A Study in Green:
Investigating Secondary Metabolite Profiles in Thirteen Grassland Plant Species

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I
Biowissenschaften

der Martin-Luther-Universität Halle-Wittenberg,

vorgelegt

von Frau Susanne Marr
geb. am 24.11.1989 in Potsdam

GUTACHTER 1: HELGE BRÜHLHEIDE
GUTACHTER 2: STEFFEN NEUMANN
GUTACHTER 3: SERGIO RASMANN

DATUM DER VERTEIDIGUNG: 11.07.2023

© 2023 S. Marr
The copyright of this thesis rests with the author of this thesis. Presented data and quotations from it should not be published without the author's prior consent. Information derived from it should be acknowledged. Chapter 1 has previously been published in a peer-review journal; all rights rest with the authors.

"ONE'S IDEAS MUST BE AS BROAD AS NATURE
IF THEY ARE TO INTERPRET NATURE" SHERLOCK HOLMES
A Study in Scarlet, by Arthur Conan Doyle; Chapter 5

A Study in Green:
Investigating Secondary Metabolite Profiles in Thirteen Grassland Plant Species

CONTENTS

SUMMARY

As ecometabolomics gains more and more interest, not just as a tool to investigate the metabolomic adaptations but also to bring them in context with the plant community level, the need for tailored and structured methods that effectively combine both the vast amount of metabolomics data and the natural diversity in ecological data increases rapidly. The standardised methods used in ecological experiments provide the opportunity to describe characteristics of plants on both the individual plant level, e.g. developmental stages and the community level, e.g. species diversity. Plant metabolite profiles directly reflect these environmental dynamics and provide the opportunity to investigate and gain insights into plant adaptations in different ecological contexts and schemes on yet another level. Especially secondary metabolites are often associated with responses to abiotic and biotic stress factors, such as seasonal changes and varying plant community neighbourhoods. However, experiments conducted in natural and semi-natural environments come with a combination of factors that can not always be accounted for. These unpredictable variations require tailored experimental designs, methods and data curation to capture the complexity of both ecological experiments and precise metabolomic data. Therefore, this study aims to combine methods and analysis strategies established for and successfully used in metabolomics and ecological studies to meet the custom requirements of ecometabolomic datasets to make metabolic profiles usable as standardised traits, especially regarding secondary metabolites.

This thesis used UPLC-ESI-Qq-TOF-MS to acquire metabolic profiles of the aboveground tissues in seven grass and six herb plant species across four seasons. The performance of different statistical methods (partial least squares, support vector machines and random forests) was tested for their usability, accuracy and predictive power on different matrices generated in this ecometabolomic study. These methods were tested on three independent data sets; two acquired in a field experiment and one in an outdoor greenhouse experiment. Plant samples were collected from neighbourhoods with different compositions, accompanied by trait records concerning both the individual plant and the plot community level. The data sets provided four different levels of metabolite profile complexity, ranging from distinct to similar profiles for comparison. Raw data and all workflows used to analyse the data sets in this thesis are published and available in the respective data repositories.

Chapter one of this thesis focuses on acquiring metabolite profile data of the aboveground tissues of thirteen grassland species in a semi-maintained field experiment (The Jena Experiment – Trait-based-experiment) from diverse plant communities across the growing season in 2017. This chapter mainly focussed on data curation, providing a comprehensive, reproducible workflow for pre-processing and curating the acquired data with missing data imputation, batch correction and validity checks on features and samples. The second chapter comprehensively demonstrates and discusses the challenges of such ecometabolomics studies. The ecometabolomic data sets used in this study demonstrate the usability of already established statistical methods on metabolomic data on four different levels with varying similarities of the profiles, starting with very distinct profiles in the functional group level and species level, down to similar profiles within the same species across the season and diversity levels. Depending on the metabolic profiles used within an experiment, guidance is provided to choose the most suitable statistical methods for the question. For the data handling, both the data amount from metabolomics analysis and dynamics from field experiments needed to be taken into account to investigate ecometabolomics data sets and to draw conclusions about underlying mechanisms reflected in the metabolite profiles. Here it was found that choosing a suitable classification method depended mainly on the complexity of the provided background for classification. Adapting existing methods to the needs of ecometabolomic data sets provided the opportunity to use more diverse data acquired for and relevant to ecological contexts. The third chapter focused on exploring the dynamics of the metabolic profiles of leaves in thirteen grassland plant species. This study investigated the chemical dynamics in the metabolomic fingerprints as a response to a changing neighbourhood regarding plant species richness and diversity caused by different seasons. The metabolite profiles across species were found to be very distinguishable compared to the within species profiles that were found to be more similar while still reflecting dynamics in seasonal and community changes.

This thesis focused on providing a road map for integrating metabolomics data to complex ecological environments and implementing reproducible, standardised methods across multiple experimental setups. Unravelling specific mechanisms reflected in the metabolic profiles will provide insights into underlying ecosystem functions.

ZUSAMMENFASSUNG

Da Ecometabolomics nicht nur als Werkzeug zur Untersuchung metabolischer Anpassungen, sondern auch zur Einordnung dieser im Kontext von Pflanzengemeinschaften immer mehr Interesse weckt, steigt der Bedarf an maßgeschneiderten und strukturierten Methoden, die sowohl die Vielzahl an metabolomischen Daten, als auch die natürliche Vielfalt an ökologischen Daten effektiv kombinieren können. Die standardisierten Methoden, die in ökologischen Experimenten verwendet werden, bieten die Möglichkeit, Eigenschaften von Pflanzen sowohl auf der individuellen Pflanzenebene, z.B. Entwicklungsstadien, als auch auf der Gemeinschaftsebene, z.B. Artenvielfalt, zu beschreiben. Die metabolomischen Profile von Pflanzen spiegeln direkt diese Umweltdynamik wider und bieten die Möglichkeit, Pflanzenanpassungen in verschiedenen ökologischen Kontexten und Schemata auf einer weiteren Ebene zu untersuchen. Insbesondere Sekundärmetabolite sind oft mit Reaktionen auf abiotische und biotische Stressfaktoren verbunden, wie z.B. saisonale Veränderungen und unterschiedliche Pflanzennachbarschaften. Experimente, die in natürlichen und semi-natürlichen Umgebungen durchgeführt werden, bringen jedoch Kombinationen von Faktoren mit sich, die nicht immer berücksichtigt werden können. Diese unvorhersehbaren Variationen erfordern maßgeschneiderte experimentelle Designs, Methoden und Datenpflege, um die Komplexität sowohl ökologischer Experimente als auch präziser metabolomischer Daten zu erfassen. Diese Dissertation zielt daher darauf ab, Methoden und Analysestrategien, die in der Metabolomik und in ökologischen Studien bereits etabliert und erfolgreich eingesetzt wurden, zu kombinieren, um den speziellen Anforderungen von Ecometabolomic Datensätzen gerecht zu werden und metabolische Profile als standardisierte Charakteristik nutzbar zu machen; insbesondere im Hinblick auf Sekundärmetabolite.

In dieser Arbeit wurde UPLC-ESI-Qq-TOF-MS verwendet, um metabolomische Profile der oberirdischen Gewebe von sieben Gras- und sechs Kräuterpflanzenarten über vier Jahreszeiten hinweg zu erfassen. Die Leistungsfähigkeit verschiedener statistischer Methoden (partial least squares, support vector machines and random forests) wurde auf unterschiedlichen Matrizen, die im Rahmen der vorliegenden Ecometabolomik-Studie generiert wurden, auf ihre Nutzbarkeit, Genauigkeit und Vorhersagekraft getestet. Diese Methoden wurden an drei unabhängigen Datensätzen getestet, von denen zwei in einem Feldexperiment und einer in einem Freiland-Gewächshaus-Experiment erhoben wurden. Pflanzenproben wurden aus Nachbarschaften mit unterschiedlicher Zusammensetzung entnommen und von Merkmalsaufzeichnungen begleitet, die sowohl die individuelle Pflanze als auch die Gemeinschaftsebene des Versuchs betrafen. Die Datensätze lieferten vier verschiedene Ebenen von metabolomischer Profilkomplexität, die von *sehr unterschiedlich* bis hin zu *sehr ähnlich* reichten. Die Rohdaten, sowie die erstellten Workflows für die Analyse der Daten sind in den jeweiligen Datenarchiven auf MetaboLights verfügbar.

Kapitel eins dieser Dissertation beschreibt die Erhebung von metabolomischen Profildaten der oberirdischen Gewebe von dreizehn Graslandarten in einem halbgepflegten Feldexperiment (Das Jena Experiment - Trait-based-experiment) aus verschiedenen Pflanzengemeinschaften während der Wachstumsperiode 2017. Dieses Kapitel konzentriert sich hauptsächlich auf die Datenkuration und lieferte einen umfassenden, reproduzierbaren Workflow zur Vorverarbeitung und Handhabung der erworbenen Daten mit Imputation von fehlenden Daten, Chargenbedingter-Messkorrektur und Validitätsprüfungen von Merkmalen und Proben. Im zweiten Kapitel werden die Herausforderungen von Ecometabolomics Studien umfassend demonstriert und diskutiert. Die hier verwendeten Ecometabolom Datensätze demonstrieren die Verwendbarkeit bereits etablierter statistischer Methoden auf metabolomische Daten auf vier verschiedenen Ebenen mit zunehmender Ähnlichkeit der Profile über die Ebenen hinweg, beginnend mit sehr unterschiedlichen Profilen zwischen funktionellen Gruppen und Arten bis hin zu ähnlichen Profilen innerhalb derselben Art über verschiedene Jahreszeiten und Diversitätsstufen. Abhängig von den verwendeten metabolomischen Profilen innerhalb eines Experiments wird ein Vorschlag zur Auswahl der am besten geeigneten statistischen Methoden für die gestellte Frage bereitgestellt. Für die Datenverarbeitung mussten sowohl die Datenmenge aus der Metabolomik-Analyse als auch die Dynamik aus Feldexperimenten berücksichtigt werden, um Ecometabolomik-Datensätze zu untersuchen und Schlüsse über zugrunde liegende Mechanismen zu ziehen, die in den metabolomischen Profilen widergespiegelt werden. Die Auswahl einer geeigneten Klassifizierungsmethode hing hauptsächlich von der Komplexität des bereitgestellten Hintergrunds für die Klassifizierung ab. Die Anpassung bestehender Methoden an die Bedürfnisse spezifizierter Datensätze bot die Möglichkeit, vielfältigere Daten zu verwenden, die für ökologische Zusammenhänge relevant sind. Das dritte Kapitel konzentriert sich darauf, die Dynamik der metabolischen Profile von Blättern in dreizehn Graslandpflanzenarten zu erforschen. Diese Studie untersuchte die chemischen Dynamiken in den metabolomischen Fingerabdrücken als Reaktion auf eine sich ändernde Nachbarschaft in Bezug auf die

Artenvielfalt und Diversität, die durch verschiedene Jahreszeiten ausgelöst wurde. Die metabolomischen Profile zwischen den Arten waren im Vergleich zu den innerhalb der Arten gefundenen Profilen klar unterscheidbar, während sie dennoch die Dynamik saisonaler und gemeinschaftlicher Veränderungen widerspiegelten. Die vorliegende Dissertation hatte zum Ziel,einen Fahrplan für die Integration von Metabolomik-Daten in komplexe ökologische Umgebungen zu liefern und reproduzierbare, standardisierte Methoden über verschiedene experimentelle Setups hinweg umzusetzen. Die Entschlüsselung spezifischer Mechanismen, die in den metabolomischen Profilen widergespiegelt werden, wird Einblicke in die zugrundeliegenden ökologischen Funktionen ermöglichen.

INTRODUCTION

> "Criminal cases are continually hinging upon that one point. A man is suspected of a crime months perhaps after it has been committed. His linen or clothes are examined, and brownish stains discovered upon them. Are they blood stains, or mud stains, or rust stains, or fruit stains, or what are they? That is a question which has puzzled many an expert, and why? Because there was no reliable test…"
>
> Sherlock Holmes, A Study in Scarlet by A. C. Doyle, 1887

… except, now there is!

Metabolomics has become a powerful tool to identify compounds and unravel underlying metabolomic variations within biological systems (Weston et al. 2015). Describing the entirety of metabolites in a living organism (Fang et al. 2019, Walker et al. 2022), metabolomics usually refers to the systematic study of chemical compounds regulated by cellular biochemical processes (Weston et al. 2015). By reflecting the current state of a plant (Weston et al. 2015), metabolites serve as a biochemical bridge between a plants genotype, phenotype, and the environment (Fiehn 2002, Weston et al. 2015, Berini et al. 2018, Fang et al. 2019), balancing resources between fitness, reproduction and defence strategies (Walker et al. 2022). While primary metabolites support growth, reproduction, and stability (Obata & Fernie 2012, Sulpice & McKeown 2015, Weston et al. 2015, Fernandez et al. 2016, Sardans et al. 2020), secondary metabolites are mainly involved in defence mechanisms (Fang et al. 2019, Ober & Hartmann 2000, Quinn et al. 2014). Secondary metabolites are also of interest to food, health (Rai et al. 2017) and the cosmetics industry (Rochfort 2005), enhancing the general interest in elucidating and predicting underlying mechanisms (Gromski et al. 2015, Fang et al. 2019).

Since secondary metabolites hold a wide range of effects on the interaction of plants with their environment (Berini et al. 2018), using metabolomics within ecological experimental setups has the potential to improve the predictive power of ecological traits (Walker et al. 2022). Classical ecological traits, such as plant height, developmental stage and carbon/nitrogen contents, are commonly used to evaluate plant performance, particularly in connection with dynamics in plant species richness and diversity within plant communities (Scherling et al. 2010). Species loss and changing plant diversity, in general, are increasingly becoming one of the biggest drivers and motivations for ecosystem research (Hooper et al. 2005) and the urge to understand underlying processes that shape the networks in which plants interact with the constantly changing abiotic environment and the biotic community (Vandenkoornhuyse et al. 2015). Across these environmental dynamics, plants continuously need to balance the resources between their primary metabolism – dealing with growth and biomass production – and secondary metabolism, used for defence against herbivores and pathogens (Wright et al. 2004, Díaz et al. 2016, Gargallo-Garriga et al. 2020). The effect of environmental conditions on the metabolite profiles of plants is mainly influenced by water, light and nutrient availability, leading to plant neighbourhood composition dynamics and, as a result, pathogen and herbivore interactions (Scherling et al. 2010). Ristok et al. (2022), for example, showed that increasing plant species richness led to decreasing rates of herbivory in some species and showed a positive effect of diversity on primary productivity. While many diversity studies have shown positive relationships between species richness and primary productivity (Balvanera et al. 2006, Cardinale et al. 2007), Scherling

et al. (2010) point out that these findings can be highly species specific in certain compositions and biogeographic contexts and are not necessarily transferable to other systems.

Ecological questions are often answered by comparing functional traits (Walker et al. 2022) in field experiments (e.g. plant height, leaf area and carbon/nitrogen contents) to help understand drivers of temporal community stability and enable predictability of ecosystem services. Plant functional traits are standardised sets of characteristics that provide a metric for measuring and comparing species (Violle et al. 2007). These classical physiological traits are the tool of choice to describe and characterise organisms and their interactions with the environment, which might be limited in elucidating the underlying mechanisms of ecosystems. Visible functional traits are a combination of underlying physio-chemical processes. They are limited in their information content compared to high-resolution and high-throughput analysis techniques, e.g. metabolomics, making it challenging to predict them as single traits (Walker et al. 2022). Metabolic profiles, on the other hand, represent plant fitness between and within generations and are a great addition to classical traits (Walker et al. 2022). Hence, combining visible ecological traits with metabolomics is essential to elucidate the underlying mechanisms of plants' interaction with their environment. Ecometabolomics studies include metabolomics to answer ecological questions (van Dam & van der Meijden 2011) and support classical physiological traits (Walker et al. 2022). Previous studies showed the change in secondary metabolite abundances as a response to environmental changes in different species (Weston et al. 2015, Berini et al. 2018). These include both abiotic and biotic factors. For instance, plant metabolite abundances can change rapidly or progressively (Schuman & Baldwin 2016). Therefore, the distinction between short- and long-term responses is of particular interest. Assuming that plants have a fixed set of defence strategies (Zanne et al. 2014), we would also expect to find similar responses to similar environmental pressure factors. Scherling et al. (2010) were already able to show that similar herb species respond with similar strategies and changes to their metabolic profiles to species richness in the direct neighbourhood. For instance, plant neighbourhood diversity affects plant performance (Scherling et al. 2010, Chiapusio et al. 2018) and metabolite profiles in terms of richness and composition (Scherling et al. 2010, Ristok et al. 2022). One reason for these changes might lie in inducing the production of allelopathic compounds (Baldwin et al. 2006, Fernandez et al. 2016). Plant species diversity and metabolite diversity affect herbivory rates (M-25: Ebeling & Meyer et al. 2014, Ristok et al. 2022), broadening the range of potential host plants and enabling herbivores to locate potential host plants through specific metabolites (Agrawal & Weber 2015). Shifts in metabolite profiles can also occur due to resource limitations that reflect and alter plant species richness and diversity (Scherling et al. 2010) and influence fitness and productivity (Walker et al. 2022), which leads to environment-dependent profiles (Berini et al. 2018). Furthermore, secondary metabolite profiles change along temperature gradients (Berini et al. 2018, Defossez et al. 2021) and can depend on plant height (Scherling et al. 2010) and developmental stage (Mandal et al. 2010).

Furthermore, ecometabolomic studies often focus on model plant species and intra-species study designs in controlled environmental conditions, e.g. greenhouse setups (Walker et al. 2022). Those experiments usually aim at a targeted approach to understanding specific factors (metabolites) that are influencing or are influenced by their environment and to be able to confirm the specific function of certain compounds (Weston et al. 2015, Berini et al. 2018, Fang et al. 2019). These targeted techniques support the identification of single components that are actively involved in responses to the environment. However, the identification of metabolites requires the elucidation of the stereochemical and elemental composition of features and compounds (Fiehn 2002) and requires profound knowledge about the metabolites that are present in the analysed species and a good estimation of which metabolites will be affected. Comprehensive reviews of metabolomics and techniques commonly used can be found elsewhere in the literature (Ellis & Goodacre 2006, Nguyen et al. 2012, Tugizimana et al. 2013).

Some of the identified secondary metabolites are unique to certain plant families, for example glucosinolates (Grubb & Abel 2006, Agerbirk & Olsen 2012, Fang et al. 2019), which are unique to Brassicales (Fahey et al. 2001), while some serve a specific function, phenolic compounds (Schuman & Baldwin 2016), alkaloids, terpenoids and flavonoids, which are known to be involved in addressing drought, and temperature stress (Yonekura-Sakakibara et al. 2014, Fernandez et al. 2016, Yang et al. 2018), and UV tolerance (Peng et al. 2017), to name a few examples. However, the exact function of many metabolites (Weston et al. 2015, Alseekh & Fernie 2018) and how abiotic and biotic factors trigger responses on the metabolite profiles remain unexplained (Berini et al. 2018). Although targeted metabolomics has already revealed numerous functions of specific metabolomics and how they are involved in the complexity of environmental changes, untargeted metabolomics allows capturing the full range of metabolomic

compound classes in the context of the plant's physiological stage within a certain environmental condition (Fiehn 2002).

Many approaches and techniques are being made available to study metabolomics in complex systems and convert knowledge already gained from complex ecological experiments and highly precise metabolomics experiments to understand underlying mechanisms on a biochemical level. However, these proposed workflows and approaches contain different priorities for different objectives and differ in their statistical methods and analysis power. Tailored and conclusive techniques for identifying compounds and metabolite families are also not yet available for untargeted approaches with big data sets. In contrast to targeted analysis, untargeted analysis does not require each feature or compound to be annotated for statistical analyses (Scherling et al. 2010). Commonly for untargeted metabolomics data, feature richness and feature abundance are valuable data that can be acquired and used for statistics. The untargeted metabolomics approach allows quantifying a plant's whole measurable feature profile snapshot without the need for identification or classification (Scherling et al. 2010). Compared to targeted analysis of metabolites, the advantage of untargeted analysis is the power of untargeted analysis to investigate global changes in the metabolic profile without making assumptions about underlying mechanisms (Berini et al. 2018). Detected features in untargeted metabolomics have high statistical relevance, which enables automated data processing (Hoehenwarter et al. 2008). Untargeted metabolomics is a powerful tool to investigate links between the environment and biochemical processes (Walker et al. 2022), but it also holds some obstacles that need to be addressed adequately to be able to draw funded conclusions from the data (Allard et al. 2017).

Walker et al. (2022) point out that metabolite profiles, especially in the context of ecometabolomics, serve as a repository of functional traits, which reflect the plant state at a certain point in time (Weston et al. 2015), and can therefore serve as a valuable trait themselves (Walker et al. 2022). In eco-metabolomics, selecting relevant features from metabolite profiles can provide insights into plant phenotype adaptations to environmental dynamics. It might be, for example, a reliable predictor of herbivory pressures in plant communities (Ristok et al. 2022). Although ecometabolomics has become a major interest in plant research, there is no standard available to yet to unify study design, data acquisition and handling (Walker et al. 2022). Although metabolomic is widely used in ecometabolomics studies, there is not always taken special care in properly handling the data, including quality controls, blanks, randomisation strategies and data handling techniques that ensure proper data used for conclusion drawing. The collection of metabolomics data requires a careful study design and considered inclusion of environmental factors (van Dam & Heil 2011, Walker et al. 2022) and can easily lead to false assumptions when not acquired with certain care and level of background information, e.g developmental and environmental components (Fang et al. 2019). Metabolomics data also require quite some caution when interpreted, as effects can easily occur as artefacts during sample preparation and data acquisition (Fiehn 2002). One example is induced responses, e.g triggered by herbivory attack may change metabolic profiles throughout the plant or only in the attacked tissue (van Dam & Heil 2011); therefore, a special collection of material is required.

Furthermore, during sample collection, it is also important to conserve the current state of the metabolite profile, e.g. by snap freezing in the field (Marr et al. 2021). This ensures that all metabolomics processes are stopped in their tracks, and the sample presents an accurate picture of the current state of being. In contrast to classical plant traits that shift and change at a slower pace, the exact time of sample collection is an important element of the ecometabolomics study design and the research question (Weston et al. 2015). The sampling timepoint across the year strongly affects herbivory pressure on the plants and, thereby, the richness and composition of secondary metabolites (Ristok et al. 2022). To overcome the incomplete snapshot phenomenon, including multiple samplings across the year does improve the interpretation of the metabolomics data (Ristok et al. 2022).

As Weston et al. (2015) pointed out, one of the main challenges in metabolomics is the sheer number of secondary metabolites and their related compounds that can be measured in plants. Given the enormous amount of produced metabolites within plants (Fang et al. 2019), handling the measured data and gaining reliable insights is still a challenging task within ecological experiments and studies (e.g. Marr et al. 2021). While metabolomics measurements, including sampling and extract preparation, is relatively easy and quick, handling the vast amount of data generated with these methods can be challenging on a multitude of levels (Weston et al. 2015). Single plant species only produce a subset of known secondary metabolites (Fernie et al. 2004). Hence, multispecies studies can result in zero-inflated data matrices. The extent of the effects of plant species diversity on general performance can be highly species specific (Scherling et al.

2010). Therefore, changes in the metabolite profiles to change environmental conditions to be similarly species specific can be expected. Integration of multiple species within the same experiment can, therefore, lead to an increased chance of zero-inflated data matrices. Currently available statistical tools are not yet tailored to capture the complexity of ecological experiments, especially regarding the growing interest in the role of secondary metabolites. As Walker et al. (2022) discussed, extensive research has been done to understand the interplay of metabolite responses to environmental dynamics. Therefore, a common strategy that allows comparison and reuses collected data in broader studies is important.

Expecting that environmental dynamics, especially season related changes in the direct neighbourhood, will be reflected in the secondary metabolite profiles regarding the richness and composition of the features and compounds. Expecting that similar species respond with similar metabolites and strategies to environmental pressure, we tested thirteen species in the two functional groups (FG) grass and herb in this thesis. Therefore, this study aims to compare the performance of classical metabolomics feature selection and classification methods in eco-metabolomics data sets, which provide challenges regarding sample size and background noise; the combination of ecological questions in open study systems (Jena Experiments) with biochemical analysis methods (metabolomics) and bioinformatic tools (automated processing) to gain interpretation potentials on diverse levels; identify suitable and stable statistical methods that are commonly used in „classical“ metabolomics data and how they can be adjusted to be used with ecometabolomics data and to answer ecological motivated questions. This thesis aims to introduce an automated and reproducible workflow for metabolomics data handling, curation and analysis to enable the automated analysis of multiple species with highly diverse metabolite profiles across environmental dynamics.

CHAPTER 1

LC-MS based plant metabolic profiles of thirteen grassland species grown in diverse neighbourhoods

Title: LC-MS based plant metabolic profiles of thirteen grassland species grown in diverse neighbourhoods

Authors:
Sue Marr [1,2,3*], Jos A. Hageman[4], Ron Wehrens[4], Nicole M. van Dam[3,5], Helge Bruelheide[2,3] & Steffen Neumann[1,3]

1 Bioinformatics & Scientific Data, Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120, Halle, Germany.
2 Institute of Biology/Geobotany and Botanical Garden, Martin Luther University Halle-Wittenberg, Am Kirchtor 1, 06108, Halle, Germany.
3 German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstr. 4, 04103, Leipzig, Germany.
4 Biometris, Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB, Wageningen, The Netherlands.
5 Molecular Interaction Ecology, Institute of Biodiversity, Friedrich-Schiller University Jena, Dornburger-Str. 159, 07743, Jena, Germany.
*corresponding author

Author contribution:

Author contributions (ctb) to the manuscript in percent.

chapter	status	author	ctb
(1) LC-MS based plant metabolic profiles of thirteen grassland species grown in diverse neighbourhoods	published	S. Marr	70
		J. A. Hageman	12
		R. Wehrens	4
		N. M. van Dam	2
		H. Bruelheide	2
		S. Neumann	10

Published in: www.nature.com/scientificdata/; *peer-reviewed*

Link: https://www.nature.com/articles/s41597-021-00836-8

Citation: Marr, S., Hageman, J. A., Wehrens, R., van Dam, N. M., Bruelheide, H., & Neumann, S. (2021). LC-MS based plant metabolic profiles of thirteen grassland species grown in diverse neighbourhoods. *Scientific data*, *8*(1), 1-12.

LC-MS based plant metabolic profiles of thirteen grassland species grown in diverse neighbourhoods

Authors
Sue Marr[1,2,3], Jos A. Hageman[4], Ron Wehrens[4], Nicole M. van Dam[3,5], Helge Bruelheide[2,3], Steffen Neumann[1,3]

Affiliations
1 Bioinformatics & Scientific Data, Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle, Germany
2 Institute of Biology/Geobotany and Botanical Garden, Martin Luther University Halle-Wittenberg, Am Kirchtor 1, 06108 Halle, Germany
3 German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstr 4, 04103 Leipzig, Germany
4 Biometris, Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB, Wageningen, The Netherlands
5 Molecular Interaction Ecology, Institute of Biodiversity, Friedrich-Schiller University Jena, Dornburger-Str. 159, 07743 Jena, Germany

Corresponding author: S. Marr (smarr@ipb-halle.de, orcid.org/0000-0003-3978-0836)

Abstract

In plants, secondary metabolite profiles provide a unique opportunity to explore seasonal variation and responses to the environment. These include both abiotic and biotic factors. In field experiments, such stress factors occur in combination. This variation alters the plant metabolic profiles in yet uninvestigated ways. This data set contains trait and mass spectrometry data of thirteen grassland species collected at four time points in the growing season in 2017. We collected above-ground vegetative material of seven grass and six herb species that were grown in plant communities with different levels of diversity in the Jena Experiment. For each sample, we recorded visible traits and acquired shoot metabolic profiles on a UPLC-ESI-Qq-TOF-MS. We performed the raw data pre-processing in Galaxy-W4M and prepared the data for statistical analysis in R by applying missing data imputation, batch correction, and validity checks on the features. This comprehensive data set provides the opportunity to investigate environmental dynamics across diverse neighbourhoods that are reflected in the metabolomic profile.

Background & Summary

Plants respond and adapt to environmental changes in many ways. Some plant species, for example, possess physical defences to cope with herbivores and abiotic stress factors[1]. In addition, plants also produce chemicals as defence strategies. These plant metabolites provide a unique opportunity to explore these adaptations as the metabolic profile is known to reflect environmental changes[2-4]. Both the primary and the secondary metabolome are involved in the responses to biotic[5-6] and abiotic factors[7-9]. However, especially secondary metabolites, which are not directly involved in the primary metabolism, play a key role in plant defence strategies[5-6,10-12].

Furthermore, compared to primary metabolite profiles, secondary metabolite profiles are more species specific even in varying environments[13]. Previous studies showed that plants change the composition of their metabolic profile and alter the abundance and the number of specific compounds, such as phenolics and terpenoids[7,14] while maintaining their distinctive profiles[13,15]. In field experiments, the impact of abiotic and biotic factors vary across the season[16-17]. These factors include, for instance, light, nutrients, water and herbivory[18].

Changes in these conditions may affect the plants' metabolic fingerprint in yet uninvestigated ways. The investigation of these changes may provide insights into the mechanisms behind plant adaptation strategies.

Grasslands are an ideal study system to investigate the effects of plant community compositions on the plant metabolomic profiles. In these ecosystems, we find a relatively high number of fast-growing grass and herb species[19]. Species that share similar characteristics form functional groups (FG). Here, we distinguish between the two FG: grasses and herbs. Most studies focus on visible traits when investigating these two FG[20-22]. Visible traits, for example, are a useful tool to understand and predict ecological strategies and functions. They are also supporting the investigation of relationships between functional traits – that describe all measurable characteristics of a plant individual - and the individual plant performance[23-26]. However, the investigative power of combining metabolomics data with such trait data has already been demonstrated in other studies[14-15,27-29].

In this data set, we collected plant material for metabolomic analysis in the field experiment "The Jena Experiment: Trait-Based-Experiment", Germany[30]. An overview of the data set is provided in Table 1 and Fig. 1, including the experimental setup (①-⑤), metabolomic analysis (⑥-⑩), and the data processing (⑪-⑯). We recorded both visible traits and metabolomic profiles to investigate species specific responses of thirteen grassland species to the composition of their neighbourhoods. For the metabolomic analysis, we collected shoot material across the growing season in 2017 at four time points: May (A), July (B), August (C), October (D). We chose these time points to cover the whole growing season (May to October; Fig. 1 ①). The sown (target) species belonged to the FGs grasses and herbs (Fig. 1 ②, Fig. 2 a). We investigated plants grown in communities with diversity levels (DL) composed of one (DL1), two (DL2), four (DL4) and eight (DL8) different species (Fig. 1 ③, Fig. 2). We collected shoots of two replicates per DL and species (Fig. 1 ④-⑤). For each species, we recorded characteristics of their surrounding neighbourhoods, including the number of plant species and their abundances per plot. For each sample, we recorded visible traits, such as the plant height, number of leaves and the level of damage caused by herbivory or pathogens.

In total, we collected 512 samples. For each sample, we acquired the metabolic profiles of methanolic extracts of the shoots on an Ultra Performance Liquid Chromatography coupled with an Electrospray Ionisation Quadrupole Time-of-Flight Mass Spectrometry (UPLC-ESI-Qq-TOF-MS; abbreviated to LC-MS in the following; Fig. 1 ⑥-⑩). We used quality controls (blanks and pooled extracts) to ensure data quality. We converted the acquired raw LC-MS data to an open file format (Fig. 1 ⑪) and processed them on the Galaxy-W4M infrastructure[31]. In Galaxy-W4M, we performed the feature detection, grouping and feature annotation (Fig. 1 ⑫). After this pre-processing, we prepared the data for statistical analysis. In R[32], we performed missing data imputation, batch correction and validity checks on the LC-MS feature (Fig. 1 ⑬-⑯). In this data descriptor, we provide a detailed description of the analytical steps performed on the acquired LC-MS data and provide the comprehensive data set in the MetaboLights repository MTBLS679[33].

Methods

Experimental Setup. *Experimental Design* The Jena Experiment[34] is a biodiversity ecosystem functioning experiment, designed to study plant and trait diversity effects on plant communities. The Jena Experiment is located in Jena, Germany, and includes the Trait-Based-Experiment[30] (TBE; Fig. 2 a). We collected plant material in the plots of the TBE. In the TBE, eight species selected from the functional groups (FG) grass and herb form a species pool. These Pools include four grass and four herb species[30]. Pool 1 (P1) comprises the grass species *Avenula pubescens* (AVEPUB), *Festuca rubra* (FESRUB), *Phleum pratense* (PHLPRA) and *Poa pratensis* (POAPRA) and the herbs *Centaurea jacea* (CENJAC), *Knautia arvensis* (KNAARV), *Leucanthemum vulgare* (LEUVUL) and *Plantago lanceolata* (PLALAN). Pool 2 comprises the grasses *Anthoxanthum odoratum* (ANTODO), *Dactylis glomerata* (DACGLO), *Holcus lanatus* (HOLLAN) and *Phleum pratense* (PHLPRA) and the herbs *Geranium pratense* (GERPRA), *Leucanthemum vulgare* (LEUVUL), *Plantago lanceolata* (PLALAN) and *Ranunculus acris* (RANACR). The target species of this study belonged to either P1 or P2 (Table 1, Fig. 2 b). The three species *Leucanthemum vulgare*, *Phleum pratense*, and *Plantago lanceolata* were part of both pools.

In the TBE, the plant species are grown in plots with different diversity levels (DL): one (DL 1), two (DL 2), four (DL 4), and eight (DL 8) different species per plot (Fig. 2 a). The plots are randomly distributed across the experimental site. P1 and P2 determine the plant species composition for each DL. Hence, all DL were composed of the species belonging to the respective Pool. For example, DL8 (P1) was composed of the following species: grass: AVEPUB, FESRUB, POAPRA, PHLPRA, herb: CENJAC, KNAARV, LEUVUL, PLALAN, while DL8 (P2) comprises these species: grass: ANTODO, DACGLO, HOLLAN, PHLPRA, herb: GERPRA, RANACR, LEUVUL, PLALAN. We collected the above-ground vegetative tissues of the thirteen target species. Per plot, we collected two plant individuals (replicates) at four time points in 2017. We chose dates across the growing season: May (A), July (B),

August (C) and October (D). In total, we sampled 512 *study samples*: 4 Season x 4 DL x 2 Pools x 8 species x 2 replicates (Fig. 1 ①-④, Fig. 2).

Traits & Sampling. Prior to plant biomass collection in each season, we surveyed each plot to record the actual number of present species (species richness), both sown (target) and weed (not deliberately cultivated) species. We also estimated the abundance of each species (Shannon diversity) in relation to the plot size. In each season, we collected the above-ground tissue of two replicates per plot and species (Fig. 2 b). In each plot, we randomly chose two plant individuals as replicates from specimens with a similar phenological stage according to the BBCH[35] scale. We recorded the following traits of these plant individuals: phenological stage (BBCH[35]), the number of leaves and inflorescences, plant height, and the proportional damage inflicted by either pathogen or mechanically.

The plants were cut 3 cm above the ground (Fig. 1 ⑤). An aliquot of shoot (leaf and stem) tissues was collected in plastic vials, snap-frozen on dry ice and stored for LC-MS analysis (referred to as *study sample*). The remaining biomass, including the inflorescences, was stored in plastic bags for biomass measurements. We collected the samples following the order of plots in the TBE (randomised DLs and Pools across the experimental site), starting at the southern end of the TBE[30]. We also recorded the exact time of the sampling for each sample to account for possible time-related shifts in the metabolic profile (sampling between 1 pm and 8 pm). We collected the samples within a single day to reduce the environmental influences to a minimum (for the exact dates see the MTBLS679[33] data repository). We applied the following labelling scheme to ensure the randomisation for sample extraction and LC-MS data acquisition. For each season, we assigned a number between 001 and 128 to each sample. These Lab-IDs were chosen randomly for each sample while collecting the biomass. For example, the Lab-ID *013_2017_A* refers to the sample 2017_A_*PHLPRA_A002_a*: collected in season 2017_A; *Phleum pratense*, in plot A002, which is referring to DL2 in P1, replicate a; and *013_2017_C* refers to the sample 2017_C_*FESRUB_B067_b*: collected in season 2017_C; *Festuca rubra;* in plot B067, which is referring to DL4 in P1; replicate b. The plot numbers (e.g. A002 and B067) and the corresponding DLs (e.g. DL2 and DL4) are specified in the sample metadata in the data records MTBLS679[33]. The sample preparation and extraction for the LC-MS data acquisition were conducted in the order of the respective Lab-IDs to ensure the equal distribution of seasons and full randomisation across the species, DL and replicates. Details on the randomisation can be found in the section "Sequence of LC-MS Measurements". All details concerning the sampling strategy are included in the sample table in the MTBLS679[33] data repository.

LC-MS Data Acquisition. *Cryo Sample Preparation.* We prepared the 511 *study samples* of frozen shoot material, collected in 20 mL vials, by adding two steel balls (7 mm) to the tubes. One sample tube (2017_B: FESRUB (P1): DL1_b) broke prior to analysis and was, therefore, excluded from further analysis. We used a cryo ball mill equipped with an autosampler (Labman IPB Cryogrinder Ball Mill, Labman Automation, Middlesbrough, UK) to grind the material at -75 °C for 150 s (5 cycles: 30 s grinding, 30 s pausing). We ground the samples according to their Lab-IDs and the season they were collected in (Fig. 1 ⑥).

Methanolic Extraction. We transferred aliquots (100 mg ± 50 mg) of the fine frozen powder to extraction tubes and added extraction beads (Rimax/Zircosil, 1.2 - 1.7 mm). For the extraction, we used methanol/water (80/20 v/v; HPLC-grade, Honeywell, Seelze, Germany) as the *extraction solvent*. We added the following internal standards at a 5 mM concentration to the *extraction solvent*: Kinetin (Roth, Karlsruhe, Germany), IAA-Val (Sigma-Aldrich, St. Louis, USA) and Biochanin A (Sigma-Aldrich, St. Louis, USA). The *extraction solvent* was added in a weight-specific five-fold surplus (Fig. 1 ⑦) to the frozen powder (e.g. 500 µL added to 100 mg powder), which we kept on liquid nitrogen. We thawed the prepared samples for 3 minutes at room temperature before extracting them in a homogeniser (Precellys® 24 Tissue Homogenizer, Bertin Technologies, Montigny-le-Bretonneux, France) for 90 s (2 cycles: 45 s run, 15 s pausing) at 6500 rpm. We centrifuged the extracts at 16168 g for 15 min and collected the supernatants in fresh extraction tubes (Fig. 3 a). After an additional extraction of the remaining pellet, 160 µL of the combined supernatants were added to 40 µL of water/formic acid (99.9/0.1 v/v) (formic acid: VWR International, Radnor, USA) and stored at -20 °C for at least 48 hours (Fig. 3 a). To prepare the samples for mass spectrometry, we centrifuged the *sample extracts* at 16168 g for 15 minutes to remove particles. We transferred 160 µL of the resulting supernatant to vials equipped with 300 µL glass inserts (*analytical sample*). We extracted all *study samples* in batches of 44 samples, in the order of their Lab-IDs (e.g. one analytical batch contains *analytical samples* with the Lab-IDs 001 to 011 of season 2017_A, 2017_B, 2017_C, and 2017_D).

Quality Controls. We used two types of blanks to account for possible contamination or inconsistency during extraction. The *field blanks* (plastic vials used for sampling) were included in the sampling, transportation and grinding steps. After the sampling in season 2017_A, we used a new shipment of plastic vials. We, therefore, labelled the *field blanks* "old" and "new" for the vials either used in 2017_A or 2017_B to 2017_D, respectively. We used the *extraction blanks* (eX01 - 03) to capture contaminations introduced in the methanolic extraction steps. For each replacement of *extraction solvent,* a new *extraction blank* was used. Both *field blanks* and *extraction blanks* were processed according to the extraction protocol applied to the *study samples*. Furthermore, we pooled 10 µL aliquots of the 511 *sample extracts*, which we used as Quality Control (*QC*) throughout the LC-MS measurements (Fig. 1 ⑧; Fig. 3 b).

Sequence of LC-MS Measurements. We measured the 511 *analytical samples* in 12 analytical batches. Each batch was composed of an acetonitrile aliquot, a blank, a *QC* aliquot and 44 *analytical samples* (Fig. 1 ⑨, Fig. 3 c). We distributed the *analytical samples* equally across the batches in the order of their Lab-IDs, and the season they were collected in (Lab-IDs were assigned to the samples randomly while sampling; see "Traits & Sampling"). For example, the *analytical samples* 2017_A_001 to 011, 2017_B_001 to 011, 2017_C_001 to 011, and 2017_D_001 to 011 were measured in batch "pos01". We started the batch measurement sequence with three acetonitrile runs followed by the *QC*. After this run-in sequence, we measured the *QC* again, to equilibrate both the LC-column and MS-system, followed by one blank and a block of 11 *analytical samples* (Fig. 3 c). We used the different blanks to detect potential systematic contaminations that were either introduced during sampling, extraction or the LC-MS measurements. After each block of *analytical samples*, we measured the *QC* again. The samples measured within one block were chosen randomly from the 44 samples assigned to the batch. After each batch, the MS ion source was cleaned, and the MS was recalibrated.

Analytical Setup & Data Acquisition. We performed the data acquisition (Fig. 1 ⑩) on a liquid chromatography system (UPLC; ACQUITY UPLC System, Waters Corporation, Milford, USA) coupled with a mass spectrometer (ESI-Qq-TOF-MS; ESI-micrOTOF-Q-II, Bruker Daltonics, Bremen, Germany). Aliquots (2 µL) of the *analytical samples* were separated at 40 °C on an HSS T3 C_{18}-column (1.8 µm, 1.0 x 100 mm, RP, Waters Corporation, Milford, USA) with the elution binary gradient at 0.15 mL min^{-1} flow rate: Solvent A (water/formic acid 99.9/0.1 v/v)/ Solvent B (acetonitrile/formic acid 99.9/0.1 v/v; acetonitrile: Merck, Darmstadt, Germany); initial: A 95%, 3 minutes linear A 82.7%, 10 minutes linear A 76%, 17 minutes linear A 5%, 18 minutes A 5%, 18.1 minutes linear A 95%, 20 min A 95%. We measured the ions in positive mode from 100 – 1000 m/z using the following instrument settings: capillary voltage 5000 V; nebuliser gas nitrogen; nebuliser 1.4 bar; dry gas nitrogen; dry gas temperature 190 °C; dry gas flow 6 L min^{-1}; spectra rate 3 Hz; endplate offset: -500 V; Funnel 1 RF: 200 Vpp; Funnel 2 RF: 200 Vpp; in-source CID energy 0 eV; hexapole RF 100 Vpp; quadrupole ion energy 3 eV; collision gas nitrogen; collision energy 7 eV; collision RF 200/200 Vpp (timing 50/50); transfer time 58.3 µs; pre pulse storage 5 µs. We used an internal calibration (lithium formate clusters, 10 mM lithium hydroxide in isopropanol/water/formic acid, 49.9/49.9/0.2 v/v/v, at 18 min) for the normalisation of the measurements.

LC-MS Data Pre-processing. We exported the vendor-specific data files (Bruker ".d") using CompassXport (Bruker, version 3.0.9, http://www.bruker.com). The conversion of LC-MS raw data files to the open data format (".mzML")[36] enables the data analysis in vendor-independent environments (Fig. 1 ⑪). We pre-processed the raw LC-MS spectra of the *analytical samples* and the quality controls (blanks and *QC*) on the Galaxy-W4M infrastructure[31] (based on XCMS 3.0). The workflow (DOI: 10.15454/1.5640497789529167E12) includes the following analytical and processing steps: feature detection, grouping and retention time correction (Fig. 1 ⑫). A detailed description of parameter settings and tool versions used in the workflow is also shown in Table 2.

The initial step in the workflow is feature detection. The parameters were set in order to separate measured peaks from background noise (Table 2). We then grouped the features across samples and corrected them for retention time shifts. We grouped the corrected spectra again and annotated adducts and isotopes of the measured features. After these pre-processing steps, we filtered the detected features for the region of interest (ROI). We cut features with retention times between 0 s to 80 s (injection peak and very polar compounds) and from 840 s to 1080 s (very nonpolar compounds). We exported the pre-processed data as separate data tables for sample metadata (*sampleMetadata*), variable metadata (*variableMetadata*) and the data matrix (*dataMatrix*), containing the measured intensities. These data matrices are also available in the associated metadata records MTBLS679[33] (https://www.ebi.ac.uk/metabolights/MTBLS679). The number of detected features per species is shown in Table 3.

Data Records

A detailed description of the experimental setup, the performed analysis and the metadata of both *study samples* and the quality controls are available as MTBLS679 "From Field to Feature in Ecometabolomics – LC-MS Based Metabolite Profiles of Thirteen Grassland Plant Species Reflecting Environmental Dynamics" on https://www.ebi.ac.uk/metabolights/MTBLS679. Raw data files of LC-MS analysis are also available in the repository. Furthermore, we provide data matrices of all stages of the processing steps (see Table 1). The W4M-Galaxy history (DOI: 10.15454/1.5640497789529167E12) that was used for data pre-processing is available at https://workflow4metabolomics.usegalaxy.fr/histories/list_published. All processing steps used for the data clean up are explained in the Supplementary File 1.

Technical Validation

Data Processing. A detailed tutorial of the processing steps performed in R[32] and the complete code used for data processing are provided as PDF and as R script in the MTBLS679[33] repository. The tutorial PDF is also made available as supplemental material (**Supplementary File 1**).

Missing Data Imputation. In this study, the pre-processing of highly diverse LC-MS spectra lead to a data matrix with 90% zero values. This high number of zeros is a result of the data matrix containing all detected features, of which only small fractions belonged to a particular species (Table 3). Hence, features that are not part of the metabolic fingerprint in this species were not detected and are recognised as true zeros. Within a species, some features are only detected in a few specimens. These absences either occur due to variations in the technical performance or are indicators of actual biological adaptations to environmental changes. These are NA values, as the reason for their absence is uncertain at this stage of analysis. In the following, we refer to any missing values as *missing data*. In order to prepare the data matrix for further data cleaning and to make it accessible to processing and statistical analysis, we replaced the *missing data* with imputed values. Here, we imputed the *missing data* with random values (noise) by drawing absolute values from a normal distribution with mean 70 and a standard deviation of 20. We chose these values as they are below the threshold initially set for our data set, which equals 100 (Fig. 1 ⑬, see Table 2: feature detection). This choice is instrument specific and based on the prefilter parameters used in the pre-processing steps.

Batch Correction. We performed a batch correction on the imputed data matrix. Splitting the 511 *analytical samples* into 12 analytical batches enhanced the chance of technical performance variability due to cleaning, recalibration and solvent replacements. These batch effects are mostly reflected in changes of intensities of the features across different batches. To account for these intensity shifts, the *QC*, which was measured multiple times across all batches (see "Sequence of LC-MS Measurements"), was used to determine the unwanted variation within (intra-batch distance) and between (inter-batch distance) batches. Ideally, the intensity profiles of the *QC* in all batches are identical. However, systematic variation between and within batches was present. Here, we used the *RUVs* function in the RUVSeq package (version 1.20.0)[37], which is based on a principal component analysis (PCA), and applied it to the *QC* measurements (referred to as *pool* in *dataMatrix*). *RUVs* creates a PCA model of the systematic part of the variation of the *QC*. This PCA model describes unwanted systematic variation. In the next step, it substracts the PCA model from the study samples; thereby eliminating any unwanted systematic variation. A detailed description of the underlying calculations can be found in Risso et al.[37]. The performance of the batch correction mainly depends on the number of components used for the analysis. We determined the optimal number of components to be used for the correction with a scree plot. In this scree plot, we compared the remaining inter-batch distances (Supplementary File 1 Fig. 1) after correction for different numbers of components. In this data set, the knee (or elbow) in the plot was reached after 6 components, as the inter-batch distances did not decrease anymore after 6 components (see Supplementary File 1 Table 3.2). After the batch correction, the calculated inter-batch distances for the *QC* measurements showed a strong decline (Table 4; Fig. 1 ⑭). The score plots before the batch correction show apparent batch effects in PC 1 and PC 2 (Fig. 1 ⑭). This shows that the batches, in which the *QC* has been measured, are the largest systematic source of variation for the *QC* measurements. After correction, the pattern in the PCs related to the different batches was no longer distinguishable. This shows that the huge variation of the feature intensities present in the original measurements related to the batches is removed and does not influence any consequent (statistical) analysis. After performing the batch correction, the *QC* measurements are removed from both the metadata and data matrix (Table 1).

Blank Removal. We checked the validity of the features before using them in the statistical analysis. We assigned a feature as valid when it was derived from an *analytical sample*. Here, we used the blanks as a reference for the validity check. Blanks did not contain a biological sample but were handled and processed like the

analytical samples. Hence, we considered all features that were detected in blanks to be systematic contaminations introduced during sampling, extraction or the LC-MS analytical process. We removed all features that were detected in at least one blank from the data matrix and excluded them from any further analysis (Fig. 1 ⑮; see Table 3 for the number of features before and after the blank removal). Following this feature validity check, we also removed the blank samples from the sample metadata (Table 1).

Sample Validity Check. The amount of biological variation in the metabolomic profiles within a species differed across the species. This intra-species variation was found to be lower than the inter-species variation. To check the validity of each sample and, thereby, ensuring that the sample was not contaminated, we compared their metabolomic profiles to the average composition of their species. Here, we defined a feature as belonging to a species when it was detected in at least 8 of the samples (25%) in that species (Table 3). Note that for assigning a feature to the respective species, we used the data matrix without the imputed values (see "Missing Data Imputation"). As a quality measure, for each sample, we calculated Mahalanobis distances (Fig. 1 ⑯). We compared the distance of each sample to the average distance of the remaining samples in the respective species. For example, we calculated distances for the 32 samples in the species *Holcus lanatus* and compared the distance of the sample "HOLLAN (P2): 2017_A (DL4_b)" to the average distance of the other 31 samples. We kept only those samples that were closer than three times the average distance and shared over 25% of their features with their species (Table 1). Consequently, we excluded the following samples from further analysis as they did not pass the validity check: ANTODO (P2): 2017_B (DL8_b), 2017_C (DL1_b); HOLLAN (P2): 2017_A (DL4_b); LEUVUL (P2): 2017_D (DL4_a, DL4_b); PHLPRA (P1): 2017_D (DL1_b, DL8_b); PHLPRA (P2): 2017_D (DL8_a, DL8_b), POAPRA (P1): 2017_D (DL2_b, DL8_a, DL8_b).

Preparation for Statistical Analysis. After performing validity checks on the data, we prepared the cleaned and processed data matrix to be used for statistical analysis. The data matrix can be accessed in three different stages, with 1) imputed values or 2) zeros or 3) NAs for missing values (see "Missing Data Imputation"). Depending on the nature of the planned analysis, either one of the matrices can be used for statistical analysis and conclusion drawing.

Usage Notes

This comprehensive data set provides the opportunity to investigate the metabolomic profiles on the feature level of thirteen grassland species grown in diverse neighbourhoods. The profiles were acquired from plants collected at different time point across the growing season. Therefore, relevant features and seasonality can be investigated within this eco-metabolomic dataset. Additionally, the mass spectrometry raw data are available in an open file format (mzML) and provide the opportunity to be re-processed with common metabolomics tools, such as xcms, OpenMS and MS-Dial.

Code availability

The raw data files and processed data matrices are available in the online repository MTBLS679 (https://www.ebi.ac.uk/metabolights/MTBLS679). The complete history of the used workflow for the raw LC-MS data pre-processing is available in Galaxy-W4M[31] from https://doi.workflow4metabolomics.org/W4M00008. We provide the complete R[32] script used to process the data along with a detailed tutorial in the supplemental material (Supplementary File 1).

Acknowledgements

The authors gratefully acknowledge the support of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – DFG–FZT 118, 202548816. The authors also thank the technical staff of the Jena Experiment (DFG, FOR 456/1451), the coordinator Anne Ebeling as well as many student helpers for maintaining the experimental field site and their support during measurements. This research was also supported by the International Research Training Group TreeDì jointly funded by the DFG (GRK2324) and the University of Chinese Academy of Sciences (UCAS). This paper has been conducted in the framework of the iDiv-Flexpool – the internal funding mechanism of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – DFG–FZT 118, 202548816.

Author contributions

SM co-designed the study, led the data collection, performed the data acquisition, conducted the raw data pre-processing, conceptualised and designed the data analysis, designed the workflow for data processing, wrote the tutorial for data analysis, managed the data and repository, designed and created the figures, and wrote the manuscript. JAH conceptualised and designed the data analysis, designed the workflow for data processing, and contributed to the tutorial provided for data analysis. RW contributed to the concept and design of the data analysis, assisted with the design of the workflow for data processing and the tutorial provided for data analysis. NMD designed the study, coordinated and managed activities leading to this paper and acquired the financial support. HB designed the study, provided the infrastructure for data acquisition, coordinated and managed activities leading to this paper and acquired the financial support. SN designed the study, helped with the raw data pre-processing, contributed to the concept and design of the data analysis, assisted with the design of the workflow for data processing and the tutorial provided for data analysis, contributed to data management and repository administration, provided the infrastructure for data acquisition, coordinated and managed activities leading to this paper and acquired financial support. All authors revised and edited the manuscript.

Competing interests

The authors declare no competing interest.

Tables

Table 1: Steps of analysis performed on the thirteen target species and the quality controls. Species belonging to the functional groups (FG) grass and herb were assembled in two groups of eight species (Pool). The Pools included four species per FG. Three of the species were represented in both pools (*). Shoots were collected at four time points (Season: A, B, C, D) in four diversity levels (DL1, DL2, DL4, DL8). A detailed list of the study samples can be found in the associated Metadata Record (MTBLS679[33]). For details of the experimental setup, see Fig. 1, and Ebeling et al. [30] for a plot overview. *Study samples* are processed in the respective analysis step (+). One sample was excluded from the analysis due to the loss of the sampled material, and some samples did not pass the final validation check (±; see section "Cryo Sample Preparation" and "Sample Validity Check"). This overview also indicates where the quality controls were used for the analysis.

				Exp.-Setup		LC-MS Data Acquisition			LC-MS Data Pre-processing				Data Processing			
Pool	FG	Species	Spec. Code	Trait Recording	Plant Material	Grinding	Extraction	LC-MS Measurements	Peak Picking	Grouping	Retention time correction	Feature Annotation	Missing Data Imputation	Batch Correction	Blank Removal	Sample Validity Check
P1	*grass*	*Avenula pubescens*	AVEPUB	+	+	+	+	+	+	+	+	+	+	+	+	+
P1	*grass*	*Festuca rubra*	FESRUB	+	+	±	±	±	±	±	±	±	±	±	±	±
P1	*grass*	*Phleum pratense* *	PHLPRA	+	+	+	+	+	+	+	+	+	+	+	+	±
P1	*grass*	*Poa pratensis*	POAPRA	+	+	+	+	+	+	+	+	+	+	+	+	±
P1	*herb*	*Centaurea jacea*	CENJAC	+	+	+	+	+	+	+	+	+	+	+	+	+
P1	*herb*	*Knautia arvensis*	KNAARV	+	+	+	+	+	+	+	+	+	+	+	+	+
P1	*herb*	*Leucanthemum vulgare* *	LEUVUL	+	+	+	+	+	+	+	+	+	+	+	+	+
P1	*herb*	*Plantago lanceolata* *	PLALAN	+	+	+	+	+	+	+	+	+	+	+	+	+
P2	*grass*	*Anthoxanthum odoratum*	ANTODO	+	+	+	+	+	+	+	+	+	+	+	+	±
P2	*grass*	*Dactylis glomerata*	DACGLO	+	+	+	+	+	+	+	+	+	+	+	+	+
P2	*grass*	*Holcus lanatus*	HOLLAN	+	+	+	+	+	+	+	+	+	+	+	+	±
P2	*grass*	*Phleum pratense* *	PHLPRA	+	+	+	+	+	+	+	+	+	+	+	+	±
P2	*herb*	*Geranium pratense*	GERPRA	+	+	+	+	+	+	+	+	+	+	+	+	+
P2	*herb*	*Leucanthemum vulgare* *	LEUVUL	+	+	+	+	+	+	+	+	+	+	+	+	±
P2	*herb*	*Plantago lanceolata* *	PLALAN	+	+	+	+	+	+	+	+	+	+	+	+	+
P2	*herb*	*Ranunculus acris*	RANACR	+	+	+	+	+	+	+	+	+	+	+	+	+
Quality Controls		field blanks ("new", "old")				+	+	+	+	+	+	+	+	+	+	
		extraction blanks ("eX01 - 03")					+	+	+	+	+	+	+	+	+	
		QC (QC2017AtoD)						+	+	+	+	+	+	+		

Table 2: Tools and Parameter used for pre-processing the LCMS raw data. The complete workflow is available in Galaxy-W4M (https://doi.workflow4metabolomics.org/W4M00008).

Tool name	Description	Version	Parameter	Value
MSnbase readMSData	Import mass-spectrometry data files	2.8.2.1		
findChromPeaks	feature detection	3.4.4.1	extraction method	40
			peak width (s)	5, 20
			signal to noise ratio	5
			prefilter	3, 100
			noise filter	100
xcms findChromPeaks Merger	merging xcms findChromPeaks	3.4.4.0		
xcms groupChromPeaks (group)	grouping of chromatographic peaks	3.4.4.0	method	PeakDensity
			bandwidth	6
			minimum fraction	0.75
			minimum number	1
			width of m/z slices	0.005
xcms adjustRtime (retcor)	retention time correction	3.4.4.1	method	PeakGroups
			minimum fraction	0.75
			maximum number	1
			smooth method	Loess – non-linear alignment
			degree smoothing	0.2
			family	gaussian
xcms groupChromPeaks (group)	grouping of chromatographic peaks	3.4.4.0	method	PeakDensity
			bandwidth	6
			minimum fraction	0.75
			minimum number	1
			width of m/z slices	0.005
CAMERA.annotate	Annotation of putative compounds	2.2.4	multiplier of sd	6
			general ppm error	5
			general abs error	0.005
			maximum ion charge	3
			maximum number	4
			isotope annotation	0.5
			correlation threshold	0.75
			grouping into pseudospectra	hcs
			correlation threshold	0.05
Check Format	Checking/formatting the sample and variable names	3.0.0		
Generic_Filter	Deleting samples and/or variables	2017.06	remove in "…" values upper	"rt", 840 (s)
			remove in"…" values lower	"rt", 80 (s)

Table 3: Number of unique LC-MS features (Fmeas) measured in both the *analytical samples* and the quality controls (Smeas). A feature is counted as part of the species when it is detected in at least 25% of the samples belonging to this particular species. After processing and blank removal, the remaining number of samples (Sval) and features (Fval) is used for analytical statistics.

		Pre-processed		Validated	
FG	Species Code	Smeas	Fmeas	Sval	Fval
	total	596	10252	499	10126
grass	ANTODO	32	1430	30	1310
grass	AVEPUB	32	1455	32	1358
grass	DACGLO	32	1281	32	1178
grass	FESRUB	31	1120	31	1020
grass	HOLLAN	32	1428	31	1317
grass	PHLPRA	64	1118	60	1054
grass	POAPRA	32	1046	29	924
herb	CENJAC	32	1711	32	1614
herb	GERPRA	32	1543	32	1464
herb	KNAARV	32	1708	32	1621
herb	LEUVUL	64	1384	62	1280
herb	PLALAN	64	1673	64	1581
herb	RANACR	32	1446	32	1348
Quality Controls	blank	12	126	0	0
	QC	73	5236	0	0

Table 4: Inter-batch distances calculated for both the *QC* (multiple measurements) and the *analytical samples* (single measurements). Distances are calculated before (pre BC) and after (post BC) applying the batch correction.

Data	pre BC	post BC
QC	16.845	0.720
analytical samples	0.056	0.058

Figures

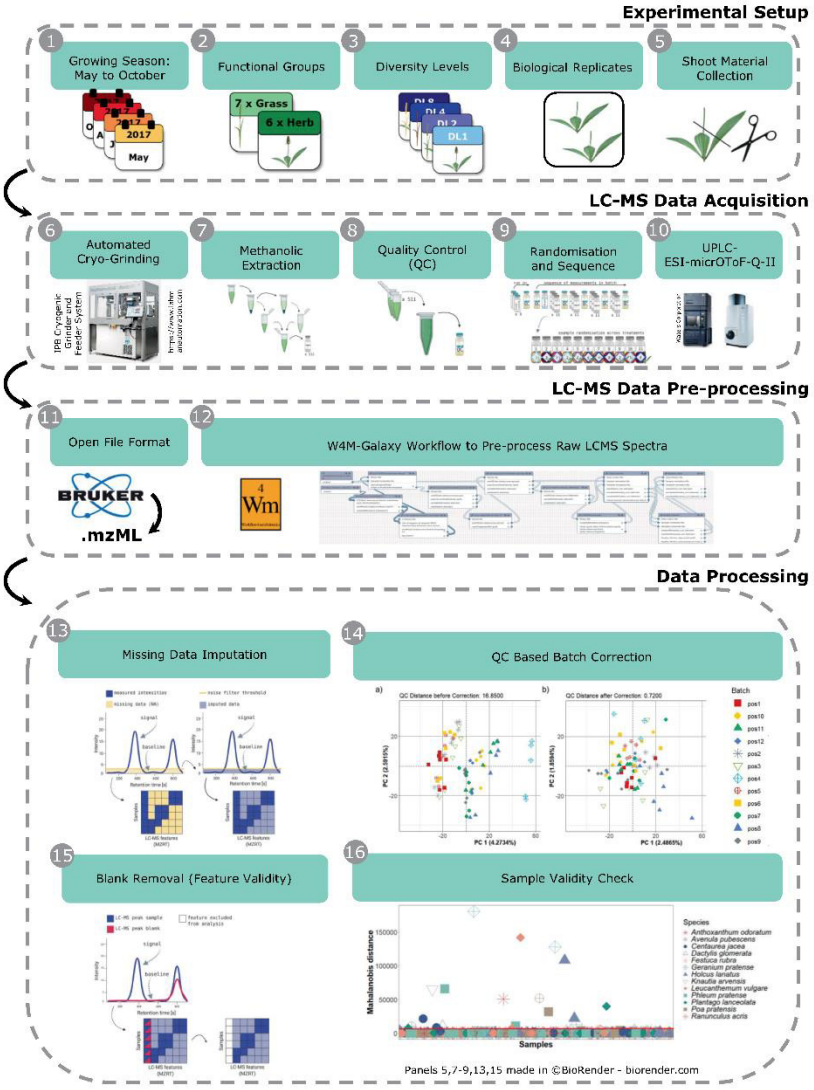


Figure 1: **Data set overview.** The data set includes the metadata of the experimental setup for the plant material collected in the TBE plots of the Jena Experiment (①-⑤), LC-MS raw data acquisition (⑥-⑩), data pre-processing steps (⑪-⑫), as well as data cleaning and validation (⑬-⑯). Created with BioRender.com

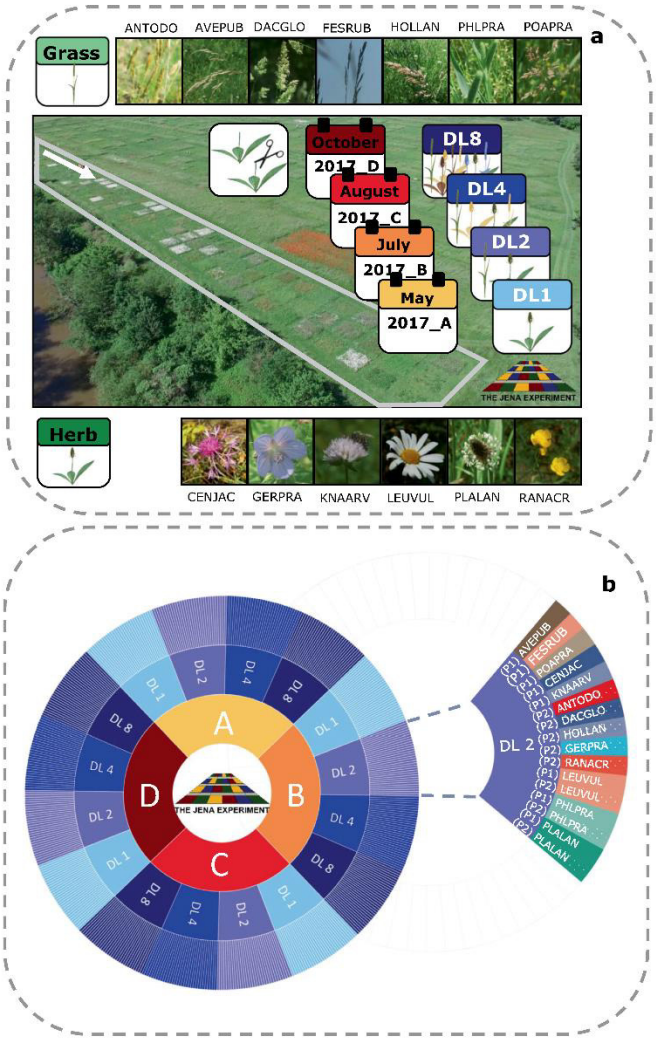


Figure 2: **Experimental Design.** (**a**) Plot and species overview. Plant material was collected in the plots of the TBE (grey borders) in the Jena Experiment. We collected shoots of seven grass (light green) and six herb species (dark green) in plots with four different diversity levels (DL). Here, either one (DL1), two (DL2), four (DL4) or eight (DL8) different species were grown per plot. In each plot, we harvested shoots of two replicates. The white arrow indicates the sampling direction, starting at the south end of the TBE. (**b**) Design overview. Plant material of species in both P1 and P2 were collected at four time points across the growing season in 2017 (May: A, July: B, August: C, October: D). The species pools P1 and P2 were each composed of four grass and four herb species. The three species LEUVUL, PHLPRA and PLALAN, were part of both pools. In total, we collected 512 *study samples*: 4 Season x 4 DL x 2 Pool x 8 species x 2 replicates. For a detailed list of the species codes see Table 1. Created with BioRender.com

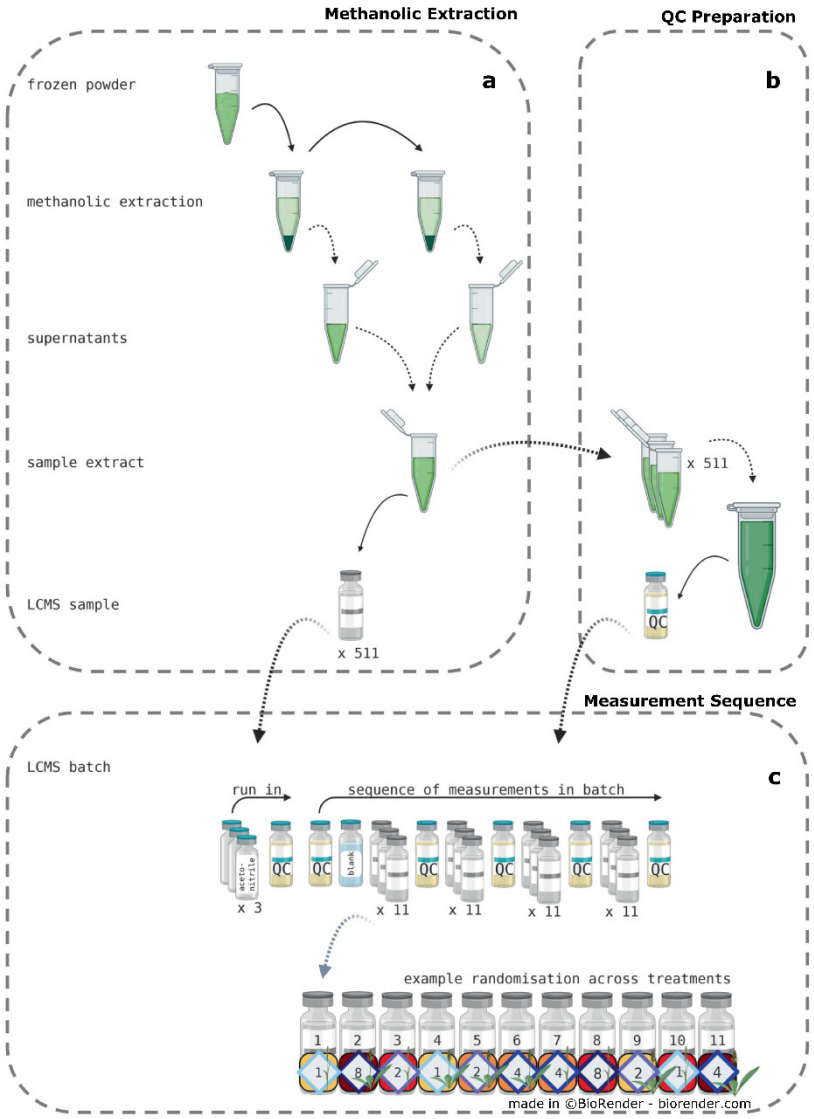


Figure 3: **LC-MS sample extraction and sequence of measurements**. a) We prepared the 511 frozen study samples by grinding and extracting the resulting fine powder with methanol (note: one sample was lost prior to analysis, see "Cryo Sample Preparation"). For each sample, we combined the supernatants of two extraction steps to the sample extract. b) We pooled aliquots of all 511 sample extracts and used them as the *Quality Control* (*QC*). c) The LC-MS measurements were split into 12 *analytical batches*. Here, each batch measurement was led by a run-in sequence: 3 x acetonitrile, 1 x QC measurement. Per batch, samples were measured in four blocks, consisting of eleven *analytical samples*. The sample measurements were preceded by one *QC* measurement and one blank, and flanked by *QC* measurements. We randomised the 511 LC-MS samples (13 species, 4 seasons, 4 diversity levels) equally across the 12 batches. For treatment colour codes, see Fig. 2. Solid black arrows mark processing steps, while dashed black arrows indicate the transfer to another process. The dashed grey arrow indicates a zoom-in for clarification purposes. Created with BioRender.com

References

1. Eichenberg, D., Purschke, O., Ristok, C., Wessjohann, L. & Bruelheide, H. Trade-offs between physical and chemical carbon-based leaf defence: of intraspecific variation and trait evolution. *J. Ecol* **103** (6), 1667-1679 (2015).
2. Fiehn, O. Metabolomics—the link between genotypes and phenotypes. *Plant Mol. Biol.* **48**, 155-171 (2002).
3. Fernie, A. R., Trethewey, R. N., Krotzky, A. J. & Willmitzer, L. Metabolite profiling: from diagnostics to systems biology. *Nat. Rev. Mol.* **5** (9), 763-769 (2004).
4. Weir, T. L., Park, S. W. & Vivanco, J. M. Biochemical and physiological mechanisms mediated by allelochemicals. *Curr. Opin. Plant Biol.* **7** (4), 472-479 (2004).
5. Rosenthal, G. A. & Berenbaum, M. R. *Herbivores: their interactions with secondary plant metabolites: ecological and evolutionary processes.* Vol. 2 (Academic Press, 2012).
6. Schweiger, R., Heise, A. M., Persicke, M. & Müller, C. Interactions between the jasmonic and salicylic acid pathway modulate the plant metabolome and affect herbivores of different feeding types. *Plant Cell Environ.* **37** (7), 1574-1585 (2014).
7. Arbona, V., Manzi, M., Ollas, C. D. & Gómez-Cadenas, A. Metabolomics as a tool to investigate abiotic stress tolerance in plants. *I. J. Mol. Sci.* **14** (3), 4885-4911 (2013).
8. Bais, H. P., Park, S., Weir, T. L., Callaway, R. M. & Vivanco, J. M. How plants communicate using the underground information superhighway. *Trends Plant Sci.* **9** (1), 26-32 (2004).
9. Badri, D. V. & Vivanco, J. M. Regulation and function of root exudates. *Plant Cell Environ.* **32** (6), 666-681 (2009).
10. Treutter, D. Significance of flavonoids in plant resistance: a review. *Environ. Chem. Lett.* **4** (3), 147-157 (2006).
11. Wurst, S., Wagenaar, R., Biere, A. & van der Putten, W. H. Microorganisms and nematodes increase levels of secondary metabolties in roots and root exudates of Plantago lanceolata. *Plant Soil* **329** (1-2), 117-126 (2010).
12. van Dam, N. M. Belowground herbivory and plant defenses. *Annu. Rev. Ecol. Evol. Syst.* **40**, 373-391 (2009).
13. Dixon, R. A. Natural products and plant disease resistance. *Nature* **411** (6839), 843-847 (2001).
14. Ristok, C., Poeschl, Y., Dudenhöffer, J.-H., Ebeling, A., Eisenhauer, N., Vergara, F., Wagg, C., van Dam, N. M. & Weinhold, A. Plant species richness elicits changes in the metabolome of grassland species via soil biotic legacy. *J. Ecol*. **107** (5), 2240-2254 (2019).
15. Macel, M., de Vos, R. C., Jansen, J. J., van der Putten, W. H. & van Dam, N. M. Novel chemistry of invasive plants: exotic species have more unique metabolomic profiles than native congeners. *Ecol. and Evol.* **4** (13), 2777-2786 (2014).
16. Atkinson, N. J. & Urwin, P. E. The interaction of plant biotic and abiotic stresses: from genes to the field. *J. Exp. Bot.* **63** (10), 3523-3543 (2012).
17. Suzuki, N., Rivero, R. M., Shulaev, V., Blumwald, E. & Mittler, R. Abiotic and biotic stress combinations. *New Phytol.* **203** (1), 32-43 (2014).
18. Breitschwerdt, E., Jandt, U., Bruelheide, H. Trait-performance relationships of grassland plant species differ between common garden and field conditions. *Ecol. Evol*. **9**, 1691–1701 (2019).
19. Díaz, S. & Cabido, M. Vive la différence: plant functional diversity matters to ecosystem processes. *Trends Ecol. Evol.* **16** (11), 646-655 (2001).
20. Barry, K. E., Weigelt, A., Ruijven, J., de Kroon, H., Ebeling, A., Eisenhauer, N., Gessler, A., Ravenek, J. M., Scherer-Lorenzen, M., Oram, N. J., Vogel, A., Wagg, C. & Mommer, L. Above-and belowground overyielding are related at the community and species level in a grassland biodiversity experiment. *Adv. Ecol. Res.* **61**, 55-89 (2019).
21. Barry, K. E., van Ruijven, J., Mommer, L., Bai, Y., Beierkuhnlein, C., Buchmann, N., de Kroon, H., Ebeling, A., Eisenhauer, N., Guimaraes-Steinicke, C., Hildebrandt, A., Isbell, F., Milcu, A., Neßhöver, C., Reich, P. B., Roscher, C., Sauheitl, L., Scherer-Lorenzen, M., Schmid, B., Tilman, D., von Felten, S. & Weigelt, A. Limited evidence for spatial resource partitioning across temperate grassland biodiversity experiments. *Ecology* **101** (1), e02905 (2020).
22. Roscher, C., Schumacher, J., Gubsch, M., Lipowsky, A., Weigelt, A., Buchmann, N., Schmid, B. & Schulze, E.-D. Using plant functional traits to explain diversity–productivity relationships. *PloS one* **7** (5), e36760 (2012).
23. Violle, C., Navas, M. L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. Let the concept of trait be functional!. *Oikos* **116** (5), 882-892 (2007).

24. Ackerly, D. D., Dudley, S. A., Sultan, S. E., Schmitt, J., Coleman, J. S., Linder, C. R., Sandquist, D. R., Geber, M. A., Evans, A. S., Dawson, T. E. & Lechowicz, M. J. The Evolution of Plant Ecophysiological Traits: Recent Advances and Future Directions: New research addresses natural selection, genetic constraints, and the adaptive evolution of plant ecophysiological traits. *Bioscience* **50** (11), 979-995 (2000).
25. Herz, K., Dietz, S., Haider, S., Jandt, U., Scheel, D. & Bruelheide, H. Predicting individual plant performance in grasslands. *Ecol. and Evol.* **7** (21), 8958-8965 (2017).
26. Gross, N., Kunstler, G., Liancourt, P., De Bello, F., Suding, K. N. & Lavorel, S. Linking individual response to biotic interactions with community structure: a trait-based framework. *Funct. Ecol.* **23** (6), 1167-1178 (2009).
27. Herz, K., Dietz, S., Gorzolka, K., Haider, S., Jandt, U., Scheel, D. & Bruelheide, H. Linking root exudates to functional plant traits. *PloS one* **13** (10), e0204128 (2018).
28. Fernie, A. R. & Schauer, N. Metabolomics-assisted breeding: a viable option for crop improvement?. *Trends Genet.* **25** (1), 39-48 (2009).
29. Stitt, M., Sulpice, R. & Keurentjes, J. Metabolic networks: how to identify key components in the regulation of metabolism and growth. *Plant Physiol.* **152** (2), 428-444 (2010).
30. Ebeling, A., Pompe, S., Baade, J., Eisenhauer, N., Hillebrand, H., Proulx, R., Roscher, C., Schmid, B., Wirth, C. & Weisser, W. W. A trait-based experimental approach to understand the mechanisms underlying biodiversity–ecosystem functioning relationships. *Basic Appl. Ecol.* **15** (3), 229-240 (2014).
31. Giacomoni, F., Le Corguillé, G., Monsoor, M., Landi, M., Pericard, P., Pétéra, M., Duperier, C., Tremblay-Franco, M., Martin, J.-F., Jacob, D., Goulitquer, S., Thévenot, E. A. & Caron, C. Workflow4Metabolomics: a collaborative research infrastructure for computational metabolomics. *Bioinformatics* **31** (9), 1493-1495 (2014).
32. R Core Team R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria* https://www.R-project.org/ (2019).
33. Marr, S., Hageman, J. A., Wehrens, R., van Dam, N. M., Bruelheide, H., Neumann, S. From Field to Feature in Ecometabolomics: LC-MS Based Metabolite Profiles of Thirteen Grassland Plant Species Reflecting Environmental Dynamics. *MetaboLights* http://identifiers.org/metabolights:MTBLS679. (2020).
34. Roscher, C., Schumacher, J., Baade, J., Wilcke, W., Gleixner, G., Weisser, W. W., Schmid, B. & Schulze, E.-D. The role of biodiversity for element cycling and trophic interactions: an experimental approach in a grassland community. *Basic Appl. Ecol.* **5** (2), 107-121 (2004).
35. Hess, M., Barralis, G., Bleiholder, H., Buhr, L., Eggers, T. H., Hack, H. & Stauss, R. Use of the extended BBCH scale—general for the descriptions of the growth stages of mono; and dicotyledonous weed species. *Weed Res.* **37** (6), 433-441 (1997).
36. Martens, L., Chambers, M., Sturm, M., Kessner, D., Levander, F., Shofstahl, J., Tang, W. H., Römpp, A., Neumann, S., Pizarro, A. D., Montecchi-Palazzi, L., Tasman, N., Coleman, M., Reisinger, F., Souda, P., Hermjakob, H., Binz, P.-A. & Deutsch, E. W. mzML—a Community Standard for Mass Spectrometry Data. *Mol. Cell. Proteomics* **10** (1), R110.000133 (2010).
37. Risso, D., Ngai, J., Speed, T. P. & Dudoit, S. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat. Biotechnol.* **32** (9), 896-902 (2014).

Supplemental Material – Workflow for MTBLS679
See Appendix – Chapter 1

CHAPTER 2

A CASE OF INVESTIGATION: SELECTING STATISTICAL TOOLS FOR DIFFERENT LEVELS IN AN ECO-METABOLOMICS EXPERIMENTAL SETUP

Title: A Case of Investigation: Selecting Statistical Tools for Different Levels in an Eco-Metabolomics Experimental Setup

Authors:
Sue Marr [1,2,3*], Jos A. Hageman[4] & Steffen Neumann[1,3]

1 Bioinformatics & Scientific Data, Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120, Halle, Germany.
2 Institute of Biology/Geobotany and Botanical Garden, Martin Luther University Halle-Wittenberg, Am Kirchtor 1, 06108, Halle, Germany.
3 German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstr. 4, 04103, Leipzig, Germany.
4 Biometris, Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB, Wageningen, The Netherlands.
*corresponding author

Author contribution:

Author contributions (ctb) to the manuscript in percent.

chapter	status	author	ctb
(2) A Case of Investigation: using common metabolomics statistical tools in an eco-metabolomics experimental setup	draft	S. Marr	90
		J. A. Hageman	5
		S. Neumann	5

A Case of Investigation: using common metabolomics statistical tools in an eco-metabolomics experimental setup

S Marr [1, 2, 3, *], JA Hageman[4] and S Neumann[1,3]

[1] Bioinformatics & Scientific Data, Leibniz Institute of Plant Biochemistry, Halle (Saale), Germany
[2] Institute of Biology, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany
[3] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig
[4] Biometris, Wageningen UR, Wageningen, The Netherlands
[*] Correspondence: smarr@ipb-halle.de

Abstract:

Introduction In eco-metabolomics, selecting relevant features from metabolomic profiles can provide insights into plant phenotype adaptations to environmental dynamics. Currently available statistical tools are not yet tailored to capture the complexity of ecological experiments, especially regarding the growing interest in the role of secondary metabolites.

Objectives Therefore, this study aims to compare the performance of classical metabolomics feature selection and classification methods in eco-metabolomics data sets, which provide challenges regarding sample size and background noise.

Methods We tested the performance of *Partial Least Squares Discriminant Analysis* (*PLS*), *Support Vector Machines* (*SVM*) and *Random Forests* (*RF*) in terms of accuracies and their predictive power on three different pre-processed data matrices in three independent data sets. The data sets provided four different levels of metabolomic profile complexity, ranging from distinct to similar profiles for comparison.

Results In general, the data matrices that either used measured intensities and zeros (*ZeroInt*) or imputed log-transformed data (*ImpLog*) were found to have the highest accuracies. In contrast, accuracy calculated on measured intensities with imputed values for missing data (ImpInt) was low overall. Comparing the model performance, we found the highest accuracies in *PLS* and *RF. PLS*, *SVM* and *RF* worked equally well on very distinguishable metabolomic profiles, while *PLS* and *RF* provided a greater classification power when the background was more similar.

Conclusion Choosing a suitable classification method depended mainly on the complexity of the provided background for classification. The choice of the data matrix depending on the model also improved the performance.

Keywords:
grassland, eco-metabolomics, classification, partial least square discriminant analysis, support vector machines, random forest

1 Introduction:

In eco-metabolomics, one primary goal is to provide insights into phenotypes and metabolic profiles of biological systems. Combining metabolomics tools (e.g. mass spectrometry) with ecological experimental designs (Fiehn, 2020) allows using chemical properties to compare metabolic profiles and identify relevant compounds in the dynamics of environmental changes. In mass spectrometry, different measured features, defined by their specific mass-to-charge ratio (m/z) and retention time (RT), can derive from the same metabolic compound as they break down into multiple ion species during the analysis (Shi et al., 2019). The enormous amount of produced data can add major obstacles to statistical analysis, such as high dimensionality due to the imbalanced number of detected features that often exceed the number of samples by far and an increased risk of classification by chance due to a missing control level (Broadhurst & Kell, 2006; Gromski et al., 2015). As this research area is recently gaining more importance, so does the need for methods used for data handling, classification, and selecting important features. However, according to Smilde et al. (2005), the currently available statistical tools, which are developed mainly for clean experimental designs with a focus on case-control-samples, are not yet at a point where the complexity of ecological experiments is captured adequately (compare Broadhurst & Kell, 2006; Tang et al., 2014).

Furthermore, Shi et al. (2019) point out that only a limited number of algorithms include redundant features while noisy features are being removed. A high variability, often introduced in field or greenhouse experiments data, requires tailored analytical tools. In classical metabolomics studies, common tools for species classification and feature selection are, among others, *Partial Least Squares Discriminant Analysis* (*PLS*), *Support Vector Machines* (*SVM*) and *Random Forests* (*RF*). As discussed by Tang et al. (2014), for building proper models, it is essential to realise that all feature selection methods work under different assumptions and are, therefore, not compatible with all classification problems per se. As Mendez et al. (2019) discussed, getting any model is granted while getting an almost perfect model reflecting natural phenomena is almost impossible. The most reliable method for a data set can depend on its size, data types, and noise level (Broadhurst & Kell, 2006). According to Brereton and Lloyd (2014), observatory methods like PLS can play a significant role in exploratory studies. It is mainly used to classify complex data due to its ability to handle high dimensionality and noisy data (Brereton & Lloyd, 2014; Gromski et al., 2015). However, one major limitation in *PLS* methods is the assumption of linear relationships between even-sized classes and the equal importance of all measurements, which is not necessarily the case in eco-metabolomics studies (Brereton & Lloyd, 2014; Gromski et al., 2015; Mendez et al., 2019).

In contrast to *PLS*, *SVM* can be used on both linear and non-linear classification problems. However, as discussed by Statnikov et al. (2008), *SVM* is sensitive to an unbalanced distribution of classes. Furthermore, *SVM* is robust to outliers and less likely to overfit while performing well with noisy and high-dimensional data. On the other hand, *Random Forests (RF)* only require prior knowledge about the class labels and select the smallest set of features that can be used for classification. Generally, *RF* is comparable with *SVM* in terms of robustness towards outliers, noisy data handling and high dimensionality (Cutler et al., 2007; Statnikov et al., 2008). However, an advantage of *RF* over *PLS* and *SVM* is the ability to handle multiclass problems without any prior assumptions about the structure of the data and the robustness towards overfitting. Furthermore, *RF* has been shown to be a powerful tool, especially for ecological data, mainly due to its high classification accuracy and the ability to handle missing data (Cutler et al., 2007; Statnikov et al., 2008).

In this study, we want to compare the performance of *PLS*, *SVM* and *RF* in terms of their suitability in eco-metabolomics studies. We use three eco-metabolomics data sets that incorporate some of the most common obstacles occurring in these types of experiments, i.e., few observations, uneven distribution of samples and noisy data. The data sets we use in this study are derived from a long-term field experiment (MTBLS679 and MTBLS1224) as well as from a greenhouse setup (MTBSL2140), each including secondary metabolomic profiles of above-ground tissue of grassland plant species. We used plant species (species) from the functional groups (FG), grasses and herbs and collected samples at different time points (season) from plots with varying stages of neighbourhood diversity (diversity levels). Hence, these data sets provide four analysis levels (FG, species, season, diversity) ranging from a very clear separation across multiple species with very distinctive metabolomic profiles (FG and species) to a more challenging separation in the analysis level of season and diversity where metabolomic profiles are compared within each species and, therefore, have major overlaps in their profiles. This eco-metabolomics study did not include a control level and, therefore, relied on the other species in the respective data set for classification background.

Consequently, the comparison and discrimination of one species is only as good as the reference data it is compared to. Furthermore, the environmental conditions were not controlled for, which introduced a high level of unexplained variation and an increased number of redundant or noisy features that do not directly contribute to the classification, which increases the chance of selecting some features just by chance (John et al., 1994).

Analysing multiple grassland plant species with different profiles within each data set also created high amounts of missing values. Performing the analysis on either zero or imputed data for these missing values is a highly discussed challenge when working with eco-metabolomic data sets. Both strategies might lead to misinterpretation of the data.

Therefore, we want to test different classification models and how well they perform on the different levels of analysis and investigate which matrices (zero vs imputed) are the most suitable for each of the models within a certain analysis level. Hence, our three motivations are: 1) do selection models use other than the unique features within each class for the classification of the samples; 2) how do the models perform on different data matrices with missing values being either treated as zeros or imputed values; 3) which model performs best on which level of analysis, including the increasing overlap of metabolic profiles within the samples. This study aims to compare and evaluate the usability of each feature selection method on the three data sets in terms of the number of selected features and their classification performance and stability.

2 Methods:

2.1 Data sets

The three data sets we used for the model calculations share a similar experimental design (Fig. 1, Table 1). All data sets (MTBLS679, MTBLS1224, MTBL2140) include thirteen grassland species (Species) in the functional group (FG) herb: *Centaurea jacea* L. (CENJAC), *Geranium pratense* L. (GERPRA), *Knautia arvensis* (L.) Coult. (KNAARV), *Leucanthemum vulgare* (Vaill.) Lam. (LEUVUL), *Plantago lanceolata* L. (PLALAN), *Ranunculus acris* L. (RANACR); and grass: *Anthoxanthum odoratum* L. (ANTODO), *Avenula pubescens* (Huds.) Dumort. (AVEPUB; species is now referred to as *Helictotrichon pubescens* (Huds.) Schult. & Schult.f.), *Dactylis glomerata* L. (DACGLO), *Festuca rubra* L. (FESRUB), *Holcus lanatus* L. (HOLLAN), *Phleum pratense* L. (PHLPRA), *Poa pratensis* L. (POAPRA). The dataset MTBLS2140 additionally contained the four grass species: *Arrhenatherum elatius* (L.) P.Beauv. ex J.Presl & C.Presl. (ARRELA), *Festuca pratensis* Huds. (FESPRA), *Poa trivialis* L. (POATRI), and *Trisetum flavescens* (L.) P.Beauv. (TRIFLA). A detailed description for the MTBLS679 data set, regarding the experimental design, sampling, measurements and processing steps can be found in Marr et al. (2021). Detailed descriptions for the data sets MTBLS1224 and MTBLS2140 can be found in the MetaboLights repositories MTBLS1224 (*MetaboLights*: MTBLS1224) and MTBLS2140 (*MetaboLights*: MTBLS2140), respectively. Raw data and sample metadata of all data sets are available in the respective repository: MTBLS679 (http://identifiers.org/metabolights:MTBLS679), MTBLS1224 (http://identifiers.org/metabolights:MTBLS1224) and MTBLS2140 (http://identifiers.org/metabolights:MTBLS2140). In data set MTBLS2140, seed material was grown in a phytocabinet and then transferred to an outdoor greenhouse. The plants were maintained, and the pots were kept free of weeds. The plant material for this data set was collected within 3 days in October 2019. The samples in the data sets MTBLS679 and MTBLS1224 were collected at four time points (season) across the growing in 2017 and 2018, respectively, from plots with different diversity levels (DL) in the TBE-Jena-Experiment (Ebeling et al. 2014). The diversity levels were compositions of either one species (DL 1), two (DL 2), four (DL 4) or eight different species (DL 8; Fig. 1). The collected above-ground material (MTBLS679, MTBLS1224, MTBLS2140) was snap-frozen, extracted with Methanol/Water (80/20 v/v) and analysed on a UPLC (Acquity, Waters, USA) coupled with an ESI-micrOToF-Q-II (Bruker, USA) in positive ion mode (LC-MS analysis).

We pre-processed the obtained raw data in the Galaxy-W4M (Giacomoni et al. 2015) instance using the workflow provided by Marr (https://doi.workflow4metabolomics.org/W4M00008) with adapted parameters for each data set. The generated data was then processed in R, performing a batch correction, blank removal, and sample validity check (Marr et al. 2021). The processed and cleaned data matrices are available as three matrices for each data set. One matrix contains the measured intensities and zero values, where no intensity was measured (*ZeroInt*; e.g. MTBLS679_dataMatrixZ_processed), measured intensities and imputed data for missing values (*ImpInt*; e.g. MTBLS679_dataMatrixImp_processed) and imputed and log-transformed data (*ImpLog*; e.g. MTBLS679_dataMatrixImpLog_processed). We are aware that there are several strategies for imputing missing data; however, we chose randomly selected small values for these data sets, assuming that there is no traceable relationship between the features. We analysed the data sets MTBLS679 and MTBLS1224 on four analysis levels (FG, species, season, diversity) and MTBLS2140 on two levels (FG and species; Table 1).

2.2 Classification methods

Many tools are already available for metabolomics data that can be used for feature selection and classification problems, all well implemented and described in detail. Hence, we avoid in-depth mathematics and

focus on their general functionality, working principles, and differences between their approaches. Reviews of feature selection methods and their requirements may be found elsewhere (Tang et al., 2014). For each model, we used repeated 10-fold cross-validation. The data set is split randomly into 10 parts, building the model on nine of them and using the remaining part for validation. The average of the error terms of each validation is calculated. This procedure is then repeated 10 times, and the overall average is calculated. We implemented all methods in R (version 4.0.3; R Core Team 2019) and used the caret package (version 6.0-86; Kuhn, 2020) for model building. These methods are also available in the MetaboAnalyst software (MetaboAnalyst 5.0; Pang et al., 2021).

Partial Least Squares Discriminant Analysis (PLS) PLS is a linear classifier used as a multivariate regression method for calibration (Barker & Rayens 2003). It is a widely used tool in the fields of chemometrics (Wold et al. 2001), proteomics (Christin et al. 2013) and metabolomics (Gromski et al. 2015) that, although designed to solve binary case-control problems, can also be adapted to multiclass problems by using latent variables on a one-vs-one or one-vs-rest approach. The *PLS* algorithm gives the likelihood for a sample to belong to either class while providing information about the features that contribute to the separation. *PLS*, therefore, can also be used for biomarker identification. In *PLS*, the data are projected to a new dimensional space, finding a linear hyperplane alongside the discriminating features spanning a maximum space between the two classes. The weighted sums of the regression coefficient are used to calculate the variable importance for each feature (Kuhn 2020). Features that provide less information, therefore, are being downweighed. Features assigned relevant, however, are being selected based on their corresponding VIP values. A detailed description of the underlying *PLS* algorithm and VIP calculation can be found here (Barker & Rayens 2003, Brereton & Lloyd 2014). We used the pls package (version 2.7-3; Mevik et al. 2019) in the caret implementation in R for building the *PLS* model.

Support Vector Machines (SVM) SVM are commonly used in metabolomics (Mahadevan et al. 2008) and also many other research fields (e.g. Kriegl et al. 2005, Sattlecker et al. 2010). In contrast to *PLS*, which performs best for linearly separable data (linear kernel; *SVMl*), *SVM* can also solve non-linear problems (Sanz et al. 2018). Different kernels in the training phase enable *SVM,* for example, to deal with polynomial problems (poly kernel,. *SVMp*) and infinite dimensions (radial kernel; *SVM*). *SVM* are generally robust and effective predictors, suitable for both classification and regression, and are, therefore, used to select relevant features for classification and prediction (Burges, 1998; Statnikov et al. 2008). In *SVM*, the classifier is built by projecting a hyperplane into a higher dimensional space, separating the two classes with a maximal margin. The support vectors, being the features closest to the boundaries of the margin, are selected as the most relevant features giving the best separating and classification accuracy (Mahadevan et al. 2008, Tang et al. 2014). A large margin, therefore, also means a good generalisation. For prediction, *SVM* focuses on which side of the support vector a sample falls. The variable importance is assessed by ranking the features according to their relevance, separating the classes and the subset with the highest classification accuracy is chosen (Sanz et al 2018). A detailed explanation of the working principles and underlying math may be found elsewhere (Mahadevan et al. 2008). For a linear *SVM*, we use a recursive feature elimination to calculate the variable importance of each feature. For building the *SVM* model, we used the e1071 package (version 1.7-6; David et al. 2019) in the caret package implementation.

Random Forests (RF) In order to cope with high dimensional and complex data, *RF* – a collection of single classification trees – are an effective statistical tool for classification and feature selection in metabolomics studies. *RF* can be used on complex and high dimensional data and is, therefore, also a popular tool amongst ecologists (e.g. Cutler et al. 2007). However, the selection of features is not implemented as a direct output. In general, *RF* is defined as a supervised classification method consisting of a collection of classification trees that randomly splits the features in trainings and test sets to grow full and unpruned trees that are aggregated to a single data set (forest) including the predictions from all classification trees (Breiman 2001, Genuer et al. 2010). At each node, the binary partitioning is performed on randomly selected subsets of features within the bootstrap samples. One tree is grown per bootstrap sample, where each node, being equally distributed vectors, casts a single vote for the most common feature. Prediction accuracy is recorded for each tree. The variable importance can also be used for clustering, multidimensional scaling and missing value imputation (Strobl et al. 2007). Detailed information about the underlying algorithm and statistical backgrounds are explained in Breiman (2001) and Genuer et al. (2010) and are furthermore discussed in (Cutler et al. 2007). Here, we used the randomForest package (version 4.6-14; Liaw & Wiener 2002) in the caret implementation in R.

3 Results:

This study used three eco-metabolomics data sets that included common obstacles such as few observations, uneven distribution of samples, and noisy data. In these data sets, metabolomic profiles were compared at four analysis levels that ranged from very distinct metabolomic profiles across multiple species (FG and species), and hence, easy classification, to very similar profiles within the same species (season and diversity) and, therefore, more challenging classification problems. Based on the provided metabolomic background, we defined features present in more than one factor per analysis level in their data set as non-unique. Hence, features present in only one factor were defined as unique. We compared those features that were selected by the respective models with the selected unique features per analysis levels. At the FG level, we selected features that are either present in grass or herb species. At the species level, we selected features present in only one species and no other species within each data set (one-vs-rest). For the levels of season and diversity (MTBLS679 and MTBLS1224) we selected the features within each species. Here, we selected features present in only one season or diversity level. For model comparison, we used accuracy and Cohen's Kappa values, where the accuracy gives the rate of correctly classified species out of all classified species in a data set, and the Kappa value gives an estimate on the reliability of the calculated accuracy for unbalanced data (some species have more features than others). The Kappa value ranges from -1 to 1, where values above 0.8 are assumed to be strong and above 0.9 as almost perfect (McHugh, 2012). We assessed the prediction performance of each model using the true positive and true negative prediction rates.

Depending on the applied threshold on the variable importance (in all levels in all models: 65, except species level's *PLS*: 10), the presented features (Fig. 2) are the variables with the highest variable importance selected by the models. Therefore, the total number of presented features would change accordingly with a lower or higher applied threshold. However, in the analytical level FG, we found the majority of unique features not to be selected by the models (Fig. 2 a). In this level, *PLS* selected most features from the pool of non-unique features and shared a major part with the SVMs models. *RF* only selected 5% of the features that PLS and SVMs selected. At this level, all models selected their features mainly from the collection of non-unique features. However, two features were selected by all models from the pool of unique features. In the level of species, most of the unique features were not selected by any of the models (Fig. 2 b). However, *PLS* and *SVM* took a majority of their features from the pool of unique features, while *RF* took less than 1% of these features. *RF* was found to draw almost equally from both pools unique and non-unique. There were no features in common between all models and the pool of unique features. On the analytical level season, *SVM*s drew almost as much of their features from the pool of unique features as from the pool of non-unique features (Fig. 2 c), while *PLS* and *RF* selected a majority of the features from the pool of unique features. In general, *RF* selected only 3% compared to the number of *PLS* selected features. In total, 36 features were selected from all models within the pool of unique features. In the analytical level diversity (Fig. 2 d), most of the unique features were not selected by the models. However, all three models selected a majority of their features from the pool of unique features. In total, 14 features were selected by all models from the pool of unique features in this analysis level.

Here, we compared the model performances based on their accuracies, i.e. the rate of correctly identified samples and the Kappa value. High Kappa values (>0.8 strong, >0.9 almost perfect) support the reliability of the calculated accuracies (Fig. 3, Table 2). Models with an accuracy level above 0.8 were considered reliable classification models. In general, the data matrices *ImpLog* and *ZeroInt* were found to have the highest accuracies. In the levels FG and Species, the overall mean accuracy was very high and supported greatly by the Kappa value. The differences between the data matrices in these two levels were found to be rather small (Table 2). ImpLog and ZeroInt showed the highest mean accuracies in FG, with 0.984 to 1.0 for *ZeroInt* and 0.997 to 1.0 for *ImpLog*. At the species level, the accuracies had a greater range between the data matrices. *ImpLog* had accuracies between 0.975 and 0.982, followed by *ZeroInt* with 0.955 to 0.978. In the levels of season and diversity, the mean overall accuracy was relatively low compared to the FG and species levels. Here the accuracies ranged from 0.332 to 0.877 and 0.203 to 0.395 for season and diversity, respectively. In these two levels, the *ZeroInt* data matrix had the highest accuracies. In the season level, the accuracies calculated ob the *ZeroInt* and *ImpLog* data matrices were relatively close, with 0.824 to 0.844 for ZeroInt (except *SVMr*: 0.332) and 0.743 to 0.877 for *ImpLog* (except *SVMr*: 0.669). In contrast, accuracies calculated on the *ImpInt* data matrix were overall low at 0346 to 0.787. In season, the accuracies differed greatly between the data matrices, e.g. *SVMr* performed best on the *ImpLog* data matrix (mean accuracy: 0.669), and was very weak with ZeroInt (0.33) and *ImpInt* (0.35) data matrices (Table 2). The accuracies for each species for *SVMr* on *ImpLog* ranged between 0.393 and 0.873, *ZeroInt* between 0.199 and 0.970, and *ImpInt* between 0.210 and 0.531 (Supplement Table A). The models *SVMl* and *SVMp* performed in the same range between accuracies of the different data matrices, with the highest accuracies on the *ZeroInt* data matrix (*SVMl*: 0.832, *SVMp*: 0.839), middle performance on *ImpLog* (*SVMl*: 0.746, *SVMp*: 0.743) and lowest accuracy for the *ImpInt* data matrix (*SVMl*: 0.554, *SVMp*: 0.565). In the

analysis level diversity, the highest accuracies across all models were reached in the *ZeroInt* data matrix (0.243 to 0.395), followed by *ImpLog* (0.203 to 0.380) and *ImpInt* (0.206 to 0.362). Furthermore, in the levels of season and diversity, the Kappa values were found to be rather low (season: 0.087 to 0.784, diversity: 0.013 to 0.174), except for *PLS* on the *ImpLog* data matrix in season (0. 827). Kappa values below 0.8 are considered not very strong and do not support the calculated accuracies (Table 2).

Comparing the model performance, we found *PLS* and *RF* generally with the highest accuracies compared to the *SVMs*. All models generally showed high accuracy in FG and species' analytical levels (Table 2). In FG, *PLS* had the highest accuracy (0.998 to 1.000) along with *SVMl* (0.998 to 1.000), followed by *SVMp* (0.997 to 1.000) and *RF* (0.996 to 1.000). *SVMr* showed the lowest accuracy (0.984) within this level. We found the highest accuracy and widest range in the analysis level species in *PLS* (0.943 to 0.982). *RF* (0.975 to 0.976), *SVMp* (0.970 to 0.978), and *SVMl* (0.968 to 0.977) performed almost equally well. *SVMr* had a slightly lower accuracy (0.955 to 0.979). In both levels of FG and species, the mean Kappa was above 0.9. In season, *PLS* was found to have the highest accuracy with the data matrices *ZeroInt* (0.824) and *ImpLog* (0.877), whereas *ImpInt* had a rather low accuracy (0.695). This performance was followed by *RF* (0.787 to 0.844), *SVMl* (0.746 to 0.832, except *ImpInt*: 0.554) and *SVMp* (0.743 to 0.839, except *ImpInt*: 0.565). Here, *SVMr* also showed very low accuracy levels (0.332 to 0.669). In this analysis level, the Kappa values were only slightly above 0.7 in *RF* (all matrices), *SVMl* (*ZeroInt*) and *SVMp* (*ZeroInt*), and between 0.784 and 0.877 for *PLS* on the *ZeroInt* and *ImpLog* data matrix, respectively. All other models had a Kappa below 0.7 (Table 2). All models were found to have generally low accuracies at the diversity level. *PLS* (0.362 to 0.395) were slightly higher than *RF* (0.296 to 0.351), *SVMp* (0.258 to 0.378) and *SVMl* (0.241 to 0.355). The accuracies in *SVMr* even stayed below 0.3 (0.203 to 0.243). All Kappa values in the diversity analysis level were below 0.18 (Table 2).

We compared the predictive power of each model by using the rates for true positive and negative predictions (Table 3). In the analysis level FG, both the true positive and true negative rates were at 1.0, meaning all samples were predicted correctly, with no false positives or false negatives. Only in *RF* we found a less accurate prediction (0.995). In the level of species, *PLS* (pos: 0.981, neg: 0.999), *RF* (pos: 0.974, neg: 0.998) and *SVMl* (pos: 0.808, neg: 0.966) had a very high prediction rate for both true positives and negatives (Table 3). However, in *SVMr* and *SVMp*, we found a slightly less accurate prediction rate for true positives (*SVMr*: 0.742, *SVMp*: 0.792) but an equally good rate for the true negatives (*SVMr*: 0.964, *SVMp*: 0.968). In the analysis level season, we found the *PLS* model to predict a true positive rate above 0.8, followed by *RF* with 0.795. The *SVM* models in this analysis level ranged from 0.649 to 0.693. The best true negative rates were also found in *PLS* (0. 960) and *RF* (0.934), while this rate ranged from 0.886 and 0.898 in the *SVM* models (Table 3). The true negative rates in the analysis level diversity were all above 0.7. However, the positive prediction rates in this level were below 0.4, with the best rates in *PLS* (0. 358) and *SVMr* (0.314). *SVMl*, *SVMp* and *RF* showed a prediction rate of true positives below 0.3 (Table 3).

4 Discussion:

The key goal of this study was to compare the performance of the three statistical tools, *PLS*, *RF* and *SVM,* in terms of their suitability for eco-metabolomics data sets.

The first question we addressed was if classification models are generally needed for a reliable classification on different levels of distinctiveness in metabolomic profiles. In general, we found that the overall composition of the metabolomic profiles was most important for classification as the pool of unique features that the models could choose from strongly depended on the provided background of metabolomic profiles (either across or within species). For example, in the levels FG and species where we compared distinct metabolomic profiles across species, we found that non-unique features were used throughout the classification process. On the other hand, the models selected more unique features in season and diversity, where the classification was performed within each species and hence on a less distinct background. As Gromski et al. (2015) discussed, using more than one feature selection method, especially when used on a data set with unknown structures or a combination of them, might improve the comprehensive interpretation of the results. In our study, a small number of features were also globally selected by all three models, which makes it likely that they, even if they might not play a key role in the classification, are of potential interest and worth a second look. The annotated and identified corresponding metabolites can then be used in more in-depth studies to prove their relevance within the ecosystem services.

The second main question concerned the handling of missing values. As discussed by Mendez et al. (2019), data sets should be curated well and pass thorough quality control checks before they are used to build a classification model. In this study, the data matrix used significantly influenced the classification performance in

all methods. In general, we found that in the levels that compared profiles across multiple species and therefore had relatively little overlap in the metabolic profiles, *ImpLog* (using log-transformed intensities and randomly imputed low values for missing values) was the preferred choice. In contrast, in the levels that compared profiles within one species and were more challenging to separate based on their profiles, classification models worked better on the *ZeroInt* (using measured intensities and zeros for missing values) data matrix. The *ImpInt* matrix, which used measured intensities with randomly imputed low values for missing values, showed generally lower accuracies across all analysis levels. Although imputed with very low values, in the *ImpInt* data matrices, the differences between the measured intensities and the imputed data are most likely not determined enough to provide a good classification base. In contrast, in *ZeroInt*, intensity values are distinguished with zeros from missing values. In the log-transformed imputed data matrix, all values are compelled in a normal distribution, making it easier to separate and classify.

However, in the levels of season and diversity, the differentiation between the plant samples across the growing season (season) and between different biogeographic legacies (diversity levels) depended on the respective species. Some profiles of the tested plant species were comparably easy to classify across different seasons and diversity levels, where these environmental dynamics could not be classified in some other species. In season, both *ZeroInt* and *ImpLog* data matrices showed the best overall accuracies. Although generally low in classification performance, the *ZeroInt* data matrix showed the best classification rate in the diversity levels. The *ZeroInt* data matrix was the most suitable at this level, probably due to the great overlap in metabolic profiles and the more explicit distinction between them based on the zeros for missing values.

We found that PLS seemingly worked best with the *ImpLog* data matrix as model-matrix combinations, followed by *ZeroInt*. *RF*, across all levels of analysis, worked best on the *ZeroInt* data matrix. *SVMl* and *SVMp* worked equally well with *ZeroInt* and *ImpLog* data matrices. *SVMr*, although showing only a poor performance in general, worked best with the *ImpLog* data matrix. Both model-matrix strategies, *ZeroInt* (SVMl, SVMp, RF) and *ImpLog* (*PLS, SVMr*), worked well but were also highly dependent on the different plant species in the seasonal and diversity levels, with some species being reliably classified and some showing almost no differences in their profiles. The choice of models and data matrix influenced the classification performance in season and diversity significantly. In season, for example, *PLS* on *ImpLog* and *RF* on *ZeroInt* showed a rather good performance, whereas, in the diversity level, both *PLS* and *RF* worked best on the *ZeroInt* data matrix. In the analysis level FG and species, the accuracies of all model-data matrix combinations varied very little. However, at the Species level, where the species-specific metabolic profiles are distinctive but generally share some more features between closely related species, we found that the choice of data matrix did change the performance of the tested models to a greater deal than in the FG levels. *ImpLog* seemingly worked best for *PLS* and *SVM*, while *RF* performed better with the *ZeroInt* data matrix.

The third major question in this study was to compare the model performances on each analysis level. Both the FG and Species analysis levels showed great overall accuracy in all tested models with only slight variations. The accuracy of this one-vs-rest approach in these levels, although enabling the models to distinguish between metabolic profiles, highly depends on the profiles they are compared to. In all tested analysis levels, we found the linear classifier *PLS* to perform best in terms of accuracies and Kappa values. *RF* showed very similar values compared to *PLS*, except in FG, where *SVMl* and *SVMp* performed slightly better. RF showed a comparatively good performance in the analysis level season, although without a supporting Kappa value. In the analysis level of diversity, we found no suitable classification strategy in most species. In general, the classification in the level of diversity is most challenging due to the underlying experimental setup. The sampled plant communities are not replicated in the Trait-Based-Experiment. Each diversity level has a different plant community as a reference and, therefore, unknown diverse effects on the metabolomic profiles. Additionally, at the diversity level, the models have to classify across the seasonal level, which comes with a greater variety in the metabolomic profiles. Here, *RF* was found to be the second-best model after *PLS*.

Across all four analysis levels, both the linear and poly kernel models in *SVM* were always close to each other in terms of accuracies and Kappa values. Our analysis found that *SVMp* often showed the best results when using one dimension for the data separation, which makes it a linear classification model. This most likely occurred due to the ability of the polynomial kernel to test different numbers of dimensions (degree) that worked best for the classification of each group (Burges, 1998; Statnikov et al., 2008). *SVMl* and *SVMp* are very close in their performance but are not as accurate as *PLS* and *RF*. *SVMr* seemed the least suitable for these data types in terms of performance across all data sets and within the analysis level (Burges, 1998).

Addressing the underlying structures of the measured features could significantly improve the classification performance (Tang et al., 2014; Hageman et al., 2017). However, in this study, we found excellent classifications at FG and species level, especially in PLS models, which might result from the data's underlying structure. The

data sets we used in this study included species whose secondary metabolite patterns are highly diverse and allowed for an excellent classification per se (Marr et al., 2021). Another reason for getting those almost perfect classifications in the FG and species levels may occur due to the over-optimistic prediction of the eco-metabolomics data. For example, the imbalance between the number of features and samples can increase the ratio of features that correlate with the initial research question just by chance or lead to overfitting models. An overfitting model would represent the training data nearly perfectly while lacking the ability of generalisation (Broadhurst & Kell 2006, Tang et al. 2014, Gromski et al. 2015). Furthermore, as discussed by Mendez et al. (2019), carefully choosing parameters while training the model is essential to avoid drawing wrong conclusions. Therefore, the outcome and interpretation of results can be very different depending on the methods used.

5 Conclusion:

In eco-metabolomics, choosing a suitable classification method relies on multiple conditions. Therefore, which method is better suited for a specific problem may be based on the combination of strengths and limitations of the classification and feature selection method concerning the individual data sets. Depending on the goal of the classification and feature selection, we found *RF* to be a great choice for selecting only a few features that are needed for the classification. In contrast, *PLS* selected a greater number of features that were involved to some degree in the classification of the classes. For the usage of the latter method, the feature selection goal could be to find a fixed number of the most relevant features for separating the groups. However, due to its underlying structure, *PLS* has a tendency to overfit the provided data and might not be the best choice as a prediction model.

The careful handling of the data matrix can also improve the performance of the chosen model. This study found that the more similar the metabolic profiles are, the better the *ZeroInt* data matrix worked with the models. In general, we found that the overall composition of the metabolomic profiles was most important for classification. Depending on the complexity of the provided background for classification, *PLS*, *SVM,* and *RF* work comparably reliably when distinguishable metabolomics profiles are used, e.g. when classifying across different species. PLS and RF provide a greater classification power when the provided background is similar, e.g. when comparing different treatments within one species. We, therefore, suggest using the introduced feature selection tools with care and as observatory tools for eco-metabolomics data sets to reduce the amount of available data and retrieve relevant information. Selected features and identified metabolites can then be used in more in-depth studies to prove their relevance within ecosystem services.

Supplementary information

Model calculations of all species and factors are available in the comprehensive supplemental table 1 (see Appendix – Chapter 2).

Funding

This paper has been conducted in the framework of the iDiv-Flexpool – the internal funding mechanism of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – DFG–FZT 118, 202548816. This manuscript is part of the *Metabolite Changes in Biodiversity Levels and Seasonal Shifts* project (MacBeSSt; W47013118Ü).

Conflict of interest

The authors declare no conflict of interests.

Availability of Data and Code

The data sets MTBLS679, MTBLS1224, MTBLS2140, including raw data files and metadata, are available at MetaboLights as http://identifiers.org/metabolights:MTBLS679, http://identifiers.org/metabolights:MTBLS1224 (in curation) and http://identifiers.org/metabolights:MTBLS2140 (in curation). The code, used for calculation and to produce the figures is available as downloadable R-script in the files of the MTBLS1224 repository (http://identifiers.org/metabolights:MTBLS1224).

Author contribution

SM, JAH and SN conceived and designed research. SM provided data sets. SM and SN uploaded and curated the data repositories. SM and JAH performed the analysis. SM wrote the manuscript. All authors read and approved the manuscript.

Acknowledgements

The authors gratefully acknowledge the support of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – DFG–FZT 118, 202548816. The authors also thank the technical staff of the Jena Experiment (DFG, FOR 456/1451) and the coordinator Anne Ebeling for maintaining the experimental field site in which the data sets MTBLS679 and MTBLS1224 were collected. Furthermore, many thanks to the gardeners in charge at the IPB Halle, for the support of growing the plant individuals used for the data set MTBLS2140. This research was supported by the International Research Training Group TreeDì jointly funded by the DFG (GRK2324) and the University of Chinese Academy of Sciences (UCAS).

References

Barker, M., & Rayens, W. (2003). Partial least squares for discrimination. Journal of Chemometrics: A Journal of the Chemometrics Society, 17(3), 166-173.

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Brereton, R. G., & Lloyd, G. R. (2014). Partial least squares discriminant analysis: taking the magic away. Journal of Chemometrics, 28(4), 213-225.

Broadhurst, D. I., & Kell, D. B. (2006). Statistical strategies for avoiding false discoveries in metabolomics and related experiments. Metabolomics, 2(4), 171-196.

Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. Data mining and knowledge discovery, 2(2), 121-167.

Christin, C., Hoefsloot, H. C., Smilde, A. K., Hoekman, B., Suits, F., Bischoff, R., & Horvatovich, P. (2013). A critical assessment of feature selection methods for biomarker discovery in clinical proteomics. Molecular & Cellular Proteomics, 12(1), 263-276.

Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. Ecology, 88(11), 2783-2792.

David, M., Evgenia, D., Kurt, H., Andreas, W., & Friedrich, L. (2019). e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1. 7–6.

Ebeling, A., et al. & Weisser, W. W. (2014). A trait-based experimental approach to understand the mechanisms underlying biodiversity–ecosystem functioning relationships. Basic and Applied Ecology, 15(3), 229-240.

Fiehn, O. (2002). Metabolomics—the link between genotypes and phenotypes. In Functional genomics (pp. 155-171). Springer, Dordrecht.

Genuer, R., Poggi, J. M., & Tuleau-Malot, C. (2010). Variable selection using random forests. Pattern recognition letters, 31(14), 2225-2236.

Giacomoni, F., Le Corguillé, G., Monsoor, M., Landi, M., Pericard, P., Pétéra, M., ... & Goulitquer, S. (2015). Workflow4Metabolomics: a collaborative research infrastructure for computational metabolomics. Bioinformatics, 31(9), 1493-1495.

Gromski, P. S., Muhamadali, H., Ellis, D. I., Xu, Y., Correa, E., Turner, M. L., & Goodacre, R. (2015). A tutorial review: Metabolomics and partial least squares-discriminant analysis–a marriage of convenience or a shotgun wedding. Analytica chimica acta, 879, 10-23.

Hageman, J. A., Engel, B., de Vos, R. C., Mumm, R., Hall, R. D., Jwanro, H., ... & van Eeuwijk, F. A. (2017). Robust and confident predictor selection in metabolomics. In Statistical analysis of proteomics, metabolomics, and lipidomics data using mass spectrometry (pp. 239-257). Springer, Cham.

John, G. H., Kohavi, R., & Pfleger, K. (1994). Irrelevant features and the subset selection problem. In Machine Learning Proceedings 1994 (pp. 121-129). Morgan Kaufmann.

Kriegl, J. M., Arnhold, T., Beck, B., & Fox, T. (2005). A support vector machine approach to classify human cytochrome P450 3A4 inhibitors. Journal of computer-aided molecular design, 19(3), 189-201.

Kuhn, M. (2020). caret: Classification and Regression Training. R package version 6.0-86. https://CRAN.R-project.org/package=caret

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. R news, 2(3), 18-22.

Mahadevan, S., Shah, S. L., Marrie, T. J., & Slupsky, C. M. (2008). Analysis of metabolomic data using support vector machines. Analytical chemistry, 80(19), 7562-7570.

Marr, S., Hageman, J. A., Wehrens, R., van Dam, N. M., Bruelheide, H., & Neumann, S. (2021). LC-MS based plant metabolic profiles of thirteen grassland species grown in diverse neighbourhoods. Scientific data, 8(1), 1-12.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. Biochemia medica, 22(3), 276-282.

Mendez, K. M., Reinke, S. N., & Broadhurst, D. I. (2019). A comparative evaluation of the generalised predictive ability of eight machine learning algorithms across ten clinical metabolomics data sets for binary classification. Metabolomics, 15(12), 150.

MetaboLights http://identifiers.org/metabolights:MTBLS679. (2017)

MetaboLights http://identifiers.org/metabolights:MTBLS1224. (2018)

MetaboLights http://identifiers.org/metabolights:MTBLS2140. (2019)

Mevik, B. H., Wehrens, R., & Liland, K. H. (2019). pls: Partial least squares and principal component regression. R package version 2.7-2.

Pang, Z., Chong, J., Zhou, G., Morais D., Chang, L., Barrette, M., Gauthier, C., Jacques, PE., Li, S., and Xia, J. (2021) MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights Nucl. Acids Res. (doi: 10.1093/nar/gkab382)

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sanz, H., Valim, C., Vegas, E., Oller, J. M., & Reverter, F. (2018). SVM-RFE: selection and visualisation of the most relevant features through non-linear kernels. BMC bioinformatics, 19(1), 1-18.

Sattlecker, M., Bessant, C., Smith, J., & Stone, N. (2010). Investigation of support vector machines and Raman spectroscopy for lymph node diagnostics. Analyst, 135(5), 895-901.

Shi, L., Westerhuis, J. A., Rosén, J., Landberg, R., & Brunius, C. (2019). Variable selection and validation in multivariate modelling. Bioinformatics, 35(6), 972-980.

Smilde, A. K., Jansen, J. J., Hoefsloot, H. C., Lamers, R. J. A., Van Der Greef, J., & Timmerman, M. E. (2005). ANOVA-simultaneous component analysis (ASCA): a new tool for analysing designed metabolomics data. Bioinformatics, 21(13), 3043-3048.

Statnikov, A., Wang, L., & Aliferis, C. F. (2008). A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. BMC bioinformatics, 9(1), 319.

Strobl, C., Boulesteix, A. L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC bioinformatics, 8(1), 25.

Tang, J., Alelyani, S., & Liu, H. (2014). Feature selection for classification: A review. Data classification: Algorithms and applications, 37.

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. Chemometrics and intelligent laboratory systems, 58(2), 109-130.

Figures

Figure. 1 Data set overview. All data sets are composed of species (pink) from the functional groups (FG; green) grasses (light green panel) and herbs (dark green panel). The data set MTBLS2140 (orange panel) includes four additional grass species. In the data sets MTBLS679 and MTBLS1224 (grey panel) each species was sampled in four different seasons (late spring, early summer, late summer, early autumn; peach) from plots with four different diversity levels (DL; pale purple).

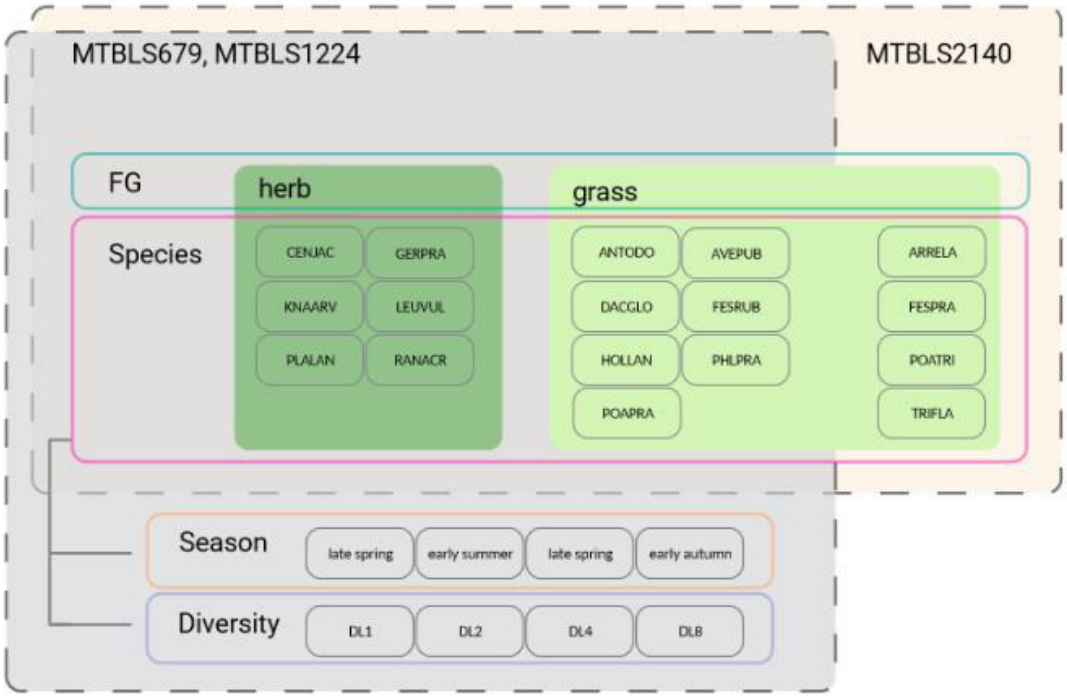
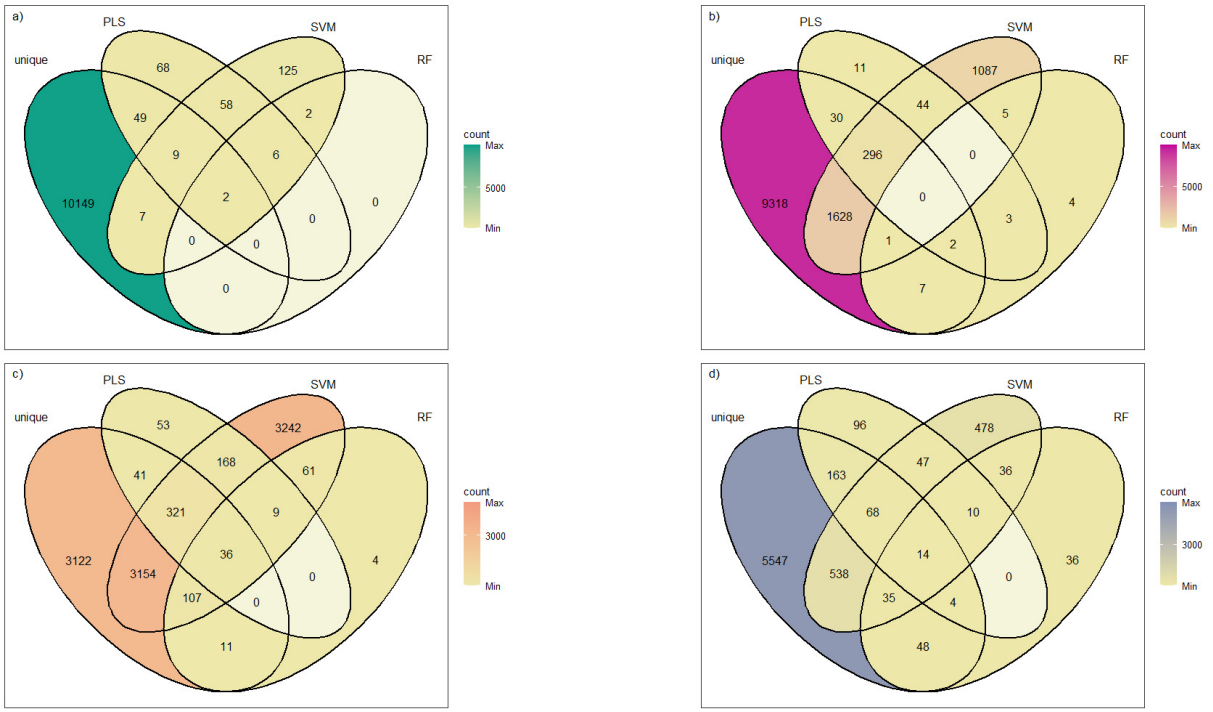
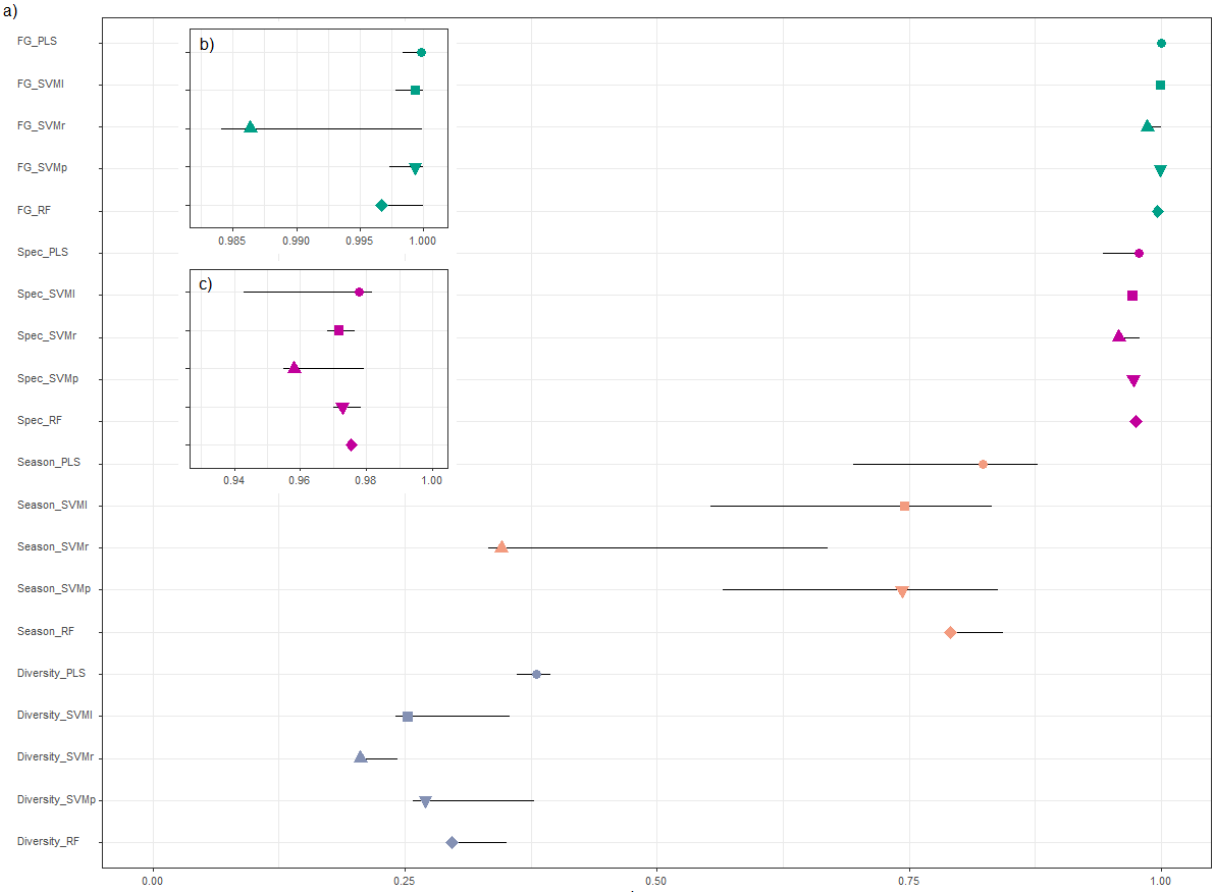


Figure 2 Features selected by classification methods, compared to unique features of each analysis level. For plotting purposes, we combined model selected features from all calculated factors within the three data sets (MTBLS679, MTBLS1224, MTBLS2140). *SVMl*, *SVMr*, and *SVMp* models are combined to *SVM* to increase the readability of the figure. In all analysis levels, we used 65 as the variable importance threshold for *PLS*, *SVM*, and *RF* (except in Species, *PLS*: 10). In each analysis level, we selected unique features separately and combined them across the three data sets (MTBLS679, MTBLS1224, MTBLS2140). a) In the analysis level FG (MTBLS679, MTBLS1224, MTBLS2140), features were selected as unique that only occurred in either one of the groups grass or herb. b) In the species level (MTBLS679, MTBLS1224, MTBLS2140), we selected unique features per species based on the one-vs-rest concept. c) Unique features in the Season level (MTBLS679, MTBLS1224) and d) Diversity level were selected within each species for each Season and Diversity level (MTBLS679, MTBLS1224), respectively, using the one-vs-rest approach.



,

Figure 3 Model performances compared based on their accuracy. Calculated accuracies of each data matrix (*ZeroInt, ImpInt, ImpLog*) are used to show the range within one model in one analysis level. The analysis levels are FG (green), species (deep pink), season (peach) and diversity (pale purple). The calculated models are PLS (circle), *SVMl* (square), *SVMr* (up pointing triangle), *SVMp* (down pointing triangle), and *RF* (diamond). The data matrices are arranged to show the minimal, medium and maximal accuracy value, which was reached by the respective model in each analysis level. Compare **Table 2**. The levels FG and species are additionally zoomed in the partial plots b (FG) and c (Spec) to increase the visibility and comparability. Note the different ranges on the x axis in both plots.



Tables

Table 1 Data set overview. All data sets (MTBLS679, MTBLS1224, MTBLS2140) are composed of species belonging to the functional groups (FG) grasses and herbs. The data set MTBLS2140, additionally, contained four grass species. Samples in the data sets MTBLS679 and MTBLS1224 were collected in four different seasons from plots with four different diversity levels.

Level	Factors	Abb	Data Set
FG	grass, herb		MTBLS679, MTBLS1224, MTBLS2140
Species	herb:		MTBLS679, MTBLS1224, MTBLS2140
	Centaurea jacea,	CENJAC	
	Geranium pratense,	GERPRA	
	Knautia arvensis,	KNAARV	
	Leucanthemum vulgare,	LEUVUL	
	Plantago lanceolata,	PLALAN	
	Ranunculus acris	RANACR	
	grass:		
	Anthoxanthum odoratum,	*ANTODO*	
	Avenula pubescens,	AVEPUB	
	Festuca rubra,	FESRUB	
	Holcus lanatus,	HOLLAN	
	Phleum pratense,	PHLPRA	
	Poa pratensis,	POAPRA	
	Arrhenaterum elatius,	*ARRELA*	MTBLS2140
	Festuca pratensis,	FESPRA	
	Poa trivialis,	POATRI	
	Trisetum flavescens	TRIFLA	
Season	late spring, early summer, late summer, early autumn		MTBLS679, MTBLS1224
Diversity	DL 1, DL2, DL,4, DL8		MTBLS679, MTBLS1224

Table 2 Model performances compared across levels of analysis and data matrices. Models were calculated for each data set (MTBLS679, MTBLS1224, MTBLS2140) separately. In the data sets MTBLS679 and MTBLS1224, the levels season and diversity were calculated for each species (Spec) separately. We compare the model performances based on their accuracy and the corresponding Kappa. In order to increase readability, the table shows the mean accuracy (mAc) and mean Kappa (mKa), averaged across the data sets (MTBLS679, MTBLS1224, MTBLS2140). Additionally, for the analysis levels season and diversity, the mAc and mKa were averaged across all species. Reliable mAc (mKa > 0.6) are highlighted in bold font. A comprehensive table, including all calculations and the best model choices, are available in Supplemental Table 1 (Online Resource 1).

	Level	FG		Spec		Season		Diversity	
Model	Data matrix	mAc	mKa	mAc	mKa	mAc	mKa	mAc	mKa
PLS	*ZeroInt*	**1.000**	1.000	**0.978**	0.976	**0.824**	0.784	0.395	0.146
	ImpInt	**0.998**	0.995	**0.943**	0.940	0.695	0.584	0.362	0.154
	ImpLog	**1.000**	1.000	**0.982**	0.980	**0.877**	0.827	0.380	0.174
SVMl	*ZeroInt*	**0.999**	0.999	**0.971**	0.969	**0.832**	0.765	0.355	0.140
	ImpInt	**0.998**	0.996	**0.968**	0.965	0.554	0.398	0.253	0.023
	ImpLog	**1.000**	1.000	**0.977**	0.974	**0.746**	0.655	0.241	0.042
SVMr	*ZeroInt*	**0.984**	0.963	**0.955**	0.951	0.332	0.087	0.243	0.016
	ImpInt	**0.986**	0.969	**0.958**	0.954	0.346	0.159	0.206	0.013
	ImpLog	**1.000**	1.000	**0.979**	0.977	0.669	0.570	0.203	0.047
SVMp	*ZeroInt*	**0.999**	0.999	**0.973**	0.970	**0.839**	0.767	0.378	0.162
	ImpInt	**0.997**	0.994	**0.970**	0.967	0.565	0.416	0.270	0.065
	ImpLog	**1.000**	1.000	**0.978**	0.976	**0.743**	0.653	0.258	0.065
RF	*ZeroInt*	**1.000**	1.000	**0.976**	0.974	**0.844**	0.782	0.351	0.150
	ImpInt	**0.996**	0.991	**0.975**	0.984	**0.787**	0.710	0.296	0.116
	ImpLog	**0.997**	0.992	**0.975**	0.973	**0.791**	0.715	0.297	0.115

Table 3 Mean values of true positive prediction rate (pos rate) and true negative prediction rate (neg rate) for all levels in the respective data sets. Predictions are calculated on each level separately. For the level FG the rates are predicted for grass vs herb species. For the species level, the rates are calculated per species vs rest, for the levels season and diversity, the rates are calculated in each species per season and per diversity level, respectively.

Level	Model	MTBLS	pos rate	neg rate
FG	*PLS*	679, 1224, 2140	1.000	1.000
	SVMl	679, 1224, 2140	1.000	1.000
	SVMr	679, 1224, 2140	1.000	1.000
	SVMp	679, 1224, 2140	1.000	1.000
	RF	679, 1224, 2140	0.995	1.000
Species	*PLS*	679, 1224, 2140	0.981	0.999
	SVMl	679, 1224, 2140	0.808	0.966
	SVMr	679, 1224, 2140	0.742	0.964
	SVMp	679, 1224, 2140	0.792	0.968
	RF	679, 1224, 2140	0.974	0.998
Season	*PLS*	679, 1224	0.880	0.960
	SVMl	679, 1224	0.661	0.892
	SVMr	679, 1224	0.649	0.886
	SVMp	679, 1224	0.693	0.898
	RF	679, 1224	0.795	0.934
Diversity	*PLS*	679, 1224	0.358	0.796
	SVMl	679, 1224	0.239	0.749
	SVMr	679, 1224	0.314	0.771
	SVMp	679, 1224	0.275	0.761
	RF	679, 1224	0.290	0.770

CHAPTER 3

The sign of the thirteen: metabolite profile dynamics across season and diversity in thirteen grassland species

Title: The sign of the thirteen: metabolite profile dynamics across season and diversity in thirteen grassland species

Authors:
Sue Marr [1,2,3*] & Steffen Neumann[1,3]

[1]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, 04103, Germany
[2]Martin Luther University Halle-Wittenberg, Institute of Biology/Geobotany and Botanical Garden, Halle, 06108, Germany
[3]Leibniz Institute of Plant Biochemistry, Bioinformatics & Scientific Data, Halle, 06120, Germany
*corresponding author

Author contribution:

Author contributions (ctb) to the manuscript in percent.

chapter	status	author	ctb
(3) The sign of the thirteen: investigating metabolite fingerprints across season and diversity in thirteen grassland plant species	draft	Sue Marr	98
		Steffen Neumann	2

The sign of the thirteen: metabolite profile dynamics across season and diversity in thirteen grassland species

Sue Marr[1,2,3*] and Steffen Neumann[1,3]

[1]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, 04103, Germany
[2]Martin Luther University Halle-Wittenberg, Institute of Biology/Geobotany and Botanical Garden, Halle, 06108, Germany
[3]Leibniz Institute of Plant Biochemistry, Bioinformatics & Scientific Data, Halle, 06120, Germany
[*]smarr@ipb-halle.de

Abstract

The impact of plant species diversity on productivity in plant communities has already been shown for multiple ecosystems. Secondary metabolites and how they shape the interaction within these plant networks, on the other hand, are increasingly becoming an important part of ecological studies. However, the level and extent to which these interactions occur are not well understood yet. In this study, we investigated the LC-MS profiles of thirteen grassland species grown in diverse communities across the growing season within the Jena experiment. We found a unique metabolite composition for all species in this study, shaping a species-specific fingerprint across the year and environmental dynamics. The specific part of the metabolite profiles actively responding to environmental pressures was found to directly correlate with plant species richness and composition in some of the species. These dynamic defence profiles generally showed similar changes in closely related plant species.

Introduction

Species loss in natural ecosystems has increased the motivation for studies investigating the relationships between plant diversity on the productivity of ecosystems (Rasmann & Turlings, 2016). Beneficial effects of biodiversity are reported throughout the literature (Lorentzen et al. 2008, Marquard et al. 2009), such as increased plant performance, resource management and overall resilience, and plants' extreme capability to adapt rapidly to changing environmental conditions. As pointed out by Scherling et al. (2010), biodiversity can trigger a multitude of effects on the performance of plant species or species communities. For example, in monocultures, plants would have to invest many resources in the same specific and precise defence strategies to tackle biotic and abiotic pressures, while diverse plant communities can benefit from shared defence mechanisms among the different species without having them present in all plants simultaneously. Representing the direct link between genotype, and phenotype (Fiehn 2002, Weston et al. 2015, Fang et al. 2019), including central (primary) and specialised (secondary) metabolism, metabolites serve as the reflection of the chemical response of plants to environmental changes (Fiehn 2002, Weston et al. 2015). Especially the interplay of secondary metabolites with ecological traits (Walker et al. 2022), in combination with proteomics and transcriptomics, are the focus of studying complex plant-plant and plant-herbivore interactions within their environment (Weston et al. 2015) and evolutionary concepts (Fang et al. 2019). As pointed out by Gargallo-Garriga et al. (2020), metabolomic variation can be driven by herbivores, as plants have to balance between primary (growth, reproduction) and secondary metabolism (defence, communication). When studying stress coping strategies, including biotic and abiotic interactions with environmental changes, metabolites are a valuable addition to traditional ecological traits as they reflect environmental conditions in their richness and composition throughout the plant kingdom (Scherling et al. 2010, Berini et al. 2018, Fang et al. 2019). Responses in plant metabolite profiles, as response to changing diversity in plant communities, can give insights into the underlying mechanisms that shape the herbivory pressure in those communities Ristok et al. (2022). For example, identifying species-specific profiles could narrow the number of metabolites relevant to defence and stress-coping strategies. For example, previous studies already showed the importance of specialised secondary metabolites in plant-insect interactions (Tikunov et al. 2005), leading to the assumption that some metabolites serve a specialised purpose in multiple plant species (Fiehn 2002, Weston et al. 2012, Quinn et al. 2014, Sardans et al. 2020), i.e. an active involvement in plant-plant and plant-herbivore interactions. However, most secondary metabolites' functions remain unexplained and only little understood (Fiehn 2002, Fang et al. 2019). Moreover,

previous studies showed that many metabolites are species specific (Weston et al. 2012, Quinn et al. 2014). Leading to the assumption that some parts of the metabolite profiles will be conserved and inherent to specific plant families. In order to investigate the function of specific metabolites, we first need to understand the interplay of the metabolomic profiles with biotic and abiotic environmental changes. This can also enable the research to understand the role of compounds as responses to these pressures (Berini et al. 2018).

Especially plant-plant interactions can influence the plant metabolome in specific ways, for example, volatiles (Tikunov et al. 2005) and other allelopathic compounds (Fernandez et al. 2016), which reflect both short and long-term defence mechanisms in plant species. Some studies are already investigating plant biodiversity effects on metabolomic profiles and physiological responses of plants (Scherling et al. 2010). A positive influence of phytochemical diversity on plant species richness and composition has also been reported in the literature (Richards et al 2015). Changes to the metabolomic profile is also associated with changing visible traits, especially along the developmental gradient. Along with the developmental stage, plant height per se can influence the metabolite profiles, for instance, due to the availability of light, and thereby imposed stress levels (Scherling et al. 2010). As discussed by Walker et al. (2022), the responses to environmental dynamics can be measured in the richness and composition of secondary metabolites (Fang et al. 2019), as the metabolite abundance is driven by abiotic (Sardans et al. 2020) and biotic pressures (Weston et al. 2015, Ristok et al. 2022). Previous studies already showed that certain biotic and abiotic stress factors, e.g. temperature (Yang et al. 2018), change specific metabolites in terms of richness and abundance, which leads to the assumption that the same stress will cause the same groups of metabolites to respond within similar plant species. For some of the secondary metabolites, such as phenolic compounds, alkaloids, and terpenoids, studies have already shown that their production is increased in different plant species in response to herbivory pressure (Fang et al. 2019). This study investigates the metabolite profiles of thirteen grassland plant species belonging to either functional group (FG) herb or grass. Thereby, functional groups are assembled with plant species that share similar characteristics, such as growth form and reproduction strategies.

Furthermore, in this study, we try to show the variation in metabolite profiles based on physiological traits and developmental stages and on the changes under the influences of seasonal dynamics (i.e. abiotic factors, such as light, water, nutrient availability and biotic factors, such as herbivory pressure), especially those triggered by neighbourhood richness and species composition. As metabolomic profiles represent both the developmental stage of a plant optimising performance across seasonal dynamics and at the same time investing in defence strategies to minimise stress caused by biotic and abiotic factors, we expect that these two strategies are also visible in the metabolomic profile. Across the season, environmental factors change dramatically, including temperature, water and light availability and nutrient demand depending on the plant species that grow within the community (Scherling et al. 2010). The variation in these biotic and abiotic factors will be reflected in the metabolomic fingerprints of the plant species by increased production of defence compounds (Weston et al. 2013, Berini et al. 2018). Analysing environmental changes across the growing season can help to capture the specific changes in metabolomic profiles that reflect a stress response within the plant species. Ristok et al. (2022) showed in their study that collecting samples from the same environments in different seasons can greatly influence the metabolomic profiles in terms of richness and composition. However, in contrast to visible traits that can be monitored over time (Flynn et al. 2008), metabolomic profiles only represent snapshots of a specific metabolomic profile at the time of sampling, affecting the ability to draw general conclusions from responses to the environment (Weston et al. 2015). What remains unknown is how the whole metabolomic profile is changing for different species under the same conditions, here across the season, especially across the changing neighbourhood that is changing across the growing season. Therefore, we sampled multiple times in this study to capture the variation in metabolite profiles across the growing season.

Plant species have highly specialised defence mechanisms in place that result in the production of specific metabolites. Assuming that in plant communities with balanced species diversity environmental pressures are significantly lower than in uneven neighbourhoods, and thereby, the production of defence compounds are not required in the same amount. Therefore, in plant species communities with lower or uneven diversity we expect to find more specialised metabolites, while the protective diversity shield will result in balanced metabolomic profiles. Having said that, our three main questions, in this study, are 1) plant species richness and 2) plant species diversity of the plant community reflected in the metabolite profiles, and 3) do we find similarities in species of the same functional group.

Methods

Plant Material Collection

Experimental Design. In this study, we used thirteen grassland plant species as target species that belong to the functional groups (FG) *grass* and *herb*. The seven *FG grass* species belonged to the family *Poaceae* (*Anthoxanthum odoratum* L., *Avenula pubescens* (Huds.) Dumort. (according to http://www.theplantlist.org/, now listed as Helictotrichon pubescens (Huds.) Schult. & Schult.f.), *Dactylis glomerata* L., *Festuca rubra* L., *Holcus lanatus* L., *Phleum pratense* L., *Poa pratensis* L.). The *FG herb* included five genera: *Asteraceae* (*Centaurea jacea* L., *Leucanthemum vulgare* (Vaill.) Lam.), Caprifoliaceae (*Knautia arvensis* (L.) Coult.), *Geraniaceae* (*Geranium pratense* L.), *Plantaginaceae* (*Plantago lanceolata* L.), and *Ranunculaceae* (*Ranunculus acris* L.). The plant samples were collected in the Jena Experiment (Roscher et al. 2004) from plots in the Trait-Based Experiment (Ebeling et al. 2014). We chose 48 plots (3.5 m x 3.5 m), initially planted in 2010 as Pool 1 and Pool 2 (Supple Figure 1), with different plant species compositions (compare Ebeling et al. 2014). Each species was collected from four plots with different initial compositions. The plots were designed with the initial composition of diversity levels (DL) with either one (DL1), two (DL2), four (DL4) or eight (DL8) different species (Figure SP2 (new01)). *Phleum pratense*, *Leucanthemum vulgare*, and *Plantago lanceolata* were collected in additional four plots, resulting in eight plots for each species (Supple Table 1). Note that the plots in the Trait-Based-Experiment are semi-maintained, with two cutting events per growing season. Two specimens of each target species were collected as one replicate, with a total of two replicates in each plot on four occasions (season) across the growing season 2018 (for further details, see also Marr et al. 2021). Due to the design of the Trait-Based Experiment, some plots were used to collect multiple target species. The seasons (S) were chosen as snapshots representing late spring (E; May 2018), early summer (F; June 2018), late summer (G; July 2018) and early autumn (H; August 2018), respectively.

Recording Plant Diversity. Plot-specific factors were recorded prior to each sampling. We recorded the current number of species (richness) for each plot in each season (spring, early and late summer, and autumn). We estimated the abundance of each species as relative coverage in relation to the plot size. For the abundance estimation, we applied a decimal scale modified after the Londo scale using the following scheme: coverage (c) <= 1% equals (→) scale value (sv) 1, c >1% to <= 3% → sv 2, c >2% to <=5% → sv 4, >5% to <=15% → sv 10, >15% to <= 25% → sv 20, >25% to <= 35% → sv 30, et cetera, > 95% → sv 100. Note that the whole coverage of each species was taken into account. Hence, also values above sv 100 were occasionally recorded. These values were used to calculate the Shannon index per plot (species Shannon). The index uses the number of species and their abundances, denoting an individual's probability of belonging to the respective species (Izsàk & Papp 2000). A higher value of this index will indicate a higher diversity, whereas a 0 would imply that only one species is present in the plot. Note that this resulted in replicates of the same plot having the same richness and Shannon indices. Measurements and calculations can be found in Supple Table 1.

Recording Plant Traits (Plant individual factors) and sampling. Plant individual factors were recorded for each collected replicate separately. For each replicate, we measured the height of both plant individuals and recorded the higher value. Following the BBCH scale by Hess et al. (1997), the developmental stage was recorded for the further developed plant individual. However, only those two plants were combined into one replicate, showing both an equal BBCH and similar height. The mechanical and pathogen-inflicted damage was collectively estimated in relation to the biomass of both individuals (for details, see Marr et al. 2021).

In this study, we collected two individual plants with similar developmental stages as one combined replicate. Leaf tissues of each plant were cut, collected in plastic vials, immediately snap-frozen, and stored on dry ice. The samples were collected within one day between midday and sunset for each season (2018_E, 2018_F, 2018_G, 2018_H) to keep possible time-related shifts in the metabolome to a minimum. Each replicate was assigned a random lab-ID as described by Marr et al. (2021). Therefore, a number between 1 and 128 per season was drawn randomly. These lab-IDs then determined the sequence of the following analysis.

Liquid-Chromatography-Mass-spectrometry (LC-MS)

Sample Extraction. This study based the extraction process on the workflow described by Marr et al. (2021). Modifications to this method are stated where applicable. Frozen leaf material was ground (20 mL plastic Vials; 2 x 7 mm steel beads) at -75 °C for 150 seconds (5 x cycles: 30 s grinding, 30 s pausing) using the automated Labman IPB Cryogrinder Ball Mill (Labman Automation, UK). Zircosil beads (1.2 - 1.7 mm) and 500 µL extraction solvent (Methanol/Water (80/20 v/v) were added to 100 mg aliquots (± 50 mg) of the ground frozen samples in 2 mL Extraction-Tubes, while kept on liquid nitrogen. Samples were thawed for 3 minutes and then extracted in a Homogenizer (Precellys® 24 Tissue Homogenizer, Bertin Technologies, USA) for 90 seconds (2 x cycles: 45 s run,

15 s pausing) at 6500 rpm. Samples were then centrifugated at 16168 x g for 15 minutes, and the supernatants were collected in a new 2 mL Extraction-Tube. In contrast to the method used by Marr et al. (2021), after an additional extraction of the remaining pellet, the combined supernatants were evaporated to dryness and then redissolved in a weight-specific five times surplus of the extraction solvent. Prior to the LC-MS measurements, we mixed aliquotes of 160µL pr sample with 40 µL of Water/Formic acid (99.9/0.1 v/v) with added internal standard N-alkyl pyridinium-3-sulfonate (NAPS) (1/100 v/v). The extracts were stored at -20 °C for at least 24 hours before supernatants were centrifugated at 16168 x g for 15 minutes and transferred to HPLC-Vials (Glasinlet, 300 µl). Note that two samples (2018_E_POAPRA_B075_b and 2018_F_ANTODO_A018_a) had to be excluded from further measurements as the amount of collected biomass was too little for the extraction.

Quality Control (QC) and Species Pools. In this study, we used fieldBlanks that underwent the entire process of sampling and transportation, including the grinding preparation in the ball mill. Furthermore, We used blank samples throughout the extraction process to keep track of possible contaminations during this process. Using a 10 µL aliquot of each sample per species, we created a pooled sample for each of the thirteen species (MSMS species pool). Aliquots of these pools were then further combined into functional-group-specific (FG) pools QC_herb and QC_grass, combining all herb and grass species. We prepared the QC_pool by combining aliquots of the two FG specific pools.

LC-MS Measurements. Liquid-Chromatography-Mass-Spectrometry measurements were conducted following the methods proposed by Marr et al. (2021). The analysis was performed on an ACQUITY UPLC (Waters Corporation, Milford, USA) coupled to ESI-micrOToF-Q-II (Bruker Daltonics, Bremen, Germany). Aliquots of 2 µL per sample were separated at 40 °C on an HSS T3 C18 column (1.8 µm, 1.0 x 100 mm, RP, Waters Corporation). The binary elution gradient at 0.15 mL min-1 flow rate was increased linearly: Solvent A (water/formic acid 99.9/0.1 v/v)/Solvent B (acetonitrile/formic acid 99.9/0.1 v/v); initial: A 95%, 3 minutes A 82.7 %,10 minutes A 76 %, 17 minutes A 5 %, 18 minutes A 5 %, 18.1 minutes A 95 %, 20 minutes A 95%. In contrast to the methods used by Marr et al. (2021), features were measured in positive ion mode using internal calibration (sodium formate clusters, 10 mM sodium hydroxide in isopropanol/water/formic acid, 49.9/ 49.9/0.2 (v/v/v), at 18 min) and following instrument settings: nebuliser gas: nitrogen; nebuliser: 1.6 bar; dry gas: nitrogen; dry gas temperature: 190 °C; dry gas flow: 6 L min-1; spectra rate: 3 Spec sec-1 (Hz); in-source CID energy: 0 eV; hexapole RF: 120 Vpp; collision gas: nitrogen; pre pulse storage 7 µs; MSMS collision 100/100 (timing 50/50); capillary voltage 4500 V; scan from 50 to 1600 m/z; endplate 1072 nA; endplate offset: -500 V; Funnel 1 RF: 200 Vpp; Funnel 2 RF: 220 Vpp; quadrupole ion energy 4 eV; collision energy 10 eV; collision RF 160/350 Vpp (timing 70/30); transfer time 90 µs. The measurement of each batch was initiated by a run in sequence using three runs of acetonitrile, one run of a blank, QC_grass, and QC_herb, followed by three runs of the QC_pool to equilibrate the column. Within a batch, we randomised the run order of the samples and measured the QC_pool after each 11th sample again. The measurement of each batch was finalised with one run of QC_pool, QC_herb, QC_grass, and blank. All samples were measured in 12 batches, splitting them by their assigned lab-IDs after each 11th sample per season. Therefore, batch "pos01", for example, included the samples 001 to 011 of 2018_E, 2018_F, 2018_G, and 2018_H, and "pos02" samples 12 to 22 of 2018_E, 2018_F, 2018_G, and 2018_H, et cetera. The thirteen MSMS species pools were measured in a separate batch using the following changed or additional parameters: endplate 1122 nA, number of precursors 3, mass tolerance 0.20 m/z, CID energy for MS/MS 5.0 eV, collision energy for MS/MS 70.0 eV.

LC-MS raw data preparation

Pre-processing. We converted the obtained Bruker files to the open data format '.mzML' using CompassXport (Bruker, version 3.0.9, http://www.bruker.com). The raw LC-MS data were pre-processed in Galaxy-W4M (based on XCMS 3.0) using the workflow provided by Marr et al. 2021: https://doi.workflow4metabolomics.org/W4M00008. The tool versions and parameters used to analyse this study's data set are listed in Suppl. Table 2. Peaks in the raw spectra are detected and grouped. In the second step, the peaks are corrected for retention time shifts. After a second grouping, isotopes and adducts are annotated using the CAMERA tool (Kuhl et al. 2012). The region of interest was filtered to 80 and 880 seconds.

Validity Check. To ensure the quality of the data, we applied different validity checks on both the LC-MS features and the samples. Following the example by Marr et al. (2021), in this study, we used batch correction to remove equipment-related variation from the measured intensities of each feature. These signal drifts are likely introduced to the mass spectrometer over time and cause slight changes in the measured intensities. To correct for signal drops within and across batches, we used the multiple measurements of the QC_pool sample, assuming that ideally, the measured intensities for each feature within this sample would rather be the same.

We calculated and subtracted the variation between the QC_pool measurements from all features. To ensure the quality of each feature and that they were not contaminations but derived from a biological sample, we removed all features detected in the blanks and excluded them from further analyses. In this study, we tested the metabolomic fingerprints of each sample for their similarity to the fingerprints of their species to ensure that the collected plant samples were not contaminated. Applying the proposed method by Marr et al. (2021), we calculated Mahalanobis distances for the samples within each species. We used these distances to flag outliers with a distance three times the average distance to the respective species. These samples were then tested for the number of features that they shared with the species. Flagged samples that shared less than 25% of their features with their species were excluded from further analysis. Resulting in 6 samples being excluded from further analysis (*Anthoxanthum odoratum*: 2018_F_ANTODO_A044_a, 2018_H_ANTODO_A018_a; *Phleum pratensis*: 2018_F_PHLPRA_B073_a, 2018_G_PHLPRA_B073_a; *Poa pratensis*: 2018_F_POAPRA_B073_b; *Ranunculus acris*: 2018_G_RANACR_A018_b). The complete code used for the data validation is available in the MetaboLights repository MTBLS1224 as "SciData_ProcessingTutorial_Rscript.R".

Feature diversity and selection

Feature indices. We used the R package *vegan* (version 2.5-7, Oksanen et al. 2013) to calculate the feature diversity indices richness (fRich), Shannon (fShan), and Evenness (fEven) for each sample (504 samples used for this analysis). The feature richness was calculated by counting the number of features with measured intensities. The feature Shannon diversity was calculated based on the Shannon index (Shannon 1948) using the feature identity as species and the measured intensity as abundance. We used the richness and the Shannon index to calculate the feature Evenness. This index accounts for the abundance of features – ranging from 0 to 1, with 1 accounting for the equal abundance of all detected features within a sample. Therefore, a low Evenness accounts for only a few highly abundant features with an increased number of low-intensity features.

Feature selection. Variance partitioning Prior to the feature selection, we used variance partitioning to identify the experimental factors that explain most of the variance in the feature's intensities. To calculate the variance partitioning with the *AnDec* function of the *StatTools* package (version 0.0.916, Brunius 2022), we used the prepared data matrix, filled missing values with small random numbers and log-transformed the matrix (compare Missing Data Imputation in Marr et al. 2021).

PLS pre-selection In this study, we used partial least square (PLS) models with the *pls* package (version 2.7-3, Mevik 2020) used by the *caret* package (version 6.0-88, Kuhn 2015) to select relevant features within each species. The PLS was performed for each species separately for the factors: species richness, species Shannon (Shan_total), plant height, plant BBCH, and mechanical- and pathogen-inflicted damage. The selected features were assigned a putative metabolite family ID (pMetFam_ID) that encodes the level of identification (Table 2, Suppl. Table 3).

Feature annotation. To putatively annotate the measured compound families, we used the MetFamily web service (Treutler et al. 2016). To use the MetFamily web service, we converted the raw LC-MS/MS data of QC_pool and the MSMS species pools using the ABF converter (AbfConverter_1.3.7815, Reifycs Inc.) for Bruker files (CompassXtract_3.2.201_64_bit, Bruker Corporation). We pre-processed them with MS-Dial (MSDIAL ver 4.70, Tsugawa et al. 2015). The parameter used for the processing were set as follows: MS1 tolerance 0.01 Da, MS2 tolerance 0.05 Da, retention time 1 to 15 min, MS1 mass range 100 - 1500 Da, min peak height 2000 amplitude, mass slice width 0.05 Da, min peak width 4 scans, including all adduct species suggested by MetFamily. We used the libraries linked to MetFamily: Leibniz_positive_mode_20170801_2021-09-02, lib_spectra_pos_new_2021-09-02, MoNA-export-LC-MS-MS_Spectra_2021-09-02, MSMS-GNPS-Curated-Pos_2021-09-02, and MSMS-Respect-Curated-Pos_2021-09-02 (Suppl. Table 3). We only used these features for the annotation that statistically showed reliability ($p<=0.05$). We used the annotations provided by MetFamily to annotate the selected features (see Feature selection).

Linear Model Calculation

For calculating linear models, we used the recorded plant species richness (pRich) and plant Shannon diversity (pShan) as predictor variables (preV). The metabolic fingerprint richness and composition indices (fRich, fShan, fEven) were used as response variables (resV). Linear models were calculated for each species, including all samples per species across all four seasons. Models calculated across all species included all 504 samples (all). We calculated the models as simple linear regression with one preV as fixed effect per resV, resulting in 84

models (Suppl. Table 4). To compare the linear models and evaluate their reliability and predictive power, we calculated the coefficient of determination (R^2), the root-mean-square error (RMSE), confidence intervals (CI) and p-values as reference. The models were calculated with the *stats package* (R version 4.1.1 2021-08-10, R Core Team 2013) as used by the *caret* package (version 6.0-88, Kuhn 2015). Additionally, we calculated the p-values with the package *lmerTest* (version 3.1-3, Kuznetsova et al. 2017). All R packages and version information are available in Suppl Table 6.

Results

Environmental traits on plot level In this study, we focussed on the dynamics in metabolite fingerprints of grassland plant species as a response to changing plant neighbourhoods across seasonal shifts. We compared the species richness designed for the TBE plots with the recorded richness (pRich) in the plots prior to sampling (Figure 01). We found that the species diversity in the plots did not represent the initially planted biodiversity (Suppl Table 4). In general, species richness ranged from 5 to 27, with an average of 13 (Supple Table 4, Figure 01). In plots with the target species belonging to FG grass, pRich ranged from 5 to 24, and in plots with target species FG herb, from 5 to 27. On average, observed species richness ranged from 13 (DL2 and DL4) to 14 (DL1 and DL8) across the year. We recorded 15 (Block A), 13 (Block B) and 12 (Block C) species on average across the seasons in the TBE plots we sampled our target species in (Supple Figure 1). We also found the most differences in plant species richness across the season and not across the plots. As averages per season, we recorded 16 to 18 species for spring, 7 to 11 species in early summer, 12 to 15 species in late summer and 14 to 15 species on average in autumn (Figure 1). We recorded the plant species Shannon index (pShan) of the surrounding plant community for each target plant. In this study, pShan ranged from 0.43 to 2.64 for FG grass plots, with an average of 1.88 and from 0.77 to 2.59, with an average of 2 in FG herb plots. Across the season, the Shannon diversity (pShan) followed the pattern as the species richness within the Blocks of the TBE. In Block A, the Shannon diversity was recorded with an average of 1.93, in Block B with 1.89 and in C with 1.73. In each Block, we found pShan to be the lowest in early summer (season code F), followed by late summer (season code G). The highest pShan values were recorded in autumn (season code H) and spring (season code E). Across all Blocks, pShan for seasons, in general, was found to be 2.06 in spring (E), 1.58 in early summer (F), 1.81 in late summer (G) and 2.03 in autumn (H). We use the recorded plant species exclusively for linear models and observation interpretation to investigate the relationship between plant species richness and diversity in different neighbourhoods on the metabolomic fingerprint of the thirteen different species. Because we found the species richness to be dependent on the season we sampled in, we treated each sample as independent replicated and independent of the plot designed diversity, plot location in the TBE and season.

Environmental Traits on plant level In order to investigate the dynamics in the plant metabolite profiles, we also recorded environmental traits on the plant level. We measured height, recorded the developmental stage (according to the BBCH scale by Hess et al. 1997) and estimated mechanical- and pathogen-inflicted damage on each plant. Measurements for each sample are available in Suppl. Table 4. In our study, we recorded heights from 6.9 cm to 102.1 cm with an average of 29.9 cm (Table 1). The general average across all seasons was found to be 31.7cm in FG grass species and 27.8 in FG herb. In FG grass, the plant heights ranged from 57.9 to 102 cm, in spring, from 10.0 to 23 cm, in early summer, from 9.1 to 35.5 cm in late summer and from 9.6 to 21.5 cm in autumn. For FG herb, the individual plant heights were recorded from 23.5 to 76.3 cm in E, 8.5 to 38.8 cm in F, 11 to 38 cm in G and 6.9 to 25.3 cm in H. The collected samples' plant developmental stages (BBCH scale) ranged from 30 to 85 in FG grass and 25 to 85 in FG herb (Table 1). In this study, we found the highest mechanical damage on leaves and stems averaged across *Plantago lanceolata* samples (15.3 %). Species of FG grass showed an average damage of 3.2 %, while the species in FG herb were recorded with 4.6 % on average. Across the seasons, we found the lowest mechanical damage for FG grass in E (0.4%), followed by F (4.1%), G (4.0%) and H (4.1%). While in FG herb, the lowest mechanical damage was found in G (3.6%), followed by E (3.8%), F (4.9%) and H (6.0%). Pathogen-inflicted damage, including signs of senescence, was found with up to 22.5% (Table 1). In FG grass species, the lowest average pathogen damage was recorded in H (7.7%), followed by E (10.4%), F (15%) and G (18.4%). In the species of FG herb, we recorded the lowest damage in F (8.7%) and H (9.8%), with higher rates in E (12.6%) and G (13.8%).

Metabolite richness and composition In this study, we measured the metabolite richness as feature richness (fRich), i.e. the number of recorded metabolite fragments and the metabolite composition, by taking the abundance of each feature into account, as Shannon diversity index of features (fShan) and feature Evenness (fEven). Across all species and seasons, we found fRich to range from 321 to 642 in FG grass species and from 452 to 800 in FG herb species (Suppl Table 4). On average, we found in *Festuca rubra* (347) and *Poa pratensis* (360) the lowest fRich across the seasons, while *Anthoxanthum odoratum* (579) and *Avenula pubescens* (532)

had the highest fRich in FG grass (Table 1a). In FG herb, we found the lowest fRich in *Ranunculus acris* (465) and *Geranium pratensis* (484) and the highest in *Knautia arvensis* (728) and *Plantago lanceolata* (738) on average (Table 1b). FG grass samples collected in autumn (H) and spring (E) showed the lowest fRich, while FG herb species showed the lowest when collected in autumn and early summer. Furthermore, when collected in late summer, most FG herb species showed the highest fRich on average. We found fRich to be highest on average across all seasons in the plots of Block A, followed by B and C (Suppl Table 4). fShan ranged from 0.11 to 5.52 in FG grass and 0.69 to 5.89 in FG herb (Suppl Table 4). On average for FG grass species, fShan was lowest in Poa pratensis (1) and *Festuca rubra* (1) and highest in *Anthoxanthum odoratum* (4) and *Avenula pubescens* (5) (Table 1 a). In FG herb, fShan was found to be lowest in *Ranunculus acris* (1) and similarly high in *Leucanthemum vulgare* (5), *Knautia arvensis* (5), *Geranium pratensis* (5), *Centaurea jacea* (5) and *Plantago lanceolata* (6) (Tabel 1b). Across the seasons, fShan was found to follow the same trend in FG grass species, with low values on average in spring (E), and high values in late summer (G) and autumn (H). For species of FG herb, we found the lowest fShan values for most species in early summer (F) and autumn (H), while in late summer (G), higher values were recorded (Table 1 b). Generally, slightly higher fShan values were measured in plants collected in Block C, compared to plots in Blocks A and B (Supple Table 4). Across the seasons and species, we found fEven to range from 0.02 to 0.87 in FG grass and from 0.11 to 0.88 (Suppl Table 4). In FG grass species, *Festuca rubra* (0.10) and *Poa pratensis* (0.13) were found to have the lowest fEven values. In contrast, *Anthoxanthum odoratum* (0.60) and *Avenula pubescens* (0.78) had the highest values on average (Table 1 a). However, only Ranunculus acris (0.20) in FG herb showed comparably low values. In contrast, *Leucanthemum vulgare* (0.80), *Knautia arvensis* (0.80), *Geranium pratensis* (0.80), *Centaurea jacea* (0.80) and *Plantago lanceolata* (0.83) all showed a similar fEven value (Table 1 b). However, the change in Evenness was only minor across the seasons within the FGs. For both FG, we found the general trend in fEven with low values in spring (E), followed by early summer (F) and high values in late summer (G) and autumn (H) (Suppl Table 4), across the blocks of the TBE. The fEven values were found to be similar, with samples collected from Block C having a slightly higher fEven across the seasons (Supple Table 4).

Metabolite profiles dynamics We accessed the impact of the plot (season, plot diversity as either desDiv or pRich and pShan) and plant level traits (height, BBCH, mechDam, pathDam) on the metabolite profile dynamics via variance partitioning Figure 2). When calculated across all samples using the designed plot diversity of the TBE, we found that about 66% of the dynamics of the fingerprint were explained by species identity (65%) and plot level factors (1%). Plant-level traits added about 8%, leaving 25% as unexplained residuals (Figure 2 a). However, when using pRich and pShan in the calculation rather than the designed diversity, 80% are explained by species identity (65%) and plot level traits (15%), with 5% added through plant level traits and 15% residuals (Figure 2 c). When focusing on the plant identity and calculating the variance partitioning per species, we found that, on average, about 33% of the profile dynamics are explained by plot level traits (Figure 2 b) when using designed diversity as a plot trait. Plant level traits were found to account for an additional 51%, resulting in 15% unexplained residuals. On the contrary, when taking pRich and pShan into account as plot level traits, 62% of the dynamics of the profiles could be explained on average. In comparison, plant level traits were found to account fan or an additional 29% of the variation of the metabolomic fingerprint, with only 9% remaining residuals (Figure 2 d).

Metabolite family annotation In this study, we used the measured LC-MS features to putatively annotate metabolite families (pMetFam) found in the thirteen target plant species (Figure 03, Table 2, Suppl. Table 3a and 3b). We used PLS models for pre-selecting features that showed significant relevance for plot or plant level traits in at least one species. The selection was performed with a threshold, so not being selected does not mean that the feature is not relevant in the species, but it means that for the tested factors, it was not selected with a high enough variable importance. In total, we were able to annotate 305 compounds within 8 metabolite families (Table 2). We putatively annotated compounds in the metabolite families Organic nitrogen compounds (pMetFam01), Organopnictogen compounds (pMetFam02), Benzenoids (pMetFam03), Organic acids and derivatives (pMetFam04), Lipids and lipid-like molecules (pMetFam05), Organoheterocyclic compounds (pMetFam06), Organic oxygen compounds (pMetFam07), and Phenylpropanoids and polyketides (pMetFam08). For example, for pMetFam_05, we found reliable annotations to the fourth and for Benzenoids (pMetFam_03) to the third level. The features for compound family pMetFam_02 were only annotated on the first level (Figure 03, Table 2, Suppl. Table 3a). A detailed list of features, their specifications (retention time, exact mass, etc.), and annotations are available in Supple. Table 3a. For each selected feature, we counted the number of samples (QCs and species pools not included) that had a measurement, even though the feature was only selected as relevant in one species-trait combination (feature count). For example, 20 selected features were annotated as Coumaric acid esters (pMetFam_08-F2.2a), adding up to a total count of 1703 (Table 2). The measured intensities for the selected feature are available in Supple. Table 3b. The lowest count was found for pMetFam01 (114), with only

two features annotated within this metabolite family (Figure 3). In pMetFam02, pMetFam03, pMetFam04 and pMetFam05 were also annotated for less than 10 features each. The highest count was found in pMetFam08 (11797) for 118 annotated features, followed by pMetFam07 and pMetFam06 with a count of 3411 in 37 features and 2034 in 23 features, respectively. When looking at the species, we found the most selected and annotated features in samples of PA (3813) for FG grass and KA (2712) for FG herb. In the herb species RA (513) and grass species HL (710), and AO (843), we found the lowest feature counts. The FG herb species PL (1283) and CJ (1839) had also a smaller feature count. The species AP (2138), PP (2505), DG (2617), and FR (3049) in FG grass and GP (2409) and LV (2459) in FG herb shared a relatively equal feature count.

Linear correlations of plot and feature traits To further investigate the relationship of plot traits with the dynamics in the metabolomic fingerprints, we used simple linear models with one feature trait as the response variable (resV) and one plot level trait as the predictor variable (preV). Models were calculated for each species separately and as a full model across all 504 samples using fRich as a response to pRich or pShan, fShan as the response to pRich or pShan and fEven as a response to pRich and pShan (Table 3, Figure 4). A detailed list of all model results is available in Suppl. Table 5. All linear models calculated across all thirteen species (504 samples) were found to be negatively correlated, with R^2 ranging only from 0.25 to 0.28 (Table 3). For the linear models that used pRich as the predictor variable, p values were found as slightly significant (0.03 to 0.05), while the models using pShan only had p values > 0.1. For the linear models using pRich as preV and fRich as resV, we found in FG grass species PP (R^2 = 0.19; p = 0.82) and FR (R^2 = 0.46; p = 0.63) to be positively correlated (Figure 4 a). Only HL and PA generally showed higher R^2 values, with 0.64 and 0.65, respectively. Significant p values were only found for AO (p = 0.03) and HL (p = 0.01; Table 3). For species of FG herb, we found three species with a positive and three with a negative correlation. Only CJ was found to have a slightly higher R^2 value at 0.57 (p = 0.15). Significant p values were only found for LV (0.01) and PL (0.01) but with low R^2 values between 0.22 and 0.23 (Table 3). We only found FR positively correlated for models with resV fShan and preV pRich (R^2 = 0.74, p = 0.02). PA showed a higher R^2 value (0.79). Most species of FG herb showed a positive correlation, while only PL (R^2 = 0.25, p = 0.08) and GP (R^2 = 0.48, p = 0.26) were negatively correlated. Here, only CJ was found with a higher R^2 value of 0.69 (p = 0.11). Furthermore, the models using preV pRich and resV fEven were found with a positive correlation in FR (R^2 = 0.75, p = 0.02) of FG grass, and LV (R^2 = 0.28, p = 0.02), KA (R^2 = 0.42, p = 0.58), CJ (R^2 = 0.46, p = 0.24) and RA (R^2 = 0.56, p = 0.24) for FG herb. The highest R^2 values were found for FR and PA (R^2 = 0.78, p = 0.17). Linear models calculated with pShan as preV and fRich as resV were found to be negatively correlated in all species of FG grass, but PP (R^2 = 0.16, p = 1). The highest R^2 was found for HL at 0.51 (p = 0.02) and DG at 0.52 (p = 0.02). In FG herb, the two species RA (R^2 = 0.52, p = 0.2) and CJ (R^2 = 0.55, p = 0.3) were found with a positive correlation of the factors. The highest R^2 was found for KA (R^2 = 0.56). Models that used fShan as resV and pShan as preV were negatively correlated in all FG grass species, with R^2 ranging from 0.22 to 0.68. In FG herb, the only models for RA (R^2 = 0.43, p = 0.69) and CJ (p = 0.23) showed a positive correlation for the factors. The highest R^2 was found in CJ with 0.64. For the models using fEven as resV and pShan as preV, all FG grass species were also found with a negative correlation and R^2 values ranging from 0.21 to 0.68. The highest R^2 values were found for AP (0.62) and DG (0.68) models. Moreover, we found three species of FG herb with positive and three with negative correlations. The highest R^2 value, here, was found in GP (0.68, p = 0.05).

Discussion

In this study, our three main questions focused on 1) how plant species richness (pRich) and 2) species diversity (pShan) are reflected in metabolomic profiles of thirteen grassland plant species in terms of richness (fRich) and composition (fShan), and evenness (fEven) of their LC-MS measured features, and 3) if we can find similarities in these profiles changes in species belonging to the same functional groups (FG). Although the TBE, as part of the Jena Experiment, provided plots with different levels of plant species diversity, composed of different species pools that included 4 grass and 4 herb species each, the underlying concept was unsuitable for our study. As sampling had to take place multiple times throughout the year and, therefore, could not be timed according to the maintenance procedures happening twice a year in these plots, the plant species composition across the full growing season that we focussed on for our metabolite analysis was not effectively reflected by the plant community present in the plots. Therefore, we separately recorded plant species richness (pRich) and diversity (pShan) for each plot on each sampling occasion.

Previous studies were able to show that seasonal dynamics directly (temperature, water and light availability) and indirectly (plant-plant and plant-herbivore) influence the production of metabolites and, in the dynamic defence profile (Weston et al. 2013], Berini et al. 2018). In this study, we define the term “season” as the interplay of weather dynamics (light, water, nutrient availability), growth and reproduction cycles of plants, and semi-maintenance occurring twice within the growing season. In our study, the snapshots of the metabolomic profiles

across the year do not comprehensively reflect the dynamics within the profiles but rather give an idea of the overall picture. Moreover, the semi-maintenance, i.e. mowing and weeding, in all TBE plots, included increased mechanical damage to the plants and general disturbance to the plot and soil structure. Depending on the set species composition in each plot, weeding was accomplished by taking out the majority of plants, leaving only a few individuals of the sown species in the plot with large patches of cleared bare ground. In contrast, in other plots, mainly grass species plots, only a few weed plants were removed, leaving those plots nearly undisturbed. Since we assume that all those external changes also directly influence the metabolite profiles by increasing stress and, thereby, triggering an increased production of defence compounds, analysing the thirteen target species in a single correlation model would have prevented us from drawing reliable conclusions. As shown in the results, integrating all species into a single correlation model would overshadow the correlations in single species with strong correlations in only a few species. Furthermore, we only used the plot level traits, pRich and pShan, for our correlation models. We assume that the recorded traits for the individual plant samples are both a direct and indirect result of the community dynamics and are, therefore, already reflected in the plot level traits.

Throughout the literature, the term metabolomic profile and profiling are often used to refer to the entirety of all, e.g. plant, produced primary and secondary metabolites (Weston et al. 2015) or as a synonym to targeted metabolomics (Fiehn 2022). However, here, we use the term *metabolomic profile* to refer to all LC-MS measured features deriving from mainly secondary metabolites, regardless of their origin (plant, microorganisms, etc.), as the interaction between plants and microorganisms is seamlessly intertwined. To be able to confirm the specific function of compounds, their identification is often a top priority in metabolomics studies (Fang et al. 2019); however, in this study, we focus instead on the big picture, presented as the profile of measured features, rather than the identification of single compounds. Many studies report that environmental changes are reflected in the secondary metabolite profile (e.g. Berini et al. 2018), leading to the assumption that all or at least major parts of the profiles are responding to environmental changes in a brief time frame; however, in our study, we were able to show that only parts of the secondary metabolite profile is directly representing the changes induced by biotic and abiotic factors. Already mentioned by Walker et al. (2022), metabolomic profiles are most likely composed of two sets of metabolites, including metabolites produced throughout the life cycle of a plant, and metabolites, which abundance is induced by environmental changes and pressures. In our study, we found that species identity accounted for 65% of the variation in the metabolomic profiles, leading to the assumption that all species maintain several metabolites, mainly representing the species identity. We identified these parts as the metabolomic fingerprint of a plant accounting for its biogeographic legacy. At the same time, the remaining metabolite represented the dynamic defence profile influenced by environmental changes and seasonal dynamics. These fingerprints also seemed to be mainly independent of direct plant features such as developmental stage, height and level of damage. We could show those two types of secondary metabolites for all thirteen species. All species showed another set of metabolites that represents changes caused by environmental dynamics.

Furthermore, we found some of the annotated metabolite families to play an important role in the defence mechanisms of some of the thirteen target species. In contrast, they are not found to undergo considerable changes in other species or haven't been measured. This finding also contributed to the decision to analyse the relationships of plant species richness and composition per species and not in a global model. As discussed by Zanne et al. (2014) and Díaz et al. (2016), plants commonly have a set of available strategies to maximise their fitness, which should make it possible to predict certain responses to environmental changes. This would support the assumption that specific metabolites are engaged in responses to specific environmental pressures, for instance, Jasmonic acid for herbivory pressure (Wasternack & Hause 2013) and p-coumaroylagmatine for UV stress (Kusano et al. 2011). However, the responses may be broader and less specific for similar pressure that is less specific, such as weather conditions. In the TBE Jena Experiment, different seasons come with different combinations of species composition in various abundance levels. The different neighbourhoods, abundances and diversity levels influence the performance, appearance and, hence, the metabolomic fingerprints of each target plant species resulting in altered metabolite richness and composition. Since the species richness present in a plot is a direct result of the time point defining the sampling season, including all environmental factors but also maintenance-related factors, the effects and correlations we find may not represent a true image of natural diversity settings.

We generally found plant species richness to be mostly season dependent, with relatively equal numbers for FG grass and FG herb target species. However, species richness was not found according to the general expectation of finding an increasing number of species over the year. Still, it mainly depended on the semi-maintenance in the TBE plots. In spring, we found the highest number of species, as this sampling was done before the first weeding and mowing action. The species had about 8 to 9 months to recover, undergo the winter

season and grow back in a natural composition with little to no disturbance, allowing many species to grow big enough for identification. On the other hand, the lowest species richness was found in early summer. Here the sampling was conducted only a few weeks after mowing and weeding, leaving the plants mostly small-statured and some plots with large bare patches of soil. For the diversity assessed by the Shannon index, we found the most range in early and late summer, while the smallest range was recorded in spring and autumn. Across the FGs, the diversity levels were mostly stable, leading to the assumption that the changes are mainly related to the initial species composition. Using species richness as an independent factor regardless of the season allowed us to investigate the influence of plant species richness and composition on LC-MS features' richness and composition.

Plant species richness (pRich) was mainly negatively correlated with featuring richness (fRich). Here we found the correlations according to our assumption that higher plant species richness would result in lower counts of features per sample confirmed. Assuming that a high species number and hence high diversity would result in less pressure, such as less herbivory, better shading, better nutrient availability, and less production of defence metabolites. For the model using plant species Shannon (pShan) index to predict feature richness (fRich), we also found a negative correlation for most species in FG herb and grass with moderately reliable R^2 values. This seems to confirm the assumption that in a neighbourhood with more species that are equally distributed, the stress levels are reduced and metabolite production can be reduced to a necessary minimum. However, some species also showed a positive correlation, which supports the assumption that the responses to environmental changes can be highly species specific. For fShan and pRich, we would have expected to find lower fShan values in communities with a lower species number (small pRich), as this would account for high pressure to abiotic and biotic factors are being answered with very specific metabolites that are produced as defence mechanisms. However, in our study, we found most species of FG grass to be negatively correlated, with the highest feature Shannon measured in plots with a high species richness. On the other hand, plant species of FG herb all showed a slightly positive correlation, underpinning the initial assumption. When calculated across all thirteen species, the trend clearly is negative, although only grass species showed this tendency in our study. According to the theory that plant species diversity acts as a protective shield within the plant community and that plants produce highly specialised defence compounds in higher amounts depending on the pressure, we would expect to see a positive correlation between feature Shannon (fShan) and plant community composition (pShan). However, in this study, we found only slight correlations but hardly any clear direction for FG herb, while grass species were mostly negatively correlated. The overall model reliability was very low when calculated across all species. When testing the correlation between plant species richness (pRich) and the feature evenness (fEven), we found that the FG herb species showed a slightly positive correlation which would be according to the expectation that with higher species richness, we have a more evenly distributed metabolite profile. In comparison, with less richness, we would see more metabolites in overproduction and, hence, less feature evenness.

However, FG grass species showed a clear negative correlation between plant species richness and feature evenness. For models with fEven and pShan, we see a similar picture. FG herb showed a slightly positive but mostly no correlation between plant diversity and feature evenness, while FG grass species were negatively correlated. We found the changes to the metabolomic profiles for the thirteen tested species to be highly species-dependent. No common change was found in the profiles for all species, but similarities in the closely related species, i.e. FG grass. The species tested here in FG herb are probably phylogenetically too far from one another to show the same changes in their metabolomic profiles. In general, differences between FGs point to different defence strategies.

In this study, we found, on average most residuals in herb species compared to grass species, highlighting the different strategies of incorporating defence mechanisms to survival and reproduction strategies on limited resources in a diverse neighbourhood. But, these species also had the most residuals remaining unexplained, pointing to other influences that weren't accounted for in the models. One explanation for the differences found in FG grass and FG herb species might be related to the fact that species richness is also influencing resource availability, which can have negative effects on plant productivity (Marquard et al. 2009, Scherling et al. 2010) and is thereby also effecting the metabolite profiles (Flynn et al. 2008). The fact that in FG grass species, we found a negative correlation between feature richness and diversity with plant species richness and diversity more often might be due to the fact that grass species in the TBE were more exposed to beneficial plant species mixtures. For example, soil-related effects, such as mycorrhiza, microorganisms, or legumes' beneficial effect. For example, Hooper et al. (2005) argue that legumes present in plant communities provide beneficial effects to plant productivity. While (Forbey and Hunter 2012) point out that herbivores can also strategically use certain secondary metabolites, which would alter the herbivore pressure in these plant communities, beneficial effects of mycorrhiza are also more associated with FG herb species than grass species, which can also affect the production of metabolites as more resources are made available by the fungi (Ristok et al. 2019)

Another explanation points toward the different strategies used by FG herb and FG grass species, including the more mechanical defence structures in grasses, which lead to the overall lower production of metabolites in FG grass species. When investigating the influence of seasons, temperature changes are one of the most apparent influences that can trigger changes to the metabolomic profiles. However, as discussed by Berini et al. (2018) the findings are not consistent across the studies. Interpreting the findings can be very challenging as even small changes in environmental conditions can dramatically affect metabolite profiles. Moreover, changes in plant species richness and plant species Shannon are not only related to the season where weather conditions change and developmental stages are different but also depend on the semi-maintenance of the plots in the TBE. Therefore, our finding that low pRich and low pShan were correlated with higher metabolite richness can also be a result of the semi-maintenance procedures being applied in the plots of the TBE Jena Experiment, where plots of FG grass species are mainly undisturbed in contrast to the highly disturbed FG herb plots across the year, resulting in sampling timepoints with high response to disturbance in cleaned and cut plants (less species abundance, and richness). Plants that are under pressure from plot disturbance and resource collection to grow back and reproduce might be more prone to biotic and abiotic factors due to free spaces in the plot (less dense community), less protection from herbivores, less beneficial effects of legumes (get weeded out), mechanical damage from mowing. As shown by (Ristok et al. 2022), metabolite diversity as a response to resource availability was found to vary in FG grass and FG herb. This might be related to resource availability, which can be more limited in communities with a higher number of species as the resource pools, such as light, water and nutrient availability are more dependent than in communities with fewer species. As discussed by (Berini et al. 2018), changes in secondary metabolite profiles might also have beneficial effects on the plant community by contributing to defence strategies against various biotic and abiotic pressures. As discussed by (Walker et al. 2022), the different strategies presented by plant species in terms of ecological traits, i.e. fast-growing (acquisitive) and slow-growing (conservative) strategies, will be reflected in their metabolite profiles. Therefore, using the recorded plant height across the season as an indicator of potential changes in the metabolite profile might provide insights into the bio-geo-legacy strategies of different plant species. Scherling et al. (2010) found in their study that plants with shorter growing habitus showed a correlation in their metabolomics profiles as a response to higher plant species richness in contrast to plants with a taller habitus, which seemingly did not show a response to different plant species richness in their metabolomic profiles. The authors concluded that small plant species are less competitive and, hence, more affected by resource limitations imposed upon the species by a higher richness.

In contrast, tall plants, which had reportedly become dominant in the tested communities, were not found to respond to the changes with alterations to their metabolomic profiles. Also, the correlation between biomass production and the metabolomic profile structure showed that larger species with more complex reproductive organs commonly produce more metabolites. Therefore, altering plant metabolite profiles likely depends on individual survival strategies of the plant species, as metabolite production must come to a balance and a trade-off between defence mechanisms and productivity on a limited resource (Díaz et al. 2016). Metabolites are the result of biochemical mechanisms and processes that shape plants' physiology and function and might, therefore, not only be influenced and shaped by the surrounding plant community but actively influence the diversity of the plant community by influencing the attractiveness to herbivores in the community (Richards et al. 2015).

Conclusion

Assuming that plant communities can benefit from higher diversity in multiple ways, including shared production of defence metabolites, in our study, we expected that the tested plant species would regulate their metabolomic profiles according to changes in plant species richness and diversity within their community in similar patterns. Communities with high species richness and evenly distributed abundances, hence, would result in metabolite profiles with higher numbers of features that are evenly distributed. In contrast, on the other hand, in communities with only a few dominant species, we would mainly find higher abundant specialised defence metabolites. Although finding these expectations to some degree resembled in two herb and one grass species, most of the tested species were negatively correlated or showed varying patterns for the tested diversity indices. Our finding indicates that the response to specific plant species compositions is not only FG specific but highly depends on each species individually. While biodiversity might have an inherent beneficial effect on some plant species, at the same time, it can present other species with increased resource pressure or act less beneficial due to changed attractiveness for a broader range of herbivores and other parties in the community. In our study, for FG grass, which only included species from the *Poaceae* family, we found similar responses to plant richness and diversity more often, leading to the assumption that grass species use similar strategies to respond to environmental changes, as they share similar physiology. As secondary metabolites are produced constantly

throughout the lifecycle of a plant, resource limitation and the tradeoff between growth and survival across environmental dynamics is a substantial part of defence strategies. We found that FG herb species produced more metabolites throughout the year than FG grass. These differences likely stem from the different defence strategies of grass and herb species. While grasses invest more resources in physical defence structures such as silica structures (Massey & Hartley 2009), herbs tend to invest more in their complex reproduction organs' chemical protection (Cooke & Leishman 2012). We conclude that the two FG have different coping mechanisms and rely on different strategies for their response to environmental dynamics. FG grass species were generally producing fewer metabolites than FG herb species, leading to the conclusion that FG herb species are investing more in chemical defence. In contrast, FG grass species rely more on their mechanical defence structures.

In our study, we found the distinction between two profile parts within the secondary metabolomic profiles, making up the species specific fingerprints, identifying each species across environmental and lifecycle changes and the dynamic defence profiles, which show strong responses to external dynamics. The proportion of the secondary metabolites used for environmental responses and those that are not is not investigated well and might alter greatly between the different species. However, in this study, we were able to show that the secondary metabolite fingerprint can be a useful tool for building up a database that can be used for species identification. In general, the underlying experimental design plays a significant role when drawing conclusions about global patterns. When investigating pattern similarities for multiple species, the selection of the species included in the study substantially impacts the comparability. Here, we used a diverse selection of herb species that differ vastly from one another in terms of physiological traits. The grass species, all Poaceae, were significantly closer to one another, allowing the drawing of broader conclusions about response patterns to specific environmental changes. However, including all species in a global model can easily lead to wrong conclusions, as many features are only detected in a few samples or are even unique to one of the tested species, making it challenging to draw a reliable conclusion.

Furthermore, the impact of metabolite profile changes also varies on multiple scales, including the plant individuum, species, populations and environment, as they are influenced by above and belowground microbial realms within their functional networks. Changes to the metabolomic profiles can, on the other hand, also alter biotic and abiotic interaction and, thereby, influence the ecosystem as a whole (Berini et al. 2018). Furthermore, the plants we used in our study grow mainly undisturbed throughout the year under natural conditions. The plants have included metabolites produced by endophytes as much as they produced their own physical and chemical defence structures. Studying defence dynamics in field experiments allows for including all natural interactions between the plants and their environment, increasing the reliability of potential pattern findings; however, the number of environmental factors that cannot be controlled for increases drastically. Furthermore, compared with greenhouse studies, where design-unrelated factors are kept to a minimum, measuring artefacts in rapidly changing metabolomic profiles from plants collected in the field can drastically alter the interpretability of the data.

In our study, we found that the reliability of linear models differed greatly between species, depending on the number of features measured for each species. The tested models, therefore, do not necessarily represent a true picture of underlying mechanisms, but they serve to get an idea of the correlations between plant species diversity and metabolite profile changes across the year. In order to be fully able to test the influences of developmental stages and environmental pressures on the defence metabolite production, the experimental design would need to be adapted accordingly. We agree with the valuation by Berini et al. (2018) that due to the different influences of abiotic and biotic factors on the metabolomic profile, predicting concrete responses to novel environmental conditions might be challenging. Although metabolite studies are still at the beginning of reliably contributing to understanding underlying processes, these studies have great potential to examine and predict plant fitness in present species loss and future climate scenarios. Narrowing down the parts of metabolite profiles that are actually involved in environmental response will also facilitate reliable identification of metabolites, and the prediction of defence-facilitating compounds might also contribute to bio-protected plant breeding environments in agricultural contexts. For instance, untargeted metabolomics might allow the investigation of changes in the metabolomic profiles across environmental dynamics without needing time and cost-intensive key compound identification. Based on their fingerprints, plant samples can be identified and grouped into their functional groups. The fingerprint is conserved across seasonal dynamics or changing environmental conditions, facilitating low-cost analysis of plant samples and forming a fingerprint-based database for ecological experiments.

Data availability

The data set, including raw LC-MS measurements, field records and metadata tables, are available in the MetaboLights repository http://identifiers.org/metabolights:MTBLS1224.

References

Berini, J. L., Brockman, S. A., Hegeman, A. D., Reich, P. B., Muthukrishnan, R., Montgomery, R. A., & Forester, J. D. (2018). Combinations of abiotic factors differentially alter production of plant secondary metabolites in five woody plant species in the boreal-temperate transition zone. Frontiers in plant science, 9, 1257.

Brunius, C. (2022). StatTools: Nifty Tools for Random Statistical Tasks. R package version 0.0.916.

Cooke, J., & Leishman, M. R. (2012). Tradeoffs between foliar silicon and carbon-based defences: evidence from vegetation communities of contrasting soil types. Oikos, 121(12), 2052-2060. https://doi.org/10.1111/j.1600-0706.2012.20057.x

Díaz, Kattge, J., Cornelissen, J. H., Wright, I. J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Prentice, I. C., Garnier, E., Bönisch, G., Westoby, M., Poorter, H., Reich, P. B., Moles, A. T., Dickie, J., Gillison, A. N., Zanne, A. E., Chave, J., Wright, S. J., Sheremet'ev, S. N., Jactel, H., Baraloto, C., Cerabolini, B., Pierce, S., Shipley, B., Kirkup, D., Casanoves, F., Joswig, J. S., Günther, A., Falczuk, V., Rüger, N., Mahecha, M. D., Gorné, L. D. (2016) The global spectrum of plant form and function. Nature, 529(7585), 167-171.

Ebeling, A., Pompe, S., Baade, J., Eisenhauer, N., Hillebrand, H., Proulx, R., ... & Weisser, W. W. (2014). A trait-based experimental approach to understand the mechanisms underlying biodiversity–ecosystem functioning relationships. Basic and Applied Ecology, 15(3), 229-240.

Fang, C., Fernie, A. R., & Luo, J. (2019). Exploring the diversity of plant metabolism. Trends in Plant Science, 24(1), 83-98.

Fernandez, C., Monnier, Y., Santonja, M., Gallet, C., Weston, L. A., Prévosto, B., Saunier, A., Baldy, V., & Bousquet-Mélou, A. (2016). The impact of competition and allelopathy on the trade-off between plant defense and growth in two contrasting tree species. Frontiers in Plant Science, 7, 594. https://doi.org/10.3389/fpls.2016.00594

Fiehn, O. (2002). Metabolomics—the link between genotypes and phenotypes. Functional genomics, 155-171.

Flynn, D. F., Schmid, B., He, J. S., Wolfe-Bellin, K. S., & Bazzaz, F. A. (2008). Hierarchical reliability in experimental plant assemblages. Journal of Plant Ecology, 1(1), 59-65.

Forbey, J. S., & Hunter, M. D. (2012). The herbivore's prescription: a pharm-ecological perspective on host-plant use by vertebrate and invertebrate herbivores. The ecology of plant secondary metabolites: From genes to global processes, 78.

Gargallo-Garriga, A., Sardans, J., Granda, V., Llusià, J., Peguero, G., Asensio, D., Ogaya, R., Urbina, I., Van Langenhove, L., Verryckt, L. T., Chave, J., Courtois, E. A., Stahl, C., Grau, O., Klem, K., Urban, O., Janssens, I. A., & Peñuelas, J. (2020). Different 'metabolomic niches' of the highly diverse tree species of the French Guiana rainforests. Scientific Reports, 10(1), 1-10. https://doi.org/10.1038/s4159 8-020-63891-y

Hess, M., Barralis, G., Bleiholder, H., Buhr, L., Eggers, T. H., Hack, H., & Stauss, R. (1997). Use of the extended BBCH scale—general for the descriptions of the growth stages of mono; and dicotyledonous weed species. Weed research, 37(6), 433-441.

Hooper, D. U., Chapin III, F. S., Ewel, J. J., Hector, A., Inchausti, P., Lavorel, S., ... & Wardle, D. A. (2005). Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecological monographs, 75(1), 3-35.

Izsák, J., & Papp, L. (2000). A link between ecological diversity indices and measures of biodiversity. Ecological Modelling, 130(1-3), 151-156.

Kuhl, C., Tautenhahn, R., Böttcher, C., Larson, T. R., & Neumann, S. (2012). CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. Analytical chemistry, 84(1), 283-289. https://doi.org/10.1021/ac202450g

Kuhn, M. (2015). Caret: classification and regression training. Astrophysics Source Code Library, ascl-1505. R package version 6.0-88. https://CRAN.R-project.org/package=caret

Kusano, M., Tohge, T., Fukushima, A., Kobayashi, M., Hayashi, N., Otsuki, H., ... & Fernie, A. R. (2011). Metabolomics reveals comprehensive reprogramming involving two independent metabolic responses of Arabidopsis to UV-B light. The Plant Journal, 67(2), 354-369.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. Journal of statistical software, 82 (13), 1-26. doi: 10.18637/jss.v082.i13

Lorentzen, S., Roscher, C., Schumacher, J., Schulze, E. D., & Schmid, B. (2008). Species richness and identity affect the use of aboveground space in experimental grasslands. Perspectives in Plant Ecology, Evolution and Systematics, 10(2), 73-87.

Marquard, E., Weigelt, A., Roscher, C., Gubsch, M., Lipowsky, A., & Schmid, B. (2009). Positive biodiversity–productivity relationship due to increased plant density. Journal of Ecology, 97(4), 696-704.

Marr, S., Hageman, J. A., Wehrens, R., van Dam, N. M., Bruelheide, H., & Neumann, S. (2021). LC-MS based plant metabolic profiles of thirteen grassland species grown in diverse neighbourhoods. Scientific data, 8(1), 1-12.

Massey, F. P., Hartley, S. E. (2009). Physical defences wear you down: progressive and irreversible impacts of silica on insect herbivores. Journal of Animal Ecology, 78(1), 281-291. https://doi.org/10.1111/j.1365-2656.2008.01472.x

MetaboLights http://identifiers.org/metabolights:MTBLS1224. (2019)

Mevik, B.-H., Wehrens, R. and Hovde Liland, K. (2020). pls: Partial Least Squares and Principal Component Regression. R package version 2.7-3. https://CRAN.R-project.org/package=pls

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B., ... & Oksanen, M. J. (2013). Package 'vegan'. Community ecology package, version, 2(9), 1-295. R package version 2.5-7. https://CRAN.R-project.org/package=vegan

Quinn, J. C., Kessell, A., Weston, L. A. (2014). Secondary plant products causing photosensitization in grazing herbivores: Their structure, activity and regulation. International Journal of Molecular Sciences, 15(1), 1441-1465.

Rasmann, S., & Turlings, T. C. (2016). Root signals that mediate mutualistic interactions in the rhizosphere. Current Opinion in Plant Biology, 32, 62-68. https://doi.org/10.1016/j.pbi.2016.06.017

R Core Team (2013). R: A language and environment for statistical computing. URL https://www.R-project.org/.

Richards, L. A., Dyer, L. A., Forister, M. L., Smilanich, A. M., Dodson, C. D., Leonard, M. D., & Jeffrey, C. S. (2015). Phytochemical diversity drives plant–insect community diversity. Proceedings of the National Academy of Sciences, 112(35), 10973-10978. doi: 10.1073/pnas.1504977112

Ristok, C., Poeschl, Y., Dudenhöffer, J.-H., Ebeling, A., Eisenhauer, N., Vergara, F., Wagg, C., van Dam, N. M., & Weinhold, A. (2019). Plant species richness elicits changes in the metabolome of grassland species via soil biotic legacy. Journal of Ecology, 107(5), 2240–2254. https://doi.org/10.1111/1365-2745.13185

Ristok, C., Weinhold, A., Ciobanu, M., Poeschl, Y., Roscher, C., Vergara, F., Eisenhauer, N., & van Dam, N. M. (2022). Plant diversity effects on herbivory are related to soil biodiversity and plant chemistry. Journal of Ecology, 00, 1– 16. https://doi.org/10.1111/1365-2745.14032

Roscher, C., Schumacher, J., Baade, J., Wilcke, W., Gleixner, G., Weisser, W. W., ... & Schulze, E. D. (2004). The role of biodiversity for element cycling and trophic interactions: an experimental approach in a grassland community. Basic and Applied Ecology, 5(2), 107-121.

Sardans, J., Gargallo-Garriga, A., Urban, O., Klem, K., Walker, T. W. N.,Holub, P., Janssens, I. A., & Peñuelas, J. (2020). Ecometabolomics for a better understanding of plant responses and acclimation to abiotic factors linked to global change. Metabolites, 10(6), 239. https://doi.org/10.3390/metab o1006 0239

Scherling, C., Roscher, C., Giavalisco, P., Schulze, E. D., Weckwerth, W. (2010). Metabolomics unravel contrasting effects of biodiversity on the performance of individual plant species. PLoS One, 5(9), e12569.

Tikunov, Y., Lommen, A., De Vos, C. R., Verhoeven, H. A., Bino, R. J., Hall, R. D., & Bovy, A. G. (2005). A novel approach for nontargeted data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles. Plant physiology, 139(3), 1125-1137. doi: 10.1104/pp.105.068130

Treutler, H., Tsugawa, H., Porzel, A., Gorzolka, K., Tissier, A., Neumann, S., & Balcke, G. U. (2016). Discovering regulated metabolite families in untargeted metabolomics studies. Analytical chemistry, 88(16), 8082-8090.

Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., ... & Arita, M. (2015). MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nature methods, 12(6), 523-526.

Walker, T. W., Alexander, J. M., Allard, P. M., Baines, O., Baldy, V., Bardgett, R. D., ... & Salguero-Gómez, R. (2022). Functional Traits 2.0: The power of the metabolome for ecology. Journal of Ecology, 110(1), 4-20.

Wasternack, C. and Hause, B. (2013) Jasmonates: biosynthe-sis, perception, signaltransduction and action in plant stress response, growth and development. An update to the 2007 review in Annals of Botany. Annals of Botany 111,1021–1058

Weston, P. A., Weston, L. A., & Hildebrand, S. (2012). Environmental impact on biocontrol agents and secondary chemistry of Paterson's curse (Echium plantagineum). In Proceedings of the 18th Australasian Weeds Conference (pp. 203-207).

Weston, P. A., Weston, L. A., & Hildebrand, S. (2013). Metabolic profiling in Echium plantagineum: presence of bioactive pyrrolizidine alkaloids and napthoquinones from accessions across southeastern Australia. Phytochemistry reviews, 12(4), 831-837. doi: 10.1007/s11101- 013 -9306 - 4

Weston, L. A., Skoneczny, D., Weston, P. A., & Weidenhamer, J. D. (2015). Metabolic profiling: An overview—New approaches for the detection and functional analysis of biologically active secondary plant products. J. Allelochem. Interact, 1, 15-27.

Yang, L., Wen, K. S., Ruan, X., Zhao, Y. X., Wei, F., & Wang, Q. (2018). Response of plant secondary metabolites to environmental factors. Molecules, 23(4), 762. doi: 10.3390/molecules2 3040762

Zanne, A. E., Tank, D. C., Cornwell, W. K., Eastman, J. M., Smith, S. A., FitzJohn, R. G., ... & Beaulieu, J. M. (2014). Three keys to the radiation of angiosperms into freezing environments. Nature, 506(7486), 89-92.

Figures

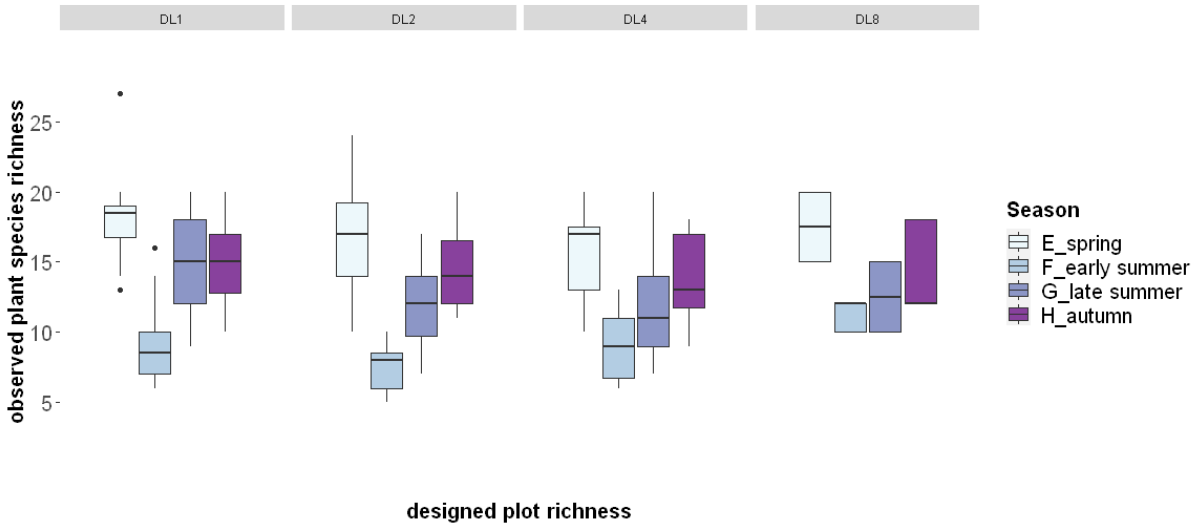


Figure 01: Observed plant species richness in plots of the TBE. TBE plot richness included diversity levels (DL) with one (DL1), two (DL2), four (DL4), and eight (DL8) initial species. Richness was recorded in 48 plots in total (DL1 = 16 plots, DL2 = 16 plots, DL4 = 14 plots, DL8 = 2 plots), representing one set of DL plots for thirteen target plant species, including three species with two sets of DL plots. To cover changes across the growing season, plant species richness was recorded for each plot in spring (E), early summer (F), late summer (G) and autumn (H).

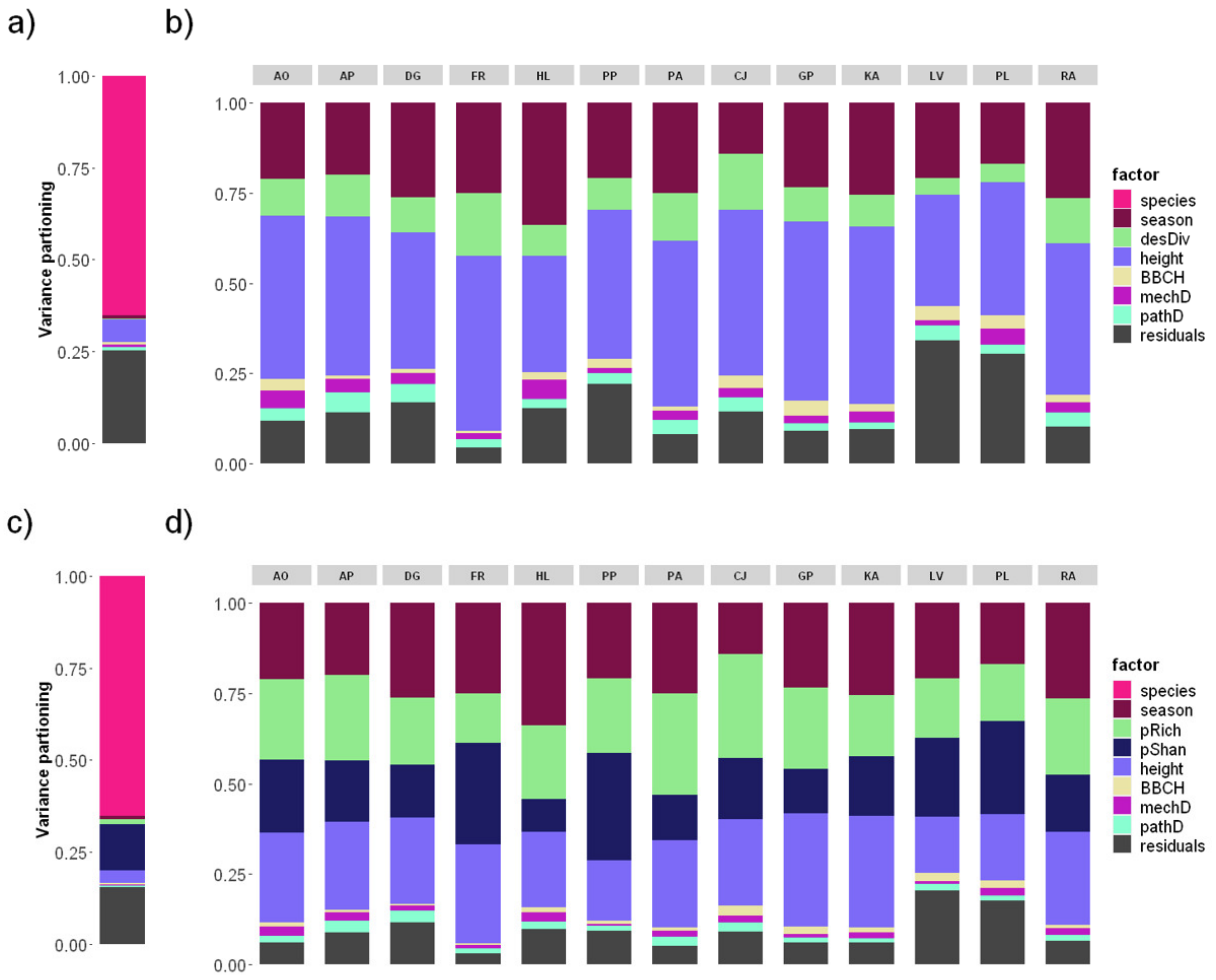


Figure 02: Explained variation in metabolite profiles. Explained variation of feature intensities based on different experimental design factors. Experimental design factors were recorded either at the plot or at the individual plant level. Factors at plot level: time point of sampling (season), number of species per plot as designed (desDiv), number of species recorded per plot (pRich), species diversity recorded per plot (pShan). Factors at the individual plant level: aboveground height (plant height), developmental stage (plant BBCH), inflicted visible damage (mechanical and pathogen damage). The remaining, unexplained variation is shown as residuals. Variance partitioning was calculated a) across all species and b) for each species separately using the designed richness per plot. The variance partitioning using both the recorded plant species richness and species Shannon is shown c) across all species and d) species specific.

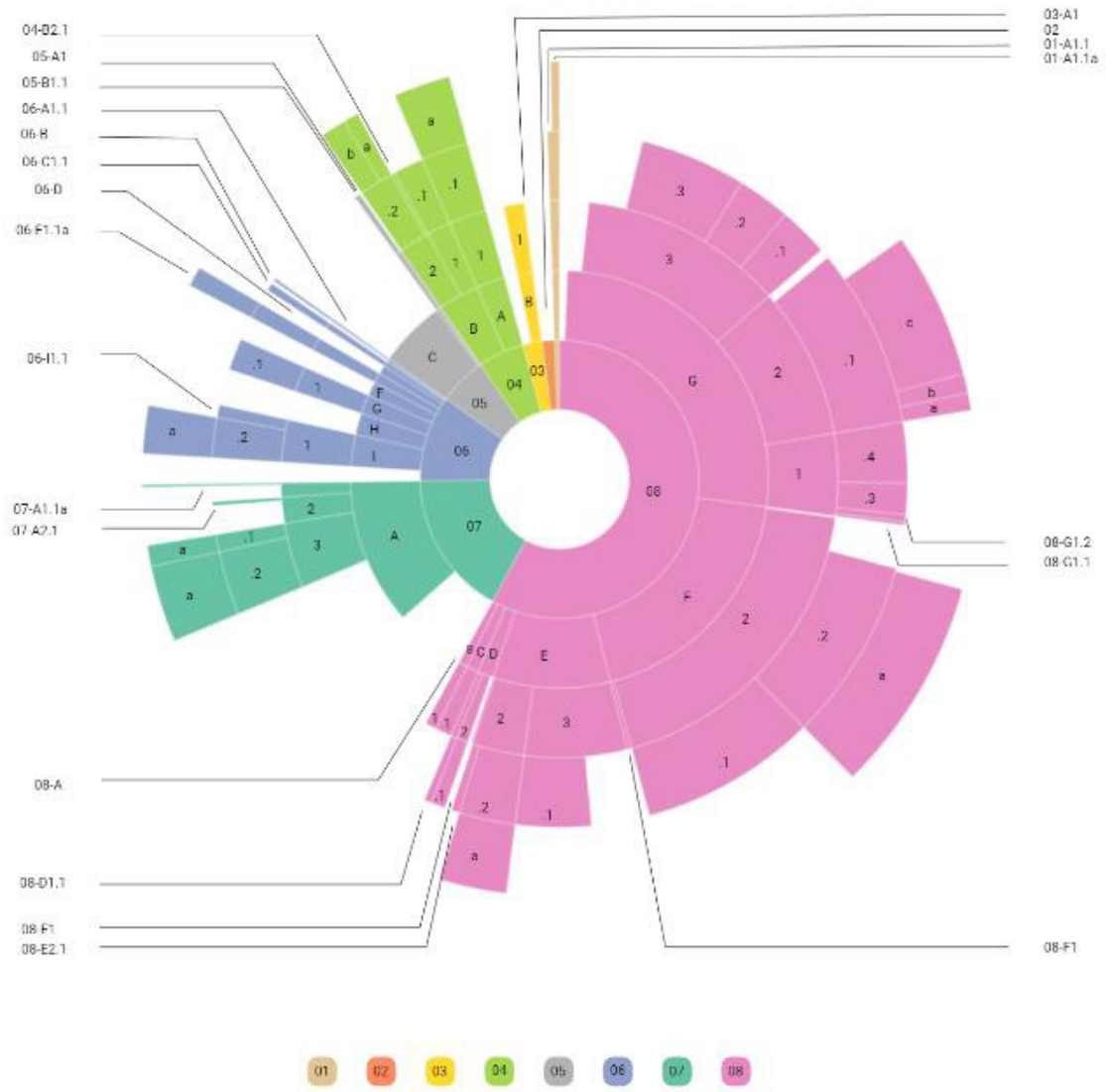
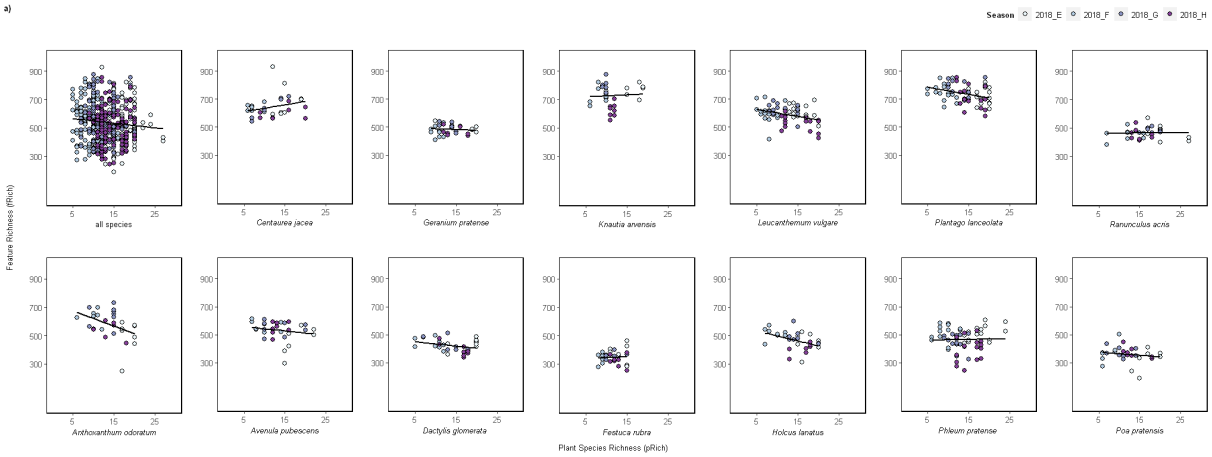
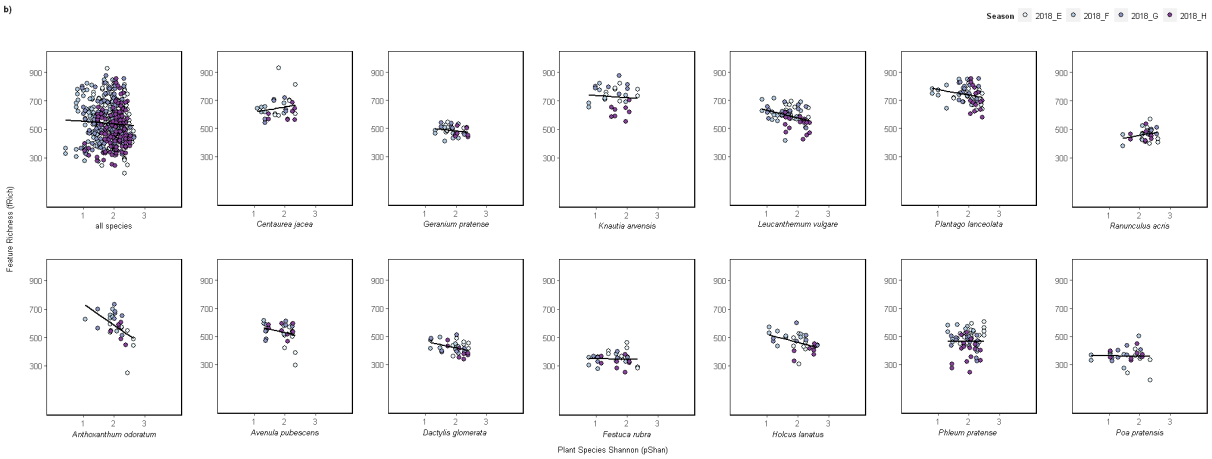
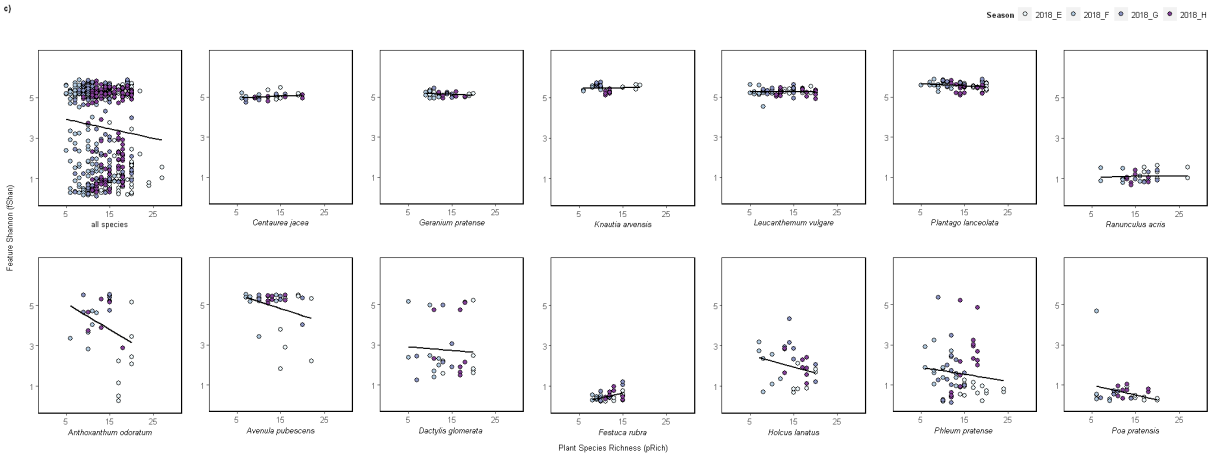


Figure 03: Relative abundance of putatively annotated metabolite families. The features are shown were pre-selected for each experimental factor and separately within each species. The colours refer to the eight annotated kingdoms *Organic nitrogen compounds* (01, gold), *Organopnictogen compounds* (02, orange), *Benzenoids* (03, yellow), *Organic acids and derivatives* (04, light green), *Lipids and lipid-like molecules* (05, grey), *Organoheterocyclic compounds* (06, blue), *Organic oxygen compounds* (07, cyan), and *Phenylpropanoids and polyketides* (08, pink). Please refer to the pMetFam_ID identification table (Table 2) for the subfamilies. The plot was calculated in excel and designed and annotated with BioRender.com.


a)
Season 2018_E 2018_F 2018_G 2018_H
Feature Richness (fRich)
all species
Centaurea jacea
Geranium pratense
Knautia arvensis
Leucanthemum vulgare
Plantago lanceolata
Ranunculus acris
Anthoxanthum odoratum
Avenula pubescens
Dactylis glomerata
Festuca rubra
Holcus lanatus
Phleum pratense
Poa pratensis
Plant Species Richness (pRich)


b)
Season 2018_E 2018_F 2018_G 2018_H
Feature Richness (fRich)
all species
Centaurea jacea
Geranium pratense
Knautia arvensis
Leucanthemum vulgare
Plantago lanceolata
Ranunculus acris
Anthoxanthum odoratum
Avenula pubescens
Dactylis glomerata
Festuca rubra
Holcus lanatus
Phleum pratense
Poa pratensis
Plant Species Shannon (pShan)


c)
Season 2018_E 2018_F 2018_G 2018_H
Feature Shannon (fShan)
all species
Centaurea jacea
Geranium pratense
Knautia arvensis
Leucanthemum vulgare
Plantago lanceolata
Ranunculus acris
Anthoxanthum odoratum
Avenula pubescens
Dactylis glomerata
Festuca rubra
Holcus lanatus
Phleum pratense
Poa pratensis
Plant Species Richness (pRich)

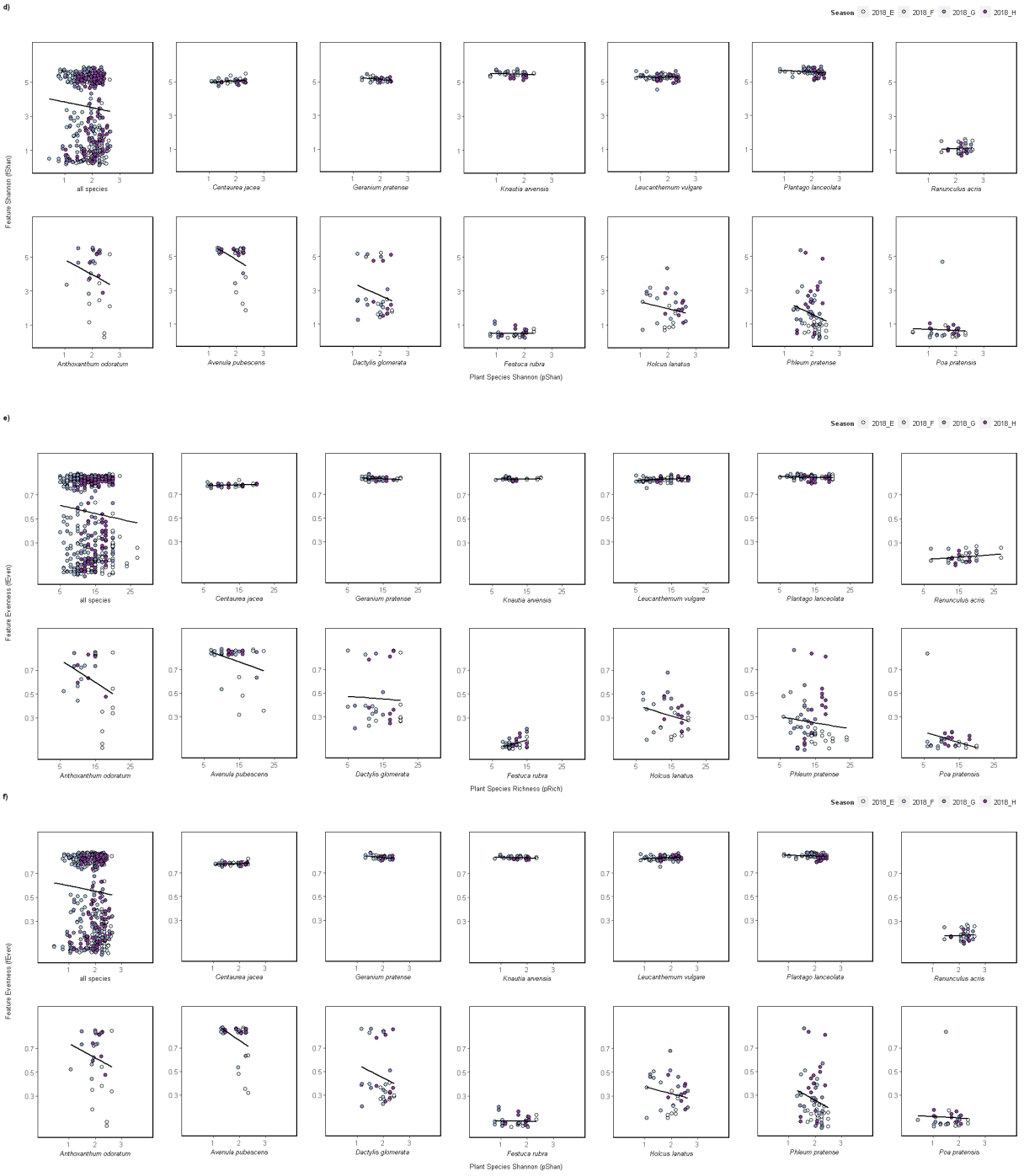


Figure 04: Linear models and correlation of richness and composition indices of metabolomic LC-MS features and plant species. Correlation is shown for each species separately and all species (13 species) combined. Plots show true measurements and the fitted linear model (black line). Values are colour coded by season: spring (E; white), early summer (F; light blue), late summer (G; dark blue), and autumn (H; purple). To investigate the correlation, we used feature richness (fRich) as a response variable to a) plant species richness (pRich) and plant species Shannon (pShan), feature Shannon (fShan) as response to c) pRich and d) pShan and feature Evenness (fEven) as response to e) pRich and f) pShan.

Tables

Table 1a: Overview of plot traits, plant traits and metabolite feature diversity for species of FG grass. Recorded traits data on plot and plant level, including feature diversity across four seasons (S; E: spring, F: early summer, G: late summer, H: autumn) for species in the functional group (FG) grass. Data are presented as average per season and species. Recorded traits on the plot level include the plant species richness (pRich) and plant species Shannon index (pShan) for each plot, and height (in cm), developmental stage (BBCH scale), as well as mechanical (mechD in % above ground biomass) and pathogen inflicted damage (pathD in % above ground biomass) on the plant level. For each sample, the chemical composition was calculated as LC-MS feature richness (fRich), feature Shannon diversity (fShan) and feature Evenness (fEven). Detailed results for all 504 samples can be found in Supple Table 1 for plot level traits and in Suppl. Table 4 for plant level traits.

FG grass			plot trait			plant trait				feature diversity		
Species	plots	code	S	pRich	pShan	height	BBCH	mechD	pathD	fRich	fShan	fEven
Anthoxanthum odoratum		AO	E	18.5	2.3	57.9	78.8	0.0	16.4	499.5	2.2	0.3
Anthoxanthum odoratum	*A018, A044, B060, C115*	AO	F	10.0	1.8	10.0	50.0	4.7	15.8	633.7	3.9	0.6
Anthoxanthum odoratum		AO	G	13.5	1.9	9.1	51.9	10.4	21.3	642.0	5.2	0.8
Anthoxanthum odoratum		AO	H	13.4	2.1	9.6	37.1	3.7	4.9	541.6	4.4	0.7
Avenula pubescens		AP	E	18.0	2.2	99.0	56.9	1.3	16.5	471.8	4.0	0.7
Avenula pubescens	*B073, B092, C103, C110*	AP	F	9.8	1.7	17.4	40.6	4.0	11.0	561.0	5.1	0.8
Avenula pubescens		AP	G	13.8	1.8	15.4	42.5	1.6	22.5	545.6	5.2	0.8
Avenula pubescens		AP	H	13.3	1.9	13.4	40.0	3.5	8.3	549.0	5.3	0.8
Dactylis glomerata		DG	E	18.3	2.1	102.1	65.0	0.6	13.8	435.3	2.3	0.4
Dactylis glomerata	*A005, A018, C097, C137*	DG	F	9.5	1.7	23.8	55.0	5.1	10.6	432.5	2.8	0.5
Dactylis glomerata		DG	G	13.3	1.8	24.9	46.3	1.9	18.1	433.0	2.9	0.5
Dactylis glomerata		DG	H	15.8	2.1	21.5	41.3	4.4	9.8	390.9	3.0	0.5
Festuca rubra		FR	E	13.5	1.9	70.1	60.0	0.3	5.3	371.1	0.4	0.1
Festuca rubra	*B064, B067, B073, C114*	FR	F	9.0	1.4	23.3	46.3	4.5	15.3	334.6	0.4	0.1
Festuca rubra		FR	G	11.8	1.4	21.5	50.0	4.4	18.1	362.3	0.6	0.1
Festuca rubra		FR	H	12.8	1.7	12.1	38.8	4.6	5.6	321.9	0.6	0.1
Holcus lanatus		HL	E	17.3	2.1	76.1	62.5	0.6	10.6	449.6	1.3	0.2
Holcus lanatus	*A018, A035, A040, B080*	HL	F	9.3	1.5	12.6	51.9	3.1	21.9	515.1	1.9	0.3
Holcus lanatus		HL	G	15.5	2.0	10.8	55.0	3.0	17.5	489.5	2.7	0.4
Holcus lanatus		HL	H	16.5	2.3	11.1	46.3	3.4	10.6	407.4	2.0	0.3
Phleum pratense		PP	E	17.8	2.2	57.9	61.6	0.1	3.3	509.7	0.8	0.1
Phleum pratense	*A002, A018, A045, A046, B051, B059, B073, C133*	PP	F	9.6	1.8	23.4	44.7	3.2	15.8	507.9	1.7	0.3
Phleum pratense		PP	G	11.9	2.0	35.3	50.0	4.0	15.1	452.5	1.7	0.3
Phleum pratense		PP	H	15.3	2.0	15.3	38.4	3.8	10.1	395.0	2.2	0.4
Poa pratensis		PA	E	16.7	2.0	60.1	62.1	0.0	7.1	321.4	0.4	0.1
Poa pratensis	*A026, B073, B075, C131*	PA	F	7.4	1.1	18.3	50.0	4.3	14.6	372.3	1.0	0.2
Poa pratensis		PA	G	10.5	1.5	20.1	45.0	2.9	16.3	372.8	0.5	0.1
Poa pratensis		PA	H	13.5	1.7	16.3	40.0	5.1	4.4	371.9	0.7	0.1

Table 1b: Overview of plot traits, plant traits and metabolite feature diversity for species of FG herb. Recorded traits data on plot and plant level, including feature diversity across four seasons (S; E: spring, F: early summer, G: late summer, H: autumn) for species in the functional group (FG) herb. Data are presented as average per season and species. Recorded traits on the plot level include the plant species richness (pRich) and plant species Shannon index (pShan) for each plot, and height (in cm), developmental stage (BBCH scale), as well as mechanical (mechD in % above ground biomass) and pathogen inflicted damage (pathD in % above ground biomass) on the plant level. For each sample, the chemical composition was calculated as LC-MS feature richness (fRich), feature Shannon diversity (fShan) and feature Evenness (fEven). Detailed results for all 504 samples can be found in Supple Table 1 for plot level traits, and in Suppl. Table 4 for plant level traits.

FG herb			plot trait			plant trait				feature diversity		
Species		code	S	R	H	height	BBCH	mechD	pathD	fRich	fShan	fEven
Centaurea jacea		CJ	E	15.0	2.0	30.0	40.0	1.3	6.5	691.0	5.1	0.8
Centaurea jacea	*A013, A027, A030, B073*	CJ	F	7.5	1.4	28.9	42.5	2.3	11.3	637.8	5.0	0.8
Centaurea jacea		CJ	G	11.8	1.8	25.3	44.4	2.5	8.4	649.6	5.0	0.8
Centaurea jacea		CJ	H	14.3	2.0	21.0	38.8	3.5	4.8	601.0	5.0	0.8
Geranium pratense		GP	E	15.0	1.8	45.0	62.5	4.6	11.6	496.5	5.2	0.8
Geranium pratense	*A018, B081, C109, C121*	GP	F	10.5	1.7	31.8	38.1	5.1	11.8	463.9	5.1	0.8
Geranium pratense		GP	G	13.5	1.9	28.6	44.4	3.5	18.5	502.1	5.2	0.8
Geranium pratense		GP	H	14.8	2.2	12.6	40.0	3.9	10.6	471.8	5.1	0.8
Knautia arvensis		KA	E	15.5	1.8	76.3	83.1	2.9	11.9	762.9	5.5	0.8
Knautia arvensis	*A010, A020, A042, B073*	KA	F	8.5	1.3	38.8	48.1	5.0	7.6	724.4	5.5	0.8
Knautia arvensis		KA	G	9.5	1.5	38.0	51.9	5.0	17.5	800.1	5.6	0.8
Knautia arvensis		KA	H	11.5	1.8	25.3	40.0	6.3	12.5	625.4	5.3	0.8
Leucanthemum vulgare		LV	E	16.1	1.9	60.9	65.9	1.1	10.9	604.3	5.3	0.8
Leucanthemum vulgare	*A018, B048, B051, B073, B085, B090, C135, C136*	LV	F	7.9	1.5	8.5	43.4	1.0	3.8	602.9	5.2	0.8
Leucanthemum vulgare		LV	G	11.6	1.6	11.0	53.1	0.6	2.8	616.5	5.3	0.8
Leucanthemum vulgare		LV	H	15.3	2.0	6.9	40.6	1.9	4.4	528.6	5.2	0.8
Plantago lanceolata		PL	E	18.3	2.1	23.5	59.7	9.0	13.4	720.6	5.6	0.8
Plantago lanceolata	*A009, A011, A018, A043, B054, B057, B073, C115*	PL	F	8.9	1.6	26.1	46.6	13.8	13.1	748.0	5.6	0.8
Plantago lanceolata		PL	G	13.5	2.0	23.9	51.3	7.9	14.7	771.3	5.7	0.9
Plantago lanceolata		PL	H	15.5	2.2	9.3	41.6	15.3	9.1	712.6	5.4	0.8
Ranunculus acris		RA	E	21.0	2.4	54.6	80.0	4.0	21.5	470.3	1.3	0.2
Ranunculus acris	*A003, A016, A018, B071*	RA	F	11.8	1.9	16.3	42.5	2.0	4.4	452.5	1.1	0.2
Ranunculus acris		RA	G	17.9	2.4	14.1	51.4	2.1	20.7	483.7	1.0	0.2
Ranunculus acris		RA	H	15.0	2.1	11.6	36.3	5.0	17.3	453.9	1.0	0.2

Table 2: pMetFam_ID identification table for putatively annotated metabolite families. Features for annotation were selected by PLS models performed for each species on both plot-level (season, pRich, pShan) and plant-level traits (BBCH, height, mechanical, and pathogen inflicted damage). The selected features were putatively annotated using databases available with MetFamily. The number of samples that had an intensity measured for a feature annotated in one of the metabolite families is given as count across all 504 samples. Note that in some metabolite families multiple features were annotated, resulting for the count to exceed the number of 504 samples. A detailed list of annotated features is available in Suppl Table 3a (feature selection overview) and Suppl Table 3b (feature intensities).

pMetFam_ID	Kingdom	Superclass	Class	Subclass	Parent	count
pMetFam_01-A1.1	(01) Organic nitrogen compounds	(A) Organonitrogen compounds	(1) Amines	(.1) Primary amines		43
pMetFam_01-A1.1a	(01) Organic nitrogen compounds	(A) Organonitrogen compounds	(1) Amines	(.1) Primary amines	(a) Monoalkylamines	71
pMetFam_02	(02) Organopnictogen compounds					232
pMetFam_03	(03) Benzenoids					210
pMetFam_03-A1	(03) Benzenoids	(A) Benzene and substituted derivatives	(1) Phenylpropanes			14
pMetFam_03-B1	(03) Benzenoids	(B) Phenols	(1) Methoxyphenols			235
pMetFam_04-A1.1a	(04) Organic acids and derivatives	(A) Organic phosphoric acids and derivatives	(1) Phosphate esters	(.1) Alkyl phosphates	(a) Monoalkyl phosphates	448
pMetFam_04-B1.1	(04) Organic acids and derivatives	(B) Carboxylic acids and derivatives	(1) Amino acids peptides and analogues	(.1) Peptides		282
pMetFam_04-B2.1	(04) Organic acids and derivatives	(B) Carboxylic acids and derivatives	(2) Carboxylic acid derivatives	(.1) Carboxylic acid esters		38
pMetFam_04-B2.2a	(04) Organic acids and derivatives	(B) Carboxylic acids and derivatives	(2) Carboxylic acid derivatives	(.2) Carboxylic acid amides	(a) Secondary carboxylic acid amides	121
pMetFam_04-B2.2b	(04) Organic acids and derivatives	(B) Carboxylic acids and derivatives	(2) Carboxylic acid derivatives	(.2) Carboxylic acid amides	(b) Tertiary carboxylic acid amides	203
pMetFam_05-A1	(05) Lipids and lipid-like molecules	(A) Steroids and steroid derivatives	(1) Spirostanes and derivatives			22
pMetFam_05-B1.1	(05) Lipids and lipid-like molecules	(B) Fatty Acyls	(1) Fatty acyl glycosides	(.1) Fatty acyl glycosides of mono- and disaccharides		54
pMetFam_05-C	(05) Lipids and lipid-like molecules	(C) Prenol lipids				1058
pMetFam_06	(06) Organoheterocyclic compounds					241
pMetFam_06-A1.1	(06) Organoheterocyclic compounds	(A) Quinolines and derivatives	(1) Quinolones and derivatives	(.1) Hydroquinolones		27
pMetFam_06-B	(06) Organoheterocyclic compounds	(B) Tetrahydroisoquinolines				32
pMetFam_06-C1.1	(06) Organoheterocyclic compounds	(C) Pyridines and derivatives	(1) Pyridinecarboxylic acids and derivatives	(.1) Pyridinecarboxylic acids		74
pMetFam_06-D	(06) Organoheterocyclic compounds	(D) Tetrahydrofurans				140
pMetFam_06-E1.1a	(06) Organoheterocyclic compounds	(E) Benzopyrans	(1) 1-benzopyrans	(.1) Chromones	(a) 3-methoxychromones	157
pMetFam_06-F	(06) Organoheterocyclic compounds	(F) Oxanes				222
pMetFam_06-G1.1	(06) Organoheterocyclic compounds	(G) Diazines	(1) Pyrimidines and pyrimidine derivatives	(.1) Hydropyrimidines		302
pMetFam_06-H	(06) Organoheterocyclic compounds	(H) Pyrrolidines				340
pMetFam_06-I1.1	(06) Organoheterocyclic compounds	(I) Imidazopyrimidines	(1) Purines and purine derivatives	(.1) 6-aminopurines		121
pMetFam_06-I1.2a	(06) Organoheterocyclic compounds	(I) Imidazopyrimidines	(1) Purines and purine derivatives	(.2) Purinones	(a) 6-oxopurines	378
pMetFam_07	(07) Organic oxygen compounds					1110
pMetFam_07-A	(07) Organic oxygen compounds	(A) Organonitrogen compounds				1058
pMetFam_07-A1	(07) Organic oxygen compounds	(A) Organonitrogen compounds	(1) Carbonyl compounds			148
pMetFam_07-A1.1a	(07) Organic oxygen compounds	(A) Organonitrogen compounds	(1) Carbonyl compounds	(.1) Ketones	(a) Aryl ketones	21
pMetFam_07-A2	(07) Organic oxygen compounds	(A) Organonitrogen compounds	(2) Alcohols and polyols			268
pMetFam_07-A2.1	(07) Organic oxygen compounds	(A) Organonitrogen compounds	(2) Alcohols and polyols	(.1) Polyols		37
pMetFam_07-A3.1a	(07) Organic oxygen compounds	(A) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(.1) Monosaccharides	(a) Pentoses	166
pMetFam_07-A3.2a	(07) Organic oxygen compounds	(A) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(.2) Glycosyl compounds	(a) C-glycosyl compounds	603
pMetFam_08	(08) Phenylpropanoids and polyketides					123
pMetFam_08-A	(08) Phenylpropanoids and polyketides	(A) Stilbenes				7
pMetFam_08-B1	(08) Phenylpropanoids and polyketides	(B) Diarylheptanoids	(1) Linear diarylheptanoids			139
pMetFam_08-C1	(08) Phenylpropanoids and polyketides	(C) Tannins	(1) Hydrolyzable tannins			192
pMetFam_08-D	(08) Phenylpropanoids and polyketides	(D) Coumarins and derivatives				39
pMetFam_08-D1.1	(08) Phenylpropanoids and polyketides	(D) Coumarins and derivatives	(1) Pyranocoumarins	(.1) Angular pyranocoumarins		71
pMetFam_08-D2.1	(08) Phenylpropanoids and polyketides	(D) Coumarins and derivatives	(2) Hydroxycoumarins	(.1) 7-hydroxycoumarins		135
pMetFam_08-E1	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(1) O-methylated isoflavonoids			24
pMetFam_08-E2.1	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(2) Isoflavans	(.1) Isoflavanols		80
pMetFam_08-E2.2a	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(2) Isoflavans	(.2) Isoflavanones	(a) 6-prenylated isoflavanones	530
pMetFam_08-E3	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(3) Furanoisoflavonoids			496
pMetFam_08-E3.1	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(3) Furanoisoflavonoids	(.1) Pterocarpans		713
pMetFam_08-F	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives				21
pMetFam_08-F1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(1) Cinnamic acid esters			52
pMetFam_08-F2	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives			388
pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		1669
pMetFam_08-F2.2	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters		26
pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	1703
pMetFam_08-G	(08) Phenylpropanoids and polyketides	(G) Flavonoids				221
pMetFam_08-G1.1	(08) Phenylpropanoids and polyketides	(G) Flavonoids	(1) O-methylated flavonoids	(.1) 6-O-methylated flavonoids		32
pMetFam_08-G1.2	(08) Phenylpropanoids and polyketides	(G) Flavonoids	(1) O-methylated flavonoids	(.2) 3'-O-methylated flavonoids		67
pMetFam_08-G1.3	(08) Phenylpropanoids and polyketides	(G) Flavonoids	(1) O-methylated flavonoids	(.3) 7-O-methylated flavonoids		295
pMetFam_08-G1.4	(08) Phenylpropanoids and polyketides	(G) Flavonoids	(1) O-methylated flavonoids	(.4) 8-O-methylated flavonoids		573
pMetFam_08-G2	(08) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides			91
pMetFam_08-G2.1	(08) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides		255
pMetFam_08-G2.1a	(08) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(a) Anthocyanins	136
pMetFam_08-G2.1b	(08) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(b) Flavonoid O-glucuronides	152
pMetFam_08-G2.1c	(08) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	1154
pMetFam_08-G3	(08) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids			437
pMetFam_08-G3.1	(08) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.1) 5-hydroxyflavonoids		470
pMetFam_08-G3.2	(08) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.2) 4'-hydroxyflavonoids		505
pMetFam_08-G3.3	(08) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.3) 7-hydroxyflavonoids		1001

Table 3: Linear models calculated for the prediction variables (preV): plant species richness (pRich) and plant species Shannon (pShan) and the response variables (resV): LC-MS feature richness (fRich), LC-MS feature Shannon (fShan) and LC-MS feature Evenness (fEven). For each model, the coefficient of determination (R^2), root means square error (RMSE) and p-values (p) are shown. The positive and negative values shown in square brackets indicate the direction of the relationship between the two variables. Linear models were calculated as a combination of all species (all) and for each species separately. Species name abbreviations can be found in Table 1 a (grass species) and b (herb species). Detailed results are also available in Suppl. Table 5.

	resV	fRich			fShan			fEven		
preV	spec	R^2	RMSE	p	R^2	RMSE	p	R^2	RMSE	p
pRich	all	[-] 0.28	135.90	0.03	[-] 0.28	2.05	0.04	[-] 0.25	0.32	0.05
pRich	AO	[-] 0.37	83.06	0.03	[-] 0.53	1.44	0.09	[-] 0.55	0.22	0.11
pRich	AP	[-] 0.45	62.51	0.30	[-] 0.43	0.88	0.10	[-] 0.58	0.13	0.10
pRich	DG	[-] 0.41	41.66	0.08	[-] 0.56	1.32	0.76	[-] 0.56	0.22	0.80
pRich	FR	[+] 0.46	41.93	0.63	[+] 0.74	0.23	0.02	[+] 0.75	0.04	0.02
pRich	HL	[-] 0.64	56.34	0.01	[-] 0.42	0.76	0.16	[-] 0.41	0.12	0.18
pRich	PP	[+] 0.19	79.31	0.82	[-] 0.20	1.11	0.38	[-] 0.20	0.18	0.42
pRich	PA	[-] 0.65	45.03	0.43	[-] 0.79	0.54	0.17	[-] 0.78	0.10	0.17
pRich	CJ	[+] 0.57	65.53	0.15	[+] 0.69	0.12	0.11	[+] 0.46	0.01	0.24
pRich	GP	[-] 0.49	34.57	0.58	[-] 0.48	0.13	0.26	[-] 0.53	0.01	0.22
pRich	KA	[+] 0.48	80.65	0.73	[+] 0.50	0.16	0.64	[+] 0.42	0.01	0.58
pRich	LV	[-] 0.22	60.62	0.01	[+] 0.26	0.16	0.90	[+] 0.28	0.02	0.02
pRich	PL	[-] 0.23	62.16	0.01	[-] 0.25	0.18	0.08	[-] 0.30	0.02	0.35
pRich	RA	[+] 0.50	37.05	0.93	[+] 0.56	0.28	0.23	[+] 0.56	0.05	0.24
pShan	all	[-] 0.25	136.33	0.21	[-] 0.27	2.06	0.13	[-] 0.25	0.32	0.16
pShan	AO	[-] 0.41	79.89	0.00	[-] 0.54	1.52	0.32	[-] 0.56	0.24	0.39
pShan	AP	[-] 0.51	60.01	0.13	[-] 0.59	0.80	0.04	[-] 0.62	0.12	0.04
pShan	DG	[-] 0.52	42.29	0.02	[-] 0.68	1.29	0.26	[-] 0.68	0.22	0.29
pShan	FR	[-] 0.51	40.49	0.79	[-] 0.58	0.23	0.88	[-] 0.59	0.04	0.94
pShan	HL	[-] 0.51	57.21	0.02	[-] 0.48	0.78	0.25	[-] 0.48	0.13	0.29
pShan	PP	[+] 0.16	79.01	1.00	[-] 0.22	1.10	0.10	[-] 0.21	0.18	0.11
pShan	PA	[-] 0.49	46.30	0.86	[-] 0.51	0.47	0.81	[-] 0.50	0.08	0.82
pShan	CJ	[+] 0.55	68.06	0.30	[+] 0.64	0.13	0.23	[+] 0.55	0.01	0.35
pShan	GP	[-] 0.54	33.25	0.23	[-] 0.53	0.12	0.04	[-] 0.68	0.01	0.05
pShan	KA	[-] 0.56	81.60	0.63	[-] 0.55	0.16	0.54	[-] 0.48	0.01	0.50
pShan	LV	[-] 0.25	60.30	0.00	[-] 0.17	0.16	0.59	[+] 0.31	0.02	0.16
pShan	PL	[-] 0.26	64.16	0.05	[-] 0.22	0.18	0.10	[-] 0.21	0.02	0.26
pShan	RA	[+] 0.52	36.75	0.20	[+] 0.43	0.29	0.69	[+] 0.41	0.05	0.74

Competing interests

The authors declare no competing interests.

Author contributions

SM designed the experiment, conducted sampling and measurements, prepared and curated the data set, conducted the raw data pre-processing, conceptualised and designed the data analysis performed the statistical analysis, designed and created figures and tables, and wrote the manuscript. SN contributed to data management and repository administration, provided the infrastructure for data acquisition. All authors reviewed and approved the manuscript.

Acknowledgements

The authors gratefully acknowledge the support of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – DFG–FZT 118, 202548816. The authors also thank the technical staff of the Jena Experiment (DFG, FOR 456/1451), the coordinator Anne Ebeling as well as many student helpers for maintaining the experimental field site and their support during measurements. Furthermore, the authors thank the technical staff of IPB Halle for their support and maintenance of the technical equipment used during the measurements. This paper has been conducted in the framework of the iDiv-Flexpool – the internal funding mechanism of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – DFG–FZT 118, 202548816.

Appendix

See Appendix – Chapter 3

DISCUSSION

The main objectives in this thesis included the development of an automated workflow to enable the analysis of ecometabolomics data across highly diverse species and environments, the evaluation of available statistical tools and their power to be integrated with ecometabolomics experiments, and the investigation of secondary plant metabolite profile changes across *season* and *diversity*.

Season, in this study, describes a time point of sampling within the growing phase of grassland plant species, ranging from spring across summer to autumn. This includes the collective factors of weather dynamics, including light, water, and nutrient availability, that impact plant growth and productivity without being recorded separately. Moreover, given the Jena-Experiment-specific conditions, *season*, here, also includes changes to the immediate plant community in the plots through maintenance actions, such as weeding and mowing that are completed up to twice a year, resulting in infrequently influenced plant communities on a non-reproducible level. For example, in plots with mainly grass species accounting for the community richness, usually only a few weed species, e.g. larger herbs, are weeded out, leaving the majority of the plot undisturbed. In contrast, plots with dominant herb communities got weeded out thoroughly, leaving large patches of dug-up bare soil in the plots behind, thereby changing soil profiles and nutrient availability and herbivory pressure. Hence, plants will have to invest all available resources to grow back to reproduce within their growing period; in contrast to the spring sample collection, where the plant had about 6 months of the undisturbed growing period, early summer sampling took place right after the first weeding and mowing action in the year, resulting in an overall lower plant species richness. Late summer samples were collected before each year's second mowing and weeding. Hence, the richness of plant species was generally higher with more flowering species. Autumn samples were collected a few weeks after the second mowing and weeding. At this point, most plants had already fully matured. Therefore, the metabolite data collected in this thesis are not comprehensively representative of general dynamics in the metabolite profiles of grassland plant species but gave the unique opportunity to use a complex data set collected from a none-uniform environment to create a workflow that is capable of handling metabolite profiles with a high biological variance. Therefore, the *diversity* that was used to draw conclusions about the interaction of the plant species communities and the metabolite profiles, and is discussed in this study, is the actual species richness and abundance recorded in each plot prior to each sample collection, rather than the designed TBE diversity levels (Chapter 3).

Performing ecometabolomics studies across multiple species includes a range of obstacles that need to be addressed properly to ensure reliable results and conclusion drawing (Marr et al. 2021). If these measures are not taken into account, for instance, the use of quality controls, blanks and randomisation strategies for both sampling and measurements, it may result in over-optimistic results that inevitably lead to conclusions that are not or not entirely supported by the data. Most studies also do not include details about their sampling protocols, including information about the use of blank samples to trace contaminations, randomisation strategies of their sample collection, and measurements (Scherling et al. 2010, Berini et al. 2018, Ristok et al. 2019, Ristok et al. 2022). With few exceptions (Ristok et al. 2019), reproducible strategies regarding data preprocessing, including blank subtraction or removal, batch correction and the handling of missing values, are also often overlooked or not given in detail (Scherling et al. 2010, Berini et al. 2018, Ristok et al. 2019, Ristok et al. 2022). For example, Scherling et al. (2010) reported a significant correlation between plant species diversity and metabolic profiles in some of the tested herb species. However, in this study, the authors do not comprehensively report how these data were acquired or which blanks, quality controls or randomisation strategies were used. This does not mean that their findings are incorrect; it only challenges the data to the point that the conclusion drawing should be done with the uttermost care. Moreover, even dedicated reviews stating key rules for the successful acquisition of metabolomic data do not take into account the relevance of blanks taken while sampling or sample extraction, randomisation while sample collection and measurements, nor the use of QC samples for signal drift control and batch correction (Defossez et al. 2023). Hence, the need for a comprehensive handbook that includes standardised protocols for ecometabolomics studies is one of the highest priorities in developing ecometabolomics experiments and observational studies (Walker et al. 2022).

In general, the amount of data generated in ecometabolomic studies requires careful handling and curation strategies. Biases, for example, are commonly introduced by biological variance and technical deficiencies. Increasing efforts are being taken to improve available data analysis tools (Wang et al. 2016, Marr et al. 2021, Walker et al. 2022). However, no common standard is available yet. This thesis focused on providing a comprehensive workflow that includes sample handling, data curation and preparation steps to ensure the

quality of ecometabolomics data collected from field experiments. One important point is that data handling not only needs to consider the data amount but also the biology-prone dynamics in samples obtained from field experiments in order for the scientist to be able to draw conclusions about underlying mechanisms altering and reflecting the metabolomics profiles. It was already shown that plant metabolite profiles could show a large biological fluctuation under environmental pressure (Scherling et al. 2010). As Berini et al. (2018) discussed, in ecometabolomics, findings and interpretations can be very challenging and small changes in environmental conditions can dramatically affect metabolite profiles. In this thesis, the biological variance was taken into account by introducing a comprehensive randomisation strategy during each sampling in the fields, ensuring that samples were collected across FG and plot specifications. Furthermore, given the rapid changes in the dynamic metabolite profiles across the diurnal circle and environmental conditions, keeping sampling collection in a narrow time frame and recording the exact time of sampling can contribute to the reliability of the ecometabolomics data. However, these data are still only rarely stated in ecometabolomics studies (Scherling et al. 2010, Berini et al. 2018, Marr et al. 2021, Ristok et al. 2022). Another often overlooked obstacle is using blank samples, not only in the lab or as run-in samples while starting measurements, but also directly in the fields to trace back contaminations. These include volatiles of unknown origins, abrasion from rubber gloves or sample tubes and organic materials that are introduced while sampling. Tracing all these sources of contamination and excluding features from the associated measurement is an essential step towards reliability increase and data quality enhancement.

Moreover, the inclusion of quality control samples (QCs) has been shown to be an important step in ecometabolomics LC-MS-based studies, to overcome technical limitations. For example, the signal shift is a common issue when using LC-MS analysis methods, as the instruments get clogged with sample extract, and the measurement capacity decreases over time. Addressing signal drift and other machine-dependent biases is crucial in metabolomic analysis, as the importance of a certain metabolite is defined by its relative changes in abundance. Therefore, it must be ensured that any metabolite's intensity can be directly compared to the measured intensities in other samples (Fiehn 2002). To be able to address these issues and keep samples comparable across multiple batches, QC samples, which are aliquots of the same sample that gets measured multiple times within a batch, are used to correct for unwanted shifts of the measured intensities (Chapter 1: Marr et al. 2021). As Walker et al. (2022) pointed out, underlying chemical structures, including degrees of chemical similarity, are not represented in the commonly used statistical methods. Hence, the consequences of environmental stress on metabolite profile composition and abundance remain mainly unpredictable (Berini et al. 2018). In order to be able to link changes in plant metabolite profiles to classical functional traits, significant overlap of these data is required, with harmonised strategies for metabolomic data acquisition across studies, species and ecosystems (Walker et al. 2022).

One of the major obstacles in multi-species studies is the huge number of missing values in the resulting data matrix, as many features are unique to certain species and are, therefore, not measured in the other species in the experiment. PLS, RF and SVM are among the common choices to analyse zero-inflated data sets (Defossez et al. 2021, Marr et al. 2021) as they are less biased than other ordinations (Walker et al. 2022). Especially in multi-species studies with very distinctive secondary metabolite fingerprints, the generated data matrices can include a high number of missing values (see Chapter 1 for an explanation) that should be considered before using the processed data in statistical analysis. Therefore, in this thesis, the usability of differently prepared matrices with different statistical methods was tested to show how reliable classifications were based on the level of background similarity. Statistical methods that are already successfully used in classical metabolomics experiments were tested: random forest (RF), support vector machines (SVM) and partial least squares (PLS). Performance tests and evaluations were performed on three levels of distinctiveness: functional-group-wise, species wise and across treatment within a species. Here, the data matrix was either used with the original zeros as present after the pre-processing step or with distinctive small values to replace zeros in the data matrix. Choosing a suitable classification method mainly depended on the complexity of the provided background for classification. The choice of the data matrix depending on the model also improved the performance. For example, comparing the model performance, we found the highest accuracies in *PLS* and *RF. PLS*, *SVM* and *RF* worked equally well on very distinguishable metabolite profiles, while *PLS* and *RF* provided a greater classification power when the background spectra were more similar. The best “statistical tool-data matrix” combination highly depended on the complexity level used as a reference for classification (see discussion Chapter 2), indicating that secondary metabolite profiles include highly species-specific parts – the secondary metabolite fingerprint –, and parts responding to environmental changes – the dynamic profile.

All species tested in this thesis showed a distinctive pattern of species-specific metabolite fingerprints and dynamic profiles (see Chapter 1). The terms metabolomic profiling and fingerprinting are vastly used in the research community with an often slightly different focus. For example, Weston et al. (2015) used metabolomic

profiling to describe quantity changes in a specific subset of metabolites that respond to environmental changes. According to Fiehn (2002), identifying and quantifying pre-selected metabolites in a sample is referred to as metabolite profiling, for instance, the investigation of a specific metabolite family, such as isoprenoids and carbohydrates. In comparison, others argue that the term metabolite profiling is used to describe the investigation of any metabolite changes as a response to a certain environmental factor or treatment (Kim and Verpoorte 2010, Obata & Fernie 2012). Furthermore, in the literature, the term *fingerprinting* is used to describe the process of classifying samples "according to the origin or their biological relevance" (Fiehn 2002) and to describe detected metabolomic shifts by LC-MS and GC-MS as a result of responses to environmental pressures (Scherling et al. 2010), which also corresponds with metabolite profiling as per definition. Therefore, to avoid confusion and to use clear identifiers to distinguish between methods, findings, and biological characteristics, in this thesis, the terms *metabolite fingerprint* and *dynamic profile* are used to describe the two different areas that the secondary *metabolite profile* of plant leaves is composed of (see discussions Chapters 2 and 3).

As demonstrated in chapter 2 of the thesis, the classification of plant species works reliably across all seasons and plant community dynamics, even in closely related species, although they share a greater overlapping proportion in their secondary metabolite fingerprints. When using the same statistical methods within the same species to distinguish between environmental changes – using the dynamic profile – the results strongly depend on the combination of statistical methods and the choice of a suitable data matrix format. All changes in the environment, such as the season of sample collection, plant species richness and composition in the plots, were reflected in the metabolite profile, mostly in the dynamic part. For metabolites representing conserved phylogenetic reproduction and coping strategies, it can be expected that such metabolites maintain the same importance over the life cycle of an individual plant across seasonal dynamics (Walker et al. 2022). It is unknown to what extent metabolites serve an explicit function, i.e. the proportion between fingerprint and dynamic profile is unknown and might vary greatly between species (Fiehn 2002). Those changes could be accessed by selecting the proportion of features that account for changes in the profiles across environmental dynamics, including classical traits such as developmental stages and plant height. However, separating the effects of each trait or environmental factor will increase the chance of overinterpretation as most of these factors are interconnected.

The findings in this thesis suggest that metabolite profile changes are directly influenced by changes in the plant neighbourhood. These changes in the plant community can be, for example, triggered by seasonal changes, including changing day length and, hence, changing light and water availability, and changing temperatures. Furthermore, fertilization and nutrient availability change throughout the year, causing changes in herbivory activity, terrestrial fauna, and metabolite profiles (Scherling et al. 2010, Richards et al. 2015). Previous studies showed that the plant metabolome also shapes the herbivore pressure within plant communities (Fernandez-Conradi et al. 2021, Philibin et al. 2021, Ristok et al. 2022). As described above, the semi-maintenance, i.e. disturbance of plots up to twice a year, also contributes to changes in the plant community, as some plots are weeded out entirely. In contrast, others remain mainly undisturbed throughout the year. Although the responses to the changing plant diversity were mainly species-specific, we also found similarities in closely related species, i.e. FG grass. In this study, the species combined to FG herb are probably phylogenetically too far and dissimilar to show the same changes in their metabolite profiles. This bio-geo-legacy response was mainly independent of direct plant features such as developmental stage and inflicted damage. Similar to Ristok et al. (2022), in this study, the metabolite richness and diversity was found to differ between FG grass and FG herb species, with grasses generally having fewer metabolites. While grass species mostly rely on physical structures for defence (Massey & Hartley 2009), herb species mostly invest in chemical defence (Cooke & Leishman 2012). As Ristok et al. (2022) pointed out, grass and herb species also vary in their interaction with mycorrhizal fungi, which might enhance the differences between herb and grass metabolite richness and composition.

Variations of physical traits between plant species can support a successful plant community (Scherling et al. 2010). Following this, it can be speculated that the comprehensive coverage of secondary metabolites that effect herbivory and other environmental pressures is beneficial for diverse plant communities as they can be prepared for a broad range of stress factors while not having to invest all resources on their own. Plants may benefit from sharing recourses for secondary metabolites used as a defence against pests and herbivores. Our findings of metabolites in each single species, hence, might be a result of the specific species composition present in each plot. As discussed in chapters 2 and 3 of this thesis, the feature selection revealed that some metabolite families were found to be important for the metabolite profile changes as a response to the environment. In contrast, the same metabolite families were present in other species but seemingly did not contribute to the dynamic profile responses.

As demonstrated in this thesis, the fingerprints in the LC-MS features profiles can be used to identify plant species across environmental dynamics. However, using secondary metabolite profiles for species identification

is usually based on unique metabolites that accurately identify a species, though this can be challenging. For example, a unique feature-based identification requires identifying these unique compounds first and establishing a broad database as a reference that explicitly shows that this feature is unique to only one species. Secondly, these compounds and metabolites might be produced throughout closely related species, making it challenging to distinguish these species solely based on one unique compound. In this thesis, using secondary metabolite profiles as a whole was proposed to be used for species identification, not requiring the identification of each compound or metabolite, but using full fingerprints for out of the raw data. With the fingerprint being conserved across the growing season and across changing community neighbourhoods, as demonstrated in this thesis, here it is also assumable that the secondary metabolite fingerprint can be used to classify species across environmental changes and treatments and that this classification can also be used in multiple species assembles with closely related species. Unlike in most ecometabolomics studies, identification of marker compounds and metabolites unique to a single species is not required. Using the whole fingerprint provides the advantage that the profiles can be used for classification based on the raw data and do not require cost- and time-intensive identifications of single relevant compounds and metabolites.

There are some possible applications for using secondary metabolite fingerprint-based identification. For example, metabolite-assisted crop breeding and fodder for the livestock industry, as a broad collection of weed plants, would automatically identify different species and screen for potentially toxic plants (Fang et al. 2019). Automated identification of species-specific profiles can also help narrow down the number of potentially interesting features and compounds involved in stress responses and identify only those of interest. Possible areas of application could be metabolomic engineering, for instance, to increase the nutritional value of foods of defence metabolites to decrease the need for pesticides and improve plant performance in general (Fiehn 2002). Fast identification of plants can be of interest to farmers to test a field of unknown plants for potentially toxic plants to livestock or beneficial plant species for crop production. Especially in industrial contexts, the benefit from reliable predictions of metabolite abundance and composition in certain plant species will be tremendous, for example, when using such predictions to estimate potential growth and resistance (Lipka et al. 2015, Tohge et al. 2016, Peng et al. 2017).

In the context of global change, it is inevitable to understand the underlying processes of ecosystem function (Berini et al. 2018). The metabolome represents the collective result of all biotic and abiotic interactions of the plant with its environment. Extending ecometabolomics studies across multiple species and across multiple ecosystems will drastically improve the understanding of the role of specialised metabolites in ecosystem services (Walker et al. 2022). Predicting how plants will respond to environmental changes can be challenging as not all parts of the metabolic profiles are equally influenced by different factors (Berini et al. 2018). Metabolomics and prediction of the abundance of certain metabolites can be of interest for example to check whether weeds growing in the field could potentially be toxic to livestock (Weston et al. 2013). Understanding how plants respond to climate change is highly interesting to the research community. However, testing the influence of temperature changes alone is nearly impossible in field experiments (Berini et al. 2018), as other biotic and abiotic interactions cause constant changes to metabolite profiles. On the other hand, conducting such an experiment under controlled greenhouse conditions will also take away these interactions and won't allow for a precise interpretation and estimation of changes to the metabolite profile. Hence, performing these types of field studies, where all interactions with the environment are included in the experimental design, is an important step towards unravelling underlying mechanisms and reliable predictions for future climates.

This thesis showed that secondary metabolite profiles maintain several metabolites as species-specific fingerprints, representing species identity, compared to the dynamic metabolites representing the environmental bio-geo-legacy. The fast profile-based identification of plant species enables the researcher to narrow down the number of potentially interesting compounds and metabolites that take part in stress response. The fast identification will also be of interest in industrial contexts, progressing in the direction of metabolite-assisted breeding strategies. Furthermore, gaining new insights into biochemical diversity interactions and environment-related production of secondary metabolites finds special interest in the exploration of future climate scenarios. There is a growing interest in understanding how warming and changing environmental conditions, in general, will influence certain plant species and their secondary metabolite profiles. In the context of climate change, being able to use LC-MS feature-based databases as a classification tool will have a major impact on the understanding and potential prediction of plant responses. Being able to predict plant responses, especially changes to metabolite profiles, and phytochemistry in general under different conditions will also enable the understanding and prediction of changes in plants when changing the environment into the unknown, such as space stations and other planets.

In this thesis, the discussion was initiated what reliable data handling and curation need in terms of tools, methods, and strategies to enable reliable research in ecometabolomic experiments and to include the natural diversity brought in by plant species into the data handling strategies. Furthermore, it was discussed how established statistical methods could be adapted and made available for the big data sets that incorporate massive amounts of feature data from different backgrounds and contexts without overestimating the power and dependency of independent data. Given that metabolomics profiles can change rapidly under environmental dynamics, including daytime and weather changes, it is exceptionally difficult to acquire reliable data from samples collected in the field. Therefore, handling samples with uttermost care is crucial to ecometabolomics experiments. However, in many studies, using blanks or other quality controls of any kind, including randomisation strategies for sampling and data acquisition, is still not integrated as standard in the studies protocols. Drawing conclusions from those studies challenges the reliability of the findings used to elucidate underlying mechanisms and artefacts that belong to biases introduced by design and lab protocols. As ecometabolomics is rapidly gaining interest in the research community, the need for a handbook that comprehensively describes the obstacles and steps needed to be taken in ecometabolomics studies would drastically increase the data quality and enhance the reliability of the conclusions. Till there is no common consensus found within the community, drawing conclusions from ecometabolomic studies should be done carefully and with an open mind.

REFERENCES

Agerbirk, N., & Olsen, C. E. (2012). Glucosinolate structures in evolution. *Phytochemistry*, *77*, 16-45.

Agrawal, A. A., & Weber, M. G. (2015). On the study of plant defence and herbivory using comparative approaches: how important are secondary plant compounds. *Ecology Letters*, *18*(10), 985-991.

Allard, P. M., Genta-Jouve, G., & Wolfender, J. L. (2017). Deep metabolome annotation in natural products research: towards a virtuous cycle in metabolite identification. *Current Opinion in Chemical Biology*, *36*, 40-49.

Alseekh, S., & Fernie, A. R. (2018). Metabolomics 20 years on: what have we learned and what hurdles remain?. *The Plant Journal*, *94*(6), 933-942.

Baldwin, I. T., Halitschke, R., Paschold, A., Von Dahl, C. C., & Preston, C. A. (2006). Volatile signaling in plant-plant interactions:" talking trees" in the genomics era. *science*, *311*(5762), 812-815.

Balvanera, P., Pfisterer, A. B., Buchmann, N., He, J. S., Nakashizuka, T., Raffaelli, D., & Schmid, B. (2006). Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology letters*, *9*(10), 1146-1156.

Berini, J. L., Brockman, S. A., Hegeman, A. D., Reich, P. B., Muthukrishnan, R., Montgomery, R. A., & Forester, J. D. (2018). Combinations of abiotic factors differentially alter production of plant secondary metabolites in five woody plant species in the boreal-temperate transition zone. *Frontiers in Plant Science*, *9*, 1257.

Cardinale, B. J., Wright, J. P., Cadotte, M. W., Carroll, I. T., Hector, A., Srivastava, D. S., ... & Weis, J. J. (2007). Impacts of plant diversity on biomass production increase through time because of species complementarity. *Proceedings of the National Academy of Sciences*, *104*(46), 18123-18128.

Chiapusio, G., Jassey, V. E., Bellvert, F., Comte, G., Weston, L. A., Delarue, F., ... & Binet, P. (2018). Sphagnum species modulate their phenolic profiles and mycorrhizal colonization of surrounding Andromeda polifolia along peatland microhabitats. *Journal of chemical ecology*, *44*, 1146-1157.

Cooke, J., & Leishman, M. R. (2012). Tradeoffs between foliar silicon and carbon-based defences: evidence from vegetation communities of contrasting soil types. *Oikos*, *121*(12), 2052-2060.

van Dam, N. M., & Heil, M. (2011). Multitrophic interactions below and above ground: en route to the next level. *Journal of Ecology*, *99*(1), 77-88.

van Dam, N. M., & van der Meijden, E. (2011). A role for metabolomics in plant ecology. *Annual Plant Reviews Volume 43: Biology of Plant Metabolomics*, *43*, 87-107.

Defossez, E., Pitteloud, C., Descombes, P., Glauser, G., Allard, P. M., Walker, T. W., ... & Rasmann, S. (2021). Spatial and evolutionary predictability of phytochemical diversity. *Proceedings of the National Academy of Sciences*, *118*(3), e2013344118.

Defossez, E., Bourquin, J., von Reuss, S., Rasmann, S., & Glauser, G. (2023). Eight key rules for successful data-dependent acquisition in mass spectrometry-based metabolomics. *Mass Spectrometry Reviews*, *42*(1), 131-143.

Díaz, S., J. Kattge, J. H. C. Cornelissen, I. J. Wright, S. Lavorel, S. Dray, B. Reu, M. Kleyer, C. Wirth, I. C. Prentice, E. Garnier, G. Bönisch, M. Westoby, H. Poorter, P. B. Reich, A. T. Moles, J. Dickie, A. N. Gillison, A. E. Zanne, J. Chave, S. J. Wright, S. N. Sheremet'ev, H. Jactel, C. Baraloto, B. Cerabolini, S. Pierce, B. Shipley, D. Kirkup, F. Casanoves, J. S. Joswig, A. Günther, V. Falczuk, N. Rüger, M. D. Mahecha, Gorné, L. D. (2016). The global spectrum of plant form and function. *Nature*, *529*(7585), 167-171.

Ebeling, A., Meyer, S. T., Abbas, M., Eisenhauer, N., Hillebrand, H., Lange, M., ... & Weisser, W. W. (2014). Plant diversity impacts decomposition and herbivory via changes in aboveground arthropods. *PloS one*, *9*(9), e106529.

Ellis, D. I., & Goodacre, R. (2006). Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy. *Analyst*, *131*(8), 875-885.

Fahey, J. W., Zalcmann, A. T., & Talalay, P. (2001). The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry*, *56*(1), 5-51.

Fang, C., Fernie, A. R., & Luo, J. (2019). Exploring the diversity of plant metabolism. *Trends in Plant Science*, *24*(1), 83-98.

Fernandez, C., Monnier, Y., Santonja, M., Gallet, C., Weston, L. A., Prévosto, B., ... & Bousquet-Mélou, A. (2016). The impact of competition and allelopathy on the trade-off between plant defense and growth in two contrasting tree species. *Frontiers in Plant Science*, *7*, 594.

Fernandez-Conradi, P., Defossez, E., Delavallade, A., Descombes, P., Pitteloud, C., Glauser, G., ... & Rasmann, S. (2022). The effect of community-wide phytochemical diversity on herbivory reverses from low to high elevation. *Journal of Ecology*, *110*(1), 46-56.

Fernie, A. R., Trethewey, R. N., Krotzky, A. J., & Willmitzer, L. (2004). Metabolite profiling: from diagnostics to systems biology. *Nature reviews molecular cell biology*, *5*(9), 763-769.

Fiehn, O. (2002). Metabolomics—the link between genotypes and phenotypes. *Functional genomics*, 155-171.

Gargallo-Garriga, A., Sardans, J., Granda, V., Llusià, J., Peguero, G., Asensio, D., ... & Peñuelas, J. (2020). Different "metabolomic niches" of the highly diverse tree species of the French Guiana rainforests. *Scientific Reports*, *10*(1), 6937.

Gromski, P. S., Muhamadali, H., Ellis, D. I., Xu, Y., Correa, E., Turner, M. L., & Goodacre, R. (2015). A tutorial review: Metabolomics and partial least squares-discriminant analysis–a marriage of convenience or a shotgun wedding. *Analytica chimica acta*, *879*, 10-23.

Grubb, C. D., & Abel, S. (2006). Glucosinolate metabolism and its control. *Trends in plant science*, *11*(2), 89-100.

Hoehenwarter, W., van Dongen, J. T., Wienkoop, S., Steinfath, M., Hummel, J., Erban, A., ... & Weckwerth, W. (2008). A rapid approach for phenotype-screening and database independent detection of cSNP/protein polymorphism using mass accuracy precursor alignment. *Proteomics*, *8*(20), 4214-4225.

Hooper, D. U., Chapin III, F. S., Ewel, J. J., Hector, A., Inchausti, P., Lavorel, S., ... & Wardle, D. A. (2005). Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological monographs*, *75*(1), 3-35.

Kim, H. K., & Verpoorte, R. (2010). Sample preparation for plant metabolomics. *Phytochemical Analysis: An International Journal of Plant Chemical and Biochemical Techniques*, *21*(1), 4-13.

Lipka, A. E., Kandianis, C. B., Hudson, M. E., Yu, J., Drnevich, J., Bradbury, P. J., & Gore, M. A. (2015). From association to prediction: statistical methods for the dissection and selection of complex traits in plants. *Current Opinion in Plant Biology*, *24*, 110-118.

Mandal, S. M., Chakraborty, D., & Dey, S. (2010). Phenolic acids act as signaling molecules in plant-microbe symbioses. *Plant signaling & behavior*, *5*(4), 359-368.

Marr, S., Hageman, J. A., Wehrens, R., van Dam, N. M., Bruelheide, H., & Neumann, S. (2021). LC-MS based plant metabolic profiles of thirteen grassland species grown in diverse neighbourhoods. *Scientific data*, *8*(1), 52.

Massey, F. P., & Hartley, S. E. (2009). Physical defences wear you down: progressive and irreversible impacts of silica on insect herbivores. *Journal of Animal Ecology*, *78*(1), 281-291.

Nguyen, Q. T., Merlo, M. E., Medema, M. H., Jankevics, A., Breitling, R., & Takano, E. (2012). Metabolomics methods for the synthetic biology of secondary metabolism. *FEBS letters*, *586*(15), 2177-2183.

Obata, T., & Fernie, A. R. (2012). The use of metabolomics to dissect plant responses to abiotic stresses. *Cellular and Molecular Life Sciences*, *69*, 3225-3243.

Ober, D., & Hartmann, T. (2000). Phylogenetic origin of a secondary pathway: the case of pyrrolizidine alkaloids. *Plant molecular biology*, *44*, 445-450.

Peng, M., Shahzad, R., Gul, A., Subthain, H., Shen, S., Lei, L., ... & Luo, J. (2017). Differentially evolved glucosyltransferases determine natural variation of rice flavone accumulation and UV-tolerance. *Nature communications*, *8*(1), 1975.

Philbin, C. S., Dyer, L. A., Jeffrey, C. S., Glassmire, A. E., & Richards, L. A. (2022). Structural and compositional dimensions of phytochemical diversity in the genus Piper reflect distinct ecological modes of action. *Journal of Ecology*, *110*(1), 57-67.

Quinn, J. C., Kessell, A., & Weston, L. A. (2014). Secondary plant products causing photosensitization in grazing herbivores: Their structure, activity and regulation. *International Journal of Molecular Sciences*, *15*(1), 1441-1465.

Rai, A., Saito, K., & Yamazaki, M. (2017). Integrated omics analysis of specialized metabolism in medicinal plants. *The Plant Journal*, *90*(4), 764-787.

Richards, L. A., Dyer, L. A., Forister, M. L., Smilanich, A. M., Dodson, C. D., Leonard, M. D., & Jeffrey, C. S. (2015). Phytochemical diversity drives plant–insect community diversity. *Proceedings of the National Academy of Sciences*, *112*(35), 10973-10978.

Ristok, C., Poeschl, Y., Dudenhöffer, J. H., Ebeling, A., Eisenhauer, N., Vergara, F., ... & Weinhold, A. (2019). Plant species richness elicits changes in the metabolome of grassland species via soil biotic legacy. *Journal of Ecology*, *107*(5), 2240-2254.

Ristok, C., Weinhold, A., Ciobanu, M., Poeschl, Y., Roscher, C., Vergara, F., ... & van Dam, N. M. (2022). Plant diversity effects on herbivory are related to soil biodiversity and plant chemistry. *Journal of Ecology*.

Rochfort, S. (2005). Metabolomics reviewed: a new "omics" platform technology for systems biology and implications for natural products research. *Journal of natural products*, *68*(12), 1813-1820.

Sardans, J., Gargallo-Garriga, A., Urban, O., Klem, K., Walker, T. W., Holub, P., ... & Peñuelas, J. (2020). Ecometabolomics for a better understanding of plant responses and acclimation to abiotic factors linked to global change. *Metabolites*, *10*(6), 239.

Scherling, C., Roscher, C., Giavalisco, P., Schulze, E. D., & Weckwerth, W. (2010). Metabolomics unravel contrasting effects of biodiversity on the performance of individual plant species. *PLoS One*, *5*(9), e12569.

Schuman, M. C., & Baldwin, I. T. (2016). The layers of plant responses to insect herbivores. *Annual review of entomology*, *61*, 373-394.

Sulpice, R., & McKeown, P. C. (2015). Moving toward a comprehensive map of central plant metabolism. *Annual review of plant biology*, *66*, 187-210.

Tohge, T., Wendenburg, R., Ishihara, H., Nakabayashi, R., Watanabe, M., Sulpice, R., ... & Fernie, A. R. (2016). Characterization of a recently evolved flavonol-phenylacyltransferase gene provides signatures of natural light selection in Brassicaceae. *Nature Communications*, *7*(1), 12399.

Tugizimana, F., Piater, L., & Dubery, I. (2013). Plant metabolomics: A new frontier in phytochemical analysis. *South African Journal of Science*, *109*(5-6), 01-11.

Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A., & Dufresne, A. (2015). The importance of the microbiome of the plant holobiont. *New Phytologist*, *206*(4), 1196-1206.

Violle, C., Navas, M. L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E. (2007). Let the concept of trait be functional!. *Oikos*, *116*(5), 882-892.

Walker, T. W., Alexander, J. M., Allard, P. M., Baines, O., Baldy, V., Bardgett, R. D., ... & Salguero-Gómez, R. (2022). Functional Traits 2.0: The power of the metabolome for ecology. *Journal of Ecology*, *110*(1), 4-20.

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., ... & Bandeira, N. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature biotechnology*, *34*(8), 828-837.

Weston, P. A., Weston, L. A., & Hildebrand, S. (2013). Metabolic profiling in Echium plantagineum: presence of bioactive pyrrolizidine alkaloids and napthoquinones from accessions across southeastern Australia. *Phytochemistry reviews*, *12*, 831-837.

Weston, L. A., Skoneczny, D., Weston, P. A., & Weidenhamer, J. D. (2015). Metabolic profiling: An overview—New approaches for the detection and functional analysis of biologically active secondary plant products. *J. Allelochem. Interact*, *1*, 15-27.

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J. H. C., Diemer, M., Flexas, J., Garnier, E., Groom, P. K. Gulias, J., Hikosaka, K., Lamont, B. B., Lee, T., Lee, W., Lusk, C., Midgley, J. J., Navas, M.-L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V. I., Roumet, C., Thomas, S. C., Tjoelker, M. G., Veneklaas, E. J., Villar, R. (2004). The worldwide leaf economics spectrum. *Nature*, *428*(6985), 821-827.

Yang, L., Wen, K. S., Ruan, X., Zhao, Y. X., Wei, F., & Wang, Q. (2018). Response of plant secondary metabolites to environmental factors. *Molecules*, *23*(4), 762.

Yonekura-Sakakibara, K., Nakabayashi, R., Sugawara, S., Tohge, T., Ito, T., Koyanagi, M., ... & Saito, K. (2014). A flavonoid 3-O-glucoside: 2 ''-O-glucosyltransferase responsible for terminal modification of pollen-specific flavonols in A rabidopsis thaliana. *The Plant Journal*, *79*(5), 769-782.

Zanne, A. E., D. C. Tank, W. K. Cornwell, J. M. Eastman, S. A. Smith, R. G. FitzJohn, D. J. McGlinn, B. C. O'Meara, A. T. Moles, P. B. Reich, D. L. Royer, D. E. Soltis, P. F. Stevens, M. Westoby, I. J. Wright, L. Aarssen, R. I. Bertin, A. Calaminus, R. Govaerts, F. Hemmings, M. R. Leishman, J. Oleksyn, P. S. Soltis, N. G. Swenson, L. Warman, J. M. Beaulieu (2014). Three keys to the radiation of angiosperms into freezing environments. *Nature*, *506*(7486), 89-92.

APPENDIX

Content

APPENDIX – CHAPTER 1

Workflow LC-MS data processing and curation

From Field to Feature in Eco-metabolomics:

LC-MS based plant metabolic profiles of thirteen grassland species grown in diverse neighbourhoods

Marr, S., Hageman, J. A., Wehrens, R., van Dam, N. M., Bruelheide, H., Neumann, S.

December 12, 2020

Contents

List of Figures

1 Abstract

In plants, secondary metabolite profiles provide a unique opportunity to explore seasonal variation and responses to the environment. These include both abiotic and biotic factors. In field experiments, such stress factors occur in combination. This variation alters the plant metabolic profiles in yet uninvestigated ways. This data set contains trait and mass spectrometry data of thirteen grassland species collected at four time points in the growing season in 2017. We collected above-ground vegetative material of seven grass and six herb species that were grown in plant communities with different levels of diversity in the Jena Experiment. For each sample, we recorded visible traits and acquired shoot metabolic profiles on a UPLC-ESI-Qq-TOF-MS. We performed the raw data pre-processing in Galaxy-W4M and prepared the data for statistical analysis in R by applying missing data imputation, batch correction, and validity checks on the features (detailed tutorial is included). This comprehensive data set provides the opportunity to investigate environmental dynamics across diverse neighbourhoods that are reflected in the metabolomic profile.

Our workflow includes the sample preparation, LC-MS measurements, the pre-processing of raw spectra and the data preparation for statistical analysis. The complete workflow is published in Marr et al. [1]

2 Data Repository & Data Import

2.1 Data Repository

The metabolomics data was acquired on a liquid chromatography system (ACQUITY UPLC System, Waters Corporation, Milford, USA; LC) coupled with a mass spectrometer (ESI-micrOToF-Q-II, Bruker Daltonics, Bremen, Germany; MS). In addition to the 512 analytical samples, this study includes quality controls, i.e. blanks and a *Quality Control* (QC) that was pooled of all samples. All raw data files are available in *MTBLS679 "From Field to Feature in Ecometabolomics: LC-MS Based Metabolite Profiles of Thirteen Grassland Plant Species Reflecting Environmental Dynamics" on https://www.ebi.ac.uk/metabolights/MTBLS679*. Detailed descriptions of the experimental setup, sample preparation, quality controls and raw data acquisition can be found in the related publication Marr et al. [1]

2.2 Pre-processing in Galaxy-W4M

Vendor-specific raw data files (.d) were converted to an open file format (.mzML) using CompassXport (version 3.0.9), enabling the usage across different data analysis procedures in vendor-independent environments. In this study, we used the Galaxy-W4M web service [2] (based on XCMS 3.0) to pre-process the raw data spectra as provided in *MTBLS679* in an automated workflow (Figure 1). The workflow enables both the processing of large data sets composed of diverse LC-MS spectra and the resulting diverse data matrix (*https://doi.workflow4metabolomics.org/W4M00008*). For all samples, including blanks and the *QC*, features were picked, grouped, corrected for retention time shifts and grouped again. Adducts and isotopes of the measured features were annotated using CAMERA. A detailed description of the workflow, including the parameter settings, is available in the corresponding manuscript by Marr et al.[1].

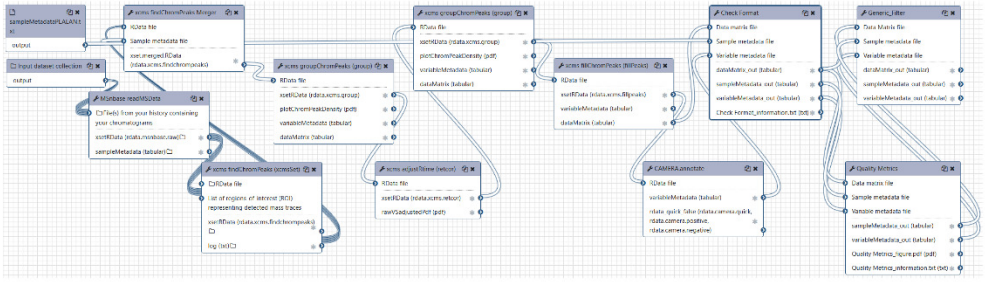


Figure 1: Galaxy-W4M workflow for raw LC-MS pre-processing. The pre-processing includes the following steps: peak picking, grouping, retention time correction, feature annotation, filtering for the region of interest. Parameter settings and tool versions are listed in Table 2 of the corresponding paper by Marr et al. [1]

2.3 Data Import

R Library

Load R libraries for data processing and visualisation. All libraries are referenced in section 8 References [5-14].

```
# ---- Processing ----
library(BatchCorrMetabolomics)
library(magrittr)
library(RUVSeq)
library(StatTools)
library(vegan)

# ---- Visualisation ----
library(FactoMineR)
library(factoextra)
library(gridExtra)
library(ggsci)
library(ggtext)
```

Import

We import the data matrix (*dataMatrix*), including pre-processed LC-MS data of secondary metabolites in the shoots of grassland plant species, and the sample metadata (*sampleMetadata*), including environmental conditions, plant traits and LC-MS data acquisition, as provided in MTBLS679 *https://www.ebi.ac.uk/metabolights/MTBLS679* and *https://doi.workflow4metabolomics.org/W4M00008*. Note that we will refer to the data matrix and the sample metadata as *dataMatrix* and *sampleMetadata*, respectively, throughout the tutorial; regardless of their state of processing.

```
# ---- Import Data from MTBLS679 ----
dataMatrix     <- data.frame(t(read.table(file   = "MTBLS679_dataMatrix.tsv",
                                          header = T, row.names = 1)))
sampleMetadata <- data.frame(read.table(file   = "MTBLS679_sampleMetadata.txt",
                                        header = T, row.names = 1, sep = "\t"))
```

```
# ---- Matrix Example ----
dataMatrix[50:53, 405:408]
```

```
                                    M176T732 M176T296 M176T262 M177T237
pos_012_2017_A_GERPRA_B081_b_2.A.5_01_8474    0.000    0.000    0.000     0.00
pos_012_2017_B_CENJAC_B073_a_2.B.8_01_8466 2750.832    0.000    0.000 12730.26
pos_012_2017_C_KNAARV_B073_a_2.D.3_01_8494    0.000 4118.352    0.000     0.00
pos_012_2017_D_RANACR_B071_b_2.E.6_01_8455    0.000    0.000 1001.952 28523.63
```

```
# ---- Metadata Example ----
sampleMetadata[50:53, c("injectionOrder", "batch", "Campaign", "desDiv")] #2, 4, 22
```

```
                                        injectionOrder batch Campaign desDiv
pos_012_2017_A_GERPRA_B081_b_2.A.5_01_8474           8474  pos2   2017_A      2
pos_012_2017_B_CENJAC_B073_a_2.B.8_01_8466           8466  pos2   2017_B      8
pos_012_2017_C_KNAARV_B073_a_2.D.3_01_8494           8494  pos2   2017_C      8
pos_012_2017_D_RANACR_B071_b_2.E.6_01_8455           8455  pos2   2017_D      2
```

For this example, we need only a few of the available factors in the sample metadata. Hence, we choose only relevant factors for processing and visualisation.

```
# ---- Select Relevant Metadata: Factor ----
l1 <- c("class", "sampleType", "batch", "FuncGroup", "SpecCode", "desDiv", "Campaign",
    "Season")

# ---- Select Relevant Metadata: Numeric ----
l2 <- c("injectionOrder")
```

```
# ---- Remove Remaining Variables From Metadata ----
delCol <- NULL
for (relVar in 1:ncol(sampleMetadata)) {
    lbl <- colnames(sampleMetadata)[relVar]

    if (lbl %in% l1) {
        idx1 <- which(colnames(sampleMetadata) == lbl)
        sampleMetadata[, idx1] <- as.factor(sampleMetadata[, idx1])
    } else {
        if (lbl %in% l2) {
            idx2 <- which(colnames(sampleMetadata) == lbl)
            sampleMetadata[, idx2] <- as.numeric(sampleMetadata[, idx2])
        } else {
            idx3 <- which(colnames(sampleMetadata) == lbl)
            delCol <- c(delCol, idx3)
        }
    }
}

sampleMetadata <- sampleMetadata[, -delCol]

# ---- Set Variable Types: Factor ----
w1 <- which(colnames(sampleMetadata) %in% intersect(l1, colnames(sampleMetadata)))

# ---- Set Variable Types: Numeric ----
w2 <- which(colnames(sampleMetadata) %in% intersect(l2, colnames(sampleMetadata)))
```

3 Missing Data Imputation

In this example, the simultaneous pre-processing of multiple species with diverse metabolomic profiles, cause the *dataMatrix* to incorporate a huge number of missing values. Those missing data can result from either variation in the instrument performances or actual biological reasons, such as species-specificity and adaptation to the environment. We prepare the *dataMatrix* for data cleaning and analysis by imputing all missing intensities, assuming that these features are independent of each other. We use random values picked from a normal distribution with a mean of 70 and an SD of 20. This range of values (70 ± 20) is instrument-specific and chosen in accordance with the prefilter for feature detection used in the pre-processing step (threshold were set to 100). For a detailed explanation also see section *Missing Data Imputation* in the corresponding manuscript by Marr et al. [1].

3.1 Missing Data Index

In order to preserve the missing intensities for the statistical analysis and investigation of the underlying biological question, we create an index for later reference, replace these zero values with NAs and create the *filterMatrix*.

```
# ---- Save Reference ----
metaPre  <- sampleMetadata
Xall     <- dataMatrix

# ---- Define Index for Missing Data ----
NAidx    <- Xall == 0

# ---- Replace Zeros with NAs ----
XallZero <- Xall
XallNA   <- Xall
XallNA[Xall == 0] <- NA
```

3.2 Filter Matrix

We use the *filterMatrix* to subset the *dataMatrix* species-wise. Only those features are assumed to be present in a species that were detected in at least 25% of the samples belonging to this species. The values are calculated as the proportional occurrence of each feature in a particular species. Thereby, the number of samples with measured intensities for each feature is counted in relation to the total number of samples within this species (e.g. 0 = not present in this species, 0.25 = present in 25% of all samples in the species, 1 = present in all samples of the species). Note that in the *filterMatrix* the *QC* is referred to as *pool*.

```
# ---- filterMatrix Calculation ----
uniqLvl      <- unique(metaPre$SpecCode) %>% as.character() %>% sort()
nSpecies     <- length(uniqLvl)
filterMatrix <- matrix(ncol=ncol(XallNA), nrow=nSpecies)

colnames(filterMatrix) <- colnames(XallNA)
rownames(filterMatrix) <- uniqLvl

for (s in 1:nSpecies) {
  sp               <- uniqLvl[s]
  Xsp              <- XallNA[metaPre$SpecCode == sp,]
  metaSp           <- metaPre[metaPre$SpecCode == sp,]
  filterMatrix[s,] <- apply(Xsp, 2, function(x) sum(!is.na(x)))/nrow(Xsp)
}

# ---- filterMatrix Example ----
round(filterMatrix[ ,1:5], digits = 4)
```

```
       M127T743 M131T466 M132T149 M133T212 M133T467
ANTODO   0.0312   0.0000   0.0000   0.0000   0.0000
AVEPUB   0.0312   0.0000   0.0000   0.0000   0.0000
blank    0.0000   0.0000   0.0000   0.0000   0.0000
CENJAC   0.2812   0.0000   0.0000   0.8125   0.0312
DACGLO   0.0312   0.0000   0.0000   0.0000   0.0000
FESRUB   0.0645   0.0000   0.0000   0.0000   0.0000
GERPRA   0.6875   0.0000   0.0312   0.0312   0.2188
HOLLAN   0.0000   0.0000   0.0000   0.0000   0.0312
KNAARV   0.0000   0.0000   0.0000   0.0000   0.0000
LEUVUL   0.5781   0.0000   0.0000   0.0000   0.5469
PHLPRA   0.0156   0.0000   0.0000   0.0000   0.0000
PLALAN   0.4062   0.8281   0.0000   0.0000   0.0000
POAPRA   0.1562   0.0000   0.0000   0.0000   0.0000
pool     0.1644   0.8082   0.0000   0.0000   0.0000
RANACR   0.0000   0.0000   0.6875   0.0000   0.0000
```

3.3 Imputation and Log Scaling

After creating the *filterMatrix*, we impute missing values with random absolute values. The imputation values are chosen accordingly to the threshold set while peak picking (intensity threshold = 100) in the data pre-processing to assure that all imputed values stay below the measured intensities. After the missing value imputation, the *dataMatrix* is scaled to log10.

```
# ---- Impute Missing Data ----
set.seed(1891)
impnum           <- abs(rnorm(sum(NAidx), mean = 70, sd = 20))
dMnoisImp        <- XallZero
dMnoisImp[NAidx] <- impnum

# ---- Log Scale ----
dMnoisImp        <- log10(dMnoisImp)
```

dataMatrix example: *before imputation: measured intensities and zero values*

	M136T92	M137T89	M137T156
pos_005_2017_B_POAPRA_B073_a_1.C.4_01_8385	0.0000	0.000	0.000
pos_005_2017_C_LEUVUL_B048_b_1.D.7_01_8396	0.0000	7001.498	0.000
pos_005_2017_D_GERPRA_C121_b_1.F.2_01_8405	0.0000	0.000	4854.355
pos_006_2017_A_RANACR_A016_b_1.B.2_01_8402	866.3226	0.000	0.000
pos_006_2017_B_RANACR_A016_b_1.C.5_01_8407	1839.2631	0.000	0.000

after imputation: measured intensities and random values below threshold

	M136T92	M137T89	M137T156
pos_005_2017_B_POAPRA_B073_a_1.C.4_01_8385	88.1253	64.4568	56.1964
pos_005_2017_C_LEUVUL_B048_b_1.D.7_01_8396	55.8652	7001.4980	92.4326
pos_005_2017_D_GERPRA_C121_b_1.F.2_01_8405	79.2178	62.1435	4854.3551
pos_006_2017_A_RANACR_A016_b_1.B.2_01_8402	866.3226	88.9791	63.4586
pos_006_2017_B_RANACR_A016_b_1.C.5_01_8407	1839.2631	53.5748	61.0608

log scaled: measured intensities and random values are log10 scaled

	M136T92	M137T89	M137T156
pos_005_2017_B_POAPRA_B073_a_1.C.4_01_8385	1.9451	1.8093	1.7497
pos_005_2017_C_LEUVUL_B048_b_1.D.7_01_8396	1.7471	3.8452	1.9658
pos_005_2017_D_GERPRA_C121_b_1.F.2_01_8405	1.8988	1.7934	3.6861
pos_006_2017_A_RANACR_A016_b_1.B.2_01_8402	2.9377	1.9493	1.8025
pos_006_2017_B_RANACR_A016_b_1.C.5_01_8407	3.2646	1.7290	1.7858

4 Batch Correction

Due to a large number of samples (total: 511), we split the LC-MS measurements into 12 analytical batches. Thereby, one batch included 44 samples (randomly picked from the 511 samples), one blank and several *QC* measurements. After running the batch, the MS system was cleaned and recalibrated. Besides the intensity shifts within the batches (intra-batch) caused by the slightly unstable performance of the instrument, this procedure led to further shifts across the batches (inter-batch). These intensity shifts are corrected for by calculating the introduced variance of the *QC* measurements based on principal component analysis (PCA) calculation[7]. A more detailed explanation can also be found in the corresponding manuscript by Marr et al. [1].

4.1 Preparation

First, we calculate the optimal number of components.

```
# ---- Define QC Measurements ----
idxQC                <- which(metaPre$class == "pool")
replicates.ind       <- matrix(-1, nrow(dMnoisImp) - length(idxQC) + 1, length(idxQC))
replicates.ind[1,]   <- idxQC
replicates.ind[-1,1] <- (1:nrow(dMnoisImp))[-idxQC]

# ---- Optimal Number of Components ----
allidbafterQC <- sapply(1:15, function(NumOfComp) {
                        sprintf("Doing Batch correction using %d number of components\n",
                                NumOfComp)

                        dMnoisImpBC <- t(RUVs(x     = t(dMnoisImp),
                                              cIdx  = 1:ncol(dMnoisImp),
                                              k     = NumOfComp,
                                              scIdx = replicates.ind,
                                              round = FALSE,
                                              isLog = TRUE)$normalizedCounts)

# ---- Metadata for Evaluation ----
                BatchCorrMetaVar    <- data.frame(SeqNr = metaPre$injectionOrder,
                                                  Batch = metaPre$batch,
                                                  SCode = metaPre$class)

                idx <- which(metaPre$sampleType == "pool")
                choiceDist_postBCQC <- evaluateCorrection(X    = dMnoisImpBC[idx,],
                                                          Y    = BatchCorrMetaVar[idx,],
                                                          what = "PCA",
                                                          plot = FALSE)

                        return(choiceDist_postBCQC)
                })
```

In order to choose the optimal number of components that we use for the correction, we compare the remaining inter-batch distances on the QC measurements (Figure 2).

```
# ---- Interbatch Distances vs Number of Components ----
BCdist       <- as.data.frame(allidbafterQC)
BCdist$nComp <- 1:15

compChoice   <- ggplot (BCdist, (aes(x = BCdist$nComp, y = BCdist$allidbafterQC))) +
                geom_line(size = 1) +
                geom_point(size = 6, shape = 18) +
                labs(x = "number of components", y = "inter-batch distance") +
                theme_bw()+
                theme(axis.text        = element_text(size = 25),
                      axis.title       = element_text(size = 30, face = "bold"),
                      legend.text      = element_text(size = 30),
                      panel.background = element_rect(fill = "white"),
                      panel.grid.major = element_line(colour = "grey50")
                      )
compChoice
```

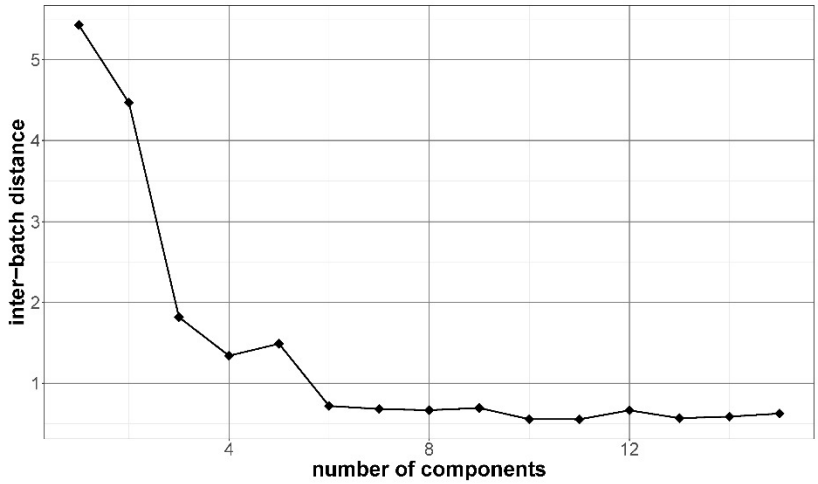


Figure 2: Number of principal components used for the batch correction keeping the inter-batch distance at a minimum.

4.2 Correction

In this example, the optimal number of components (*nComp*) is six, as the remaining inter-batch distance (*allidbafterQC*) decreases only slightly with more components.

```
print(BCdist)
```

```
   allidbafterQC nComp
1      5.4277283     1
2      4.4684877     2
3      1.8185237     3
4      1.3409498     4
5      1.4896308     5
6      0.7199777     6
7      0.6824285     7
8      0.6677630     8
9      0.6949238     9
10     0.5562601    10
11     0.5543296    11
12     0.6660080    12
13     0.5690315    13
14     0.5878562    14
15     0.6262616    15
```

Hence, we perform the batch correction with six components and save the corresponding metadata for the evaluation.

```
# ----- Batch Correction ----
NumOfComp <-  6

dMnoisImpBC <- t(RUVs(x  = t(dMnoisImp),
                 cIdx  = 1:ncol(dMnoisImp),
                 k     = NumOfComp,
                 scIdx = replicates.ind,
                 round = FALSE,
                 isLog = TRUE)$normalizedCounts)

# ---- Metadata for Evaluation ----
BatchCorrMetaVar <- data.frame(SeqNr = metaPre$injectionOrder,
                               Batch = metaPre$batch,
                               SCode = metaPre$class)
```

4.3 Evaluation

To evaluate the applied batch correction, we calculate the inter-batch distances before the correction and after the correction for all measurements (including quality controls) and the QC measurements separately (Figure 3).

```
# ---- Batch Correction Evaluation ----
Dist_preBCQC  <- evaluateCorrection(X    = dMnoisImp[idxQC,],
                                    Y    = BatchCorrMetaVar[idxQC,],
                                    what = "PCA",
                                    plot = FALSE)
Dist_postBCQC <- evaluateCorrection(X    = dMnoisImpBC[idxQC,],
                                    Y    = BatchCorrMetaVar[idxQC,],
                                    what = "PCA",
                                    plot = FALSE)
Dist_preBC    <- evaluateCorrection(X    = dMnoisImp,
                                    Y    = BatchCorrMetaVar,
                                    what = "PCA",
                                    plot = FALSE)
Dist_postBC   <- evaluateCorrection(X    = dMnoisImpBC,
                                    Y    = BatchCorrMetaVar,
                                    what = "PCA",
                                    plot = FALSE)
```

We used PCA to visualise the effects of the correction on all measurement and QC measurements separately.

```
# ---- Calculate PCA Before and After Correction: QC ----
idxQC <- which(metaPre$sampleType == "pool")

pca_Dist_preBCQC  <- FactoMineR::PCA(dMnoisImp[idxQC,],
                                     quali.sup  = w1,
                                     quanti.sup = w2,
                                     scale.unit = TRUE,
                                     ncp        = 6,
                                     graph      = FALSE)
pca_Dist_postBCQC <- FactoMineR::PCA(dMnoisImpBC[idxQC,],
                                     quali.sup  = w1,
                                     quanti.sup = w2,
                                     scale.unit = TRUE,
                                     ncp        = 6,
                                     graph      = FALSE)

# ---- Calculate PCA Before and After Correction: Samples and Quality Controls ----
pca_Dist_preBC    <- FactoMineR::PCA(dMnoisImp,
                                     quali.sup  = w1,
                                     quanti.sup = w2,
                                     scale.unit = TRUE,
                                     ncp        = 6,
                                     graph      = FALSE)
pca_Dist_postBC   <- FactoMineR::PCA(dMnoisImpBC,
                                     quali.sup  = w1,
                                     quanti.sup = w2,
                                     scale.unit = TRUE,
                                     ncp        = 6,
                                     graph      = FALSE)
```

```r
# ---- Plot PCA Before Correction: QC ----
idxQC <- which(metaPre$sampleType == "pool")

preBCQC <- fviz_pca_ind(pca_Dist_preBCQC,
                        axes        = c(1, 2),
                        geom        = c("point"),
                        addEllipses = F,
                        pointsize   = 6,
                        habillage   = metaPre[idxQC,]$batch,
                        mean.point  = FALSE,
                        title       = paste0("QC Distance before Correction: ",
                                             format(signif(Dist_preBCQC, 4), nsmall = 4))) +
          labs(x = paste0("PC 1 (",
                          format(round(pca_Dist_preBCQC$eig[1, 2], 4),
                                 nsmall = 4, digits = 4), "%)"),
               y = paste0("PC 2 (",
                          format(round(pca_Dist_preBCQC$eig[2, 2], 4),
                                 nsmall = 4, digits = 4), "%)"),
               tag = "a)") +
  scale_x_continuous(limits = c(-50, 100), breaks = seq(-100, 100, by = 40)) +
  scale_y_continuous(limits = c(-40,  40), breaks = seq(-100, 100, by = 40)) +
  scale_shape_manual(values = c(15, 16, 17, 18, 8, 25, 9, 10, 15, 16, 17, 18),
                     name   = "Batch") +
  scale_color_manual(values = c("#CC0C00FF", "#FFCD00FF", "#00AF66FF",
                                "#5C88DAFF", "#7C878EFF", "#84BD00FF",
                                "#00B5E2FF", "#CC0C00FF", "#FFCD00FF",
                                "#00AF66FF", "#5C88DAFF", "#7C878EFF"),
                     name   = "Batch") +
  theme_bw()+
    theme(axis.text         = element_text(size = 25),
          axis.title        = element_text(size = 30, face = "bold"),
          panel.background  = element_rect(fill = "white"),
          panel.grid.major  = element_line(colour = "grey50"),
          plot.tag          = element_text(size = 30),
          plot.title        = element_text(size = 25),
          legend.position   = "none",
          legend.title      = element_text(size = 30),
          legend.text       = element_text(size = 30)
          )
```

```r
# ---- Plot PCA After Correction: QC ----
postBCQC <- fviz_pca_ind(pca_Dist_postBCQC,
                         axes        = c(1, 2),
                         geom        = c("point"),
                         addEllipses = F,
                         pointsize   = 6,
                         habillage   = metaPre[idxQC, ]$batch,
                         mean.point  = FALSE,
                         title       = paste0("QC Distance after Correction: ",
                                              format(Dist_postBCQC, nsmall = 4, digits = 4))) +
            labs(x = paste0("PC 1 (",
                            format(round(pca_Dist_postBCQC$eig[1, 2], 4),
                                   nsmall = 4, digits = 4), "%)"),
                 y = paste0("PC 2 (",
                            format(round(pca_Dist_postBCQC$eig[2, 2], 4),
                                   nsmall = 4, digits = 4), "%)"),
                 tag  = "b)",
                 fill = "Batch") +
            scale_x_continuous(limits = c(-50, 100), breaks = seq(-100, 100, by = 40)) +
            scale_y_continuous(limits = c(-40,  40), breaks = seq(-100, 100, by = 40)) +
            scale_shape_manual(values = c(15, 16, 17, 18, 8, 25, 9, 10, 15, 16, 17, 18),
                               name   = "Batch") +
            scale_color_manual(values = c("#CC0C00FF", "#FFCD00FF", "#00AF66FF",
                                          "#5C88DAFF", "#7C878EFF", "#84BD00FF",
                                          "#00B5E2FF", "#CC0C00FF", "#FFCD00FF",
                                          "#00AF66FF", "#5C88DAFF", "#7C878EFF"),
                               name   = "Batch") +
            theme_bw() +
            theme(axis.text            = element_text(size = 25),
                  axis.title           = element_text(size = 30, face = "bold"),
                  panel.background     = element_rect(fill = "white"),
                  panel.grid.major     = element_line(colour = "grey50"),
                  plot.tag             = element_text(size = 30),
                  plot.title           = element_text(size = 25),
                  legend.direction     = "vertical",
                  legend.justification = c("top"),
                  legend.margin        = margin(8, 8, 8, 8),
                  legend.position      = c(0.85, 0.90),
                  legend.title         = element_text(size = 30),
                  legend.text          = element_text(size = 25)
                  )
```

```
# ---- Plot PCA Before Correction: Samples and Quality Controls ----
preBC <- fviz_pca_ind(pca_Dist_preBC,
                      axes        = c(1, 2),
                      geom        = c("point"),
                      addEllipses = FALSE,
                      habillage   = metaPre$batch,
                      mean.point  = FALSE,
                      pointsize   = 6,
                      title       = paste0("Distance before Correction: ",
                                           format(round(Dist_preBC, 4),
                                                  digits = 4, nsmall = 4))) +
        labs(x = paste0("PC 1 (",
                        format(round(pca_Dist_preBC$eig[1, 2], 4),
                               nsmall = 4, digits = 4),
                        "%)"),
             y = paste0("PC 2 (",
                        format(round(pca_Dist_preBC$eig[2, 2], 4),
                               nsmall = 4, digits = 4),
                        "%)"),
             tag = "c)") +
        scale_x_continuous(limits = c(-50, 100), breaks = seq(-100, 100, by = 40)) +
        scale_y_continuous(limits = c(-60,  60), breaks = seq(-100, 100, by = 40)) +
        scale_shape_manual(values = c(15, 16, 17, 18, 8, 25, 9, 10, 15, 16, 17, 18),
                           name   = "Batch") +
        scale_color_manual(values = c("#CC0C00FF", "#FFCD00FF", "#00AF66FF",
                                      "#5C88DAFF", "#7C878EFF", "#84BD00FF",
                                      "#00B5E2FF", "#CC0C00FF", "#FFCD00FF",
                                      "#00AF66FF", "#5C88DAFF", "#7C878EFF"),
                           name   = "Batch") +
        theme_bw() +
        theme(axis.text        = element_text(size = 25),
              axis.title       = element_text(size = 30, face = "bold"),
              panel.background = element_rect(fill = "white"),
              panel.grid.major = element_line(colour = "grey50"),
              plot.tag         = element_text(size = 30),
              plot.title       = element_text(size = 25),
              legend.position  = "none",
              legend.title     = element_text(size = 30),
              legend.text      = element_text(size = 30)
              )
```

```
# ---- Plot PCA After Correction: Samples and Quality Controls ----
postBC <- fviz_pca_ind(pca_Dist_postBC,
                       axes        = c(1, 2),
                       addEllipses = F,
                       geom        = c("point"),
                       pointsize   = 6,
                       habillage   = metaPre$batch,
                       mean.point  = FALSE,
                       title       = paste0("Distance after Correction: ",
                                            format(round(Dist_postBC, 4),
                                                   nsmall = 4, digits = 4))) +
         labs(x = paste0("PC 1 (",
                         format(round(pca_Dist_postBC$eig[1, 2], 4),
                                nsmall = 4, digits = 4),
                         "%)"),
              y = paste0("PC 2 (",
                         format(round(pca_Dist_postBC$eig[2, 2], 4),
                                nsmall = 4, digits = 4),
                         "%)"),
              tag = "d)") +
         scale_x_continuous(limits = c(-50, 100), breaks = seq(-100, 100, by = 40)) +
         scale_y_continuous(limits = c(-60,  60), breaks = seq(-100, 100, by = 40)) +
         scale_shape_manual(values = c(15, 16, 17, 18, 8, 25, 9, 10, 15, 16, 17, 18),
                            name   = "Batch") +
         scale_color_manual(values = c("#CC0C00FF", "#FFCD00FF", "#00AF66FF",
                                       "#5C88DAFF", "#7C878EFF", "#84BD00FF",
                                       "#00B5E2FF", "#CC0C00FF", "#FFCD00FF",
                                       "#00AF66FF", "#5C88DAFF", "#7C878EFF"),
                            name   = "Batch") +
         theme_bw() +
         theme(axis.text        = element_text(size = 25),
               axis.title       = element_text(size = 30, face = "bold"),
               panel.background = element_rect(fill = "white"),
               panel.grid.major = element_line(colour = "grey50"),
               plot.tag         = element_text(size = 30),
               plot.title       = element_text(size = 25),
               legend.position  = "none",
               legend.title     = element_text(size = 30),
               legend.text      = element_text(size = 30)
               )

# ---- Arrange Batch Correction Plots ----
# grid.arrange(preBC, postBC, preBCQC, postBCQC, nrow = 2, ncol = 2)
```

The inter-batch distances in the QC measurements before (16.8450445) and after (0.7199777) the batch correction show that the data set is subject to unwanted variations. However, looking at all measurements before (0.056425) and after (0.057719) the correction, the inter-batch distances do not change significantly, revealing that the variation between the samples is much higher than the variation in the QC measurements that were introduced by the batches. PCA plots visualise the effects of batch correction on both QC measurements and analytical samples.

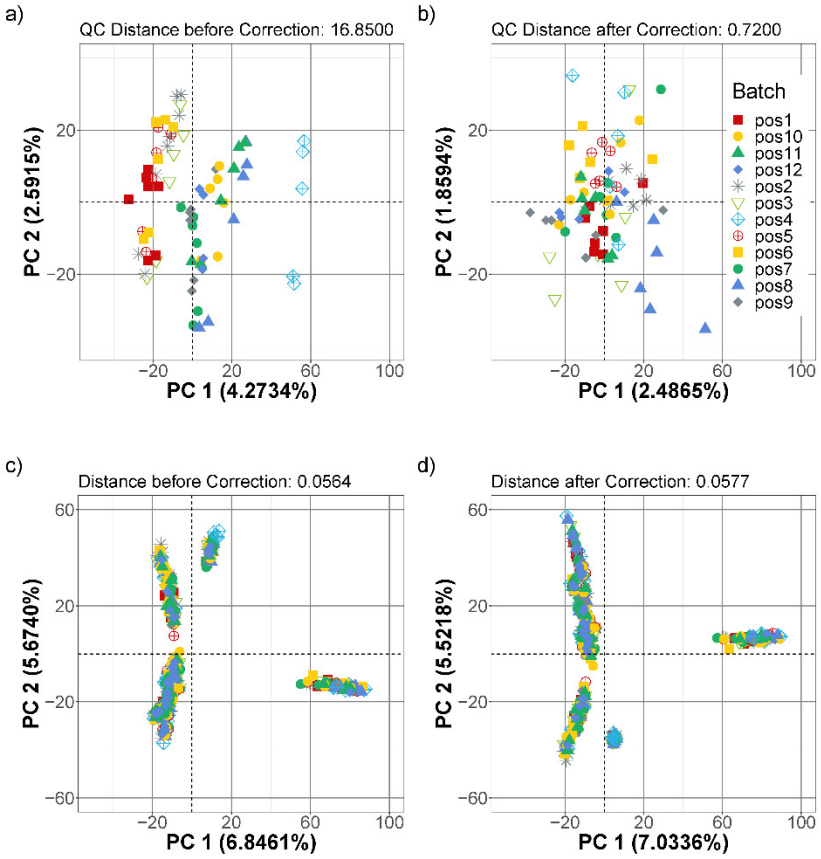


Figure 3: Inter-batch distances before and after batch correction. Distances are calculated across 12 analytical LC-MS batches. Distances before and after correction are compared for the QC measurements and for the sample measurements, respectively. The QC measurements show a high inter-batch distance (a) that is reduced significantly after correction (b) , while the distances detected for the analytical samples (c) only change slightly after correction (d). The proportion of the total explained variance is shown for each component in brackets.

5 Validity Check

5.1 Feature Validity

Blank Removal

All features were checked for their validity. We only kept those features in the *dataMatrix* that were not detected in the blanks, as these features are likely to have been introduced unintendedly through extraction or while measurements. Therefore, all features detected in blanks were removed from the *dataMatrix* and excluded from further analysis.

```
# ---- Keep Processed Data ----
dMnoisImpSBC  <- dMnoisImpBC
dMnoisImpS    <- dMnoisImp
meta          <- metaPre
NAidxS        <- NAidx

# ---- Remove Blank-Features ----
mzrtnames     <- names(which(filterMatrix["blank", ] >0))
tempColnames  <- setdiff(colnames(dMnoisImpSBC), mzrtnames)
NAidxS        <- NAidxS[ , tempColnames]
filterMatrixS <- filterMatrix[ , tempColnames]
dMnoisImpSBC  <- dMnoisImpSBC[ , tempColnames]
dMnoisImpS    <- dMnoisImpS[ ,  tempColnames]
```

QC Removal

In this example, we used the quality controls *QC* and blanks for batch correction and feature validity checks, respectively. These quality controls are not needed for further analysis and are, therefore, removed from both the *dataMatrix* and *sampleMetadata*.

```
# ---- Remove QC Measurements ----
idxQC         <- meta$SpecCode != "pool"

meta          <- meta[idxQC, ]
NAidxS        <- NAidxS[idxQC, ]
dMnoisImpSBC  <- dMnoisImpSBC[idxQC, ]
dMnoisImpS    <- dMnoisImpS[idxQC, ]

# ---- Remove Blanks ----
idxB          <- meta$SpecCode != "blank"
meta          <- meta[idxB, ]
NAidxS        <- NAidxS[idxB, ]
dMnoisImpSBC  <- dMnoisImpSBC[idxB, ]
dMnoisImpS    <- dMnoisImpS[idxB, ]

# ---- Drop Unused Levels ----
meta[ , w1]    <- data.frame(lapply(w1, function(ii){
                    base::droplevels(meta[ , ii])
                  }))
```

5.2 Sample Validity

As the last step of processing, we also checked the analytical samples for their validity. In our example, the data set is derived from a field experiment that is incorporating a high level of variance on both the inter-species and intra-species level. We, therefore, test all samples within their respective species for feature composition similarity, ensuring they are not contaminated or corrupted. A sample is considered valid when its Mahalanobis distance to its species cluster does not exceed the distance of the 3-fold mean (Figure 4) and when it shares at least 0.25 of features with the other samples in its respective species. Therefore, we calculate the Mahalanobis distances for all species subsets separately.

```
# ---- Mahalanobis Distances ----
lbls             <- levels(meta$SpecCode)
OutlierDist_list <- list()

for (lbl in lbls) {
  idxSpec            <- meta$SpecCode == lbl
  zeroBCdata         <- dMnoisImpSBC
  zeroBCdata[NAidxS] <- 0
  idxF               <- colSums(zeroBCdata[idxSpec, ], na.rm = TRUE) != 0
  sdat               <- zeroBCdata[idxSpec, idxF]
  idxFS              <- which( apply(sdat, 2,
                                     function(i) { sum( i != 0) /length(i)}) >= 0.25)
  sdatMin            <- sdat[, idxFS]
  fShared            <- specnumber(sdatMin)/ncol(sdatMin)
  share              <- 0.25
  idxfShared         <- fShared < share
  pcamdl             <- prcomp(sdat, center = TRUE, scale.= TRUE)
  sdat               <- pcamdl$x[ , 1:20]
  myLOOMDdist        <- sapply(1:nrow(sdat), function(i) {
                          mahalanobis(sdat[i, ], colMeans(sdat[-i, ], na.rm = TRUE),
                                      cov(sdat[-i, ]))
                        })
  maxDist            <- 3 * mean(myLOOMDdist)
  idxLOOMD           <- myLOOMDdist > maxDist


  OD                 <- matrix(myLOOMDdist,
                               nrow = length(myLOOMDdist),1,
                               dimnames = list(rownames(sdat), c("Mdist")))
  OutlierDist        <- merge(sdat, OD, by = "row.names")

  OutlierDist_list[[lbl]] <- OutlierDist

}

OutlierDist_all  <- do.call(rbind, OutlierDist_list)
```

We plot all samples indicating the distance to their respective species.

```
# ---- Set Variables ----
SampleNames <- as.factor(OutlierDist_all$Row.names)
maxDist       <- 3 * mean(OutlierDist_all$Mdist)
```

```r
# ---- Plot Mahalanobis Distances for All Samples ----
MD_allSamples <- ggplot(data       = OutlierDist_all,
                        aes(x      = SampleNames,
                            y      = Mdist, #OutlierDist_all$Mdist,
                            colour = meta$SpecCode,
                            shape  = meta$SpecCode
                            )) +
                 geom_point(size = 6) +
                 labs(x = paste0("samples"),
                      y = paste0("mahalanobis distance")) +
                 theme_bw() +
                 theme(axis.text        = element_text(size = 25),
                       axis.title       = element_text(size = 30, face = "bold"),
                       axis.text.x      = element_blank(),
                       axis.ticks.x     = element_blank(),
                       legend.text      = element_markdown(size = 30),
                       legend.position  = "right",
                       legend.title     = element_text(size = 30),
                       panel.background = element_rect(fill = "white"),
                       panel.grid.major = element_blank(),
                       panel.grid       = element_blank()
                       ) +
                 scale_color_manual(name   = "Species",
                                    values = c("#DC0000FF", "#7E6148FF", "#3C5488FF",
                                               "#3C5488FF", "#F39B7FFF", "#4DBBD5FF",
                                               "#8491B4FF", "#8491B4FF", "#F39B7FFF",
                                               "#91D1C2FF", "#00A087FF", "#B09C85FF",
                                               "#E64B35FF"),
                                    labels = c("*Anthoxanthum odoratum*",
                                               "*Avenula pubescens*",
                                               "*Centaurea jacea*",
                                               "*Dactylis glomerata*",
                                               "*Festuca rubra*",
                                               "*Geranium pratense*",
                                               "*Holcus lanatus*",
                                               "*Knautia arvensis*",
                                               "*Leucanthemum vulgare*",
                                               "*Phleum pratense*",
                                               "*Plantago lanceolata*",
                                               "*Poa pratensis*",
                                               "*Ranunculus acris*")) +
                 scale_shape_manual(name   = "Species",
                                    values = c(8, 10, 16, 24, 8, 9, 17,
                                               25, 18, 15, 18, 15, 9),
                                    labels = c("*Anthoxanthum odoratum*",
                                               "*Avenula pubescens*",
                                               "*Centaurea jacea*",
                                               "*Dactylis glomerata*",
                                               "*Festuca rubra*",
                                               "*Geranium pratense*",
                                               "*Holcus lanatus*",
                                               "*Knautia arvensis*",
                                               "*Leucanthemum vulgare*",
                                               "*Phleum pratense*",
                                               "*Plantago lanceolata*",
                                               "*Poa pratensis*",
                                               "*Ranunculus acris*")) +
                 geom_abline(intercept = maxDist, color = "red", size = 1.5)

MD_allSamples
```

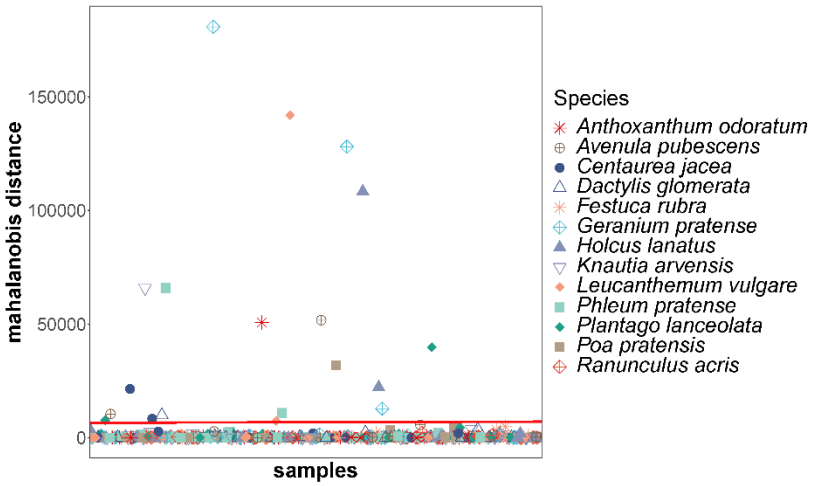


Figure 4: Sample validity check. Calculating the Mahalanobis distances for each species separately. Thresholds (3-fold mean distance within species) are calculated species specific. The indicated threshold (red line) represente the 3-fold distance of the total mean and is included for demonstration purposes only. Only samples above their respective threshold and < 0.25 shared features with their respective species are excluded from further analysis.

Mahalanobis distances and the corresponding feature composition is calculated for all samples within their species separately.

```
# ---- Species Specific Validity Check ----
metaOut   <- meta
corrSamp  <- unlist(sapply(levels(meta$SpecCode), function(lbl) {
  idxSpec            <- meta$SpecCode == lbl

  zeroBCdata         <- dMnoisImpSBC
  zeroBCdata[NAidxS] <- 0


  # ---- Species Subset ----
  idxF        <- colSums(zeroBCdata[idxSpec,], na.rm = TRUE)!= 0
  sdat        <- zeroBCdata[idxSpec, idxF]

  idxFS       <- which( apply(sdat, 2, function(i) { sum( i != 0) /length(i)}) >= 0.25)
  sdatMin     <- sdat[, idxFS]

  # ---- Feature Composition ----
  fShared     <- specnumber(sdatMin)/ncol(sdatMin)
  share       <- 0.25
  idxfShared  <- fShared < share

  pcamdl      <- prcomp(sdat, center = TRUE, scale.=TRUE)
  sdat        <- pcamdl$x[,1:20]

  # ---- Mahalanobis Distances ----
  myLOOMDdist <- sapply(1:nrow(sdat), function(i) {
                   mahalanobis(sdat[i,], colMeans(sdat[-i,], na.rm = TRUE),  cov(sdat[-i,]))
  })

  # ---- Threshold Distance ----
  maxDist     <- 3 * mean(myLOOMDdist)
  idxLOOMD    <- myLOOMDdist > maxDist

  return(rownames(sdat)[idxLOOMD & idxfShared])

}))
```

In this data set, 12 samples did not pass the validity check. We, therefore, excluded them from the *dataMatrix* and further analysis.

```
 [1] "pos_066_2017_C_ANTODO_B060_b_1.A.6_01_9086"
 [2] "pos_078_2017_B_ANTODO_A018_b_1.A.5_01_9085"
 [3] "pos_036_2017_A_HOLLAN_A035_b_2.A.7_01_9037"
 [4] "pos_110_2017_D_LEUVUL_B051_a_1.F.8_01_8877"
 [5] "pos_116_2017_D_LEUVUL_B051_b_1.F.3_01_8966"
 [6] "pos_018_2017_D_PHLPRA_B073_b_2.F.4_01_8481"
 [7] "pos_049_2017_D_PHLPRA_A018_b_1.F.2_01_8529"
 [8] "pos_053_2017_D_PHLPRA_B059_b_1.F.6_01_8559"
 [9] "pos_070_2017_D_PHLPRA_A018_a_1.F.1_01_8688"
[10] "pos_016_2017_D_POAPRA_A026_b_2.F.2_01_8484"
[11] "pos_057_2017_D_POAPRA_B073_a_2.E.7_01_8626"
[12] "pos_073_2017_D_POAPRA_B073_b_1.F.4_01_8682"
```

```
# ---- Remove Samples ----
idxVfyd         <- which(!(rownames(dMnoisImpSBC) %in% corrSamp))
dMnoisImpSBCOut <- dMnoisImpSBC[idxVfyd, ]
metaOut         <- metaOut[idxVfyd, ]
```

6 Processed Data Output

6.1 Data Saving

After performing the validity checks on the data, we save all processed data matrices.

```
# ---- Save processed data matrices ----
prcdSampleMetadata <- metaOut
prcdDataMatrixImputedLog <- dMnoisImpSBCOut
prcdDataMatrixImputed <- 10^prcdDataMatrixImputedLog

prcdDataMatrixZero <- prcdDataMatrixImputed
NAidxPrcd <- NAidxS[rownames(prcdDataMatrixImputed), colnames(prcdDataMatrixImputed)]
prcdDataMatrixZero[NAidxPrcd] <- 0


# combined meta data and data matrix
prcdMetadataMatrix <- cbind(prcdSampleMetadata, prcdDataMatrixZero, make.row.names = TRUE)
```

6.2 Filter Matrix Re-calculation

In order to apply further statistics to the data, we also need to recalculate the *filterMatrix* with the processed data matrix (*prcdDataMatrixZero*) and the corresponding sample metadata (*prcdSampleMetadata*).

```
       M127T743 M131T466 M132T149 M133T212 M133T467
ANTODO   0.0333   0.0000   0.0000   0.0000   0.0000
AVEPUB   0.0312   0.0000   0.0000   0.0000   0.0000
CENJAC   0.2812   0.0000   0.0000   0.8125   0.0312
DACGLO   0.0312   0.0000   0.0000   0.0000   0.0000
FESRUB   0.0645   0.0000   0.0000   0.0000   0.0000
GERPRA   0.6875   0.0000   0.0312   0.0312   0.2188
HOLLAN   0.0000   0.0000   0.0000   0.0000   0.0323
KNAARV   0.0000   0.0000   0.0000   0.0000   0.0000
LEUVUL   0.5968   0.0000   0.0000   0.0000   0.5645
PHLPRA   0.0167   0.0000   0.0000   0.0000   0.0000
PLALAN   0.4062   0.8281   0.0000   0.0000   0.0000
POAPRA   0.1724   0.0000   0.0000   0.0000   0.0000
RANACR   0.0000   0.0000   0.6875   0.0000   0.0000
```

7 Visualisation

Here, we use PCA to visualise the processed data. The functional-group-wise calculation shows good species separation for both *grass* and *herb* species (Figure 5). For a more detailed explanation see also the corresponding manuscript by Marr et al. [1]

```
# ---- Calculate PCA for All Samples ----
pcaGlobal <- FactoMineR::PCA(dMnoisImpSBCOut,
                             quali.sup  = w1,
                             quanti.sup = w2,
                             scale.unit = TRUE,
                             ncp        = 6,
                             graph      = F)

# ---- Calculate PCA Functional Group Specific: Grass ----
idxGrass <- subset(dMnoisImpSBCOut,
                   metaOut$FuncGroup == "grass")
pcaGrass <- FactoMineR::PCA(idxGrass,
```

```
                            quali.sup  = w1,
                            quanti.sup = w2,
                            scale.unit = TRUE,
                            ncp        = 6,
                            graph      = F)

# ---- Calculate PCA Functional Group Specific: Herb ----
idxHerb <- subset(dMnoisImpSBCOut,
                  metaOut$FuncGroup == "herb")
pcaHerb <- FactoMineR::PCA(idxHerb,
                           quali.sup   = w1,
                           quanti.sup  = w2,
                           scale.unit  = TRUE,
                           ncp         = 6,
                           graph       = F)
```

```
# ---- Plot All Samples ----
pcaGlobalplot <- fviz_pca_ind(pcaGlobal,
                              axes        = c(1,2),
                              addEllipses = FALSE,
                              geom        = c("point"),
                              pointsize   = 6,
                              habillage   = metaOut$SpecCode,
                              mean.point  = FALSE,
                              title       = paste0("Functional Group: Grass & Herb")) +
                labs(x   = paste0("PC 1 (",
                                  format(round(pcaGlobal$eig[1,2], 4),
                                         nsmall = 4, digits = 4), "%)"),
                     y   = paste0("PC 2 (",
                                  format(round(pcaGlobal$eig[2,2], 4),
                                         nsmall = 4, digits = 4), "%)"),
                     tag = "a)") +
                theme_bw() +
                theme(axis.text        = element_text(size = 25),
                      axis.title       = element_text(size = 30, face = "bold"),
                      legend.position  = "bottom",
                      legend.text      = element_text(size = 30),
                      legend.title     = element_text(size = 30, colour = "white"),
                      panel.background = element_rect(fill = "white"),
                      panel.grid.major = element_line(colour = "grey50"),
                      plot.tag         = element_text(size = 30),
                      plot.title       = element_text(size = 25)
                      ) +
                scale_x_continuous(limits = c(-80,  90), breaks = seq(-100, 100, by = 40)) +
                scale_y_continuous(limits = c(-50, 100), breaks = seq(-100, 100, by = 40)) +
                scale_shape_manual(values = c(8, 10, 16, 24, 8, 9, 17,
                                              25, 18, 15, 18, 15, 9)) +
                scale_color_manual(values = c("#DC0000FF", "#7E6148FF", "#3C5488FF",
                                              "#3C5488FF", "#F39B7FFF", "#4DBBD5FF",
                                              "#8491B4FF", "#8491B4FF", "#F39B7FFF",
                                              "#91D1C2FF", "#00A087FF", "#B09C85FF",
                                              "#E64B35FF"))
```

```
# ---- Plot Functional Group Specific: Grass ----
idxGrass     <- subset(dMnoisImpSBCOut, metaOut$FuncGroup == "grass")
idxGrassMeta <- subset(metaOut, metaOut$FuncGroup == "grass")

pcaGrassplot <- fviz_pca_ind(pcaGrass,
                             axes        = c(1,2),
                             addEllipses = FALSE,
                             geom        = c("point"),
                             pointsize   = 6,
                             habillage   = idxGrassMeta$SpecCode,
                             mean.point  = FALSE,
                             title       = paste0("Functional Group: Grass")) +
                labs(x   = paste0("PC 1 (",
                                  format(round(pcaGrass$eig[1,2], 4), nsmall = 4, digits = 4),
                                  "%)"),
                     y   = paste0("PC 2 (",
                                  format(round(pcaGrass$eig[2,2], 4), nsmall = 4, digits = 4),
                                  "%)"),
                     tag = "b)") +
                theme_bw()+
                theme(axis.text        = element_text(size = 25),
                      axis.title       = element_text(size = 30, face = "bold"),
                      legend.position  = "none",
                      panel.background = element_rect(fill = "white"),
                      panel.grid.major = element_line(colour = "grey50"),
                      plot.tag         = element_text(size = 30),
                      plot.title       = element_text(size = 25)
                      ) +
                scale_x_continuous(limits = c(-80,  90), breaks = seq(-100, 100, by = 40)) +
                scale_y_continuous(limits = c(-50, 100), breaks = seq(-100, 100, by = 40)) +
                scale_shape_manual(values = c(8, 10, 24, 8, 17, 15, 15)) +
                scale_color_manual(values = c("#DC0000FF", "#7E6148FF", "#3C5488FF",
                                              "#F39B7FFF", "#8491B4FF", "#91D1C2FF",
                                              "#B09C85FF"))
```

```
# ---- Plot Functional Group Specific: Herb ----
idxHerb     <- subset(dMnoisImpSBCOut, metaOut$FuncGroup == "herb")
idxHerbMeta <- subset(metaOut, metaOut$FuncGroup == "herb")

pcaHerbplot <- fviz_pca_ind(pcaHerb,
                            axes        = c(1,2),
                            addEllipses = FALSE,
                            geom        = c("point"),
                            pointsize   = 6,
                            habillage   = idxHerbMeta$SpecCode,
                            mean.point  = FALSE,
                            title       = paste0("Functional Group: Herb")) +
               labs(x   = paste0("PC 1 (",
                                 format(round(pcaHerb$eig[1,2], 4), nsmall = 4, digits = 4),
                                 "%)"),
                    y   = paste0("PC 2 (",
                                 format(round(pcaHerb$eig[2,2], 4), nsmall = 4, digits = 4),
                                 "%)"),
                    tag = "c)") +
               theme_bw()+
               theme(axis.text        = element_text(size = 25),
                     axis.title       = element_text(size = 30, face = "bold"),
                     legend.position  = "none",
                     panel.background = element_rect(fill = "white"),
                     panel.grid.major = element_line(colour = "grey50"),
                     plot.tag         = element_text(size = 30),
                     plot.title       = element_text(size = 25)
                     ) +
               scale_x_continuous(limits = c(-80,  90), breaks = seq(-100, 100, by = 40)) +
               scale_y_continuous(limits = c(-50, 100), breaks = seq(-100, 100, by = 40)) +
               scale_shape_manual(values = c(16, 9, 25, 18, 18, 9)) +
               scale_color_manual(values = c("#3C5488FF", "#4DBBD5FF", "#8491B4FF",
                                             "#F39B7FFF", "#00A087FF", "#E64B35FF"))

# ---- Arrange PCA Plots ----

# pL <- list(pcaGlobalplot, pcaGrassplot, pcaHerbplot)
# grid.arrange(
#     grobs = pL,
#     widths = c(2, 2, 1),
#     layout_matrix = rbind(c(1, 1),
#                           c(2, 3))
# )
```

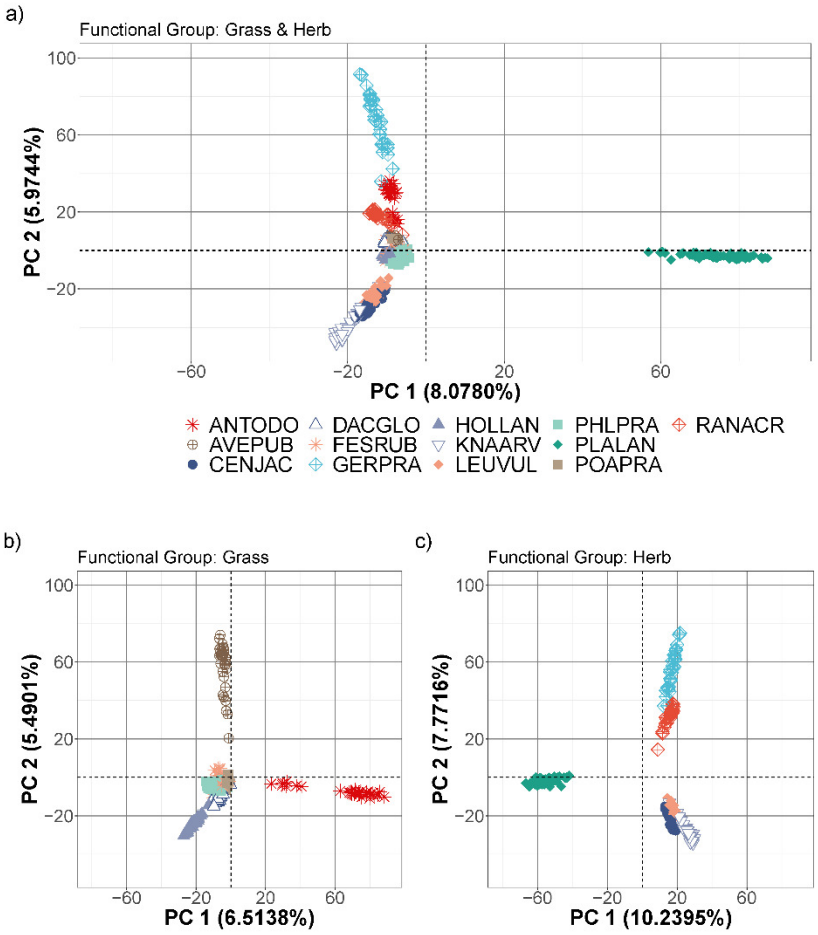


Figure 5: Sample selection for statistical analysis. The separation of the samples by principle components shows that samples belonging to the same species are grouped in clusters (a). Samples separated by functional groups *grass* (b) and *herb* (c), are still found to group in their respective species clusters. We used species abbreviations in the plot's key for clarity reasons: *Anthoxanthum odoratum* (ANTODO), *Avenula pubescens* (AVEPUB), *Centaurea jacea* (CENJAC), *Dactylis glomerata* (DACGLO), *Festuca rubra* (FESRUB), *Geranium pratense* (GERPRA), *Holcus lanatus* (HOLLAN), *Knautia arvensis* (KNAARV), *Leucanthemum vulgare* (LEUVUL), *Phleum pratense* (PHLPRA), *Plantago lanceolata* (PLALAN), *Poa pratensis* (POAPRA), *Ranunculus acris* (RANACR).

8 References

1 Marr, S., Hageman, J. A., Wehrens, R., van Dam, N. M., Bruelheide, H. & Neumann, S. LC-MS based plant metabolic profiles of thirteen grassland species grown in diverse neighbourhoods. *Sci. Data* (supplemental material).

2 Giacomoni, F., Le Corguillé, G., Monsoor, M., Landi, M., Pericard, P., Pétéra, M., Duperier, C., Tremblay-Franco, M., Martin, J.-F., Jacob, D., Goulitquer, S., Thévenot, E. A. & Caron, C. Workflow4Metabolomics: a collaborative research infrastructure for computational metabolomics. *Bioinformatics* **31** (9), 1493-1495 (2014).

3 Allaire, J. J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., Chang, W. & Iannone, R. rmarkdown: Dynamic Documents for R. R package version 2.0. URL https://rmarkdown.rstudio.com (2019).

4 Xie, Y., Allaire, J. J. & Grolemund, G. R Markdown: The Definitive Guide. URL https://bookdown.org/yihui/rmarkdown (Chapman and Hall, 2018).

5 Wehrens, R., Mumm, R., Keurentjes, J. & de Vos, R. BatchCorrMetabolomics: Complementary package to "Improved Batch Correction in Untargeted MS-Based Metabolomics". R package version 0.1.14. (2020).

6 Bache, S. M. & Wickham, H. magrittr: A Forward-Pipe Operator for R. R package version 1.5. https://CRAN.R-project.org/package=magrittr (2014).

7 Risso, D., Ngai, J., Speed, T. P. & Dudoit, S. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat. Biotechnol.* **32** (9), 896-902 (2014).

8 Brunius, C. Nifty Tools for Random Statistical Tasks. R package version 0.0.914. (2020).

9 Oksanen, J., Guillaume Blanchet, F. , Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E. & Wagner, H. vegan: Community Ecology Package. R package version 2.5-6. https://CRAN.R-project.org/package=vegan (2019).

10 Kassambara, A. & Mundt, F. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.6. https://CRAN.R-project.org/package=factoextra (2019).

11 Le, S., Josse, J. & Husson, F. FactoMineR: An R Package for Multivariate Analysis. *J. Stat. Softw.* **25** (1), 1-18 (2008).

12 Xiao, N. ggsci: Scientific Journal and Sci-Fi Themed Color Palettes for 'ggplot2'. R package version 2.9. https://CRAN.R-project.org/package=ggsci (2018).

13 Wilke, C. O. ggtext: Improved text rendering support for 'ggplot2'. R package version 0.1.0. https://wilkelab.org/ggtext (2020).

14 Auguie, B. gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3. https://CRAN.R-project.org/package=gridExtra (2017).

15 Microsoft Corporation & Weston, S. doParallel: Foreach Parallel Adaptor for the 'parallel' Package. R package version 1.0.15. https://CRAN.R-project.org/package=doParallel (2019).

Session Info:

```
R version 4.0.3 (2020-10-10)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 18363)

Matrix products: default

locale:
[1] LC_COLLATE=English_World.1252      LC_CTYPE=English_World.1252
[3] LC_MONETARY=English_World.1252     LC_NUMERIC=C
[5] LC_TIME=English_United States.1252

attached base packages:
[1] stats4    parallel  stats     graphics  grDevices utils     datasets
[8] methods   base

other attached packages:
 [1] doParallel_1.0.16           iterators_1.0.13
 [3] foreach_1.5.1               ggtext_0.1.0
 [5] ggsci_2.9                   gridExtra_2.3
 [7] factoextra_1.0.7            ggplot2_3.3.2
 [9] FactoMineR_2.3              vegan_2.5-7
[11] lattice_0.20-41             permute_0.9-5
[13] StatTools_0.0.915           RUVSeq_1.24.0
[15] edgeR_3.32.0                limma_3.46.0
[17] EDASeq_2.24.0               ShortRead_1.48.0
[19] GenomicAlignments_1.26.0    SummarizedExperiment_1.20.0
[21] MatrixGenerics_1.2.0        matrixStats_0.57.0
[23] Rsamtools_2.6.0             GenomicRanges_1.42.0
[25] GenomeInfoDb_1.26.1         Biostrings_2.58.0
[27] XVector_0.30.0              IRanges_2.24.0
[29] S4Vectors_0.28.0            BiocParallel_1.24.1
[31] Biobase_2.50.0              BiocGenerics_0.36.0
[33] magrittr_2.0.1              BatchCorrMetabolomics_0.1.14
[35] knitr_1.30

loaded via a namespace (and not attached):
  [1] readxl_1.3.1            backports_1.2.0         aroma.light_3.20.0
  [4] BiocFileCache_1.14.0    splines_4.0.3           digest_0.6.27
  [7] htmltools_0.5.0         memoise_1.1.0           cluster_2.1.0
 [10] openxlsx_4.2.3          R.utils_2.10.1          sandwich_3.0-0
 [13] askpass_1.1             prettyunits_1.1.1       jpeg_0.1-8.1
 [16] colorspace_2.0-0        blob_1.2.1              rappdirs_0.3.1
 [19] ggrepel_0.8.2           haven_2.3.1             xfun_0.19
 [22] dplyr_1.0.2             crayon_1.3.4            RCurl_1.98-1.2
 [25] kohonen_3.0.10          zoo_1.8-8               glue_1.4.2
 [28] gtable_0.3.0            zlibbioc_1.36.0         DelayedArray_0.16.0
 [31] car_3.0-10              kernlab_0.9-29          prabclus_2.3-2
 [34] DEoptimR_1.0-8          abind_1.4-5             scales_1.1.1
 [37] DBI_1.1.0               rstatix_0.6.0           Rcpp_1.0.5
 [40] progress_1.2.2          gridtext_0.1.3          crch_1.0-4
 [43] scoringRules_1.0.1      foreign_0.8-80          flashClust_1.01-2
 [46] bit_4.0.4               mclust_5.4.7            Formula_1.2-4
 [49] httr_1.4.2              RColorBrewer_1.1-2      fpc_2.2-8
 [52] modeltools_0.2-23       ellipsis_0.3.1          pkgconfig_2.0.3
 [55] XML_3.99-0.5            R.methodsS3_1.8.1       flexmix_2.3-17
 [58] farver_2.0.3            nnet_7.3-14             dbplyr_2.0.0
 [61] locfit_1.5-9.4          tidyselect_1.1.0        labeling_0.4.2
 [64] rlang_0.4.9             AnnotationDbi_1.52.0    cellranger_1.1.0
 [67] munsell_0.5.0           tools_4.0.3             generics_0.1.0
```

```
 [70] RSQLite_2.2.1           pls_2.7-3              broom_0.7.2
 [73] evaluate_0.14           stringr_1.4.0          yaml_2.2.1
 [76] bit64_4.0.5             zip_2.1.1              robustbase_0.93-6
 [79] ChemometricsWithR_0.2.0 purrr_0.3.4            nlme_3.1-149
 [82] formatR_1.7             R.oo_1.24.0            leaps_3.1
 [85] xml2_1.3.2              biomaRt_2.46.0         compiler_4.0.3
 [88] curl_4.3                png_0.1-7              ggsignif_0.6.0
 [91] tibble_3.0.4            stringi_1.5.3          GenomicFeatures_1.42.1
 [94] forcats_0.5.0           Matrix_1.2-18          markdown_1.1
 [97] vctrs_0.3.5             pillar_1.4.7           lifecycle_0.2.0
[100] data.table_1.13.2       bitops_1.0-6           rtracklayer_1.49.5
[103] R6_2.5.0                latticeExtra_0.6-29    hwriter_1.3.2
[106] rio_0.5.16              codetools_0.2-16       MASS_7.3-53
[109] assertthat_0.2.1        openssl_1.4.3          withr_2.3.0
[112] GenomeInfoDbData_1.2.4  diptest_0.75-7         mgcv_1.8-33
[115] hms_0.5.3               grid_4.0.3             tidyr_1.1.2
[118] class_7.3-17            rmarkdown_2.5          carData_3.0-4
[121] ggpubr_0.4.0            scatterplot3d_0.3-41
```

APPENDIX – CHAPTER 2

Table 1: Comprehensive model evaluation (eval.), including model accuracies (ac), settings for the best model (bm), the area under the receiver-operator curve (AUC), and Kappa value (Kpp) for the data sets MTBLS679, MTBLS1224, and MTBLS2140. In MTBLS679 and MTBLS1224, the levels season and diversity were analysed for each species separately. bm were chosen on model specific parameters: For Partial Least Squares Discriminant Analysis (PLS), the number of components (ncomp); for Support Vector Machines (SVM) with linear kernel (l) cost (c), with radial kernel (r) sigma (s) and c, with poly kernel (p) degree (d), scale (sc) and c; and for Random Forests (RF) the number of variables that is randomly collected to be sampled at each split time (mtry).

Data set			MTBLS679														
			PLS			*SVMl*			*SVMr*			*SVMp*			*RF*		
Level	species	eval.	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog
FG		**ac**	0.999	0.995	1.000	0.998	0.998	1.000	0.992	0.992	1.000	0.998	0.998	1.000	1.000	1.000	1.000
									s = 7.016e-05,	s = 7.309e-05,	s = 5.467e-05,	d = 1, sc =	d = 1, sc = 0.01, c	d = 1, sc = 0.001,			
		bm	ncomp = 11	ncomp = 16	ncomp = 2	c = 1	c = 1	c = 1	c = 16	c = 8	c = 4	0.01, c = 0.25	= 0.25	c = 0.25	mtry = 15	mtry = 148	mtry = 22
		AUC	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.968	0.998	0.998
		Kpp	0.999	0.990	1.000	0.996	0.996	1.000	0.984	0.984	0.999	0.996	0.996	1.000	1.000	1.000	1.000
Species		**ac**	0.967	0.888	0.971	0.957	0.957	0.962	0.949	0.948	0.968	0.960	0.957	0.965	0.968	0.968	0.968
									s = 7.033e-05,	s = 6.470e-05,	s = 5.907e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 21	ncomp = 30	ncomp = 29	c = 1	c = 1	c = 1	c = 2	c = 2	c = 1	0.001, c = 0.25	c = 2	c = 0.25	mtry = 7	mtry = 81	mtry = 209
		AUC	0.995	0.989	0.996	0.910	0.910	0.918	0.827	0.836	0.859	0.910	0.911	0.919	0.997	0.965	0.996
		Kpp	0.964	0.877	0.968	0.953	0.953	0.958	0.944	0.943	0.965	0.956	0.953	0.962	0.965	0.996	0.965
Season	ANTODO	**ac**	0.848	0.601	0.842	0.791	0.605	0.664	0.245	0.385	0.564	0.794	0.588	0.628	0.812	0.795	0.785
									s = 2.385e-04,	s = 4.842e-05,	s = 4.917e-05,	d = 2, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 3	ncomp = 8	ncomp = 3	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	0.001, c = 0.5	c = 0.25	c = 0.25	mtry = 148	mtry = 2147	mtry = 3030
		AUC	0.942	0.858	0.976	0.929	0.806	0.898	0.500	0.809	0.859	0.931	0.848	0.856	0.924	0.882	0.881
		Kpp	0.787	0.481	0.761	0.710	0.458	0.545	0.000	0.241	0.431	0.719	0.451	0.511	0.716	0.694	0.689
Season	AVEPUB	**ac**	0.888	0.575	0.902	0.843	0.693	0.852	0.239	0.258	0.747	0.859	0.689	0.874	0.818	0.764	0.770
									s = 1.896e-04,	s = 4.908e-05,	s = 4.915e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 8	ncomp = 7	ncomp = 5	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	0.01, c = 0.25	c = 0.25	c = 0.25	mtry = 24	mtry = 48	mtry = 383
		AUC	0.966	0.815	0.969	0.950	0.802	0.937	0.500	0.722	0.885	0.822	0.829	0.922	0.945	0.906	0.906
		Kpp	0.838	0.416	0.854	0.786	0.587	0.800	0.000	0.121	0.677	0.619	0.580	0.826	0.742	0.672	0.678
Season	DACGLO	**ac**	0.826	0.626	0.818	0.682	0.534	0.666	0.277	0.296	0.613	0.774	0.573	0.668	0.753	0.662	0.653
									s = 2.760e-04,	s = 5.082e-05,	s = 5.003e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 3	ncomp = 7	ncomp = 4	c = 1	c = 1	c = 1	c = 1	c = 2	c = 2	0.001, c = 1	c = 0.25	c = 0.25	mtry = 136	mtry = 271	mtry = 192
		AUC	0.927	0.837	0.937	0.879	0.816	0.865	0.501	0.703	0.862	0.884	0.835	0.853	0.872	0.809	0.809
		Kpp	0.749	0.491	0.747	0.566	0.372	0.557	0.036	0.115	0.500	0.679	0.424	0.577	0.663	0.541	0.540
Season	FESRUB	**ac**	0.758	0.575	0.806	0.816	0.708	0.628	0.263	0.265	0.642	0.817	0.738	0.661	0.774	0.718	0.716
										s = 4.817e-05,	s = 5.151e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 9	ncomp = 9	ncomp = 5	c = 1	c = 1	c = 1	s = 1.514, c = 1	c = 2	c = 2	0.01, c = 0.25	c = 0.25	c = 0.25	mtry = 249	mtry =7819	mtry = 9290
		AUC	0.942	0.743	0.952	0.882	0.826	0.846	0.500	0.595	0.855	0.874	0.796	0.858	0.900	0.841	0.840
		Kpp	0.659	0.409	0.706	0.742	0.603	0.503	0.000	0.021	0.457	0.746	0.624	0.526	0.690	0.621	0.622
Season	HOLLAN	**ac**	0.829	0.538	0.962	0.807	0.716	0.846	0.199	0.317	0.689	0.787	0.702	0.832	0.881	0.803	0.818
									s = 2.474e-04,	s = 5.062e-05,	s = 5.008e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 10	ncomp = 10	ncomp = 6	c = 1	c = 1	c = 1	c = 1	c = 2	c = 2	0.01, c = 0.25	c = 0.25	c = 0.25	mtry = 52	mtry = 57	mtry = 74
		AUC	0.971	0.798	0.994	0.934	0.881		0.500	0.708		0.913	0.887		0.955	0.918	0.918
		Kpp	0.763	0.353	0.944	0.709	0.609	0.781	0.000	0.137	0.599	0.688	0.606	0.768	0.826	0.724	0.735
Season	PHLPRA	**ac**	0.878	0.624	0.830	0.860	0.768	0.746	0.265	0.340	0.714	0.864	0.760	0.761	0.860	0.824	0.824
									s = 2.076e-04,	s = 5.074e-05,	s = 5.043e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 4	ncomp = 9	ncomp = 10	c = 1	c = 1	c = 1	c = 1	c = 2	c = 4	0.001, c = 0.5	c = 0.25	c = 0.25	mtry = 1077	mtry = 589	mtry = 496
		AUC	0.955	0.834	0.955	0.935	0.865	0.871	0.500	0.690	0.856	0.924	0.847	0.868	0.949	0.912	0.912
		Kpp	0.835	0.498	0.770	0.810	0.687	0.662	0.000	0.112	0.617	0.815	0.675	0.679	0.809	0.761	0.762
Season	POAPRA	**ac**	0.789	0.833	0.908	0.801	0.711	0.710	0.263	0.431	0.648	0.793	0.743	0.720	0.823	0.693	0.689
									s = 2.742e-04,	s = 4.938e-05,	s = 4.895e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 7	ncomp = 9	ncomp = 4	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	0.01, c = 0.25	c = 0.25	c = 0.25	mtry = 228	mtry = 176	mtry = 136
		AUC	0.931	0.957	0.950	0.857	0.869	0.900	0.480	0.501	0.863	0.878	0.862	0.888	0.934	0.899	0.901
		Kpp	0.708	0.774	0.858	0.703	0.598	0.593	0.000	0.211	0.539	0.683	0.658	0.595	0.755	0.587	0.578
Season	CENJAC	**ac**	0.696	0.633	0.678	0.661	0.588	0.601	0.224	0.258	0.537	0.686	0.535	0.596	0.708	0.653	0.645
							c = 1	c = 1	s = 2.227e-04,	s = 4.936e-05,	s = 4.857e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			mtry = 1395
		bm	ncomp = 3	ncomp = 4	ncomp = 8	c = 1			c = 2	c = 2	c = 2	0.01, c = 0.25	c = 0.25	c = 0.25	mtry = 3600	mtry = 52	
		AUC	0.851	0.784	0.891	0.767	0.694	0.767	0.500	0.273	0.718	0.786	0.684	0.764	0.844	0.796	0.796
		Kpp	0.565	0.493	0.552	0.519	0.439	0.477	0.000	0.138	0.433	0.553	0.367	0.470	0.594	0.534	0.516
Season	GERPRA	**ac**	0.928	0.671	0.973	0.926	0.720	0.872	0.248	0.435	0.778	0.941	0.732	0.880	0.918	0.853	0.856
									s = 2.478e-04,	s = 5.028e-05,	s = 5.050e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 7	ncomp = 10	ncomp = 6	c = 1	c = 1	c = 1	c = 4	c = 2	c = 2	0.01, c = 0.25	c = 0.25	c = 0.25	mtry = 8523	mtry = 6038	mtry = 10125
		AUC	0.990	0.865	1.000	0.947	0.848	0.961	0.521	0.798	0.913	0.954	0.830	0.948	0.964	0.936	0.936
		Kpp	0.899	0.540	0.961	0.905	0.614	0.818	0.000	0.325	0.713	0.917	0.642	0.831	0.889	0.799	0.802
Season	KNAARV	**ac**	0.818	0.543	0.900	0.773	0.699	0.750	0.243	0.310	0.666	0.793	0.696	0.719	0.780	0.707	0.713
									s = 2.149e-04,	s = 4.843e-05,	s = 4.856e-05,	d = 2, sc =	d = 1, sc = 0.001,	d= 1, sc = 0.001,			
		bm	ncomp = 7	ncomp = 10	ncomp = 10	c = 1	c = 1	c = 1	c = 4	c = 2	c = 2	0.001, c = 0.5	c = 0.25	c = 0.25	mtry = 161	mtry = 6581	mtry = 1658
		AUC	0.963	0.739	0.979	0.863	0.820	0.869	0.500	0.277	0.839	0.877	0.826	0.889	0.916	0.870	0.871
		Kpp	0.743	0.348	0.861	0.684	0.591	0.671	0.000	0.136	0.553	0.708	0.587	0.626	0.699	0.604	0.610
Season	LEUVUL	**ac**	0.757	0.492	0.793	0.763	0.668	0.767	0.292	0.427	0.715	0.783	0.669	0.752	0.873	0.883	0.885
									s = 1.873e-04,	s = 4.992e-05,	s = 4.871e-05,	d= 2, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 10	ncomp = 9	ncomp = 10	c = 1	c = 1	c = 1	c = 4	c = 2	c = 2	0.001, c = 0.5	c = 0.25	c = 0.25	mtry = 10125	mtry = 4663	mtry = 5540
		AUC	0.912	0.719	0.936	0.922	0.829	0.902	0.511	0.731	0.904	0.919	0.829	0.898	0.942	0.920	0.920
		Kpp	0.674	0.319	0.719	0.680	0.554	0.686	0.046	0.236	0.619	0.707	0.560	0.665	0.828	0.841	0.844
Season	PLALAN	**ac**	0.810	0.560	0.863	0.772	0.708	0.743	0.258	0.531	0.690	0.786	0.722	0.716	0.813	0.786	0.792
									s = 2.093e-04,	s = 4.938e-05,	s = 4.915e-05,	d = 2, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 5	ncomp = 8	ncomp = 8	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	0.001, c = 0.25	c = 0.25	c = 0.25	mtry = 161	mtry = 832	mtry = 496
		AUC	0.921	0.813	0.964	0.880	0.844	0.882	0.531	0.802	0.863	0.888	0.842	0.867	0.922	0.892	0.892
		Kpp	0.743	0.406	0.814	0.692	0.604	0.655	0.013	0.384	0.592	0.712	0.627	0.624	0.747	0.711	0.719
Season	RANACR	**ac**	0.855	0.785	0.885	0.813	0.542	0.631	0.248	0.352	0.596	0.844	0.553	0.636	0.735	0.668	0.671
									s = 2.392e-04,	s = 4.915e-05,	s = 4.885e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 5	ncomp = 8	ncomp = 6	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	0.001, c = 1	c = 0.25	c = 0.25	mtry = 37	mtry = 7174	mtry = 5540
		AUC	0.945	0.932	0.966	0.925	0.769	0.804	0.500	0.662	0.755	0.933	0.804	0.805	0.882	0.842	0.843
		Kpp	0.790	0.785	0.840	0.732	0.382	0.491	0.000	0.148	0.483	0.776	0.385	0.487	0.625	0.553	0.549

Table 1: continued – MTBLS679

			PLS			*SVMl*			*SVMr*			*SVMp*			*RF*	
Level	species	eval.	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog
Diversity	ANTODO	ac	0.163	0.211	0.103	0.146	0.115	0.059	0.180	0.103	0.080	0.259	0.012	0.192	0.093	0.198	0.198
									s = 2.346e-04,	s = 4.943e-05,	s = 4.864e-05,	d = 3, sc = 0.1,	d = 3, sc = 0.01, c	d = 3, sc = 0.01, c			
		bm	ncomp = 3	ncomp = 1	ncomp = 9	c = 1	c = 1	c = 1	c = 2	c = 0.25	c = 2	c = 0.25	= 0.25	= 0.25	mtry = 3	mtry = 10125	mtry = 10125
		AUC	0.332	0.445	0.316	0.550	0.600	0.692	0.500	0.514	0.668	0.629	0.655	0.727	0.240	0.240	0.240
		Kpp	-0.061	-0.017	-0.178	-0.095	-0.135	-0.167	0.000	-0.007	-0.099	-0.006	0.012	-0.013	-0.122	-0.021	-0.021
Diversity	AVEPUB	ac	0.291	0.279	0.288	0.291	0.167	0.156	0.242	0.219	0.127	0.238	0.227	0.148	0.280	0.179	0.179
									s = 1.987e-04,	s = 5.033e-05,	s = 4.938e-05,	d = 1, sc =	d = 3, sc = 0.1, c	d = 3, sc = 0.1, c			
		bm	ncomp = 7	ncomp = 6	ncomp = 9	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	0.01, c = 0.25	= 0.25	= 0.25	mtry = 10125	mtry = 455	mtry = 455
		AUC	0.546	0.536	0.576	0.535	0.525	0.569	0.500	0.345	0.640	0.665	0.570	0.561	0.405	0.404	0.404
		Kpp	0.095	0.050	0.054	0.041	-0.043	-0.012	0.000	0.005	0.021	0.055	0.000	-0.004	0.094	0.040	0.040
Diversity	DACGLO	ac	0.396	0.303	0.243	0.377	0.204	0.103	0.220	0.173	0.123	0.414	0.182	0.139	0.313	0.372	0.372
										s = 4.802e-05,	s = 4.890e-05,	d = 2, sc =	d = 1, sc = 0.001,	d = 2, sc = 0.001,			
		bm	ncomp = 10	ncomp = 8	ncomp = 1	c = 1	c = 1	c = 1	s = 2.848, c = 2	c = 2	c = 2	0.001, c = 0.5	c = 0.25	c = 0.25	mtry = 1077	mtry = 8523	mtry = 8523
		AUC	0.659	0.468	0.545	0.614	0.551	0.539	0.498	0.412	0.502	0.650	0.576	0.535	0.548	0.548	0.548
		Kpp	0.175	0.071	0.035	0.176	-0.034	-0.072	-0.009	-0.013	-0.049	0.210	-0.040	0.048	0.103	0.201	0.201
Diversity	FESRUB	ac	0.366	0.308	0.396	0.338	0.358	0.223	0.264	0.247	0.183	0.456	0.363	0.215	0.311	0.222	0.222
									s = 1.815e-04,	s = 5.176e-05,	s = 5.048e-05,	d = 2, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 8	ncomp = 5	ncomp = 7	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	0.01, c = 0.25	c = 0.25	c = 0.25	mtry = 10125	mtry = 7819	mtry = 7819
		AUC	0.673	0.598	0.710	0.615	0.473	0.564	0.500	0.461	0.445	0.619	0.555	0.545	0.612	0.439	0.439
		Kpp	0.147	0.087	0.176	0.108	0.140	0.014	0.000	0.016	0.012	0.270	0.158	0.049	0.087	0.037	0.037
Diversity	HOLLAN	ac	0.471	0.313	0.341	0.377	0.294	0.216	0.171	0.103	0.134	0.367	0.264	0.212	0.294	0.254	0.254
									s = 2.036e-04,	s = 4.956e-05,	s = 5.013e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 7	ncomp = 4	ncomp = 7	c = 1	c = 1	c = 1	c = 4	c = 2	c = 2	0.01, c = 0.25	c = 0.25	c = 0.25	mtry = 2551	mtry = 2551	mtry = 3303
		AUC	0.584	0.597	0.562	0.463	0.507	0.477	0.500	0.446	0.461	0.513	0.434	0.503	0.514	0.514	0.465
		Kpp	0.276	0.112	0.132	0.146	0.104	0.029	0.000	-0.013	-0.006	0.159	0.064	0.007	0.066	0.066	0.080
Diversity	PHLPRA	ac	0.525	0.449	0.536	0.506	0.393	0.427	0.225	0.260	0.436	0.507	0.421	0.428	0.462	0.392	0.392
									s = 1.580e-04,	s = 5.031e-05,	s = 5.115e-05,	d = 1, sc =	d = 2, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 10	ncomp = 10	ncomp = 9	c = 1	c = 1	c = 1	c = 0.25	c = 1	c = 4	0.001, c = 0.5	c = 0.25	c = 0.25	mtry = 11	mtry = 176	mtry = 176
		AUC	0.735	0.695	0.715	0.708	0.580	0.579	0.500	0.637	0.642	0.702	0.585	0.621	0.652	0.568	0.568
		Kpp	0.358	0.263	0.374	0.335	0.180	0.237	0.000	0.072	0.260	0.338	0.228	0.236	0.285	0.199	0.199
Diversity	POAPRA	ac	0.310	0.253	0.249	0.351	0.378	0.304	0.219	0.298	0.294	0.396	0.401	0.346	0.269	0.465	0.465
									s = 2.552e-04,	s = 5.057e-05,	s = 4.909e-05,	d = 3, sc = 1, c	d = 1, sc = 0.001,	d = 2, sc = 0.001,			
		bm	ncomp = 10	ncomp = 9	ncomp = 10	c = 1	c = 1	c = 1	c = 4	c = 2	c = 2	= 0.25	c = 0.25	c = 0.25	mtry = 8523	mtry = 8523	mtry = 8523
		AUC	0.576	0.530	0.616	0.409	0.401	0.471	0.512	0.410	0.460	0.493	0.441	0.490	0.507	0.477	0.477
		Kpp	0.089960727	0.023073147	0.028571429	0.1146744	0.1966071	0.1103013	0	0.07888408	0.1268409	0.1980685	0.1731203	0.11585665	0.058916424	0.306938432	0.306938432
Diversity	CENJAC	ac	0.423	0.345	0.362	0.342	0.293	0.179	0.252	0.286	0.187	0.350	0.273	0.195	0.418	0.317	0.317
									s = 1.984e-04,	s = 4.999e-05,	s = 5.030e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 6	ncomp = 2	ncomp = 8	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	0.01, c = 0.25	c = 0.25	c = 0.25	mtry = 4278	mtry = 1077	mtry = 1077
		AUC	0.662	0.578	0.598	0.729	0.702	0.668	0.500	0.496	0.527	0.725	0.569	0.703	0.597	0.523	0.523
		Kpp	0.18579932	0.147426647	0.14357143	0.1435185	0.0777417	-0.03380952	0	0.07335257	0.03757816	0.140833333	0.0639881	-0.01220238	0.23386905	0.15315476	0.15315476
Diversity	GERPRA	ac	0.293	0.253	0.338	0.260	0.128	0.143	0.242	0.221	0.123	0.332	0.144	0.143	0.212	0.159	0.159
									s = 2.963e-04,	s = 5.010e-05,	s = 5.164e-05,	d = 2, sc = 0.1,	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 10	ncomp = 1	ncomp = 4	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	c = 0.25	c = 0.25	c = 0.25	mtry = 18	mtry = 4	mtry = 4
		AUC	0.561	0.461	0.472	0.518	0.515	0.516	0.479	0.459	0.551	0.522	0.532	0.575	0.428	0.341	0.341
		Kpp	0.073	0.016	0.070	0.019	-0.073	-0.030	0.000	0.041	0.006	0.120	-0.062	-0.032	-0.009	0.029	0.029
Diversity	KNAARV	ac	0.372	0.352	0.275	0.292	0.220	0.144	0.250	0.150	0.091	0.330	0.213	0.160	0.293	0.134	0.134
									s = 1.965e-04,	s = 5.090e-05,	s = 5.025e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 4	ncomp = 1	ncomp = 10	c = 1	c = 1	c = 1	c = 4	c = 2	c = 2	0.01, c = 0.25	c = 0.25	c = 0.25	mtry = 114	mtry = 17	mtry = 17
		AUC	0.659	0.572	0.534	0.587	0.508	0.487	0.500	0.521	0.471	0.625	0.580	0.476	0.529	0.290	0.290
		Kpp	0.150	0.169	0.008	0.078	-0.025	-0.053	0.000	0.030	-0.055	0.119	-0.010	-0.057	0.094	-0.034	-0.034
Diversity	LEUVUL	ac	0.377	0.278	0.370	0.354	0.244	0.249	0.252	0.214	0.236	0.357	0.273	0.264	0.423	0.379	0.379
									s = 2.361e-04,	s = 5.008e-05,	s = 4.987e-05,	d = 1, sc =	d = 2, sc = 0.1, c	d = 1, sc = 0.001,			
		bm	ncomp = 3	ncomp = 7	ncomp = 10	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	0.01, c = 0.25	= 0.25	c = 0.25	mtry = 125	mtry = 7174	mtry = 7174
		AUC	0.593	0.531	0.634	0.641	0.527	0.578	0.493	0.497	0.599	0.677	0.544	0.590	0.605	0.577	0.577
		Kpp	0.170	0.037	0.164	0.138	0.003	0.013	0.010	-0.035	0.002	0.144	0.066	0.035	0.233	0.175	0.175
Diversity	PLALAN	ac	0.404	0.360	0.380	0.409	0.302	0.313	0.277	0.309	0.281	0.395	0.286	0.287	0.373	0.317	0.317
									s = 2.253e-04,	s = 4.929e-05,	s = 4.935e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 6	ncomp = 5	ncomp = 2	c = 1	c = 1	c = 1	c = 4	c = 2	c = 2	0.01, c = 0.25	c = 0.25	c = 0.25	mtry = 5540	mtry = 148	mtry = 148
		AUC	0.724	0.671	0.655	0.676	0.639	0.631	0.507	0.632	0.592	0.481	0.640	0.597	0.629	0.570	0.570
		Kpp	0.205	0.150	0.187	0.216	0.079	0.100	0.036	0.089	0.066	0.189	0.054	0.063	0.163	0.107	0.107
Diversity	RANACR	ac	0.278	0.314	0.317	0.275	0.171	0.094	0.182	0.104	0.088	0.283	0.153	0.105	0.313	0.192	0.192
									s = 2.778e-04,	s = 4.972e-05,	s = 4.903e-05,	d = 1, sc =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 8	ncomp = 1	ncomp = 9	c = 1	c = 1	c = 1	c = 2	c = 2	c = 0.25	0.001, c = 2	c = 0.25	c = 0.25	mtry = 7819	mtry = 3924	mtry = 3924
		AUC	0.573	0.635	0.636	0.453	0.570	0.507	0.500	0.646	0.555	0.228	0.551	0.465	0.521	0.397	0.397
		Kpp	0.031	0.111	0.137	0.050	-0.054	-0.097	0.000	-0.034	-0.013	0.015	-0.075	-0.088	0.108	0.023	0.023

Table 1: continued – MTBLS1224

Data set			MTBLS1224														
			PLS			*SVMl*			*SVMr*			*SVMp*			*RF*		
Level	species	eval.	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog
FG		**ac**	1.000	1.000	1.000	1.000	0.997	1.000	0.987	0.989	1.000	1.000	0.999	1.000	1.000	1.000	1.000
									s = 1.122e-04,	s = 1.081e-04,	s = 1.101e-04,	d = 1, sc = 0.001,	d = 3, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 1	ncomp = 3	ncomp = 2	c = 1	c = 1	c = 1	c = 1	c = 1	c = 0.25	c = 0.25	c = 0.25	c = 0.25	mtry = 2	mtry = 2	mtry = 2
		AUC	1.000	1.000	1.000	1.000	0.997	1.000	1.000	1.000	1.000	1.000	0.996	1.000	1.000	0.991	0.991
		Kpp	1.000	1.000	1.000	1.000	0.993	1.000	0.975	0.977	1.000	1.000	0.998	1.000	1.000	1.000	1.000
Species		**ac**	0.971	0.970	0.974	0.969	0.960	0.968	0.958	0.958	0.970	0.970	0.964	0.970	0.971	0.971	0.971
									s = 1.154e-04,	s = 1.205e-04,	s = 1.241e-04,	d = 1, sc = 0.001,	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 15	ncomp = 28	ncomp = 20	c = 1	c = 1	c = 1	c = 1	c = 1	c = 1	c = 0.5	c = 0.25	c = 0.25	mtry = 26	mtry = 4	mtry = 9
		AUC	0.996	0.996	0.996	0.883	0.882	0.905	0.790	0.781	0.796	0.884	0.882	0.911	0.999	0.998	0.998
		Kpp	0.969	0.968	0.972	0.966	0.957	0.965	0.954	0.954	0.967	0.967	0.961	0.967	0.968	0.968	0.968
Season	ANTODO	**ac**	0.785	0.645	0.770	0.798	0.263	0.548	0.281	0.284	0.451	0.802	0.281	0.583	0.764	0.622	0.633
									s = 3.448e-04,	s = 8.280e-05,	s = 8.376e-05,	d = 1, sc = 0.01, c	d = 3, sc = 0.1, c	d = 1, sc = 0.001,			
		bm	ncomp = 8	ncomp = 10	ncomp = 6	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	= 0.25	= 0.25	c = 0.25	mtry = 3541	mtry = 1462	mtry = 603
		AUC	0.931	0.911	0.974	0.922	0.516	0.718	0.512	0.522	0.652	0.909	0.535	0.688	0.913	0.831	0.832
		Kpp	0.684	0.477	0.679	0.695	0.035	0.396	0.040	0.039	0.239	0.701	0.048	0.397	0.667	0.468	0.484
Season	AVEPUB	**ac**	0.867	0.726	0.934	0.835	0.346	0.711	0.312	0.283	0.665	0.828	0.426	0.724	0.910	0.861	0.866
									s = 4.550e-04,	s = 8.457e-05,	s = 8.708e-05,	d = 1, sc = 0.01, c	d = 2, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 8	ncomp = 10	ncomp = 6	c = 1	c = 1	c = 1	c = 1	c = 2	c = 2	= 0.25	c = 0.25	c = 0.25	mtry = 68	mtry = 270	mtry = 195
		AUC	0.986	0.883	0.988	0.957	0.559	0.909	0.547	0.460	0.799	0.944	0.612	0.891	0.964	0.943	0.943
		Kpp	0.816	0.626	0.905	0.774	0.097	0.610	0.075	0.050	0.566	0.757	0.253	0.612	0.873	0.809	0.819
Season	DACGLO	**ac**	0.898	0.857	0.906	0.877	0.427	0.823	0.293	0.230	0.646	0.856	0.455	0.794	0.923	0.865	0.872
									s = 4.944e-04,	s = 9.081e-05,	s = 8.634e-05,	d = 1, sc = 0.001,	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 3	ncomp = 8	ncomp = 3	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	c = 1	c = 0.25	c = 0.25	mtry = 22	mtry = 902	mtry = 832
		AUC	0.972	0.968	0.966	0.957	0.675	0.888	0.544	0.427	0.853	0.954	0.707	0.904	0.964	0.938	0.939
		Kpp	0.854	0.814	0.856	0.824	0.234	0.758	0.072	0.059	0.560	0.799	0.259	0.731	0.891	0.818	0.830
Season	FESRUB	**ac**	0.860	0.670	0.851	0.869	0.358	0.726	0.311	0.286	0.647	0.859	0.373	0.664	0.837	0.823	0.843
									s = 6.370e-04,	s = 8.570e-05,	s = 8.739e-05,	d = 1, sc = 0.01, c	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 10	ncomp = 8	ncomp = 7	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	= 0.25	c = 0.25	c = 0.25	mtry = 270	mtry = 141	mtry = 80
		AUC	0.963	0.862	0.978	0.935	0.600	0.896	0.532	0.453	0.877	0.902	0.582	0.890	0.954	0.921	0.923
		Kpp	0.808	0.533	0.781	0.811	0.138	0.625	0.073	0.059	0.547	0.803	0.159	0.572	0.776	0.753	0.786
Season	HOLLAN	**ac**	0.938	0.885	0.998	0.931	0.642	0.885	0.368	0.336	0.823	0.956	0.663	0.897	0.951	0.945	0.944
									s = 6.653e-04,	s = 9.063e-05,	s = 8.897e-05,	d = 2, sc = 0.001,	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 5	ncomp = 8	ncomp = 6	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	c = 0.5	c = 0.25	c = 0.25	mtry = 153	mtry = 63	mtry = 94
		AUC	0.989	0.990	1.000	0.981	0.834	0.962	0.583	0.573	0.913	0.985	0.848	0.975	0.990	0.981	0.981
		Kpp	0.905	0.836	0.997	0.907	0.511	0.841	0.160	0.139	0.764	0.938	0.544	0.857	0.931	0.924	0.924
Season	PHLPRA	**ac**	0.277	0.815	0.916	0.854	0.493	0.812	0.416	0.426	0.787	0.860	0.496	0.825	0.824	0.780	0.784
									s = 3.126e-04,	s = 9.059e-05,	s = 8.493e-05,	d = 1, sc = 0.001,	d = 2, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 9	ncomp = 10	ncomp = 10	c = 1	c = 1	c = 1	c = 4	c = 2	c = 2	c = 2	c = 0.25	c = 0.25	mtry = 229	mtry = 153	mtry = 153
		AUC	0.962	0.901	0.978	0.959	0.662	0.945	0.657	0.643	0.914	0.946	0.680	0.949	0.946	0.915	0.916
		Kpp	0.814	0.748	0.885	0.801	0.324	0.746	0.223	0.237	0.716	0.810	0.322	0.766	0.763	0.704	0.710
Season	POAPRA	**ac**	0.858	0.840	0.883	0.817	0.441	0.711	0.238	0.232	0.499	0.834	0.456	0.705	0.920	0.796	0.808
									s = 5.402e-04,	s = 8.792e-05,	s = 8.619e-05,	d = 1, sc = 0.01, c	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 10	ncomp = 9	ncomp = 10	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	= 0.25	c = 0.25	c = 0.25	mtry = 437	mtry = 2017	mtry = 2567
		AUC	0.942	0.938	0.968	0.885	0.676	0.846	0.544	0.513	0.814	0.882	0.608	0.853	0.941	0.895	0.894
		Kpp	0.791	0.759	0.832	0.744	0.283	0.587	0.034	0.043	0.349	0.750	0.246	0.601	0.878	0.714	0.7264438
Season	CENJAC	**ac**	0.628	0.448	0.733	0.651	0.295	0.465	0.253	0.210	0.393	0.622	0.268	0.476	0.589	0.466	0.461
									s = 4.402e-04,	s = 8.001e-05,	s = 8.869e-05,	d = 1, sc = 0.1, c	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 9	ncomp = 2	ncomp = 8	c = 1	c = 1	c = 1	c = 2	c = 4	c = 2	= 0.25	c = 0.25	c = 0.25	mtry = 1148	mtry = 832	mtry = 195
		AUC	0.858	0.590	0.864	0.782	0.526	0.676	0.531	0.534	0.658	0.807	0.507	0.675	0.805	0.703	0.703
		Kpp	0.491	0.248	0.627	0.499	0.050	0.296	0.033	-0.015	0.211	0.482	0.030	0.316	0.440	0.295	0.301
Season	GERPRA	**ac**	0.921	0.717	0.899	0.950	0.441	0.768	0.420	0.346	0.624	0.938	0.499	0.778	0.928	0.879	0.8925
									s = 5.741e-04,	s = 8.450e-05,	s = 8.881e-05,	d = 1, sc = 0.01, c	d = 2, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 5	ncomp = 6	ncomp = 6	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	= 0.25	c = 0.25	c = 0.25	mtry = 437	mtry = 166	mtry = 343
		AUC	0.986	0.900	0.971	0.984	0.571	0.928	0.658	0.625	0.905	0.991	0.649	0.941	0.958	0.936	0.936
		Kpp	0.882	0.609	0.858	0.931	0.253	0.685	0.246	0.183	0.531	0.915	0.335	0.700	0.896	0.835	0.852
Season	KNAARV	**ac**	0.968	0.904	1.000	0.981	0.435	0.934	0.502	0.418	0.873	0.973	0.458	0.926	0.995	0.986	0.986
									s = 4.615e-04,	s = 8.850e-05,	s = 8.332e-05,	d = 2, sc = 0.001,	d = 2, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 9	ncomp = 9	ncomp = 3	c = 1	c = 1	c = 1	c = 4	c = 2	c = 2	c = 1	c = 0.25	c = 0.25	mtry = 120	mtry = 74	mtry = 102
		AUC	0.995	0.975	1.000	1.000	0.703	0.992	0.684	0.702	0.945	1.000	0.691	0.977	0.996	0.985	0.985
		Kpp	0.959	0.874	1.000	0.973	0.230	0.907	0.323	0.252	0.823	0.962	0.305	0.899	0.993	0.981	0.981
Season	LEUVUL	**ac**	0.944	0.825	0.943	0.939	0.507	0.876	0.539	0.501	0.862	0.921	0.502	0.879	0.923	0.913	0.911
									s = 5.000e-04,	s = 8.937e-05,	s = 8.724e-05,	d = 1, sc = 0.001,	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 8	ncomp = 9	ncomp = 10	c = 1	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	c = 0.25	c = 0.25	mtry = 557	mtry = 343	mtry = 229
		AUC	0.991	0.944	0.997	0.977	0.672	0.973	0.730	0.704	0.971	0.978	0.665	0.974	0.973	0.967	0.967
		Kpp	0.924	0.765	0.923	0.918	0.338	0.832	0.382	0.333	0.814	0.931	0.333	0.836	0.895	0.881	0.879
Season	PLALAN	**ac**	0.835	0.722	0.854	0.862	0.626	0.803	0.475	0.481	0.799	0.859	0.619	0.803	0.863	0.832	0.834
									s = 4.170e-04,	s = 8.772e-05,	s = 9.081e-05,	d = 1, sc = 0.01, c	d = 2, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 5	ncomp = 8	ncomp = 5	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	= 0.25	c = 0.25	c = 0.25	mtry = 22	mtry = 3267	mtry = 46
		AUC	0.957	0.870	0.969	0.964	0.742	0.935	0.737	0.745	0.942	0.958	0.727	0.929	0.956	0.940	0.940
		Kpp	0.777	0.625	0.804	0.813	0.495	0.737	0.302	0.308	0.734	0.808	0.485	0.737	0.815	0.772	0.777
Season	RANACR	**ac**	0.960	0.968	0.969	0.973	0.459	0.848	0.970	0.351	0.731	0.980	0.503	0.823	0.958	0.900	0.912
									s = 6.240e-04,	s = 8.737e-05,	s = 8.807e-05,	d = 1, scale =	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 10	ncomp = 9	ncomp = 4	c = 1	c = 1	c = 1	c = 1	c = 2	c = 2	0.01, c = 0.25	c = 0.25	c = 0.25	mtry = 46	mtry = 49	mtry = 74
		AUC	0.998	0.990	1.000	0.980	0.624	0.923	0.543	0.598	0.903	0.957	0.677	0.910	0.985	0.957	0.957
		Kpp	0.940	0.949	0.957	0.960	0.251	0.780	0.193	0.129	0.656	0.972	0.312	0.760	0.942	0.866	0.881

Table 1: continued – MTBLS1224

			PLS			*SVMl*			*SVMr*			*SVMp*			*RF*	
Level	species	eval.	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog
Diversity	ANTODO	ac	0.426	0.400	0.327	0.369	0.363	0.251	0.287	0.277	0.232	0.383	0.387	0.306	0.283	0.356	0.344
									s = 3.734e-04,	s = 8.770e-05,	s = 8.107e-05,	d = 1, sc = 0.01, c	d = 2, sc = 0.001,	d = 2, sc = 0.01, c			
		bm	ncomp = 3	ncomp = 2	ncomp = 7	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	= 0.25	c = 0.25	= 0.25	mtry = 5737	mtry = 4159	mtry = 2567
		AUC	0.332	0.588	0.546	0.521	0.419	0.376	0.470	0.482	0.505	0.577	0.494	0.434	0.465	0.419	0.420
		Kpp	-0.061	0.183	0.085	0.164	0.100	0.033	0.021	0.008	0.013	0.164	0.168	0.035	0.072	0.176	0.157
Diversity	AVEPUB	ac	0.411	0.414	0.411	0.317	0.173	0.262	0.212	0.157	0.228	0.361	0.326	0.268	0.337	0.317	0.331
									s = 3.855e-04,	s = 8.450e-05,	s = 8.773e-05,	d = 2, sc = 0.001,	d = 2, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 7	ncomp = 9	ncomp = 7	c = 1	c = 1	c = 1	c = 4	c = 2	c = 2	c = 0.25	c = 0.25	c = 0.25	mtry = 1584	mtry = 372	mtry = 3838
		AUC	0.546	0.645	0.723	0.499	0.500	0.544	0.495	0.497	0.468	0.539	0.564	0.586	0.543	0.506	0.507
		Kpp	0.095	0.225	0.222	0.106	-0.080	0.059	0.002	-0.019	0.034	0.151	0.098	0.062	0.136	0.146	0.154
Diversity	DACGLO	ac	0.342	0.269	0.447	0.257	0.259	0.276	0.250	0.215	0.153	0.306	0.255	0.246	0.257	0.243	0.233
									s = 4.651e-04,	s = 8.512e-05,	s = 8.615e-05,	d = 1, sc = 0.001,	d = 1, sc = 0.001,	d, = 1, sc =			
		bm	ncomp = 10	ncomp = 2	ncomp = 7	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	c = 2	c = 0.25	0.001, c = 0.25	mtry = 20	mtry = 2017	mtry = 1584
		AUC	0.659	0.441	0.729	0.649	0.546	0.447	0.520	0.478	0.348	0.639	0.528	0.456	0.475	0.473	0.471
		Kpp	0.175	0.055	0.262	0.020	0.008	0.101	0.004	0.000	0.018	0.053	-0.007	0.074	0.043	0.069	0.053
Diversity	FESRUB	ac	0.558	0.576	0.610	0.540	0.273	0.376	0.287	0.239	0.347	0.533	0.344	0.393	0.498	0.412	0.426
									s = 5.941e-04,	s = 8.791e-05,	s = 8.764e-05,	d = 1, sc = 0.1, c	d = 2, sc =	d = 1, sc = 0.001,			
		bm	ncomp = 8	ncomp = 8	ncomp = 9	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	= 0.25	1e+05, c = 0.25	c = 0.25	mtry = 2185	mtry = 292	mtry = 372
		AUC	0.673	0.766	0.824	0.686	0.514	0.666	0.482	0.436	0.627	0.690	0.503	0.679	0.731	0.654	0.651
		Kpp	0.147	0.448	0.461	0.354	0.047	0.202	0.045	0.030	0.211	0.366	0.144	0.255	0.312	0.254	0.258
Diversity	HOLLAN	ac	0.406	0.380	0.314	0.380	0.236	0.116	0.248	0.213	0.105	0.394	0.231	0.250	0.300	0.168	0.140
									s = 5.448e-04,	s = 9.056e-05,	s = 8.782e-05,	d = 3, sc = 0.001,	d = 3, sc = 0.01, c	d = 3, sc = 0.1, c			
		bm	ncomp = 7	ncomp = 10	ncomp = 6	c = 1	c = 1	c = 1	c = 2	c = 2	c = 2	c = 0.5	= 0.25	= 0.25	mtry = 33	mtry = 24	mtry = 17
		AUC	0.584	0.724	0.538	0.567	0.563	0.575	0.435	0.445	0.563	0.524	0.599	0.532	0.494	0.363	0.362
		Kpp	0.276	0.189	0.097	0.150	0.005	-0.078	0.052	0.001	-0.051	0.185	0.040	-0.003	0.095	0.026	-0.022
Diversity	PHLPRA	ac	0.465	0.428	0.442	0.427	0.275	0.341	0.299	0.246	0.228	0.436	0.340	0.334	0.350	0.347	0.356
									s = 3.831e-04,	s = 9.177e-05,	s = 8.663e-05,	d = 1, sc = 0.1, c	d = 2, sc = 10, c =	d = 3, sc = 0.001,			
		bm	ncomp = 10	ncomp = 10	ncomp = 10	c = 1	c = 1	c = 1	c = 4	c = 2	c = 1	= 0.25	0.25	c = 0.25	mtry = 832	mtry = 26	mtry = 195
		AUC	0.735	0.669	0.725	0.697	0.519	0.642	0.557	0.509	0.574	0.690	0.531	0.648	0.642	0.598	0.599
		Kpp	0.358	0.222	0.252	0.231	0.038	0.123	0.059	0.020	0.113	0.240	0.120	0.134	0.128	0.140	0.151
Diversity	POAPRA	ac	0.543	0.490	0.670	0.464	0.215	0.328	0.348	0.252	0.281	0.462	0.256	0.362	0.523	0.328	0.319
									s = 6.184e-04,	s = 7.964e-05,	s = 8.631e-05,	d = 1, sc = 0.01, c	d = 3, sc = 0.01, c	d = 1, sc = 0.001,			
		bm	ncomp = 10	ncomp = 10	ncomp = 9	c = 1	c = 1	c = 1	c = 8	c = 2	c = 2	= 0.25	= 0.25	c = 0.25	mtry = 1861	mtry = 5737	mtry = 5294
		AUC	0.576	0.770	0.841	0.777	0.509	0.606	0.524	0.472	0.672	0.820	0.589	0.626	0.725	0.534	0.534
		Kpp	0.089960727	0.3199708	0.51528822	0.283812	-0.05768287	0.1119529	0.1187996	0.03829966	0.13143338	0.08249158	0.038571429	0.200680272	0.33603896	0.141774892	0.110329485
Diversity	CENJAC	ac	0.654	0.425	0.680	0.671	0.425	0.619	0.238	0.233	0.473	0.683	0.393	0.574	0.725	0.678	0.691
									s = 4.364e-04,	s = 1.021e-04,	s = 8.301e-05,	d = 1, sc = 0.01, c	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 6	ncomp = 9	ncomp = 10	c = 1	c = 1	c = 1	c = 2	c = 2	c = 4	= 0.25	c = 0.25	c = 0.25	mtry = 2567	mtry = 557	mtry = 654
		AUC	0.662	0.618	0.869	0.827	0.647	0.743	0.447	0.511	0.742	0.808	0.577	0.760	0.852	0.796	0.796
		Kpp	0.18579932	0.21612666	0.5547374	0.5374325	0.2526725	0.4811207	0.009583333	0.02154762	0.3440957	0.54680451	0.22065476	0.441618567	0.6044974	0.5525493	0.5679413
Diversity	GERPRA	ac	0.400	0.421	0.455	0.313	0.184	0.193	0.174	0.102	0.153	0.313	0.261	0.201	0.418	0.212	0.243
									s = 6.170e-04,	s = 8.801e-05,	s = 8.708e-05,	d = 1, sc = 0.1, c	d = 2, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 10	ncomp = 6	ncomp = 5	c = 1	c = 1	c = 1	c = 4	c = 2	c = 2	= 0.25	c = 0.25	c = 0.25	mtry = 4885	mtry = 3838	mtry = 2017
		AUC	0.561	0.613	0.684	0.435	0.472	0.417	0.552	0.517	0.476	0.451	0.511	0.435	0.544	0.454	0.453
		Kpp	0.073	0.214	0.267	0.090	-0.067	-0.009	-0.047	-0.001	0.035	0.070	0.056	0.002	0.206	0.012	0.035
Diversity	KNAARV	ac	0.293	0.354	0.269	0.229	0.214	0.253	0.171	0.143	0.113	0.286	0.236	0.236	0.363	0.230	0.219
									s = 4.266e-04,	s = 8.949e-05,	s = 8.740e-05,	d = 2, sc = 1, c =	d = 2, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 4	ncomp = 1	ncomp = 5	c = 1	c = 1	c = 1	c = 8	c = 2	c = 0.25	0.25	c = 0.25	c = 0.25	mtry = 5737	mtry = 768	mtry = 4507
		AUC	0.659	0.554	0.541	0.435	0.524	0.362	0.520	0.461	0.545	0.611	0.547	0.367	0.476	0.445	0.445
		Kpp	0.150	0.140	0.057	-0.003	-0.013	0.069	-0.060	-0.069	0.017	0.070	0.026	0.040	0.181	0.050	0.041
Diversity	LEUVUL	ac	0.392	0.387	0.315	0.322	0.238	0.156	0.277	0.152	0.161	0.337	0.259	0.185	0.296	0.215	0.218
									s = 4.218e-04,	s = 8.432e-05,	s = 8.750e-05,	d = 1, sc = 0.001,	d = 3, sc = 0.001,	d = 3, sc = 0.001,			
		bm	ncomp = 3	ncomp = 10	ncomp = 9	c = 1	c = 1	c = 1	c = 2	c = 2	c = 0.25	c = 2	c = 0.25	c = 0.25	mtry = 141	mtry = 8	mtry = 13
		AUC	0.593	0.640	0.572	0.451	0.482	0.415	0.473	0.489	0.515	0.491	0.503	0.538	0.492	0.394	0.397
		Kpp	0.170	0.179	0.091	0.096	-0.008	-0.104	0.042	-0.092	-0.017	0.119	0.042	0.004	0.071	-0.006	-0.008
Diversity	PLALAN	ac	0.359	0.367	0.374	0.331	0.222	0.243	0.317	0.314	0.262	0.345	0.297	0.271	0.305	0.259	0.264
									s = 4.443e-04,	s = 8.543e-05,	s = 8.548e-05,	d = 2, sc = 0.01, c	d = 2, sc = 0.01, c	d = 1, sc = 0.001,			
		bm	ncomp = 6	ncomp = 2	ncomp = 1	c = 1	c = 1	c = 1	c = 2	c = 2	c = 4	= 0.25	= 0.25	c = 0.25	mtry = 102	mtry = 2	mtry = 8
		AUC	0.724	0.577	0.579	0.505	0.504	0.483	0.482	0.465	0.473	0.556	0.496	0.496	0.526	0.466	0.050
		Kpp	0.205	0.151	0.172	0.109	-0.028	0.007	0.095	0.092	0.045	0.129	0.088	0.040	0.076	0.050	0.053
Diversity	RANACR	ac	0.345	0.473	0.377	0.281	0.222	0.252	0.232	0.119	0.153	0.308	0.232	0.259	0.423	0.356	0.351
									s = 6.415e-04,	s = 9.009e-05,	s = 8.763e-05,	d = 1, sc = 0.001,	d = 1, sc = 0.001,	d = 1, sc = 0.001,			
		bm	ncomp = 8	ncomp = 7	ncomp = 1	c = 1	c = 1	c = 1	c = 8	c = 0.25	c = 2	c = 4	c = 0.25	c = 0.25	mtry = 2369	mtry = 4159	mtry = 5737
		AUC	0.573	0.696	0.688	0.591	0.530	0.554	0.504	0.495	0.504	0.589	0.568	0.624	0.640	0.545	0.544
		Kpp	0.031	0.253	0.149	0.040	-0.011	0.056	0.028	0.003	0.011	0.073	0.016	0.065	0.238	0.135	0.148

Table 1: continued – MTBLS2140

Data set			MTBLS2140																	
			PLS				*SVMl*			*SVMr*			*SVMp*				*RF*			
Level	species	eval.	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog	ZeroInt-pP	ImpInt	ImpLog
FG		**ac**	1.000	1.000	1.000	1.000	0.999	1.000	0.973	0.979	1.000	1.000	0.995	1.000	1.000	0.989	0.990			
		bm	ncomp = 6	ncomp = 18	ncomp = 3	c = 1	c = 1	c = 1	s = 4.404e-04, c = 2	s = 5.389e-04, c = 2	s = 4.276e-04, c = 1	d = 1, sc = 0.1, c = 0.25	d = 2, sc = 0.001, c = 2	d = 1, sc = 0.01, c = 0.25	mtry = 18	mtry = 268	mtry = 306			
		AUC	1.000	1.000	1.000	1.000	0.994	1.000	0.997	0.997	1.000	1.000	0.994	1.000	0.968	1.000	1.000			
		Kpp	1.000	0.995	1.000	1.000	0.997	1.000	0.932	0.947	1.000	1.000	0.988	1.000	1.000	0.973	0.976			
Species		**ac**	0.994	0.970	1.000	0.988	0.987	1.000	0.958	0.968	1.000	0.988	0.988	1.000	0.990	0.988	0.988			
		bm	ncomp = 20	ncomp = 30	ncomp = 20	c = 1	c = 1	c = 1	s = 4.743e-04, c = 8	s = 4.200e-04, c = 8.	s = 4.950e-04, c = 4	d = 1, sc = 0.1, c = 0.25	d = 1, sc = 0.1, c = 0.25	d = 1, sc = 0.1, c = 0.25	mtry = 490	mtry = 2	mtry = 3			
		AUC	1.000	1.000	1.000	0.951	0.950	0.945	0.873	0.860	0.924	0.950	0.949	0.945	1.000	1.000	1.000			
		Kpp	0.994	0.976	1.000	0.987	0.986	1.000	0.955	0.966	1.000	0.988	0.987	1.000	0.989	0.987	0.987			

APPENDIX – CHAPTER 3

Supplement Table 1: Plant species indices. Plant species richness and Shannon were recorded in each season for each plot separately. Seasons represent sampling in spring (E) , early summer (F), late summer (G) and autumn (H) of the 2018 growing season. Only identified species were included in the analysis. Species Shannon is shown as calculation for each species. Due to the number of species, the table is split across multiple parts: Shannon & Richness, FG legume, FG herb, FG grass.

(Shannon & Richness Season 2018 E)

Season	plot	observed Shannon all species per plot (pShan)	observed Shannon FG grass per plot (Shan_grass)	observed Shannon FG legume per plot (Shan_legume)	observed Shannon FG herb per plot (Shan_herb)	number of species per plot (pRich)
2018_E	A002	2.099	0.701	0.147	1.250	24
2018_E	A003	2.315	0.753	0.190	1.373	20
2018_E	A005	2.417	0.696	0.156	1.565	20
2018_E	A009	1.894	0.259	0.112	1.522	20
2018_E	A010	1.769	0.540	0.115	1.114	18
2018_E	A011	1.513	0.099	0.275	1.139	17
2018_E	A013	2.141	0.052	0.337	1.752	19
2018_E	A016	2.594	0.279	0.389	1.925	27
2018_E	A018	2.249	0.496	0.326	1.427	20
2018_E	A020	1.624	0.159	0.036	1.429	19
2018_E	A026	2.003	1.018	0.206	0.780	20
2018_E	A027	1.685	0.288	0.052	1.345	14
2018_E	A030	1.790	0.101	0.250	1.439	12
2018_E	A035	1.867	0.527	0.255	1.085	15
2018_E	A040	2.032	0.367	0.110	1.555	16
2018_E	A042	1.332	0.071	0.118	1.143	10
2018_E	A043	2.134	0.168	0.529	1.438	18
2018_E	A044	1.880	0.388	0.508	0.985	17
2018_E	A045	1.858	0.937	0.683	0.238	14
2018_E	A046	1.723	0.185	0.859	0.680	15
2018_E	B048	1.898	0.079	0.415	1.404	17
2018_E	B051	2.066	0.902	0.375	0.789	17
2018_E	B054	1.995	0.523	0.432	1.040	19
2018_E	B057	2.396	0.365	0.449	1.582	20
2018_E	B059	2.506	0.479	0.327	1.701	19
2018_E	B060	2.629	0.747	0.338	1.544	20
2018_E	B064	1.993	0.602	0.369	1.022	15
2018_E	B067	1.831	0.799	0.477	0.555	13
2018_E	B071	2.346	0.252	0.504	1.590	17
2018_E	B073	2.347	0.854	0.301	1.192	15
2018_E	B075	1.609	0.381	0.436	0.792	13
2018_E	B080	2.184	0.376	0.581	1.228	18
2018_E	B081	1.686	0.452	0.126	1.109	11
2018_E	B085	1.780	0.350	0.502	0.928	19
2018_E	B090	1.764	0.497	0.331	0.936	14
2018_E	B092	2.233	1.157	0.248	0.828	22
2018_E	C097	1.910	0.972	0.171	0.767	13
2018_E	C103	1.983	0.611	0.228	1.144	16
2018_E	C109	1.733	0.487	0.119	1.126	10
2018_E	C110	2.154	0.654	0.222	1.278	19
2018_E	C114	1.404	0.876	0.025	0.503	11
2018_E	C115	2.436	0.506	0.432	1.499	17
2018_E	C121	1.712	0.425	0.108	1.179	19
2018_E	C131	1.865	0.592	0.222	1.051	17
2018_E	C133	2.504	0.412	0.182	1.910	18
2018_E	C135	1.613	0.358	0.140	1.115	14
2018_E	C136	1.718	0.527	0.272	0.919	13
2018_E	C137	2.020	0.885	0.257	0.878	20

Supplement Table 1: Plant species indices – continued (Shannon & Richness Season 2018 F & G)

Season	plot	observed Shannon all species per plot (pShan)	observed Shannon FG grass per plot (Shan_grass)	observed Shannon FG legume per plot (Shan_legume)	observed Shannon FG herb per plot (Shan_herb)	number of species per plot (pRich)
2018_F	A002	1.625	0.576	0.085	0.965	8
2018_F	A003	2.086	0.688	0.146	1.252	12
2018_F	A005	1.994	0.549	0.138	1.307	11
2018_F	A009	1.680	0.441	0.000	1.239	8
2018_F	A010	1.118	0.427	0.072	0.620	8
2018_F	A011	1.286	0.000	0.184	1.102	9
2018_F	A013	1.232	0.162	0.162	0.908	7
2018_F	A016	2.076	0.227	0.399	1.451	16
2018_F	A018	2.090	0.560	0.070	1.460	12
2018_F	A020	1.293	0.085	0.000	1.208	10
2018_F	A026	1.046	0.518	0.061	0.467	9
2018_F	A027	1.098	0.000	0.000	1.098	7
2018_F	A030	1.345	0.000	0.000	1.345	6
2018_F	A035	1.606	0.453	0.093	1.060	10
2018_F	A040	1.333	0.294	0.000	1.040	7
2018_F	A042	0.771	0.000	0.000	0.771	6
2018_F	A043	0.822	0.000	0.280	0.542	7
2018_F	A044	1.066	0.221	0.364	0.481	6
2018_F	A045	2.156	1.396	0.242	0.519	13
2018_F	A046	1.324	0.138	0.354	0.832	10
2018_F	B048	1.333	0.197	0.197	0.938	8
2018_F	B051	1.474	0.601	0.133	0.740	6
2018_F	B054	1.008	0.332	0.314	0.362	5
2018_F	B057	1.828	0.786	0.099	0.943	9
2018_F	B059	2.272	0.310	0.436	1.526	10
2018_F	B060	1.872	0.572	0.431	0.869	10
2018_F	B064	0.768	0.152	0.159	0.458	9
2018_F	B067	1.645	0.972	0.199	0.475	9
2018_F	B071	1.454	0.491	0.295	0.668	7
2018_F	B073	1.962	0.964	0.063	0.935	10
2018_F	B075	1.496	0.681	0.196	0.619	6
2018_F	B080	1.072	0.303	0.186	0.582	8
2018_F	B081	1.638	0.675	0.144	0.819	10
2018_F	B085	1.219	0.000	0.508	0.711	7
2018_F	B090	0.838	0.284	0.168	0.386	5
2018_F	B092	1.395	1.019	0.000	0.376	8
2018_F	C097	1.481	0.316	0.000	1.165	10
2018_F	C103	1.312	0.643	0.052	0.617	7
2018_F	C109	1.914	0.570	0.184	1.160	11
2018_F	C110	2.244	0.365	0.222	1.658	14
2018_F	C114	1.051	0.879	0.000	0.172	8
2018_F	C115	1.947	0.363	0.267	1.318	11
2018_F	C121	1.328	0.000	0.000	1.328	9
2018_F	C131	0.431	0.227	0.000	0.204	6
2018_F	C133	1.894	0.368	0.218	1.308	8
2018_F	C135	1.175	0.177	0.067	0.931	7
2018_F	C136	1.597	0.627	0.238	0.732	8
2018_F	C137	1.170	0.335	0.000	0.835	5

Season	plot	observed Shannon all species per plot (pShan)	observed Shannon FG grass per plot (Shan_grass)	observed Shannon FG legume per plot (Shan_legume)	observed Shannon FG herb per plot (Shan_herb)	number of species per plot (pRich)
2018_G	A002	1.958	0.785	0.138	1.034	12
2018_G	A003	2.339	0.468	0.289	1.582	20
2018_G	A005	2.023	0.461	0.078	1.484	13
2018_G	A009	2.324	0.395	0.045	1.885	19
2018_G	A010	1.745	0.355	0.228	1.162	10
2018_G	A011	1.796	0.170	0.225	1.401	14
2018_G	A013	1.638	0.048	0.213	1.377	14
2018_G	A016	2.543	0.152	0.302	2.089	18
2018_G	A018	2.268	0.745	0.130	1.393	15
2018_G	A020	1.403	0.049	0.000	1.355	9
2018_G	A026	1.292	0.483	0.046	0.763	14
2018_G	A027	1.989	0.067	0.243	1.680	16
2018_G	A030	1.359	0.000	0.000	1.359	7
2018_G	A035	1.965	0.458	0.184	1.323	14
2018_G	A040	2.645	0.435	0.091	2.119	20
2018_G	A042	0.976	0.056	0.000	0.920	9
2018_G	A043	1.676	0.059	0.260	1.357	14
2018_G	A044	1.481	0.310	0.311	0.860	9
2018_G	A045	2.018	1.253	0.298	0.467	13
2018_G	A046	1.439	0.149	0.350	0.940	11
2018_G	B048	1.684	0.000	0.377	1.308	15
2018_G	B051	1.581	0.604	0.110	0.867	9
2018_G	B054	1.725	0.288	0.279	1.159	11
2018_G	B057	1.861	0.470	0.357	1.033	10
2018_G	B059	2.252	0.461	0.316	1.476	12
2018_G	B060	2.029	0.367	0.391	1.271	15
2018_G	B064	0.928	0.212	0.193	0.523	15
2018_G	B067	1.689	0.899	0.227	0.563	12
2018_G	B071	2.383	0.371	0.431	1.581	17
2018_G	B073	2.026	0.926	0.035	1.065	10
2018_G	B075	1.578	0.657	0.164	0.756	7
2018_G	B080	1.208	0.252	0.238	0.718	13
2018_G	B081	1.843	0.461	0.262	1.120	14
2018_G	B085	1.751	0.134	0.369	1.248	15
2018_G	B090	1.060	0.187	0.252	0.621	12
2018_G	B092	1.392	0.750	0.096	0.546	15
2018_G	C097	1.538	0.275	0.114	1.149	18
2018_G	C103	1.368	0.380	0.000	0.988	10
2018_G	C109	1.893	0.626	0.122	1.145	14
2018_G	C110	2.249	0.368	0.204	1.678	20
2018_G	C114	1.060	0.800	0.000	0.260	10
2018_G	C115	2.013	0.484	0.225	1.304	15
2018_G	C121	1.518	0.000	0.000	1.518	11
2018_G	C131	1.102	0.758	0.043	0.301	11
2018_G	C133	2.288	0.328	0.392	1.568	12
2018_G	C135	1.005	0.041	0.000	0.964	8
2018_G	C136	1.493	0.502	0.060	0.931	9
2018_G	C137	1.192	0.403	0.000	0.789	7

Supplement Table 1: Plant species indices – continued (Shannon & Richness Season 2018 H)

Season	plot	observed Shannon all species per plot (pShan)	observed Shannon FG grass per plot (Shan_grass)	observed Shannon FG legume per plot (Shan_legume)	observed Shannon FG herb per plot (Shan_herb)	number of species per plot (pRich)
2018_H	A002	1.762	0.436	0.094	1.232	14
2018_H	A003	1.696	0.513	0.078	1.105	13
2018_H	A005	2.259	0.392	0.090	1.777	17
2018_H	A009	2.440	0.000	0.260	2.180	19
2018_H	A010	1.967	0.512	0.104	1.351	11
2018_H	A011	2.194	0.083	0.288	1.823	19
2018_H	A013	2.311	0.000	0.392	1.919	20
2018_H	A016	2.181	0.235	0.108	1.837	15
2018_H	A018	2.386	0.670	0.226	1.490	18
2018_H	A020	1.622	0.117	0.112	1.394	11
2018_H	A026	1.715	0.898	0.116	0.700	18
2018_H	A027	2.249	0.112	0.112	2.025	16
2018_H	A030	1.468	0.000	0.169	1.299	9
2018_H	A035	2.539	0.631	0.256	1.652	18
2018_H	A040	2.516	0.274	0.161	2.081	17
2018_H	A042	1.498	0.000	0.117	1.381	12
2018_H	A043	2.126	0.000	0.337	1.789	15
2018_H	A044	2.159	0.465	0.313	1.381	15
2018_H	A045	1.942	1.220	0.201	0.521	17
2018_H	A046	1.848	0.028	0.424	1.396	18
2018_H	B048	2.176	0.083	0.240	1.854	20
2018_H	B051	2.202	0.409	0.337	1.456	17
2018_H	B054	2.059	0.416	0.223	1.420	14
2018_H	B057	2.311	0.322	0.495	1.495	14
2018_H	B059	1.447	0.214	0.110	1.123	12
2018_H	B060	1.902	0.363	0.092	1.447	10
2018_H	B064	1.667	0.448	0.264	0.955	13
2018_H	B067	1.935	0.821	0.291	0.823	15
2018_H	B071	2.200	0.320	0.360	1.520	14
2018_H	B073	2.078	0.840	0.141	1.098	12
2018_H	B075	1.904	0.693	0.300	0.911	11
2018_H	B080	1.884	0.409	0.213	1.262	13
2018_H	B081	2.350	0.512	0.069	1.769	16
2018_H	B085	2.298	0.136	0.443	1.718	20
2018_H	B090	1.593	0.302	0.086	1.206	12
2018_H	B092	1.464	0.774	0.100	0.590	13
2018_H	C097	2.102	0.611	0.000	1.490	17
2018_H	C103	1.893	0.729	0.113	1.051	12
2018_H	C109	1.981	0.478	0.131	1.372	13
2018_H	C110	2.268	0.438	0.148	1.682	16
2018_H	C114	1.179	0.776	0.032	0.371	11
2018_H	C115	2.225	0.341	0.209	1.674	13
2018_H	C121	2.047	0.000	0.000	2.047	12
2018_H	C131	1.038	0.328	0.071	0.638	13
2018_H	C133	2.024	0.099	0.099	1.826	14
2018_H	C135	1.719	0.176	0.166	1.377	12
2018_H	C136	1.605	0.583	0.073	0.949	11
2018_H	C137	1.745	0.352	0.225	1.167	11

Supplement Table 1: Plant species indices – continued (FG legume Season 2018 E)

	functional group	legume	legume	legume	legume	legume	legume	legume	legume	legume	legume
	family	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae
	Code	Latpra	Lotcor	Medlup	Medspe	Tridub	Trifra	Tripra	Trirep	Viccra	Vicsep
	species	Lathyrus_pratensis	Lotus_corniculatus	Medicago_lupulina	Medicago_spec.	Trifolium_dubium	Trifolium_fragiferum	Trifolium_pratense	Trifolium_repens	Vicia_cracca	Vicia_sepium
	season	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G	2018_E	2018_E,H	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G
Season	plot										
2018_E	A002	-0.068	0	-0.040	0	0	0	0	-0.040	0	0
2018_E	A003	-0.071	0	-0.118	0	0	0	0	0	0	0
2018_E	A005	-0.060	0	-0.060	0	0	0	0	0	0	-0.035
2018_E	A009	0	0	-0.071	0	0	0	0	-0.042	0	0
2018_E	A010	-0.031	-0.031	-0.053	0	0	0	0	0	0	0
2018_E	A011	0	0	-0.096	0	0	0	0	-0.179	0	0
2018_E	A013	-0.052	0	-0.167	0	0	0	0	-0.030	-0.088	0
2018_E	A016	-0.066	-0.110	-0.110	0	0	0	0	-0.038	0	-0.066
2018_E	A018	-0.051	0	-0.246	0	0	0	0	0	-0.029	0
2018_E	A020	0	0	0	0	0	0	0	0	-0.036	0
2018_E	A026	-0.065	-0.038	-0.038	0	0	0	0	0	-0.065	0
2018_E	A027	0	0	0	0	0	0	0	-0.052	0	0
2018_E	A030	0	-0.107	0	0	0	0	0	-0.037	-0.107	0
2018_E	A035	-0.038	-0.108	-0.108	0	0	0	0	0	0	0
2018_E	A040	0	0	-0.110	0	0	0	0	0	0	0
2018_E	A042	0	-0.118	0	0	0	0	0	0	0	0
2018_E	A043	-0.061	0	-0.279	0	0	0	0	0	-0.189	0
2018_E	A044	-0.069	-0.115	-0.209	0	0	0	0	0	-0.115	0
2018_E	A045	-0.102	-0.189	0	0	0	0	0	-0.035	-0.358	0
2018_E	A046	-0.142	-0.142	0	0	0	0	0	-0.220	-0.355	0
2018_E	B048	0	0	-0.245	0	0	0	-0.085	0	-0.085	0
2018_E	B051	0	-0.074	-0.074	0	0	0	0	-0.142	-0.043	-0.043
2018_E	B054	-0.031	0	-0.054	0	0	0	0	0	-0.257	-0.091
2018_E	B057	-0.128	0	-0.066	0	-0.128	0	0	0	-0.128	0
2018_E	B059	0	0	-0.076	0	-0.125	0	0	0	-0.125	0
2018_E	B060	0	0	0	0	-0.169	0	0	0	-0.169	0
2018_E	B064	-0.095	-0.178	0	0	0	0	0	-0.095	0	0
2018_E	B067	-0.043	0	-0.073	0	0	0	0	0	-0.219	-0.141
2018_E	B071	-0.252	0	0	0	0	0	0	0	-0.252	0
2018_E	B073	0	0	0	-0.029	-0.159	0	0	-0.029	-0.084	0
2018_E	B075	0	0	-0.090	0	0	0	0	-0.256	-0.090	0
2018_E	B080	-0.149	-0.230	0	-0.046	-0.078	0	0	0	-0.078	0
2018_E	B081	0	-0.032	0	0	0	0	0	-0.094	0	0
2018_E	B085	0	0	-0.094	0	0	0	0	-0.176	-0.176	-0.056
2018_E	B090	0	0	-0.165	0	0	0	0	-0.165	0	0
2018_E	B092	-0.113	0	-0.068	0	0	0	0	-0.068	0	0
2018_E	C097	0	0	0	0	0	0	0	-0.171	0	0
2018_E	C103	0	0	-0.088	0	0	0	0	-0.088	-0.052	0
2018_E	C109	0	0	0	0	0	0	0	-0.119	0	0
2018_E	C110	-0.053	0	0	0	0	0	0	0	-0.169	0
2018_E	C114	0	0	-0.025	0	0	0	0	0	0	0
2018_E	C115	-0.123	0	-0.063	0	0	0	-0.123	-0.123	0	0
2018_E	C121	0	0	-0.108	0	0	0	0	0	0	0
2018_E	C131	-0.086	0	-0.086	0	0	0	0	0	-0.051	0
2018_E	C133	-0.113	0	-0.068	0	0	0	0	0	0	0
2018_E	C135	0	0	-0.104	0	0	0	-0.036	0	0	0
2018_E	C136	0	0	-0.167	0	0	0	0	-0.053	-0.053	0
2018_E	C137	-0.035	0	0	0	0	0	0	-0.187	-0.035	0

Supplement Table 1: Plant species indices – continued (FG legume Season 2018 F)

	functional group	legume	legume	legume	legume	legume	legume	legume	legume	legume	legume
	family	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae
	Code	Latpra	Lotcor	Medlup	Medspe	Tridub	Trifra	Tripra	Trirep	Viccra	Vicsep
	species	Lathyrus_praten	Lotus_corniculat	Medicago_lupuli	Medicago_spec.	Trifolium_dubiur	Trifolium_fragife	Trifolium_praten	Trifolium_repen	Vicia_cracca	Vicia_sepium
	season	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G	2018_E	2018_E,H	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G
Season	plot										
2018_F	A002	-0.085	0	0	0	0	0	0	0	0	0
2018_F	A003	-0.146	0	0	0	0	0	0	0	0	0
2018_F	A005	-0.138	0	0	0	0	0	0	0	0	0
2018_F	A009	0	0	0	0	0	0	0	0	0	0
2018_F	A010	-0.072	0	0	0	0	0	0	0	0	0
2018_F	A011	-0.115	-0.069	0	0	0	0	0	0	0	0
2018_F	A013	0	0	0	0	0	0	0	0	-0.162	0
2018_F	A016	-0.141	0	-0.086	0	0	0	-0.086	0	-0.086	0
2018_F	A018	-0.070	0	0	0	0	0	0	0	0	0
2018_F	A020	0	0	0	0	0	0	0	0	0	0
2018_F	A026	-0.061	0	0	0	0	0	0	0	0	0
2018_F	A027	0	0	0	0	0	0	0	0	0	0
2018_F	A030	0	0	0	0	0	0	0	0	0	0
2018_F	A035	-0.046	0	0	0	0	0	-0.046	0	0	0
2018_F	A040	0	0	0	0	0	0	0	0	0	0
2018_F	A042	0	0	0	0	0	0	0	0	0	0
2018_F	A043	-0.078	0	0	0	0	0	0	0	-0.202	0
2018_F	A044	-0.221	0	0	0	0	0	0	0	-0.143	0
2018_F	A045	-0.092	0	0	0	0	0	0	0	-0.150	0
2018_F	A046	-0.138	0	0	0	0	0	0	0	-0.215	0
2018_F	B048	-0.197	0	0	0	0	0	0	0	0	0
2018_F	B051	-0.133	0	0	0	0	0	0	0	0	0
2018_F	B054	-0.136	0	0	0	0	0	0	0	-0.178	0
2018_F	B057	-0.099	0	0	0	0	0	0	0	0	0
2018_F	B059	-0.218	0	0	0	0	0	0	0	-0.218	0
2018_F	B060	0	-0.132	0	0	0	0	0	0	-0.299	0
2018_F	B064	-0.059	-0.100	0	0	0	0	0	0	0	0
2018_F	B067	-0.100	0	0	0	0	0	0	0	-0.100	0
2018_F	B071	-0.147	0	0	0	0	0	0	0	-0.147	0
2018_F	B073	0	-0.063	0	0	0	0	0	0	0	0
2018_F	B075	0	0	0	0	0	0	0	-0.052	-0.144	0
2018_F	B080	-0.116	-0.070	0	0	0	0	0	0	0	0
2018_F	B081	-0.072	0	0	0	0	0	0	-0.072	0	0
2018_F	B085	-0.212	-0.083	0	0	0	0	0	0	-0.212	0
2018_F	B090	-0.168	0	0	0	0	0	0	0	0	0
2018_F	B092	0	0	0	0	0	0	0	0	0	0
2018_F	C097	0	0	0	0	0	0	0	0	0	0
2018_F	C103	0	0	0	0	0	0	0	0	-0.052	0
2018_F	C109	-0.135	0	0	0	0	0	0	-0.049	0	0
2018_F	C110	-0.111	0	0	0	0	0	0	0	-0.111	0
2018_F	C114	0	0	0	0	0	0	0	0	0	0
2018_F	C115	-0.267	0	0	0	0	0	0	0	0	0
2018_F	C121	0	0	0	0	0	0	0	0	0	0
2018_F	C131	0	0	0	0	0	0	0	0	0	0
2018_F	C133	-0.218	0	0	0	0	0	0	0	0	0
2018_F	C135	-0.067	0	0	0	0	0	0	0	0	0
2018_F	C136	0	-0.138	0	0	0	0	-0.050	0	0	-0.050
2018_F	C137	0	0	0	0	0	0	0	0	0	0

Supplement Table 1: Plant species indices – continued (FG legume Season 2018 G)

	functional group	legume	legume	legume	legume	legume	legume	legume	legume	legume	legume
	family	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae
	Code	Latpra	Lotcor	Medlup	Medspe	Tridub	Trifra	Tripra	Trirep	Viccra	Vicsep
	species	Lathyrus_praten	Lotus_corniculat	Medicago_lupuli	Medicago_spec.	Trifolium_dubiur	Trifolium_fragife	Trifolium_praten	Trifolium_repen	Vicia_cracca	Vicia_sepium
	season	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G	2018_E	2018_E,H	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G
Season	plot										
2018_G	A002	-0.138	0	0	0	0	0	0	0	0	0
2018_G	A003	-0.185	-0.035	0	0	0	0	0	-0.035	-0.035	0
2018_G	A005	-0.078	0	0	0	0	0	0	0	0	0
2018_G	A009	0	0	0	0	0	0	0	-0.045	0	0
2018_G	A010	-0.228	0	0	0	0	0	0	0	0	0
2018_G	A011	-0.225	0	0	0	0	0	0	0	0	0
2018_G	A013	-0.132	0	0	0	0	0	0	0	-0.081	0
2018_G	A016	-0.152	0	-0.056	0	0	0	0	0	-0.094	0
2018_G	A018	-0.082	0	0	0	0	0	0	0	-0.048	0
2018_G	A020	0	0	0	0	0	0	0	0	0	0
2018_G	A026	-0.046	0	0	0	0	0	0	0	0	0
2018_G	A027	-0.177	-0.067	0	0	0	0	0	0	0	0
2018_G	A030	0	0	0	0	0	0	0	0	0	0
2018_G	A035	-0.135	-0.049	0	0	0	0	0	0	0	0
2018_G	A040	0	0	0	0	0	0	0	-0.091	0	0
2018_G	A042	0	0	0	0	0	0	0	0	0	0
2018_G	A043	-0.100	0	0	0	0	0	0	0	-0.161	0
2018_G	A044	-0.121	0	0	0	0	0	0	0	-0.191	0
2018_G	A045	-0.075	0	0	0	0	0	0	0	-0.223	0
2018_G	A046	-0.091	0	0	0	0	0	0	0	-0.258	0
2018_G	B048	-0.276	0	0	0	0	0	0	0	-0.101	0
2018_G	B051	-0.070	-0.041	0	0	0	0	0	0	0	0
2018_G	B054	-0.107	0	0	0	0	0	0	0	-0.172	0
2018_G	B057	-0.215	-0.071	0	0	0	0	0	0	-0.071	0
2018_G	B059	0	0	0	0	0	0	0	0	-0.316	0
2018_G	B060	0	0	0	0	0	0	0	-0.052	-0.339	0
2018_G	B064	-0.052	-0.088	0	0	0	0	0	-0.052	0	0
2018_G	B067	-0.055	0	0	0	0	0	0	0	-0.173	0
2018_G	B071	-0.159	0	0	0	0	0	0	0	-0.272	0
2018_G	B073	0	0	0	0	0	0	0	0	-0.035	0
2018_G	B075	0	0	0	0	0	0	0	0	-0.164	0
2018_G	B080	-0.091	-0.147	0	0	0	0	0	0	0	0
2018_G	B081	-0.094	-0.056	0	0	0	-0.056	0	-0.056	0	0
2018_G	B085	0	-0.082	0	0	0	0	0	-0.048	-0.238	0
2018_G	B090	-0.084	-0.084	0	0	0	0	0	-0.084	0	0
2018_G	B092	-0.048	-0.048	0	0	0	0	0	0	0	0
2018_G	C097	0	-0.057	0	0	0	0	0	-0.057	0	0
2018_G	C103	0	0	0	0	0	0	0	0	0	0
2018_G	C109	-0.047	-0.047	0	0	0	0	-0.027	0	0	0
2018_G	C110	-0.055	0	0	0	0	0	0	-0.055	-0.093	0
2018_G	C114	0	0	0	0	0	0	0	0	0	0
2018_G	C115	-0.225	0	0	0	0	0	0	0	0	0
2018_G	C121	0	0	0	0	0	0	0	0	0	0
2018_G	C131	-0.043	0	0	0	0	0	0	0	0	0
2018_G	C133	-0.237	0	0	0	0	0	0	-0.155	0	0
2018_G	C135	0	0	0	0	0	0	0	0	0	0
2018_G	C136	0	0	0	0	0	0	-0.030	0	0	-0.030
2018_G	C137	0	0	0	0	0	0	0	0	0	0

Supplement Table 1: Plant species indices – continued (FG legume Season 2018 H)

	functional group	legume	legume	legume	legume	legume	legume	legume	legume	legume	legume
	family	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae	Leguminosae
	Code	Latpra	Lotcor	Medlup	Medspe	Tridub	Trifra	Tripra	Trirep	Viccra	Vicsep
	species	Lathyrus_praten	Lotus_corniculat	Medicago_lupuli	Medicago_spec.	Trifolium_dubiur	Trifolium_fragife	Trifolium_praten	Trifolium_repen	Vicia_cracca	Vicia_sepium
	season	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G	2018_E	2018_E,H	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G
Season	plot										
2018_H	A002	-0.094	0	0	0	0	0	0	0	0	0
2018_H	A003	-0.049	0	0	0	0	0	0	0	-0.028	0
2018_H	A005	-0.090	0	0	0	0	0	0	0	0	0
2018_H	A009	0	0	0	0	0	0	-0.260	0	0	0
2018_H	A010	-0.104	0	0	0	0	0	0	0	0	0
2018_H	A011	-0.083	0	0	0	-0.157	0	-0.049	0	0	0
2018_H	A013	-0.058	-0.058	0	0	-0.181	0	0	0	-0.097	0
2018_H	A016	-0.081	0	0	0	0	0	0	0	-0.027	0
2018_H	A018	-0.087	0	0	0	-0.087	0	0	0	-0.052	0
2018_H	A020	-0.041	0	0	0	-0.070	0	0	0	0	0
2018_H	A026	-0.087	-0.030	0	0	0	0	0	0	0	0
2018_H	A027	-0.112	0	0	0	0	0	0	0	0	0
2018_H	A030	-0.169	0	0	0	0	0	0	0	0	0
2018_H	A035	-0.128	0	0	0	-0.128	0	0	0	0	0
2018_H	A040	0	0	0	0	0	0	-0.161	0	0	0
2018_H	A042	-0.043	-0.074	0	0	0	0	0	0	0	0
2018_H	A043	-0.093	0	0	0	-0.151	0	0	0	-0.093	0
2018_H	A044	-0.104	0	0	0	-0.104	0	0	0	-0.104	0
2018_H	A045	-0.068	0	0	0	0	0	0	0	-0.133	0
2018_H	A046	-0.157	0	0	0	0	0	-0.028	0	-0.239	0
2018_H	B048	-0.083	0	0	0	0	0	0	0	-0.157	0
2018_H	B051	-0.105	-0.037	0	0	-0.195	0	0	0	0	0
2018_H	B054	-0.085	0	0	0	0	0	0	0	-0.138	0
2018_H	B057	-0.195	0	0	0	-0.195	0	0	0	-0.105	0
2018_H	B059	0	0	0	0	0	0	0	0	-0.110	0
2018_H	B060	0	0	0	0	0	0	0	0	-0.092	0
2018_H	B064	-0.102	-0.061	0	0	-0.102	0	0	0	0	0
2018_H	B067	-0.074	0	0	0	-0.143	0	0	0	-0.074	0
2018_H	B071	-0.077	0	0	0	0	0	0	0	-0.283	0
2018_H	B073	0	0	0	0	0	0	0	0	-0.141	0
2018_H	B075	0	0	0	0	0	0	0	-0.150	-0.150	0
2018_H	B080	-0.107	-0.107	0	0	0	0	0	0	0	0
2018_H	B081	0	0	0	0	0	-0.035	0	-0.035	0	0
2018_H	B085	-0.041	0	0	0	0	0	0	-0.136	-0.266	0
2018_H	B090	-0.086	0	0	0	0	0	0	0	0	0
2018_H	B092	-0.075	-0.025	0	0	0	0	0	0	0	0
2018_H	C097	0	0	0	0	0	0	0	0	0	0
2018_H	C103	-0.071	0	0	0	0	0	-0.042	0	0	0
2018_H	C109	-0.082	-0.049	0	0	0	0	0	0	0	0
2018_H	C110	0	0	0	0	0	0	0	0	-0.148	0
2018_H	C114	-0.032	0	0	0	0	0	0	0	0	0
2018_H	C115	-0.209	0	0	0	0	0	0	0	0	0
2018_H	C121	0	0	0	0	0	0	0	0	0	0
2018_H	C131	-0.071	0	0	0	0	0	0	0	0	0
2018_H	C133	-0.099	0	0	0	0	0	0	0	0	0
2018_H	C135	0	0	0	0	-0.166	0	0	0	0	0
2018_H	C136	0	0	0	0	0	0	-0.073	0	0	0
2018_H	C137	0	0	0	0	-0.225	0	0	0	0	0

Supplement Table 1: Plant species indices – continued (FG herb species A-B Season 2018 E)

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Amaranthaceae	Apiaceae	Apiaceae	Apiaceae	Apiaceae	Apiaceae	Apiaceae	Brassicaceae	Brassicaceae	Brassicaceae
	Code	Chealb	Antsyl	Carcar	Daucar	Hersph	Passat	Pimmaj	Arahir	Capbur	Sinalb
	species	Chenopodium_a	Anthriscus_sylve	Carum_carvi cf.	Daucus_carota	Heracleum_spho	Pastinaca_sativa	Pimpinella_majo	Arabis_hirsuta	Capsella_bursa-p	Sinapis_alba
	season	2018_E,F,G,H	2018_E,F,G	2018_E	2018_E,F,G,H	2018_E,F	2018_E,F,G	2018_E,F,G,H	2018_H	2018_E,F,G,H	2018_H
Season	plot										
2018_E	A002	0	0	0	-0.040	0	0	0	0	0	0
2018_E	A003	0	0	0	0	0	-0.042	-0.042	0	0	0
2018_E	A005	0	0	0	-0.060	0	0	-0.188	0	0	0
2018_E	A009	0	0	0	0	0	0	0	0	0	0
2018_E	A010	0	0	0	0	0	0	0	0	0	0
2018_E	A011	0	0	0	0	0	0	0	0	0	0
2018_E	A013	0	0	0	-0.030	0	0	-0.030	0	0	0
2018_E	A016	0	0	0	0	0	0	-0.038	0	0	0
2018_E	A018	0	0	0	0	0	0	0	0	0	0
2018_E	A020	-0.036	0	0	0	0	0	0	0	0	0
2018_E	A026	0	0	0	0	0	0	-0.108	0	0	0
2018_E	A027	0	0	0	0	0	0	0	0	0	0
2018_E	A030	0	0	0	0	0	0	0	0	0	0
2018_E	A035	0	0	0	0	0	0	0	0	0	0
2018_E	A040	0	0	0	0	0	0	0	0	0	0
2018_E	A042	0	0	0	0	0	0	0	0	0	0
2018_E	A043	0	0	0	0	0	0	0	0	-0.102	0
2018_E	A044	-0.040	0	0	0	0	0	0	0	0	0
2018_E	A045	0	0	0	0	0	0	0	0	0	0
2018_E	A046	0	0	0	0	-0.025	0	0	0	0	0
2018_E	B048	0	0	0	0	0	0	0	0	-0.085	0
2018_E	B051	0	0	0	0	0	0	0	0	0	0
2018_E	B054	0	0	0	0	0	0	0	0	0	0
2018_E	B057	0	0	0	0	0	0	-0.022	0	0	0
2018_E	B059	-0.125	0	0	0	0	0	0	0	-0.125	0
2018_E	B060	-0.031	0	0	0	0	0	0	0	-0.169	0
2018_E	B064	0	0	0	0	0	0	0	0	0	0
2018_E	B067	0	0	0	0	0	0	0	0	0	0
2018_E	B071	-0.030	0	0	0	0	0	0	0	-0.167	0
2018_E	B073	0	0	0	0	0	0	0	0	0	0
2018_E	B075	0	0	0	0	0	0	0	0	0	0
2018_E	B080	0	0	0	0	0	0	0	0	0	0
2018_E	B081	0	0	0	0	0	0	0	0	0	0
2018_E	B085	0	0	0	0	0	0	0	0	-0.032	0
2018_E	B090	0	0	0	0	0	0	0	0	0	0
2018_E	B092	0	0	0	0	0	0	-0.040	0	0	0
2018_E	C097	0	0	0	-0.054	0	0	0	0	0	0
2018_E	C103	0	0	0	0	0	0	0	0	0	0
2018_E	C109	0	0	0	0	0	0	0	0	0	0
2018_E	C110	-0.090	0	0	0	0	0	-0.031	0	0	0
2018_E	C114	0	0	0	0	0	0	0	0	0	0
2018_E	C115	0	0	0	0	0	0	0	0	0	0
2018_E	C121	0	0	-0.038	0	0	0	0	0	0	0
2018_E	C131	-0.029	-0.029	0	0	0	0	0	0	0	0
2018_E	C133	-0.181	0	0	0	0	0	0	0	-0.068	0
2018_E	C135	0	0	0	0	0	0	0	0	0	0
2018_E	C136	0	0	0	0	0	0	0	0	-0.030	0
2018_E	C137	0	0	0	0	0	0	0	0	0	0

Supplement Table 1: Plant species indices – continued (FG herb species A-B Season 2018 F)

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Amaranthaceae	Apiaceae	Apiaceae	Apiaceae	Apiaceae	Apiaceae	Apiaceae	Brassicaceae	Brassicaceae	Brassicaceae
	Code	Chealb	Antsyl	Carcar	Daucar	Hersph	Passat	Pimmaj	Arahir	Capbur	Sinalb
	species	Chenopodium_a	Anthriscus_sylve	Carum_carvi	cf. Daucus_carota	Heracleum_spho	Pastinaca_sativa	Pimpinella_majo	Arabis_hirsuta	Capsella_bursa-p	Sinapis_alba
	season	2018_E,F,G,H	2018_E,F,G	2018_E	2018_E,F,G,H	2018_E,F	2018_E,F,G	2018_E,F,G,H	2018_H	2018_E,F,G,H	2018_H
Season	plot										
2018_F	A002	0	0	0	0	0	0	0	0	0	0
2018_F	A003	0	0	0	0	0	0	0	0	0	0
2018_F	A005	0	0	0	-0.085	0	-0.085	0	0	0	0
2018_F	A009	0	0	0	0	0	0	0	0	0	0
2018_F	A010	0	0	0	0	0	0	0	0	0	0
2018_F	A011	0	0	0	0	0	0	-0.069	0	0	0
2018_F	A013	0	0	0	0	0	0	0	0	0	0
2018_F	A016	0	0	0	0	0	0	0	0	0	0
2018_F	A018	0	0	0	0	0	0	-0.041	0	0	0
2018_F	A020	0	0	0	0	0	0	0	0	0	0
2018_F	A026	0	0	0	0	0	0	0	0	0	0
2018_F	A027	0	0	0	0	0	0	0	0	0	0
2018_F	A030	0	0	0	0	0	0	0	0	0	0
2018_F	A035	0	0	0	0	0	0	0	0	0	0
2018_F	A040	0	0	0	-0.173	0	0	0	0	0	0
2018_F	A042	0	0	0	0	0	0	0	0	0	0
2018_F	A043	0	0	0	0	0	0	0	0	0	0
2018_F	A044	0	0	0	0	0	0	0	0	0	0
2018_F	A045	0	0	0	-0.092	0	0	0	0	0	0
2018_F	A046	0	0	0	0	-0.085	0	0	0	0	0
2018_F	B048	0	0	0	0	0	0	0	0	0	0
2018_F	B051	0	0	0	0	0	0	0	0	0	0
2018_F	B054	0	0	0	0	0	0	0	0	0	0
2018_F	B057	0	0	0	0	0	0	0	0	0	0
2018_F	B059	-0.218	0	0	0	0	0	0	0	-0.218	0
2018_F	B060	0	0	0	0	0	0	0	0	0	0
2018_F	B064	0	0	0	0	0	0	0	0	0	0
2018_F	B067	0	0	0	0	0	0	0	0	0	0
2018_F	B071	0	0	0	0	0	0	0	0	0	0
2018_F	B073	0	0	0	0	0	0	0	0	0	0
2018_F	B075	0	0	0	0	0	0	0	0	0	0
2018_F	B080	0	0	0	0	0	0	0	0	0	0
2018_F	B081	0	0	0	-0.072	0	0	0	0	0	0
2018_F	B085	0	0	0	-0.083	0	0	0	0	0	0
2018_F	B090	0	0	0	0	0	0	0	0	0	0
2018_F	B092	0	0	0	0	0	0	0	0	0	0
2018_F	C097	0	-0.165	0	0	0	0	0	0	0	0
2018_F	C103	0	0	0	0	0	0	0	0	0	0
2018_F	C109	0	0	0	0	0	0	0	0	0	0
2018_F	C110	-0.111	0	0	0	0	0	0	0	0	0
2018_F	C114	0	0	0	0	0	0	0	0	0	0
2018_F	C115	0	0	0	0	0	0	0	0	0	0
2018_F	C121	0	0	0	0	0	0	0	0	0	0
2018_F	C131	0	0	0	0	0	0	0	0	0	0
2018_F	C133	0	0	0	0	0	0	0	0	0	0
2018_F	C135	0	0	0	0	0	0	0	0	0	0
2018_F	C136	0	0	0	0	0	0	0	0	0	0
2018_F	C137	0	0	0	0	0	0	0	0	0	0

Supplement Table 1: Plant species indices – continued (FG herb species A-B Season 2018 G)

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Amaranthaceae	Apiaceae	Apiaceae	Apiaceae	Apiaceae	Apiaceae	Apiaceae	Brassicaceae	Brassicaceae	Brassicaceae
	Code	Chealb	Antsyl	Carcar	Daucar	Hersph	Passat	Pimmaj	Arahir	Capbur	Sinalb
	species	Chenopodium_a	Anthriscus_sylve	Carum_carvi cf.	Daucus_carota	Heracleum_spho	Pastinaca_sativa	Pimpinella_majo	Arabis_hirsuta	Capsella_bursa-p	Sinapis_alba
	season	2018_E,F,G,H	2018_E,F,G	2018_E	2018_E,F,G,H	2018_E,F	2018_E,F,G	2018_E,F,G,H	2018_H	2018_E,F,G,H	2018_H
Season	plot										
2018_G	A002	-0.050	0	0	0	0	0	0	0	0	0
2018_G	A003	-0.035	0	0	-0.035	0	-0.035	0	0	0	0
2018_G	A005	0	0	0	-0.078	0	0	-0.046	0	0	0
2018_G	A009	0	-0.045	0	-0.045	0	0	0	0	0	0
2018_G	A010	0	0	0	0	0	0	0	0	0	0
2018_G	A011	0	0	0	0	0	0	-0.045	0	0	0
2018_G	A013	0	0	0	-0.048	0	0	0	0	0	0
2018_G	A016	-0.094	0	0	0	0	0	0	0	-0.056	0
2018_G	A018	-0.048	0	0	0	0	0	-0.048	0	0	0
2018_G	A020	0	0	0	0	0	0	0	0	0	0
2018_G	A026	0	0	0	0	0	0	0	0	-0.046	0
2018_G	A027	-0.067	0	0	0	0	0	0	0	-0.111	0
2018_G	A030	0	0	0	0	0	0	0	0	0	0
2018_G	A035	-0.083	0	0	0	0	0	0	0	0	0
2018_G	A040	-0.091	0	0	-0.091	0	0	0	0	0	0
2018_G	A042	0	0	0	0	0	0	0	0	0	0
2018_G	A043	0	0	0	0	0	0	0	0	0	0
2018_G	A044	-0.073	0	0	-0.073	0	0	0	0	0	0
2018_G	A045	0	0	0	0	0	0	0	0	0	0
2018_G	A046	0	0	0	0	0	-0.054	0	0	0	0
2018_G	B048	0	0	0	0	0	0	0	0	0	0
2018_G	B051	0	0	0	0	0	0	0	0	0	0
2018_G	B054	0	0	0	0	0	0	0	0	0	0
2018_G	B057	0	0	0	0	0	0	0	0	0	0
2018_G	B059	-0.145	0	0	0	0	0	0	0	-0.145	0
2018_G	B060	-0.087	0	0	0	0	0	0	0	-0.143	0
2018_G	B064	0	0	0	0	0	0	0	0	0	0
2018_G	B067	0	0	0	0	0	0	0	0	-0.032	0
2018_G	B071	-0.099	0	0	0	0	0	0	0	-0.159	0
2018_G	B073	0	0	0	0	0	0	0	0	0	0
2018_G	B075	0	0	0	0	0	0	0	0	0	0
2018_G	B080	0	0	0	0	0	0	-0.054	0	0	0
2018_G	B081	0	0	0	0	0	0	0	0	0	0
2018_G	B085	0	0	0	0	0	0	0	0	0	0
2018_G	B090	0	0	0	0	0	0	0	0	0	0
2018_G	B092	0	0	0	0	0	0	0	0	-0.048	0
2018_G	C097	-0.057	0	0	-0.057	0	0	0	0	-0.057	0
2018_G	C103	-0.035	0	0	0	0	0	0	0	0	0
2018_G	C109	0	0	0	0	0	0	-0.027	0	0	0
2018_G	C110	-0.055	0	0	0	0	0	0	0	-0.093	0
2018_G	C114	0	0	0	0	0	0	0	0	0	0
2018_G	C115	-0.045	0	0	0	0	0	-0.045	0	-0.125	0
2018_G	C121	0	0	0	0	0	0	0	0	-0.236	0
2018_G	C131	0	0	0	0	0	0	0	0	0	0
2018_G	C133	-0.155	0	0	0	0	0	0	0	-0.155	0
2018_G	C135	-0.041	0	0	0	0	0	0	0	-0.211	0
2018_G	C136	-0.030	0	0	0	0	0	0	0	0	0
2018_G	C137	0	0	0	0	0	0	-0.039	0	0	0

Supplement Table 1: Plant species indices – continued (FG herb species A-B Season 2018 H)

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Amaranthaceae	Apiaceae	Apiaceae	Apiaceae	Apiaceae	Apiaceae	Apiaceae	Brassicaceae	Brassicaceae	Brassicaceae
	Code	Chealb	Antsyl	Carcar	Daucar	Hersph	Passat	Pimmaj	Arahir	Capbur	Sinalb
	species	Chenopodium_a	Anthriscus_sylve	Carum_carvi cf.	Daucus_carota	Heracleum_spho	Pastinaca_sativa	Pimpinella_majo	Arabis_hirsuta	Capsella_bursa-p	Sinapis_alba
	season	2018_E,F,G,H	2018_E,F,G	2018_E	2018_E,F,G,H	2018_E,F	2018_E,F,G	2018_E,F,G,H	2018_H	2018_E,F,G,H	2018_H
Season	plot										
2018_H	A002	0	0	0	0	0	0	0	0	-0.177	0
2018_H	A003	0	0	0	0	0	0	-0.028	0	-0.049	0
2018_H	A005	0	0	0	-0.053	0	0	-0.031	0	-0.255	0
2018_H	A009	0	0	0	-0.092	0	0	-0.055	0	0	0
2018_H	A010	0	0	0	0	0	0	0	0	0	0
2018_H	A011	0	0	0	0	0	0	-0.049	-0.028	-0.157	0
2018_H	A013	0	0	0	-0.097	0	0	-0.033	0	-0.269	0
2018_H	A016	0	0	0	0	0	0	0	0	-0.327	0
2018_H	A018	0	0	0	-0.030	0	0	-0.030	0	-0.087	0
2018_H	A020	0	0	0	0	0	0	0	0	-0.117	0
2018_H	A026	0	0	0	0	0	0	0	0	-0.087	0
2018_H	A027	0	0	0	0	0	0	0	0	-0.296	0
2018_H	A030	0	0	0	0	0	0	0	0	0	0
2018_H	A035	0	0	0	0	0	0	0	0	0	-0.046
2018_H	A040	0	0	0	-0.059	0	0	0	0	-0.161	0
2018_H	A042	0	0	0	0	0	0	0	0	0	0
2018_H	A043	0	0	0	0	0	0	0	-0.151	-0.262	0
2018_H	A044	0	0	0	-0.036	0	0	0	0	-0.193	0
2018_H	A045	0	0	0	-0.023	0	0	-0.023	0	-0.068	0
2018_H	A046	0	0	0	0	0	0	-0.028	0	0	0
2018_H	B048	0	0	0	0	0	0	0	0	-0.083	0
2018_H	B051	0	0	0	0	0	0	0	0	-0.105	0
2018_H	B054	0	0	0	0	0	0	0	0	-0.244	0
2018_H	B057	0	0	0	0	0	0	0	0	0	0
2018_H	B059	0	0	0	0	0	0	0	0	-0.288	0
2018_H	B060	-0.092	0	0	0	0	0	0	0	-0.347	0
2018_H	B064	0	0	0	0	0	0	0	0	-0.189	0
2018_H	B067	0	0	0	0	0	0	0	0	0	0
2018_H	B071	-0.046	0	0	0	0	0	0	0	-0.283	0
2018_H	B073	0	0	0	-0.025	0	0	0	0	0	0
2018_H	B075	0	0	0	0	0	0	0	0	0	0
2018_H	B080	0	0	0	0	0	0	-0.037	0	0	0
2018_H	B081	0	0	0	-0.100	0	0	0	0	-0.274	0
2018_H	B085	0	0	0	0	0	0	0	0	0	0
2018_H	B090	0	0	0	0	0	0	0	0	0	0
2018_H	B092	-0.025	0	0	0	0	0	0	0	-0.277	0
2018_H	C097	-0.029	0	0	-0.050	0	0	0	0	-0.244	0
2018_H	C103	0	0	0	0	0	0	0	0	-0.138	0
2018_H	C109	0	0	0	0	0	0	-0.028	0	0	0
2018_H	C110	-0.026	0	0	0	0	0	0	0	-0.228	0
2018_H	C114	0	0	0	0	0	0	0	0	0	0
2018_H	C115	-0.040	0	0	0	0	0	0	0	-0.209	0
2018_H	C121	0	0	0	0	0	0	0	0	-0.253	0
2018_H	C131	-0.071	0	0	0	0	0	0	0	0	0
2018_H	C133	-0.099	0	0	0	0	0	0	0	-0.365	0
2018_H	C135	0	0	0	0	0	0	0	0	-0.166	0
2018_H	C136	0	0	0	0	0	0	0	0	0	0
2018_H	C137	0	0	0	0	0	0	-0.045	0	-0.225	0

Supplement Table 1: Plant species indices – continued (FG herb species C Season 2018 E)

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Campanulaceae	Caprifoliaceae	Caryophyllaceae	Caryophyllaceae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae
	Code	Campat	Knaarv	Cerhol	Silniv	Achmil	Belper	Carcri	Cenjac	Cirarv	Cirole	Crebie	Eriann	Erican
	species	**Campanula_pa**	**Knautia_arve**	**Cerastium_holoste**	**Silene_nivea**	**Achillea_millefol**	**Bellis_peren**	**Carduus_cris**	**Centaurea_ja**	**Cirsium_arve**	**Cirsium_olerac**	**Crepis_bien**	**Erigeron_ann**	**Erigeron_canad**
	season	2018_E	2018_E,F,G,H	2018_E,F,G,H	2018_E,G,H	2018_E,G,H	2018_E,F,H	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_F,G,H	2018_F	2018_E,F,G,H	2018_H
Seaso	**plot**													
2018_E	A002	0	0	-0.113	0	0	-0.040	0	0	0	0	0	0	0
2018_E	A003	0	0	-0.118	0	0	0	0	0	0	0	0	0	0
2018_E	A005	0	0	-0.101	0	0	0	0	0	0	0	0	0	0
2018_E	A009	0	-0.071	-0.118	0	0	0	0	-0.071	0	0	0	-0.042	0
2018_E	A010	0	-0.352	-0.053	0	0	-0.053	0	-0.255	0	0	0	0	0
2018_E	A011	-0.033	-0.057	-0.057	0	-0.033	0	0	-0.057	-0.033	0	0	0	0
2018_E	A013	-0.030	0	-0.167	0	0	-0.088	0	-0.340	0	0	0	0	0
2018_E	A016	0	-0.038	-0.110	0	-0.038	0	0	0	0	0	0	0	0
2018_E	A018	0	0	-0.162	0	0	-0.029	0	0	0	0	0	0	0
2018_E	A020	0	-0.314	-0.281	0	0	0	0	0	0	0	0	0	0
2018_E	A026	0	-0.038	-0.108	0	0	0	0	-0.038	0	0	0	0	0
2018_E	A027	0	-0.052	-0.144	0	0	-0.088	0	-0.312	0	0	0	0	0
2018_E	A030	0	-0.368	-0.287	0	0	0	0	-0.107	0	0	0	0	0
2018_E	A035	0	0	-0.108	0	0	0	0	0	0	0	0	0	0
2018_E	A040	0	-0.110	-0.293	0	0	0	0	0	0	0	0	0	0
2018_E	A042	0	-0.297	-0.306	0	0	0	0	-0.118	0	0	0	0	0
2018_E	A043	0	-0.189	-0.279	0	0	0	0	0	0	0	0	0	0
2018_E	A044	0	-0.040	-0.209	0	0	0	0	0	0	0	0	0	0
2018_E	A045	0	0	-0.102	0	0	0	0	0	0	0	0	0	0
2018_E	A046	-0.025	0	-0.142	-0.025	0	0	0	0	0	0	0	0	0
2018_E	B048	0	0	-0.161	-0.029	0	0	0	0	0	0	0	0	0
2018_E	B051	0	0	-0.074	0	0	0	0	0	0	0	0	0	0
2018_E	B054	0	-0.054	-0.091	0	0	0	0	-0.091	0	0	0	0	0
2018_E	B057	-0.022	0	-0.320	0	0	0	0	-0.253	0	0	0	0	0
2018_E	B059	0	-0.045	-0.359	0	0	0	0	-0.076	0	0	0	0	0
2018_E	B060	0	0	-0.309	0	0	0	0	0	0	0	0	0	0
2018_E	B064	0	0	-0.319	0	0	0	0	0	0	0	0	0	0
2018_E	B067	0	0	-0.073	0	0	0	0	-0.073	0	0	0	0	0
2018_E	B071	0	-0.052	-0.340	0	0	0	0	0	-0.052	0	0	0	0
2018_E	B073	0	-0.333	-0.243	0	0	0	0	-0.159	0	0	0	0	0
2018_E	B075	0	0	0	0	0	0	0	-0.031	0	0	0	0	0
2018_E	B080	0	0	-0.230	0	0	0	0	0	0	0	0	0	0
2018_E	B081	-0.032	0	-0.262	0	0	0	0	0	0	0	0	0	0
2018_E	B085	0	-0.032	-0.094	0	0	0	0	-0.094	0	0	0	0	0
2018_E	B090	0	0	-0.087	0	0	0	0	-0.087	0	0	0	0	0
2018_E	B092	0	-0.040	-0.068	0	0	0	0	0	-0.040	0	0	0	0
2018_E	C097	0	0	-0.257	0	0	0	0	0	0	0	0	0	0
2018_E	C103	0	-0.366	-0.088	0	0	0	0	-0.052	0	0	0	0	0
2018_E	C109	0	0	-0.189	0	0	0	0	0	0	0	0	0	0
2018_E	C110	0	0	-0.169	0	0	-0.053	0	0	0	0	0	0	0
2018_E	C114	0	-0.075	-0.145	0	0	0	0	0	0	0	0	0	0
2018_E	C115	0	-0.194	-0.194	0	0	0	0	0	0	0	0	0	0
2018_E	C121	0	-0.108	-0.108	0	0	0	0	0	-0.038	0	0	0	0
2018_E	C131	0	0	-0.162	0	0	0	0	0	0	0	0	0	0
2018_E	C133	0	-0.181	-0.068	0	0	0	0	0	-0.181	0	0	0	0
2018_E	C135	0	-0.342	-0.104	0	0	0	0	-0.104	0	0	0	0	0
2018_E	C136	0	-0.089	-0.053	0	0	0	0	-0.053	0	0	0	0	0
2018_E	C137	0	-0.101	-0.060	0	0	0	0	-0.060	0	0	0	0	0

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae
	Code	Galpar	Lacser	Leoaut	Leuvul	Scoaut	Solcan	Solvir	Sonarv	Sonasp	Sonole	Taroff
	species	**Galinsoga_parvi**	**Lactuca_serr**	**Leontodon_autum**	**Leucanthemum_vul**	**Scorzoneroides_autur**	**Solidago_canad**	**Solidago_virga**	**Sonchus_arve**	**Sonchus_asp**	**Sonchus_olera**	**Taraxacum_officinale agg.**
	season	2018_E,F,G	2018_E,G,H	2018_H	2018_E,F,G,H	2018_F,G	2018_E,G,H	2018_F	2018_G	2018_E	2018_E	2018_E,F,G,H
Seaso	**plot**											
2018_E	A002	0	0	0	-0.068	0	-0.068	0	0	0	0	-0.206
2018_E	A003	0	0	0	-0.352	0	-0.042	0	0	0	0	0
2018_E	A005	0	0	0	0	0	0	0	0	0	0	-0.101
2018_E	A009	-0.042	0	0	-0.071	0	0	0	0	-0.042	0	0
2018_E	A010	0	0	0	-0.090	0	0	0	0	0	0	0
2018_E	A011	0	0	0	-0.338	0	0	0	0	0	0	0
2018_E	A013	0	0	0	0	0	-0.252	0	0	0	0	0
2018_E	A016	0	0	0	-0.201	0	-0.038	0	0	-0.066	0	-0.342
2018_E	A018	0	0	0	-0.246	0	0	0	0	0	0	0
2018_E	A020	-0.036	0	0	-0.036	0	-0.062	0	0	-0.036	0	0
2018_E	A026	0	0	0	-0.038	0	-0.108	0	0	0	0	0
2018_E	A027	0	0	0	-0.144	0	0	0	0	0	0	-0.088
2018_E	A030	0	0	0	-0.338	0	0	0	0	0	0	0
2018_E	A035	0	0	0	0	0	0	0	0	0	0	0
2018_E	A040	0	-0.039	0	-0.202	0	-0.110	0	0	-0.039	0	-0.202
2018_E	A042	0	0	0	0	0	0	0	0	0	0	-0.071
2018_E	A043	0	0	0	-0.061	0	-0.102	0	0	0	0	-0.102
2018_E	A044	0	0	0	-0.115	0	0	0	0	0	0	0
2018_E	A045	0	0	0	0	0	0	0	0	0	0	0
2018_E	A046	0	0	0	-0.365	0	0	0	0	0	0	0
2018_E	B048	-0.050	0	0	-0.346	0	-0.161	0	0	0	0	0
2018_E	B051	0	0	0	-0.355	0	0	0	0	0	0	0
2018_E	B054	0	0	0	-0.091	0	0	0	0	0	0	0
2018_E	B057	0	0	0	-0.201	0	0	0	0	0	0	0
2018_E	B059	0	0	0	0	0	-0.125	0	0	0	-0.045	0
2018_E	B060	0	0	0	0	0	-0.031	0	0	0	-0.031	-0.090
2018_E	B064	0	-0.095	0	0	0	-0.178	0	0	0	0	0
2018_E	B067	0	0	0	-0.365	0	0	0	0	0	0	0
2018_E	B071	0	-0.030	0	-0.088	0	0	0	0	0	0	-0.052
2018_E	B073	0	0	0	-0.297	0	0	0	0	0	0	0
2018_E	B075	-0.053	0	0	-0.331	0	0	0	0	0	0	-0.031
2018_E	B080	0	-0.026	0	-0.046	0	0	0	0	0	-0.026	-0.026
2018_E	B081	0	0	0	0	0	0	0	0	0	0	0
2018_E	B085	0	0	0	-0.317	0	0	0	0	0	0	0
2018_E	B090	0	0	0	-0.339	0	0	0	0	0	0	0
2018_E	B092	-0.068	0	0	-0.206	0	0	0	0	0	0	-0.040
2018_E	C097	0	0	0	-0.171	0	-0.054	0	0	0	0	0
2018_E	C103	-0.052	0	0	-0.251	0	0	0	0	0	0	0
2018_E	C109	-0.035	0	0	-0.365	0	0	0	0	0	0	0
2018_E	C110	0	0	0	-0.090	0	-0.090	0	0	0	-0.031	-0.255
2018_E	C114	0	0	0	-0.044	0	0	0	0	0	0	0
2018_E	C115	0	0	0	0	0	0	0	0	0	0	-0.123
2018_E	C121	0	0	0	-0.108	0	-0.065	0	0	-0.065	0	0
2018_E	C131	0	0	0	-0.162	0	0	0	0	0	0	0
2018_E	C133	0	0	0	0	0	0	0	0	0	0	-0.299
2018_E	C135	0	0	0	-0.333	0	0	0	0	0	0	0
2018_E	C136	0	0	0	-0.354	0	0	0	0	0	0	0
2018_E	C137	0	0	0	-0.101	0	0	0	0	-0.035	0	0

Supplement Table 1: Plant species indices – continued (FG herb species C Season 2018 F)

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Campanulaceae	Caprifoliaceae	Caryophyllaceae	Caryophyllaceae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae
	Code	Campat	Knaarv	Cerhol	Silniv	Achmil	Belper	Carcri	Cenjac	Cirarv	Cirole	Crebie	Eriann	Erican	Galpar	Lacser	Leoaut	Leuvul	Scoaut	Solcan	Solvir	Sonarv	Sonasp	Sonole	Taroff
	species	**Campanula_pa**	**Knautia_arve**	**Cerastium_holoste**	**Silene_nivea**	**Achillea_millefol**	**Bellis_peren**	**Carduus_cris**	**Centaurea_ja**	**Cirsium_arve**	**Cirsium_olerac**	**Crepis_bien**	**Erigeron_ann**	**Erigeron_canade**	**Galinsoga_parvi**	**Lactuca_serr**	**Leontodon_autumı**	**Leucanthemum_vul**	**Scorzoneroides_autur**	**Solidago_canad**	**Solidago_virga**	**Sonchus_arve**	**Sonchus_asp**	**Sonchus_olera**	**Taraxacum_officinale agg.**
	season	2018_E	2018_E,F,G,H	2018_E,F,G,H	2018_E,G,H	2018_E,G,H	2018_E,F,H	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_F,G,H	2018_F	2018_E,F,G,H	2018_H	2018_E,F,G	2018_E,G,H	2018_H	2018_E,F,G,H	2018_F,G	2018_E,G,H	2018_F	2018_G	2018_E	2018_E	2018_E,F,G,H
Seaso	**plot**																								
2018_F	A002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.334
2018_F	A003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.343	0	0	0	0	0	0	-0.146
2018_F	A005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.085
2018_F	A009	0	0	0	0	0	0	0	0	0	0	0	-0.111	0	0	0	0	0	0	0	-0.111	0	0	0	-0.177
2018_F	A010	0	-0.240	0	0	0	0	0	-0.189	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.072
2018_F	A011	0	0	0	0	0	0	0	-0.069	0	0	0	0	0	0	0	0	-0.344	0	0	0	0	0	0	-0.115
2018_F	A013	0	0	0	0	0	0	0	-0.364	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2018_F	A016	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.086	0	0	-0.086	0	0	0	-0.218
2018_F	A018	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.303	0	0	0	0	0	0	0
2018_F	A020	0	-0.270	0	0	0	0	0	0	0	-0.085	0	0	0	0	0	0	0	0	0	-0.085	0	0	0	0
2018_F	A026	0	0	0	0	0	0	0	0	0	0	-0.061	-0.061	0	0	0	0	0	0	0	0	0	0	0	0
2018_F	A027	0	0	0	0	0	-0.087	0	-0.251	0	0	0	-0.221	0	0	0	0	0	0	0	-0.087	0	0	0	-0.143
2018_F	A030	0	-0.347	0	0	0	0	0	-0.150	0	0	0	-0.150	0	0	0	0	-0.347	0	0	0	0	0	0	0
2018_F	A035	0	0	0	0	0	0	0	0	0	-0.046	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2018_F	A040	0	0	0	0	0	0	0	0	0	-0.173	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.173
2018_F	A042	0	-0.173	0	0	0	0	0	-0.177	0	0	0	-0.111	0	0	0	0	0	0	0	0	0	0	0	0
2018_F	A043	0	0	0	0	0	0	0	0	-0.078	0	-0.078	0	0	0	0	0	0	0	0	0	0	0	0	-0.129
2018_F	A044	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.143
2018_F	A045	0	-0.092	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.092	0	0	0	0	0	0	0
2018_F	A046	0	-0.085	0	0	0	-0.085	0	0	0	0	0	0	0	0	0	0	-0.270	0	0	0	0	0	0	-0.085
2018_F	B048	0	0	0	0	0	0	0	0	0	0	0	-0.076	0	0	0	0	-0.317	0	0	-0.076	0	0	0	0
2018_F	B051	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.364	0	0	0	0	0	0	0
2018_F	B054	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2018_F	B057	0	0	0	0	0	0	0	-0.355	0	0	0	0	0	0	0	0	-0.099	0	0	0	0	0	0	-0.059
2018_F	B059	0	0	-0.218	0	0	0	0	0	0	0	0	-0.218	0	0	0	0	0	0	0	0	0	0	0	-0.218
2018_F	B060	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.132
2018_F	B064	0	0	0	0	0	0	0	0	0	0	0	-0.059	0	0	0	0	-0.161	0	0	-0.059	0	0	0	0
2018_F	B067	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.356	0	0	0	0	0	0	0
2018_F	B071	0	0	0	0	0	0	0	0	-0.091	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2018_F	B073	0	-0.362	0	0	0	0	0	-0.170	0	0	0	0	0	0	0	0	-0.170	0	0	0	0	0	0	0
2018_F	B075	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.368	0	0	0	0	0	0	0
2018_F	B080	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.070	0	0	0	0	0	0	-0.070
2018_F	B081	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.072	0	0	0	0	0	0	0
2018_F	B085	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.279	0	0	0	0	0	0	0
2018_F	B090	0	-0.063	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.218	0	0	0	0	0	0	0
2018_F	B092	0	0	0	0	0	0	0	0	0	0	0	-0.094	0	0	0	0	0	0	0	0	0	0	0	-0.094
2018_F	C097	0	0	0	0	0	0	0	0	0	-0.061	0	-0.061	0	0	0	0	0	0	0	0	0	0	0	0
2018_F	C103	0	-0.315	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.250	0	0	0	0	0	0	0
2018_F	C109	0	-0.135	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.365	0	0	0	0	0	0	0
2018_F	C110	0	0	0	0	0	0	0	0	-0.111	0	0	-0.111	0	0	0	0	0	-0.111	0	0	0	0	0	-0.264
2018_F	C114	0	0	0	0	0	0	0	0	0	-0.043	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.043
2018_F	C115	0	0	0	0	0	0	0	0	-0.057	-0.057	0	0	0	0	0	0	-0.155	0	0	0	0	0	0	-0.096
2018_F	C121	0	0	0	0	0	0	0	-0.129	-0.078	-0.078	0	-0.129	0	-0.322	0	0	0	0	0	0	0	0	0	0
2018_F	C131	0	0	0	0	0	0	0	0	0	-0.051	0	0	0	-0.051	0	0	0	0	0	0	0	0	0	-0.051
2018_F	C133	0	0	0	0	0	0	0	0	-0.218	0	0	0	0	-0.218	0	0	0	0	0	0	0	0	0	-0.218
2018_F	C135	0	-0.283	0	0	0	0	0	0	0	-0.067	0	0	0	0	0	0	-0.294	0	0	0	0	0	0	0
2018_F	C136	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.366	0	0	0	0	0	0	0
2018_F	C137	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.058	0	0	-0.269	0	0	0	0	0	0	-0.156

Supplement Table 1: Plant species indices – continued (FG herb species C Season 2018 G)

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Campanulaceae	Caprifoliaceae	Caryophyllaceae	Caryophyllaceae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae
	Code	Campat	Knaarv	Cerhol	Silniv	Achmil	Belper	Carcri	Cenjac	Cirarv	Cirole	Crebie	Eriann	Erican	Galpar	Lacser	Leoaut	Leuvul	Scoaut	Solcan	Solvir	Sonarv	Sonasp	Sonole	Taroff
	species	**Campanula_pa**	**Knautia_arve**	**Cerastium_holoste**	**Silene_nivea**	**Achillea_millefol**	**Bellis_peren**	**Carduus_cris**	**Centaurea_ja**	**Cirsium_arve**	**Cirsium_olerac**	**Crepis_bien**	**Erigeron_ann**	**Erigeron_canad**	**Galinsoga_parvi**	**Lactuca_serr**	**Leontodon_autum**	**Leucanthemum_vul**	**Scorzoneroides_autur**	**Solidago_canad**	**Solidago_virga**	**Sonchus_arve**	**Sonchus_asp**	**Sonchus_olera**	**Taraxacum_officinale agg.**
	season	2018_E	2018_E,F,G,H	2018_E,F,G,H	2018_E,G,H	2018_E,G,H	2018_E,F,H	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_F,G,H	2018_F	2018_E,F,G,H	2018_H	2018_E,F,G	2018_E,G,H	2018_H	2018_E,F,G,H	2018_F,G	2018_E,G,H	2018_F	2018_G	2018_E	2018_E	2018_E,F,G,H
Seaso	**plot**																								
2018_G	A002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.244
2018_G	A003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.356	0	0	0	0	0	0	-0.100
2018_G	A005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2018_G	A009	0	-0.125	0	0	0	0	0	-0.045	0	-0.076	0	-0.125	0	0	0	0	-0.225	-0.045	-0.076	0	0	0	0	-0.125
2018_G	A010	0	-0.349	0	0	0	0	0	-0.228	0	0	0	0	0	0	0	0	-0.127	0	0	0	0	0	0	0
2018_G	A011	0	-0.076	0	0	0	0	0	-0.125	0	0	0	-0.045	0	0	0	0	-0.368	0	-0.045	0	0	0	0	-0.125
2018_G	A013	0	-0.236	0	0	0	0	0	-0.363	0	0	0	-0.081	0	0	0	0	-0.048	0	-0.048	0	0	0	0	-0.048
2018_G	A016	0	0	0	0	-0.056	0	0	0	0	0	0	-0.056	0	0	0	0	-0.152	0	0	0	-0.056	0	0	-0.263
2018_G	A018	0	-0.048	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.329	0	0	0	0	0	0	0
2018_G	A020	0	-0.334	0	0	0	0	0	0	0	0	0	-0.049	0	0	0	0	-0.135	0	0	0	0	0	0	0
2018_G	A026	0	-0.078	0	0	0	0	0	-0.046	0	0	0	-0.046	0	0	0	0	-0.046	0	0	0	0	0	0	-0.078
2018_G	A027	0	-0.177	0	0	-0.067	0	0	-0.351	0	0	0	-0.177	0	0	0	0	0	0	-0.111	0	-0.067	0	0	-0.177
2018_G	A030	0	-0.361	0	0	0	0	0	-0.352	0	0	0	-0.042	0	0	0	0	-0.306	0	-0.042	0	0	0	0	0
2018_G	A035	0	-0.049	0	0	0	0	-0.049	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.135
2018_G	A040	0	-0.227	-0.091	0	0	0	0	0	0	-0.091	0	-0.147	0	-0.091	0	0	-0.147	0	-0.091	0	0	0	0	-0.227
2018_G	A042	0	-0.202	0	0	0	0	0	-0.152	0	0	0	-0.056	0	0	0	0	0	0	0	0	0	0	0	-0.152
2018_G	A043	0	-0.274	0	0	0	0	-0.059	-0.100	0	0	0	-0.059	0	0	-0.059	0	0	0	0	0	0	0	0	-0.161
2018_G	A044	0	-0.121	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.191
2018_G	A045	0	-0.075	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.075	0	0	0	0	0	0	-0.075
2018_G	A046	0	-0.091	-0.054	-0.149	0	0	0	0	0	0	0	0	0	0	0	0	-0.298	0	0	0	0	0	0	-0.149
2018_G	B048	0	-0.101	0	-0.060	0	0	0	0	0	0	0	-0.060	0	0	0	0	-0.323	0	-0.101	0	0	0	0	0
2018_G	B051	0	0	0	0	0	0	0	-0.041	0	0	0	0	0	0	0	0	-0.367	0	0	0	0	0	0	0
2018_G	B054	0	-0.288	0	0	0	0	0	-0.107	0	0	0	0	0	0	0	0	0	0	-0.064	0	0	0	0	-0.107
2018_G	B057	0	-0.118	0	0	0	0	0	-0.306	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2018_G	B059	0	-0.145	0	0	0	0	0	0	-0.145	0	0	0	0	0	0	0	0	-0.145	-0.145	0	0	0	0	0
2018_G	B060	0	0	-0.052	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.087	-0.052	-0.052	0	0	0	0	-0.087
2018_G	B064	0	0	-0.052	0	0	0	-0.052	0	0	0	0	-0.052	0	0	0	0	-0.052	0	-0.052	0	0	0	0	0
2018_G	B067	0	-0.032	-0.032	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.313	0	0	0	0	0	0	0
2018_G	B071	0	-0.059	0	0	0	0	0	0	-0.159	0	0	-0.059	0	-0.059	0	0	0	0	0	0	0	0	0	-0.159
2018_G	B073	0	-0.328	0	0	0	0	0	-0.187	0	0	0	0	0	0	0	0	-0.328	0	0	0	0	0	0	0
2018_G	B075	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.367	0	0	0	0	0	0	0
2018_G	B080	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.147	0	-0.054	0	0	0	0	-0.091
2018_G	B081	0	0	0	0	0	0	0	0	0	0	0	-0.056	0	0	0	0	-0.094	0	0	0	0	0	0	0
2018_G	B085	0	-0.238	-0.048	0	0	0	0	0	-0.048	0	0	-0.048	0	0	0	0	-0.336	0	0	0	0	0	0	-0.082
2018_G	B090	0	-0.084	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.202	0	0	0	0	0	0	0
2018_G	B092	0	0	-0.048	0	0	0	0	0	-0.048	0	0	-0.048	0	0	0	0	-0.081	0	0	0	0	0	0	-0.048
2018_G	C097	0	-0.057	-0.057	0	0	0	0	-0.057	-0.057	-0.057	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2018_G	C103	0	-0.320	0	0	0	0	0	0	-0.035	0	0	0	0	0	0	0	-0.278	0	0	0	0	0	0	0
2018_G	C109	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.363	0	0	0	0	0	0	0
2018_G	C110	0	-0.093	0	0	0	0	0	0	0	-0.055	0	-0.055	0	0	0	0	-0.151	-0.093	-0.055	0	0	0	0	-0.262
2018_G	C114	0	-0.043	0	0	0	0	0	0	0	-0.043	0	0	0	0	-0.043	0	0	0	0	0	0	0	0	-0.043
2018_G	C115	0	-0.125	0	0	0	0	0	0	0	-0.045	0	0	0	0	0	0	-0.045	0	-0.045	0	0	0	0	0
2018_G	C121	0	0	0	0	0	0	-0.048	0	0	-0.132	0	-0.327	0	0	0	0	0	0	-0.048	0	0	0	0	-0.081
2018_G	C131	0	-0.043	0	0	0	0	0	0	0	-0.043	0	-0.043	0	0	0	0	0	0	0	0	-0.043	0	0	0
2018_G	C133	0	0	0	0	0	0	0	0	0	-0.155	0	0	0	0	0	0	0	0	-0.155	0	0	0	0	-0.155
2018_G	C135	0	-0.256	0	0	0	0	0	0	-0.041	-0.041	0	0	0	0	0	0	-0.303	0	0	0	0	0	0	0
2018_G	C136	0	-0.088	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.366	0	0	0	0	0	0	0
2018_G	C137	0	-0.293	0	0	0	0	0	0	0	-0.066	0	-0.039	0	0	0	0	0	0	0	0	0	0	0	0

Supplement Table 1: Plant species indices – continued (FG herb species C Season 2018 H)

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Campanulaceae	Caprifoliaceae	Caryophyllaceae	Caryophyllaceae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae
	Code	Campat	Knaarv	Cerhol	Silniv	Achmil	Belper	Carcri	Cenjac	Cirarv	Cirole	Crebie	Eriann
	species	**Campanula_pa**	**Knautia_arve**	**Cerastium_holoste**	**Silene_nivea**	**Achillea_millefol**	**Bellis_peren**	**Carduus_cris**	**Centaurea_ja**	**Cirsium_arve**	**Cirsium_olerac**	**Crepis_bien**	**Erigeron_ann**
	season	2018_E	2018_E,F,G,H	2018_E,F,G,H	2018_E,G,H	2018_E,G,H	2018_E,F,H	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_F,G,H	2018_F	2018_E,F,G,H
Season	**plot**												
2018_H	A002	0	0	0	0	0	0	0	-0.094	0	0	0	-0.033
2018_H	A003	0	0	0	0	0	0	0	0	0	0	0	0
2018_H	A005	0	-0.090	0	0	0	0	0	0	0	0	0	0
2018_H	A009	0	-0.150	-0.055	0	0	-0.055	0	-0.150	-0.055	0	0	-0.092
2018_H	A010	0	-0.360	0	0	0	0	0	-0.193	0	0	0	0
2018_H	A011	0	-0.157	-0.028	0	-0.028	-0.083	0	-0.157	0	0	0	0
2018_H	A013	0	-0.181	0	0	0	-0.058	0	-0.322	0	0	0	-0.058
2018_H	A016	0	0	0	0	-0.027	0	0	0	0	0	0	0
2018_H	A018	0	-0.165	0	0	0	0	0	0	0	0	0	0
2018_H	A020	0	-0.339	0	0	0	0	0	0	0	0	0	0
2018_H	A026	0	-0.087	-0.030	0	-0.030	0	0	-0.030	0	0	0	-0.051
2018_H	A027	0	-0.205	-0.112	0	-0.039	-0.067	0	-0.345	-0.039	0	0	-0.205
2018_H	A030	0	-0.342	0	0	0	-0.031	0	-0.169	0	0	0	0
2018_H	A035	0	-0.128	0	0	0	0	0	0	-0.046	0	0	0
2018_H	A040	0	-0.274	-0.059	0	0	0	-0.059	0	0	0	0	-0.059
2018_H	A042	0	-0.327	-0.122	0	0	0	0	-0.312	0	0	0	0
2018_H	A043	0	-0.262	0	0	0	0	-0.093	-0.151	0	0	0	-0.151
2018_H	A044	0	-0.193	-0.036	0	0	0	0	-0.062	0	0	0	0
2018_H	A045	0	-0.068	0	0	0	-0.023	0	0	-0.023	0	0	0
2018_H	A046	0	-0.239	-0.028	-0.028	0	-0.049	0	0	0	0	0	0
2018_H	B048	0	-0.157	-0.083	0	0	-0.028	0	0	0	0	0	-0.083
2018_H	B051	0	-0.105	-0.105	0	0	-0.037	0	0	0	0	0	0
2018_H	B054	0	-0.334	-0.138	0	0	0	0	-0.138	0	0	0	0
2018_H	B057	0	-0.285	-0.105	0	0	-0.105	0	-0.336	0	0	0	0
2018_H	B059	0	-0.201	0	0	0	0	0	0	-0.038	0	0	0
2018_H	B060	0	0	-0.173	0	0	0	0	0	0	0	0	0
2018_H	B064	0	-0.102	-0.102	0	0	0	0	0	-0.035	0	0	0
2018_H	B067	0	-0.074	0	0	0	0	0	0	0	0	0	0
2018_H	B071	0	-0.228	0	0	0	0	-0.077	0	0	0	0	0
2018_H	B073	0	-0.310	-0.073	0	0	0	0	-0.141	0	0	0	0
2018_H	B075	0	-0.150	0	0	0	0	0	0	0	0	0	0
2018_H	B080	0	-0.107	-0.107	0	0	0	0	0	0	0	0	0
2018_H	B081	0	0	-0.100	0	0	-0.059	0	0	0	0	0	0
2018_H	B085	0	-0.212	-0.136	0	0	-0.070	0	-0.070	-0.024	0	0	-0.070
2018_H	B090	0	-0.163	0	0	0	-0.086	0	-0.029	-0.029	0	0	0
2018_H	B092	0	0	0	0	0	0	0	0	-0.025	0	0	-0.044
2018_H	C097	0	0	-0.085	0	0	0	-0.029	0	-0.085	0	0	-0.050
2018_H	C103	0	-0.333	-0.071	0	0	-0.024	0	0	0	0	0	0
2018_H	C109	0	0	-0.157	0	0	-0.049	0	0	0	0	0	0
2018_H	C110	0	-0.148	0	0	0	0	0	-0.077	-0.046	0	0	0
2018_H	C114	0	-0.094	0	0	0	0	0	-0.032	-0.032	0	0	0
2018_H	C115	0	-0.209	-0.301	0	0	0	0	0	0	0	0	0
2018_H	C121	0	-0.145	0	0	0	0	0	-0.145	-0.253	0	0	-0.145
2018_H	C131	0	-0.071	0	0	0	0	0	0	-0.071	0	0	0
2018_H	C133	0	-0.184	-0.273	0	0	0	0	-0.059	-0.099	0	0	0
2018_H	C135	0	-0.367	-0.166	0	0	-0.052	0	0	0	-0.052	0	0
2018_H	C136	0	-0.272	0	0	0	0	0	0	0	0	0	0
2018_H	C137	0	-0.225	0	0	0	0	0	-0.125	-0.076	0	0	-0.076

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae	Compositae
	Code	Erican	Galpar	Lacser	Leoaut	Leuvul	Scoaut	Solcan	Solvir	Sonarv	Sonasp	Sonole	Taroff
	species	**Erigeron_canad**	**Galinsoga_parvi**	**Lactuca_serr**	**Leontodon_autum**	**Leucanthemum_vul**	**Scorzoneroides_autur**	**Solidago_canad**	**Solidago_virga**	**Sonchus_arve**	**Sonchus_asp**	**Sonchus_olera**	**Taraxacum_officinale agg.**
	season	2018_H	2018_E,F,G	2018_E,G,H	2018_H	2018_E,F,G,H	2018_F,G	2018_E,G,H	2018_F	2018_G	2018_E	2018_E	2018_E,F,G,H
Season	**plot**												
2018_H	A002	0	0	0	0	-0.056	0	0	0	0	0	0	-0.264
2018_H	A003	0	0	0	0	-0.368	0	0	0	0	0	0	-0.084
2018_H	A005	0	0	0	0	-0.255	0	0	0	0	0	0	-0.090
2018_H	A009	0	0	0	0	-0.150	0	-0.055	0	0	0	0	-0.150
2018_H	A010	0	0	0	0	0	0	0	0	0	0	0	0
2018_H	A011	0	0	0	0	-0.368	0	-0.028	0	0	0	0	-0.157
2018_H	A013	0	0	0	0	-0.181	0	-0.058	0	0	0	0	-0.058
2018_H	A016	0	0	0	0	-0.236	0	-0.027	0	0	0	0	-0.291
2018_H	A018	0	0	0	0	-0.339	0	0	0	0	0	0	0
2018_H	A020	0	0	0	0	-0.117	0	0	0	0	0	0	0
2018_H	A026	-0.051	0	0	0	0	0	0	0	0	0	0	-0.051
2018_H	A027	0	0	0	0	0	0	-0.039	0	0	0	0	0
2018_H	A030	0	0	0	0	-0.353	0	0	0	0	0	0	0
2018_H	A035	0	0	0	0	-0.128	0	0	0	0	0	0	0
2018_H	A040	0	0	0	0	-0.274	0	0	0	0	0	0	-0.161
2018_H	A042	0	0	0	0	0	0	-0.074	0	0	0	0	-0.043
2018_H	A043	0	0	0	0	-0.055	0	0	0	0	0	0	0
2018_H	A044	0	0	0	-0.036	-0.193	0	0	0	0	0	0	0
2018_H	A045	0	0	0	0	-0.133	0	0	0	0	0	0	0
2018_H	A046	0	0	0	0	-0.352	0	0	0	0	0	0	-0.082
2018_H	B048	0	0	0	0	-0.363	0	-0.049	0	0	0	0	-0.083
2018_H	B051	0	0	0	0	-0.361	0	0	0	0	0	0	-0.105
2018_H	B054	0	0	0	0	0	0	0	0	0	0	0	-0.050
2018_H	B057	0	0	0	0	-0.195	0	0	0	0	0	0	-0.063
2018_H	B059	0	0	0	0	0	0	-0.066	0	0	0	0	-0.110
2018_H	B060	0	0	0	0	-0.173	0	0	0	0	0	0	-0.055
2018_H	B064	0	0	0	0	-0.330	0	0	0	0	0	0	0
2018_H	B067	0	0	0	0	-0.339	0	0	0	0	0	0	0
2018_H	B071	0	0	0	0	0	0	0	0	0	0	0	-0.148
2018_H	B073	0	0	0	0	-0.337	0	0	0	0	0	0	0
2018_H	B075	0	0	0	0	-0.322	0	0	0	0	0	0	0
2018_H	B080	0	0	0	0	-0.287	0	0	0	0	0	0	-0.196
2018_H	B081	0	0	0	0	0	0	-0.100	0	0	0	0	0
2018_H	B085	0	0	0	0	-0.367	0	0	0	0	0	0	-0.070
2018_H	B090	0	0	0	0	-0.343	0	0	0	0	0	0	0
2018_H	B092	0	0	0	0	-0.044	0	0	0	0	0	0	0
2018_H	C097	0	0	0	0	-0.085	0	0	0	0	0	0	-0.085
2018_H	C103	0	0	0	0	-0.306	0	0	0	0	0	0	0
2018_H	C109	0	0	0	0	-0.368	0	0	0	0	0	0	-0.082
2018_H	C110	0	0	0	0	-0.148	0	-0.026	0	0	0	0	-0.148
2018_H	C114	0	0	0	0	0	0	0	0	0	0	0	-0.032
2018_H	C115	0	0	0	0	0	0	-0.040	0	0	0	0	0
2018_H	C121	0	0	0	0	0	0	-0.145	0	0	0	0	-0.089
2018_H	C131	0	0	0	0	-0.071	0	0	0	0	0	0	-0.071
2018_H	C133	0	0	-0.034	0	0	0	0	0	0	0	0	-0.059
2018_H	C135	0	0	0	-0.030	-0.367	0	0	0	0	0	0	0
2018_H	C136	0	0	0	0	-0.368	0	0	0	0	0	0	-0.043
2018_H	C137	0	0	0	0	-0.125	0	-0.045	0	0	0	0	0

Supplement Table 1: Plant species indices – continued (FG herb species E-O Season 2018 E)

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Equisetaceae	Euphorbiaceae	Euphorbiaceae	Geraniaceae	Geraniaceae	Hypericaceae	Lamiaceae	Lamiaceae	Malvaceae	Onagraceae
	Code	Equarv	Eupspe	Merann	Gerpra	Gerspe	Hypper	Glehed	Pruvul	Malmos	Epispe
	species	**Equisetum_arvense**	**Euphorbia_spec.**	**Mercurialis_annua**	**Geranium_pratense**	**Geranium_spec.**	**Hypericum_perforatum**	**Glechoma_hederacea**	**Prunella_vulgaris**	**Malva_moschata**	**Epilobium_spec.**
	season	2018_E,F,G,H	2018_H, G	2018_G,H	2018_E,F,G,H	2018_E	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_F,G,H	2018_E
Season	**plot**										
2018_E	A002	-0.113	0	0	-0.113	0	0	-0.040	-0.040	0	0
2018_E	A003	-0.071	0	0	-0.071	0	0	-0.071	0	0	0
2018_E	A005	-0.277	0	0	0	0	0	0	0	0	0
2018_E	A009	-0.351	0	0	0	-0.042	0	0	0	0	0
2018_E	A010	-0.053	0	0	0	0	0	0	-0.031	0	0
2018_E	A011	0	0	0	0	0	0	0	-0.057	0	0
2018_E	A013	0	0	0	0	0	0	-0.030	0	0	-0.030
2018_E	A016	-0.110	0	0	0	0	0	-0.110	-0.038	0	0
2018_E	A018	-0.051	0	0	-0.367	0	0	0	-0.051	0	0
2018_E	A020	-0.191	0	0	0	0	0	0	-0.036	0	0
2018_E	A026	0	0	0	0	0	0	0	-0.065	0	0
2018_E	A027	0	0	0	0	0	0	-0.088	0	0	0
2018_E	A030	0	0	0	0	0	0	0	0	0	0
2018_E	A035	-0.108	0	0	-0.355	0	0	-0.065	0	0	0
2018_E	A040	0	0	0	0	0	0	0	-0.039	0	0
2018_E	A042	-0.071	0	0	0	0	0	0	-0.118	0	0
2018_E	A043	0	0	0	0	0	0	-0.035	0	0	0
2018_E	A044	0	0	0	0	0	0	0	0	0	0
2018_E	A045	0	0	0	0	0	0	0	0	0	0
2018_E	A046	0	0	0	0	0	0	0	0	0	0
2018_E	B048	0	0	0	0	0	0	-0.050	-0.085	0	0
2018_E	B051	0	0	0	0	0	0	0	0	0	0
2018_E	B054	0	0	0	-0.054	0	0	-0.091	-0.031	0	0
2018_E	B057	0	0	0	0	0	0	0	-0.066	0	0
2018_E	B059	0	0	0	0	0	0	0	-0.125	0	0
2018_E	B060	0	0	0	0	0	0	0	-0.169	0	0
2018_E	B064	0	0	0	0	0	0	-0.033	-0.095	0	0
2018_E	B067	0	0	0	0	0	0	0	0	0	0
2018_E	B071	0	0	0	-0.088	0	0	0	0	0	0
2018_E	B073	0	0	0	0	0	0	0	0	0	0
2018_E	B075	0	0	0	0	0	0	0	0	0	0
2018_E	B080	0	0	0	-0.026	0	0	0	0	0	0
2018_E	B081	0	0	0	-0.364	0	0	0	-0.094	0	0
2018_E	B085	0	0	0	0	0	0	-0.176	-0.032	0	0
2018_E	B090	0	0	0	-0.030	0	0	-0.165	0	0	0
2018_E	B092	0	0	0	0	0	0	0	0	0	-0.040
2018_E	C097	0	0	0	0	0	0	0	0	0	0
2018_E	C103	0	0	0	-0.030	0	0	-0.166	-0.052	0	0
2018_E	C109	0	0	0	-0.358	0	0	0	0	0	0
2018_E	C110	0	0	0	-0.090	0	0	0	-0.090	0	0
2018_E	C114	0	0	0	0	0	0	-0.025	0	0	0
2018_E	C115	0	0	0	0	0	0	0	-0.194	0	0
2018_E	C121	0	0	0	-0.296	0	0	-0.108	0	0	-0.065
2018_E	C131	0	0	0	0	0	0	-0.029	-0.086	0	0
2018_E	C133	0	0	0	0	0	0	0	-0.068	0	0
2018_E	C135	0	0	0	0	0	0	-0.036	-0.062	0	-0.036
2018_E	C136	0	0	0	0	0	0	0	0	0	0
2018_E	C137	0	0	0	-0.035	0	0	0	-0.035	0	0

Supplement Table 1: Plant species indices – continued (FG herb species E-O Season 2018 F)

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Equisetaceae	Euphorbiaceae	Euphorbiaceae	Geraniaceae	Geraniaceae	Hypericaceae	Lamiaceae	Lamiaceae	Malvaceae	Onagraceae
	Code	Equarv	Eupspe	Merann	Gerpra	Gerspe	Hypper	Glehed	Pruvul	Malmos	Epispe
	species	Equisetum_arvense	Euphorbia_spec.	Mercurialis_annua	Geranium_pratense	Geranium_spec.	Hypericum_perforatum	Glechoma_hederacea	Prunella_vulgaris	Malva_moschata	Epilobium_spec.
	season	2018_E,F,G,H	2018_H, G	2018_G,H	2018_E,F,G,H	2018_E	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_F,G,H	2018_E
Season	plot										
2018_F	A002	-0.138	0	0	0	0	0	0	0	0	0
2018_F	A003	-0.255	0	0	-0.053	0	0	0	0	-0.053	0
2018_F	A005	-0.334	0	0	0	0	0	0	0	0	0
2018_F	A009	-0.365	0	0	0	0	0	0	0	0	0
2018_F	A010	-0.119	0	0	0	0	0	0	0	0	0
2018_F	A011	0	0	0	0	0	0	0	0	0	0
2018_F	A013	0	0	0	0	0	0	0	0	0	0
2018_F	A016	-0.086	0	0	0	0	0	-0.086	0	0	0
2018_F	A018	-0.211	0	0	-0.350	0	0	0	0	0	0
2018_F	A020	-0.215	0	0	-0.085	0	0	0	0	0	0
2018_F	A026	-0.164	0	0	0	0	0	0	0	0	0
2018_F	A027	0	0	0	0	0	0	0	0	0	0
2018_F	A030	0	0	0	0	0	0	0	0	0	0
2018_F	A035	-0.232	0	0	-0.366	0	0	-0.046	0	0	0
2018_F	A040	0	0	0	0	0	0	0	0	0	0
2018_F	A042	-0.177	0	0	0	0	0	-0.067	0	0	0
2018_F	A043	0	0	0	0	0	0	0	0	0	0
2018_F	A044	0	0	0	0	0	0	0	0	0	0
2018_F	A045	0	0	0	0	0	0	0	0	0	0
2018_F	A046	0	0	0	0	0	0	-0.085	0	0	0
2018_F	B048	0	0	0	0	0	0	0	0	0	0
2018_F	B051	0	0	0	0	0	0	0	0	0	0
2018_F	B054	0	0	0	0	0	0	0	0	0	0
2018_F	B057	0	0	0	0	0	0	0	0	0	0
2018_F	B059	0	0	0	0	0	0	0	0	0	0
2018_F	B060	0	0	0	0	0	0	0	0	0	0
2018_F	B064	0	0	0	0	0	0	0	0	0	0
2018_F	B067	0	0	0	0	0	0	0	0	0	0
2018_F	B071	0	0	0	0	0	0	0	0	0	0
2018_F	B073	0	0	0	0	0	0	0	0	0	0
2018_F	B075	0	0	0	0	0	0	0	0	0	0
2018_F	B080	0	0	0	0	0	0	0	0	0	0
2018_F	B081	0	0	0	-0.368	0	0	0	0	0	0
2018_F	B085	0	0	0	0	0	0	0	0	0	0
2018_F	B090	0	0	0	0	0	0	0	0	0	0
2018_F	B092	0	0	0	0	0	0	0	0	0	0
2018_F	C097	0	0	0	0	0	0	-0.061	0	0	0
2018_F	C103	0	0	0	0	0	0	0	0	0	0
2018_F	C109	0	0	0	-0.240	0	0	0	-0.083	0	0
2018_F	C110	0	0	0	-0.111	0	0	-0.111	0	0	0
2018_F	C114	0	0	0	0	0	0	0	0	0	0
2018_F	C115	0	0	0	0	0	0	-0.155	0	0	0
2018_F	C121	0	0	0	-0.306	0	0	0	0	0	0
2018_F	C131	0	0	0	0	0	0	-0.051	0	0	0
2018_F	C133	0	0	0	0	0	0	0	0	0	0
2018_F	C135	0	0	0	0	0	0	0	0	0	0
2018_F	C136	0	0	0	0	0	0	0	0	0	0
2018_F	C137	0	0	0	0	0	0	0	0	0	0

Supplement Table 1: Plant species indices – continued (FG herb species E-O Season 2018 G)

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Equisetaceae	Euphorbiaceae	Euphorbiaceae	Geraniaceae	Geraniaceae	Hypericaceae	Lamiaceae	Lamiaceae	Malvaceae	Onagraceae
	Code	Equarv	Eupspe	Merann	Gerpra	Gerspe	Hypper	Glehed	Pruvul	Malmos	Epispe
	species	Equisetum_arver	Euphorbia_spec.	Mercurialis_ann	Geranium_prate	Geranium_spec.	Hypericum_perf	Glechoma_hede	Prunella_vulgari	Malva_moschata	Epilobium_spec.
	season	2018_E,F,G,H	2018_H, G	2018_G,H	2018_E,F,G,H	2018_E	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_F,G,H	2018_E
Season	plot										
2018_G	A002	-0.085	-0.050	0	0	0	0	0	0	0	0
2018_G	A003	-0.185	0	0	0	0	0	0	0	-0.059	0
2018_G	A005	-0.230	-0.046	0	0	0	0	0	0	0	0
2018_G	A009	-0.317	0	0	0	0	0	0	0	0	0
2018_G	A010	-0.127	0	0	0	0	0	0	0	0	0
2018_G	A011	0	0	0	0	0	0	0	0	0	0
2018_G	A013	0	0	0	0	0	0	-0.048	0	0	0
2018_G	A016	-0.094	0	0	0	0	0	-0.056	0	0	0
2018_G	A018	-0.134	0	0	-0.329	0	0	0	0	0	0
2018_G	A020	-0.331	0	0	0	0	0	0	0	0	0
2018_G	A026	-0.128	0	0	0	0	0	0	0	0	0
2018_G	A027	0	-0.067	0	0	0	0	0	0	0	0
2018_G	A030	0	0	0	0	0	0	0	0	0	0
2018_G	A035	-0.240	0	0	-0.363	0	0	-0.083	0	0	0
2018_G	A040	0	0	-0.091	0	0	0	0	0	0	0
2018_G	A042	-0.094	0	0	0	0	0	-0.056	0	0	0
2018_G	A043	0	-0.059	0	0	0	0	-0.059	0	0	0
2018_G	A044	0	0	0	0	0	0	0	0	0	0
2018_G	A045	0	0	0	0	0	0	0	0	0	0
2018_G	A046	0	0	0	0	0	0	0	0	0	0
2018_G	B048	0	0	-0.060	0	0	0	-0.060	0	-0.060	0
2018_G	B051	0	0	0	0	0	0	0	0	0	0
2018_G	B054	0	0	0	0	0	0	0	0	0	0
2018_G	B057	0	0	0	0	0	0	0	0	0	0
2018_G	B059	0	0	-0.316	0	0	0	0	0	0	0
2018_G	B060	0	0	-0.087	0	0	0	0	0	0	0
2018_G	B064	0	0	0	0	0	0	-0.052	0	0	0
2018_G	B067	0	0	0	0	0	0	0	0	0	0
2018_G	B071	0	0	-0.099	0	0	0	-0.059	0	0	0
2018_G	B073	0	0	0	0	0	0	0	0	0	0
2018_G	B075	0	0	0	0	0	0	0	0	0	0
2018_G	B080	0	0	0	0	0	0	-0.054	0	0	0
2018_G	B081	0	0	0	-0.349	0	0	-0.056	0	0	0
2018_G	B085	0	0	0	0	0	0	-0.048	-0.048	0	0
2018_G	B090	0	0	0	-0.050	0	0	0	0	0	0
2018_G	B092	0	0	0	0	0	0	-0.048	0	0	0
2018_G	C097	0	0	-0.057	0	0	-0.057	-0.057	0	0	0
2018_G	C103	0	0	0	-0.061	0	0	0	0	0	0
2018_G	C109	0	0	0	-0.326	0	0	0	-0.153	0	0
2018_G	C110	0	0	-0.055	-0.093	0	0	0	0	0	0
2018_G	C114	0	0	0	0	0	0	0	0	0	0
2018_G	C115	0	0	0	0	0	0	-0.045	0	0	0
2018_G	C121	0	0	0	-0.340	0	0	0	0	0	0
2018_G	C131	0	0	0	0	0	0	-0.043	0	0	0
2018_G	C133	0	0	-0.155	0	0	0	0	0	0	0
2018_G	C135	0	0	0	0	0	0	0	0	0	0
2018_G	C136	0	0	0	0	0	0	0	0	0	0
2018_G	C137	0	0	0	0	0	0	0	0	0	0

Supplement Table 1: Plant species indices – continued (FG herb species E-O Season 2018 H)

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Equisetaceae	Euphorbiaceae	Euphorbiaceae	Geraniaceae	Geraniaceae	Hypericaceae	Lamiaceae	Lamiaceae	Malvaceae	Onagraceae
	Code	Equarv	Eupspe	Merann	Gerpra	Gerspe	Hypper	Glehed	Pruvul	Malmos	Epispe
	species	Equisetum_arvense	Euphorbia_spec.	Mercurialis_annua	Geranium_pratense	Geranium_spec.	Hypericum_perforatum	Glechoma_hederacea	Prunella_vulgaris	Malva_moschata	Epilobium_spec.
	season	2018_E,F,G,H	2018_H, G	2018_G,H	2018_E,F,G,H	2018_E	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_F,G,H	2018_E
Season	plot										
2018_H	A002	0	0	-0.094	0	0	0	0	-0.094	0	0
2018_H	A003	-0.084	0	-0.028	0	0	0	0	0	0	0
2018_H	A005	-0.169	0	0	0	0	0	-0.090	0	0	0
2018_H	A009	-0.347	0	-0.055	-0.055	0	0	0	0	0	0
2018_H	A010	-0.104	0	0	0	0	0	0	0	0	0
2018_H	A011	0	0	0	0	0	0	0	0	0	0
2018_H	A013	0	0	-0.058	0	0	0	-0.033	0	0	0
2018_H	A016	0	0	-0.081	0	0	0	-0.048	0	0	0
2018_H	A018	-0.087	0	0	-0.165	0	0	0	0	0	0
2018_H	A020	-0.304	0	0	0	0	0	0	0	0	0
2018_H	A026	-0.087	0	-0.030	0	0	0	0	0	0	0
2018_H	A027	0	-0.067	-0.039	0	0	0	0	0	0	0
2018_H	A030	0	-0.031	0	0	0	0	0	0	0	0
2018_H	A035	-0.229	-0.046	-0.046	-0.321	0	0	-0.078	0	0	0
2018_H	A040	-0.100	-0.161	-0.059	0	0	0	0	0	0	0
2018_H	A042	-0.074	0	0	0	0	0	0	0	0	0
2018_H	A043	0	-0.093	-0.055	0	0	-0.055	0	0	0	0
2018_H	A044	0	0	0	0	0	0	0	0	0	0
2018_H	A045	0	0	0	0	0	0	0	0	0	0
2018_H	A046	0	-0.028	-0.028	0	0	0	0	-0.049	0	0
2018_H	B048	0	0	-0.028	0	0	0	0	-0.157	-0.083	0
2018_H	B051	0	0	0	0	0	0	-0.037	0	0	0
2018_H	B054	0	0	-0.050	0	0	0	-0.050	0	0	0
2018_H	B057	0	0	0	0	0	0	0	0	0	0
2018_H	B059	0	0	-0.110	0	0	0	0	0	0	0
2018_H	B060	0	0	-0.173	0	0	0	0	0	0	0
2018_H	B064	0	0	0	0	0	0	-0.035	0	0	0
2018_H	B067	0	0	0	0	0	0	0	0	0	0
2018_H	B071	0	-0.026	-0.026	0	0	0	0	0	0	0
2018_H	B073	0	0	0	0	0	0	0	0	0	0
2018_H	B075	0	0	0	0	0	0	-0.026	0	0	0
2018_H	B080	0	0	0	0	0	0	-0.064	0	0	0
2018_H	B081	0	0	-0.059	-0.274	0	0	0	-0.059	0	0
2018_H	B085	0	0	-0.024	0	0	0	-0.212	-0.070	0	0
2018_H	B090	0	0	0	-0.029	0	0	-0.163	0	-0.029	0
2018_H	B092	0	0	0	-0.025	0	0	0	0	0	0
2018_H	C097	0	0	-0.050	0	0	0	-0.085	0	0	0
2018_H	C103	0	0	0	-0.042	0	0	0	0	0	0
2018_H	C109	0	0	0	-0.294	0	0	0	-0.082	0	0
2018_H	C110	0	0	-0.026	-0.148	0	0	0	0	0	0
2018_H	C114	0	0	0	0	0	0	0	0	0	0
2018_H	C115	0	0	-0.040	0	0	0	-0.115	0	0	0
2018_H	C121	0	0	0	-0.368	0	0	-0.053	0	0	0
2018_H	C131	0	0	0	0	0	0	-0.071	0	0	0
2018_H	C133	0	0	-0.184	0	0	0	0	0	0	0
2018_H	C135	0	0	0	0	0	0	0	0	0	0
2018_H	C136	0	0	0	0	0	0	0	0	0	0
2018_H	C137	0	0	0	0	0	0	0	0	0	0

Supplement Table 1: Plant species indices – continued (FG herb species P-Z Season 2018 E & F)

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Papaveraceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Polygonaceae	Polygonaceae	Ranunculaceae	Ranunculaceae	Rosaceae	Rosaceae	Rubiaceae	Sapindaceae
	Code	Paparg	Linvul	Plalan	Plamed	Veranv	Veracf	Vercha	Polavi	Rumace	Ranacr	Ranrep	Potrep	Sagoff	Galmol	Acespa
	species	Papaver_argemone	Linaria_vulgaris	Plantago_lanceolata	Plantago_media	Veronica_arvensis	Veronica_arvensis cf.	Veronica_chamaedrys	Polygonum_aviculare	Rumex_acetosa	Ranunculus_acris	Ranunculus_repens	Potentilla_reptans	Saguisorba_officinalis	Galium_mollugo	Acer_spec.
	season	2018_E	2018_F	2018_E,F,G,H	2018_E,F,G,H	2018_F,H	2018_E	2018_E,F,G,H	2018_E	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G,H
Season	plot															
2018_E	A002	0	0	-0.068	-0.040	0	0	-0.113	0	-0.040	-0.040	0	0	0	0	-0.113
2018_E	A003	0	0	-0.071	0	0	0	-0.071	0	-0.071	-0.306	0	0	0	0	-0.042
2018_E	A005	-0.035	0	-0.329	-0.035	0	0	-0.060	0	-0.101	-0.277	0	0	0	0	0
2018_E	A009	-0.042	0	-0.361	0	0	0	-0.118	0	-0.071	0	0	0	-0.042	0	-0.042
2018_E	A010	0	0	-0.090	0	0	0	-0.053	0	-0.053	0	0	0	0	0	-0.031
2018_E	A011	0	0	-0.351	0	0	0	-0.057	0	-0.033	-0.033	0	0	0	0	0
2018_E	A013	-0.030	0	-0.360	0	0	0	-0.167	0	-0.167	0	0	0	0	0	-0.030
2018_E	A016	0	0	-0.110	-0.066	0	-0.038	-0.066	0	-0.110	-0.342	0	0	0	0	-0.066
2018_E	A018	0	0	-0.246	0	0	0	-0.051	0	-0.029	-0.086	-0.029	0	-0.029	0	-0.051
2018_E	A020	-0.062	0	-0.062	0	0	0	-0.103	0	-0.103	0	0	0	-0.036	0	-0.036
2018_E	A026	0	0	-0.108	0	0	0	-0.065	0	-0.065	0	0	0	-0.038	0	0
2018_E	A027	-0.052	0	-0.144	0	0	0	-0.088	0	-0.144	0	0	0	0	0	0
2018_E	A030	0	0	-0.107	0	0	0	-0.196	0	0	0	0	0	0	0	-0.037
2018_E	A035	0	0	-0.199	0	0	0	-0.065	0	-0.038	-0.108	0	0	0	0	-0.038
2018_E	A040	0	0	-0.202	-0.039	0	0	-0.066	0	-0.110	0	0	0	-0.066	0	-0.039
2018_E	A042	0	0	0	0	0	0	-0.042	0	0	0	0	0	0	0	-0.118
2018_E	A043	-0.061	0	-0.368	0	0	0	-0.102	0	0	0	0	0	0	0	-0.035
2018_E	A044	-0.040	0	-0.344	0	0	0	-0.115	0	-0.040	0	0	0	0	0	-0.040
2018_E	A045	0	0	-0.102	0	0	0	0	0	0	0	0	0	0	0	-0.035
2018_E	A046	0	0	0	0	0	0	-0.074	0	-0.025	0	0	0	0	0	0
2018_E	B048	0	0	-0.085	0	0	0	-0.161	0	-0.161	0	0	0	-0.029	0	0
2018_E	B051	0	0	-0.274	0	0	0	-0.043	0	-0.043	0	0	0	0	0	0
2018_E	B054	0	0	-0.363	0	0	0	-0.091	0	-0.031	-0.054	0	0	0	0	0
2018_E	B057	-0.066	0	-0.320	-0.022	0	0	-0.253	0	-0.038	0	0	0	0	0	0
2018_E	B059	0	0	-0.225	0	0	0	-0.225	0	-0.225	0	0	0	0	0	0
2018_E	B060	0	0	-0.255	0	0	0	-0.169	0	-0.169	-0.031	0	0	0	-0.090	0
2018_E	B064	-0.033	0	0	0	0	0	-0.178	0	-0.033	0	0	0	0	0	-0.057
2018_E	B067	0	0	0	0	0	0	-0.043	0	0	0	0	0	0	0	0
2018_E	B071	-0.052	0	-0.167	0	0	0	-0.167	0	0	-0.252	0	0	0	0	-0.052
2018_E	B073	0	0	-0.159	0	0	0	0	0	0	0	0	0	0	0	0
2018_E	B075	0	0	-0.256	0	0	0	-0.090	0	0	0	0	0	0	0	0
2018_E	B080	0	0	-0.361	0	0	0	-0.230	0	0	-0.230	0	0	0	0	-0.026
2018_E	B081	0	0	0	0	0	0	-0.262	0	0	-0.094	0	0	0	0	0
2018_E	B085	0	0	-0.056	0	0	0	-0.094	0	0	0	0	0	0	0	0
2018_E	B090	0	0	-0.087	0	0	0	-0.087	0	-0.052	0	0	0	0	0	0
2018_E	B092	0	0	-0.068	0	0	0	-0.113	0	-0.068	0	0	0	0	0	-0.040
2018_E	C097	0	0	0	0	0	0	-0.171	0	0	-0.031	0	0	0	0	-0.031
2018_E	C103	0	0	0	0	0	0	-0.088	0	0	0	0	0	0	0	0
2018_E	C109	0	0	0	0	0	0	-0.119	0	-0.061	0	0	0	0	0	0
2018_E	C110	0	0	0	0	0	0	-0.169	0	-0.090	0	0	0	0	0	-0.031
2018_E	C114	-0.025	0	0	0	0	0	-0.145	0	0	-0.044	0	0	0	0	0
2018_E	C115	-0.021	0	-0.335	0	0	0	0	0	-0.123	-0.314	0	0	0	0	0
2018_E	C121	0	0	0	0	0	0	-0.103	-0.037672964	0	0	0	0	-0.038	0	0
2018_E	C131	0	0	-0.086	0	0	0	-0.162	0	0	-0.246	0	0	-0.029	0	-0.029
2018_E	C133	0	0	-0.068	0	0	0	-0.113	-0.068239076	-0.299	-0.068	0	0	0	-0.181	-0.068
2018_E	C135	0	0	-0.036	0	0	0	-0.062	0	0	0	0	0	0	0	0
2018_E	C136	0	0	-0.253	0	0	0	-0.089	0	0	0	0	0	0	0	0
2018_E	C137	0	0	-0.357	0	0	0	-0.060	0	0	0	0	0	0	-0.035	0

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Papaveraceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Polygonaceae	Polygonaceae	Ranunculaceae	Ranunculaceae	Rosaceae	Rosaceae	Rubiaceae	Sapindaceae
	Code	Paparg	Linvul	Plalan	Plamed	Veranv	Veracf	Vercha	Polavi	Rumace	Ranacr	Ranrep	Potrep	Sagoff	Galmol	Acespe
	species	Papaver_argemone	Linaria_vulgaris	Plantago_lanceolata	Plantago_media	Veronica_arvensis	Veronica_arvensis cf.	Veronica_chamaedrys	Polygonum_aviculare	Rumex_acetosa	Ranunculus_acris	Ranunculus_repens	Potentilla_reptans	Saguisorba_officinalis	Galium_mollugo	Acer_spec.
	season	2018_E	2018_F	2018_E,F,G,H	2018_E,F,G,H	2018_F,H	2018_E	2018_E,F,G,H	2018_E	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G,H
Season	plot															
2018_F	A002	0	0	-0.138	-0.138	0	0	0	0	0	-0.215	0	0	0	0	0
2018_F	A003	0	0	-0.146	0	0	0	0	0	0	-0.255	0	0	0	0	0
2018_F	A005	0	0	-0.215	0	0	0	0	0	-0.085	-0.334	0	0	-0.085	0	0
2018_F	A009	0	0	-0.365	0	0	0	0	0	0	0	0	0	-0.111	0	0
2018_F	A010	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2018_F	A011	0	0	-0.367	0	0	0	0	0	-0.069	0	0	0	-0.069	0	0
2018_F	A013	0	0	-0.364	0	0	0	0	0	-0.060	0	0	0	0	-0.060	-0.060
2018_F	A016	0	-0.08600431	-0.218	-0.141	0	0	0	0	-0.086	-0.358	0	0	0	0	0
2018_F	A018	0	0	-0.303	0	0	0	0	0	-0.041	-0.211	0	0	0	0	0
2018_F	A020	0	0	-0.215	-0.085	0	0	0	0	-0.085	0	0	0	-0.085	0	0
2018_F	A026	0	0	-0.061	0	0	0	0	0	-0.061	0	0	0	-0.061	0	0
2018_F	A027	0	0	-0.221	0	0	0	0	0	-0.087	0	0	0	0	0	0
2018_F	A030	0	0	-0.260	0	0	0	0	0	-0.092	0	0	0	0	0	0
2018_F	A035	0	0	-0.323	0	0	0	0	0	-0.046	0	0	0	0	0	0
2018_F	A040	0	0	-0.173	0	0	0	0	0	-0.173	0	0	0	-0.173	0	0
2018_F	A042	0	0	0	0	0	0	0	0	-0.067	0	0	0	0	0	0
2018_F	A043	0	0	-0.179	0	0	0	0	0	0	-0.078	0	0	0	0	0
2018_F	A044	0	0	-0.251	0	0	0	0	0	-0.087	0	0	0	0	0	0
2018_F	A045	0	0	-0.150	0	0	0	0	0	0	0	0	0	-0.092	0	0
2018_F	A046	0	0	-0.138	0	0	0	0	0	0	0	0	0	0	0	0
2018_F	B048	0	0	-0.317	0	0	0	0	0	-0.076	0	0	0	0	-0.076	0
2018_F	B051	0	0	-0.328	0	0	0	0	0	-0.048	0	0	0	0	0	0
2018_F	B054	0	0	-0.279	0	0	0	0	0	-0.083	0	0	0	0	0	0
2018_F	B057	0	0	-0.272	0	0	0	0	0	0	-0.159	0	0	0	0	0
2018_F	B059	0	0	-0.218	0	0	0	0	0	0	-0.218	0	0	0	0	0
2018_F	B060	0	0	-0.132	-0.132	0	0	0	0	-0.207	-0.132	0	0	0	0	-0.132
2018_F	B064	0	0	-0.059	0	0	0	0	0	-0.059	-0.059	0	0	0	0	0
2018_F	B067	0	0	-0.059	-0.059	0	0	0	0	0	0	0	0	0	0	0
2018_F	B071	0	0	-0.227	0	0	0	0	0	0	-0.350	0	0	0	0	0
2018_F	B073	0	0	-0.170	0	0	0	0	0	-0.063	0	0	0	0	0	0
2018_F	B075	0	0	-0.252	0	0	0	0	0	0	0	0	0	0	0	0
2018_F	B080	0	0	-0.256	0	0	0	0	0	0	-0.116	0	0	-0.070	0	0
2018_F	B081	0	0	-0.189	0	0	0	0	0	0	-0.119	0	0	0	0	0
2018_F	B085	0	0	-0.212	0	0	0	0	0	-0.136	0	0	0	0	0	0
2018_F	B090	0	0	-0.105	0	0	0	0	0	0	0	0	0	0	0	0
2018_F	B092	0	0	-0.094	0	0	0	0	0	-0.094	0	0	0	0	0	0
2018_F	C097	0	0	-0.280	0	-0.103	0	-0.103	0	-0.165	-0.165	0	0	0	0	0
2018_F	C103	0	0	0	0	0	0	0	0	-0.052	0	0	0	0	0	0
2018_F	C109	0	0	0	0	0	0	0	0	-0.049	-0.240	0	0	0	-0.049	0
2018_F	C110	0	0	-0.177	0	0	0	-0.111	0	-0.177	-0.264	0	0	0	0	0
2018_F	C114	0	0	-0.043	0	0	0	0	0	0	-0.043	0	0	0	0	0
2018_F	C115	0	0	-0.351	0	0	0	0	0	-0.096	-0.351	0	0	0	0	0
2018_F	C121	0	0	-0.129	0	0	0	0	0	0	-0.078	0	0	-0.078	0	0
2018_F	C131	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2018_F	C133	0	0	0	-0.218	0	0	0	0	-0.218	0	-0.218	0	0	0	0
2018_F	C135	0	0	-0.111	0	0	0	-0.177	0	0	0	0	0	0	0	0
2018_F	C136	0	0	-0.366	0	0	0	0	0	0	0	0	0	0	0	0
2018_F	C137	0	0	-0.352	0	0	0	0	0	0	0	0	0	0	0	0

Supplement Table 1: Plant species indices – continued (FG herb species P-Z Season 2018 G & H)

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Papaveraceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Polygonaceae	Polygonaceae	Ranunculaceae	Ranunculaceae	Rosaceae	Rosaceae	Rubiaceae	Sapindaceae
	Code	Paparg	Linvul	Plalan	Plamed	Veranv	Veracf	Vercha	Polavi	Rumace	Ranacr	Ranrep	Potrep	Sagoff	Galmol	Acespe
	species	Papaver_argemone	Linaria_vulgaris	Plantago_lanceolata	Plantago_media	Veronica_arvensis	Veronica_arvensis cf.	Veronica_chamaedrys	Polygonum_aviculare	Rumex_acetosa	Ranunculus_acris	Ranunculus_repens	Potentilla_reptans	Sanguisorba_officinalis	Galium_mollugo	Acer_spec.
	season	2018_E	2018_F	2018_E,F,G,H	2018_E,F,G,H	2018_F,H	2018_E	2018_E,F,G,H	2018_E	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G,H
Season	plot															
2018_G	A002	0	0	-0.244	-0.085	0	0	-0.138	0	0	-0.138	0	0	0	0	0
2018_G	A003	0	0	-0.185	0	0	0	-0.185	0	-0.100	-0.274	0	0	0	0	-0.035
2018_G	A005	0	0	-0.361	0	0	0	-0.078	0	-0.230	-0.322	0	-0.046	-0.046	0	0
2018_G	A009	0	0	-0.359	0	0	0	-0.125	0	-0.076	0	0	0	-0.076	0	0
2018_G	A010	0	0	-0.127	0	0	0	-0.077	0	-0.127	0	0	0	0	0	0
2018_G	A011	0	0	-0.359	0	0	0	0	0	-0.125	-0.045	0	0	-0.045	0	0
2018_G	A013	0	0	-0.365	0	0	0	0	0	-0.048	0	0	0	0	-0.048	0
2018_G	A016	0	0	-0.263	-0.263	0	0	-0.263	0	-0.152	-0.263	0	0	0	0	0
2018_G	A018	0	0	-0.238	0	0	0	0	0	-0.082	-0.134	0	0	0	0	0
2018_G	A020	0	0	-0.240	-0.049	0	0	0	0	-0.135	0	0	0	-0.083	0	0
2018_G	A026	0	0	-0.128	0	0	0	0	0	-0.078	0	0	0	-0.046	0	-0.046
2018_G	A027	0	0	-0.177	-0.067	0	0	0	0	-0.067	0	0	0	0	0	0
2018_G	A030	0	0	-0.215	0	0	0	0	0	-0.042	0	0	0	0	0	0
2018_G	A035	0	0	-0.240	0	0	0	0	0	-0.083	0	0	0	0	0	0
2018_G	A040	0	0	-0.227	0	0	0	-0.147	0	-0.091	-0.091	0	0	-0.091	0	-0.091
2018_G	A042	0	0	-0.152	0	0	0	0	0	-0.056	0	0	0	0	0	0
2018_G	A043	0	0	-0.327	0	0	0	-0.100	0	-0.100	0	0	0	0	0	0
2018_G	A044	0	0	-0.331	0	0	0	0	0	-0.073	0	0	0	0	0	0
2018_G	A045	0	0	-0.124	0	0	0	-0.075	0	0	0	0	0	-0.044	0	0
2018_G	A046	0	0	0	0	0	0	0	0	-0.091	0	0	0	0	0	-0.054
2018_G	B048	0	0	-0.162	0	0	0	-0.101	0	-0.101	0	0	0	-0.060	-0.060	0
2018_G	B051	0	0	-0.349	-0.070	0	0	0	0	-0.041	0	0	0	0	0	0
2018_G	B054	0	0	-0.357	-0.064	0	0	-0.107	0	-0.064	0	0	0	0	0	0
2018_G	B057	0	0	-0.352	-0.042	0	0	0	0	0	-0.215	0	0	0	0	0
2018_G	B059	0	0	-0.145	0	0	0	0	0	-0.145	0	0	0	0	0	0
2018_G	B060	0	0	-0.250	-0.143	0	0	-0.087	0	-0.143	0	0	0	0	0	0
2018_G	B064	0	0	-0.052	0	0	0	0	0	-0.052	-0.052	0	0	0	0	-0.052
2018_G	B067	0	0	-0.092	0	0	0	0	0	-0.032	-0.032	0	0	0	0	0
2018_G	B071	0	0	-0.159	0	0	0	-0.099	0	-0.059	-0.355	0	0	0	0	0
2018_G	B073	0	0	-0.187	0	0	0	0	0	0	-0.035	0	0	0	0	0
2018_G	B075	0	0	-0.338	0	0	0	0	0	0	0	0	0	0	-0.051	0
2018_G	B080	0	0	-0.229	0	0	0	0	0	0	-0.091	0	0	0	0	0
2018_G	B081	0	0	-0.263	-0.056	0	0	-0.094	0	0	-0.152	0	0	0	0	0
2018_G	B085	0	0	-0.134	0	0	0	-0.134	0	-0.082	0	0	0	0	0	0
2018_G	B090	0	0	-0.137	0	0	0	0	0	-0.050	-0.050	0	0	0	-0.050	0
2018_G	B092	0	0	-0.081	0	0	0	0	0	-0.048	0	0	0	0	-0.048	0
2018_G	C097	0	0	-0.155	0	0	0	-0.155	0	-0.155	-0.057	0	0	0	0	0
2018_G	C103	0	0	0	0	0	0	-0.109	0	-0.035	0	0	0	0	-0.035	0
2018_G	C109	0	0	0	0	0	0	-0.047	0	-0.153	-0.047	0	0	0	-0.027	0
2018_G	C110	0	0	-0.262	-0.055	0	0	-0.151	0	-0.093	-0.055	0	0	0	0	0
2018_G	C114	0	0	-0.043	0	0	0	0	0	-0.043	0	0	0	0	0	0
2018_G	C115	0	0	-0.359	-0.076	0	0	0	0	-0.125	-0.225	0	0	0	0	0
2018_G	C121	0	0	-0.132	0	0	0	0	0	0	-0.048	-0.048	0	-0.081	0	0
2018_G	C131	0	0	0	0	0	0	0	0	-0.043	-0.043	0	0	0	0	0
2018_G	C133	0	0	0	-0.155	0	0	-0.155	0	-0.328	0	0	0	0	0	0
2018_G	C135	0	0	-0.070	0	0	0	0	0	0	0	0	0	0	0	0
2018_G	C136	0	0	-0.360	0	0	0	-0.088	0	0	0	0	0	0	0	0
2018_G	C137	0	0	-0.352	0	0	0	0	0	0	0	0	0	0	0	0

	functional group	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb	herb
	family	Papaveraceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Plantaginaceae	Polygonaceae	Polygonaceae	Ranunculaceae	Ranunculaceae	Rosaceae	Rosaceae	Rubiaceae	Sapindaceae
	Code	Paparg	Linvul	Plalan	Plamed	Veranv	Veracf	Vercha	Polavi	Rumace	Ranacr	Ranrep	Potrep	Sagoff	Galmol	Acespe
	species	Papaver_argemone	Linaria_vulgaris	Plantago_lanceolata	Plantago_media	Veronica_arvensis	Veronica_arvensis cf.	Veronica_chamaedrys	Polygonum_aviculare	Rumex_acetosa	Ranunculus_acris	Ranunculus_repens	Potentilla_reptans	Sanguisorba_officinalis	Galium_mollugo	Acer_spec.
	season	2018_E	2018_F	2018_E,F,G,H	2018_E,F,G,H	2018_F,H	2018_E	2018_E,F,G,H	2018_E	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G	2018_G,H	2018_E,F,G,H	2018_E,F,G,H	2018_E,F,G,H
Season	plot															
2018_H	A002	0	0	0	0	-0.033	0	-0.177	0	-0.033	-0.177	0	0	0	0	0
2018_H	A003	0	0	-0.084	0	0	0	0	0	-0.084	-0.296	0	0	0	0	0
2018_H	A005	0	0	-0.090	0	0	0	-0.169	0	-0.169	-0.255	0	-0.031	-0.031	0	0
2018_H	A009	0	0	-0.055	0	0	0	-0.260	0	-0.260	0	0	0	-0.092	0	0
2018_H	A010	0	0	-0.193	0	0	0	-0.334	0	-0.104	-0.062	0	0	0	0	0
2018_H	A011	0	0	-0.294	0	0	0	0	0	-0.157	-0.083	0	0	-0.049	0	0
2018_H	A013	0	0	-0.352	0	0	0	0	0	-0.097	0	0	0	-0.033	0	-0.033
2018_H	A016	0	0	-0.081	0	0	0	-0.154	0	-0.236	-0.327	0	0	0	0	0
2018_H	A018	0	0	-0.249	0	0	0	-0.087	0	-0.165	-0.087	0	0	0	0	0
2018_H	A020	0	0	-0.212	0	0	0	0	0	-0.117	-0.117	0	0	-0.070	0	0
2018_H	A026	0	0	-0.051	0	0	0	0	0	-0.087	0	0	0	-0.030	0	0
2018_H	A027	0	0	-0.296	0	0	0	0	0	-0.205	-0.067	0	0	0	0	0
2018_H	A030	0	0	-0.254	0	0	0	-0.089	0	0	-0.031	0	0	0	0	0
2018_H	A035	0	0	-0.128	0	0	0	-0.229	0	-0.229	0	0	0	0	0	0
2018_H	A040	0	0	-0.274	0	0	0	-0.161	0	-0.161	0	0	0	-0.059	0	0
2018_H	A042	0	0	-0.220	0	-0.043	0	0	0	-0.122	0	0	0	0	0	-0.043
2018_H	A043	0	0	-0.368	-0.093	0	0	0	0	0	0	0	0	0	0	0
2018_H	A044	0	0	-0.334	0	0	0	-0.193	0	-0.104	0	0	0	0	0	0
2018_H	A045	0	0	-0.068	0	0	0	0	0	-0.068	0	0	0	-0.023	0	0
2018_H	A046	0	0	-0.239	0	0	0	-0.082	0	-0.082	-0.082	0	0	0	0	0
2018_H	B048	0	0	-0.157	0	-0.028	0	-0.240	0	-0.157	0	0	0	-0.049	-0.028	0
2018_H	B051	0	0	-0.195	0	0	0	-0.195	0	-0.105	-0.105	0	0	0	0	0
2018_H	B054	0	0	-0.138	0	0	0	0	0	-0.138	-0.138	0	0	0	0	0
2018_H	B057	0	0	-0.195	0	0	0	0	0	-0.105	-0.105	0	0	0	0	0
2018_H	B059	0	0	-0.201	0	0	0	0	0	-0.110	0	0	0	0	0	0
2018_H	B060	0	0	-0.173	0	0	0	0	0	-0.260	0	0	0	0	0	0
2018_H	B064	0	0	-0.061	0	0	0	0	0	-0.102	0	0	0	0	0	0
2018_H	B067	0	0	-0.143	-0.025	0	0	-0.143	0	-0.074	-0.025	0	0	0	0	0
2018_H	B071	0	0	-0.220	0	0	0	-0.026	0	-0.148	-0.283	0	0	0	0	0
2018_H	B073	0	0	-0.141	0	0	0	0	0	-0.073	0	0	0	0	0	0
2018_H	B075	0	0	-0.230	0	0	0	0	0	-0.078	-0.026	0	0	0	-0.078	0
2018_H	B080	0	0	-0.358	0	0	0	0	0	0	-0.107	0	0	0	0	0
2018_H	B081	0	0	-0.274	0	0	0	-0.185	0	-0.100	-0.185	0	0	0	0	0
2018_H	B085	0	0	-0.136	-0.024	-0.024	0	-0.136	0	-0.070	0	0	0	0	0	0
2018_H	B090	0	0	-0.247	0	0	0	0	0	-0.086	0	0	0	0	0	0
2018_H	B092	0	0	0	0	0	0	-0.075	0	-0.075	0	0	0	0	0	0
2018_H	C097	0	0	-0.244	0	0	0	-0.161	0	-0.161	-0.050	0	0	0	0	0
2018_H	C103	0	0	0	0	0	0	0	0	-0.138	0	0	0	0	0	0
2018_H	C109	0	0	-0.157	0	0	0	0	0	-0.157	0	0	0	0	0	0
2018_H	C110	0	0	-0.283	0	0	0	-0.143	0	-0.228	0	0	0	0	0	0
2018_H	C114	0	0	-0.094	0	0	0	0	0	-0.032	-0.056	0	0	0	0	0
2018_H	C115	0	0	-0.209	0	0	0	0	0	-0.209	-0.301	0	0	0	0	0
2018_H	C121	0	0	-0.253	0	0	0	-0.145	0	0	0	0	0	-0.053	0	0
2018_H	C131	0	0	-0.071	0	0	0	0	0	-0.071	-0.071	0	0	0	0	0
2018_H	C133	0	0	-0.184	0	0	0	-0.099	0	-0.184	0	0	0	0	0	0
2018_H	C135	0	0	-0.088	0	0	0	0	0	-0.088	0	0	0	0	0	0
2018_H	C136	0	0	-0.218	-0.025	0	0	0	0	0	0	0	0	0	-0.025	0
2018_H	C137	0	0	-0.225	0	0	0	0	0	0	0	0	0	0	0	0

Supplement Table 1: Plant species indices – continued (FG grass Season 2018 E)

	functional group	grass	grass	grass	grass	grass	grass	grass	grass	grass	grass	grass	grass
	family	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae
	Code	Antodo	Avepub	Brohor	Dacglo	Fesovi	Fespra	Fesrub	Fesspe	Hollan	Lolper	Phlpra	Poapra
	species	Anthoxanthum_odoratum	Avenula_pubescens	Bromus_hordeaceus	Dactylis_glomerata	Festuca_ovina agg.	Festuca_pratensis cf.	Festuca_rubra	Festuca_spec.	Holcus_lanatus	Lolium_perenne	Phleum_pratense	Poa_pratensis
	season	2018_E,F,G,H	2018_E,F,G,H	2018_E,G,H	2018_E,F,G,H	2018_E,F,G,H	2018_F,G	2018_E,F,G,H	2018_G	2018_E,F,G,H	2018_E	2018_E,F,G,H	2018_E,F,G,H
Season	plot												
2018_E	A002	0	-0.348	-0.040	0	0	0	0	0	-0.040	0	-0.206	-0.068
2018_E	A003	-0.215	0	0	-0.306	0	0	-0.071	0	-0.042	-0.118	0	0
2018_E	A005	0	-0.060	0	-0.277	0	0	0	0	-0.101	-0.035	-0.188	-0.035
2018_E	A009	0	-0.071	0	0	0	0	0	0	-0.118	0	0	-0.071
2018_E	A010	0	-0.309	0	-0.031	0	0	0	0	-0.031	0	-0.169	0
2018_E	A011	0	-0.033	0	0	0	0	0	0	0	0	-0.033	-0.033
2018_E	A013	0	0	0	0	0	0	0	0	0	0	0	-0.052
2018_E	A016	0	0	-0.038	0	-0.066	0	0	0	0	0	-0.110	-0.066
2018_E	A018	-0.086	0	0	-0.162	0	0	0	0	-0.162	0	-0.086	0
2018_E	A020	0	0	0	0	0	0	-0.036	0	-0.062	0	0	-0.062
2018_E	A026	0	0	0	0	0	0	-0.108	0	-0.065	-0.199	-0.290	-0.355
2018_E	A027	0	0	0	0	0	0	0	0	-0.144	0	0	-0.144
2018_E	A030	0	0	0	0	0	0	0	0	0	-0.037	0	-0.064
2018_E	A035	0	0	0	0	0	0	0	0	-0.290	0	-0.199	-0.038
2018_E	A040	0	0	0	0	0	0	0	0	-0.367	0	0	0
2018_E	A042	0	0	0	0	0	0	0	0	0	0	0	-0.071
2018_E	A043	-0.035	-0.035	0	0	0	0	0	0	-0.061	0	-0.035	0
2018_E	A044	-0.209	-0.040	0	0	0	0	-0.069	0	-0.069	0	0	0
2018_E	A045	-0.035	-0.368	0	-0.035	0	0	-0.061	0	0	-0.189	-0.189	-0.061
2018_E	A046	0	-0.043	0	-0.025	0	0	0	0	0	0	-0.074	-0.043
2018_E	B048	0	0	0	0	0	0	0	0	-0.029	0	0	-0.050
2018_E	B051	-0.025	-0.074	0	-0.338	0	0	0	0	-0.142	-0.025	-0.274	-0.025
2018_E	B054	-0.091	0	-0.031	0	0	0	0	0	0	-0.091	-0.311	0
2018_E	B057	-0.022	0	-0.022	0	0	0	0	0	-0.128	0	-0.066	-0.128
2018_E	B059	0	0	-0.076	0	0	0	-0.076	0	-0.076	0	-0.125	-0.125
2018_E	B060	-0.255	0	-0.090	0	0	0	0	0	-0.090	-0.053	-0.090	-0.169
2018_E	B064	0	0	0	0	0	0	-0.367	0	-0.057	0	0	-0.178
2018_E	B067	-0.025	0	0	0	0	0	-0.311	0	0	-0.025	-0.219	-0.219
2018_E	B071	-0.252	0	0	0	0	0	0	0	0	0	0	0
2018_E	B073	0	-0.243	0	-0.050	0	0	-0.084	0	-0.159	0	-0.159	-0.159
2018_E	B075	0	0	-0.090	0	0	0	-0.090	0	-0.031	0	0	-0.170
2018_E	B080	0	0	0	-0.046	0	0	0	0	-0.284	0	-0.046	0
2018_E	B081	0	0	0	0	0	0	-0.056	0	-0.364	-0.032	0	0
2018_E	B085	0	-0.056	-0.032	0	0	0	-0.056	0	0	-0.056	-0.094	-0.056
2018_E	B090	0	0	0	-0.250	0	0	-0.030	0	0	-0.052	0	-0.165
2018_E	B092	0	-0.366	-0.040	-0.113	0	0	-0.346	0	0	-0.113	-0.068	-0.113
2018_E	C097	-0.171	-0.091	0	-0.363	0	0	-0.091	0	0	0	0	-0.257
2018_E	C103	-0.052	-0.305	0	0	0	0	-0.166	0	0	0	0	-0.088
2018_E	C109	0	0	0	-0.061	0	0	0	0	-0.308	0	0	-0.119
2018_E	C110	0	-0.364	0	-0.031	0	0	0	0	-0.169	0	-0.090	0
2018_E	C114	0	0	0	0	0	0	-0.363	0	0	-0.145	0	-0.368
2018_E	C115	-0.194	-0.063	0	-0.063	0	0	0	0	-0.063	0	-0.123	0
2018_E	C121	-0.038	-0.108	0	0	0	0	0	0	0	-0.065	-0.108	-0.108
2018_E	C131	0	0	0	0	0	0	-0.086	0	0	0	-0.162	-0.344
2018_E	C133	0	0	-0.113	0	0	0	0	0	0	0	-0.299	0
2018_E	C135	0	-0.192	0	0	0	0	0	0	0	0	-0.062	-0.104
2018_E	C136	0	-0.307	0	0	0	0	0	0	0	-0.053	-0.167	0
2018_E	C137	-0.060	-0.101	-0.035	-0.368	0	0	0	0	-0.035	0	-0.101	-0.187

Supplement Table 1: Plant species indices – continued (FG grass Season 2018 F)

	functional group	grass	grass	grass	grass	grass	grass	grass	grass	grass	grass	grass	grass	grass
	family	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae
	Code	Antodo	Avepub	Brohor	Dacglo	Fesovi	Fespra	Fesrub	Fesspe	Hollan	Lolper	Phlpra	Poapra	
	species	Anthoxanthum_odoratum	Avenula_pubescens	Bromus_hordeaceus	Dactylis_glomerata	Festuca_ovina agg.	Festuca_pratensis cf.	Festuca_rubra	Festuca_spec.	Holcus_lanatus	Lolium_perenne	Phleum_pratense	Poa_pratensis	
	season	2018_E,F,G,H	2018_E,F,G,H	2018_E,G,H	2018_E,F,G,H	2018_E,F,G,H	2018_F,G	2018_E,F,G,H	2018_G	2018_E,F,G,H	2018_E	2018_E,F,G,H	2018_E,F,G,H	
Season	plot													
2018_F	A002	0	-0.360	0	0	0	0	0	0	0	0	-0.215	0	
2018_F	A003	-0.146	0	0	-0.343	0	0	-0.146	0	0	0	0	-0.053	
2018_F	A005	0	0	0	-0.215	0	0	0	0	0	0	-0.334	0	
2018_F	A009	0	0	0	0	0	0	-0.177	0	0	0	-0.264	0	
2018_F	A010	0	-0.189	0	0	0	-0.119	0	0	0	0	-0.119	0	
2018_F	A011	0	0	0	0	0	0	0	0	0	0	0	0	
2018_F	A013	0	0	0	0	0	0	-0.162	0	0	0	0	0	
2018_F	A016	0	0	0	0	0	0	-0.141	0	0	0	-0.086	0	
2018_F	A018	-0.116	0	0	-0.211	0	0	0	0	-0.116	0	-0.116	0	
2018_F	A020	0	0	0	0	0	0	0	0	0	0	0	-0.085	
2018_F	A026	0	0	0	0	0	0	0	0	0	0	-0.278	-0.240	
2018_F	A027	0	0	0	0	0	0	0	0	0	0	0	0	
2018_F	A030	0	0	0	0	0	0	0	0	0	0	0	0	
2018_F	A035	0	0	0	0	0	0	0	0	-0.130	0	-0.323	0	
2018_F	A040	0	0	0	0	0	0	0	0	-0.294	0	0	0	
2018_F	A042	0	0	0	0	0	0	0	0	0	0	0	0	
2018_F	A043	0	0	0	0	0	0	0	0	0	0	0	0	
2018_F	A044	-0.221	0	0	0	0	0	0	0	0	0	0	0	
2018_F	A045	0	-0.230	0	-0.092	-0.150	0	-0.347	0	0	0	-0.230	-0.347	
2018_F	A046	0	0	0	0	0	0	0	0	0	0	-0.138	0	
2018_F	B048	-0.197	0	0	0	0	0	0	0	0	0	0	0	
2018_F	B051	0	0	0	-0.364	0	0	0	0	0	0	-0.237	0	
2018_F	B054	0	0	0	0	0	0	0	0	0	0	-0.332	0	
2018_F	B057	0	0	0	0	0	0	-0.272	0	0	0	-0.159	-0.355	
2018_F	B059	0	0	0	0	0	0	0	0	0	0	-0.310	0	
2018_F	B060	-0.365	0	0	0	0	0	0	0	0	0	-0.207	0	
2018_F	B064	0	0	0	0	0	0	-0.152	0	0	0	0	0	
2018_F	B067	0	0	0	0	0	-0.100	-0.356	0	0	0	-0.161	-0.356	
2018_F	B071	-0.344	0	0	0	0	0	0	0	0	0	-0.147	0	
2018_F	B073	0	-0.286	0	0	0	0	-0.286	0	0	0	-0.106	-0.286	
2018_F	B075	0	0	0	0	0	0	-0.340	0	0	0	0	-0.340	
2018_F	B080	0	0	0	0	0	0	0	0	-0.303	0	0	0	
2018_F	B081	0	-0.189	0	0	0	0	0	0	-0.368	0	-0.119	0	
2018_F	B085	0	0	0	0	0	0	0	0	0	0	0	0	
2018_F	B090	0	0	0	-0.284	0	0	0	0	0	0	0	0	
2018_F	B092	0	-0.349	0	-0.234	-0.094	0	-0.342	0	0	0	0	0	
2018_F	C097	0	0	0	-0.316	0	0	0	0	0	0	0	0	
2018_F	C103	0	-0.250	0	0	0	0	-0.250	0	0	0	-0.143	0	
2018_F	C109	0	0	0	-0.240	0	0	0	0	-0.331	0	0	0	
2018_F	C110	0	-0.365	0	0	0	0	0	0	0	0	0	0	
2018_F	C114	0	0	0	0	-0.121	-0.043	-0.357	0	0	0	0	-0.357	
2018_F	C115	-0.267	0	0	0	0	0	0	0	0	0	-0.096	0	
2018_F	C121	0	0	0	0	0	0	0	0	0	0	0	0	
2018_F	C131	0	0	0	0	0	0	0	0	0	0	-0.141	-0.087	
2018_F	C133	0	0	0	0	0	0	0	0	0	0	-0.368	0	
2018_F	C135	0	0	0	0	0	0	-0.177	0	0	0	0	0	
2018_F	C136	0	-0.244	0	-0.138	0	0	0	0	0	0	-0.244	0	
2018_F	C137	0	0	0	-0.335	0	0	0	0	0	0	0	0	

Supplement Table 1: Plant species indices – continued (FG grass Season 2018 G)

	functional group	grass	grass	grass	grass	grass	grass	grass	grass	grass	grass	grass	grass
	family	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae
	Code	Antodo	Avepub	Brohor	Dacglo	Fesovi	Fespra	Fesrub	Fesspe	Hollan	Lolper	Phlpra	Poapra
	species	Anthoxanthum_odoratum	Avenula_pubescens	Bromus_hordeaceus	Dactylis_glomerata	Festuca_ovina agg.	Festuca_pratensis cf.	Festuca_rubra	Festuca_spec.	Holcus_lanatus	Lolium_perenne	Phleum_pratense	Poa_pratensis
	season	2018_E,F,G,H	2018_E,F,G,H	2018_E,G,H	2018_E,F,G,H	2018_E,F,G,H	2018_F,G	2018_E,F,G,H	2018_G	2018_E,F,G,H	2018_E	2018_E,F,G,H	2018_E,F,G,H
Season	plot												
2018_G	A002	0	-0.366	-0.085	0	0	0	0	0	0	0	-0.334	0
2018_G	A003	-0.100	0	-0.059	-0.274	0	0	0	0	0	0	0	-0.035
2018_G	A005	0	0	0	-0.230	0	0	0	0	0	0	-0.230	0
2018_G	A009	0	0	-0.225	0	0	0	-0.125	0	0	0	-0.045	0
2018_G	A010	0	-0.228	0	0	0	0	0	0	0	0	-0.127	0
2018_G	A011	0	-0.125	-0.045	0	0	0	0	0	0	0	0	0
2018_G	A013	0	0	0	0	0	0	0	0	0	0	-0.048	0
2018_G	A016	0	0	0	0	0	0	0	0	0	0	-0.152	0
2018_G	A018	-0.134	0	0	-0.238	0	0	0	0	-0.134	0	-0.238	0
2018_G	A020	0	0	0	0	0	0	0	0	-0.049	0	0	0
2018_G	A026	0	0	0	0	0	0	0	0	0	0	-0.229	-0.254
2018_G	A027	0	0	0	0	0	0	0	-0.067	0	0	0	0
2018_G	A030	0	0	0	0	0	0	0	0	0	0	0	0
2018_G	A035	0	0	-0.083	0	0	0	0	0	-0.240	0	-0.135	0
2018_G	A040	0	0	-0.091	0	0	0	0	0	-0.344	0	0	0
2018_G	A042	0	0	0	0	0	-0.056	0	0	0	0	0	0
2018_G	A043	0	0	-0.059	0	0	0	0	0	0	0	0	0
2018_G	A044	-0.310	0	0	0	0	0	0	0	0	0	0	0
2018_G	A045	0	-0.357	0	0	-0.044	0	-0.315	0	0	0	-0.223	-0.315
2018_G	A046	0	0	0	0	0	0	0	0	0	0	-0.149	0
2018_G	B048	0	0	0	0	0	0	0	0	0	0	0	0
2018_G	B051	0	0	0	-0.302	0	0	0	0	0	0	-0.302	0
2018_G	B054	0	0	0	0	0	0	0	0	0	0	-0.288	0
2018_G	B057	0	0	0	0	0	0	0	0	0	0	-0.118	-0.352
2018_G	B059	0	0	0	0	0	0	-0.145	0	0	0	-0.316	0
2018_G	B060	-0.367	0	0	0	0	0	0	0	0	0	0	0
2018_G	B064	0	0	-0.052	0	0	0	-0.160	0	0	0	0	0
2018_G	B067	0	0	0	0	0	0	-0.363	0	0	0	-0.173	-0.363
2018_G	B071	-0.272	0	0	0	0	0	-0.099	0	0	0	0	0
2018_G	B073	0	-0.187	0	0	0	0	-0.276	0	0	0	-0.187	-0.276
2018_G	B075	0	0	0	0	0	0	-0.303	0	0	0	-0.051	-0.303
2018_G	B080	0	-0.054	-0.054	0	0	0	0	0	-0.091	0	-0.054	0
2018_G	B081	0	0	-0.094	0	0	0	0	0	-0.368	0	0	0
2018_G	B085	0	-0.134	0	0	0	0	0	0	0	0	0	0
2018_G	B090	0	0	0	-0.137	0	0	0	0	0	0	-0.050	0
2018_G	B092	0	-0.364	0	0	-0.048	0	-0.338	0	0	0	0	0
2018_G	C097	0	0	0	-0.275	0	0	0	0	0	0	0	0
2018_G	C103	0	-0.278	0	0	0	0	-0.102	0	0	0	0	0
2018_G	C109	0	0	-0.047	-0.289	0	0	0	0	-0.289	0	0	0
2018_G	C110	0	-0.368	0	0	0	0	0	0	0	0	0	0
2018_G	C114	0	0	0	0	-0.043	0	-0.357	0	0	0	-0.043	-0.357
2018_G	C115	-0.359	0	0	0	0	0	0	0	0	0	-0.125	0
2018_G	C121	0	0	0	0	0	0	0	0	0	0	0	0
2018_G	C131	0	0	0	0	0	0	-0.357	0	0	0	-0.043	-0.357
2018_G	C133	0	0	0	0	0	0	0	0	0	0	-0.328	0
2018_G	C135	0	0	0	0	0	0	0	0	0	0	0	-0.041
2018_G	C136	0	-0.251	0	0	0	0	0	0	0	0	-0.251	0
2018_G	C137	0	0	-0.039	-0.365	0	0	0	0	0	0	0	0

Supplement Table 1: Plant species indices – continued (FG grass Season 2018 H)

	functional group	grass	grass	grass	grass	grass	grass	grass	grass	grass	grass	grass	grass
	family	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae
	Code	Antodo	Avepub	Brohor	Dacglo	Fesovi	Fespra	Fesrub	Fesspe	Hollan	Lolper	Phlpra	Poapra
	species	Anthoxanthum_odoratum	Avenula_pubescens	Bromus_hordeaceus	Dactylis_glomerata	Festuca_ovina agg.	Festuca_pratensis cf.	Festuca_rubra	Festuca_spec.	Holcus_lanatus	Lolium_perenne	Phleum_pratense	Poa_pratensis
	season	2018_E,F,G,H	2018_E,F,G,H	2018_E,G,H	2018_E,F,G,H	2018_E,F,G,H	2018_F,G	2018_E,F,G,H	2018_G	2018_E,F,G,H	2018_E	2018_E,F,G,H	2018_E,F,G,H
Season	plot												
2018_H	A002	0	-0.341	0	0	0	0	0	0	0	0	-0.094	0
2018_H	A003	-0.159	0	0	-0.354	0	0	0	0	0	0	0	0
2018_H	A005	0	0	0	-0.361	0	0	0	0	0	0	-0.031	0
2018_H	A009	0	0	0	0	0	0	0	0	0	0	0	0
2018_H	A010	0	-0.283	0	0	0	0	0	0	-0.193	0	-0.036	0
2018_H	A011	0	-0.083	0	0	0	0	0	0	0	0	0	0
2018_H	A013	0	0	0	0	0	0	0	0	0	0	0	0
2018_H	A016	0	-0.154	-0.081	0	0	0	0	0	0	0	0	0
2018_H	A018	-0.249	0	0	-0.304	0	0	0	0	-0.087	0	-0.030	0
2018_H	A020	0	0	-0.117	0	0	0	0	0	0	0	0	0
2018_H	A026	0	0	0	0	0	0	-0.367	0	0	0	-0.164	-0.366
2018_H	A027	0	0	-0.112	0	0	0	0	0	0	0	0	0
2018_H	A030	0	0	0	0	0	0	0	0	0	0	0	0
2018_H	A035	0	-0.229	0	-0.128	0	0	0	0	-0.229	0	-0.046	0
2018_H	A040	0	0	0	0	0	0	0	0	-0.274	0	0	0
2018_H	A042	0	0	0	0	0	0	0	0	0	0	0	0
2018_H	A043	0	0	0	0	0	0	0	0	0	0	0	0
2018_H	A044	-0.360	-0.104	0	0	0	0	0	0	0	0	0	0
2018_H	A045	0	-0.360	0	0	-0.068	0	-0.360	0	0	0	-0.133	-0.299
2018_H	A046	0	0	0	0	0	0	0	0	0	0	-0.028	0
2018_H	B048	0	0	-0.083	0	0	0	0	0	0	0	0	0
2018_H	B051	0	-0.037	0	-0.336	0	0	0	0	0	0	-0.037	0
2018_H	B054	0	0	0	-0.366	0	0	0	0	0	0	-0.050	0
2018_H	B057	0	0	0	0	0	0	0	0	0	0	-0.037	-0.285
2018_H	B059	-0.110	0	-0.066	0	0	0	0	0	0	0	-0.038	0
2018_H	B060	-0.363	0	0	0	0	0	0	0	0	0	0	0
2018_H	B064	0	0	0	0	-0.102	0	-0.347	0	0	0	0	0
2018_H	B067	0	-0.044	0	-0.074	0	0	-0.365	0	0	0	-0.025	-0.313
2018_H	B071	-0.320	0	0	0	0	0	0	0	0	0	0	0
2018_H	B073	0	-0.272	0	0	0	0	-0.272	0	0	0	-0.025	-0.272
2018_H	B075	0	0	0	0	0	0	-0.347	0	0	0	0	-0.347
2018_H	B080	0	0	0	-0.107	0	0	-0.196	0	-0.107	0	0	0
2018_H	B081	0	-0.185	0	0	0	0	0	0	-0.327	0	0	0
2018_H	B085	0	0	-0.136	0	0	0	0	0	0	0	0	0
2018_H	B090	0	0	0	-0.302	0	0	0	0	0	0	0	0
2018_H	B092	0	-0.366	0	0	-0.044	0	-0.364	0	0	0	0	0
2018_H	C097	0	0	-0.244	-0.367	0	0	0	0	0	0	0	0
2018_H	C103	0	-0.352	0	-0.071	0	0	-0.306	0	0	0	0	0
2018_H	C109	0	0	0	-0.239	0	0	0	0	-0.239	0	0	0
2018_H	C110	0	-0.360	0	0	0	0	0	0	-0.077	0	0	0
2018_H	C114	0	0	0	0	-0.056	0	-0.360	0	0	0	0	-0.360
2018_H	C115	-0.301	0	0	0	0	0	0	0	0	0	-0.040	0
2018_H	C121	0	0	0	0	0	0	0	0	0	0	0	0
2018_H	C131	0	-0.071	-0.071	0	0	0	0	0	0	0	0	-0.187
2018_H	C133	0	0	0	0	0	0	0	0	0	0	-0.099	0
2018_H	C135	0	-0.088	-0.088	0	0	0	0	0	0	0	0	0
2018_H	C136	0	-0.364	-0.073	0	0	0	0	0	-0.073	0	-0.073	0
2018_H	C137	0	0	0	-0.352	0	0	0	0	0	0	0	0

Supplement Table 2: Tool versions and parameters used in the Galaxy-W4M workflow. The workflow for the LC-MS raw data pre-processing is available at https://doi.workflow4metabolomics.org/W4M00008 using an example dataset.

Tool name	Description	Version	Parameter	Value
MSnbase readMSData	Imports mass-spectrometry data files	2.8.2.2		
findChromPeaks	Chromatographic peak detection	3.6.1+galaxy1	Extraction method for peaks detection	CentWave
			max tolerated ppm m/z deviation	25
			Min. Max peak width	5 s, 20 s
			Signal to noise ratio cutoff	5
			Prefilter step (ROI detection)	2 scans, 1000 intensity
			Minimum difference in m/z for peaks with overlapping retention times	-0.005
			Noise filter	2000
xcms findChromPeaks Merger	Merge xcms findChromPeaks RData into a unique file to be used by group	3.6.1+galaxy0		
xcms groupChromPeaks (group)	Perform the correspondence, the grouping of chromatographic peaks within and between samples	3.6.1+galaxy2	Method	PeakDensity
			Bandwidth	4 s
			Minimum fraction of samples	0.75
			Minimum number of samples	1
			Width of overlapping m/z slices	0.25
xcms adjustRtime (retcor)	Retention Time Correction	3.6.1+galaxy1	Method	PeakGroups
			Min required fraction	0.75
			Max number of additional peaks	1
			Smooth method	Loess – non-linear alignment
			Degree of smoothing	0.2
			Family	gaussian
xcms groupChromPeaks (group)	Perform the correspondence, the grouping of chromatographic peaks within and between samples	3.6.1+galaxy2	Method	PeakDensity
			Bandwidth	3 s
			Minimum fraction of samples	0.75
			Minimum number of samples	1
			Width of overlapping m/z slices	0.005
CAMERA annotate	Returns annotation results (isotope peaks, adducts and fragments)	2.2.6+camera1.42.0-galaxy1	Multiplier of the standard deviation	6
			Percentage of FWHM width	0.6
			General ppm error	5
			General absolut error in m/z	0.005
			Max. ion charge	3
			Max. number of expected isotopes	4
			The percentage number of samples, which must satisfy the C12/C13 rule for isotope annotation	0.5
			groupCorr: correlation threshold (0..1)	0.75
			groupCorr: Method selection for grouping peaks after correlation analysis into pseudospectra	hcs
			groupCorr: significant correlation threshold	0.5319
Check Format	Checking/formatting the sample and variable names	3.0.0		
Generic_Filter	Deleting samples and/or variables	2020.01	remove in " ... " values upper	"rt", 880 (s)
			remove in " ... " values lower	"rt", 80 (s)

Supplement Table 3: Metabolite families for PLS selected features putatively annotated. The features were selected in thirteen species for seven factors. Putative annotations were based on database searches performed by MetFamily. Only those annotations that showed certain statistical reliability ($p <= 0.05$) were accepted. For each annotated family, we assigned a putative metabolite family ID (pMetFam_ID) used to indicate the features used for the statistical analysis. a) shows the annotation of each feature, for which trait in which species it was selected and which library was used for the putative annotation. b) Shows the intensity of the feature measured in each sample and the number of samples (across all species) that had an intensity measured for the feature. Note that to allow the table to be searchable by species and trait, features that had been PLS-selected for multiple species-trait combinations are listed for all these combinations.

Supplement Table 3a: Metabolite families - continued

Species	TwoLetterCode	Factor	Feature	mz	rt_s	rt_m	species	Library - database used	annotated mz	annotated rt_m	p-value	pMetFam_ID	Kingdom	Superclass	Class	Subclass	Parent	count
Pouyannia	PA	obsShan	M161T176	161.1072465	177.095	2.966	PA	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	161.103	2.999	0.040	pMetFam_01-A1-1	(01) Organic nitrogen compounds	(A) Organonitrogen compounds	(1) Amines	(1) Primary amines			43
Anthocaulon calcivum	A0	pathDam	M250T05	250.0934105	84.696	1.400	A0	i-111_Leibniz_positive_mode_20170001_2021-09-02	250.092	1.442	0.042	pMetFam_01-A1-1a	(01) Organic nitrogen compounds	(A) Organonitrogen compounds	(1) Amines	(1) Primary amines	(a) Monoalkylamines		71
Dactylis glomerata	DG	obsRich	M163T500	163.0393395	500.325	8.339	DG	L-2214_lib_spectra_pos_new_2021-09-02	163.039	8.388	0.046	pMetFam_02	(02) Organopnictogen compounds						292
Leucanthemum vulgare	LV	meshDam	M163T500	163.0393395	500.325	8.339	LV	L-2214_lib_spectra_pos_new_2021-09-02	163.039	8.388	0.046	pMetFam_02	(02) Organopnictogen compounds						292
Arenaria purpurascens	AP	BBCH	M199T175	199.0862288	175.195	2.920	AP	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	199.060	2.946	0.034	pMetFam_03	(03) Benzenoids						125
Plantago lanceolata	PL	obsShan	M177T045	177.0545564	316.273	5.271	PL	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	177.055	5.519	0.050	pMetFam_03	(03) Benzenoids						42
Plantago lanceolata	PL	BBCH	M177T316	177.0545465	316.273	5.271	PL	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	177.055	5.519	0.050	pMetFam_03	(03) Benzenoids						42
Plantago lanceolata	PL	season	M325T310	325.0931221	310.053	5.301	PL	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	325.092	5.309	0.050	pMetFam_03	(03) Benzenoids						43
Plantago lanceolata	PL	obsRich	M325T310	325.0931221	310.053	5.301	PL	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	325.092	5.309	0.050	pMetFam_03	(03) Benzenoids						43
Plantago lanceolata	PL	obsShan	M325T310	325.0931221	310.053	5.301	PL	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	325.092	5.309	0.050	pMetFam_03	(03) Benzenoids						43
Plantago lanceolata	PL	BBCH	M325T310	325.0931221	310.053	5.301	PL	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	325.092	5.309	0.050	pMetFam_03	(03) Benzenoids						43
Knautia arvensis	KA	pathDam	M215T107	215.0914279	107.195	3.120	KA	L-2214_lib_spectra_pos_new_2021-09-02	215.092	3.160	0.050	pMetFam_03-A1	(03) Benzenoids	(A) Benzene and substituted derivatives	(1) Phenylpropanes				14
Anthocaulon calcivum	A0	season	M183T100	183.0859932	107.693	1.795	A0	L-2214_lib_spectra_pos_new_2021-09-02	183.085	2.779	0.044	pMetFam_03-B1	(03) Benzenoids	(B) Phenols	(1) Methoxyphenols				17
Anthocaulon calcivum	A0	obsShan	M183T100	183.0859932	107.693	1.795	A0	L-2214_lib_spectra_pos_new_2021-09-02	183.085	2.779	0.044	pMetFam_03-B1	(03) Benzenoids	(B) Phenols	(1) Methoxyphenols				17
Anthocaulon calcivum	A0	height	M450T850	450.233122	549.736	9.162	A0	L-2214_lib_spectra_pos_new_2021-09-02	450.234	9.178	0.039	pMetFam_03-B1	(03) Benzenoids	(B) Phenols	(1) Methoxyphenols				14
Anthocaulon calcivum	A0	pathDam	M450T850	450.233122	549.736	9.162	A0	L-2214_lib_spectra_pos_new_2021-09-02	450.234	9.178	0.039	pMetFam_03-B1	(03) Benzenoids	(B) Phenols	(1) Methoxyphenols				14
Centaurea jacea	CJ	obsShan	M207T321	207.127593	321.489	5.358	CJ	L-2214_lib_spectra_pos_new_2021-09-02	207.128	5.386	0.037	pMetFam_03-B1	(03) Benzenoids	(B) Phenols	(1) Methoxyphenols				157
Dactylis glomerata	DG	BBCH	M257T232	257.1304952	232.369	3.873	DG	L-2214_lib_spectra_pos_new_2021-09-02	257.129	3.894	0.037	pMetFam_03-B1	(03) Benzenoids	(B) Phenols	(1) Methoxyphenols				47
Festuca rubra	FR	obsRich	M209T263	209.1534057	262.556	4.376	FR	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	209.154	4.404	0.023	pMetFam_04-A1-1a	(04) Organic acids and derivatives	(A) Organic phosphoric acids and derivatives	(1) Phosphate esters		(a) Monoalkyl phosphates		446
Festuca rubra	FR	season	M403T229	403.1710527	229.024	3.817	CJ	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	403.179	3.841	0.012	pMetFam_04-B1-1	(04) Organic acids and derivatives	(B) Carboxylic acids and derivatives	(1) Amino acids peptides and analogues	(1) Peptides			202
Geranium pratense	GP	obsRich	M403T229	403.1710527	229.024	3.817	GP	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	403.179	3.841	0.012	pMetFam_04-B1-1	(04) Organic acids and derivatives	(B) Carboxylic acids and derivatives	(1) Amino acids peptides and analogues	(1) Peptides			202
Geranium pratense	GP	height	M403T733	403.2117453	733.259	12.222	GP	L-2214_lib_spectra_pos_new_2021-09-02	403.216	12.241	0.011	pMetFam_04-B2-1	(04) Organic acids and derivatives	(B) Carboxylic acids and derivatives	(2) Carboxylic acid derivatives	(1) Carboxylic acid esters			38
Anthocaulon calcivum	A0	season	M161T261	161.0959744	260.943	4.349	FR	L-2214_lib_spectra_pos_new_2021-09-02	161.060	3.583	0.016	pMetFam_04-B2-2a	(04) Organic acids and derivatives	(B) Carboxylic acids and derivatives	(2) Carboxylic acid derivatives	(2) Carboxylic acid amides	(a) Secondary carboxylic acid amides		203
Anthocaulon calcivum	A0	pathDam	M271T706	271.1630853	705.571	11.760	PA	L-2214_lib_spectra_pos_new_2021-09-02	271.190	11.779	0.038	pMetFam_04-B2-2a	(04) Organic acids and derivatives	(B) Carboxylic acids and derivatives	(2) Carboxylic acid derivatives	(2) Carboxylic acid amides	(a) Secondary carboxylic acid amides		92
Polygonum viviparum	PV	season	M161T261	161.0959744	260.943	4.349	FR	L-2214_lib_spectra_pos_new_2021-09-02	161.060	3.583	0.016	pMetFam_04-B2-2b	(04) Organic acids and derivatives	(B) Carboxylic acids and derivatives	(2) Carboxylic acid derivatives	(2) Carboxylic acid amides	(b) Tertiary carboxylic acid amides		203
Festuca rubra	FR	obsRich	M161T261	161.0959744	260.943	4.349	FR	L-2214_lib_spectra_pos_new_2021-09-02	161.060	3.583	0.016	pMetFam_04-B2-2b	(04) Organic acids and derivatives	(B) Carboxylic acids and derivatives	(2) Carboxylic acid derivatives	(2) Carboxylic acid amides	(b) Tertiary carboxylic acid amides		203
Festuca rubra	FR	BBCH	M161T261	161.0959744	260.943	4.349	FR	L-2214_lib_spectra_pos_new_2021-09-02	161.060	3.583	0.016	pMetFam_04-B2-2b	(04) Organic acids and derivatives	(B) Carboxylic acids and derivatives	(2) Carboxylic acid derivatives	(2) Carboxylic acid amides	(b) Tertiary carboxylic acid amides		203
Festuca rubra	FR	BBCH	M429T800	429.2905701	807.007	13.463	A0	L-2214_lib_spectra_pos_new_2021-09-02	429.300	13.466	0.042	pMetFam_05-A1	(05) Lipids and lipid-like molecules	(A) Steroids and steroid derivatives	(1) Spirostanes and derivatives				22
Leucanthemum vulgare	LV	obsShan	M453T838	453.1332352	838.133	13.969	LV	M+21900_MoNA-export-LC-MS-HS_Spectra_2021-09-02	453.133	13.979	0.024	pMetFam_05-B1-1	(05) Lipids and lipid-like molecules	(B) Fatty Acids	(1) Fatty acyl glycosides				54
Leucanthemum vulgare	LV	height	M453T838	453.1332352	838.133	13.969	LV	M+21900_MoNA-export-LC-MS-HS_Spectra_2021-09-02	453.133	13.979	0.024	pMetFam_05-B1-1	(05) Lipids and lipid-like molecules	(B) Fatty Acids	(1) Fatty acyl glycosides	(1) Fatty acyl glycosides of mono- and disaccharides			54
Plantago lanceolata	PL	pathDam	M225T274	225.1483881	273.961	4.566	PL	i-111_Leibniz_positive_mode_20170001_2021-09-02	225.149	4.605	0.042	pMetFam_05-C	(05) Lipids and lipid-like molecules	(C) Prenol lipids					174
Plantago lanceolata	PL	height	M525T763	525.2611342	762.500	12.716	PL	L-2214_lib_spectra_pos_new_2021-09-02	525.268	12.733	0.049	pMetFam_05-C	(05) Lipids and lipid-like molecules	(C) Prenol lipids					359
Plantago lanceolata	PL	pathDam	M525T763	525.2611342	762.500	12.716	PL	L-2214_lib_spectra_pos_new_2021-09-02	525.268	12.733	0.049	pMetFam_05-C	(05) Lipids and lipid-like molecules	(C) Prenol lipids					359
Centaurea jacea	CJ	height	M473T707	473.3621214	706.613	11.777	CJ	M+21900_MoNA-export-LC-MS-HS_Spectra_2021-09-02	473.362	11.674	0.036	pMetFam_05-C	(05) Lipids and lipid-like molecules	(C) Prenol lipids					35
Centaurea jacea	CJ	meshDam	M475T695	475.1931144	695.425	11.590	CJ	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	475.194	11.620	0.045	pMetFam_05-C	(05) Lipids and lipid-like molecules	(C) Prenol lipids					26
Dactylis glomerata	DG	obsShan	M353T053	353.2670428	053.103	14.218	DG	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	353.269	14.220	0.019	pMetFam_05-C	(05) Lipids and lipid-like molecules	(C) Prenol lipids					357
Knautia arvensis	KA	obsShan	M175T340_2	175.1430005	340.393	6.340	KA	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	175.149	6.597	0.045	pMetFam_05-C	(05) Lipids and lipid-like molecules	(C) Prenol lipids					65
Knautia arvensis	KA	BBCH	M175T340_2	175.1430005	340.393	6.340	KA	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	175.149	6.597	0.045	pMetFam_05-C	(05) Lipids and lipid-like molecules	(C) Prenol lipids					65
Phleum pratense	PP	obsRich	M353T053	353.2670428	053.103	14.218	PP	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	353.269	14.220	0.019	pMetFam_05-C	(05) Lipids and lipid-like molecules	(C) Prenol lipids					357
Phleum pratense	PP	obsShan	M353T053	353.2670428	053.103	14.218	PP	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	353.269	14.220	0.019	pMetFam_05-C	(05) Lipids and lipid-like molecules	(C) Prenol lipids					357
Phleum pratense	PP	pathDam	M353T053	353.2670428	053.103	14.218	PP	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	353.269	14.220	0.019	pMetFam_05-C	(05) Lipids and lipid-like molecules	(C) Prenol lipids					357

Species	TwoLetterCode	Factor	Feature	mz	rt_s	rt_m	species	Library - database used	annotated mz	annotated rt_m	p-value	pMetFam_ID	Kingdom	Superclass	Class	Subclass	Parent	count
Dactylis glomerata	DG	season	M177T300	177.0545564	299.919	4.999	DG	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	177.055	5.025	0.037	pMetFam_06	(06) Organoheterocyclic compounds					155
Dactylis glomerata	DG	pathDam	M163T406	163.0392968	405.819	6.764	DG	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	163.039	6.868	0.035	pMetFam_06	(06) Organoheterocyclic compounds					86
Helictotrichon	HL	height	M177T300	177.0545564	299.919	4.999	HL	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	177.055	5.025	0.037	pMetFam_06	(06) Organoheterocyclic compounds					155
Helictotrichon	HL	meshDam	M177T300	177.0545564	299.919	4.999	HL	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	177.055	5.025	0.037	pMetFam_06	(06) Organoheterocyclic compounds					155
Phleum pratense	PP	obsRich	M177T300	177.0545564	299.919	4.999	PP	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	177.055	5.025	0.037	pMetFam_06	(06) Organoheterocyclic compounds					155
Centaurea jacea	CJ	height	M201T722	201.1273111	721.887	12.031	CJ	L-2214_lib_spectra_pos_new_2021-09-02	201.128	12.241	0.004	pMetFam_06-A1-1	(06) Organoheterocyclic compounds	(A) Quinolines and derivatives	(1) Quinolones and derivatives	(-1) Hydroquinolones			27
Centaurea jacea	CJ	meshDam	M201T722	201.1273111	721.887	12.031	CJ	L-2214_lib_spectra_pos_new_2021-09-02	201.128	12.241	0.004	pMetFam_06-A1-1	(06) Organoheterocyclic compounds	(A) Quinolines and derivatives	(1) Quinolones and derivatives	(-1) Hydroquinolones			27
Centaurea jacea	CJ	season	M289T390	289.1639901	390.468	6.508	CJ	L-2214_lib_spectra_pos_new_2021-09-02	289.165	6.736	0.026	pMetFam_06-B	(06) Organoheterocyclic compounds	(B) Tetrahydroisoquinolines					32
Centaurea jacea	CJ	obsShan	M289T390	289.1639901	390.468	6.508	CJ	L-2214_lib_spectra_pos_new_2021-09-02	289.165	6.736	0.026	pMetFam_06-B	(06) Organoheterocyclic compounds	(B) Tetrahydroisoquinolines					32
Centaurea jacea	CJ	pathDam	M289T390	289.1639901	390.468	6.508	CJ	L-2214_lib_spectra_pos_new_2021-09-02	289.165	6.736	0.026	pMetFam_06-B	(06) Organoheterocyclic compounds	(B) Tetrahydroisoquinolines					32
Centaurea jacea	CJ	height	M233T396	233.0810119	396.082	6.601	CJ	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	233.081	6.625	0.044	pMetFam_06-C1-1	(06) Organoheterocyclic compounds	(C) Pyridines and derivatives	(1) Pyridinecarboxylic acids and derivatives	(-1) Pyridinecarboxylic acids			74
Leucanthemum vulgare	LV	meshDam	M351T036	351.1069535	635.904	10.588	LV	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	351.107	10.630	0.031	pMetFam_06-D	(06) Organoheterocyclic compounds	(D) Tetrahydrofurans					98
Plantago lanceolata	PL	season	M329T222	329.1229381	222.210	3.703	PL	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	329.124	3.891	0.026	pMetFam_06-D	(06) Organoheterocyclic compounds	(D) Tetrahydrofurans					42
Plantago lanceolata	PL	obsShan	M329T222	329.1229381	222.210	3.703	PL	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	329.124	3.891	0.026	pMetFam_06-D	(06) Organoheterocyclic compounds	(D) Tetrahydrofurans					42
Helictotrichon	HL	meshDam	M301T726	301.0706777	725.536	12.092	HL	L-2214_lib_spectra_pos_new_2021-09-02	301.071	12.078	0.043	pMetFam_06-E1-1a	(06) Organoheterocyclic compounds	(E) Benzopyrans	(1) 1-benzopyrans				157
Centaurea jacea	CJ	height	M267T366	267.1568901	366.448	6.107	CJ	i-111_Leibniz_positive_mode_20170001_2021-09-02	267.159	6.131	0.035	pMetFam_06-F	(06) Organoheterocyclic compounds	(F) Oxanes					23
Knautia arvensis	KA	obsShan	M455T938	455.3514094	698.223	11.637	KA	L-2214_lib_spectra_pos_new_2021-09-02	455.352	11.781	0.039	pMetFam_06-F	(06) Organoheterocyclic compounds	(F) Oxanes					22
Centaurea jacea	CJ	height	M459T565	459.2212233	564.507	9.408	CJ	M+21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	459.222	9.447	0.050	pMetFam_06-F	(06) Organoheterocyclic compounds	(F) Oxanes					177
Dactylis glomerata	DG	obsShan	M459T565	459.2212233	564.507	9.408	DG	M+21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	459.222	9.447	0.050	pMetFam_06-F	(06) Organoheterocyclic compounds	(F) Oxanes					177
Dactylis glomerata	DG	height	M459T565	459.2212233	564.507	9.408	DG	M+21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	459.222	9.447	0.050	pMetFam_06-F	(06) Organoheterocyclic compounds	(F) Oxanes					177
Arenaria purpurascens	AP	height	M346T188	346.1332296	188.419	3.140	AP	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	346.113	3.173	0.044	pMetFam_06-G1-1	(06) Organoheterocyclic compounds	(G) Diazines	(1) Pyrimidines and pyrimidine derivatives	(-1) Hydropyrimidines			34
Avenula pratensis	AP	height	M348T191	348.1279962	190.811	3.180	AP	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	348.128	3.235	0.044	pMetFam_06-G1-1	(06) Organoheterocyclic compounds	(G) Diazines	(1) Pyrimidines and pyrimidine derivatives	(-1) Hydropyrimidines			35
Leucanthemum vulgare	LV	BBCH	M191T325	191.0702731	325.189	5.420	LV	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	191.071	5.452	0.047	pMetFam_06-G1-1	(06) Organoheterocyclic compounds	(G) Diazines	(1) Pyrimidines and pyrimidine derivatives	(-1) Hydropyrimidines			118
Plantago lanceolata	PL	obsShan	M271T406	271.0964713	405.866	7.764	PL	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	271.097	7.797	0.047	pMetFam_06-G1-1	(06) Organoheterocyclic compounds	(G) Diazines	(1) Pyrimidines and pyrimidine derivatives	(-1) Hydropyrimidines			58
Plantago lanceolata	PL	obsShan	M149T406	149.0599972	405.839	7.764	PL	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	149.060	1.267	0.047	pMetFam_06-G1-1	(06) Organoheterocyclic compounds	(G) Diazines	(1) Pyrimidines and pyrimidine derivatives	(-1) Hydropyrimidines			57
Dactylis glomerata	DG	height	M177T345	177.054535	345.406	5.757	DG	L-2214_lib_spectra_pos_new_2021-09-02	177.055	5.519	0.029	pMetFam_06-H	(06) Organoheterocyclic compounds	(H) Pyrrolidines					59
Geranium pratense	GP	season	M435T259	435.0913661	358.902	5.982	GP	L-2214_lib_spectra_pos_new_2021-09-02	435.091	5.533	0.022	pMetFam_06-H	(06) Organoheterocyclic compounds	(H) Pyrrolidines					20
Geranium pratense	GP	height	M435T259	435.0913661	358.902	5.982	GP	L-2214_lib_spectra_pos_new_2021-09-02	435.091	5.533	0.022	pMetFam_06-H	(06) Organoheterocyclic compounds	(H) Pyrrolidines					20
Geranium pratense	GP	meshDam	M435T259	435.0913661	358.902	5.982	GP	L-2214_lib_spectra_pos_new_2021-09-02	435.091	5.533	0.022	pMetFam_06-H	(06) Organoheterocyclic compounds	(H) Pyrrolidines					20
Centaurea jacea	CJ	meshDam	M177T324	177.0546553	323.627	5.566	KA	L-2214_lib_spectra_pos_new_2021-09-02	177.055	5.519	0.029	pMetFam_06-H	(06) Organoheterocyclic compounds	(H) Pyrrolidines					226
Leucanthemum vulgare	LV	obsShan	M313T325	313.1442894	325.455	13.758	LV	L-2214_lib_spectra_pos_new_2021-09-02	313.144	13.767	0.029	pMetFam_06-H	(06) Organoheterocyclic compounds	(H) Pyrrolidines					35
Leucanthemum vulgare	LV	BBCH	M313T325	313.1442894	325.455	13.758	LV	L-2214_lib_spectra_pos_new_2021-09-02	313.144	13.767	0.029	pMetFam_06-H	(06) Organoheterocyclic compounds	(H) Pyrrolidines					35
Plantago lanceolata	PL	obsRich	M177T345	177.054535	345.406	5.757	PL	L-2214_lib_spectra_pos_new_2021-09-02	177.055	5.519	0.029	pMetFam_06-H	(06) Organoheterocyclic compounds	(H) Pyrrolidines					59
Pedicularis	PA	pathDam	M465T261	465.1016959	262.466	4.374	PA	L-2214_lib_spectra_pos_new_2021-09-02	465.103	4.463	0.039	pMetFam_06-I-1-1	(06) Organoheterocyclic compounds	(I) Imidazopyrimidines	(1) Purines and purine derivatives	(-1) 6-aminopurines			121
Knautioranella	KA	height	M137T88	137.0233804	87.642	1.461	KA	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	137.025	1.475	0.022	pMetFam_06-I-2a	(06) Organoheterocyclic compounds	(I) Imidazopyrimidines	(1) Purines and purine derivatives	(-2) Purinones	(a) 6-oxopurines		210
Knautioranella	KA	season	M197T366	197.1171336	366.162	6.103	KA	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	197.117	6.133	0.044	pMetFam_06-I-2a	(06) Organoheterocyclic compounds	(I) Imidazopyrimidines	(1) Purines and purine derivatives	(-2) Purinones	(a) 6-oxopurines		168
Pedicularis	PA	obsRich	M197T366	197.1171336	366.162	6.103	PA	R-2760_MSMS-Respect-Curated-Pos_2021-09-02	197.117	6.133	0.044	pMetFam_06-I-2a	(06) Organoheterocyclic compounds	(I) Imidazopyrimidines	(1) Purines and purine derivatives	(-2) Purinones	(a) 6-oxopurines		168

Supplement Table 3a: Metabolite families - continued

Species	TwoLetterCode	Factor	Feature	mz	rt_s	rt_m	species	Library - database used	annotated mz	annotated rt_m	p-value	pMetFam_ID	Kingdom	Superclass	Class	Subclass	Parent	count
Centaurea jacea	CJ	pathDam	M225T43	215.1642719	743.499	12.392	CJ	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	215.164	12.403	0.044	pMetFam_07	(07) Organic oxygen compounds					224
Dactylis glomerata	DG	obsShan	M225T43	215.1642719	743.499	12.392	DG	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	215.164	12.403	0.044	pMetFam_07	(07) Organic oxygen compounds					224
Festuca rubra	FR	season	M225T43	215.1642719	743.499	12.392	DG	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	215.164	12.403	0.044	pMetFam_07	(07) Organic oxygen compounds					224
Poa pratensis	PA	season	M195T361	195.1377673	360.574	6.010	FR	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	195.138	6.061	0.044	pMetFam_07	(07) Organic oxygen compounds					98
Poa pratensis	PA	obsShan	M225T43	215.1642719	743.499	12.392	PA	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	215.164	12.403	0.044	pMetFam_07	(07) Organic oxygen compounds					224
Phleum pratense	PP	pathDam	M195T361	195.1377673	360.574	6.010	PP	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	195.138	6.061	0.044	pMetFam_07	(07) Organic oxygen compounds					98
Avenula pubescens	AP	BBCH	M191T431	191.142962	430.596	7.177	AP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	191.143	7.000	0.000	pMetFam_07	(07) Organic oxygen compounds					267
Avenula pubescens	AP	obsShan	M139T203	139.0390938	202.746	3.379	AP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	139.039	2.975	0.000	pMetFam_07	(07) Organic oxygen compounds					103
Avenula pubescens	AP	BBCH	M307T106	307.0809579	105.910	1.765	AP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	307.082	1.451	0.000	pMetFam_07	(07) Organic oxygen compounds					102
Avenula pubescens	AP	season	M353T373	353.2679383	372.673	14.545	AP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	353.268	14.958	0.017	pMetFam_07	(07) Organic oxygen compounds					80
Avenula pubescens	AP	obsShan	M353T373	353.2679383	372.673	14.545	AP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	353.268	14.958	0.017	pMetFam_07	(07) Organic oxygen compounds					80
Koeleria arenicola	KA	BBCH	M209T455	209.1535602	455.362	7.589	KA	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	209.154	7.636	0.000	pMetFam_07	(07) Organic oxygen compounds					90
Koeleria arenicola	KA	meshDam	M155T83	155.0340229	87.823	1.464	KA	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	155.034	2.639	0.001	pMetFam_07	(07) Organic oxygen compounds					146
Phleum pratense	PP	obsRich	M353T373	353.2679383	372.673	14.545	PP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	353.268	14.958	0.017	pMetFam_07	(07) Organic oxygen compounds					80
Avenula pubescens	AP	BBCH	M163T219	163.0393417	219.158	3.653	AP	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	161.060	3.593	0.039	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				323
Avenula pubescens	AP	obsShan	M163T219	163.0393417	219.158	3.653	AP	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	161.060	3.593	0.039	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				323
Dactylis glomerata	DG	obsShan	M209T220	209.1534032	219.501	3.658	DG	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	209.154	3.672	0.012	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				237
Dactylis glomerata	DG	obsShan	M227T317	227.1639345	316.931	5.282	DG	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	227.165	5.319	0.034	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				84
Festuca rubra	FR	meshDam	M503T472	503.1895541	471.719	7.862	FR	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	503.191	7.899	0.046	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				135
Festuca rubra	FR	season	M241T361	241.143377	360.709	6.012	FR	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	241.143	6.059	0.034	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				81
Festuca rubra	FR	BBCH	M241T361	241.143377	360.709	6.012	FR	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	241.143	6.059	0.034	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				81
Festuca rubra	FR	meshDam	M241T361	241.143377	360.709	6.012	FR	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	241.143	6.059	0.034	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				81
Geranium pratense	GP	BBCH	M209T220	209.1534032	219.501	3.658	GP	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	209.154	3.672	0.012	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				237
Geranium pratense	GP	pathDam	M209T220	209.1534032	219.501	3.658	GP	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	209.154	3.672	0.012	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				237
Koeleria arenicola	KA	BBCH	M503T472	503.1895541	471.719	7.862	KA	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	503.191	7.899	0.046	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				135
Koeleria arenicola	KA	height	M211T448	211.1690786	448.053	7.468	KA	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	211.170	7.141	0.046	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				27
Koeleria arenicola	KA	meshDam	M211T448	211.1690786	448.053	7.468	KA	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	211.170	7.141	0.046	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				27
Festuca rubra	FR	season	M200T83	200.0914806	83.303	1.388	FR	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	200.092	1.418	0.001	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				120
Leucanthemum vulgare	LV	meshDam	M163T537	163.0392862	537.243	8.954	LV	i-111_Leibniz_positive_mode_20170801_2021-09-02	163.040	8.986	0.013	pMetFam_07-A	(07) Organic oxygen compounds	(4) Organonitrogen compounds				51
Centaurea jacea	CJ	pathDam	M437T030	437.3400778	808.036	13.467	CJ	i-111_Leibniz_positive_mode_20170801_2021-09-02	437.340	13.498	0.044	pMetFam_07-A1	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(1) Carbonyl compounds			45
Centaurea jacea	CJ	height	M249T366	249.1491049	366.402	6.107	CJ	L-2214_lib_spectra_pos_new_2021-09-02	249.149	6.137	0.019	pMetFam_07-A1	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(1) Carbonyl compounds			28
Koeleria arenicola	KA	obsRich	M455T357	455.3151154	357.333	14.209	KA	L-2214_lib_spectra_pos_new_2021-09-02	455.316	14.310	0.018	pMetFam_07-A1	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(1) Carbonyl compounds			26
Koeleria arenicola	KA	season	M455T395	455.3506297	794.863	13.248	KA	L-2214_lib_spectra_pos_new_2021-09-02	455.352	13.542	0.030	pMetFam_07-A1	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(1) Carbonyl compounds			26
Koeleria arenicola	KA	BBCH	M455T395	455.3506297	794.863	13.248	KA	L-2214_lib_spectra_pos_new_2021-09-02	455.352	13.542	0.030	pMetFam_07-A1	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(1) Carbonyl compounds			26
Ranunculus acris	RA	pathDam	M487T671	487.2526262	670.899	11.182	RA	L-2214_lib_spectra_pos_new_2021-09-02	487.253	11.204	0.050	pMetFam_07-A1	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(1) Carbonyl compounds			23
Ranunculus acris	RA	obsRich	M313T229	313.0916934	226.506	3.808	RA	L-2214_lib_spectra_pos_new_2021-09-02	313.093	3.040	0.022	pMetFam_07-A1.1a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(1) Carbonyl compounds	(1) Ketones	(a) Aryl ketones	21
Ranunculus acris	RA	obsShan	M313T229	313.0916934	226.506	3.808	RA	L-2214_lib_spectra_pos_new_2021-09-02	313.093	3.040	0.022	pMetFam_07-A1.1a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(1) Carbonyl compounds	(1) Ketones	(a) Aryl ketones	21

Species	TwoLetterCode	Factor	Feature	mz	rt_s	rt_m	species	Library - database used	annotated mz	annotated rt_m	p-value	pMetFam_ID	Kingdom	Superclass	Class	Subclass	Parent	count
Geranium pratense	GP	pathDam	M191T265	191.1433096	265.234	4.421	GP	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	191.143	4.468	0.050	pMetFam_07-A2	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(2) Alcohols and polyols			216
Phleum pratense	PP	height	M332T816	332.2789208	816.462	13.600	PP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	332.280	13.637	0.045	pMetFam_07-A2	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(2) Alcohols and polyols			52
Leucanthemum vulgare	LV	obsShan	M247T604_1	247.0934081	604.293	10.072	LV	i-111_Leibniz_positive_mode_20170801_2021-09-02	247.092	10.093	0.030	pMetFam_07-A2.1	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(2) Alcohols and polyols	(1) Polyols		37
Leucanthemum vulgare	LV	pathDam	M247T604_1	247.0934081	604.293	10.072	LV	i-111_Leibniz_positive_mode_20170801_2021-09-02	247.092	10.093	0.030	pMetFam_07-A2.1	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(2) Alcohols and polyols	(1) Polyols		37
Poa pratensis	PA	season	M239T513	239.0814162	513.062	8.551	PA	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	239.082	8.591	0.029	pMetFam_07-A3.1a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(1) Monosaccharides	(a) Pentoses	32
Poa pratensis	PA	meshDam	M239T513	239.0814162	513.062	8.551	PA	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	239.082	8.591	0.029	pMetFam_07-A3.1a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(1) Monosaccharides	(a) Pentoses	32
Anthoxanthum odoratum	AO	BBCH	M369T270	369.1174485	270.093	4.502	AO	i-111_Leibniz_positive_mode_20170801_2021-09-02	369.118	4.515	0.030	pMetFam_07-A3.1a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(1) Monosaccharides	(a) Pentoses	117
Poa pratensis	PA	BBCH	M369T270	369.1174485	270.093	4.502	PA	i-111_Leibniz_positive_mode_20170801_2021-09-02	369.118	4.515	0.030	pMetFam_07-A3.1a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(1) Monosaccharides	(a) Pentoses	117
Centaurea jacea	CJ	obsShan	M245T645	245.1021124	645.495	10.750	CJ	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	245.103	10.799	0.025	pMetFam_07-A3.1a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(1) Monosaccharides	(a) Pentoses	17
Anthoxanthum odoratum	AO	season	M591T298	591.2539612	297.788	4.963	AO	M-21908_MoNA-export-LC-MS-MS_Spectra_2021-09-02	591.255	4.977	0.018	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	21
Anthoxanthum odoratum	AO	height	M591T298	591.2539612	297.788	4.963	AO	M-21908_MoNA-export-LC-MS-MS_Spectra_2021-09-02	591.255	4.977	0.018	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	21
Dactylis glomerata	DG	pathDam	M539T294	539.117534	294.335	4.906	DG	M-21908_MoNA-export-LC-MS-MS_Spectra_2021-09-02	539.119	5.890	0.031	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	12
Festuca rubra	FR	height	M419T431	419.0970745	430.852	7.181	FR	M-21908_MoNA-export-LC-MS-MS_Spectra_2021-09-02	419.097	7.217	0.008	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	44
Festuca rubra	FR	meshDam	M595T370	595.1647157	376.438	6.307	FR	M-21908_MoNA-export-LC-MS-MS_Spectra_2021-09-02	595.165	5.944	0.003	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	116
Poa pratensis	PA	obsShan	M419T431	419.0970745	430.852	7.181	PA	M-21908_MoNA-export-LC-MS-MS_Spectra_2021-09-02	419.097	7.217	0.008	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	44
Poa pratensis	PA	BBCH	M419T431	419.0970745	430.852	7.181	PA	M-21908_MoNA-export-LC-MS-MS_Spectra_2021-09-02	419.097	7.217	0.008	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	44
Poa pratensis	PA	pathDam	M419T431	419.0970745	430.852	7.181	PA	M-21908_MoNA-export-LC-MS-MS_Spectra_2021-09-02	419.097	7.217	0.008	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	44
Poa pratensis	PA	obsShan	M595T370	595.1647157	376.438	6.307	PA	M-21908_MoNA-export-LC-MS-MS_Spectra_2021-09-02	595.165	5.944	0.003	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	116
Poa pratensis	PA	pathDam	M595T370	595.1647157	376.438	6.307	PA	M-21908_MoNA-export-LC-MS-MS_Spectra_2021-09-02	595.165	5.944	0.003	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	116
Festuca rubra	FR	obsRich	M637T574	637.1795417	574.150	9.569	FR	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	637.177	8.757	0.050	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	48
Avenula pubescens	AP	meshDam	M681T517	681.1657454	516.959	8.616	AP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	681.167	8.644	0.050	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	28
Avenula pubescens	AP	obsRich	H303T415	303.0496968	415.318	6.922	AP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	303.050	7.156	0.050	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	67
Geranium pratense	GP	season	H303T415	303.0496968	415.318	6.922	GP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	303.050	7.156	0.050	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	67
Poa pratensis	PA	obsRich	H507T512	507.1124078	512.264	8.538	PA	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	507.112	8.379	0.050	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	78
Poa pratensis	PA	pathDam	H507T512	507.1124078	512.264	8.538	PA	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	507.112	8.379	0.050	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	78
Phleum pratense	PP	BBCH	H465T379	465.1016541	379.324	6.322	PP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	465.103	6.333	0.050	pMetFam_07-A3.2a	(07) Organic oxygen compounds	(4) Organonitrogen compounds	(3) Carbohydrates and carbohydrate conjugates	(2) Glycosyl compounds	(a) C-glycosyl compounds	139

Supplement Table 3a: Metabolite families - continued

Species	TwoLetterCode	Factor	Feature	mz	rt_s	rt_m	species	Library - database used	annotated mz	annotated rt_m	p-value	pMetFam_ID	Kingdom	Superclass	Class	Subclass	Parent	count
Avenula pubescens	AP	obsRich	M263T174	263.1381818	174.082	2.901	AP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	263.138	2.275	0.045	pMetFam_08	(08) Phenylpropanoids and polyketides					52
Festuca rubra	FR	pathDam	M499T194	499.1426252	194.149	3.236	FR	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	499.144	3.263	0.042	pMetFam_08	(08) Phenylpropanoids and polyketides					71
Phleum pratense	PP	pathDam	M499T194	499.1426252	194.149	3.236	PP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	499.144	3.263	0.042	pMetFam_08	(08) Phenylpropanoids and polyketides					71
Geranium pratense	GP	height	M369T742	369.1324414	742.124	12.369	GP	L-2214_lib_spectra_pos_new_2021-09-02	369.132	12.679	0.036	pMetFam_08-A	(08) Phenylpropanoids and polyketides	(A) Stilbenes				7
Centaurea jacea	CJ	season	M161T305	161.0599157	304.517	5.075	CJ	L-2214_lib_spectra_pos_new_2021-09-02	161.060	3.583	0.026	pMetFam_08-B1	(08) Phenylpropanoids and polyketides	(B) Diarylheptanoids	(1) Linear diarylheptanoids			36
Festuca rubra	FR	season	M223T361	223.1327032	360.552	6.009	FR	L-2214_lib_spectra_pos_new_2021-09-02	223.133	6.059	0.037	pMetFam_08-B1	(08) Phenylpropanoids and polyketides	(B) Diarylheptanoids	(1) Linear diarylheptanoids			103
Leucanthemum vulgare	LV	pathDam	M341T379	341.1362548	379.036	6.317	LV	L-2214_lib_spectra_pos_new_2021-09-02	341.136	6.350	0.033	pMetFam_08-C1	(08) Phenylpropanoids and polyketides	(C) Tannins	(1) Hydrolyzable tannins			192
Plantago lanceolata	PL	height	M381T476	381.1173599	475.541	7.926	PL	L-2214_lib_spectra_pos_new_2021-09-02	381.117	7.962	0.024	pMetFam_08-D	(08) Phenylpropanoids and polyketides	(D) Coumarins and derivatives				39
Plantago lanceolata	PL	BBCH	M381T476	381.1173599	475.541	7.926	PL	L-2214_lib_spectra_pos_new_2021-09-02	381.117	7.962	0.024	pMetFam_08-D	(08) Phenylpropanoids and polyketides	(D) Coumarins and derivatives				39
Knautia arvensis	KA	height	M463T161	463.1798641	160.975	2.683	KA	L-2214_lib_spectra_pos_new_2021-09-02	463.180	2.698	0.027	pMetFam_08-D1.1	(08) Phenylpropanoids and polyketides	(D) Coumarins and derivatives	(1) Pyranocoumarins	(.1) Angular pyranocoumarins		71
Anthoxanthum odoratum	AO	BBCH	M207T291	207.0654284	291.373	4.856	AO	L-2214_lib_spectra_pos_new_2021-09-02	207.065	4.886	0.011	pMetFam_08-D2.1	(08) Phenylpropanoids and polyketides	(D) Coumarins and derivatives	(2) Hydroxycoumarins	(.1) 7-hydroxycoumarins		108
Ranunculus acris	RA	BBCH	M207T634	207.0651206	633.799	10.563	RA	L-2214_lib_spectra_pos_new_2021-09-02	207.065	10.601	0.008	pMetFam_08-D2.1	(08) Phenylpropanoids and polyketides	(D) Coumarins and derivatives	(2) Hydroxycoumarins	(.1) 7-hydroxycoumarins		27
Centaurea jacea	CJ	obsRich	M347T500	347.0757352	500.193	8.337	CJ	L-2214_lib_spectra_pos_new_2021-09-02	347.077	8.025	0.045	pMetFam_08-E1	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(1) O-methylated isoflavonoids			24
Centaurea jacea	CJ	height	M347T500	347.0757352	500.193	8.337	CJ	L-2214_lib_spectra_pos_new_2021-09-02	347.077	8.025	0.045	pMetFam_08-E1	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(1) O-methylated isoflavonoids			24
Phleum pratense	PP	height	M273T170	273.0755207	169.788	2.830	PP	L-2214_lib_spectra_pos_new_2021-09-02	273.076	3.029	0.049	pMetFam_08-E2.1	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(2) Isoflavans	(.1) Isoflavanols		80
Geranium pratense	GP	BBCH	M181T758	181.1222403	758.152	12.636	GP	L-2214_lib_spectra_pos_new_2021-09-02	181.122	12.652	0.030	pMetFam_08-E2.2a	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(2) Isoflavans	(.2) Isoflavanones	(a) 6-prenylated isoflavanones	415
Poa pratensis	PA	obsRich	M175T295	175.1480374	295.339	4.922	PA	L-2214_lib_spectra_pos_new_2021-09-02	175.148	4.384	0.030	pMetFam_08-E2.2a	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(2) Isoflavans	(.2) Isoflavanones	(a) 6-prenylated isoflavanones	115
Geranium pratense	GP	obsShan	M181T854	181.1222311	854.346	14.239	GP	L-2214_lib_spectra_pos_new_2021-09-02	181.122	14.258	0.048	pMetFam_08-E3	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(3) Furanoisoflavonoids			496
Anthoxanthum odoratum	AO	obsShan	M373T276	373.1127427	276.406	4.607	AO	L-2214_lib_spectra_pos_new_2021-09-02	373.113	4.642	0.047	pMetFam_08-E3.1	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(3) Furanoisoflavonoids	(.1) Pterocarpans		184
Dactylis glomerata	DG	mechDam	M373T276	373.1127427	276.406	4.607	DG	L-2214_lib_spectra_pos_new_2021-09-02	373.113	4.642	0.047	pMetFam_08-E3.1	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(3) Furanoisoflavonoids	(.1) Pterocarpans		184
Dactylis glomerata	DG	season	M339T299	339.1068074	299.274	4.988	DG	L-2214_lib_spectra_pos_new_2021-09-02	339.107	5.011	0.050	pMetFam_08-E3.1	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(3) Furanoisoflavonoids	(.1) Pterocarpans		304
Dactylis glomerata	DG	mechDam	M339T299	339.1068074	299.274	4.988	DG	L-2214_lib_spectra_pos_new_2021-09-02	339.107	5.011	0.050	pMetFam_08-E3.1	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(3) Furanoisoflavonoids	(.1) Pterocarpans		304
Poa pratensis	PA	obsRich	M339T299	339.1068074	299.274	4.988	PA	L-2214_lib_spectra_pos_new_2021-09-02	339.107	5.011	0.050	pMetFam_08-E3.1	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(3) Furanoisoflavonoids	(.1) Pterocarpans		304
Phleum pratense	PP	height	M291T175	291.0861592	174.684	2.911	PP	L-2214_lib_spectra_pos_new_2021-09-02	291.086	2.989	0.033	pMetFam_08-E3.1	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(3) Furanoisoflavonoids	(.1) Pterocarpans		84
Phleum pratense	PP	height	M291T203	291.0861657	203.273	3.388	PP	L-2214_lib_spectra_pos_new_2021-09-02	291.086	3.413	0.026	pMetFam_08-E3.1	(08) Phenylpropanoids and polyketides	(E) Isoflavonoids	(3) Furanoisoflavonoids	(.1) Pterocarpans		141
Centaurea jacea	CJ	mechDam	M261T695	261.1483944	695.297	11.588	CJ	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	261.148	11.610	0.048	pMetFam_08-F	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives				21
Centaurea jacea	CJ	height	M261T695	261.1483944	695.297	11.588	CJ	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	261.148	11.610	0.048	pMetFam_08-F	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives				21
Plantago lanceolata	PL	obsShan	M433T466	433.1490664	465.840	7.764	PL	L-2214_lib_spectra_pos_new_2021-09-02	433.149	7.797	0.035	pMetFam_08-F1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(1) Cinnamic acid esters			52

Species	TwoLetterCode	Factor	Feature	mz	rt_s	rt_m	species	Library - database used	annotated mz	annotated rt_m	p-value	pMetFam_ID	Kingdom	Superclass	Class	Subclass	Parent	count
Dactylis glomerata	DG	obsShan	M341T382	341.0650215	382.031	6.367	DG	L-2214_lib_spectra_pos_new_2021-09-02	341.066	6.860	0.042	pMetFam_08-F2	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives			20
Holcus lanatus	HL	season	M207T270	207.0653517	270.213	4.504	HL	M-21908_MoNA-export-LC-MS-MS_Spectra_2021-09-02	207.065	4.886	0.044	pMetFam_08-F2	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives			135
Holcus lanatus	HL	pathDam	M207T270	207.0653517	270.213	4.504	HL	M-21908_MoNA-export-LC-MS-MS_Spectra_2021-09-02	207.065	4.886	0.044	pMetFam_08-F2	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives			135
Poa pratensis	PA	BBCH	M207T270	207.0653517	270.213	4.504	PA	M-21908_MoNA-export-LC-MS-MS_Spectra_2021-09-02	207.065	4.886	0.044	pMetFam_08-F2	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives			135
Leucanthemum vulgare	LV	obsRich	M163T183	163.0394705	182.529	3.042	LV	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	163.040	3.178	0.046	pMetFam_08-F2	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives			233
Poa pratensis	PA	BBCH	M207T249	207.0651231	248.703	4.145	PA	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	207.065	4.049	0.037	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		181
Poa pratensis	PA	pathDam	M207T249	207.0651231	248.703	4.145	PA	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	207.065	4.049	0.037	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		181
Dactylis glomerata	DG	mechDam	M515T441	515.1535289	441.168	7.353	DG	L-2214_lib_spectra_pos_new_2021-09-02	515.154	7.401	0.026	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		39
Dactylis glomerata	DG	pathDam	M515T441	515.1535289	441.168	7.353	DG	L-2214_lib_spectra_pos_new_2021-09-02	515.154	7.401	0.026	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		39
Festuca rubra	FR	pathDam	M369T287	369.117543	286.546	4.776	FR	L-2214_lib_spectra_pos_new_2021-09-02	369.119	4.814	0.037	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		189
Festuca rubra	FR	obsShan	M207T398	207.0652593	397.657	6.628	FR	M-21908_MoNA-export-LC-MS-MS_Spectra_2021-09-02	207.065	6.306	0.039	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		44
Plantago lanceolata	PL	pathDam	M147T189	147.0443204	189.349	3.156	PL	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	147.044	3.998	0.039	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		309
Avenula pubescens	AP	obsRich	M251T133	251.1388214	132.532	2.209	AP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	251.140	2.272	0.021	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		52
Geranium pratense	GP	BBCH	M233T375	233.081045	375.232	6.254	GP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	233.081	6.281	0.027	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		119
Knautia arvensis	KA	BBCH	M163T145	163.0392152	144.551	2.409	KA	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	163.040	2.442	0.007	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		386
Poa pratensis	PA	mechDam	M233T375	233.081045	375.232	6.254	PA	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	233.081	6.281	0.027	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		119
Phleum pratense	PP	pathDam	M147T239	147.0443784	238.730	3.979	PP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	147.044	3.998	0.039	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		142
Phleum pratense	PP	pathDam	M147T520	147.0441969	519.882	8.665	PP	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	147.045	8.215	0.016	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		53
Ranunculus acris	RA	obsShan	M137T113	137.0599368	113.322	1.889	RA	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	137.060	2.493	0.014	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		91
Ranunculus acris	RA	obsRich	M207T260	207.0652892	259.814	4.330	RA	R-2768_MSMS-Respect-Curated-Pos_2021-09-02	207.065	4.886	0.007	pMetFam_08-F2.1	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.1) Coumaric acids and derivatives		64
Holcus lanatus	HL	mechDam	M307T451	307.1019816	450.942	7.516	HL	L-2214_lib_spectra_pos_new_2021-09-02	307.102	8.001	0.028	pMetFam_08-F2.2	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters		26

Species	TwoLetterCode	Factor	Feature	mz	rt_s	rt_m	species	Library - database used	annotated mz	annotated rt_m	p-value	pMetFam_ID	Kingdom	Superclass	Class	Subclass	Parent	count
Anthoxanthum odoratum	AO	season	M275T288	275.1500777	288.180	4.803	AO	L-2214_lib_spectra_pos_new_2021-09-02	275.151	4.836	0.022	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	23
Leucanthemum vulgare	LV	obsShan	M351T604	351.1068321	604.117	10.069	LV	L-2214_lib_spectra_pos_new_2021-09-02	351.107	10.095	0.009	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	42
Leucanthemum vulgare	LV	height	M351T604	351.1068321	604.117	10.069	LV	L-2214_lib_spectra_pos_new_2021-09-02	351.107	10.095	0.009	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	42
Leucanthemum vulgare	LV	pathDam	M351T604	351.1068321	604.117	10.069	LV	L-2214_lib_spectra_pos_new_2021-09-02	351.107	10.095	0.009	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	42
Avenula pubescens	AP	season	M370T156	370.1099301	155.697	2.595	AP	L-2214_lib_spectra_pos_new_2021-09-02	370.110	2.626	0.016	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	30
Avenula pubescens	AP	height	M370T156	370.1099301	155.697	2.595	AP	L-2214_lib_spectra_pos_new_2021-09-02	370.110	2.626	0.016	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	30
Avenula pubescens	AP	season	M370T191	370.1097726	191.191	3.187	AP	L-2214_lib_spectra_pos_new_2021-09-02	370.110	3.202	0.018	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	27
Avenula pubescens	AP	height	M370T191	370.1097726	191.191	3.187	AP	L-2214_lib_spectra_pos_new_2021-09-02	370.110	3.202	0.018	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	27
Centaurea jacea	CJ	season	M341T269	341.137102	269.286	4.488	CJ	L-2214_lib_spectra_pos_new_2021-09-02	341.138	4.260	0.007	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	83
Dactylis glomerata	DG	mechDam	M361T435	361.0911591	434.921	7.249	DG	L-2214_lib_spectra_pos_new_2021-09-02	361.091	7.876	0.024	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	22
Festuca rubra	FR	BBCH	M163T387	163.0594029	386.689	6.445	FR	L-2214_lib_spectra_pos_new_2021-09-02	163.060	9.571	0.024	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	32
Festuca rubra	FR	season	M177T361_2	177.1274201	360.616	6.010	FR	L-2214_lib_spectra_pos_new_2021-09-02	177.127	6.070	0.015	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	55
Festuca rubra	FR	obsRich	M493T300	493.1329542	299.696	4.995	FR	L-2214_lib_spectra_pos_new_2021-09-02	493.134	5.020	0.014	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	49
Festuca rubra	FR	obsShan	M339T220	339.1062572	220.081	3.668	FR	L-2214_lib_spectra_pos_new_2021-09-02	339.106	3.531	0.013	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	118
Geranium pratense	GP	BBCH	M425T238	425.1410326	237.704	3.962	GP	L-2214_lib_spectra_pos_new_2021-09-02	427.196	9.332	0.019	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	209
Holcus lanatus	HL	obsShan	M323T598	323.1387758	598.066	9.968	HL	L-2214_lib_spectra_pos_new_2021-09-02	323.140	9.980	0.022	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	21
Knautia arvensis	KA	obsRich	M341T269	341.137102	269.286	4.488	KA	L-2214_lib_spectra_pos_new_2021-09-02	341.138	4.260	0.007	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	83
Leucanthemum vulgare	LV	BBCH	M183T338	183.0654658	337.916	5.632	LV	L-2214_lib_spectra_pos_new_2021-09-02	183.065	5.930	0.024	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	35
Leucanthemum vulgare	LV	mechDam	M147T162	147.0442974	161.641	2.694	LV	L-2214_lib_spectra_pos_new_2021-09-02	147.044	2.858	0.009	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	55
Leucanthemum vulgare	LV	season	M351T604	351.1068321	604.117	10.069	LV	L-2214_lib_spectra_pos_new_2021-09-02	351.107	10.095	0.009	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	42
Leucanthemum vulgare	LV	pathDam	M351T667	351.1068941	667.489	11.125	LV	L-2214_lib_spectra_pos_new_2021-09-02	351.108	11.884	0.019	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	40
Poa pratensis	PA	pathDam	M227T262	227.1641291	262.432	4.374	PA	L-2214_lib_spectra_pos_new_2021-09-02	227.164	4.401	0.024	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	411
Poa pratensis	PA	pathDam	M321T337	321.2168411	336.819	5.614	PA	L-2214_lib_spectra_pos_new_2021-09-02	321.218	5.665	0.019	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	31
Phleum pratense	PP	pathDam	M177T361_2	177.1274201	360.616	6.010	PP	L-2214_lib_spectra_pos_new_2021-09-02	177.127	6.070	0.015	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	55
Phleum pratense	PP	pathDam	M147T225	147.0443427	224.810	3.747	PP	L-2214_lib_spectra_pos_new_2021-09-02	147.044	3.930	0.008	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	234
Ranunculus acris	RA	obsRich	M147T324	147.0443179	324.379	5.406	RA	L-2214_lib_spectra_pos_new_2021-09-02	147.044	5.498	0.009	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	74
Ranunculus acris	RA	mechDam	M147T324	147.0443179	324.379	5.406	RA	L-2214_lib_spectra_pos_new_2021-09-02	147.044	5.498	0.009	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	74
Ranunculus acris	RA	pathDam	M147T324	147.0443179	324.379	5.406	RA	L-2214_lib_spectra_pos_new_2021-09-02	147.044	5.498	0.009	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	74
Ranunculus acris	RA	obsShan	M309T419	309.0967691	418.708	6.978	RA	L-2214_lib_spectra_pos_new_2021-09-02	309.097	6.476	0.006	pMetFam_08-F2.2a	(08) Phenylpropanoids and polyketides	(F) Cinnamic acids and derivatives	(2) Hydroxycinnamic acids and derivatives	(.2) Hydroxycinnamic acid esters	(a) Coumaric acid esters	112

Supplement Table 3a: Metabolite families – continued

Species	TwoLetterCode	Factor	Feature	mz	rt_s	rt_m	species	Library - database used	annotated mz	annotated rt_m	p-value	pMetFam_ID	Kingdom	Superclass	Class	Subclass	Parent	count
Dactylis glomerata	DG	obsShan	M611T260	611.15062	260.165	4.336	DG	R-2760_MSMS-Respect-curated-Pos_2021-09-02	611.161	13.371	0.041	pMetFam_06-G	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 6-O-methylated flavonoids			221
Festuca rubra	FR	obsShan	M325T419	325.0920691	418.812	6.980	FR	R-2760_MSMS-Respect-curated-Pos_2021-09-02	325.093	12.373	0.048	pMetFam_06-G1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 6-O-methylated flavonoids			32
Festuca rubra	FR	height	M369T311	369.1175323	311.475	5.191	FR	R-2760_MSMS-Respect-curated-Pos_2021-09-02	369.118	12.373	0.028	pMetFam_06-G12	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) 3-O-methylated flavonoids			67
Poa pratensis	PA	obsRich	M369T311	369.1175323	311.475	5.191	PA	R-2760_MSMS-Respect-curated-Pos_2021-09-02	369.118	12.373	0.028	pMetFam_06-G12	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) 3-O-methylated flavonoids			67
Festuca rubra	FR	pathDam	M109T258	109.1273272	257.804	4.286	FR	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	109.128	4.322	0.029	pMetFam_06-G13	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 7-O-methylated flavonoids			202
Plikweria primrose	PP	BBCH	M109T258	109.1273272	257.804	4.286	PP	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	109.128	4.322	0.029	pMetFam_06-G13	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 7-O-methylated flavonoids			202
Festuca rubra	FR	season	M625T410_1	625.1754202	410.015	6.967	FR	L-2214_lib_spectra_pos_new_2021-09-02	625.176	7.212	0.033	pMetFam_06-G13	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 7-O-methylated flavonoids			13
Anthoxanthum odoratum	AO	pathDam	M579T457	579.1700155	457.215	7.620	AO	L-2214_lib_spectra_pos_new_2021-09-02	579.170	7.905	0.048	pMetFam_06-G14	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 0-methylated flavonoids			126
Anthoxanthum odoratum	AO	height	M282T285	282.1591756	285.075	4.751	AO	L-2214_lib_spectra_pos_new_2021-09-02	282.159	4.819	0.050	pMetFam_06-G14	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 0-methylated flavonoids			21
Avenula pubescens	AP	obsShan	M137T157	137.0590533	157.411	2.624	AP	L-2214_lib_spectra_pos_new_2021-09-02	137.060	2.779	0.017	pMetFam_06-G14	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 0-methylated flavonoids			257
Avenula pubescens	AP	season	M333T349	333.1229394	349.152	5.819	AP	L-2214_lib_spectra_pos_new_2021-09-02	333.123	5.858	0.017	pMetFam_06-G14	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 0-methylated flavonoids			9
Avenula pubescens	AP	height	M333T349	333.1229394	349.152	5.819	AP	L-2214_lib_spectra_pos_new_2021-09-02	333.123	5.858	0.017	pMetFam_06-G14	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 0-methylated flavonoids			9
Centaurea jacea	CJ	obsShan	M331T745	331.0809405	744.947	12.414	CJ	L-2214_lib_spectra_pos_new_2021-09-02	331.081	12.408	0.006	pMetFam_06-G14	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 0-methylated flavonoids			43
Festuca rubra	FR	obsShan	M281T770	281.1020946	776.436	12.974	FR	L-2214_lib_spectra_pos_new_2021-09-02	281.103	12.991	0.017	pMetFam_06-G14	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 0-methylated flavonoids			13
Geranium pratense	GP	pathDam	M133T467	133.1014792	466.934	7.775	GP	L-2214_lib_spectra_pos_new_2021-09-02	133.102	7.803	0.017	pMetFam_06-G14	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 0-methylated flavonoids			75
Leucanthemum vulgare	LV	BBCH	M579T457	579.1700155	457.215	7.620	LV	L-2214_lib_spectra_pos_new_2021-09-02	579.170	7.905	0.048	pMetFam_06-G14	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 0-methylated flavonoids			126
Poa pratensis	PA	obsRich	M137T157	137.0590533	157.411	2.624	PA	L-2214_lib_spectra_pos_new_2021-09-02	137.060	2.779	0.017	pMetFam_06-G14	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 0-methylated flavonoids			257
Poa pratensis	PA	season	M217T042	217.1501447	042.361	14.099	PA	L-2214_lib_spectra_pos_new_2021-09-02	217.159	14.043	0.017	pMetFam_06-G14	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 0-methylated flavonoids			29
Poa pratensis	PA	meshDam	M217T042	217.1501447	042.361	14.099	PA	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	217.159	14.043	0.017	pMetFam_06-G14	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(1) 0-methylated flavonoids			29
Poa pratensis	PA	BBCH	M559T270	559.1552594	270.124	4.502	PA	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	559.155	4.523	0.037	pMetFam_06-G2	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides			91

Species	TwoLetterCode	Factor	Feature	mz	rt_s	rt_m	species	Library - database used	annotated mz	annotated rt_m	p-value	pMetFam_ID	Kingdom	Superclass	Class	Subclass	Parent	count
Geranium pratense	GP	pathDam	M467T142	467.0300177	142.101	2.368	GP	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	467.062	2.849	0.018	pMetFam_06-G2.1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides		15
Poa pratensis	PA	pathDam	M449T474	449.1065634	474.274	7.905	PA	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	449.107	7.941	0.023	pMetFam_06-G2.1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides		171
Plikweria primrose	PP	obsRich	M449T474	449.1065634	474.274	7.905	PP	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	449.107	7.941	0.023	pMetFam_06-G2.1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides		171
Leucanthemum vulgare	LV	BBCH	M449T474	449.1065634	474.274	7.905	LV	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	449.107	7.941	0.023	pMetFam_06-G2.1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides		171
Knautia arvensis	KA	height	M449T554	449.106972	553.885	9.231	KA	L-2214_lib_spectra_pos_new_2021-09-02	449.107	9.271	0.038	pMetFam_06-G2.1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides		69
Leucanthemum vulgare	LV	BBCH	M387T274	387.2002741	273.972	4.566	LV	M-21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	387.200	4.633	0.050	pMetFam_06-G2.1.9	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(a) Anthocyanins	136
Plantago lanceolata	PL	pathDam	M387T274	387.2002741	273.972	4.566	PL	M-21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	387.200	4.633	0.050	pMetFam_06-G2.1.9	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(a) Anthocyanins	136
Geranium pratense	GP	season	M495T445	495.0969912	445.334	7.423	GP	L-2214_lib_spectra_pos_new_2021-09-02	495.091	6.895	0.042	pMetFam_06-G2.1.8	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(b) Flavonoid O-glucuronides	24
Leucanthemum vulgare	LV	BBCH	M519T594	519.1125512	594.242	9.904	LV	L-2214_lib_spectra_pos_new_2021-09-02	519.114	9.945	0.042	pMetFam_06-G2.1.8	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(b) Flavonoid O-glucuronides	39
Plikweria primrose	PP	meshDam	M447T685	447.1270578	685.026	11.417	PP	L-2214_lib_spectra_pos_new_2021-09-02	447.126	11.446	0.016	pMetFam_06-G2.1.8	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(b) Flavonoid O-glucuronides	89
Anthoxanthum odoratum	AO	season	M448T103	448.1271116	103.267	3.054	AO	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	448.126	3.088	0.040	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	15
Anthoxanthum odoratum	AO	season	M295T593	295.1022493	592.694	9.212	AO	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	295.103	9.686	0.042	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	75
Avenula pubescens	AP	meshDam	M262T033	262.1202672	33.407	1.390	AP	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	262.120	1.415	0.042	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	22
Anthoxanthum odoratum	AO	height	M262T033	262.1202672	33.407	1.390	AO	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	262.120	1.415	0.042	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	42
Avenula pubescens	AP	BBCH	M448T103	448.1271116	103.267	3.054	AO	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	448.126	3.088	0.040	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	15
Avenula pubescens	AP	pathDam	M461T573	461.1068196	573.185	9.553	AP	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	461.107	9.553	0.035	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	20
Dactylis glomerata	DG	season	M339T532	339.1063433	532.144	8.869	DG	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	339.107	8.905	0.037	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	35
Festuca rubra	FR	pathDam	M495T773	495.1270503	772.707	12.875	FR	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	495.129	12.892	0.013	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	83
Leucanthemum vulgare	LV	BBCH	M463T665	463.1224315	664.713	11.079	LV	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	463.122	11.091	0.012	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	245
Leucanthemum vulgare	LV	BBCH	M611T301	611.1596254	301.345	5.022	LV	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	611.160	5.576	0.013	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	37
Poa pratensis	PA	BBCH	M345T338	345.1176637	337.946	5.632	LV	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	345.119	5.917	0.037	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	37
Poa pratensis	PA	obsRich	M495T773	495.1270503	772.707	12.875	PA	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	495.129	12.892	0.013	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	83
Plantago lanceolata	PL	meshDam	M309T491	309.0967664	490.590	8.177	PL	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	309.098	7.945	0.037	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	90
Plantago lanceolata	PL	meshDam	M325T449	325.0917507	448.421	7.480	PL	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	325.092	7.542	0.037	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	37
Plikweria primrose	PP	meshDam	M309T491	309.0967664	490.590	8.177	PP	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	309.098	7.945	0.037	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	90
Leucanthemum vulgare	LV	BBCH	M449T432	449.1069703	432.106	7.202	LV	L-2214_lib_spectra_pos_new_2021-09-02	449.108	7.238	0.043	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	65
Avenula pubescens	AP	obsShan	M611T370	611.1596341	377.930	6.299	AP	M-21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	611.161	6.374	0.040	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	133
Avenula pubescens	AP	pathDam	M491T573	491.1174118	572.714	9.545	AP	M-21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	491.117	9.583	0.034	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	47
Avenula pubescens	AP	BBCH	M653T485	653.1705556	485.014	8.004	AP	M-21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	653.171	7.517	0.031	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	13
Holcus lanatus	HL	season	M449T398	449.1441108	397.966	6.633	HL	M-21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	449.144	6.684	0.035	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	17
Leucanthemum vulgare	LV	BBCH	M551T402	551.1495544	402.805	6.705	LV	M-21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	551.150	6.757	0.035	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	108
Poa pratensis	PA	pathDam	M611T370	611.1596341	377.930	6.299	PA	M-21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	611.161	6.374	0.040	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	133
Geranium pratense	GP	season	M465T519	465.1026946	518.612	8.644	GP	R-2760_MSMS-Respect-curated-Pos_2021-09-02	465.102	8.677	0.035	pMetFam_06-G2.1.6	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(2) Flavonoid glycosides	(.1) Flavonoid O-glycosides	(c) Flavonoid-3-O-glycosides	33

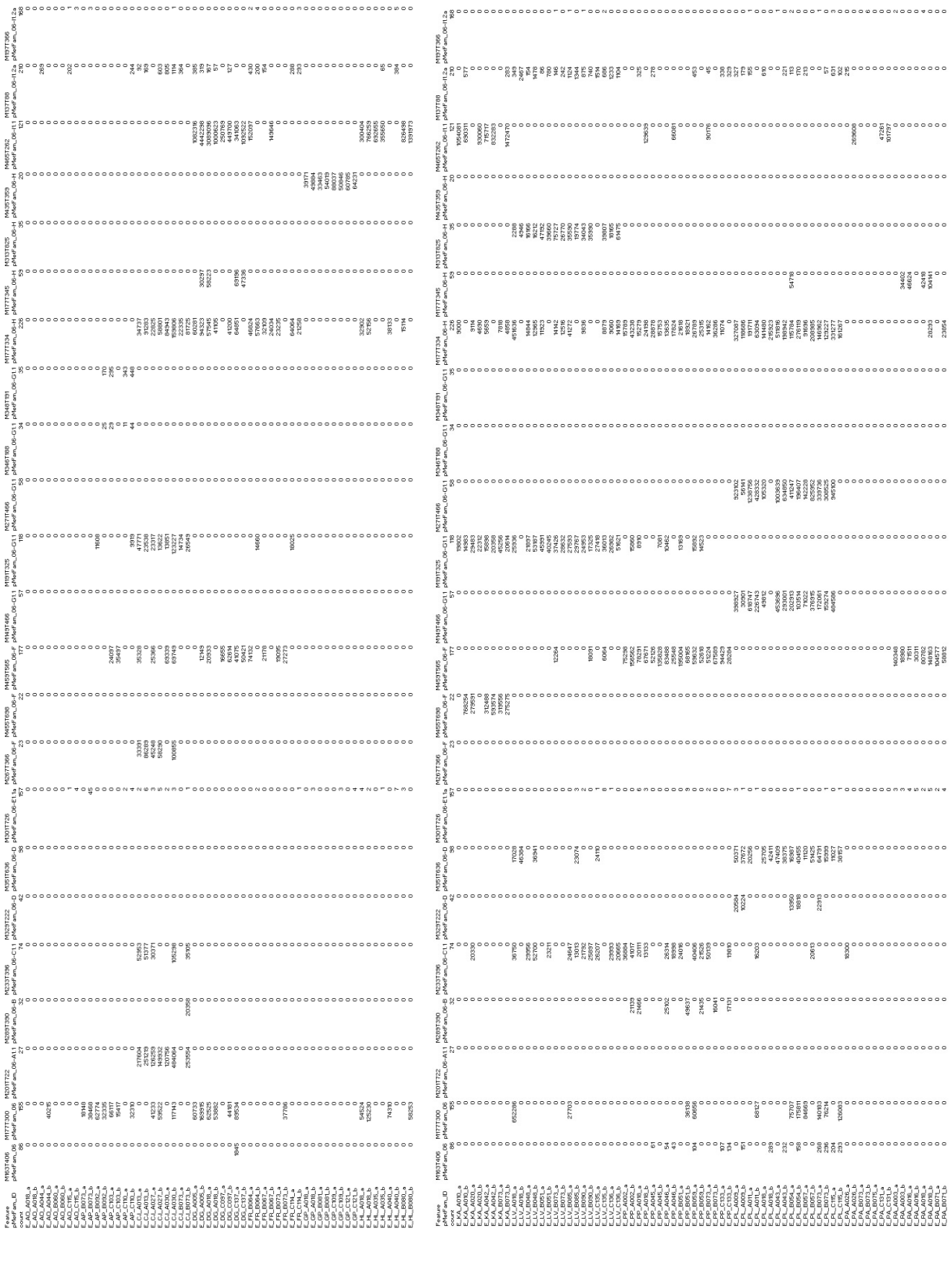
Species	TwoLetterCode	Factor	Feature	mz	rt_s	rt_m	species	Library - database used	annotated mz	annotated rt_m	p-value	pMetFam_ID	Kingdom	Superclass	Class	Subclass	Parent	count
Avenula pubescens	AP	meshDam	M477T445_1	477.1021371	445.260	7.421	AP	R-2760_MSMS-Respect-curated-Pos_2021-09-02	477.102	6.907	0.048	pMetFam_06-G3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids			29
Avenula pubescens	AP	pathDam	M477T445_1	477.1021371	445.260	7.421	AP	R-2760_MSMS-Respect-curated-Pos_2021-09-02	477.102	6.907	0.048	pMetFam_06-G3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids			29
Festuca rubra	FR	obsShan	M325T419	325.0920691	418.812	6.980	FR	R-2760_MSMS-Respect-curated-Pos_2021-09-02	325.093	7.019	0.021	pMetFam_06-G3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids			141
Holcus lanatus	HL	season	M609T665	609.1804949	664.640	11.077	HL	R-2760_MSMS-Respect-curated-Pos_2021-09-02	609.180	10.411	0.043	pMetFam_06-G3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids			16
Holcus lanatus	HL	BBCH	M609T665	609.1804949	664.640	11.077	HL	R-2760_MSMS-Respect-curated-Pos_2021-09-02	609.180	10.411	0.043	pMetFam_06-G3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids			16
Holcus lanatus	HL	season	M289T260	289.07071	259.963	4.333	HL	R-2760_MSMS-Respect-curated-Pos_2021-09-02	289.072	4.399	0.036	pMetFam_06-G3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids			20
Holcus lanatus	HL	height	M339T247	339.1066095	246.706	4.112	HL	R-2760_MSMS-Respect-curated-Pos_2021-09-02	339.107	4.131	0.021	pMetFam_06-G3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids			163
Holcus lanatus	HL	meshDam	M339T247	339.1066095	246.706	4.112	HL	R-2760_MSMS-Respect-curated-Pos_2021-09-02	339.107	4.131	0.021	pMetFam_06-G3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids			163
Knautia arvensis	KA	height	M325T419	325.0920691	418.812	6.980	KA	R-2760_MSMS-Respect-curated-Pos_2021-09-02	325.093	7.019	0.021	pMetFam_06-G3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids			141
Plantago lanceolata	PL	obsRich	M325T558	325.0920725	537.616	8.960	PL	R-2760_MSMS-Respect-curated-Pos_2021-09-02	325.093	8.496	0.021	pMetFam_06-G3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids			68
Centaurea jacea	CJ	obsShan	M337T451	337.0916393	450.904	7.515	CJ	L-2214_lib_spectra_pos_new_2021-09-02	337.093	7.560	0.038	pMetFam_06-G3.1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.1) 5-hydroxyflavonoids		141
Centaurea jacea	CJ	pathDam	M337T451	337.0916393	450.904	7.515	CJ	L-2214_lib_spectra_pos_new_2021-09-02	337.093	7.560	0.038	pMetFam_06-G3.1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.1) 5-hydroxyflavonoids		141
Festuca rubra	FR	season	M463T397	463.0864342	396.702	6.612	FR	L-2214_lib_spectra_pos_new_2021-09-02	463.087	6.650	0.013	pMetFam_06-G3.1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.1) 5-hydroxyflavonoids		138
Festuca rubra	FR	BBCH	M463T397	463.0864342	396.702	6.612	FR	L-2214_lib_spectra_pos_new_2021-09-02	463.087	6.650	0.013	pMetFam_06-G3.1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.1) 5-hydroxyflavonoids		138
Geranium pratense	GP	obsShan	M449T230	449.1070958	230.229	3.837	GP	L-2214_lib_spectra_pos_new_2021-09-02	449.107	3.568	0.043	pMetFam_06-G3.1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.1) 5-hydroxyflavonoids		156
Geranium pratense	GP	BBCH	M449T230	449.1070958	230.229	3.837	GP	L-2214_lib_spectra_pos_new_2021-09-02	449.107	3.568	0.043	pMetFam_06-G3.1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.1) 5-hydroxyflavonoids		156
Knautia arvensis	KA	height	M449T230	449.1070958	230.229	3.837	KA	L-2214_lib_spectra_pos_new_2021-09-02	449.107	3.568	0.043	pMetFam_06-G3.1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.1) 5-hydroxyflavonoids		156
Poa pratensis	PA	meshDam	M449T230	449.1070958	230.229	3.837	PA	L-2214_lib_spectra_pos_new_2021-09-02	449.107	3.568	0.043	pMetFam_06-G3.1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.1) 5-hydroxyflavonoids		156
Plantago lanceolata	PL	height	M457T430	457.1339054	429.706	7.162	PL	R-2760_MSMS-Respect-curated-Pos_2021-09-02	457.133	7.104	0.032	pMetFam_06-G3.1	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.1) 5-hydroxyflavonoids		35
Avenula pubescens	AP	pathDam	M449T401_1	449.1070416	400.885	6.681	PA	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	449.107	6.064	0.041	pMetFam_06-G3.2	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.2) 4'-hydroxyflavonoids		292
Knautia arvensis	KA	BBCH	M449T401_1	449.1070416	400.885	6.681	KA	L-2214_lib_spectra_pos_new_2021-09-02	449.107	6.064	0.041	pMetFam_06-G3.2	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.2) 4'-hydroxyflavonoids		292
Leucanthemum vulgare	LV	BBCH	M595T390	595.1645907	390.237	6.504	LV	M-21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	595.165	6.946	0.048	pMetFam_06-G3.2	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.2) 4'-hydroxyflavonoids		169
Leucanthemum vulgare	LV	obsShan	M289T493	289.0915029	492.861	8.214	LV	R-2760_MSMS-Respect-curated-Pos_2021-09-02	289.092	8.232	0.041	pMetFam_06-G3.2	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.2) 4'-hydroxyflavonoids		44
Avenula pubescens	AP	height	M565T321	565.1545962	320.788	5.346	AP	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	565.155	5.301	0.043	pMetFam_06-G3.3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.3) 7-hydroxyflavonoids		241
Avenula pubescens	AP	meshDam	M565T321	565.1545962	320.788	5.346	AP	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	565.155	5.301	0.043	pMetFam_06-G3.3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.3) 7-hydroxyflavonoids		241
Festuca rubra	FR	BBCH	M611T245	611.1597322	245.232	4.087	FR	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	611.161	3.671	0.043	pMetFam_06-G3.3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.3) 7-hydroxyflavonoids		203
Festuca rubra	FR	obsShan	M611T245	611.1597322	245.232	4.087	FR	G-3771_MSMS-GNPS-Curated-Pos_2021-09-02	611.161	3.671	0.043	pMetFam_06-G3.3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.3) 7-hydroxyflavonoids		203
Centaurea jacea	CJ	height	M449T320	449.1073228	320.209	5.337	CJ	L-2214_lib_spectra_pos_new_2021-09-02	449.108	5.002	0.030	pMetFam_06-G3.3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.3) 7-hydroxyflavonoids		312
Centaurea jacea	CJ	pathDam	M449T320	449.1073228	320.209	5.337	CJ	L-2214_lib_spectra_pos_new_2021-09-02	449.108	5.002	0.030	pMetFam_06-G3.3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.3) 7-hydroxyflavonoids		312
Dactylis glomerata	DG	season	M359T349	359.1210164	348.802	5.813	DG	L-2214_lib_spectra_pos_new_2021-09-02	359.113	5.837	0.041	pMetFam_06-G3.3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.3) 7-hydroxyflavonoids		62
Dactylis glomerata	DG	obsRich	M359T349	359.1210164	348.802	5.813	DG	M-21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	359.113	5.837	0.041	pMetFam_06-G3.3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.3) 7-hydroxyflavonoids		62
Plantago lanceolata	PL	meshDam	M465T210	465.1006219	210.042	3.501	PA	L-2214_lib_spectra_pos_new_2021-09-02	465.103	3.555	0.008	pMetFam_06-G3.3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.3) 7-hydroxyflavonoids		41
Plantago lanceolata	PL	obsRich	M449T490	449.1425307	490.223	8.170	PL	M-21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	449.144	8.231	0.023	pMetFam_06-G3.3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.3) 7-hydroxyflavonoids		60
Plantago lanceolata	PL	meshDam	M559T269	559.1552594	268.943	4.465	PA	M-21900_MoNA-export-LC-MS-MS_Spectra_2021-09-02	559.155	4.512	0.044	pMetFam_06-G3.3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.3) 7-hydroxyflavonoids		45
Plikweria primrose	PP	BBCH	M559T269	559.1552594	269.053	3.868	PP	R-2760_MSMS-Respect-curated-Pos_2021-09-02	559.155	3.101	0.044	pMetFam_06-G3.3	(00) Phenylpropanoids and polyketides	(G) Flavonoids	(3) Hydroxyflavonoids	(.3) 7-hydroxyflavonoids		41

Supplement Table 3b: Metabolite intensities – showing Season E as example (pMetFam01 to 05)

Feature	M161T178	M250T85	M163T500	M177T316	M199T175	M325T318	M215T187	M183T108	M207T321	M257T232	M450T550	M209T263	M408T229	M403T733	M271T706	M373T457	M161T261	M429T808	M453T838	M175T380_2	M225T274	M353T853	M473T707	M475T695	M515T844	M525T763
pMetFam_ID	pMetFam_01-A'	pMetFam_01-A1.	pMetFam_0	pMetFam_0	pMetFam_0	pMetFam_0	pMetFam_03-/	pMetFam_03-I	pMetFam_03-I	pMetFam_03-I	pMetFam_03-I	pMetFam_04-A1.	pMetFam_04-B'	pMetFam_04-B2	pMetFam_04-B2.:	pMetFam_04-B2.:	pMetFam_04-B2.:	pMetFam_05-/	pMetFam_05-B'	pMetFam_05-	pMetFam_05-	pMetFam_05-	pMetFam_05-	pMetFam_05-	pMetFam_05-	pMetFam_05-C
count	43	71	292	42	125	43	14	17	157	47	14	448	282	38	92	29	203	22	54	65	174	357	35	26	359	42
E_AQ_A018_a	0	0	0	0	0	0	0	1845	0	0	90823	27505	0	0	0	179406	0	50926	0	0	0	66	0	0	5386	0
E_AQ_A018_b	0	61523	0	0	0	0	0	28754	0	0	97956	34956	0	0	0	36545	8983	0	0	0	90032	20	0	0	1266	0
E_AQ_A044_a	0	191911	0	0	0	0	0	0	0	0	263566	23996	0	0	0	0	0	17678	0	0	83434	26	0	0	1740	0
E_AQ_A044_b	0	73221	0	0	0	0	0	0	0	0	229947	44822	0	0	0	155263	0	12569	0	0	0	58	0	0	5098	0
E_AQ_B060_a	0	89765	0	0	0	0	0	0	0	0	0	36057	0	0	0	22910	0	0	0	0	0	0	0	0	2389	0
E_AQ_B060_b	0	0	0	0	0	0	0	0	0	0	165318	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0
E_AQ_C115_a	0	0	0	0	12819	0	0	3834	0	0	0	158738	0	0	0	101458	0	107458	0	0	37604	0	0	0	3403	0
E_AQ_C115_b	0	0	0	0	0	0	0	0	0	0	55667	12818	0	0	0	0	0	584583	0	0	0	7	0	0	0	0
E_AP_B073_a	0	0	0	0	0	0	0	0	0	0	0	93021	0	0	0	0	0	0	0	12576	0	0	0	0	1935	0
E_AP_B073_b	0	0	0	0	10644	0	0	0	0	0	0	53858	0	0	0	0	0	0	0	0	0	57	0	0	0	0
E_AP_B092_a	0	0	0	0	24239	0	0	0	0	0	0	81759	12856	0	0	0	12595	0	0	0	0	31	0	0	5284	0
E_AP_B092_b	0	0	0	0	11541	0	0	0	0	0	0	103465	11626	0	0	0	13237	0	0	0	39001	32	0	0	11801	0
E_AP_C103_a	0	0	0	0	21901	0	0	0	0	0	0	222233	0	0	0	0	6033	0	0	0	13816	35	0	0	7187	0
E_AP_C103_b	0	0	0	0	25387	0	0	0	0	0	0	240019	0	0	0	0	0	0	0	0	133302	59	0	0	35371	0
E_AP_C110_a	0	0	0	0	8276	0	0	0	0	0	0	51577	37628	0	0	0	7296	0	0	0	0	46	0	0	10088	0
E_AP_C110_b	0	0	0	0	0	0	0	0	0	0	0	92020	12439	0	0	0	19594	0	0	0	0	124	0	0	26688	0
E_CJ_A013_a	0	0	90795	0	0	0	0	0	32146	0	0	75802	0	1959298	0	0	10432	0	0	0	0	115	0	73139	21220	0
E_CJ_A013_b	0	0	97665	0	0	0	0	0	0	0	0	101197	3593	2447998	0	0	4888	0	0	0	0	102	0	147723	48121	0
E_CJ_A027_a	0	0	58509	0	0	0	0	0	0	0	0	89084	0	1277052	0	0	0	0	0	0	0	76	0	63844	10447	0
E_CJ_A027_b	0	0	57968	0	0	0	0	0	0	0	0	60913	0	1498350	0	0	0	0	0	0	0	130	0	83674	37059	0
E_CJ_A030_a	0	0	31828	0	0	0	0	0	27927	0	0	63016	0	1180044	0	0	0	0	0	0	0	109	0	49057	50068	0
E_CJ_A030_b	0	0	226174	0	12709	0	0	0	0	30654	0	218167	15262	4378932	0	0	26788	0	0	0	0	0	0	169254	42269	16767
E_CJ_B073_a	0	0	241867	0	0	0	8577	0	0	0	0	168063	0	0	0	0	4993	0	0	0	0	25	922880	0	7990	0
E_CJ_B073_b	0	0	51756	0	0	0	0	0	0	0	0	26138	6226	2004894	0	0	6398	0	0	0	0	152	0	58154	26737	0
E_DG_A005_a	0	0	16377	0	0	0	0	0	0	17881	0	0	8537	0	39390	0	0	0	0	0	0	0	0	0	4582	0
E_DG_A005_b	0	0	16706	0	0	0	0	0	11071	21023	0	50869	0	0	14497	0	0	0	0	0	32519	0	0	0	2160	0
E_DG_A018_a	0	0	7991	0	0	0	0	0	0	13388	0	44696	0	0	18017	0	0	0	0	16750	0	54	0	0	4685	0
E_DG_A018_b	0	0	0	0	0	0	0	0	14647	8797	0	85169	0	0	11881	0	0	0	0	0	0	56	0	0	6310	0
E_DG_C097_a	0	0	56889	0	14096	0	0	0	0	9586	0	19904	5642	0	0	0	0	0	0	0	0	0	0	0	5600	0
E_DG_C097_b	0	0	154934	0	0	0	0	0	0	0	0	73797	12525	0	0	0	6622	0	0	17728	0	59	0	0	5319	0
E_DG_C137_a	0	0	139930	0	9115	0	0	0	0	48350	0	113873	16285	0	9740	0	0	0	0	0	37257	0	0	0	6022	0
E_DG_C137_b	0	0	59941	0	0	0	0	0	15798	19791	0	51260	23340	0	6478	0	0	0	0	0	39263	0	0	0	2878	0
E_FR_B064_a	22670	0	21347	0	6351	0	0	0	0	0	0	79465	0	0	9164	0	6134	0	0	0	0	28	0	0	17654	0
E_FR_B064_b	9754	0	12514	0	0	0	0	0	0	0	0	79758	0	0	0	0	23636	0	0	0	0	28	0	0	8701	0
E_FR_B067_a	0	0	10956	0	0	0	0	0	0	0	0	51875	0	0	8193	0	6431	0	0	0	0	0	0	0	6424	0
E_FR_B067_b	0	63695	0	0	0	0	0	0	0	0	0	32393	0	0	0	0	5249	0	0	0	0	0	0	0	5453	0
E_FR_B073_a	0	0	0	0	35988	0	0	0	0	0	0	56516	0	0	0	0	4993	0	0	0	0	50	0	0	6167	0
E_FR_B073_b	0	0	0	0	12263	0	0	0	0	0	0	116288	8738	0	0	0	0	0	0	0	0	50	0	0	8748	0
E_FR_C114_a	40376	27194	8130	0	0	0	0	0	0	0	0	70233	8983	0	0	0	22022	0	0	11852	0	18	0	0	9820	0
E_FR_C114_b	0	55327	0	0	7493	0	0	0	0	0	0	24176	0	0	0	0	9566	0	0	0	0	0	0	0	5880	0
E_GP_A018_a	0	0	0	0	0	0	0	0	188796	0	0	0	15735	0	0	0	0	0	0	0	0	15	0	0	1910	0
E_GP_A018_b	0	0	0	0	0	0	0	0	262883	0	0	0	11683	0	0	0	0	0	0	0	0	32	0	0	3679	0
E_GP_B081_a	0	0	0	0	0	0	0	0	247538	0	0	0	13824	0	0	0	0	0	0	0	0	28	0	0	1596	0
E_GP_B081_b	0	0	0	0	0	0	0	0	323536	0	0	0	7372	0	0	0	0	0	0	0	0	43	0	0	2987	0
E_GP_C109_a	0	0	0	0	0	0	0	0	247228	0	0	40614	16783	0	0	0	0	0	0	0	0	39	0	0	5135	0
E_GP_C109_b	0	0	0	0	0	0	0	0	221676	0	0	27609	12677	0	0	0	0	0	0	0	0	40	0	0	3184	0
E_GP_C121_a	0	0	0	0	0	0	0	0	317883	0	0	0	23218	0	0	0	0	0	0	0	0	32	0	0	2997	0
E_GP_C121_b	0	0	0	0	0	0	0	0	238674	0	0	17330	19665	0	0	0	0	0	0	0	0	40	0	0	5420	0
E_HL_A018_a	0	0	0	0	31868	0	0	0	0	0	0	93931	0	0	0	0	0	0	0	0	0	0	0	0	2176	0
E_HL_A018_b	0	0	0	0	24581	0	0	0	0	0	0	172130	0	0	0	0	0	0	0	0	0	26	0	0	3685	0
E_HL_A035_a	0	0	0	0	56385	0	0	0	0	0	0	90900	0	0	0	0	0	0	0	0	0	0	0	0	4407	0
E_HL_A035_b	28145	0	0	0	110192	0	0	0	0	0	0	116142	0	0	0	0	0	0	0	0	0	38	0	0	7614	0
E_HL_A040_a	45283	0	10753	0	98544	0	0	0	0	0	0	76857	0	0	0	0	0	0	0	0	0	60	0	0	19560	0
E_HL_A040_b	26720	0	0	0	31755	0	0	0	6068	0	0	82792	0	0	0	0	0	0	0	0	0	224	0	0	0	0
E_HL_B080_a	0	0	0	0	33542	0	0	0	0	0	0	104374	0	0	0	0	0	0	0	0	0	0	0	0	4609	0
E_HL_B080_b	0	0	0	0	41980	0	0	0	0	0	0	140030	4419	0	0	0	0	0	0	0	0	38	0	0	11505	0

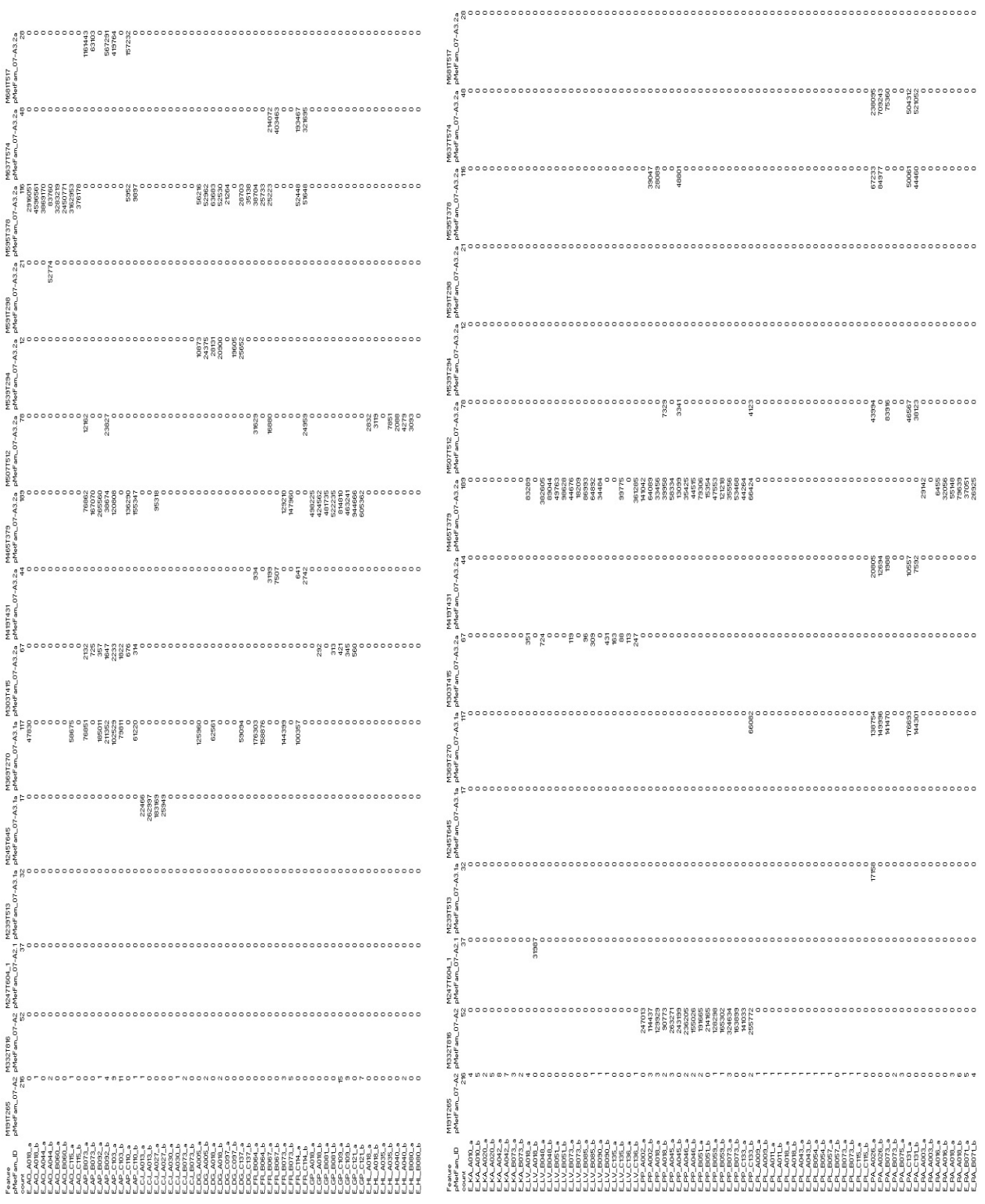
Feature	M161T178	M250T85	M163T500	M177T316	M199T175	M325T318	M215T187	M183T108	M207T321	M257T232	M450T550	M209T263	M408T229	M403T733	M271T706	M373T457	M161T261	M429T808	M453T838	M175T380_2	M225T274	M353T853	M473T707	M475T695	M515T844	M525T763
pMetFam_ID	pMetFam_01-A1.1	pMetFam_01-A1.1a	pMetFam_02	pMetFam_03	pMetFam_03	pMetFam_03	pMetFam_03-A1	pMetFam_03-B1	pMetFam_03-B1	pMetFam_03-B1	pMetFam_03-B1	pMetFam_04-A1.1a	pMetFam_04-B1.1	pMetFam_04-B2.1	pMetFam_04-B2.2a	pMetFam_04-B2.2a	pMetFam_04-B2.2b	pMetFam_05-A1	pMetFam_05-B1.1	pMetFam_05-C	pMetFam_05-C	pMetFam_05-C	pMetFam_05-C	pMetFam_05-C	pMetFam_05-C	pMetFam_05-C
count	43	71	292	42	125	43	14	17	157	47	14	448	282	38	92	29	203	22	54	65	174	357	35	26	359	42
E_KA_A010_a	0	0	322716	0	0	0	0	0	0	0	0	633900	0	0	0	0	6173	0	0	0	0	0	192712	0	785	0
E_KA_A010_b	0	0	268490	0	0	0	0	0	0	0	0	194676	0	0	0	0	3288	0	0	27053	0	0	168655	0	0	0
E_KA_A020_a	0	0	310118	0	0	0	0	0	0	9350	0	277242	0	0	0	0	9269	0	0	17880	0	0	190420	0	3683	0
E_KA_A020_b	0	0	274830	0	0	0	0	0	0	0	0	371397	10410	0	0	0	5669	0	0	30438	0	16	265269	0	3030	0
E_KA_A042_a	0	0	250565	0	0	0	0	0	0	11876	0	154711	0	0	0	0	14318	0	0	35153	0	0	253229	0	1666	0
E_KA_A042_b	0	0	314672	0	0	0	0	0	0	0	0	232768	0	0	0	0	7779	0	0	27165	0	0	220650	0	2121	0
E_KA_B073_a	0	0	361622	0	0	0	13178	0	0	0	0	196705	0	0	0	0	11233	0	0	13137	0	0	229517	0	1013	0
E_KA_B073_b	0	0	158489	0	0	0	0	0	0	0	0	272411	0	0	0	0	7895	0	0	21981	0	0	272149	0	2979	0
E_LV_A018_a	0	0	67468	0	0	0	0	0	0	0	0	45534	9948	0	0	0	16366	0	14264	0	68401	97	0	0	9039	0
E_LV_A018_b	0	0	94474	0	0	0	0	0	0	0	0	38495	0	0	0	0	8875	0	30342	0	0	116	0	0	3722	0
E_LV_B048_a	0	0	294400	0	0	0	0	0	27927	0	0	69607	8783	0	0	0	30343	0	14277	0	105519	146	0	0	6098	0
E_LV_B048_b	0	0	133533	0	0	0	0	0	42225	0	0	104494	0	0	0	0	32244	0	0	0	0	119	0	0	8378	0
E_LV_B051_a	0	0	85499	0	0	0	0	0	37551	0	0	80840	11936	0	0	0	28528	0	0	0	127752	103	0	0	6078	0
E_LV_B051_b	0	0	113423	0	0	0	0	0	26392	0	0	116806	0	0	0	0	31239	0	0	0	253096	100	0	0	6934	0
E_LV_B073_a	0	0	111147	0	0	0	0	0	33438	0	0	87778	0	0	0	0	30387	0	7282	0	183553	65	0	0	3727	0
E_LV_B073_b	0	0	251068	0	0	0	0	24482	31037	0	0	38679	16236	0	0	0	15133	0	22335	0	0	184	0	0	8297	0
E_LV_B085_a	0	0	116253	0	10467	0	0	0	39569	0	0	47867	19300	0	0	0	16432	0	142404	0	78832	133	0	0	6124	0
E_LV_B085_b	0	0	75560	0	0	0	0	0	28855	0	0	113465	7689	0	5990	0	19530	0	0	0	68213	139	0	0	6512	0
E_LV_B090_a	0	0	97809	0	0	16435	0	0	38844	0	0	78081	6155	0	0	0	20903	0	0	0	121391	152	0	0	8090	0
E_LV_B090_b	0	0	75706	0	0	0	0	0	17951	0	0	98596	25800	0	0	0	23332	0	0	0	88749	181	0	0	6192	0
E_LV_C135_a	0	0	80032	0	0	0	0	0	18997	0	0	90635	11461	0	0	0	30530	0	0	0	150835	161	0	0	6722	0
E_LV_C135_b	0	0	147594	0	0	0	0	0	60536	0	0	114670	14926	0	0	0	27829	0	14468	0	0	138	0	0	5024	0
E_LV_C136_a	0	0	134897	0	0	0	0	0	0	0	0	101261	46381	0	0	0	18629	0	0	0	156557	201	0	0	8677	0
E_LV_C136_b	0	0	117768	0	0	13265	0	0	0	0	0	54504	14409	0	0	0	34396	0	17491	0	77432	255	0	0	6448	0
E_PP_A002_a	0	0	34952	0	0	0	0	0	29548	0	0	64855	3960	0	0	0	7527	0	0	0	0	19	0	0	4268	0
E_PP_A002_b	0	30151	34050	0	0	0	0	0	63566	0	0	117193	32231	0	0	0	13326	0	0	0	46411	0	0	0	1826	0
E_PP_A018_a	0	0	26022	0	0	0	0	0	33942	0	0	117291	3894	0	0	0	9699	0	0	0	85310	14	0	0	1808	0
E_PP_A018_b	0	38883	42648	0	0	0	0	0	35240	0	0	65725	0	0	0	0	10259	0	0	0	25032	14	0	0	0	0
E_PP_A045_a	0	31542	14650	0	0	0	0	0	55383	0	0	88367	4869	0	0	0	8572	0	0	0	60122	0	0	0	0	0
E_PP_A045_b	0	46256	45688	0	0	0	0	0	35031	0	0	77847	10209	0	0	0	14778	0	0	0	60562	20	0	0	3194	0
E_PP_A046_a	0	0	22487	0	0	0	0	0	56106	0	0	126135	18354	0	0	0	12511	0	0	0	87129	23	0	0	3817	0
E_PP_A046_b	0	17837	32307	0	0	0	0	0	0	0	0	36708	9160	0	0	0	16987	0	0	0	25651	16	0	0	1218	0
E_PP_B051_a	0	28169	52566	0	0	0	0	0	58640	0	0	97176	12918	0	0	0	12498	0	0	0	0	19	0	0	3815	0
E_PP_B051_b	0	17968	38456	0	0	0	0	0	0	0	0	68797	0	0	0	0	20338	0	0	0	0	0	0	0	2398	0
E_PP_B059_a	0	0	43907	0	0	0	0	0	33518	0	0	70930	21439	0	0	0	27534	0	0	0	45118	30	0	0	5636	0
E_PP_B059_b	0	0	55061	0	0	0	0	0	28041	0	0	102842	34596	0	0	0	21530	0	0	0	0	37	0	0	5966	0
E_PP_B073_a	0	0	42999	0	0	0	0	0	26136	0	0	88074	8676	0	0	0	15099	0	0	0	48959	0	0	0	0	0
E_PP_B073_b	0	19568	21006	0	0	0	0	0	32226	0	0	67426	13730	0	0	0	7144	0	0	0	41058	23	0	0	6617	0
E_PP_C133_a	0	0	20422	0	0	0	0	0	38984	0	0	116866	48203	0	0	0	8470	0	0	0	40519	33	0	0	3928	0
E_PP_C133_b	0	0	30671	0	0	0	0	0	24958	0	0	207287	25942	0	0	0	11809	0	0	0	40241	30	0	0	4786	0
E_PL_A009_a	0	0	24741	0	0	156655	0	0	0	0	0	210622	0	0	0	0	0	0	0	0	65466	17	0	0	0	14728
E_PL_A009_b	0	0	50680	0	0	319671	0	0	0	0	0	227791	6385	0	0	0	0	0	0	0	56306	7	0	0	0	7638
E_PL_A011_a	0	0	47788	0	0	199782	0	0	0	0	0	170883	12705	0	0	0	0	0	0	0	31379	4	0	0	1335	0
E_PL_A011_b	0	0	42640	0	0	163111	0	0	0	0	0	258702	22662	0	0	0	0	0	0	0	53589	4	0	0	0	9307
E_PL_A018_a	0	0	37941	0	0	163924	0	0	0	0	0	204462	0	0	0	0	0	0	0	0	26781	7	0	0	0	13144
E_PL_A018_b	0	0	53356	0	0	133880	0	0	0	0	0	220960	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E_PL_A043_a	0	0	27819	0	0	132549	0	0	0	0	0	207698	8917	0	0	0	0	0	0	0	48670	7	0	0	0	0
E_PL_A043_b	0	0	33462	0	0	134112	0	0	0	0	0	248605	0	0	0	0	0	0	0	7306	47424	7	0	0	0	10666
E_PL_B054_a	0	0	41354	0	0	231707	0	0	0	0	0	279964	0	0	0	0	0	0	0	0	0	0	0	0	3501	12151
E_PL_B054_b	0	0	42828	93400	0	136882	0	0	0	0	0	258406	0	0	0	0	0	0	0	6601	58134	0	0	0	0	0
E_PL_B057_a	0	0	0	0	0	188961	0	0	0	0	0	261731	15752	0	0	0	0	0	0	0	0	9	0	0	0	0
E_PL_B057_b	0	0	39408	0	0	180135	0	0	0	0	0	128662	9268	0	0	0	0	0	0	0	26590	7	0	0	0	0
E_PL_B073_a	0	0	42166	0	0	630367	0	0	0	0	0	259648	0	0	0	0	0	0	0	7322	59291	8	0	0	1267	36130
E_PL_B073_b	0	0	71520	0	0	253789	0	0	0	0	0	220719	5985	0	0	0	0	0	0	0	53979	0	0	0	0	8678
E_PL_C115_a	0	0	0	0	0	138709	0	0	0	0	0	222495	16005	0	0	0	0	0	0	6234	28252	16	0	0	1949	0
E_PL_C115_b	0	0	39029	0	0	118864	0	0	0	0	0	178067	0	0	0	0	0	0	0	0	0	9	0	0	0	12911
E_PA_A026_a	23279	117608	0	0	0	0	0	0	0	0	0	0	8194	0	10186	0	0	0	0	0	33523	0	0	0	5563	0
E_PA_A026_b	32908	150651	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	34822	0	0	0	5708	0
E_PA_B073_a	0	177171	0	0	0	0	0	0	0	0	0	43448	4531	0	10097	0	15277	0	0	0	0	26	0	0	11137	0
E_PA_B073_b	0	0	0	0	11240	0	0	0	0	0	0	52871	0	0	0	0	0	0	0	0	34973	54	0	0	0	0
E_PA_B075_a	0	0	0	0	6199	0	0	0	0	0	0	37494	0	0	0	0	0	0	0	0	0	0	0	0	2087	0
E_PA_C131_a	22981	185066	0	0	0	0	0	0	0	0	0	0	13943	0	5432	0	6559	0	0	0	0	25	0	0	9488	0
E_PA_C131_b	0	152947	0	0	0	0	0	0	0	0	0	0	18375	0	13970	0	5934	0	0	0	0	0	0	0	5750	0
E_RA_A003_a	0	0	235093	0	7750	0	0	0	0	120099	0	128419	0	0	0	0	0	0	0	17190	0	30	0	0	0	0
E_RA_A003_b	0	0	421251	0	0	0	0	0	0	118872	0	44739	0	0	0	0	0	0	0	0	0	38	0	0	0	0
E_RA_A016_a	0	0	280502	0	18900	0	0	0	0	76810	0	37169	0	0	0	0	0	0	0	0	0	43	0	0	0	0
E_RA_A016_b	0	0	445571	0	4269	0	0	0	0	88328	0	54210	0	0	0	0	0	0	0	0	0	46	0	0	1532	0
E_RA_A018_a	0	0	37261	0	0	0	0	0	0	72952	0	212367	0	0	0	0	0	0	0	56054	0	20	0	0	0	0
E_RA_A018_b	0	0	304343	0	0	0	0	0	0	110102	0	183296	0	0	0	0	0	0	0	20008	0	0	0	0	0	0
E_RA_B071_a	0	0	314577	0	0	0	0	0	0	113322	0	169798	0	0	0	0	0	0	0	15293	0	38	0	0	2973	3637
E_RA_B071_b	0	0	264859	0	0	0	0	0	0	52163	0	113659	0	0	0	0	0	0	0	14339	0	32	0	0	0	0

Supplement Table 3b: Metabolite intensities – showing Season E as example (pMetFam06 to i1.2a)

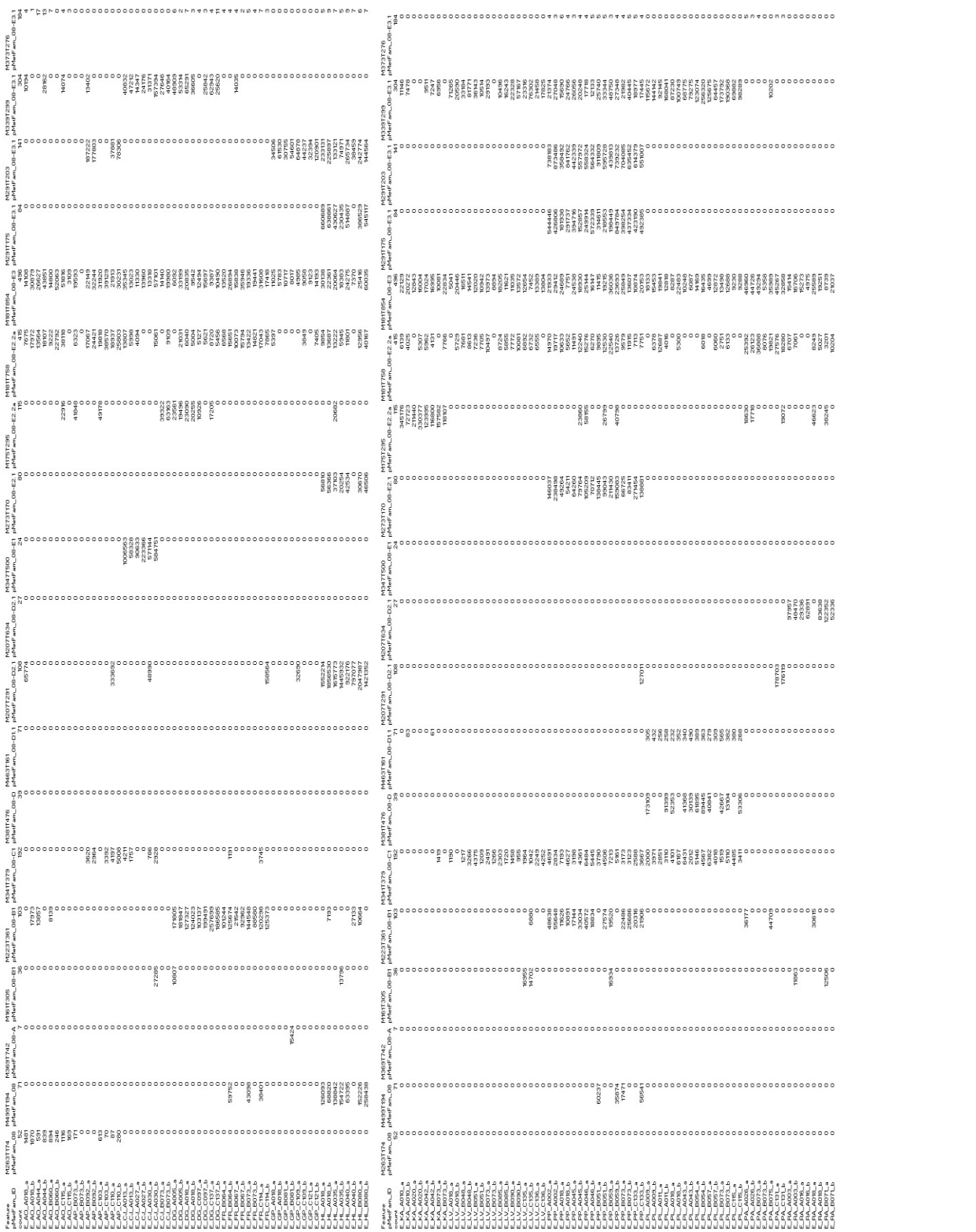


Supplement Table 3b: Metabolite intensities – showing Season E as example (pMetFam07 to A1.1a)

Supplement Table 3b: Metabolite intensities – showing Season E as example (pMetFam07-A2 to A3.2a)

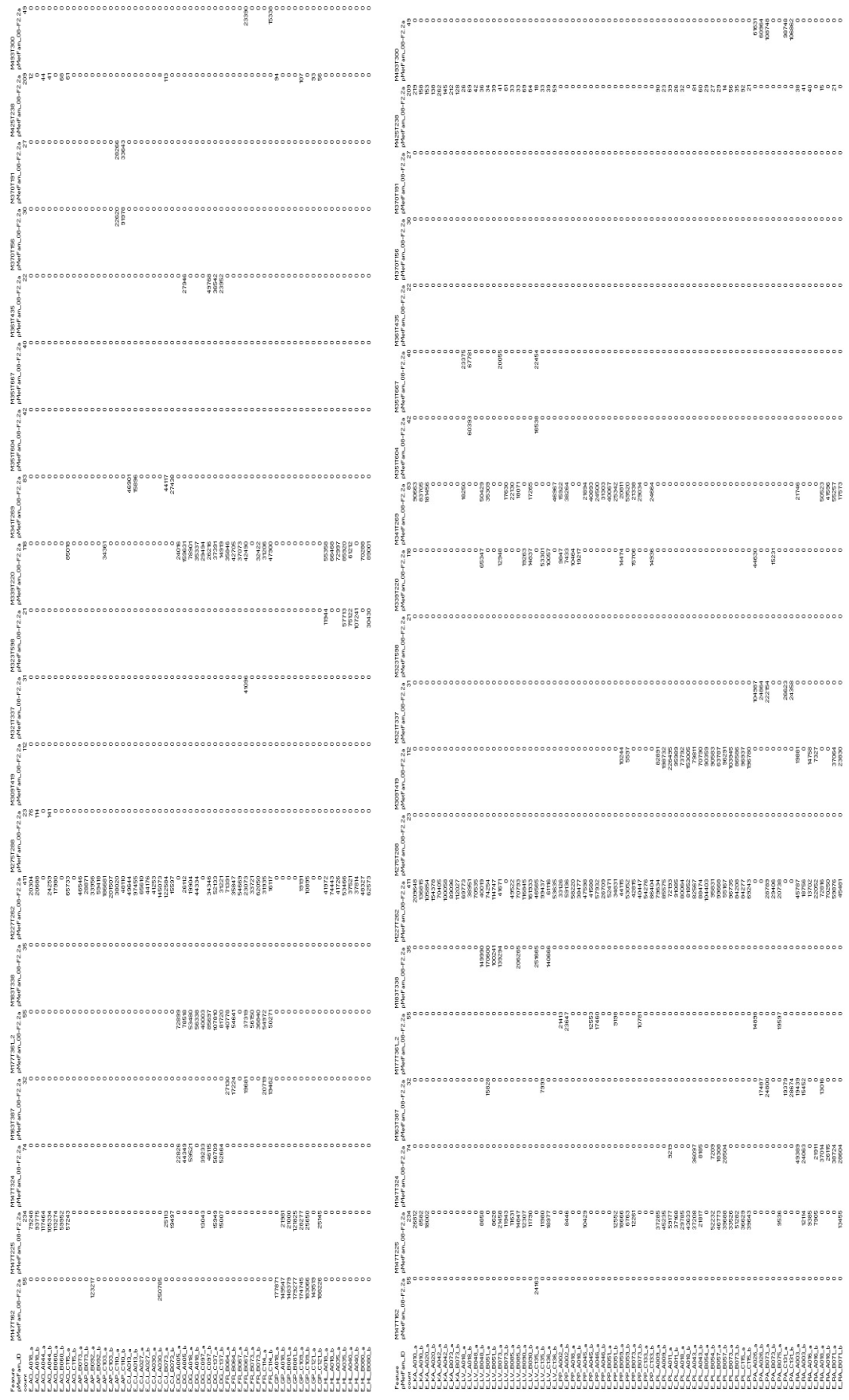


Supplement Table 3b: Metabolite intensities – showing Season E as example (pMetFam08 to E3.1)



Supplement Table 3b: Metabolite intensities – showing Season E as example (pMetFam08-F1 to F2.2)

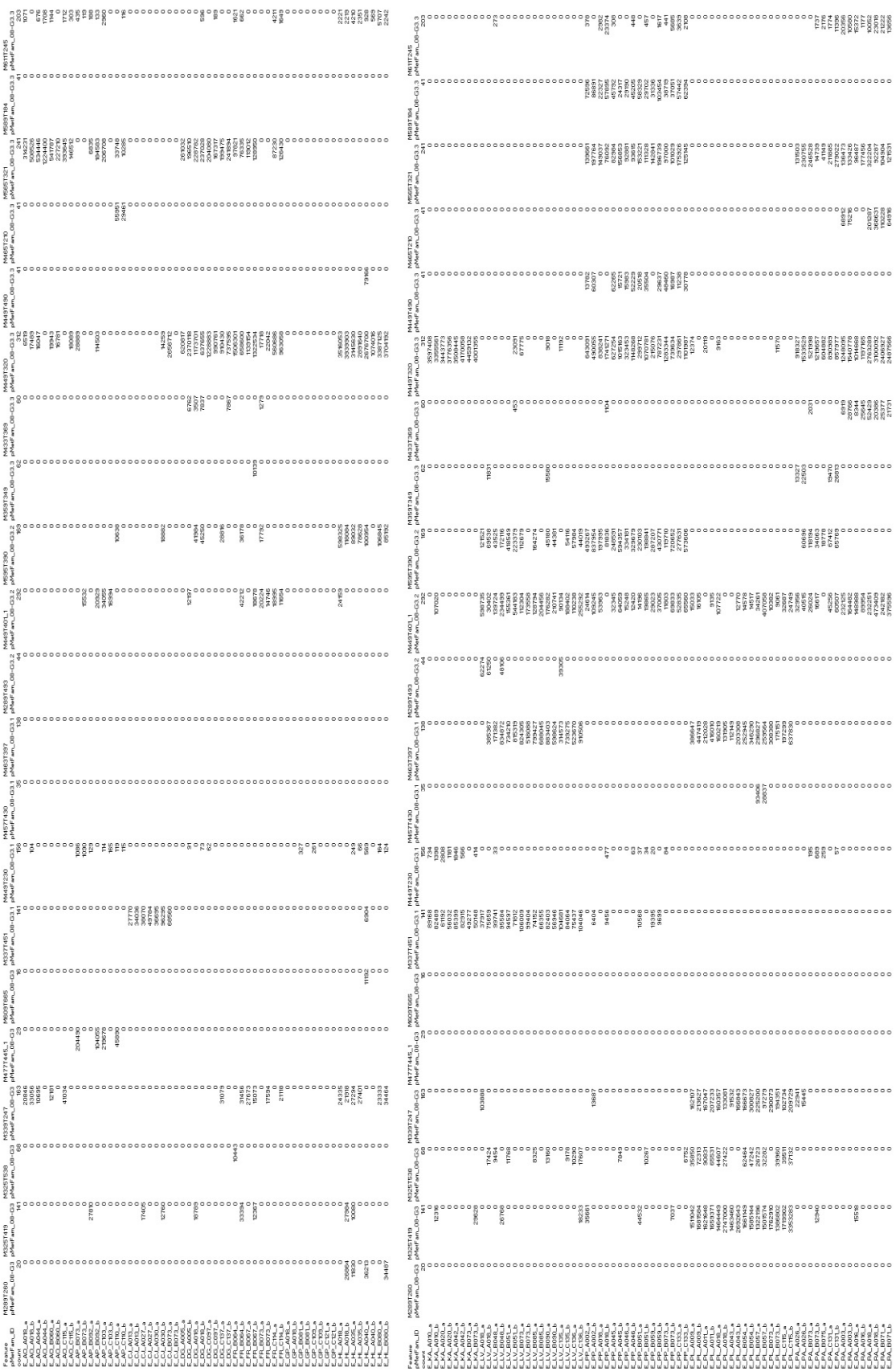
Supplement Table 3b: Metabolite intensities – showing Season E as example (pMetFam08-F2.2a)



Supplement Table 3b: Metabolite intensities – showing Season E as example (pMetFam08-G to G2.1b)

Supplement Table 3b: Metabolite intensities – showing Season E as example (pMetFam08-G2.1c)

Supplement Table 3b: Metabolite intensities – showing Season E as example (pMetFam08-G3)



Supplemental Table 4: Recorded traits on plot and plant level, and feature diversity indices across four seasons and thirteen plant species. The plant species belong to the functional groups (FG) grass and herb and were collected in plots of the Trait-Based Experiment located in the Jena Experiment. Trait recorded on plot level included plant species richness (pRich) and plant species Shannon diversity (pShan) per plot and season. On the plant level we recorded height, developmental stage (BBCH scale), mechanical (mechD) and pathogen inflicted damage (pathD) for each replicate. For each sample, LC-MS feature richness (fRich), feature Shannon diversity (fShan) and feature Evenness (fEven) was measured and calculated. Data is shown for 504 validated samples.

Sample ID							plot level trait		plant level trait				feature diversity		
sampleID	Species	2-letter code	FG	plot	Season	Season letter code	pRich [#]	pShan [scale]	height [cm]	BBCH [scale]	mechD [%]	pathD [%]	fRich	fShan	fEven
2018_E_ANTODO_A018_a	*Anthoxanthum odoratum*	AO	grass	A018	spring	E	20	2.25	52	85	0	20	568	3.44	0.54
2018_E_ANTODO_A018_b	*Anthoxanthum odoratum*	AO	grass	A018	spring	E	20	2.25	42	85	0	40	571	2.44	0.38
2018_E_ANTODO_A044_a	*Anthoxanthum odoratum*	AO	grass	A044	spring	E	17	1.88	66	80	0	15	533	1.17	0.19
2018_E_ANTODO_A044_b	*Anthoxanthum odoratum*	AO	grass	A044	spring	E	17	1.88	60	75	0	1	595	2.24	0.35
2018_E_ANTODO_B060_a	*Anthoxanthum odoratum*	AO	grass	B060	spring	E	20	2.63	59	65	0	0	488	2.09	0.34
2018_E_ANTODO_B060_b	*Anthoxanthum odoratum*	AO	grass	B060	spring	E	20	2.63	58	80	0	15	442	5.14	0.84
2018_E_ANTODO_C115_a	*Anthoxanthum odoratum*	AO	grass	C115	spring	E	17	2.44	70	80	0	20	549	0.52	0.08
2018_E_ANTODO_C115_b	*Anthoxanthum odoratum*	AO	grass	C115	spring	E	17	2.44	56	80	0	20	250	0.27	0.05
2018_E_AVEPUB_B073_a	*Avenula pubescens*	AP	grass	B073	spring	E	15	2.35	90	70	0	5	383	3.76	0.63
2018_E_AVEPUB_B073_b	*Avenula pubescens*	AP	grass	B073	spring	E	15	2.35	100	70	5	30	296	1.80	0.32
2018_E_AVEPUB_B092_a	*Avenula pubescens*	AP	grass	B092	spring	E	22	2.23	85	60	0	30	500	5.31	0.85
2018_E_AVEPUB_B092_b	*Avenula pubescens*	AP	grass	B092	spring	E	22	2.23	84	60	0	2	538	2.19	0.35
2018_E_AVEPUB_C103_a	*Avenula pubescens*	AP	grass	C103	spring	E	16	1.98	104	40	0	10	506	5.30	0.85
2018_E_AVEPUB_C103_b	*Avenula pubescens*	AP	grass	C103	spring	E	16	1.98	122	40	0	20	419	2.87	0.48
2018_E_AVEPUB_C110_a	*Avenula pubescens*	AP	grass	C110	spring	E	19	2.15	96	60	5	20	567	5.49	0.87
2018_E_AVEPUB_C110_b	*Avenula pubescens*	AP	grass	C110	spring	E	19	2.15	111	55	0	15	565	5.43	0.86
2018_E_CENJAC_A013_a	*Centaurea jacea*	CJ	herb	A013	spring	E	19	2.14	43	40	3	5	700	5.10	0.78
2018_E_CENJAC_A013_b	*Centaurea jacea*	CJ	herb	A013	spring	E	19	2.14	27	40	0	5	696	5.17	0.79
2018_E_CENJAC_A027_a	*Centaurea jacea*	CJ	herb	A027	spring	E	14	1.68	38	40	0	20	603	5.00	0.78
2018_E_CENJAC_A027_b	*Centaurea jacea*	CJ	herb	A027	spring	E	14	1.68	36	40	0	20	594	4.86	0.76
2018_E_CENJAC_A030_a	*Centaurea jacea*	CJ	herb	A030	spring	E	12	1.79	21	40	2	0	591	4.99	0.78
2018_E_CENJAC_A030_b	*Centaurea jacea*	CJ	herb	A030	spring	E	12	1.79	34	40	0	0	929	5.36	0.78
2018_E_CENJAC_B073_a	*Centaurea jacea*	CJ	herb	B073	spring	E	15	2.35	15	40	5	0	813	5.48	0.82
2018_E_CENJAC_B073_b	*Centaurea jacea*	CJ	herb	B073	spring	E	15	2.35	26	40	0	2	602	4.94	0.77
2018_E_DACGLO_A005_a	*Dactylis glomerata*	DG	grass	A005	spring	E	20	2.42	112	75	0	25	413	1.73	0.29
2018_E_DACGLO_A005_b	*Dactylis glomerata*	DG	grass	A005	spring	E	20	2.42	101	75	5	15	453	1.81	0.30
2018_E_DACGLO_A018_a	*Dactylis glomerata*	DG	grass	A018	spring	E	20	2.25	98	75	0	30	445	1.59	0.26
2018_E_DACGLO_A018_b	*Dactylis glomerata*	DG	grass	A018	spring	E	20	2.25	100	60	0	5	461	2.45	0.40
2018_E_DACGLO_C097_a	*Dactylis glomerata*	DG	grass	C097	spring	E	13	1.91	82	75	0	5	357	1.57	0.27
2018_E_DACGLO_C097_b	*Dactylis glomerata*	DG	grass	C097	spring	E	13	1.91	107	80	0	5	437	2.20	0.36
2018_E_DACGLO_C137_a	*Dactylis glomerata*	DG	grass	C137	spring	E	20	2.02	96	40	0	10	431	1.60	0.26
2018_E_DACGLO_C137_b	*Dactylis glomerata*	DG	grass	C137	spring	E	20	2.02	121	40	0	15	485	5.21	0.84
2018_E_FESRUB_B064_a	*Festuca rubra*	FR	grass	B064	spring	E	15	1.99	55	55	0	5	424	0.54	0.09
2018_E_FESRUB_B064_b	*Festuca rubra*	FR	grass	B064	spring	E	15	1.99	58	55	0	5	465	0.31	0.05
2018_E_FESRUB_B067_a	*Festuca rubra*	FR	grass	B067	spring	E	13	1.83	60	55	0	5	353	0.29	0.05
2018_E_FESRUB_B067_b	*Festuca rubra*	FR	grass	B067	spring	E	13	1.83	72	55	0	5	365	0.43	0.07
2018_E_FESRUB_B073_a	*Festuca rubra*	FR	grass	B073	spring	E	15	2.35	88	55	0	10	291	0.59	0.10
2018_E_FESRUB_B073_b	*Festuca rubra*	FR	grass	B073	spring	E	15	2.35	99	60	2	5	283	0.79	0.14
2018_E_FESRUB_C114_a	*Festuca rubra*	FR	grass	C114	spring	E	11	1.40	64	65	0	2	404	0.24	0.04
2018_E_FESRUB_C114_b	*Festuca rubra*	FR	grass	C114	spring	E	11	1.40	65	80	0	5	384	0.24	0.04
2018_E_GERPRA_A018_a	*Geranium pratense*	GP	herb	A018	spring	E	20	2.25	23	45	5	8	463	5.16	0.84
2018_E_GERPRA_A018_b	*Geranium pratense*	GP	herb	A018	spring	E	20	2.25	39	40	10	10	503	5.19	0.83
2018_E_GERPRA_B081_a	*Geranium pratense*	GP	herb	B081	spring	E	11	1.69	70	60	5	15	472	5.10	0.83
2018_E_GERPRA_B081_b	*Geranium pratense*	GP	herb	B081	spring	E	11	1.69	62	60	8	15	522	5.23	0.83
2018_E_GERPRA_C109_a	*Geranium pratense*	GP	herb	C109	spring	E	10	1.73	38	80	2	20	504	5.12	0.82
2018_E_GERPRA_C109_b	*Geranium pratense*	GP	herb	C109	spring	E	10	1.73	35	80	2	20	545	5.36	0.85
2018_E_GERPRA_C121_a	*Geranium pratense*	GP	herb	C121	spring	E	19	1.71	44	65	5	5	479	5.05	0.82
2018_E_GERPRA_C121_b	*Geranium pratense*	GP	herb	C121	spring	E	19	1.71	49	70	0	0	484	5.15	0.83
2018_E_HOLLAN_A018_a	*Holcus lanatus*	HL	grass	A018	spring	E	20	2.25	83	55	0	3	419	1.68	0.28
2018_E_HOLLAN_A018_b	*Holcus lanatus*	HL	grass	A018	spring	E	20	2.25	79	75	0	7	464	1.83	0.30
2018_E_HOLLAN_A035_a	*Holcus lanatus*	HL	grass	A035	spring	E	15	1.87	48	65	0	5	437	0.70	0.11
2018_E_HOLLAN_A035_b	*Holcus lanatus*	HL	grass	A035	spring	E	15	1.87	71	65	5	20	495	0.88	0.14
2018_E_HOLLAN_A040_a	*Holcus lanatus*	HL	grass	A040	spring	E	16	2.03	77	60	0	5	526	2.13	0.34
2018_E_HOLLAN_A040_b	*Holcus lanatus*	HL	grass	A040	spring	E	16	2.03	92	60	0	30	314	0.88	0.15
2018_E_HOLLAN_B080_a	*Holcus lanatus*	HL	grass	B080	spring	E	18	2.18	81	60	0	10	433	0.89	0.15
2018_E_HOLLAN_B080_b	*Holcus lanatus*	HL	grass	B080	spring	E	18	2.18	78	60	0	5	509	1.14	0.18

Supplemental Table 4: continued

Sample ID					plot level trait				plant level trait				feature diversity		
sampleID	Species	2-letter code	FG	plot	Season	Season letter code	pRich [#]	pShan [scale]	height [cm]	BBCH [scale]	mechD [%]	pathD [%]	fRich	fShan	fEven
2018_E_KNAARV_A010_a	*Knautia arvensis*	KA	herb	A010	spring	E	18	1.77	32	85	0	5	821	5.63	0.84
2018_E_KNAARV_A010_b	*Knautia arvensis*	KA	herb	A010	spring	E	18	1.77	77	85	3	10	695	5.41	0.83
2018_E_KNAARV_A020_a	*Knautia arvensis*	KA	herb	A020	spring	E	19	1.62	59	85	0	10	780	5.62	0.84
2018_E_KNAARV_A020_b	*Knautia arvensis*	KA	herb	A020	spring	E	19	1.62	68	85	0	10	788	5.60	0.84
2018_E_KNAARV_A042_a	*Knautia arvensis*	KA	herb	A042	spring	E	10	1.33	65	85	10	25	759	5.58	0.84
2018_E_KNAARV_A042_b	*Knautia arvensis*	KA	herb	A042	spring	E	10	1.33	82	85	5	15	744	5.44	0.82
2018_E_KNAARV_B073_a	*Knautia arvensis*	KA	herb	B073	spring	E	15	2.35	90	80	5	15	735	5.50	0.83
2018_E_KNAARV_B073_b	*Knautia arvensis*	KA	herb	B073	spring	E	15	2.35	77	75	0	5	781	5.51	0.83
2018_E_LEUVUL_A018_a	*Leucanthemum vulgare*	LV	herb	A018	spring	E	20	2.25	56	75	3	7	550	5.21	0.83
2018_E_LEUVUL_A018_b	*Leucanthemum vulgare*	LV	herb	A018	spring	E	20	2.25	55	60	0	5	507	5.19	0.83
2018_E_LEUVUL_B048_a	*Leucanthemum vulgare*	LV	herb	B048	spring	E	17	1.90	58	70	0	20	568	5.20	0.82
2018_E_LEUVUL_B048_b	*Leucanthemum vulgare*	LV	herb	B048	spring	E	17	1.90	70	70	5	15	649	5.43	0.84
2018_E_LEUVUL_B051_a	*Leucanthemum vulgare*	LV	herb	B051	spring	E	17	2.07	70	60	0	15	585	5.28	0.83
2018_E_LEUVUL_B051_b	*Leucanthemum vulgare*	LV	herb	B051	spring	E	17	2.07	71	60	0	10	572	5.41	0.85
2018_E_LEUVUL_B073_a	*Leucanthemum vulgare*	LV	herb	B073	spring	E	15	2.35	43	60	5	5	646	5.37	0.83
2018_E_LEUVUL_B073_b	*Leucanthemum vulgare*	LV	herb	B073	spring	E	15	2.35	70	60	2	8	578	5.29	0.83
2018_E_LEUVUL_B085_a	*Leucanthemum vulgare*	LV	herb	B085	spring	E	19	1.78	60	60	0	5	691	5.52	0.84
2018_E_LEUVUL_B085_b	*Leucanthemum vulgare*	LV	herb	B085	spring	E	19	1.78	62	60	0	5	584	5.20	0.82
2018_E_LEUVUL_B090_a	*Leucanthemum vulgare*	LV	herb	B090	spring	E	14	1.76	57	75	0	10	595	5.26	0.82
2018_E_LEUVUL_B090_b	*Leucanthemum vulgare*	LV	herb	B090	spring	E	14	1.76	67	75	0	5	569	5.30	0.84
2018_E_LEUVUL_C135_a	*Leucanthemum vulgare*	LV	herb	C135	spring	E	14	1.61	64	60	0	20	670	5.17	0.79
2018_E_LEUVUL_C135_b	*Leucanthemum vulgare*	LV	herb	C135	spring	E	14	1.61	49	60	0	20	627	5.42	0.84
2018_E_LEUVUL_C136_a	*Leucanthemum vulgare*	LV	herb	C136	spring	E	13	1.72	70	75	2	10	669	5.39	0.83
2018_E_LEUVUL_C136_b	*Leucanthemum vulgare*	LV	herb	C136	spring	E	13	1.72	53	75	0	15	608	5.25	0.82
2018_E_PHLPRA_A002_a	*Phleum pratense*	PP	grass	A002	spring	E	24	2.10	53	60	0	0	525	0.64	0.10
2018_E_PHLPRA_A002_b	*Phleum pratense*	PP	grass	A002	spring	E	24	2.10	59	60	0	5	593	0.80	0.13
2018_E_PHLPRA_A018_a	*Phleum pratense*	PP	grass	A018	spring	E	20	2.25	60	70	0	5	461	0.58	0.09
2018_E_PHLPRA_A018_b	*Phleum pratense*	PP	grass	A018	spring	E	20	2.25	52	55	2	10	444	0.72	0.12
2018_E_PHLPRA_A045_a	*Phleum pratense*	PP	grass	A045	spring	E	14	1.86	66	75	0	5	440	0.53	0.09
2018_E_PHLPRA_A045_b	*Phleum pratense*	PP	grass	A045	spring	E	14	1.86	67	75	0	0	495	0.87	0.14
2018_E_PHLPRA_A046_a	*Phleum pratense*	PP	grass	A046	spring	E	15	1.72	74	75	0	5	519	0.45	0.07
2018_E_PHLPRA_A046_b	*Phleum pratense*	PP	grass	A046	spring	E	15	1.72	59	75	0	5	447	1.02	0.17
2018_E_PHLPRA_B051_a	*Phleum pratense*	PP	grass	B051	spring	E	17	2.07	77	60	0	2	547	1.12	0.18
2018_E_PHLPRA_B051_b	*Phleum pratense*	PP	grass	B051	spring	E	17	2.07	72	60	0	5	502	0.85	0.14
2018_E_PHLPRA_B059_a	*Phleum pratense*	PP	grass	B059	spring	E	19	2.51	45	60	0	0	605	0.20	0.03
2018_E_PHLPRA_B059_b	*Phleum pratense*	PP	grass	B059	spring	E	19	2.51	37	40	0	0	564	0.95	0.15
2018_E_PHLPRA_B073_a	*Phleum pratense*	PP	grass	B073	spring	E	15	2.35	57	50	0	5	462	1.24	0.20
2018_E_PHLPRA_B073_b	*Phleum pratense*	PP	grass	B073	spring	E	15	2.35	51	50	0	5	506	0.91	0.15
2018_E_PHLPRA_C133_a	*Phleum pratense*	PP	grass	C133	spring	E	18	2.50	52	60	0	0	508	0.60	0.10
2018_E_PHLPRA_C133_b	*Phleum pratense*	PP	grass	C133	spring	E	18	2.50	45	60	0	0	537	0.60	0.09
2018_E_PLALAN_A009_a	*Plantago lanceolata*	PL	herb	A009	spring	E	20	1.89	14	50	15	25	747	5.70	0.86
2018_E_PLALAN_A009_b	*Plantago lanceolata*	PL	herb	A009	spring	E	20	1.89	24	85	10	20	783	5.64	0.85
2018_E_PLALAN_A011_a	*Plantago lanceolata*	PL	herb	A011	spring	E	17	1.51	19	40	5	10	717	5.53	0.84
2018_E_PLALAN_A011_b	*Plantago lanceolata*	PL	herb	A011	spring	E	17	1.51	27	40	15	10	714	5.57	0.85
2018_E_PLALAN_A018_a	*Plantago lanceolata*	PL	herb	A018	spring	E	20	2.25	19	40	15	30	706	5.69	0.87
2018_E_PLALAN_A018_b	*Plantago lanceolata*	PL	herb	A018	spring	E	20	2.25	30	55	6	0	627	5.31	0.82
2018_E_PLALAN_A043_a	*Plantago lanceolata*	PL	herb	A043	spring	E	18	2.13	31	75	10	10	669	5.49	0.84
2018_E_PLALAN_A043_b	*Plantago lanceolata*	PL	herb	A043	spring	E	18	2.13	26	85	15	25	698	5.44	0.83
2018_E_PLALAN_B054_a	*Plantago lanceolata*	PL	herb	B054	spring	E	19	1.99	22	75	10	5	782	5.78	0.87
2018_E_PLALAN_B054_b	*Plantago lanceolata*	PL	herb	B054	spring	E	19	1.99	18	75	10	15	819	5.81	0.87
2018_E_PLALAN_B057_a	*Plantago lanceolata*	PL	herb	B057	spring	E	20	2.40	32	80	5	15	670	5.52	0.85
2018_E_PLALAN_B057_b	*Plantago lanceolata*	PL	herb	B057	spring	E	20	2.40	20	55	5	10	669	5.39	0.83
2018_E_PLALAN_B073_a	*Plantago lanceolata*	PL	herb	B073	spring	E	15	2.35	22	40	5	10	773	5.68	0.85
2018_E_PLALAN_B073_b	*Plantago lanceolata*	PL	herb	B073	spring	E	15	2.35	16	40	3	10	772	5.60	0.84
2018_E_PLALAN_C115_a	*Plantago lanceolata*	PL	herb	C115	spring	E	17	2.44	39	65	10	5	636	5.48	0.85
2018_E_PLALAN_C115_b	*Plantago lanceolata*	PL	herb	C115	spring	E	17	2.44	17	55	5	15	747	5.53	0.84
2018_E_POAPRA_A026_a	*Poa pratensis*	PA	grass	A026	spring	E	20	2.00	51	60	0	5	368	0.24	0.04
2018_E_POAPRA_A026_b	*Poa pratensis*	PA	grass	A026	spring	E	20	2.00	54	55	0	5	344	0.31	0.05
2018_E_POAPRA_B073_a	*Poa pratensis*	PA	grass	B073	spring	E	15	2.35	64	60	0	0	332	0.34	0.06
2018_E_POAPRA_B073_b	*Poa pratensis*	PA	grass	B073	spring	E	15	2.35	62	60	0	20	190	0.47	0.09
2018_E_POAPRA_B075_a	*Poa pratensis*	PA	grass	B075	spring	E	13	1.61	68	80	0	15	242	0.92	0.17
2018_E_POAPRA_C131_a	*Poa pratensis*	PA	grass	C131	spring	E	17	1.86	67	60	0	0	364	0.39	0.07
2018_E_POAPRA_C131_b	*Poa pratensis*	PA	grass	C131	spring	E	17	1.86	55	60	0	5	410	0.23	0.04
2018_E_RANACR_A003_a	*Ranunculus acris*	RA	herb	A003	spring	E	20	2.31	54	80	0	20	495	1.66	0.27
2018_E_RANACR_A003_b	*Ranunculus acris*	RA	herb	A003	spring	E	20	2.31	71	80	0	20	398	0.99	0.17
2018_E_RANACR_A016_a	*Ranunculus acris*	RA	herb	A016	spring	E	27	2.59	47	80	1	30	408	1.03	0.17
2018_E_RANACR_A016_b	*Ranunculus acris*	RA	herb	A016	spring	E	27	2.59	57	80	1	25	431	1.56	0.26
2018_E_RANACR_A018_a	*Ranunculus acris*	RA	herb	A018	spring	E	20	2.25	35	80	30	2	514	1.32	0.21
2018_E_RANACR_A018_b	*Ranunculus acris*	RA	herb	A018	spring	E	20	2.25	62	80	0	30	488	1.32	0.21
2018_E_RANACR_B071_a	*Ranunculus acris*	RA	herb	B071	spring	E	17	2.35	79	80	0	20	571	1.34	0.21
2018_E_RANACR_B071_b	*Ranunculus acris*	RA	herb	B071	spring	E	17	2.35	32	80	0	25	457	1.56	0.25

Supplemental Table 4: continued

Sample ID							plot level trait		plant level trait				feature diversity		
sampleID	Species	2-letter code	FG	plot	Season	Season letter code	pRich [#]	pShan [scale]	height [cm]	BBCH [scale]	mechD [%]	pathD [%]	fRich	fShan	fEven
2018_F_ANTODO_A018_b	*Anthoxanthum odoratum*	AO	grass	A018	early summer	F	12	2.09	5	60	0	5	541	4.63	0.74
2018_F_ANTODO_A044_b	*Anthoxanthum odoratum*	AO	grass	A044	early summer	F	6	1.07	9	40	0	10	626	3.35	0.52
2018_F_ANTODO_B060_a	*Anthoxanthum odoratum*	AO	grass	B060	early summer	F	10	1.87	13	35	3	10	655	3.65	0.56
2018_F_ANTODO_B060_b	*Anthoxanthum odoratum*	AO	grass	B060	early summer	F	10	1.87	11	35	10	20	636	2.85	0.44
2018_F_ANTODO_C115_a	*Anthoxanthum odoratum*	AO	grass	C115	early summer	F	11	1.95	13	65	10	30	700	4.72	0.72
2018_F_ANTODO_C115_b	*Anthoxanthum odoratum*	AO	grass	C115	early summer	F	11	1.95	9	65	5	20	644	4.03	0.62
2018_F_AVEPUB_B073_a	*Avenula pubescens*	AP	grass	B073	early summer	F	10	1.96	18	50	0	3	576	3.40	0.53
2018_F_AVEPUB_B073_b	*Avenula pubescens*	AP	grass	B073	early summer	F	10	1.96	16	35	0	10	518	5.17	0.83
2018_F_AVEPUB_B092_a	*Avenula pubescens*	AP	grass	B092	early summer	F	8	1.39	19	35	5	10	535	5.16	0.82
2018_F_AVEPUB_B092_b	*Avenula pubescens*	AP	grass	B092	early summer	F	8	1.39	18	35	5	5	543	5.39	0.86
2018_F_AVEPUB_C103_a	*Avenula pubescens*	AP	grass	C103	early summer	F	7	1.31	17	65	10	10	592	5.50	0.86
2018_F_AVEPUB_C103_b	*Avenula pubescens*	AP	grass	C103	early summer	F	7	1.31	13	35	5	5	613	5.35	0.83
2018_F_AVEPUB_C110_a	*Avenula pubescens*	AP	grass	C110	early summer	F	14	2.24	20	35	5	25	589	5.49	0.86
2018_F_AVEPUB_C110_b	*Avenula pubescens*	AP	grass	C110	early summer	F	14	2.24	18	35	2	20	522	5.17	0.83
2018_F_CENJAC_A013_a	*Centaurea jacea*	CJ	herb	A013	early summer	F	7	1.23	40	40	0	0	615	5.02	0.78
2018_F_CENJAC_A013_b	*Centaurea jacea*	CJ	herb	A013	early summer	F	7	1.23	38	40	0	5	648	4.97	0.77
2018_F_CENJAC_A027_a	*Centaurea jacea*	CJ	herb	A027	early summer	F	7	1.10	23	55	5	20	641	5.01	0.77
2018_F_CENJAC_A027_b	*Centaurea jacea*	CJ	herb	A027	early summer	F	7	1.10	41	35	5	20	640	5.02	0.78
2018_F_CENJAC_A030_a	*Centaurea jacea*	CJ	herb	A030	early summer	F	6	1.34	25	35	3	0	615	4.93	0.77
2018_F_CENJAC_A030_b	*Centaurea jacea*	CJ	herb	A030	early summer	F	6	1.34	25	35	5	40	654	5.20	0.80
2018_F_CENJAC_B073_a	*Centaurea jacea*	CJ	herb	B073	early summer	F	10	1.96	13	35	0	0	609	4.99	0.78
2018_F_CENJAC_B073_b	*Centaurea jacea*	CJ	herb	B073	early summer	F	10	1.96	26	65	0	5	680	5.12	0.79
2018_F_DACGLO_A005_a	*Dactylis glomerata*	DG	grass	A005	early summer	F	11	1.99	20	65	3	10	447	1.37	0.23
2018_F_DACGLO_A005_b	*Dactylis glomerata*	DG	grass	A005	early summer	F	11	1.99	22	65	3	15	409	1.70	0.28
2018_F_DACGLO_A018_a	*Dactylis glomerata*	DG	grass	A018	early summer	F	12	2.09	20	65	3	5	379	2.31	0.39
2018_F_DACGLO_A018_b	*Dactylis glomerata*	DG	grass	A018	early summer	F	12	2.09	32	35	5	10	427	1.98	0.33
2018_F_DACGLO_C097_a	*Dactylis glomerata*	DG	grass	C097	early summer	F	10	1.48	24	35	10	10	494	2.45	0.39
2018_F_DACGLO_C097_b	*Dactylis glomerata*	DG	grass	C097	early summer	F	10	1.48	21	65	10	15	417	4.99	0.83
2018_F_DACGLO_C137_a	*Dactylis glomerata*	DG	grass	C137	early summer	F	5	1.17	30	70	2	5	474	2.37	0.39
2018_F_DACGLO_C137_b	*Dactylis glomerata*	DG	grass	C137	early summer	F	5	1.17	21	40	5	15	413	5.16	0.86
2018_F_FESRUB_B064_a	*Festuca rubra*	FR	grass	B064	early summer	F	9	0.77	20	35	5	15	305	0.36	0.06
2018_F_FESRUB_B064_b	*Festuca rubra*	FR	grass	B064	early summer	F	9	0.77	16	65	5	15	358	0.50	0.08
2018_F_FESRUB_B067_a	*Festuca rubra*	FR	grass	B067	early summer	F	9	1.65	25	60	2	25	382	0.35	0.06
2018_F_FESRUB_B067_b	*Festuca rubra*	FR	grass	B067	early summer	F	9	1.65	18	35	3	20	334	0.41	0.07
2018_F_FESRUB_B073_a	*Festuca rubra*	FR	grass	B073	early summer	F	10	1.96	23	70	0	2	328	0.51	0.09
2018_F_FESRUB_B073_b	*Festuca rubra*	FR	grass	B073	early summer	F	10	1.96	25	35	1	10	331	0.55	0.09
2018_F_FESRUB_C114_a	*Festuca rubra*	FR	grass	C114	early summer	F	8	1.05	22	35	10	5	278	0.31	0.05
2018_F_FESRUB_C114_b	*Festuca rubra*	FR	grass	C114	early summer	F	8	1.05	37	35	10	30	361	0.58	0.10
2018_F_GERPRA_A018_a	*Geranium pratense*	GP	herb	A018	early summer	F	12	2.09	23	35	0	10	532	5.16	0.82
2018_F_GERPRA_A018_b	*Geranium pratense*	GP	herb	A018	early summer	F	12	2.09	30	35	20	15	428	5.00	0.83
2018_F_GERPRA_B081_a	*Geranium pratense*	GP	herb	B081	early summer	F	10	1.64	34	35	0	20	408	4.93	0.82
2018_F_GERPRA_B081_b	*Geranium pratense*	GP	herb	B081	early summer	F	10	1.64	36	35	5	15	506	5.28	0.85
2018_F_GERPRA_C109_a	*Geranium pratense*	GP	herb	C109	early summer	F	11	1.91	37	35	3	1	428	5.09	0.84
2018_F_GERPRA_C109_b	*Geranium pratense*	GP	herb	C109	early summer	F	11	1.91	34	35	10	20	464	4.92	0.80
2018_F_GERPRA_C121_a	*Geranium pratense*	GP	herb	C121	early summer	F	9	1.33	33	35	3	10	483	5.24	0.85
2018_F_GERPRA_C121_b	*Geranium pratense*	GP	herb	C121	early summer	F	9	1.33	27	60	0	3	462	5.08	0.83
2018_F_HOLLAN_A018_a	*Holcus lanatus*	HL	grass	A018	early summer	F	12	2.09	9	35	5	15	522	1.37	0.22
2018_F_HOLLAN_A018_b	*Holcus lanatus*	HL	grass	A018	early summer	F	12	2.09	12	40	5	20	502	1.35	0.22
2018_F_HOLLAN_A035_a	*Holcus lanatus*	HL	grass	A035	early summer	F	10	1.61	19	40	0	30	502	2.58	0.41
2018_F_HOLLAN_A035_b	*Holcus lanatus*	HL	grass	A035	early summer	F	10	1.61	17	65	0	30	509	1.11	0.18
2018_F_HOLLAN_A040_a	*Holcus lanatus*	HL	grass	A040	early summer	F	7	1.33	8	70	5	20	543	3.19	0.51
2018_F_HOLLAN_A040_b	*Holcus lanatus*	HL	grass	A040	early summer	F	7	1.33	10	65	10	30	440	2.74	0.45
2018_F_HOLLAN_B080_a	*Holcus lanatus*	HL	grass	B080	early summer	F	8	1.07	12	65	0	15	574	2.37	0.37
2018_F_HOLLAN_B080_b	*Holcus lanatus*	HL	grass	B080	early summer	F	8	1.07	14	35	0	15	529	0.72	0.12

Supplemental Table 4: continued

Sample ID					plot level trait				plant level trait				feature diversity		
sampleID	Species	2-letter code	FG	plot	Season	Season letter code	pRich [#]	pShan [scale]	height [cm]	BBCH [scale]	mechD [%]	pathD [%]	fRich	fShan	fEven
2018_F_KNAARV_A010_a	*Knautia arvensis*	KA	herb	A010	early summer	F	8	1.12	34	65	15	3	824	5.58	0.83
2018_F_KNAARV_A010_b	*Knautia arvensis*	KA	herb	A010	early summer	F	8	1.12	49	65	3	10	778	5.49	0.83
2018_F_KNAARV_A020_a	*Knautia arvensis*	KA	herb	A020	early summer	F	10	1.29	39	35	5	0	709	5.51	0.84
2018_F_KNAARV_A020_b	*Knautia arvensis*	KA	herb	A020	early summer	F	10	1.29	53	55	0	3	716	5.43	0.83
2018_F_KNAARV_A042_a	*Knautia arvensis*	KA	herb	A042	early summer	F	6	0.77	34	65	5	20	684	5.36	0.82
2018_F_KNAARV_A042_b	*Knautia arvensis*	KA	herb	A042	early summer	F	6	0.77	40	30	2	10	655	5.32	0.82
2018_F_KNAARV_B073_a	*Knautia arvensis*	KA	herb	B073	early summer	F	10	1.96	28	35	0	0	737	5.57	0.84
2018_F_KNAARV_B073_b	*Knautia arvensis*	KA	herb	B073	early summer	F	10	1.96	33	35	10	15	692	5.38	0.82
2018_F_LEUVUL_A018_a	*Leucanthemum vulgare*	LV	herb	A018	early summer	F	12	2.09	12	40	0	15	688	5.20	0.80
2018_F_LEUVUL_A018_b	*Leucanthemum vulgare*	LV	herb	A018	early summer	F	12	2.09	5	70	5	15	625	5.38	0.84
2018_F_LEUVUL_B048_a	*Leucanthemum vulgare*	LV	herb	B048	early summer	F	8	1.33	10	35	0	2	553	5.28	0.84
2018_F_LEUVUL_B048_b	*Leucanthemum vulgare*	LV	herb	B048	early summer	F	8	1.33	10	35	5	2	592	5.21	0.82
2018_F_LEUVUL_B051_a	*Leucanthemum vulgare*	LV	herb	B051	early summer	F	6	1.47	7	35	0	2	592	5.16	0.81
2018_F_LEUVUL_B051_b	*Leucanthemum vulgare*	LV	herb	B051	early summer	F	6	1.47	6	65	0	0	613	5.27	0.82
2018_F_LEUVUL_B073_a	*Leucanthemum vulgare*	LV	herb	B073	early summer	F	10	1.96	7	35	0	2	564	5.17	0.82
2018_F_LEUVUL_B073_b	*Leucanthemum vulgare*	LV	herb	B073	early summer	F	10	1.96	7	65	0	3	582	5.27	0.83
2018_F_LEUVUL_B085_a	*Leucanthemum vulgare*	LV	herb	B085	early summer	F	7	1.22	7	35	2	5	593	5.11	0.80
2018_F_LEUVUL_B085_b	*Leucanthemum vulgare*	LV	herb	B085	early summer	F	7	1.22	7	35	2	5	712	5.60	0.85
2018_F_LEUVUL_B090_a	*Leucanthemum vulgare*	LV	herb	B090	early summer	F	5	0.84	12	35	0	5	703	5.62	0.86
2018_F_LEUVUL_B090_b	*Leucanthemum vulgare*	LV	herb	B090	early summer	F	5	0.84	7	35	2	5	625	5.20	0.81
2018_F_LEUVUL_C135_a	*Leucanthemum vulgare*	LV	herb	C135	early summer	F	7	1.17	11	35	0	0	617	5.18	0.81
2018_F_LEUVUL_C135_b	*Leucanthemum vulgare*	LV	herb	C135	early summer	F	7	1.17	7	35	0	0	560	5.17	0.82
2018_F_LEUVUL_C136_a	*Leucanthemum vulgare*	LV	herb	C136	early summer	F	8	1.60	13	35	0	0	412	4.53	0.75
2018_F_LEUVUL_C136_b	*Leucanthemum vulgare*	LV	herb	C136	early summer	F	8	1.60	8	70	0	0	615	5.36	0.84
2018_F_PHLPRA_A002_a	*Phleum pratense*	PP	grass	A002	early summer	F	8	1.62	20	40	10	15	583	3.22	0.51
2018_F_PHLPRA_A002_b	*Phleum pratense*	PP	grass	A002	early summer	F	8	1.62	16	65	5	10	487	1.02	0.16
2018_F_PHLPRA_A018_a	*Phleum pratense*	PP	grass	A018	early summer	F	12	2.09	16	65	2	10	486	2.39	0.39
2018_F_PHLPRA_A018_b	*Phleum pratense*	PP	grass	A018	early summer	F	12	2.09	16	35	0	5	486	2.70	0.44
2018_F_PHLPRA_A045_a	*Phleum pratense*	PP	grass	A045	early summer	F	13	2.16	42	40	5	15	534	1.35	0.21
2018_F_PHLPRA_A045_b	*Phleum pratense*	PP	grass	A045	early summer	F	13	2.16	21	35	0	25	541	0.93	0.15
2018_F_PHLPRA_A046_a	*Phleum pratense*	PP	grass	A046	early summer	F	10	1.32	22	35	5	15	532	1.84	0.29
2018_F_PHLPRA_A046_b	*Phleum pratense*	PP	grass	A046	early summer	F	10	1.32	21	35	5	20	580	1.85	0.29
2018_F_PHLPRA_B051_a	*Phleum pratense*	PP	grass	B051	early summer	F	6	1.47	16	65	0	5	481	1.85	0.30
2018_F_PHLPRA_B051_b	*Phleum pratense*	PP	grass	B051	early summer	F	6	1.47	22	65	0	15	462	2.90	0.47
2018_F_PHLPRA_B059_a	*Phleum pratense*	PP	grass	B059	early summer	F	10	2.27	17	35	5	17	394	0.25	0.04
2018_F_PHLPRA_B059_b	*Phleum pratense*	PP	grass	B059	early summer	F	10	2.27	15	35	2	15	405	0.18	0.03
2018_F_PHLPRA_B073_b	*Phleum pratense*	PP	grass	B073	early summer	F	10	1.96	18	35	2	20	569	1.37	0.22
2018_F_PHLPRA_C133_a	*Phleum pratense*	PP	grass	C133	early summer	F	8	1.89	52	50	2	20	553	1.60	0.25
2018_F_PHLPRA_C133_b	*Phleum pratense*	PP	grass	C133	early summer	F	8	1.89	37	35	5	30	526	1.72	0.27
2018_F_PLALAN_A009_a	*Plantago lanceolata*	PL	herb	A009	early summer	F	8	1.68	31	60	15	5	739	5.60	0.85
2018_F_PLALAN_A009_b	*Plantago lanceolata*	PL	herb	A009	early summer	F	8	1.68	31	65	20	5	849	5.88	0.87
2018_F_PLALAN_A011_a	*Plantago lanceolata*	PL	herb	A011	early summer	F	9	1.29	28	55	5	10	808	5.53	0.83
2018_F_PLALAN_A011_b	*Plantago lanceolata*	PL	herb	A011	early summer	F	9	1.29	31	55	10	10	642	5.26	0.81
2018_F_PLALAN_A018_a	*Plantago lanceolata*	PL	herb	A018	early summer	F	12	2.09	25	35	10	10	709	5.54	0.84
2018_F_PLALAN_A018_b	*Plantago lanceolata*	PL	herb	A018	early summer	F	12	2.09	18	65	20	10	668	5.41	0.83
2018_F_PLALAN_A043_a	*Plantago lanceolata*	PL	herb	A043	early summer	F	7	0.82	30	30	10	15	781	5.72	0.86
2018_F_PLALAN_A043_b	*Plantago lanceolata*	PL	herb	A043	early summer	F	7	0.82	23	40	15	20	748	5.60	0.85
2018_F_PLALAN_B054_a	*Plantago lanceolata*	PL	herb	B054	early summer	F	5	1.01	23	65	20	20	734	5.57	0.84
2018_F_PLALAN_B054_b	*Plantago lanceolata*	PL	herb	B054	early summer	F	5	1.01	21	35	15	20	771	5.58	0.84
2018_F_PLALAN_B057_a	*Plantago lanceolata*	PL	herb	B057	early summer	F	9	1.83	29	35	20	5	688	5.47	0.84
2018_F_PLALAN_B057_b	*Plantago lanceolata*	PL	herb	B057	early summer	F	9	1.83	29	35	15	15	787	5.63	0.84
2018_F_PLALAN_B073_a	*Plantago lanceolata*	PL	herb	B073	early summer	F	10	1.96	20	35	5	10	739	5.76	0.87
2018_F_PLALAN_B073_b	*Plantago lanceolata*	PL	herb	B073	early summer	F	10	1.96	21	35	5	10	726	5.66	0.86
2018_F_PLALAN_C115_a	*Plantago lanceolata*	PL	herb	C115	early summer	F	11	1.95	34	65	20	30	730	5.50	0.83
2018_F_PLALAN_C115_b	*Plantago lanceolata*	PL	herb	C115	early summer	F	11	1.95	23	35	15	15	849	5.73	0.85
2018_F_POAPRA_A026_a	*Poa pratensis*	PA	grass	A026	early summer	F	9	1.05	17	65	0	20	375	0.29	0.05
2018_F_POAPRA_A026_b	*Poa pratensis*	PA	grass	A026	early summer	F	9	1.05	20	50	3	5	388	0.19	0.03
2018_F_POAPRA_B073_a	*Poa pratensis*	PA	grass	B073	early summer	F	10	1.96	19	35	0	2	502	0.57	0.09
2018_F_POAPRA_B075_a	*Poa pratensis*	PA	grass	B075	early summer	F	6	1.50	19	40	5	5	372	0.30	0.05
2018_F_POAPRA_B075_b	*Poa pratensis*	PA	grass	B075	early summer	F	6	1.50	10	65	2	20	274	4.69	0.84
2018_F_POAPRA_C131_a	*Poa pratensis*	PA	grass	C131	early summer	F	6	0.43	17	55	10	20	329	0.51	0.09
2018_F_POAPRA_C131_b	*Poa pratensis*	PA	grass	C131	early summer	F	6	0.43	26	40	10	30	366	0.51	0.09
2018_F_RANACR_A003_a	*Ranunculus acris*	RA	herb	A003	early summer	F	12	2.09	17	40	3	3	470	0.81	0.13
2018_F_RANACR_A003_b	*Ranunculus acris*	RA	herb	A003	early summer	F	12	2.09	22	40	2	2	466	1.00	0.16
2018_F_RANACR_A016_a	*Ranunculus acris*	RA	herb	A016	early summer	F	16	2.08	10	35	0	10	429	0.86	0.14
2018_F_RANACR_A016_b	*Ranunculus acris*	RA	herb	A016	early summer	F	16	2.08	15	40	0	5	463	1.13	0.18
2018_F_RANACR_A018_a	*Ranunculus acris*	RA	herb	A018	early summer	F	12	2.09	12	65	3	0	523	0.94	0.15
2018_F_RANACR_A018_b	*Ranunculus acris*	RA	herb	A018	early summer	F	12	2.09	8	35	2	0	426	1.51	0.25
2018_F_RANACR_B071_a	*Ranunculus acris*	RA	herb	B071	early summer	F	7	1.45	25	35	4	5	382	0.88	0.15
2018_F_RANACR_B071_b	*Ranunculus acris*	RA	herb	B071	early summer	F	7	1.45	21	50	2	10	461	1.52	0.25

Supplemental Table 4: continued

Sample ID							plot level trait		plant level trait				feature diversity		
sampleID	Species	2-letter code	FG	plot	Season	Season letter code	pRich [#]	pShan [scale]	height [cm]	BBCH [scale]	mechD [%]	pathD [%]	fRich	fShan	fEven
2018_G_ANTODO_A018_a	*Anthoxanthum odoratum*	AO	grass	A018	late summer	G	15	2.27	6	40	0	0	512	5.21	0.84
2018_G_ANTODO_A018_b	*Anthoxanthum odoratum*	AO	grass	A018	late summer	G	15	2.27	10	40	0	30	652	5.38	0.83
2018_G_ANTODO_A044_a	*Anthoxanthum odoratum*	AO	grass	A044	late summer	G	9	1.48	7	65	10	30	697	5.51	0.84
2018_G_ANTODO_A044_b	*Anthoxanthum odoratum*	AO	grass	A044	late summer	G	9	1.48	9	60	10	15	562	4.64	0.73
2018_G_ANTODO_B060_a	*Anthoxanthum odoratum*	AO	grass	B060	late summer	G	15	2.03	14	30	5	30	667	4.75	0.73
2018_G_ANTODO_B060_b	*Anthoxanthum odoratum*	AO	grass	B060	late summer	G	15	2.03	9	40	5	30	682	5.38	0.82
2018_G_ANTODO_C115_a	*Anthoxanthum odoratum*	AO	grass	C115	late summer	G	15	2.01	12	70	3	15	633	5.47	0.85
2018_G_ANTODO_C115_b	*Anthoxanthum odoratum*	AO	grass	C115	late summer	G	15	2.01	6	70	50	20	731	5.52	0.84
2018_G_AVEPUB_B073_a	*Avenula pubescens*	AP	grass	B073	late summer	G	10	2.03	11	70	3	10	580	5.32	0.84
2018_G_AVEPUB_B073_b	*Avenula pubescens*	AP	grass	B073	late summer	G	10	2.03	13	35	5	20	608	5.36	0.84
2018_G_AVEPUB_B092_a	*Avenula pubescens*	AP	grass	B092	late summer	G	15	1.39	11	40	0	25	583	5.42	0.85
2018_G_AVEPUB_B092_b	*Avenula pubescens*	AP	grass	B092	late summer	G	15	1.39	17	40	3	20	482	5.25	0.85
2018_G_AVEPUB_C103_a	*Avenula pubescens*	AP	grass	C103	late summer	G	10	1.37	14	40	0	30	471	5.27	0.86
2018_G_AVEPUB_C103_b	*Avenula pubescens*	AP	grass	C103	late summer	G	10	1.37	22	40	2	25	537	5.49	0.87
2018_G_AVEPUB_C110_a	*Avenula pubescens*	AP	grass	C110	late summer	G	20	2.25	15	40	0	20	570	3.99	0.63
2018_G_AVEPUB_C110_b	*Avenula pubescens*	AP	grass	C110	late summer	G	20	2.25	20	35	0	30	534	5.32	0.85
2018_G_CENJAC_A013_a	*Centaurea jacea*	CJ	herb	A013	late summer	G	14	1.64	40	40	3	5	706	5.09	0.78
2018_G_CENJAC_A013_b	*Centaurea jacea*	CJ	herb	A013	late summer	G	14	1.64	32	25	2	5	698	5.15	0.79
2018_G_CENJAC_A027_a	*Centaurea jacea*	CJ	herb	A027	late summer	G	16	1.99	30	75	0	15	717	5.14	0.78
2018_G_CENJAC_A027_b	*Centaurea jacea*	CJ	herb	A027	late summer	G	16	1.99	31	35	2	20	717	5.02	0.76
2018_G_CENJAC_A030_a	*Centaurea jacea*	CJ	herb	A030	late summer	G	7	1.36	11	35	5	5	539	4.73	0.75
2018_G_CENJAC_A030_b	*Centaurea jacea*	CJ	herb	A030	late summer	G	7	1.36	17	35	5	10	565	4.87	0.77
2018_G_CENJAC_B073_a	*Centaurea jacea*	CJ	herb	B073	late summer	G	10	2.03	20	40	0	2	634	4.93	0.76
2018_G_CENJAC_B073_b	*Centaurea jacea*	CJ	herb	B073	late summer	G	10	2.03	21	70	3	5	621	4.86	0.76
2018_G_DACGLO_A005_a	*Dactylis glomerata*	DG	grass	A005	late summer	G	13	2.02	40	70	5	30	510	2.13	0.34
2018_G_DACGLO_A005_b	*Dactylis glomerata*	DG	grass	A005	late summer	G	13	2.02	22	65	5	20	393	4.97	0.83
2018_G_DACGLO_A018_a	*Dactylis glomerata*	DG	grass	A018	late summer	G	15	2.27	16	40	0	10	415	3.05	0.51
2018_G_DACGLO_A018_b	*Dactylis glomerata*	DG	grass	A018	late summer	G	15	2.27	27	35	0	20	394	1.90	0.32
2018_G_DACGLO_C097_a	*Dactylis glomerata*	DG	grass	C097	late summer	G	18	1.54	24	40	0	15	388	5.12	0.86
2018_G_DACGLO_C097_b	*Dactylis glomerata*	DG	grass	C097	late summer	G	18	1.54	22	40	0	25	398	2.14	0.36
2018_G_DACGLO_C137_a	*Dactylis glomerata*	DG	grass	C137	late summer	G	7	1.19	22	40	5	15	481	1.24	0.20
2018_G_DACGLO_C137_b	*Dactylis glomerata*	DG	grass	C137	late summer	G	7	1.19	26	40	0	10	485	2.43	0.39
2018_G_FESRUB_B064_a	*Festuca rubra*	FR	grass	B064	late summer	G	15	0.93	21	40	5	15	367	1.08	0.18
2018_G_FESRUB_B064_b	*Festuca rubra*	FR	grass	B064	late summer	G	15	0.93	15	70	5	15	366	1.23	0.21
2018_G_FESRUB_B067_a	*Festuca rubra*	FR	grass	B067	late summer	G	12	1.69	24	40	10	30	400	0.42	0.07
2018_G_FESRUB_B067_b	*Festuca rubra*	FR	grass	B067	late summer	G	12	1.69	19	40	0	15	356	0.44	0.08
2018_G_FESRUB_B073_a	*Festuca rubra*	FR	grass	B073	late summer	G	10	2.03	18	65	3	5	362	0.25	0.04
2018_G_FESRUB_B073_b	*Festuca rubra*	FR	grass	B073	late summer	G	10	2.03	21	35	2	5	357	0.75	0.13
2018_G_FESRUB_C114_a	*Festuca rubra*	FR	grass	C114	late summer	G	10	1.06	28	35	5	30	359	0.44	0.07
2018_G_FESRUB_C114_b	*Festuca rubra*	FR	grass	C114	late summer	G	10	1.06	26	75	5	30	331	0.28	0.05
2018_G_GERPRA_A018_a	*Geranium pratense*	GP	herb	A018	late summer	G	15	2.27	19	35	3	8	467	5.16	0.84
2018_G_GERPRA_A018_b	*Geranium pratense*	GP	herb	A018	late summer	G	15	2.27	31	40	5	20	455	4.97	0.81
2018_G_GERPRA_B081_a	*Geranium pratense*	GP	herb	B081	late summer	G	14	1.84	24	40	5	25	512	5.23	0.84
2018_G_GERPRA_B081_b	*Geranium pratense*	GP	herb	B081	late summer	G	14	1.84	25	60	10	30	537	5.13	0.82
2018_G_GERPRA_C109_a	*Geranium pratense*	GP	herb	C109	late summer	G	14	1.89	22	40	0	15	493	5.18	0.84
2018_G_GERPRA_C109_b	*Geranium pratense*	GP	herb	C109	late summer	G	14	1.89	24	40	0	20	501	5.15	0.83
2018_G_GERPRA_C121_a	*Geranium pratense*	GP	herb	C121	late summer	G	11	1.52	46	35	5	15	511	5.43	0.87
2018_G_GERPRA_C121_b	*Geranium pratense*	GP	herb	C121	late summer	G	11	1.52	38	65	0	15	541	5.45	0.87
2018_G_HOLLAN_A018_a	*Holcus lanatus*	HL	grass	A018	late summer	G	15	2.27	13	40	0	15	513	2.87	0.46
2018_G_HOLLAN_A018_b	*Holcus lanatus*	HL	grass	A018	late summer	G	15	2.27	7	40	0	5	480	2.34	0.38
2018_G_HOLLAN_A035_a	*Holcus lanatus*	HL	grass	A035	late summer	G	14	1.96	12	35	2	15	601	4.32	0.68
2018_G_HOLLAN_A035_b	*Holcus lanatus*	HL	grass	A035	late summer	G	14	1.96	11	75	2	15	470	3.17	0.51
2018_G_HOLLAN_A040_a	*Holcus lanatus*	HL	grass	A040	late summer	G	20	2.64	13	75	0	10	448	1.23	0.20
2018_G_HOLLAN_A040_b	*Holcus lanatus*	HL	grass	A040	late summer	G	20	2.64	12	75	5	15	441	2.07	0.34
2018_G_HOLLAN_B080_a	*Holcus lanatus*	HL	grass	B080	late summer	G	13	1.21	9	60	10	35	471	2.95	0.48
2018_G_HOLLAN_B080_b	*Holcus lanatus*	HL	grass	B080	late summer	G	13	1.21	9	40	5	30	492	2.84	0.46

Supplemental Table 4: continued

Sample ID					plot level trait				plant level trait				feature diversity		
sampleID	Species	2-letter code	FG	plot	Season	Season letter code	pRich [#]	pShan [scale]	height [cm]	BBCH [scale]	mechD [%]	pathD [%]	fRich	fShan	fEven
2018_G_KNAARV_A010_a	*Knautia arvensis*	KA	herb	A010	late summer	G	10	1.74	36	65	10	15	876	5.74	0.85
2018_G_KNAARV_A010_b	*Knautia arvensis*	KA	herb	A010	late summer	G	10	1.74	47	65	3	20	746	5.55	0.84
2018_G_KNAARV_A020_a	*Knautia arvensis*	KA	herb	A020	late summer	G	9	1.40	36	35	5	20	798	5.54	0.83
2018_G_KNAARV_A020_b	*Knautia arvensis*	KA	herb	A020	late summer	G	9	1.40	28	70	5	15	796	5.62	0.84
2018_G_KNAARV_A042_a	*Knautia arvensis*	KA	herb	A042	late summer	G	9	0.98	44	70	5	20	800	5.67	0.85
2018_G_KNAARV_A042_b	*Knautia arvensis*	KA	herb	A042	late summer	G	9	0.98	34	35	2	20	789	5.72	0.86
2018_G_KNAARV_B073_a	*Knautia arvensis*	KA	herb	B073	late summer	G	10	2.03	36	35	5	20	784	5.58	0.84
2018_G_KNAARV_B073_b	*Knautia arvensis*	KA	herb	B073	late summer	G	10	2.03	43	40	5	10	812	5.59	0.83
2018_G_LEUVUL_A018_a	*Leucanthemum vulgare*	LV	herb	A018	late summer	G	15	2.27	8	35	0	2	503	4.94	0.79
2018_G_LEUVUL_A018_b	*Leucanthemum vulgare*	LV	herb	A018	late summer	G	15	2.27	7	70	0	10	656	5.61	0.87
2018_G_LEUVUL_B048_a	*Leucanthemum vulgare*	LV	herb	B048	late summer	G	15	1.68	17	60	0	0	616	5.39	0.84
2018_G_LEUVUL_B048_b	*Leucanthemum vulgare*	LV	herb	B048	late summer	G	15	1.68	14	75	2	10	594	5.21	0.82
2018_G_LEUVUL_B051_a	*Leucanthemum vulgare*	LV	herb	B051	late summer	G	9	1.58	8	75	5	3	585	5.29	0.83
2018_G_LEUVUL_B051_b	*Leucanthemum vulgare*	LV	herb	B051	late summer	G	9	1.58	25	80	0	5	688	5.33	0.82
2018_G_LEUVUL_B073_a	*Leucanthemum vulgare*	LV	herb	B073	late summer	G	10	2.03	6	40	0	0	636	5.34	0.83
2018_G_LEUVUL_B073_b	*Leucanthemum vulgare*	LV	herb	B073	late summer	G	10	2.03	5	75	0	0	667	5.43	0.83
2018_G_LEUVUL_B085_a	*Leucanthemum vulgare*	LV	herb	B085	late summer	G	15	1.75	6	60	0	5	615	5.48	0.85
2018_G_LEUVUL_B085_b	*Leucanthemum vulgare*	LV	herb	B085	late summer	G	15	1.75	8	40	2	10	624	5.43	0.84
2018_G_LEUVUL_B090_a	*Leucanthemum vulgare*	LV	herb	B090	late summer	G	12	1.06	24	40	0	0	631	5.37	0.83
2018_G_LEUVUL_B090_b	*Leucanthemum vulgare*	LV	herb	B090	late summer	G	12	1.06	17	40	0	0	569	5.31	0.84
2018_G_LEUVUL_C135_a	*Leucanthemum vulgare*	LV	herb	C135	late summer	G	8	1.00	7	40	0	0	610	5.29	0.83
2018_G_LEUVUL_C135_b	*Leucanthemum vulgare*	LV	herb	C135	late summer	G	8	1.00	7	40	0	0	655	5.37	0.83
2018_G_LEUVUL_C136_a	*Leucanthemum vulgare*	LV	herb	C136	late summer	G	9	1.49	7	40	0	0	622	5.16	0.80
2018_G_LEUVUL_C136_b	*Leucanthemum vulgare*	LV	herb	C136	late summer	G	9	1.49	10	40	0	0	593	5.14	0.80
2018_G_PHLPRA_A002_a	*Phleum pratense*	PP	grass	A002	late summer	G	12	1.96	35	40	5	10	482	2.47	0.40
2018_G_PHLPRA_A002_b	*Phleum pratense*	PP	grass	A002	late summer	G	12	1.96	38	40	5	20	494	1.70	0.27
2018_G_PHLPRA_A018_a	*Phleum pratense*	PP	grass	A018	late summer	G	15	2.27	23	80	2	15	508	0.98	0.16
2018_G_PHLPRA_A018_b	*Phleum pratense*	PP	grass	A018	late summer	G	15	2.27	19	35	0	2	325	1.60	0.28
2018_G_PHLPRA_A045_a	*Phleum pratense*	PP	grass	A045	late summer	G	13	2.02	47	40	5	20	499	2.24	0.36
2018_G_PHLPRA_A045_b	*Phleum pratense*	PP	grass	A045	late summer	G	13	2.02	43	60	5	20	430	1.58	0.26
2018_G_PHLPRA_A046_a	*Phleum pratense*	PP	grass	A046	late summer	G	11	1.44	19	60	0	15	503	1.98	0.32
2018_G_PHLPRA_A046_b	*Phleum pratense*	PP	grass	A046	late summer	G	11	1.44	25	40	3	10	434	1.28	0.21
2018_G_PHLPRA_B051_a	*Phleum pratense*	PP	grass	B051	late summer	G	9	1.58	53	80	0	5	492	5.35	0.86
2018_G_PHLPRA_B051_b	*Phleum pratense*	PP	grass	B051	late summer	G	9	1.58	74	80	5	15	473	1.25	0.20
2018_G_PHLPRA_B059_a	*Phleum pratense*	PP	grass	B059	late summer	G	12	2.25	24	40	5	10	346	0.38	0.06
2018_G_PHLPRA_B059_b	*Phleum pratense*	PP	grass	B059	late summer	G	12	2.25	24	35	5	15	433	0.11	0.02
2018_G_PHLPRA_B073_b	*Phleum pratense*	PP	grass	B073	late summer	G	10	2.03	19	40	5	5	460	0.23	0.04
2018_G_PHLPRA_C133_a	*Phleum pratense*	PP	grass	C133	late summer	G	12	2.29	44	40	5	15	471	1.01	0.16
2018_G_PHLPRA_C133_b	*Phleum pratense*	PP	grass	C133	late summer	G	12	2.29	42	40	10	50	438	3.45	0.57
2018_G_PLALAN_A009_a	*Plantago lanceolata*	PL	herb	A009	late summer	G	19	2.32	27	65	5	15	800	5.73	0.86
2018_G_PLALAN_A009_b	*Plantago lanceolata*	PL	herb	A009	late summer	G	19	2.32	22	65	5	20	855	5.89	0.87
2018_G_PLALAN_A011_a	*Plantago lanceolata*	PL	herb	A011	late summer	G	14	1.80	20	65	3	10	756	5.73	0.86
2018_G_PLALAN_A011_b	*Plantago lanceolata*	PL	herb	A011	late summer	G	14	1.80	29	40	3	5	690	5.60	0.86
2018_G_PLALAN_A018_a	*Plantago lanceolata*	PL	herb	A018	late summer	G	15	2.27	20	40	15	10	698	5.47	0.83
2018_G_PLALAN_A018_b	*Plantago lanceolata*	PL	herb	A018	late summer	G	15	2.27	30	65	10	10	711	5.55	0.85
2018_G_PLALAN_A043_a	*Plantago lanceolata*	PL	herb	A043	late summer	G	14	1.68	25	35	5	20	702	5.59	0.85
2018_G_PLALAN_A043_b	*Plantago lanceolata*	PL	herb	A043	late summer	G	14	1.68	26	65	5	15	823	5.72	0.85
2018_G_PLALAN_B054_a	*Plantago lanceolata*	PL	herb	B054	late summer	G	11	1.73	16	75	5	15	828	5.54	0.82
2018_G_PLALAN_B054_b	*Plantago lanceolata*	PL	herb	B054	late summer	G	11	1.73	22	40	5	15	753	5.77	0.87
2018_G_PLALAN_B057_a	*Plantago lanceolata*	PL	herb	B057	late summer	G	10	1.86	26	40	10	20	798	5.86	0.88
2018_G_PLALAN_B057_b	*Plantago lanceolata*	PL	herb	B057	late summer	G	10	1.86	24	40	10	15	783	5.77	0.87
2018_G_PLALAN_B073_a	*Plantago lanceolata*	PL	herb	B073	late summer	G	10	2.03	23	40	10	5	805	5.80	0.87
2018_G_PLALAN_B073_b	*Plantago lanceolata*	PL	herb	B073	late summer	G	10	2.03	21	40	15	10	845	5.83	0.87
2018_G_PLALAN_C115_a	*Plantago lanceolata*	PL	herb	C115	late summer	G	15	2.01	24	70	15	30	745	5.63	0.85
2018_G_PLALAN_C115_b	*Plantago lanceolata*	PL	herb	C115	late summer	G	15	2.01	27	35	5	20	748	5.50	0.83
2018_G_POAPRA_A026_a	*Poa pratensis*	PA	grass	A026	late summer	G	14	1.29	23	70	5	30	352	0.34	0.06
2018_G_POAPRA_A026_b	*Poa pratensis*	PA	grass	A026	late summer	G	14	1.29	19	40	5	25	400	0.31	0.05
2018_G_POAPRA_B073_a	*Poa pratensis*	PA	grass	B073	late summer	G	10	2.03	18	35	5	10	404	0.52	0.09
2018_G_POAPRA_B073_b	*Poa pratensis*	PA	grass	B073	late summer	G	10	2.03	17	35	3	5	353	0.69	0.12
2018_G_POAPRA_B075_a	*Poa pratensis*	PA	grass	B075	late summer	G	7	1.58	27	35	0	15	369	0.35	0.06
2018_G_POAPRA_B075_b	*Poa pratensis*	PA	grass	B075	late summer	G	7	1.58	16	50	0	10	434	0.34	0.06
2018_G_POAPRA_C131_a	*Poa pratensis*	PA	grass	C131	late summer	G	11	1.10	20	55	5	15	346	0.42	0.07
2018_G_POAPRA_C131_b	*Poa pratensis*	PA	grass	C131	late summer	G	11	1.10	21	40	0	20	324	0.66	0.11
2018_G_RANACR_A003_a	*Ranunculus acris*	RA	herb	A003	late summer	G	20	2.34	16	35	5	20	478	1.41	0.23
2018_G_RANACR_A003_b	*Ranunculus acris*	RA	herb	A003	late summer	G	20	2.34	14	70	3	10	471	0.90	0.15
2018_G_RANACR_A016_a	*Ranunculus acris*	RA	herb	A016	late summer	G	18	2.54	14	35	3	15	465	1.32	0.21
2018_G_RANACR_A016_b	*Ranunculus acris*	RA	herb	A016	late summer	G	18	2.54	10	40	0	15	511	0.93	0.15
2018_G_RANACR_A018_a	*Ranunculus acris*	RA	herb	A018	late summer	G	15	2.27	11	65	0	20	467	0.97	0.16
2018_G_RANACR_B071_a	*Ranunculus acris*	RA	herb	B071	late summer	G	17	2.38	19	40	2	40	501	0.87	0.14
2018_G_RANACR_B071_b	*Ranunculus acris*	RA	herb	B071	late summer	G	17	2.38	15	75	2	25	493	0.84	0.14

Supplemental Table 4: continued

Sample ID					plot level trait				plant level trait				feature diversity		
sampleID	Species	2-letter code	FG	plot	Season	Season letter code	pRich [#]	pShan [scale]	height [cm]	BBCH [scale]	mechD [%]	pathD [%]	fRich	fShan	fEven
2018_H_ANTODO_A018_b	*Anthoxanthum odoratum*	AO	grass	A018	autumn	H	18	2.39	7	30	2	10	447	2.90	0.48
2018_H_ANTODO_A044_a	*Anthoxanthum odoratum*	AO	grass	A044	autumn	H	15	2.16	8	40	5	2	573	5.19	0.82
2018_H_ANTODO_A044_b	*Anthoxanthum odoratum*	AO	grass	A044	autumn	H	15	2.16	11	40	2	5	588	5.16	0.81
2018_H_ANTODO_B060_a	*Anthoxanthum odoratum*	AO	grass	B060	autumn	H	10	1.90	12	40	5	5	547	3.74	0.59
2018_H_ANTODO_B060_b	*Anthoxanthum odoratum*	AO	grass	B060	autumn	H	10	1.90	11	30	5	5	542	4.65	0.74
2018_H_ANTODO_C115_a	*Anthoxanthum odoratum*	AO	grass	C115	autumn	H	13	2.22	9	40	5	5	607	5.31	0.83
2018_H_ANTODO_C115_b	*Anthoxanthum odoratum*	AO	grass	C115	autumn	H	13	2.22	9	40	2	2	487	3.90	0.63
2018_H_AVEPUB_B073_a	*Avenula pubescens*	AP	grass	B073	autumn	H	12	2.08	18	40	3	8	517	5.18	0.83
2018_H_AVEPUB_B073_b	*Avenula pubescens*	AP	grass	B073	autumn	H	12	2.08	9	40	3	8	467	5.06	0.82
2018_H_AVEPUB_B092_a	*Avenula pubescens*	AP	grass	B092	autumn	H	13	1.46	12	40	3	5	560	5.26	0.83
2018_H_AVEPUB_B092_b	*Avenula pubescens*	AP	grass	B092	autumn	H	13	1.46	11	40	5	10	585	5.42	0.85
2018_H_AVEPUB_C103_a	*Avenula pubescens*	AP	grass	C103	autumn	H	12	1.89	14	40	5	10	543	5.43	0.86
2018_H_AVEPUB_C103_b	*Avenula pubescens*	AP	grass	C103	autumn	H	12	1.89	14	40	2	5	594	5.47	0.86
2018_H_AVEPUB_C110_a	*Avenula pubescens*	AP	grass	C110	autumn	H	16	2.27	14	40	5	10	534	5.21	0.83
2018_H_AVEPUB_C110_b	*Avenula pubescens*	AP	grass	C110	autumn	H	16	2.27	15	40	2	10	592	5.48	0.86
2018_H_CENJAC_A013_a	*Centaurea jacea*	CJ	herb	A013	autumn	H	20	2.31	31	40	2	2	642	5.11	0.79
2018_H_CENJAC_A013_b	*Centaurea jacea*	CJ	herb	A013	autumn	H	20	2.31	20	30	5	2	561	4.94	0.78
2018_H_CENJAC_A027_a	*Centaurea jacea*	CJ	herb	A027	autumn	H	16	2.25	30	30	2	2	684	5.01	0.77
2018_H_CENJAC_A027_b	*Centaurea jacea*	CJ	herb	A027	autumn	H	16	2.25	33	40	2	2	622	5.01	0.78
2018_H_CENJAC_A030_a	*Centaurea jacea*	CJ	herb	A030	autumn	H	9	1.47	12	40	5	5	604	4.94	0.77
2018_H_CENJAC_A030_b	*Centaurea jacea*	CJ	herb	A030	autumn	H	9	1.47	19	40	5	15	567	5.01	0.79
2018_H_CENJAC_B073_a	*Centaurea jacea*	CJ	herb	B073	autumn	H	12	2.08	8	30	5	5	565	5.02	0.79
2018_H_CENJAC_B073_b	*Centaurea jacea*	CJ	herb	B073	autumn	H	12	2.08	15	60	2	5	563	4.78	0.75
2018_H_DACGLO_A005_a	*Dactylis glomerata*	DG	grass	A005	autumn	H	17	2.26	25	30	5	10	336	1.90	0.33
2018_H_DACGLO_A005_b	*Dactylis glomerata*	DG	grass	A005	autumn	H	17	2.26	25	70	5	15	376	1.63	0.27
2018_H_DACGLO_A018_a	*Dactylis glomerata*	DG	grass	A018	autumn	H	18	2.39	17	40	3	10	368	2.12	0.36
2018_H_DACGLO_A018_b	*Dactylis glomerata*	DG	grass	A018	autumn	H	18	2.39	21	30	2	10	379	5.08	0.86
2018_H_DACGLO_C097_a	*Dactylis glomerata*	DG	grass	C097	autumn	H	17	2.10	22	40	5	15	349	4.74	0.81
2018_H_DACGLO_C097_b	*Dactylis glomerata*	DG	grass	C097	autumn	H	17	2.10	24	40	5	5	414	1.49	0.25
2018_H_DACGLO_C137_a	*Dactylis glomerata*	DG	grass	C137	autumn	H	11	1.74	17	40	5	10	473	2.31	0.38
2018_H_DACGLO_C137_b	*Dactylis glomerata*	DG	grass	C137	autumn	H	11	1.74	21	40	5	3	432	4.74	0.78
2018_H_FESRUB_B064_a	*Festuca rubra*	FR	grass	B064	autumn	H	13	1.67	12	30	10	0	330	0.99	0.17
2018_H_FESRUB_B064_b	*Festuca rubra*	FR	grass	B064	autumn	H	13	1.67	12	40	2	5	282	0.79	0.14
2018_H_FESRUB_B067_a	*Festuca rubra*	FR	grass	B067	autumn	H	15	1.93	10	40	5	5	379	0.33	0.05
2018_H_FESRUB_B067_b	*Festuca rubra*	FR	grass	B067	autumn	H	15	1.93	13	30	5	10	252	0.49	0.09
2018_H_FESRUB_B073_a	*Festuca rubra*	FR	grass	B073	autumn	H	12	2.08	4	60	5	5	319	0.56	0.10
2018_H_FESRUB_B073_b	*Festuca rubra*	FR	grass	B073	autumn	H	12	2.08	21	30	2	10	331	0.71	0.12
2018_H_FESRUB_C114_a	*Festuca rubra*	FR	grass	C114	autumn	H	11	1.18	11	40	3	5	366	0.46	0.08
2018_H_FESRUB_C114_b	*Festuca rubra*	FR	grass	C114	autumn	H	11	1.18	14	40	5	5	316	0.37	0.06
2018_H_GERPRA_A018_a	*Geranium pratense*	GP	herb	A018	autumn	H	18	2.39	8	40	2	5	438	4.99	0.82
2018_H_GERPRA_A018_b	*Geranium pratense*	GP	herb	A018	autumn	H	18	2.39	15	40	3	5	450	5.01	0.82
2018_H_GERPRA_B081_a	*Geranium pratense*	GP	herb	B081	autumn	H	16	2.35	13	40	5	15	506	5.27	0.85
2018_H_GERPRA_B081_b	*Geranium pratense*	GP	herb	B081	autumn	H	16	2.35	12	40	10	5	500	5.03	0.81
2018_H_GERPRA_C109_a	*Geranium pratense*	GP	herb	C109	autumn	H	13	1.98	15	40	2	10	457	4.94	0.81
2018_H_GERPRA_C109_b	*Geranium pratense*	GP	herb	C109	autumn	H	13	1.98	14	40	5	20	441	5.09	0.84
2018_H_GERPRA_C121_a	*Geranium pratense*	GP	herb	C121	autumn	H	12	2.05	12	40	2	15	465	5.19	0.85
2018_H_GERPRA_C121_b	*Geranium pratense*	GP	herb	C121	autumn	H	12	2.05	12	40	2	10	517	5.22	0.84
2018_H_HOLLAN_A018_a	*Holcus lanatus*	HL	grass	A018	autumn	H	18	2.39	12	40	5	15	423	1.56	0.26
2018_H_HOLLAN_A018_b	*Holcus lanatus*	HL	grass	A018	autumn	H	18	2.39	13	40	5	15	421	1.89	0.31
2018_H_HOLLAN_A035_a	*Holcus lanatus*	HL	grass	A035	autumn	H	18	2.54	12	40	3	5	433	2.43	0.40
2018_H_HOLLAN_A035_b	*Holcus lanatus*	HL	grass	A035	autumn	H	18	2.54	11	60	5	15	453	1.13	0.19
2018_H_HOLLAN_A040_a	*Holcus lanatus*	HL	grass	A040	autumn	H	17	2.52	12	70	2	5	381	1.94	0.33
2018_H_HOLLAN_A040_b	*Holcus lanatus*	HL	grass	A040	autumn	H	17	2.52	10	40	5	20	408	2.32	0.39
2018_H_HOLLAN_B080_a	*Holcus lanatus*	HL	grass	B080	autumn	H	13	1.88	7	40	0	5	335	1.66	0.29
2018_H_HOLLAN_B080_b	*Holcus lanatus*	HL	grass	B080	autumn	H	13	1.88	12	40	2	5	405	2.85	0.47

Supplemental Table 4: continued

Sample ID					plot level trait				plant level trait				feature diversity		
sampleID	Species	2-letter code	FG	plot	Season	Season letter code	pRich [#]	pShan [scale]	height [cm]	BBCH [scale]	mechD [%]	pathD [%]	fRich	fShan	fEven
2018_H_KNAARV_A010_a	*Knautia arvensis*	KA	herb	A010	autumn	H	11	1.97	22	40	5	10	650	5.32	0.82
2018_H_KNAARV_A010_b	*Knautia arvensis*	KA	herb	A010	autumn	H	11	1.97	27	50	5	5	556	5.11	0.81
2018_H_KNAARV_A020_a	*Knautia arvensis*	KA	herb	A020	autumn	H	11	1.62	44	40	5	15	638	5.33	0.82
2018_H_KNAARV_A020_b	*Knautia arvensis*	KA	herb	A020	autumn	H	11	1.62	26	50	10	10	592	5.19	0.81
2018_H_KNAARV_A042_a	*Knautia arvensis*	KA	herb	A042	autumn	H	12	1.50	24	40	5	10	587	5.20	0.82
2018_H_KNAARV_A042_b	*Knautia arvensis*	KA	herb	A042	autumn	H	12	1.50	23	30	10	20	653	5.33	0.82
2018_H_KNAARV_B073_a	*Knautia arvensis*	KA	herb	B073	autumn	H	12	2.08	28	40	5	10	622	5.25	0.82
2018_H_KNAARV_B073_b	*Knautia arvensis*	KA	herb	B073	autumn	H	12	2.08	8	30	5	20	705	5.44	0.83
2018_H_LEUVUL_A018_a	*Leucanthemum vulgare*	LV	herb	A018	autumn	H	18	2.39	5	30	0	2	469	5.01	0.82
2018_H_LEUVUL_A018_b	*Leucanthemum vulgare*	LV	herb	A018	autumn	H	18	2.39	4	40	3	5	540	5.22	0.83
2018_H_LEUVUL_B048_a	*Leucanthemum vulgare*	LV	herb	B048	autumn	H	20	2.18	6	40	2	3	550	5.34	0.85
2018_H_LEUVUL_B048_b	*Leucanthemum vulgare*	LV	herb	B048	autumn	H	20	2.18	11	40	0	3	421	4.92	0.81
2018_H_LEUVUL_B051_a	*Leucanthemum vulgare*	LV	herb	B051	autumn	H	17	2.20	10	80	5	0	564	5.37	0.85
2018_H_LEUVUL_B051_b	*Leucanthemum vulgare*	LV	herb	B051	autumn	H	17	2.20	7	30	0	0	538	5.37	0.85
2018_H_LEUVUL_B073_a	*Leucanthemum vulgare*	LV	herb	B073	autumn	H	12	2.08	5	30	3	3	545	5.33	0.85
2018_H_LEUVUL_B073_b	*Leucanthemum vulgare*	LV	herb	B073	autumn	H	12	2.08	5	40	5	5	583	5.37	0.84
2018_H_LEUVUL_B085_a	*Leucanthemum vulgare*	LV	herb	B085	autumn	H	20	2.30	5	40	2	10	538	5.28	0.84
2018_H_LEUVUL_B085_b	*Leucanthemum vulgare*	LV	herb	B085	autumn	H	20	2.30	7	40	0	2	450	5.10	0.83
2018_H_LEUVUL_B090_a	*Leucanthemum vulgare*	LV	herb	B090	autumn	H	12	1.59	7	40	0	3	610	5.47	0.85
2018_H_LEUVUL_B090_b	*Leucanthemum vulgare*	LV	herb	B090	autumn	H	12	1.59	10	40	0	0	529	5.18	0.83
2018_H_LEUVUL_C135_a	*Leucanthemum vulgare*	LV	herb	C135	autumn	H	12	1.72	6	40	2	5	571	5.23	0.82
2018_H_LEUVUL_C135_b	*Leucanthemum vulgare*	LV	herb	C135	autumn	H	12	1.72	7	40	2	10	500	5.06	0.81
2018_H_LEUVUL_C136_a	*Leucanthemum vulgare*	LV	herb	C136	autumn	H	11	1.61	6	40	2	5	470	5.03	0.82
2018_H_LEUVUL_C136_b	*Leucanthemum vulgare*	LV	herb	C136	autumn	H	11	1.61	9	40	5	15	579	5.11	0.80
2018_H_PHLPRA_A002_a	*Phleum pratense*	PP	grass	A002	autumn	H	14	1.76	13	40	5	20	512	5.21	0.84
2018_H_PHLPRA_A002_b	*Phleum pratense*	PP	grass	A002	autumn	H	14	1.76	12	40	5	20	414	2.15	0.36
2018_H_PHLPRA_A018_a	*Phleum pratense*	PP	grass	A018	autumn	H	18	2.39	13	40	5	5	403	4.85	0.81
2018_H_PHLPRA_A018_b	*Phleum pratense*	PP	grass	A018	autumn	H	18	2.39	8	40	3	5	331	2.22	0.38
2018_H_PHLPRA_A045_a	*Phleum pratense*	PP	grass	A045	autumn	H	17	1.94	14	30	10	20	528	2.92	0.47
2018_H_PHLPRA_A045_b	*Phleum pratense*	PP	grass	A045	autumn	H	17	1.94	13	40	2	5	321	2.30	0.40
2018_H_PHLPRA_A046_a	*Phleum pratense*	PP	grass	A046	autumn	H	18	1.85	16	40	0	2	454	1.97	0.32
2018_H_PHLPRA_A046_b	*Phleum pratense*	PP	grass	A046	autumn	H	18	1.85	13	40	5	10	429	2.68	0.44
2018_H_PHLPRA_B051_a	*Phleum pratense*	PP	grass	B051	autumn	H	17	2.20	14	30	0	2	444	3.01	0.49
2018_H_PHLPRA_B051_b	*Phleum pratense*	PP	grass	B051	autumn	H	17	2.20	36	40	5	20	409	3.23	0.54
2018_H_PHLPRA_B059_a	*Phleum pratense*	PP	grass	B059	autumn	H	12	1.45	19	40	5	10	276	0.63	0.11
2018_H_PHLPRA_B059_b	*Phleum pratense*	PP	grass	B059	autumn	H	12	1.45	18	40	3	10	304	0.44	0.08
2018_H_PHLPRA_B073_a	*Phleum pratense*	PP	grass	B073	autumn	H	12	2.08	10	40	3	3	482	0.38	0.06
2018_H_PHLPRA_B073_b	*Phleum pratense*	PP	grass	B073	autumn	H	12	2.08	15	35	5	5	353	0.47	0.08
2018_H_PHLPRA_C133_a	*Phleum pratense*	PP	grass	C133	autumn	H	14	2.02	12	40	2	15	243	0.91	0.17
2018_H_PHLPRA_C133_b	*Phleum pratense*	PP	grass	C133	autumn	H	14	2.02	18	40	3	10	417	1.47	0.24
2018_H_PLALAN_A009_a	*Plantago lanceolata*	PL	herb	A009	autumn	H	19	2.44	3	75	10	5	576	5.20	0.82
2018_H_PLALAN_A009_b	*Plantago lanceolata*	PL	herb	A009	autumn	H	19	2.44	5	40	30	5	695	5.47	0.84
2018_H_PLALAN_A011_a	*Plantago lanceolata*	PL	herb	A011	autumn	H	19	2.19	14	40	15	15	789	5.59	0.84
2018_H_PLALAN_A011_b	*Plantago lanceolata*	PL	herb	A011	autumn	H	19	2.19	13	40	15	10	621	5.16	0.80
2018_H_PLALAN_A018_a	*Plantago lanceolata*	PL	herb	A018	autumn	H	18	2.39	9	30	0	0	689	5.43	0.83
2018_H_PLALAN_A018_b	*Plantago lanceolata*	PL	herb	A018	autumn	H	18	2.39	8	60	5	10	635	5.48	0.85
2018_H_PLALAN_A043_a	*Plantago lanceolata*	PL	herb	A043	autumn	H	15	2.13	12	30	15	10	806	5.53	0.83
2018_H_PLALAN_A043_b	*Plantago lanceolata*	PL	herb	A043	autumn	H	15	2.13	12	40	10	10	685	5.20	0.80
2018_H_PLALAN_B054_a	*Plantago lanceolata*	PL	herb	B054	autumn	H	14	2.06	7	40	15	20	603	5.08	0.79
2018_H_PLALAN_B054_b	*Plantago lanceolata*	PL	herb	B054	autumn	H	14	2.06	8	40	15	10	697	5.17	0.79
2018_H_PLALAN_B057_a	*Plantago lanceolata*	PL	herb	B057	autumn	H	14	2.31	11	40	20	10	756	5.44	0.82
2018_H_PLALAN_B057_b	*Plantago lanceolata*	PL	herb	B057	autumn	H	14	2.31	10	40	25	5	755	5.52	0.83
2018_H_PLALAN_B073_a	*Plantago lanceolata*	PL	herb	B073	autumn	H	12	2.08	6	30	20	10	855	5.81	0.86
2018_H_PLALAN_B073_b	*Plantago lanceolata*	PL	herb	B073	autumn	H	12	2.08	9	40	15	10	828	5.67	0.84
2018_H_PLALAN_C115_a	*Plantago lanceolata*	PL	herb	C115	autumn	H	13	2.22	9	40	15	10	746	5.63	0.85
2018_H_PLALAN_C115_b	*Plantago lanceolata*	PL	herb	C115	autumn	H	13	2.22	12	40	20	5	665	5.22	0.80
2018_H_POAPRA_A026_a	*Poa pratensis*	PA	grass	A026	autumn	H	18	1.71	21	50	3	5	331	0.64	0.11
2018_H_POAPRA_A026_b	*Poa pratensis*	PA	grass	A026	autumn	H	18	1.71	14	30	3	3	342	0.80	0.14
2018_H_POAPRA_B073_a	*Poa pratensis*	PA	grass	B073	autumn	H	12	2.08	21	40	0	5	356	0.75	0.13
2018_H_POAPRA_B073_b	*Poa pratensis*	PA	grass	B073	autumn	H	12	2.08	16	40	0	0	373	0.33	0.06
2018_H_POAPRA_B075_a	*Poa pratensis*	PA	grass	B075	autumn	H	11	1.90	15	40	5	10	446	0.42	0.07
2018_H_POAPRA_B075_b	*Poa pratensis*	PA	grass	B075	autumn	H	11	1.90	13	40	10	5	380	0.94	0.16
2018_H_POAPRA_C131_a	*Poa pratensis*	PA	grass	C131	autumn	H	13	1.04	15	40	5	2	349	0.72	0.12
2018_H_POAPRA_C131_b	*Poa pratensis*	PA	grass	C131	autumn	H	13	1.04	15	40	15	5	398	1.04	0.17
2018_H_RANACR_A003_a	*Ranunculus acris*	RA	herb	A003	autumn	H	13	1.70	14	40	3	10	428	0.99	0.16
2018_H_RANACR_A003_b	*Ranunculus acris*	RA	herb	A003	autumn	H	13	1.70	16	30	3	10	465	1.06	0.17
2018_H_RANACR_A016_a	*Ranunculus acris*	RA	herb	A016	autumn	H	15	2.18	12	30	3	30	412	1.20	0.20
2018_H_RANACR_A016_b	*Ranunculus acris*	RA	herb	A016	autumn	H	15	2.18	9	40	3	35	418	1.40	0.23
2018_H_RANACR_A018_a	*Ranunculus acris*	RA	herb	A018	autumn	H	18	2.39	12	60	10	10	453	1.07	0.18
2018_H_RANACR_A018_b	*Ranunculus acris*	RA	herb	A018	autumn	H	18	2.39	8	30	3	3	434	0.84	0.14
2018_H_RANACR_B071_a	*Ranunculus acris*	RA	herb	B071	autumn	H	14	2.20	12	30	5	20	539	0.80	0.13
2018_H_RANACR_B071_b	*Ranunculus acris*	RA	herb	B071	autumn	H	14	2.20	10	30	10	20	482	0.69	0.11

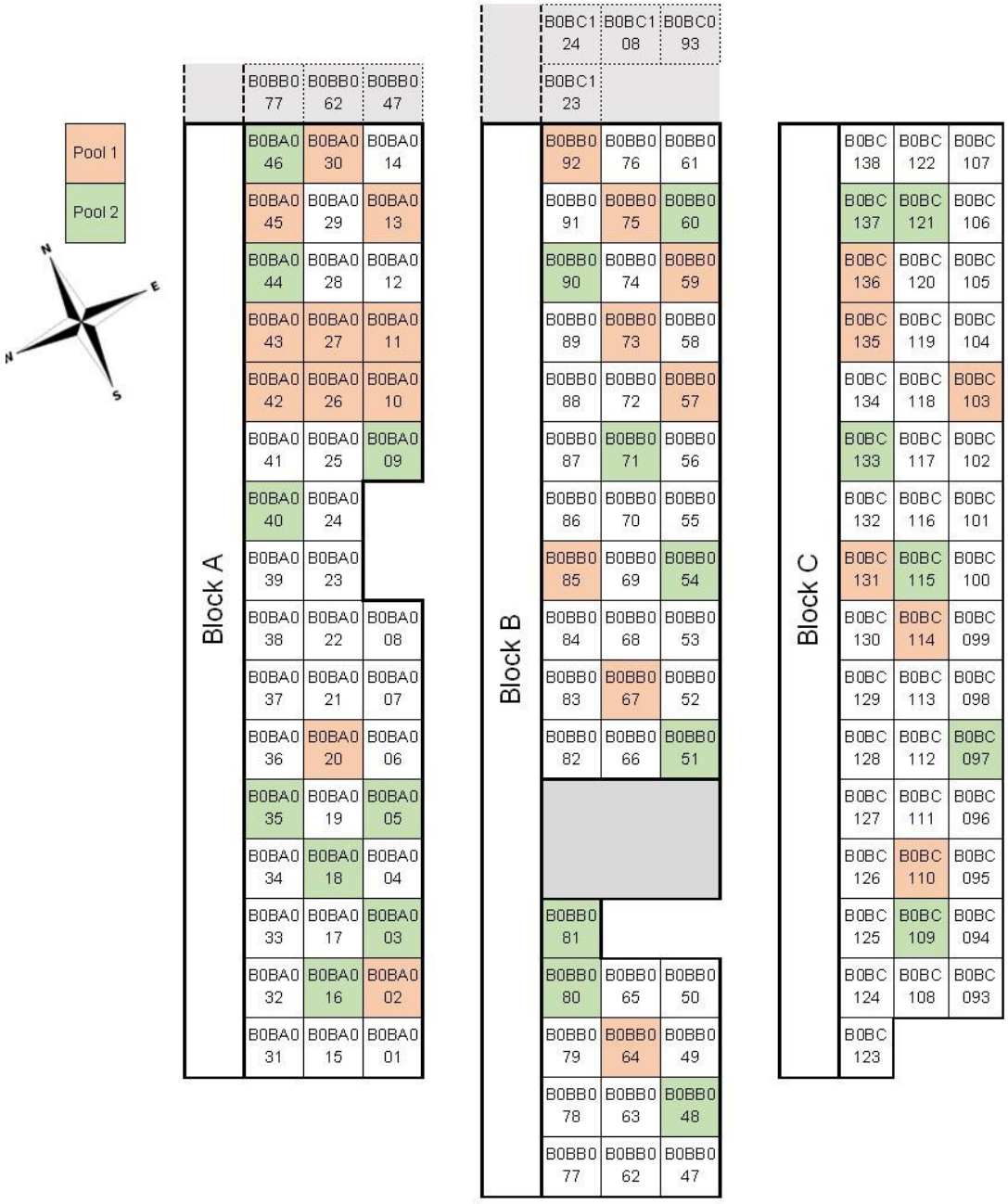
Supplement Table 5: Linear models were calculated for all 504 samples across thirteen species (all samples) and species specific for all species separately. We calculated the models using the prediction variables (preV): plant species richness (pRich) and plant species Shannon (pShan) and the response variables (resV): LC-MS feature richness (fRich), LC-MS feature Shannon (fShan) and LC-MS feature Evenness (fEven). We use the model intercept and coefficient (coef) to show the direction of the model that indicates the relationship. To compare the models' predictive power, we calculated the coefficient of determination (R^2), root means square error (RMSE), and confidence intervals upper bound (CI_ub) and lower bound (CI_lb) and included the p-value as reference.

		prediction variable (preV)	plant species richness (pRich)						plant species Shannon (pShan)					
resV	species	2-letter code	intercept	coef	R^2	RMSE	CI_lb	CI_ub	p value	intercept	coef	R^2	RMSE	CI_lb	CI_ub	p value
fRich	all samples	all	581.60	-3.18	0.28	135.90	522.45	555.59	0.03	573.56	-18.44	0.25	136.33	522.50	555.54	0.21
fRich	*Anthoxanthum odoratum*	AO	727.25	-10.65	0.37	83.06	529.41	624.04	0.03	885.90	-150.10	0.41	79.89	533.41	620.03	0.00
fRich	*Avenula pubescens*	AP	574.02	-3.08	0.45	62.51	498.25	565.44	0.30	627.18	-50.23	0.51	60.01	498.46	565.23	0.13
fRich	*Centaurea jacea*	CJ	586.92	4.78	0.57	65.53	606.34	683.35	0.15	577.00	37.80	0.55	68.06	605.51	684.18	0.30
fRich	*Dactylis glomerata*	DG	466.63	-3.08	0.41	41.66	401.29	444.52	0.08	511.45	-45.93	0.52	42.29	402.14	443.67	0.02
fRich	*Festuca rubra*	FR	327.30	1.72	0.46	41.93	324.86	370.08	0.63	355.02	-4.73	0.51	40.49	324.82	370.12	0.79
fRich	*Geranium pratense*	GP	498.27	-1.09	0.49	34.57	465.76	501.37	0.58	533.20	-25.90	0.54	33.25	466.09	501.04	0.23
fRich	*Holcus lanatus*	HL	568.08	-7.02	0.64	56.34	437.06	493.75	0.01	576.10	-55.62	0.51	57.21	436.81	494.00	0.02
fRich	*Knautia arvensis*	KA	711.68	1.47	0.48	80.65	689.73	766.65	0.73	754.21	-16.31	0.56	81.60	688.92	767.46	0.63
fRich	*Leucanthemum vulgare*	LV	652.27	-5.05	0.22	60.62	567.18	608.92	0.01	684.86	-55.27	0.25	60.30	567.43	608.67	0.00
fRich	*Phleum pratense*	PP	457.73	0.59	0.19	79.31	438.69	492.99	0.82	465.46	0.19	0.16	79.01	438.51	493.16	1.00
fRich	*Plantago lanceolata*	PL	809.54	-5.09	0.23	62.16	716.24	759.95	0.01	819.50	-41.30	0.26	64.16	716.38	759.81	0.05
fRich	*Poa pratensis*	PA	385.63	-2.09	0.65	45.03	330.60	390.27	0.43	366.43	-3.79	0.49	46.30	330.36	390.50	0.86
fRich	*Ranunculus acris*	RA	461.83	0.16	0.50	37.05	443.12	485.84	0.93	387.73	34.92	0.52	36.75	443.86	485.11	0.20
fShan	all samples	all	4.16366	-0.05	0.28	2.05	3.29	3.78	0.04	4.16	-0.33	0.27	2.06	3.29	3.78	0.13
fShan	*Anthoxanthum odoratum*	AO	5.74	-0.13	0.53	1.44	3.13	4.68	0.09	5.72	-0.88	0.54	1.52	3.13	4.69	0.32
fShan	*Avenula pubescens*	AP	5.86	-0.07	0.43	0.88	4.42	5.37	0.10	6.79	-1.00	0.59	0.80	4.43	5.37	0.04
fShan	*Centaurea jacea*	CJ	4.90	0.01	0.69	0.12	4.95	5.10	0.11	4.88	0.08	0.64	0.13	4.95	5.10	0.23
fShan	*Dactylis glomerata*	DG	2.97	-0.02	0.56	1.32	2.04	3.43	0.76	4.15	-0.73	0.68	1.29	2.06	3.41	0.26
fShan	*Festuca rubra*	FR	0.00	0.04	0.74	0.23	0.40	0.63	0.02	0.54	-0.01	0.58	0.23	0.39	0.64	0.88
fShan	*Geranium pratense*	GP	5.25	-0.01	0.48	0.13	5.08	5.21	0.26	5.45	-0.16	0.53	0.12	5.08	5.20	0.04
fShan	*Holcus lanatus*	HL	2.82	-0.06	0.42	0.76	1.55	2.40	0.16	2.75	-0.39	0.48	0.78	1.54	2.40	0.25
fShan	*Knautia arvensis*	KA	5.43	0.00	0.50	0.16	5.39	5.55	0.64	5.54	-0.04	0.55	0.16	5.39	5.55	0.54
fShan	*Leucanthemum vulgare*	LV	5.26	0.00	0.26	0.16	5.21	5.33	0.90	5.32	-0.03	0.17	0.16	5.21	5.33	0.59
fShan	*Phleum pratense*	PP	2.05	-0.03	0.20	1.11	1.16	1.99	0.38	3.17	-0.80	0.22	1.10	1.17	1.98	0.10
fShan	*Plantago lanceolata*	PL	5.70	-0.01	0.25	0.18	5.50	5.63	0.08	5.76	-0.10	0.22	0.18	5.50	5.63	0.10
fShan	*Poa pratensis*	PA	1.24	-0.05	0.79	0.54	0.24	1.04	0.17	0.75	-0.07	0.51	0.47	0.23	1.05	0.81
fShan	*Ranunculus acris*	RA	0.90	0.01	0.56	0.28	0.99	1.25	0.23	0.96	0.07	0.43	0.29	0.98	1.26	0.69
fEven	all samples	all	0.65	-0.01	0.25	0.32	0.52	0.59	0.05	0.64	-0.05	0.25	0.32	0.52	0.59	0.16
fEven	*Anthoxanthum odoratum*	AO	0.88	-0.02	0.55	0.22	0.49	0.73	0.11	0.86	-0.12	0.56	0.24	0.49	0.73	0.39
fEven	*Avenula pubescens*	AP	0.92	-0.01	0.58	0.13	0.71	0.85	0.10	1.06	-0.15	0.62	0.12	0.71	0.85	0.04
fEven	*Centaurea jacea*	CJ	0.77	0.00	0.46	0.01	0.77	0.78	0.24	0.77	0.01	0.55	0.01	0.77	0.78	0.35
fEven	*Dactylis glomerata*	DG	0.49	0.00	0.56	0.22	0.34	0.57	0.80	0.67	-0.11	0.68	0.22	0.34	0.57	0.29
fEven	*Festuca rubra*	FR	0.00	0.01	0.75	0.04	0.07	0.11	0.02	0.09	0.00	0.59	0.04	0.07	0.11	0.94
fEven	*Geranium pratense*	GP	0.85	0.00	0.53	0.01	0.82	0.84	0.22	0.87	-0.02	0.68	0.01	0.82	0.84	0.05
fEven	*Holcus lanatus*	HL	0.45	-0.01	0.41	0.12	0.25	0.39	0.18	0.44	-0.06	0.48	0.13	0.25	0.39	0.29
fEven	*Knautia arvensis*	KA	0.83	0.00	0.42	0.01	0.83	0.84	0.58	0.84	0.00	0.48	0.01	0.83	0.84	0.50
fEven	*Leucanthemum vulgare*	LV	0.81	0.00	0.28	0.02	0.82	0.83	0.02	0.81	0.01	0.31	0.02	0.82	0.83	0.16
fEven	*Phleum pratense*	PP	0.33	-0.01	0.20	0.18	0.19	0.32	0.42	0.50	-0.12	0.21	0.18	0.19	0.32	0.11
fEven	*Plantago lanceolata*	PL	0.85	0.00	0.30	0.02	0.84	0.85	0.35	0.86	-0.01	0.21	0.02	0.84	0.85	0.26
fEven	*Poa pratensis*	PA	0.22	-0.01	0.78	0.10	0.04	0.18	0.17	0.13	-0.01	0.50	0.08	0.04	0.18	0.82
fEven	*Ranunculus acris*	RA	0.15	0.00	0.56	0.05	0.16	0.20	0.24	0.16	0.01	0.41	0.05	0.16	0.20	0.74

Supplement Table 6: R session info. The list of R packages, including their versions, used for data analysis and plot creation.

R package	version
abind	1.4-5
AnnotationDbi	1.56.2
aroma.light	3.24.0
assertthat	0.2.1
backports	1.4.1
BatchCorrMetabolomics	0.1.14
Biobase	2.54.0
BiocFileCache	2.2.0
BiocGenerics	0.40.0
BiocIO	1.4.0
BiocManager	1.30.16
BiocParallel	1.28.3
biomaRt	2.50.1
Biostrings	2.62.0
bit	4.0.4
bit64	4.0.5
bitops	1.0-7
blob	1.2.2
broom	0.7.11
cachem	1.0.6
callr	3.7.0
car	3.0-12
carData	3.0-5
caret	6.0-90
cellranger	1.1.0
ChemometricsWithR	0.2.0
class	7.3-19
cli	3.1.0
cluster	2.1.2
codetools	0.2-18
colorspace	2.0-2
compiler	4.1.2
cowplot	1.1.1
crayon	1.4.2
crch	1.0-4
curl	4.3.2
data.table	1.14.2
DBI	1.1.2
dbplyr	2.1.1
DelayedArray	0.20.0
DEoptimR	1.0-10
digest	0.6.29
diptest	0.76-0
doParallel	1.0.16
dplyr	1.0.7
DT	0.2
e1071	1.7-9
EDASeq	2.28.0
edgeR	3.36.0
ellipsis	0.3.2
factoextra	1.0.7
FactoMineR	2.4
fansi	0.5.0
farver	2.1.0
fastmap	1.1.0
filelock	1.0.2
flashClust	1.01-2
flexmix	2.3-17
forcats	0.5.1
foreach	1.5.1
Formula	1.2-4
fpc	2.2-9
fs	1.5.2
future	1.23.0
future.apply	1.8.1
generics	0.1.1
GenomeInfoDb	1.30.0
GenomeInfoDbData	1.2.7
GenomicAlignments	1.30.0
GenomicFeatures	1.46.3
GenomicRanges	1.46.1
ggplot2	3.3.5
ggpubr	0.4.0
ggrepel	0.9.1
ggsignif	0.6.3
gtext	0.1.1
globals	0.14.0
glue	1.6.0
gower	0.2.2
grid	4.1.2
gridExtra	2.3
gridtext	0.1.4
gtable	0.3.0
haven	2.4.3
hms	1.1.1
htmltools	0.5.2
htmlwidgets	1.5.4
httr	1.4.2
hwriter	1.3.2
ipred	0.9-12
IRanges	2.28.0
iterators	1.0.13
jpeg	0.1-9
jsonlite	1.7.2
KEGGREST	1.34.0
kernlab	0.9-29
knitr	1.37
kohonen	3.0.10
labeling	0.4.2
lattice	0.20-45
latticeExtra	0.6-29
lava	1.6.10
leaps	3.1
lifecycle	1.0.1
limma	3.50.0
listenv	0.8.0
lmerTest	3.1-3
locfit	1.5-9.4
lubridate	1.8.0
magrittr	2.0.1
markdown	1.1
MASS	7.3-54
Matrix	1.3-4
MatrixGenerics	1.6.0
matrixStats	0.61.0
mclust	5.4.9
memoise	2.0.1
mgcv	1.8-38
MLmetrics	1.1.1
ModelMetrics	1.2.2.2
modelr	0.1.8
modeltools	0.2-23
munsell	0.5.0
nlme	3.1-153
nnet	7.3-16
parallel	4.1.2
parallelly	1.30.0
permute	0.9-5
pillar	1.6.4
pkgbuild	1.3.1
pkgconfig	2.0.3
pls	2.8-0
plyr	1.8.6
png	0.1-7
prabclus	2.3-2
prettyunits	1.1.1
pROC	1.18.0
processx	3.5.2
prodlim	2019.11.13
progress	1.2.2
proxy	0.4-26
ps	1.6.0
purrr	0.3.4
R.methodsS3	1.8.1
R.oo	1.24.0
R.utils	2.11.0
R6	2.5.1
rappdirs	0.3.3
RColorBrewer	1.1-2
Rcpp	1.0.7
RCurl	1.98-1.5
readr	2.1.1
readxl	1.3.1
recipes	0.1.17
remotes	2.4.2
reprex	2.0.1
reshape2	1.4.4
restfulr	0.0.13
rjson	0.2.20
rlang	0.4.12
robustbase	0.93-9
ROCR	1.0-11
rpart	4.1-15
rprojroot	2.0.2
Rsamtools	2.10.0
RSQLite	2.2.9
rstatix	0.7.0
rstudioapi	0.13
rtracklayer	1.54.0
RUVSeq	1.28.0
rvest	1.0.2
S4Vectors	0.32.3
sandwich	3.0-1
scales	1.1.1
scatterplot3d	0.3-41
scoringRules	1.0.1
ShortRead	1.52.0
splines	4.1.2
stats4	4.1.2
StatTools	0.0.916
stringi	1.7.6
stringr	1.4.0
SummarizedExperiment	1.24.0
survival	3.2-13
tibble	3.1.6
tidyr	1.1.4
tidyselect	1.1.1
tidyverse	1.3.1
timeDate	3043.102
tools	4.1.2
tzdb	0.2.0
utf8	1.2.2
vctrs	0.3.8
vegan	2.5-7
vroom	1.5.7
withr	2.4.3
xfun	0.29
XML	3.99-0.8
xml2	1.3.3
XVector	0.34.0
yaml	2.2.1
zlibbioc	1.40.0
zoo	1.8-9
R	version 4.1.2 (2021-11-01)

Supplement Figure 1: Schematic overview of plots in the Trait-Based Experiment (TBE) in the Jena Experiment. In this study, we used the two plant species pools Pool1 (orange) and Pool2 (green) each containing four selected grass and four selected herb species. In the TBE, plots with different initial diversity levels (DL) were arranged randomly across three blocks: Block A, Block B and Block C, which are located next to each other along the river Saale, where each plot measured 3.5 m by 3.5 m. Due to an overlap in the species pools Pool1 and Pool2, three target species were collected from both pools. Which target species were collected from which plots can be found in tables 1a and 1b of the main manuscript. In the TBE A009, A016, A020, A027, A040, A043, B048, B059, B060, B064, B085, C097, C110, C121, C131, and C133 were designed as DL 1, being initially planted with only the target species; A002, A011, A013, A026, A042, A044, A046, B054, B071, B080, B081, B090, B092, C114, C135, C137 were designed as DL 2 with the target species and one additional plant species randomly assigned from the species pool; A003, A005, A010, A030, A035, A045, B051, B057, B067, B075, C103, C109, C115, and C136 were designed as DL4 with the target species and three additional species; A018 and B073 were designed as DL8 with all target species from the species pool.



CURRICULUM VITAE

Career

Doctoral researcher: Biology and Bioinformatics since 2017
Martin-Luther-University Halle-Wittenberg (in collaboration with the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig and Leibniz Institute of Plant Biochemistry (IPB)
Master of Science: Biology – 2016
University Hamburg (in cooperation with Universidade de Federal de Pernambuco – INNOVATE)
Bachelor of Science: Biology – University Hamburg 2012

Certification

MicroMaster in Bioinformatics, online at University of Maryland Global Campus and University System of Maryland (in collaboration with edX) 2020
Data Science Professional Certificate, online at HarvardX (in collaboration with edX) 2021

Publications

Ferlian, O., …, **Marr, S.**, …, Eisenhauer, N. (2020). Soil chemistry turned upside down: a meta-analysis of invasive earthworm effects on soil chemical properties. Ecology, 101(3), e02936.
Marr, S., et al. (2021). LC-MS based plant metabolic profiles of thirteen grassland species grown in diverse neighbourhoods. Scientific Data
Jurburg, SD, …, **Marr,S.**, …, Heintz-Buschart, A. (2022). The community ecology perspective of omics data. Microbiome, 10(1), 1-13.
Walker, TWN, …, **Marr, S.**, …, Salguero-Gómez, R. (2022). Functional Traits 2.0: The power of the metabolome for ecology. Journal of Ecology, 110 (1), 4-20

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – DFG–FZT 118, 202548816 – iDiv-Flexpool, Martin Luther University Halle-Wittenberg (MLU) and Leibniz Institute of Plant Biochemistry (IPB) for funding and supporting the research. I am particularly grateful to the initiators of the project, Dierk Scheel, Steffen Neumann, and Helge Bruelheide, who acquired the funds and provided facilities and infrastructure. In addition, I would like to thank the defence committee as well as all members of the PhD advisory committee, including Nicole M. van Dam and Nico Eisenhauer. Also, special thanks to yDiv, for endless moral support, guidance and inspiration.

I also want to thank the gardeners and lab teams of the MLU and IPB, especially SD, and SK, for all their guidance, feedback and outstanding support. Many thanks to the gardeners of the Jena Experiment (DFG, FOR 456/1451) and numerous student helpers for maintaining the experimental side and plots. Special thanks also to MF, who went above and beyond to ensure a smooth sampling collection even in heavy thunderstorms. I am grateful to everyone I've collaborated with externally, such as the Brunius lab at Epihubben, Uppsala Universitet, Roessner lab at the University of Melbourne and TreeDì jointly funded by the DFG (GRK2324) and the University of Chinese Academy of Sciences (UCAS). I would also like to highlight the Biometris team at Wageningen University & Research, especially Jos A. Hageman and Ron Wehrens, for their remarkable support, mentoring and encouragement, which far exceeded all my expectations.

Finally, an immense thank you to CU and SS for their outstanding support, valuable discussions, feedback and unlimited bioinformatics superpowers and to UH, who has been and always will be mentioned – after all, you are the master of the library! The most special thanks to team Ilmenau for being a constant source of encouragement, and unconditional support, especially to PsR and his terrace and SAS for her unbreakable friendship. Special thanks also to MetraLabs, who provided me not only with funds and an office to finish the writing but also with secret knowledge about autonomous robots.

EIDESSTATTLICHE ERKLÄRUNG

Hiermit erkläre ich, Susanne Marr, an Eides statt, dass ich die vorliegende Dissertationsschrift mit dem Titel:

„A Study in Green: Investigating Secondary Metabolite Profiles in Thirteen Grassland Plant Species“

wie eingereicht an der Martin-Luther-Universität Halle-Wittenberg, selbstständig und ohne fremde Hilfe verfasst und alle Quellen und Hilfsmittel für wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe. Die jeweiligen Beiträge der anderen Autoren zu den Manuskripten der jeweiligen Kapitel ist am Anfang derselbigen für jedes Kapitel einzeln aufgeschlüsselt und am Ende jedes Kapitel einzeln genannt. Die vorliegende Dissertation wurde bis zu diesem Zeitpunkt an keiner Hochschule, auch nicht an der Naturwissenschaftlichen Fakultät I – Biowissenschaften der Martin-Luther-Universität Halle-Wittenberg, zum Zwecke der Promotion eingereicht oder vorgelegt. Eine Bewerbung um den Doktorgrad erfolgte bisher nicht.

Susanne Marr