From the root of variation: A metabolomics perspective to plant soil-feedback

Huberty, M.D.

Chapter 1

General introduction

Soils are a critical component for most life on earth and are not only a substrate for plants to grow in, but also harbour a plethora of macro and microorganism, that influence plants (Van der Heijden et al., 2008). These interactions greatly affect plants in many ways and are decisive factors for a plant's survival. For the last decades, differences in soil abiotic factors, such as humidity and nutrients, and their effect on plant performance have been thoroughly investigated (Chadha et al., 2019; Patterson, 1995; Schulze, 1986). More recently, this focus has shifted to examining the effects of soil microbes on plants. Organisms such as plant growth promoting bacteria or mycorrhizal fungi can enhance plant growth or induce resistance against plant antagonists such as insect herbivores or pathogens (Pineda et al., 2010; Van der Putten et al., 2016). On the other hand, soil also contains detrimental organisms such as pathogenic fungi and bacteria which negatively influence plant growth. The interplay between these "positive" and "negative" soil-born organisms is crucial for the performance of a plant (Van der Putten et al., 2016)

Plant soil feedbacks

Interactions of plants and soil microbes are complex and the growth of a plant in a soil leads to species-specific, abiotic and biotic changes in the soil. These species-specific changes remain present in the soil and can thereby influence the performance of plants growing later in the same soil (Bezemer et al., 2010, 2006; Van de Voorde et al., 2011). This process is called plant-soil feedback (PSF) (Bever, 1994; Van der Putten et al., 1993). If the biomass of a plant increases, relative to a control situation, when grown in a soil conditioned by a plant species, we refer to this as positive plant soil feedback. In contrast, if the biomass is lower it is referred to as negative feedback (Bever, 1994; Van der Putten et al., 2013). Most plant species exhibit negative conspecific PSFs, i.e. grow less well in "own" soil conditioned by an individual of the same species than in "foreign" soil (Bever, 2003; Kulmatiski et al., 2008; Petermann et al., 2008). The scientific interest in PSF has increased rapidly in the last decade (Forero et al., 2019; Smith-Ramesh and Reynolds, 2017). Nevertheless, so far the main focus has been on the effect of conspecific PSF although heterospecific PSFs may as well be of great importance in mixed plant communities (Van der Putten et al., 2013; van de Voorde et al., 2012a). While the traditional definition of PSF is strongly related to biomass, other traits such as herbivore resistance or the concentration of secondary compounds in the roots or shoots can also be largely influenced by PSF (Huberty et al., 2020b; Ristok et al., 2019; Zhu et al., 2018).

The mechanisms that can cause PSFs are numerous. For example, plant pathogens can accumulate in soils conditioned by certain plants species and this can lead to a reduction in biomass in the feedback phase (Mills and Bever, 1998; Nijjer et al., 2007; Van der Putten et al., 2013). Alternatively, plants may accumulate symbionts, which can cause an increase in biomass of the later growing plant e.g. through releasing enzymes cause an increase in nutrient availability in the soil (Reynolds et al., 2003; Van der Putten et al., 2013). Chemicals released by plant roots can cause shifts in the microbial community in the soil, which can influence later growing plants, while these chemicals in the soil can also directly influence the growth of plants in the feedback phase (Bais et al., 2003; Hu et al., 2018; Inderjit et al., 2011; Vivanco et al., 2004). PSF can also arise from extracellular self-DNA which can reduce plant growth (Mazzoleni et al., 2015). Furthermore, while plants grow in soils they deplete nutrients in the soil and this can negatively affect later growing plants (Lepinay et al., 2018). Of all these mechanisms shifts in microbial communities are among the primary causes for PSF effects.

While many studies focus on PSF effects on the biomass of plants, to better understand the impacts of PSFs on deleterious and beneficial organism the plant is associated to (e.g. pathogens above and belowground, herbivores, beneficial microbes, pollinators), it is of great importance to also understand how PSFs influence the chemical composition of a plant. The overarching aim of this thesis is to examine how plant-mediated changes in the soil, influence the metabolome of other, later growing plants.

Microbial influence on the metabolome of plants

Several studies have examined the effects of soil microorganisms on plant metabolomes. In a study published in 2014, Schweiger et al. showed that the leaf metabolome of several plants (*Plantago lanceolata*, *P. major, Veronica chamaedrys*, *Medicago truncatula* and *Poa annua)* differed when these plants were inoculated with a generalist arbuscular mycorrhizal fungus (AMF). However, the responsiveness of the metabolome to mycorrihizal inoculation largely differed between plant species. *M. truncatula* was the most responsive to AMF inoculation with 14.7% of chemical features (putative compounds) changed. On the other hand, the grass *P. annua* showed a minimal response with only 1.7% of the features being modulated by AMF. In all other species around 5% of features were modulated due to AMF. Almost all changes associated with AMF inoculation were quantitative (i.e. variation in concentration) and only

fewer than 1% of the features were only present in either the inoculated or the non-inoculated plants (qualitative). The direction of the quantitative changes as well as the individual features that changed varied between species.

Hill and colleagues (2018), investigated the metabolomic responses of root and shoot tissues of the plant *Jacobaea vulgaris* after inoculation with the AMF *Rhizophagus irregularis*. They showed that 33 compounds detected in the roots, most of them blumenols, significantly increased in plants that were exposed to AMF. The concentration of four pyrrolizidine alkaloids (PAs) increased in the roots after exposure to AMF but not in the shoots. This study showed that changes in the soil, e.g. presence or absence of AMF, can have distinguished effects on the metabolome of plants above and belowground.

Other studies investigated the effects of rhizobacteria and endophytes on plant chemistry, however those studies did not use soil that was inoculated but directly inoculated the plants under investigation. Pandey and colleagues (2016) showed that the inoculation of *Catharanthus roseus* with fungal endophytes upregulated terpenoid indole alkaloid (TIA) pathway genes potentially leading to changes in indole alkaloid concentrations above ground. Another study showed that flavonoids and anthocyanins in the leaves were found in higher concentrations in *Epichloe festucae*, when the plant was infected with endophytes (Dupont et al. 2016). Inoculation of sterile in vitro-grown poplar plants with *Paenibacillus sp*. changed the metabolomic profile of the shoots (Scherling et al. 2009). Specifically, eleven metabolites were different between non-inoculated and inoculated plants. Concentrations of asparagine, urea and threitol increase upon inoculation and organic acids as well as amino acids and fructose-6-phosphate were found in lower concentrations.

Effects of soil inoculation on the metabolome of plants

Badri and colleagues (2013) inoculated *Arabidopsis thaliana* with different soil slurries and analysed how these influenced the growth of the plant, the leaf metabolome, and the feeding behaviour of the caterpillar *Trichoplusia ni*. As inocula the autors used soil slurries collected from fields with different agricultural histories that were either filtered or non-filtered in order to differentiate between the influence of microbial and chemical factors of the soil. Arabidopsis biomass above and belowground as well as the leaf metabolome changed depending on the

legacy of the soil. The authors pyrosequenced the soils and analysed the nutrient content of the inoculates. Inoculation altered the concentrations of different chemical groups (e.g. sugars, amino acids, phenolics, sugar alcohols) and some of the inocula caused a decrease in the larval weight of *T.ni* when they fed from these plants. Since the nutrients in the inoculates did only vary mildly the observed changes were mostly likely due to differences in the composition of microbial communities within the inoculates. The decreased larval weight was negatively correlated with an increase in the abundance of bacterial taxa (e.g. Balneimonas, Skermanella, Nocardioides) in a principal component analysis. This study showed for the first time that the microbial composition of the soil is important for plant growth, herbivore performance and metabolome composition.

Kos and colleagues (2015c) studied how soil conditioned by different plant species influenced the growth, the chemical composition in the phloem and aphid performance on the plant *J. vulgaris.* They showed that growth by different plant species caused changes in soil fungal communities and that the biomass, the concentration of amino acids and PAs in the phloem, as well as the performance of a generalist aphid and a specialist aphid all varied depending on the soil in which *J. vulgaris* was growing. In this study the chemical analysis focused on specific compounds. However, as only targeted compounds were analysed, these studies may have overlook changes that occur in the non-targeted part of the metabolome of plants. Changes in these plant compounds could also influence interactions of these plants with other organisms. Furthermore, targeted approaches require previous knowledge of the metabolomic composition of the plant species under investigation which is not always available. Ecological studies often encompass different plant species or genotypes. Variation in targeted compounds such as secondary compounds might even vary considerably between different genotypes of one species although variation in primary compounds which are likely also involved in the interactions between plants and other organisms, is less pronounced among genotypes and more universal among plant species (Castro-Moretti et al., 2020). While this seems generally true for plants grown under controlled conditions, more variation is observed for plants collected in the field (Nagler et al., 2018). Interestingly changes in the metabolome of the plants growing in soil with particular microbes are not limited to secondary plant metabolites. Changes in the microbial communities in the soil can result in different concentrations of primary compounds in plants (Sampaio, Edrada-Ebel, & Costa, 2016). Technical advances

make it now possible to examine entire metabolomic profiles and therefore can provide a more complete picture of the response of a plant to changes in its environment.

The plant metabolism

The metabolism of a plant consists of a complex of dynamic processes, such as photosynthesis, respiration and the synthesis and degradation of organic compounds and all of these processes are linked to each other (Ratcliffe and Shachar-Hill, 2001; Vickery and Vickery, 1981). The foundation of all these processes is photosynthesis which produces substrates for respiration. By that it provides the building blocks of organic compounds such as amino acids, nucleic acids, proteins, carbohydrates, organic acids, lipids, and secondary compounds. Traditionally the plant metabolome is divided into the primary metabolism, which is required for a plant's survival and the secondary metabolism which main purpose is the defence of the plant (Berenbaum, 1995). Secondary plant compounds are chemically highly diverse and differ largely between different plant species (Moghe and Last, 2015; Weng et al., 2012) Primary metabolites, in contrast, are traditionally thought to be conserved across different plant species (Pichersky and Lewinsohn, 2011). However, primary metabolites can also play a role in defence, for example against pathogens and herbivores (Berenbaum, 1995). Plant hormones often act as signalling molecules and prime primary pathways in the plant. Furthermore, they can regulate plant development and its metabolic status (Beckers et al., 2016). The composition of the chemical defence compounds in a plant depends on the environment of the plant. For example, microorganisms in the soil can trigger the production of defence compounds in above and belowground plant tissues (Schweiger et al., 2014; Zhu et al., 2018). Defences are typically divided into constitutive and induced defences (Hulten et al., 2006). Constitutive defences describe defences with compounds which are present in the plant all time. Induced defences describe compounds of which production increases upon a stress (e.g. herbivores or pathogens) leading to higher concentrations of these compounds in the plant.

Variation in space

Plants protect themselves from pathogens and herbivores above and belowground by producing chemical compounds and adapting their chemical composition to the local environment. However, the threats and the beneficial organisms, that plants can deter or attract with certain compounds vary considerably over space. Since plants cannot escape those unfavourable conditions, they must cope with the situation the face, and they do so by adapting their chemical composition to the local circumstances. For example, Kleine an Müller (2011) showed that the chemical composition of *Tanacetum vulgare* varied strongly within a small spatial scale of 1.5 meters, probably due to selection pressures of different specialist and generalist herbivores in the area, and these chemical changes can lead to species-specific distribution of herbivores over space. Not only herbivores differ on small spatial scales but also the soil microbial community varies largely between sites and within single locations. The microbiome of the soil can differ at scales from millimetres to centimetres (Ettema and Wardle, 2002; Fierer, 2017; O'Brien et al., 2016). Soil microbial communities are known to affect the growth of plants and the composition of plant communities (Heinen et al., 2020; Wang et al., 2019b). However, if and how strongly these spatial differences in the microbiome are related to differences in chemical composition of plants remains unknown so far.

Variation of PSF in time

All the mentioned causes for PSF, such as changes in the microbial communities in the soil, changes in soil chemistry, or changes in the concentration of plant exudates and self DNA, vary over time (Hannula et al., 2019a; Hu et al., 2018; Mazzoleni et al., 2015; Waring et al., 2015). We may expect that the longer a specific plant grows in the soil, the stronger the influence of this plant on the soil in which it is growing. These soil legacies can then influence properties of plants which grow in the same soil afterwards. Through these changes in the soil over time, the duration that a plant conditions the soil could potentially lead to alterations of the metabolome of plants grown in theses soils and the performance of insects on these plants. This is relevant for predicting the effect of PSFs on ecosystems and to understand how plants change the growth of other plants via the soil. A current debate among ecologists focuses on the reproducibility of PSF (De Long et al., 2019). Experiments that study the strengths of PSFs at different time points are urgently needed. Recently, we sampled the soils of six forb and

grass species over the course of one year and followed bacterial and fungal communities over time (Hannula et al., 2019a). In that study we show that fungal communities in the soil are more stable over time than bacterial communities, which showed considerable changes between time points. Fungal communities also varied, but these changes were more related to the functional groups (grass/forb) of the monocultures that were growing in the soil than to temporal changes (Hannula et al., 2019a). In this thesis, I used these soils conditioned by monocultures to investigate how the soil legacies influence the metabolome and the biomass of *J. vulgaris* plants growing in these soils over the course of a year to examine how stable PSF patterns are over time.

Methods to analyse metabolomes

The effect of bioprocessing is reflected in the metabolic pool, the so-called metabolome. Traditionally, responses of plants to environmental challenges, such as changes in abiotic or biotic factors in the environment, are measured by recording the changes of a few targeted metabolites focusing on some plausible biosynthetic pathways. However, these targeted approaches can miss alterations in the plant simply because the change caused by the environmental challenge might not or not only occur in the targeted compounds. In contrast to these targeted analyses that require at least some prior knowledge about the species under investigation, untargeted metabolomics enable to explore changes in the whole metabolome without prior chemical knowledge. Up to today most plant metabolomics research focuses on model plant species (e.g. *Arabidopsis thaliana*, tomato and maize) (Kuhlisch and Pohnert, 2015) and a few secondary groups of compounds (Pyrrolizidine alkaloids, Glucosinolates). However, for many plant species no prior knowledge about their chemical composition is available. Therefore, untargeted metabolomics represent an ideal tool to investigate the chemical composition of non-model species and by that opens up a new tool to ecologists to analyse how interactions in nature influence the metabolome of plants.

Metabolomics is a field that emerged due to the technological advances in chromatography during the last decades with the ultimate goal to extract and detect all compounds in an organism. It is currently used in a broad range of life sciences ranging from medicine to biology. It is the omics technique which investigates the end product of the whole cellular

machineries and is therefore the omics technique which gives the most realistic picture of the situation that an organism interacting with a plant will encounter.

Ecometabolomics is a young research branch of metabolomics and describes the use of metabolomics techniques to assess the effect of an environmental gradient or stress on the metabolic state of an organism. The application of metabolomics techniques in ecology focus on understanding the biochemical mechanisms that drive species interactions with other organisms and with the environment (Bundy et al., 2009; Peñuelas and Sardans, 2009; Sardans et al., 2011; Van Dam and Meijden, 2011). Today it is also commonly used to explore the metabolic response of one or more organisms which interact and their variation in time and space (Peters et al., 2018). Metabolites within plants are needed to survive, reproduce, and interact with other individuals and organisms in the environment and are adjusted to the environmental conditions continuously. They represent the language that plants speak and adjustments to the current specific situation can be measured and identified with metabolomics techniques. Metabolomic analyses can be based on metabolic fingerprinting represents an untargeted approach supplying an overview of the differences in the metabolome but not necessarily identifying compounds within the organism, or on a semi-targeted approach, metabolite profiling, which focuses on specific groups of metabolites within samples (Hall, 2011).

Despite the continuous advancements in techniques and analysis used in metabolomics this young research field is still facing multiple challenges. The ultimate goal of metabolomics is to extract and detect all compounds in a certain organism at a given time (Stobiecki and Kachlicki, 2005). However up to today the field is a long way from reaching this ultimate goal. The first analytical step in metabolomics studies, untargeted and targeted, is the extraction of the compounds within the organism with an extraction solvent. However, no solvent can truly extract all compounds. Water is a universal solvent which, in plant metabolomics, is often used for the extraction of primary compounds such as sugars amino acids, organic acids and phosphorylated compounds (Cocuron et al., 2014). Other solvents such a methanol are more effective in the extraction of secondary compounds from plant leaves (Cocuron et al., 2019). A solution that is often applied is to extract with mixtures of methanol and water (Kim et al., 2010). The next step is to choose the suitable method to analyse the extract sample. The most common methods used for metabolomics analysis are liquid chromatography coupled with

mass spectrometry (LC-MS), gas chromatography coupled with MS (GC-MS) and nuclear magnetic resonance spectroscopy (NMR) (Peters et al., 2018). All of the methods bear advantages and disadvantages. LC-MS allows for detection of polar to medium polar compounds, but often lacks a suitable database and absolute quantification of compounds is not possible. In GC-MS the samples need to be derivatized and heated in order to be detectable and injectable. This process can alter compounds in the samples, and detected compounds might differ from the ones originally contained in the sample. The size of compounds analysed with GC-MS is also limited, since they need to be dissolvable in the gas phase, which excludes large compounds, because of their high molecular weight. NMR covers a broad range of compounds which can be detected, but the identification of individual signals is challenging and requires experience due to overlapping signals in mixtures. As opposed to other methods absolute quantification is possible with NMR.

Especially ecometabolomics requires the analyses of complex chemical samples, and many different factors might influence the composition of the plant under investigation. In the light of this complexity multiplatform approaches can be useful to cover a large spectrum of compounds with different chemical properties.

Thesis outline

In this thesis, I examine if plant-induced changes in the soil influence the metabolome of plants which grow later in these soils, and how these effects differ in time and space. Furthermore, I examine how multiplatform metabolomics can be used for ecometabolomics studies.

A graphical abstract of the research questions addressed in the experimental chapters 2 to 5 can be found in Fig. 1.1.

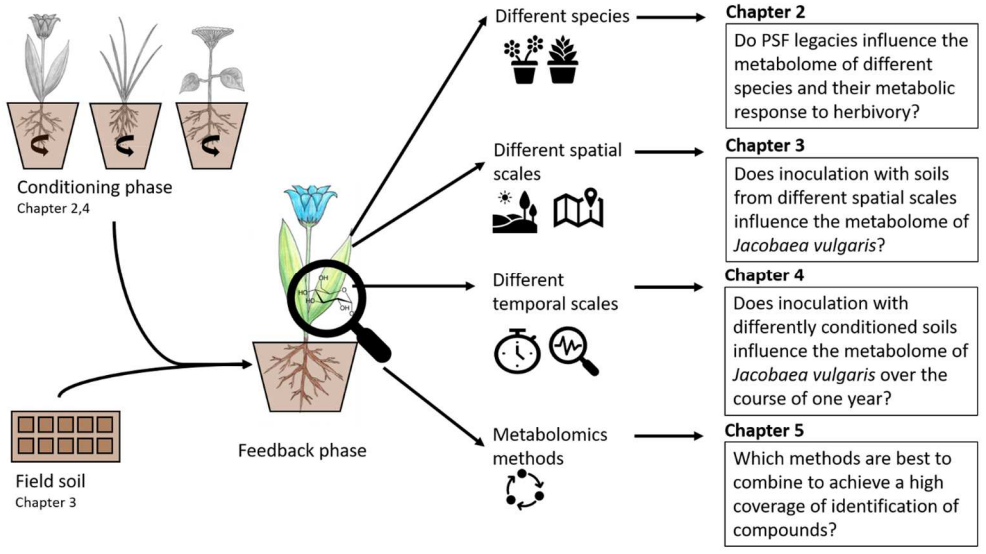

Figure 1.1 Schematic overview of the experimental chapters of this thesis. Plant soil feedback (PSF) experiments were carried out by first conditioning the soil in the conditioning phase. This led to abiotic and biotic changes in the soil with which sterilised soils were inoculated in the feedback phase. We evaluated how these changes influenced the metabolome of 12 different species and their response to herbivory in chapter 2. In chapter 3 the variation of inoculation with field soil collected across different spatial scales on a model species (*Jacobaea vulgaris*) was accessed. In chapter 4 temporal variation of PSF and their effects on metabolomes of *Jacobaea vulgaris* was accessed. In chapter 5 the possible methods for metabolomics are compared and combined on the example of *Taraxacum officinale*

In **Chapter 2**, I examine the effect of PSFs on the metabolomics of a number of forb and grass species. In a fully crossed design, we grew 12 species in soils inoculated with conditioned soils of all 12 plants species. We then exposed the plants to herbivory by a caterpillar (*Mamestra brassciae*) or not. The plants were then harvested, and the shoots were analysed with untargeted metabolomics with NMR. The metabolic response to the soils was species specific. Surprisingly, the soil conditioning predominately influenced the concentration of primary compounds such as sugars and amino acids. Remarkably, the soil treatment explained more of the variation of the metabolome of the plants than herbivory, for 7 of the 12 plant species, even though herbivory is generally known to have a large impact on the metabolome of plants. This chapter proves that PSFs influence the metabolome of plants. However, the sensitivity of the metabolome and the compounds that change are species specific. Moreover, even though I expected strong interactions between the effects of the soil and herbivory treatments, in this study, only for one species did the metabolic response to herbivory depend on the soil treatment.

For the study presented in **Chapter 3**, I collected soil from four different natural grasslands in the Netherlands. In each grassland, multiple soil samples were collected according to a spatial gradient. I then grew identical clones of *J. vulgaris* in sterilized soil inoculated with the different soil samples. Based on current knowledge I expected the soil samples to inhabit different microbial communities and I hypothesized that metabolomic profiles of these *J. vulgaris* plants will depend on the soil origin of the inocula and differ between sites. Within each site I expected that the metabolomic profile of soil samples collected close together (30 cm) would be more similar than samples collected further apart (100 m). Inoculation with soils from one of the four grasslands altered the metabolome of *J. vulgaris*, compared to plants grown in sterile soil but I did not detect a distance relationship. However, similarity analysis revealed that the metabolomes of plants grown with soil inocula that originated from the same site were more similar than the metabolomes of plants grown with inocula collected from different sites.

In **Chapter 4**, I investigate the temporal dynamics of PSF effects on plant metabolomics by repeatedly examining plant growth, and metabolomic composition of genetically identical *J. vulgaris* plants and by determining herbivore performance on these plants for an entire year. We set up an experiment in which we conditioned soils with monocultures of six different

plant species (including *J. vulgaris*) in large containers outdoors in a common garden. Every two months I collected soil from these containers and grew genetically identical *J. vulgaris* plants in sterile soil inoculated with soil from the containers in a climate-controlled growth chamber. After 6 weeks of growth I determined the shoot and root biomass and the performance of *M. brassicae* on leaves of the plants and analysed the shoot metabolome with untargeted metabolomics. The relative response, in terms of biomass and herbivore performance, to the soil conditioning treatments was similar over the course of the year. However, the effects of soil conditioning on the metabolome varied greatly over time. Furthermore, we show that the changes in the metabolome are more strongly correlated to changes in the bacterial communities than the fungal communities of the soil. The changes in the microbiome of the soil and the metabolome were correlated strongest after 8 months of conditioning.

In **Chapter 5**, I used three commonly used methods in metabolomics. ^{1}H NMR, LC-MS and HPTLC to do metabolic fingerprinting of *Taraxacum officinale* plants that either experienced herbivory of not. All three methods revealed that the metabolome clearly differed between plants which experienced herbivory or not. However, the identification of compounds was hindered by the chemical complexity of the extracts. Therefore, in a next step we used HPTLC as a preparative tool, enabling us to separate the extracts into fractions with distinct chemical properties. In the next step these fractions were analysed with LC-MS and GC-MS. By that we were able to detect campesterol which changed depending on the treatment and could not be detected with any of the conventional methods. We show that each chemical method bears advantages and disadvantages and that HPTLC can be a useful tool to concentrate compounds from mixtures and therefore enabling the identification of compounds, which would have been overlooked otherwise.

In **Chapter 6**, I discuss the results of the experimental chapters in a broader context and highlight a number of potential areas for future research.