Jasmonates and Salicylates Signaling in Plants* 115–127. Springer, Cham.

Atkin OK, Tjoelker MG. 2003. Thermal acclimation and the dynamic response of plant respiration to temperature. *Trends in Plant Science* **8**: 343–351.

Babu AK, Kumaresan G, Raj VAA, Velraj R. **2018**. Review of leaf drying: Mechanism and influencing parameters, drying methods, nutrient preservation, and mathematical models. *Renewable and Sustainable Energy Reviews* **90**: 536–556.

Bağcıoğlu M, Kohler A, Seifert S, Kneipp J, Zimmermann B. **2017**. Monitoring of plant–environment interactions by high-throughput FTIR spectroscopy of pollen. *Methods in Ecology and Evolution* **8**: 870–880.

Bailey JP. **2003**. Japanese Knotweed s.l. at home and abroad. *Plant invasions: ecological threats and management solutions*: 183–196.

Bailey JP. **2013**. The Japanese knotweed invasion viewed as a vast unintentional hybridisation experiment. *Heredity* **110(2):**105-10.

Bailey JP, Bímová K, Mandák B. **2009**. Asexual spread versus sexual reproduction and evolution in Japanese Knotweed s.l. sets the stage for the 'battle of the Clones'. *Biological Invasions* **11(5):**1189-203.

Bailey JP, Conolly AP. **2000**. Prize-winners to pariahs -A history of Japanese Knotweed s.l. (Polygonaceae) in the British Isles. *Watsonia* **23**: 93–110.

Baird R. **2022**. Rics scraps Japanese knotweed 7-metre rule. *Mortgage Finance Gazette*.

Bajwa AA, Mahajan G, Chauhan BS. **2015**. Nonconventional Weed Management Strategies for Modern Agriculture. *Weed Science* **63**: 723–747.

Baker MJ, Trevisan J, Bassan P, Bhargava R, Butler HJ, Dorling KM, Fielden PR, Fogarty SW, Fullwood NJ, Heys KA, *et al.* 2014. Using Fourier transform IR spectroscopy to analyze biological materials. *Nature Protocols* **9**: 1771–1791.

Ballabio D. **2015**. A MATLAB toolbox for Principal Component Analysis and unsupervised exploration of data structure. *Chemometrics and Intelligent Laboratory Systems* **149**: 1–9.

Ballabio D, Consonni V. **2013**. Classification tools in chemistry. Part 1: linear models. PLS-DA. *Analytical Methods* **5**: 3790–3798.

Ballaré CL, Pierik R. **2017**. The shade-avoidance syndrome: multiple signals and ecological consequences. *Plant, Cell & Environment* **40**: 2530–2543.

Banavali NK, Roux BB. **2005**. Free Energy Landscape of A-DNA to B-DNA Conversion in Aqueous Solution. *Journal of the American Chemical Society* **127(18)**: 6866-76.

Banerjee AK, Guo W, Huang Y. **2019**. Genetic and epigenetic regulation of phenotypic variation in invasive plants - Linking research trends towards a unified framework. *NeoBiota* **49**: 77–103.

Banta JA, Richards CL. **2018**. Quantitative epigenetics and evolution. *Heredity* **121**: 210–224.

Barney JN, Tharayil N, DiTommaso A, Bhowmik PC. **2006**. The Biology of Invasive Alien Plants in Canada. 5. *Polygonum cuspidatum* Sieb. & Zucc. [=

Fallopia japonica (Houtt.) Ronse Decr.]. *Canadian Journal of Plant Science* **86**: 887–906.

Baronio CM, Barth A. **2020**. The Amide i Spectrum of Proteins - Optimization of Transition Dipole Coupling Parameters Using Density Functional Theory Calculations. *Journal of Physical Chemistry B* **124**: 1703–1714.

Bashtanova UB, Beckett KP, Flowers TJ. **2009**. Review: Physiological Approaches to the Improvement of Chemical Control of Japanese Knotweed (*Fallopia japonica*). *Weed Science* **57**: 584–592.

BBC News. **2015**. Kenneth McRae killed wife Jane and then himself over knotweed fears. *BBC News Birmingham & Black Country*.

Beaury EM, Patrick M, Bradley BA. **2021**. Invaders for sale: the ongoing spread of invasive species by the plant trade industry. *Frontiers in Ecology and the Environment* **19**: 550–556.

Beć KB, Grabska J, Huck CW. **2020**. Near-Infrared Spectroscopy in Bio-Applications. *Molecules 2020, Vol. 25, Page 2948* **25**: 2948.

Beerling DJ, Bailey JP, Conolly AP. **1994**. *Fallopia Japonica* (Houtt.) Ronse Decraene. *The Journal of Ecology* **82**: 959.

Belfer S, Purinson Y, Kedem O. **1998**. Surface modification of commercial polyamide reverse osmosis membranes by radical grafting: An ATR-FTIR study. *Acta Polymerica* **49**: 574–582.

Bellard C, Jeschke JM, Leroy B, Mace GM. **2018**. Insights from modeling studies on how climate change affects invasive alien species geography. *Ecology and Evolution* **8**: 5688–5700.

Bellard C, Thuiller W, Leroy B, Genovesi P, Bakkenes M, Courchamp F. **2013**. Will climate change promote future invasions? *Global Change Biology* **19**: 3740–3748.

Bellon-Maurel V, McBratney A. **2011**. Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils – Critical review and research perspectives. *Soil Biology and Biochemistry* **43**: 1398–1410.

Belz RG, Duke SO. **2014**. Herbicides and plant hormesis. *Pest Management Science* **70**: 698–707.

Ben-Ari G, Lavi U. **2012**. Marker-assisted selection in plant breeding. In: Plant Biotechnology and Agriculture. Elsevier Inc., 163–184.

Bensemmane N, Bouzidi N, Daghbouche Y, Garrigues S, de la Guardia M, El Hattab M. **2021**. Quantification of phenolic acids by partial least squares Fourier-transform infrared (PLS-FTIR) in extracts of medicinal plants. *Phytochemical Analysis* **32**: 206–221.

Berardo N, Pisacane V, Battilani P, Scandolara A, Pietri A, Marocco A. **2005**. Rapid detection of kernel rots and mycotoxins in maize by near-infrared reflectance spectroscopy. *Journal of Agricultural and Food Chemistry* **53**: 8128–8134.

Berman HM. **1991**. Hydration of DNA. *Current Opinion in Structural Biology* **1**: 423–427.

Bharath P, Gahir S, Raghavendra AS. **2021**. Abscisic Acid-Induced Stomatal Closure: An Important Component of Plant Defense Against Abiotic and Biotic Stress. *Frontiers in Plant Science* **12**: 324.

Bielczynski LW, Łącki MK, Hoefnagels I, Gambin A, Croce R. **2017**. Leaf and Plant Age Affects Photosynthetic Performance and Photoprotective Capacity. *Plant Physiology* **175**: 1634.

Bímová K, Mandák B, Pyšek P. **2003**. Experimental study of vegetative regeneration in four invasive Reynoutria taxa (Polygonaceae). *Plant Ecology* **166**: 1–11.

Blázquez MA, Nelson DC, Weijers D. **2020**. Evolution of Plant Hormone Response Pathways. *https://doi.org/10.1146/annurev-arplant-050718-100309* **71**: 327–353.

Bolan NS, Curtin D, Adriano DC. **2004**. Acidity. In: Encyclopedia of Soils in the Environment. Elsevier Inc., 11–17.

Booy O, Wade M, Roy H. **2015**. *Field guide to invasive plants and animals in Britain*. Bloomsbury Publishing.

Bosco R, Daeseleire E, Van Pamel E, Scariot V, Leus L. **2014**. Development of an Ultrahigh-Performance Liquid Chromatography–Electrospray Ionization–Tandem Mass Spectrometry Method for the Simultaneous Determination of Salicylic Acid, Jasmonic Acid, and Abscisic Acid in Rose Leaves. *Journal of Agricultural and Food Chemistry* **62**: 6278–6284.

Bossdorf O, Auge H, Lafuma L, Rogers WE, Siemann E, Prati D. **2005**. Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia* **144**: 1–11.

Bourchier RS, Van Hezewijk BH. **2010**. Distribution and Potential Spread of Japanese Knotweed (*Polygonum cuspidatum*) in Canada Relative to Climatic Thresholds. *Invasive Plant Science and Management* **3**: 32–39.

Boyer M, Brasier W. **2019**. Traitement mécanique de volumes importants de terres infestées par des rhizomes de renouée du Japon: technique par criblage-concassage. *Sciences Eaux Territoires*: 68–73.

Bradshaw CJA, Leroy B, Bellard C, Roiz D, Albert C, Fournier A, Barbet-Massin M, Salles JM, Simard F, Courchamp F. **2016**. Massive yet grossly underestimated global costs of invasive insects. *Nature Communications* **7**: 1–8.

Bram MR, Mcnair JN, Patrick CA. Seed Germinability and Its Seasonal Onset of Japanese Knotweed (*Polygonum cuspidatum*). *Source: Weed Science Weed Science* **52**.

Brereton RG. **2007**. *Applied chemometrics for scientists*. John Wiley & Sons.

Bridelli MG. **2017**. Fourier Transform Infrared Spectroscopy in the Study of Hydrated Biological Macromolecules. In: Fourier Transforms - High-tech Application and Current Trends. InTech.

Bro R, Smilde AK. **2014**. Principal component analysis. *Analytical methods* **6**: 2812–2831.

Broadbent A, Stevens CJ, Peltzer DA, Ostle NJ, Orwin KH. **2018**. Belowground competition drives invasive plant impact on native species regardless of nitrogen availability. *Oecologia* **186**: 577–587.

Brown CS, Anderson VJ, Claassen VP, Stannard ME, Wilson LM, Atkinson SY, Bromberg JE, Grant TA, Munis MD. **2008**. Restoration Ecology and Invasive Plants in the Semiarid West. *Invasive Plant Science and Management* **1**: 399–413.

Van Bruggen AHC, He MM, Shin K, Mai V, Jeong KC, Finckh MR, Morris JG. **2018**. Environmental and health effects of the herbicide glyphosate. *Science of The Total Environment* **616–617**: 255–268.

Brunerová A, Brožek M. **2017**. Potential of wild growing Japanese knotweed (*Reynoutria japonica*) for briquette production Development and verification of manual low-and medium-pressure briquetting presses (with particular use in rural areas of Southeast Asia) View project Green Energy View project.

Buitrago MF, Groen TA, Hecker CA, Skidmore AK. **2016**. Changes in thermal infrared spectra of plants caused by temperature and water stress. *ISPRS Journal of Photogrammetry and Remote Sensing* **111**: 22–31.

Bureau S, Cozzolino D, Clark CJ. **2019**. Contributions of Fourier-transform mid infrared (FT-MIR) spectroscopy to the study of fruit and vegetables: A review. *Postharvest Biology and Technology* **148**: 1–14.

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

Butler HJ, Martin FL, Roberts MR, Adams S, McAinsh MR. **2020**. Observation of nutrient uptake at the adaxial surface of leaves of tomato (*Solanum lycopersicum*) using Raman spectroscopy. *Analytical Letters* **53**: 536–562.

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

Butler HJ, Smith BR, Fritzsch R, Radhakrishnan P, Palmer DS, Baker MJ. **2018**. Optimised spectral pre-processing for discrimination of biofluids via ATR-FTIR spectroscopy. *Analyst* **143**: 6121–6134.

Bzdega K, Janiak A, Tarłowska S, Kurowska M, Tokarska-Guzik B, Szarejko I. **2012**. Unexpected genetic diversity of *Fallopia japonica* from Central Europe revealed after AFLP analysis. *Flora: Morphology, Distribution, Functional Ecology of Plants* **207(9):** 636-645.

Cao Z, Wang Z, Shang Z, Zhao J. **2017**. Classification and identification of Rhodobryum roseum limpr. And its adulterants based on fourier-transform infrared spectroscopy (FTIR) and chemometrics. *PLoS ONE* **12(2):** e0172359.

Cappuccino N, Arnason JT. **2006**. Novel chemistry of invasive exotic plants. *Biology Letters* **2**: 189–193.

Carpentier SC, Dens K, Van den houwe I, Swennen R, Panis B. **2007**. Lyophilization, a Practical Way to Store and Transport Tissues Prior to Protein Extraction for 2DE Analysis? *PROTEOMICS* **7**: 64–69.

Cedergreen N, Hansen NKK, Arentoft BW. **2016**. The influence of nitrogen and phosphorous status on glyphosate hormesis in Lemna minor and Hordeum vulgare. *European Journal of Agronomy* **73**: 107–117.

Centre for Agriculture and Bioscience International (CABI). **2018**. Data Sheet: *Fallopia japonica* (Japanese knotweed).

Chan KLA, Kazarian SG. **2006**. ATR-FTIR spectroscopic imaging with expanded field of view to study formulations and dissolution. *Lab on a Chip* **6**: 864–870.

Chan KLA, Kazarian SG. **2016**. Attenuated total reflection Fourier-transform infrared (ATR-FTIR) imaging of tissues and live cells. *Chemical Society Reviews* **45**: 1850–1864.

Chen J, Guo B, Yan R, Sun S, Zhou Q. **2017**. Rapid and automatic chemical identification of the medicinal flower buds of Lonicera plants by the benchtop and hand-held Fourier transform infrared spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **182**: 81–86.

Cheng F, Cheng Z. **2015**. Research Progress on the use of Plant Allelopathy in Agriculture and the Physiological and Ecological Mechanisms of Allelopathy. *Frontiers in Plant Science* **6**: 1020.

Cheng JX, Xie XS. **2015**. Vibrational spectroscopic imaging of living systems: An emerging platform for biology and medicine. *Science* **350**.

Chiang C, Olsen JE, Basler D, Bånkestad D, Hoch G. **2019**. Latitude and weather influences on sun light quality and the relationship to tree growth. *Forests* **10(8)**: 610.

Chinchio E, Crotta M, Romeo C, Drewe JA, Guitian J, Ferrari N. **2020**. Invasive alien species and disease risk: An open challenge in public and animal health (LJ Knoll, Ed.). *PLOS Pathogens* **16**: e1008922.

Ciriminna R, Fidalgo A, Ilharco LM, Pagliaro M. **2019**. Herbicides based on pelargonic acid: Herbicides of the bioeconomy. *Biofuels, Bioproducts and Biorefining* **13**: 1476–1482.

Cirino de Carvalho L, de Lelis Medeiros de Morais C, Gomes de Lima KM, Cunha Júnior LC, Martins Nascimento PA, Bosco de Faria J, Henrique de Almeida Teixeira G. **2016**. Determination of the geographical origin and ethanol content of Brazilian sugarcane spirit using near-infrared spectroscopy coupled with discriminant analysis. *Analytical Methods* **8**: 5658–5666.

Clements DR, DiTommaso A. **2012**. Predicting weed invasion in Canada under climate change: Evaluating evolutionary potential. *Canadian Journal of Plant Science* **92**: 1013–1020.

Clements DR, Larsen T, Grenz J. **2016**. Knotweed Management Strategies in North America with the Advent of Widespread Hybrid Bohemian Knotweed, Regional Differences, and the Potential for Biocontrol Via the Psyllid Aphalara itadori Shinji. *Invasive Plant Science and Management* **9**: 60–70.

Coleman GRY, Stead A, Rigter MP, Xu Z, Johnson D, Brooker GM, Sukkarieh S, Walsh MJ. **2019**. Using energy requirements to compare the suitability of alternative methods for broadcast and site-specific weed control. *Weed Technology* **33**: 633–650.

Conservation Dogs Ireland. **2022**. Dogs Saving Wildlife. *Conservation Dogs Ireland webpage- https://conservationdogsireland.ie/*.

Cortes C, Vapnik V. **1995**. Support-vector networks. *Machine learning* **20**: 273–297.

da Costa VA, Tom Cothren J. **2011**. Drought Effects on Gas Exchange, Chlorophyll, and Plant Growth of 1-Methylcyclopropene Treated Cotton. *Agronomy Journal* **103**: 1230–1241.

Courbier S, Grevink S, Sluijs E, Bonhomme P-O, Kajala K, Wees SCM Van, Pierik R. **2020**. Far-red light promotes Botrytis cinerea disease development in tomato leaves via jasmonate-dependent modulation of soluble sugars. *Plant, Cell & Environment* **43**: 2769–2781.

Dassonville N, Vanderhoeven S, Gruber W, Meerts P. **2007**. Invasion by *Fallopia japonica* increases topsoil mineral nutrient concentrations. *Écoscience* **14**: 230–240.

Davidson AM, Jennions M, Nicotra AB. **2011**. Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. *Ecology Letters* **14**: 419–431.

Davies PJ. **2010**. The Plant Hormones: Their Nature, Occurrence, and Functions. *Plant Hormones: Biosynthesis, Signal Transduction, Action!*: 1–15.

Davis MA, Grime JP, Thompson K. **2000**. Fluctuating resources in plant communities: A general theory of invasibility. *Journal of Ecology* **88**: 528–534.

Dayan FE. **2019**. Current Status and Future Prospects in Herbicide Discovery. *Plants 2019, Vol. 8, Page 341* **8**: 341.

Dejonghe W, Russinova E. **2017**. Plant Chemical Genetics: From Phenotype-Based Screens to Synthetic Biology. *Plant Physiology* **174**: 5–20.

Delabre I, Rodriguez LO, Smallwood JM, Scharlemann JPW, Alcamo J, Antonarakis AS, Rowhani P, Hazell RJ, Aksnes DL, Balvanera P, *et al.* **2021**. Actions on sustainable food production and consumption for the post-2020 global biodiversity framework. *Science Advances* **7**: 8259.

Demir P, Onde S, Severcan F. **2015**. Phylogeny of cultivated and wild wheat species using ATR-FTIR spectroscopy. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy* **135**: 757–763.

Desjardins SD. **2015**. Evolutionary studies in subtribe Reynoutriineae (Polygonaceae). With contributions to the study of hybridisation in Helosciadium and Berula (Apiaceae) included as appendices. *Doctoral dissertation, University of Leicester.*

Dhawale NM, Adamchuk VI, Prasher SO, Viscarra Rossel RA, Ismail AA, Kaur J. **2015**. Proximal soil sensing of soil texture and organic matter with a prototype portable mid-infrared spectrometer. *European Journal of Soil Science* **66**: 661–669.

Diago MP, Bellincontro A, Scheidweiler M, Tardaguila J, Tittmann S, Stoll M. **2017**. Future opportunities of proximal near infrared spectroscopy approaches to determine the variability of vineyard water status. *Australian Journal of Grape and Wine Research* **23**: 409–414.

Diago MP, Fernández-Novales J, Gutiérrez S, Marañón M, Tardaguila J. **2018**. Development and validation of a new methodology to assess the vineyard water status by on-the-go near infrared spectroscopy. *Frontiers in Plant Science* **9**: 59.

Diez JM, D'Antonio CM, Dukes JS, Grosholz ED, Olden JD, Sorte CJB, Blumenthal DM, Bradley BA, Early R, Ibáñez I, *et al.* 2012. Will extreme climatic events facilitate biological invasions? *Frontiers in Ecology and the Environment* **10**: 249–257.

Dodd IC, Egea G, Davies WJ. **2008**. Abscisic acid signalling when soil moisture is heterogeneous: decreased photoperiod sap flow from drying roots limits abscisic acid export to the shoots. *Plant, cell & environment* **31**: 1263–1274.

Dommanget F, Evette A, Spiegelberger T, Gallet C, Pacé M, Imbert M, Navas M-L. **2014**. Differential allelopathic effects of Japanese knotweed on willow and cottonwood cuttings used in riverbank restoration techniques. *Journal of Environmental Management* **132**: 71–78.

Dommanget F, Spiegelberger T, Cavaillé P, Evette A. **2013**. Light availability prevails over soil fertility and structure in the performance of Asian knotweeds on riverbanks: New management perspectives. *Environmental Management* **52**: 1453–1462.

Du C, Ma Z, Zhou J, Goyne KW. **2013**. Application of mid-infrared photoacoustic spectroscopy in monitoring carbonate content in soils. *Sensors and Actuators B: Chemical* **188**: 1167–1175.

Dubis EN, Dubis AT, Popławski J. **2001**. Determination of the aromatic compounds in plant cuticular waxes using FT-IR spectroscopy. *Journal of Molecular Structure* **596**: 83–88.

Duff T, Porteous MG, Goodger JQD. **2022**. Wildfires, Flammable Oils, and Eucalyptus Trees: The Persistence and Volatility of Terpenes in Excised Leaves. *SSRN Electronic Journal*. 4120325.

Durak T, Depciuch J. **2020**. Effect of plant sample preparation and measuring methods on ATR-FTIR spectra results. *Environmental and Experimental Botany* **169**: 103915.

Durand M, Murchie EH, Lindfors A V., Urban O, Aphalo PJ, Robson TM. **2021**. Diffuse solar radiation and canopy photosynthesis in a changing environment. *Agricultural and Forest Meteorology* **311**: 108684.

Dusz MA, Martin FM, Dommanget F, Petit A, Dechaume-Moncharmont C, Evette A. **2021**. Review of Existing Knowledge and Practices of Tarping for the Control of Invasive Knotweeds. *Plants 2021, Vol. 10, Page 2152* **10**: 2152.

Ellstrand NC, Schierenbeck KA. **2000**. Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences of the United States of America* **97**: 7043–7050.

ElMasry G, Wang N, ElSayed A, Ngadi M. **2007**. Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry. *Journal of Food Engineering* **81**: 98–107.

Emmett B, Frogbrook Z, Chamberlain P, Giffiths R, Pickup R, Poskitt J, Reynolds B, Rowe E, Spurgeon D, Rowland P, *et al.* 2008. CS Technical Report No 3/07 - Soils Manual Volume 1. *Centre for Ecology and Hydrology, Wallingford, UK (2008)*.

Emsley NEM, Holden CA, Guo S, Bevan RS, Rees C, McAinsh MR, Martin FL, Morais CLM. **2022**. Machine Learning Approach Using a Handheld Near-Infrared (NIR) Device to Predict the Effect of Storage Conditions on Tomato Biomarkers. *ACS Food Science & Technology* **2**: 187–194.

Van den Ende W. **2013**. Multifunctional fructans and raffinose family oligosaccharides. *Frontiers in Plant Science* **4**: 247.

Engler J, Abt K, Buhk C. **2011**. Seed characteristics and germination limitations in the highly invasive *Fallopia japonica s.l.* (Polygonaceae). *Ecological Research* **26**: 555–562.

Environment Agency. **2013**. *Managing Japanese knotweed on development sites: the knotweed code of practice.*

Environment Agency. **2017**. Agreement to use herbicides in or near water Agreement to use herbicides in or near water Guidance notes. *Guidance notes AqHerb01*: 1–3.

Eriksson BK, Hillebrand H. **2019**. Rapid reorganization of global biodiversity. *Science* **366**: 308–309.

Estoup A, Ravigné V, Hufbauer R, Vitalis R, Gautier M, Facon B. **2016**. Is There a Genetic Paradox of Biological Invasion? *Annual Review of Ecology, Evolution, and Systematics* **47**: 51–72.

Euring D, Löfke C, Teichmann T, Polle A. **2012**. Nitrogen fertilization has differential effects on N allocation and lignin in two Populus species with contrasting ecology. *Trees - Structure and Function* **26**: 1933–1942.

Falcão L, Araújo MEM. **2013**. Tannins characterization in historic leathers by complementary analytical techniques ATR-FTIR, UV-Vis and chemical tests. *Journal of Cultural Heritage* **14**: 499–508.

Fan P, Hay A, Marston A, Hostettmann K. **2009**. Allelopathic potential of phenolic constituents from *Polygonum cuspidatum* Sieb. & Zucc (Polygonaceae). *Planta Medica* **75**: 09.

Farber C, Mahnke M, Sanchez L, Kurouski D. **2019**. Advanced spectroscopic techniques for plant disease diagnostics. A review. *TrAC Trends in Analytical Chemistry* **118**: 43–49.

Fennell M, Wade M, Bacon KL. **2018**. Japanese knotweed (*Fallopia japonica*): an analysis of capacity to cause structural damage (compared to other plants) and typical rhizome extension. *PeerJ* **6**: e5246.

Fernández-Novales J, Tardaguila J, Gutiérrez S, Marañón M, Diago MP. **2018**. In field quantification and discrimination of different vineyard water regimes by on-the-go NIR spectroscopy. *Biosystems Engineering* **165**: 47–58.

Festa G, Andreani C, Baldoni M, Cipollari V, Martínez-Labarga C, Martini F, Rickards O, Rolfo MF, Sarti L, Volante N, *et al.* **2019**. First analysis of ancient burned human skeletal remains probed by neutron and optical vibrational spectroscopy. *Science Advances* **5**.

Fischer C, Speth V, Fleig-Eberenz S, Neuhaus G. **1997**. Induction of Zygotic Polyembryos in Wheat: Influence of Auxin Polar Transport. *The Plant Cell* **9**: 1767–1780.

Fogarty AC, Laage D. **2014**. Water dynamics in protein hydration shells: The molecular origins of the dynamical perturbation. *Journal of Physical Chemistry B* **118**: 7715–7729.

Forman J, Kesseli RV. **2003**. Sexual Reproduction in the Invasive Species Fallopia japonica (Polygonaceae). *American Journal of Botany American Journal of Botany* **90**: 586–592.

Fowler D, Coyle M, Skiba U, Sutton MA, Cape JN, Reis S, Sheppard LJ, Jenkins A, Grizzetti B, Galloway JN, *et al.* **2013**. The global nitrogen cycle in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368(1621):** 20130164.

Fraser EDG. **2020**. The challenge of feeding a diverse and growing population. *Physiology & Behavior* **221**: 112908.

Gandolfo DS, Mortimer H, Woodhall JW, Boonham N. **2016**. Fourier transform infra-red spectroscopy using an attenuated total reflection probe to distinguish between Japanese larch, pine and citrus plants in healthy and diseased states. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **163**: 181–188.

Gant RE, Clebsch EEC. **1975**. The Allelopathic Influences of Sassafras albidum in Old-field Succession in Tennessee. *Ecology* **56**: 604–615.

Gaskin JF, Schwarzländer M, Grevstad FS, Haverhals MA, Bourchier RS, Miller TW. **2014**. Extreme differences in population structure and genetic diversity for three invasive congeners: knotweeds in western North America. *Biological Invasions* **16**: 2127–2136.

Ge L, Peh CYC, Yong JWH, Tan SN, Hua L, Ong ES. **2007**. Analyses of gibberellins by capillary electrophoresis–mass spectrometry combined with solid-phase extraction. *Journal of Chromatography A* **1159**: 242–249.

Gendrin C, Roggo Y, Collet C. **2008**. Pharmaceutical applications of vibrational chemical imaging and chemometrics: A review. *Journal of Pharmaceutical and Biomedical Analysis* **48**: 533–553.

Geng YP, van Klinken RD, Sosa A, Li B, Chen J, Xu C-Y. **2016**. The Relative Importance of Genetic Diversity and Phenotypic Plasticity in Determining Invasion Success of a Clonal Weed in the USA and China. *Frontiers in Plant Science* **7**: 216.

Geng YP, Pan XY, Xu CY, Zhang WJ, Li B, Chen JK, Lu BR, Song ZP. **2007**. Phenotypic plasticity rather than locally adapted ecotypes allows the invasive alligator weed to colonize a wide range of habitats. *Biological Invasions* **9**: 245–256.

Gerber E, Krebs C, Murrell C, Moretti M, Rocklin R, Schaffner U. **2008**. Exotic invasive knotweeds (Fallopia spp.) negatively affect native plant and invertebrate assemblages in European riparian habitats. *Biological Conservation* **141**: 646–654.

Gfeller A, Dubugnon L, Liechti R, Farmer EE. **2010**. Jasmonate biochemical pathway. *Science Signaling* **3(109)**: cm3.

Di Giambattista L, Pozzi D, Grimaldi P, Gaudenzi S, Morrone S, Castellano AC. **2011**. New marker of tumor cell death revealed by ATR-FTIR spectroscopy. *Analytical and Bioanalytical Chemistry* **399**: 2771–2778.

Gillies S, Clements DR, Grenz J. **2016**. Knotweed (Fallopia spp.) Invasion of North America Utilizes Hybridization, Epigenetics, Seed Dispersal (Unexpectedly), and an Arsenal of Physiological Tactics. *Invasive Plant Science and Management* **9**: 71–80.

Gniwotta F, Vogg G, Gartmann V, Carver TLW, Riederer M, Jetter R. **2005**. What Do Microbes Encounter at the Plant Surface? Chemical Composition of Pea Leaf Cuticular Waxes. *Plant Physiology* **139(1)**:519-30.

Godfray HCJ, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, Pretty J, Robinson S, Thomas SM, Toulmin C. **2010**. Food security: The challenge of feeding 9 billion people. *Science* **327**: 812–818.

Gomes MP, Le Manac'h SG, Hénault-Ethier L, Labrecque M, Lucotte M, Juneau P. **2017**. Glyphosate-Dependent Inhibition of Photosynthesis in Willow. *Frontiers in Plant Science* **8**: 207.

Gordon R, Chapman J, Power A, Chandra S, Roberts J, Cozzolino D. **2019**. Mid-infrared spectroscopy coupled with chemometrics to identify spectral variability in Australian barley samples from different production regions. *Journal of Cereal Science* **85**: 41–47.

Gorzsas A. **2020**. ATR-FTIR Microspectroscopy Brings a Novel Insight Into the Study of Cell Wall Chemistry at the Cellular Level. In: Proceedings of IPSC 2019-2nd International Plant Spectroscopy Conference. Frontiers Media SA.

Götz A, Nikzad-Langerodi R, Staedler Y, Bellaire A, Saukel J. **2020**. Apparent penetration depth in attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy of Allium cepa L. epidermis and cuticle. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **224**: 117460.

Grevstad FS, Andreas JE, Bourchier RS, Shaw R, Winston RL, Randall CB. **2018**. Biology and biological control of knotweeds. *United States Department of Agriculture, Forest Health Assessment and Applied Sciences Team.*

Grevstad F, Shaw R, Bourchier R, Sanguankeo P, Cortat G, Reardon RC. **2013**. Efficacy and host specificity compared between two populations of the psyllid Aphalara itadori, candidates for biological control of invasive knotweeds in North America. *Biological Control* **65(1):** 53-62.

Grime JP, Pierce S. **2012**. *The evolutionary strategies that shape ecosystems*. John Wiley & Sons.

Grimsby JL, Tsirelson D, Gammon MA, Kesseli R. **2007**. Genetic diversity and clonal vs. sexual reproduction in Fallopia spp. (Polygonaceae). *American Journal of Botany* **94**: 957–964.

Groãÿkinsky DK, Albacete A, Jammer A, Krbez P, Van der Graaff E, Pfeifhofer H, Roitsch T. **2014**. A Rapid Phytohormone and Phytoalexin Screening Method for Physiological Phenotyping. *Molecular Plant* **7**: 1053–1056.

Groeneveld E, Belzile F, Lavoie C. **2014**. Sexual reproduction of Japanese knotweed (*Fallopia japonica s.l.*) at its northern distribution limit: new evidence of the effect of climate warming on an invasive species. *American journal of botany* **101**: 459–66.

Grossmann K. **2003**. Mediation of Herbicide Effects by Hormone Interactions. *Journal of Plant Growth Regulation* **22**: 109–122.

Gudi G, Krähmer A, Krüger H, Schulz H. **2015**. Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy on Intact Dried Leaves of Sage (Salvia officinalis L.): Accelerated Chemotaxonomic Discrimination and Analysis of Essential Oil Composition. *Journal of Agricultural and Food Chemistry* **63**: 8743–8750.

Gurevitch J, Fox GA, Wardle GM, Inderjit, Taub D. **2011**. Emergent insights from the synthesis of conceptual frameworks for biological invasions. *Ecology Letters* **14**: 407–418.

Hagenblad J, Hülskötter J, Acharya KP, Brunet J, Chabrerie O, Cousins SAO, Dar PA, Diekmann M, De Frenne P, Hermy M, *et al.* 2015. Low genetic diversity despite multiple introductions of the invasive plant species Impatiens glandulifera in Europe. *BMC Genetics* **16**: 103.

Hammer B, Gersmann K. **2003**. A note on the universal approximation capability of support vector machines. *Neural Processing Letters* **17**: 43–53.

Head L. **2017**. The social dimensions of invasive plants. *Nature Plants* **3**: 1–7.

Heap, B, Holden, CA, Taylor, JE and McAinsh, M 2020. Crosstalk in Signalling Pathways. In *eLS*. Wiley. 1–9.

Hejda M, Pyšek P, Jarošík V. **2009**. Impact of invasive plants on the species richness, diversity and composition of invaded communities. *Journal of Ecology* **97**: 393–403.

Henderson A. **2015**. CLIRSPEC: Clinical Infrared and Raman Spectroscopy for Medical Diagnosis. *https://zenodo.org/communities/clirspec/*.

Heredia-Guerrero JA, Benítez JJ, Domínguez E, Bayer IS, Cingolani R, Athanassiou A, Heredia A. **2014**. Infrared and Raman spectroscopic features of plant cuticles: a review. *Frontiers in plant science* **5**: 305.

Herman JJ, Sultan SE. **2011**. Adaptive transgenerational plasticity in plants: Case studies, mechanisms, and implications for natural populations. *Frontiers in Plant Science* **2**: 102.

Herrera CM, Bazaga P. **2010**. Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet Viola cazorlensis. *New Phytologist* **187**: 867–876.

Hirose T, Tateno M. **1984**. Soil nitrogen patterns induced by colonization of *Polygonum cuspidatum* on Mt. Fuji. *Oecologia* **61**: 218–223.

Hoad SP, Leakey RRB. **1994**. Effects of light quality on gas exchange and dry matter partitioning in Eucalyptus grandis W. Hill ex Maiden. *Forest Ecology and Management* **70**: 265–273.

Holalu S V., Reddy SK, Finlayson SA. **2021**. Low Red Light:Far Red Light Inhibits Branching by Promoting Auxin Signaling. *Journal of Plant Growth Regulation* **40**: 2028–2036.

Holden CA, Bailey JP, Taylor JE, Martin F, Beckett P, McAinsh M. **2022**. Know your enemy: Application of ATR-FTIR spectroscopy to invasive species control (D Changwen, Ed.). *PLOS ONE* **17**: e0261742.

Holden CA, Morais CLM, Taylor JE, Martin FL, Beckett P, McAinsh M. **2021**. Regional differences in clonal Japanese knotweed revealed by chemometrics-linked attenuated total reflection Fourier-transform infrared spectroscopy. *BMC Plant Biology 2021 21:1* **21**: 1–20.

Hollingsworth ML, Bailey JP. **2000a**. Evidence for massive clonal growth in the invasive weed *Fallopia japonica* (Japanese Knotweed). *Botanical Journal of the Linnean Society* **133**: 463–472.

Hollingsworth ML, Bailey JP. **2000b**. Hybridisation and clonal diversity in some introduced Fallopia species (Polygonaceae). *Watsonia* **23**: 111–121.

Hollingsworth ML, Hollingsworth PM, Jenkins GI, Bailey JP, Ferris C. **1998**. The use of molecular markers to study patterns of genotypic diversity in some invasive alien Fallopia spp. (Polygonaceae). *Molecular Ecology* **7**: 1681–1691.

Holm AK, Elameen A, Oliver BW, Brandsæter LO, Fløistad IS, Brurberg MB. **2018**. Low genetic variation of invasive Fallopia spp. in their northernmost European distribution range. *Ecology and Evolution* **8**: 755–764.

Hoogsteen MJJ, Lantinga EA, Bakker EJ, Groot JCJ, Tittonell PA. **2015**. Estimating soil organic carbon through loss on ignition: effects of ignition conditions and structural water loss. *European Journal of Soil Science* **66**: 320–328.

Hou Q, Ufer G, Bartels D. **2016**. Lipid signalling in plant responses to abiotic stress. *Plant, Cell & Environment* **39**: 1029–1048.

Hou M, Wu D, Li Y, Tao W, Chao L, Zhang Y. **2021**. The role of auxin in nitrogen-modulated shoot branching. *Plant Signaling & Behavior* **16(4):**1885888.

Houhou R, Bocklitz T. **2021**. Trends in artificial intelligence, machine learning, and chemometrics applied to chemical data. *Analytical Science Advances* **2**: 128–141.

House of Commons Select Committee on Science and Technology. **2019**. Section 4: Information and guidance to support lending decisions. In: Japanese knotweed and the built environment. *Science and Technology Select Committe Report*.

Huang Y, Lu R, Chen K. **2018**. Prediction of firmness parameters of tomatoes by portable visible and near-infrared spectroscopy. *Journal of Food Engineering* **222**: 185–198.

Hulme PE. **2009**. Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology* **46**: 10–18.

Hulme PE. **2020**. Plant invasions in New Zealand: global lessons in prevention, eradication and control. *Biological Invasions* **22**: 1539–1562.

Hussain S, Iqbal N, Rahman T, Liu T, Brestic M, Safdar ME, Asghar MA, Farooq MU, Shafiq I, Ali A, *et al.* 2019. Shade effect on carbohydrates dynamics and stem strength of soybean genotypes. *Environmental and Experimental Botany* **162**: 374–382.

Hutengs C, Seidel M, Oertel F, Ludwig B, Vohland M. **2019**. In situ and laboratory soil spectroscopy with portable visible-to-near-infrared and mid-infrared instruments for the assessment of organic carbon in soils. *Geoderma* **355**: 113900.

Inamura A, Ohashi Y, Sato E, Yoda Y, Masuzawa T, Ito M, Yoshinaga K. **2000**. lntraspecific Sequence Variation of Chloroplast DNA Reflecting Variety and Geographical Distribution of *Polygonurn cuspidatum* (Polygonaceae) in Japan. *Journal of Plant Research* **113(4)**: 419-26.

Islam T, Gulzar R, Singh G, Nawchoo IA, Khuroo AA. **2022**. Misidentification impedes invasion management: report of Matricaria discoidea DC., an invasive alien species in Kashmir Himalaya. *Vegetos* **35**: 551–557.

Izquierdo L, Martín I, Albornos L, Hernández-Nistal J, Hueso P, Dopico B, Labrador E. **2018**. Overexpression of Cicer arietinum βIII-Gal but not βIV-Gal in arabidopsis causes a reduction of cell wall β-(1,4)-galactan compensated by an increase in homogalacturonan. *Journal of Plant Physiology* **231**: 135–146.

Jeger M, Beresford R, Bock C, Brown N, Fox A, Newton A, Vicent A, Xu X, Yuen J. **2021**. Global challenges facing plant pathology: multidisciplinary approaches to meet the food security and environmental challenges in the mid-twenty-first century. *CABI Agriculture and Bioscience 2021 2:1* **2**: 1–18.

Jha UC, Nayyar H, Jha R, Khurshid M, Zhou M, Mantri N, Siddique KHM. **2020**. Long non-coding RNAs: emerging players regulating plant abiotic stress response and adaptation. *BMC Plant Biology 2020 20:1* **20**: 1–20.

Ji W, Adamchuk VI, Biswas A, Dhawale NM, Sudarsan B, Zhang Y, Viscarra Rossel RA, Shi Z. **2016**. Assessment of soil properties in situ using a prototype portable MIR spectrometer in two agricultural fields. *Biosystems Engineering* **152**: 14–27.

Jin N, Semple KT, Jiang L, Luo C, Zhang D, Martin FL. **2018**. Spectrochemical analyses of growth phase-related bacterial responses to low (environmentally-relevant) concentrations of tetracycline and nanoparticulate silver. *Analyst* **143**: 768–776.

Johnson KL, Gidley MJ, Bacic A, Doblin MS. **2018**. Cell wall biomechanics: a tractable challenge in manipulating plant cell walls 'fit for purpose'! *Current Opinion in Biotechnology* **49**: 163–171.

Jones D, Bruce G, Fowler MS, Law-Cooper R, Graham I, Abel A, Street-Perrott FA, Eastwood D. **2018**. Optimising physiochemical control of invasive Japanese knotweed. *Biological Invasions* **20**: 2091–2105.

Jones D, Fowler MS, Hocking S, Eastwood D. **2020a**. Please don't mow the Japanese knotweed! *NeoBiota* **60**: 19–23.

Jones IM, Smith SM, Bourchier RS. **2020b**. Establishment of the biological control agent Aphalara itadori is limited by native predators and foliage age. *Journal of Applied Entomology* **144**: 710–718.

Jordan MI, Mitchell TM. **2015**. Machine learning: Trends, perspectives, and prospects. *Science* **349**: 255–260.

Josephs EB, Van Etten ML, Harkess A, Platts A, Baucom RS. **2021**. Adaptive and maladaptive expression plasticity underlying herbicide resistance in an agricultural weed. *Evolution letters* **5**: 432–440.

Kabat TJ, Stewart G, Pullin A. **2006**. Are Japanese Knotweed (*Fallopia japonica*) control and eradication interventions effective? Systematic Review n 21. *Center for Evidence-Based Conservation, Birmingham, UK*.

Kacuráková M, Capek P, Sasinková V, Wellner N, Ebringerová A. **2000**. FT-IR study of plant cell wall model compounds: Pectic polysaccharides and hemicelluloses. *Carbohydrate Polymers* **43**: 195–203.

Kang S, Amarasiriwardena D, Xing B. **2008**. Effect of dehydration on dicarboxylic acid coordination at goethite/water interface. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **318**: 275–284.

Kanter U, Heller W, Durner J, Winkler JB, Engel M, Behrendt H, Holzinger A, Braun P, Hauser M, Ferreira F, *et al.* 2013. Molecular and Immunological Characterization of Ragweed (Ambrosia artemisiifolia L.) Pollen after Exposure of the Plants to Elevated Ozone over a Whole Growing Season. *PLoS ONE* **8(4):** e61518.

Karasov TL, Chae E, Herman JJ, Bergelson J. **2017**. Mechanisms to Mitigate the Trade-Off between Growth and Defense. *The Plant Cell* **29**: 666–680.

Kaszowska Z, Malek K, Pańczyk M, Mikołajska A. **2013**. A joint application of ATR-FTIR and SEM imaging with high spatial resolution: Identification and distribution of painting materials and their degradation products in paint cross sections. *Vibrational Spectroscopy* **65**: 1–11.

Kato-Noguchi H. **2022**. Allelopathy of Knotweeds as Invasive Plants. *Plants* **11**: 3.

Kazarian SG, Chan KLA. **2013**. ATR-FTIR spectroscopic imaging: recent advances and applications to biological systems. *The Analyst* **138**: 1940.

Kebbas S, Lutts S, Aid F. **2015**. Effect of drought stress on the photosynthesis of Acacia tortilis subsp. raddiana at the young seedling stage. *Photosynthetica* **53(2)**: 288–298.

Kelly LT, Giljohann KM, Duane A, Aquilué N, Archibald S, Batllori E, Bennett AF, Buckland ST, Canelles Q, Clarke MF, *et al.* 2020. Fire and biodiversity in the Anthropocene. *Science* **370(6519):** eabb0355.

Khairudin K, Sukiran NA, Goh HH, Baharum SN, Noor NM. **2014**. Direct discrimination of different plant populations and study on temperature effects by Fourier transform infrared spectroscopy. *Metabolomics* **10**: 203–211.

Kharrat F, Khlif M, Hilliou L, Haboussi M, Covas JA, Nouri H, Bradai C. **2020**. Minimally processed date palm (Phoenix dactylifera L.) leaves as natural fillers and processing aids in poly(lactic acid) composites designed for the extrusion film blowing of thin packages. *Industrial Crops and Products* **154**: 112637.

Kidd H, Shaw D. **2000**. Japanese knotweed - The world's largest female! *Pesticide Outlook* **11**: 99–100.

Kim SW, Ban SH, Chung H, Cho S, Chung HJ, Choi PS, Yoo OJ, Liu JR. **2004**. Taxonomic discrimination of flowering plants by multivariate analysis of Fourier transform infrared spectroscopy data. *Plant Cell Reports* **23**: 246–250.

van Kleunen M, Bossdorf O, Dawson W. **2018**. The Ecology and Evolution of Alien Plants. *Annual Review of Ecology, Evolution, and Systematics* **49**: 25–47.

Van Kleunen M, Weber E, Fischer M. **2010**. A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecology Letters* **13**: 235–245.

Kogami H, Hanba YT, Kibe T, Terashima I, Masuzawa T. **2001**. CO_2 transfer conductance, leaf structure and carbon isotope composition of Polygonum cuspidatum leaves from low and high altitudes. *Plant, Cell & Environment* **24**: 529–538.

Köhn H-F, Hubert LJ. **2015**. Hierarchical Cluster Analysis. In: Wiley StatsRef: Statistics Reference Online. John Wiley & Sons, Ltd, 1–13.

Kong J, Yu S. **2007**. Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures. *Acta Biochimica et Biophysica Sinica* **39**: 549–559.

Kos G, Lohninger H, Krska R. **2003**. Development of a method for the determination of Fusarium fungi on corn using mid-infrared spectroscopy with attenuated total reflection and chemometrics. *Analytical Chemistry* **75**: 1211–1217.

Kotilainen T, Aphalo PJ, Brelsford CC, Böök H, Devraj S, Heikkilä A, Hernández R, Kylling A, Lindfors A V., Robson TM. **2020**. Patterns in the spectral composition of sunlight and biologically meaningful spectral photon ratios as affected by atmospheric factors. *Agricultural and Forest Meteorology* **291**: 108041.

Krähmer A, Böttcher C, Gudi G, Stürtz M, Schulz H. **2021**. Application of ATR-FTIR spectroscopy for profiling of non-structural carbohydrates in onion (*Allium cepa L.*) bulbs. *Food Chemistry* **360**: 129978.

Kuhlgert S, Austic G, Zegarac R, Osei-Bonsu I, Hoh D, Chilvers MI, Roth MG, Bi K, TerAvest D, Weebadde P, *et al.* 2016. MultispeQ Beta: A tool for large-scale plant phenotyping connected to the open photosynQ network. *Royal Society Open Science* **3(10):** 160592.

Lan W, Renard CMGC, Jaillais B, Leca A, Bureau S. **2020**. Fresh, freeze-dried or cell wall samples: Which is the most appropriate to determine chemical, structural and rheological variations during apple processing using ATR-FTIR spectroscopy? *Food Chemistry* **330**: 127357.

Larsen DH, Woltering EJ, Nicole CCS, Marcelis LFM. **2020**. Response of Basil Growth and Morphology to Light Intensity and Spectrum in a Vertical Farm. *Frontiers in Plant Science* **11**: 1893.

Lavoie C. **2017**. The impact of invasive knotweed species (Reynoutria spp.) on the environment: review and research perspectives. *Biological Invasions* **19**: 2319–2337.

Lee CE. **2002**. Evolutionary genetics of invasive species. *Trends in Ecology & Evolution* **17**: 386–391.

Lee R. **2011**. The outlook for population growth. *Science* **333**: 569–573.

Legris M, Nieto C, Sellaro R, Prat S, Casal JJ. **2017**. Perception and signalling of light and temperature cues in plants. *The Plant Journal* **90**: 683–697.

Lei Y, Hannoufa A, Christensen D, Shi H, Prates LL, Yu P. **2018**. Molecular Structural Changes in Alfalfa Detected by ATR-FTIR Spectroscopy in Response to Silencing of TT8 and HB12 Genes. *International Journal of Molecular Sciences 2018, Vol. 19, Page 1046* **19**: 1046.

Lein W, Börnke F, Reindl A, Ehrhardt T, Stitt M, Sonnewald U. **2004**. Target-based discovery of novel herbicides. *Current Opinion in Plant Biology* **7**: 219–225.

Lemoine R. **2000**. Sucrose transporters in plants: update on function and structure. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1465**: 246–262.

Lew TTS, Sarojam R, Jang I-C, Park BS, Naqvi NI, Wong MH, Singh GP, Ram RJ, Shoseyov O, Saito K, Chua NH. 2020. Species-independent analytical tools for next-generation agriculture. *Nature Plants* **6(612)**: 1408–1417.

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

Liao H, D'Antonio CM, Chen B, Huang Q, Peng S. **2016**. How much do phenotypic plasticity and local genetic variation contribute to phenotypic divergences along environmental gradients in widespread invasive plants? A meta-analysis. *Oikos* **125**: 905–917.

Lincoln Soil and Water Conservation District. Bohemian Knotweed Page - Lincoln Soil and Water Conservation District. *http://www.lincolnswcd.org/bohemian-knotweed-page.html*.

Liu Y, Oduor AMO, Dai ZC, Gao FL, Li J, Zhang X, Yu FH. **2021**. Suppression of a plant hormone gibberellin reduces growth of invasive plants more than native plants. *Oikos* **130**: 781–789.

Liu X, Renard CMGC, Bureau S, Le Bourvellec C. **2021**. Revisiting the contribution of ATR-FTIR spectroscopy to characterize plant cell wall polysaccharides. *Carbohydrate Polymers* **262**: 117935.

Liu N, Zhao L, Tang L, Stobbs J, Parkin I, Kunst L, Karunakaran C. **2020**. Mid-infrared spectroscopy is a fast screening method for selecting Arabidopsis genotypes with altered leaf cuticular wax. *Plant, Cell & Environment* **43**: 662–674.

Long F, Liu W, Jiang X, Zhai Q, Cao X, Jiang J, Xu J. **2021**. State-of-the-art technologies for biofuel production from triglycerides: A review. *Renewable and Sustainable Energy Reviews* **148**: 111269.

Lopes J, Sousa C. **2018**. Vibrational spectroscopy for plant varieties and cultivars characterization. In: Comprehensive Analytical Chemistry. 2–299.

López-Ramos M, Perruccio F. **2010**. HPPD: Ligand- and target-based virtual screening on a herbicide target. *Journal of Chemical Information and Modeling* **50**: 801–814.

Lowe S, Browne M, Boudjelas S, De Poorter M. **2000**. *100 of the world's worst invasive alien species: a selection from the global invasive species database*. Invasive Species Specialist Group Auckland, New Zealand.

Lowry DB, Popovic D, Brennan DJ, Holeski LM. **2019**. Mechanisms of a locally adaptive shift in allocation among growth, reproduction, and herbivore resistance in Mimulus guttatus*. *Evolution* **73**: 1168–1181.

Luo Y, Liu H, Wu C, Paraskevaidi M, Deng Y, Shi W, Yuan Y, Feng R, Martin FL, Pang W. **2021**. Diagnostic Segregation of Human Breast Tumours Using Fourier-transform Infrared Spectroscopy Coupled with Multivariate Analysis: Classifying Cancer Subtypes. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **255**: 119694.

Ma J, Xu L, Wang S. **2002**. A quick, simple, and accurate method of screening herbicide activity using green algae cell suspension cultures. *Weed science* **50**: 555–559.

Ma C, Zhang HH, Wang X. **2014**. Machine learning for Big Data analytics in plants. *Trends in Plant Science* **19**: 798–808.

Macchioni V, Picchi V, Carbone K. **2022**. Hop leaves as an alternative source of health-active compounds: Effect of genotype and drying conditions. *Plants* **11**: 99.

Macfarlane JS, Stuart J. **2011**. Development of Strategies for the Control and Eradication of Japanese Knotweed. *MPhil Dissertation, University of Exeter.*

MacLeod M, Arp HPH, Tekman MB, Jahnke A. **2021**. The global threat from plastic pollution. *Science* **373**: 61–65.

Madden MJL, Young RG, Brown JW, Miller SE, Frewin AJ, Hanner RH. **2019**. Using DNA barcoding to improve invasive pest identification at U.S. ports-of-entry. *PLOS ONE* **14**: e0222291.

Mafakheri A, Siosemardeh A, Bahramnejad B, Struik PC, Sohrabi Y. **2010**. Effect of Drought Stress on Yield, Proline and Chlorophyll Contents in Three Chickpea Cultivars. *Australian Journal of Crop Science* **4**: 580–585.

Mahlein AK. **2016**. Plant disease detection by imaging sensors – Parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease* **100**: 241–254.

Mallamace F, Corsaro C, Mallamace D, Vasi S, Vasi C, Dugo G. **2015**. The role of water in protein's behavior: The two dynamical crossovers studied by NMR and FTIR techniques. *Computational and Structural Biotechnology Journal* **13**: 33–37.

Mandák B, Pyšek P, Lysák M, Suda JAN, Krahulcová A, Bímová K. **2003**. Variation in DNA-ploidy Levels of Reynoutria Taxa in the Czech Republic. *Annals of Botany* **92**: 265–272.

Manoharan B, Qi SS, Dhandapani V, Chen Q, Rutherford S, Wan JSH, Jegadeesan S, Yang HY, Li Q, Li J, *et al.* 2019. Gene Expression Profiling Reveals Enhanced Defense Responses in an Invasive Weed Compared to Its Native Congener During Pathogenesis. *International Journal of Molecular Sciences* **20**: 4916.

Marks S, Paravicini G. **2017**. French and Italians sense golden opportunity in glyphosate ban. *POLITICO Europe, Axel Springer SE, Berlin.*

Martin FM, Dommanget F, Evette A. **2020a**. Improving the management of Japanese knotweed s.l.: a response to Jones and colleagues. *NeoBiota* **63**: 147–153.

Martin FM, Dommanget F, Janssen P, Spiegelberger T, Clément Viguier·, Evette A. **2019**. Could knotweeds invade mountains in their introduced range? An analysis of patches dynamics along an elevational gradient. *Alpine Botany* **129**: 33–42.

Martin FM, Dommanget F, Lavallée F, Evette A. **2020b**. Clonal growth strategies of *Reynoutria japonica* in response to light, shade, and mowing, and perspectives for management. *NeoBiota* **56**: 89–110.

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

Martin FM, Müllerová J, Borgniet L, Dommanget F, Breton V, Evette A. **2018**. Using Single- and Multi-Date UAV and Satellite Imagery to Accurately Monitor Invasive Knotweed Species. *Remote Sensing 2018, Vol. 10, Page 1662* **10**: 1662.

Maruri-López I, Aviles-Baltazar NY, Buchala A, Serrano M. **2019**. Intra and Extracellular Journey of the Phytohormone Salicylic Acid. *Frontiers in Plant Science* **0**: 423.

Mason MG, Ross JJ, Babst BA, Wienclaw BN, Beveridge CA. **2014**. Sugar demand, not auxin, is the initial regulator of apical dominance. *Proceedings of the National Academy of Sciences of the United States of America* **111**: 6092–6097.

McManamen C, Nelson CR, Wagner V. **2018**. Timing of seeding after herbicide application influences rates of germination and seedling biomass of native plants used for grassland restoration. *Restoration Ecology* **26**: 1137–1148.

McNear DH, Chaney RL, Sparks DL. **2010**. The hyperaccumulator Alyssum murale uses complexation with nitrogen and oxygen donor ligands for Ni transport and storage. *Phytochemistry* **71**: 188–200.

Mehmood T, Liland KH, Snipen L, Sæbø S. **2012**. A review of variable selection methods in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems* **118**: 62–69.

Mejia AF, Nebel MB, Eloyan A, Caffo B, Lindquist MA. **2017**. PCA leverage: Outlier detection for high-dimensional functional magnetic resonance imaging data. *Biostatistics* **18**: 521–536.

de Mendiburu F. **2020**. agricolae. *Universidad Nacional Agraria La Molina*.

Met Office. **2018**. Regional values - Met Office. *https://www.metoffice.gov.uk/research/climate/maps-and-data/regional-values. Date Accessed 01-06-2019*.

Met Office. **2019**. UK Regional Climates. *https://www.metoffice.gov.uk/research/climate/maps-and-data/regional-climates/index*.

Michalet S, Rouifed S, Pellassa-Simon T, Fusade-Boyer M, Meiffren G, Nazaret S, Piola F. **2017**. Tolerance of Japanese knotweed s.l. to soil artificial polymetallic pollution: early metabolic responses and performance during vegetative multiplication. *Environmental Science and Pollution Research* **24**: 20897–20907.

Miller AJ. **2014**. Plant Mineral Nutrition. In: eLS. Chichester, UK: John Wiley & Sons, Ltd.

Miller JH, Manning ST, Enloe SF. **2013**. A management guide for invasive plants in southern forests. *Gen. Tech. Rep. SRS–131. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 120 p.* **131**: 1–120.

Milne RI, Abbott RJ. **2000**. Origin and evolution of invasive naturalized material of Rhododendron ponticum L. in the British Isles. *Molecular Ecology* **9**: 541–556.

Minnehaha Creek Watershed District. **2013**. Hybrid Milfoil Study | MCWD: Minnehaha Creek Watershed District. *Minnehaha Creek Watershed District Invasive Species Projects*.

Minnes R, Nissinmann M, Maizels Y, Gerlitz G, Katzir A, Raichlin Y. **2017**. Using Attenuated Total Reflection–Fourier Transform Infra-Red (ATR-FTIR) spectroscopy to distinguish between melanoma cells with a different metastatic potential. *Scientific Reports* **7**: 1–7.

Mishra P, Roger JM, Marini F, Biancolillo A, Rutledge DN. **2021**. FRUITNIR-GUI: A graphical user interface for correcting external influences in multi-batch near infrared experiments related to fruit quality prediction. *Postharvest Biology and Technology* **175**: 111414.

Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G, *et al.* 2015. Human-level control through deep reinforcement learning. *Nature 2015 518:7540* **518**: 529–533.

Monaghan RM, Paton RJ, Smith LC, Drewry JJ, Littlejohn RP. **2005**. The impacts of nitrogen fertilisation and increased stocking rate on pasture yield, soil physical condition and nutrient losses in drainage from a cattle-grazed pasture. *New Zealand Journal of Agricultural Research* **48**: 227–240.

Mondal R, Choudhuri MA. **1981**. Role of Hydrogen Peroxide in Senescence of Excised Leaves of Rice and Maize. *Biochemie und Physiologie der Pflanzen* **176**: 700–709.

Moody ML, Les DH. **2007**. Geographic distribution and genotypic composition of invasive hybrid watermilfoil (Myriophyllum spicatum × M. sibiricum) populations in North America. *Biological Invasions* **9**: 559–570.

Morais CL. **2020**. Novel chemometric approaches towards handling biospectroscopy datasets. *(PhD Thesis) University of Central Lancashire, Preston*.

Morais CL, Butler HJ, McAinsh MR, Martin FL. **2019a**. Plant Hyperspectral Imaging. *eLS*: 1–12.

Morais CLM, Costa FSL, Lima KMG. **2017**. Variable selection with a support vector machine for discriminating Cryptococcus fungal species based on ATR-FTIR spectroscopy. *Analytical Methods* **9**: 2964–2970.

Morais CLM, Lima KMG. **2018**. Principal Component Analysis with Linear and Quadratic Discriminant Analysis for Identification of Cancer Samples Based on Mass Spectrometry. *Article J. Braz. Chem. Soc* **29**: 472–481.

Morais CLM, Lima KMG, Singh M, Martin FL. **2020**. Tutorial: multivariate classification for vibrational spectroscopy in biological samples. *Nature Protocols* **15**: 2143–2162.

Morais CLM, Paraskevaidi M, Cui L, Fullwood NJ, Isabelle M, Lima KMG, Martin-Hirsch PL, Sreedhar H, Trevisan J, Walsh MJ, *et al.* 2019b. Standardization of complex biologically derived spectrochemical datasets. *Nature Protocols* **14**: 1546–1577.

Morris TL, Esler KJ, Barger NN, Jacobs SM, Cramer MD. **2011**. Ecophysiological traits associated with the competitive ability of invasive Australian acacias. *Diversity and Distributions* **17**: 898–910.

Moskal P, Wesełucha-Birczyńska A, Łabanowska M, Filek M. **2019**. Adaxial and abaxial pattern of Urtica dioica leaves analyzed by 2DCOS ATR-FTIR as a function of their growth time and impact of environmental pollution. *Vibrational Spectroscopy* **104**: 102948.

Mounger J, Ainouche M, Bossdorf O, Cavé-Radet A, Li B, Parepa M, Salmon A, Yang J, Richards C. **2021**. Epigenetics and the success of invasive plants. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376(1826)**: 20200117.

Moura RF, Queiroga D, Vilela E, Moraes AP. **2021**. Polyploidy and high environmental tolerance increase the invasive success of plants. *Journal of Plant Research* **134**: 105–114.

Movasaghi Z, Rehman S, ur Rehman I. **2008**. Fourier Transform Infrared (FTIR) Spectroscopy of Biological Tissues. *Applied Spectroscopy Reviews* **43**: 134–179.

Mullerscharer H, Schaffner U, Steinger T. **2004**. Evolution in invasive plants: implications for biological control. *Trends in Ecology & Evolution* **19**: 417–422.

Murphy BW. **2015**. Impact of soil organic matter on soil properties - A review with emphasis on Australian soils. *Soil Research* **53**: 605–635.

Murrell C, Gerber E, Krebs C, Parepa M, Schaffner U, Bossdorf O. **2011**. Invasive knotweed affects native plants through allelopathy. *American Journal of Botany* **98**: 38–43.

Nadeem MA, Amjad Nawaz M, Shahid MQ, Doğan Y, Comertpay G, Yıldız M, Hatipoğlu R, Ahmad F, Alsaleh A, Labhane N, *et al.* **2017**. DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. *Biotechnology & Biotechnological Equipment* **32(2)**: 261-285.

Naqvi SMZA, Zhang Y, Ahmed S, Abdulraheem MI, Hu J, Tahir MN, Raghavan V. **2022**. Applied surface enhanced Raman Spectroscopy in plant hormones detection, annexation of advanced technologies: A review. *Talanta* **236**: 122823.

Nawrot-Hadzik I, Slusarczyk S, Granica S, Hadzik J, Matkowski A. **2019**. Phytochemical diversity in rhizomes of three Reynoutria species and their antioxidant activity correlations elucidated by LC-ESI-MS/MS analysis. *Molecules* **24**.

Nentwig W, Bacher S, Kumschick S, Pyšek P, Vilà M. **2018**. More than "100 worst" alien species in Europe. *Biological Invasions* **20**: 1611–1621.

Netting AG, Theobald JC, Dodd IC. **2012**. Xylem sap collection and extraction methodologies to determine in vivo concentrations of ABA and its bound forms by gas chromatography-mass spectrometry (GC-MS). *Plant Methods* **8**: 1–14.

Nguyen L. **1997**. The Value of Japanese Knotweed in Phytoremediation of Contaminated Soils Along the Woonasquatucket River. *Doctoral dissertation, Brown University.*

Nicotra AB, Atkin OK, Bonser SP, Davidson AM, Finnegan EJ, Mathesius U, Poot P, Purugganan MD, Richards CL, Valladares F, *et al.* **2010**. Plant phenotypic plasticity in a changing climate. *Trends in Plant Science* **15**: 684–692.

Nishinari K, Takemasa M, Zhang H, Takahashi R. **2007**. 2.19 Storage Plant Polysaccharides: Xyloglucans, Galactomannans, Glucomannans. In: Comprehensive glycoscience. Elsevier Oxford, 613–652.

Nogales-Bueno J, Baca-Bocanegra B, Rooney A, Miguel Hernández-Hierro J, José Heredia F, Byrne HJ. **2017**. Linking ATR-FTIR and Raman features to phenolic extractability and other attributes in grape skin. *Talanta* **167**: 44–50.

Nozahic V, Amziane S. **2012**. Influence of sunflower aggregates surface treatments on physical properties and adhesion with a mineral binder. *Composites Part A: Applied Science and Manufacturing* **43**: 1837–1849.

de Oliveira GA, de Castilhos F, Renard CMGC, Bureau S. **2014**. Comparison of NIR and MIR spectroscopic methods for determination of individual sugars, organic acids and carotenoids in passion fruit. *Food Research International* **60**: 154–162.

Olsen SR, Cole C V, Watanabe FS, Dean LA, States. U, Agriculture. D of. **1954**. *Estimation of available phosphorus in soils by extraction with sodium bicarbonate*. Washington, D.C.: U.S. Dept. of Agriculture.

Orcelli T, di Mauro E, Urbano A, Valezi DF, da Costa ACS, Zaia CTB V., Zaia DAM. **2018**. Study of Interaction Between Glyphosate and Goethite Using Several Methodologies: an Environmental Perspective. *Water, Air, & Soil Pollution* **229**: 150.

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

Orwin KH, Buckland SM, Johnson D, Turner BL, Smart S, Oakley S, Bardgett RD. **2010**. Linkages of plant traits to soil properties and the functioning of temperate grassland. *Journal of Ecology* **98**: 1074–1083.

Osakabe Y, Osakabe K, Shinozaki K, Tran LSP. **2014**. Response of plants to water stress. *Frontiers in Plant Science* **5**: 86.

Paini DR, Sheppard AW, Cook DC, De Barro PJ, Worner SP, Thomas MB. **2016**. Global threat to agriculture from invasive species. *Proceedings of the National Academy of Sciences of the United States of America* **113**: 7575–7579.

Palacio-López K, Gianoli E. **2011**. Invasive plants do not display greater phenotypic plasticity than their native or non-invasive counterparts: a meta-analysis. *Oikos* **120**: 1393–1401.

Pandey P, Ge Y, Stoerger V, Schnable JC. **2017**. High throughput in vivo analysis of plant leaf chemical properties using hyperspectral imaging. *Frontiers in Plant Science* **8**: 1348.

Pandit MK, White SM, Pocock MJO. **2014**. The contrasting effects of genome size, chromosome number and ploidy level on plant invasiveness: a global analysis. *New Phytologist* **203**: 697–703.

Parepa M, Bossdorf O. **2016**. Testing for allelopathy in invasive plants: it all depends on the substrate! *Biological Invasions* **18**: 2975–2982.

Parepa M, Fischer M, Bossdorf O. **2013**. Environmental variability promotes plant invasion. *Nature Communications* **4**: 1–4.

Parepa M, Fischer M, Krebs C, Bossdorf O. **2014**. Hybridization increases invasive knotweed success. *Evolutionary Applications* **7**: 413–420.

Parepa M, Kahmen A, Werner RA, Fischer M, Bossdorf O. **2019**. Invasive knotweed has greater nitrogen-use efficiency than native plants: evidence from a 15N pulse-chasing experiment. *Oecologia* **191**: 389–396.

Parepa M, Schaffner U, Bossdorf O. **2012**. Sources and modes of action of invasive knotweed allelopathy: the effects of leaf litter and trained soil on the germination and growth of native plants. *NeoBiota* **13**: 15–30.

Park Y, Runkle ES. **2018**. Spectral effects of light-emitting diodes on plant growth, visual color quality, and photosynthetic photon efficacy: White versus blue plus red radiation. *PLoS ONE* **13(8)**: e0202386.

Parker IM, Rodriguez J, Loik ME. **2003**. An Evolutionary Approach to Understanding the Biology of Invasions: Local Adaptation and General-Purpose Genotypes in the Weed Verbascum thapsus. *Conservation Biology* **17**: 59–72.

Pashley CH. **2003**. The use of molecular markers in the study of the origin and evolution of Japanese Knotweed *sensu lato*. *Doctoral dissertation, University of Leicester.*

Pashley CH, Bailey JP, Ferris C. **2007**. Clonal diversity in British populations of the alien invasive Giant Knotweed, Fallopia sachalinensis (F. Schmidt) Ronse Decraene, in the context of European and Japanese plants. *Watsonia* **26(3)**: 359-72.

Payne T, Hoxley M. **2012**. Identifying and eradicating Japanese knotweed in the UK built environment. *Structural Survey* **30**: 24–42.

Pejchar L, Mooney HA. **2009**. Invasive species, ecosystem services and human well-being. *Trends in Ecology and Evolution* **24**: 497–504.

Peng W, Qin R, Li X, Zhou H. **2013**. Botany, phytochemistry, pharmacology, and potential application of Polygonum cuspidatum Sieb.et Zucc.: A review. *Journal of Ethnopharmacology* **148**: 729–745.

Pennisi G, Pistillo A, Orsini F, Cellini A, Spinelli F, Nicola S, Fernandez JA, Crepaldi A, Gianquinto G, Marcelis LFM. **2020**. Optimal light intensity for sustainable water and energy use in indoor cultivation of lettuce and basil under red and blue LEDs. *Scientia Horticulturae* **272**: 109508.

Peters S. **2019**. Public Outreach on Japanese Knotweed in Washington Township, Michigan. *Dissertation, The Honors College Oakland University.*

Pevsner A, Diem M. **2001**. Infrared spectroscopic studies of major cellular components. Part II: The effect of hydration on the spectra of nucleic acids. *Applied Spectroscopy* **55**: 1502–1505.

Pichancourt J-B, van Klinken RD. **2012**. Phenotypic Plasticity Influences the Size, Shape and Dynamics of the Geographic Distribution of an Invasive Plant (G Bonaventure, Ed.). *PLoS ONE* **7**: e32323.

Pierik R, Cuppens MLC, Voesenek LACJ, Visser EJW. **2004**. Interactions between Ethylene and Gibberellins in Phytochrome-Mediated Shade Avoidance Responses in Tobacco. *Plant Physiology* **136**: 2928–2936.

Pilling M, Gardner P. **2016**. Fundamental developments in infrared spectroscopic imaging for biomedical applications. *Chemical Society Reviews* **45**: 1935–1957.

PLS_Toolbox with MIA_Toolbox 8.2.1. **2016**. PLS_Toolbox with MIA_Toolbox 8.2.1. *Eigenvector Research, Inc., Manson, WA USA 98831*.

Porfírio S, Sonon R, Gomes da Silva MDR, Peixe A, Cabrita MJ, Azadi P. **2016**. Quantification of free auxins in semi-hardwood plant cuttings and microshoots by dispersive liquid–liquid microextraction/microwave derivatization and GC/MS analysis. *Analytical Methods* **8**: 6089–6098.

Pradko AG, Litvinovskaya RP, Sauchuk AL, Drach S V., Baranovsky A V., Zhabinskii VN, Mirantsova T V., Khripach VA. **2015**. A new ELISA for quantification of brassinosteroids in plants. *Steroids* **97**: 78–86.

Price EAC, Gamble R, Williams GG, Marshall C. **2002**. Seasonal patterns of partitioning and remobilization of 14C in the invasive rhizomatous perennial Japanese knotweed (*Fallopia japonica* (Houtt.) Ronse Decraene). In: Ecology and Evolutionary Biology of Clonal Plants. Dordrecht: Springer Netherlands, 125–140.

Professional Agricultural Analysis Group. **2013**. Routine soil samples. *PAAG soil sampling guide*.

Pyšek P. **1997**. The ecology and evolution of clonal plants. *Backhuys Publishers, Leiden*.

Pyšek P, Hulme PE, Simberloff D, Bacher S, Blackburn TM, Carlton JT, Dawson W, Essl F, Foxcroft LC, Genovesi P, *et al.* 2020. Scientists' warning on invasive alien species. *Biological Reviews* **95**: 1511–1534.

Qing H, Yanlin H, Fenlin S, Zuyi T. **1996**. Effects of pH and metal ions on the conformation of bovine serum albumin in aqueous solution: An Attenuated Total Reflection (ATR) FTIR spectroscopic study. *Spectrochimica Acta - Part A Molecular Spectroscopy* **52**: 1795–1800.

R Core Team. **2014**. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.*

Raba DN, Poiana MA, Borozan AB, Stef M, Radu F, Popa MV. **2015**. Investigation on crude and high-temperature heated coffee oil by ATR-FTIR spectroscopy along with antioxidant and antimicrobial properties. *PLoS ONE* **10(9)**: e0138080.

Rahman MM, Feng X, Zhang H, Yan X, Peng Q, Yu P. **2019**. Using vibrational ATR-FTIR spectroscopy with chemometrics to reveal faba CHO molecular spectral profile and CHO nutritional features in ruminant systems. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **214**: 269–276.

Rana R, Herz K, Bruelheide H, Dietz S, Haider S, Jandt U, Pena R. **2018**. Leaf Attenuated Total Reflection Fourier Transform Infrared (ATR-FTIR) biochemical profile of grassland plant species related to land-use intensity. *Ecological Indicators* **84**: 803–810.

Rana R, Müller G, Naumann A, Polle A. **2008**. FTIR spectroscopy in combination with principal component analysis or cluster analysis as a tool to distinguish beech (Fagus sylvatica L.) trees grown at different sites. *Holzforschung* **62**: 530–538.

Ribeiro da Luz B. **2006**. Attenuated total reflectance spectroscopy of plant leaves: a tool for ecological and botanical studies. *New Phytologist* **172**: 305–318.

Richards CL, Alonso C, Becker C, Bossdorf O, Bucher E, Colomé-Tatché M, Durka W, Engelhardt J, Gaspar B, Gogol-Döring A, *et al.* 2017. Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecology Letters* **20**: 1576–1590.

Richards CL, Bossdorf O, Muth NZ, Gurevitch J, Pigliucci M. **2006**. Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecology Letters* **9**: 981–993.

Richards CL, Pigliucci M. **2020**. Epigenetic Inheritance. A Decade into the Extended Evolutionary Synthesis. *Paradigmi* **XXXVIII**: 463–494.

Richards CL, Schrey AW, Pigliucci M. **2012**. Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation (M Velend, Ed.). *Ecology Letters* **15**: 1016–1025.

Richards CL, Walls RL, Bailey JP, Parameswaran R, George T, Pigliucci M. **2008**. Plasticity in salt tolerance traits allows for invasion of novel habitat by Japanese knotweed s. l. (*Fallopia japonica and F.x bohemica*, Polygonaceae). *American Journal of Botany* **95**: 931–942.

Ringelmann A, Riedel M, Riederer M, Hildebrandt U. **2009**. Two sides of a leaf blade: Blumeria graminis needs chemical cues in cuticular waxes of Lolium perenne for germination and differentiation. *Planta* **230**: 95–105.

Rogowska A, Szakiel A. **2020**. The role of sterols in plant response to abiotic stress. *Phytochemistry Reviews* **19**: 1525–1538.

Romera-Fernández M, Berrueta LA, Garmón-Lobato S, Gallo B, Vicente F, Moreda JM. **2012**. Feasibility study of FT-MIR spectroscopy and PLS-R for the fast determination of anthocyanins in wine. *Talanta* **88**: 303–310.

Rouifed S, Byczek C, Laffray D, Piola F. **2012**. Invasive Knotweeds are Highly Tolerant to Salt Stress. *Environmental Management* **50**: 1027–1034.

Rouifed S, Puijalon S, Bardon C, Meiffren G, Buonomo A, Sebei N, Poussineau S, Vallier F, Shimoda M, Piola F. **2018**. Comparison of defence and performance traits between one widespread clone and native populations in a major invasive plant species (I Ibáñez, Ed.). *Diversity and Distributions* **24**: 297–312.

Roy HE, Bacher S, Essl F, Adriaens T, Aldridge DC, Bishop JDD, Blackburn TM, Branquart E, Brodie J, Carboneras C, *et al.* 2018. Developing a list of invasive alien species likely to threaten biodiversity and ecosystems in the European Union. *Global Change Biology* **25**: 1032–1048.

Ruiz-Sola MÁ, Rodríguez-Concepción M. **2012**. Carotenoid Biosynthesis in Arabidopsis: A Colorful Pathway. *The Arabidopsis Book* **10**: e0158.

Ruoff K, Luginbu¨hl W, Luginbu¨hl L, Ku¨nzli R, Ku¨nzli K, Äa M, Iglesias T, Bogdanov S, Bosset JO, Von Der Ohe K, von der Ohe, W. and Amadò, R. 2006. Authentication of the Botanical and Geographical Origin of Honey by Mid-Infrared Spectroscopy. *Journal of agricultural and food chemistry* **54(18)**: 6873-6880.

Saleem M, Fariduddin Q, Castroverde CDM. **2021**. Salicylic acid: A key regulator of redox signalling and plant immunity. *Plant Physiology and Biochemistry* **168**: 381–397.

Sankaran S, Maja J, Buchanon S, Ehsani R. **2013**. Huanglongbing (Citrus Greening) Detection Using Visible, Near Infrared and Thermal Imaging Techniques. *Sensors* **13**: 2117–2130.

Santo P. **2017**. Assessing diminution in value of residential properties affected by Japanese Knotweed. *Journal of Building Survey, Appraisal & Valuation* **6(11)**: 211-221.

Schenone M, Dančík V, Wagner BK, Clemons PA. **2013**. Target identification and mechanism of action in chemical biology and drug discovery. *Nature chemical biology* **9**: 232.

Schmidt P, Dybal J, Trchová M. **2006**. Investigations of the hydrophobic and hydrophilic interactions in polymer-water systems by ATR FTIR and Raman spectroscopy. *Vibrational Spectroscopy* **42**: 278–283.

Schoen DJ, Schultz ST. **2019**. Somatic Mutation and Evolution in Plants. *Annual Review of Ecology, Evolution, and Systematics* **50**: 49–73.

Schuster TM, Reveal JL, Bayly MJ, Kron KA. **2015**. An updated molecular phylogeny of Polygonoideae (Polygonaceae): Relationships of Oxygonum, Pteroxygonum, and Rumex, and a new circumscription of Koenigia. *Taxon* **64**: 1188–1208.

Science and Technology Committee - House of Commons. **2019**. Japanese knotweed and the built environment. *Science and Technology Committee Report*: 13.

Seppälä M, Laine A, Rintala J. **2013**. Screening of novel plants for biogas production in northern conditions. *Bioresource Technology*. **139**: 355-62.

Shackleton RT, Shackleton CM, Kull CA. **2019**. The role of invasive alien species in shaping local livelihoods and human well-being: A review. *Journal of Environmental Management* **229**: 145–157.

Sharma S, Uttam KN. **2017**. Rapid analyses of stress of copper oxide nanoparticles on wheat plants at an early stage by laser induced fluorescence and attenuated total reflectance Fourier transform infrared spectroscopy. *Vibrational Spectroscopy* **92**: 135–150.

Sharma S, Uttam KN. **2018**. Early Stage Detection of Stress Due to Copper on Maize (Zea mays L.) by Laser-Induced Fluorescence and Infrared Spectroscopy. *Journal of Applied Spectroscopy* **85**: 771–780.

Shaw RH, Bryner S, Tanner R. **2009**. The life history and host range of the Japanese knotweed psyllid, Aphalara itadori Shinji: Potentially the first classical biological weed control agent for the European Union. *Biological Control* **49**: 105–113.

Shiroma C, Rodriguez-Saona L. **2009**. Application of NIR and MIR spectroscopy in quality control of potato chips. *Journal of Food Composition and Analysis* **22**: 596–605.

Shivu B, Seshadri S, Li J, Oberg KA, Uversky VN, Fink AL. **2013**. Distinct β-Sheet Structure in Protein Aggregates Determined by ATR–FTIR Spectroscopy. *Biochemistry* **52(31)**:5176-83.

De Silva DLR, Cox RC, Hetherington AM, Mansfield TA. **1986**. The role of abscisic acid and calcium in determining the behaviour of adaxial and abaxial stomata. *New phytologist* **104**: 41–51.

Silvertown J. **2008**. The evolutionary maintenance of sexual reproduction: Evidence from the ecological distribution of asexual reproduction in clonal plants. *International Journal of Plant Sciences* **169**: 157–168.

Simsek Ozek N, Tuna S, Erson-Bensan AE, Severcan F. **2010**. Characterization of microRNA-125b expression in MCF7 breast cancer cells by ATR-FTIR spectroscopy. *The Analyst* **135**: 3094.

Šimura J, Antoniadi I, Široká J, Tarkowská D, Strnad M, Ljung K, Novák O. **2018**. Plant Hormonomics: Multiple Phytohormone Profiling by Targeted Metabolomics. *Plant Physiology* **177**: 476.

Singh S, Kaur I, Kariyat R. **2021**. The Multifunctional Roles of Polyphenols in Plant-Herbivore Interactions. *International Journal of Molecular Sciences* **22**: 1442.

Skinner RH, Van Der Grinten M, Gover AE. **2012**. Planting Native Species to Control Site Reinfestation by Japanese Knotweed (*Fallopia japonica*). *Ecological Restoration* **30**: 192–199.

Skolik P, McAinsh MR, Martin FL. **2018**. Biospectroscopy for Plant and Crop Science. In: Comprehensive Analytical Chemistry. Elsevier B.V., 15–49.

Skolik P, McAinsh MR, Martin FL. **2019a**. ATR-FTIR spectroscopy non-destructively detects damage-induced sour rot infection in whole tomato fruit. *Planta* **249**: 925–939.

Skolik P, Morais CLM, Martin FL, McAinsh MR. **2019b**. Determination of developmental and ripening stages of whole tomato fruit using portable infrared spectroscopy and Chemometrics. *BMC Plant Biology* **19**: 236.

Smith H. **1982**. Light quality, photoperception, and plant strategy. *Annual review of plant physiology* **33**: 481–518.

Smith B. **2018**. Infrared spectral interpretation: A systematic approach. *CRC Press.* 1–304.

Smith JMD, Ward JP, Child LE, Owen MR. **2007**. A simulation model of rhizome networks for *Fallopia japonica* (Japanese knotweed) in the United Kingdom.

Ecological Modelling **200**: 421–432.

Snow AA, Andersen B, Jorgensen RB. **1999**. Costs of transgenic herbicide resistance introgressed from Brassica napus into weedy *B. rapa*. *Molecular Ecology* **8**: 605–615.

Soll J. **2004**. Controlling knotweed in the Pacific Northwest. *URL: http://tncweeds. ucdavis. edu/esadocs. html (accessed 30 Nov 2004). The Nature Conservancy*.

Sołtysiak J. **2020**. Heavy metals tolerance in an invasive weed (*Fallopia japonica*) under different levels of soils contamination. *Journal of Ecological Engineering* **21**: 81–91.

Soriano-Disla JM, Janik LJ, McLaughlin MJ. **2018**. Assessment of cyanide contamination in soils with a handheld mid-infrared spectrometer. *Talanta* **178**: 400–409.

Spalding K, Bonnier F, Bruno C, Blasco H, Board R, Benz-de Bretagne I, Byrne HJ, Butler HJ, Chourpa I, Radhakrishnan P, *et al.* 2018. Enabling quantification of protein concentration in human serum biopsies using attenuated total reflectance – Fourier transform infrared (ATR-FTIR) spectroscopy. *Vibrational Spectroscopy* **99**: 50–58.

Springer NP, Duchin F. **2014**. Feeding nine billion people sustainably: Conserving land and water through shifting diets and changes in technologies. *Environmental Science and Technology* **48**: 4444–4451.

Stasiak MA, Hofstra G, Fletcher RA. **1992**. Physiological changes induced in birch seedlings by sublethal applications of glyphosate. *Canadian Journal of Forest Research* **22**: 812–817.

Stordrange L, Libnau FO, Malthe-Sørenssen D, Kvalheim OM. **2002**. Feasibility study of NIR for surveillance of a pharmaceutical process, including a study of different preprocessing techniques. *Journal of Chemometrics* **16**: 529–541.

Van Straalen NM, Legler J. **2018**. Decision-making in a storm of discontent. *Science* **360**: 958–960.

Strgulc Krajšek S, Buh P, Zega A, Kreft S. **2008**. Identification of Herbarium Whole-Leaf Samples ofEpilobium Species by ATR-IR Spectroscopy. *Chemistry & Biodiversity* **5**: 310–317.

Strong R, Martin FL, Jones KC, Shore RF, Halsall CJ. **2017**. Subtle effects of environmental stress observed in the early life stages of the Common frog, Rana temporaria. *Scientific Reports* **7**: 1–13.

Stroud S, Fennell M, Mitchley J, Lydon S, Peacock J, Bacon KL. **2022**. The botanical education extinction and the fall of plant awareness. *Ecology and Evolution* **12**: e9019.

Sultan SE. **2015**. *Organism and environment: ecological development, niche construction, and adaption*. Oxford University Press, USA.

Taiz L, Zeiger E, Møller IM, Murphy A. **2015**. Plant physiology and development. *Plant physiology and development.*

Talari ACS, Martinez MAG, Movasaghi Z, Rehman S, Rehman IU. **2017**. Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Applied Spectroscopy Reviews* **52**: 456–506.

Tamura M, Tharayil N. **2014**. Plant litter chemistry and microbial priming regulate the accrual, composition and stability of soil carbon in invaded ecosystems. *New Phytologist* **203**: 110–124.

Tarone RE. **2018**. On the International Agency for Research on Cancer classification of glyphosate as a probable human carcinogen. *European Journal of Cancer Prevention* **27**: 82–87.

Tehrani FM, Alexandrou A, Adhikari D, Mahoney M, Maestas R. **2014**. Energy inputs and carbon dioxide emissions during construction of a golf course. *International Journal of Engineering Research & Innovation* **6**: 78–86.

Thalmann M, Santelia D. **2017**. Starch as a determinant of plant fitness under abiotic stress. *New Phytologist* **214**: 943–951.

Therriault TW, Weise AM, Higgins SN, Guo Y, Duhaime J. **2013**. Risk assessment for three dreissenid mussels (Dreissena polymorpha, Dreissena rostriformis bugensis, and Mytilopsis leucophaeata) in Canadian freshwater ecosystems. *Canadian Science Advisory Secretariat = Secrétariat canadien de consultation scientifique.*

Thiébaut G, Tarayre M, Rodríguez-Pérez H. **2019**. Allelopathic Effects of Native Versus Invasive Plants on One Major Invader. *Frontiers in Plant Science* **10**: 854.

Thum R, Newman R, Fieldseth E. **2017**. Occurrence and Distribution of Eurasian, Northern and Hybrid Watermilfoil in Lake Minnetonka and Christmas Lake: Genetic Analysis Phase II. *Minnehaha Creek Watershed District Invasive Species Project Report.*

Tiwari G, Slaughter DC, Cantwell M. **2013**. Nondestructive maturity determination in green tomatoes using a handheld visible and near infrared instrument. *Postharvest Biology and Technology* **86**: 221–229.

Traoré M, Kaal J, Martínez Cortizas A. **2018**. Differentiation between pine woods according to species and growing location using FTIR-ATR. *Wood Science and Technology* **52**: 487–504.

Del Tredici P. **2017**. The introduction of Japanese knotweed, *Reynoutria japonica* , into North America. *The Journal of the Torrey Botanical Society* **144**: 406–416.

Tretter L, Patocs A, Chinopoulos C. **2016**. Succinate, an intermediate in metabolism, signal transduction, ROS, hypoxia, and tumorigenesis. *Biochimica et Biophysica Acta - Bioenergetics* **1857**: 1086–1101.

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

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

Trujillo-González A, Thuo DN, Divi U, Sparks K, Wallenius T, Gleeson D. **2022**. Detection of Khapra Beetle Environmental DNA Using Portable Technologies in Australian Biosecurity. *Frontiers in Insect Science* **2**. 795379.

Tucker Serniak L. **2016**. Comparison of the allelopathic effects and uptake of *Fallopia japonica* phytochemicals by *Raphanus sativus* (K Bailey, Ed.). *Weed Research* **56**: 97–101.

Tuia D, Kellenberger B, Beery S, Costelloe BR, Zuffi S, Risse B, Mathis A, Mathis MW, van Langevelde F, Burghardt T, *et al.* 2022. Perspectives in machine learning for wildlife conservation. *Nature Communications 2022 13:1* **13**: 1–15.

UK Soil Observatory. **2020**. UKSO. *http://mapapps2.bgs.ac.uk/ukso/home.html. Date Accessed 21-05-2020*.

Urcelay C, Austin AT. **2020**. Exotic plants get a little help from their friends. *Science (New York, N.Y.)* **368**: 934–936.

Urgenson LS. **2006**. The Ecological Consequences of Knotweed Invasion into Riparian Forests. *Master's Thesis, University of Washington*.

USDA Plants Database. Plant Profile: Polygonum ×bohemicum (J. Chrtek & Chrtková) Zika & Jacobson [cuspidatum × sachalinense]. *United States Department of Agriculture Plants Database: https://plants.sc.egov.usda.gov/core/profile?symbol=POBO10*.

Usman K, Al-Ghouti MA, Abu-Dieyeh MH. **2019**. The assessment of cadmium, chromium, copper, and nickel tolerance and bioaccumulation by shrub plant Tetraena qataranse. *Scientific Reports* **9**: 1–11.

Valliere JM, Escobedo EB, Bucciarelli GM, Sharifi MR, Rundel PW. **2019**. Invasive annuals respond more negatively to drought than native species. *New Phytologist* **223**: 1647–1656.

Valsecchi C, Grisoni F, Consonni V, Ballabio D. **2020**. Consensus versus Individual QSARs in Classification: Comparison on a Large-Scale Case Study. *Journal of Chemical Information and Modeling* **60**: 1215–1223.

Vanderklein DW, Galster J, Scherr R. **2014**. The impact of Japanese knotweed on stream baseflow. *Ecohydrology* **7**: 881–886.

VanWallendael A, Hamann E, Franks SJ. **2018**. Evidence for plasticity, but not local adaptation, in invasive Japanese knotweed (Reynoutria japonica) in North America. *Evolutionary Ecology* **32**: 395–410.

Van Der Velde G, Platvoet D. **2007**. Quagga mussels Dreissena rostriformis bugensis (Andrusov, 1897) in the Main River (Germany). *Aquatic Invasions* **2**: 261–264.

Venyaminov SY, Kalnin NN. **1990**. Quantitative IR spectrophotometry of peptide compounds in water (H_2O) solutions. II. Amide absorption bands of polypeptides and fibrous proteins in α-, β-, and random coil conformations. *Biopolymers* **30**: 1259–1271.

Veste M, Freese D, Lebzien S. **2015**. Effects of nitrogen and phosphate fertilization on leaf nutrient content, photosynthesis, and growth of the novel bioenergy crop Fallopia sachalinensis cv. 'Igniscum Candy' Changes in soil phosphorus under different fertilisation treatments in chernozem soils, Germany View project FarmImpact-Development of sustainable water and energy solutions for farms in South Africa View project. *Article in Journal of Applied Botany and Food Quality* **88(1)**.

Vohland M, Ludwig B, Seidel M, Hutengs C. **2022**. Quantification of soil organic carbon at regional scale: Benefits of fusing vis-NIR and MIR diffuse reflectance data are greater for in situ than for laboratory-based modelling approaches. *Geoderma* **405**: 115426.

Vohra K, Marais EA, Bloss WJ, Schwartz J, Mickley LJ, Van Damme M, Clarisse L, Coheur PF. **2022**. Rapid rise in premature mortality due to anthropogenic air pollution in fast-growing tropical cities from 2005 to 2018. *Science Advances* **8**: 4435.

de Vries FT, Manning P, Tallowin JRB, Mortimer SR, Pilgrim ES, Harrison KA, Hobbs PJ, Quirk H, Shipley B, Cornelissen JHC. **2012**. Abiotic drivers and plant traits explain landscape-scale patterns in soil microbial communities. *Ecology letters* **15**: 1230–1239.

Vyankatrao N, Arts NT, Commerce VS. 2014. Effect of drying methods on nutritional value of some vegetables. *Proceeding of the National Conference on Conservation of Natural Resources & Biodiversity for Sustainable Development* **6**: 72-79.

Wadoux AMJC, Minasny B, McBratney AB. **2020**. Machine learning for digital soil mapping: Applications, challenges and suggested solutions. *Earth-Science Reviews* **210**: 103359.

Wagner H, Liu Z, Langner U, Stehfest K, Wilhelm C. **2010**. The use of FTIR spectroscopy to assess quantitative changes in the biochemical composition of microalgae. *Journal of Biophotonics* **3**: 557–566.

Waldron KW, Faulds CB. **2007**. Cell Wall Polysaccharides: Composition and Structure. *Comprehensive Glycoscience: From Chemistry to Systems Biology* **1**–**4**: 181–201.

Walls RL. **2010**. Hybridization and Plasticity Contribute to Divergence Among Coastal and Wetland Populations of Invasive Hybrid Japanese Knotweed *s.l.* {*Fallopia* spp.). *Estuaries and Coasts* **33**: 902–918.

Walther GR, Roques A, Hulme PE, Sykes MT, Pyšek P, Kühn I, Zobel M, Bacher S, Botta-Dukát Z, Bugmann H, *et al.* 2009. Alien species in a warmer world: risks and opportunities. *Trends in Ecology and Evolution* **24**: 686–693.

Wang CY, Cheng SH, Kao CH. **1982**. Senescence of Rice Leaves VII. PROLINE ACCUMULATION IN SENESCING EXCISED LEAVES. *Plant Physiology* **69**: 1348–1349.

Wang Q, He H, Li B, Lin H, Zhang Y, Zhang J, Wang Z. **2017**. UV–Vis and ATR–FTIR spectroscopic investigations of postmortem interval based on the changes in rabbit plasma. *PLoS ONE* **12(7):** e0182161.

Wang C, Liu Y, Li S-S, Han G-Z. **2015**. Insights into the Origin and Evolution of the Plant Hormone Signaling Machinery. *Plant Physiology* **167**: 872–886.

Warren LM. **2019**. Is Japanese Knotweed inherently damaging? *Network Rail Infrastructure Ltd v Williams and Waistell* [2018] EWCA Civ 1514. *Environmental Law Review* **21**: 226–234.

Van De Weert M, Haris PI, Hennink WE, Crommelin DJA. **2001**. Fourier transform infrared spectrometric analysis of protein conformation: Effect of sampling method and stress factors. *Analytical Biochemistry* **297**: 160–169.

van der Weijde T, Huxley LM, Hawkins S, Sembiring EH, Farrar K, Dolstra O, Visser RGF, Trindade LM. **2017**. Impact of drought stress on growth and quality of miscanthus for biofuel production. *GCB Bioenergy* **9**: 770–782.

Wickham H. **2009**. ggplot2: elegant graphics for data analysis. *Springer-Verlag New York*.

Willett W, Rockström J, Loken B, Springmann M, Lang T, Vermeulen S, Garnett T, Tilman D, DeClerck F, Wood A, *et al.* **2019**. Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems. *The Lancet* **393**: 447–492.

Williams F, Eschen R, Harris A, Djeddour D, Pratt C, Shaw RS, Varia S, Lamontagne-Godwin J, Thomas SE, Murphy ST. **2010**. The economic cost of invasive non-native species on Great Britain. *CABI Proj No VM10066*: 1–99.

Williams P, Geladi P, Fox G, Manley M. **2009**. Maize kernel hardness classification by near infrared (NIR) hyperspectral imaging and multivariate data analysis. *Analytica Chimica Acta* **653**: 121–130.

Wilson SD, Tilman D. **1991**. Components of plant competition along an experimental gradient of nitrogen availability. *Ecology* **72**: 1050–1065.

Wolters H, Jürgens G. **2009**. Survival of the flexible: Hormonal growth control and adaptation in plant development. *Nature Reviews Genetics* **10**: 305–317.

Wu H-C, Bulgakov VP, Jinn T-L. **2018**. Pectin Methylesterases: Cell Wall Remodeling Proteins Are Required for Plant Response to Heat Stress. *Frontiers in Plant Science* **9**: 1612.

Yang Y, Benning C. **2018**. Functions of triacylglycerols during plant development and stress. *Current Opinion in Biotechnology* **49**: 191–198.

Yang J, Duan G, Li C, Liu L, Han G, Zhang Y, Wang C. **2019**. The Crosstalks Between Jasmonic Acid and Other Plant Hormone Signaling Highlight the Involvement of Jasmonic Acid as a Core Component in Plant Response to Biotic and Abiotic Stresses. *Frontiers in Plant Science* **10**: 1349.

Yang L, Wen KS, Ruan X, Zhao YX, Wei F, Wang Q. **2018**. Response of Plant Secondary Metabolites to Environmental Factors. *Molecules* **23**: 762.

Yang H, Yang S, Kong J, Dong A, Yu S. **2015**. Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy. *Nature Protocols* **10**: 382–396.

Yuan W, Pigliucci M, Richards CL. 2022. Rapid phenotypic differentiation and local adaptation in Japanese knotweed s.l. (*Reynoutria japonica and R. × bohemica*, Polygonaceae) invading novel habitats. *bioRxiv*: 2022.03.07.483296.

Zeghoud S, Rebiai A, Hemmami H, Seghir B Ben, Elboughdiri N, Ghareba S, Ghernaout D, Abbas N. **2021**. ATR–FTIR Spectroscopy, HPLC Chromatography, and Multivariate Analysis for Controlling Bee Pollen Quality in Some Algerian Regions. *ACS Omega* **6**: 4878–4887.

Zhang ZM, Chen S, Liang YZ, Liu ZX, Zhang QM, Ding LX, Ye F, Zhou H. **2010**. An intelligent background-correction algorithm for highly fluorescent samples in Raman spectroscopy. *Journal of Raman Spectroscopy* **41**: 659–669.

Zhang X, Liu F, Feng ZM, Jiang JS, Yang YN, Zhang PC. **2021**. Bioactive amides from Polygonum cuspidatum. *Journal of Asian Natural Products Research* **23(3)**: 228-34.

Zhang Y-Y, Parepa M, Fischer M, Bossdorf O. **2016**. Epigenetics of colonizing species? A study of Japanese knotweed in Central Europe. In: Barrett SCH, Colautti RI, Dlugosch KM, Rieseberg LH (Eds) Invasion Genetics. Chichester, UK: John Wiley & Sons, Ltd, 328–340.

Zheng W, Wang SY. **2001**. Antioxidant Activity and Phenolic Compounds in Selected Herbs. *Journal of Agricultural and Food Chemistry* **49**: 5165–5170.

Zhou C, Jiang W, Via BK, Fasina O, Han G. **2015**. Prediction of mixed hardwood lignin and carbohydrate content using ATR-FTIR and FT-NIR. *Carbohydrate Polymers* **121**: 336–341.

Zhu J, Agyekum AA, Kutsanedzie FYH, Li H, Chen Q, Ouyang Q, Jiang H. **2018**. Qualitative and quantitative analysis of chlorpyrifos residues in tea by surface-enhanced Raman spectroscopy (SERS) combined with chemometric models. *LWT* **97**: 760–769.

Zika PF, Jacobson AL. **2003**. An overlooked hybrid Japanese knotweed (*Polygonum cuspidatum × sachalinense*; Polygonaceae) in North America. *Rhodora* 143-152.

Zou T, Huang C, Wu P, Ge L, Xu Y. **2020**. Optimization of Artificial Light for Spinach Growth in Plant Factory Based on Orthogonal Test. *Plants* **9**: 490

Appendix A. PCA

PCA decomposition takes the form:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathbf{T}} + \mathbf{E}$$

(1)

where $\mathbf{X}$ represents the pre-processed spectral data, $\mathbf{T}$ is the PCA scores, $\mathbf{P}$ the loadings, and $\mathbf{E}$ the residuals (Bro & Smilde, 2014).

Appendix B.   Ward's Method

$$TD_{c_1 c_2} = \sum_{x \in c_1 \cup c_2} (x, \mu_{c_1 \cup c_2})^2$$

(2)

where c1 and c2 are separate clusters, $\in$ is an element of, and $\cup$ is a union (Lavrenko, 2015).

Appendix C.   PCA-LDA

PCA-LDA classification scores (L_{ik}) are calculated as follows (Morais & Lima, 2018):

$$\mathbf{L_{ik}} = (\mathbf{x_i} - \bar{\mathbf{x}}_\mathbf{k})^\mathbf{T} \mathbf{C}^{-1}_\mathbf{pooled}(\mathbf{x_i} - \bar{\mathbf{x}}_\mathbf{k}) - \mathbf{2\log_e \pi_k}$$

(3)

in which $\mathbf{x}_i$ is the vector containing the classification variables (i.e., PCA scores) for sample i; $\bar{\mathbf{x}}_k$ is the mean vector of class k; $\mathbf{C}_{\text{pooled}}$ is the pooled covariance matrix; and π_k is the prior probability of class k. $\mathbf{C}_{\text{pooled}}$ and π_k are calculated as follows:

$$\mathbf{C_{pooled}} = \frac{\mathbf{1}}{\mathbf{n}} \sum_{\mathbf{k=1}}^{\mathbf{K}} \mathbf{n_k C_k}$$

(4)

$$\mathbf{C_k} = \frac{\mathbf{1}}{\mathbf{n_k - 1}} \sum_{\mathbf{i=1}}^{\mathbf{n_k}} (\mathbf{x_i} - \bar{\mathbf{x}}_\mathbf{k})(\mathbf{x_i} - \bar{\mathbf{x}}_\mathbf{k})^\mathbf{T}$$

(5)

$$\mathbf{\pi_k} = \frac{\mathbf{n_k}}{\mathbf{n}}$$

(6)

where n is the total number of samples in the training set; K is the number of classes; n_k is the number of samples in class k; and $\mathbf{C}_k$ is the variance-covariance matrix of class k.

Appendix D. SVM

SVM takes the form of:

$$f(x) = \text{sign}\left(\sum_{i=1}^{N_{SV}} \alpha_i y_i K(\mathbf{x}_i, \mathbf{z}_j) + b\right)$$

(7)

where N_{SV} is the number of support vectors; α_i is the Lagrange multiplier; y_i is the class membership (± 1); $K(\mathbf{x}_i, \mathbf{z}_j)$ is the kernel function; and b is the bias parameter (Morais *et al.*, 2017). The RBF kernel function is calculated by:

$$K(\mathbf{x}_i, \mathbf{z}_j) = \exp\left(-\gamma \|\mathbf{x}_i - \mathbf{z}_j\|^2\right)$$

(8)

in which $\mathbf{x}_i$ and $\mathbf{z}_j$ are sample measurement vectors, and γ is the parameter that determines the RBF width (Cortes *et al.*, 1995).

Appendix E.   Accuracy, Sensitivity and Specificity

Accuracy (AC), sensitivity (SENS), and specificity (SPEC) are important factors in the evaluation of model success. These metrics are calculated as follows:

$$AC\% = \left(\frac{TP + TN}{TP + FP + TN + FN}\right) \times 100$$

(9)

$$SENS\% = \left(\frac{TP}{TP + FN}\right) \times 100$$

(10)

$$SPEC\% = \left(\frac{TN}{TN + FP}\right) \times 100$$

(11)

where TP stands for true positives, TN for true negatives, FP for false positives, and FN for false negatives.

Appendix F.   Crosstalk in Signalling Pathways

Brittany Heap*, Claire Holden*, Jane Taylor, Martin McAinsh

Lancaster Environnent Centre, Lancaster University.

Corresponding Author: Brittany Heap, b.heap2@lancaster.ac.uk

*Authors have contributed equally to this work.

Abstract

Plants have evolved complex signalling networks in order to respond to the plethora of environmental and developmental stimuli to which they are exposed. Within these networks, a diverse range of signals are integrated, allowing the plant to formulate an appropriate response to the prevailing conditions. An important feature of signalling pathways into networks is the potential for pathways to interact with each other thereby influencing the responses observed. This is known as crosstalk. Signalling networks act through a hub and spoke model where the hub is a key determinant of the pathways that will crosstalk with one another. Reactive oxygen species (ROS) are an example of such a signalling hub that play an intrinsic role in crosstalk within plant stress signalling networks. ROS have positive or negative effects, depending on their abundance within the cell. Here, we discuss crosstalk focusing on the role of ROS within the plant signalling network.

Key concepts
- Plants employ complex signalling networks in their responses to environmental and developmental stimuli, transducing a wide range of signals from the point of perception, typically via a receptor, to a diverse array of effectors.

- Signalling can occur via a range of intermediates (chemical, electrical, or hydraulic) including well characterised secondary messenger compounds.
- Crosstalk can occur between signalling pathways when one or more of these components interacts with each other.
- ROS are important signalling molecules, essential for plant cell homeostasis and the regulation of vital plant responses, such as reaction to abiotic and biotic stress, hormone signalling, and the regulation of growth.
- ROS can have a positive or negative effect within a cell, acting as either a toxic by-product of aerobic respiration within cells or essential signalling molecules through their role as an indispensable hub in plant signalling networks.

Keywords

Crosstalk, reactive oxygen species (ROS), signal transduction, abscisic acid (ABA), Ca^{2+} channels, guard cells, root tip.

Introduction

As sessile organisms, plants have evolved responses to stress that allow them to tolerate and survive adverse conditions, many of which are inducible and are activated only once a plant senses an environmental stimuli (See also: Abiotic Stress, DOI.org/10.1002/9780470015902.a0020087; Pathogen Resistance Signalling in Plants, DOI.org/10.1002/9780470015902.a0020119.pub2; Plant Stress Physiology, DOI.org/10.1002/9780470015902.a0001297.pub2, Flooding Stress in Plants, DOI.org/10.1002/9780470015902.a0001317.pub3; Temperature Stress in Plants, DOI.org/10.1002/9780470015902.a0001320.pub2; Plant Salinity Tolerance, DOI.org/10.1002/9780470015902.a0001300.pub3; Plant Light Stress, DOI.org/10.1002/9780470015902.a0001319.pub3). Frequently, these responses have an associated fitness cost and therefore plants have also evolved complex, targeted regulatory mechanisms to efficiently coordinate these responses to stress

in order to minimise this cost. These include multi-level signalling networks which are diverse in their forms and scales, both spatial and temporal. Signalling pathways comprise a number of fundamental elements (See also: Cell Signalling Mechanisms in Plants, DOI.org/10.1002/9780470015902.a0026507):

1. **Sensing:** A cell detects an exogenous signalling molecule. This frequently, but not exclusively, occurs via a ligand binding to a receptor protein on the cell surface or inside the cell.

2. **Signal transduction:** Binding of the signalling molecule to the receptor elicits a change in structure that initiates the process of transduction.

3. **Signal amplification and response:** Protein modulation and/or post-translational modifications trigger a specific cellular response.

Historically, signalling pathways in plants were described as linear sequences and treated as independent, parallel processes. However, this has been replaced by a molecular network-structured model with a much higher level of complexity in which crosstalk occurs between signalling pathways. Crosstalk is an integration of signals from more than one response input within a network that affects the same output (Vert and Chory, 2011). Once considered an unwanted transfer of signals between cellular communication channels, crosstalk is now recognised as an integral aspect of cell function, with intracellular signalling pathways frequently integrating and conferring benefits. In particular, as plants commonly face multiple stresses simultaneously, crosstalk may prepare a plant to respond quickly to a response by upregulating plant defence responses, a process known as priming.

Networks are defined by the number of interactions between each component, or node, in the network and form either random (also referred to as exponential), or scale-free networks (Vandereyken et al., 2018). Most biological networks are scale-free. The nodes in random networks all have the same number of connections, whilst in scale-free networks there are nodes that interact with many other nodes; these are referred to as hubs, signalling components which play a central role in a signalling network. Due to their high connectivity, hubs are critical within networks

such that removal or interference with a network hub has a marked effect on the functioning of the network and plant fitness (He and Zhang, 2006). Hub loss of function mutants have been shown to exhibit abnormal stress responses, highlighting the essential role they play in the detection and transduction of stress signals in plant stress response networks (Dietz et al., 2010). Signalling molecules, and the extent of their interaction, are still being discovered and can take many different forms such as molecular, chemical, electrical or hydraulic signals (Huber and Bauerle, 2016; Table E.1). For example, ROS are implicated as an important hub in plant stress signalling networks, allowing crosstalk between signalling pathways.

Table E.1. The nature of signal transduction pathways.

Class	Signal transduction pathway	Reference
Second messengers	Reactive oxygen species (ROS) Calcium (Ca^{2+}) Potassium (K^+) Reactive nitrogen species (RNS)	Gill and Tuteja (2010) Tuteja and Mahajan (2007) Amtmann et al. (2005) del Río et al. (2006)
Chemical	Volatile compounds Plant hormones Mitogen-activated protein kinase (MAPK)	Wu and Baldwin (2009) Jaillais and Chory (2010) Pitzschke et al. (2009)
Electro-chemical	Action potentials Slow wave potentials Wound potentials System potentials pH	Volkov et al. (2007) Stahlberg et al. (2006) Huber and Bauerle (2016) Zimmermann et al. (2009) Felle (2001)
Hydraulic	Changes in cell turgor upon water stress Changes in leaf thickness upon wounding Long distance signalling as a precursor to ABA signalling during water shortage	Buckley and Mott (2002) Malone (1994) Christmann et al. (2007)

ROS: A key hub for crosstalk in signalling

ROS form a group of signalling molecules which are common to many different pathways (Figure E.1) and are crucial regulators of fitness in living organisms (Apel and Hirt, 2004). ROS include superoxide (O_2^-), hydrogen peroxide (H_2O_2), and singlet oxygen (O_2^*) and are produced as part of normal aerobic metabolic processes inside plant cells. Only a relatively small change in their equilibrium is required to induce a cell response. As such, ROS are particularly interesting molecules in that they can behave either as toxins, resulting in oxidative stress, or can be involved in signalling in response to stress (Halliwell, 2006) (See also: Oxidative stress, DOI.org/10.1002/9780470015902.a0001376.pub2). Specific mechanisms have evolved for maintaining ROS homeostasis, making a group of potentially toxic molecules into one of evolutionary benefit to the plant (See also: Oxidative Stress and Redox Signalling in Plants, DOI.org/10.1002/9780470015902.a0026508). ROS are produced in response to many abiotic stresses including drought, extreme temperatures, high salinity, high light, osmotic stress, metal toxicity, and the presence of xenobiotics such as herbicides and ozone (O_3) (See also: Whole-Plant Physiological Responses to Water-Deficit Stress, DOI.org/10.1002/9780470015902.a0001298.pub3; Plant Salt Stress, DOI.org/https://doi.org/10.1002/9780470015902.a0001300.pub3; Temperature Stress in Plants, DOI.org/10.1002/9780470015902.a0001320.pub2; Ozone and Reactive Oxygen Species, DOI.org/10.1002/9780470015902.a0001299.pub3). They are also produced in response to biotic stresses during an oxidative burst phase and have the potential to accumulate within cells to cytotoxic levels, causing cell damage through lipid and protein oxidation and nucleic acid degradation (Mühlenbock et al., 2007) (See also, Free Radicals and Other Reactive Species in Disease, DOI: https://doi.org/10.1002/9780470015902.a0002269.pub3).
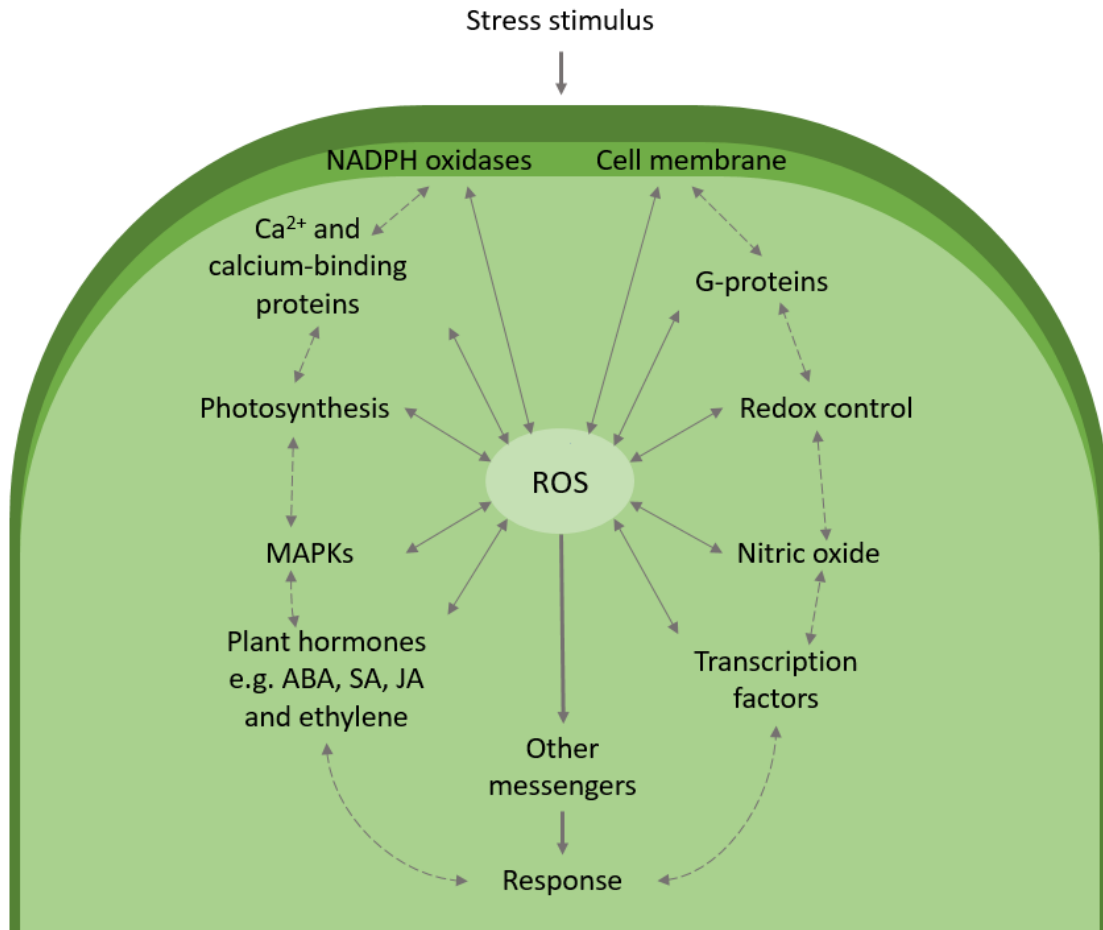
Figure E.1: ROS as a signalling hub common to different pathways, allowing points of crosstalk between signalling events. ROS work upstream and downstream of other signalling components, such as membranes, NADPH oxidases, G-proteins, Ca^{2+}, MAPKs, transcription factors, and plant hormones; salicylic acid (SA), jasmonic acid (JA), abscisic acid (ABA) and ethylene. These components interrelate with processes such as redox homeostasis and photosynthesis. Solid arrows indicate direct ROS interactions with other signalling components, and dashed arrows for expected indirect interactions. Sewelam et al. (2016). Licensed under CC BY.

ROS levels within cells are tightly regulated in order to prevent oxidative damage. ROS are rapidly detoxified by enzymatic and non-enzymatic mechanisms, located in defined subcellular compartments, in order to maintain homeostasis within a cell (Halliwell, 2006). Oxidative stress triggers the production of antioxidants such as the lipid-soluble, carotenoids and tocopherols, which are associated with the hydrophobic cell and organelle membranes, and the water-soluble anthocyanins,

ascorbate and glutathione within the cytosol, chloroplasts, and other subcellular compartments. ROS-scavenging enzymes including superoxide dismutase (SOD), ascorbate peroxidase (APX), and catalase (CAT), Glutathione catalase and Glutathione S-transferases (GSTs), thioredoxins and peroxiredoxins are also induced (Apel and Hirt, 2004; Perez and Brown, 2014). Specific isozymes of APX and SOD localize to the cytosol, chloroplasts, apoplast, peroxisomes and mitochondria, whereas CAT localizes exclusively to peroxisomes (Perez and Brown, 2014) (See also: Antioxidants, DOI.org/10.1002/9780470015902.a0023204). However, whilst plants have sophisticated mechanisms to combat stress-induced increases in ROS there are also circumstances when plants generate ROS as signalling molecules to control various processes including pathogen defence, programmed cell death, and stomatal behaviour (Apel and Hirt, 2004) (See also: Pathogen Resistance Signalling in Plants, DOI.org/10.1002/9780470015902.a0020119.pub2; Plant Programmed Cell Death, DOI.org/10.1002/9780470015902.a0001689.pub3; Stomata, DOI.org/10.1002/9780470015902.a0002075.pub3, DOI.org/10.1002/9780470015902.a0026526).

ROS production is modulated by a cooperative system controlling the redox state of the cell. Whilst ROS acts as a key secondary messenger in stress signalling, under conditions of severe or prolonged stress, ROS production exceeds the ROS scavenging capacity of antioxidants and ROS become cytotoxic, causing oxidative damage (Das and Roychoudhury, 2014). The cellular redox state provides several points of crosstalk between the ROS signalling network and the hormone auxin and can affect the cell's response. Plant hormones are critical to the coordination of many physiological and developmental process and play an intrinsic role in plant responses to stress (See also: Plant Hormones, DOI.org/10.1002/9780470015902.a0002091.pub2). There is frequent crosstalk between hormone signalling pathways that deliver a range of outcomes, from regulating another hormone's abundance, transport, or signalling efficacy, to modulating the activity of a shared signalling node or hub. An increase in ROS is often a feature of hormone signalling. For example, in abscisic acid (ABA) signalling

ROS acts as an important signalling intermediate (See also: Abscisic Acid (ABA), DOI.org/10.1002/9780470015902.a0020088.pub3). During wounding of plant tissue, jasmonic acid accumulates in the damaged tissue, but also in undamaged leaves. This is a consequence of ROS acting as a long distance signal, priming plant defences against possible attack or damage (Gilroy et al., 2016). Two important regulators of the redox homeostasis, reduced nicotinamide adenine dinucleotide phosphate (NADPH)-thiorexin reductase and glutathione (GSH), alter auxin transport and metabolism. Auxins can indirectly modulate ROS homeostasis by affecting the stability of DELLA proteins, or directly by inducing ROS detoxification enzymes, such as glutathione (GSH) S-transferases and quinine reductases (Tognetti et al., 2012). Another molecule whose primary task is ROS homeostasis, the hydrogen peroxide-scavenging enzyme APX6, mediates cross talk between ROS, abscisic acid, and auxin (Chen et al., 2014).

Phosphatases and heat shock transcription factors (HSFs) directly modify cellular redox state (Apel and Hirt, 2004) (See also: Oxidative Stress and Redox Signalling in Plants, DOI.org/10.1002/9780470015902.a0026508). HSFs play important roles in sensing environmental stresses including high temperature stress which has been shown to induce ROS in plants. Crosstalk between stress-response signalling pathways via HSFs is supported by the increased resistance to both drought and bacterial infection following the overexpression of the HSF-A1b, which is independent of hormone signalling but dependent on H_2O_2 signalling (Bechtold et al., 2013).

The range of abiotic and biotic stresses that stimulate ROS production, and hence the potential for crosstalk within stress signalling networks when stresses are perceived concurrently, raises the question of how specificity is maintained within ROS signalling. Specificity may occur through the generation of a unique ROS signature, that is unique to spatiotemporal changes in ROS, which is characteristic of the stress combination. The ROS signature varies with cell type, developmental stage and stress level within each cellular compartment, as each compartment

controls its own ROS homeostasis. Decoding these signatures via different ROS sensors can create a stress-specific signal that will tailor the acclimation response to the type of stress affecting the plant (Choudhury et al., 2017). The major ROS-producing sites during abiotic stress are chloroplasts, mitochondria, peroxisome and apoplast (Zorov et al., 2014; Choudhury et al., 2017). Metabolic processes occurring in the chloroplasts and mitochondria act as a continual source of ROS which are produced as by-products of electron leakage. Some forms of ROS, such as H_2O_2 and O_2^{-}, can diffuse from these organelles to the nucleus to act as a form of communication, an effect also thought to occur with the reactive nitrogen species (RNS) (See also: Nitric Oxide Signalling in Plants, DOI.org/10.1002/9780470015902.a0020109.pub2). Consequently, ROS are frequently found alongside other major signalling molecules such as cytosolic calcium ions, as hubs in the signalling networks by which plants respond to hormones and environmental stresses.

Crosstalk between calcium and ROS signalling

Ca^{2+} regulates many signalling pathways involved in growth, development and stress tolerance (See also: Calcium Signalling in Plants, DOI.org/10.1002/9780470015902.a0023722). The concentration of Ca^{2+} in the cytosol is maintained by balancing Ca^{2+} influx and efflux, the latter being achieved by P-type Ca^{2+}ATPases or antiporters (Tuteja and Mahajan, 2007) . Overexpression of members of the P-type 2B Ca^{2+}ATPase family has been shown to confer tolerance to salinity and drought stress and to increase cadmium stress tolerance (Huda et al., 2013) reducing ROS accumulation and oxidative stress, respectively, and enhancing the expression of stress-responsive genes, providing clear evidence of crosstalk between Ca^{2+} and ROS signalling. Crosstalk between ROS, Ca^{2+} and NO signalling is also observed in plant hormone signalling networks including ABA signalling (Baxter et al., 2014). Additionally, It has been suggested that crosstalk between Ca^{2+} and other signalling pathways, such as ROS, could contribute to the generation of specificity within signalling networks (McAinsh and Hetherington, 1998). It is also

emerging that signals occuring concurrently provides a potential for signal integration that ensures coordination of many distinct environmental responses with each other and with plant development (Kudla et al., 2018).

Signal amplification is a property common to signal transduction pathways. In ROS signalling, a small family of membrane-bound proteins known as respiratory burst oxidase homologues (Rbohs) play a major role in the creation and amplification of ROS signals through the propagation of a ROS wave. Importantly, Rbohs have been shown to act as important mediators in ROS interactions with Ca^{2+} gradients and hormones such as ABA enabling signalling crosstalk (see Figure E.1 for a summary of crosstalk between ROS and signalling pathways). Processes affected include Ca^{2+} -dependent tip-growth (Gilroy et al., 2016), and ABA-induced stomatal closing (Kwak et al., 2003; Klusener et al., 2002). Interestingly, different isoforms of Rboh may contribute to the specificity of a plant response through the production of stimulus-specific ROS signatures. For example, RbohD and RbohF function synergistically in signalling cascades to regulate stomatal closure, seed germination, root elongation, and Na^+/K^+ homeostasis under salt stress (Kwak et al., 2003). Abiotic stress induces ROS production via the Ca^{2+}-dependent activation of Rbohs in the apoplast. This process is regulated by interaction with Ca^{2+} through Ca^{2+}-binding EF-hand motifs in the N-terminus (Baxter et al., 2014) providing a point of crosstalk between Ca^{2+} and ROS-mediated signalling (See also: Calcium Signalling in Plants, DOI.org/10.1002/9780470015902.a0023722). Accumulation of the Rboh-produced ROS is sensed by neighbouring cells, activating ROS and/or Ca^{2+}-activated Ca^{2+} channels in the plasma membrane. The resultant Ca^{2+} fluxes activate ROS production via Rboh function in the neighbouring cell, triggering the ROS wave. ROS and electrochemical waves are connected through RbohD and glutamate receptor (GLR) function, and Ca^{2+} and electric waves are linked via channels such as GLRs (Figure E.2). The ROS wave is auto-propagated throughout the plant resulting in the activation of systemic acquired acclimation (SAA) or systemic aquired resistance (SAR) in response to abiotic stress and biotic stress, including microbial elicitors or

artificial chemical stimuli, respectively (Gilroy et al., 2016). Wave propagation within SAA and SAR provides a further point of signalling crosstalk though their regulation of each other.
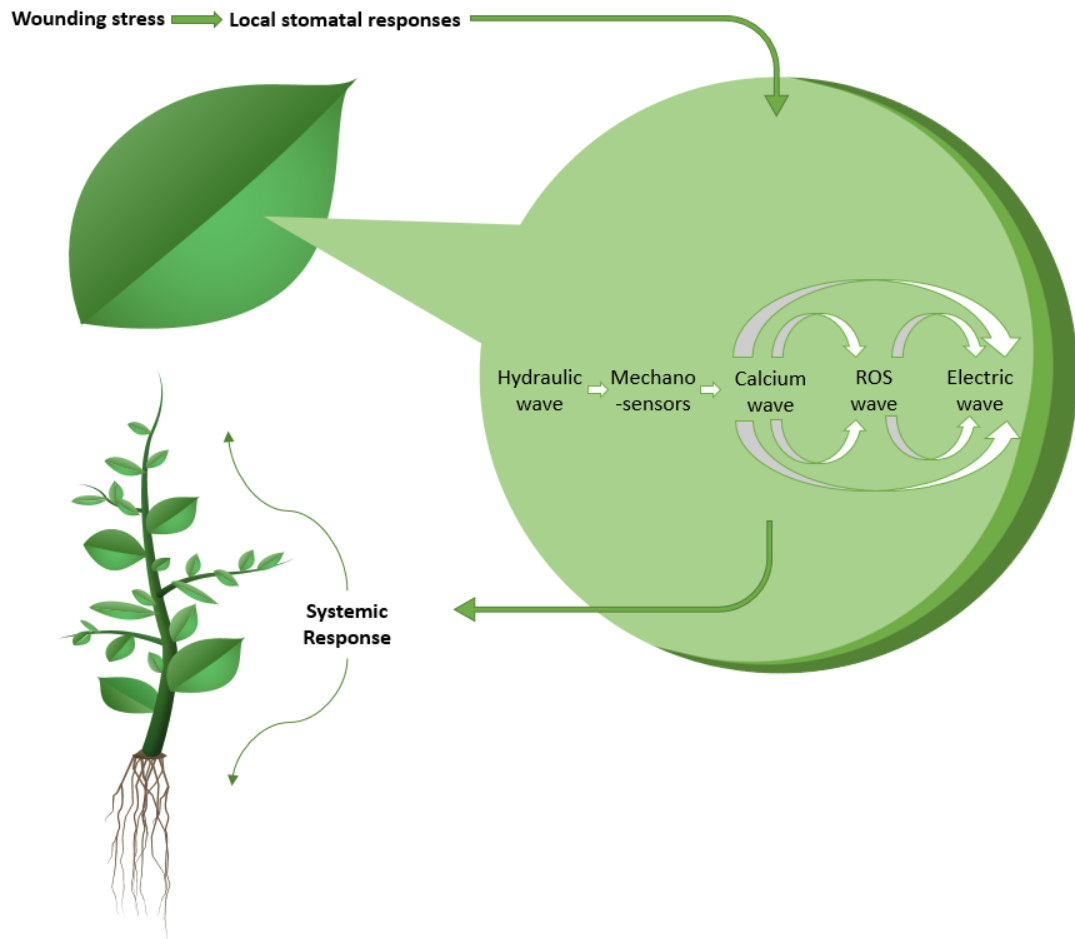
Figure E.2: Signal propagation from local stimuli by a ROS wave throughout the plant resulting in the activation of systemic acquired acclimation (SAA) to reach distant tissues and activate a response. Integration of the different waves that mediate rapid systemic signaling during SAA. Local stimuli are shown to trigger the ROS/calcium/electric wave, as well as a hydraulic wave that in turn triggers the calcium wave via mechano-sensors. Reproduced with permission from Gilroy et al. (2016) © 2016 American Society of Plant Biologists.

Given the complexity of plant signalling networks and the potential for crosstalk within them, it is essential that plants possess mechanisms for encoding specificity within stimulus-induced signals in order to prevent the 'unwanted' transfer of information and to ensure the generation of an appropriate response to a given stimulus. In plant Ca^{2+} signalling, it has been suggested that this occurs through the generation of a stimulus-specific 'Ca^{2+} signature' (McAinsh and Hetherington, 1998) in which signalling information is encoded in the varying spatio-temporal dynamics of plant Ca^{2+} signals (McAinsh and Pittman, 2009). It is hypothesized that the information encoded in these Ca^{2+} signatures is subsequently decoded and relayed to downstream responders, including Ca^{2+}-dependent protein kinases (CDPKs) and transcription factors, by an array of Ca^{2+}-binding sensor proteins such as the binding protein calmodulin (CaM) (See also: See also: Calcium Signalling in Plants, DOI.org/10.1002/9780470015902.a0023722).

Crosstalk in guard cell closure signalling

Stomatal closure has been widely used to study plant stress response signalling. Various stimuli including light, carbon dioxide (CO_2) concentration, humidity and plant hormones including ABA, ethylene (ET) and methyl jasmonate (MeJA) act as effectors for stomatal closure (See also: Stomata, DOI.org/10.1002/9780470015902.a0002075.pub3, DOI.org/10.1002/9780470015902.a0026526). Stomatal closure requires the coordinated action of ion transport in both the vacuolar and plasma membranes in order to bring about a reduction in guard cell turgor. Guard cells therefore possess an exquisite signalling network for integrating the complex signalling information from the diverse range of stimuli to which they are exposed (see Kim et al. (2010)).

Although many of the signalling intermediates involved in guard cell closure are known, their mechanisms of crosstalk are still being discovered (Li et al., 2017; Munemasa et al., 2019) and several mathematical models have been developed in an attempt to capture these interactions. Ethylene signalling has been shown to

inhibit jasmonate signalling as well as ABA signalling in guard cells via suppression of ROS production and S-type anion channel activity (Munemasa et al., 2019). Despite these advances, there are still many connections yet to be elucidated, for example the crosstalk between ABA and MeJA signalling cascades leading to guard cell closure. These are known to share some common components, however, mutant experiments show that there are also some constituents unique to each pathway (Khokon et al., 2015).

ROS and nitric oxide (NO) (See also: Oxidative Stress and Redox Signalling in Plants, DOI.org/10.1002/9780470015902.a0026508; Nitric Oxide Signalling in Plants, DOI.org/10.1002/9780470015902.a0020109.pub2) are central components of the ABA signalling network that regulate stomatal movement and may provide an indispensable point of crosstalk with Ca^{2+} in the ABA signalling cascade in guard cells (Li *et al.* (2017). ABA activates the synthesis of H_2O_2 in guard cells via the SnRK2 kinase OST1 (OPEN STOMATA1)-mediated phosphorylation of a plasma-membrane-associated NADPH oxidase RbohF (Sirichandra *et al.*, 2009) activating ROS, ABA signalling (creating a feed-forward loop), and ABA-dependent Ca^{2+} channels in the plasma membrane (Wang Y *et al.*, 2013) to control stomatal closure. H_2O_2-induces ABA accumulation but does not induce salicylic acid (SA) accumulation. This suggests that the implied crosstalk with ABA is achieved through ABA accumulation, while the crosstalk with SA is thought to involve the benzoic acid pathway (the precursor of SA), as activation of benzoic acid 2-hydroxylase is H_2O_2 dependent (Hieno *et al.*, 2019). ABA-induced ROS production and ABA induced cytosolic Ca^{2+} increase, whilst exogenous H_2O_2 application restores Ca^{2+} channel activation and stomatal closure. This supports a central role for ROS production in ABA signal transduction (Kwak *et al.*, 2003). Additionally, mutants lacking phospholipase D alpha (PLD) fail to activate ROS production in response to ABA whilst phosphatidic acid (PA), the product of PLD, activates of NADPH oxidase in these mutants. This provides evidence of crosstalk between ABA-ROS-Ca^{2+} signalling with another signalling mechanism in which PA acts as a second

messenger mediating stomatal closure. Importantly, interaction of PA with the 2C protein phosphatase ABI1 (ABA INSENSITIVE 1), inhibition of which by ABA leads to the activation of OST1 and RbohF, appeared necessary for stomatal closure suggesting that the complementary signalling mechanisms interact (Zhang et al., 2009). Interestingly, ROS levels also increase in guard cells in response to the plant defence elicitors chitosan and oligo-GalUA resulting in stomatal closure (Klusener *et al*., 2002). Additionally, crosstalk is thought to exist between ABA, ROS and CO_2 (Qi *et al.*, 2018).

The role of NO in guard cell signalling is complex although there is evidence for its involvement in ABA-induced stomatal closure. In plants, NO is most likely produced through the action of the enzyme nitrate reductase (NR) with NR1 appearing to be the most important in the control of stomatal closure (Wilson *et al.*, 2009). Importantly, ABA-induced stomatal closure is enhanced by the NO donor sodium nitroprusside (SNP) (García-Mata and Lamattina, 2002) but inhibited by NO with the 2-Phenyl-4,4,5,5-tetramethylimidazoline-1-oxyl 3-oxide (PTIO) (See also: Nitric Oxide Signalling in Plants, DOI.org/10.1002/9780470015902.a0020109.pub2). Endogenous NO levels increase in response to ABA-induced reactive oxygen production although NO does not regulate directly ABA-mediated ion transport suggesting that the action of NO may be mediated at least in part by changes in cytosolic Ca^{2+} contributing further to the crosstalk within the guard cell signalling network.

Crosstalk in root growth responses

Root hairs are single cells that absorb water and nutrients from the soil, their growth is regulated by a network of cellular regulators that combine to focus the activity of the growth machinery on polar cell expansion (See also: DOI.org/ https://doi.org/10.1002/9780470015902.a0023746). Polar growth of root hairs is a complex process and questions remain about the molecular mechanisms involved. However, it is established that their morphology is strongly regulated by

environmental signals such as: nutrient availability, water, and CO_2 levels; along with hormonal signalling within a plant (notably; ethylene, auxin and cytokinin).

Polar growth of the root hair is maintained by feedback loops consisting of ROS, Ca^{2+} and pH crosstalk (Figure E.3). During root hair cell expansion there are high levels of cytosolic Ca^{2+} in the tip, triggering production of apoplastic ROS in a reaction that is catalsyed and controlled by NADPH oxidases (Mangano *et al.,* 2016). In turn, high levels of ROS elevate cytosolic Ca^{2+}, creating a feed-forward loop (mechanism unknown). To control concentration gradients and signal transduction, Ca^{2+} ions can be transferred from outside of the cell or stored in cell wall polymers until mobilisation is required.

The rate of Ca^{2+} influx determines the rate of cell elongation, whilst the Ca^{2+} gradient directs the polar secretion necessary for new cell wall material, along with transportation of organelles and biochemical activity related to enzymes (Gilroy *et al.,* 2016). Interestingly, cytosolic ROS, cytosolic pH and apoplastic pH oscillate during root hair cell growth. Due to the nature of oscillations of ROS and Ca^{2+}, it is thought that polar cell growth is preceded (and possibly repressed) by high levels of cytosolic Ca^{2+}, followed by high ROS concentrations and a higher pH. Changes in pH during root hair growth are regulated by membrane H^+-ATPases in addition to cation (H^+)/anion (OH^-)-permeable channels and antiporters (Ca^{2+}/H^+ exchangers) (Mangano *et al.,* 2016). The AHA gene family regulates apoplastic pH, which manipulates enzymes responsible for modification of the cell wall during expansion, AHA2 has particularly high expression in root hairs during root growth. It is thought that AHA-induced acidification within the apoplast releases Ca^{2+}. Whilst it is known that ROS homeostasis is closely connected with Ca^{2+} activity and pH during cell expansion for root hair growth, definitive links are still lacking regarding the mechanisms involved.
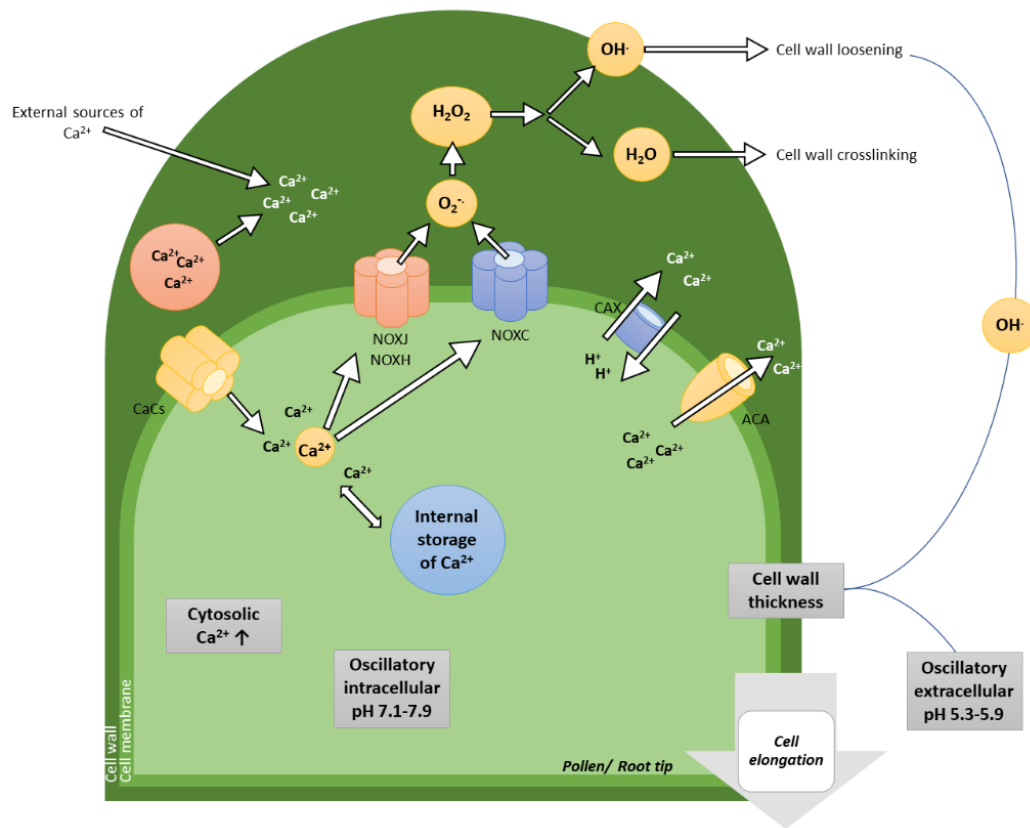
Figure E.3: Depiction of the role of Ca²⁺, pH and ROS as key signalling hubs for elongation of root/pollen cells. When external sources of Ca²⁺ are taken up by the cell through calcium channels (CaCs), cytosolic Ca²⁺ homeostasis is maintained by the transportation of Ca²⁺ back to the apoplast via autoinhibited Ca²⁺-ATPases (ACAs). Ca²⁺ is additionally translocated back to the apoplast by the H⁺/Ca²⁺ antiporter activity of calcium exchangers (CAX), which also import protons into the cytoplasm. An increase in cytosolic Ca²⁺ activates apoplastic ROS production, through activation of NADPH oxidases (NOX-J, H and C). Apoplastic superoxide is dismutated by superoxide dismutase (SOD) to H_2O_2. The reactive hydroxyl radical (•OH) is produced through the reaction of apoplastic H_2O_2 and oxygen O_2. •OH then catalyses nonenzymatic cleavage of polysaccharides, thereby allowing tip growth. Singh *et al.* (2016). Licensed under CC BY.

In legumes, ROS has been shown to have a symbiotic interaction with Nod factors (NFs), signalling molecules produced by rhizobia during the initation of nodules on roots. Legume roots excrete flavonoids which induce expression of nodulation genes in bacteria which synthesise NFs. These NFs are perceived by the root and initiate ion fluxes, cytosolic alkination and Ca^{2+} oscillations, which leads to nodule formation through bacterial invasion (Cardenas and Quinto, 2008).

The Arabidopsis mutant *root hair defective 2 (rhd2)*, which lacks NADPH oxidase RbohC, has poorly developed root hairs and lack a ROS gradient at the root tip. This suggests that RHD2 synthesis of ROS is required for root hair growth (Jones MA *et al.,* 2007); ROS accumulation at the root tip is also inhibited by the NADPH oxidase inhibitor diphenylene iodonium chloride (DPI). Since NADPH oxidase contains Ca^{2+} binding EF-hand domains, the ROS gradient at the root tip provides an important point of crosstalk in root hair signalling. However, the exact role of Ca^{2+} in root hair development requires further clarification (Cardenas and Quinto, 2008).

Crosstalk: friend or foe?

As discussed above, signalling intermediates and pathways are organized into complex networks and therefore rarely operate independently. Although crosstalk within such networks is an issue in terms of maintaining stimulus specificity it also has the potential to confer benefits to the plant, for example in terms of acclimation responses, particularly in cases where two stimuli occur simultaneously and/or initiate the same response. ROS are crucial mediators of plant defence systems against both biotic and abiotic stresses. ROS acts as a signalling hub, they are therefore an intrinsic part of crosstalk interactions that occur as plants respond to external stimuli and can be both beneficial or harmful to plant health, with plants maintaining ROS homeostasis for optimum growth and development (See also: Oxidative Stress and Redox Signalling in Plants, DOI.org/10.1002/9780470015902.a0026508; Nitric Oxide Signalling in Plants, DOI.org/10.1002/9780470015902.a0020109.pub2).

Due to this interconnected nature of signalling pathways, crosstalk research is shifting to incorporate systems biology approaches to address these issues. Crosstalk can be predicted using reverse engineering and network exploration tools to analyse genetic data on a larger scale (Friedel *et al.,* 2012). The traditional small-scale 'wet laboratory' approach, which employs the application of specific stimulae to monitor downstream responses, is complemented by bioinformatics which can give an overview of the system as a whole.

Interestingly, crosstalk between stress pathways may also offer a unique opportunity for improving agricultural sustainability though the exploitation of ROS signals. ROS signalling pathways might be exploited to induce a fatal oxidative stress in plants, or to generate a controlled oxidative stress that can prime and upregulate plant defence responses. Hormesis is the phenomenon whereby the dose-response to a chemical is biphasic, with a low dose inducing a stimulatory or beneficial effect and a high dose inducing an inhibitory or toxic effect; herbicide-induced hormesis is acknowledged for a range of herbicides (Belz *et al.,* 2011). To date, lack of mechanistic understanding has posed a lack of predictability needed for commercial use, though recent developments in the field have improved our knowledge. Therefore, low-dose herbicide-induced priming of stress response through signalling crosstalk offers a potential for the development of innovative new crop protection strategies for reducing the levels of crop protection products necessary for weed-control with ROS having the ability to act as both a priming agent and herbicide.

Further reading

Lui, J, Moore, S, Chen C, & Lindsey K. 2017. Crosstalk Complexities between Auxin, Cytokinin, and Ethylene in Arabidopsis Root Development: From Experiments to Systems Modeling, and Back Again. *Molecular Plant,* 10, 1480-1496.

Glossary

Crosstalk - The interaction of shared signalling components from more than one signal transduction pathway, allowing information to pass between pathways.

Reactive oxygen species – A group of molecules containing oxygen which are chemically reactive, including peroxides, superoxide, the hydroxyl radical, and singlet oxygen.

Signal transduction - A series of molecular events which transmit a signal from the exterior to the interior of a cell.

Abscisic acid (ABA) - A plant hormone which acts as a growth inhibitor and is involved in numerous processes such as guard cell closure, bud dormancy, seed maturation, and stress responses.

Calcium channel - An ion channel which is selectively permeable to calcium.

Guard cell - A specialised plant cell found in pairs in the leaf epidermis. The pair of guard cells form an opening called the stomatal pore which can be opened and closed by tutor pressure to control gas exchange and transpiration.

Root tip - The apical growing part of a root.

References

Amtmann A, Hammond JP, Armengaud P, White PJ. 2005. Nutrient Sensing and Signalling in Plants: Potassium and Phosphorus. *Advances in Botanical Research* **43**: 209-257.

Apel K, Hirt H. 2004. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu Rev Plant Biol* **55**: 373-99.

Baxter A, Mittler R, Suzuki N. 2014. ROS as key players in plant stress signalling. *J Exp Bot* **65**: 1229-40.

Bechtold U, Albihlal WS, Lawson T, Fryer MJ, Sparrow PA, Richard F, Persad R, Bowden L, Hickman R, Martin C, Beynon JL. 2013. Arabidopsis HEAT SHOCK TRANSCRIPTION FACTORA1b overexpression enhances water productivity, resistance to drought, and infection. *Journal of experimental botany* **64(11)**: 3467-3481.

Belz RG, Cedergreen N, Duke SO. 2011. "Herbicide hormesis–can it be useful in crop production?" *Weed Research* **51(4)**: 321-332.

Buckley TN, Mott KA. 2002. Stomatal Water Relations and the Control of Hydraulic Supply and Demand. In: Esser, K., Lüttge, U., Beyschlag, W. & Hellwig, F. (eds.) *Progress in Botany: Genetics. Physiology. Ecology.* Berlin, Heidelberg: Springer Berlin Heidelberg.

Cardenas L, Quinto C. 2008. Reactive oxygen species (ROS) as early signals in root hair cells responding to rhizobial nodulation factors. *Plant Signal Behavior* **3**: 1101-2.

Chen C, Letnik I, Hacham Y, Dobrev P, Ben-Daniel B-H, Vanková R, Amir R, Miller G. 2014. ASCORBATE PEROXIDASE6 Protects Arabidopsis Desiccating and Germinating Seeds from Stress and Mediates Cross Talk between Reactive Oxygen Species, Abscisic Acid, and Auxin. *Plant Physiology* **166**: 370-383.

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

Christmann A, Weiler EW, Steudle E, Grill E. 2007. A hydraulic signal in root-to-shoot signalling of water shortage. *Plant J* **52**: 167-74.

Das K, Roychoudhury A. 2014. Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. *Frontiers in Environmental Science* **2**: 53.

Del Río LA, Sandalio LM, Corpas FJ, Palma JM, Barroso JB. 2006. Reactive Oxygen Species and Reactive Nitrogen Species in Peroxisomes. Production, Scavenging, and Role in Cell Signaling. *Plant Physiology* **141**: 330-335.

Dietz KJ, Jacquot JP, Harris G. 2010. Hubs and bottlenecks in plant molecular signalling networks. *New Phytologist* **188**: 919-38.

Felle H. 2001. pH: Signal and Messenger in Plant Cells. *Plant Biology* **3(6)**: 577-591.

Friedel S, Usadel B, Von Wirén N, Sreenivasulu N. 2012. Reverse Engineering: A Key Component of Systems Biology to Unravel Global Abiotic Stress Cross-Talk. *Frontiers in Plant Science* **3**: 294

García-Mata C, Lamattina L. 2002. Nitric Oxide and Abscisic Acid Cross Talk in Guard Cells. *Plant Physiology* **128**: 790-792.

Gill SS, Tuteja N. 2010. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiology and Biochemistry* **48**: 909-930.

Gilroy S, Białasek M, Suzuki N, Górecka M, Devireddy AR, Karpiński S, Mittler R. 2016. ROS, Calcium, and Electric Signals: Key Mediators of Rapid Systemic Signaling in Plants. *Plant Physiology* **171**: 1606.

Halliwell B. 2006. Reactive species and antioxidants. Redox biology is a fundamental theme of aerobic life. *Plant physiology* **141**: 312-322.

He X, Zhang J. 2006. Why Do Hubs Tend to Be Essential in Protein Networks? *PLOS Genetics* **2(6)**: e88.

Hieno A, Naznin HA, Inaba-Hasegawa K, Yokogawa T, Hayami N, Nomoto M, Tada Y, Yokogawa T, Higuchi-Takeuchi M, Hanada K, Matsui M. 2019. Transcriptome analysis and identification of a transcriptional regulatory network in the response to H2O2. *Plant physiology*, **180(3)**: 1629-1646.

Huber AE, Bauerle TL. 2016. Long-distance plant signaling pathways in response to multiple stressors: the gap in knowledge. *Journal of Experimental Botany* **67**: 2063-2079.

Huda KM, Banu MS, Garg B, Tula S, Tuteja R, Tuteja N. 2013. OsACA6, a P-type IIB Ca(2)(+) ATPase promotes salinity and drought stress tolerance in tobacco by ROS scavenging and enhancing the expression of stress-responsive genes. *Plant J* **76**: 997-1015.

Jaillais Y, Chory J. 2010. Unraveling the paradoxes of plant hormone signaling integration. *Nature structural & molecular biology* **17**: 642-645.

Jones MA, Raymond MJ, Yang Z, Smirnoff N. 2007. NADPH oxidase-dependent reactive oxygen species formation required for root hair growth depends on ROP GTPase. *Journal of Experimental Botany* **58**: 1261-1270.

Khokon MAR, Salam MA, Jammes F, YE W, Hossain MA, Uraji M, Nakamura Y, Mori IC, Kwak JM, Murata Y. 2015. Two guard cell mitogen-activated protein kinases, MPK9 and MPK12, function in methyl jasmonate-induced stomatal closure in Arabidopsis thaliana. *Plant Biology* **17**: 946-952.

Kim TH, Bohmer M, Hu H, Nishimura N, Schroeder JI. 2010. Guard cell signal transduction network: advances in understanding abscisic acid, CO_2, and Ca2+ signaling. *Annu Rev Plant Biol* **61**: 561-91.

Klusener B, Young JJ, Murata Y, Allen GJ, Mori IC, Hugouvieux V, Schroeder JI. 2002. Convergence of calcium signaling pathways of pathogenic elicitors and abscisic acid in Arabidopsis guard cells. *Plant Physiol* **130**: 2152-63.

Kudla J, Becker D, Grill E, Hedrich R, Hippler M, Kummer U, Parniske M, Romeis T, Schumacher K. 2018. Advances and current challenges in calcium signaling. *New Phytologist* **218**: 414-431.

Kwak JM, Mori IC, Pei ZM, Leonhardt N, Torres MA, Dangl JL, Bloom RE, Bodde S, Jones JD, Schroeder JI. 2003. NADPH oxidase AtrbohD and AtrbohF genes function in ROS-dependent ABA signaling in Arabidopsis. *Embo Journal* **22**: 2623-33.

Li Q, Wang Y-J, Liu C-K, Pei Z-M, Shi W-L. 2017. The crosstalk between ABA, nitric oxide, hydrogen peroxide, and calcium in stomatal closing of Arabidopsis thaliana. *Biologia* **72(10):** 1140-1146.

Malone M. 1994. Wound-Induced Hydraulic Signals and Stimulus Transmission in Mimosa pudica L. *New Phytologist* **128**: 49-56.

Mangano S, Juárez SPD, Estevez JM. 2016. ROS Regulation of Polar Growth in Plant Cells. *Plant Physiology* **171**: 1593-1605.

McAinsh MR, Hetherington AM. 1998. Encoding specificity in Ca^{2+} signalling systems. *Trends in Plant Science* **3**: 32-36.

McAinsh MR, Pittman JK. 2009. Shaping the calcium signature. *New Phytologist* **181**: 275-294.

Mühlenbock P, Karpinska B, Karpinski S. 2007. Oxidative Stress and Redox Signalling in Plants. *eLS.*

Munemasa S, Hirao Y, Tanami K, Mimata Y, Nakamura Y, Murata Y. 2019. Ethylene Inhibits Methyl Jasmonate-Induced Stomatal Closure by Modulating Guard Cell Slow-Type Anion Channel Activity via the OPEN STOMATA 1/SnRK2.6 Kinase-Independent Pathway in Arabidopsis. *Plant and Cell Physiology* **60(10)**: 2263-2271.

Perez IB, Brown PJ. 2014. The role of ROS signaling in cross-tolerance: from model to crop. *Frontiers in plant science* **5**: 754-754.

Pitzschke A, Schikora A, Hirt H. 2009. MAPK cascade signalling networks in plant defence. *Current Opinion in Plant Biology* **12**: 421-426.

Qi J, Song C-P, Wang B, Zhou J, Kangasjärvi J, Zhu J-K, Gong Z. 2018. Reactive oxygen species signaling and stomatal movement in plant responses to drought stress and pathogen attack. *Journal of Integrative Plant Biology* **60**: 805-826.

Sewelam N, Kazan K, Schenk PM. 2016. Global Plant Stress Signaling: Reactive Oxygen Species at the Cross-Road. *Frontiers in Plant Science* **7**: 187.

Singh R, Singh S, Parihar P, Mishra RK, Tripathi DK, Singh VP, Chauhan DK, Prasad SM. 2016. Reactive Oxygen Species (ROS): Beneficial Companions of Plants' Developmental Processes. *Frontiers in Plant Science* **7**: 1299.

Sirichandra C, Gu D, Hu HC, Davanture M, Lee S, Djaoui M, Valot B, Zivy M, Leung J, Merlot S, Kwak JM. 2009. Phosphorylation of the Arabidopsis AtrbohF NADPH oxidase by OST1 protein kinase. *FEBS Lett* **583**: 2982-6.

Stahlberg R, Cleland RE, Van Volkenburgh E. 2006. Slow wave potentials - a propagating electrical signal unique to higher plants. *Communication in Plants: Neuronal Aspects of Plant Life*, 291-308. Springer, Berlin, Heidelberg.

Tognetti VB, Mühlenbock P, Van Breusegem F. 2012. Stress homeostasis – the redox and auxin perspective. *Plant, Cell & Environment* **35**: 321-333.

Tuteja N, Mahajan S. 2007. Calcium signaling network in plants: an overview. *Plant signaling & behavior* **2**: 79-85.

Vandereyken K, Van Leene J, DE Coninck B, Cammue BPA. 2018. Hub Protein Controversy: Taking a Closer Look at Plant Stress Response Hubs. *Frontiers in Plant Science* **9**: 694

Vert G, Chory J. 2011. Crosstalk in cellular signaling: background noise or the real thing? *Dev Cell* **21**: 985-91.

Volkov AG, Lang RD, Volkova-Gugeshashvili MI. 2007. Electrical signaling in Aloe vera induced by localized thermal stress. *Bioelectrochemistry* **71**: 192-197.

Wang Y, Chen ZH, Zhang B, Hills A, Blatt MR. 2013. PYR/PYL/RCAR abscisic acid receptors regulate K^+ and Cl^- channels through reactive oxygen species-mediated activation of Ca^{2+} channels at the plasma membrane of intact Arabidopsis guard cells. *Plant Physiol* **163**: 566-77.

Wilson ID, Ribeiro DM, Bright J, Confraria A, Harrison J, Barros RS, Desikan R, Neill SJ, Hancock JT. 2009. Role of nitric oxide in regulating stomatal apertures. *Plant signaling & behavior* **4**: 467-469.

Wu J, Baldwin IT. 2009. Herbivory-induced signalling in plants: perception and action. *Plant, Cell & Environment* **32**: 1161-1174.

Zhang Y, Zhu H, Zhang Q, Li M, Yan M, Wang R, Wang L, Welti R, Zhang W, Wang X. 2009. Phospholipase dalpha1 and phosphatidic acid regulate NADPH oxidase activity and production of reactive oxygen species in ABA-mediated stomatal closure in Arabidopsis. *The Plant cell* **21**: 2357-2377.

Zimmermann MR, Maischak H, Mithöfer A, Boland W, Felle HH. 2009. System Potentials, a Novel Electrical Long-Distance Apoplastic Signal in Plants, Induced by Wounding. *Plant Physiology* **149**: 1593-1600.

Zorov DB, Juhaszova M, Sollott SJ. 2014. Mitochondrial reactive oxygen species (ROS) and ROS-induced ROS release. *Physiological reviews* **94**: 909-950.

Appendix G.   A machine learning approach using a handheld near-infrared (NIR) device to predict the effect of storage conditions on tomato biomarkers

Natalia E M Emsley[1,¶], Claire A. Holden[2,¶], Sarah Guo[1], Rhiann S Bevan[1], Christopher Rees[3], Martin R. McAinsh[2], Francis L Martin[4,*], Camilo L M Morais[5,*]

[1] *Sixth Form, Lancaster Girls' Grammar School, Lancaster, Lancashire, UK*

[2] *Lancaster Environment Centre, Library Avenue, Lancaster University, Lancaster LA1 4YQ, UK*

[3] *Chemistry Department, Lancaster Girls' Grammar School, Lancaster, Lancashire, UK*

[4] *Biocel Ltd, Hull HU10 7TS, UK*

[5] *School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston, Lancashire, UK*

Corresponding authors: Francis L Martin (flm13@biocel.uk); Camilo L M Morais (camilomorais1@gmail.com)
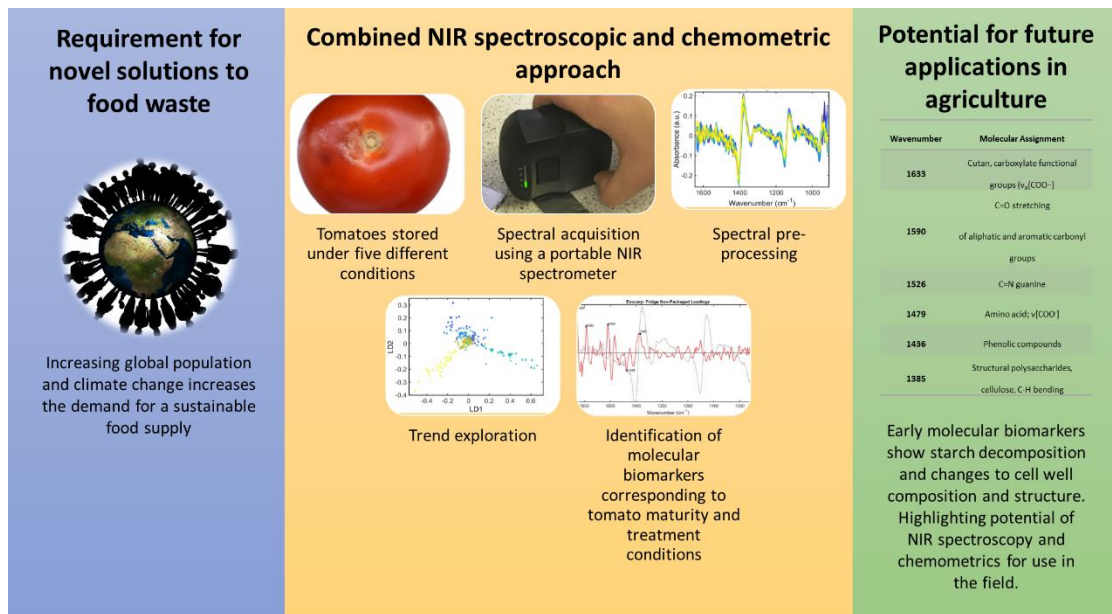
¶, Joint first authors

F.1 Abstract

Minimising food waste critical to future global food security. This study aimed to assess the potential of near-infrared (NIR) spectroscopy combined with machine learning to monitor the stability of tomato fruit during storage. Freshly harvested UK-grown tomatoes (n=135) were divided into five equally sized groups, each stored in different conditions. Absorbance spectra were obtained from both the tomato exocarp and locular gel using a portable NIR spectrometer, capable of connecting to a mobile phone, before subsequent chemometric analysis. Results show that support vector machines can predict the storage conditions and time-after-harvest of tomatoes. Molecular biomarkers highlighting key wavenumber and molecular changes due to time and storage conditions were also identified. This method shows potential for development of this approach for use in the field to help mitigate the environmental and economic impacts of food waste.

F.2 Keywords

Tomato, infrared spectroscopy, food security, machine learning, chemometrics, food storage

F.3 Table of Contents Graphic

Requirement for novel solutions to food waste	Combined NIR spectroscopic and chemometric approach	Potential for future applications in agriculture

F.4 Introduction

Sustainable food supply chains are required to provide adequate nutrition for the rising global population, which is projected to reach approximately 10 billion by 2050 (Food and Agriculture Organization of the United Nations, 2009). Achieving food security and ending malnutrition are priorities in the United Nations sustainable development agenda for 2030 (Nino, 2015). To achieve these goals using the same amount of available land, it is important that to maximise the efficiency of current supply chains; this includes the reduction of food waste. In the UK, food waste within the home is largely avoidable with 2.9 million metric tons, equating to £6.3 billion, of food spoiling before consumption (Parfitt *et al.*, 2010). Tomatoes are a popular crop globally, representing a versatile and nutrient-dense superfood, offering an excellent source of lycopene and other antioxidant molecules which benefit human health (Story *et al.*, 2010). In 2017, 9.2 kg of fresh tomatoes and 33.2 kg of processed tomato products were consumed per capita in the United States of America (Wu *et al.*, 2021). However, fresh ripe tomatoes are prone to a high

rate of food waste, with a total loss of 53.8% collectively from production to consumption (Boz & Sand, 2020). This results from a 20% loss in agricultural production, 7% loss in processing and packaging, 10% distribution and retail loss and 31% loss by consumers (Boz & Sand, 2020). The importance of reducing this food waste is emphasised during the current climate crisis, particularly as the production of tomatoes comes with a high environmental impact. They have been recognized as one of the most carbon-intensive food products due to electricity and fertiliser usage (Xue *et al.*, 2021), and greenhouse grown tomatoes are estimated to have a carbon footprint of up to 10.1 CO_2e kg (Ntinas *et al.*, 2017). Improving the efficiency of the supply chain for tomato, and agri-food in general, would reduce the economic and environmental impacts of food waste, and help us to meet the challenge of nourishing a growing population. Technical innovations to increase shelf-life have the potential to reduce the unnecessary disposal of home produce and hence reduce food waste (Quested *et al.*, 2013). Factors that affect the marketability and shelf-life of fresh ripe tomatoes include the duration that the fruits are safe for consumption, but also aesthetic qualities such as colour and firmness. As a climacteric fruit, the quality of fresh tomatoes can be improved after harvest. The ripening process involves changes to cell wall composition and thickness (Segado *et al.*, 2016), and the conversion of storage carbohydrates into monosaccharides (Fan *et al.*, 2018). These processes increase the softness and sweetness of the fruit, which in turn increases its appeal to consumers. However, if damaged or left to ripen for too long, tomatoes are prone to infections such as sour rot (Skolik *et al.*, 2019a). Early biomarker determination to rapidly predict fruit ripening, and hence its potential shelf-life before the fruit are no longer saleable, could help to achieve the optimum balance in post-harvest ripening under different climacteric conditions.

Near-infrared (NIR; wavelengths from 800 to 2,500 nm) spectroscopy has been used extensively in various recent studies to quantitatively determine fruit quality or maturity in tomatoes (Clément *et al.*, 2008; Sirisomboon *et al.*, 2012; Tiwari *et al.*, 2013; Saad *et al.*, 2014; Huang *et al.*, 2018; Skolik *et al.*, 2019a,b). The portability of

new miniaturised spectrometers has opened up new applications for spectroscopic studies, such as agriculture, see (Nolasco Perez *et al.*, 2018; Kaufmann *et al.*, 2019; Oliveira *et al.*, 2020; Beć *et al.*, 2021). Near infrared spectroscopy uses changes in molecular vibrations upon absorption of infrared light to gain information about the chemical composition of a sample. Biological materials preferentially absorb light in the fingerprint region (1800-900 cm^{-1}), which gives information about key biomolecules (Morais *et al.*, 2020). These include: lipids (C=O symmetric stretching at ~1,750 cm^{-1} and CH$_2$ bending at ~1,470 cm^{-1}), proteins (amide I at ~1,650 cm^{-1}, amide II at ~1,550 cm^{-1} and amide III at ~1,260 cm^{-1}), carbohydrates (CO-O-C symmetric stretching at ~1,155 cm^{-1}), nucleic acid (asymmetric phosphate stretching at ~1,225 cm^{-1} and symmetric phosphate stretching at ~1,080 cm^{-1}), glycogen (C-O stretching at ~1,030 cm^{-1}) and protein phosphorylation (~970 cm^{-1}) (Morais *et al.*, 2020). The current study utilised a handheld NIR spectroscopy device to determine tomato maturity and explore the effects of five different storage conditions over twenty days, varying in temperature and packaging. Spectral measurements were taken from both the exocarp and the locular gel within, allowing comparison of results from destructive and conservative techniques. A novel approach combining NIR using a handheld spectrometer and machine learning was used for classification and identification of key biomarkers. Chemometric methods included principal component analysis (PCA), PCA coupled with linear discriminant analysis (PCA-LDA) and support vector machines (SVM). Our results highlight key wavenumber changes associated with ripening and the effects of post-harvest climatic conditions, raising opportunity of developing this combined method for use in the field.

F.5 Materials and Methods

F.5.1 Samples and data acquisition

This study measured tomatoes over a period of twenty days under five different storage conditions (ambient packaged, ambient non-packaged, fridge packaged, fridge non-packaged, and incubated non-packaged). A total of 135 tomatoes were used in this study, twenty-seven per storage condition.

All tomatoes were F1 hybrids, a first generation cross between two varieties, grafted onto a strong disease resistant rootstock and grown in Rockwool, UK (supplied by John Lane). They were harvested at ten weeks old on Friday 13th September 2019. The tomatoes used in this study were selected for uniformity. At the start of the study, they were all graded to be of similar size, colour, and ripeness, were undamaged, and had a consistent colour across the whole surface. They were then randomly sorted into treatment groups. Three tomatoes from each treatment condition were selected for exocarp sampling and labelled X, Y and Z. Each treatment group was stored in different conditions over the course of the experiment, including ambient packaged, ambient non-packaged, fridge packaged, fridge non-packaged and incubated non-packaged. The packaging used consisted of a sealed plastic freezer bag to simulate modified atmosphere packaging (MAP) although without the introduction of a protective gas mix into the bag. The ambient tomatoes were stored at room temperature (18°C), the fridge tomatoes at 3°C and the incubated tomatoes at 25°C.

Spectral absorbances were measured using a hand-held NIR spectrometer NIR-S-G1 (Allied Scientific Pro, Gatineau, Quebec, Canada) using ISC NIRScan software (raw spectra are shown in the Supporting Information, Figure SF.1). Spectra were measured from two types of sample; exocarp (the tomato surface), and locular gel (a gel that develops prior to ripening of the pericarp and exhibits a liquid-like consistency towards the terminal stage of ripening, see (Atta-Aly *et al.*, 2000)). The

spectrometer crystal was cleaned between measurements using isopropyl alcohol wipes (Bruker Optics, Coventry, UK), and each time background spectra were taken to account for ambient atmospheric conditions. For locular gel samples spectral acquisition took place on days 3, 5, 7, 10, 12, 14, 17, and 20. Each day three different tomatoes were destructively analysed to extract the locular gel, and two spectral replicates per tomato were taken. This resulted in six locular gel spectra per day, forty-eight per treatment group, and 240 spectra in total. The locular gel was extracted by incision at two approximately equidistant locations around the equator. Fifty μL of this gel was collected using a Gilson pipette and transferred onto glass slides covered in aluminium foil. Before spectral acquisition of samples an aluminium foil standard spectrum was measured. For the exocarp measurements, the same three tomatoes were analysed throughout the whole time-course and subsequently returned to their original storage conditions. For exocarp samples, measurements were taken on days 1, 3, 5, 7, 10, 12, 14, 17, and 20. Ten spectral replicates of each tomato were taken, resulting 270 spectra per treatment group and 1350 spectra in total.

F.5.2 Data analysis and validation

The reflectance spectral data were imported and processed within MATLAB R2014b (MathWorks, Inc., Natick, MA, USA). Pre-processing and data analysis were performed using the PLS Toolbox version 7.9.3 (Eigenvector Research, Inc., Manson, WA, USA). The raw spectra were pre-processed by Savitzky-Golay smoothing (window of 7 points, 2nd order polynomial fitting) to improve the signal-to-noise ratio and standard normal variate (SNV) to correct for light scattering. The pre-processed data were also mean-centred before multivariate analysis.

Principal component analysis (PCA) was applied to the pre-processed spectral data for exploratory analysis. PCA is an exploratory analysis technique that reduces the spectral dataset into a small number of principal components (PCs) responsible for the majority of the original data variance (Bro & Smilde, 2014). Each PC is

orthogonal to each other, and they are formed in a decreasing order of explained variance, so PC1 covers more variance than PC2, and so on. Each PC is composed of scores and loadings. The scores represent the variance on sample direction, thus being used to identify patterns of similarity between the samples, and the loadings represent the variance on wavelength direction, thus being used to identify possible spectral markers responsible for the scores pattern. Although being a robust exploratory analysis technique, PCA was not able to classify samples in an objective fashion, therefore, a supervised classifier was added and the samples were analysed by principal component analysis with linear discriminant analysis (Morais & Lima, 2018). For more rigorous classification, support vector machines (SVM) algorithm was applied to estimate the tomatoes time-after-harvest. SVM is a linear classifier with a non-linear step called the kernel transformation (Cortes & Vapnik, 1995). The kernel function transforms the data space to a feature space where samples can be better discriminated. Herein, the radial basis function (RBF) kernel was used and optimised via cross-validation. The pre-processed spectra were then randomly split into training (70%) and test (30%) sets, and a supervised classification model was constructed using support vector machines (SVM) to systemically predict the time-after-harvest regardless the storage condition. The training set was used to build the SVM training model, and the test set to evaluate its predictive ability. The SVM model was optimised by cross-validation venetian blinds with 10 data splits.

Metrics such as accuracy, sensitivity and specificity were calculated for the test set. For more than two-classes, these metrics are calculated individually per class; herein, the average for all classes is reported. The accuracy (AC) represents the total number of samples correctly classified considering true and false negatives; the sensitivity (SENS) represents the proportion of positives that are correctly classified; and the specificity (SPEC) represents the proportion of negatives that are correctly identified (Morais & Lima, 2017). These metrics are calculated as follows:

$$AC(\%) = \left(\frac{TP+TN}{TP+FP+TN+FN}\right) \times 100$$

$$(1)$$

$$SENS(\%) = \left(\frac{TP}{TP+FN}\right) \times 100$$

$$(2)$$

$$SPEC(\%) = \left(\frac{TN}{TN+FP}\right) \times 100$$

$$(3)$$

where TP stands for true positives, TN for true negatives, FP for false positives, and FN for false negatives.

F.6 Results

Figure F.1 shows the fingerprint spectra for a) exocarp and b) locular gel absorbances, and pre-processed spectra for c) exocarp and d) locular gel. The pre-processed spectra for the exocarp measurements in Figure F.1c contain bands around 1000 nm (C-H stretching 3rd overtone), 1200 nm (C-H stretching 2nd overtone in fibre parameters such as cellulose and lignin), 1360 nm (small arm, R-O-H stretching 1st overtone in alcohol), 1450 nm (O-H stretching 1st overtone in water), and spectral differences at around 1700 nm (C-H stretching 1st overtone in glucose/lignin) (Jerome & Workman, 2006; Talari *et al.*, 2017). The locular gel fingerprint spectra, Figure F.1d, have their bands compressed by the strong water peak at 1450 nm (Morais *et al.*, 2020).
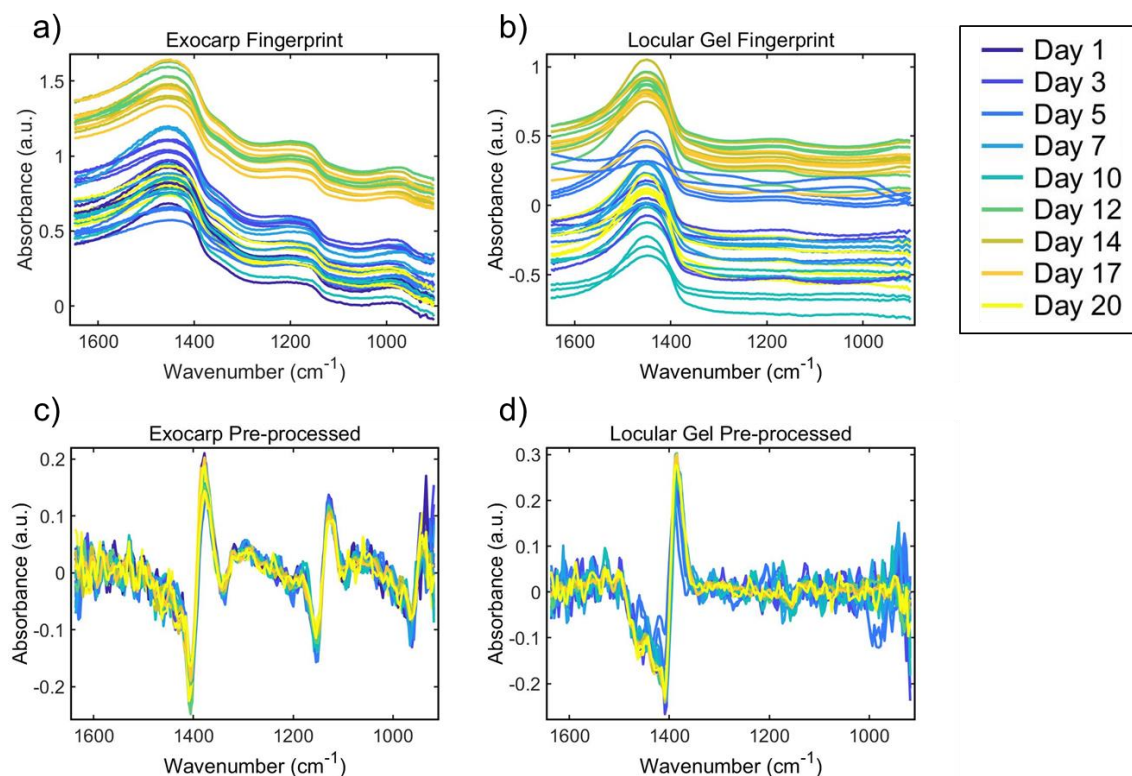
Figure F.1: Fingerprint spectra for **a)** exocarp and **b)** locular gel absorbances, and pre-processed spectra for **c)** exocarp and **d)** locular gel absorbances. Each line is a class mean of spectra from a specific day and treatment condition. The colours represent the day the spectra were taken.

The pre-processed spectral data initially underwent unsupervised exploratory analysis by PCA, where overall segregation trends were observed in the data. Figure F.2 shows PCA scatter plots for a) exocarp and b) locular gel spectra. The colours represent the time in days and the marker shapes represent the treatment and numbers inside parenthesis represent the explained variance for each PC. PCA did not provide separation between the samples (Figures F.2a and F.2b). For PCA scores of each storage condition individually, see supporting information Figure SF.2 and S3, for exocarp and locular gel measurements respectively.

Following the initial PCA, a supervised method, linear discriminant analysis, was also applied. PCA-LDA scatter plots are shown in Figure F.2c) for exocarp and 2d) for locular gel spectra. Although there is some evidence of a time evolution trend for

the exocarp measurements (see Figure F.2c), the samples cannot be differentiated into clear clusters. Some clustering was however achieved for locular gel samples. Figure F.2d shows that days 1, 10 and 20 are separated out best in the locular gel PCA-LDA scatter plot along the axes LD1 and LD2.
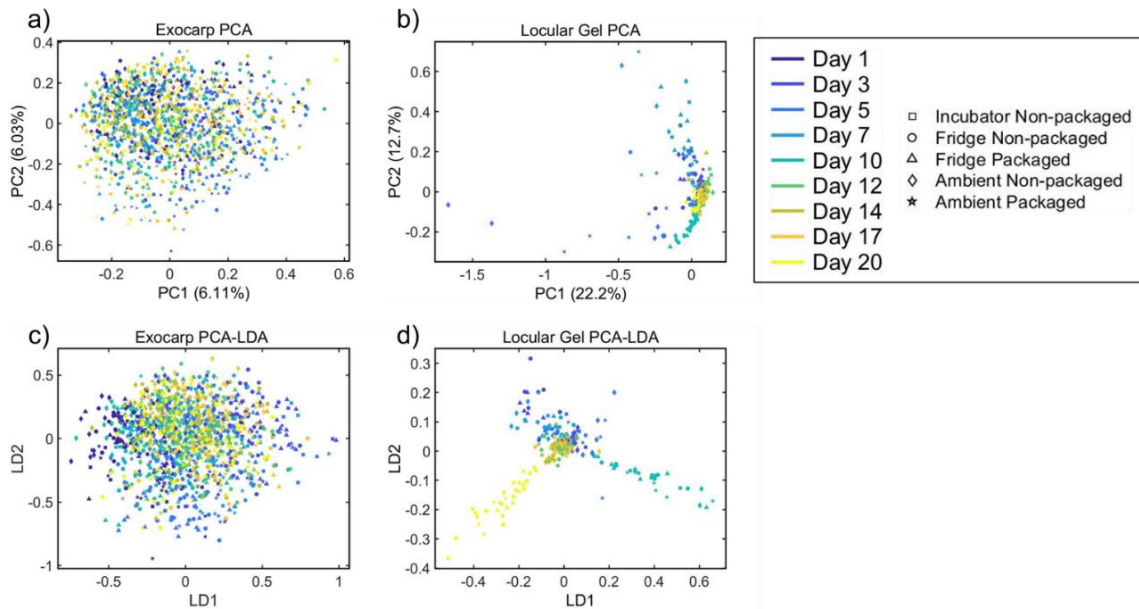

Figure F.2: PCA scatter plots for **a)** exocarp and **b)** locular gel spectra. Numbers inside parenthesis represent the explained variance for each PC. PCA-LDA scatter plots for **c)** exocarp and **d)** locular gel spectra. The colours represent the time in days and the marker shapes represent the treatment.

Figure F.3 shows the SVM model prediction results for each time and treatment sample type grouped separately. Overall, exocarp achieved higher accuracy sensitivity and specificity than locular gel spectra. For all forty-five exocarp and forty-one local gel categories, the SVM model achieved poor results at 51% and 54% accuracy, respectively. The SVM model to predict the time-after-harvest regardless the storage condition for the exocarp spectra achieved 92% accuracy, 86% sensitivity and 98% specificity in test sets. For locular gel spectra, the SVM model achieved 84% accuracy, 74% sensitivity and 95% specificity in test sets. Sensitivity of the exocarp spectra was - high, with many samples correctly classified. For locular gel samples, packaged fridge-stored samples at days 3 and 20, and ambient stored non-packaged samples at day 3 achieved higher true positive rates than the other

groups. To give an overall picture of the trend over time, despite treatment, spectra from different storage conditions were also grouped together. The SVM results for this grouping can be viewed in the supporting information: see Figure SF.4 for optimisation parameters, Table SF.1 for classification metrics, and Table SF.2 for test set confusion matrices.
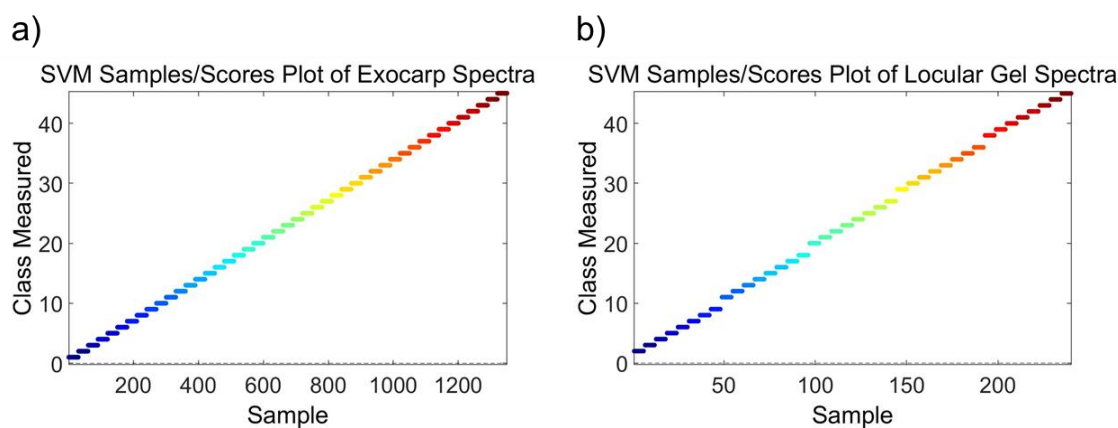

Figure F.3: SVM predicted results for the test set for **a)** exocarp and **b)** locular gel. For all forty-five exocarp categories, the SVM model achieved 92% accuracy, 86% sensitivity and 98% specificity in test sets. For all forty-one locular gel categories, the SVM model achieved 84% accuracy, 74% sensitivity and 95% specificity in test sets. Each predicted category is shown in a different colour.

PCA loadings were used to determine the key wavenumbers associated with biochemical changes over time for each treatment, in both exocarp and locular gel spectra (see Tables F.1 and F.2, respectively). These spectral changes related to their biological origin by seeking molecular assignments from existing literature. Spectra from each treatment group were considered separately so that the changes highlighted in Tables F.1 and F.2 relate only to time.

Table F.1: Loadings showing the key wavenumber changes for locular gel spectra associated with biochemical changes over time for each treatment.

Sample Type	Treatment Type	Loadings	Assignment	Reference
Locular Gel	INP	1636	CHO stretching of carbonyl group, typical saccharide absorption	(Talari *et al.*, 2017)
		1456	CH_3 bending vibration (lipids and proteins)	(Talari *et al.*, 2017)
		1170	C-O bands from glycomaterials and proteins	(Talari *et al.*, 2017)
		1095	Stretching PO_2 symmetric	(Talari *et al.*, 2017)
		981	Phosphodiester region	(Talari *et al.*, 2017)
		918	Polysaccharides	(Di Giambattista *et al.*, 2011)
Locular Gel	FP	1605	v_{as} (COO$^-$) (polysaccharides, pectin)	(Talari *et al.*, 2017)
		1571	C=N adenine	(Talari *et al.*, 2017)
		1381	Amide II	(Luo *et al.*, 2021)
		1075	Symmetric phosphate stretching modes or v(PO$_2^-$) sym.	(Talari *et al.*, 2017)
		1012	Starch	(Luo *et al.*, 2021)
		945	D-(-)-Arabinose	(Liu X *et al.*, 2021)

Locular Gel	ANP	1636	C=O stretching of carbonyl group, typical saccharide absorption	(Talari *et al.*, 2017)
		1596	Methylated nucleotides	(Talari *et al.*, 2017)
		1443	δ(CH) (polysaccharides, pectin), δ(CH$_2$), lipids, fatty acids	(Talari *et al.*, 2017)
		1170	C-O bands from glycomaterials and proteins	(Talari *et al.*, 2017)
		993	Arabinoxylans	(Demir *et al.*, 2015)
		930	Polysaccharides	(Di Giambattista *et al.*, 2011)
Locular Gel	AP	1636	C=O stretching of carbonyl group, typical saccharide absorption	(Talari *et al.*, 2017)
		1493	Protein	(Meinen & Rauber, 2015)
		1422	Protein and lipids	(Meinen & Rauber, 2015)
		1124	Polysaccharides	(Butler *et al.*, 2015)
		977	C-O-C stretching at the β-(1 $\rightarrow$ 4)-glycosidic linkages of amorphous cellulose	(Bekiaris *et al.*, 2015)
		918	Polysaccharides	(Di Giambattista *et al.*, 2011)
Locular Gel	FNP	1609	Adenine vibration in DNA	(Talari *et al.*, 2017)

		1574	C=N adenine	(Talari *et al.*, 2017)
		1500	In-plane CH bending vibration from the phenyl rings, or Amide II (an N-H bending vibration coupled to C-N stretching)	(Talari *et al.*, 2017)
		1236	Amide III and asymmetric phosphodiester stretching mode, $v_{as}(PO_2^-)$ mainly from the nucleic acids	(Talari *et al.*, 2017)
		993	Arabinoxylans	(Demir *et al.*, 2015)
		926	Polysaccharides	(Di Giambattista *et al.*, 2011)

Table F.2: Loadings showing the key wavenumber changes for exocarp spectra associated with biochemical changes over time for each treatment.

Sample Type	Treatment Type	Loadings	Assignment	Reference
Exocarp	INP	1636	C=O stretching of carbonyl group, typical saccharide absorption	(Talari *et al.*, 2017)
		1590	C=O stretching of aliphatic and aromatic carbonyl groups	(Talari *et al.*, 2017)
		1542	Amide II	(Talari *et al.*, 2017)
		1483	Protein and lipids	(Meinen & Rauber, 2015)
		962	Polysaccharides	(Di Giambattista *et al.*, 2011)
		918	Polysaccharides	(Di Giambattista *et al.*, 2011)
Exocarp	FP	1627	Phenolic compounds	(Heredia-Guerrero *et al.*, 2014)
		1584	Amide II	(Talari *et al.*, 2017)
		1526	C=N guanine	(Talari *et al.*, 2017)
		1456	CH_3 bending vibration (lipids and proteins)	(Talari *et al.*, 2017)

		981	Phosphodiester region	(Talari *et al.*, 2017)
		918	Polysaccharides	(Di Giambattista *et al.*, 2011)
Exocarp	ANP	1636	C=O stretching of carbonyl group, typical saccharide absorption	(Talari *et al.*, 2017)
		1574	C=N adenine	(Talari *et al.*, 2017)
		1532	Stretching C=N, C=C	(Talari *et al.*, 2017)
		1479	Amino acid; v[COO$^-$]	(Luo *et al.*, 2021)
		1381	Amide II	(Luo *et al.*, 2021)
		918	Polysaccharides	(Di Giambattista *et al.*, 2011)
Exocarp	AP	1627	Phenolic compounds	(Heredia-Guerrero *et al.*, 2014)
		1581	Ring C-C stretch of phenyl	(Talari *et al.*, 2017)
		1509	Hemicellulose, C-C and C=C	(Zhou *et al.*, 2015)
		1473	Glycerolipids, wax hydrocarbons, $\delta(CH_2)$ scissoring	(Skolik *et al.*, 2019b)
		1409	Succinic acid	(Kang *et al.*, 2008)

		933	Z type DNA	(Talari *et al.*, 2017)
Exocarp	FNP	1630	Amide I	(Talari *et al.*, 2017)
		1593	Protein	(Butler *et al.*, 2015)
		1509	Hemicellulose, C-C and C=C	(Zhou *et al.*, 2015)
		1436	Phenolic compounds	(Heredia-Guerrero *et al.*, 2014)
		1385	Structural polysaccharides, cellulose, C-H bending	(Rana *et al.*, 2018)
		926	Polysaccharides	(Di Giambattista *et al.*, 2011)

To compare the effects of different storage treatments on spectral absorbances, data from the middle of the time course prior to any visible change in tomato quality were selected. Spectra from day 10 were used to derive loadings showing which wavenumbers were associated with differing treatment conditions; these were subsequently connected with their corresponding biomarkers. To view loading graphs from which the information in Tables F.1-F.3 were derived, see the supporting information Figure SF.5 for locular gel, Figure SF.6 for exocarp and Figure SF.7 for day ten.

The different treatment conditions resulted in tomatoes of different qualities after 20 days (Figure F.4). Visually, fridge non-packaged and ambient packaged treatments best preserved the quality of the tomatoes, whereas tomatoes from the incubator non-packaged and ambient non-packaged treatments showed signs of infection, and fridge packaged tomatoes showed signs of a possible chilling stress.

Table F.3: Loadings showing the key wavenumber differences between treatments at day 10 for locular gel and exocarp spectra.

Sample Type	Day	Wavenumber	Molecular Assignment	Reference
Exocarp	Day 10	1633	Cutan, carboxylate functional groups (ν_a(COO$^-$)	(Heredia-Guerrero *et al.*, 2014)
		1590	C=O stretching of aliphatic and aromatic carbonyl groups	(Talari *et al.*, 2017)
		1526	C=N guanine	(Talari *et al.*, 2017)
		1479	Amino acid; ν[COO$^-$]	(Luo *et al.*, 2021)
		1436	Phenolic compounds	(Heredia-Guerrero *et al.*, 2014)
		1385	Structural polysaccharides, cellulose, C-H bending	(Rana *et al.*, 2018)
Locular Gel	Day 10	1602	Pectin, phenolic compounds; C-C aromatic stretching; C-O-O$^-$ asymmetric stretching	(Moskal *et al.*, 2019)
		1486	Proteins	(Jin *et al.*, 2018)

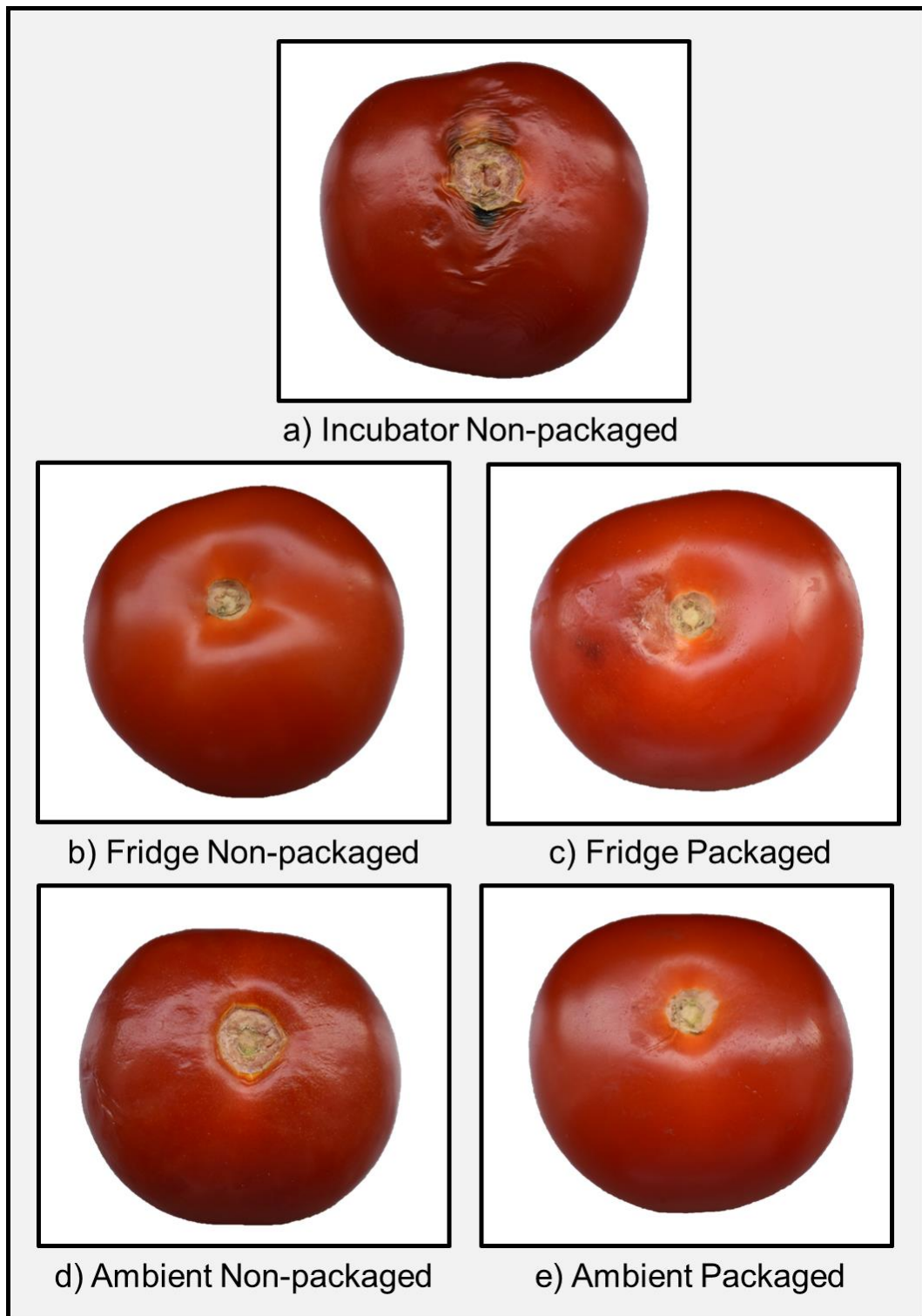|		1440	Phenolic compounds; ν(C-C) aromatic (conjugated with C=C)	(Heredia-Guerrero *et al.*, 2014)
		1064	C–O stretching, C–C stretching (mannose- containing hemicellulose)	(Liu X *et al.*, 2021)
		993	Arabinoxylans	(Demir *et al.*, 2015)
		933	Z type DNA	(Talari *et al.*, 2017)

Figure F.4: Images illustrating relative tomato quality between treatments on day 20.

F.7 Discussion

SVM achieved outstanding classification accuracy across exocarp (92%) and locular gel (84%) samples to estimate the time-after-harvest despite the storage condition. The accuracy to predict all forty-five and forty-one categories of samples (including differences between storage conditions) could be improved if the model was built to differentiate between fewer categories, or a larger dataset was used. When more than ten classes are involved in the classification model its performance tends to reduce substantially since the dataset size is relatively reduced by a fraction of the number of classes when building the classifier on a one-against-the-others categories basis. Therefore, a very large dataset would be required to perform this type of classification. The models were better at identifying where not to place a spectrum but often failed to find the correct category to sort it into, with high specificities (60% and 91%) but low sensitivities (3% and 20%). When spectra from all storage condition treatments were grouped together, classification by time achieved excellent results with SVM; 92% average accuracy, 86% average sensitivity, and 98% average specificity in the test set for exocarp measurements, and 84% average accuracy, 74% average sensitivity, and 95% average specificity for the test set of locular gel samples, see supporting information Table SF.2.

PCA loadings identified key wavenumbers which allow discrimination between spectral absorbances from tomatoes of different storage conditions and maturities, see Tables F.1, 2 and 3. These were connected to biomarkers using existing literature databases, allowing insight into the biochemical changes taking place. Many of the identified changes related to the progression of starch degradation, where the starch stored during development and is converted to soluble sugars, a key process during post-harvest fruit ripening (Fan *et al.*, 2018). Wavenumber 1012 cm^{-1} in locular gel fridge-stored spectra identified starch specifically as an important indicator of tomato age (Luo *et al.*, 2021). All treatment conditions for both locular gel and exocarp spectra identified that peaks for polysaccharides (Di Giambattista

et al., 2011) and/ or saccharides (Talari *et al.*, 2017) were used for the differentiation between tomatoes of different ages. This indicates that all fruits underwent some level of post-harvest ripening, irrespective of storage conditions.

Other key peaks identified across different conditions were those relating to the structural and compositional development of the cuticle and cell wall. Compositional changes in key compounds such as pectin, cellulose, and other polysaccharides, as well as changes in cell wall thickness, are part of the ripening process (Segado *et al.*, 2016). Tomato maturity was indicated in the spectra of ambient packaged exocarp samples at peak 1473 cm^{-1}, which relates to the $\delta(CH_2)$ scissoring of cuticle glycerolipids and wax hydrocarbons (Skolik *et al.*, 2019b). Pectin was identified as a measure of tomato maturity in ambient non-packaged and fridge packaged samples from locular gel spectra (Talari *et al.*, 2017), and also as a differentiator between storage conditions (Moskal *et al.*, 2019). Cellulose was an indicator of tomato maturity in the locular gel of ambient packaged fruit (Bekiaris *et al.*, 2015). However cellulose and hemicellulose was more commonly associated with exocarp spectra where these compounds differentiated between maturity in ambient packaged and fridge non-packaged samples (Zhou *et al.*, 2015). Cellulose was a key differentiator between treatment conditions in exocarp spectra (Rana *et al.*, 2018).

Dissolution of the cell wall is key to another ripening process, softening. Arabinogalactan proteins, which act as a cross-linker for pectin and arabinoxylan, are thought to be important in the alteration of cell wall assembly (Leszczuk *et al.*, 2020). In this study, locular gel arabinoxylan levels were a key differentiator between treatment conditions and allowed maturity determination in fridge and ambient non-packaged fruit (Demir *et al.*, 2015). These results suggest that the packaging is altering the environment experienced by the tomatoes, and having an impact on arabinoxylan, and consequently fruit-softness. The texture changes can be seen in Figure F.4, where the ambient non-packaged fruit appears softer than its packaged equivalent. Expression of arabinogalactan proteins in tomato fruit has also

been linked to possible involvement in stress adaptations (Leszczuk *et al.*, 2020), which may be why the packaged tomatoes suffered from a chilling stress in the fridge whilst their unpacked equivalents appeared to be in better condition, see Figure F.4. Fleshy fruit such as tomatoes are particularly vulnerable to chilling injury, is a type of oxidative stress that occurs during storage below 10 °C (Stevens *et al.*, 2008). Despite this chilling, fridge non-packaged tomatoes were amongst the best-preserved tomatoes alongside ambient packaged tomatoes, see Figure F.4. The least well-preserved tomatoes were those not protected with packaging, stored under ambient room temperature or in the incubator, which showed signs of infection. The lack of packaging likely exposed the tomatoes to any pathogens present, and the high temperatures in the incubator conditions promoted growth. Bacteria and fungi grow best at temperatures between 25-30 °C (Jay *et al.*, 2008; Adams & Moss, 2020), and the incubated tomatoes were stored at 25°C.

This study has identified biomarkers, including those indicating tomato ripening, which are modified by storage and packaging conditions and that have the potential to be used to target reductions in the unnecessary disposal of tomatoes through spoilage in the supply chain. The rapid and non-destructive nature of exocarp scanning and the portability of the spectrometer, capable of connection to a mobile phone, renders this method particularly suitable for use within the food industry. With further development, this user-friendly and non-destructive technology displays potential for a wider range of applications within and beyond the food industry.

F.8 Acknowledgements

We would like to acknowledge John Lane, who supplied the tomatoes and Lancaster Girls' Grammar School for their support for the project.

F.9 Author Contributions

NEME, FLM and CLMM conceived the study; NEME, SG, RSB, FLM and CLMM jointly designed the experiment; NEME, SG, RSB conducted the experiment and collected the data; CLMM and CAH undertook the chemometric analyses; NEME, CAH, SG, RSB, CLMM jointly drafted the manuscript, MRM advised on the plant biology and all authors approved the final version; CR provided support and supervision for the conduct of the experiment; FLM was principal investigator.

F.10 Declaration of Interest

FLM is a major shareholder in Biocel UK Ltd, a company with interests in developing tools described herein for commercial gain.

F.11 References

Adams MR, Moss MO. **2020**. The microbiology of food preservation.

Atta-Aly MA, Brecht JK, Huber DJ. **2000**. Ripening of tomato fruit locule gel tissue in response to ethylene. *Postharvest Biology and Technology* **19**: 239–244.

Beć KB, Grabska J, Huck CW. **2021**. Principles and Applications of Miniaturized Near-Infrared (NIR) Spectrometers. *Chemistry – A European Journal* **27**: 1514–1532.

Bekiaris G, Lindedam J, Peltre C, Decker SR, Turner GB, Magid J, Bruun S. **2015**. Rapid estimation of sugar release from winter wheat straw during bioethanol production using FTIR-photoacoustic spectroscopy. *Biotechnology for Biofuels 2015 8:1* **8**: 1–12.

Boz Z, Sand CK. **2020**. A systematic analysis of the overall nutritional contribution of food loss and waste in tomatoes, spinach, and kidney beans as a function of processing. *Journal of Food Process Engineering* **43**: e13509.

Bro R, Smilde AK. **2014**. Principal component analysis. *Analytical methods* **6**: 2812–2831.

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

Clément A, Dorais M, Vernon M. **2008**. Multivariate Approach to the Measurement of Tomato Maturity and Gustatory Attributes and Their Rapid Assessment by Vis–NIR Spectroscopy. *Journal of Agricultural and Food Chemistry* **56**: 1538–1544.

Cortes C, Vapnik V. **1995**. Support-vector networks. *Machine learning* **20**: 273–297.

Demir P, Onde S, Severcan F. **2015**. Phylogeny of cultivated and wild wheat species using ATR-FTIR spectroscopy. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy* **135**: 757–763.

Fan Z-Q, Ba L-J, Shan W, Xiao Y-Y, Lu W-J, Kuang J-F, Chen J-Y. **2018**. A banana R2R3-MYB transcription factor MaMYB3 is involved in fruit ripening through modulation of starch degradation by repressing starch degradation-related genes and MabHLH6. *The Plant Journal* **96**: 1191–1205.

Food and Agriculture Organization of the United Nations. **2009**. High-level expert forum—How to feed the world in 2050.

Di Giambattista L, Pozzi D, Grimaldi P, Gaudenzi S, Morrone S, Castellano AC. **2011**. New marker of tumor cell death revealed by ATR-FTIR spectroscopy. *Analytical and Bioanalytical Chemistry* **399**: 2771–2778.

Heredia-Guerrero JA, Benítez JJ, Domínguez E, Bayer IS, Cingolani R, Athanassiou A, Heredia A. **2014**. Infrared and Raman spectroscopic features of plant cuticles: a review. *Frontiers in Plant Science* **0**: 305.

Huang Y, Lu R, Chen K. **2018**. Prediction of firmness parameters of tomatoes by portable visible and near-infrared spectroscopy. *Journal of Food Engineering* **222**: 185–198.

Jay JM, Loessner MJ, Golden DA. **2008**. Low-temperature food preservation and characteristics of psychrotrophic microorganisms. In: Modern food microbiology. Springer Science & Business Media.

Jerome J, Workman J. **2006**. Interpretive Spectroscopy for Near Infrared. *http://dx.doi.org/10.1080/05704929608000571* **31**: 251–320.

Jin N, Semple KT, Jiang L, Luo C, Zhang D, Martin FL. **2018**. Spectrochemical analyses of growth phase-related bacterial responses to low (environmentally-relevant) concentrations of tetracycline and nanoparticulate silver. *Analyst* **143**: 768–776.

Kang S, Amarasiriwardena D, Xing B. **2008**. Effect of dehydration on dicarboxylic acid coordination at goethite/water interface. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **318**: 275–284.

Kaufmann K, Favero F, de Vasconcelos M, Godoy H, Sampaio K, Barbin D. **2019**. Portable NIR Spectrometer for Prediction of Palm Oil Acidity. *Journal of food science* **84**: 406–411.

Leszczuk A, Kalaitzis P, Blazakis KN, Zdunek A. **2020**. The role of arabinogalactan proteins (AGPs) in fruit ripening—a review. *Horticulture Research 2020 7:1* **7**: 1–12.

Liu X, Renard CMGC, Bureau S, Le Bourvellec C. **2021**. Revisiting the contribution of ATR-FTIR spectroscopy to characterize plant cell wall polysaccharides. *Carbohydrate Polymers* **262**: 117935.

Luo Y, Liu H, Wu C, Paraskevaidi M, Deng Y, Shi W, Yuan Y, Feng R, Martin FL, Pang W. **2021**. Diagnostic Segregation of Human Breast Tumours Using Fourier-transform Infrared Spectroscopy Coupled with Multivariate Analysis: Classifying Cancer Subtypes. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **255**: 119694.

Meinen C, Rauber R. **2015**. Root discrimination of closely related crop and weed species using FT MIR-ATR spectroscopy. *Frontiers in Plant Science* **6**: 765.

Morais CLM, Lima KMG. **2017**. Comparing unfolded and two-dimensional discriminant analysis and support vector machines for classification of EEM data. *Chemometrics and Intelligent Laboratory Systems* **170**: 1–12.

Morais CLM, Lima KMG. **2018**. Principal Component Analysis with Linear and Quadratic Discriminant Analysis for Identification of Cancer Samples Based on Mass Spectrometry. *Article J. Braz. Chem. Soc* **29**: 472–481.

Morais CLM, Lima KMG, Singh M, Martin FL. **2020**. Tutorial: multivariate classification for vibrational spectroscopy in biological samples. *Nature Protocols* **15**: 2143–2162.

Moskal P, Wesełucha-Birczyńska A, Łabanowska M, Filek M. **2019**. Adaxial and abaxial pattern of Urtica dioica leaves analyzed by 2DCOS ATR-FTIR as a function of their growth time and impact of environmental pollution. *Vibrational Spectroscopy* **104**: 102948.

Nino FS. **2015**. Sustainable Development Goals—United Nations. *United Nations Sustainable Development*.

Nolasco Perez I, Badaró A, Barbon S, Barbon A, Pollonio M, Barbin D. **2018**. Classification of Chicken Parts Using a Portable Near-Infrared (NIR) Spectrophotometer and Machine Learning. *Applied spectroscopy* **72**: 1774–1780.

Ntinas GK, Neumair M, Tsadilas CD, Meyer J. **2017**. Carbon footprint and cumulative energy demand of greenhouse and open-field tomato cultivation systems under Southern and Central European climatic conditions. *Journal of cleaner production* **142**: 3617–3626.

Oliveira MM, Cruz-Tirado JP, Roque J V, Teófilo RF, Barbin DF. **2020**. Portable near-infrared spectroscopy for rapid authentication of adulterated paprika powder. *Journal of Food Composition and Analysis* **87**: 103403.

Parfitt J, Barthel M, MacNaughton S. **2010**. Food waste within food supply chains: Quantification and potential for change to 2050. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**: 3065–3081.

Quested T, Ingle R, Parry A. **2013**. Household food and drink waste in the United Kingdom 2012, Waste & Resources Action Programme (WRAP). *Banbury, UK*.

Rana R, Herz K, Bruelheide H, Dietz S, Haider S, Jandt U, Pena R. **2018**. Leaf Attenuated Total Reflection Fourier Transform Infrared (ATR-FTIR) biochemical profile of grassland plant species related to land-use intensity. *Ecological Indicators* **84**: 803–810.

Saad A, Gawad Saad A, Jaiswal P, Narayan Jha S. **2014**. Non-destructive quality evaluation of intact tomato using VIS-NIR spectroscopy Related papers Non-destructive quality evaluation of intact tomato using VIS-NIR spectroscopy. *International Journal of Advanced Research* **2**: 632–639.

Segado P, Domínguez E, Heredia A. **2016**. Ultrastructure of the Epidermal Cell Wall and Cuticle of Tomato Fruit (Solanum lycopersicum L.) during Development. *Plant Physiology* **170**: 935.

Sirisomboon P, Tanaka M, Kojima T, Williams P. **2012**. Nondestructive estimation of maturity and textural properties on tomato 'Momotaro' by near infrared spectroscopy. *Journal of Food Engineering* **112**: 218–226.

Skolik P, McAinsh MR, Martin FL. **2019a**. ATR-FTIR spectroscopy non-destructively detects damage-induced sour rot infection in whole tomato fruit. *Planta* **249**: 925–939.

Skolik P, Morais CLM, Martin FL, McAinsh MR. **2019b**. Determination of developmental and ripening stages of whole tomato fruit using portable infrared spectroscopy and Chemometrics. *BMC Plant Biology* **19**: 236.

Stevens R, Page D, Gouble B, Garchery C, Zamir D, Causse M. **2008**. Tomato fruit ascorbic acid content is linked with monodehydroascorbate reductase activity and tolerance to chilling stress. *Plant, cell & environment* **31**: 1086–1096.

Story EN, Kopec RE, Schwartz SJ, Harris GK. **2010**. An Update on the Health Effects of Tomato Lycopene. *http://dx.doi.org/10.1146/annurev.food.102308.124120* **1**: 189–210.

Talari ACS, Martinez MAG, Movasaghi Z, Rehman S, Rehman IU. **2017**. Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Applied Spectroscopy Reviews* **52**: 456–506.

Tiwari G, Slaughter DC, Cantwell M. **2013**. Nondestructive maturity determination in green tomatoes using a handheld visible and near infrared instrument. *Postharvest Biology and Technology* **86**: 221–229.

Wu X, Yu L, Pehrsson PR. **2021**. Are Processed Tomato Products as Nutritious as Fresh Tomatoes? Scoping Review on the Effects of Industrial Processing on Nutrients and Bioactive Compounds in Tomatoes. *Advances in Nutrition*. **13(1):** 138-51.

Xue L, Cao Z, Scherhaufer S, Östergren K, Cheng S, Liu G. **2021**. Mapping the EU tomato supply chain from farm to fork for greenhouse gas emission mitigation strategies. *Journal of Industrial Ecology* **25**: 377–389.

Zhou C, Jiang W, Via BK, Fasina O, Han G. **2015**. Prediction of mixed hardwood lignin and carbohydrate content using ATR-FTIR and FT-NIR. *Carbohydrate Polymers* **121**: 336–341.
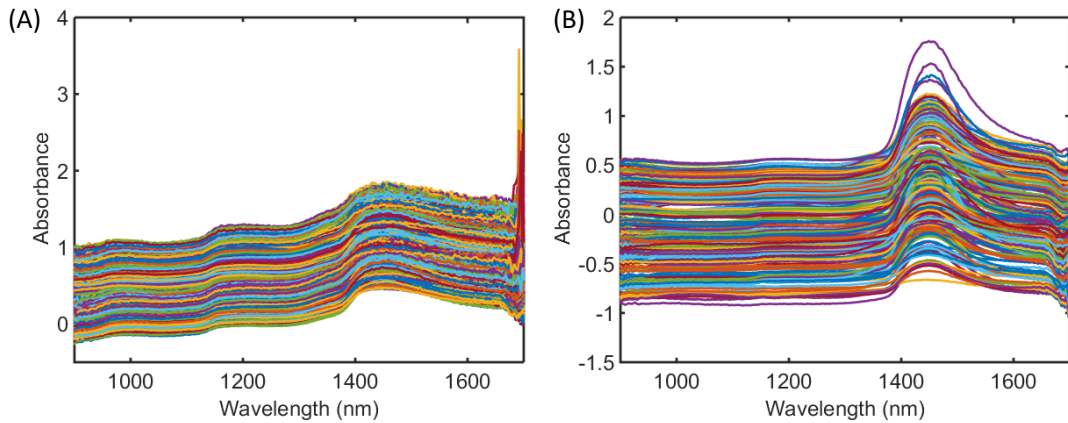
F.12 Supporting Information

Figure SF.1: Raw NIR spectra of **a)** exocarp measurements and **b)** locular gel.
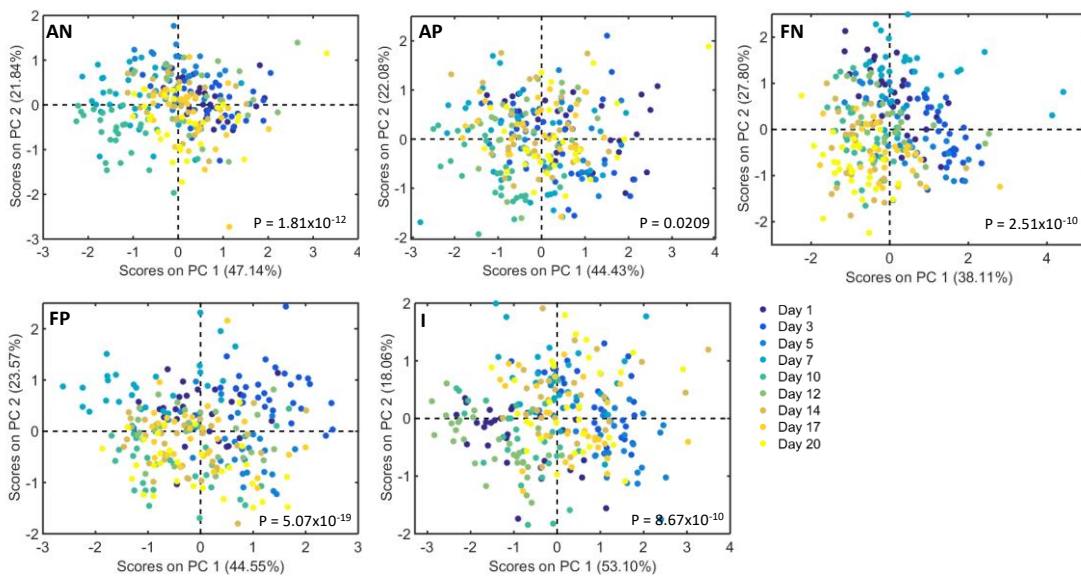

Figure SF.2: PCA scores (PC1 x PC2) for exocarp measurements by storage condition: AN (ambient non-packaged), AP (ambient packaged), FN (fridge non-packaged), FP (fridge packaged), I (incubated non-packaged).
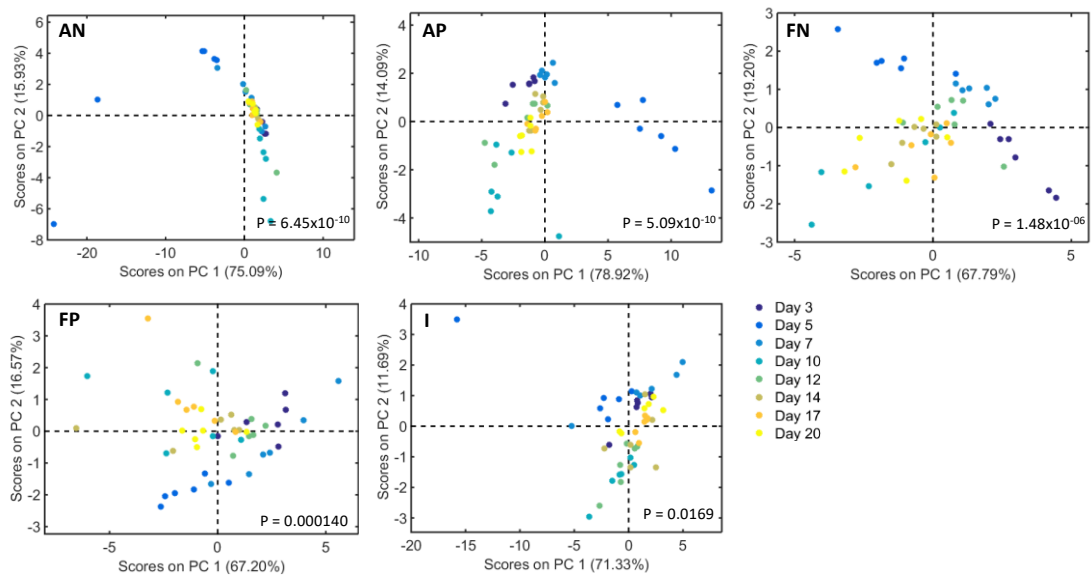
Figure SF.3: PCA scores (PC1 x PC2) for locular gel by storage condition: AN (ambient non-packaged), AP (ambient packaged), FN (fridge non-packaged), FP (fridge packaged), I (incubated non-packaged).
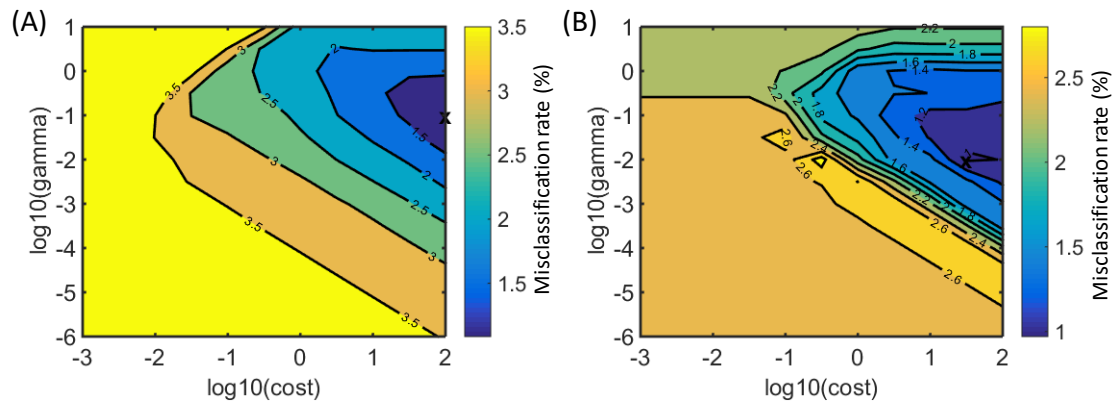

Figure SF.4: SVM parameters optimisation using cross-validation for all treatment groups grouped by time.

Table SF.1: Classification metrics for SVM: average accuracy (AC), average sensitivity (SENS) and average specificity (SPEC).

Dataset	Subset	AC (%)	SENS (%)	SPEC (%)
Surface measurement	Training	99	98	100
	Cross-validation	93	88	98
	Test	92	86	98
Liquid biopsies	Training	93	82	98
	Cross-validation	85	73	97
	Test	84	74	95

Table SF.2: SVM confusion matrices in the test set. Green shading indicates correct classification.

|									
Exocarp									
	Day 1	**Day 3**	**Day 5**	**Day 7**	**Day 10**	**Day 12**	**Day 14**	**Day 17**	**Day 20**
Day 1	97%	0%	0%	0%	0%	0%	0%	0%	3%
Day 3	0%	89%	2%	2%	7%	0%	0%	0%	0%
Day 5	0%	0%	98%	2%	0%	0%	0%	0%	0%
Day 7	0%	10%	0%	90%	0%	0%	0%	0%	0%
Day 10	0%	0%	0%	0%	98%	0%	0%	0%	2%
Day 12	0%	0%	0%	0%	0%	78%	16%	6%	0%
Day 14	0%	0%	0%	0%	0%	9%	75%	16%	0%
Day 17	0%	0%	0%	0%	0%	0%	31%	69%	0%
Day 20	8%	0%	0%	0%	6%	0%	0%	0%	86%
Locular gel									
Day 3		82%	0%	18%	0%	0%	0%	0%	0%
Day 5		0%	86%	14%	0%	0%	0%	0%	0%
Day 7		13%	0%	87%	0%	0%	0%	0%	0%
Day 10		0%	0%	0%	100%	0%	0%	0%	0%
Day 12		0%	0%	13%	0%	7%	53%	27%	0%
Day 14		0%	0%	9%	0%	0%	46%	45%	0%
Day 17		0%	0%	0%	0%	0%	13%	83%	0%
Day 20		0%	0%	0%	0%	0%	0%	0%	100%
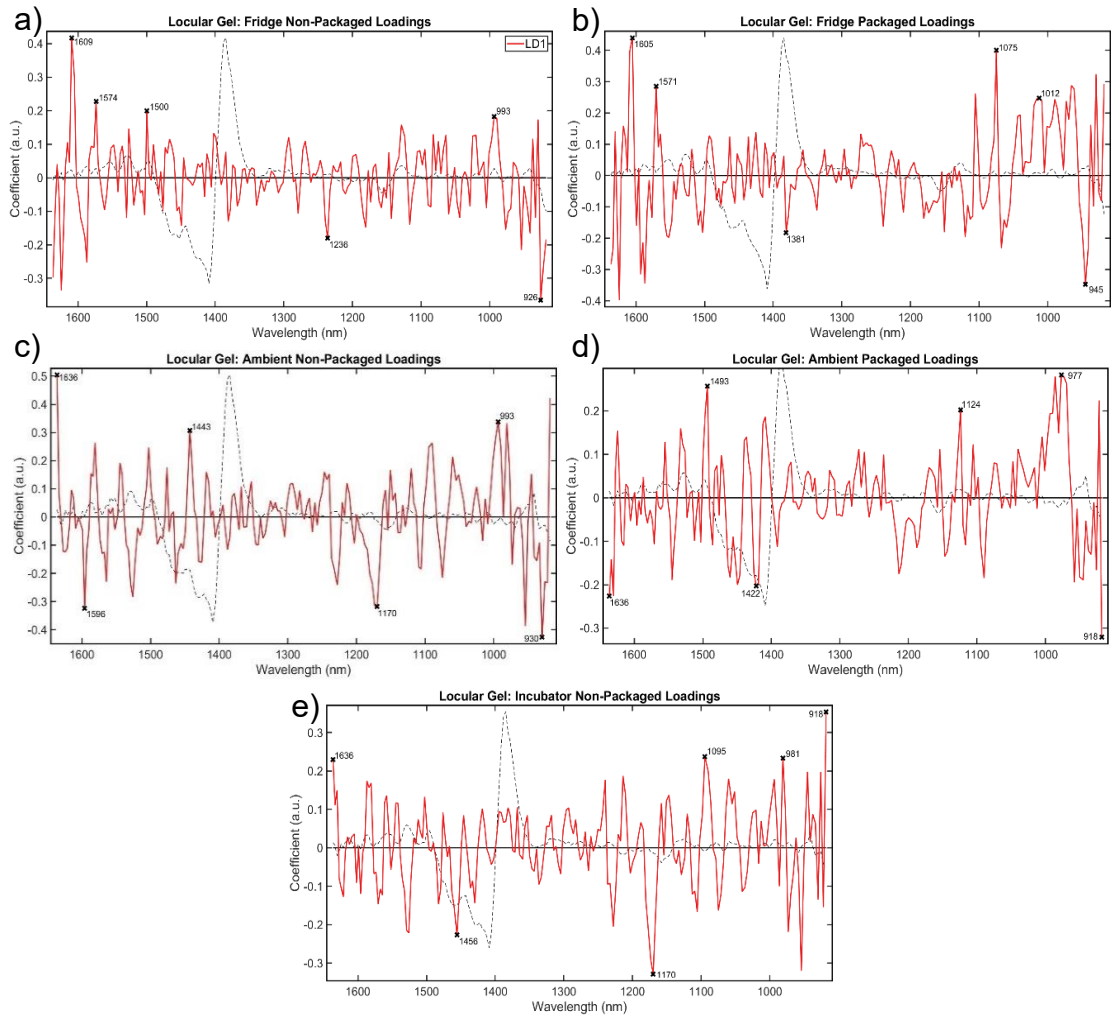
Figure SF.5: PCA loadings for locular gel measurements by storage condition: **a)** fridge non-packaged, **b)** fridge packaged, **c)** ambient non-packaged, **d)** ambient packaged, **e)** incubated non-packaged.
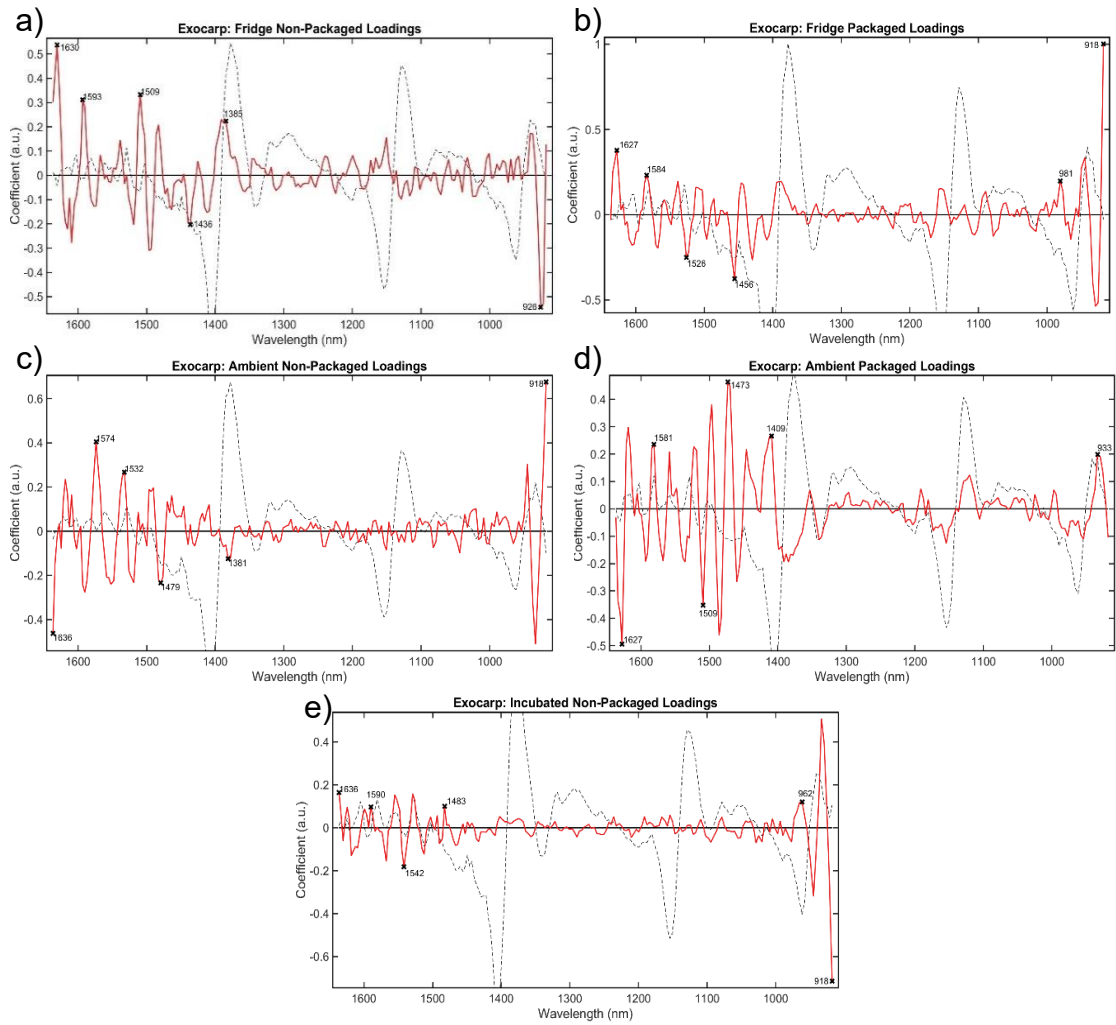
Figure SF.6: PCA loadings for exocarp measurements by storage condition: **a)** fridge non-packaged, **b)** fridge packaged, **c)** ambient non-packaged, **d)** ambient packaged, **e)** incubated non-packaged.
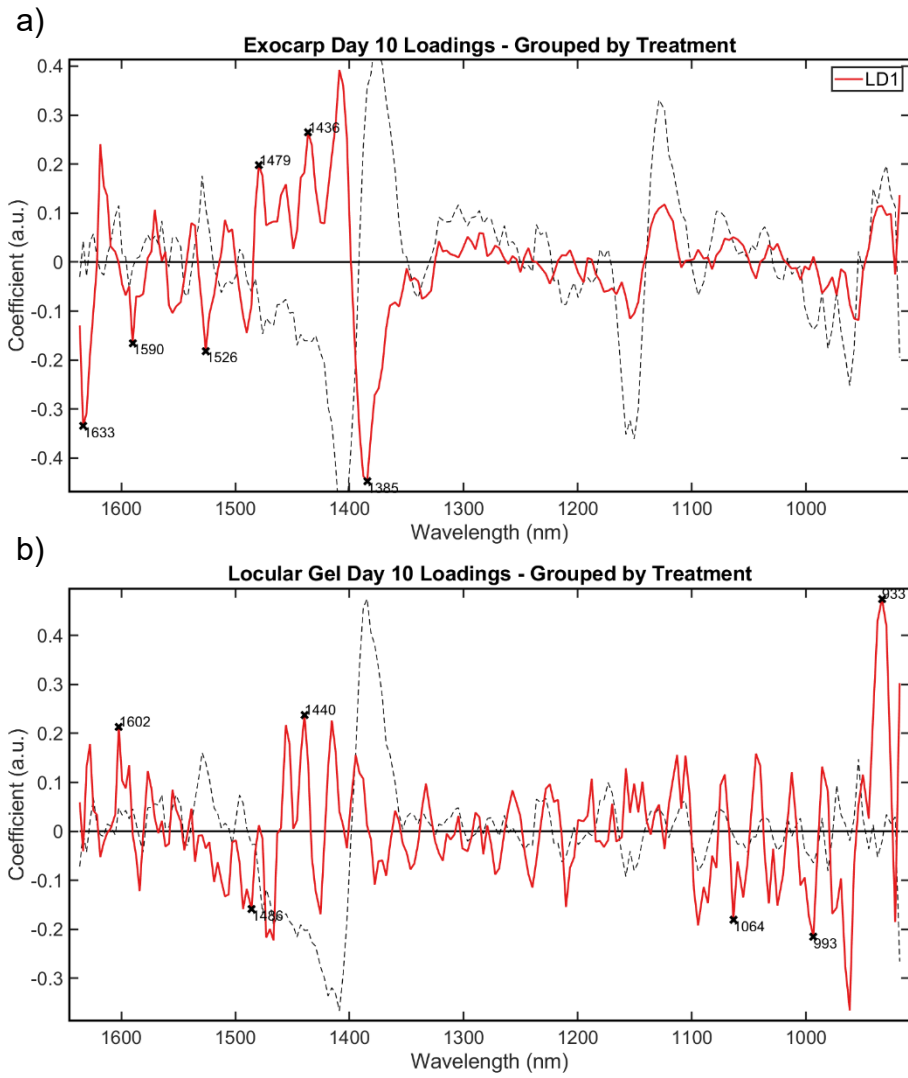
a)

b)

Figure SF.7: PCA loadings for locular gel and exocarp measurements at day ten for **a)** exocarp and **b)** locular gel.