Technische Universität München
Ingenieurfakultät Bau Geo Umwelt
Lehrstuhl für Siedlungswasserwirtschaft

Pathway Effect Studies of Different Environmental Pollutants on *Lemna minor* and *Phragmites australis* Metabolism Using Polarity-Extended Chromatographic Separation with Mass Spectrometric Detection

Rofida Mostafa Ali Wahman

Vollständiger Abdruck der von der Ingenieurfakultät Bau Geo Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFT (DR. RER. NAT.)

genehmigten Dissertation.

Vorsitzende/-r: Dr. rer. nat. Christian Wurzbacher

Prüfer der Dissertation:

1. Priv-Doz. Dr. Thomas Letzel
2. Prof. Dr. Jörg E. Drewes
3. apl. Prof. Dr. Peter Schröder

Die Dissertation wurde am 01.07.2021 bei der Technischen Universität München eingereicht und durch die Ingenieurfakultät für Bau, Geo und Umwelt am 27.08.2021 angenommen.

Acknowledgments

This dissertation at hand is based on my research carried out at the Chair of Urban Water Systems Engineering, Technical University of Munich (TUM), supervised by PD. Dr. Thomas Letzel. I would like to express my deepest gratitude to PD. Dr. Thomas Letzel for his constant guidance, helpful comments, advice, continued efforts, and support during this work. His guidance helped throughout the whole period of this work. I could not have imagined having a wonderful, flexible, and open-minded supervisor for my Ph.D. work like him.

I would also like to express my deep and sincere gratitude to Prof. Dr. Jörg Drewes for the valuable advice, support, and opportunity to prepare this dissertation at the Chair. In addition, Prof. Drewes gave me the chance to study in the Hydrosciences Montpellier department, faculty of Pharmacy, Montpellier, France, which enhance my acknowledgment and this work.

Furthermore, the international atmosphere of the institute that he is heading to gives me the chance to learn from different cultures, and consequently, have friends/colleagues from various countries around the world. All of these made me feel lucky to have worked at this prestigious institute. I would like to express my sincere thanks to PD Dr. Johanna Graßmann for providing invaluable guidance.

I would like to express my heartfelt gratitude to Prof. Dr. Peter Schröder for being the third examiner of this thesis. He is one of the most respectable specialists in the field of plant physiology and biochemistry. With the guidance, support, and incredible knowledge of Prof. Dr. Peter Schröder, this work has been accomplished. It is my great privilege and honor to have Prof. Dr. Peter Schröder as my co-supervisor. I wish we may have more cooperation in my future career and life. Also, I would like to my deep and sincere gratitude to Dr. Catarina Cruzeiro for providing invaluable guidance, encouragement, and patience. Thank you for always being there. Her help and guidance particularly on how to transfer ideas and results precisely into readable (high quality) publications will benefit me forever.

My sincere thanks go to Prof. Dr. Elena Gomez, the head of Hydrosciences Montpellier department, faculty of Pharmacy, and Dr. Serge Chiron for their hospitalization, support, and great guidance. Moreover, I would like to express my great thanks to Dr. Andrés Sauvêtre, Dr. Geoffroy Duporté, Prof. Dr. Frédérique Courant, as well as, the whole team there for their support, guidance, and incredible knowledge in Metabolomics analysis. Furthermore, the cooperation with them was used for validating this work is highly appreciated.

I am also greatly indebted to my colleagues on the Chair of Urban Water Systems Engineering, Technical University of Munich (TUM), and at AFIN-TS GMbH Company. I would like to thank administrative and technical staff members of the chair who have been kind enough to advise and help in their respective roles. Further thanks to Ms. Andrea Vogel for her help and cooperation. Further thanks to Mr. Stefan Moser for his help and cooperation in the statistical data analysis.

I am grateful to the Egyptian Ministry of Higher Education and Scientific Research (MHESR) for funding my Ph.D. project from 29.06.2018 to 8.10.2020.

My deepest gratitude goes to my parents for forgetting themselves and dedicating their life to us. My father, you are the greatest man I have ever known taking care of all the annoying details in our life. My Mother, your blessing and prayer for me is what keeps me going. I am also grateful to my lovely brothers and sister for their support and kindness.

Finally, but most importantly, I would like to dedicate this dissertation to my husband Mohamed for his love, patience, sacrifice, and understanding. I would like to deeply thank him for listening, helpful discussions, and encouraging me to finish this work on time.

Special thanks to my boys, Ibrahim and Hamza, thanks for your understanding of my work. You are partners in this work, without your help and understanding, I cannot finish it.

Abstract

A plant metabolomics approach is used to acquire fundamental comprehensive, nonbiased, and high-throughput analyses of complex metabolite mixtures in plant extracts to provide a functional screening of the cellular state. Plant metabolomes consist of a wide range of chemical species with very diverse physicochemical properties. Therefore, recent advances in mass spectrometry (MS) technology drive primarily metabolome analysis, which requires powerful analytical tools for the separation, characterization, and quantification of this vast compound diversity in plant matrices. The metabolome composition analysis is very sophisticated and requires establishing optimal sample extraction, metabolite separation/detection/identification, and automated data gathering/handling/analysis. Thus, this work provides a systematic investigation of Plant metabolomics changes under the effect of different pharmaceuticals.

Initially, the extraction method is optimized for *Lemna minor* and *Phragmites australis*. For this purpose, plant samples were extracted with different solvents (solvent mixtures) and analyzed with RPLC-HILIC-ESI-TOF-MS. From this study, fresh samples were considered the optimum samples to investigate the changes in plant metabolic under pharmaceuticals stress to avoid any false positive and false negative interpretations. The ultimate decision to use 100 % MeOH, 50% MeOH, and 100 % H_2O extracts paves the way to new views of plant metabolomes.

Secondly, the data handling and analysis (statistically), the metabolite separation, and data analysis workflows were conducted for *Lemna minor* and *Phragmites australis* using different analytical systems. The data handling was changed depending on the MS. The results provide important insight on the effect of diclofenac and carbamazepine on the *Phragmites australis* metabolic profile. The characterization could be realized with two partially newly developed, complementary workflows using novel statistical analysis.

The results showed that diclofenac activated the glutathione pathway in *Phragmites australis*; however, the unsaturated fatty acid pathway was enhanced due to carbamazepine incubation. *Lemna minor* incubation with diclofenac showed an enhancement in flavonoids and chlorophyll degradation pathways. Thus, *Lemna* and *Phragmites* could be a source of antioxidant compounds.

Thirdly, particular attention was paid to the phenylalanine, tyrosine, and tryptophan biosynthesis pathway and phenylpropanoids pathway. For this purpose, *Lemna* minor was incubated with diclofenac and carbamazepine for 4, 8, and 12 days. Also, it was incubated with sulfamethoxazole, trimethoprim, diclofenac, carbamazepine, and a mixture of all of them at environmentally relevant concentrations. The results showed a difference in *Lemna's* response to each pharmaceutical. The response was based on the dosage of the drug.

The use of versatile MS-based techniques, in addition to chemometric analysis techniques, enabled the investigation of the effect of different pharmaceuticals on *Lemna minor* and *Phragmites australis* metabolic profiles. With a strong connection with the analytical and (statistical and biochemical) research, this work provides a comprehensive investigation of plant metabolic pathways under different pharmaceuticals effects.

Zusammenfassung

Ein pflanzlicher Metabolomics-Ansatz wird verwendet, um grundlegende, umfassende, unverzerrte und Hochdurchsatz-Analysen von komplexen Metabolit-Mischungen in Pflanzenextrakten zu erhalten, um ein funktionelles Screening des zellulären Zustands zu ermöglichen. Pflanzliche Metabolome bestehen aus einer Vielzahl von chemischen Spezies mit sehr unterschiedlichen physikochemischen Eigenschaften. Daher treiben die jüngsten Fortschritte in der Massenspektrometrie (MS)-Technologie vor allem die Metabolomanalyse voran, die leistungsstarke Analysewerkzeuge für die Trennung, Charakterisierung und Quantifizierung dieser enormen Verbindungsvielfalt in Pflanzenmatrices erfordert. Die Analyse der Metabolom-Zusammensetzung ist sehr anspruchsvoll und erfordert die Etablierung einer optimalen Probenextraktion, Metabolitentrennung/Detektion/Identifizierung und eine automatisierte Datenerfassung/Bearbeitung/Analyse. Daher bietet diese Arbeit eine systematische Untersuchung der Metabolomveränderungen von Pflanzen unter der Wirkung verschiedener Pharmazeutika.

Zunächst wird die Extraktionsmethode für *Lemna minor* und *Phragmites australis* optimiert. Dazu wurden die Pflanzenproben mit verschiedenen Lösungsmitteln (Lösungsmittelgemischen) extrahiert und mit RPLC-HILIC-ESI-TOF-MS analysiert. In dieser Studie wurden frische Proben als die optimalen Proben für die Untersuchung der Veränderungen des Pflanzenstoffwechsels unter Pharmastress angesehen, um jegliche falsch positive und falsch negative Interpretationen zu vermeiden. Die endgültige Entscheidung, 100 % MeOH, 50 % MeOH und 100 % H_2O-Extrakte zu verwenden, ebnete den Weg zu neuen Ansichten der Pflanzenmetabolome.

Zweitens wurden das Datenhandling und die Datenanalyse (statistisch), die Metabolitentrennung und die Datenanalyse-Workflows für *Lemna minor* und *Phragmites australis* mit unterschiedlichen Analysesystemen durchgeführt. Das Datenhandling wurde in Abhängigkeit vom MS geändert. Die Ergebnisse liefern wichtige Erkenntnisse über die Wirkung von Diclofenac und Carbamazepin auf das metabolische Profil von *Phragmites australis*. Die Charakterisierung konnte mit zwei teilweise neu entwickelten, sich ergänzenden Workflows unter Verwendung neuartiger statistischer Analysen realisiert werden.

Die Ergebnisse zeigten, dass Diclofenac den Glutathion-Stoffwechselweg in *Phragmites australis* aktiviert; der ungesättigte Fettsäure-Stoffwechselweg wurde jedoch durch die Inkubation mit Carbamazepin verstärkt. Die Inkubation von *Lemna minor* mit Diclofenac zeigte eine Verstärkung der Flavonoide und des Chlorophyll-Abbauweges. Somit könnten *Lemna* und *Phragmites* eine Quelle für antioxidative Verbindungen sein.

Drittens wurde besonderes Augenmerk auf den Phenylalanine, Tyrosine und Tryptophane Biosyntheseweg und den Phenylpropanoid Weg gelegt. Zu diesem Zweck wurde *Lemna minor* mit Diclofenac und Carbamazepin für 4, 8 und 12 Tage inkubiert. Außerdem wurde es mit Sulfamethoxazol, Trimethoprim, Diclofenac, Carbamazepin und einer Mischung aus allen in umweltrelevanten Konzentrationen inkubiert. Die Ergebnisse zeigten einen Unterschied in der Reaktion von *Lemna* auf die einzelnen Arzneimittel. Die Reaktion war abhängig von der Dosierung des Medikaments.

Die Verwendung vielseitiger MS-basierter Techniken, zusätzlich zu chemometrischen Analyseverfahren, ermöglichte die Untersuchung der Wirkung verschiedener Pharmazeutika auf die Stoffwechselprofile von *Lemna minor* und *Phragmites australis*. Mit einer starken Verbindung zur analytischen und (statistischen und biochemischen) Forschung bietet diese Arbeit eine umfassende Untersuchung der pflanzlichen Stoffwechselwege unter den Auswirkungen verschiedener Pharmazeutika.

Table of Contents

List of abbreviations

AAAs	Aromatic Amino Acids
HMG-CoA	β-Hydroxy-β-Methylglutaryl-CoA
CE-MS	Capillary Electrophoresis-Mass Spectrometry
CBZ	Carbamazepine
CM	Chorismate Mutase
DCF	Diclofenac
DMF	Differentiating Metabolic Profile
DHFR	Dihydrofolate Reductase
DHPS	Dihydropteroate Synthase
EI	Electron Ionization
ESI	Electrospray Ionization
GC-MS	Gas Chromatography-Mass Spectrometry
GNPS	Global Natural Products Social Molecular Networking resource
GMD	Golm Metabolome Database
HMDB	Human Metabolome Database
HILIC	Hydrophilic Interaction Liquid Chromatography
LC-MS	Liquid Chromatography-Mass Spectrometry
MoNA	Mass Bank of North America
MS	Mass Spectrometer
MVA	Mevalonic Acid
NADPH	Nicotinamide-Adenine Dinucleotide Phosphate
NSAID	Non-Steroidal Anti-Inflammatory Drug
NMR	Nuclear Magnetic Resonance
OPLS-DA	Orthogonal Partial Least Squares Discriminant Analysis
PLS	Partial Least Squares Regression
PCM	Pathway of Carbamazepine Metabolism
PAL	Phenylalanine Ammonia-Lyase
PCs	Phytochelatins
PCB	Polychlorinated Biphenyls
PCA	Principal Component Analysis
RP	Reversed-Phase
SMX	Sulfamethoxazole
SFC	Supercritical Fluid Chromatography
SAR	Systemic Acquired Resistance
TPs	Transformation Products
TRIM	Trimethoprim
UPLC	Ultra-Performance LC

List of Figures

List of Tables

1. Introduction

1.1. Plant metabolomics

Metabolomics is the approach that deals with the investigation of a biological system (cell, tissue, or organism) by determining its overall metabolite profile at a given time point with the specified set of conditions. According to (Toyo'oka, 2008), metabolomics is defined as the quantitative measurement of the dynamic multi-parametric metabolic response of living systems to pathophysiological stimuli or genetic modification through subsequent data analysis via a range of multivariate statistical approaches. He defined the important aspect of metabolomics as the precise and significant identification of biomarkers (Toyo'oka, 2008). In the 1950s, metabolite profiling appeared for the first time and continues in development throughout the next three decades. The concept of metabolomics was firstly proposed in 1999 using the same idea as genomics and proteomics (Yang *et al.*, 2019). Despite this, metabolomics has developed slowly and has only recently become an area of major research interest (Rochfort, 2005). Plant metabolomics is one of the "Metabolomics" fields, which have witnessed a huge development in recent years (Heyman and Dubery, 2016). In the plant research community, metabolomics research has been extremely active due to a large number of metabolites with important functions in plant ecology and the protection against stress conditions (Rochfort, 2005).

It has a fundamental prospect in improving the current understanding of plant natural products (plant metabolites), enabling the development of new investigation strategies, which has matured as a valuable tool for advancing our understanding of plant biology and physiology. The basic goal of metabolomics is to provide a comprehensive qualitative and/or quantitative analysis of all metabolites present in a living system. Interestingly, this concept could be extended in natural product drug discovery via studying the relationship between the whole metabolome of nature-derived remedies and their biological effects (Sumner *et al.*, 2015). Also, it provides a broader insight into the biochemical status and gene functions of the studied organisms. Furthermore, the relation between the compounds and their activity in the metabolome could be revealed with the advanced bioinformatics tools, which can be useful in pharmacological standardization and fingerprint investigations (Ernst *et al.*, 2014; Sumner *et al.*, 2015; Heyman and Dubery, 2016).

Plants contain a wide variety of metabolites differing in molecular size, polarity, functionality, etc.; the total number in the plant kingdom is estimated to be up to 200,000. Therefore, plant metabolomics has a wide range of applications in different fields such as fingerprinting for taxonomic or biochemical purposes, the influence on a metabolite profile by external stimuli, interactions between plants and herbivores, quality control of medicinal herbs, and activity

determinations of medicinal plants (Ernst *et al.*, 2014). Plant Metabolomics studies involve sample preparation, sample analysis, data analysis, and metabolic pathway analysis (Yang *et al.*, 2019).

1.2. Analytical methods

As a consequence of the fundamental importance of plant metabolites, analytical methods have been developed for decades to investigate their importance in the human diet and biochemical pathways, besides their usage as a biomarker for a wide range of biological conditions. Recently, the analytical strategies are gas chromatography-mass spectrometry (GC-MS), capillary electrophoresis-mass spectrometry (CE-MS), liquid chromatography-mass spectrometry (LC-MS), and nuclear magnetic resonance (NMR) spectroscopy (Salem *et al.*, 2020). GC-MS has been one of the most popular techniques to analyze volatile and semi-volatile organic compounds in a wide variety of samples. A great advantage of GC-MS is that it is both relatively sensitive and highly robust, and can routinely and reproducibly measure hundreds of analytes across thousands of samples with the help of various databases. The GC-MS method also has several applications in plant, pharmacological, and medical metabolomics studies, and is considered as one of the most suitable techniques for the accurate determination of primary metabolites, however, it is severely compromised in measuring thermolabile secondary metabolites. Also, the thermal stability of the stationary phase, metabolites, and their derivatives, which might introduce artifacts, limit the metabolome coverage derived by GC-MS (Aretz and Meierhofer, 2016; Salem *et al.*, 2020). Also, CE-MS separates polar and charged compounds based on their charge-to-mass ratio from small injection volumes, including nucleotides and highly charged metabolites. The use of CE-MS in plant metabolism remains relatively rare due to time-consuming, poor migration time reproducibility, and a lack of reference libraries (Salem *et al.*, 2020).

Liquid chromatography-mass spectrometry (LC-MS) has become the most comprehensive technique to measure a wide range of diverse metabolites. Unlike GC-MS, it does not require prior sample treatment, and crude extracts obtained by simple extraction can be introduced directly to the LC-MS. It is a unique method for measuring plant secondary metabolites such as flavonoids and alkaloids and primary metabolites such as amino acids (Salem *et al.*, 2020). Metabolomics is fundamentally based on the rapid and efficient qualitative and quantitative analysis of a large number of samples. The simplest approach available to shorten the analytical run opens up the possibility of a relatively high throughput (number or variety of metabolites retained and separated by the system) screening for samples containing multiple components by decrease the column particle size as in Ultra-Performance LC (UPLC) (Toyo'oka, 2008).

The development of UPLC rendered the technique even more powerful concerning resolution, sensitivity, and throughput for samples containing multiple components. Also, the different

stationary phase mechanism of separation is a powerful tool in metabolomics analysis. Tolstikov and Fiehn have proven that hydrophilic interaction liquid chromatography (HILIC) is a potent technique for polar metabolites profiling (Tolstikov and Fiehn, 2002). Then, the next approach was the coupling of different small size column such as; two coupled monolithic silica columns of different polarity were applied to verify the presence of salicylic compounds and other phenolic derivatives in the bark of six species from the genus *Salix* (Pobłocka-Olech *et al.*, 2007), and coupling of reversed-phase (RP) and HILIC to analyze food constituents (Hemmler *et al.*, 2018). Moreover, several improvements like serial coupling, column switching, two-dimensional LC, and supercritical fluid chromatography (SFC) have extended the coverage of the metabolome throughput (Haggarty and Burgess, 2017).

Still, the poor availability of standard compounds for plant metabolites is considered the limitation of MS-based metabolomics usage. Recently, web resources enable overcoming the grand challenge of MS-based metabolomics concerning peak identification and alignment or in the structural identification of compounds and/or pathways (Tohge and Fernie, 2009).

The high sensitivity, low sample consumption, fast analysis, and advantages of separation and identification of mass spectrometers enhanced their usage widely, and how the method is considered as the method of choice in metabolomics analysis (Yang *et al.*, 2019). Recently the number of MS metabolomics studies has increased including the effect of drugs (Cuperlovic-Culf and Culf, 2016), toxins (Zhao *et al.*, 2012), and various xenobiotics (Mishra *et al.*, 2017; Wahman *et al.*, 2021) on the metabolites.

The choice of ionization mode is based on the compound's chemical nature, matrix, and pH of the experiment. Doubtless, the analysis of plant samples in both positive and negative ionization modes will provide the most comprehensive insight into their metabolic composition. Nevertheless, the analysis in just a single ionization mode is already sufficient for an untargeted screening strategy to obtain a global overview of the differences and similarities between samples and to relate metabolic variation to traits, such as genetic variation, gene function analysis, plant development, and abiotic stress (De Vos *et al.*, 2012). Concerning LC, the purification technique is inadequate and unknowns' identification as in NMR. Hence, NMR spectroscopy is a powerful analytical tool and a very useful technique for the structure elucidation of novel and/or unexpected compounds. Unlike MS, NMR is less biased as the results do not rely on the type of ionization condition. Also, NMR methods are highly reproducible, noninvasive, nondestructive, and require minimal sample preparation. Nevertheless, the limitation of NMR is the low sensitivity compared to MS in plant metabolomics due to plant metabolites' low concentrations (Zhang *et al.*, 2016; Salem *et al.*, 2020). In 2017, Matsuo and coworkers reported interesting attempts in GC-MS structural elucidation without NMR (Matsuo *et al.*, 2017).

1.3. Data evaluation and Statistical analysis

GC-MS, LC-MS, and NMR are the most relevant techniques within the context of natural product discovery. The GC-MS high resolution of the chromatographic separation and stability of ionization method is adequate for providing both stable separation and mass spectra fragmentation via its most common ionization source electron ionization (EI). Mass features deconvolution and several tools are available which are accomplished of performing all steps from the raw data to peak picking and metabolite annotation, besides, matching against databases such as the Golm Metabolome Database (GMD) and MassBank of North America (MoNA), and others.

LC-MS is the most versatile technique in terms of metabolome coverage and relevant for the rapid detection of natural products in complex mixtures due to high sensitivity, the multitude of stationary phases with different chemistries, and ionization sources. The greatest challenges faced by LC-MS are the reproducibility of the results due to retention time's shift across multiple runs. Many factors are influencing separation and ionization efficiencies such as changes in solvent composition, column stability, electrospray formation, ion suppression, and matrix effects. The lack of reproducibility in both retention and spectra makes the identification of known compounds based on database matches significantly less efficient than for GC.

Data processing is also a more demanding task due to the ionization instability combined with the lower resolution of LC separation, resulting in much more complex datasets with a large number of mass features. Despite all of these challenges, the great flexibility of LC-MS can lead to comprehensive developments for this technique such as coupling of different columns (Navarro-Reig *et al.*, 2017; Wahman *et al.*, 2020; Wahman *et al.*, 2021). Recently, several approaches are available to process, analyze, and interpret LC-MS-based data. Considering that plant metabolomics is mainly investigated using the untargeted metabolomics analysis strategy, which is based on the detection of all signals within the dataset. Various software packages are currently available for the processing of mass chromatograms, which work on similar principles, identifying m/z signals above a certain threshold along with connecting these signals over time (retention time) to create a chromatography (for features) representation such as XCMS, MZmine 2, OpenMS data, etc. Those features have to be aligned across multiple samples before data analysis. One of the first challenges in processing the resulting data matrices is the absolute number of features, usually in ten of thousands, most of which are background noise, contaminations, and in-source fragmentation (Salem *et al.*, 2020). The extraction blanks and technical replicates of quality controls (standards mixtures) are useful for the identification of features that do not originate from the samples or with low reproducibility. The second challenge is the annotation of different features. This process can significantly facilitate the next steps of data analysis, particularly for the identification of biomarkers and differentiating features between experimental groups, which is still considered the major bottleneck. Due to the lack of reference standards for the majority of the metabolome, the identification of mass spectrometer datasets is a vital obstacle. Consequently, the

database has significantly expanded its coverage by including in silico generated MS/MS spectra such as the Metfrag tool. The identification of compounds by mass spectrometry includes now the experiments with tandem MS fragmentation, which are used for compound identification based on matching spectra signals against databases such as Metlin, Mass Bank, MoNA, European Mass Bank, Global Natural Products Social Molecular Networking resource (GNPS), and Human Metabolome Database (HMDB). Moreover, other databases adopt specific organisms such as ECMDB for *E. coli*, YMDB for yeast, ResPect, and PLANT-IDENT for plants.

Recently, untargeted metabolomics depends on high resolution and highly accurate ion mass measurements provided mainly by QTOF and Orbitrap-based systems. Putative molecular ion masses matching is not as reliable as MS/MS spectral matching, however, can still bring valuable insights into putative structures and provide a significantly smaller search space in databases. Unfortunately, this information is still often not well structured, with few databases providing extensive collections of curated data, e.g. PLANT-IDENT.

1.4. Application of plant metabolomics

The plant metabolomics approach usually is performed using GC and/or LC coupled to a high-resolution mass spectrometer. The spectra are acquired under definitive conditions using sufficient ionization such as electrospray ionization (ESI) in positive and/or negative modes on tandem mass spectrometer such as quadrupole time-of-flight (QTOF). The produced fragmentation data enable the identification of different metabolites using the previously mentioned software and databases. The limitations of database identification are the lack of chromatographic data and differences between MS. The databases contain spectral information, and some have chromatographic information, but there are no retention times (or indices) data except the PLANT-IDENT database in the FOR-IDENT platform. However, the main usage of these databases is the annotation of the spectra with taxonomical information about the structural class and species from which metabolites have been identified such as GNPS. Consequently, the untargeted metabolic strategy is performed to assign the differentiating metabolic profile between different groups (Wahman *et al.*, 2021).

Moreover, many plant metabolic pathways have been elucidated traditionally by radioactive labeling of primary metabolites and subsequently following their fate. With the advent of recombinant DNA technology, these approaches were combined with the isolation of the respective genes and purification of candidate enzymes such as flavonoids (Nabavi *et al.*, 2020). Recently, targeted molecular approaches or, the combination of genomics metabolic profiling are changing dramatically the field of the plant metabolites biosynthesis pathways studies with new knowledge concerning the changes in the synthesis of the plant metabolites (Scossa *et al.*, 2018). The genomes of *Lemna minor* and *Phragmites australis* have been investigated in (Neuhaus *et al.*,

1993; Van Hoeck *et al.*, 2015), respectively. Consequently, the investigation of pathway changes in both plants can be performed with the targeted analysis and/or the metabolic profile studies.

1.5. *Lemna minor*

Lemna minor is an aquatic monophyletic plant known as duckweed, which belongs to the Lemnaceae/Araceae family. There are about 40 species in five genera: *Lemna*, *Landoltia*, *Spirodela*, *Wolffia*, and *Wolffiella* (Ekperusi *et al.*, 2019).

Lemna minor composes of one or a few leaves called fronds and a single root or rootlet with no stem. It reproduces vegetatively by simply dividing to form separate individual fronds. *L. minor* is about 2-4 mm across. It aggregates together forming colonies on surface waters. The frond doubling time for *L. minor* is about 1.4 days. Duckweed cultured in the laboratory can grow indefinitely if nutrients, light, and water are provided, thus producing unlimited duckweed specimens for use at any moment (Ekperusi *et al.*, 2019). *Lemna spp.* are abundant in the tropical and subtropical countries; growing profusely in still, nutrient-rich small ponds, ditches, and swamps or in slowly moving water bodies, which have relatively high levels of N, P, and K. There, it concentrates the minerals and synthesizes protein (Chakrabarti *et al.*, 2018b).

Of all the Lemnaceae species, *Lemna* species are probably the best known because of their extensive use in lab-based tests due to their small size. Therefore, *L. minor* is considered as a model system to study fundamental plant research in the photoperiodic control of flowering, and the discovery of auxin biosynthesis and sulfur assimilation pathways. The sporadic distribution and invasive nature of the plant increased its potential to withstand stressful environmental conditions including pollution or wastewater. Therefore, *L. minor* is an efficient detoxifier of pollutants (organic pollutants, heavy metals, agrochemicals, radioactive waste, nanomaterials, and hydrocarbons) in surface waters. Duckweeds are also used for the treatment of wastewater and the production of bio-fuel. Moreover, *Lemna* varieties show sensitivity to toxicants allowing their usage in ecotoxicological research (Aliferis *et al.*, 2009; Ekperusi *et al.*, 2019). Hence, *L. minor* is a successful organic pollutants phytoremediator, which can result in a 94.45 and 79.39% reduction in biochemical oxygen demand and phosphate with an increased level of dissolved oxygen due to an improvement of nutrients load (Mohedano *et al.*, 2012). Moreover, duckweeds are capable of metal bioaccumulation in a substantial amount into their tissue. In 2008, Razinger and coworkers have reported that a high accumulation of cadmium resulted in the inhibition of *L. minor*. However, duckweed was able to recover within days after high exposure to copper, nickel, and cadmium (Drost *et al.*, 2007). Moreover, *L. minor* exposure to sulfadimethoxine for 7 days results in increasing in guaiacol peroxidase and catalase activity. Additionally, the total frond area and chlorophyll b content seem to be the most sensitive endpoints with EC50 of 478 and 554 μg/L,

respectively (Drobniewska *et al.*, 2017). Ciprofloxacin, lomefloxacin, ofloxacin, levofloxacin, and norfloxacin have all been screened for phytotoxicity in *L. gibba* and *L. minor*, with 7-day fresh weight EC50 values ranging from 97 to 913 μg/L and 53 to 2,470 μg/L, respectively. Further, *Lemna sp*. EC50 values for tetracyclines were 219- 4,920 μg/L (Brain *et al.*, 2008a) due to chloroplastic type II topoisomerases and DNA replication inhibition (Brain *et al.*, 2009).

L. minor contains crude protein 16-45%, fat 4.4-4.0%, *p*-coumaric acid 0.015%, fiber 8-10%, and ash 4-5% of the dry weight. The flavonoid content as luteolin-7-glucoside (cinnaroside) is 0.38 ± 0.01%. Further, the common duckweed containes 14 elements: calcium 4990 mg/100 g, potassium, silicon 2495 mg/100 g, sodium 1870 mg/100 g, magnesium 155 mg/100 g, iron 934 mg/100 g, phosphorus 515 mg/100 g, aluminum 0.93 mg/100 g, manganese 935 mg/100 g, nickel 0.93 mg/100 g, copper 0.78 mg/100 g, lead 0.03 mg/100 g, molybdenum 0.02 mg/100 g and zinc 0.01 mg/100 g (Al-Snafi, 2019).

The fatty acid composition of *Lemna minor* equals 60-63% of total fatty acids, largely α-linolenic acid 41 to 47% and linoleic acid 17-18%, and total triacylglycerols are 0.03 ± 0.01% of dry weight. Three fatty acids palmitic, linoleic acid, and α-linolenic acid comprise more than 80% of total fatty acids (Al-Snafi, 2019). Additionally, unsaturated fatty acids (76.7 %) oleinic and linoleic have been reported. Furthermore, the various essential (arginine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, valine, and tyrosine) and non-essential amino acids, polyunsaturated fatty acids, β-carotene, and xanthophylls have made *Lemna spp*. a potential feed source for livestock. Moreover, it is used as food, in the treatment of flu-like illness, and for reducing body response to allergy (Zhao *et al.*, 2014b). *Lemna minor* has been used traditionally as antipruritic, antiscorbutic, astringent, depurative, diuretic, febrifuge, and soporific (Al-Snafi, 2019). Additionally, *L. minor* has been manipulated to produce monoclonal antibodies for the treatment of human diseases such as tissue inflammation, an autoimmune condition, and cancerous cells (Ekperusi *et al.*, 2019).

1.6. *Phragmites australis*

Phragmites australis, known as “common reed”, belongs to the Poaceae family (Park and Blossey, 2008). Reed is a typical swamp and aquatic plant species; it inhabits both aquatic and terrestrial ecosystems. *Phragmites australis* has annual cane-like stems up to 6 m in height, from 4 to 10 mm in diameter, as well as an extensive rhizome system (Milke *et al.*, 2020).

The lifespan of individual rhizomes is, on average, about six years, and they can grow within a radius of 10 m, at a rate of 1 per year. The annual upright stems spread from vertical rhizomes. Its roots develop from rhizomes and other submerged parts of shoots. Rhizomes form the largest densities at a depth of 0.5 m. The leaves are smooth, alternate with narrow-lanceolate laminae 20-

70 cm long and 1-5 cm wide. They are closely nerved, and they taper to long slender points. The inflorescence is a terminal panicle, often 30 cm long, lax, with color from dull purple to yellow, and the main branches bear many spikelets. The smooth branches usually have scattered groups of long silky hairs (Milke *et al.*, 2020).

Phragmites australis can be found all over the world except in Antarctica, but it is mainly present in Europe, the Middle East, and America (Köbbing *et al.*, 2013). Owing to its wide distribution, as well as its phenotypic plasticity, the plant exhibits a high capacity for adaptation to environmental conditions which are considered harmful (Milke *et al.*, 2020). According to (Sauvêtre and Schröder, 2015), the *Phragmites* root zone is rich in dissolved oxygen and organic carbon providing suitable conditions for the colonization of microorganisms. Bacteria and mycorrhiza fungi found in the rhizosphere play an important role in phytoremediation through degrading metals, organic pollutants, radionuclides, and xenobiotic compounds.

Kastratović *et al.* (2013) reported that the concentrations of five heavy metals (Cd, Cu, Mn, Zn, and Sr) were found to be higher in the *Phragmites australis* than in sediment and water. Prica and coworker (2019) reported that plant responds differently depending on the plant species and defense mechanisms and not a function of the heavy metal total concentrations (Milke *et al.*, 2020). It has been used for a long time in wetlands to remove pollutants, reduce nitrogen loads and provide oxygen to the rhizosphere. Furthermore, it is used as a resource for traditional crafts and fodder. In some regions like Northern China, it is grown as a crop and its leaves are used in the treatment of bronchitis and cholera (Park and Blossey, 2008).

1.7. Xenobiotics

Human pharmaceuticals are enteries into the aquatic environment, through the discharge of wastewater treatment plants (WWTPs) effluent. Despite the removal of a large proportion of many pharmaceuticals through various treatment processes, they are not completely eliminated. Consequently, pharmaceuticals are occurring in detectable amounts (Brain *et al.*, 2008a). Among the problematic xenobiotics in worldwide wastewater treatment plants are diclofenac, carbamazepine, sulfamethoxazole, and trimethoprim since they are associated with adverse ecotoxicological effects. Their concentrations in the influent, effluent of the wastewater plant, and surface water are listed in Table 1.

Diclofenac (DCF) is a widely distributed non-steroidal anti-inflammatory drug that can be found in surface waters in concentrations up to ng/L level and is considered an environmental risk factor (Huber *et al.*, 2012). Further, in 37 countries (out of 50), including Germany, the average concentration of DCF exceeded yearly the defined predicted no-effect concentration with values up to 1 µg/L. The estimated removal efficacy for DCF using conventional wastewater treatment

plants varies from 39% over 43% to 70% (Fischer *et al.*, 2020). Furthermore, DCF's photo-transformation products have high potential toxicity at concentration levels close to environmental concentrations (Schmitt *et al.*, 2007). Moreover, DCF is toxic at a low dose (100 µg/L); for example, it changed the algae and bacteria structures in rivers biofilms as it was shown in Huber *et al.* (2012) and Thelusmond *et al.* (2018). Alkimin *et al.* (2019) reported that DCF incubation in *L. minor* caused a decrease in the content of photosynthetic pigments, relative fluorescence decay values of chlorophyll, and oxidoreductase and dehydrogenase activities. However, it increased in non-photochemical quenching, amount of reactive nitrogen and oxygen species in roots, lipid peroxidation, oxidized ascorbate and thiols, and glutathione-reductase activity (Alkimin *et al.*, 2019). Bigott and co-workers, (2021) reported that DCF (20 µg/L) can act as a signal, which affects the circadian expression of reactive oxygen species (ROS) response selected genes in lettuce (Bigott *et al.*, 2021). These findings induce concerns regarding the chronic exposure of plants in constructed wetlands. Further, carbamazepine (CBZ) has been used widely as an antiepileptic and mood stabilizer since the 1970s. Both DCF and CBZ are taken up and translocated into the aerial parts of plants, where they can be accumulated or metabolized into more or less toxic products (Huber *et al.*, 2012; Sauvêtre and Schröder, 2015).

Sulfamethoxazole (SMX) and trimethoprim (TRIM) are among the most frequently detected antibiotics in the environment. SMX belongs to sulfonamides, which are the first class of antimicrobial agents widely used to treat diseases and infections for both humans and livestock. Moreover, the use of antibiotics for veterinary purposes is approximately five times higher than that for humans and makes approximately 10-23% of total antibiotic use in some EU countries and South Korea (Cheong *et al.*, 2020). Trimethoprim (TRIM) is an antibiotic usually administered in combination with SMX. Compounds concentrations are listed in (Table 1). These two antibiotics target two parts of the bacterial tetrahydrofolic acid synthesis pathway, which is an important cofactor, e.g. for DNA synthesis. Both of them showed a toxic effect on soil and plant growth, in addition to seed germination (Liu *et al.*, 2009). Also, in 2008, Brain and coworkers reported that *p*-aminobenzoic acid increased in *Lemna gibba* upon exposure to SMX. These results suggest that, as in bacteria, sulfonamide antibiotics specifically disrupt folate biosynthesis via inhibition of the dihydropteroate synthase enzyme. Exact estimates of SMX usage are unknown, though approximately 23,000 tons of antibiotics are produced each year in the U.S. alone. Consequently, SMX tends to enter the environment via two main routes: sewage effluent from human sources and agricultural runoff from animal husbandry operations. Sulfamethoxazole readily contaminates both surface water and groundwater due to its physicochemical properties. Furthermore, the half-life of SMX in aquatic systems is 19 days. Hence, SMX displays a considerable propensity for contaminating aquatic environments (Brain *et al.*, 2008b). For TRIM, it has been shown a slow degradation rate with direct photolysis in water matrices. Also, very stable photo-transformation products are formed (Sirtori *et al.*, 2010). In several studies, it has

been found that TRIM is quite resistant to biological wastewater treatment techniques. While other studies reported a partial removal of TRIM ranging from 40 to 50% or even almost complete removal, the concentration is listed in (Table 1). So far, different transformation products (TPs) of TRIM have been identified, resulting from hydroxylation reactions and ring-opening (Jewell *et al.*, 2016).

Table 1. Concentration of diclofenac, carbamazepine, sulfamethoxazole, and trimethoprim in the influent, effluent, and surface water in µg/L. Also, the toxicity (EC50) to *Lemna spp* in (mg/L) was listed.

Compound Name	Influent (µg/L)	Effluent (µg/L)	Surface water (µg/L)	Toxicity (EC50) (mg/L)
Diclofenac	0.44-7.1 (Vieno and Sillanpää, 2014)	0.12-4.7 (Vieno and Sillanpää, 2014)	0.06-15 in Germany rivers (Alkimin *et al.*, 2019)	7.5 (Cleuvers, 2003)
Carbamazepine	0.21-2.62 (Santos *et al.*, 2007)	1.08-6.3 (Hai *et al.*, 2018)	0.08-1.10 (Hai *et al.*, 2018)	25.5 (Hai *et al.*, 2018)
Sulfamethoxazole	3 (Giebułtowicz *et al.*, 2020)	0.1-956.4 (Kairigo *et al.*, 2020)	96.9-142.6 (Kairigo *et al.*, 2020)	0.13 (Brain *et al.*, 2008b)
Trimethoprim	0.29 (Kairigo *et al.*, 2020)	15.8 (Kairigo *et al.*, 2020)	3.8-4.4 (Kairigo *et al.*, 2020)	2 (Kortesmäki *et al.*, 2020)

1.8. Specific metabolites pathway applications to evaluate xenobiotic stress

Xenobiotics have been reported to modify important physiological processes such as photosynthesis (reduction of photosynthetic pigments), respiration, carbon allocation, increased leaf thickness, stomatal functioning, metabolic and enzyme activity (such as antioxidants) (Sharma *et al.*, 2020). Different xenobiotics may produce similar metabolic reactions, analysis of changes in metabolic pathways can provide a stronger correlation between the measured biomarker and the biological effect (Singer *et al.*, 2003). *Betula pendula* exposure to ozone showed an increase of the total sugars and reduced the amount of phenolic glucoside. Besides, it showed an increase in the catechin pentoside, hyperoside, and papyriferic acid in the stems (Lavola *et al.*, 1994). Sugars are precursors of terpenes biosynthesis. A study has suggested a disturbance in sugar metabolism

concerning the terpene biosynthesis following SO_2 exposure in *Picea albies and Pinus sylvestris*. The SO_2 exposure caused an increase in palustric and neoabietic acids, besides glucose and fructose in the first, however, decreased their concentrations in the second, respectively (Kainulainen *et al.*, 1995). Furthermore, Bartha and coworkers, in 2010 reported that acetaminophen caused changes in antioxidant defense enzyme activities towards reactive oxygen species in Indian mustard. In 2019, Sivaram and coworkers reported a high impact on the galactose metabolism, aminoacyl tRNA biosynthesis, and TCA cycle in maize leaves exposed to benzo[a]pyrene, and pyrene, individually, and their mixture. Moreover, they induced the phytochemicals which have a defense function and improved the protection against biotic and abiotic stresses. They are divided into phytoanticipins and phytoalexins. The former are low molecular weight compounds present in the plant before or produced after exposure from existed precursors, while the latter is synthesized and accumulated after exposure to abiotic agents (Iriti and Faoro, 2009).

For example, phytochelatins (PCs) are a class of heavy-metal binding peptides in all plants, which are involved in detoxification and homeostasis of heavy metals, and are functionally analogous to metallothioneins in animals and some fungi. Several studies have shown that PCs are induced after exposure to metals representing a unique biomarker of metal exposure and, potentially, of metal effect. In 1987, Grill and coworkers reported that PCs are selectively induced in *Rauvolfia serpentina* cell cultures after exposure to Ni^{2+}, Au^{+}, Bi^{3+}, Sn^{2+}, and the anions AsO_4^{-3} and SeO_3^{-2}, which was not known to induce metallothionein synthesis in humans and animals. Further, in 2006, Srivastava *et al.*, PC2, and PC3 were induced by 5.6 and 9.4 fold in *Hydrilla verticillata* after exposure to copper. Additionally, Gupta *et al.* 1995, reported the induction of PC2 and PC3 due to exposure to lead (2.5 µM). Furthermore, Pawlik-Skowron´ska´s (2001) PCs production was increased in *Stigeoclonium spp.* as a result of exposure to mining water containing a mixture of heavy metals (17 µM). Yin and coworkers reported that PC2 and PC3 concentrations increased in *Lemna aequioctialis* exposed to 20 µM of cadmium. Also, *Arabidopsis thaliana* was exposed to poly-vinylpyrrolidone-coated AgNPs (5 mg/L) and silver ions (Ag^{+}) 5 mg/L for 10 days resulting in an upregulation of 286 and 84 genes and downregulation of 81 and 53 genes by reference to control plants, respectively. The upregulated genes were primarily associated with the response to metals and oxidative stress while downregulated genes were more associated with response to pathogens and hormonal stimuli (Tehrani *et al.*, 2012).

Moreover, heavy metals could induce a stress effect on different metabolites pathways, which depends on the concentration and tolerance of heavy metals. For instance, the growth of Chives (*Allium schoenoprasum*) was inhibited on exposure to Ni, Co, and Cd at 0.25 mM concentrations. Further, Pb and Cr inhibited the seed germination and decreased the dry weight of chickpea (*Cicer auratinum*), with the increase in metal concentrations and time intervals. However, Cd did not significantly affect dry weight, leaf number, and total chlorophyll (a & b) concentration of a pea

at 20 μM concentration. Furthermore, Zn changed chlorophyll-a fluorescence and antioxidant system parameters in *Zea mays* with an increase in Zn doses (Sumiahadi and Acar, 2017).

Xenobiotic effects on metabolic pathways could be evaluated via measuring the activity of a target enzyme directly or through variations in metabolite concentrations (either upstream or downstream). Further, changes in metabolite concentrations provide the foundation for stressor-induced effects (Brain and Cedergreen, 2009). Furthermore, metabolites generated from root turnover induce microorganisms to biodegrade xenobiotics (Singer *et al.*, 2003).

A- Phenylpropanoid Pathway

Phenylpropanoids are a class of phenylalanine (essential amino acid) derivatives, which have simple phenols (e.g. organic acid) and polyphenolic compounds (e.g. stilbenes, and flavonoids). Phenylalanine, tyrosine, and tryptophan synthesize from the shikimate pathway. Phenylalanine and tyrosine are synthesized from prephenic and arogenic acid, whereas tryptophan from anthranilic acid. Shikimic and chorismic acid is the intermediate of phosphoenolpyruvate, from glycolysis, and erythrose 4-phosphate from the pentose phosphate pathway (Iriti and Faoro, 2009).

Flavonoids (phenylpropanoid) are a class of phenolic natural products comprising more than 4,500 compounds, which are catalyzed by the chalcone synthase enzyme. Their production was enhanced as a response to a variety of environmental stresses, including high light/UV, pathogen attack, wounding, low temperatures as well as deficiencies in nitrogen, phosphate, and iron.

Regarding simple phenolics, salicylate induces systemic acquired resistance (SAR) in plants, which has been linked to the microbial degradation of naphthalene, polycyclic aromatic hydrocarbons (Singer *et al.*, 2003). In 1998, Chen and Aitken showed that salicylate enhances the rate of removal of fluoranthene, pyrene, benz[a]anthracene, chrysene, and benz[a]pyren. Interestingly, ozone exposure (120-170 nmol/mol for 5 h) caused enhanced emission of methyl salicylate in tobacco (*Nicotiana tabacum*), which is a volatile salicylic acid derivative (Iriti and Faoro, 2009). Concluding, ozone activates at least two different signaling pathways, including a salicylic acid-dependent pathway that resembles the pathogen defense reactions, and a second one salicylic-independent, as a protective response to ozone.

The phenylpropanoids play a vital role in plant tolerance against different xenobiotic stresses due to their antioxidant power against free radicals (ROS) and as precursors of structural biopolymers, such as lignin. Lignin protects the membrane damages from lipid peroxidation (Iriti and Faoro, 2009). The key enzymes of lignin biosynthesis, shikimate dehydrogenase, phenylalanine ammonia-lyase (PAL), and cinnamyl alcohol dehydrogenase in the shikimate pathway increased up to 15 fold and 23 fold the control values in poplar (*Populus tremula* x *alba*) leaves under ozone exposure (60-120 nL/L, during a 14 h light period, for 1 month). Interestingly, the newly

synthesized lignin structurally differed from the control lignin. In tomatoes exposed to ozone (150 nL/L for 3 h), the biosynthesis of shikimate dehydrogenase, and PAL enzymes were exacerbated in the plant, while, cinnamyl alcohol dehydrogenase activity diminished. This was due to the different herbaceous and woody plant responses and acute or chronic ozone doses.

Despite decreasing the psoralen, bergapten, and other furanocoumarins in celery (*Apium graveolens*) after ozone (0.2 ppm for 24 h) exposure, their level in treated leaves increased rapidly at 120 h. Moreover, the powerful antioxidant flavonol kaempferol glycosides increased in soybeans (*Glycine max*), as well as resistance to manganese (Mn) toxicity in one soybean line. Also, a 16.2% increase in totals phenylpropanoid and a corresponding 9.9% increase of 10 compounds such as chlorogenic acid and catechin were reported in silver birch (*Betula pendula*) chronically exposed to ozone (Iriti and Faoro, 2009).

Furthermore, ozone caused the accumulation of isoflavonoids in soybeans. Additionally, it stimulated the induction of defense-related genes and up-regulated the key enzymes in the phenylpropanoid pathway in *Arabidopsis thaliana* exposed to (300 ppb ozone daily for 6 h), which increased phenylalanine ammonia-lyase (PAL) enzyme 3 fold than control plants. Also, ozone increased in PAL, and chalcone synthase early 3 fold and 1.2 fold, respectively, in *Petroselinum crispum* at (200 nL/L for 10 h) and in bean (120 nL/L for 4 h), which resulted in a 2 fold increase of total leaf furanocoumarins and flavone glycosides. However, it caused an increase in chalcone synthase and not in PAL due to competition for the same substrate. Also, it caused a decrease in coumaric acid, ferulic acid, gallic acid, and catechin, however, an increase in caffeic acid (Iriti and Faoro, 2009).

The production of isoflavonoid phytoalexins, including phaseollin, and kievitone protected beans from pathogen attack by *C. lindemuthianum.* The same was observed for medicarpin and pisatin, two isoflavonoid phytoalexins from alfalfa (*Medicago sativa*) and pea (*Pisum sativum*), respectively (Iriti and Faoro, 2009). Further, resveratrol protects the grapevine (*Vitis vinifera*) against *Botrytis cinerea* infection (grey mold).

The flavonoids can be affected by heavy metals (Brain and Cedergreen, 2009). In 2003, Babu *et al.* found that chalcone synthase was upregulated by exposure to a combination of copper and photosynthetically active radiation. Hence, flavonoids are known to act as scavengers of reactive oxygen species (ROS). Moreover, the hydroxyl groups in flavonoids can chelate heavy metals. Thus, flavonoids show promise as biomarkers of heavy metal exposure. Donnelly and coworkers reported that flavonoids could support the growth of some polychlorinated biphenyls (PCB)-degrading microorganisms and enhance PCB metabolism. Naringin provided the best growth substrate for *R. eutropha* H850 and fostered the greatest metabolic activity towards the 13 PCB tested. Myricetin also induced the16 of the 19 PCB degradation in *B. cepacia* LB400 (Donnelly *et al.*, 1994). Besides, coumarin induced degradation of 13 PCB by *Corynebacterium sp.* MB1.Thus,

the plant roots network is capable of distributing metabolites into the rhizosphere and inducing PCB degradation in microorganisms over long periods.

The cucumber was exposed to 2,4,6-trichlorophenol, chlorpyrifos, and oxytetracycline in the absence or presence of 24-epibrassinolide, a potent regulator of plant growth and stress tolerance for 10 days. They induced accumulations of both hydrogen peroxide and nitric oxide in cucumber root tips increasing the malondialdehyde content, which is an indicator of membrane lipid peroxidation. In presence of 24-epibrassinolide, it increased the content of reduced glutathione, free radical scavenging capacity, flavonoid content, and the activity and transcription of secondary metabolism-related enzymes (Ahammed *et al.*, 2017).

B- Alkaloids pathway

Alkaloids have toxicological, pharmacological, nutritional, and cosmetic activities, as well as ecological importance for plants e.g. cocaine, atropine, noradrenaline (or norepinephrine), adrenaline (epinephrine), papaverine, and morphine. They don't synthesize from a common pathway, however, the majority of them originate from amino acids e.g. ornithine, leucine, lysine, tyrosine, tryptophan, histidine, and phenylalanine (Hughes and Shanks, 2002). Ozone reduced the total alkaloids content in tobacco plants, especially nicotine, which resulted in the development of hornworm (*Manduca sexta*) larvae (Jackson *et al.*, 2000). Moreover, polyamines, important alkaloid precursors, play an important role in plant response to both biotic and abiotic stresses (Liu *et al.*, 2007). It has been reported that free and conjugated polyamines respond to ozone exposure with two different mechanisms: by inhibiting the ethylene biosynthesis and by direct ROS scavenging. Ethylene and polyamines share the same biosynthetic precursor, thus, the metabolic shift to polyamine biosynthesis can enhance ozone tolerance (Langebartels *et al.*, 2002). Generally, plants can shift pathways to produce alkaloids as a response to environmental stress.

C- Terpenoids (isoprenoids) pathway

Terpenoids represent the chemically and functionally most diversified class (both primary and secondary metabolites) in plants such as electron carriers (quinones), membrane constituents (sterols), saponins, vitamins (A, D, E, and K), plant hormones (side chain of cytokinins, abscisic acid, gibberellins, and brassinosteroids), photosynthetic pigments (chlorophyll, phytol, and carotenoids) and essential oils. The effects of the pollutant on either the sterol concentration and composition have been early reported in several plants due to their importance as cell membrane stabilizers (Grunwald and Endress, 1985). Tobacco leaves exposed to ozone showed increased lipid concentrations, but decreased levels of free sterols and triglycerides which are responsible for stabilizing the cell membranes. It has been reported that the β-hydroxy-β-methylglutaryl (HMG) - CoA synthase is increased in *Pinus sylvestris* exposed to ozone (0.3 μL/L for 8 h), which

is a key enzyme of isoprenoid biosynthesis (Wegener *et al.*, 1997; Iriti and Faoro, 2009). However, pine seedlings enhanced the biosynthesis of plastidial isopentenyl diphosphate and not the mevalonate synthesis in the cytosol due to long period treatment with ozone (Iriti and Faoro, 2009). Statins (e.g., lovastatin and atorvastatin) are inhibitors of 3-hydroxy-3-methyl glutaryl coenzyme-A reductase, the rate-determining step in the mevalonic acid (MVA) pathway in higher plants. Statins have demonstrated phytotoxicity and sterol reductions in *L. gibba*. Recently, studies have reported that sterol responses were in a concentration-dependent manner (Brain *et al.*, 2008a).

Saponins are glycosylated compounds widely distributed among plant families. They divide into three major groups, based on the structure of the aglycone: a triterpenoid, steroid, or steroidal glycoalkaloid (Iriti and Faoro, 2009). Avenacin is a saponin, which causes resistance root-infecting fungus *Gaeumannomyces graminis* in oat.

Chlortetracycline and oxytetracycline have been reported to induce chlorosis in higher plants via disruption of transcription/translation in the chloroplasts. Further, the treatment of cucumber cotyledons with 5 mM isoxaflutole (a carotenoid biosynthetic inhibitor) reduced carotenoid levels by nearly twice the proportion of fresh weight as a result of (p-hydroxyphenyl pyruvate dioxygenase) enzyme inhibition.

Biogenic volatile organic compounds consist mainly of isoprenoids (particularly hemi-, mono- and sesquiterpenes) synthesized during cell growth, in response to several kinds of stresses (such as light intensity, temperature, water supply, and pollutants), and mediating plant-arthropod interactions (Kesselmeier and Staudt, 1999). Their emission can be triggered by exposure to high ozone concentrations due to their antioxidant activity (Grote and Niinemets, 2008). Besides, their composition could be modified due to chronic exposure. Thus, ozone weakens directly plant defense responses against arthropods (Freiwald *et al.*, 2008).

Further, in 2001, Hernandez, *et al* reported that orange peel, ivy leaves, pine needles, or eucalyptus leaves supplement resulted in 105 times more biphenyl utilizers (10^8 / g) than unmodified soils (10^3 / g) simultaneously inducing Aroclor 1242 degradation. Also, pine needles and orange peel induced Aroclor 1248 loss in rhizospheres of various plant species by 44-55% and 54-59%, respectively (Dzantor and Woolston, 2001). Carvone and limonene were used to stimulate biodegradation of Delor 103 by *Pseudomonas stutzeri* (Tandlich *et al.*, 2001). Park and coworkers (1999), concluded that carvone might induce a different degradative pathway.

Dust pollution causes stoma closure and diminished CO_2 flux. The essential oil yield and composition of *Cistus albicans* and *Pinus pinea* from rural fields showed an increase in the yield but unaltered in composition than quarry work (Figueiredo *et al.*, 2008).

There is still a gap of knowledge on the effect of xenobiotics on the plant metabolites pathway. Hence, the reference standards of plant metabolites are unavailable. The untargeted metabolomics approach is a good choice to enrich this topic and find new insights.

2. Research significance, goals, and hypotheses

Recently, research on plant metabolomics has been extremely active due to a large number of metabolites with important functions in plant ecology and the protection against stress conditions (Rochfort, 2005). Furthermore, the assimilated pharmaceuticals are not excreted by the plants but rather stored in vacuoles inside their tissues. Phytoremediation research provided clear knowledge about the uptake of pollutants by the plant and their detoxification (Adamia *et al.*, 2018; Ohlbaum *et al.*, 2018). For example, in 2014, Bartha and coworkers reported the uptake of diclofenac by *Typha* root. They revealed a new transformation product exclusively in plants, not in bacteria, fungi, or even mammals (Bartha *et al.*, 2014). Another study that was done by Huber and *et al.*, 2009, conveyed that cell cultures of horseradish root could take up acetaminophen and degrade it (Huber *et al.*, 2009). However, a major problem related to this kind of research was not concerning the plant metabolites pathways. Knowledge about the changes in the plant metabolomics, as well as whether the pollutants change the plant biosynthesis cycle and whether the pollutants can push the biosynthesis cycle in a specific direction is still a growing research field. Recently, a study showed that incubation with DCF (20 µg/L) can already act as a signal, and affect the circadian expression of reactive oxygen species response selected genes in lettuce (Bigott *et al.*, 2021). Also, Sivaram and coworkers, 2019, reported a high impact on the galactose metabolism, aminoacyl tRNA biosynthesis, and TCA cycle in maize leaves exposed to benzo[a]pyrene, pyrene, individually, and their mixture.

Moreover, higher vascular plants used in phytoremediation (CW), have been described to have nutrition and medicinal values (Sánchez *et al.*, 2018; Xi *et al.*, 2018), which might introduce another usage possibilities of the plant's large biomass than biofuels (bioenergy crops e.g. *Salix spp*) (Aronsson and Perttu, 2001). In natural Indian wetlands, 51 edible plant species were grown, which were used by indigenous people for food and medicinal purposes (Jain *et al.*, 2011). Furthermore, ornamental plants have been used in constructed wetlands to enable wastewater treatment with the additional benefit of commercial value (Rahman *et al.*, 2020).

These new approaches require optimizing a novel untargeted screening strategy, which investigates the pathway changes of plant metabolites. This could be accomplished by studying the influence of pharmaceuticals (e.g. diclofenac and carbamazepine, as well as antibiotics) on the plant metabolites and metabolomes. Accordingly, new insights are given using accurate and high-resolution mass spectrometer as well as novel data evaluation workflows. This thesis will afford a comprehensive untargeted screening strategy to observe the changes in the metabolomics of the plants. Additionally, it will allow the monitoring of constructed wetlands plant metabolomics, and extend the knowledge about their biosynthetic pathways changes, as well as their usage as a source for antioxidant compounds. Through the following objectives and hypotheses:

4. Isolation and identification of the secondary metabolites from plants

In-plant metabolomics, sample preparation is a fundamental step because it greatly affects the precision of the results. Also, the storage condition of samples affected the quality and quantity of the metabolites due to the enzymatic activity. Samples have to be harvested uniformly and rapidly to avoid enzymatic changes (Salem *et al.*, 2020). According to the Metabolomics Standards Initiative (MSI), plant extract should be obtained with minimum pretreatment, to prohibit the loss of metabolites. Thus, we hypothesized that

Hypothesis #1 Extraction strategy for *Lemna minor* and *Phragmites australis* can influence their metabolites chromatographic fingerprints.

To test this hypothesis, *Lemna minor* and *Phragmites australis* were extracted with different solvent(s) (mixtures). Also, the storage effect was investigated on *Lemna minor* samples. The extracts were analyzed using the extend-polarity chromatography RPLC-HILIC-TOF-MS to investigate the changes due to the storage condition and to determine the appropriate extraction solvent for a wide polarity range of metabolites. The corresponding study is presented in **Chapter 4.**

5. Analytical method development and statistical data analysis for plant metabolites using RPLC-HILIC-MS

The plant metabolites number is between 100,000 and one million metabolites with differences in function depending on their physicochemical properties. There is a wide range of analytical metabolomics approaches using different systems for instrumental analysis such as GC-MS, LC-MS, and NMR. Recently, systems with different stationary phases were used in plant metabolomics analysis which allows the separation of different metabolites in single runs (Hemmler *et al.*, 2018). QTOF-mass spectrometers are the most widely used mass spectrometer in metabolomics because they offer the highest scan rates, high efficiency, resolution, and mass accuracy (Gika *et al.*, 2010).

The RPLC-HILIC serial coupling-QTOF provides a large amount of data, consequently adequate workflow to process and analyze it. Although metabolomics data processing is a straightforward process, the metabolite identification process continues to be critical. From peak picking to identify the metabolites, the workflow contains numerous steps, which require to be precise, compatible, and with less effort and time-consuming. All the steps require quality control procedures for trustworthy metabolomics analysis outcomes, therefore, we hypothesized that

Hypothesis #2 Analytical workflows, applied in the determination of trace organic compounds in the aquatic environment, can be used for plant metabolites detection.

To test the hypothesis, different workflows were applied to investigate *Lemna minor* samples according to the mass spectrometer type. The different data sets (obtained with different MS) were compared and categorized statistically. The workflow's different steps were validated using

internal and/or external standards. The obtained features were characterized and/or identified using the different databases, e.g. the PLANT-IDENT database in the FOR-IDENT platform and others. The responding study is presented in **Chapter 5**.

6. Study the effect of the different environmental pollutants on plant metabolism

Metabolomics studies are essential to characterize significantly the metabolites profiles changes during biochemical reactions. Owing to the fact that, metabolomics is the approach for the overall investigation of metabolite variations in biological systems. Consequently, metabolic profiling can be used as a robust tool to discuss the metabolic response of plants to environmental disorders, such as xenobiotics stress

7. Identification of changed different secondary metabolites biosynthetic pathways

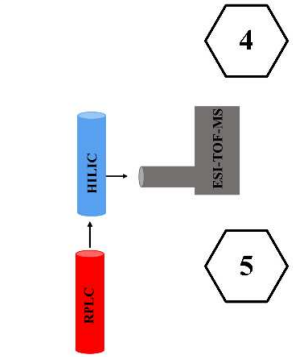
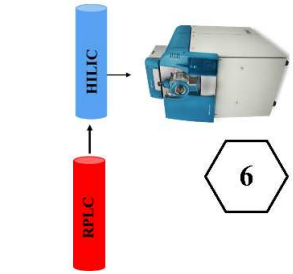
Moreover, pharmaceuticals such as DCF and CBZ, as well as antibiotics, were detected at a high level in the influent, effluent, and surface water (Table 1). They were incompletely degraded. Additionally, their transformed products were more toxic than the parent compounds. Thus, they were considered recalcitrant xenobiotics (Fischer *et al.*, 2020). These findings induce concerns regarding the chronic exposure of plants in constructed wetlands, which might lead to an alteration in its physiological characteristics, especially the molecular interaction and biochemical properties due to the adverse effect caused due to the acute exposure results reported recently in literature (Alkimin *et al.*, 2019; Kostopoulou *et al.*, 2020). These changes in plant metabolic profile could be investigated through two independent workflows: (1) the revalent metabolites identification, and (2) the statistical identification of the metabolomics indicators. Univariate and multivariate statistical strategies are typically used in a chemometrics style. They can be used as standard approaches to extract relevant information from complex datasets (Worley and Powers, 2013; Gromski *et al.*, 2014; Lamichhane *et al.*, 2018; Prinsloo and Vervoort, 2018). Thus, we hypothesized that

Hypothesis # 3 The presence of pharmaceuticals such as diclofenac and carbamazepine in the aquatic environment can affect the biosynthetic pathways of *Lemna minor and Phragmites australis*.

To test the hypothesis, different experiments were conducted using *Lemna minor* and *Phragmites australis*. *Lemna minor* and *Phragmites australis* were incubated with DCF and CBZ for 4 days, individually. Also, *Lemna minor* was incubated with both of them for 4, 8, 12 days under the same concentrations, individually**.**

Moreover, another experiment was conducted with *Lemna minor*, which was incubated with DCF, CBZ, SMX, TRIM, and a mixture of them at environmental concentration for 5 days to mimic the condition in the constructed wetlands. The corresponding studies are presented in **Chapters 6 and 7**.

Table 2. Structure of the dissertation: Research objectives, methodology, hypotheses, and publications.

Chapter	Objectives	Hypothesis	Publication
	Isolation and identification of the secondary metabolites from plants	Extraction strategy for *Lemna minor* and *Phragmites australis* can influence their metabolites chromatographic fingerprints	**Paper II** Wahman et al., 2020, *Lemna minor* Studies under various storage periods using extended-polarity extraction and metabolite non-target screening analysis, Journal of Pharmaceutical and Biomedical Analysis 188:113362
	Analytical method development and statistical data analysis for plant metabolites using RPLC-HILIC-MS	Analytical workflows, applied in the determination of trace organic compounds in the aquatic environment, can be used for plant metabolites detection	**Paper I** Wahman et al., 2019, Plant Metabolomics Workflows Using Reversed-Phase LC and HILIC with ESI-TOF-MS Current Trends in Mass Spectrometry, LCGC. **Paper IV** Wahman et al., 2021, Untargeted analysis of *Lemna minor* metabolites workflow and priorization strategy comparing highly confident features between extracts and different mass spectrometers (submitted to Metaboltites).
	Study the effect of the different environmental pollutants on plant metabolism	The presence of pharmaceuticals such as diclofenac and carbamazepine in the aquatic environment can affect the biosynthetic pathways of *Lemna minor* and *Phragmites australis*	**Paper III** Wahman et al., 2021, Untargeted Metabolomics Studies on Drug-Incubated *Phragmites australies* Profiles, Metaboltites 11(1):2 **Paper V** Wahman et al., 2021, The changes in *Lemna minor* metabolomic profile: A response to diclofenac iuncubation, Chemosphere 287, 13207.
	Identification of changed different secondary metabolites biosynthetic pathways	The presence of pharmaceuticals such as diclofenac and carbamazepine in the aquatic environment can affect the biosynthetic pathways of *Lemna minor* and *Phragmites australis*	

3. Materials and Methods

3.1. Reagents and chemicals

LC-MS grade methanol and water were obtained from VWR, Darmstadt, Germany. Formic acid and amino acids standards were obtained from Sigma-Aldrich, Steinheim, Germany. Ethanol was purchased from AppliChem, Darmstadt, Germany. Gallic acid and sodium carbonate were obtained from Acros Organics, Niederau, Germany for reducing content determination. Folin-Ciocalteau reagent was purchased from Merck Chemicals, Darmstadt, Germany. Quercetin dehydrate (≥ 95%, high-performance liquid chromatography (HPLC)) was purchased from Enzo Life Sciences GmbH, Lörrach, Germany. Diclofenac (>99%) was obtained from Cayman Chemical Company, Ann Arbor, Michigan, USA. Glyphosate (100 quality level, HPLC), gabapentin (200 quality level, HPLC), monuron (100 quality level, HPLC), chloridazon (100 quality level, HPLC), carbetamide (100 quality level, HPLC), metobromuron (100 quality level, HPLC), sotalol (≥98%), quinoxyfen (100 quality level, HPLC), metconazol (100 quality level, HPLC) and fenofibrate (≥99%) 6-Amino-1,3-dimethyl-5-(formylamino)uracil (certified reference material), chlortoluron (100 quality level, HPLC), famotidine (100 quality level, HPLC), vidarabine (certified reference material), etilefrine (pharmaceutical primary standard), and 2,4-diamino-6-(hydroxymethyl)pteridine hydrochloride (100 quality level, HPLC), apigenin (≥ 97%), vitexin (≥ 95%), flavone (≥ 99%), DL-Ala-DL-Ala (98%), nicotinic acid (≥ 99.5%), nicotinamide (≥ 99%), and galangin (≥ 97%) were obtained from Sigma, Darmstadt, Germany. Metformin (300 quality level, HPLC) and acacetin were obtained from Fluka, Buchs, Switzerland. Furthermore, chlorbromuron (99.24%) and diazinon (99.53%) were obtained from Dr. Ehrenstorfer, Augsburg, Germany. Carbamazepine, 2,3-dihydro-2,3-dihydroxycarbamazepine, 10,11-dihydro-10,11-dihydroxy-carbamazepine, 10,11-dihydro-10-hydroxy-carbamazepine, 9-acridine carboxaldehyde and carbamazepine-10,11-epoxide were kindly provided by the German Research Center for Environmental Health, Comparative Microbiome Analysis (COMI), Helmholtz Centrum of Munich, Munich, Germany. Apigenin-7-glucoside, Apigenin-5-glucoside, apigenin-6-arabinoside-8-glucoside, Apigenin-6,8-di-glucoside, apiin, orientin, kaempferol-7-rhamnoside, peonidin, norwogonin, luteolin-3,7-di-glucoside, 6-methoxy flavone, 4-methoxy cinnamic acid, naringenin-7-glucoside, isovitexin, tricin, quercetin-3-glucoside, saponarin, myricetin, 5-hydroxy-6-methoxy flavone, chrysoeriol, isovitexin, robinetin, umbelliferone were kindly provided by Center of Life and Food Science Weihenstephan, Biotechnology of Natural Products, Technical University of Munich.

3.2. Plant material

A-*Lemna minor* L. was cultivated in aquaria according to (Obermeier *et al.*, 2015) with minor modifications. Plants were grown at 23 °C with a photoperiod of 8-16 h and an average light intensity of 43 µ mol/m/s. *Lemna* fronds were subcultured every two weeks in 24 L of Steinberg

medium made up of (in mg/L) 85 $NaNO_3$, 13.4 KH_2PO_4, 75 $MgSO_4 \cdot 7\ H_2O$, 36 $CaCl_2 \cdot 2\ H_2O$, 20 Na_2CO_3, 1 H_3BO_3, 0.2 $MnCl_2 \cdot 4\ H_2O$, 0.01 $Na_2MoO_4 \cdot 2\ H_2O$, 0.05 $ZnSO_4 \cdot 7\ H_2O$, 0.005 $CuSo_4 \cdot 5\ H_2O$, 0.01 $Co(NO_3)_2 \cdot 6\ H_2O$, 0.84 $FeCl_3 \cdot 6\ H_2O$ and 1.4 Na_2-EDTA$\cdot 2\ H_2O$. The plants were then harvested, rinsed briefly in tap water, dried with lint-free tissue paper, and finally frozen in liquid nitrogen. Samples were kept at -80°C until further processing. The fresh samples that were frozen for several days were from the same plant generation.

B-*Phragmites australis* (Cav.) Trin. ex STEUD. plants were grown in semi-hydroponic conditions in the greenhouse as described by Sauvêtre and Schröder earlier, in 2015. Plants were grown in Hoagland solution made of (in mg/L) 472.30 $Ca(NO_3)_2 \cdot 4H_2O$, 202.22 KNO_3, 492.96 $MgSO_4 \cdot 7H_2O$, 68.04 KH_2PO_4, 80.04 NH_4NO_3, H_3BO_3, 1.8 $MnCl_2 \cdot 4H_2O$, 0.2 $ZnSO_4 \cdot 7H_2O$, 0.1 $CuSO_4 \cdot 5H_2O$, 0.025 $NaMoO_4$ and 3.67 FeNa-Ethylenediaminetetraacetic acid. Plants (approximately 0.8 m in height) of uniform size were selected and placed into individual pots containing 2 L of spiked Hoagland solution Each pot contained one plant and was arranged in the greenhouse following a completely randomized design. The nutrient medium was spiked with a stock solution to reach the desired final concentration (diclofenac, 10 and 100 µM and carbamazepine, 10 and 50 µM, respectively). Control plants growing in Hoagland solution (spiked with the same amount of solvent as incubated plants) were used to obtain a reference plant matrix. Two pots were set up for each of the four exposure concentrations. Each assay consisted of duplicates arranged in the greenhouse following a randomized design. To compensate for water losses by evapotranspiration, distilled water was added daily to the pots for a final volume of 2 L. Plants were exposed for 4 days before they were harvested. Harvested material was divided into roots, rhizomes, and leaves, frozen in liquid nitrogen, and stored at -80 °C until further processing. *P. australis* was kindly provided by the German Research Center for Environmental Health, Comparative Microbiome Analysis (COMI), Helmholtz Center Munich, Munich, Germany.

3.3. Extraction

A- *Lemna minor* samples were extracted: 1) fresh plants LF, 2) plants had been frozen for several days at -80ºC, LFD, and 3) plants frozen for 6 months at -80ºC (frozen months; LFM). For each sample, 500 mg of freeze-dried and milled powder was extracted with a) 100% methanol (MeOH), b) acidic 90% MeOH (MeOH-water-formic acid (FAC) (90:9.5:0.5, v/v/v)), c) MeOH-water (50:50, v/v), and d) 100% water (H_2O), respectively. Additionally, the frozen plant material LFM was extracted with 100% ethanol. Extracts with plant powder were sonicated (Sonorex super RK 106, Bandelin, Berlin, Germany) for 10 minutes at 4ºC. Although this extraction method uses an ultrasonic bath, extracts with plant powder were sonicated for 10 minutes at 4ºC and a frequency of 35 kHz to protect the samples from heating and loss of thermolabile compounds. The samples were then centrifuged (Z 200 A Universal Compact Centrifuge, Hermle LaborTechnik GmbH, Germany) at 1500 rpm for 20 min and the supernatants transferred to clean glass test tubes. The extraction process was triplicated in

identical experimental conditions. Finally, the extracts were evaporated to dryness (using a SpeedVac, Fischer Scientific, Göteborg, Sweden) and dissolved in MeOH-H_2O (50:50, v/v) (Kaufmann *et al.*, 2016; Kaufmann *et al.*, 2017), as shown in (Fig. 1).

B- *P. australis* leaves, rhizomes, and roots, respectively, were collected, frozen under liquid nitrogen. Then, the samples were freeze-dried and milled (Retsch S1 planetary ball mill, Retsch GmbH, Haan, Germany). Duplicates of each plant part were extracted with (a) 100% MeOH, (b) acidic 90% MeOH, (c) 50% MeOH and (d) 100% water (H_2O), respectively. The solvents containing 500 mg plant powder were sonicated for 10 min at 4 °C with a 35 kHz frequency. Then, samples were centrifuged at 1500 rpm/261.6× *g* for 20 min and the supernatants were transferred to clean glass test tubes. The extraction process was triplicated in identical experimental conditions. Finally, the extracts were evaporated to dryness and dissolved in (50:50 (*v/v* %)) MeOH: H_2O. Samples and standards were filtered (with 22-µm filter, Analytics Shops, Munich, Germany)

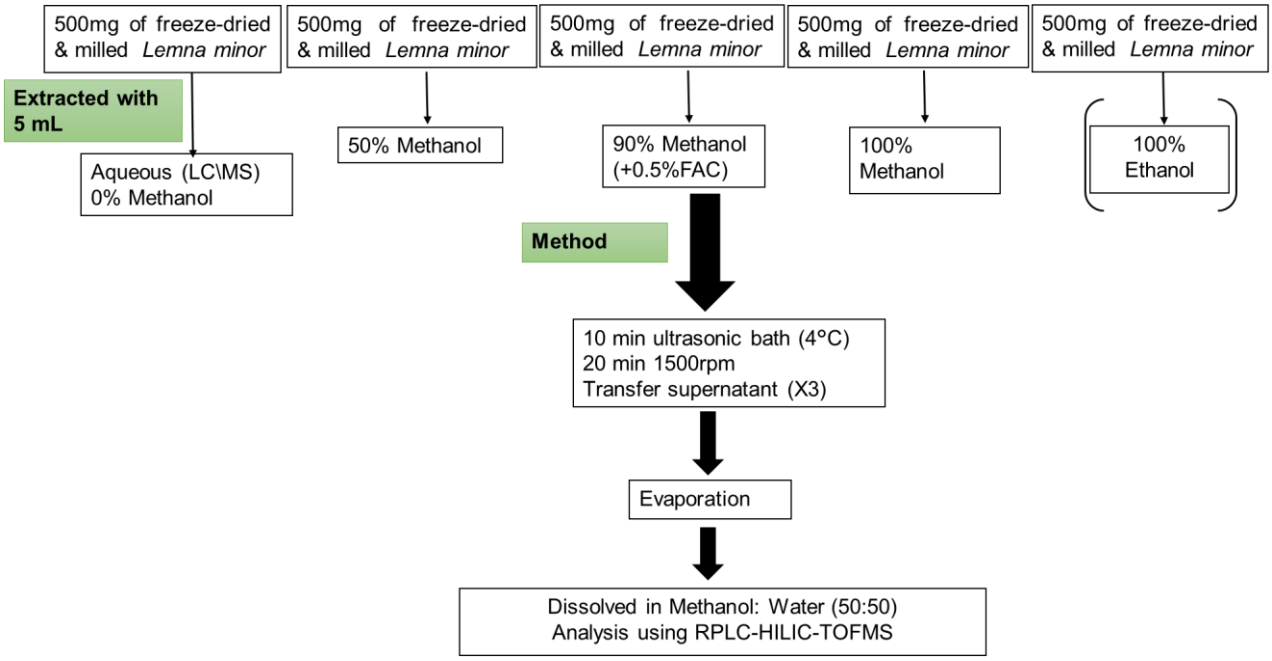


Figure 1. The extraction method

C- Steinberg media was extracted with SPE OASIS HLB cartridges (Water, Milford, MA, USA) (6 cc/200 mg) with 4 mL of 100% MeOH. After evaporation, the residue was suspended in (50:50 (*v/v* %)) MeOH: H_2O. Samples and standards were filtered (with a 22-µm filter, Analytics Shops, Munich, Germany).

3.4. Reducing contents

100 µL of Folin-Ciocalteau reagent-H_2O (1:9, v/v) added to the 20 µL sample (Gallic acid at different concentrations (to plot the calibration curve) or plant extracts dissolved in MeOH-H_2O (50:50, v/v)) with and mixed well. 100 µL of sodium carbonate-H_2O (7.5 g: 100, w/v)

solution were then added and the mixture incubated for 60 min in darkness at 25°C. The absorbance was measured in triplicate at 765nm. The reducing content was calculated from the calibration curve as gallic acid equivalent (GAE) (Kaufmann *et al.*, 2017).

3.5. Instruments

3.5.1. Chromatographic system

The Agilent 1260 Infinity LC-systems comprised an autosampler, column oven, two columns, and two binary pumps. An online degasser, a mixing chamber, and a UV detector were used to perform reversed-phase (RPLC) and zwitterion hydrophilic interaction liquid chromatography (HILIC) in its serial coupling. The first binary pump and the autosampler were connected to the reversed-phase separation column, a Poroshell 120 EC-C18 (50.0 × 3.0 mm, 2.7 μm; Agilent Technologies). The outlet of this column was connected to the HILIC column, which was a ZIC-HILIC column (150 × 2.1 mm, 5 μm, 200 Å; Merck Sequent, Umea, Sweden). The columns were coupled through a T-piece (Upchurch, IDEX Europe GmbH, Erlangen, Germany). The third port of the T-piece was connected to the HILIC flow pump. The injection volume was 10 μL. The mobile phase of the serial coupling (RPLC-HILIC-ToF-MS) was employed as follows: Solvent A: 10 mM ammonium acetate in 90:10 (v/v) water–acetonitrile; solvent B: 10 mM ammonium acetate in 10:90 (v/v) water–acetonitrile; solvent C: acetonitrile; solvent D: water. Further mobile phase conditions are summarized in Table 3 (Greco *et al.*, 2013)

Table 3. Mobile phase condition of RPLC-HILIC-TOF-MS

Binary Pump 1				**Binary Pump 2**			
Time (min)	**Flow rate (mL/min)**	**A%**	**B%**	**Time (min)**	**Flow rate (mL/min)**	**C%**	**D%**
0	0.05	100	0	0	0.4	100	0
7	0.05	100	0	6	0.4	100	0
12	0.05	50	50	13	0.4	60	40
13	0.1	50	50	32	0.4	60	40
22	0.1	0	100	33	0.8	100	0
32	0.1	0	100	53	0.8	100	0
33	0.1	100	0	54	0.4	100	0
53	0.1	100	0	58	0.4	100	0
54		0.05		100		0	
58		0.05		100		0	

3.5.2. Mass spectrometric systems

A-System A

Samples were detected with a 'time-of-flight' mass spectrometer (Agilent Technologies, Waldbronn, Germany), equipped with the Jet Stream ESI interface. Ions were detected in positive ionization mode with a mass range of 50-2100 Dalton. The resolution of the instrument was better than 10,000 at m/z 922. The parameters were as follows: 325 ℃ gas temperature, 10 L/min drying gas flow, 325 ℃ sheath gas temperature, 7.5 L/min sheath gas flow, 45-psi nebulizer operating pressure, and 100 V fragmentor voltage.

B-System B

Lemna metabolic profile mass spectra were obtained on a Triple TOF 5600 system (AB SCIEX triple TOF 5600, Darmstadt, Germany) with a Duospray ion source and a Turbolonspary ESI probe in positive ion mode. Ions were detected in positive ionization mode with a mass range up to *m/z* 40,000. The mass spectrometric parameters were set as the following parameters: ISVF (ion spray voltage floating) 2000 kV and turbo spray temperature, 650°C. DP clustering potential (DP), and collision energy (CE) were set to 46, and 40 V,

respectively. The nebulizer and the auxiliary gas were both nitrogen. Also, the nebulizer gas (gas 1), the heater gas (gas 2), and the curtain gas were set to 44, 50, and 29 psi, respectively.

C-System C

System C consisted of RPLC reverse-phase Discovery column (25cm× 4.6 mm; 5 µm particle size; Sigma Aldrich) connected to Q Exactive Focus Orbitrap (Thermo Fischer Scientific), equipped with a heated electrospray ionization probe (HESI) source. The MS was turned to a mass resolution of 50,000 (FWHM, m/z 200) with a mass spectrum range of 50-600 m/z. The analyses in positive and negative ionization mode were performed with a 3.40 kV and 3.40 kV, 45 V and -50 V, 90 V, and -120 V, and 26 V and -25 V for spray, capillary, a tube lens, and skimmer voltages, respectively. The capillary and heater temperature was 300 °C.

3.6. Quality control of the RPLC-HILIC-ESI-TOF-MS System

3.6.1. External standards

The robustness and reproducibility of the RPLC-HILIC-ESI-TOF-MS system were tested with a different standard mixture containing 13 different reference standards in a 20 µM final concentration. The mixture consisted of metformin, glyphosate, gabapentin, monuron, chloridazon, carbetamide, metobromuron, sotalol, chlorbromuron, diazinon, quinoxyfen, metconazol, and fenofibrate. Also, the mixture containing kaempferol, rutin, taxifolin, apigenin, resveratrol, galangin, diclofenac, flavone, vitexin, and quercetin was injected. *Lemna* and *Phragamites* metabolic profile was obtained using a serial coupling of reversed-phase liquid chromatographic column and the hydrophilic interaction liquid chromatographic column, which was connected to time-of-flight' mass spectrometer (TOF-MS). The injection volume was 10 µL (Wahman *et al.*, 2021). The mixtures were injected at the beginning/end of the experiment series and fixed intervals during the experiment. It was injected after each extraction batch. The absolute variation between the literature monoisotopic mass and the mean of measured isotopic masses (Δ ppm) was computed according to the following equation:

$$\Delta ppm = (monoisotopic\ mass\ of\ standard - mean\ of\ standard\ masses) \div monoisotpoic\ mass\ of\ standard \times 10^6 \quad (1)$$

Moreover, the standard deviations (SD) of RT and Relative standard deviations (RSD) were calculated:

$$\%of\ RSD = SD\ of\ compound\ RTs\ in\ different\ injection \div mean\ of\ compound\ RTs \quad (2)$$

The results are summarized in (Tables 4-7). The absolute mass deviation ranged from 0.2 (Da) to 7 (Da). The RT standard deviation was less than 1%. Moreover, the relative standard deviation (%RSD) ranged from 0.5% to 3.6%. Thus, the results indicate the accuracy, repeatability, and reproducibility of the LC system as previously reported in the investigation of plant metabolites in *Lemna minor* samples (Wahman *et al.*, 2019).

Table 4. List of standards used in the quality control of the analytical system. Monoisotopic mass in the literature, the monoisotopic mass of the first, second, and third injections, mean monoisotopic of the three injections, and variation *(*Δppm) between them are listed.

Name	Mono isotopic Mass (Da) (L)	Mono isotopic Mass (Da) (1st inj.)	Δppm	Mono isotopic Mass (Da) (2nd inj.)	Δppm	Mono isotopic Mass (Da) (3rd inj.)	Δppm
Flavone	222.0680	222.0684	-1.44	222.0684	-1.44	222.0685	-1.89
Resveratrol	228.0786	228.0784	1.07	228.0782	1.95	228.0779	3.26
Apigenin	270.0528	270.0522	2.37	270.0524	1.57	270.0525	1.2
Galangin	270.0528	270.0517	4.16	270.0518	3.79	270.0524	1. 6
Kaempferol	286.0477	286.0475	0.83	286.0474	1.18	286.0472	1.88
Diclofenac	295.0167	295.0159	2.66	295.0158	2.99	295.0158	2.99
Quercetin	302.0427	302.0419	2.49	302.0419	2.49	302.0422	1.5
Taxifolin	304.0583	304.0576	2.31	304.0576	2.31	304.0575	2.64
Vitexin	432.1056	432.1062	-1.28	432.1062	-1.28	432.1064	-1.74
Rutin	610.1534	610.1529	0.79	610.1532	0.30	610.1533	0.14

Table 5. Standards used in the quality control of the analytical system. RT in the first, second and third injections, mean RT of the three injections, standard deviation, and relative standard deviation are listed

Name	Mean RT (Min) (1^{st} inj.)	Mean RT (Min) (2^{nd} inj.)	Mean RT (Min) (3^{rd} inj.)	SD	Mean RT (Min)	%RSD
Flavone	29.4	29.5	29.4	0.02	29.4	0.09
Resveratrol	25.6	25.6	25.6	0.04	25.6	0.14
Apigenin	26.5	26.5	26.5	0.014	26.5	0.06
Galangin	26.5	26.6	26.5	0.02	26.5	0.09
Kaempferol	26.7	26.8	26.8	0.03	26.8	0.12
Diclofenac	26	26	25.9	0.05	26	0.18
Quercetin	26.	26	26	0.04	26	0.14
Taxifolin	24.2	24.3	24.2	0.08	24.2	0.31
Vitexin	25.6	25.6	25.6	0.04	25.6	0.18
Rutin	26.5	26.5	26.5	0.01	26.5	0.14

Table 6. The standard compounds of the quality control external calibration mixture, monoisotopic mass in the literature (L), monoisotopic in different injection and the mean of them, the variation between monoisotopic mass in the literature (L), and the mean of measured monoisotopic mass, and mean mass standard deviation (SD) are listed.

Name	Mono isotopic Mass (Da) (L)	Mono isotopic Mass (Da) (1^{st} inj.)	Mono isotopic Mass (Da) (2^{nd}inj.)	Mono isotopic Mass (Da) (3^{rd}inj.)	Mono isotopic Mass (Da) (4^{th}inj.)	Mono isotopic Mass (Da) (5^{th} inj.)	Mono isotopic Mass (Da) (6^{th}inj.)	Mean Mono isotopic Mass (Da)	Δ ppm	SD
Metformin	129.1014	129.1006	129.1001	129.1000	129.1005	129.1015	129.1014	129.1007	5.78	0.001
Glyphosat	169.0140	169.0132	169.0134	169.0138	169.0123	169.0120	169.0136	169.0131	5.62	0.001
Gabapentin	171.1259	171.1252	171.1254	171.1254	171.1231	171.1231	171.1253	171.1246	7.81	0.001
Monuron	198.0560	198.0566	198.0568	198.0569	198.0563	198.0564	198.0564	198.0566	-2.96	0
Chloridazon	221.0356	221.0359	221.0362	221.0361	221.0355	221.0359	221.0355	221.0359	-1.22	0
Carbetamid	236.1161	236.1176	236.1179	236.1180	236.1164	236.1169	236.1169	236.1173	-5.1	0.001
Metobromuron	258.0004	258.0012	258.0014	258.0010	258.0006	258.0025	258.0060	258.0021	-6.54	0.002
Sotalol	272.1195	272.1193	272.1193	272.1192	272.1192	272.1211	272.1195	272.1196	-0.55	0.001
Chlorbromuron	291.9615	291.9616	291.9614	291.9613	291.9613	291.9614	291.9614	291.9614	0.21	0
Diazinon	304.1010	304.1010	304.1019	304.1014	304.1004	304.1001	304.1022	304.1012	-0.42	0.001
Quinoxyfen	306.9967	306.9961	306.9961	306.9976	306.9948	306.9964	306.9967	306.9963	1.32	0.001
Metconazol	319.1451	319.1456	319.1452	319.1443	319.1452	319.1462	319.1451	319.1453	-0.46	0.001
Fenofibrat	360.1128	360.1135	360.1141	360.1137	360.1132	360.1141	360.1139	360.1138	-2.58	0

Table 7. The standard compounds of the quality control external calibration mixture, the single RT, mean RT of the different injections, mean RT standard deviation (SD), and relative standard deviation (RSD) of the standards are listed.

Name	RT (Min) (1^{st} inj.)	RT (Min) (2^{nd}inj.)	RT (Min) (3^{rd}inj.)	RT (Min) (4^{th}inj.)	RT (Min) (5^{th} inj.)	RT (Min) (6^{th}inj.)	Mean RT (Min)	SD	% RSD
Metformin	14.3	14.3	14.	14.3	14.3	14.3	14.3	0.09	0.7
Glyphosat	13.7	13.7	14.4	13.9	14.03	13.8	13.9	0.25	1.8
Gabapentin	7.6	7.5	7.5	7.3	7.85	7.6	7.6	0.19	2.5
Monuron	24.3	24.2	24.3	24.6	24.6	24.2	24.4	0.17	0.7
Chloridazon	22	21.9	22	21.5	21.4	21.85	21.8	0.28	1.3
Carbetamid	23.9	23.8	23.8	22.7	22.7	23.9	23.5	0.62	2.6
Metobromuron	26.3	26.2	26.2	25.7	25.9	26.1	26	0.22	0.9
Sotalol	9.6	9.2	9.1	9.2	9.8	9.9	9.5	0.34	3.6
Chlorbromuron	26.9	27.6	27.7	27.4	27.7	28.	27.6	0.37	1.3
Diazinon	33.5	33.5	33.4	33.1	33.4	34.3	33.5	0.42	1.3
Quinoxyfen	35.2	35.5	35.2	35.2	34.5	34.5	35	0.41	1.2
Metconazol	28.6	28.6	28.6	28.7	28.5	28.3	28.5	0.14	0.5
Fenofibrat	32.9	33.1	32.9	33.2	33.2	33.4	33.1	0.19	0.6

3.6.2. Internal standards

Each sample and blanks were spiked with a standard mixture of 12 substances. The mixture consists of 6-amino-1,3-dimethyl-5-(formylamino) uracil, chlortoluron, famotidine, vidarabine, etilefrine, monuron, carbetamide, metobromuron, sotalol, chlorbromuron, metconazol, and 2,4-diamino-6-(hydroxymethyl) pteridine hydrochloride to obtain a final concentration of 5µM each. The absolute variation between literature monoisotopic mass and the mean of measured isotopic masses is shown as Δ ppm. The results are summarized in (Table 7). The absolute mass deviation ranged from 0.09 (ppm) to 2.81 (ppm). The RT standard deviation was less than 0.3. Moreover, the relative standard deviation (%RSD) was less than 2 % as shown in (Table 8). These parameters were used to correlate the features in different samples.

Table 8. The compounds of the internal standard calibration mixture. Mean monoisotopic mass of the different compounds in Daltons**.** Mean mass standard deviation (SD). Mean RT of the different compounds in minutes, mean RT standard deviation (SD), and relative standard deviation (RSD) of the standards were listed.

Component Name	Mean Mass (Da)	SD	RT Mean (Min.)	SD RT	%RSD
2,4-Diamino-6-(hydroxymethyl)pteridine	193.0837	0.0003	8.92	0.09	1.04
6-Amino-1,3dimethyl-5-(formylamino)uracil	199.0831	0.0004	6.53	0.09	1.31
Carbetamid	237.1239	0.0004	26.18	0.03	0.09
Chloridazon	222.0437	0.0006	24.9	0.03	0.12
Chlorotoluron	213.0798	0.0007	27.66	0.03	0.10
Etilefrine	182.1181	0.0004	11.71	0.05	0.42
Famotidine	338.0530	0.0007	15.44	0.15	0.94
Metconazol	320.1539	0.0008	30.98	0.06	0.18
Metobromuron	259.0082	0.0007	28.60	0.03	0.11
Monuron	199.0639	0.0006	26.65	0.02	0.09
Sotalol	273.1276	0.0007	15.09	0.23	1.51
Vidarabine	268.1048	0.0006	9.69	0.11	1.17

3.7. Data evaluation

3.7.1. Non-Spectrometric Data Evaluation

The extraction yield was calculated according to the equation:

$$\text{Extraction yield} = \text{evaporated extractable matter(g)} \div 0.5\ \text{g} \times 100 \quad (3)$$

The residue of the extracts was weighed and divided by the weight of the milled powder.

3.7.2 Mass spectrometric data evaluation

3.7.2.1. System A

The data was acquired with MassHunter Workstation LC/MS Data Acquisition software B 05.00, (Agilent Technologies, Waldbronn, Germany) and processed with Agilent Profinder B.06.00 (Agilent Technologies) to extract the RT and determine the exact mass of various *Lemna minor* extracts for the triplicate injections of each sample, after detaching the features found in the blank samples. The extracted ion chromatogram was smoothed with a Gaussian function using 9 points and 5000 points Gaussian width. In a second step of the peak picking workflow, the "expected RT" was set to a ±3.00 min window, which determined the retention time correction (and alignment). Furthermore, the "expected RT" was set to +0.5 min and the mass window was set to 5.6 ppm for adducts and isotopes, which are important factors to prevent the false positive peaks calculated from the same molecular ion adducts/isotope. Hence, the isotope abundance window was 7.5 ppm. Furthermore, the adducts were set to "positive ions" with H^+, Na^+, K^+, and $NH4^+$, and "charge state" to 1. All of the previous parameters contributed to score the peaks as accepted (>75) or rejected (<75). This limits the results to the 2000 group of compounds. The features were conducted via untargeted screening to detect the important *m/z* of the *L. minor* metabolomic profile. The compounds were classified according to Schymanski's (2014) confidence level. The Agilent MassHunter ID Browser (B.07.00) gave a molecular formula to some compounds (level 4). Then, with the help of the local database and different databases, some of them were upgraded into level 3 and level 2 as well as some compounds were upgraded into level 1 with the confirmation using the reference standards. The extracted features with their mass and RT were plotted. After that, the common features between the respective solvent blank and the samples were deleted. The RT (in minutes) was plotted against their masses (Da). This limits the result to 2000 compound groups, which were subsequently exported as .cef files to Mass Profiler Professional (MPP) 13.1.1 for further data evaluation. In MPP, the retention time correction was done subsequently by the regression curve without standard delta RT corresponding to each compound's RT. The compounds were then aligned according to the following parameters: RT window of 0.15 min and mass window of 5 ppm. The cut-off was set twofold and all entities with fold change values larger or equal to two were displayed. The different extracts are represented on the scatter diagrams in accordance with the fold change (FC), which is the absolute ratio between

the normalized averaged intensities of the compound between two extracts. The fold change was calculated according to the following equation, in which the data is considered without the logarithm:

$$y = (FC)x, y = x, y = (1 \div FC)x \quad (4)$$

$$y = x + \log(FC), y = x, y = x - \log(FC), \quad (5)$$

The pathway analysis experiment was performed using a single experiment analysis using the pathway databases (e.g. Wikipathways, BioCyc, and KEGG) provided by the software package.

3.7.2.2. System B

3.7.2.2.1. Peak picking

Data collection and preprocessing are the first and the important step in metabolomics analysis because they determine the quality of the data for the next steps. The data was obtained using the Analyst Software version (TF 1.7.1). Then, it was preprocessed with MarkerView Software (version 1.3.1). The parameter was optimized according to the target analysis of the internal standards in all samples and blanks. Also, transformation and/or normalization of the data acquired in this step, which is required for univariate and multivariate statistical analysis. The minimum and maximum retention times were 5.0 and 34.0 min respectively. The subtraction offset was 15 scans. The noise threshold and subtraction multiplication factors were 50 and 1, respectively. Further, the minimum spectral peak width was 2.00 ppm (Fig. 2). The noise threshold and subtraction multiplication factors have a fundamental impact on peak picking especially for the low abundance metabolites due to the matrix effect of unfractionated extracts. This was achieved using internal standards.

3.7.2.2.2. Alignment and filtering

Three injections of each extract feature list were compared and aligned. The variation between different injections was determined according to tolerance in RT and mass. The retention time tolerance is 2%, which is assigned according to the target analysis of the internal standards. Further, mass tolerance was 5.00 ppm based on QTOF specification (Nürenberg *et al.*, 2015). Then, the background was deleted (i.e. the features found in the blank corresponding to each solvent were deleted from the same extract) (Fig. 2). *Lemna minor* metabolic profile investigation demands alignments of features, which were found in different extracts and the different injections of similar extracts. Further, the features were deleted, which were found in the corresponding blank. This step was fundamental in the untargeted metabolomics analysis to avoid misleading statements.

Also, the retention time index and mass tolerance were determined according to the results of the internal standard, which led to a decrease in the number of false-positive, as well as negative peaks,

and reduce the number of features of the same metabolite. The minimum variation in the noise threshold affected the peak extraction by increasing the number of false-positive peaks or subtract real peaks. The isotopic, adduct (e.g., $[M+Na]^+$, $[M+K]^+$ or $[M+NH4]^+$), and fragments ions were removed or combined to produce one metabolite corresponding to one ion (Zhao *et al.*, 2018).

Data and statistical analysis were conducted in Origin 2017 (Origin Lab Corporation, Academic) with SIMCA statistical software (Malmö, Sweden), respectively. Further analysis and data evaluation was performed with Microsoft Excel 2016 (Washington, USA).

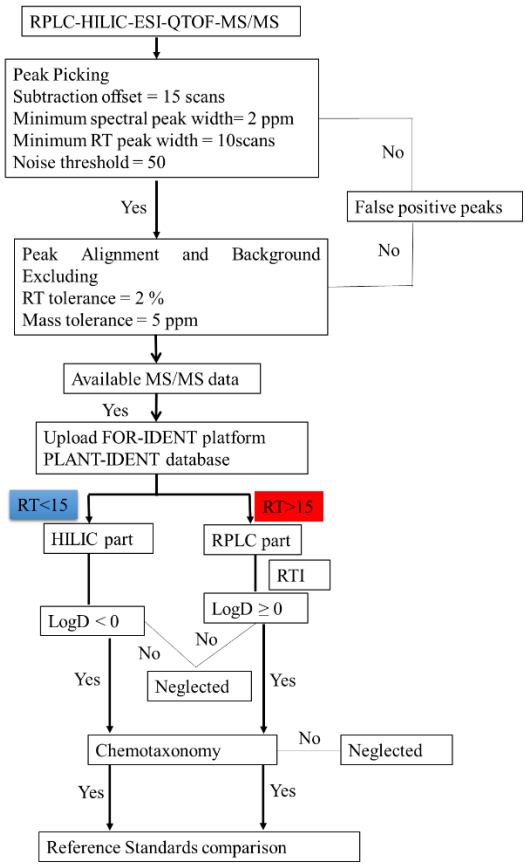


Figure 2. The untargeted workflow flowchart to evaluate the MSMS data.

3.7.2.3. System C

Data from system C was acquired by Xcalibur software 4.0.27.19 (Thermo Fisher Scientific Inc., USA). The extracted ion chromatograms (EIC) were obtained from FMS-ddMS2 (full scan data-dependent acquisition) with an inclusion list for each relevant compound *m/z* using Qual Browser software 4.0.27.19 (Thermo Fisher Scientific Inc., USA). Identification was based on the calculated exact mass of the molecular ions, which were listed in the inclusion list.

The mass deviation of 5 ppm was set. The signal was divided by the signal of the respective deuterated standard. Relative signal intensities were calculated by normalization of the analyte/internal standard response ratio. Further data processing was conducted with Microsoft Excel 2016 (Washington, USA).

3.7.3. DCF and CBZ transformation products detection

DCF and CBZ transformation product standards were analyzed using Agilent Profinder B.06.00 (Agilent Technologies). They were detected in the *P. australis* and *Lemna minor* extracts, with masses estimated at ±10 ppm and RT ±0.3 min of the exact mass and RT of the standards, respectively. The metabolites were identified and suspected (when analytical standards were not available) after RPLC-HILIC-ESI-TOF-MS separation in the suspect analysis. (Suspects screening typically is performed with accurate and high-resolution mass spectrometers to observe the empirical formula of each molecule present and/or with tandem-mass spectrometry to observe specific fragment spectra). A local database was built using MassHunter PCDL Manager B.04.00 (Agilent Technologies, Waldbronn, Germany). Further, the logD (pH7) was the third parameter used to certify the identity of metabolites. The highly polar to polar compounds eluted from the HILIC column at RT < 15 min, with logD values below zero. The nonpolar compounds were eluted from the RP column at RT > 15 min, with logD values above zero. Metabolites within the criteria of mass, RT, and logD (pH7) in the suspect analysis were considered.

3.7.4. Statistical data analysis

3.7.4.1. Data organization

The experiment setup was inserted into the access database. Each run has a unique numerical ID from 1 to 432. Masses, RTs, and an abundance of compounds from each run were merged to the ID by inserting them into the access database. The internal Access programming helped to visualize and check that all the data was correctly inserted. This was done by reviewing the graphs and the corresponding pre-documented additional data like (Solvent, part of the plant …..). After this from the dataset, a pivot matrix data table was calculated.

The pivot data table was arranged that masses@Rt are used as rows while the columns documented the accompanied abundance. The algorithm to attach the abundance to mass@Rt used the first value of abundance. Different approaches of mapping were tested such as (Mean, Average, First Value, Last Value…) but turned out to process a similar outcome in further analysis. Therefore, combination mass@Rt has occurred only in a unique combination with abundance.

Because of the limited possible number of columns (IDs) in Access the data table needed to be split into single excel files via a script before reunion it in SIMCA.

After the Excel, files with all the runs were merged in SIMCA the data table needed to be transposed to treat the different mass@RT as variables (Columns) and the different runs as observations (Rows).

After the union and transposing the data, the additional documentation like (Solvent, part of the plant …..) was pasted into SIMCA as well.

In the statistical software (SIMCA), the additional information (Solvent, part of the plant …..) was defined as secondary observations. This means that this information is not used within the developed models as variables. Some variables Mass@RT weren't found in all observation runs. Accordingly, the pivot table used to merge the data did document this with missing data. This is not very helpful in analyzing the data because the statistical software would see this data as "missing". Instead of "no occurrence". To put this right the empty cells were replaced with zeros

After this, the data was used to build the first model. In this starting PCA analysis, the untreated data was stored as a reference and to start the basic analysis with further models.

This is an especially important step to get an overview of the overall pattern in the data. The most important tools are a.) the score scatters plot which presents the consistency of the data using the uses the hosteling ellipse. This ellipse represents a 95% confidence interval in the multidimensional space. Observation outside this ellipse is remarkably interesting and needs to be investigated. Sometimes those observations could also be identified as outliers.

Also the DModX "Distance to the Model in X" could give insights about the portion of the Variance (Predicted – Observed) which couldn't be described by the model, to get a better understanding of what the model is capable of and what might be very unlikely and need more detailed investigation.

In the data, no anomalies or outliers have been found.

Within the first analysis, the underlying correlation pattern is represented with clusters in the score scattered plot, which summarized the information of all investigation runs in each one data point. With the help of the 2nd observation, the data set can be colored accordingly to check if the for-instance solvent or part of the plant does have significant uniqueness to expose the observation in one of the clusters. This is an easy way to analyze the clusters using the secondary information of documentation without considering the extra information to build the model.

3.7.4.2. Data analysis strategy

Data statistical analyses were conducted with SIMCA 16 software (Malmö, Sweden). Further analysis and data evaluation were performed with Microsoft Access and Excel 2016 (Redmond, Washington, USA) and OriginPro 2019, Origin Lab cooperation, Northampton, MA, USA.

The preprocessed data is a matrix. The rows are the exact masses, retention times (RTs), and abundances of each sample, which were listed in a Microsoft Excel Sheet. For statistical analysis, it is common to handle the data matrices with rows as observations and columns as compounds. Therefore, the data was organized in the Microsoft Access Database file (DBF), which was exclusively built to be suitable for SIMCA 16. The main advantage of the data matrix is the inherent support to align metadata (plant part, plant number, extraction solvent, and drug incubation) along with related quantitative data (i.e., feature annotations/abundance as columns and sample annotations as rows). In DBF, the RTs, masses, and abundances were connected to the corresponding plant part (i.e., leaf, rhizome, and root), plant number (i.e., plants 1 and 2), extraction solvent, and drug incubation. Once the matrix was created, comprehensive statistical analyses could be performed by using the vast range of functions provided by the software. The matrix consisted of 432 observations (i.e., the incubation with/without DCF or CBZ) and 11,442 variables (features). Furthermore, the plant part, plant number, and extraction solvent were used as secondary observations. The data was not transformed and centered; however, it was scaled. The data was analyzed according to the following two strategies, considering the statistical analyses in untargeted metabolomics. By default, SIMCA provides an algorithm called "cross-validation" to get the most valid model by calculating the adequate number of principal components to prevent overfitting of the data in the model. Furthermore, the software provides a large number of visual diagrams (score plots (with Hotelling's ellipse), DModX (Distant to Model), and statistic tables to assess the quality of the model in addition to the R^2 and Q^2.

Metabolite fingerprinting was used to capture metabolite patterns across metabolite profiles. They are characterized without further identification steps (i.e., without the need for standard reference material). Partial Least Squares (PLS) and Orthogonal Partial Least Squares regression-Discriminant Analysis (OPLS-DA) were used to relate sets of X-variables (such as plant part, plant number, extraction solvent, and drug incubation) to the metabolites matrix. SIMCA 16 has a tool called Multiblock Orthogonal Component Analysis (MOCA). MOCA's concept is used to accomplish a fast and accurate analysis of multiple blocks of data (variables) registered for the same set of observations. MOCA aims at extracting the information in complex multi-block data analytics. Furthermore, it will extract two sets of components: the joint and the unique components. The quality of the models is described by R^2 and Q^2 values, where R^2 is the proportion of variance in the data explained by the models and indicates the goodness of fit and Q^2 is the proportion of

variance in the data predictable by the model and expresses predictability (Löfstedt and Trygg, 2011).

Metabolite profiling which uses sets of predefined metabolites were studied in different plant samples which were usually related to the incubation with DCF or CBZ. Metabolite/variable selection was conducted to observe only the most significant metabolite candidates that explain the differences between the samples using S- and contribution plots. The statistical models were built with confidence limits of 95%. Also, the differentiating metabolic profile (DMF) was chosen based on their contribution to the variation and correlation within the data sets. The related metabolic pathways were analyzed using MetaboAnalyst 4.0. Moreover, their contributions and biological clarifications were described based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The KEGG pathway analysis tool was used by the *Arabidopsis thaliana* database. The pathway analysis module combines enrichment analysis and topology analysis based on KEGG. Fisher's test was used to generate *p* values. The *p*-value was equal to 0.05, which indicates the fundamental connection of the identified metabolite with their respective metabolite and not due to the random chance (Chong *et al.*, 2019a; Chong *et al.*, 2019b).

3.7.5. Metabolomics data analysis

The DMF of *Phragmites australis* assigned with the OPLS-DA and S-plots were extracted. The extracted data was returned to the original data. It is impossible to identify a pathway depend on just a mass. To overcome this issue, a key concept is to shift the unit of analysis from individual compounds to individual pathways or a group of functionally related compounds. The mummichog algorithm is the first implementation of this concept to infer pathway activities from a ranked list of MS peaks identified by untargeted metabolomics. The original algorithm implements an over-representation analysis (ORA) method to evaluate pathway-level enrichment based on significant features assigned with the statistical analysis. Users need to specify a pre-defined cutoff based on *p*-values. For further details about the original implementation, please refer to Li *et al.* 2013. The mass accuracy was set to 5 ppm on the positive mode. The *p*-value cutoff was assigned to 0.05 to delineate between significantly enriched and non-significantly enriched features.

4. Isolation and identification of the secondary metabolites from plants

This chapter has been previously published with editorial changes as follows:

Wahman, R., Graßmann, J., Sauvêtre, A., Schröder, P., Letzel, T. (2020). *Lemna minor* studies under various storage periods using extended-polarity extraction and metabolite non-target screening analysis. Journal of Pharmaceutical and Biomedical Analysis 188:113362.

Author contributions:

Rofida Wahman and Thomas Letzel conceptualized the research objective and designed the methodology Andrés Sauvêtre, and Peter Schröder created the test set and performed the plant growing and incubation together with Rofida Wahman. Rofida Wahman collected and analyzed the data. Rofida Wahman applied and validated the models and wrote the paper. Johanna Graßmann, Andrés Sauvêtre, Peter Schröder, and Thomas Letzel reviewed the manuscript. All authors approved the final version of the manuscript.

4.1. Introduction

In-plant metabolomics, sample storage, and preparation are considered the most important steps because they determine the quality of the acquired data. Also, they greatly affect the reliability of the metabolomics results. A small alteration in the sample extraction and storage greatly impacts the metabolite stability and leads to major changes in the metabolic profile. The ultimate goal is to minimize the biologically irrelevant changes resulting from sample processing, which is considered the most important bias in metabolomics studies. Hence in this chapter, the metabolic profile of *Lemna minor* and *Phragmites australis* have been performed using an extraction method with minimal pretreatment via extended polarity chromatography. The extraction method was developed to be capable of providing the metabolic profile of both plants and other plants. The univariate and multivariate statistical analyses have been utilized. Furthermore, the effect of storage conditions has been investigated. The parameters related to experimental design, sample extraction to data analysis follow the Metabolomics Standards Initiative (MSI) (Sumner *et al.*, 2007).

Hypothesis # 1 Extraction strategy for *Lemna minor* and *Phragmites australis* can influence their metabolites chromatographic fingerprints

The field of metabolomics represents a relatively new approach to the systematic study of metabolites and enables their presence and content to be determined in such biological samples as plants. The information gained from metabolomics can make a profound contribution to understanding the interactive nature of secondary metabolic networks in plants and their responses to environmental and genetic changes. It can also provide unique insights into the fundamental nature of plant phenotypes concerning development, physiology, tissue identity, resistance, and biodiversity. Metabolomics fingerprinting can therefore be beneficial in drug discovery, gene-function analysis, and multiple diagnostic methods in phytomedicines (Fukushima and Kusano, 2013). The plant extract should be obtained with minimum pretreatment, to avoid the loss of metabolites. Consequently, the extract can be analyzed based on an untargeted screening strategy. However, challenges such as massive differences in abundance and polarities limit the analytical power of plant metabolites (Schäfer *et al.*, 2016). Thus, sample preparation (extraction method) is one of the most important challenges in the analysis of plant metabolites. In recent years, several extraction methods of plant materials have been presented in the literature. The amount of extractable components is mainly influenced by the strength of the extraction method as well as the efficiency of the solvent in dissolving these compounds. Azwanida enumerated the limitations of both conventional and modern techniques of plant extraction in 2015. The conventional methods of extraction, such as maceration, infusion, diffusion, and decoction, waste many solvents, which is a serious environmental issue (Azwanida, 2015). Also, the Soxhlet extraction method causes

environmental pollution and is not suitable for the extraction of thermolabile compounds. In 2009, Sultana *et al.* revealed that the reflux extraction method reduces the scavenging activity of extracts, due to the degenerative reactions of antioxidant compounds. The heat accelerates the oxidation of antioxidant compounds in the determination of extraction yield and antioxidant activity in different parts of a plant (Sultana *et al.*, 2009). There is a consensus among scientists that alongside their advantages, modern techniques such as solid-phase extraction, microwave-assisted extraction, accelerated solvent extraction, ultrasound-assisted extraction, and supercritical fluid extraction also display weaknesses (Ameer *et al.*, 2017). What is required today is a type of polarity-extended extraction that can be utilized in new polarity-extended chromatographic separation systems.

By employing various separation methods along with mass spectrometric detection, metabolomics has enabled the modification of the chemistry in phytomedicines as well as nutrition and toxicology fields. It comprises two classes: targeted and untargeted metabolomics (Bao *et al.*, 2018). Targeted screening, formerly known as 'quantitative analysis', observes analytes using isotopically labeled reference substances. Untargeted metabolomics aims to assign as many compounds as possible in a plant sample. The analytical tools employed in untargeted screening of plant metabolomes with LC-ESI-MS/MS techniques can yield 'big data'. However, these metabolite classes are difficult to retard and separate analytically (both from each other and the matrix components) using the traditional and most commonly used method of reversed-phase liquid chromatography (RPLC), due to their great polarity variation (Joshi, 2002). Thus, the plant's polar metabolites have low adsorption on non-polar surfaces such as C18 RPLC material. Such molecules are often involved in a multitude of metabolic pathways and play an important role in the recognition of crucial metabolic changes (Iwasaki *et al.*, 2012). Liu and Rochfort (2014), discussed the application of different stationary phases to the separation of polar compounds (Liu and Rochfort, 2014), such as HILIC, which is employed to characterize and separate (very) polar compounds in biological samples (Bucar *et al.*, 2013). Therefore, it is essential to develop a separation technique with an extended polarity range that allows simultaneous monitoring of non-polar, polar, and (very) polar plant metabolites. Serial coupling of RPLC and HILIC is used to separate compounds with extended-polarity chromatography in a single run. Moreover, serial RPLC-HILIC can be coupled with accurate high-resolution time-of-flight mass spectrometers (TOF-MS), which provide accurate detection of non-polar to (very) polar compounds, including the empirical formula for each one (Bieber *et al.*, 2017). Systems with "parallel column coupling" are also being developed for extended-polarity separation (Hemmler *et al.*, 2018).

4.2. Result and Discussion

Many factors can influence the analysis of plant metabolites. This study investigates the effects of extraction and storage procedures on the (secondary) metabolite content of exemplary plant samples using an untargeted screening strategy. The extended-polarity extraction method using *Lemna minor* enabled analysis of its metabolites with RPLC-HILIC-ESI-TOF-MS. Initially, five different solvents were used to ensure the effective extraction of molecules with a broad polarity variation. It was found that increasing the aqueous content of the solvent increased the content of extracted (very) polar compounds. On the other hand, it decreased the solubility of moderately polar to non-polar compounds. The converse behavior could be observed using less polar solvents such as ethanol. The untargeted screening strategy reflects the differentiation between various dissolved metabolites, in addition to, the behavior of *Lemna minor* metabolites under different storage periods.

4.2.1 The extraction method and yields of Lemna minor

This study applied the extraction method previously approved by (Kaufmann *et al.*, 2016; Kaufmann *et al.*, 2017). The method provides a high extraction yield with minimal pretreatment. As stated in the literature, the extraction method was modified to suit the requirements of extended-polarity chromatography in untargeted screening analysis. Figure 3a shows the extraction yield plot of *Lemna minor* expressed as g/kg *Lemna* powder). Different trends in extraction yields were obtained with the different solvents, although the weight of the *Lemna minor* powder and the extraction process remained the same for all extraction solvents. A comparison of the extraction yields of *Lemna* fresh (LF), frozen for days (LFD), and frozen for months (LFM) shows that the different *Lemna* samples have a range of between 0 and 457.9 g/Kg of LF in acidic 90% MeOH extract. LF has, in general, high extraction yields 129.9, 457.9, and 245.1 g/Kg in100% MeOH, acidic 90% MeOH and 100% H_2O extract, respectively, but LFD also gives significantly high yields, whereas the measured yields of LFM seem to be rather low. However, as reported in (Nostro *et al.*, 2000), acidified MeOH was applied as a solvent to improve the extraction of the plant metabolites. In general, good extraction efficiencies for *Lemna minor* were obtained with acidic 90% MeOH, but also in 50% MeOH and 100% H_2O solvents. In general, LF produces higher extraction yields than LFD and LFM in 100% MeOH, acidic 90% MeOH and 100% H_2O extract, respectively. The disparity in the extraction yield of different extracts may be due to the different solubilities of extractable components resulting from their varied chemical compositions. Previous authors were generally more interested in questions concerning 100% MeOH and 100% EtOH extraction power. Sultana *et al.* (2009), concluded that the 100% MeOH extract produced higher extraction yields of total phenolic and flavonoids from various organs of different plants than the 100% EtOH extract. Furthermore, Eloff, 1998, pointed out that 100% MeOH in extraction was more efficient than 100% EtOH in extraction for eluting antimicrobial compounds in different plants based on different parameters (Eloff, 1998). A survey conducted by Miliauskas

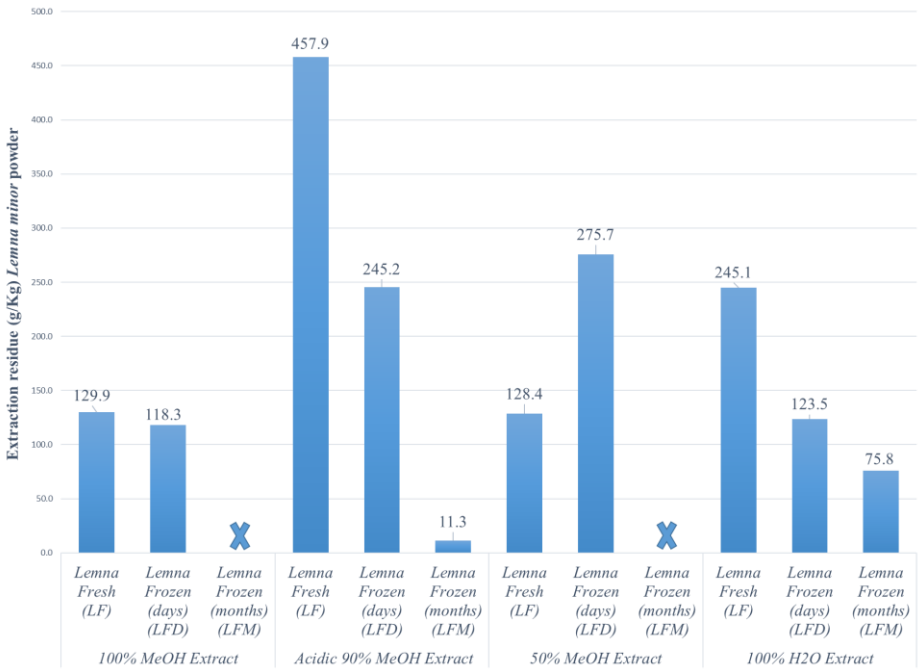
et al. (2004) showed that the 100% MeOH extract of twelve different plants had a high level of radical scavenging activity. This strongly suggested that 100% MeOH extracts contained the highest number of antioxidant compounds attributed to the defense mechanism of the plant that protects it against stress factors (Miliauskas *et al.*, 2004). In the present study, 100% EtOH as an extraction solvent has a significantly lower efficient extraction yield than 100% MeOH as an extraction solvent, as a result of which EtOH was not applied any further in this study. It can be concluded from the above findings that different solvent systems dissolve components differently and that a variety of solutions with a wide range of polarities may be used for maximum extraction. However, it should be kept in mind that a higher extraction yield does not necessarily result in a high reducing content or a higher number of different compounds, which is indicated by the intrinsic nature of the components (such as the acidic 90% MeOH LFM).

4.2.2. Reducing contents of Lemna minor

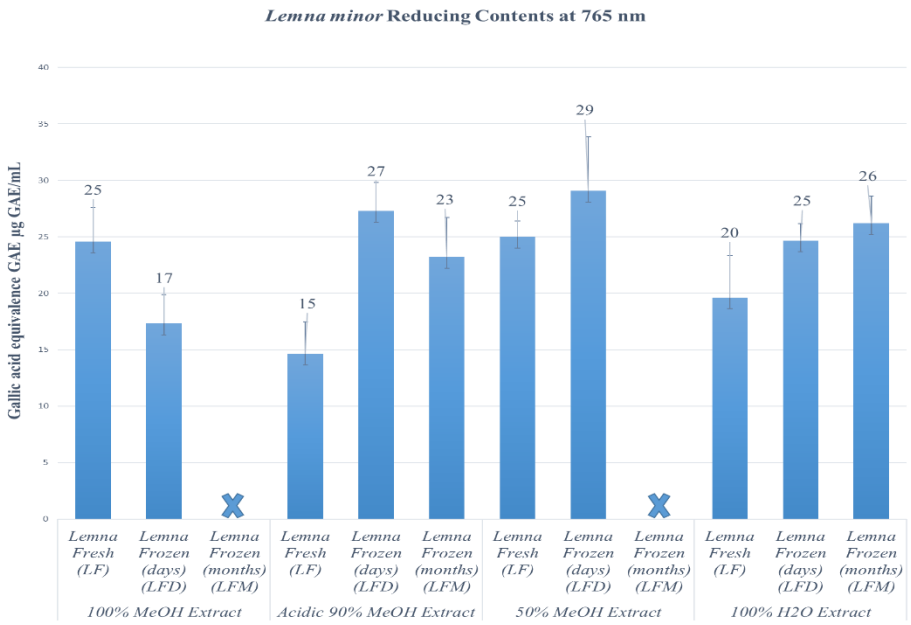
The reducing contents were expressed as gallic acid equivalents (GAE), measured by molecular absorbance at 765nm; these are presented in (Fig. 3b). According to the results, the extracts contained a mixture of reducing compounds with different polarities. The reducing contents for LF, LFD, and LFM ranged from 15 µg GAE/mL in an acidic 90% MeOH extract to LFD with 29 µg GAE/mL in a 50% MeOH extract of LFD. In general, LFD in the extracts often creates a significantly higher reducing content compared to the same LF extract. The freezing of *Lemna minor* increases the reducing contents in the extracts (excluding the 100% MeOH extract).

Lemna minor contains different compounds of various natures, such as fatty acids, saturated carboxylic acids, aldehydes, ketones, aromatic alcohols, and phytosterols (Vladimirova and Georgiyants, 2014). Moreover, duckweed contains protocatechuic aldehyde, *p*-hydroxyl benzaldehyde, truxillic acid, and *p*-coumaric acid (Zhao *et al.*, 2014a). On the other hand, the reducing contents of the *Lemna minor* increased with freezing, which can change their phenolic behavior during the freezing period by way of molecular transformation, such as oxidation.

In conclusion, there is no direct relation between the extraction yield and the reducing contents of different *Lemna minor* samples in the different extracts.



a



b

Figure 3. The extraction yield is expressed in g/Kg of *Lemna minor* fresh (LF), frozen for days (LFD), and frozen for months (LFM). In 100% MeOH, acidic 90% MeOH, 50% MeOH, and 100% H_2O extracts, respectively; (b) The reducing contents of *Lemna minor* fresh, frozen for days, and frozen for months in 100% MeOH, acidic 90% MeOH, 50% MeOH, and 100% H_2O extracts, respectively. The reducing contents are expressed as Gallic acid equivalent measured at the absorbance of 765nm.

4.2.3. RPLC-HILIC-MS analysis of different extracts of Lemna minor samples

Fresh *Lemna minor* samples were extracted with the four solvents 100% MeOH, acidic 90% MeOH, 50% MeOH, and 100% H_2O, producing the results presented above. The extraction method is illustrated in detail in (Fig. 1). The four fresh *Lemna minor* LF extracts were injected into a RPLC-HILIC-ESI-TOF-MS system. The mass spectrometric total ion chromatogram (TIC) was analyzed to determine the feature (i.e. molecule) numbers after excluding the experimental background. The TIC was interpreted to obtain extracted ion chromatograms (EICs), which enabled the extraction of the retention time/mass plot for each extract (see Fig. 4). In the RT/mass plot, the features RT are plotted on the x-axis in minutes while molecular masses are plotted on the y-axis in Dalton units. The polar to (very) polar molecules separated by HILIC are marked by blue triangles and the polar to non-polar molecules separated by RPLC are marked by red circles. The reason why the retention time of 15 minutes can be translated into a logD value of zero is that the break between the molecules eluting from HILIC and RPLC was in a retention window of 16 and 21 minutes. A hard limit was defined at the RT for metformin of 15.0 ± 0.3 minutes, as it was the last very polar standard compound eluting from the HILIC column. A detailed definition of the polarity classification used in this study was published in Bieber *et al.* [14]. In (Table 9), the feature values for the four extracts of LF are ordered in decreasing solvent polarity: 100% H_2O > 50% MeOH > 90% MeOH (+ 0.5% FAC) > 100% MeOH. The 0.5% formic acid lowered the pH of the solvent, resulting in the lowest number of compounds in the HILIC component. Increasing the water content to 50% led to a minor increase in the compounds found in the HILIC and RPLC components. Although 100% aqueous extract has a similar quantity of total compounds, the highest significant compound number of polar and (very) polar molecules appeared in the HILIC component. On the other hand, the molecule number in the RPLC component decreased significantly.

Table 9. The number of compounds in the four extracts of *Lemna minor* fresh (LF) separated by HILIC with RT (0-15min) and RPLC columns with RT (16-33min).

Sample	*Lemna minor* fresh (LF)		
	HILIC	RPLC	Total
% H_2O	806	292	1098
50% MeOH	701	410	1111
Acidic 90% MeOH	595	502	1097
100% MeOH	686	383	1069

So far, the comparison of the four extracts has been presented in terms of general values, such as extraction yield, reducing content, and the number of compounds. However, the molecules in various extracts may be different or the same. For this reason, the features from the retention time/mass plot are evaluated concerning their precise mass (in other words, their empirical formula, i.e. their identical atomic composition). The empirical formula for most of the compounds was determined using the mass spectrometric supplier software Mass Profiler Professional (MPP). This was the basis on which the comparison between the compounds of the different extraction samples was performed. Due to the significantly different pH values and resulting in other charging properties, the acidic 90% MeOH extracts will be neglected from this point on in the study. The results for the 100% MeOH, 50% MeOH, and 100% H_2O extracts of *Lemna minor* fresh LF are summarized in the Venn diagrams in Fig. 4. The Venn diagrams show the unique, overlapping compounds between the extracts (inside the cycles) as well as the total number of compounds in each extract (outside of the cycles). The blue numbers again represent the HILIC-separated compounds, while the red numbers represent the RPLC-separated compounds. As can be seen in Figure 5, the total amount of extracted molecules remains similar, and the absolute number of detected molecules is always in a HILIC/RPLC ratio larger than 1. Moreover, the 100% H_2O extract has the highest number of HILIC-separated molecules.

Although the 50% MeOH extract has a large number of different molecules in the HILIC component, it has a less unique number compared to 100% MeOH and 100% H_2O extracts (Fig. 5a). On the other hand, 50% MeOH extract overlapped in a large number of compounds with 100% MeOH and 100% H_2O extracts, as a consequence of the polarity median of both solvents. However, the latter two extracts with a wide polarity discrepancy contain 36 common compounds, and all three extracts contain at least 35 common molecules. The same is true for the RPLC component (in an inverse manner).

Although the 50% MeOH extract has a large number of different molecules in the RPLC component, it has a less unique number compared to 100% MeOH (Fig. 5b). On the other hand, 50% MeOH extract overlapped in a large number of compounds with 100% MeOH and 100% H_2O extracts because of the polarity median of both solvents. However, the latter two extracts with a wide polarity discrepancy contain nine common compounds, and all three extracts contain at least 26 common molecules. The small quantity of non-polar molecules in the MeOH extraction observed with RPLC confirms the previous polarity argumentation. The higher yield of LF extracts was analyzed using a RPLC-HILIC-ESI-TOF-MS system. 100% H_2O displays the highest number of features, while 100% MeOH has the lowest number. Although the 100% MeOH extract is the most non-polar solvent in this study, it generally has a higher number of compounds in the HILIC component compared to the RPLC component. The same is true of extractions containing MeOH. MeOH is an amphiphilic compound with a polarity index of 5.1. This means that MeOH can dissolve non-polar and (very) polar compounds. The 0.5% formic acid reduced the number of compounds in the HILIC component. This may be due to the

neutralization of acidic compounds that elute, later on, resulting in the highest number of observed molecules (in the RPLC section). Based on this argumentation, it can be expected that acidic extraction will produce a small number of molecules with basic functional groups. Moreover, the presence of formic acid in the acidic 90% MeOH extract may cleave the glucoside bonds and release the phenolic compounds (Siddiqui *et al.*, 2011). The project supported the untargeted screening (NTS) strategy, which aims to assign as many compounds as possible. Thus, the combined results from the various extracts may be a precise reflection of the NTS strategy. Subsequently, the acidic 90% MeOH extract was eliminated due to the differences in pH values. The different compounds of 100% MeOH, 50% MeOH, and 100% H_2O extracts of *Lemna minor* fresh LF were compared in Venn diagrams, as shown in Fig. 3. The 50% MeOH extract overlapped in a large number of compounds with 100% MeOH and 100% H_2O extracts in the HILIC and RPLC components, as a consequence of the polarity median of both solvents. The combination of 100% MeOH, 50% MeOH, and 100% H_2O extracts (with similar pH-values) is appropriate for *Lemna minor* extraction and determining its chromatographic fingerprint with RPLC-HILIC-ESI-TOF-MS. Concerning the three solvents, a very large number of *Lemna minor* metabolites were extracted. An overall picture of the metabolites was obtained.

It is convenient to use an untargeted screening strategy as a basis for the extended-polarity chromatographic fingerprint workflow with plant extracts. In particular, the large number of (very) polar molecules dissolved in 100% H_2O extract and detected in the HILIC component highlights its applicability for future measurements in the (very) polar region.

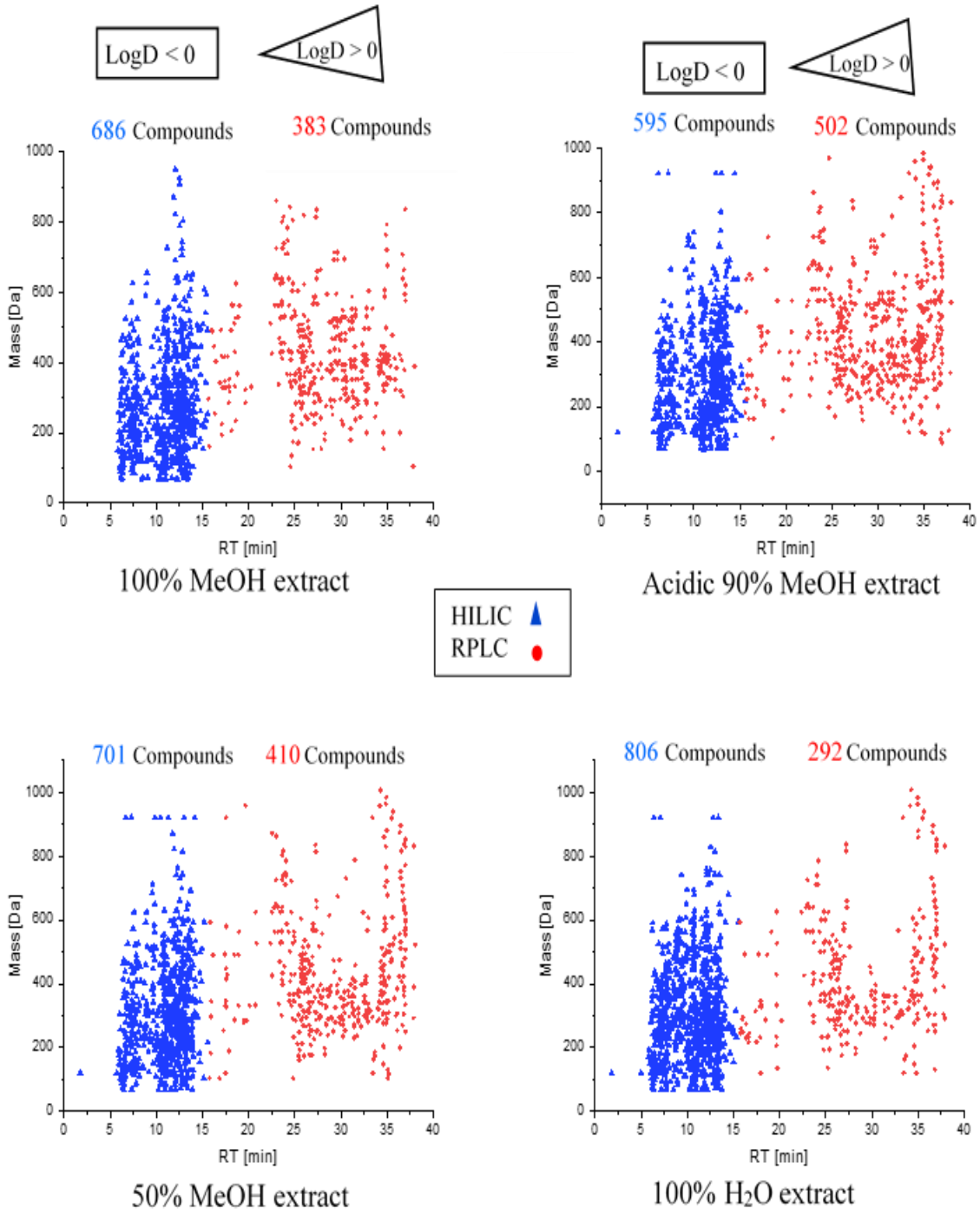


Figure 4. Retention time/mass plot of *Lemna minor* fresh sample (LF) 100% MeOH, acidic 90% MeOH, 50% MeOH, and 100% H2O extracts, respectively. The numbers above each plot represent the total number of molecules separated with HILIC in blue and RPLC in red. Blue, triangles represent HILIC retarded molecules eluting (0-15) min, which have LogD values below zero and red, circles represent RPLC retarded molecules eluting (16-38) min, which have LogD values above zero.

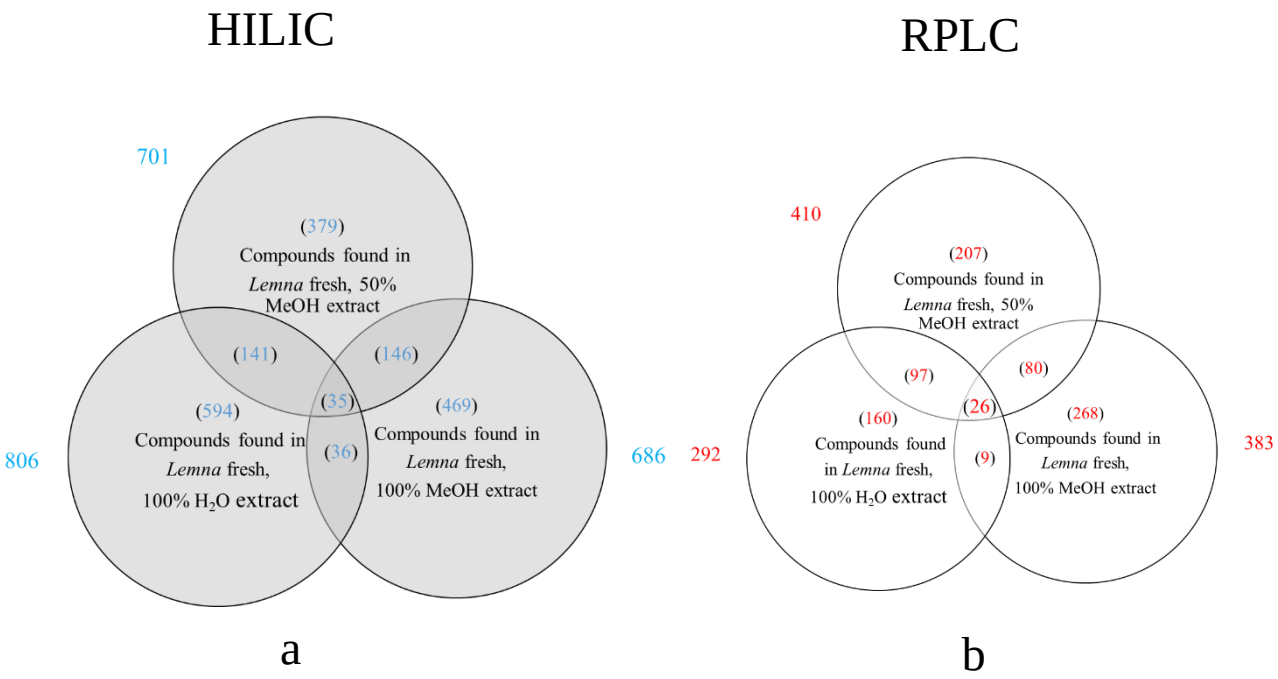
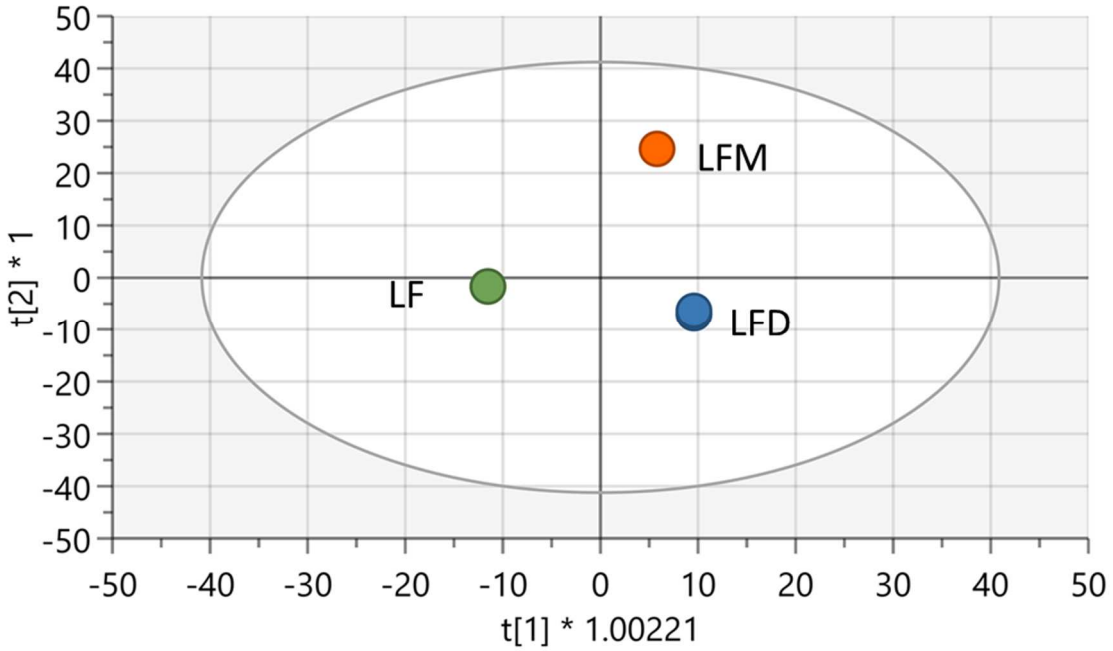


Figure 5. Venn diagrams summarizing the number of shared and unique compounds in three extracts of *Lemna minor* fresh sample. (a) Compounds measured in the HILIC column (with a RT≤ 15) and are blue colored; (b) Compounds measured in the RPLC column (with a RT > 1 and are red-colored. The number of the total compounds found outside each circle, unique compounds of the extracts on the inside, and the shared compounds.

4.2.4. Comparison of Lemna minor samples: fresh, frozen (days), and frozen (months)

The chemical and bioactive nature of plant metabolites necessitates the use of methodologies that do not alter their concentration or structure. Previous studies have shown that the metabolite yield could be affected by changing the solvent used for extraction, the duration of extraction, and the storage periods of the plant material or extracts, as well as the biochemical activity, heat, light, vacuum, and drying procedures (Jin *et al.*, 2011). It is therefore important when investigating molecular fingerprints using an untargeted screening strategy, to study the effect of storage on plant metabolites. It has recently been shown that storage can alter the plant's metabolites (Sharma and Lee, 2016). Three samples of *Lemna minor* were investigated in this study: fresh (LF), stored for days at -80 ºC (LFD) and stored for months at -80 ºC (LFM). The extraction yield and the reducing contents of the samples are described above and shown in (Fig. 2a and b) respectively. The acidic 90% MeOH has a considerable reducing content, although, it has a lower feature number in LFM compared to the 100% H_2O extract. The three 100% H_2O extract samples (analyzed with RPLC-HILIC-ESI-TOF-MS) were shown as examples in this study and the results discussed in terms of their presence in the samples. An orthogonal partial least squares discriminant analysis (OPLS-DA) model was drawn up with two components. The horizontal component distinguishes between groups, while, the vertical one differentiates within a group. Although it resembles PCA, the data is guided by known

class information taken from the prediction model. Consequently, OPLS-DA is a supervised modeling approach. In Figure 6, the OPLS-DA model is at a significance level α= 0.05. The OPLS-DA model characterized 86 % of the variation in X (R^2X (cum) = 0.86); 100% in response Y (R^2Y(cum) = 1). The Y was predicted by the 7 fold cross-validation (Q^2(cum) =1). The high-value parameters illustrate the good classification and prediction ability of the OPLS-DA model. Furthermore, the model demonstrates the significant strength and reproducibility of the data. The model indicates that the LFD and LFM are clustered separately from the LF. The vertical component distinguishes between LFD and LFM. Moreover, it shows that LF and LFD are more correlated. This supports the argument that the features in LF and LFD are highly correlated. The comparison of LF, LFD, and LFM metabolites using the OPLS-DA model showed that the data was significantly separated, also LF and LFD were more correlated than LFM. In a bilateral comparison, the metabolites of *Lemna minor* were located differently according to the storage periods. Some metabolites were degraded due to having been stored for a long period at -80ºC. In addition, the relative intensities of the metabolites changed due to storage.



a

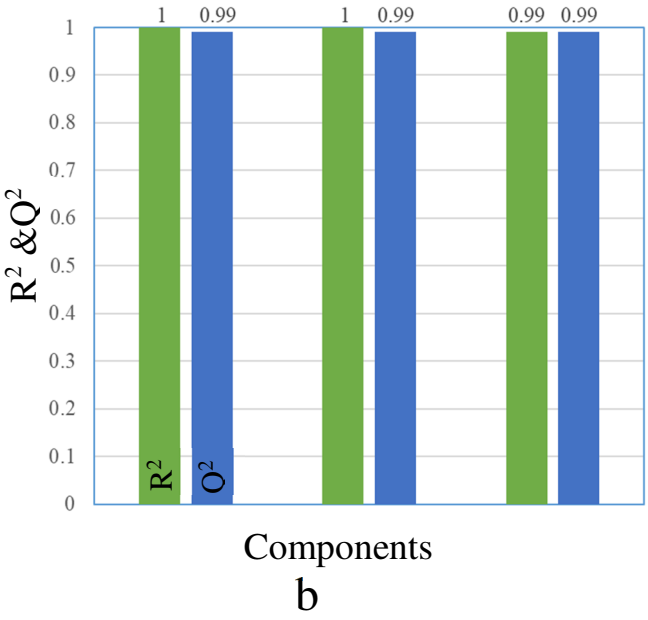


Figure 6. The Orthogonal partial least squares discriminant analysis (OPLS-DA) model of *Lemna minor* fresh sample (LF), frozen for days (LFD) and frozen for months (LFM) 100% H_2O extracts. Scaled proportionally to R^2X, $R^2X[1]$ = 0.238, $R^2X[2]$ = 0.243, Ellipse: Hotelling's T2 (95%). (b) X/Y overview plot displays the individual cumulative R^2 (green columns) and Q^2 (blue columns) for the goodness of fits and cross-validation parameters.

The features were extracted in Venn diagrams containing the LF, LFD, and LFM 100% H_2O extracts, as presented in (Fig. 7). The unique, overlapped and total compound numbers are presented as described above. The blue color (on the left-hand side) represents compounds separated with the HILIC column, whereas the red color (on the right-hand side) represents compounds separated with the RPLC column. The direct feature comparison between the LF, LFD, and LFM 100% H_2O extracts shows that LFM has a significantly lower amount of molecules in the HILIC component, whereas there is a significantly higher amount of molecules in LFM that cannot be found in LF and LFD. Accordingly, (very) polar compounds might be degraded into less polar compounds by freezing, these having been retarded on the RPLC. Furthermore, these compound structures may be alternated into a less polar compound. In addition, LFM has the highest number of unique compounds in the RPLC component, which confirms the polarity argumentation.

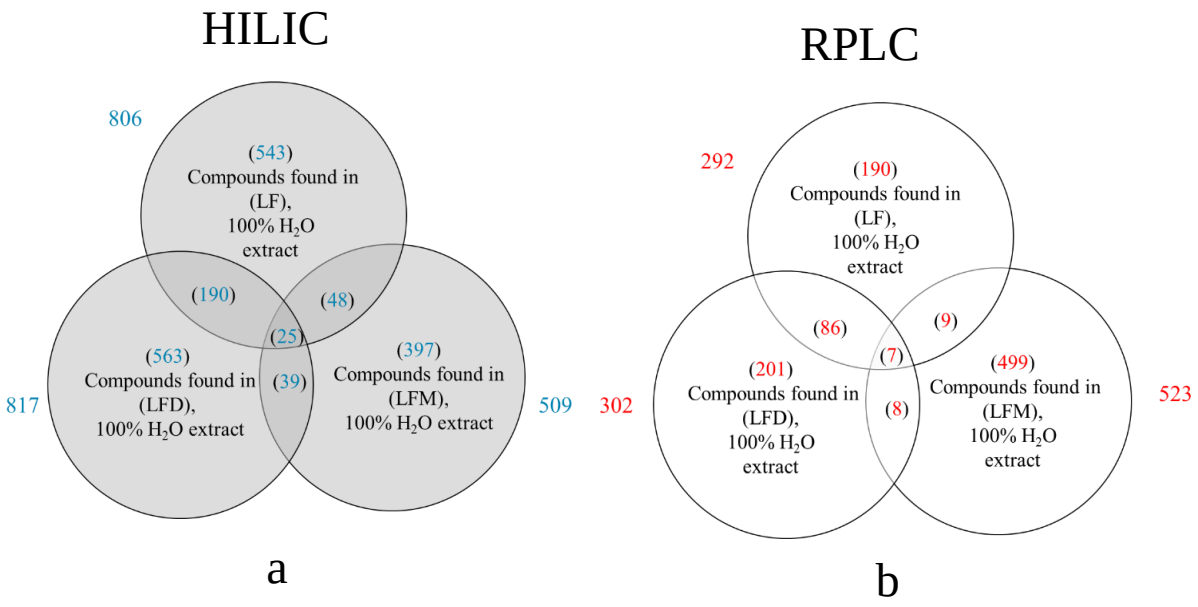
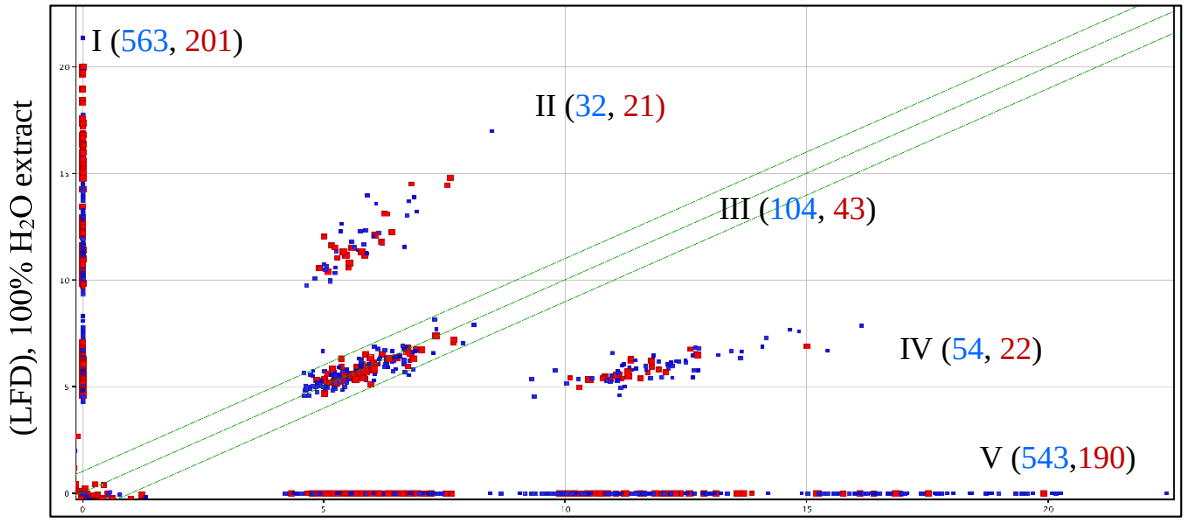


Figure 7.Venn diagrams summarizing the number of shared and unique compounds in 100% H_2O extracts of *Lemna minor* fresh sample (LF), frozen for days (LFD), and frozen for months (LFM); (a) Compounds measured in the HILIC column (with a RT ≤ 15) and are blue colored; (b) Compounds measured in the RPLC column (with a RT > 15) and are red-colored. The number of the total compounds found outside each circle, unique compounds of the extracts on the inside, and the shared compounds.

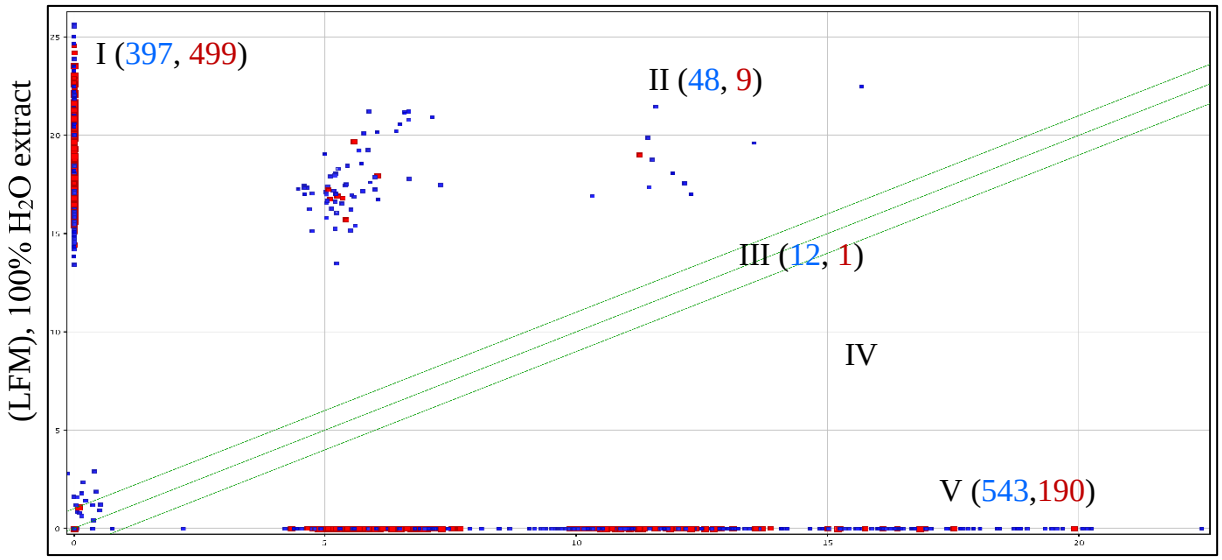
The number of overlapping compounds in the HILIC and RPLC components between LFD and LF is significantly higher than between LFM and LF. However, the total amount of overlapping compounds in the three extracts is very small in comparison with the number of unique compounds. In particular, the LFM sample differs significantly (in the overall polarity region) from the other two samples.

Therefore, the data was considered without the logarithm, according to the fold change. The data was plotted between two samples, as shown in (Fig. 8). Figure 8a showed a bilateral comparison plot of an LFD fingerprint versus an LF fingerprint. Compounds exclusively found in an LFD 100% H_2O extract were located on the y-axis (Fig. 8a) (i.e. plot region I), whereas LF 100% H_2O extract compounds are exclusively located on the x-axis (Fig. 8a), (i.e. plot region V). The compounds presented in region II (Fig. 8a) are found in both extracts with higher signal intensities in LFD 100% H_2O extract. Compounds close to the axis halving line (region III) of (Fig. 8a) were found in both extracts obtained with similar signal intensities. However, the compounds in region IV (Fig. 8a) have higher signal intensities in the LF 100% H_2O extract. The results reflect that *Lemna minor* might be degrading its original metabolites due to the longer storage period at -80ºC. In Figure 8b, the unique compounds of LF 100% H_2O extract were located on the x-axis and the compounds unique to the LFM 100% H_2O extract were located on the y-axis. However, Figure 8c presents unique compounds in the LFD

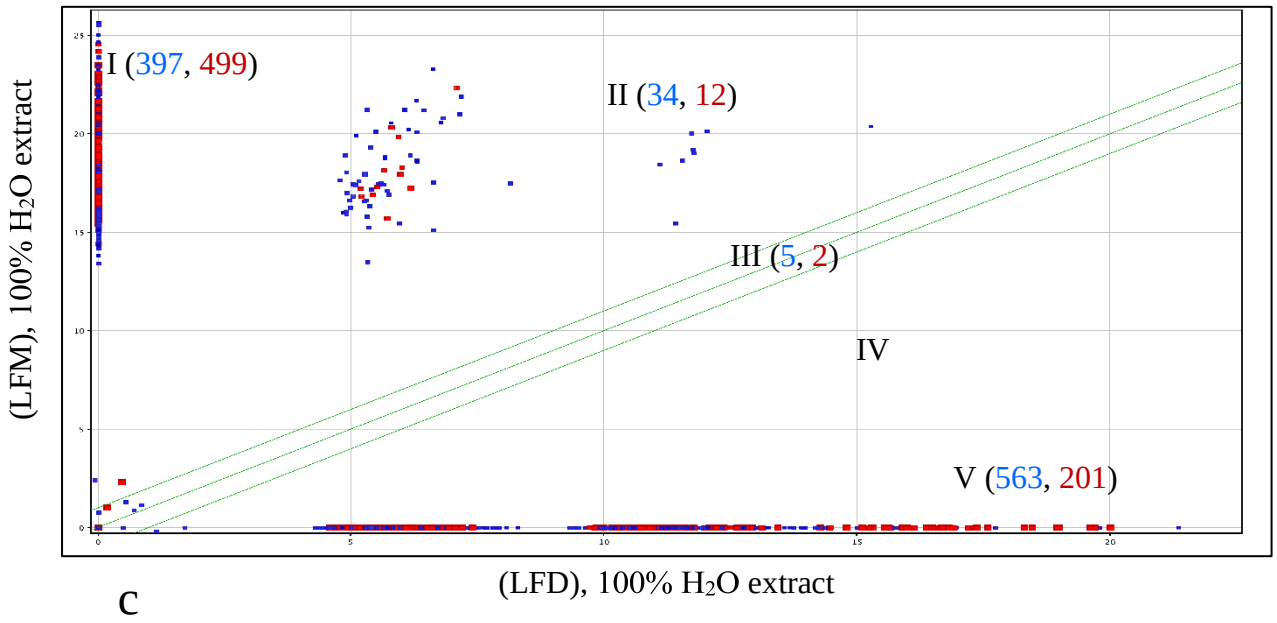
100% H_2O extract on the x-axis versus unique LFM 100% H_2O extract compounds on the y-axis. The compounds on the y-axis could not be found in LF and LFD, which may be due to the degradation of the *Lemna minor* metabolites. Moreover, in a direct bilateral comparison between LFD and LFM, the compounds of both samples have higher signal intensities in the LFM, as shown in (Fig. 8c) region II.



a



b



C

Figure 8. (a) Feature (signal intensity) comparison plot of *Lemna minor* fresh (LF) 100% H_2O extract versus *Lemna minor* frozen for days (LFD) 100% H_2O extract. (b) *Lemna minor* fresh (LF) 100% H_2O extract versus *Lemna minor* frozen for months (LFM) 100% H_2O extract. (b) *Lemna minor* fresh (LF) 100% H2O extract versus *Lemna minor* frozen for months (LFM) 100% H_2O extract. (c) *Lemna minor* frozen for days (LFD) 100% aqueous extract versus *Lemna minor* for months (LFM) 100% aqueous extract. Blue, small squares represent RT (0-15min), and red, big squares represent the RT (16-33min).

I. Compounds exclusively found on y-axes.
II. Compounds were found in both samples with increased feature intensity on the y-axis compared to the x-axis.
III. Compounds were found in both samples.
IV. Compounds were found in both samples with decreased feature intensity on the y-axis compared to the x-axis.
V. Compounds were found exclusively on x-axes.

Fourteen compounds in Table 10 could be suggested in the LF 100% H_2O extract by comparing measured masses with molecular masses of compounds found in the Lemnaceae and/or Araceae family. On the one hand, there are the four amino acids L-histidine, L-leucine, L-isoleucine, and L-tyrosine, which were previously identified in *Lemna minor*, and, on the other hand, the three flavonoids luteolin 7-rutinoside, heptamethoxyflavone, and kaempferol-3-O-rutinoside. Furthermore, dimethyl malate, nicotinic acid, adenine, 2,5 dihydroxymethyl-3,4-dihydroxypyrrolidin (DMDP), uracil-1-beta-D-arabinofuranoside, 3,4-dihydroxycinnamic acid methyl ester, and 3,4,5-trimethoxyallylbenzene were also characterized.

L-histidine, L-leucine, L-isoleucine were also identified in varying intensities in the LFD sample (see Table 11).

Moreover, in the LFD 100% H_2O extract, a small amount of L-methionine was identified in the *Lemna minor,* (Table 10). Also, the L-phenylalanine signal intensity increased, although the L-tyrosine signal intensity decreased below the measured limit. Consequently, the signal intensity of heptamethoxyflavone increased (Table 11). Isorhamnetin-3-galactoside was also characterized. However, there was a slight decrease in schaftoside, isoschaftoside, and isovitexin-7-O-xyloside signal intensities.

In addition, 2-aminohexanedioic acid, a metabolite of lysine metabolism, was suggested. in addition to this, hexose and apiose were also characterized. Consequently, 2,6-dideoxy-2,6-iminoheptitol, which is a potent α-glucosidase inhibitor, could also be characterized.

In the LFM 100% H_2O extract, there was a significant increase in the expected L-methionine signals. However, the signal intensities of other amino acids decreased below the detection limit. A flavanoid scoparin 2"-xyloside was potentially characterized, (Table 10). Furthermore, schaftoside, isoschaftoside, and isovitexin-7-*O*-xyloside signals increased significantly in LFM (Table 11). Signal intensities also increased for 2,5 dihydroxymethyl-3,4-dihydroxypyrrolidin (DMDP) and 2,6-dideoxy-2,6-iminoheptitol. The sugars fell below the detection limit. Various compounds were detected and identified. Fourteen compounds could be suggested in LF 100% H_2O extract and three amino acids (L-histidine, L-leucine, and L-isoleucine) were identified by reference material. In the LFD sample, the L-phenylalanine signal intensity increased, which is a precursor of the flavonoids' biosynthesis pathway in the plant. Hence, the signal intensity of heptamethoxyflavone was increased and isorhamnetin-3-galactoside could be characterized. On the other hand, there was a slight decrease in the signal intensities for schaftoside, isoschaftoside, and isovitexin-7-*O*-xyloside. New flavonoids were characterized in the LFM sample. Moreover, there was a significant increase in schaftoside, isoschaftoside, and isovitexin-7-*O*-xyloside signals. Also, there was a significant increase in L-methionine. It increases especially in plants suffering from stress conditions (Hacham *et al.*, 2017).

The intensities of polar carbohydrates (such as hexose) fell below the measured detection limit. Thus, LFM had the lowest number in the HILIC component, which may be due to the increase in reducing sugar ('low-temperature sweetening') (Hammond *et al.*, 1990). The same effect was detected with exposure of Scots pine to SO_2 whereby the small concentration increased the production of carbohydrates and the larger one reduced it (Kainulainen *et al.*, 1995). Hence, the freezing of *Lemna minor* for short periods generally decreases the intensity of some compounds. However, freezing for long periods generates new compounds. Freezing under -80°C did not serve to arrest the enzymatic activity in *Lemna minor*. It is possible that some enzymes were still active and degraded metabolites. Furthermore, they altered the biosynthetic pathways in different ways. Similar findings were reported for onion storage, i.e. the quercetin concentration decreased but its conjugates increased (Sharma and Lee, 2016). Bilia *et al.* reported in 2002 that the concentration of flavonoids in *Calendula*, milk thistle, and passionflower tinctures resulted in different behaviors during storage depending on the type of flavonoids (Bilia *et al.*, 2002). Researchers reported that the concentration of secondary metabolites was reduced due to freezing at different temperatures (Kapcum and Uriyapongson,

2018). In conclusion, it is essential to use fresh samples to perform correct analytical measurements in (plant) metabolomics (i.e. untargeted screening) studies, as plant freezing changes the molecular content.

Table 10. List of molecules characterized in *Lemna minor* fresh (LF), frozen for days (LFD), and frozen for months (LFM) 100% H_2O extract with a TOF-MS along with the monoisotopic masses values with respective deviations in ppm, logD values at pH 7, respective the intensities and retention times of compounds. References are marked as (*) for literature dealing with *Lemna minor*, (x) for literature discussing compounds from the Lemnaceae family and/or Araceae family, and (+) for literature in which the core-flavonoid was found without glycosylation or glucuronidation.

Name	Measured Mass (Da)	Mono-isotopic mass (Da)	Δppm	RT (Min)	LogD (pH7)	Intensity	Refer-ences
LF							
Dimethyl Malate[X]	162.0533	162.0530	-1.85	7.72	-0.82	6.91	(Xie *et al.*, 2013)
L-Histidine*	155.0689	155.0690	-6.45	14.98	-3.70	12.99	(Chakrabarti *et al.*, 2018a)
Nictonic acid	123.0314	123.0320	3.25	9.24	-2.76	12.56	(Chua *et al.*, 2010)
Adenin	135.0544	135.0540	-2.96	6.81	-0.58	16.42	(Zhang *et al.*, 2013)
2,5-Dihydroxymethyl-3,4-dihydroxypyrrolidin (DMDP)[X]	163.0835	163.0840	3.077	6.37	-4.76	6.48	(Watson *et al.*, 2001)
Luteolin-7-rutinoside[X]	594.1601	594.1580	-3.53	15.40	-0.77	6.02	(Ferreres *et al.*, 2012)
1-β-Ribofuranosyluracil*	244.0701	244.0695	2.46	8.97	-2.42	6.80	(Kandeler, 2019)
3,4-Dihydroxycinnamic acid methyl ester[X]	194.0584	194.0580	-2.06	25.41	1.91	5.60	(Le Moullec *et al.*, 2015)
L-Leucine L-Isoleucine*	131.0952	131.0950	-1.53	10.88	-1.59	7.20	(Chakrabarti *et al.*, 2018a)
Heptamethoxyflav-one*	432.1413	432.1420	-1.62	23.74	1.89	5.79	(Harborne and Williams, 1982)
Kaempferol-3-*O*-rutinoside[+]	594.1601	594.1585	2.69	15.4	-1.72	6.02	(McClure and Alston, 1966)

Name	Measured Mass (Da)	Mono-isotopic mass (Da)	Δppm	RT (Min)	LogD (pH7)	Intensity	Refer-ences
3,4,5-Trimethoxyallylbenzene[X]	208.11	208.1099	0.48	8.80	2.60	5.49	(Appenroth *et al.*, 2018)
L-Tyrosin*	181.074	181.0740	9.39	11.82	-1.49	5.95	(Chakrabarti *et al.*, 2018a)
LFD							
Isorhamnetin-3-galactoside[X]	478.1109	478.1110	0.21	13.48	-0.74	5.20	(Champagne *et al.*, 2011)
D-hexose*	180.0632	180.0633	-0.56	7.07	-2.93	7.65	(Zhao *et al.*, 2014a)
Apiose*	150.0526	150.0530	2.67	6.29	-2.44	5.60	(Hart and Kindel, 1970)
2,6-Dideoxy-2,6-iminoheptitol[X]	193.0945	193.095	2.60	6.46	-4.21	5.55	(Watson *et al.*, 2001)
L-Methionine*	149.0505	149.0510	3.35	9.99	-2.19	6.63	(Chakrabarti *et al.*, 2018a)
2-Aminohexanedioic acid[X]	161.0702	161.0690	-7.45	13.23	-5.29	5.18	(Smith and Meeuse, 1966)
L-Phenylalanine *	165.0791	165.0790	0.61	10.58	-1.19	7.46	(Bao *et al.*, 2018)
LFM							
L-Methionine*	149.0505	149.0510	3.35	9.99	-2.19	19.68	(Chakrabarti *et al.*, 2018a)
Scoparin 2"-xyloside [X]	594.1594	594.1580	-2.36	16.82	-2.25	20.80	(Muhit *et al.*, 2016)
Found in all samples							
Isovitexin-7-xylosideX	564.1536	564.1480	-9.93	23.60	-2.06	------	(Mayo *et al.*, 1998)
Isoschaftoside * SchaftosideX	564.1492	564.1479	2.30	19.15	-3.91	------	(Valant-Vetschera, 1985) (Muñoz-Cuervo *et al.*, 2016)

Table 11. List of molecules characterized in *Lemna minor* fresh (LF), frozen for days (LFD), and frozen for months (LFM) 100 % H_2O extract with a TOF-MS along with respective intensities. *It is the absolute ratio between the normalized averaged intensities of the compound between two extracts.

Name	Intensities*		
	LF	LFD	LFM
L-Methionine	----	6.63	19.68
2,6-Dideoxy-2,6-iminoheptitol	----	5.55	16.73
Isovitexin-7-O-xyloside	6.04	5.97	17.92
2,5-Dihydroxymethyl-3,4 dihydroxypyrrolidin (DMDP)	6.48	----	18.1
Heptamethoxyflavone	5.79	6.07	
Isoschaftoside Schaftoside	6.05	6.04	19.45
L-Leucine L-Isoleucine	7.20	15	-----
L-Histidine	12.99	6.27	----

4.2.5. The total extraction yield of Lemna minor

After the method optimization, the whole study was performed with fresh plant samples, which were extracted with 100% MeOH, 50% MeOH, and 100% H_2O to investigate the changes in *Lemna's* metabolic profile due to different treatments. The extraction yield of fresh *Lemna minor* different samples during the whole study was summarized in (Fig. 9). Concluding the extraction yield expressed as g/Kg *Lemna* powder ranging from 769.49 to 1046.40 g/Kg. Moreover, the order of the extraction yield was 100% H_2O > 100% MeOH > 50% MeOH.

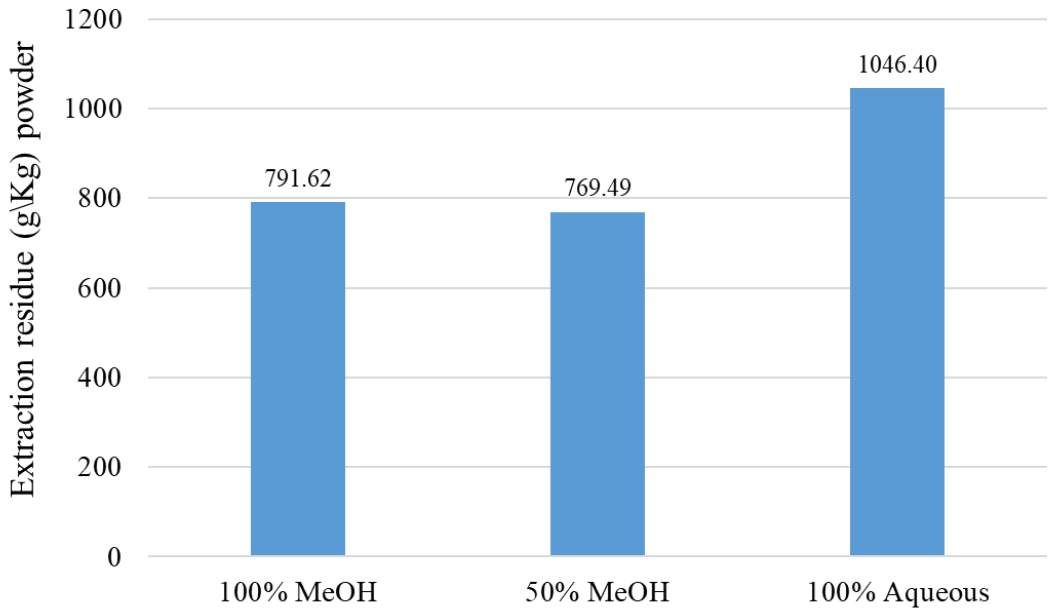


Figure 9.The extraction yield is expressed in g/Kg of *Lemna minor* in 100% MeOH, acidic 90% MeOH, 50% MeOH, and 100% H_2O extracts (n=20).

4.2.6. Metabolic profiling elucidation in Phragmites australis extracts with RPLC-HILIC-ESI-TOF-MS

Different parts of *Phragmites australis* were extracted with four different solvents100% MeOH, acidic 90% MeOH, 50% MeOH, and 100% H_2O. The 100% MeOH extract of the leaf has the highest extraction yield with 234.20g/Kg and the lowest concentration in root extract 82.16 g/Kg (Fig. 10). The four different extracts 100% MeOH, acidic 90% MeOH, 50% MeOH, and 100% H_2O of *P. australis* leaf, rhizome, and root were analyzed similarly with RPLC-HILIC-ESI-TOF-MS coupling as previously described (Greco *et al.*, 2013; Bieber *et al.*, 2017; Wahman *et al.*, 2020). The obtained mass spectrometric total ion chromatograms (TICs) were interpreted to extract the feature (extracted masses, RT, and signal abundance) according to the parameters mentioned in data processing in the material and methods chapter. Background signals (i.e., all the peaks detected in the corresponding blank) were deleted to avoid false positives. Lastly, the features found in the triplicate injections were considered for further analysis. Exemplarily, Figure 11 represents (RT)/Mass plots for 100% MeOH extracts of *P. australis* leaf, rhizome, and root in positive ion mode. The highly polar to polar compounds eluted at RT < 15 minutes, with logD values below zero (HILIC part). The unpolar compounds were eluted at RT >15 minutes, with logD values above zero (RPLC part). The differences in chromatographic fingerprints reflected the variability in metabolite profiles (and composition) in the leaf, rhizome, and root samples. Detailed information and a description of

data analysis have already been discussed in previous publications (Wahman *et al.*, 2019; Wahman *et al.*, 2020).

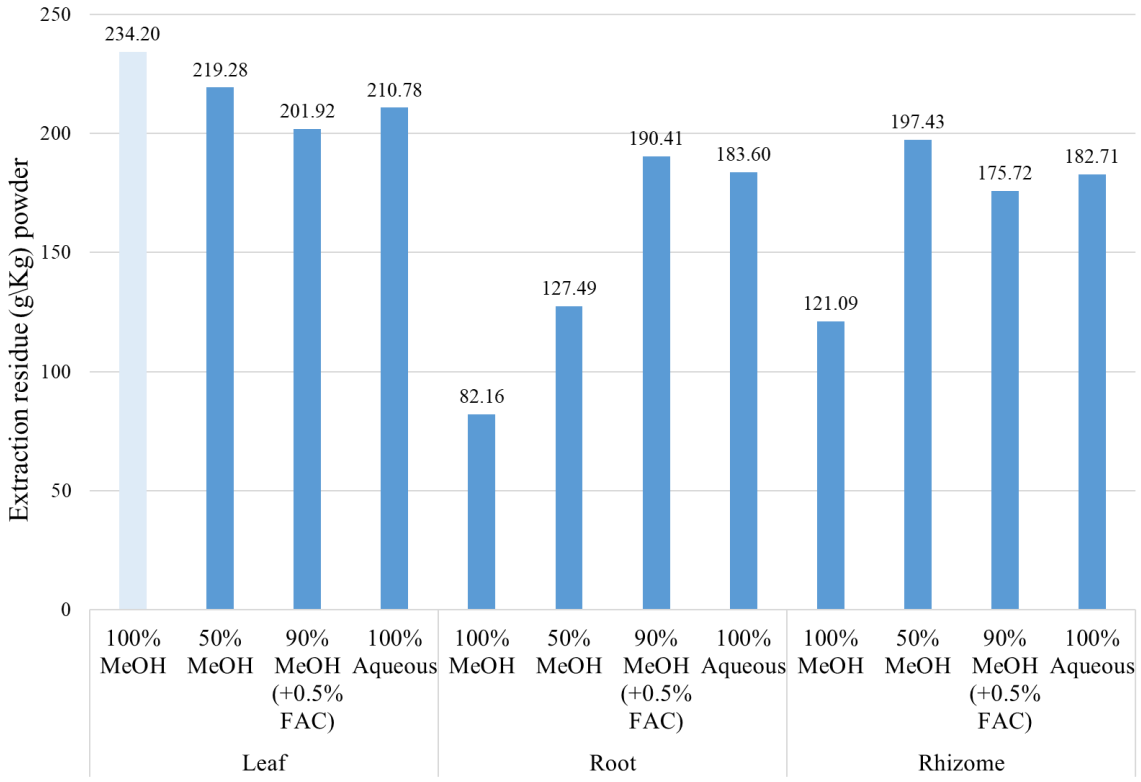


Figure 10. The extraction yield is expressed in g/Kg of *Phragmites australis* leaf, root, and rhizome in 100% MeOH, acidic 90% MeOH, 50% MeOH, and 100% H_2O extracts (n=20).

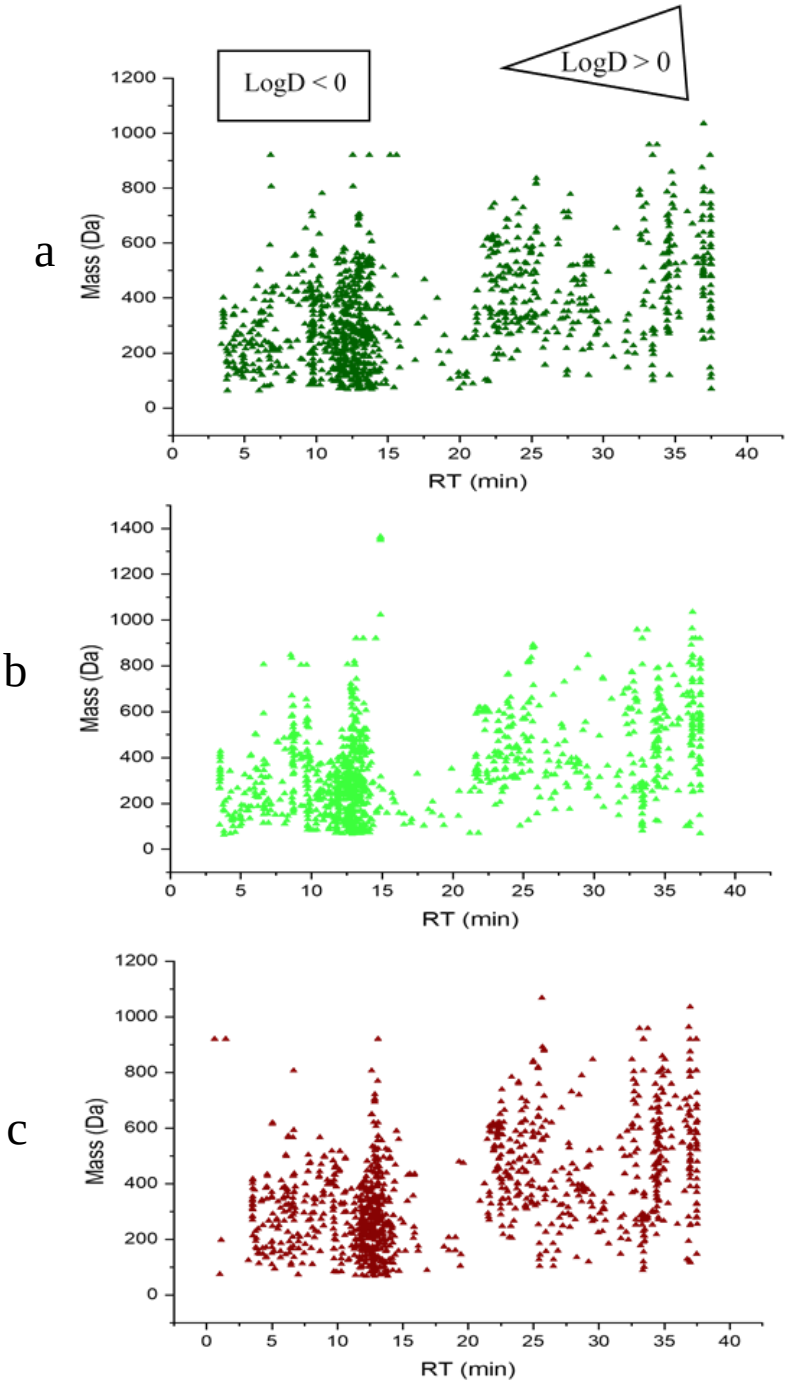
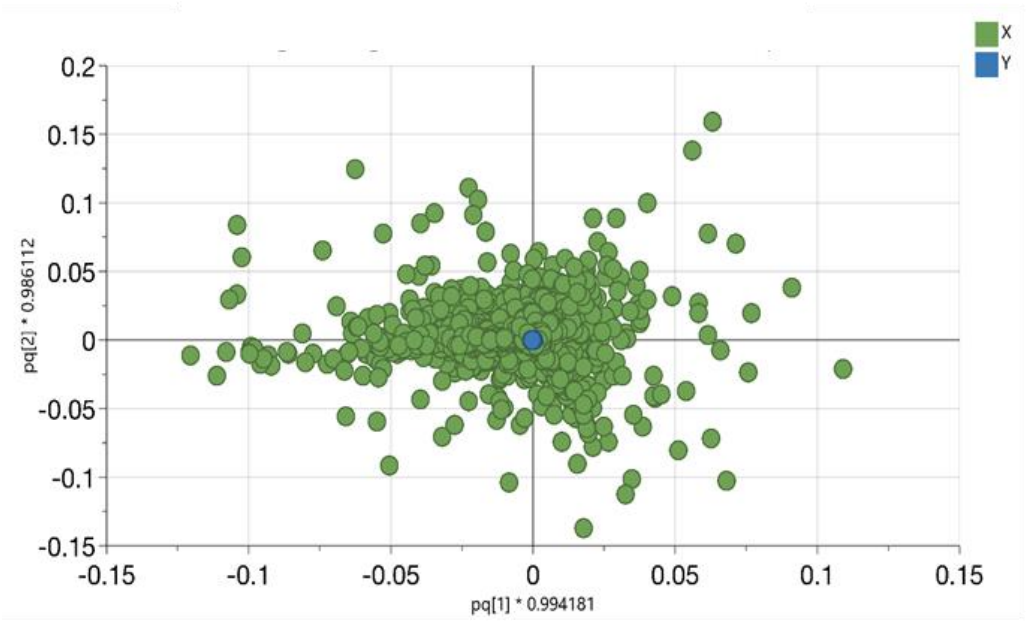


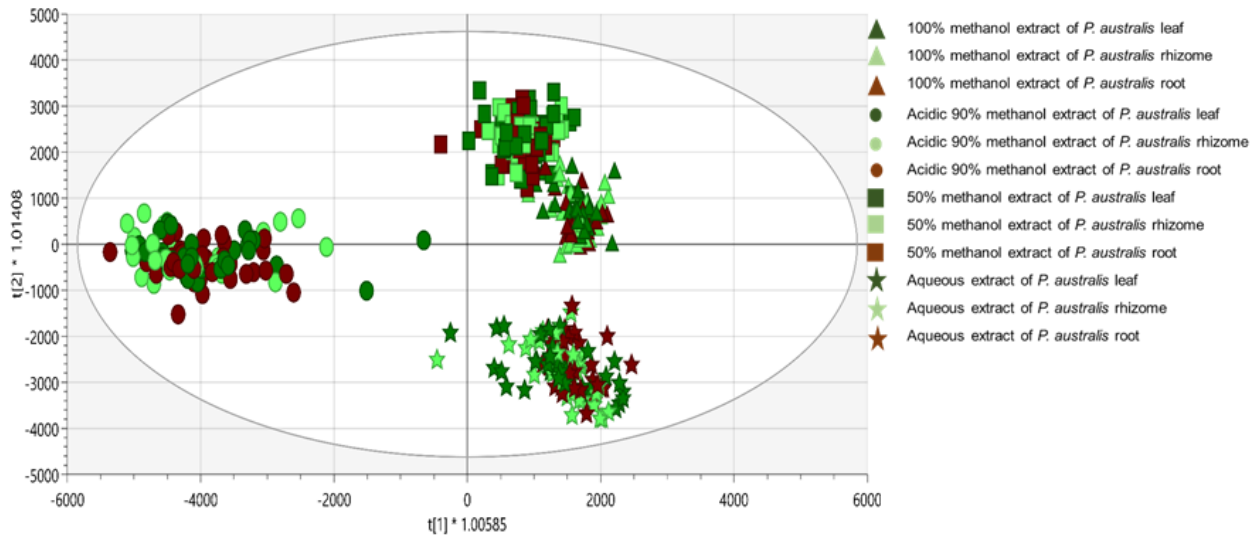
Figure 11. Retention time (RT)/Mass plot of *Phragmites australis* 100% MeOH extracts analyzed by RPLC-HILIC-ESI-TOF-MS in positive electrospray ionization mode. (a) Leaf; (b) Rhizome; (c) Root, which showed the features' separation according to their polarity and detected according to their m/z.

4.2.7. Different extracts of Phragmites australis's metabolic fingerprints

Untargeted metabolomics analysis demands minimal pretreatment methods to allow the detection of almost all the sample metabolites. Different leaf, rhizome, and root samples were extracted with the described four solvents (i.e., 100% MeOH, acidic 90% MeOH, 50% MeOH, and 100% H_2O), individually, which ensured the extraction of a wide range of metabolites from unpolar to highly polar metabolites relative to the solvents, which were used in the extraction process. All the extracts were analyzed along with a metabolites' fingerprint strategy. In OPLS-DA, the variables were *P. australis* metabolites, which were plotted in the loading score plot (as in Fig. 11a). The OPLS-DA model described the variables according to the solvents of the extraction class. Of the data variations, 18.3% ($R^2X(cum)$) are responsible to distinguish between the classes that were previously established based on solvents selections. The rest of the variation (orthogonal components) describes the variation within the solvent classes. The high value of those parameters indicates that the OPLS-DA model had a good classification and prediction efficiency to distinguish between different extracts, even though it described one variation (i.e., solvent type). This is owing to the accuracy of variables, which were separated with a robust and reproducible LC system. In Figure 11b, the samples are distributed according to t1 (predicative component) and t2 (orthogonal component). The predictive component (t1) separated the samples into two groups, the first group (negative side) contained the acidic 90% MeOH extracts and the second group (positive side) consisted of 100% MeOH, 50% MeOH, and 100% H_2O extracts (Fig. 11b). Moreover, the orthogonal component (t2) described the differences within the group. Consequently, it separated the second group into 100%MeOH and 50%MeOH in the positive part and the 100% H_2O extracts in the negative part (Fig. 10B). Further, in Figure 11c, the good separation was reached according to the cumulative goodness of fit and the cross-validation parameters of each variable R^2 and Q^2, respectively. Therefore, the model had no risk of overfitting. *P.australis* metabolomics' statistical analysis of the four different extracts reveals that the 90% acidic MeOH extracts are very different from the other extracts of various *P. australis* leaf, rhizome, and root samples. However, the 100% MeOH, 50% MeOH, and 100% H_2O extracts are more related to each other. 100% MeOH was generally located in the middle between 50% MeOH and 100% H_2O extracts, however, representing a similar solution behavior to all three extract types. A large t2 range value was observed mostly in acidic 90% MeOH extracts, that is, a value far above the critical limits in the score space. Hence, this is likely to be an outlying observation. The elimination of acidic 90% MeOH extracts might improve the model. Also, most of the metabolites extracted with acidic 90% MeOH have been detected in three other solvents. The same observation was made for *Lemna minor* extracts and can be described with the different occurrence of charged molecules caused by different pH values.


X
Y
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
pq[2] * 0.986112
-0.15
-0.1
-0.05
0
0.05
0.1
0.15
pq[1] * 0.994181

a


5000
4000
3000
2000
1000
0
-1000
-2000
-3000
-4000
-5000
t[2] * 1.01408
-6000
-4000
-2000
0
2000
4000
6000
t[1] * 1.00585
100% methanol extract of P. australis leaf
100% methanol extract of P. australis rhizome
100% methanol extract of P. australis root
Acidic 90% methanol extract of P. australis leaf
Acidic 90% methanol extract of P. australis rhizome
Acidic 90% methanol extract of P. australis root
50% methanol extract of P. australis leaf
50% methanol extract of P. australis rhizome
50% methanol extract of P. australis root
Aqueous extract of P. australis leaf
Aqueous extract of P. australis rhizome
Aqueous extract of P. australis root

b

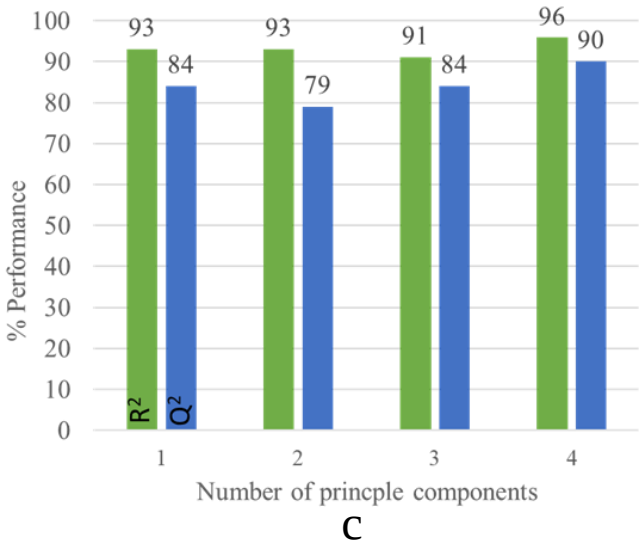


Figure 11. (a) The OPLS-DA score plot of different *Phragmites australis* extracts with a confidence limit of 95% discriminating according to the solvent used in the extraction. The variables were plotted according to the first principal component (t1) and the orthogonal component (t2). The triangles represent 100% MeOH extracts, circles represent acidic 90% MeOH extracts, the squares represent 50% MeOH extracts, and the stars represent 100% H_2O extracts, respectively. The green color represents leaf samples, the light green color represents rhizome samples, and the brown color represents root samples, respectively. Each symbol represents one observation of *P. australis* leaf, rhizome, and root plant part; (b) loading scatter plot for the selected principal components; (c) the Q^2/R^2 Overview plot displays the individual cumulative R^2 (green columns) and Q^2 (blue columns) for the goodness of fits and cross-validation parameters.

4.3. Conclusions

Three types of *Lemna minor* samples were extracted using different solvents. The extracts were then dried and redissolved residues analyzed using extended-polarity RPLC-HILIC-TOF-MS to determine their chromatographic/mass spectrometric fingerprint. This study produced an extended-polarity untargeted screening workflow that reflects the great importance of detecting (very) polar molecules separated by HILIC in the plant metabolome (in addition to the classic non-polar to polar molecules separated by RPLC). The ultimate decision to use 100% MeOH extracts, 50% MeOH, and 100% H_2O extracts pave the way to new views of plant metabolomes for (very) polar molecules. The untargeted screening was applied to study the aging behavior of *Lemna minor* under different freezing periods. The comparison of compounds in *Lemna minor* fresh and frozen samples showed that freezing under -80 ºC of *Lemna* changes the intensity and presence of compounds. To avoid any false positive and false negative interpretations, it is necessary to extract *Lemna minor* freshly to determine its metabolomic fingerprint. The results can then be used in phytomedicine research.

Furthermore, the features of unknown identity in this study can be subjected to further investigation in the form of various univariate and/or multivariate statistical analyses. This will then lead to mutual correlations of features. Where features relevant to a process are characterized by this analytical strategy, compound and/or analytical libraries can be used to identify the compounds. Therefore, the hypothesis **extraction strategy for *Lemna minor* and *Phragmites australis* can influence their metabolites chromatographic fingerprints can be accepted.**

5. Analytical method development and statistical data analysis for plant metabolites using RPLC-HILIC-MS

This chapter has been previously published with editorial changes as follows:

Wahman, R., Graßmann, J., Schröder, P., Letzel, T. (2019). Plant metabolomics workflows using reversed-phase LC and HILIC with ESI-TOF-MS

Current Trends in Mass Spectrometry, LCGC. N. Am., 37 (3), 8-15.

Author contributions:

Rofida Wahman and Thomas Letzel conceptualized the research objective and designed the methodology. Rofida Wahman and Peter Schröder created the test set and performed the plant growing. Rofida Wahman collected and analyzed the data. Rofida Wahman applied and validated the models and wrote the paper. Johanna Graßmann, Peter Schröder, and Thomas Letzel reviewed the manuscript. All authors approved the final version of the manuscript.

This chapter also had been submitted to Metabolites Journal as follows:

Wahman, R., Moser, S., Bieber, S., Cruzeiro, C., Schröder, P., Gilg, A., Lesske, F., Letzel, T. (2021). Untargeted analysis of *Lemna minor* metabolites: workflow and prioritization strategy comparing highly confident features between different mass spectrometers. Metobiltes. To be submitted.

Author contributions:

Rofida Wahman and Thomas Letzel designed the metabolomics study. Rofida Wahman prepared the samples, performed the untargeted metabolomics analysis, and ran the analyses on the metabolomics platform. Stefan Moser created the statistical design for data evaluation and data interpretation and performed the realization together with Rofida Wahman and Thomas Letzel. Catarina Cruzeiro and Peter Schröder created the test set and performed the plant growing together with Rofida Wahman. Rofida Wahman and Thomas Letzel conceived and drafted the manuscript. August Gilg and Frank Lesske are the IT specialists of the FOR-IDENT platform. All the authors contributed with critical intellectual input. All authors have read and agreed to the published version of the manuscript.

Hypothesis #2. Analytical workflows, applied in the determination of trace organic compounds in the aquatic environment, can be used for plant metabolites detection

5.1. Introduction

The metabolomics approach aims to contact the identification of metabolites and changes in their concentrations. The modern development of analytical techniques expanded the use of metabolomics in biological systems investigations. This approach permits remarkable insights into regulation mechanisms as well as studying responses to different perturbations. Over the decades, target and untargeted metabolomics analysis has become a routine application in different fields. The most common proposed workflow applies to the food and nutrition sciences. These fields adopt metabolomics- especially the untargeted strategy -as an analytical tool for decades (Consonni and Cagliani, 2019). Also, in medical research (Zhang *et al.*, 2020) metabolomics provide new insights that are correlated with other clinical variables (Bedia *et al.*, 2018) revealing new interactions on the molecular level between drugs and the human body. Besides, In environmental sciences metabolomics was considered as a technique to evaluate the physiological responses of the drug action, toxic effects, and metabolic disorders on the organisms from the initial chemical interaction to the final adverse outcome (adverse outcome pathways (AOP)) (Bedia *et al.*, 2018).

Usually, the first step toward metabolomes analysis is the separation technique, which provides a reproducible, precise, and wide range of polarity separation. A serial coupling of two different columns with different polarities secured the separation of different compounds from the complex matrix (e.g., unfractionated plant extracts) and detection using a high-resolution mass spectrometer (HRMS/MS). The LC system-MSMS provides a large amount of data, consequently, adequate workflow to process and analyze besides interpret it. Although the metabolomics data processing continues to be problematic, especially the metabolite identification process. From peak picking to identify the metabolites, the workflow contains numerous steps, which require to be precise, compatible, and with less effort and time-consuming. All the steps required quality control procedures for trustworthy metabolomics analysis outcomes. The workflow has several preprocessing steps, starting from peak picking and signal to noise threshold detection. Following by deleting the background and alignments of different data sets. The workflow ends with automated metabolite identification algorithms and the biological interpretation of the data. The identification of the metabolites is the fundamental step to transform the analytical data into biological knowledge, which is still considered the major bottleneck. The number of identified metabolites in untargeted metabolomics studies is in general below 50%. In LC-HRMS-based approaches, metabolites can be identified if the retention time (RT), mass, MS/MS fragment spectra are successfully matched with an authentic reference standard measured at the same instrument, because there are over 200,000 to 1,000,000 different metabolites in the plant kingdom. Many plant secondary metabolites have been demonstrated to have ‘specialized’ roles for adaptive significance in protection against predator and microbial infection. Further, the identification of specialized metabolites is still difficult due to the unavailability of commercial standards.

Several MS/MS databases have been established to facilitate metabolite annotation, such as MassBank (http://www.massbank.jp/index.html?lang=en), METLIN (http://metlin.scripps.edu/index.php), LipidBlast (http://fiehnlab.ucdavis.edu/projects/LipidBlast/), and ReSpect (http://spectra.psc.riken.jp). Moreover, plant metabolomics databases have been developed such as KNApSAcK (http://kanaya.naist.jp/KNApSAcK/), MetaCyc (http://metacyc.org), PlantMetabolomics.org (http://www.plantmetabolomics.org) , KEGG (http://www.genome.jp/kegg/), and PRIMe (http://prime.psc.riken.jp/). The choice of databases can have an essential influence on the study's interpretation. A small, well-defined database can suggest more meaningful hits compared to searches in huge ones, which is containing several (hundred-) thousand compounds from many different fields that are not related to the study scientific and biological question (Yang *et al.*, 2014; Doppler *et al.*, 2019). PLANT-IDENT(https://water.for-ident.org/#!home) is a relatively small database for the metabolites of the Lemnaceae family in addition to Poaceae, Brassicaceae, Nymphaeaceae, and, compounds from different families, which is organized for rapid search retrieval by a computer. The FOR-IDENT platform contains different parameters retention time (RT), mass, and MS/MS *in silico* fragmentation tool, in addition to retention time index (RTI), and logD values. The PLANT-IDENT also contains a chemotaxonomic parameter. Hence, a large number of the peaks were annotated, which could enhance the identification of biomarkers and indicators in untargeted metabolomics analysis. However, reference substances have to be used to prove the results.

In this chapter, the suspects and untargeted LC-MS/MS screening strategic relevance of plant samples were proven. A workflow is described for the identification of plant metabolites from the theoretical predictions to the final analyses in *Lemna* metabolic profile using reference materials. Further, the metabolomics data was analyzed with predictive methodology Orthogonal Partial Least Squares-Discriminant Analysis (OPLS-DA). We adopted the discrimination method to validate the features, which were extracted with the workflow. Consequently, the standard statistical methods of metabolomics data investigation were used in the identification of relevant variables (i.e. conditional attributes to each solvent), which related to the discrimination analysis (Blasco *et al.*, 2015), the steps of the workflow resembling the different ones. However, the usage of the open-access compound PLANT-IDENT database to identify the different metabolites represents an addition to plant metabolomics. A PLANT-IDENT database contains plant metabolites that will encourage untargeted screening in plant metabolomics including exact masses as well as other analytical results and makes use of the retention index tool.

5.2. Results and Discussion

5.2.1. OPLS-DA analysis of Lemna minor metabolic profile obtained with system A and B

Regarding the untargeted analysis, it is fundamental to assign the common metabolic profile between the different data sets of the same organism and/or different organisms obtained from

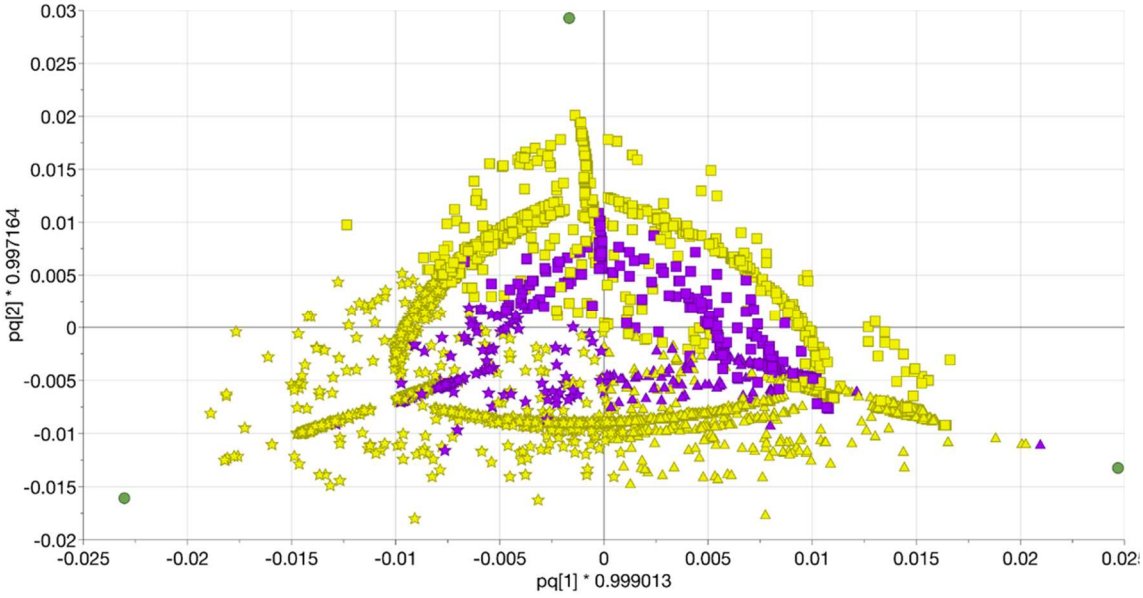
the same and/or different laboratories. The untargeted analysis concept depends mainly on the databases (mainly spectral). Thus, the datasets were investigated statistically to show the differences and the commons between the two MS (as a first additional quality control step).

The processed data was statistically investigated to test the data quality or (workflow reliability). The OPLS-DA was built to the three extracts of three *Lemna* samples, which were analyzed with single TOF-MS (system A), and QTOF-MS/MS (system B), respectively. Each extract with three injections was plotted individually. The Score plot of OPLS-DA explained 99.2% of the variations in the various extracts (R^2Y (cum)) with a higher predictive value (Q^2 (cum) = 0.877) (Fig. 13). The first component (t1) separated the 100% H_2O and 50%MeOH extracts in the negative part and the 100%MeOH extracts in the positive part. The orthogonal component (t2) distinguished between the 50%MeOH and 100% H_2O in the negative and positive parts, respectively. The loading scatters plot showed which variables (features) were expressed differently between the different extracts. The variable is responsible for the discrimination analysis. Hence, coloring the score plot according to the variable corresponding to the extracts emphasizes the separation between the different extracts (Fig. 13). In the loading score, the features were colored according to their main solvent, which was used to extract them from the *Lemna minor*. Those molecules could be extracted with the other two solvents. However, this indicated that this solvent is the best choice to extract them from the plant with high intensity (quantity). Furthermore, the suitability of different mass spectrometers could also be distinguished. Concluding, from the statistical analysis, the *L. minor* metabolic profile has small common features between the single and QTOF-MS. However, the different extracts could be significantly discriminated in both MS. Regarding the solvent used for the extraction; there is no big difference because it depends on the pKa of metabolites. Further, QTOF was able to detect more features than single TOF-MS due to its sensitivity.

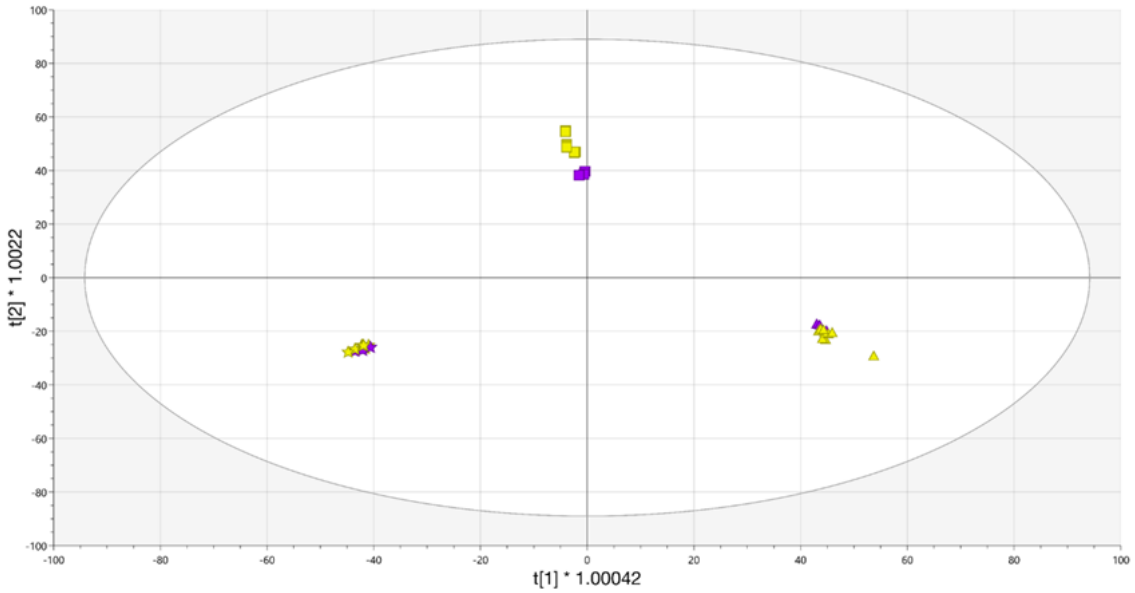
The prediction model was also built using the data from systems A and B to assign the MS to the unknown data using SIMCA 16 software (Fig. 13c). The predicated data sets were uploaded with unknown MS (blue and orange). The model could predict significantly the corresponding MS to each data set according to the calculated Y predicted response values and confidence intervals (Fig. 13). The feature number was the main reason for the discrimination. Further, each dataset was obtained with a different workflow of different software.

The processed data was statistically investigated to test the data quality or workflow reliability as a second additional quality control step. The important variables were responsible for the discrimination of the different samples. For example, the features (271.06@26.6) and (565.15@23.3) were detected in 100% MeOH and 50% MeOH extracts, respectively. Hence, the heat maps of both features were plotted in the score plot, which was colored according to the abundance of these two features (Fig. 14). The correlation statistical analysis showed that (271.06@26.6) belongs to the 100% MeOH extract cluster, in addition, the (565.15@23.3) feature belongs to the 50% MeOH extract cluster. They were identified using the PLANT-IDENT database. Then, they were confirmed with the reference standards. The first compound

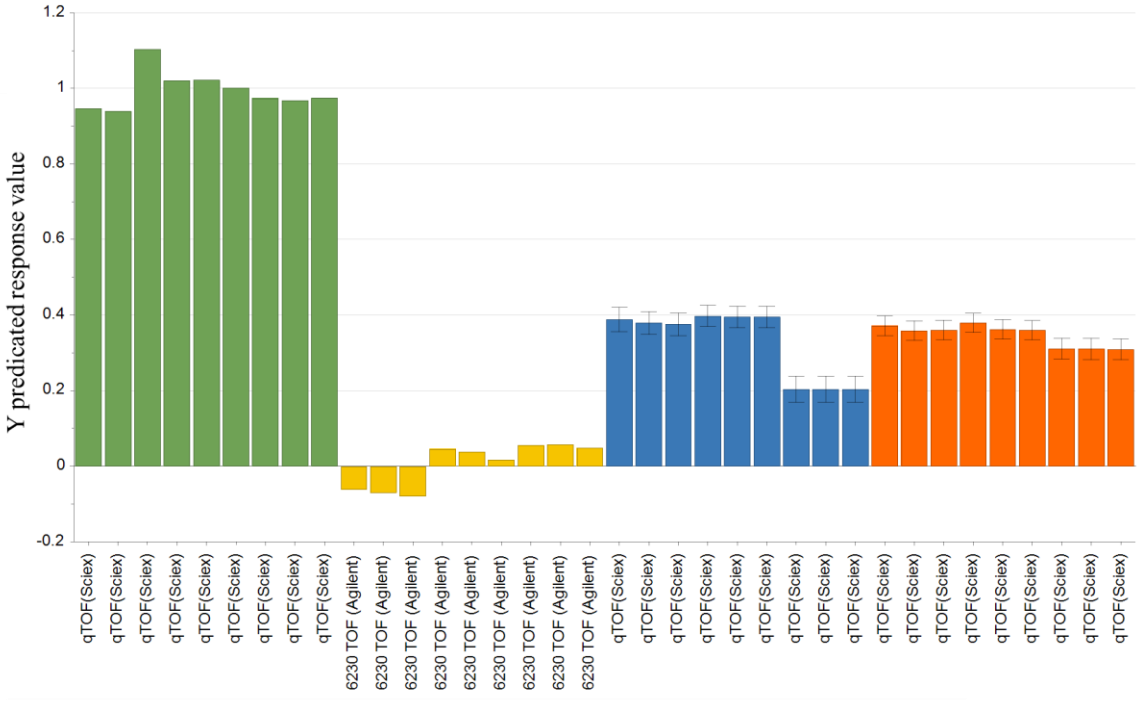
was apigenin and the latter was Apigenin-6-arabinopyranoside-8-glucopyranose, respectively (Table 14). Thus, the workflow provided statistically significant features (variables).



a



b



c

Figure 12. (a) The loading plot displays the relation between the different *Lemna minor* samples 100% MeOH (triangular), 50% MeOH (stars), and 100% H_2O (squares) extracts analyzed with TOF (violet) and QTOF (yellow); (b) OPLS-DA score scatter plot of *Lemna minor* samples. The confidence limit is 95%; (c) The prediction column plot of *Lemna minor* was analyzed with systems A and B with confidence intervals. The original data of QTOF single TOF was colored in green and yellow respectively. The unknown data sets were colored in blue and orange.

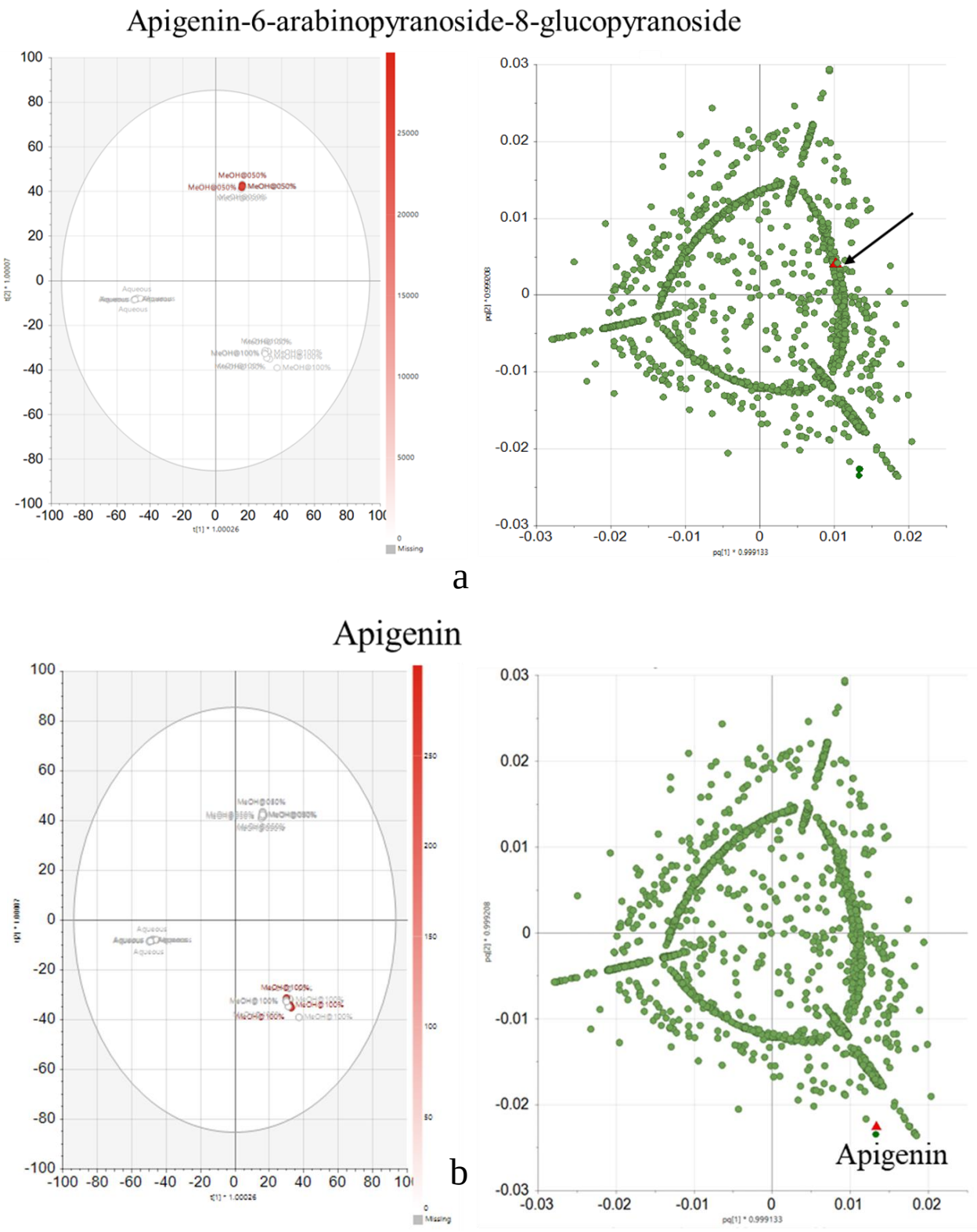


Figure 13. (a) OPLS-DA scores scatter plot of *Lemna minor* metabolic profile colored according to feature (271.06@26.6) apigenin, which was marked as red triangular in the loading plot of *Lemna minor* the same variable; (b) OPLS-DA score scatter plot of *Lemna minor* metabolic profile colored according to feature (565.15@23.3) apigenin-6-arabinopyranoside-8-glucopyranoside, which was marked as red triangular in the loading plot.

5.2.2. The strategy of Lemna minor metabolites identification based on the STOFF-IDENT database

First, the untargeted screening of *Lemna minor* data was evaluated by exporting them to the open-access platform FOR-IDENT (FI; https://water.for-ident.org). There, the features and RT were normalized, as well as an accurate mass was uploaded. They were compared with compounds in the database STOFF-IDENT (containing anthropogenic compounds expected in the 100% aqueous environment). This search gives several hits with suggested compounds by respective empirical formula and logD value. Applying this strategy resulted in hits and suggested amino acids and organic acids (see Fig. 12-molecules labeled with stars). For instance, glutamic acid and cinnamic acid were exclusive hits in STOFF-IDENT for 147.0532 Da and 148.052 Da, respectively. Thus, one can prove the identity of both very fast and easily by measuring the respective reference substance for each compound. However, this is not always resulting in hits if the corresponding molecule is not stored in the compound database. Furthermore, it is not always easy to differentiate clearly. For example, the mass of 165.079 gives 10 matches in FI, of which nine have a positive logD value and one a negative one. Here, the nine matches could easily be deleted because the feature eluted in the negative logD region, i.e. it was retarded on the HILIC column. The remaining hit was the amino acid phenylalanine as could further be evaluated by reference material. Moreover, sometimes the help of compound databases does not lead to unequivocal results. For example, a feature at 89.048 Da resulted in five hits. After excluding the two positive logD values, three hits remained with negative logD values. L-alanine was proven by a reference standard also eluting in the HILIC region. In general, this strategy leads to explicit and filtered suggestions for compound identities. Furthermore, tryptophan, leucine, and isoleucine (as shown in Fig.12) were identified similarly. After excluding the positive logD results, the remaining hits with negative logD values were either the corresponding amino acid or various synthetic compounds. The latter are typically anthropogenic (brought into the environment) and cannot be originated from the plant, thus later hits were neglected in this study. The reference standards for the remaining amino acids were injected to prove the identity and presence in the *Lemna minor* extracts, which also were reported previously in the literature (Maciejewska-Potapczyk *et al.*, 1975). For a successful hidden target screening strategy (Letzel *et al.*, 2015) using the compound database STOFF-IDENT, amino acid standards were available proving directly the hits of expected amino acid in *Lemna minor* methanol extracts. The identified amino acids are phenylalanine, L-leucine, L-isoleucine, tryptophan, glutamic acid, and L-alanine, as marked and labeled in (Fig. 12). The amino acid standards were injected in triplicate and the mean of each standard mass (S) and RTs were calculated. The identification was done by comparing the RT deviations between standard and extract RT means, which was in the range of 0 min (phenylalanine) to 0.15 (tryptophan).

Moreover, the mass variation between the standard mass (S) and the extracted mass is less than 5 ppm. In Table 12, the six identified amino acids are listed with their mean isotopic mass and mean RT. Also, standards mean isotopic mass and RT were calculated for the three

injections. The identification was done by the following workflow. The differences in RT (Δ RT) between the standard and the extract were in the range of ≈ 0 to 0.15 min. The variations between the extract and the standard mean isotopic masses (Δ ppm) were calculated for the six amino acids. Δ ppm was in the range of -3.05 for L-leucine to 0.08 for phenylalanine. All the amino acids were in the same RT and mass range as formerly reported for the serial coupling HILIC-RPLC in a 100% aqueous environment (Bieber *et al.*, 2017). Also without using a tandem-MS (and its significant fragment spectra), the hits result in category 1 of the identification scheme published earlier (Letzel *et al.*, 2014; Schymanski *et al.*, 2014).

Table 12. List of exemplarily found amino acids in *Lemna minor* extract with the mean monoisotopic mass of reference standard (S), extract (M), the variation between them, mean RT of the reference standard (S), and the extract (M) and the variation between them, mean RT of the reference standard (S), and the extract (M) and the variation between them.

Compound Name	Mean Monoisotopic Mass (Da) (S)	Mean Monoisotopic Mass (Da) (M)	Δ ppm	Mean RT (Min) (S)	Mean RT (Min) (M)	Δ RT
L-Isoleucine	131.0944	131.0947	-2.29	10.82	10.83	-0.01
L-Leucine	131.0941	131.0945	-3.05	10.79	10.83	-0.04
Phenylalanine	165.0787	165.0789	0.081	10.71	10.72	0
Tryptophan	204.0898	204.0899	-0.26	11.45	11.61	-0.15
Glutamic acid	147.0529	147.0525	2.49	12.63	12.71	-0.08
L-Alanine	89.0477	89.0476	1.12	12.84	12.93	-0.09

5.2.3. Strategy of Lemna minor metabolites identification based on Metlin database

Generally, untargeted screening data from plant extracts can be compared with available metabolites databases like Metlin (https://metlin.scripps.edu/landing_page.php?pgcontent=batch_search) to find expected molecular hits in the same hidden target strategy as shown above. A disadvantage compared to the FOR-IDENT platform is that other databases like Metlin have no automated RT and mass comparison functionality. Applying the Metlin database manually, four compounds were identified through comparing the RT and logD of the compounds with the literature, which are pongamoside B, eleganin, 2-aminooctadeconic acid, and 2,6-nonadienoic acid (Fig. 12). These compounds and cinnamic acid (later additionally observed from FOR-IDENT) were not proven by standards so far, thus these compounds remain on category 2 like suggested by (Letzel *et al.*, 2014; Schymanski *et al.*, 2014). However, the four compounds presumably identified through Metlin and cinnamic acid were listed in Table 13 by the same workflow. The variation between measured monoisotopic mass and mass found in literature is less than ± 4 ppm.

Moreover, the standard deviation (SD) of RT and the relative standard deviation (RSD) were calculated. The RSDs for all compounds were less than 0.1% for the three injections, which indicates the reproducibility of the LC system and as formerly reported for the serial coupling HILIC-RPLC (Bieber *et al.*, 2017). Consequently, the RT for each compound could be used in the standards RT indices (RTI) calibration curve to calculate the logD (pH 7) (Sylvia Grosse, 2017). The lower polarity limit of polar compounds was set to a logD (pH 7) value of zero because of the applied RPLC column, thus it can likely retain compounds above this polarity. The polarity region below a logD zero in this study is restricted to HILIC. Pongamoside B, eleganin, 2-aminooctadeconic acid, and 2,6-nonadienoic acid were retained by RPLC because their RT was higher than 15 min. The logDs of suspects (experimental) were calculated from the twelve standards calibration curve of logD (pH7) and RTI. The difference between experimental and predicted logD (pH 7) was in the range of 0.5 (cinnamic acid) to 1.6 (2-Aminooctadecanoic acid). Subsequently, for all compounds, the calculated logD is compatible with the literature logD, which is predicated using ChemAxon software. In consideration of supporting parameters: variation between masses, RT, and logD values of cinnamic acid, pongamoside B, eleganin, 2-aminooctadeconic acid, and 2,6-nonadienoic acid were presumably identified through Metlin database using hidden target screening strategy.

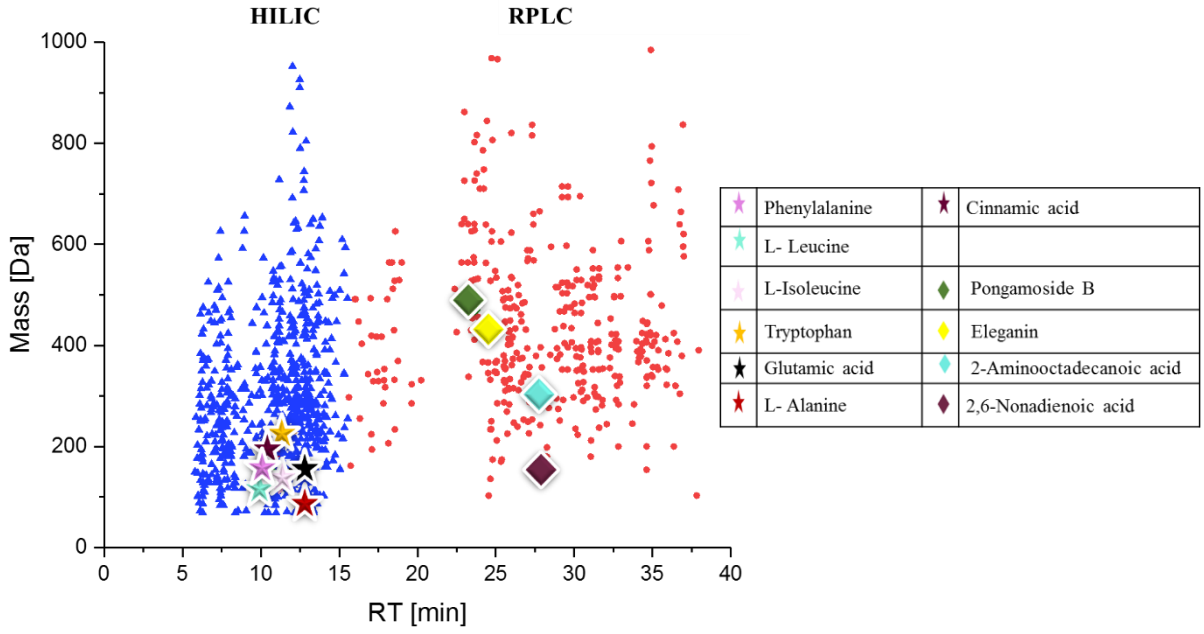


Figure 14. Retention time/mass plot of *Lemna minor* extract (of an untreated sample). Blue, triangles represent HILIC retarded molecules eluting (0-15) min and red, circles represent RPLC retarded molecules eluting (16-38) min. Different colored and shaped dots represent examples of identified (stars) and expected (diamonds) compounds (as in a table).

Table 13. Exemplary compounds found in *Lemna minor* extract using a hidden target screening strategy. Monoisotopic mass in the literature (L), mean monoisotopic in extract (M) and the variation between them, mean RT of the extract (M), standard deviation, relative standard deviation, log D is predicted (P), and log D experimental (E) are listed.

Compound Name	Mean Monoisotopic Mass (Da) (Ph)	Mean Monoisotopic Mass (Da) (L)	Δ ppm	Mean RT (Min) (Ph)	Mean RT (Min) (L)	Δ RT
L-Isoleucine	131.0951	131.0947	2.67	11.21	10.83	0.38
L-Leucine	131.0951	131.0945	4.83	10.88	10.83	0.05
Phenylalanine	165.0788	165.0789	-0.40	10.71	10.72	0.33
Glutamic acid	147.0537	147.0525	8.27	12.47	12.71	-0.24
L-Alanine	89.0476	89.0476	0	12.79	12.93	-0.14

5.2.4. Strategy of Lemna minor metabolites identification based on PLANT-IDENT database

A comprehensive strategy was applied to identify metabolites in untargeted metabolomics. According to the statistical results of *Lemna* metabolic profile using systems A and B. the system A data was identified using STOFF-IDENT and Metlin databases. The data of system B were identified using the PLANT-IDENT database. The metabolites identification via Sciex OS software was performed to achieve precursor ions and retention time. Later, database searching and scoring were done automatically. The general strategy framework for metabolites identification is shown in Figure 2. Metabolites identification from known databases depends substantiality on the following: (1) comprehensive data integration and weighting of retention time, mass, and MSMS fragments, and (2) employment or ignoring of MSMS intensity information (Zhao *et al.*, 2018). For PLANT-IDENT, RT and mass are applied simultaneously as a prerequisite with MSMS data for scoring evaluation with the help of user-defined weights. Two stationary phases, the HILIC and RPLC (additional RTI), the weighting of RT, mass, and MSMS data are provided for optimal choice and comparison of each feature in the PLANT-IDENT. The determination scoring of metabolites is calculated based on the accurate finding of the features RT, mass, MSMS data extracted from samples, and PLANT-IDENT, which is achieved by using the predetermined differences values Δppm. Then, to reduce the false fragments determination, the intensity threshold is defined. The threshold differentiates the line between the signals and noise level. After this, the stationary phase (reversed-phase (RP) and HILIC) searching are simultaneously employed for feature searching. The fundamentally different principle between the two columns is (RTI) to calibrate the RT of the RP eluted features.

The scoring ranges from zero to 100, in which the candidate with a higher possibility should have a larger value of scoring and be marked in the "look at" column in the FOR-IDENT platform. Those candidates are prioritized by the platform and considered for validation using reference standards.

5.2.5. Application of Lemna minor metabolites identification

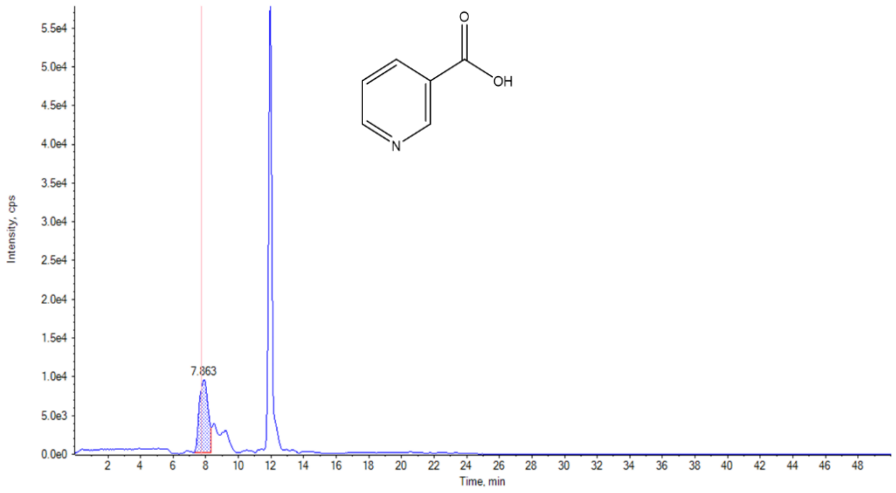
Lemna minor untargeted metabolomics study was performed using the proposed workflow of metabolites identification (Fig. 2). The three extracts (100% MeOH, 50% MeOH, and 100% H_2O) were analyzed with RPLC-HILIC-QTOF-MS/MS. The MSMS data were collected from the eight parallel experiments at a collision energy of 40 ± 20 eV in the full cycle scan. The precursor ion was aligned and filtered with Marker View Software. The obtained data was uploaded to Sciex OS software to extract the available MSMS data for each feature, which was divided according to the elution column. First, the features eluted with the HILIC column with RT< 15 minutes. Features were uploaded into the PLANT-IDENT database. The parameters set according to the experimental pH at pH=7, the absolute mass deviation of the precursor ion and fragments was 5 ppm. The intensity threshold was 5 with a positive ion (H^+). The search was scored according to the mass screening, and MS/MS.

The results were filtered by different filters in a succession process. First the filtration according to the polarity (or hydrophilicity) using the LogD as an indicator. The candidates who have LogD ≥ 0 were eliminated. Second, the candidates were filtered according to the score of the mass scoring and MS/MS (if present). To avoid false-positive and negative results, the candidates were moved to the next filtration step, which has priority "look at" by the database. Third, the fragments of candidates were also compared against the Metfrag (*in silico* fragmentation). Finally, in the chemotaxonomic filter, the candidates were considered only if they were reported in the Lemnacea and/or Araceae family. Accordingly, forty-four compounds were identified with the reference standards and upgraded into level 1, which were listed in (Table 14). The compounds have RT deviation of less than 1 min. and a mass deviation of less than 5ppm between the reference standard mass and the *Lemna minor* candidate feature mass.

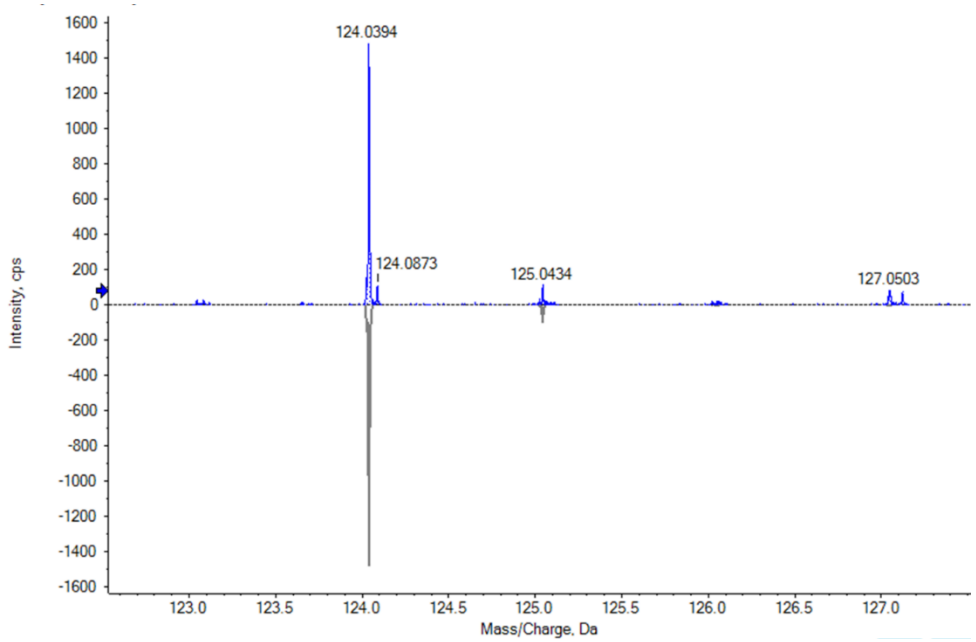
The PLANT-IDENT database suggested pegamine, tryptophan, aspartic acid, alanine, valine, vitexin, 4-methoxy cinnamic acid, di-L-alanine, niacin, nicotinamide, and betaine (trimethylglycine) candidates matching features. They have LogD ≥ 0 at (pH= 7) except pergamine, which was eliminated. They also have priority "look at" and standards reference injected for validation except for betaine. Subsequently, the filtration parameters decreased the number of the results into 16 metabolites. Those metabolites were annotated and classified in the second level and confirmed with standards reference injections. The amino acids (phenylalanine, proline, tryptophan, alanine, tyrosine, aspartic acid, isoleucine, serine, and valine) were identified in *L. minor* using system A using STOFF-IDENT and reference standards. The extracted ion chromatogram, mass spectrum, and MS/MS spectrum of niacin were shown as an example in (Fig. 15).

Secondly, the RP part with RT > 15 minutes was uploaded into the PLANT-IDENT database. The retention time of the calibration mixture was also uploaded to normalize the RT. The normalization was performed according to the calibration curve of the target analysis of the standard mixture. The mixture consists of compounds, which have affixed logD at different

pHs. Here the scoring of suspected compounds depends on the same parameters in addition to RTI screening. Each is 25% of the total score. The 28 compounds were separated from the RP column: apigenin-6,8-diglucopyranosid, apigenin-7-glucoside, apigenin-5-glucoside, apigenin-6-arabinopyranoside-8-glucopyranoside, apigenin, robinetin, quercetin, luteolin, kaempferol, acacetin, orientin, Isoorientin, peonidin, 6-methoxy-flavon, flavon, naringenin-7-O-glucoside, quercetin-3-glucoside, saponarin, 5-hydroxy-6-methoxy-flavon, luteolin-3',7-di-O-glucoside, apiin, chrysoeriol, umbelliferone, norwogonin, isovitexin, tricin, galangin, and myricetin. Further, the MSMS fragments of the reference standard and the identified peaks in the *Lemna* metabolic profile were compared with the literature, as shown in (Table 14). Then, the data was stored in the Sciex OS software library. The extracted ion chromatogram, mass spectrum, and MS/MS spectrum of apiin were shown as an example in (Fig. 16). According to chemotaxonomy criteria, results, 42 compounds were separated previously from Lemnaceae or Araceae family. Further, robinetin and norwogonin were not reported from Lemnaceae or Araceae family according to the available literature. Those two compounds were not filtered with the last filter of the chemotaxonomy because they are hydroxyl-flavone derivatives. Also, luteolin, quercetin, kaempferol, apigenin, and myricetin and their glucosides had been identified in the family before, which are hydroxyl flavones derivatives (AKHTAR *et al.*, 2010; Tsolmon *et al.*, 2020). Besides, robinetin is known as 5-deoxy-myricetin (pentahydroxy-flavone), and norwogonin is a 5,7,8-trihydroxy flavone. Therefore, robinetin and norwogonin reference standards were injected.



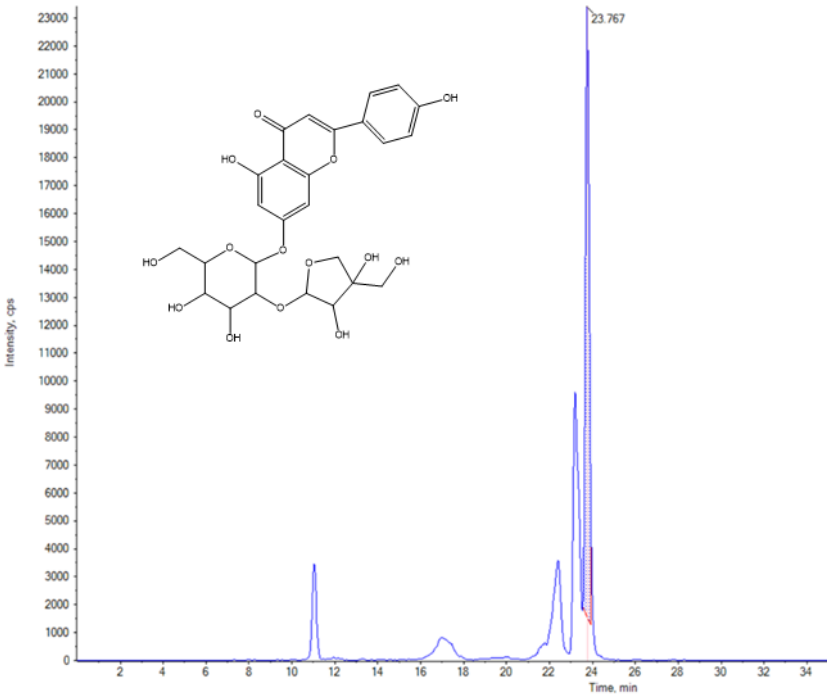
a



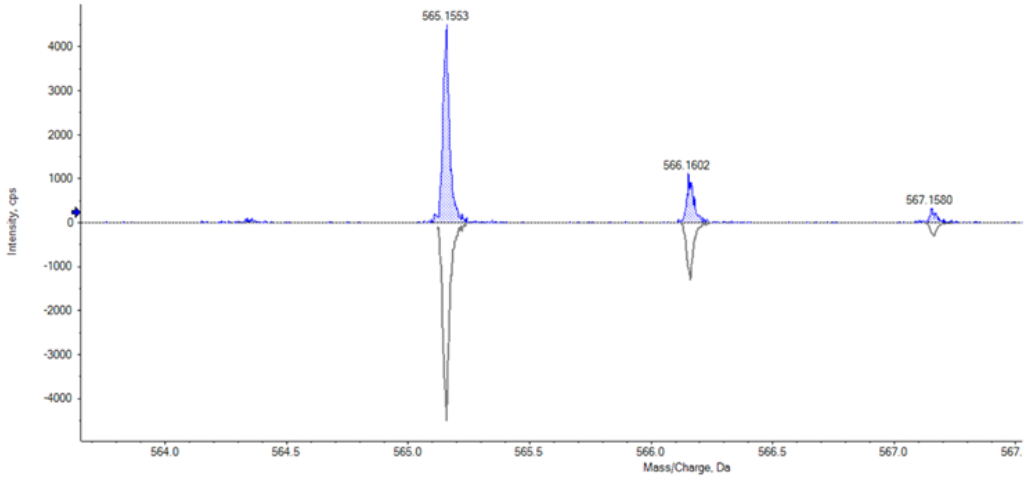
b

c

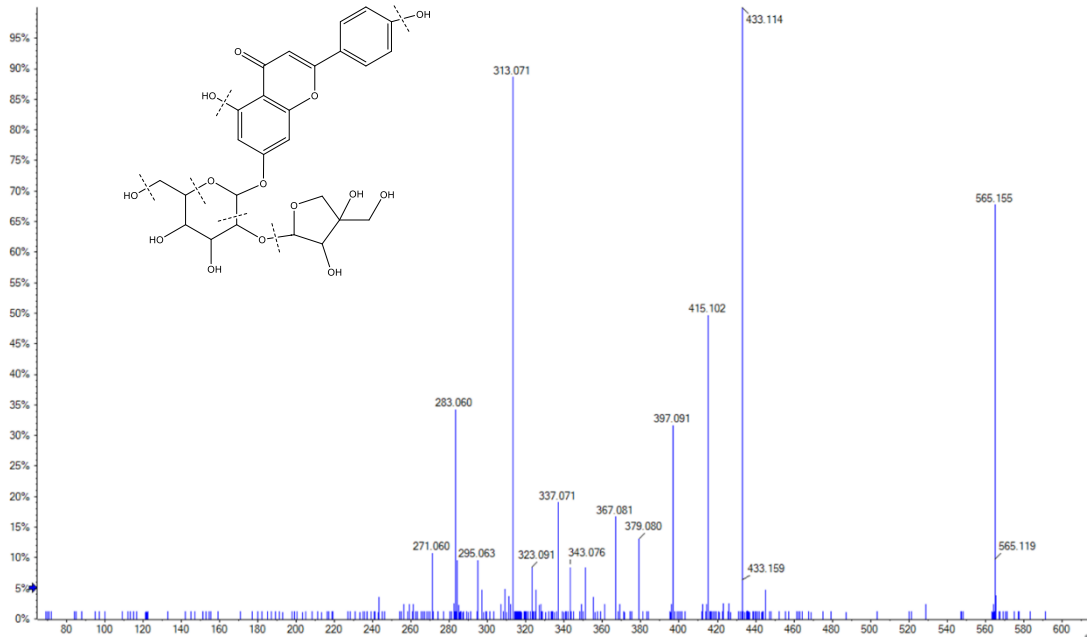
Figure 15. (a) Extracted ion chromatogram (EIC) of niacin shows intensity in counts per second (cps), appears on the y-axis, while RT appears on the x-axis; (b) Mass Spectrum of niacin represents the masses and intensity of the ions with particular mass-to-charge (m/z) values in Daltons; (c) MS/MS spectrum of niacin shows the fragments and relative intensity (%).


23.767
Intensity, cps
Time, min

a


565.1553
566.1602
567.1580
Intensity, cps
Mass/Charge, Da

b



C

Figure 16. (a) Extracted ion chromatogram (EIC) of apiin shows intensity in counts per second (cps), appears on the y-axis, while RT appears on the x-axis; (b) Mass Spectrum of apiin represents the masses and intensity of the ions with particular mass-to-charge (m/z) values in Daltons; (c) MS/MS spectrum of apiin shows the fragments and relative intensity (%).

Table 14. The compounds were identified in *Lemna minor* metabolic profile identified by PLANT-IDENT. Retention time (RT) means of standards (S), measured (M), and the RT deviation. Also, the mass means of standards (S), measured (M), and the deviation between them. The mean fragments of standards and measured were compared with the literature and listed with the references. Compounds could be detected in system B (single TOF-MS) marked with *

Compound Name	RT (S) [Min]	RT (M) [Min]	ΔRT [Min]	Mass (S) [Da]	Mass (M) [Da]	Δppm	MSMS fragments	References
Vitexin	7.5	7.3	0.2	433.1133	433.1129	0.8	433;415;397;379;337;313;283	(Waridel *et al.*, 2001)
Niacin*	7.6	7.8	-0.2	124.0394	124.0393	0.7	124;96;80;78	(Ndolo *et al.*, 2015)
Nicotinamide	7.8	7.6	0.1	123.0554	123.0553	0.8	123;106;80;78	(Hau *et al.*, 2001)
Phenylalanine*	11.0	11.1	-0.1	166.0866	166.08627	2.0	120;103;77	MassBank of North America (MoNA)
Leucine/Isoleucine*	11.2	11.2	-0.1	132.1018	132.1020	-1.4	86;69;44;30	(MoNA)
Tryptophan*	11.7	11.7	0.0	205.0973	205.0970	1.8	188;146;144	(MoNA)
Valine*	12.1	11.9	0.1	118.0863	118.0862	0.8	72;71;55	(MoNA)
Tyrosine*	12.3	12.2	0.1	182.0811	182.0810	1.9	136;123;119	(MoNA)
Proline*	12.4	12.4	0.0	116.0705	116.0707	0.3	70;68;43	(MoNA)
Glutamic acid*	12.5	12.6	-0.2	147.0434	147.0430	3.0	130;102;84	(MoNA)
Aspartic acid*	12.7	12.7	0.0	134.0447	134.0447	-0.2	134;115	(MoNA)
Di-L-Alanine	12.7	12.8	-0.1	161.0928	161.0920	4.9	161;115;90	(MoNA)
4-Methoxy cinnamic acid	13.4	13.1	0.3	179.0706	179.0708	-0.9	147;137	(Wang *et al.*, 2020)
Alanine*	13.4	13.2	0.2	90.0550	90.0548	2.1	44;28	(MoNA)
Threonine*	13.6	13.4	0.2	120.0656	120.0653	2.8	73;56	(MoNA)
Serine*	14.0	13.8	0.2	106.0500	116.0499	0.9	60;42;43	(MoNA)

Compound Name	**RT (S) [Min]**	**RT (M) [Min]**	**ΔRT [Min]**	**Mass (S) [Da]**	**Mass (M) [Da]**	**Δppm**	**MSMS fragments**	**References**
Apigenin-6,8-diglucopyranoside*	15.8	15.7	0.1	595.1659	595.1658	0.2	595; 383	(Sakalem *et al.*, 2012)
Robinetin	15.8	15.9	-0.1	303.0494	303.0493	0.3	285;267;147	(MoNA)
Quercetin	24.8	24.9	-0.1	303.0549	303.0544	1.7	303;285;257;229;165	(Tsimogiannis *et al.*, 2007)
Luteolin	24.8	24.6	0.2	287.0562	287.0557	1.6	287;269;241;153	(Tsimogiannis *et al.*, 2007)
Acacetin	28.8	29.1	-0.3	285.0759	285.0760	-0.4	285;242;153	(Xu *et al.*, 2017)
Apigenin 7-glucoside	24.8	24.7	0.1	433.1130	433.1132	-0.4	433;271	(Lee *et al.*, 2018)
Orientin	25.1	25.2	-0.1	449.1123	449.1134	-2.6	449; 329	(Sakalem *et al.*, 2012)
Peonidin	25.6	25.2	0.4	302.0785	302.0792	-2.4	302;283;197	(MoNA)
6-Methoxyflavone	30.4	30.6	-0.2	253.0879	253.0881	-0.7	253; 238; 210	NIST
Luteolin-3',7-di-*O*-glucoside	23.8	23.6	0.3	611.1640	611.1622	2.8	611;449;287	(MoNA)
Kaempferol	29.0	29.1	-0.1	287.0531	287.0540	-3.1	287;269;231;165;153;133	(Tsimogiannis *et al.*, 2007)
Apigenin	26.8	26.7	0.1	271.0603	271.0604	-0.6	271;253;153	(Tsimogiannis *et al.*, 2007)

Compound Name	**RT (S) [Min]**	**RT (M) [Min]**	**ΔRT [Min]**	**Mass (S) [Da]**	**Mass (M) [Da]**	**Δppm**	**MSMS fragments**	**References**
Flavon	29.9	29.6	0.2	223.0756	223.0748	3.6	223;178;152;121	(MoNA)
Naringenin-7-*O*-glucoside	25.0	24.1	0.9	435.1298	435.1285	2.9	435;273	(Le Gall *et al.*, 2003)
Quercetin-3-glucoside	24.2	24.3	-0.1	465.1018	465.1022	-0.7	465; 303	(Jang *et al.*, 2018)
Saponarin	23.8	24.0	-0.2	595.1638	595.1663	-4.2	433;415;397;367;337;283;271	(Akhtar *et al.*, 2020)
5-Hydroxy-6-Methoxyflavon	31.3	31.1	0.1	269.0823	269.0819	1.3	269;254;104	(MoNA)
Apiin	24.6	23.8	0.9	565.1566	565.1559	1.3	433;313	(Crow *et al.*, 1986)
Chrysoeriol	26.9	26.8	0.1	301.0731	301.0722	2.9	286;121	(Xu *et al.*, 2017)
Isoorientin	23.8	23.6	0.2	449.1085	449.1095	-2.1	499;329;299;165	(Pereira *et al.*, 2005)
Umbelliferone	24.7	24.4	0.2	163.0396	163.0391	2.9	135;107	(Wang *et al.*, 2020)
Apigenin-5-glucoside	24.2	23.9	0.3	433.1127	433.1127	0.1	433;271	(Lee *et al.*, 2018)
Apigenin-6-arabopyranoside-8-glucopyranose	23.4	23.3	0.2	565.1550	565.1557	-1.2	565;547;379;337;325;295;121	(Simirgiotis *et al.*, 2013)
Norwogonin	24.2	24.0	0.2	271.0604	271.0599	1.8	271;253;241;225	(Gao *et al.*, 2014)
Isovitexin	24.1	23.9	0.2	433.1125	433.1134	-2.0	313;295;284;283;267	(Waridel *et al.*, 2001)

Compound Name	**RT (S) [Min]**	**RT (M) [Min]**	**ΔRT [Min]**	**Mass (S) [Da]**	**Mass (M) [Da]**	**Δppm**	**MSMS fragments**	**References**
Tricin	26.8	26.3	0.6	331.0811	331.0796	4.7	331;315	(Yang *et al.*, 2014)
Galangin	29.4	29.4	-0.1	271.0602	271.0608	2.3	271;253	(Hughes *et al.*, 2001)
Myricetin	25.1	25.1	0.0	319.0440	319.0453	4.0	301;283;265;111	(Ma *et al.*, 1997)

5.3. Conclusion

The advances in MS are driving untargeted metabolomics analysis through generating an accurate empirical formula, using tandem mass spectrometry with structural information observed by molecule fragmentation and normalizing retention times, and correlating with logD values.

Thus, the application of data evaluation platforms and compound databases can be helpful in the identification of compounds due to the lack of reference standards by searching for the exact masses in suspect and untargeted screening. The PLANT-IDENT database will help researchers, which utilize highly sensitive mass spectrometers in the identification of plant metabolomics. This database will lead to clear hits as stated above via different parameters like mass, RT, and logD values. Reference substances will be used to prove the results, which were obtained with untargeted screening strategies like hidden target screening. PLANT-IDENT database contains up to 3019 naturally occurring compounds with their physicochemical properties, in addition to, *Silico* fragmentation and chemotaxonomic data. The workflow considered the features represented in triplicates to minimize the false-positive results. The identification of *Lemna* metabolites proceeded according to different filters: LogD, mass deviation, MSMS fragment comparison, and chemotaxonomy filter. Moreover, compounds were identified using reference standards

The applied workflow in *Lemna minor* metabolites analysis will be a cornerstone for subsequent research. Despite using different mass spectrometers (systems A and B), researchers can apply this workflow with/or without reference materials to identify suspect and hidden targets.

The statistical investigation of the workflow was conducted through different statistical tests to monitor the metabolite differences between different samples, which were obtained with the different mass spectrometer. These steps provided more quality control steps on the applied workflow.

The results showed that there were some differences in *Lemna's* metabolic profile, which were obtained with the system A and B. Detailed information of the applied workflow including all parameter settings, and criteria should be required in all studies. Open-access tools, in addition to software (compound and spectral databases), could be a way for the enhancement of untargeted screening.

Thus, the hypothesis of **analytical workflows, applied in the determination of trace organic compounds in the aquatic environment, can be used for plant metabolites detection** could be accepted.

6. Study the effect of the different environmental pollutants on plant metabolism

The *Phragmites australis* study has been previously published with editorial changes as follows:

Wahman, R., Graßmann, J., Sauvêtre, A., Schröder, P., Moser, S., Letzel, T. (2021). Untargeted metabolomics studies on drug-incubated *Phragmites australis* profiles. Metabolites 11(1):2.

Author contributions:

Rofida Wahman and Thomas Letzel designed the metabolomics study. Rofida Wahman prepared the samples, performed the untargeted metabolomics analysis, and ran the analyses on the metabolomics platform. Stefan Moser created the statistical design for data evaluation and data interpretation and performed the realization together with Rofida Wahman and Thomas Letzel. Andrés Sauvêtre, and Peter Schröder created the test set and performed the plant growing and incubation together with Rofida Wahman. They also supported Rofida Wahman in biological data interpretation. Rofida Wahman and Thomas Letzel conceived and drafted the manuscript. All the authors contributed with critical intellectual input. All authors have read and agreed to the published version of the manuscript.

Also, the *Lemna minor* study using system A has been previously published with editorial changes as follows:

Wahman, R., Cruzeiro, C., Grassmann, J., Schröder, P., Letzel, T. (2021). The changes in *Lemna minor* metabolomic profile: A response to diclofenac incubation. Chemosphere 287, 13207.

Author contributions:

Rofida Wahman, Catarina Cruzeiro, Peter Schröder and Thomas Letzel conceptualized the research objective. Rofida Wahman prepared the samples, performed the untargeted metabolomics analysis, and ran the analyses on the metabolomics platform. Rofida Wahman wrote the paper. Catarina Cruzeiro, and Peter Schröder created the test set and performed the plant growing and incubation together with Rofida Wahman. Rofida Wahman and Peter Schröder conceptualized and analyzed the biochemical part. Johanna Grassmann reviewed the manuscript. Rofida Wahman, Catarina Cruzeiro, Peter Schröder and Thomas Letzel critically reviewed the manuscript.

The study Also, the *Lemna minor* study using system B is under-preparation for submission into Metabolomics Journal as follows:

Wahman, R., Schröder, P., Cruzeiro, C., Drewes, J., Letzel, T. (2021). The effect of prolonged diclofenac incubation on *Lemna minor* metabolomic profile: flavonoids content. Metabolomics.

Author contributions:

Rofida Wahman, Peter Schröder, Catarina Cruzeiro, and Thomas Letzel conceptualized the research objective. Rofida Wahman prepared the samples, performed the untargeted metabolomics analysis, and ran the analyses on the metabolomics platform. Rofida Wahman wrote the paper. Catarina Cruzeiro, Peter Schröder created the test set and performed the plant growing and incubation together with Rofida Wahman. Rofida Wahman and Peter Schröder conceptualized and analyzed the biochemical part. Jörg Drewes reviewed the manuscript. Rofida Wahman, Peter Schröder, Catarina Cruzeiro, and Jörg Drewes Thomas Letzel will critically review the manuscript.

Hypothesis #3. The presence of pharmaceuticals such as diclofenac and carbamazepine in the aquatic environment can affect the biosynthetic pathways of *Lemna minor* and *Phragmites australis*

6.1. Introduction

System A

Metabolomics is the approach for the overall investigation of metabolite variations in biological systems. Metabolomics studies are essential to characterize these profiles as the metabolites change significantly during biochemical reactions. Consequently, metabolic profiling can be used as a robust tool to discuss the metabolic response of plants to environmental disorders, such as xenobiotics, nutrient deficiency, high salinity, and temperature stress (Kralova *et al.*, 2012). As one knows, primary (e.g. amino acids (AA)) and specialized (e.g. fatty acids, and flavonoids) metabolites reflect the plant's functional and physiological states of the cell and organism, respectively (Wu *et al.*, 2020).

Thus, the changes in plant metabolic profile could be investigated through two independent workflows: (1) the revalent metabolites identification, and (2) the statistical identification of the metabolomics indicators. Recently, the awareness of untargeted metabolomics analysis has increased due to its capabilities in the assessment of xenobiotics exposure/specific biomarkers and the risk of contaminants to living organisms. For example, the metabolic fingerprints of *Plantago lanceolata* showed various chemical changes as a response to different stresses (Riach *et al.*, 2015). Metabolic fingerprinting experiments aim to determine relative differences between two or more systems elucidating a biological relationship. Therefore, statistical strategies are typically used in a chemometrics style. Univariate and multivariate statistics can be used as standard approaches to extract relevant information from complex datasets (Worley and Powers, 2013; Gromski *et al.*, 2014; Lamichhane *et al.*, 2018; Prinsloo and Vervoort, 2018).

One environmental problem is water pollution, especially with slowly transformed and/or non/transformed pollutants. A good example is a diclofenac (DCF), which is a widely distributed non-steroidal anti-inflammatory drug that can be found in surface waters and is considered an environmental risk factor (Huber *et al.*, 2012). In 2019, Alkimin *et al.* reported that DCF incubation in *Lemna minor* caused a decrease in the content of photosynthetic pigments, relative fluorescence decay values of chlorophyll, and oxidoreductase and dehydrogenase activities. However, it led to increases in non-photochemical quenching, amounts of reactive nitrogen and oxygen species in roots, lipid peroxidation, oxidized ascorbate and thiols, and glutathione-reductase activity (Alkimin *et al.*, 2019). Further compound, carbamazepine (CBZ) is one of the most commonly detected pharmaceuticals in municipal sewage-treatment plant effluents, which is an antiepileptic drug and a relatively lipophilic compound (K_{ow} = 2.2), thus, CBZ is not degraded in freshwater environments in concentrations vary between 0.7 and 6.3 μg/L (Vernouillet *et al.*, 2010). It was reported that glutathione S-transferase activity was increased in *Lemna minor* incubated with psychoactive drug mixture (valproic acid, citalopram, carbamazepine, cyamemazine, hydroxyzine,

oxazepam, norfluoxetine, lorazepam, fluoxetine, and sertraline) (Bourioug *et al.*, 2018). Furthermore, the incubation of *Lemna minor* with carbamazepine, oxcarbazepine, and acridine 9-carboxylic acid and their mixture altered the nitrogen balance and the chlorophyll indices. Besides, the phenolic compound index depended on pharmaceuticals and time of exposure with no specific trend (Desbiolles *et al.*, 2020).

One of the most common plants in wetlands is *P. australis*, known as "common reed." *P. australis* belongs to the Poaceae family and is an invasive plant that spreads worldwide (Park and Blossey, 2008). It has been used for a long time in wetlands to remove pollutants, reduce nitrogen loads and provide oxygen to the rhizosphere (Gray and Biddlestone, 1995; Villette *et al.*, 2019). Furthermore, it is used as a resource for traditional crafts and fodder. In some regions like Northern China, it is grown as a crop and its leaves are used in the treatment of bronchitis and cholera (Park and Blossey, 2008).

These findings induce concerns regarding the chronic exposure of plants in constructed wetlands, which might lead to an alteration in its physiological characteristics, especially the molecular interaction and biochemical properties, during exposure to the recalcitrant DCF and CBZ. Therefore, this chapter discusses the different analytical workflow to evaluate the changes in the metabolic profile of *Lemna minor* and *P. australis* when incubated with two concentrations of DCF and CBZ 10 & 100 and 10 & 50 µM, respectively. The study was conducted by an untargeted metabolomics analysis strategy to characterize the metabolic profile of *Lemna* and reveal the changes due to DCF and CBZ incubation, individually. Also, a statistical workflow was used to discriminate against the changes in *Lemna minor* and *P. australis* metabolites and detect the differentiating metabolic profiles.

6.2. Experimental design

Lemna minor and *P. australis* were incubated individually with DCF and CBZ 10 & 100 and 10 & 50 µM, respectively. *Lemna minor* was analyzed with serial coupling RPLC-HILIC and connected to a single TOF-MS (system A). The analytical workflow depends on the type of data. For further instrumentation and methodological details, see chapter 3.

6.3. Results and Discussion

A-Lemna minor

6.3.1 RPLC-HILIC-ESI-TOF-MS analysis

Lemna minor samples were extracted with 100% MeOH, 50% MeOH, and 100% H_2O. The solvents guaranteed the extraction of a wide range of polarity of *Lemna* metabolites. The large difference in the polarity ensured the extraction of a wide range of metabolites. Further, the LC-system consisted of two-column in serial coupling which allowed separation of a wide range of different metabolites from (very-polar to non-polar) (Wahman *et al.*, 2019; Wahman *et al.*, 2020) in a positive ionization mode, which is already sufficient for an untargeted screening strategy to obtain a global overview of the differences and similarities between samples (De Vos *et al.*, 2012). Consequently, the data will fit the untargeted analysis

metabolomics concept using the features from the three extracts. The %RSDs of the RT of (quality control) standards were within the range of 0.06-0.31%, which represented the RT shift during the experiment. Further, the mass deviation was less than 5 ppm. The results are summarized in Tables 4 & 5. Then, the RT and mass windows were assigned to perform the untargeted analysis of *Lemna minor* different treatment. *Lemna minor* fingerprint showed low %RSD values (2%) indicating the robustness and the reproducibility of the method.

For the comprehensible investigation of different treatment features, a scatter plot between control and incubated samples (10 & 100 µM DCF) was drawn according to the average intensity of different features, respectively (Fig. 17). The blue small squares represent the features that were separated with the HILIC column (i.e., have RT < 15 min, and LogD < 0), while the red large squares represent features, which were separated with the RPLC column (i.e., have RT > 15 min, and LogD ≥ 0). Besides, the scatter plot shows a change in *Lemna minor* metabolites intensities and metabolomics due to 10 & 100 µM DCF. For example, in Figure 17 parts I & V, the compounds disappeared from the control samples and new ones appeared in the treated samples; detailed information can be seen in Table 15. On one hand, the compounds have a high intensity in the incubated samples with DCF (Fig. 17; part II), where located in the upper part of the scatter plot. On the other hand, compounds in the lower part of the plot (Fig. 17; part IV) have a low intensity in the incubated samples. Further, the compounds that have the same intensity in both samples are located in the middle line (Fig. 17; part III). The scatter plot shows an alteration in the metabolic profile as expected; hence, some compounds disappeared and others appeared after DCF incubation; these results show that plants were capable of dealing with DCF, even at higher concentrations. Also, it is compatible with the reported results of *Lemna minor* exposed to DCF (Alkimin *et al.*, 2019).

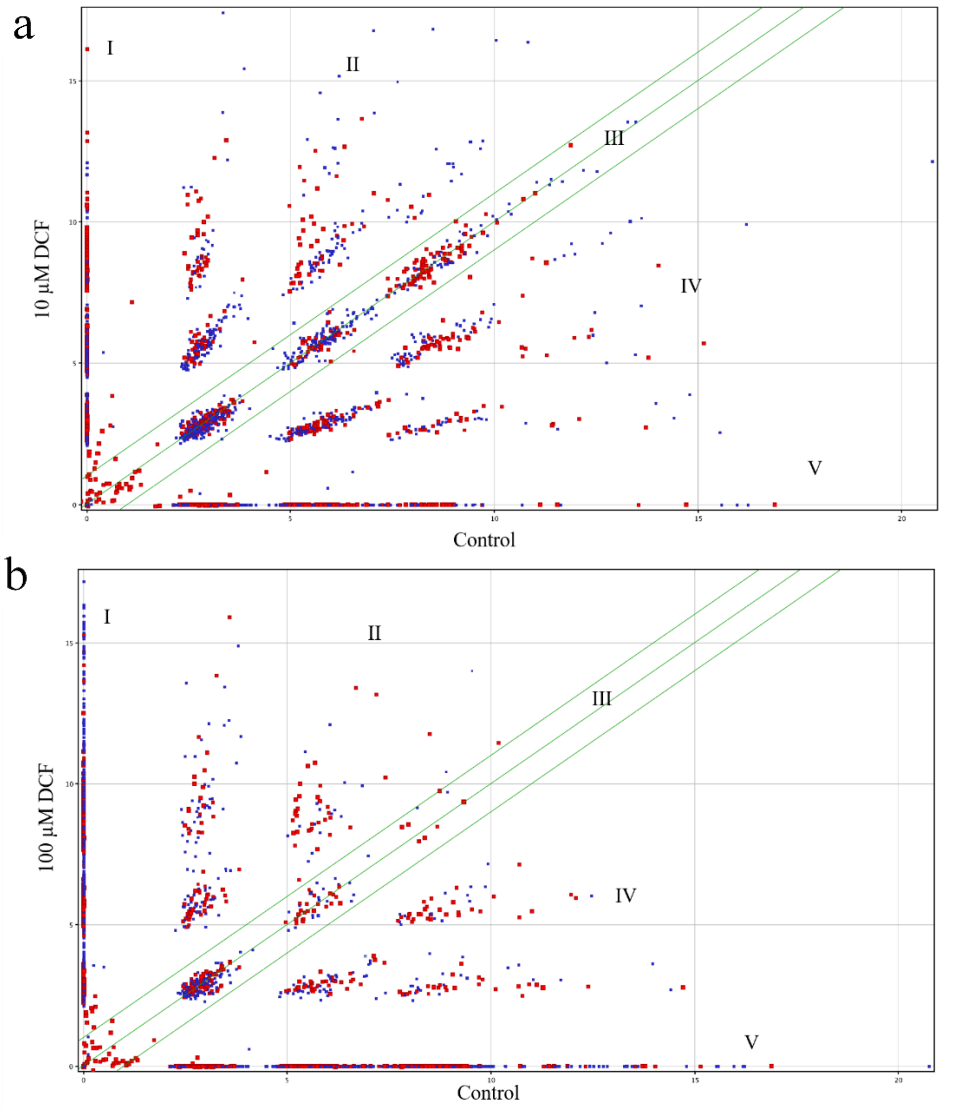


Figure 17. (a) Feature (signal intensity) comparison plot of *Lemna minor* control versus 10 µM DCF metabolic profile; (b) *Lemna minor* control versus 100 µM DCF metabolic profile. Blue, small squares represent RT (0-15min), and red, big squares represent RT (0-15min), and red, big squares represent the RT (16-33min).

I. Compounds were exclusively found on y-axes.
II. Compounds were found in both samples with increased feature intensity on the y-axis compared to the x-axis.
III. Compounds were found in both samples.
IV. Compounds were found in both samples with decreased feature intensity on the y-axis compared to the x-axis.
V. Compounds were found exclusively on x-axes.

Table 15. Compounds that are common in *Lemna* incubated with 10 &100 µM diclofenac 100% MeOH, 50% MeOH and 100% H_2O extracts. Also, the number of compounds has a relative standard deviation (%RSD) less than 2. The blue color represents the compounds, which separated with the HILIC column (i.e. have RT>15min.). The red color represents the compounds, which separated with the RPLC column (i.e. have RT<15min.).

Lemna minor extract	100 µM DCF, 100% MeOH	100 µM DCF, 50% MeOH	100 µM DCF, 100% H_2O	10 µM DCF, 100% MeOH	10 µM DCF, 50% MeOH	10 µM DCF, 100% H_2O
100 µM DCF, 100% MeOH	279 (205+74) 50 compounds with (% RSD) less than 2.	16 (12+4)	7	12 (5+7)	6 (2+4)	----------
100 µM DCF, 50% MeOH	16 (12+4)	484 (336+148) 52 compounds with RSD (%) less than 2.	92 (67+25)	1	31(13+18)	9 (7+2)
100 µM DCF, 100% H_2O	7	92 (67+25)	402 (346+56) 50 compounds with (% RSD) less than 2.	6 (3+3)	7 (5+2)	6 (5+1)
10 µM DCF, 100% MeOH	12 (5+7)	1	----------	252 (145+107) 100 compounds with (% RSD) less than 2.	9 (5+4)	8 (4+4)
10 µM DCF, 50% MeOH	6 (2+4)	31(13+18)	7 (5+2)	9 (5+4)	265 (162+103) 78 compounds with (% RSD) less than 2.	7 (3+4)
10 µM DCF, 100% H_2O	6 (3+3)	9 (7+2)	6 (5+1)	8 (4+4)	7 (3+4)	367 (278+89) 100 compounds with (% RSD) less than 2.

6.3.2. Total reducing contents

Considering the extracts independently, we observed a clear increasing pattern for the 100% MeOH and 50% MeOH extracts, as control < 10 µM DCF< 100 µM DCF while for the 100% H_2O extracts, *Lemna* control, and the one incubated with 10 µM samples presented approximately the same gallic acid equivalent content (ca. 730 µg GAE/g DW). Furthermore, the reducing contents increased significantly (about 1.8 fold) in samples incubated with 100 µM (Fig. 18). *Lemna minor* increased the production of stress-defensive compounds such as fatty acids, saturated carboxylic acids, phenolics, and flavonoids (Vladimirova and Georgiyants, 2014), as expected (H3), which was already mentioned in other studies using *Lemna minor*. For example, in 2012, Forni and coworkers revealed that *Lemna minor* has higher production of phenols at the first three days of treatment with sodium dodecyl sulfate (an anionic surfactant) with a higher concentration (up to 50 mg/L) than when exposed for seven days to a lower dosage (25 mg/L), being the phenol content even lower than the control for the last one (Forni *et al.*, 2012). Further, in 2013, Varga and coworkers showed that the total phenolic of *Lemna minor* changed significantly when exposed to Hg in comparison to Cd or Cr in a concentration ranging from 0.02 to 20 mg/L for 24 hours (Varga *et al.*, 2013). Furthermore, in 2020, Kostopoulou, *et al.*, reported that *Lemna minor* increases the production of aromatic amino acids (AAAs) after incubation with drugs such as glyphosate, metribuzin, and their mixture (Kostopoulou *et al.*, 2020).

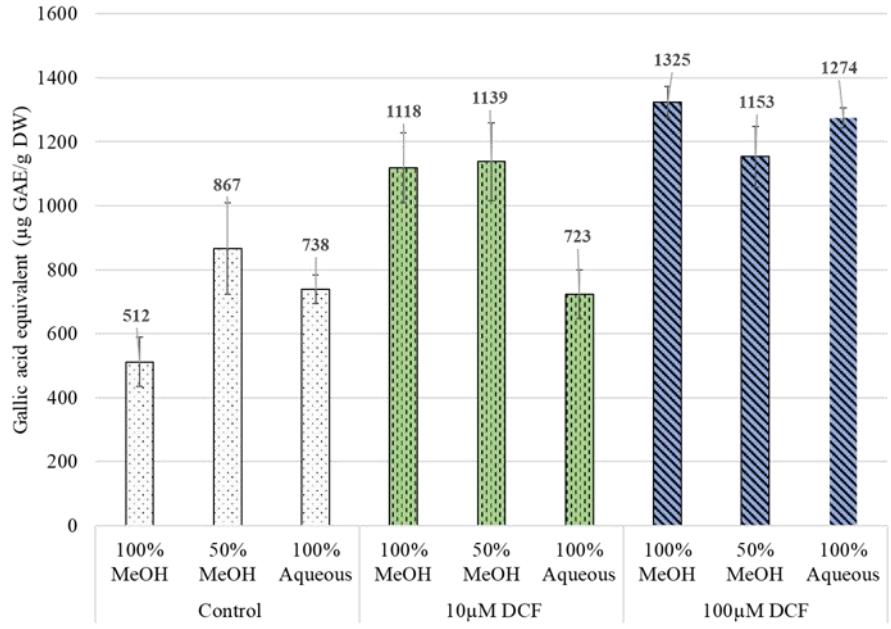


Figure 18. Reducing contents of 100% MeOH, 50% MeOH, 100% H_2O extracts in control (white) and incubated groups (10 & 100 µM DCF, green and blue, respectively), which is expressed as Gallic acid equivalent (µg GAE/ g of dry weight (DW)); (n=9) ± SD.

6.3.3. Target and suspect screening analysis of Lemna profile

The changes in *Lemna minor* metabolites profile were investigated with target and suspect screening to follow the changes in the intensities after exposure to 10 & 100 µM DCF. Several amino acids (serine, tyrosine, aspartic acid, glutamic acid, proline, histidine, phenylalanine, tryptophan, valine, alanine, isoleucine, leucine, and proline) were identified in100% MeOH extracts in duckweed after treatments. The intensity of each amino acid was compared in the different samples incubated with 10 & 100 µM DCF against control (Table 16 and Figs. 19&20). Serine decreased in *Lemna minor* incubated with 10 µM DCF and kept decreasing until it dropped below the LOD in the 100 µM incubated sample. Further, aspartic acid, glutamic acid, histidine, valine, and alanine increased in *Lemna minor* incubated with 10 µM DCF and decreased in *Lemna minor* incubated with 100 µM DCF. In plants, serine biosynthesis proceeds by different pathways. As known, plants use the phosphorylated pathway in response to an infection, and to environmental and abiotic stresses (Igamberdiev and Kleczkowski, 2018). In this pathway, plants produce glutamate, which synthesizes in its turn, proline. In plants, intracellular proline levels have been found to increase by more than 100 fold during stress, as we observed in 100 µM DCF treatment. Rhodes *et al.*, (1986), suggested that the proline increased after the exposure of *Lemna minor* to methionine sulfoximine. Later, it has been proven that the concentration of proline increased because the glutamate pool served as its precursor, which explains the increase of proline in the two groups (Delauney and Verma, 1993). Also, the glutamic acid concentration decreased in *Lemna minor* incubated with metribuzin, glyphosate, and their mixtures for 72 hours (Kostopoulou *et al.*, 2020). This may explain the decrease in the concentration of glutamic acid in our data for 100 µM DCF. For, phenylalanine, leucine, and isoleucine intensities increased in both incubated samples. They were directly proportional to DCF concentration. However, tryptophan decreased in *Lemna minor* incubated with 10 µM DCF and increased in the 100 µM incubated sample. However, tyrosine concentration increased in *Lemna minor* incubated with 10 µM DCF, while it decreased under the measurable limit in 100 µM DCF incubation. For organic acids, *p*-coumaric and sinapic acids were identified in all the samples extracted in 100% MeOH. In 10 µM DCF exposure, these compounds presented opposite behaviors; i.e., *p*-coumaric acid decreased while the sinapic acid increased. For 100 µM DCF treatment, both were under the measurement limit. For cinnamic acid, this one was increased and decreased in the10 µM DCF and 100 µM DCF exposure, respectively as shown in (Table 16) and displayed (Fig. 19b). This seems to be the explanation of the reducing contents in incubated samples with DCF. Tryptophan, tyrosine, and phenylalanine are aromatic amino acids (AAAs), which are required for protein biosynthesis in all living cells. In plants, AAAs serve as precursors of a wide variety of plant natural products that play crucial roles in plant growth, development, reproduction, defense, and environmental responses such as alkaloids, phytoalexins, and indole glucosinolates as well as the numerous phenolic compounds (Maeda and Dudareva, 2012). The increase in phenylalanine was observed with the DCF incubation which was accompanied by an increase of cinnamic acid, supporting once again the increase in the reducing contents.

Cinnamic acid is the first compound in the phenylpropanoid pathway that begins with the deamination of phenylalanine (Maeda and Dudareva, 2012). Consequently, *Lemna minor* enhanced the phenylpropanoid pathway, which protects it from oxidative stress against DCF. However, at 100 µM, their concentrations decreased. The flavonoids content also decreased, when *Lemna gibba* was exposed to several environmental challenges (stressor); according to Akhtar *et al.*, 2010, this might be a result of the promotion of the photosynthetic electron transport chain reduction, causing flavonoid reduction. Moreover, *p*-coumaric and sinapic acid decreased for the 100 µM treatment, which can be related to the alternative usage of NADPH to degrade DCF instead of producing these organic acids (Huber *et al.*, 2012). The same profile was also reported by Kostopoulou *et al.*, (2020). Aspartic acid increased in 10 µM DCF, which might be due to increasing isoleucine and valine synthesis. They increased despite their common precursor (aspartic acid) decreased. Rhodes *et al.* (1986) expected an increase of isoleucine due to protein degradation, in *Lemna* exposed to methionine sulfoximine, for 24 hours. However, a recent study indicated that the increase is due to aspartic acid catabolism, such as in the study where *Arabidopsis* was subjected to bacterial infection (Yang and Ludewig, 2014). However, for the 100 µM DCF, the opposite was observed for aspartic acid. This might be a consequence of the 100 µM DCF where aspartic acid is used to provide energy to the tricarboxylic acid cycle (TCA), as a response to abiotic stress (Galili, 2011).

Alanine increased significantly in *Lemna minor* incubated with 10 µM DCF when compared to control and 100 µM DCF. The same was observed in plant response to hypoxic stress, which was accompanied by higher rates of glycolysis and ethanol fermentation causing the fast depletion of sugar stores and carbon stress (Limami *et al.*, 2008).

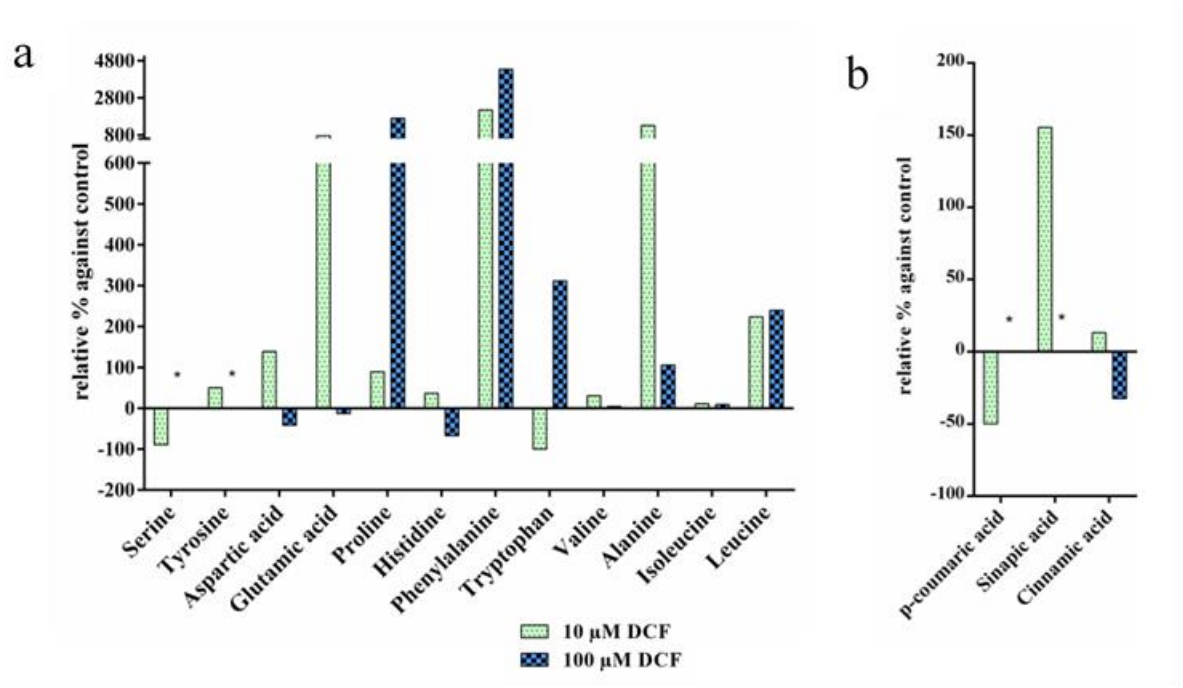


Figure 19. Chart of the intensities of different metabolites identified in the 10 & 100 µM DCF treatments against the control; (a) amino acids (AA); (b) organic acids; the asterisk symbol (*) means intensities below the LOD. (note: the broken y-axis is used to demonstrate all the values)

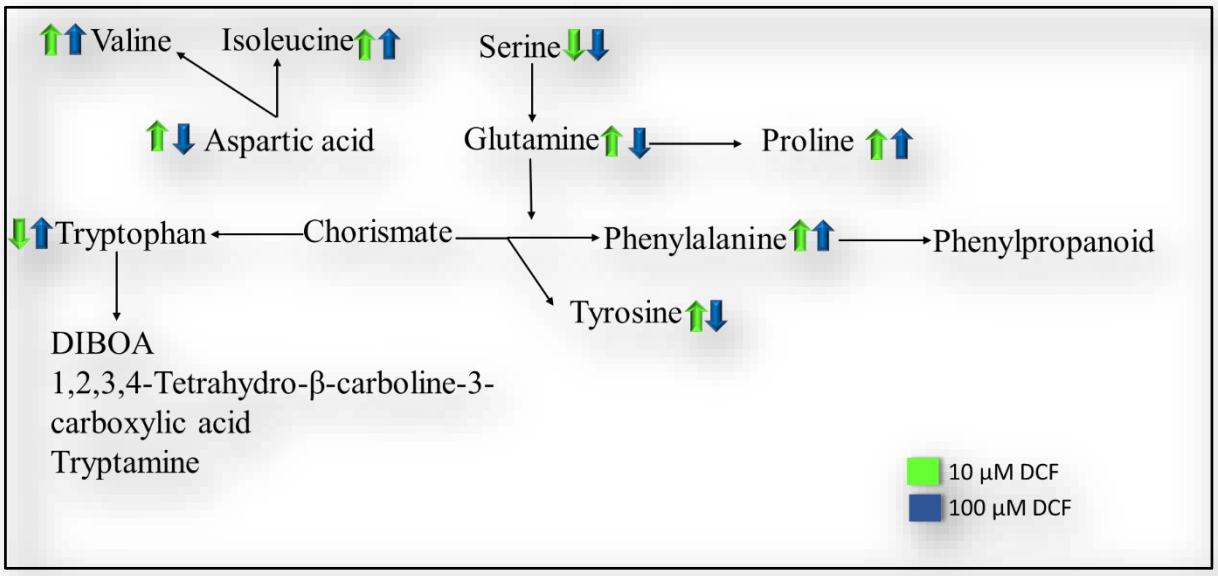


Figure 20. The amino acids biosynthesis pathways in *Lemna minor*; the green and blue colors represent the *Lemna* incubated with 10 µM DCF and 100 µM DCF, respectively.

Table 16. List of amino acids (AA) and organic acids detected in the different *Lemna minor* treatments (control, 10 & 100 µM DCF); details such as name, the monoisotopic mean mass and the corresponding mean RT, the absolute deviation in mass (Δppm), and RT (ΔRT), the intensities (inten.) of the standards and the treated samples.

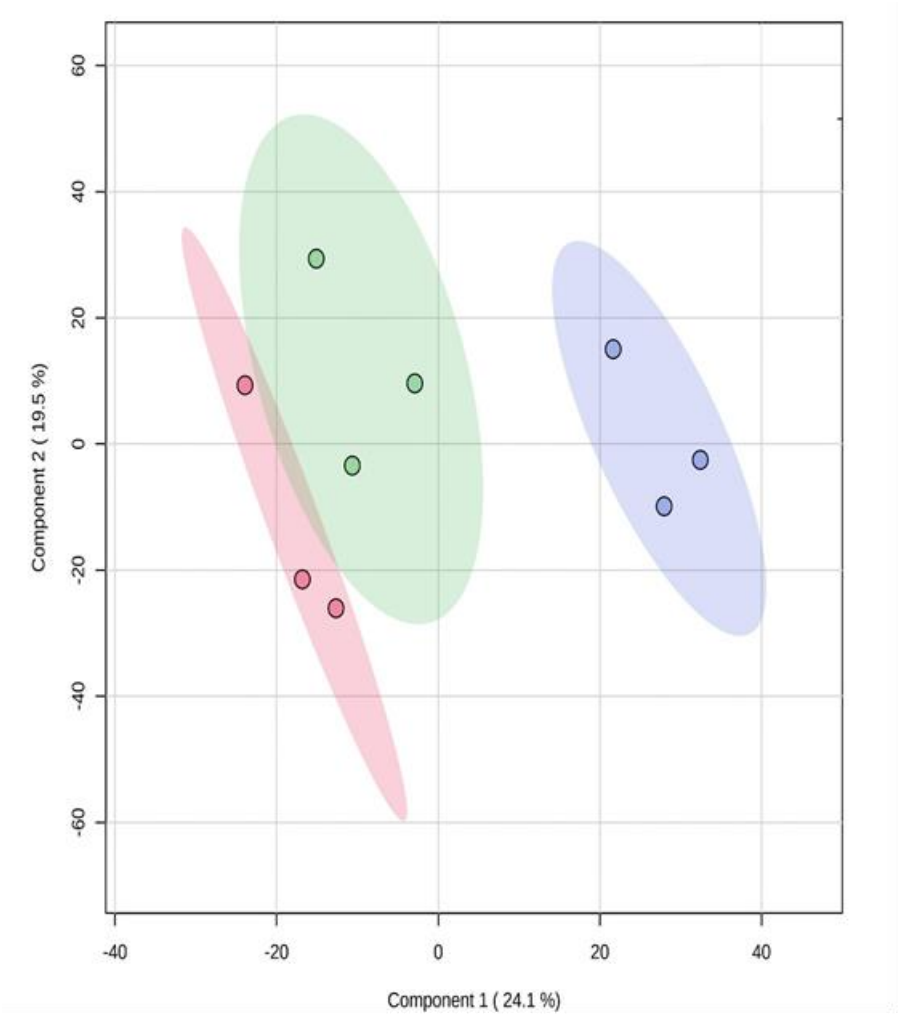
Name	**Standard**		**Control**					**10 µM DCF**					**100 µM DCF**				
	Mean Mono isotopic Mass (Da)	Mean RT (min)	Mean Mono isotopic Mass (Da)	Δppm	Mean RT (min)	ΔRT	Inten.	Mean Mono isotopic Mass (Da)	Δppm	Mean RT (min)	ΔRT	Inten.	Mean Mono isotopic Mass (Da)	Δppm	Mean RT (min)	ΔRT	Inten.
Serine	105.0428	13.44	105.0433	4.13	13.45	0.02	13.74	105.0433	4.76	13.49	0.05	1.45	—	—	—	—	—
Tyrosine	181.0730	11.96	181.0726	2.49	11.94	0.02	5.67	181.0725	1.75	11.96	0.001	8.54	—	—	—	—	—
Aspartic acid	133.0375	12.71	133.0369	4.76	12.88	0.17	5.41	133.0375	0	12.83	0.12	12.93	133.0376	0.5	12.93	0.04	3.16
Glutamic acid	147.0529	12.63	147.0522	2.04	12.75	0.04	3.00	147.0520	1.81	12.73	0.02	25.90	147.0529	0	12.77	0.04	2.60
Proline	115.0636	12.17	115.0633	2.35	12.03	0.14	1.76	115.0632	3.48	12.40	-0.24	3.33	115.0634	-0.32	12.36	-0.19	31.75
Histidine	155.0687	15.35	155.0696	6.02	14.65	0.7	9.42	155.0692	3.44	15.05	0.3	12.83	155.0704	3.44	15.23	0.11	3.11
Phenylalanine	165.0787	10.71	165.0786	2.30	10.70	0.01	0.16	165.0783	4.32	10.70	0.01	3.68	165.0799	5.77	10.67	0.04	7.27
Tryptophan	204.0899	11.45	204.0900	0.82	11.43	0.03	0.84	204.0898	0.33	11.43	0.02	0	204.0908	4.25	11.46	0.01	3.46
Valine	117.079	11.64	117.0790	0.28	11.60	0.03	3.52	117.0791	0.85	11.62	0.02	4.60	117.0796	5.12	11.71	0.07	3.68
Alanine	89.0477	12.84	89.0478	1.12	13.11	0.27	0.89	89.0475	2.25	12.99	0.15	12.46	89.0479	1.87	12.89	0.06	1.823
Isoleucine	131.0944	10.82	131.0949	4.07	10.87	0.04	12	131.0947	2.8	10.83	0.01	13.33	131.0947	2.8	10.67	0.16	13.12
Leucine	131.0941	10.79	131.0943	1.78	10.98	0.19	3.07	131.0947	4.83	10.68	0.11	9.94	131.0947	4.83	10.67	0.12	10.43
Cinnamic acid	148.0519	10.61	148.0521	1.35	10.6	0.01	4.55	148.0520	0.90	10.52	0.09	5.15	148.0523	2.70	10.50	0.11	3.08
p-coumaric acid	164.0468	7.72	164.0479	6.50	7.75	0.03	2.00	164.0472	2.64	7.80	0.05	1.00	—	—	—	—	—
Sinapic acid	224.0683	11.95	224.0677	2.83	12.14	0.19	2.69	224.0679	1.79	12.17	0.22	6.87	—	—	—	—	—

The symbol (—) means below the LOD

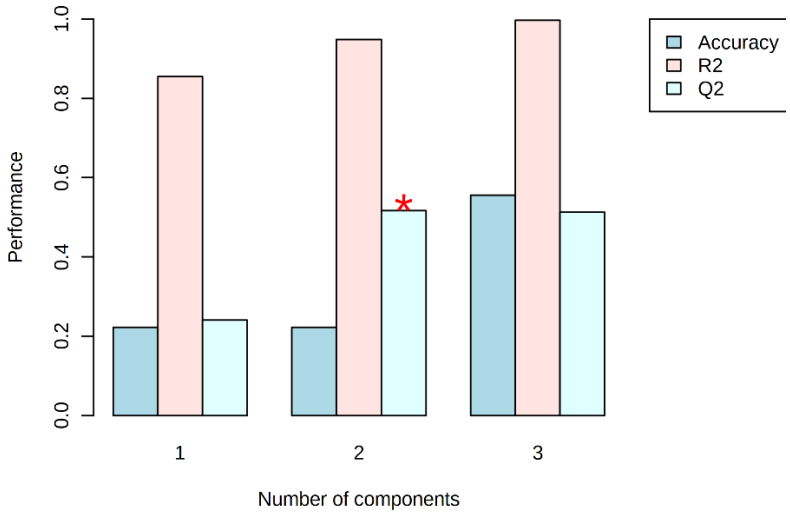
6.3.4. Alterations in Lemna minor metabolic profile

The present study was based on the LC-MS metabolomics analysis to investigate changes in the metabolic profile of *Lemna* under the influence of 10 & 100 µM DCF. After assigning an accurate mass, 6989 metabolites were selected and used to differentiate between control and the 10 & 100 µM DCF samples, which are represented in separated clouds in the PLS-DA plot (Fig. 21).

The first and second significant PLS components explained 24.1% and 19.5% of the total variance, respectively. The clear separation between treatments was observed as being more evident for the 100 µM DCF (violet cloud) treatment. Within treatments, variation among extracts was observed due to their different types of solvent (100%MeOH, 50%MeOH, and 100% H_2O) and due to biological variability.



a



b

Figure 21. (a) PLS-DA scores plot of *Lemna* control and 10 & 100 µM DCF treatments between the first and second components. The explained variances are shown in brackets (95% confidence level); red represents control, green represents 10 µM DCF, and violet represents 100 µM DCF treatment samples (n=6); (b) The plot of PLS-DA components and their performances. The red star indicates the best classifier. R^2 parameter is known as the "goodness of fit", and the Q^2 parameter is termed "goodness of prediction" or "cross-validation".

The contribution of metabolites to the discrimination between the *Lemna* control and incubated fronds with DCF metabolic profiles is based on the coefficients of the PLS-DA, which were given in (Table 17). One hundred and thirteen compounds were selected, through the coefficients of PLS-DA, as important metabolites in the discrimination between the different treatments. 23% of them were known and the rest were unknown. They are organic acids, lignin, sugars, AA, dipeptide, flavonoids, bioflavonoids, fatty acids, and miscellaneous (Table 17).

The concentrations varied among treatments, which were calculated according to the Metaboanalyst software:

- Gallic acid, nicotinic acid, 1-monopalmitoylglycerol, oleic acid (fatty acids), and 1-β-Ribofuranosyluracil decreased due to incubation with 100 µM DCF;
- Glutamic acid, hexose, ribose and hydroxymethyl furfural (sugars), 5-methoxyflavone, 3,5,7,4'-tetramethoxyflavone (flavonoids), 3,7-dimethyl-2,6-octadienoic acid, and 1-pentadecane carboxylic acid (fatty acids), 5-methoxy-N,N'-dimethyltryptamine, and 2,4-dihydroxy-1,4-benzoxazine-one (DIBOA) increased in the 10 µM DCF treatment;

- Leucine, syringaresinol, 7-hydroxy-4-methoxy-5-methylcoumarin, 3-heptadecyl-5-methoxyphenol, kaurenoic acid, , 9,12-octadecadienoic acid, Taraxer-14-en-3-one, 3,9-dimethyl-undecane, and 6,10-dimethyl- 2-oxo-5,9-undecadiene increased in 100 µM DCF mainly to control.

Table 17. The relevant coefficients of compounds produced according to PLS indicating the effect of DCF incubation; the concentration (intensities) of each compound was represented as a color, which is indicated in the bar

Low  High

Compound Name	Coefficient			
	Mean	Control	10 µM DCF	100µM DCF
Organic acids				
Gallic acid	11.266	4.87	4.99	23.94
Nicotinic acid	21.031	7.80	8.18	47.12
Truxillic acid	6.8703	9.28	7.90	3.43
Sugars				
Hexose	13.696	6.78	7.70	26.61
Ribose	6.8234	9.72	7.30	3.45
Hydroxymethyl furfural	24.944	15.26	15.86	43.71
Amino acids				
Glutamic acid	27.508	21.81	21.74	38.97
Leucine	33.212	46.05	35.37	18.22
Fatty acids				
1-Monopalmitoylglycerol	12.006	3.99	7.0	25.022
1-Pentadecanecarboxylic acid	48.224	69.40	56.47	18.8
Taraxer-14-en-3-one	13.808	20.6	13.65	7.18
3-Heptadecyl-5-methoxyphenol	13.133	5.63	5.60	28.17
3,9-Dimethyl-Undecane	7.917	3.64	4.84	15.27
9,12-Octadecadienoic acid	17.671	30.12	21.55	1.346
3,7-Dimethyl-2,6-Octadienoic acid	15.091	7.73	8.27	29.28
6,10-Dimethyl- 2-oxo-5,9-undecadiene	13.774	15.24	11.1	14.98
Oleic acid	20.64	20.17	17.86	23.89
Flavanoids				
Isorhamnetin-3-O-galactoside	14.695	20.72	15.98	7.38
3,5,7,4'-Tetramethoxyflavone	12.73	12.92	11.44	13.84
5-Methoxyflavone	13.953	18.17	16.56	7.13
Miscellaneous				
1-β-Ribofuranosyluracil	29.586	14.46	15.63	58.67
Syringaresinol	29.409	36.90	27.75	23.58
7-Hydroxy-4-methoxy-5-methylcoumarin	20.555	9.34	10.98	41.35
5-Methoxy-N,N'-dimethyltryptamine	30.112	35.28	26.51	28.55
2,4-Dihydroxy-1,4-benzoxanzin-34H-one (DIBOA)	6.3403	8.57	7.294	3.16
Kaurenoic acid	51.172	75.07	52.08	26.37

The targeted analysis showed an increase in phenylalanine, as well as cinnamic acid which enhanced the phenylpropanoid pathway, which protects *Lemna* from oxidative stress against DCF, increasing the patterns intensities of 3,5,7,4'-tetramethoxy flavone, and 5-methoxy flavone. However, at 100 µM, their concentrations decreased and the 7-hydroxy-4-methoxy-5-methyl coumarin concentration is increased. Thus, it is possible that *Lemna* could generally enhance the phenylpropanoid pathway as a defense mechanism. Phenylpropanoids consist of a huge number of metabolites with different structures and biosynthesis pathways. Thus, it is difficult to conclude which compounds were inhibited or enhanced without reference standards. Besides, the flavonoids content also decreased, when *Lemna gibba* was exposed to several environmental challenges (stressor); according to Akhtar *et al.*, 2010, this might be a result of the promotion of the photosynthetic electron transport chain reduction, causing flavonoid reduction. Further, in 10 µM DCF, gallic, nicotinic, have a higher concentration than 100 µM DCF.

Gallic acid plays also an important role in plant defense against stress conditions, which can synthesize in the plant from phenylalanine → cinnamic acid → protocatechuic acid (β-oxidative pathway) and directly from 5-dehydroshikimic acid. This finding is illustrated by the increase in phenylalanine and the concomitant decrease in *p*-coumaric, sinapic, and truxillic acids in DCF incubated samples. It seems that *Lemna* preferred the formation of phenylalanine → protocatechuic acid → gallic acid through the β-oxidative pathway when incubated with 10 µM DCF. This might be due to the inhibition in the 5-dehydroshikimic acid pathway. This pathway was observed in the mutant strain of *Neurospora crassa* blocked in the conversion of 5-dehydroshikimic acid into shikimic acid (Dewick and Haslam, 1969). The syringaresinol that is biosynthesized from sinapic alcohol by peroxidase enzymes (Habib *et al.*, 2018), increased in DCF treatment; this is an expected result of peroxidase and oxidase enzymes activation in *Lemna* by DCF (Alkimin *et al.*, 2019).

Upon treatment with 10 µM DCF, hexoses such as (glucose, galactose, mannose, and raffinose) and ribose potential increased. This might be due to higher energy demand to tolerate the stress effect of DCF incubation through exhaustion of sugar stores. Also, it has been reported that non-soluble sugars have osmoprotectant and antioxidant activities (Sivaram *et al.*, 2019). However, in 100 µM DCF, their concentrations decreased which might be due to the fluctuations of sugars mechanisms, which are affected by changes in the genotype (e.g. INV, SuSy, ATB2 bZIP, and α-amylase gene) and stress factors (Rosa *et al.*, 2009). The fast depletion of sugar stores and carbon stress was observed in plant response to hypoxic stress, which was accompanied by higher rates of glycolysis and ethanol fermentation causing (Limami *et al.*, 2008). Moreover, in 10 µM DCF, the decrease in the tryptophan concentration is accompanied by an increase in DIBOA, which is synthesized by tryptophan synthase α that catalyzes its formation from indole (as a precursor) instead of tryptophan (Frey *et al.*, 1997). Interestingly, the opposite behavior was observed for the 100 µM DCF, which might be due to overexpression of tryptophan synthase β gene like in plants treated with DCF (Sanjaya *et al.*, 2008). Moreover, 5-methoxy-N, N'-dimethyltryptamine showed the same pattern as DIBOA

because they originate also from indole (tryptophan metabolism). Additionally, the 1-monopalmitoylglycerol pattern decreased with 10 µM DCF incubation compared to control, however, was up-modulated compared to treatment with 100 µM DCF. It is one of the phosphatidic acid-binding proteins, which may be a predominant signaling component enhanced in multiple biotic and abiotic stress response pathways (Lim *et al.*, 2017). Also, the 1-pentadecanecarboxylic acid and 3,7-dimethyl-2,6-octadecadienoic acid were increased, however, many unsaturated fatty acids were increased in 100 µM DCF. This might eventually protect the *Lemna* against the ROS as a result of fatty acid degradation under stress, which are produced from DCF and/or its transformation products under a high dose of DCF (Alkimin *et al.*, 2019). In addition, the oleic acid, undecanes, and 9,12-octadecadienoic acid concentrations also were increased. Since, saturated and unsaturated fatty acids induce broad-spectrum resistance against infections in the plant, such as *Pseudomonas syringae* in tomatoes. Also, unsaturated fatty acids are particularly important in plant defense (Lim *et al.*, 2017). Some of them (stearic, oleic, and palmitic acids) increased in *Lemna* incubated with glyphosate, and glyphosate metribuzin mixture for 72 hrs (Kostopoulou *et al.*, 2020); this might explain the increase in reducing potential, in 100 µM DCF, as a result of unsaturated fatty acid formation/accumulation. Also, kaurenoic acid (diterpenes) was increased in DCF higher concentration.

6.3.5. Impact of DCF on Lemna metabolic pathway

Metabolic pathway analysis was performed to identify the pathways that were induced upon the incubation of *Lemna minor* with DCF via MetaboAnalyst 5.0 software based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. For the investigation, the mummichog algorithm was the implementation of this concept to infer pathway activities from a ranked list of MS peaks identified by untargeted metabolomics. This was applied to test their participation in pathways using a combination of network analysis and functional enrichment analysis. The pathways exhibiting $p < 0.05$ were considered a statistically significant metabolic pathway, meaning that they were affected with DCF or its transformation products. KEGG pathway analysis using the pathway data set of *A. thaliana* matched 48 pathways in *Lemna* incubated with DCF. They were filtered according to the number of significant hits and 21 pathways were considered (Fig. 22 and Table 18). They were identified as glucosinolate biosynthesis, aminoacyl-tRNA biosynthesis, porphyrin and chlorophyll metabolism, tryptophan metabolism, sphingolipid metabolism, phenylalanine, tyrosine and tryptophan biosynthesis, cyanoamino acid metabolism, arachidonic acid metabolism, nicotinate and nicotinamide metabolism, phenylpropanoid biosynthesis, biosynthesis of unsaturated fatty acids, purine metabolism, pyrimidine metabolism, cysteine and methionine metabolism, valine, leucine and isoleucine degradation, valine, leucine and isoleucine biosynthesis, arginine and proline metabolism, pyruvate metabolism, monoterpenoid biosynthesis, flavonoids biosynthesis, and anthocyanin biosynthesis. The same data was investigated with MPP software to detect the altered pathways in *Lemna* incubated with DCF at two different concentrations. The results showed three

changed pathways, which were glucosinolate, porphyrin and chlorophyll metabolism, and flavonoids biosynthesis (Table 18). The data investigation with two software packages (Metaboanalyst and MPP), which have different algorithms, confirmed the results, as well as indicating the accuracy of the data.

The incubation of *Lemna* with DCF caused distinguished changes in its biological activity because 5 (glucosinolate biosynthesis, aminoacyl-tRNA biosynthesis, phenylalanine, tyrosine and tryptophan biosynthesis, phenylpropanoid biosynthesis, valine, leucine, and isoleucine biosynthesis) of the identified pathways were verified as key factors in response to energy demands (Chen *et al.*, 2019). The glucosinolates pathway is connected with other metabolic pathways which play fundamental roles in plant growth, development, and interaction with the environment because some of them are antioxidants. The glucosinolate pathway has a link with the phenylpropanoid pathway due to flavone 3'-*O*-methyltransferase 1, which is responsible for methylation of the 3'-hydroxy group in flavonoids, which exhibits also an activity for methylation of hydroxyl-indole glucosinolates (Chhajed *et al.*, 2020). Additionally, the enhancement of the flavonoid pathway and phenylpropanoid pathway is expected from the target analysis results (phenylalanine up-modulation), which produced the antioxidant compounds. Furthermore, unsaturated fatty acid biosynthesis and the related pyruvate metabolism were induced after incubation with DCF to protect the *Lemna* against the ROS especially at higher concentrations (Table 17). The decrease in oleic acid in 10, 100 µM DCF, however, the increase of 9,12-octadecadienoic acid (linoleic acid) at higher concentration confirmed the enhancement in arachidonic acid metabolism, which is the main constituent of *Lemna* fatty acid content (Al-Snafi, 2019). Arachidonic acid is also considered a potent elicitor of programmed cell death and defense responses, as well as induces resistance to viruses via producing plants oxylipins (Savchenko *et al.*, 2010). Also, pyrimidine metabolism could be proposed to be a result of glucosinolate biosynthesis pathway alteration, which up-regulates the biosynthesis of other stress-induced pathways (Tantikanjana *et al.*, 2004). Pyrimidine metabolism is enhanced and leads to products that could be used in the case of salvage i.e. recovery of infections and subsequent synthesis of secondary products with specific functions in defense mechanisms. It was reported that DCF decreased the content of photosynthetic pigments, relative fluorescence decay values of chlorophyll in *Lemna* (Alkimin *et al.*, 2019). Consequently, the porphyrin and chlorophyll metabolism was enhanced in *Lemna* incubated with DCF. The pathway was identified with the two software.

Finally, *Lemna minor* altered its metabolomics after incubation with DCF increased the antioxidant compounds (phenolic, unsaturated fatty acids). Those compounds could be used as a source for phytochemicals from the plants in a constructed wetland, which can be used in industrial concepts.

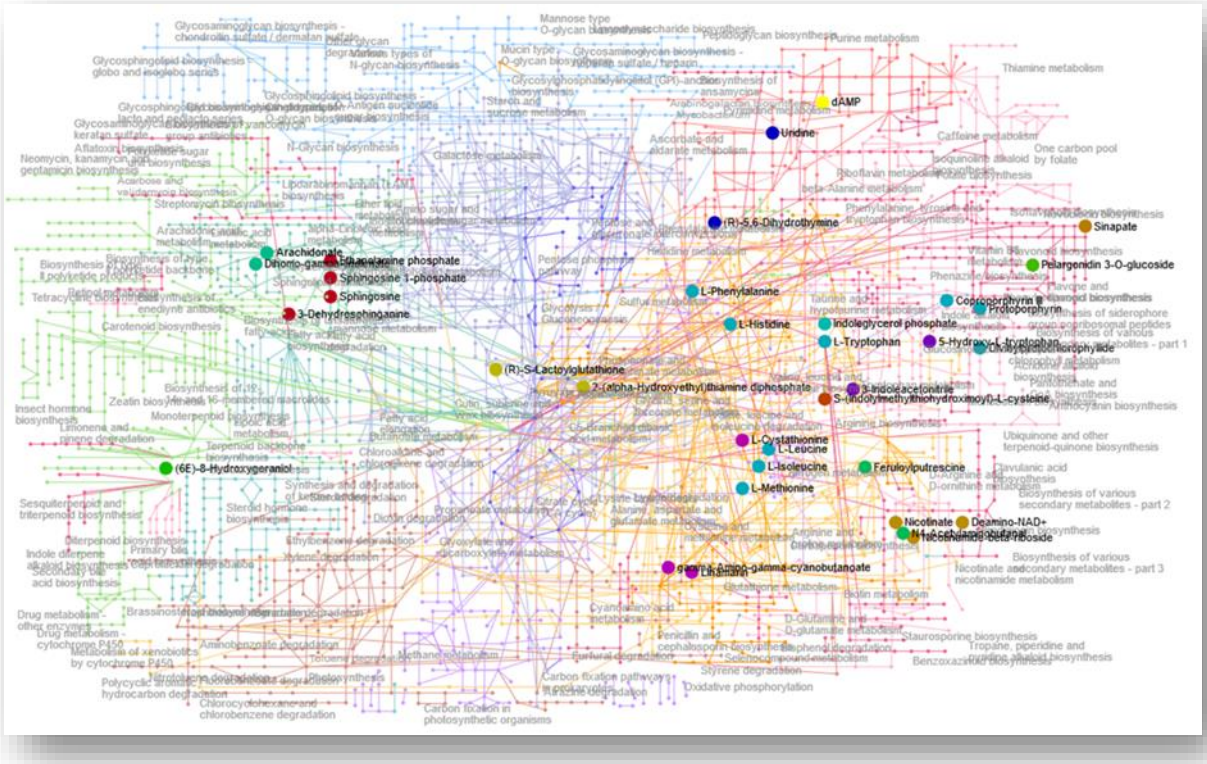


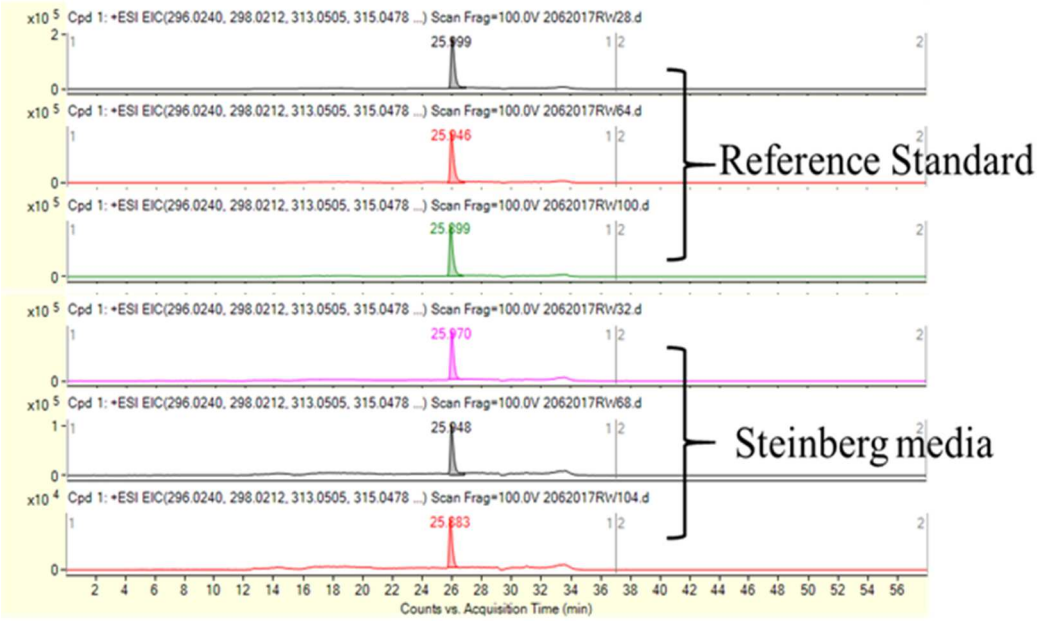
Figure 22. The network of the statistically significant metabolic pathways of *Lemna minor*, which was enhanced due to incubation with DCF and/or its transformation products. They were predicated with the Metaboanalyst software. Some of them were investigated with target analysis and literature comparison. The significant hits of each pathway were color-coded.

Table 18. The output of the mummichog analysis containing ranked pathways that were enriched in the *Lemna* incubated with DCF at two different concentrations. The table includes the total number of hits per pathway (all, significant, and expected), and the color-coding. (*) represents the pathway that was detected with MPP software.

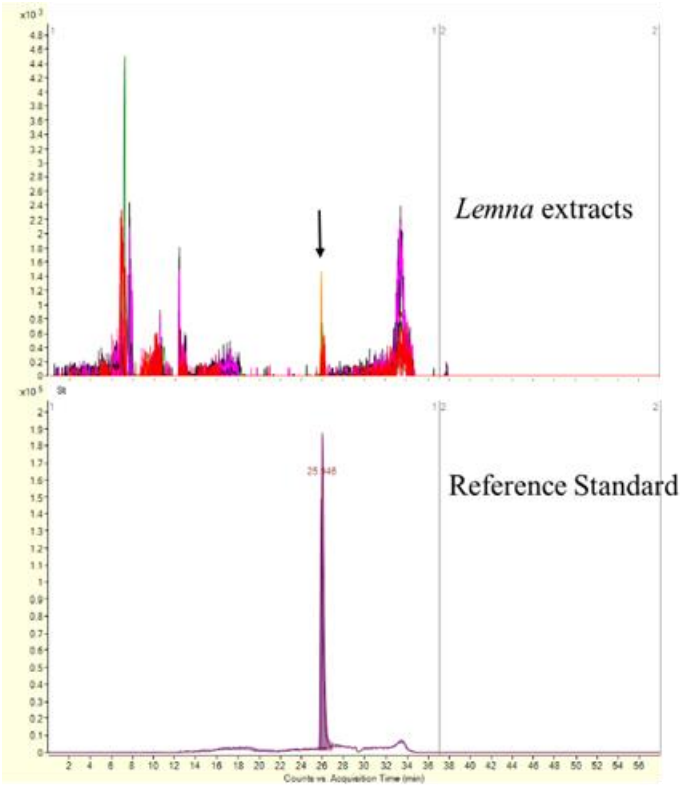
Pathway name	Total hits	Significant hits	Expected hits	Color coding
Glucosinolate biosynthesis*	6	6	3.1468	
Aminoacyl-tRNA biosynthesis	6	6	1.0651	
Porphyrin and chlorophyll metabolism*	5	4	2.0817	
Tryptophan metabolism	4	3	1.3556	
Sphingolipid metabolism	4	4	0.43571	
Phenylalanine, tyrosine, and tryptophan biosynthesis	3	3	1.0651	
Cyanoamino acid metabolism	3	3	1.404	
Arachidonic acid metabolism	3	2	0.53254	
Nicotinate and nicotinamide metabolism	3	3	0.62937	
Phenylpropanoid biosynthesis	3	3	2.0333	
Biosynthesis of unsaturated fatty acids	3	3	1.0651	
Purine metabolism	2	2	3.05	
Pyrimidine metabolism	2	2	1.8397	
Cysteine and methionine metabolism	2	2	2.227	
Valine, leucine and isoleucine degradation	2	2	1.5492	
Valine, leucine, and isoleucine biosynthesis	2	2	1.0651	
Arginine and proline metabolism	2	2	1.5976	
Pyruvate metabolism	2	2	0.91984	
Monoterpenoid biosynthesis	2	2	0.43571	
Flavonoids biosynthesis*	2	1	2.227	
Anthocyanin biosynthesis	2	2	0.53254	

6.3.6. Metabolism of DCF in Lemna minor

Lemna minor was exposed for DCF at two different concentrations (10 and 100 µM DCF) for 96 hrs., individually. The extracts and the Steinberg media were analyzed with RPLC-HILIC-ESI-TOF-MS to detect the DCF and its transformed products. Diclofenac was identified at 295.0161 Da with a 4 ppm deviation from the monoisotopic mass, which eluted at 26 minutes (logD (pH7) > 0). The EICs of diclofenac in extracts and standard as well as in the Steinberg media were shown in (Fig. 23). DCF structure is an aniline ring with 2 chlorine atoms and a phenylacetic group. The first transformed product (DM) of DCF can not be detected, which has a hydroxyl group at positions 3´, 4`, and 5´(aniline ring) (Huber *et al.*, 2012). However, *Lemna* transformed product (LTP) with hydroxyl derivative at position 3 in the phenyl group, as well as the loss of acetic moiety and oxidation of the phenyl group was detected (Fig. 24). The transformed product is 282.0107 Da, which was identified before using the same system (TP_13) by comparing the *m/z* and RT (Rajab *et al.*, 2013). Also, the hydroxylation of the phenyl group was detected in the human liver cell culture as minor Phase I metabolites, however, it is the major here in *Lemna* (Sarkar *et al.*, 2017). Moreover, the transformed products DM_2, DM_3, DM_7 were detected in the *Lemna minor* metabolic profile (Table 19). Thus, uptake and transformed DCF by *Lemna* and the changes, which were detected in *Lemna* metabolic profile is a result of exposure to DCF and/or its transformation products (Phase I) (Huber *et al.*, 2012).



a



b

Figure 23. The extracted ion chromatogram of diclofenac was identified in (a) Steinberg media; (b) *Lemna* extract after incubation for 96hrs

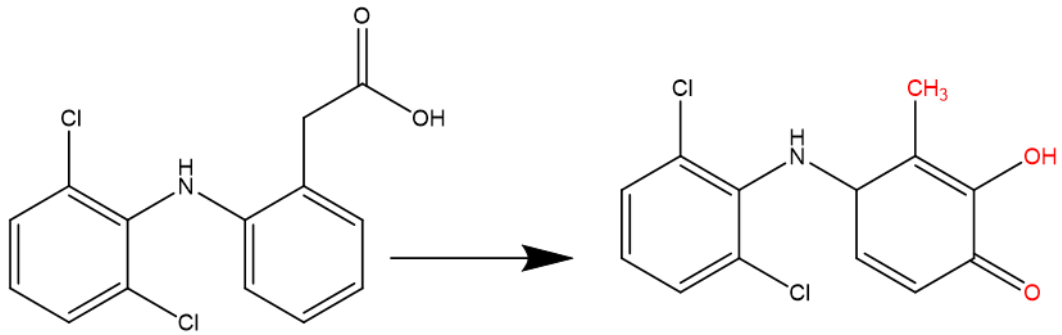


Figure 24. The transformation of diclofenac in *Lemna minor* was observed and included the hydroxyl derivative *Lemna* transformed product (LTP) as Phase I metabolites.

B-Phragmites australis

6.3.7. Metabolic profiling elucidation in Phragmites australis extracts with RPLC-HILIC-ESI-TOF-MS

In the *phragmites australis* study, a mixture of 13 reference standards was injected at the beginning/end and fixed intervals during the experimental sets (i.e., after each extract batch). The results of the standard mixture proved the accuracy, repeatability, and reproducibility of the analytical system. Mass and RT of the standards during the experiment showed an acceptable deviation of less than 8 ppm (with a TOF system from the year 2012) and % RSD less than 2 except for gabapentin, carbamide, and sotalol (Tables 6&7). For more information, readers are referred to Wahman *et al.*, 2019.

The four different extracts 100% MeOH, acidic 90% MeOH, 50% MeOH, and 100% H_2O of *P. australis* leaf, rhizome, and root were analyzed similarly with RPLC-HILIC-ESI-TOF-MS coupling as previously described (Wahman *et al.*, 2020). The obtained mass spectrometric total ion chromatograms (TICs) were interpreted to extract the feature (extracted masses, RT, and signal abundance) according to the parameters mentioned in chapter 3. The differences in chromatographic fingerprints reflected the variability in metabolite profiles (and composition) in the leaf, rhizome, and root samples. Detailed information and a description of data analysis have already complied in previous publications (Wahman *et al.*, 2019; Wahman *et al.*, 2020).

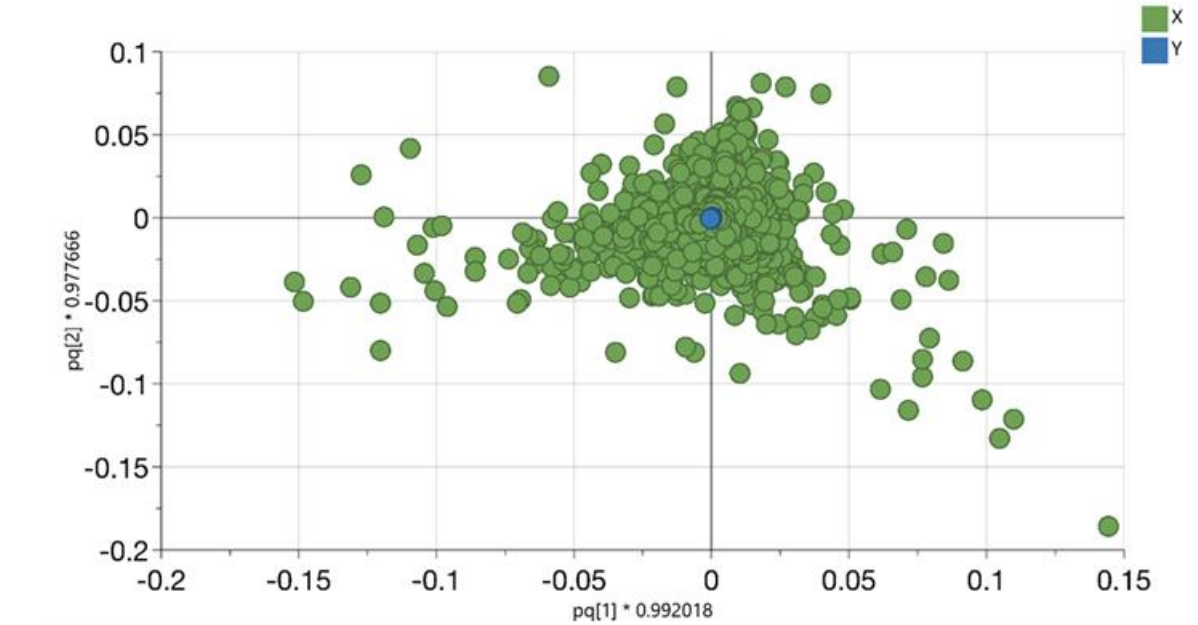
6.3.8. Phragmites australis leaf, rhizome, and root metabolic fingerprints

In this section, the 11,442 variables were differentiated according to the plant part that they originated from. The variables were in the loading score plot (Fig. 25a). This step was performed to improve the statistical significance of the dataset through additional cross-validation. Also, the OPLS-DA analysis was conducted to investigate the structure of the data. Another 18% of the data was investigated through the variation between X (metabolites) and Y (plant parts) given by $R^2X(cum)$ and explained approximately 94% of the variations in the various samples ($R^2Y(cum)$). We found the predictive value of the model was ($Q^2(cum)$ = 74%), which was created by OPLS-DA. The cross-validation performance was confirmed by analysis of variance (ANOVA). OPLS-DA discriminates the different samples of *P. australis* leaf, rhizome, and root regardless of incubation, with or without DCF or CBZ, and regardless of the extraction solvent composition according to the plant part (Fig. 25b).

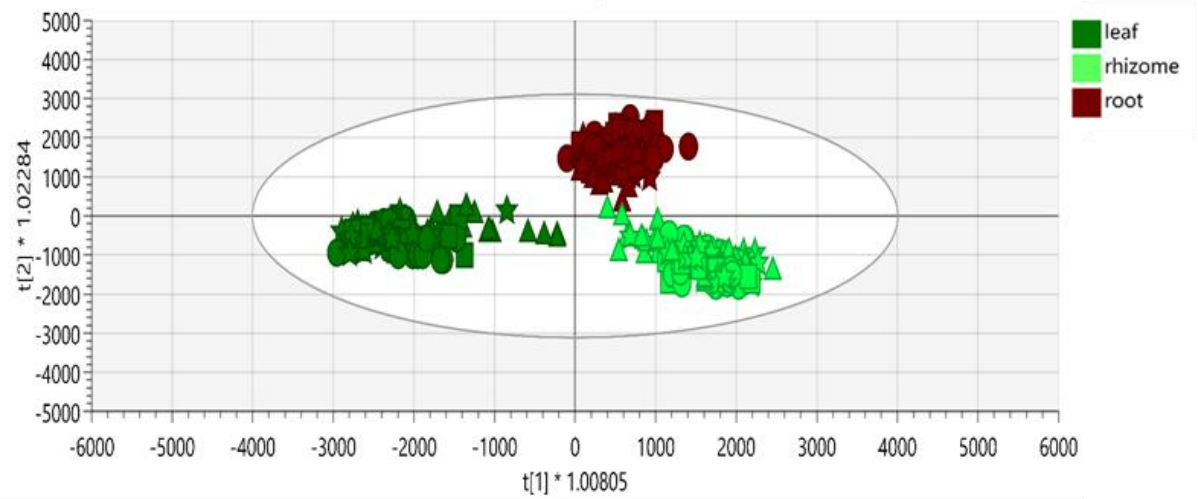
The predictive component t1 differentiated between leaf extracts in one group (negative part) and the root and rhizomes extracts in the second group (positive part), as shown in (Fig. 25e). However, the t2 (orthogonal component) differentiated rhizomes sample in the negative part and the roots in the positive one.

The quality of the OPLS-DA module was expressed by the cumulative value of the goodness of fit and the cross-validation for each value R^2 and Q^2, respectively, as shown in the Q^2/R^2 Overview plot (Fig. 25c). The high values of the previous parameters indicated a good classification and prediction efficiency to distinguish between different plant parts.

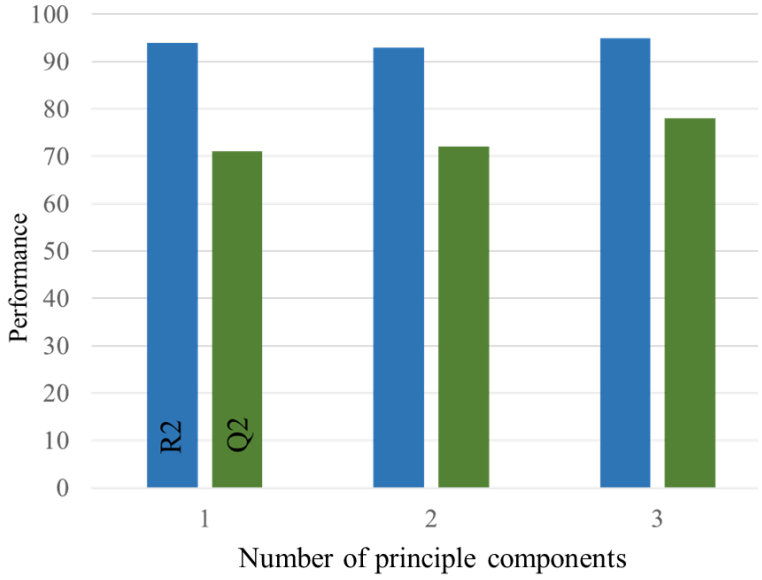
The statistical software allowed cross-validation between the datasets. Furthermore, an additional step of cross-validation of the data was done through the differentiation between the x (metabolites) and the plant parts. The results reveal the significance of the data organization. Hence, the compounds extracted from roots and rhizomes are more correlated than compounds in leaves, which supports the physiological similarities between them (Hess, 1975). Also, this emphasized the distribution of DCF and CBZ transformed products.



a



b



C

Figure 25. (a) The loading scatter plot for the selected principal components displays the relation between the different *Phragmites australis* samples and the chosen metabolites; (b) The OPLS-DA score plot of the different parts of *Phragmites australis* with a confidence limit of 95%, discriminating according to the plant part. The variables were plotted according to the first principal component (t1) and the orthogonal component (t2). The triangles represent 100% MeOH extracts, circles represent acidic 90% MeOH extracts, the squares represent 50% MeOH extracts, and the stars represent 100% aqueous extracts, respectively. The green color represents leaf samples, the light green color represents rhizome samples, and the brown color represents root samples, respectively. Each symbol represents one observation of *P. australis* leaf, rhizome, and root plant part; (c) The Q^2/R^2 Overview plot displays the individual cumulative R^2 (green columns) and Q^2 (blue columns) for the goodness of fits and cross-validation parameters.

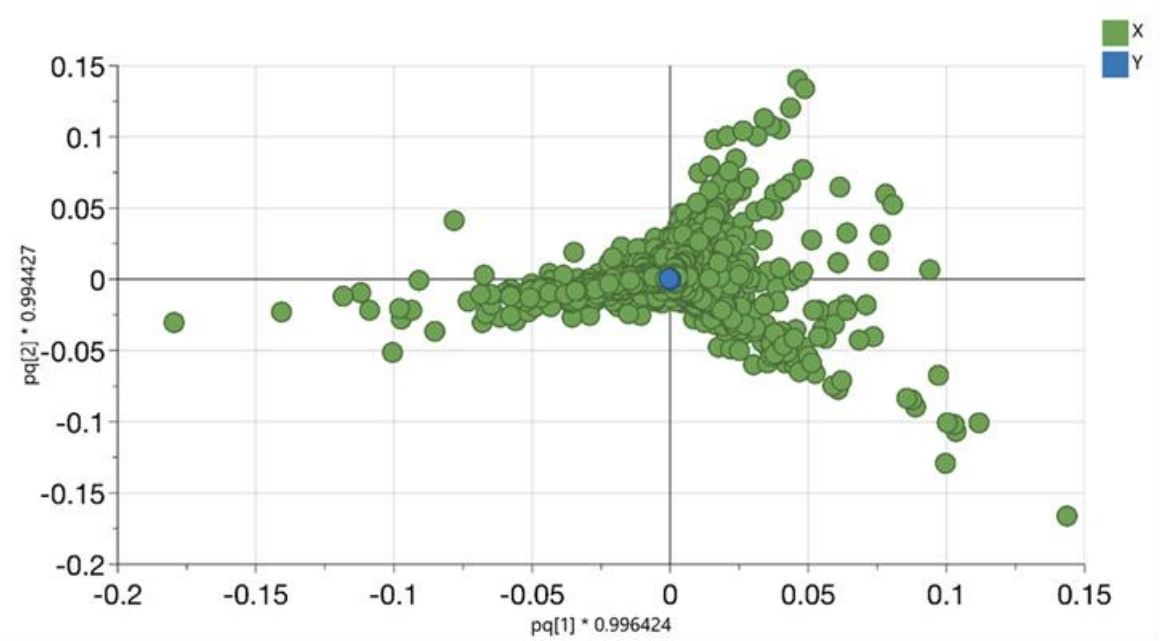
6.3.9. Untargeted Metabolomics Analysis of Phragmites australis Incubated with DCF or CBZ

The untargeted metabolomics analysis of *P. australis* incubated with/without DCF or CBZ, respectively, was performed using OPLS-DA to assign the changes in its metabolic fingerprint. The *P. australis* metabolic fingerprints of different extraction solvents with various plant parts were investigated previously, to test the organization and reliability of the data. Then, the large data set was used to perform the untargeted analysis and assign a list of metabolites that determined the distance between different groups. Also, the metabolic markers of *P. australis* incubated with/without DCF or CBZ were plotted each by the OPLS-DA, which represented the variability in metabolic patterns due to the different incubation (Fig. 26, that is, the loading plot in Fig. 26a and the OPLS-DA score plot in 26b). OPLS-DA analysis showed the identified and unidentified metabolites, which distinguished the different clusters according to the characteristic change of control or incubated samples metabolite profiles. It was used to enhance the quality of pairwise classification analysis. The relatively high goodness of fit indicated the good separation of different incubation groups of *P. australis* R^2 and Q^2, respectively, (Fig. 26c). OPLS-DA is a supervised approach that tends to improve the

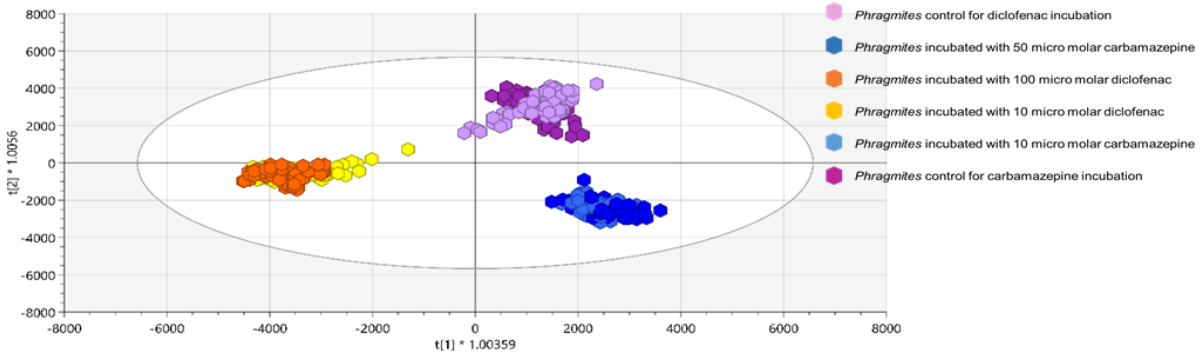
separation between (two or more) groups of samples. For this reason, it is widely used for classification purposes and biomarker identification and/or differentiating metabolic profiles (DMF) in metabolomics studies (Lamichhane *et al.*, 2018). To facilitate the interpretation and visualization of OPLS-DA, the S-plot was drawn to illustrate the model's influence with accuracy in the search for differentiating metabolic profile (DMF). S-plot analysis represented the highest contributing signals for the control and incubation of *P. australis* with DCF or CBZ (Fig. 27 a,b, respectively). The DMF was extracted from S-plots, which were marked in red color. They were selected based on their contribution to the variation and correlation within the data set between the X-variables and the predictive component t1 (p (corr) vector). Hence, they were considered functionally in the combined form of a metabolic profile, which was distinguished between the control and the incubated *P. australis*. In Figure 27 c,d, the contribution plots extend the data to better visualize and indicate regions responsible for sample clustering. The contribution plot summarizes the changing trends in the metabolites in the pairwise groups (i.e., control and incubated samples) by expressing the fold change of the metabolites between the control and incubated samples. As examples, succinic acid (DM_4), propane-1,2,3-triol (DM_6), and 2- hydroxypropanoic acid (DM_7) were detected as transformation products of DCF (DMs, see Table 19), thus they appeared significant in the positive part of the plot (Fig. 26c). Also, quercetin (compound 1 in Fig. 27c,d) was found in elevated signal heights after *P. australis* exposure to DCF indicating that they originated partially from diclofenac. The same was in CBZ transformation products, which were not produced by *P. australis.* Quercetin is a flavonoid that protects the plant against reactive oxygen species (ROS) during exposure to pharmaceuticals. Hence, the increase of succinic acid (DM_4) might affect the TCA cycle (tricarboxylic acid cycle) of *P. australis*.

In CBZ incubation, the contribution plot showed that in incubated samples 2,3-dihydro-2,3-dihydroxy-carbamazepine (compound 2 in Fig. 27d) and carbamazepine-10,11-epoxide (compound 3 in Fig. 27d) could be identified as metabolites (evaluated and proven with reference standards) and appeared with significant intensity. The carbamazepine-10,11-epoxide was statistically significant variable important (VIP) > 1 and $p < 0.05$. Carbamazepine-10,11-epoxide is considered a main transformed product of carbamazepine in plants (compound 3 in Fig. 27d). It is also the first metabolite of carbamazepine in tomato plants and *Armoracia rusticana* root cultures (Riemenschneider *et al.*, 2017b; Sauvêtre *et al.*, 2018).

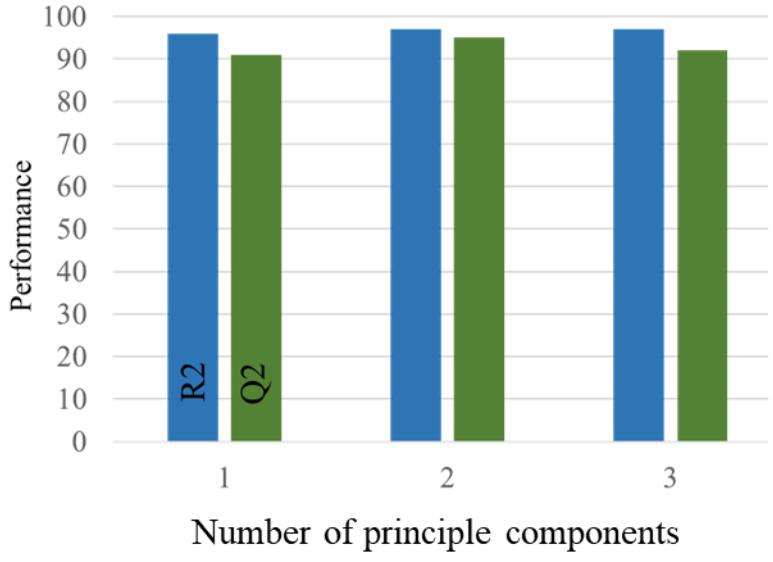
However, in this case, quercetin (compound 1 in Fig. 27d) intensity was found to decrease significantly. *P. australis* responds to the incubation with different classes of compounds, with the potential to protect it against ROS, such as polyphenols and flavonoids. Also, it might be that CBZ does not cause oxidative stress to the same extent as DCF.


X
Y
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
-0.2
pq[2] * 0.994427
-0.2
-0.15
-0.1
-0.05
0
0.05
0.1
0.15
pq[1] * 0.996424

a

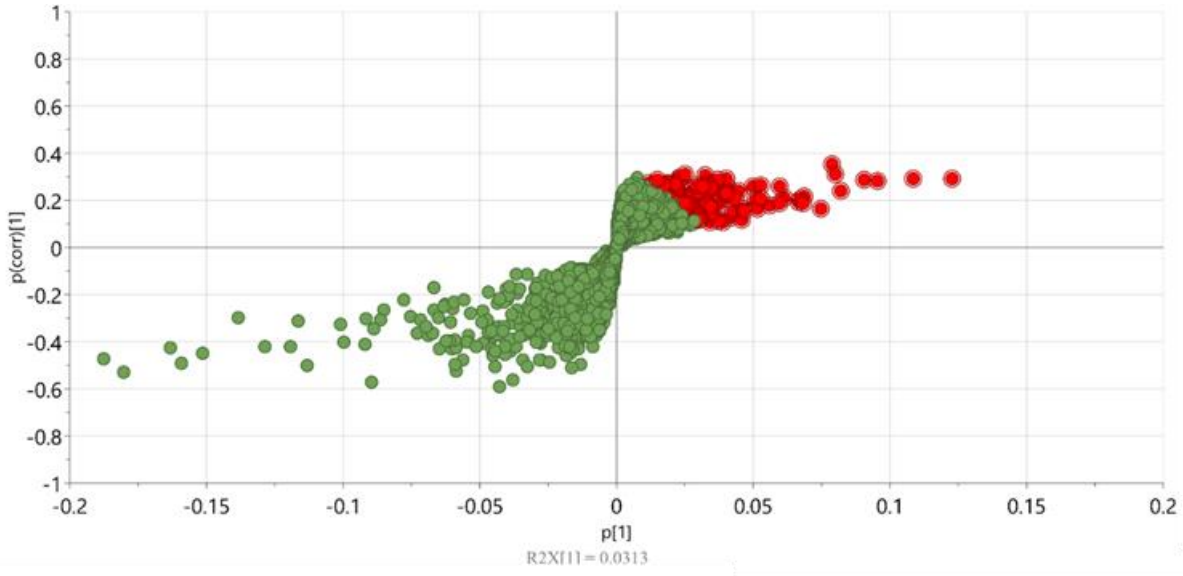

Phragmites control for diclofenac incubation
Phragmites incubated with 50 micro molar carbamazepine
Phragmites incubated with 100 micro molar diclofenac
Phragmites incubated with 10 micro molar diclofenac
Phragmites incubated with 10 micro molar carbamazepine
Phragmites control for carbamazepine incubation
t[2] * 1.0056
t[1] * 1.00359

b

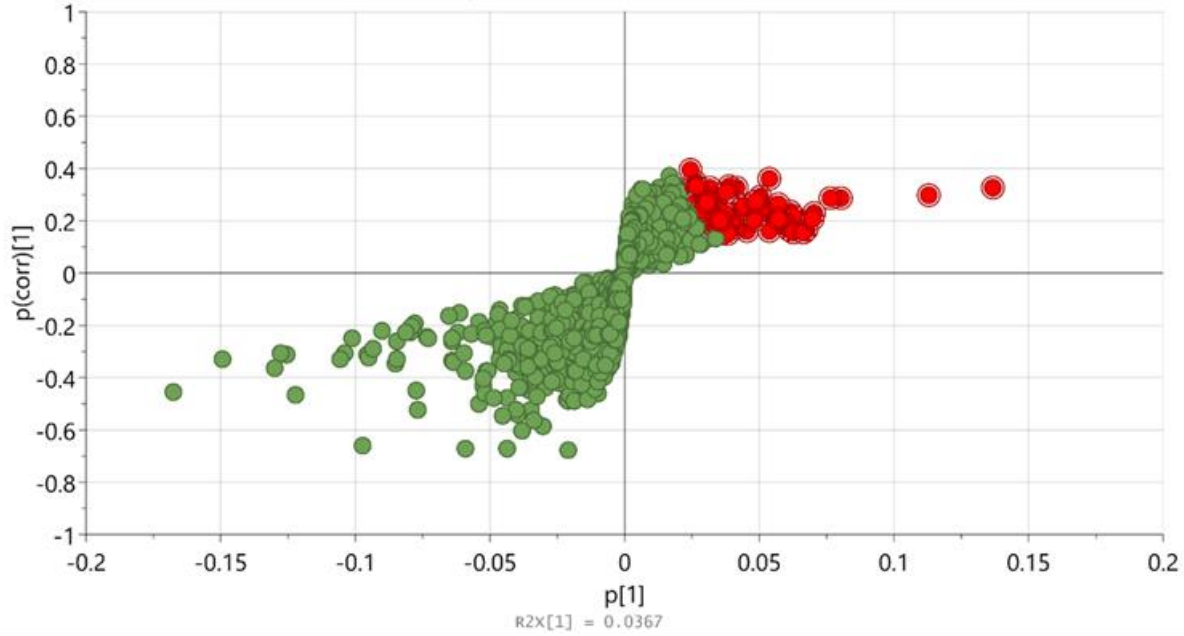

Performance
R2
Q2
1
2
3
Number of principle components

c

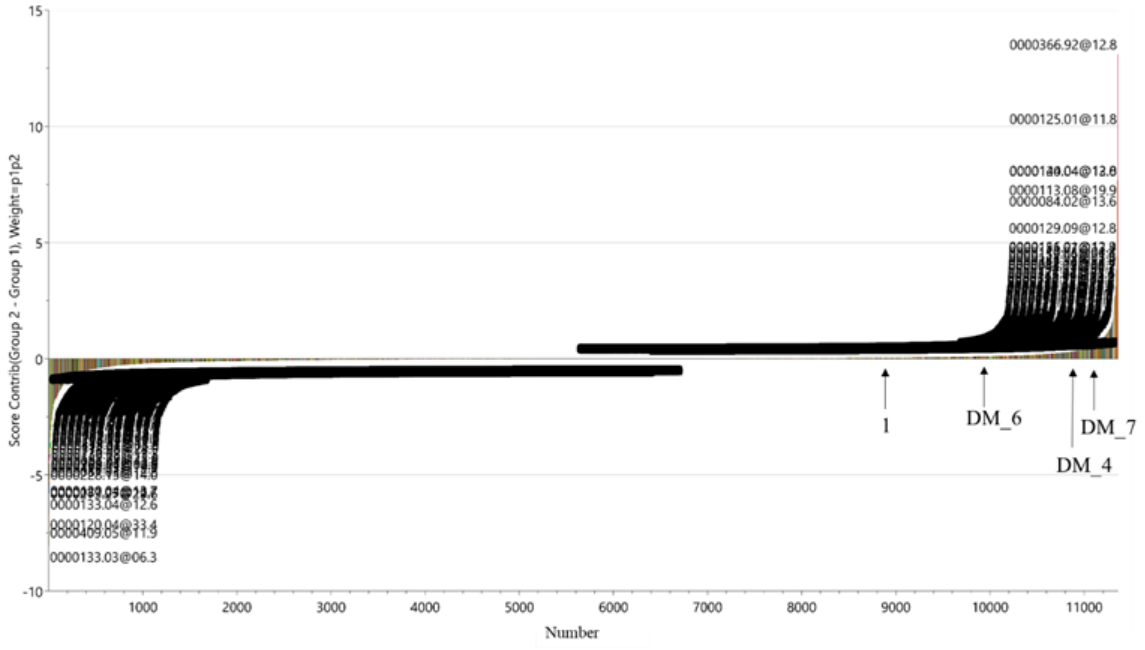
Figure 26. (a) The loading plot displays the relation between the different *Phragmites australis* samples and the chosen metabolite; (b) The OPLS-DA score plot of different *Phragmites australis* samples incubated with 10, and 50 µM carbamazepine, 10, and 100 µM diclofenac, individually. The confidence limit is 95%. For carbamazepine incubation: the purple color represents the control group, the light blue represents a sample incubated with 10 µM carbamazepine, and the blue color represents samples incubated with 50 µM carbamazepine. For diclofenac incubation: the light purple color represents the control group, the yellow color represents samples incubated with 10 µM diclofenac, and the orange color represents samples incubated with 100 µM diclofenac; (c) The Q^2/R^2 Overview plot displays the individual cumulative R^2 (green columns) and Q^2 (blue columns) for the goodness of fits and cross-validation parameters.



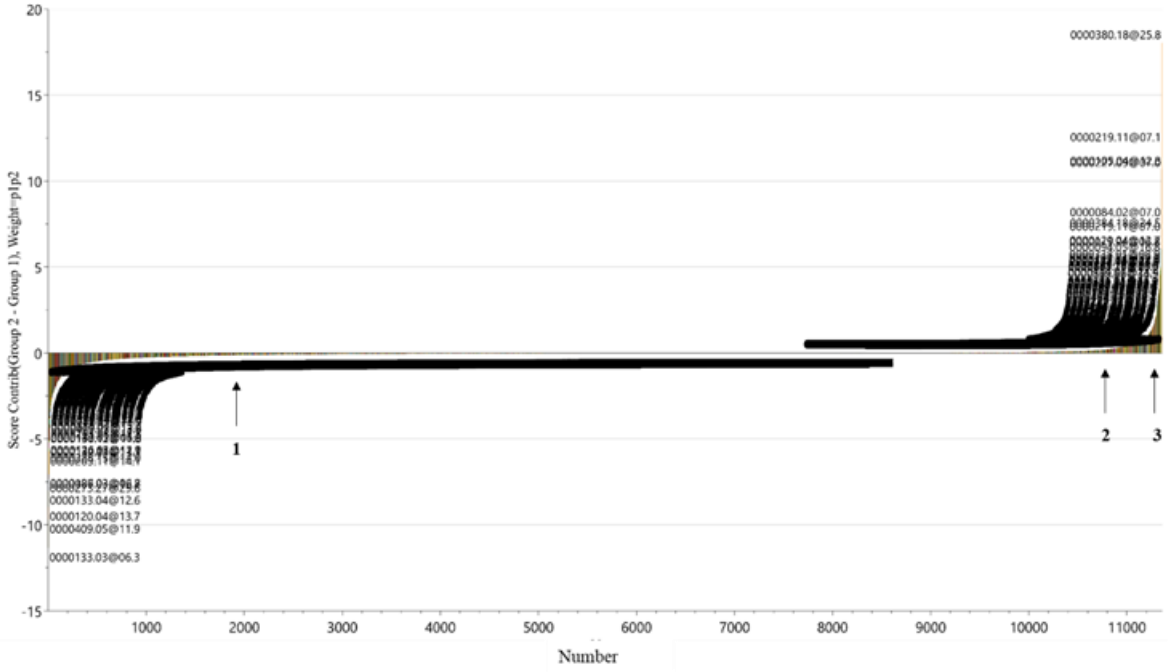
a



b



c



d

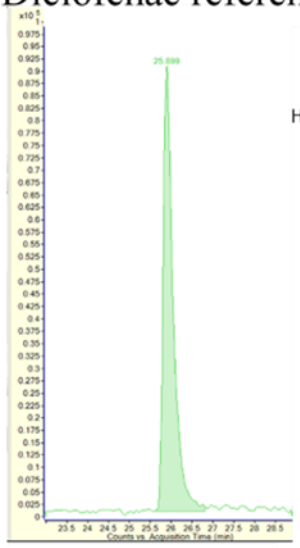
Figure 27. (a) S-plot of *Phragmites australis* control and incubated with 10 and 100 µM diclofenac samples; (b) S-plot of *Phragmites australis* control and incubated with 10 and 50 µM carbamazepine samples. The S-plot provides the visualization of the loading components of OPLS-DA to enable the interpretation of the data. The red-labeled compounds represent the differentiating metabolic profile (DMF) of each incubation; (c) The contribution plot shows the up and down-regulated compounds due to the incubation of *Phragmites australis* with 10 and 100 µM diclofenac, individually. Down-regulated compounds have negative values, while up-regulated compounds have positive values. (1) Quercetin, DM_6, DM_4, and DM_7 have up-regulated in *Phragmites australis* due to incubation with diclofenac at (8820, 0.0385581),

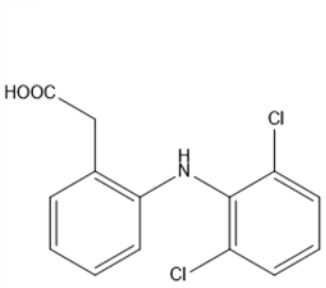
(9968, 0101411), (10817, 0.279365), and (10985, 0.389), respectively; (d) The contribution plot of *Phragmites australis* with 10 and 50 µM carbamazepine, individually. Down-regulated compounds have negative values as compound 1 which was quercetin at (2033,-0.118706). Up-regulated compounds have positive values as compounds 2 and 3 which were 2,3-dihydro-2,3-dihydroxycarbamazepine at (10921, 0.288274) and carbamazepine-10,11-epoxide at (11270, 1.27941) respectively.

6.3.10. Metabolism of DCF in Phragmites australis

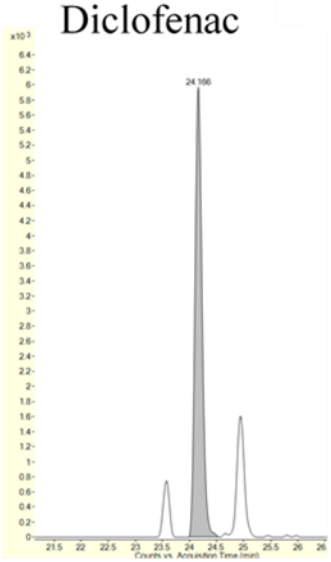
P. australis was exposed for 96 h to diclofenac in the two concentrations of 10 and 100 µM, respectively. Thereafter, the data was processed with MassHunter Workstation Software Profinder B.06.00 (Agilent Technologies, Waldbronn, Germany) to detect the DCF molecule and its metabolites (including background subtraction). Diclofenac and its hydroxylated metabolites were detected in the roots and rhizomes of *P. australis*. Diclofenac was identified at 295.0173 Da with a 2 ppm deviation from the monoisotopic mass, which eluted at 26 min (logD (pH7) > 0). The EICs of diclofenac in extracts and standard are shown in (Fig. 28). Further, hydroxylated metabolites were suspected according to the previously mentioned criteria at 311.0114 Da with 0.6 ppm deviation from the monoisotopic mass, which eluted at 24 min (logD (pH7) > 0) (Rajab *et al.*, 2013). Moreover, the analysis of different extracts of *P. australis* samples based on the mass, RT, and LogD (pH7) revealed seven proposed metabolites of diclofenac, as summarized in (Table 19 and Fig. 28). The comparison of mass spectra from leaf, rhizome, and root resulted in evidence for five metabolites of diclofenac in all parts of the treated samples of *P. australis*. However, DM_6 and DM_7 were detected in leaf extracts (Table 19). Also, diclofenac and its hydroxylated metabolites were identified in *P. australis* extracts. It has been reported that the formation of hydroxylated metabolites is the first step in the detoxification of diclofenac in the plant. These results reveal that P450 monooxygenases or peroxidases were involved to detoxify diclofenac (Huber *et al.*, 2012).

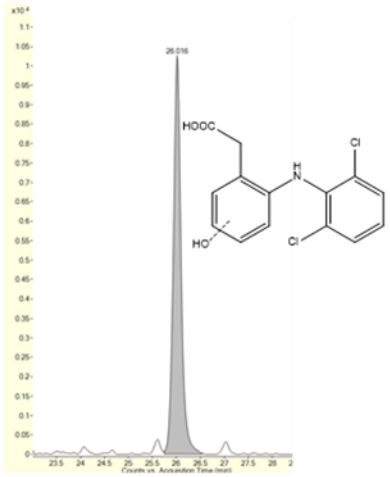
Diclofenac reference standards

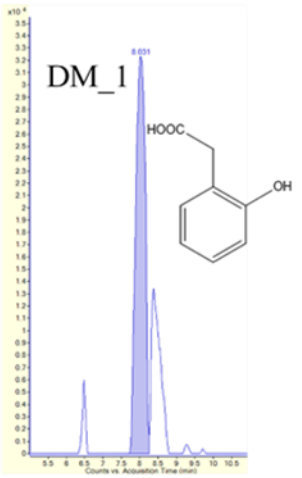


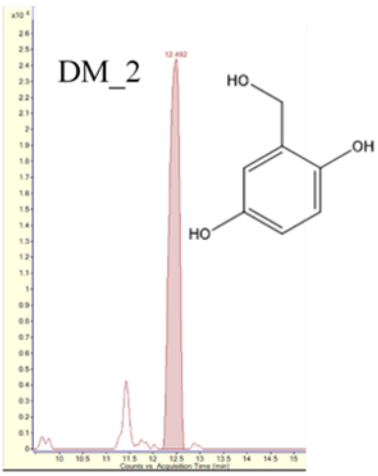


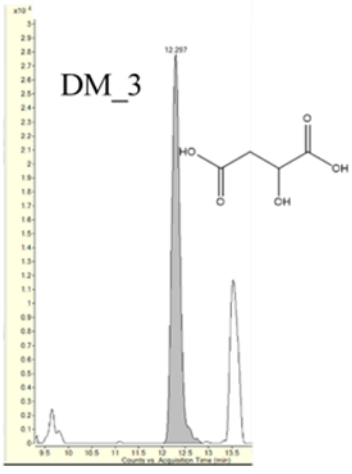
Diclofenac

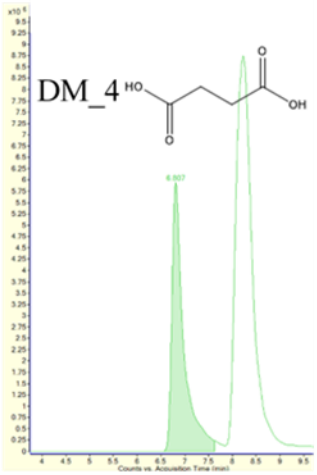


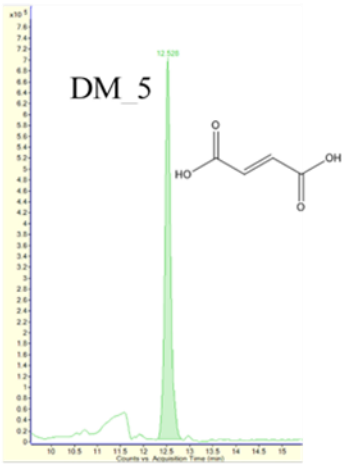












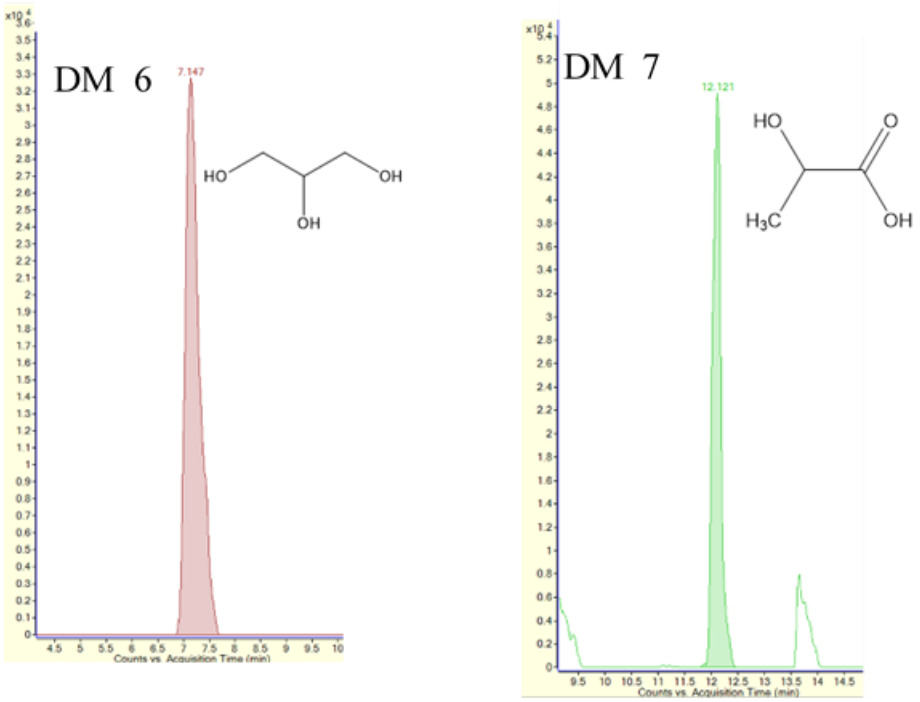


Figure 28. EICs were corresponding to measured diclofenac (right) and the reference standard (left), which were identified in the extracts of *Phragmites australis* leaf, rhizome, and roots incubated with 10 and 100 µM diclofenac. Also, EICs relative to transformed products are suspected in the extracts of *Phragmites australis* leaf, rhizome, and roots incubated with 10 and 100 µM diclofenac.

Table 19. List of diclofenac (DCF) transformation products detected in different *Phragmites australis* leaf, rhizome and roots and *Lemna minor* incubated with 10 and 100 µM diclofenac samples with the monoisotopic mass in the literature (L), the mean of the measured monoisotopic mass (M), the variation between them, mean RT (M), standard deviation, relative standard deviation and LogD (pH=7.4) were listed. The logD values were predicted from ChemAxon software (https://disco.chemaxon.com/apps/demos/logd/)

DCF Transformed Products	Name	Mono isotopic Mass (Da) (L) (Rajab *et al.*, 2013)	Mean Mono isotopic Mass (Da) (M)	Δ ppm	Mean RT (Min) (M)	SD Of RT (Min)	%RSD	LogD (pH=7.4)	Leaf	Rhizome	Root	*Lemna*
DM_ 1	2-Hydroxypropanoic acid	152.0473	152.0475	-0.99	8.1	0.05	0.67	-1.86	√	√	√	×
DM_ 2	2-(Hydroxymethyl)benzene-1,4-diol	140.0473	140.0472	0.71	12.5	0.06	0.51	0.60	√	√	√	√
DM_ 3	2-Hydroxysuccinic acid	134.0215	134.0215	0.07	12.6	0.03	0.23	-6.81	√	√	√	√
DM_ 4	Succinic acid	118.0266	118.0271	-3.95	6.8	0.04	0.55	-1.99	√	√	√	×
DM_ 5	Fumaric acid	116.0101	116.01	0.49	12.6	0.05	0.42	-2.00	√	√	√	×
DM_ 6	Propane-1,2,3-triol	92.0473	92.04703	2.9	7.2	0.07	0.99	-1.84	√	×	×	×
DM_ 7	2-Hydroxypropanoic acid	90.0317	90.03207	-4.07	12.1	0.04	0.32	-1.00	√	×	×	√

6.3.11 Metabolism of CBZ in Phragmites australis

P. australis was exposed for 96 h to carbamazepine in two CBZ concentrations of 10 and 50 µM, respectively. Thereafter, the data was processed with MassHunter Workstation Software Profinder B.06.00 to detect the compound and its metabolites (including background subtraction). According to Sauvêtre and co-workers (2018), four different pathways for carbamazepine metabolism (PCM) were investigated in the plant (Sauvêtre *et al.*, 2018).

The different transformation products were identified in different *P. australis* samples using reference standards. The mean monoisotopic mass of standards and different samples and their absolute variation were tabulated, as well as the retention time in both standards and samples and the variation between them (Table 20, Fig. 29). Carbamazepine-10,11-epoxide, 10,11-dihydro-10,11-dihydroxy-carbamazepine, 10,11-dihydro-10-hydroxy-carbamazepine, 9-acridine carboxaldehyde, and 2,3-dihydro-2,3-dihydroxy-carbamazepine were identified in *P. australis* incubated samples. The mass deviation was less than 5 ppm and the deviation in the RT was less than 0.3 min. Carbamazepine-10,11-epoxide and 9-acridine carboxaldehyde have been identified in leaf, rhizome, and root in both incubation levels 10 and 50 µM. 10,11-dihydro-10,11-dihydroxy-carbamazepine and 2,3-dihydro-2,3-dihydroxy-carbamazepine were identified in root and rhizome extracts in both incubation concentrations. Moreover, the first was identified in leaf extracts of 50 µM carbamazepine; however, the latter was not detectable in leaf extracts, respectively. They originated from different metabolism pathways (PCM). The 10, 11-diOH pathway (PCM) has been investigated comprehensively in plants like cucumber, tomato, sweet potato, lettuce, carrot, and horseradish (Riemenschneider *et al.*, 2017a; Sauvêtre *et al.*, 2018). The main transformation product is carbamazepine-10,11-epoxide (Martínez-Piernas *et al.*, 2019), which has been identified in leaf, rhizome, and root in both incubations levels 10 and 50 µM, as shown in Table 13 (Martínez-Piernas *et al.*, 2019). This is the first oxidation step, which was observed in different organisms from bacteria and fungi to mammals. This step is either conducted by cytochrome P450 or peroxidases, leading to different sub-pathways (PCM).

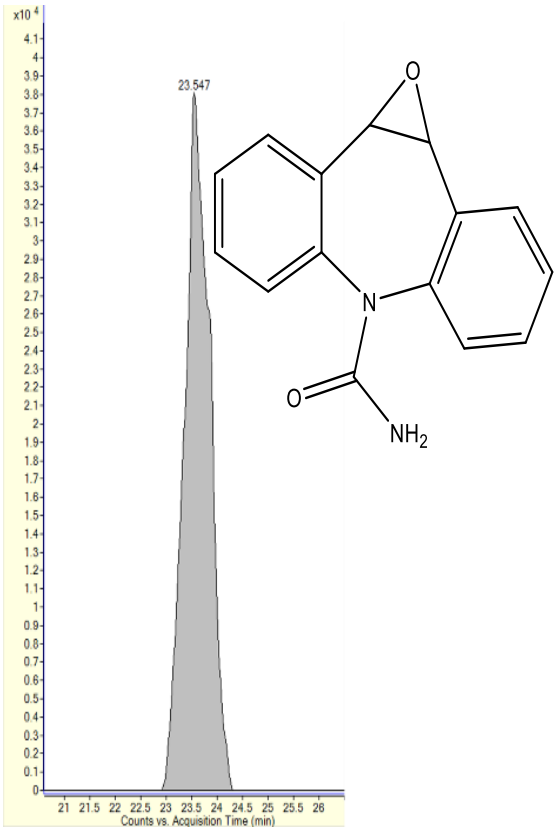
The first sub-pathway (PCM) starts with cleavage of the epoxide bond and the hydroxylation to form 10,11-dihydro-10,11-dihydroxy-carbamazepine. This step is catalyzed by epoxide hydrolase enzymes (Tybring *et al.*, 1981). A recent study reported further metabolism of 10,11-dihydro-10,11-dihydroxy-carbamazepine through hydroxylation of the benzene ring (Martínez-Piernas *et al.*, 2019). 10,11-dihydro-10,11-dihydroxy-carbamazepine was identified in root and rhizome extracts in both incubation concentrations, as shown in (Table 20). However, 10,11-dihydro-10-hydroxy-carbamazepine was identified in leaf extracts in the 50 µM carbamazepine treatment. It seems that the 10,11-dihydro-10,11-dihydroxy-carbamazepine was transferred to the rhizome and leaf for further metabolism. This explains why it reaches the leaves in high concentration from the highly loaded rhizome tissue. Thus, *P. australis* can degrade carbamazepine through a 10,11-diOH pathway (PCM), initiated in the root and completed in leaf and rhizome.

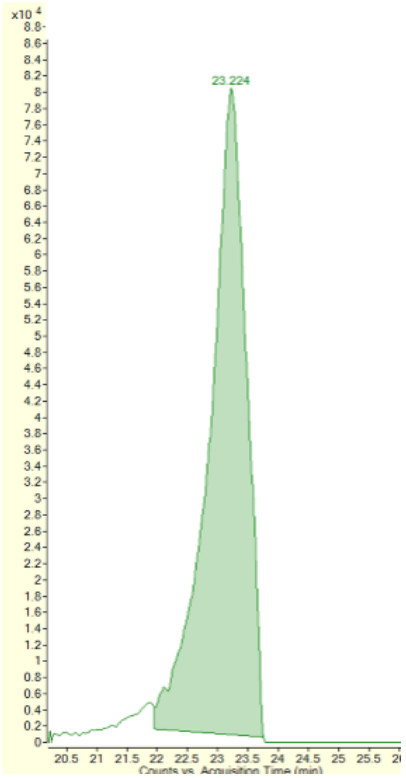
Further, *P. australis* metabolized the carbamazepine-10,11-epoxide into 9-acridine carboxaldehyde, which is analogous to other studies (Riemenschneider *et al.*, 2017a; Riemenschneider *et al.*, 2017b; Sauvêtre *et al.*, 2018). The 9-acridine carboxaldehyde is a reactive compound that was identified in leaf, rhizome, and root extracts of plants exposed to 10 and 50 µM carbamazepine. Further metabolites downstream from 9-acridine carboxaldehyde could not be detected. 9-acridine carboxaldehyde is typically transformed into acridone, which is a non-toxic compound with the formation of the two intermediates acridine and 9-hydroxy acridine (Sauvêtre *et al.*, 2018). In our conditions, they were below the detectable limit or did not arise. Because 9-acridine carboxaldehyde was identified in leaf, rhizome, and root extracts in both concentrations, it may be assumed that there are further steps in the acridine metabolic pathway (PCM) that could be detected, like in lettuce, with longer incubation periods (Martínez-Piernas *et al.*, 2019).

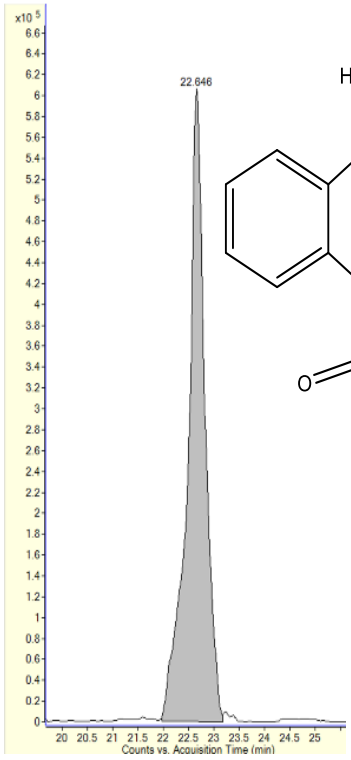
The last sub-pathway (PCM) is similar to the 10, 11-diOH pathway in the first steps. It includes a consecutive oxidation reaction on the aromatic group. Carbamazepine was metabolized to 2,3-dihydro-2,3-dihydroxy-carbamazepine which comprises two steps with the participation of a cytochrome P450 or a peroxidase and an epoxide hydrolase (Tybring *et al.*, 1981). 2,3-dihydro-2,3-dihydroxy-carbamazepine was identified in the *P. australis* rhizome and root in both incubation levels. However, 2,3-dihydro-2,3-dihydroxy-carbamazepine was not detectable in leaf extracts. It is moreover possible that the leaf metabolizes CBZ to 4-hydroxy-carbamazepine as in the leaves of the tomato plant. It was identified formerly in lettuce after 4 days of exposure; however, the researchers spiked the lettuce with 1 mg/L carbamazepine (Riemenschneider *et al.*, 2017b; Martínez-Piernas *et al.*, 2019).

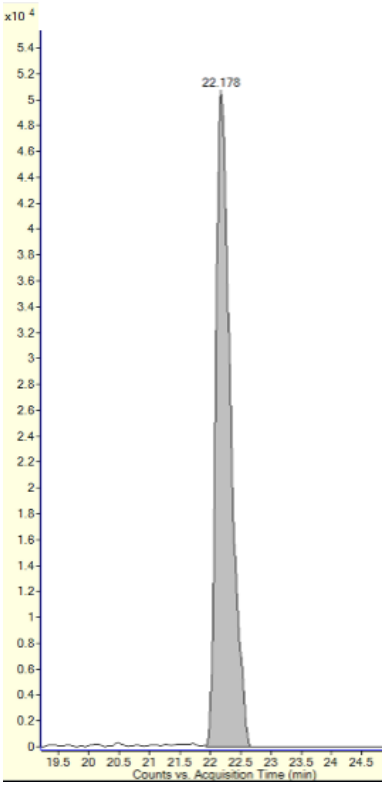
Thus, it is possible that *P. australis* takes up and metabolizes carbamazepine by different mechanisms. Still, all the metabolites resemble phase I metabolism in plant metabolism and not the glycosylated transformation products (Phase II) (Schröder *et al.*, 2001)

	Standard reference EICs	*Phragmites* EICs
Carbamazepine -10,11-epoxide	Mass= 252.0903 Da	Mass= 252.0891 Da
10,11-Dihydro-10,11-dihydroxy-carbamaz-epine	Mass=270.10 Da	Mass=270.0994 Da



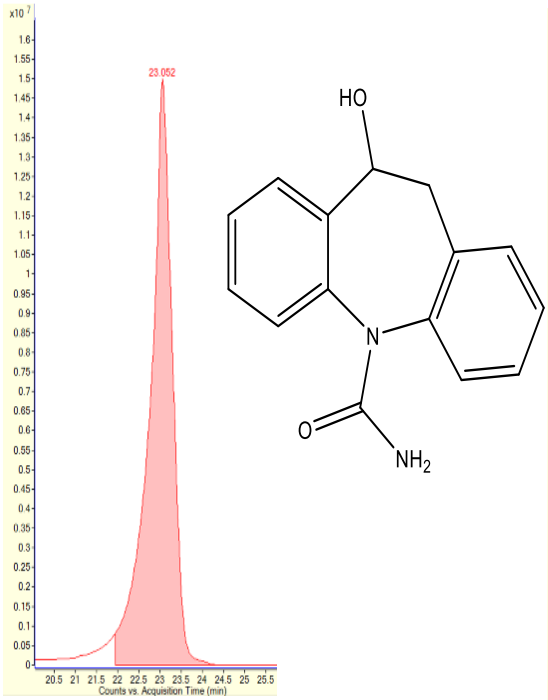




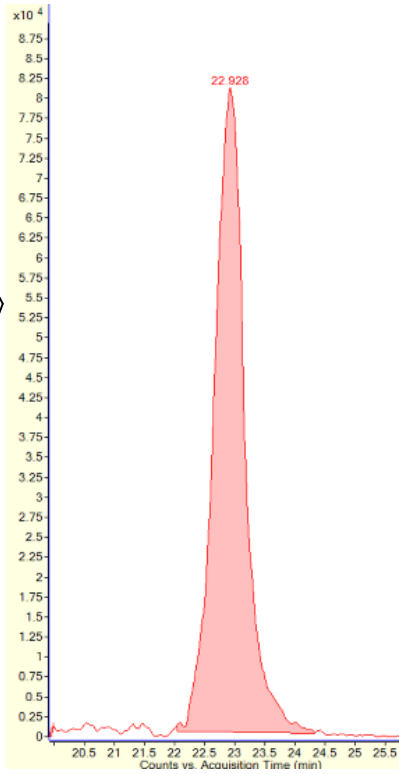


10,11-Dihydro-10-hydroxy-carbamazepine

Mass= 254.1055 Da

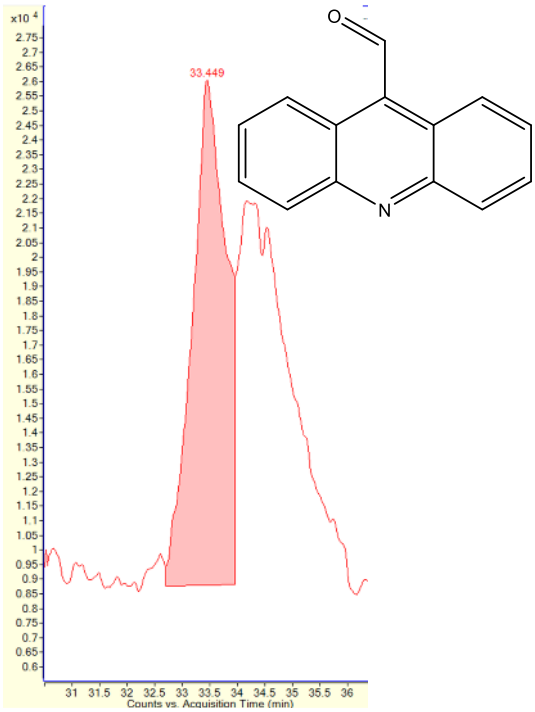


Mass= 254.1044 Da

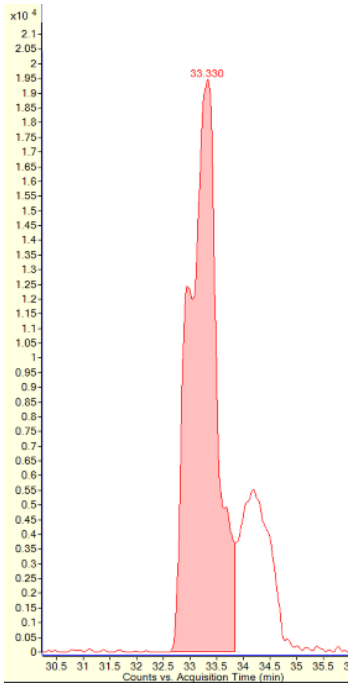


9-Acridine carboxaldehyde

Mass= 207.0684 Da

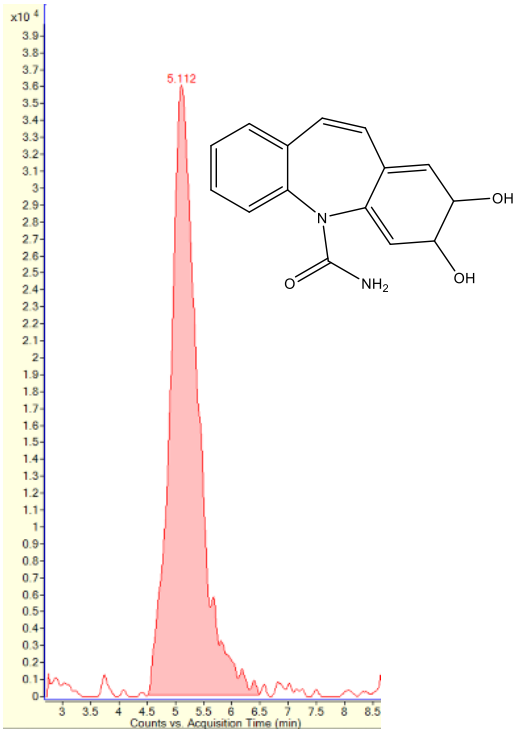


Mass=207.0682



2,3-Dihydro-2,3-dihydroxy-carbamazepine

Mass= 270.1 Da



Mass= 270.0994 Da

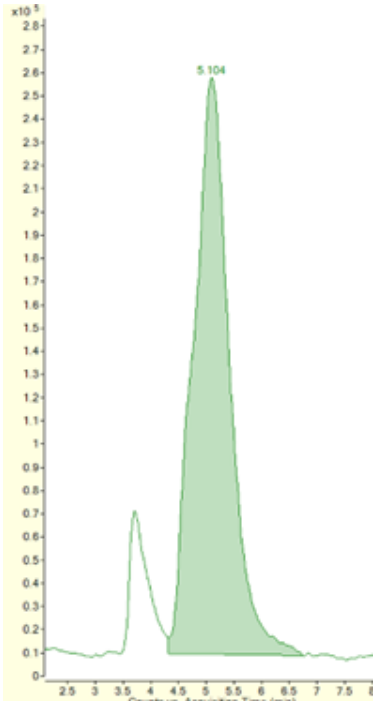


Figure 29. EICs were corresponding to carbamazepine (CBZ) and its transformed product standards (left), which were identified in the extracts of *Phragmites australis* leaf, rhizome, and roots incubated with 10 and 50µM carbamazepine (measured right).

Table 20. List of carbamazepine (CBZ) transformed product identified in *Phragmites australis* different samples with the mean monoisotopic mass in the standards (S), the mean monoisotopic mass of *Phragmites australis* (Ph), the variation between them, mean RT of standards (S), mean RT of *Phragmites australis* (Ph), and the variation between them were listed. The logD values were predicated from ChemAxon software (https://disco.chemaxon.com/apps/demos/logd/)

CBZ Transformed Products	Mean Mono isotopic Mass (Da) (S)	Mean Mono isotopic Mass (Da) (Ph)	Δ ppm	Mean RT (Min) (S)	Mean RT (Min) (Ph)	Δ RT	LogD (pH=7.4)
Carbamazepine-10,11-epoxide	252.0903	252.0891	4.81	23.57	23.39	0.18	1.97
10,11-Dihydro-10,11-dihydroxy-carbamazepine	270.10	270.0994	2.16	22.34	22.12	0.22	0.81
10,11-Dihydro-10-hydroxy-carbamazepine	254.1055	254.1044	4.33	23.06	22.93	0.13	1.73
9-Acridine carboxaldehyde	207.0684	207.0682	1.10	33.46	33.27	-0.19	2.98
2,3-Dihydro-2,3-dihydroxy-carbamazepine	270.1	270.0994	2.16	5.37	5.20	0.17	-0.13

6.3.12. Impacts of DCF and CBZ on Phragmites australis metabolic pathways

Metabolic pathway analysis was performed to identify the pathways that were induced upon the incubation of *P. australis* with DCF or CBZ via MetaboAnalyst 4.0 software (Montreal, Canada) based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The results are established on OPLS-DA and S-plot analysis of different metabolites with/without DCF or CBZ incubation depending on the p (corr) vector in the OPLS-DA module. Consequently, these metabolites represent the metabolic profile (DMF) which differentiate between control and incubated samples.

For the investigation, the mummichog algorithm was the first implementation of this concept to infer pathway activities from a ranked list of MS peaks identified by untargeted metabolomics. The DMF was applied to test their participation in pathways using a combination of network analysis and functional enrichment analysis. The pathways exhibiting $p < 0.05$ were considered a statistically significant metabolic pathway, meaning that they were affected with DCF or CBZ or their transformation products via MetaboAnalyst 4.0 software based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.

KEGG pathway analysis of the analyzed metabolites using the pathway data set of *A. thaliana* matched 27 pathways in *P. australis* incubated with DCF. Further, 11 were significantly biologically active and are listed in the table in (Fig. 30a). They were identified as glycolysis/gluconeogenesis, ascorbate and alternate metabolism, fructose and mannose metabolism, galactose metabolism, the pentose phosphate pathway, arginine biosynthesis, alanine, aspartate and glutamate metabolism, purine metabolism, pyrimidine metabolism, glutathione metabolism, and phenylalanine metabolism.

After CBZ incubation, 22 pathways were significantly altered. Twelve pathways have a more significant *p*-value, relating them to the effect of CBZ and/or its transformation products (Fig. 30b). They were the pentose phosphate pathway, purine metabolism, pyrimidine metabolism, fatty acid biosynthesis, arachidonic acid metabolism, tyrosine metabolism, tryptophan metabolism, β-alanine metabolism, arginine and proline metabolism, pantothenate and CoA biosynthesis, carbon fixation in photosynthetic organisms, and folate biosynthesis. The previous pathways were induced after incubation with DCF or CBZ.

The glycolysis biosynthesis pathway provides the plant with defense compounds against biting insects and worms. Furthermore, the cytochrome P450 enzyme controls glucosinolate biosynthesis, which is involved in the metabolism of DCF and CBZ during phase I (Huber *et al.*, 2012; Lee *et al.*, 2016). The decline in sugars after plant exposure to DCF and CBZ indicates an increase in energy consumption (Hurtado *et al.*, 2017).

Pyrimidine metabolism could be proposed to be a result of glycolysis biosynthesis pathway alteration, which up-regulates the biosynthesis of other stress-induced pathways (Tantikanjana *et al.*, 2004). Pyrimidine metabolism is enhanced and leads to products that could be used in the case of salvage, that is, recovery of infections and subsequent synthesis of secondary products with specific functions in defense mechanisms. Additionally, pyrimidine metabolism provides

a source of β-alanine or β-aminobutyrate, which might be an important source for the pantothenate of coenzyme A (Zrenner *et al.*, 2006).

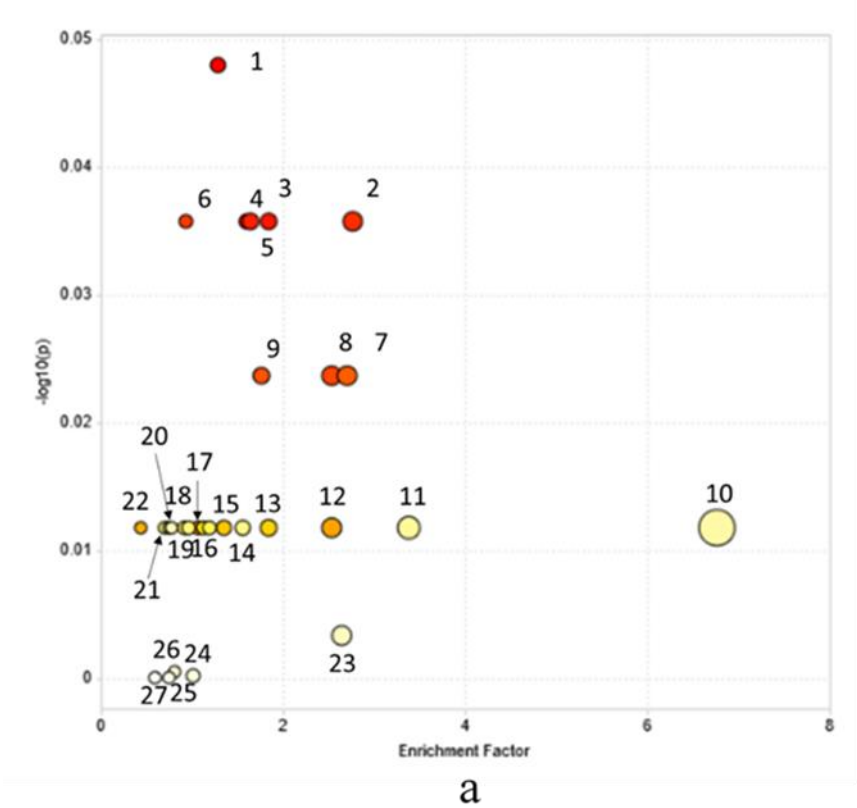
Furthermore, fatty acid biosynthesis and the related pyruvate metabolism were induced after incubation of *P. australis* with DCF and CBZ. After incubation with DCF, *P. australis* exhibits some responses differently from incubation with CBZ, since it showed an increase in phenylalanine metabolism which aligned with the increase of quercetin content. The flavonoid 3′-monooxygenase and the flavonoid 3′,5′-hydroxylase enzymes responsible for the conversion of kaempferol to quercetin are cytochrome P450 plant types, which are triggered after incubation with DCF. However, the quercetin levels were lower after CBZ incubation (Fig. 27d).

Glutathione metabolism was affected in DCF incubation but not under the influence of CBZ. Glutathione metabolism is a central part of the antioxidative ascorbate-glutathione cycle. Further, glutathione is critical for the detoxification of xenobiotics, environmental stress tolerance, and, in the form of phytochelatins, also the retention of heavy metals (Gong *et al.*, 2018). The metabolism of DCF required nicotinamide-adenine dinucleotide phosphate (NADPH) as reductants (Huber *et al.*, 2012). Upon DCF incubation, succinic acid formed from the degradation of diclofenac might have fueled the TCA cycle, which has been detected significantly in extracts incubated with DCF (Fig. 27c).

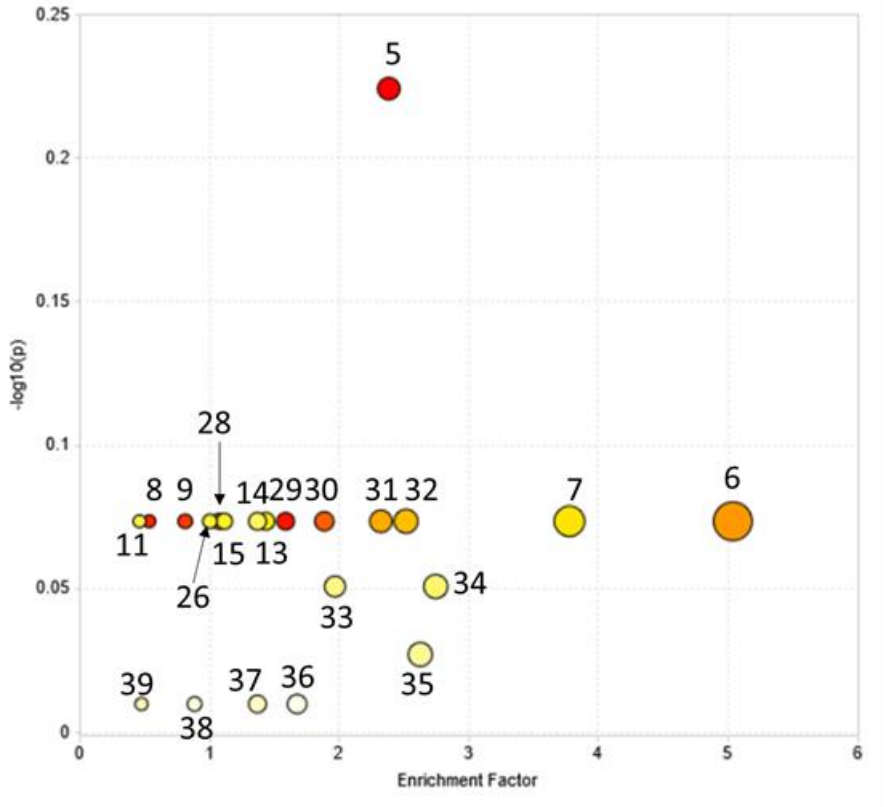
In 2019, Sivaram and coworkers reported a high impact on the TCA cycle in maize leaves exposed to pyrene. Consequently, the γ-aminobutyric acid (GABA) shunt bypasses two steps of the TCA cycle. Moreover, it has an important pathway under stress conditions and it is associated with numerous physiological responses, including the regulation of cytosolic pH, carbon fluxes into the TCA cycle, nitrogen metabolism, osmoregulation, and plant-pathogen interaction. Increased GABA levels also occur in response to changing environmental conditions and represent another adaptive mechanism in the attempt to maintain the rate of respiration under certain harmful conditions (Araújo *et al.*, 2012).

The same result was observed in lettuce crops, which were exposed to different contaminants of emerging concern (CEC) concentrations (Hurtado *et al.*, 2017). Also, it was reported that aromatic hydrocarbons altered the osmotic balance in maize (Sivaram *et al.*, 2019). The alteration of previously mentioned biosynthetic pathways enhances the *P. australis* defense mechanisms. They also seem to be involved in the transformation of DCF and CBZ.

Finally, in *P. australis* the glutathione metabolism pathway was induced upon exposure to DCF and the unsaturated fatty acid pathway was induced, leading to protection during the incubation with CBZ. Hence, *P. australis* responded differently to the DCF and CBZ through changing its metabolic pathway regardless of the type of drug to some extent. Consequently, specific changes in several common metabolic pathways can be considered as a marker for pollutant exposure in *P. australis*. However, each drug has fingerprints of the alteration of distinct metabolic pathways, which might be connected to its metabolites and the enzymes involved in metabolism. Therefore, the induced or changed pathways could be used as indicators for the exposure of the plant to DCF or CBZ.


-log10(p)
Enrichment Factor
0.05
0.04
0.03
0.02
0.01
0
0
2
4
6
8
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

a


-log10(p)
Enrichment Factor
0.25
0.2
0.15
0.1
0.05
0
0
1
2
3
4
5
6
5
6
7
8
9
11
13
14
15
26
28
29
30
31
32
33
34
35
36
37
38
39

b

1	Glycolysis / Gluconeogenesis	20	Glyoxylate and dicarboxylate metabolism
2	Ascorbate and aldarate metabolism	21	Cyanoamino acid metabolism
3	Fructose and mannose metabolism	22	Starch and sucrose metabolism
4	Galactose metabolism	23	Amino sugar and nucleotide sugar metabolism
5	Pentose phosphate pathway	24	Inositol phosphate metabolism
6	Arginine biosynthesis	25	Sphingolipid metabolism
7	Alanine, aspartate, and glutamate metabolism	26	Pyruvate metabolism
8	Purine metabolism	27	Nicotinate and nicotinamide metabolism
9	Pyrimidine metabolism	28	One carbon pool by folate
10	Phosphatidylinositol signaling system	29	Arachidonic acid metabolism
11	Aminoacyl-tRNA biosynthesis	30	Fatty acid biosynthesis
12	Phenylalanine metabolism	31	Tyrosine metabolism
13	Glycine, serine, and threonine metabolism	32	Tryptophan metabolism
14	Valine, leucine, and isoleucine degradation	33	Pantothenate and COA biosynthesis
15	Valine, leucine, and isoleucine biosynthesis	34	Carbon fixation in photosynthetic organisms
16	Lysine degradation	35	Folates biosynthesis
17	Glutathione metabolism	36	Biosynthesis of unsaturated fatty acids
18	Histidine metabolism	37	Taurine hypotaurine metabolism
19	Phenylalanine, tyrosine, and tryptophan biosynthesis	38	Terpenoid backbone biosynthesis
		39	β-alanine metabolism

Figure 30. Overview of pathway analysis by using Metaboanalyst 4.0. For metabolite set enrichment analysis of *Phragmites australis* differentiating metabolic profile (DMF) after incubation with: (a) 10 and 100 µM diclofenac; (b) 10 and 50 µM carbamazepine. The overview displays all matched pathways as circles. The color and size of each circle are based on the p-value and the impact of the pathway value, respectively. The names of the metabolic pathways are listed in the table.

6.4. Introduction

System B

For decades, plant metabolomics studies have been performed with different analytical strategies because of the fundamental importance of metabolites as components of biochemical pathways, and their use as diagnostic markers for a wide range of biological conditions such as response to xenobiotics. The most common analytical tool is LC-MS (Salem *et al.*, 2020). During the first part of this chapter, the study of *Lemna minor* was performed using the serial coupling of RPLC-HILIC, which was coupled to TOF-MS (system A). In this part, *Lemna* was analyzed using the serial coupling connected to QTOF-MS to investigate the changes in *Lemna's* metabolic profile due to incubation with 10 &100µM DCF and 10, 50µM CBZ, individually.

The previous results in sections (6.3.3 and 6.3.4) showed that DCF enhanced the phenylpropanoid pathway including the flavonoid biosynthesis pathway. This has been proven with target and untargeted strategies. Thus, in this part, the phenylpropanoid pathway was investigated using different metabolites. In addition, the untargeted investigation of the data was done using different statistical analyses.

6.5. Experimental design

Lemna minor was incubated with 10 and 100µM DCF for 96 hrs, individually. Moreover, *Lemna* was incubated with the same concentration of DCF for 4, 8, and 12 days to investigate the time effect on *Lemna's* metabolic profile. Also, an independent experiment has been performed for the CBZ, which lasted for 96 hrs. *Lemna* was grown under the same conditions. Each experiment was extracted with three different solvents (100% MeOH, 50% MeOH, 100% H_2O), separately, and analyzed using RPLC-HILIC-ESI-QTOF-MS. The data analysis workflow and further instrumentation and methodological details, see chapter 3.

6.6. Results and Discussion

A- Diclofenac

6.6.1. Identification of Lemna minor metabolites

Different extracts of *Lemna minor* and standards were analyzed with RPLC-HILIC-QTOF-MS/MS in positive mode. The identified compounds and the deviation in their masses and RTs were listed in chapter 4. The choice of the compounds depends on the prioritization with the PLANT-IDENT database through control data investigation. The phenolic compounds were chosen due to their role in plant defense, biological importance, as well as the previous results of the reducing contents, target, and untargeted analysis.

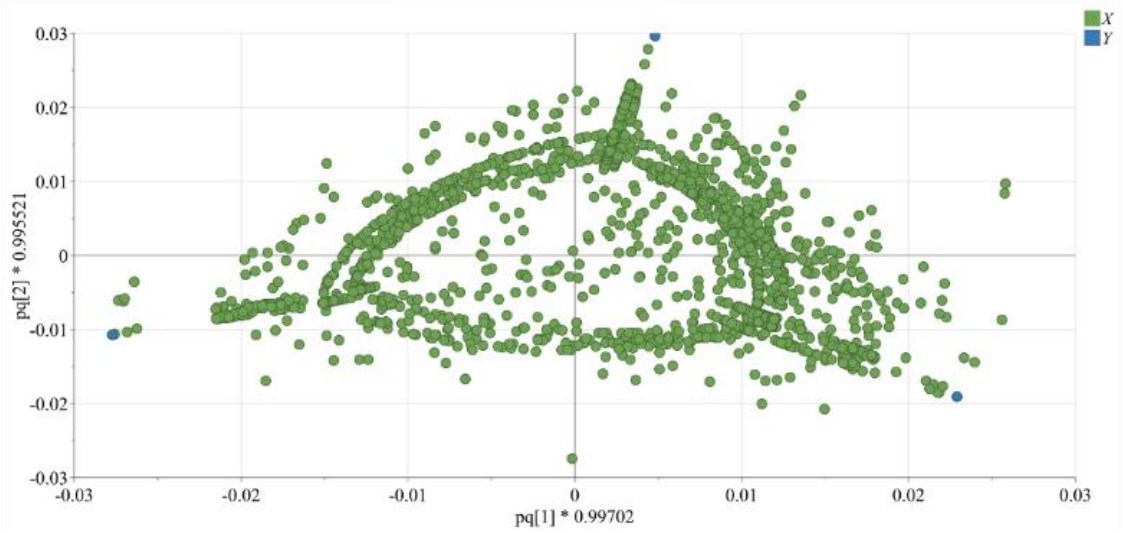
6.6.2. The untargeted analysis of Lemna minor incubated with DCF for 4 days

The metabolic profile of *Lemna* control, 10 & 100 µM DCF was performed using RPLC-HILIC-QTOF-MS/MS resulting in 8936 features. The analytical reliability and accuracy were detected using the internal standards (see Chapter 3). The features (variables) were plotted in the loading scatter plot, which is responsible for the discrimination between the different treatments. They showed which variables (features) were expressed differently between the

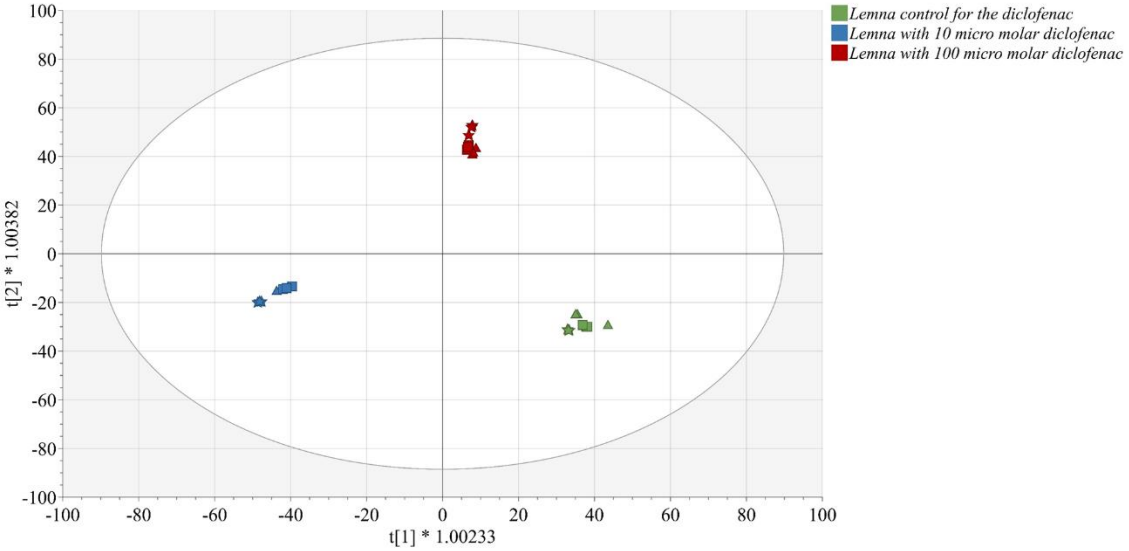
different treatments. The Score plot of OPLS-DA explained 49.5% of the variations in the various extracts ($R^2Y(cum)= 0.992$) with a higher predictive value ($Q^2(cum) = 0.984$) (Fig. 31). The first component (t1) separated the control in the positive part and the 10 µM DCF samples in the negative part. Moreover, the orthogonal component (t2) distinguishes between the control and 10 µM DCF and 100 µM DCF in the negative and positive parts, respectively.

This classification was according to the variables, which were highlighted by OPLS-DA regression and the derived S-plot (Fig. 31d). The variables showing the highest *p* and *p(corr)* values are considered the most relevant metabolites for the classification between samples. The compounds were marked below in (Fig. 31d &e) and considered the differentiating metabolic profile (DMF) of *Lemna minor* due to DCF treatment. The same concept has been applied before to *Phragmites australis* incubated with DCF and CBZ, which showed the changes in its metabolic pathways. This chapter was designed to investigate the phenylpropanoid pathway using the isolated different metabolites.

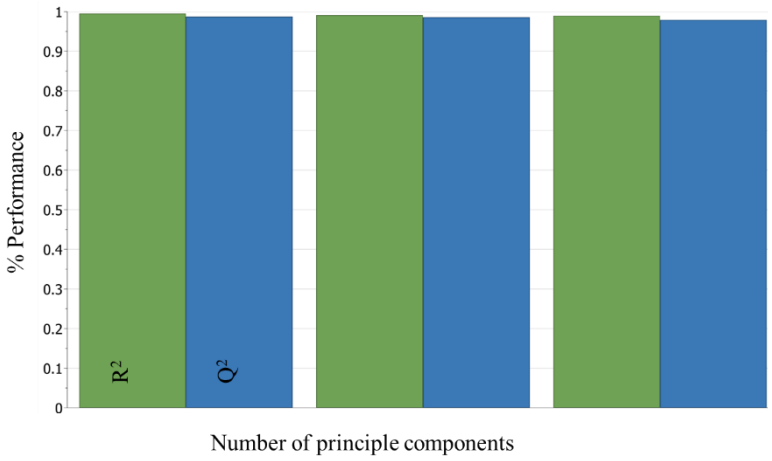
Moreover, the contribution plot of the same study showed that quercetin (phenolic metabolites) was plotted not at the far end such as DCF and CBZ transformed products. Hence, the abundance of apigenin-6-arabinopyranoside-8-glucopyranoside was investigated using the OPLS-DA score plot. Apigenin-6-arabinopyranoside-8-glucopyranoside was identified before using the PLANT-IDENT database. The OPLS-DA score was colored according to the abundance showing an increase in apigenin-6-arabinopyranoside-8-glucopyranoside in 100 µM DCF (Fig. 32). This indicates that there were changes in the flavonoids pathway, which have been detected with system A due to the treatment, not to the fluctuations between the measurement or an error in sample preparation.



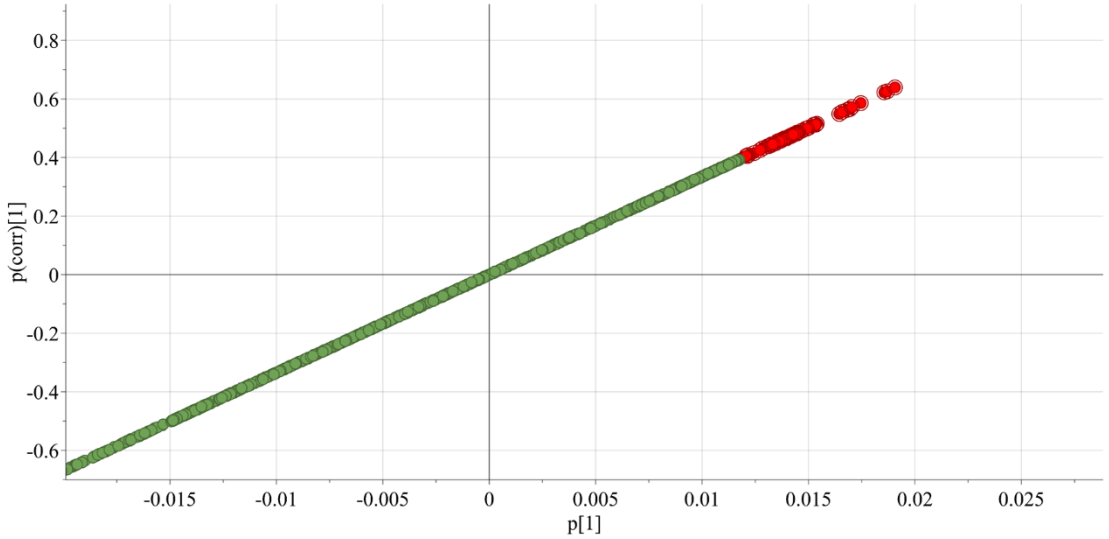
a


Lemna control for the diclofenac
Lemna with 10 micro molar diclofenac
Lemna with 100 micro molar diclofenac
t[2] * 1.00382
t[1] * 1.00233

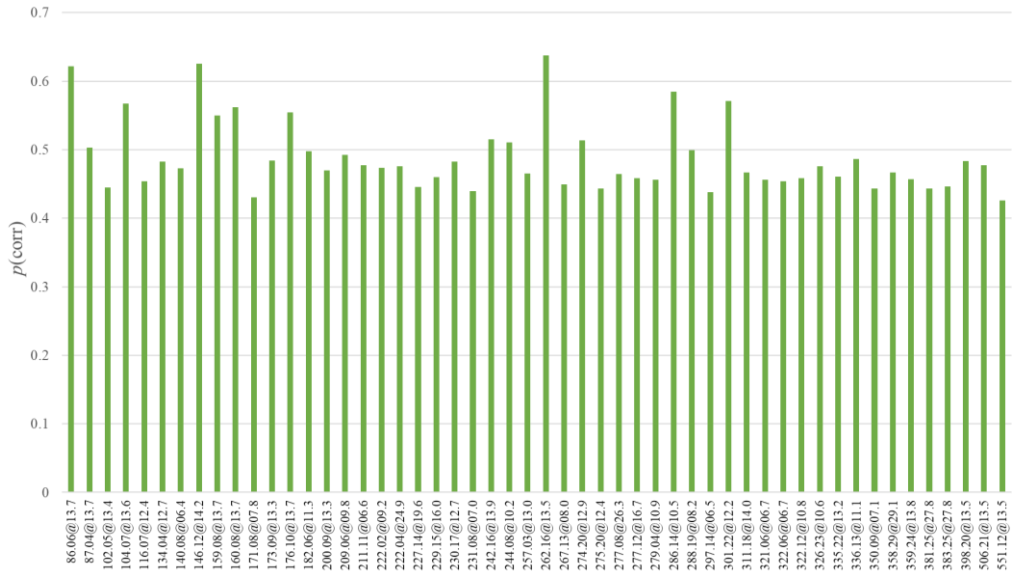
b


% Performance
R2
Q2
Number of principle components

c



d



e

Figure 31. (a) The loading plot displays the relation between the different *Lemna minor* samples and the chosen metabolite; The variables were plotted according to the first principal component (t1) and the orthogonal component (t2); (b) The OPLS-DA score plot of different *Lemna minor* extracts control as well as incubated with, 10, and 100 µM diclofenac, individually with a confidence limit of 95% discriminating. The triangles represent 100% MeOH extracts, the squares represent 50% MeOH extracts, and the stars represent 100% H_2O extracts, respectively. The green color represents the control, the blue color represents 10 µM diclofenac, and the red color represents 100 µM diclofenac groups, respectively. Each symbol represents one observation of *Lemna minor*; (c) the Q^2/R^2 Overview plot displays the individual cumulative R^2 (green columns) and Q^2 (blue columns) for the goodness of fits and cross-validation parameters; (d) S-plot of *Lemna minor* control and incubated with 10 and 100 µM diclofenac samples. The S-plot provides the visualization of the loading components of OPLS-DA to enable the interpretation of the data. The red-labeled compounds represent the differentiating

metabolic profile (DMF) of each incubation; (e) The chart of DMF with the corresponding p (corr) values.

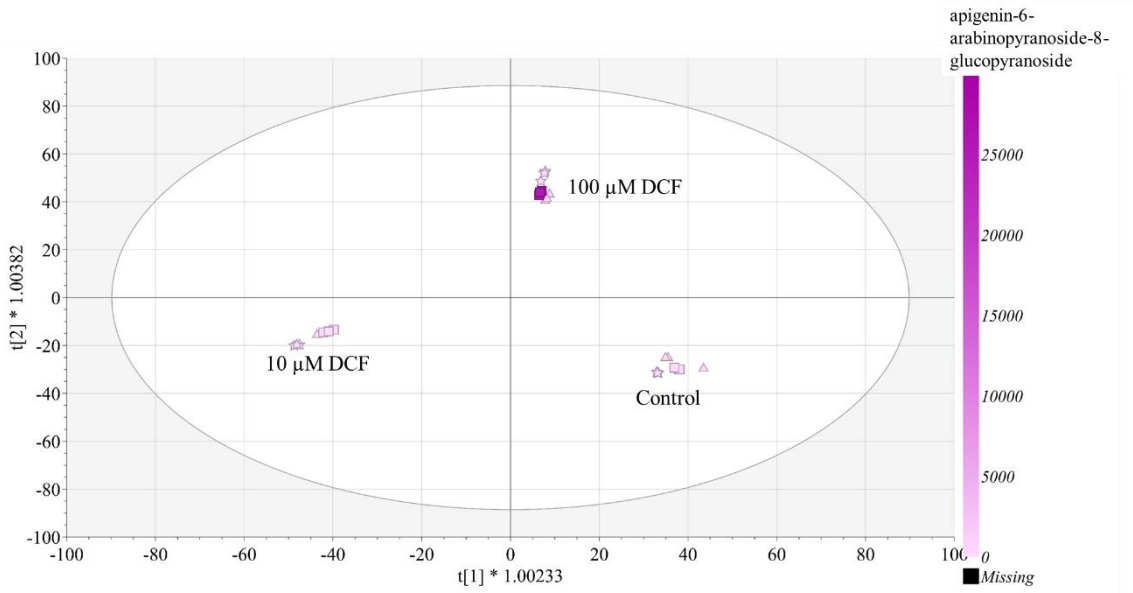


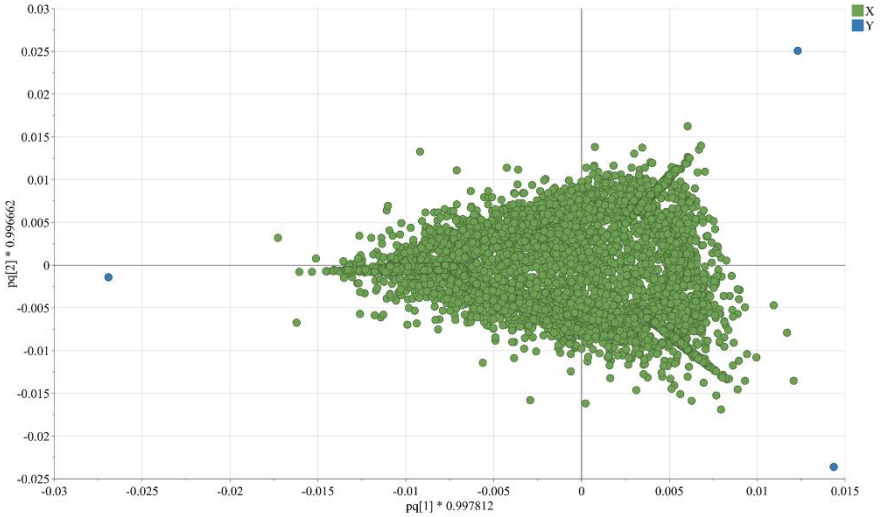
Figure 32. The OPLS-DA score plot of different *Lemna minor* extracts control as well as incubated with, 10, and 100 µM diclofenac, individually with a confidence limit of 95% discriminating. The triangles represent 100% MeOH extracts, the squares represent 50% MeOH extracts, and the stars represent 100% H_2O extracts, respectively. The different samples were colored according to the abundance of apigenin-6-arabinopyranoside-8-glucopyranoside in a gradient manner. The darker the color is the higher abundance. The black color represents the abundance equal to zero.

6.6.3. The untargeted analysis of Lemna's metabolic profile after incubation with DCF for 4 8, and 12 days

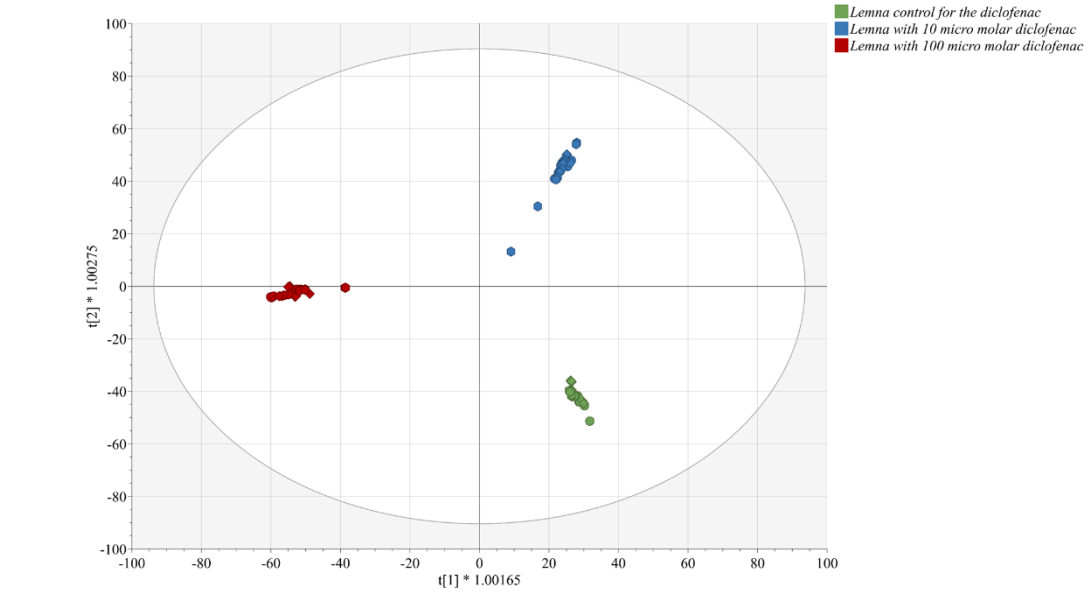
After spiking *Lemna minor* with 10 and 100 µM DCF, individually, it was harvested at the fourth, eighth and twelfth day of exposure. Each time point was extracted with the three different solvents and analyzed with RPLC-HILIC-QTOF-MS/MS. The metabolic profile consisted of the three extracts features (34122) after excluding the background ones (Fig. 33a). The data was analyzed using SIMCA 17 software. The OPLS-DA sore plot ware explained the variations of the different exposure periods ($R^2Y(cum)=0.984$) with a higher predictive value ($Q^2(cum)= 0.969$) (Fig. 33b). The data was clustered into 3 clusters control, 10 µM DCF, and 100 µM DCF. Also, the distances between the control and 10 µM DCF were increased, however, both treatments were more correlated than 4 days exposure. This indicates that the DCF affected the *Lemna* metabolic profile depending on the exposure period. Hence, it was expected that there were changes in the *Lemna* metabolic pathways with the increase in the exposure period. Therefore, apigenin-6-arabinopyranoside-8-glucopyranoside was investigated statistically in *Lemna minor* exposed for DCF for 4,8,12 days (Fig. 33c). Apigenin-6-arabinopyranoside-8-glucopyranoside was increased in *Lemna* incubated with 100 µM DCF

and with 10 µM DCF for 4 and 8 days, respectively, which indicated that *Lemna* metabolic system is resistant to ROS species.

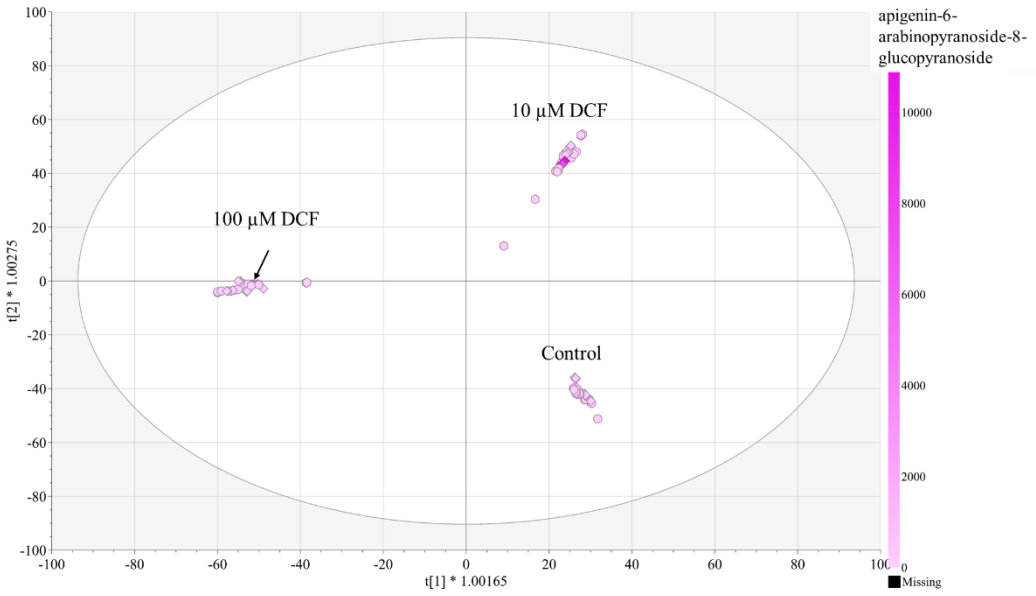
Moreover, the normalized diclofenac concentration in *Lemna minor* metabolic profiles were listed in (Fig. 34). The results showed that *Lemna* could uptake and accumulate the DCF. The ability of *Lemna* to accumulate DCF was decreased by 32% after 12 exposure days at 100 µM. The longer the duration of exposure, the lower was the detectable DCF concentration, and the lower were the differences between the two concentrations of DCF. The same was shown in *Populus alba* L. incubated with DCF at 1 mg/L (Pierattini *et al.*, 2018). However, in 10 µM, *Lemna* preferred to transform diclofenac over bioaccumulation. This might be due to the small concentration of DCF which could be handled by the *Lemna* metabolic. Moreover, the abiotic stress caused by DCF exposure triggers systemic responses such as long-distance electrical and reactive oxygen species (ROS)waves such as in pathogenic infections and wounding (Pierattini *et al.*, 2018). Thus, DCF was expected to upregulate *Lemna's* antioxidative defense



a



b



c

Figure 33. (a) The loading plot displays the relation between the different *Lemna minor* samples at different exposure periods; (b) The OPLS-DA score plot of different *Lemna minor* samples control as well as incubated with, 10, and 100 µM diclofenac at 4, 8, and 12 days, individually with a confidence limit of 95% discriminating. The circles represent 4 treatment days, the diamonds represent 8 treatment days, and the hexagons represent 12 treatment days, respectively. The green color represents the control, the blue color represents 10 µM diclofenac, and the red color represents 100 µM diclofenac groups, respectively. Each symbol represents one observation of *Lemna minor*; (c) The OPLS-DA score plot of different *Lemna minor* samples control as well as incubated with, 10, and 100 µM diclofenac at 4, 8, and 12 days, individually with a confidence limit of 95% discriminating. The different samples were colored according to the abundance of apigenin-6-arabinopyranoside-8-glucopyranoside in a gradient manner. The darker the color is the higher abundance. The black color represents the abundance equal to zero. The arrow shows the increase in *Lemna* incubated with 100 µM for 4 days.

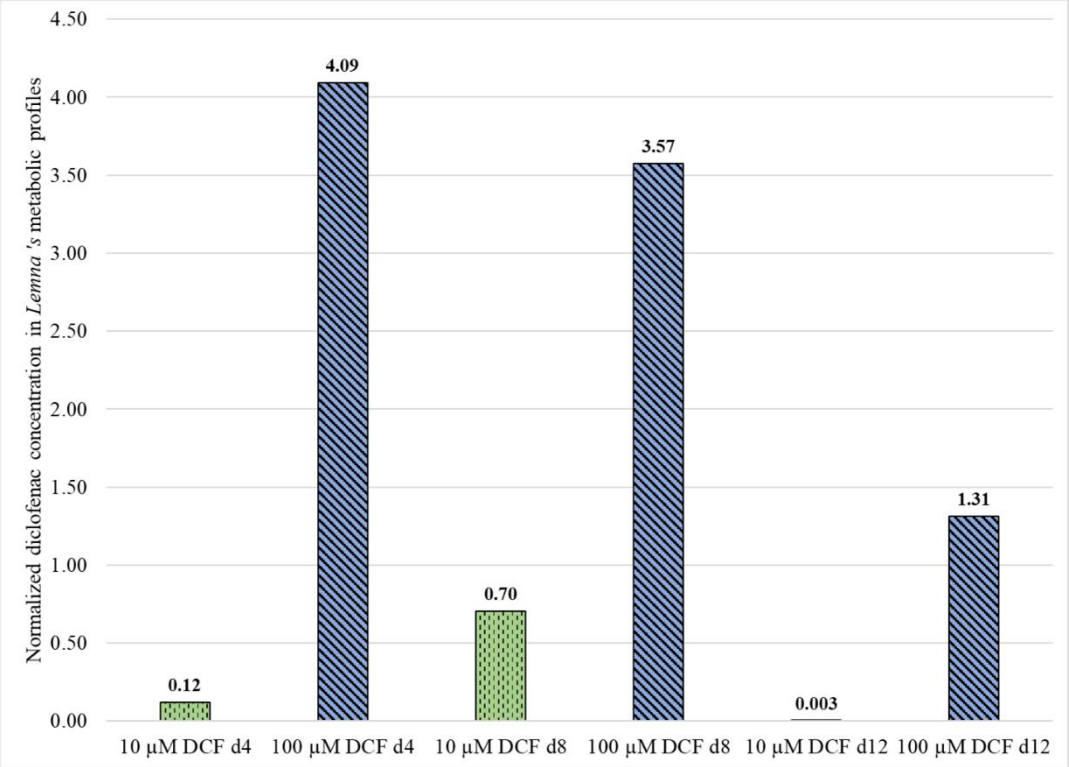


Figure 34. The chart of normalized diclofenac concentration in *Lemna's* metabolic profile incubated with DCF at 10 and 100 µM for 4, 8, and 12 days (d).

6.6.4. The target analysis of Lemna minor flavonoids

The previous results of *Lemna* with system A, in addition to the statistical data analysis, were used to determine the target analysis compounds list. The target analysis was performed in *Lemna minor* metabolic profiles for the following metabolites: acacetin, apigenin, galangin, myricetin, naringin, apigenin-7-*O*-glucoside, chrysoeriol, isoorientin, apiin, umbelliferone, isovitexin, quercetin-3-*O*-glucoside, tricin, cyanin, dihydrokaempferol, dihydromyricetin, apigenin-6-arabinopyranoside-8-glucopyranoside, kaempferol-7-*O*-rhamnoside, naringenin-7-*O*-glucoside, and 4-methoxycinnamic acid were identified in *Lemna minor* metabolic profile using the PLANT-IDENT database (chapter 4). Their relative percentage against control was determined for each compound after normalized their peak area against the internal standard, which was eluted with the same column (i.e., for the RPLC column eluted compounds, chloridazon was used (eluted at 27.7 min) and vice versa). The relative percentages against the control were listed in (Fig. 35&36).

The precursor of the phenylpropanoid biosynthesis pathway is phenylalanine, which was increased in *Lemna* incubated with DCF. Thus, the pathway was investigated. The pathway is divided to synthesize the simple (such as organic acids and phenolic) and polyphenolic compounds (e.g. flavonoids and stilbenes). The first compound coming from phenylalanine in cinnamic acid. The 4- methoxycinnamic acid was identified in *Lemna* incubated with 10 and 100 µ DCF and showed an increased pattern in both concentrations because the 4-methoxycinnamic acids bound tightly to CYP199A4 and were better substrates for CYP199A4

than cinnamic acid itself to synthesis the phenylpropanoid. The activity of the cytochrome P450 family increases with DCF incubation (Huber *et al.*, 2012).

Regarding the simple phenolics, umbelliferone is a coumarin, was decreased in *Lemna* incubated with 10µ DCF, however, was increased in higher concentration (Fig. 35&36). The same was observed in (section 6.3.4) in the untargeted study of *Lemna*, which was analyzed with system A. The previous results showed an increase in the 7-hydroxy-4-methoxy-5-methyl coumarin. Coumarins have antioxidant properties protecting the plants against pathogens infections through ROS scavenging role and restrict cell death (Stringlis *et al.*, 2019). Coumarin biosynthesis inside the plant is activated with cytochrome P450 (CYP) enzyme, which was induced with DCF in incubated *Lemna* (Huber *et al.*, 2012). It was reported that the exposure of roots of chamomile (*Matricaria chamomilla*) to salicylic acids for 24 hrs. has resulted in the accumulation of coumarins umbelliferone and herniarin in the leaves (Stringlis *et al.*, 2019). Moreover, coumarins were released as a response to iron and phosphate deficiency in *Arabidopsis* (Chutia *et al.*, 2019). Therefore, DCF caused a deficiency in the minerals uptake in *Lemna minor*.

The first compound of the flavonoids pathway is naringenin, which is the general precursor for flavonols, anthocyanins, proanthocyanidins, flavones, and isoflavones. The naringin and naringenin-7-*O*-glucoside were identified in *Lemna,* which was synthesized from naringenin. Naringenin-7-*O*-glucoside was increased in both treatments, however, with a higher rate in 10 µM DCF than 100 µM DCF (Fig. 35&36). Moreover, naringin is more or less equal to the control in10 µM DCF, however, it was increased in 100 µM DCF. This indicates the increase in naringenin and its glucosides in *Lemna* exposed to DCF, consequently, and enhanced the flavonoids pathway.

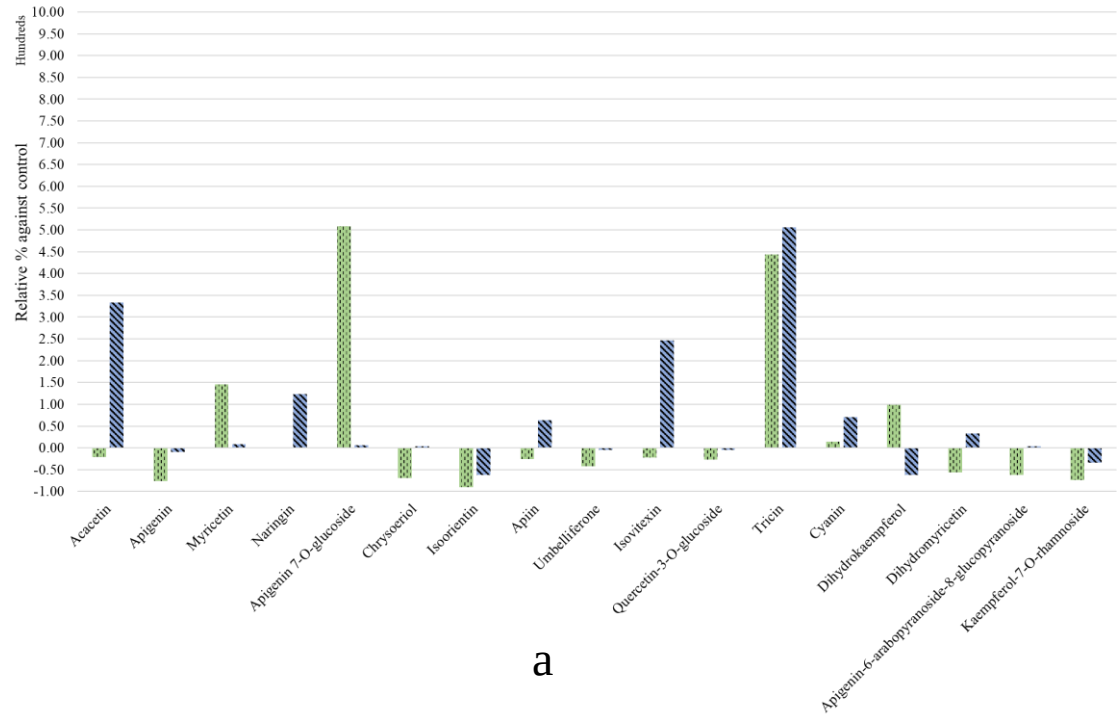
Then, the pathway is branched to synthesis flavone and flavonol. The first compound in the flavone class is apigenin which synthesizes from naringenin by flavone synthase I or flavone synthase II. Apigenin, apiin, and apigenin-6-arabopyranoside-8-glucopyranoside were decreased in 10 µM DCF and increased in 100 µM DCF (Fig. 35&36) which seemed to be due to an increase in the activity of flavone synthase I and II. Apiin showed a remarkable increase when compared with others. However, in 10 µM DCF, apigenin- 7-*O*-glucoside was increased. Interestingly, the vicenin 1 was increased in *Lemna* with 10 µM DCF, which was shown by the untargeted analysis. This indicated the increase in the activity of glycosyltransferase enzymes. DCF caused limitations in electron transport in *Lemna minor,* which might be caused by the reduction in flavonoids at 10 µM DCF (Hájková *et al.*, 2019). Moreover, other flavones (acacetin, isovitexin, isoorientin, and chrysoeriol), synthesize from apigenin, were decreased in 10 µM DCF, however, increased in 100 µM DCF (Fig. 35&36). Moreover, apigenin-6,8-diglucopyranoside (isovitexin 8-*C*-β-glucoside), which derives from an isovitexin, showed the same pattern of isovitexin.

Tricin showed an increase in both concentrations with an approximately equal value. Tricin plays a vital role in cell wall lignification as a nucleation site for lignification (Tohge *et al.*, 2017). It was reported that tricin and *O*-methyltransferases enzyme were increased in wheat due

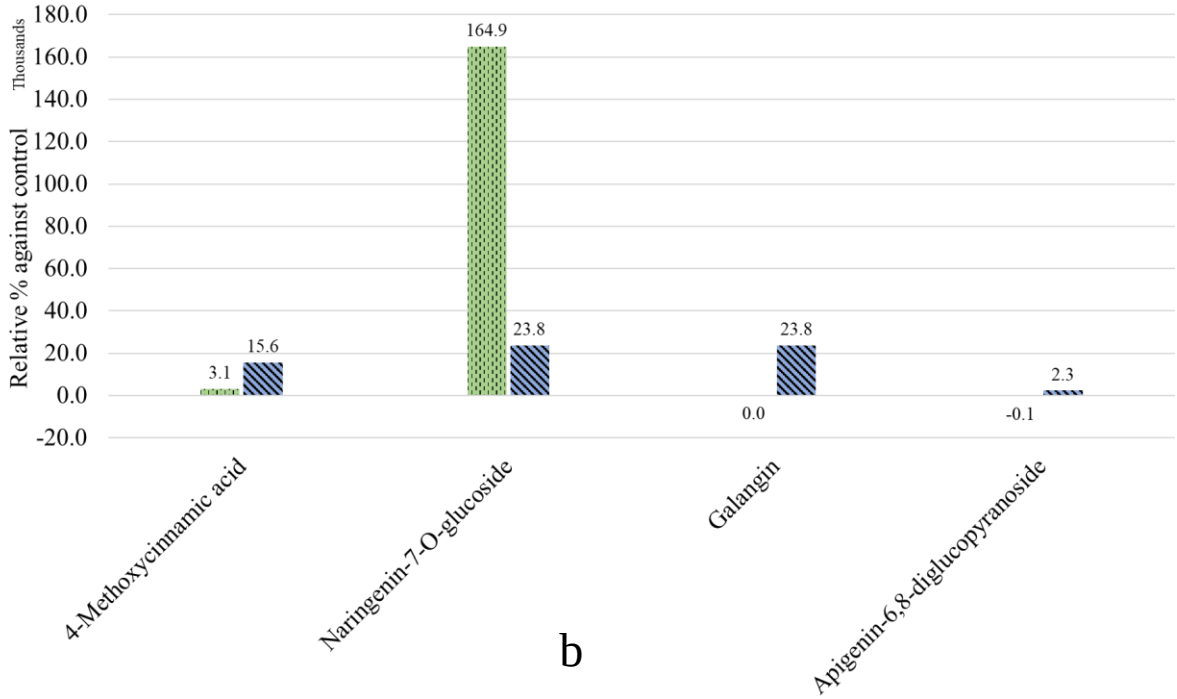
to abiotic stresses (e.g. cold, salt and drought) to protect the plant and enhance its defense mechanism (Moheb *et al.*, 2013). The results were compatible with the apigenin results and showed the same pattern of precursor, intermediate, and final compounds.

On the other hand, the flavonol biosynthesis starts with dihydrokaempferol, which synthesizes from naringenin by flavanone 3-hydroxylase. Dihydrokaempferol was increased in 10 µM DCF, however, a decrease in 100 µM DCF. This led to a decrease of galangin, kaempferol-7-*O*-rhamnoside, quercetin-3-*O*-glucoside in *Lemna* incubated with 10 µM DCF, besides, to increase their relative % in 100 µM DCF. Moreover, dihydrokaempferol was converted to dihydroquercetin by dihydroflavonol 4-reductase, which is further converted to cyanin by leucoanthocyanidin dioxygenase/anthocyanidin synthase. Cyanin was increased in both concentrations, which might lead to the enhancement of the anthocyanin pathway (see section 6.3.5).

Further, flavonoid F3′5′H-hydroxylase catalyzes the hydroxylation of the C3′ and C3′/C5′ positions of dihydrokaempferol to synthesize dihydromyricetin. The dihydromyricetin showed the same pattern as galangin, kaempferol-7-*O*-rhamnoside, quercetin-3-*O*-glucoside. Myricetin is synthesized from dihydromyricetin using the flavonol synthase enzyme, which belongs to oxidoreductases. Also, DCF activates the oxidoreductase activity in *Lemna* (Alkimin *et al.*, 2019). Therefore, myricetin increased in 10 µM DCF (Fig. 35&36). However, it decreased in 100 µM DCF treatment with the accumulation of dihydromyricetin. This might be due to the loss of oxidoreductase activity at higher DCF concentrations. Also, flavone synthase II, Flavonoid 3′-hydroxylase (F3′H) and flavonoid F3′5′H-hydroxylase are cytochrome P450 family member which was increased with DCF incubation. Moreover, the DCF seems to activate the glycosyltransferases (GTs), and methyltransferases. Also, flavonoids protect the plants against the light (UV) stress factor, which induces the synthesis of B-ring-substituted forms such as kaempferol, quercetin, and myricetin which perform antioxidant roles, thus contributing to ROS-detoxification (Petrussa *et al.*, 2013). Additionally, the incubation of *Lemna* with DCF affected the chlorophyll contents. Hence, *Lemna* enhances the flavonoids pathway as an indirect response to DCF incubation. Concluding, DCF enhanced the flavanoids and anthocyanin pathways in *Lemna minor* as a direct and indirect response. The type of metabolites was depended on the dosage of DCF. Thus, *Lemna minor* could resist to the stress effect of DCF.



a



b

Figure 35. Relative intensities (%) of different metabolites identified in *Lemna minor* incubated with 10 & 100 µM DCF against the control.

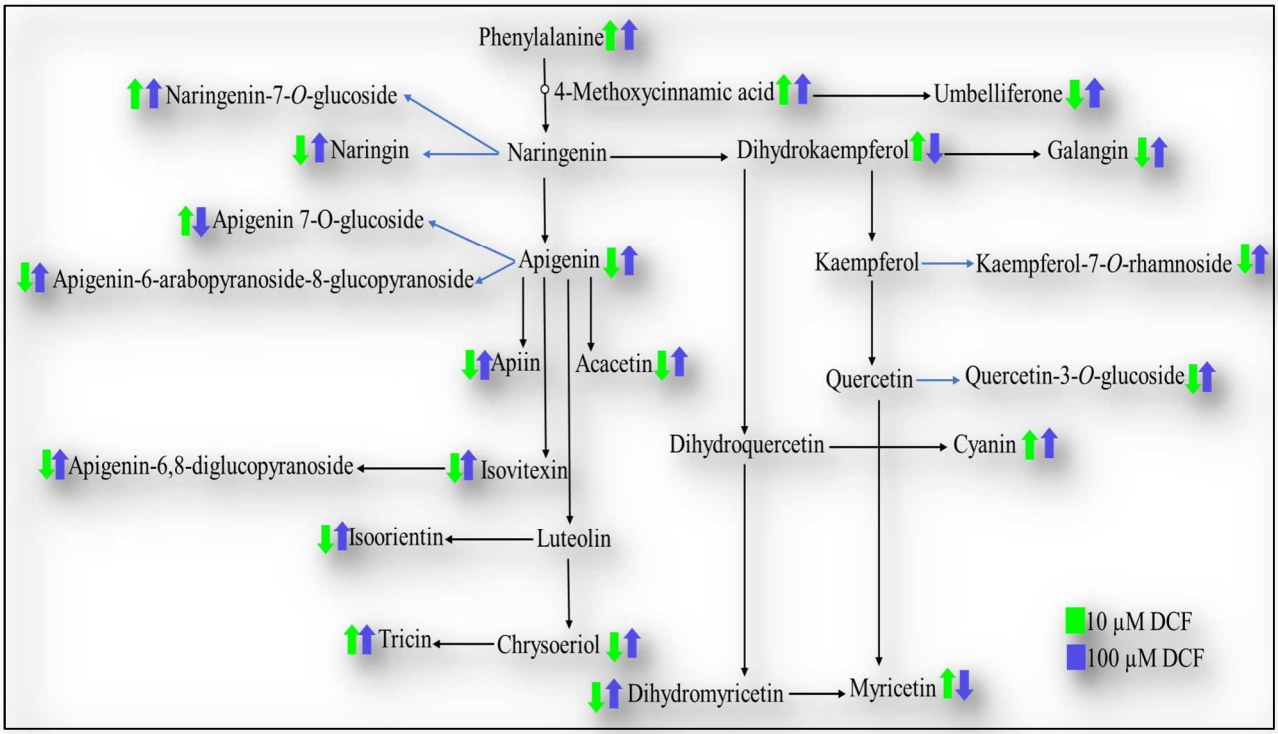


Figure 36. The flavonoids biosynthesis pathways in Lemna minor*; the green and blue color represent the* Lemna *incubated with 10 µM DCF and 100 µM DCF, respectively.*

6.6.5. The flavonoids pathway in Lemna minor incubated with DCF for 8 and 12 days

The flavonoids, as well as their precursors, were investigated in the *Lemna* incubated with 10 and 100 µM DCF for 8 and 12 days. Each treatment was compared with its corresponding control to determine the relative intensities of the compounds.

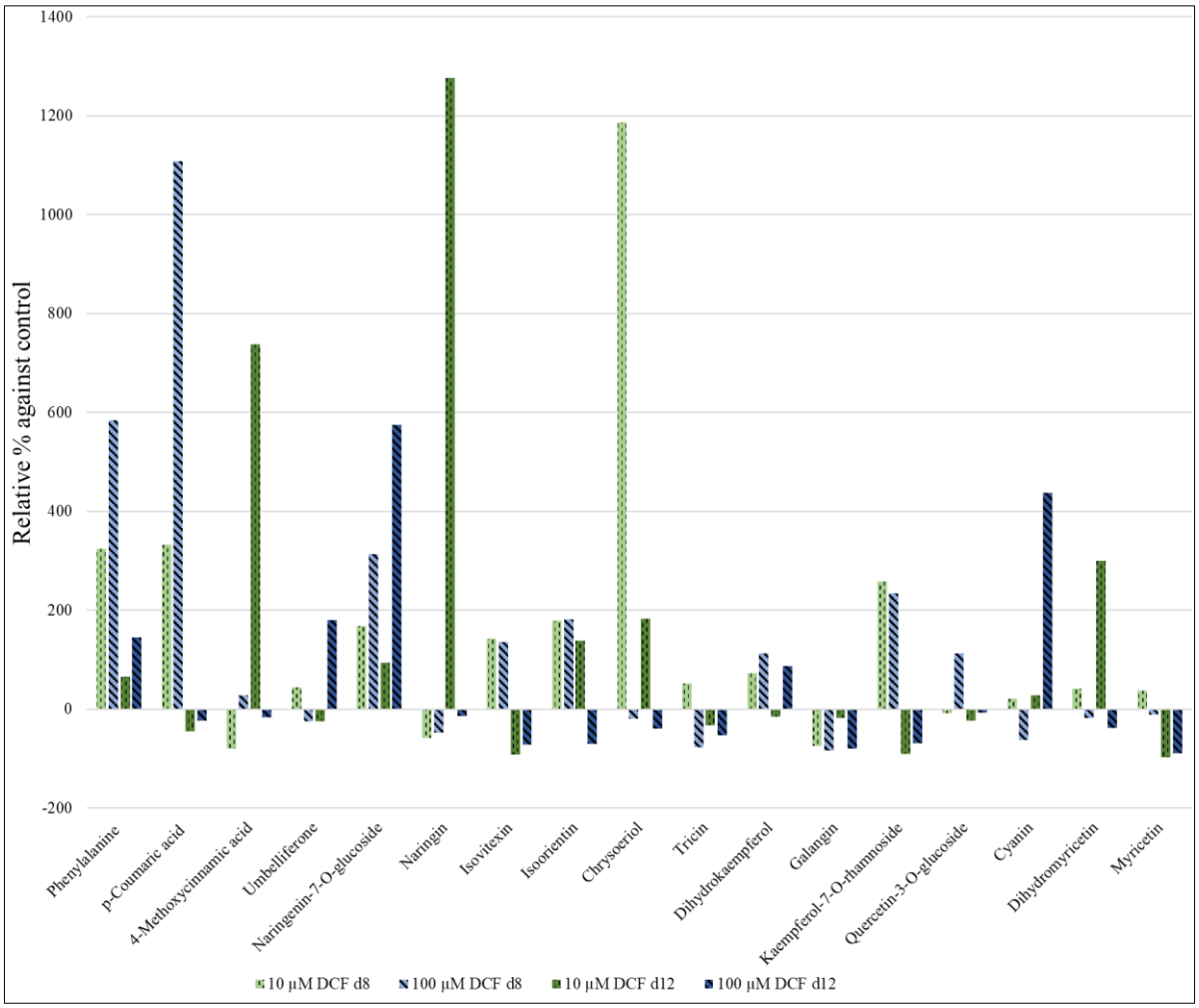
The results showed that *Lemna* continuously synthesizes phenylalanine, naringenin 7-*O*-glucoside with a higher rate than control in both exposure periods at 4 days; however, the first was increased in incubation for 8 days, and the latter in incubation for 12 days (Fig. 37a). Phenylalanine is deaminated to form the cinnamic acid which syntheses the *p*-coumaric acid. 4-Methoxycinnamic acid showed a decrease in 10 and 100 µM DCF on the 8th and 12th days, respectively, however, it increased in the 100 µM DCF of the first and 10 µM DCF of the second. In addition, 4-methoxy cinnamic acid is approximately equal to the control in the 100 µM DCF on the 12th day. For *p*-coumaric acid, an increase in the relative intensity after the incubation for 8 days is recorded with higher intensity in 100 µM DCF. However, it decreased on the 12th exposure day. *p*-Coumaric acid was dependent on phenylalanine concentration because it synthesizes from cinnamic acid, which synthesis from phenylalanine. Thus, the decrease in phenylalanine was accompanied by a decrease in *p*-coumaric acid. Therefore, they are positively correlated and the decrease in phenylalanine was expected to show a decrease in *p*-coumaric acid. Hence, the results were due to the effect of DCF on *Lemna* metabolic

pathways. Umbelliferone was increased significantly in *Lemna* incubated with 100 µM DCF for 12 days. Consequently, *Lemna* induced the simple phenolic synthesis after incubation for a long time.

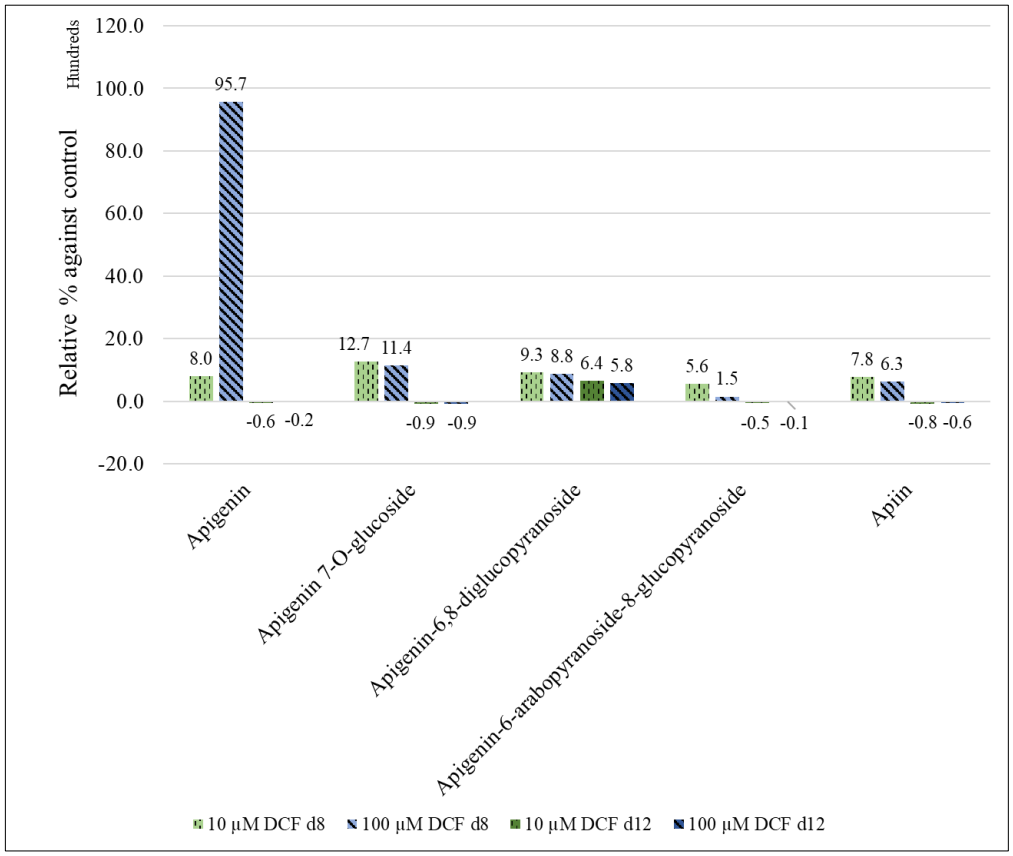
Further, the naringenin-7-*O*-glucoside decrease in 10 µM DCF on the 8th day was accompanied by a significant increase in naringin. This indicated the enhancement of the pathways as naringenin (the aglycone of both of them) is the general precursor for flavonols, anthocyanins, proanthocyanidins, flavones, and isoflavones. Furthermore, apigenin, apigenin-7-*O*-glucoside, apiin, and apigenin-6-arabopyranoside-8-glucopyranoside were increased at eighth exposure days for both concentrations. However, they were decreased on the 12th exposure days for both concentrations (Fig. 37b).

All of them except apigenin were higher in 10 µM DCF than 100µM DDCF. Interestingly, apigenin-6,8-diglucopyranoside was increased on the 8th and 12th exposure days, as well as at both concentrations. Isovitexin also was increased and decreased in incubation for 8 and 12 days, respectively. However, it had a higher concentration in 100 µM DCF in incubation for 12 days. Isoorientin also showed the same pattern as isovitexin except for the increase in its intensity at 10 µM DCF after 12 days of incubation. The chrysoeriol and tricin were directly proportional because tricin is synthesized from chrysoeriol. The dihydrokaempferol was increasing in all treatments except at 10 µM DCF after incubation for 12 days, its intensity is approximately equal to the control. Interestingly, galangin were decreased in all treatments. On the other hand, kaempferol-7-*O*-rhamnoside was increased and decreased after incubation for 8 and 12 days, respectively. The increase and decrease were higher in 10 µM DCF. Further, quercetin-3-*O*-glucoside showed an intensity less or equal to the control for all the treatments except at 100 µM DCF, which had higher intensity (Fig. 36a). Dihydromyricetin and cyanin showed the same pattern except at the 100 µM DCF after incubation for 12 days. This might due to the activation of another pathway and/or changes in the intensity of leucocyanidin. The increase and decrease in dihydromyricetin were accompanied by the same pattern in myricetin except 10 µM DCF after incubation for 12 days (Fig. 37a).

Concluding, the increase in the treatment period enhances the flavonoids pathway; hence increase the activities of the responsible enzymes. The precursors (phenylalanine, 4-methoxycinnamic acid, *p*-coumaric acid, and naringenin) intensities were started to fade after incubation of 12 days when compared to the 8 days of incubation. After a long period of 12 days of incubation, most of the flavonoids were decreased. Thus, *Lemna's* metabolism was induced to synthesize more of the simple phenols. Their synthesis is less complicated as it requires lower energy and less involved enzymes and cell organelle, which could be the reason. Also, *Lemna* metabolic system was exhausted to degrade and accumulate the higher concentration of DCF. Moreover, the DCF activation to the glycosyltransferases and methyltransferases seems to be decreased with increasing the incubation period. *Lemna's* metabolic system was resistant until the 12th day of treatments; however, it showed a decrease in the final products.


1400
1200
1000
800
600
400
200
0
-200
Relative % against control
Phenylalanine
p-Coumaric acid
4-Methoxycinnamic acid
Umbelliferone
Naringenin-7-O-glucoside
Naringin
Isovitexin
Isoorientin
Chrysoeriol
Tricin
Dihydrokaempferol
Galangin
Kaempferol-7-O-rhamnoside
Quercetin-3-O-glucoside
Cyanin
Dihydromyricetin
Myricetin
10 µM DCF d8
100 µM DCF d8
10 µM DCF d12
100 µM DCF d12

a



b

Figure 37. Relative intensities percentage (%) of different metabolites identified in *Lemna minor* incubated with 10 & 100 µM DCF for 8 and 12 days against the control; (a) precursors and identified flavonoids; (b) apigenin and its derivatives.

6.6.6. Glucose profile in Lemna minor different samples

The untargeted analysis of *Lemna minor* incubated with DCF for 4 days showed an increase in the hexose and ribose concentrations as well as flavonoid glycosides. Thus, glucose as one of the hexoses in plants was investigated in *Lemna minor* different samples. The glucose was identified in the aqueous extracts of all treatments using the reference standard with variation in the mass and RT equal to -1.41 ppm and less than -0.1 minutes, respectively.

The relative intensities were determined in all samples after being normalized against the internal standard. The results showed a decrease in all the samples at different time points. The order was as follows 100 µM DCF d8 < 100 µM DCF d12 < 10 µM DCF d4 < 100 µM DCF d4 < 10 µM DCF d12 < 10 µM DCF d8 (Fig. 38). The decline in sugars after plant exposure to DCF indicates an increase in energy consumption, which was reported in Lettuce after exposure to DCF (Hurtado *et al.*, 2017). After *Betula pendula* exposure to ozone, the total sugars were increased, and the phenolic glucosides contents were reduced. They were irreversibly proportional (Lavola *et al.*, 1994). The untargeted analysis showed high content in hexose in 10 µM DCF. Also, the targeted analysis showed an increase in the glycosylic flavonoids in 100 µM DCF, which have more solubility than the aglycone to facilitates their transportation to the cell membrane, as well as prevents auto-oxidation of hydroxyl groups (Nabavi *et al.*, 2020).

Thus, *Lemna* utilized the free sugars to synthesis the phenolic glucosides in DCF incubation. This was the reason for the glucose decrease in 100 µM DCF. Flavonoids activate the Phase II enzymes such as glutathione-*S*-transferase (Moon *et al.*, 2006), which have been reported to be increased due to DCF incubation in *Lemna minor* (Alkimin *et al.*, 2019).

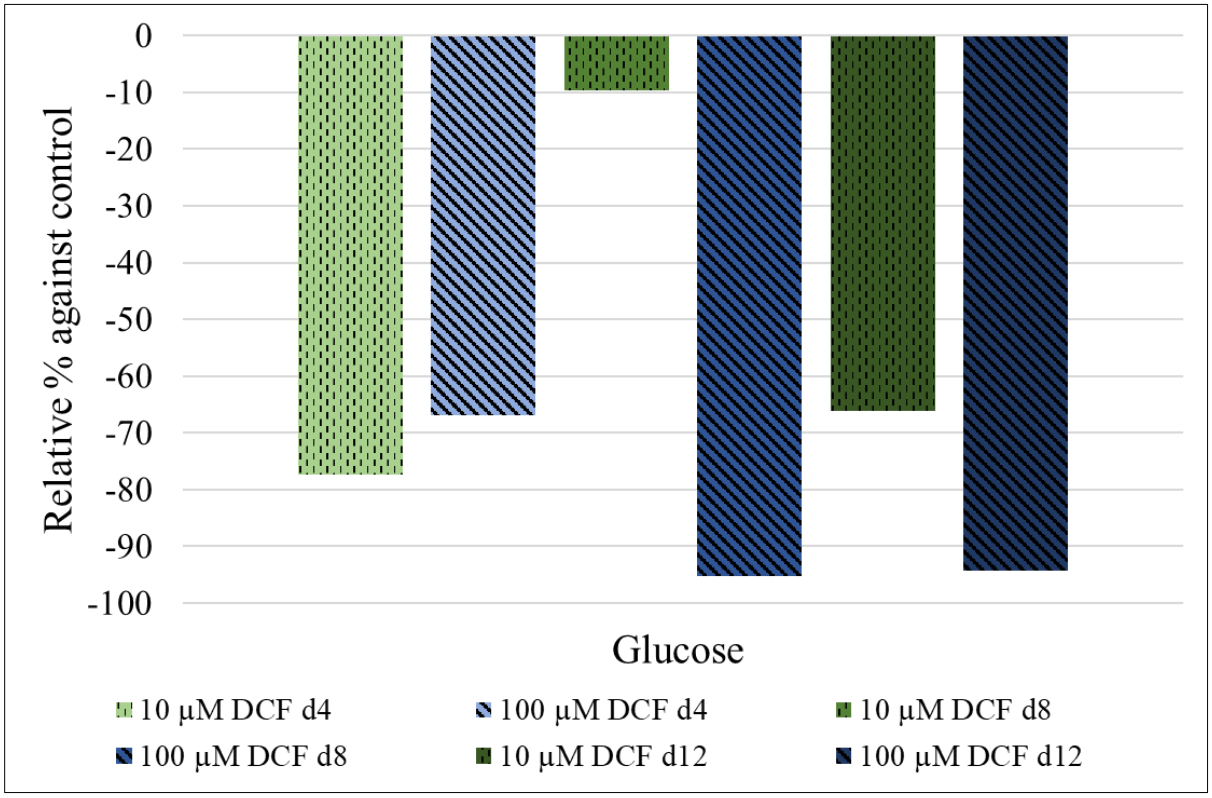


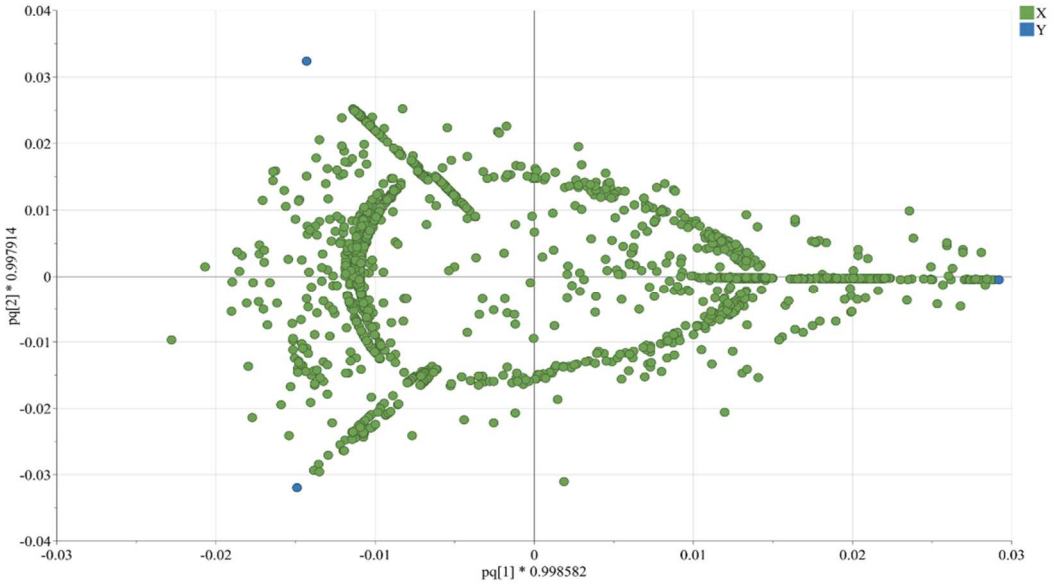
Figure 38. . Relative intensities percentage (%) of glucose in the different *Lemna minor* samples incubated with 10 & 100 µM DCF for 4, 8, and 12 days against the control

B- Carbamazepine

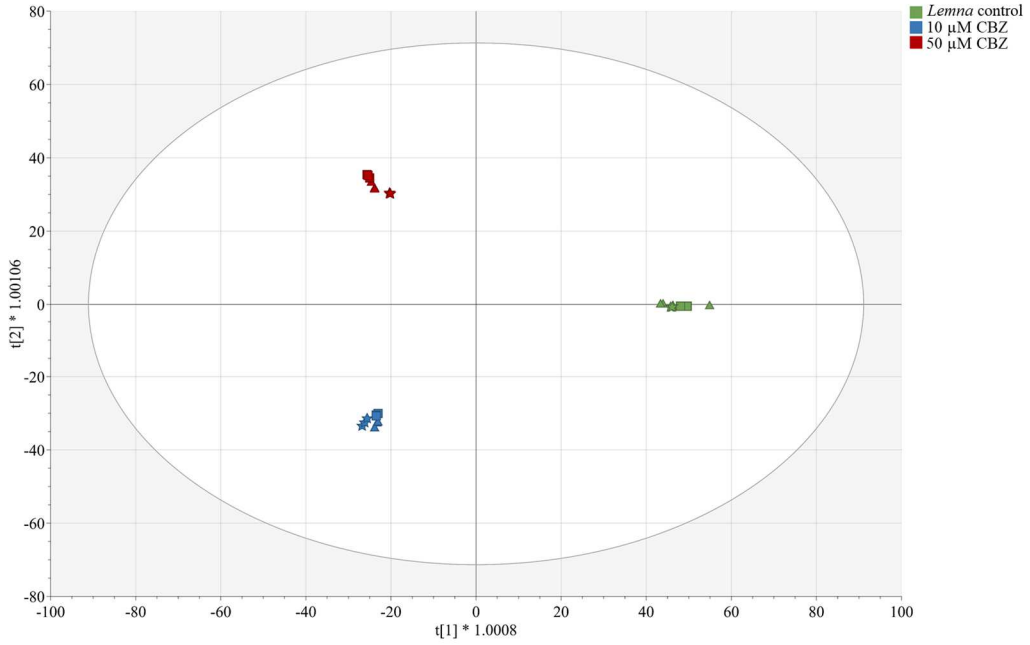
6.6.7. The untargeted analysis of Lemna minor incubated with CBZ

Lemna minor was incubated with 10 and 50 µM CBZ, individually. *Lemna* control, 10 µM, and 50 µM CBZ were analyzed with RPLC-HILIC-QTOF-MS/MS in triplicate. The features were obtained using the MarkerView Software (version 1.3.1). The background features were deleted from the corresponding extract. This resulted in 7456 features, which are organized in a data access file to upload into SIMCA 17 software. The data was investigated using multivariate statistical analysis. The features were plotted in the loading plot (Fig. 39a), which were the variables used to build the OPLS-DA. The OPLS-DA was able to explain 54.3% of the variations between the control, 10 and 50 µM CBZ samples (R^2Y (cum) = 0.996) with a higher predictive value (Q^2 (cum) = 0.986) (Fig. 39b). The first component (t1) separated the control in the positive part and the treatment samples in the negative part. Moreover, the orthogonal component (t2) discriminated between the two concentrations of CBZ. Thus, *Lemna* incubation with CBZ caused statistically significant changes in its metabolites (variables), which could discriminate between the control and treatments. These changes could be due to

an increase or decrease in the intensities of *Lemna* metabolites and/or the CBZ degraded products in the incubated samples. Hence, some metabolite intensities were investigated in *Lemna* control, incubated with 10 and 50 µM CBZ.



a



b

Figure 39. (a) The loading plot displays the relation between the different *Lemna minor* samples; (b) The OPLS-DA score plot of different *Lemna minor* samples control as well as incubated with, 10, and 50 µM carbamazepine for 4 days, individually with a confidence limit of 95% discriminating. The triangles represent 100% MeOH extracts, the squares represent 50% MeOH extracts, and the stars represent 100% H_2O extracts, respectively. The green color represents the control, the blue color represents 10 µM carbamazepine, and the red color represents 50 µM carbamazepine, respectively. Each symbol represents one observation of *Lemna minor*.

6.6.8. Lemna's metabolites investigations in control and incubated samples

Different reference standards were analyzed with RPLC-HILIC-QTOF-MS/MS in positive mode. Their intensity was normalized against the internal standards. Then the relative percentage against the control was calculated and shows in (Fig. 40). Phenylalanine and *p*-coumaric acid were directly proportional. They increased in 10 µM CBZ and decreased at higher concentrations. Further, 4- methoxycinnamic acid was decreased in both treatments. Furthermore, naringin and naringenin-7-*O*-glucoside were increased at higher concentrations. Thus, at 50 µM CBZ, CBZ enhanced the *Lemna* flavonoid pathways. However, vitexin was increased and decreased in 10 and 50 µM CBZ, respectively. Vitexin has two pathways from naringenin or apigenin. The apigenin was decreased to half at CBZ's higher concentration. This might indicate that *Lemna* synthesized the vitexin through the apigenin. Further, apigenin-6-arabopyranoside-8-glucopyranoside and acacetin were increased at 10 and 50 µM CBZ. Furthermore, apiin, apigenin-6,8-diglucosides, and isoorientin were decreased and increased in 10 and 50 µM CBZ, respectively. However, apigenin-7-*O*-glucoside was decreased at both concentrations. Furthermore, chrysoeriol and tricin were remarkably increased in 50 µM CBZ, however, the first was decreased and the latter was increased at 10 µM CBZ, respectively. Regarding the flavonol, dihydrokaempferol was increased at 10 µM CBZ, which was decreased with 5.9 fold in 50 µM CBZ than 10 µM CBZ. Further, galangin, kaempferol-7-*O*-rhamnoside, quercetin-3-*O*-glucoside, dihydromyricetin, and myricetin were increased in *Lemna* incubated with 10 and 50 µM CBZ, except galangin and myricetin were approximately equal to control at 10 µM CBZ. Also, peonidin (anthocyanin) was increased in *Lemna* incubated with 10 and 50 µM CBZ. Finally, the increase of flavonoids incubated with 50 µM CBZ was much higher than 10 µM CBZ. According to the available reference standards, incubation of *Lemna* with CBZ showed an enhancement in the flavonoids pathway especially in the flavonol and flavone biosynthesis, as well as anthocyanin. Flavonoids, antioxidants, play a vital role in the defense mechanism of plants against ROS. The induction in the synthesis of B-ring-substituted seems to be a result of the CBZ effect on *Lemna* chlorophyll, which protects the plants against the light (UV) stress factor.

The glucose was identified in *Lemna minor* control and incubated with 10 and 50 µM CBZ. The glucose relative intensity percentage showed a decrease in its pattern in both concentrations. This might be due to the energy required to degrade CBZ and to withstand the stress effect.

DCF and CBZ could enhance the intensities of flavonoids in *Lemna minor* as a response to the stress effect of both of them in a direct way as an antioxidant against the ROS species, as well as to activate the Phase II enzymes, or because of their effect on the chlorophyll contents of *Lemna minor* (like the response to light stress).

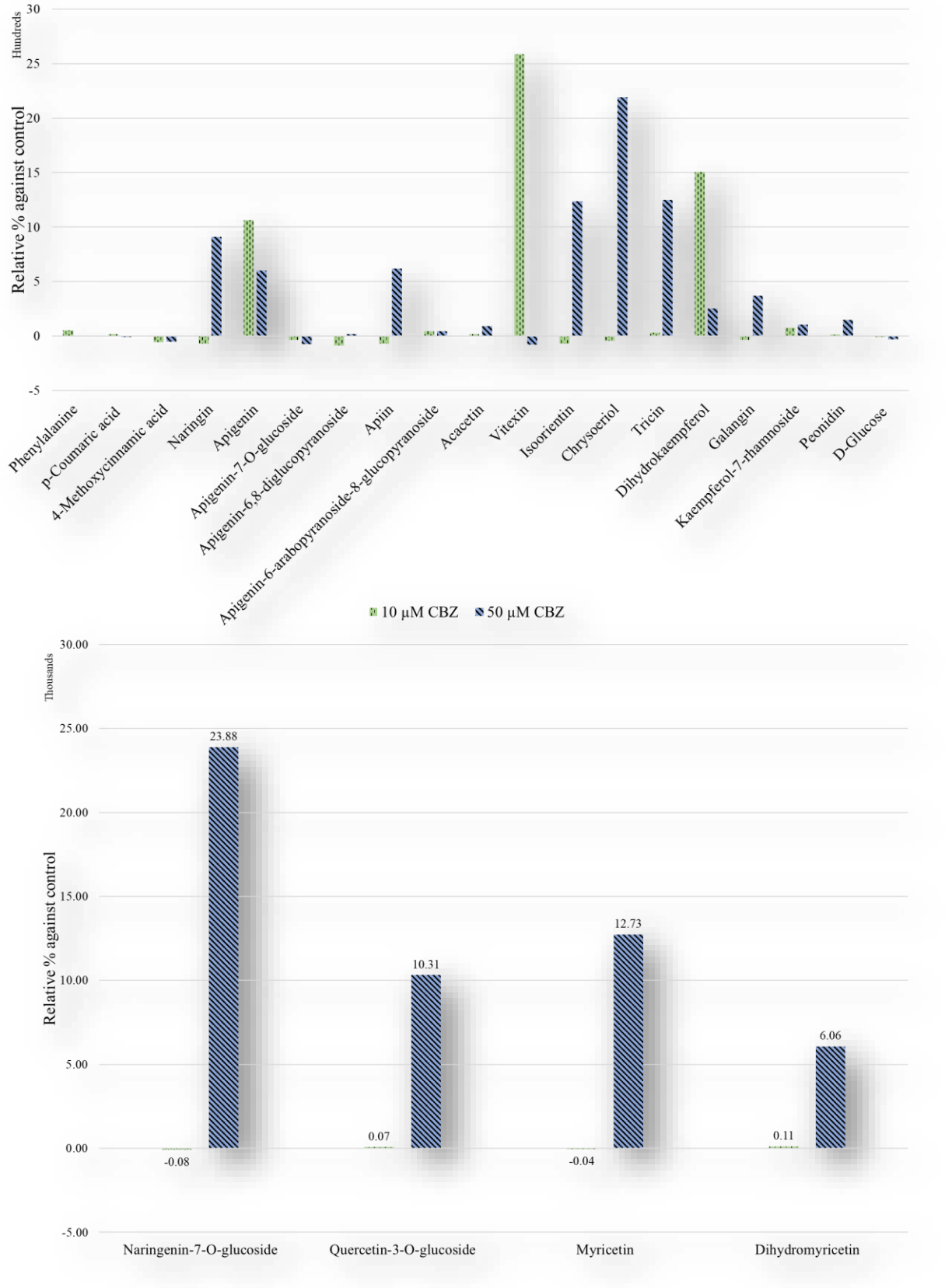


Figure 40. Relative intensities percentage (%) of different metabolites identified in *Lemna minor* incubated with 10 & 50 µM CBZ for 4 days against the control

6.7. Conclusion

Metabolomics analysis of *Lemna minor* and *Phragmites australis* incubated with 10 and 100 µM DCF and 10 & 50 µM CBZ was performed using RPLC-HILIC-ESI-TOF-MS and QTOF-MS/MS. Also, a prolonged experiment has been done via incubating *Lemna minor* with DCF for 4, 8, 12 days to investigate the time effect on the targeted metabolites. The spectral and statistical results showed changes in the metabolic profile of *Lemna and Phragmites* due to DCF and CBZ incubation after 4 days. The PLS-DA and OPLS-DA analysis identified the significant differences between the controls and incubated samples. The metabolites exhibited changes in their intensities as a response to the incubations, which were related to stress defense mechanisms.

For *Phragmites australis,* significant DMF was determined after the incubation with DCF or CBZ, individually. Different metabolic pathways were predicted from the statically identified DMF. These pathways were related mainly to the defense of the plant against stressful environmental conditions. *Phragmites australis* adapted to each drug differently. *Phragmites australis* could putatively use the glutathione metabolism pathway and unsaturated fatty acid pathway to protect itself during the incubation with DCF and CBZ, respectively.

In *Phragmites* and *Lemna*, glucose content was decreased to provide the required energy for both of them to adapt and degraded them.

Further, DCF could enhance the intensities of flavonoids in *Lemna minor* as a response to the stress effect of both of them in a direct way as an antioxidant against the ROS species, as well as to activate the Phase II enzymes, or because of their effect on the chlorophyll contents of *Lemna minor* (like the response to light stress). Also, *Lemna* could accumulate DCF for 12 days of incubation.

Hence, untargeted metabolomics has a fundamental function in determining the metabolic changes in plants due to xenobiotics exposure. These results provided insights into untargeted metabolomics as they serve as a workflow to monitor the changes in *Lemna minor* metabolic profile. Also, the mass spectrometric untargeted metabolomics strategy has a substantial role in investigating the biochemical changes and metabolic adaptation of plants in xenobiotics exposure cases. Thus, the hypothesis that **the presence of pharmaceuticals such as diclofenac and carbamazepine in the aquatic environment can affect the biosynthetic pathways of *Lemna minor* and *Phragmites australis*** can be accepted using the isolated metabolites and multivariate statistical analysis. However, more knowledge regarding genetic activity is required to understand the changes in the plant biosynthetic pathway due to xenobiotics.

7. Identification of changed different secondary metabolites biosynthetic pathways

Hypothesis #3. The presence of pharmaceuticals such as diclofenac and carbamazepine in the aquatic environment can affect the biosynthetic pathways of *Lemna minor*

7.1. Introduction

Pharmaceutical compounds (such as antibiotics: e.g. sulfamethazine, sulfamethoxazole, trimethoprim, ciprofloxacin, and amoxicillin) can be absorbed and accumulated by plants (Zhang *et al.*, 2017). Additionally, the annual usage of sulfonamide drugs (antibiotics) in veterinary medicine is approximately five times higher than the total antibiotic use in some EU countries and South Korea due to their broad-spectrum antimicrobial activity (Cheong *et al.*, 2020). Sulfonamides target and interfere with the folate biosynthetic pathway and inhibit the growth of microorganisms, in addition, plants change their growth and development to respond to the external environments, which affected their yield and quality. Sulfonamide antibiotics resemble and substitute the *p*-aminobenzoic acid, which is involved in a wide variety of metabolic processes and possesses antioxidant, anti-mutagenic, protective, and reparative properties. Thus, they are considered a competitive inhibitor of dihydropteroate synthetase (DHPS) of folate metabolism. Folate metabolism is involved in nucleic acids, proteins, lipids, and other biomolecules biosynthesis, as well as epigenetic controls via one-carbon unit transfer. Hussain and coworkers reported that ofloxacin, ciprofloxacin, levofloxacin, oxytetracycline, and doxycycline were detected in soil and vegetable samples in surrounding areas of the pharmaceutical industry (Hussain *et al.*, 2016). Further, wheat could uptake sulfamethoxazole, trimethoprim, ofloxacin, and carbamazepine when spray-irrigated with WWTP effluent (Franklin *et al.*, 2016). Furthermore, Pan and coworkers found an accumulation of antibiotics in the different edible parts of various crops (Pan *et al.*, 2014).

In 2020, it was reported that seedling growth inhibition was observed in lentil bean (*Lens culinaris*), rice (*Oryza sativa*), and Napa cabbage plants upon sulfonamide exposure (Cheong *et al.*, 2020). Also, the effect of the most common antibiotics (ciprofloxacin HCl, oxytetracycline HCl, and sulfamethazine) in wastewater have been studied in *Phragmites australis* at concentrations of 0.1-1000 μg/L for 62 days. *Phragmites* could accumulate the antibiotics via passive absorption.

They have a toxic effect on the root structure and chlorophyll constituents of the leaf. Moreover, the antibiotics induce hormesis because they decreased significantly superoxide dismutase and catalase activity (5-55% and 9-58%), while increased significantly peroxidase activity (Liu *et al.*, 2013a). Moreover, Sulfa drugs (e.g. sulfamethoxazole) and dihydrofolate reductase (DHFR) inhibitors (e.g. trimethoprim) are usually used in combination as antibacterial drugs to avoid the drug-resistance problems, shown in Figure 41 (Capasso and Supuran, 2014). Hence, Sulfamethoxazole was consistently the most toxic sulfonamide to

aquatic plants. Also, experiments with sulfamethoxazole and trimethoprim in *Lemna gibba* showed no significant difference between exposure to sulfamethoxazole alone and combinatorial exposures to sulfamethoxazole (SMX) and trimethoprim (TRIM) at equimolar concentrations (Brain *et al.*, 2009). These results have suggested differential uptake, translocation, and/or metabolism since receptors exist for both sulfamethoxazole and trimethoprim in plants (Brain *et al.*, 2009).

Despite the low human annual exposure to antibiotics through the consumption of edible crops (1.10 to 7950 μg/y), several potential adverse impacts, including allergic reactions, chronic toxic effects as a result of prolonged exposure, and even the disruption of digestive system functions, have been speculated (Zhang *et al.*, 2017). Thus, there is a concern regarding antibiotic pollution in plants and its effect along the food chain.

The majority of studies have been focused on evaluating the toxicity, up taking, and accumulation of antibiotics on the plants. Moreover, the potential effects of antibiotics on the development and spread of antibiotic-resistant bacteria and antibiotic-resistant genes (ARGs), in-plant endophytic systems such as in pakchoi (*Brassica chinensis* L.) (Zhang *et al.*, 2017). Recently, the usage of endophytic bacteria as antibacterial substances was investigated (Pal and Paul, 2013). Moreover, the biological and physiological effects of pharmaceuticals especially antibiotics on plant growth and development remain unclear (Cheong *et al.*, 2020).

The researches have been performed to improve the knowledge of the effect of pharmaceuticals in constructed wetland plants. However, the effluent of WWTP contains a mixture of different pharmaceuticals classes such as antibiotics, anti-inflammatory, cosmetics, etc.

However, in this chapter, the effect of the different pharmaceutical mixture on *L. minor* will be discussed. Also, the exposure concentration is close to the environmental concentration. The experiment was designed to exposure *L. minor* to DCF, CBZ, SMX, TRIM, and a mixture of all of them, individually. This worked aimed to study the individual effect of each drug on the *L. minor* metabolic pathway, besides, mimicking the wastewater plant effect on the lab scale.

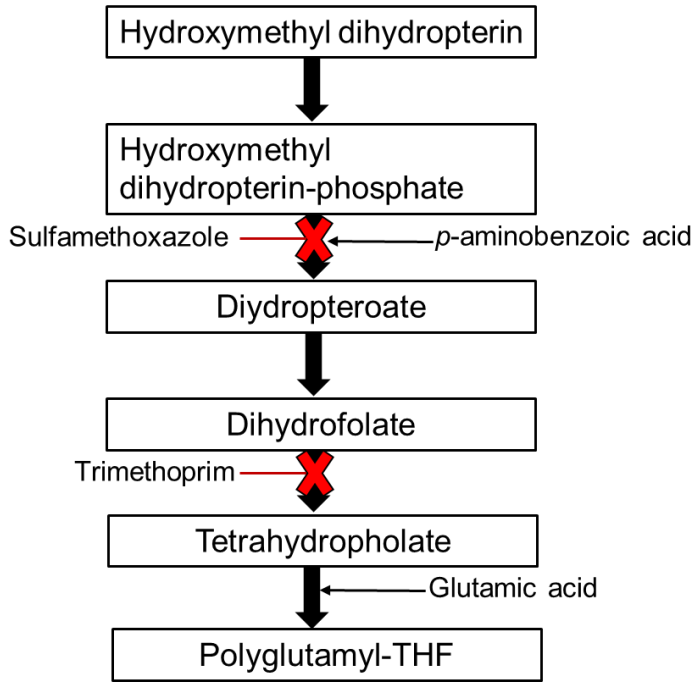


Figure 41.The folate pathway in the mitochondria of a plant cell which are suspected to be inhibited with sulfamethoxazole and trimethoprim. Sulfamethoxazole targets dihydropteroate synthase (DHPS), and trimethoprim targets dihydrofolate reductase (DHFR) in the folate biosynthetic pathway.

7.2. Experimental design

7.2.1. Plant material

60 g of *Lemna minor* was divided into (2g) and exposed to SMX, TRIM, DCF, CBZ, and the mixture of them (MIX) at 5ppb (5µg/L) in Steinberg media for 5 days. Each exposure has five biological replicates. In addition, the solvent control was prepared in identical environmental conditions.

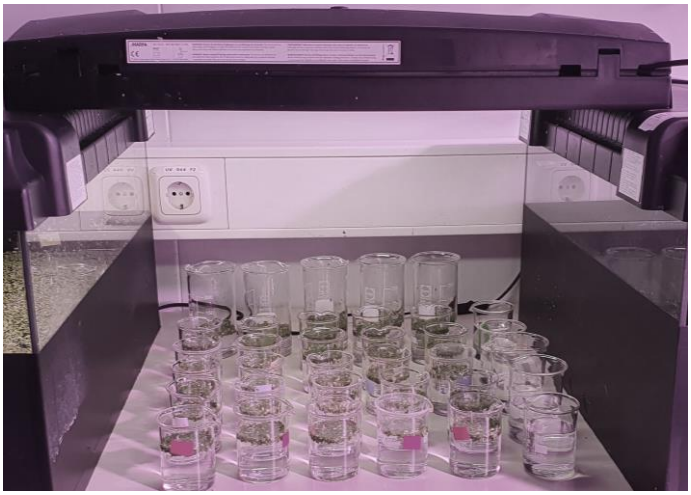


Figure 42.Photo of *Lemna minor* exposed for SMX, TRIM, DCF, CBZ, and mixture of them (MIX) at 5ppb (5µg/L) in Steinberg media for 5 days.

7.2.2. Extraction method

After freeze-drying and grinding the *Lemna* different treatments, 42 mg was sonicated for 10 minutes at 4˚C with 35 kHz frequency (Heating Ultrasonic Water Bath, Advantage-Lab™ AL04-04, Thermo-Scientific, USA) in 1 mL of 100% MeOH. Then, each extract was centrifuged at 1500 rpm for 20 min at 4˚C (ST16R Refrigerated Centrifuge, Thermo-Scientific, North Hampton, NH 03862, USA) and the supernatants were transferred to clean glass test tubes. The extraction process was triplicated in identical experimental conditions. Finally, the extracts were evaporated to dryness using liquid nitrogen and resuspended in 0.5 mL of 100% MeOH, separately.

For protein precipitation, 10% of 5-sulfosalicylic acid was added to each extract. The extract with 5-sulfosalicylic acid was mixed and centrifuged at 16100 x g for 10 min at 4°C (ST16R Refrigerated Centrifuge, Thermo-Scientific, North Hampton, NH 03862, USA). The upper phase was collected and filtered using chromafil PTFE filters.

The samples were diluted (1:20% v/v) in the mobile phase (H_2O: ACN; 95:5% v/v) before injection.

7.2.3. Instrument

The samples were randomly separated with a reverse-phase Discovery column (25cm× 4.6 mm; 5 μm particle size; Sigma Aldrich). The mobile phase for the positive mode was water A, acetonitrile as solvent B both with 0.1% formic acid. For the negative mode, water as solvent C, acetonitrile as solvent D were used. The flow rate was 250 μL/min. the volume of injection was 10 μL. The metabolites identification was performed using Q Exactive Focus Orbitrap (Thermo Fischer Scientific), equipped with a heated electrospray ionization probe (HESI) source. Samples were analyzed in both positive and negative electrospray ionization modes (ESI+ and ESI-). For further instrumentation and methodological details, see chapter 3.

7.2.4. Quality control

For repeatable metabolic analyses, three features of the analytical system must be stable retention time, signal intensity, and mass accuracy. Therefore, quality control (QC) samples were prepared and injected at regular intervals during sample analysis to control analytical repeatability and sensitivity (Want *et al.*, 2013). In addition, QC was injected at the beginning of the analysis because it is important to both conditions the column and assess the performance of metabolites within the same sample matrix. The QC is a pooled sample corresponding to a pool of 10 μL collected from each sample, which provided a true representation of the breadth of metabolites present in the sample set. Additionally, another QC sample was spiked with a 50 ppb mixture of metabolites. The mixture consists of phenylalanine, tryptophan, indole, anthranilic acid, 4-aminobenzoic acid, apigenin, glutamic acid, chrosimic acid, prephenic acid, shikimic acid, *p*-coumaric acid, *o*-coumaric acid, phenylpyruvate, tyrosine, benzoic acid, salicylic acid, caffeic acid, cinnamic acid, and 4-hydroxyphenyl pyruvate. The RT standard deviation was less than 0.2. Moreover, the relative standard deviation (%RSD) was less than 2

% as shown in (Table 21). The signal intensity was investigated in the QC sample, which showed %RSD less than 30 indicating the repeatability of the system according to (Want *et al.*, 2013). In addition, this indicated that the change in the intensities was due to the incubation with different drugs.

Table 21. List of the metabolites modulated by exposure to sulfamethoxazole, trimethoprim, diclofenac, carbamazepine, and a mixture of them in *Lemna minor*

Name	Molecular ion species	Mean RT (Min.) QC	SD	%RSD	Mean RT (Min.) S	ΔRT	%RSD Of intensity
Anthranilic acid	$[M+H]^+$	2.74	0.04	1.51	2.73	-0.01	25.3
4-Amino benzoic acid	$[M+H]^+$	9.41	0.04	0.45	9.42	0.01	25.9
Phenylalanine	$[M+H]^+$	5.58	0.04	0.67	5.59	0.01	10.3
Tryptophan	$[M+H]^+$	10.77	0.02	0.16	10.76	-0.01	4.8
Indole	$[M+H]^+$	14.53	0.18	1.22	14.56	0.03	8.4
Glutmatic acid	$[M+H]^+$	1.66	0.00	0.28	1.65	-0.01	8.6
Apigenin	$[M-H]^-$	14.02	0.01	0.05	14.03	0.01	21.9
Chrosimic/Prephenic acid	$[M-H]^-$	2.20	0.02	0.80	2.24	0.04	17.4
Shikimic acid	$[M-H]^-$	21.62	0.12	0.54	21.6	-0.02	2.6
Ferulic acid	$[M-H]^-$	8.65	0.12	1.43	8.61	-0.04	17.1
Dihydrokamepferol	$[M-H]^-$	13.24	0.02	0.16	13.25	0.01	12.2
Phenylpyruvate	$[M-H]^-$	8.39	0.02	0.26	8.40	0.01	4.8
p-Coumaric acid	$[M-H]^-$	2.24	0.03	1.12	2.23	-0.01	6.2
O-Coumaric acid	$[M-H]^-$	9.63	0.18	1.91	9.62	-0.01	14.5
Tyrosine	$[M-H]^-$	4.77	0.08	1.59	4.80	0.03	6.5
Naringenin	$[M-H]^-$	14.15	0.03	0.19	14.17	0.02	9.8
Taxifolin	$[M-H]^-$	11.42	0.02	0.15	11.44	0.02	25.6
Benzoic acid	$[M-H]^-$	10.55	0.03	0.25	10.57	0.02	6
Salicylic acid	$[M-H]^-$	5.60	0.09	1.66	5.30	-0.30	20.8
4-Hydroxy-phenylpyruvate	$[M-H]^-$	1.89	0.03	1.43	1.92	0.03	13.6
Caffeic acid	$[M-H]^-$	8.00	0.12	1.55	7.95	-0.05	19.6
Cinnamic acid	$[M-H]^-$	12.09	0.16	1.35	12.14	0.05	28.3

7.2.5. Internal Standard

Each sample was spiked with a standards mixture of 5 substances. The mixture consists of carbamazepine d10, sulfamethoxazole d4, diclofenac d4, trimethoprim d3, and ibuprofen to obtain a final concentration of 50 ppb each. It was used to normalize the peak area of the corresponding sample. The mass absolute variation (Δ ppm) was less than 5 ppm. The RT standard deviation was less than 0.2. Moreover, the relative standard deviation (%RSD) was less than 1 % as shown in (Table 22). These parameters were used to correlate the features in different samples.

Table 22. The mean RT of internal standards of the different injections, mean RT standard deviation (SD), and relative standard deviation (RSD) of the standards are listed (n=30).

Name	Mean RT (Min.)	SD	%RSD	Molecular ion species
Carbamazepine d10	12.84	0.01	0.09	$[M+H]^+$
Sulfamethoxazole d4	11.96	0.02	0.15	$[M+H]^+$
Diclofenac d4	15.06	0.07	0.49	$[M+H]^+$
Trimethoprim d3	16.04	0.01	0.08	$[M+H]^+$
Ibuprofen	15.90	0.10	0.66	$[M-H]^-$

7.3. Results and Discussion

7.3.1. Metabolites identification

The compounds were identified by comparing their masses, RT, and fragments with that of the reference standards. The compounds were identified using Qual Browser software 4.0.27.19 (Thermo Fisher Scientific Inc., USA). The pooled samples were used to determine the RT of the metabolites, in addition to, system repeatability and sensitivity (Table 21). The data was compared with the reference standard fragments, in addition to the literature. The metabolites were identified using their RT and fragments when they have the same masses, such as in *p*-coumaric acid, *o*-coumaric acid, and phenyl pyruvic acid. The ratio of the two fragments 116.9 and 119 was used to distinguish between *p*-coumaric acid, *o*-coumaric acid. Further, the phenyl pyruvic acid can be identified with the presence of 117.3 fragments (peak) and the base ion peak at 119.05 (Fig. 42). The naringenin and naringenin chalcone have the same mass and RT. They can not be separated using the applied method. Therefore, the peak area is calculated as the total of both of them. In addition, it was the case for the chrosimic acid and prephenic acid

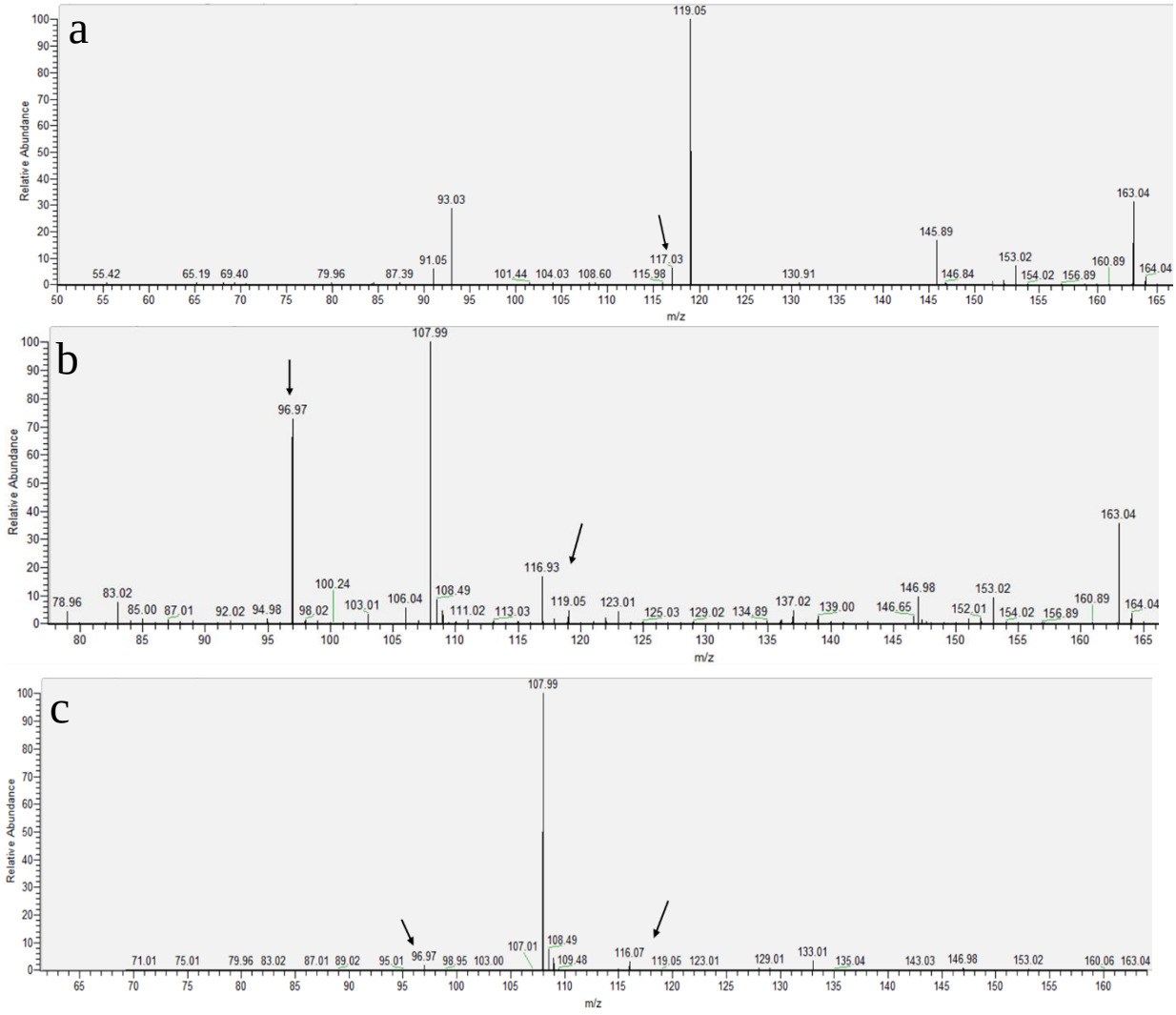


Figure 43. Mass spectrum of (a) phenylpyruvate; (b) *o*-coumaric acid, (c) *p*-coumaric acid

7.3.2. The selection of metabolic pathways

Flavonoids are important metabolites to plants playing many roles as attractors, feeding repellents, photoreceptors, oviposition stimulants, phytoalexins, antioxidants, and antimicrobials, which synthesis through phenylpropanoid pathway as well as different phenolic compounds. Additionally, they are used in pharmaceutical and cosmetic preparations for humans due to their anti-allergenic, anti-microbial, anti-inflammatory, vasodilatory, anti-mutagenic, and anti-carcinogenic activities. The antioxidant property of flavonoids is important to humans and plants. On one hand for humans, the antioxidant activity is higher than those of vitamin E and vitamin C (Nabavi *et al.*, 2020). Moreover, flavonoids regulate the movement and catabolism of auxins, indole-3-acetic acid oxidase affecting morpho-anatomical characters of plants and modulate protein activity implicated in cell growth, respectively (Nabavi *et al.*, 2020). Flavonoids synthesize from *p*-coumaroyl-CoA, derived from the (phenylpropanoid) and malonyl-CoA from the acetate-malonate pathway. Those precursors are the substrates to form naringenin chalcone, which is catalyzed by chalcone synthase. Then, naringenin chalcone cyclizes stereospecifically to form naringenin, which is considered the general precursor of the flavonoids and anthocyanin (Yonekura-Sakakibara *et al.*, 2019). Furthermore, cinnamic acid is synthesized from phenylalanine, which is synthesized from shikimic acid in the shikimate pathway. DCF and CBZ show an effect of flavonoids and anthocyanin pathways in *Leman minor* and *Phragmites australis*. Also, SMX and TRIM affect the folate pathway (Fig. 44). Thus, the phenylalanine tyrosine and tryptophan biosynthesis, folate biosynthesis, phenylpropanoid (including flavonoids and anthocyanin pathways) pathways were chosen to investigate in different *Lemna minor* samples incubated with DCF, CBZ, SMX, TRIM, and a mixture of all of them at 5 ppb.

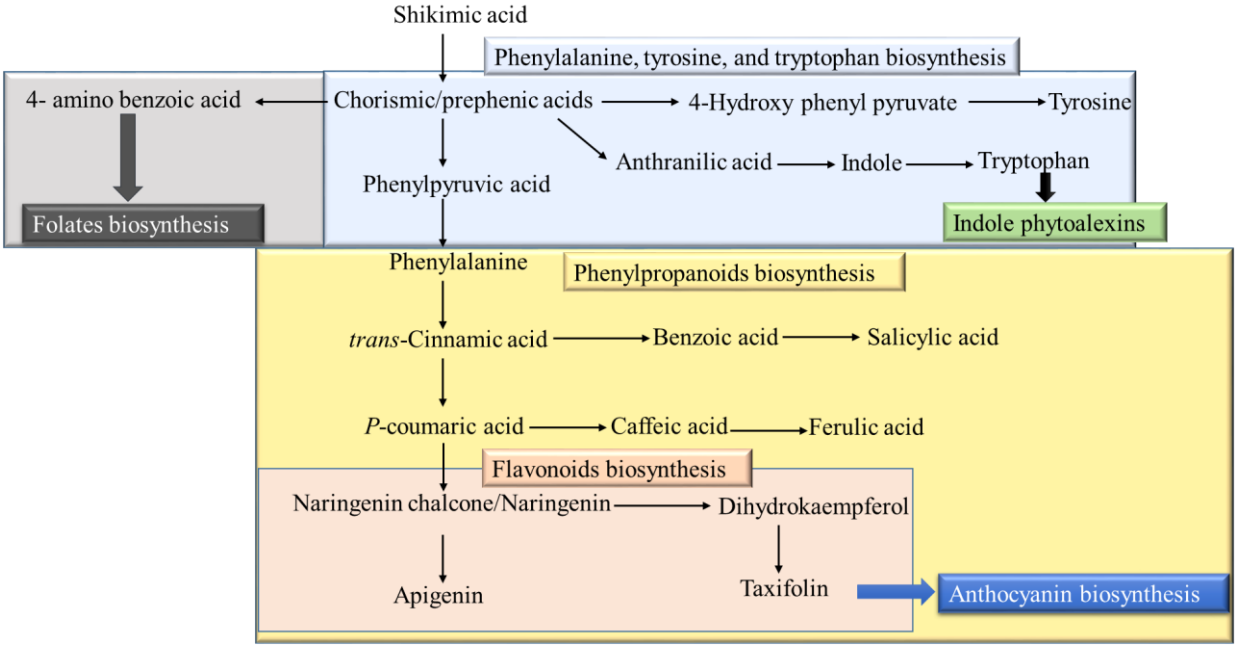


Figure 44. The schematic of the investigated biosynthetic pathway in *Lemna minor*.

7.3.3. Lemna incubated with at 10 and 50 µM CBZ for 4 days

The aromatic amino acids biosynthesis (phenylalanine, tryptophan, and tyrosine). They are the precursors for many metabolites in plants, which play a role in plant defense and resistance to pathological and environmental stress.

AAAs are produced from the final product of the shikimate pathway, chorismate. Anthranilate synthase and chorismate mutase (CM) enzymes are inhibited by the final product(s) feedback of the corresponding pathways (tryptophan and phenylalanine/tyrosine, respectively) in both microbes and plants. In addition, tryptophan activates CM to redirect flux from tryptophan to phenylalanine/tyrosine biosynthesis (Maeda and Dudareva, 2012).

Shikimic acid, chrosimic/prephenic acid, phenyl pyruvic acid, and phenylalanine were increased after 10 µM CBZ exposure and decreased after 50 µM CBZ exposure (Fig. 44). Thus, the phenylalanine biosynthesis is enhanced in *Lemna minor* incubated with 10 µM CBZ but not at a higher concentration.

However, tryptophan biosynthesis was enhanced in both concentrations. The anthranilic acid, indole, and tryptophan concentrations were increased in both treatments (Fig. 45). Consequently, the glutamic acid concentration was decreased in both concentrations. Also, anthranilic acid was increased into (2 fold) in higher concentration, however, indole and tryptophan were increased less than 10 µM CBZ. Despite this, the 4-hydroxy phenylpyruvate was increased in 10 µM CBZ. It was decreased after 50µM CBZ incubation (Fig. 45). This was accompanied by a decrease in tyrosine concentration in both concentrations. This was due to the reflux mechanism of tryptophan and *Lemna* induces phytoalexins production. Tryptophan effectively activates chorismate with activation constants (Ka) of (1.2–2.4 µM) and can reverse phenylalanine/tyrosine-mediated chorismate inhibition (0.3-1.1 µM). The tight regulation of tryptophan and tyrosine biosynthesis in plants ensures that the major carbon flux is directed toward phenylalanine biosynthesis for the production of abundant phenolic compounds (Maeda and Dudareva, 2012). Hence, Maeda and Dudareva, 2012 reported that the feedback mechanism of phenylalanine increased the production of phenylalanine and tyrosine products whereas the products of tryptophan decreased (Maeda and Dudareva, 2012). In the folate pathway, chorismate and glutamine are converted to aminodeoxychorismate and glutamate. The first is converted to 4-aminobenzoic acid catalyzed by aminodeoxychorismate lyase (Gorelova *et al.*, 2017).

The last compound, which is formed from chrosimic acid, is 4- aminobenzoic acid, which was increased with 3.5 fold in 50 µM CBZ compared to 10 µM CBZ (Fig. 45).

The folate's roles in stress response and resistance have a big knowledge gap. The transcriptome and metabolome analyses studies confirmed that folate metabolism is dramatically and differentially affected by various stress conditions. In the *Arabidopsis* suspension cell cultures, menadione caused an up-regulation of genes involved in folate biosynthesis and consumption (Baxter *et al.*, 2007). However, the proteomic analysis of rice under cold stress has demonstrated a down-regulation of folate biosynthesis (Neilson *et al.*,

2011). Also, a down-regulation of folate biosynthesis-related genes was observed in plants under salt stress (Storozhenko *et al.*, 2007).

In the phenylpropanoid pathway, the simple phenolic compounds, cinnamic, benzoic, salicylic, *p*-coumaric, caffeic acids were increased in *Lemna* incubated with 10 µM CBZ. In 50 µM CBZ, they were decreased except *p*-coumaric, benzoic, and salicylic acids. Salicylic acid showed a higher increase (up to 4.8 fold) than the lower concentration of CBZ. Salicylic acid was synthesized from *trans*-cinnamic acid leads to the formation of benzoic acid through a β-oxidation and non-β-oxidation pathway. In the first one, benzoic acid is formed from *trans*-cinnamic acid through cinnamoyl-CoA and benzoyl-CoA intermediates. Also, it can be synthesized using *o*-coumaric acid as an intermediate. In the non-β-oxidation pathway, The benzoic acid formation involves *para*-hydroxybenzaldehyde as the intermediate (Kohli *et al.*, 2017). Therefore, the *o*-coumaric acid concentration was investigated in *Lemna minor* showing a decrease in 10 µM CBZ, however, a remarkable increase at higher concentration, which explains the higher increase of salicylic acid at higher concentration despite the slight increase in benzoic acid concentration. Salicylic acid was increased to protect against oxidative stress caused by heavy metals (Kohli *et al.*, 2017). Moreover, *Lemna* incubated with DCF for 4 days used the β-oxidative pathway to produce gallic acid (section 6.3.4).

On the other hand, the flavonoids pathway, naringenin chalcone/naringenin, apigenin, dihydrokaempferol concentrations were increased in 10 µM CBZ, however, decreased at higher concentration. Further, naringenin chalcone/naringenin and apigenin were approximately equal to control at 50 µM CBZ (Fig. 45). In the previous study of *Lemna* with system B, the flavonoid glycosides were detected with higher concentrations at 50 µM CBZ.

Regarding the anthocyanin pathway, the taxifolin (dihydroquercetin) is considered the linkage between flavonoids and anthocyanin biosynthesis. It is concentration was increased at both concentrations (Fig. 45). In the previous study of *Lemna* with system B, the peonidin (anthocyanin) was increased due to CBZ incubation.

7.3.4. Lemna incubated with at 10 and 50 µM CBZ for 4,8,12 days

Lemna minor was incubated with CBZ for long period to investigate the prolonged stress effect of CBZ on *Lemna*, in addition to the ability of *Lemna* to accumulate CBZ. The experimental design resembled that of DCF in the previous chapter (section 6.6.5).

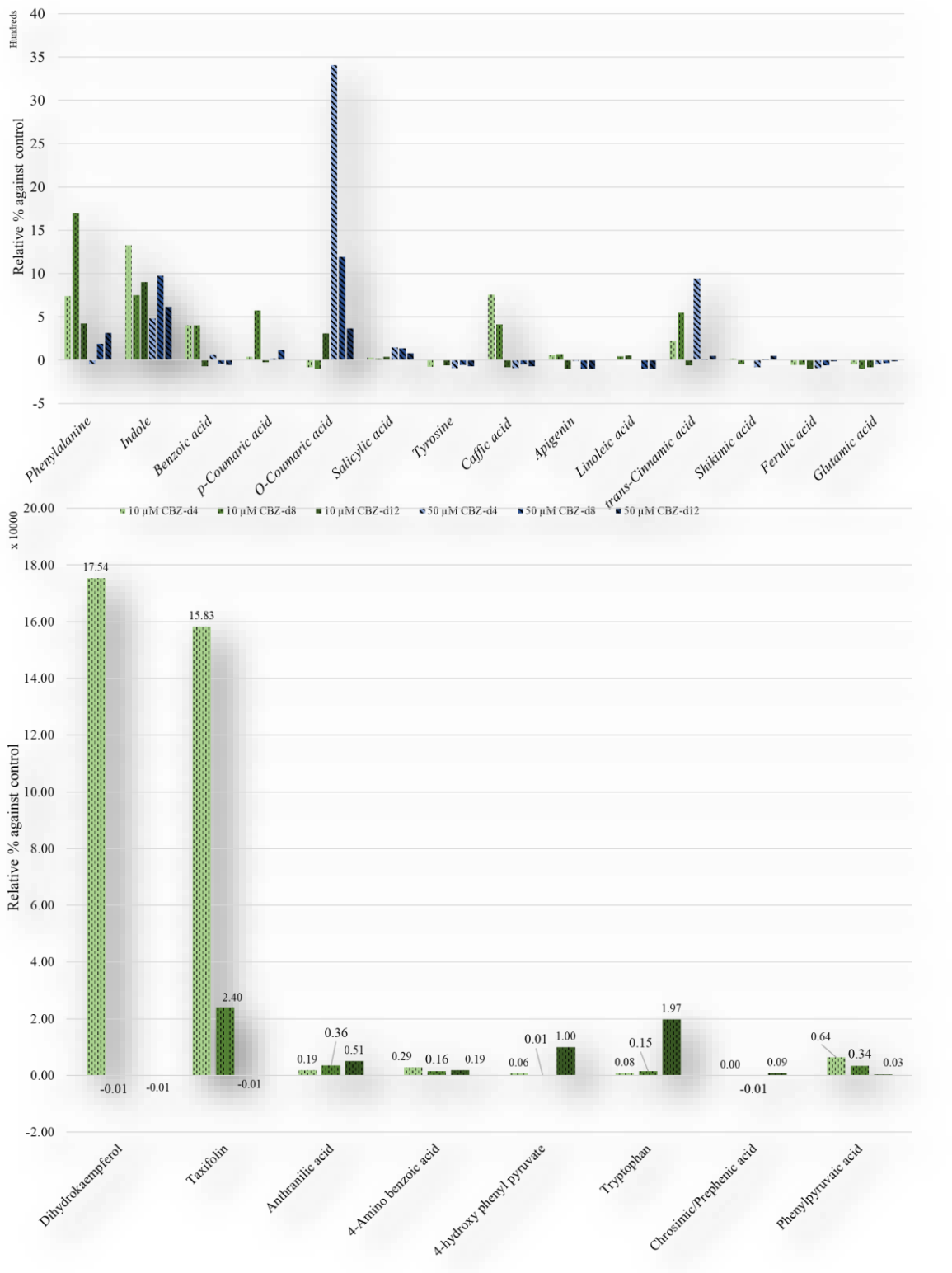
Regarding *Lemna* incubation for 8 days, *Lemna* was exposed to 10 and 50 µM CBZ under the same experimental conditions, individually. The phenylalanine biosynthesis is enhanced after incubation for 8 days at both concentrations due to an increase in phenyl pyruvic acid and phenylalanine, Despite the decrease of shikimic acid and chrosimic/prephenic acid at 10 µM CBZ, their concentration was increased at 50 µM CBZ (Fig. 45). Further, the enhancement of tryptophan biosynthesis is continuous to produce the phytoalexins with an increase in anthranilic acid, indole, and tryptophan concentrations. Furthermore, the 4-hydroxy phenylpyruvate was increased at both concentrations. Also, it increased at 50 µM CBZ with 261 fold than 10 µM CBZ. The tyrosine was increased and decreased at 10 and 50 µM CBZ,

respectively. This might be due to the activation of the tyrosine aminotransferase enzyme which induces the reversible deamination of tyrosine to form 4-hydroxy phenylpyruvate, which is an initial step in the biosynthesis of many tyrosine-derived metabolites such as tocopherols and plastoquinone, betalains, benzylisoquinolines, and rosmarinic acid (Xu *et al.*, 2018).

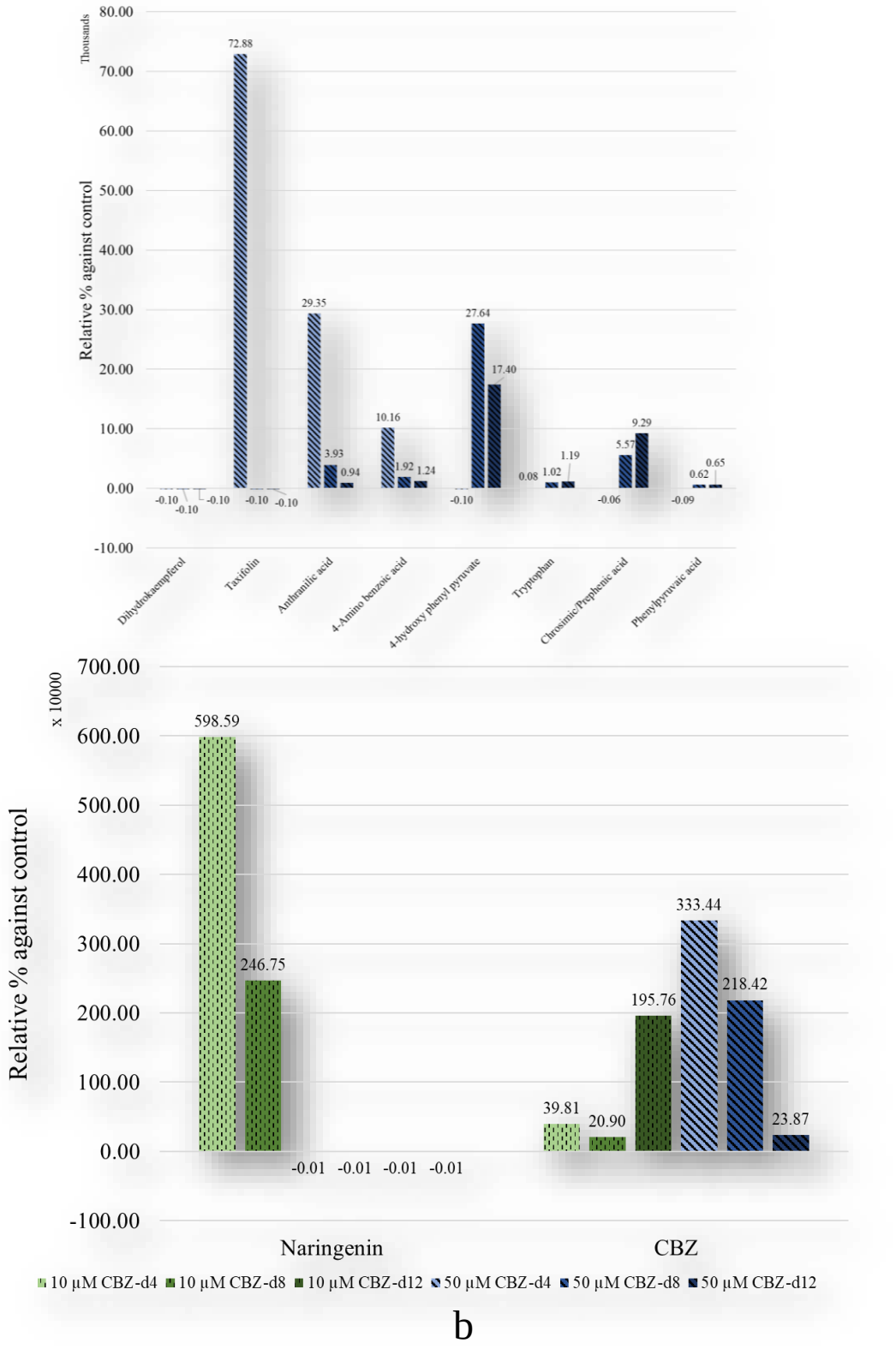
Regarding the phenylpropanoid pathway, the trans-cinnamic acid and *p*-coumaric acid were increased in both concentrations, however, in 10 µM CBZ, they increased by 41.6 and 4.8 fold than 50 µM CBZ, respectively (Fig. 45). Further, by the same fold increased the benzoic acid as *Lemna was* incubated for 4 days at 10 µM CBZ concentration. Also, its concentration was decreased at a higher CBZ concentration. However, salicylic acid was decreased to half at 10 µM CBZ than *Lemna* incubated for 4 days. Also, it kept the same increases at 50 µM CBZ. This was due to the decrease and increase in *o*-coumaric acid at 10 and 50 µM CBZ, respectively. This indicates that *Lemna* under incubation with CBZ synthesis the salicylic acid using the *o*-coumaric acid. Despite the increase in the *p*-coumaric concentration, the ferulic acid decreased approximability at the same level after 10 and 50 µM CBZ exposure. Moreover, the caffeic acid showed the same pattern as incubation for 4 days. However, it decreased with 0.5 fold than 10 µM CBZ incubated for 4 days. The flavonoids pathway showed the same pattern for the compounds as *Lemna* incubated for 4 days. However, dihydrokaempferol was decreased at 50 µM CBZ and its concentration was approximately equal to the control. Moreover, the taxifolin concentration was increased and decreased at 10 and 50 µM CBZ, respectively (Fig. 45). The folate pathway showed the same pattern as at 4-day incubation due to an increase in 4-aminobenzoic acid and a decrease in glutamic acid concentrations (Fig. 45).

Regarding *Lemna* incubation for 12 days, shikimic acid and chrosimic/prephenic acid, phenyl pyruvic acid, phenylalanine, anthranilic acid, indole, and tryptophan concentrations were increased at 10 and 50 µM CBZ (Fig. 45). This showed the enhancement of phenylalanine, tyrosine, and tryptophan biosynthesis in *Lemna minor* due to incubation with CBZ. Also, the tyrosine and 4-hydroxy phenyl pyruvic acid showed the same pattern as in exposure for 8 days. The production of salicylic acid was increased with a decrease in benzoic acid and an increase in *o*-coumaric acid. Also, the ferulic acid and caffeic acid concentrations were decreased by 1.8, 5, and 0.2, 0.7 fold at 10 and 50 µM CBZ, respectively than *Lemna* incubated for 8 days (Fig. 45). Furthermore, the folate pathway showed the same pattern as at 4 and 8-day incubation due to an increase in 4-aminobenzoic acid and a decrease in glutamic acid concentrations.

Concluding, *Lemna* metabolic pathway has been changed due to incubation with 10 and 50 µM CBZ. The increase in the incubation period showed in some aspects the same response as in tryptophan and tyrosine biosynthesis, folates, and salicylic acid biosynthesis. Besides, the long period of incubation increases the phenylalanine biosynthesis. The flavonoids pathway showed an enhancement mainly at 10 µM CBZ. Finally, all of the effects due to the CBZ as shown in Figure 45 *Lemna* could uptake and accumulate CBZ. In 10 µM CBZ, the concentration of CBZ is increased with increasing the period of incubation. However, in 50 µM CBZ, the concentration is decreased with prolonged periods of incubation.


Hundreds
Relative % against control
40
35
30
25
20
15
10
5
0
-5
Phenylalanine
Indole
Benzoic acid
p-Coumaric acid
O-Coumaric acid
Salicylic acid
Tyrosine
Caffic acid
Apigenin
Linoleic acid
trans-Cinnamic acid
Shikimic acid
Ferulic acid
Glutamic acid
10 µM CBZ-d4
10 µM CBZ-d8
10 µM CBZ-d12
50 µM CBZ-d4
50 µM CBZ-d8
50 µM CBZ-d12
x 10000
Relative % against control
18.00
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
-2.00
17.54
-0.01
-0.01
15.83
2.40
-0.01
0.19
0.36
0.51
0.29
0.16
0.19
0.06
0.01
1.00
0.08
0.15
1.97
0.00
0.09
-0.01
0.64
0.34
0.03
Dihydrokaempferol
Taxifolin
Anthranilic acid
4-Amino benzoic acid
4-hydroxy phenyl pyruvate
Tryptophan
Chrosimic/Prephenic acid
Phenylpyruvaic acid

a



b

Figure 45. Relative intensities percentage (%) of different metabolites identified in *Lemna minor* incubated with 10 & 50 µM CBZ for 4, 8, and 12 days against the control. The green color represents 10 µM CBZ treatment. The blue color represents the 50 µM CBZ treatment. The darker the color, the longer the incubation period.

7.3.5. The effect of SMX on Lemna metabolic pathway

Lemna minor was incubated with SMX for 5 days at 5ppb concentration. Different metabolic pathways were investigated in *Lemna* extract. The phenylalanine, tyrosine, and tryptophan biosynthesis are up-modulated. Hence, phenylalanine, tyrosine, and tryptophan (final products) were increased by 5, 4.3, and 4.3 fold than the control samples, respectively. Moreover, the precursor (chorsimic/prephenic acid) was up-modulated (1.4 fold) than the control extracts (Fig. 46b). Further, the intermediate phenylpyruvate (1.2 fold), 4-hydroxy phenylpyruvate (1.9 fold), anthranilic acid (2.4 fold), and indole (34 fold) were up-modulated (Fig. 46b&c), however, the shikimic acid was down-modulated in *Lemna minor* 100% MeOH extracts (Fig. 46a). This may lead to an increase in the biosynthesis of different metabolites. The feedback inhibition loops of the three aromatic amino acids were inhibited. The inhibition of tryptophan feedback was demonstrated with the increase in its intermediate anthranilic acid (Tzin and Galili, 2010). Besides, tryptophan induces the production of phenylalanine and tyrosine. *Lemna minor* increases also the production of aromatic amino acids (AAAs) after incubation with drugs such as glyphosate, metribuzin, and their mixture (Kostopoulou *et al.*, 2020). Chorismate is the initial branch point metabolite in the synthesis of all three aromatic acids, as well as the wide range of aromatic secondary metabolites such as folates (Tzin and Galili, 2010). The incubation of *Lemna* with SMX increased the 4-aminobenzoic acid into (1.9 fold) than control extracts. Also, glutamic acid was up-regulated (1.6 fold) in incubated samples (Fig. 46b&c). *Lemna gibba* showed an increase in 4-aminobenzoic acid after exposure to SMX (Brain *et al.*, 2008b). Thus, SMX seems to block the folate biosynthesis pathway in *Lemna minor* as in bacteria. The up-modulation in phenylalanine in *Lemna minor* resulted in an up-regulation of the phenylpropanoid pathway and hence the flavonoids pathway. The first compound in the phenylpropanoid pathway is the trans-cinnamic acid, which was increased by 2.7 fold compared to the control extracts (Fig. 46a). Further, *p*-coumaric (intermediate) was increased 4.5 fold (Fig. 46b). The simple phenolic compounds were increased such as benzoic acid (1.3 fold), and caffeic acid (1.3 fold), however, the ferulic acid and salicylic acid had the same and lesser concentration than the control extracts, respectively (Fig. 46b&d). The data showed an up-modulation in the phenylpropanoid pathway. Consequently, an enhancement in the flavonoid pathway was expected. The results showed an increase of naringenin, which is the general precursor of all flavonoids. This lead to a rise in apigenin (1.6 fold) and dihydrokaempferol (2 fold) compared to control extracts (Fig. 46c). They are the precursors for flavone and flavonol biosynthesis. However, the taxifolin was down-modulated which is considered the precursor for the anthocyanin (Fig. 46c). The up-regulated flavonoid pathway can protect the plant against various stressful conditions such as freezing or nutrient limitation, cold treatment, and nitrogen depletion (Olsen *et al.*, 2009). Further, Liu and coworkers showed an increase in flavanone 3-hydroxylase gene (F3H), a key enzyme in flavonoid biosynthetic pathway under UV-B radiation and drought stress in *Reaumuria soongorica* (Liu *et al.*, 2013b). Furthermore, heavy metals stimulate the phenylpropanoid pathway in plants by up-regulating the activities of key biosynthetic enzymes. Additionally, the phenylpropanoid scavenges the ROS species and

enhances the plant's resistance to abiotic stress conditions (Sharma *et al.*, 2019). Thus, the accumulation of phenolic compounds is due to the up-regulation of the biosynthesis of phenylpropanoid enzymes including phenylalanine ammonia-lyase, chalcone synthase, shikimate dehydrogenase, cinnamyl alcohol dehydrogenase, and polyphenol oxidase.

7.3.6. The effect of TRIM on Lemna metabolic pathway

Lemna minor was incubated with trimethoprim (TRIM) with the same concentrations as SMX, individually. It has the same effect as SMX on bacteria and causes inhibition of the folic acid but in different positions of the folates biosynthesis pathway. In *Lemna minor*, TRIM caused up-modulation of phenylalanine, tyrosine, and tryptophan biosynthesis pathway. Its effect resembled that of SMX. The precursors and intermediate, and final product concentrations were increased as follows: chrosimic/prephenic acid (1.7 fold), phenylpyruvate (1.5 fold), 4-hydroxy phenylpyruvate (1.8 fold), anthranilic acid (1.2 fold), indole (4.1 fold), phenylalanine (3.9 fold), tyrosine (4.6 fold), and tryptophan (4 fold) compared to control extracts (Fig. 46).

Regarding the folate pathway, the accumulation of 4-amino benzoic acid (1.8 fold) and glutamic acid (1.7 fold) indicated that *Lemna* attempted to synthesize the intermediates of the folate pathways to overcome the TRIM effect (Fig. 46b&c).

Furthermore, the phenylpropanoid pathway was up-regulated. The *trans*-cinnamic acid (3.9 fold), *p*-coumaric (2.2 fold), benzoic acid (1.6 fold), salicylic acid (1.2 fold), caffeic acid (3.2 fold), and ferulic acid (2.2 fold) (Fig. 46b,c &d). The *o*-coumaric acid was increased by 2.6 fold compared to control extracts (Fig. 46b). This indicated that *Lemna* used it as a precursor for the synthesis of salicylic acid. In TRIM incubation, *Lemna* accumulates caffeic acid and ferulic acids more than in SMX incubation, although the tyrosine concentration is the same in SMX and TRIM treatments. Interestingly, the phenylalanine and *trans*-cinnamic acid concentrations were decreased and increased in TRIM treatments, respectively. This might be due to enhancement in the phenylalanine ammonia-lyase activity. Consequently, the lignin pathway was up-regulated. Lignin is a well-known defense polymer, which has antimicrobial activity, providing structural support, transports water, and acts as a physical barrier (Xie *et al.*, 2018). It was reported recently that the accumulation of lignin through the activation of quinate/shikimate *p*-hydroxycinnamoyltransferase is regulated by pathogens in plants (e.g. *Populous*). Cinnamate 4-hydroxylase and caffeic acid *O*-methyltransferase also are considered the key enzymes for the downstream metabolites. The enhancement of the three enzymes caused an accumulation in lignin. This might be the reason for lignin accumulation in *Lemna* incubated with TRIM. Consequently, the flavonoids pathway was upmodulated due to the increase in naringenin and apigenin. Apigenin was increased by 2.4 fold than the control extracts and *Lemna* incubated with SMX. However, dihydrokaempferol and taxifolin were decreased (Fig. 46a&c). This explains the increase in apigenin more than in *Lemna* incubated with SMX due to shifting the phenylpropanoid pathway in flavonoids and lignin direction but not in anthocyanin one.

7.3.7. The effect of DCF on Lemna metabolic pathway

Lemna minor was incubated with DCF at 10 and 100 µM. The different extracts were analyzed with systems A and B. The higher doses were expected to cause a more vigorous effect, which could be detected using system A, and defining the optimum workflow for data analysis. In this chapter, *Lemna* was incubated with 5 ppb individually to investigate the effect of the dosage on *Lemna's* metabolic pathway. The results showed that phenylalanine, tyrosine, and tryptophan biosynthesis pathway were all up-regulated. The precursors and intermediate and final product concentrations were increased. Chorsimic/prephenic acid (1.8 fold), and tryptophan (4.1 fold) were up-modulated as in *Lemna* incubated with TRIM (Fig. 45b). Furthermore, phenylpyruvate (2.4 fold), and anthranilic acid (2 fold) concentrations were less than their concentration in TRIM extracts (Fig. 46b&c).

4-Hydroxy phenylpyruvate (3.4 fold), indole (6.4 fold), phenylalanine (5.1 fold), tyrosine (5.5 fold) were higher than TRIM extracts (Fig. 46). However, indole concentrations in SMX extracts were higher than both of them. Hence, the plants possess other enzymes besides tryptophan synthase to produce indole from indole-3-glycerol phosphate. In maize indole-3-glycerolphosphate lyase is responsible to form the volatile indole under herbivore attack and tryptophan synthase-like enzyme (BX 1) produces natural benzoxazinoid pesticides ((DIBOA). The production of indole or benzoxazinoids sometimes exceeds tryptophan (Buchanan *et al.*, 2015). *Lemna* showed the same results after incubated with SMX and DCF. This explained the decrease in tryptophan relative intensity at incubation with DCF higher dosage and (DIBOA) detection.

Regarding the folate pathway, the accumulation of 4-amino benzoic acid (1.6 fold) and glutamic acid (1.3) were less compared to extracts incubated with SMX and TRIM (Fig. 46b&c). this indicates that SMX and TRIM block the folate pathway in *Lemna.*

Furthermore, the phenylpropanoid pathway was up-modulated. The *trans*-cinnamic acid (2.7 fold), *p*-coumaric (3.3 fold) concentrations were the opposite of TRIM extracts, which have *trans*-cinnamic acid (3.9 fold), and *p*-coumaric (2.2 fold) concentrations compared to the control.

Benzoic acid (1.5 fold), salicylic acid (1.3 fold) were the same as in TRIM samples due to the decrease in *o*-coumaric acid concentration (Fig. 46). Moreover, the caffeic acid (2.6 fold), and ferulic acid (2.8 fold) concentrations were decreased and increased than TRIM extracts, respectively. Also, tyrosine concentration was higher than TRIM extracts. Consequently, the lignin pathway was up-modulated. Despite that naringenin, concentration was approximately equal to the control samples, apigenin (1.2 fold) was up-modulated. Moreover, the dihdyrokaempferol and taxifolin were down-modulated (Fig. 46c). The higher is the DCF concentration, the higher is the stress effect. Also, the tryptophan inhibition mechanism on phenylalanine products is less at a higher concentration. Thus, *Lemna* enhanced the phenylpropanoid pathway under the influence of 10 and 100 µM DCF.

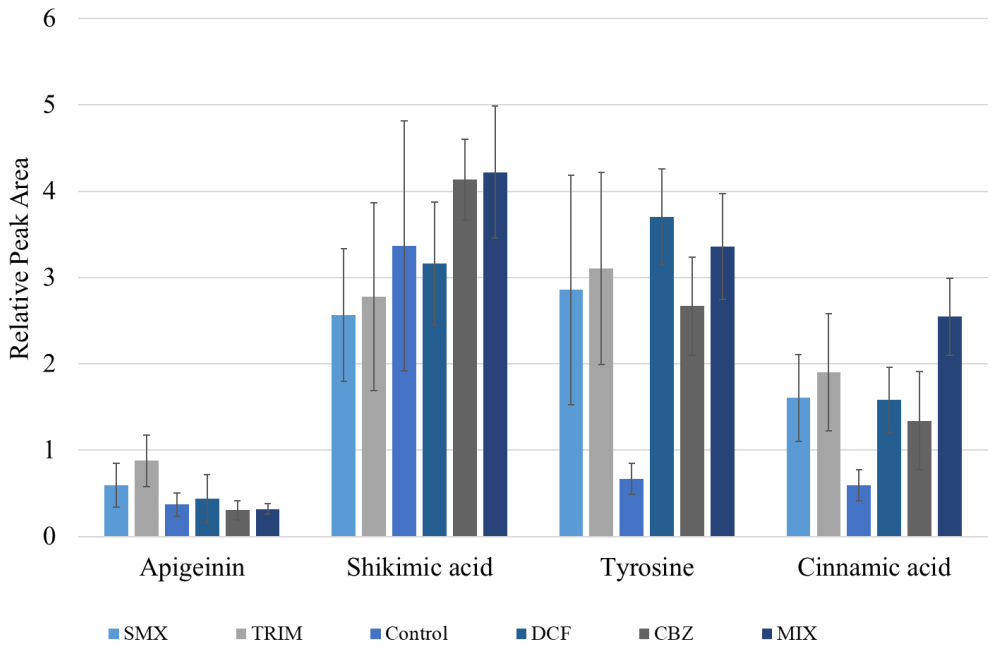
7.3.8. The effect of CBZ on Lemna metabolic pathway

The CBZ showed the same effect as SMX, TRIM, and DCF on *Lemna's* metabolic profile. However, shikimic acid (1.2 fold) was up-modulated, which resulted in the increase of phenylalanine with a concentration higher than in *Lemna* incubated with SMX, TRIM, and DCF (Fig. 46a). Despite the increase in indole (4.5 fold) above control extracts, tryptophan was up-modulated with a concentration less than *Lemna* incubated with SMX, TRIM, and DCF. Interestingly, 4-aminobenzoic acid was accumulated in *Lemna* incubated with SMX and TRIM. This is might be due to the stress effect of CBZ on *Lemna* (Fig. 46c). Thus, *Lemna* required an enhancement in the carbon flux mechanism (one-carbon unit transfer) and/or the effect of CBZ on *Lemna* DNA.

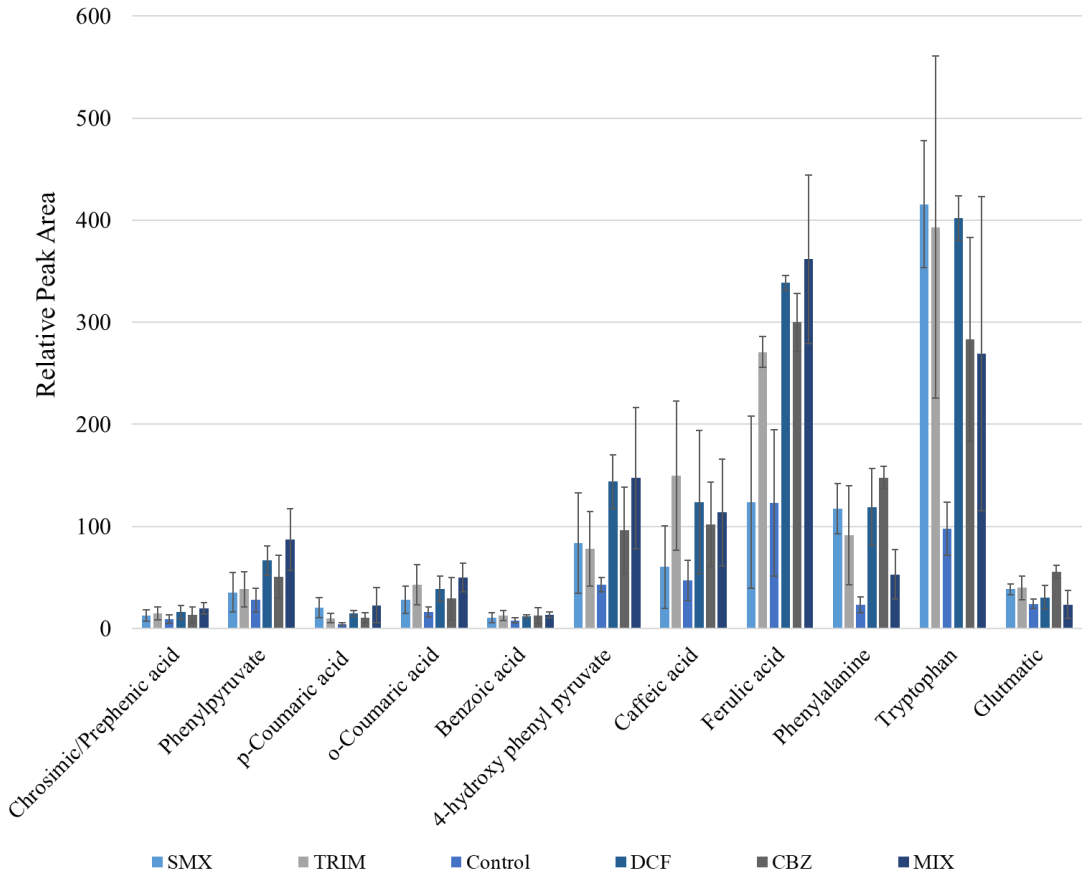
Regarding the phenylpropanoid pathway, an increase of lignin is also observed as in *Lemna* incubated with TRIM, and DCF. Further, the salicylic acid was down-modulated, although the up-modulation of *o*-coumaric acid. This indicated that the concentration of *o*-coumaric acid less than 2 fold did not enhance the salicylic acid synthesis. Furthermore, the lignin pathway was modulated (Fig. 46). However, the data showed a down-modulation in flavonoids and anthocyanin pathways due to the down-modulation of apigenin, dihydrokaempferol, taxifolin. Moreover, naringenin was approximately equal to the control extracts (Fig. 46).

7.3.9. The effect of SMX, TRIM, DCF, and CBZ mixture on Lemna metabolic pathway

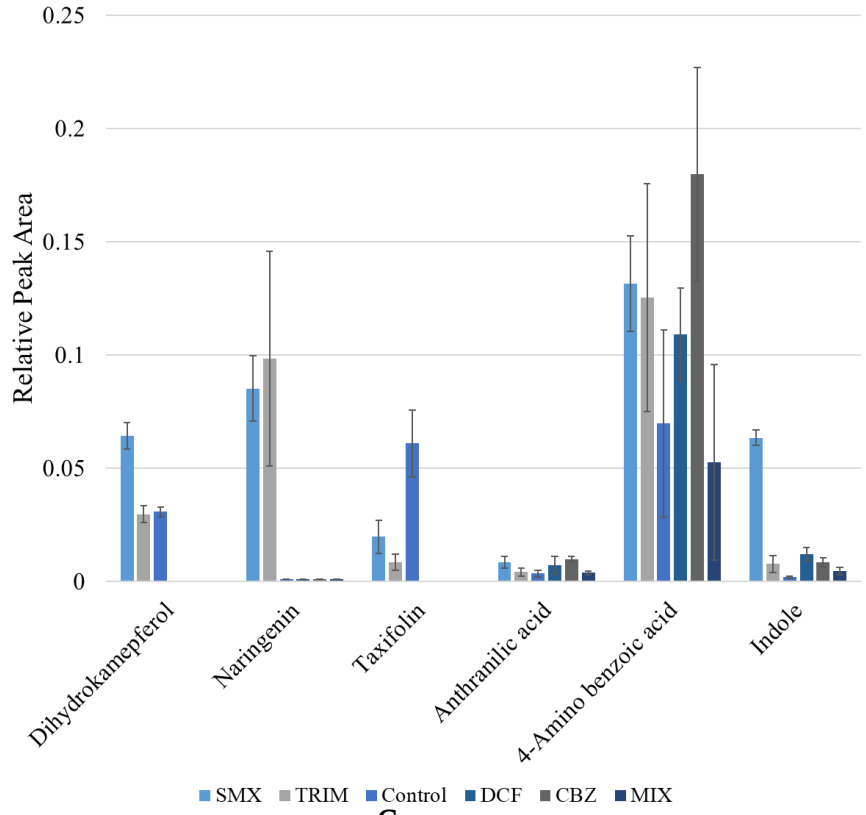
Lemna was incubated with a 5 ppb mixture of (SMX, TRIM, DCF, and CBZ) to investigate the synergistic effect of different xenobiotics on *Lemna's* metabolic profile. The shikimate pathway was up-modulated due to the up-modulation of shikimic acid (1.3 fold) (Fig. 46a). This is related to the CBZ incubation effect on *Lemna.* Hence, CBZ is the only compound in the four pharmaceuticals that caused up-modulation in shikimic acid in this study. The results also showed up-modulation of phenylalanine, tyrosine, and tryptophan biosynthesis pathway. However, the concentration of indole was less than in single treatments (Fig. 46). This indicated that indole was trapped inside the tryptophan synthase tunnel to condense with serine to form the tryptophan. Regarding the folate pathway, 4-aminobenzoic acid and glutamic acid were down-modulated, which was expected due to the inhibition effect of SMX and TRIM on the uptake of each other. Moreover, the up-modulation of the lignin pathway was observed due to the up-modulation of *p*-coumaric acid, caffeic, and ferulic acid (Fig. 46b). Additionally, the flavonoids and anthocyanin pathway were down-modulated (Fig. 46). This could be due to the effect of DCF and CBZ on *Lemna minor*. In figure 47, the concentration of the different drugs inside *Lemna* extract was listed. The uptake of SMX and TRIM was decreased in the MIX compared with single-drug treatment. The results showed that *Lemna* could uptake the drugs. Moreover, the DCF could not be detected in the *Lemna* extracts incubated with the mixture (MIX). Thus, This might be due to drug competing in the *Lemna minor* uptake. The CBZ inhibition of flavonoids and anthocyanin pathways was observed. Moreover, the mixture effect on *Lemna's* metabolic pathway resembled that of CBZ. Based on this activity, *Lemna* can be taken into consideration as a tool for *in situ* remediation of drug-contaminated waters.


Relative Peak Area
6
5
4
3
2
1
0
Apigeinin
Shikimic acid
Tyrosine
Cinnamic acid
SMX
TRIM
Control
DCF
CBZ
MIX

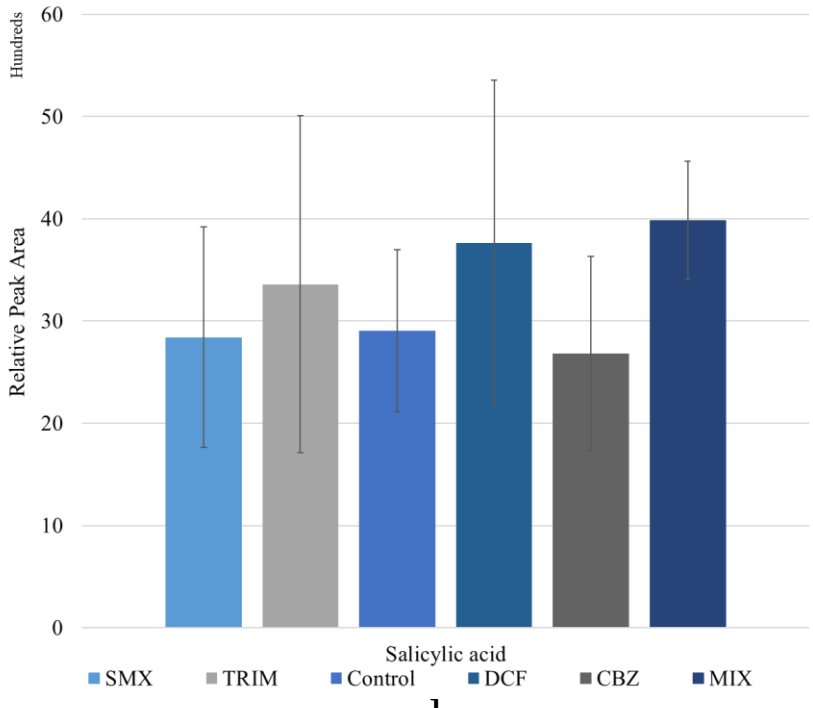
a


Relative Peak Area
600
500
400
300
200
100
0
Chrosimic/Prephenic acid
Phenylpyruvate
p-Coumaric acid
o-Coumaric acid
Benzoic acid
4-hydroxy phenyl pyruvate
Caffeic acid
Ferulic acid
Phenylalanine
Tryptophan
Glutmatic
SMX
TRIM
Control
DCF
CBZ
MIX

b



c



d

Figure 46.The relative peak area of different metabolites of the phenylalanine, tyrosine, and tryptophan biosynthesis, folate biosynthesis, phenylpropanoid pathways, which identified in *Lemna minor* incubated with 5ppb SMX, TRIM, DCF, CBZ, and a mixture of them, besides control for 5 days.

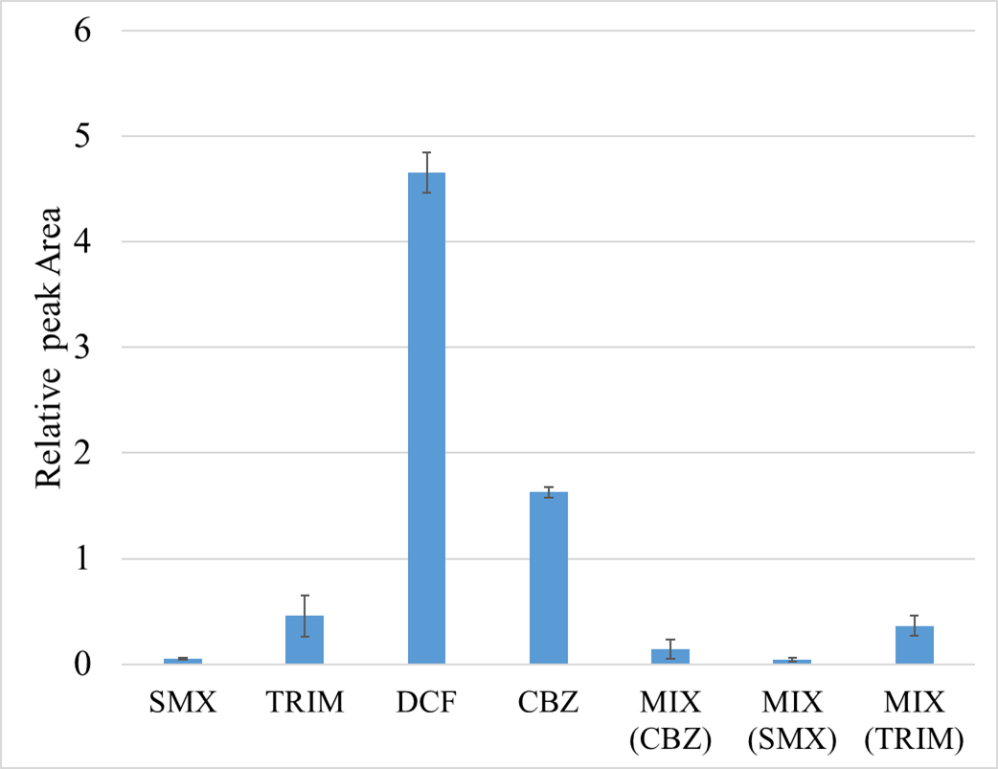


Figure 47. The relative peak area of SMX, TRIM, DCF, CBZ, and a mixture of them identified in *Lemna minor* extracts.

7.3.10. The effect of SMX, TRIM, DCF, CBZ and their mixture on Lemna fatty acid contents

The fatty acid composition of *Lemna minor* was α-linolenic acid 41 to 47% and linoleic acid 17-18% of dry weight. Three fatty acids, palmitic, linoleic acid, and α-linolenic acid comprised more than 80% of total fatty acids (Al-Snafi, 2019). The linoleic acid concentration was investigated quantitatively in incubated samples to demonstrate the effect of different pharmaceuticals on *Lemna's* fatty acid content. The linoleic acid concentration in different samples were in the following order TRIM> MIX > SMX > DCF > control > CBZ (Fig. 48). The linoleic acid was up-modulated in all incubated samples except in CBZ treatment. This indicated the SMX, TRIM, and DCF enhance the fatty acid production in *Lemna*. Further, DCF at higher concentrations caused activation of arachidonic acid metabolism and biosynthesis of unsaturated fatty acids. Furthermore, the untargeted analysis showed an increase of 9,12-octadecadienoic acid (linoleic acid) at a higher concentration confirmed the enhancement in arachidonic acid metabolism. Arachidonic acid is also considered a potent elicitor of programmed cell death and defense responses, as well as induces resistance to viruses via producing plants oxylipins (Savchenko *et al.*, 2010).

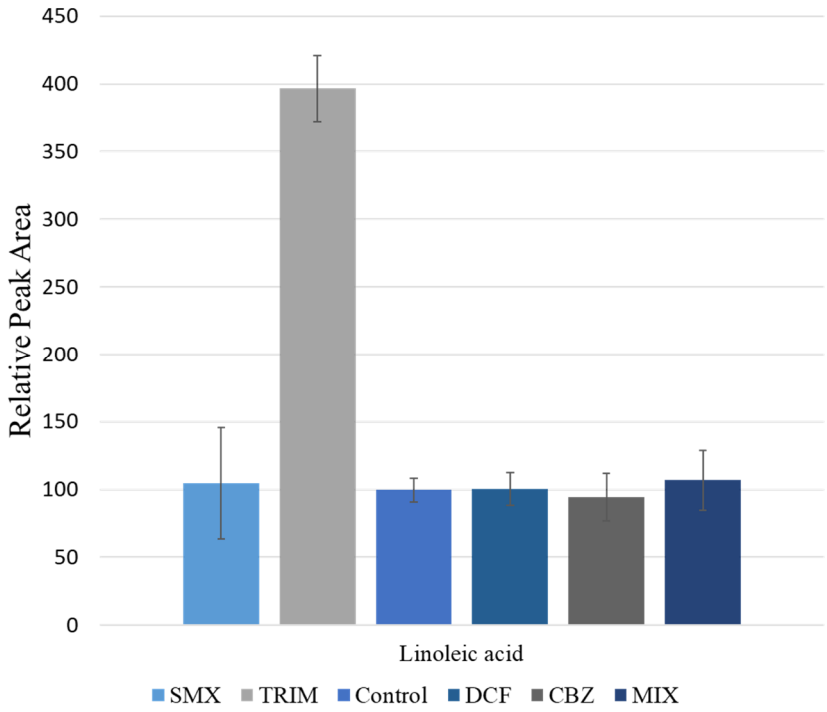


Figure 48. The relative peak area of linoleic acid identified in *Lemna minor* was incubated with 5ppb SMX, TRIM, DCF, CBZ, and a mixture of them, besides control for 5 days.

7.4. Conclusion

The plant metabolomics approach is used for specific metabolic pathways change investigation due to the pharmaceutical's stress effect. The evaluation of the different metabolites in the specific pathway indicates the changes in this pathway. The phenylalanine, tyrosine, and tryptophan biosynthesis, phenylpropanoid pathway were chosen according to the results of the previous chapters. *Lemna* was incubated with SMX, TRIM, DCF, CBZ, and a mixture of them at 5 ppb, individually. The used concentration equals approximately the environmental concentrations reported in the literature.

After 5 days of incubation, *Lemna* samples were extracted with 100%MeOH and analyzed via system C. The results showed the changes in the chosen pathway in incubated *Lemna*. The response of *Lemna* was specific to each drug. SMX and TRIM blocked the folate biosynthesis pathway in *Lemna minor* as in bacteria. Interestingly, 4-aminobenzoic acid was accumulated in *Lemna* incubated with CBZ than SMX and TRIM. This is might be due to the stress effect of CBZ on *Lemna*. Thus, *Lemna* required an enhancement in the carbon flux mechanism and/or the effect of CBZ on *Lemna* DNA. *Lemna* incubated with TRIM, DCF, CBZ, and MIX up modulated the lignin pathway. Hence, the accumulation of lignin is due to the activation of quinate/shikimate p-hydroxycinnamoyltransferase, which is regulated by pathogens in plants

(e.g. *Populus*). Cinnamate 4-hydroxylase and caffeic acid O-methyltransferase also are considered the key enzymes for the downstream metabolites. The enhancement of those three enzymes caused an accumulation in lignin. This might be the reason for lignin accumulation in *Lemna*. SMX, TRIM, and DCF were up-regulated the flavonoids pathway in incubated *Lemna*. Upon DCF treatment, naringenin concentration was approximately equal to the control samples, and apigenin was increased. Moreover, the dihydrokaempferol and taxifolin concentrations were less than the control extracts (Fig. 45c). The higher is the DCF concentration, the higher is the stress effect. Thus, in *Lemna,* an enhancement led to the increase in the phenylpropanoid pathway at 10 and 100 µM DCF treatment.

Thus, the accumulation of phenolic compounds is due to the up-regulation of the biosynthesis of phenylpropanoid enzymes including phenylalanine ammonia-lyase, chalcone synthase, shikimate dehydrogenase, cinnamyl alcohol dehydrogenase, and polyphenol oxidase. The pharmaceuticals competed on the uptake by *Lemna minor*, which could result in the decrease in the uptake of one drug in the presence of the other.

Therefore, the hypothesis **"The presence of pharmaceuticals such as diclofenac and carbamazepine in the aquatic environment can affect the biosynthetic pathways of *Lemna minor*"** could be accepted considering the metabolomics evaluation in different samples analyzed with different analytical systems. However, genetic data analysis will be required to determine the up-regulation and the downregulation of the different genes.

8. Overall conclusions, prospects, and future research challenges

8.1. Conclusions

In this thesis, three hypotheses were discussed, which related to the isolation and identification of plant metabolites with different analytical methods studying the effect of various environmental pollutants on plant metabolism. The hypotheses were investigated using different LC systems coupled to different mass spectrometers. The main findings were presented in chapters (4- 7) based on peer-reviewed publications.

The plant metabolomics approach is a holistic analysis of the plant extracts metabolites, to provide a functional screening of the cellular state. The metabolites pool is a result of the biochemical and physiological state of the plant cell. Additionally, the pharmaceuticals such as diclofenac, carbamazepine, and antibiotics (such as sulfamethoxazole, and trimethoprim) are regularly be detected in the surface water in high concentrations due to incomplete removal using wastewater treatment plants and surface runoffs. Thus, this work provides a systematic study of the effect of different pharmaceuticals on the plant metabolic profile, in consequence, the changes in metabolic pathways.

The method of extraction and the sample preparation optimization parameters were investigated in Chapter 4. The plant extracts were analyzed using different mass spectrometers and workflows, which were discussed in Chapter 5. In Chapter 6, the effects of diclofenac and carbamazepine on the *Lemna minor* and *Phragmites australis* were investigated using uni- and multivariate statistical analysis to determine the changes in their metabolic pathways. Finally, the determined pathways using different statistical analyses were investigated in *Lemna minor* in a small concentration approximately equal to the environmental one in Chapter 7.

The comprehensive discussion of the fundamental thesis objectives illustrates the hypotheses, the designed experiments, and the obtained results. furthermore, future outlooks are given to deeply investigate the changes in the metabolic pathway due to environmental pollutants on the proteomic and genetic levels.

8.1.1. Isolation and Identification of the secondary metabolites from plants

Plants can synthesize a large diversity of organic compounds using inorganic ones, which are over 1,000,000 different metabolites in the plant kingdom. The plan metabolomics is divided according to their function into primary, and secondary (specialized metabolites). Primary metabolites are required for the growth and development of plants. Secondary metabolites support the plant to interact with the biotic and abiotic environment such as phenolics, terpenes, and nitrogen-containing compounds (Erb and Kliebenstein, 2020). Despite that, the role of the secondary metabolites is to protect the plant, the primary metabolites are affected by the different stressors. Plant metabolites are widely different in their physicochemical characters. Moreover, the plants have enzymes that can degrade the metabolites during the extraction and storage of the samples. Consequently, they require an extended-polarity method for extraction

and analysis. The sample preparation step is important in the plant metabolomics analysis, which is performed according to the Metabolomics Standards Initiative (MSI) (Salem *et al.*, 2020). Consequently, hypothesis #1 "Extraction strategy for *Lemna minor* and *Phragmites australis* can influence their metabolites chromatographic fingerprints" was developed and tested using different extraction solvents (mixtures). Moreover, different storage period was used to optimize the sample storage condition.

To test the hypothesis,

a- *Lemna minor* was extracted with five different solvents (mixtures) at three different storage periods. The reducing content and the number of features were compared.

b- Consequently, *Phragmites australis* leaf, rhizome, and root were extracted with four different solvents according to the preliminary results of *Lemna minor*. The reducing content and the number of features were compared. Moreover, a multivariate statistical analysis was performed.

Various *Lemna minor* samples were extracted using different solvents and analyzed using extended-polarity RPLC-HILIC-ESI-TOF-MS determining their chromatographic/mass spectrometric fingerprints.

The results displayed that 100% EtOH has a significantly lower efficient extraction yield than 100% MeOH, as a result of which EtOH was not applied any further in the study. Previous authors were generally more interested in questions concerning 100% MeOH and 100% EtOH extraction power (Eloff, 1998; Miliauskas *et al.* 2004; Sultana *et al.* 2009).

The good extraction efficiencies for *Lemna minor* were obtained with acidic 90% MeOH, but also in 50% MeOH and 100% H_2O solvents. Fresh samples produce higher extraction yields than frozen ones in 100% MeOH, acidic 90% MeOH and 100% H2O extract, respectively. The disparity in the extraction yield of different extracts may be due to the different solubilities of extractable components resulting from their varied chemical compositions. It can be concluded from the above findings that different solvent systems dissolve components differently and that a variety of solutions with a wide range of polarities may be used for maximum extraction. However, it should be borne in mind that a higher extraction yield does not necessarily result in a high reducing content or a higher number of different compounds such as the acidic 90% MeOH extracts of the frozen sample.

Furthermore, the comparison of compounds in *Lemna minor* fresh and frozen samples showed that freezing under -80 ºC of *Lemna* changes the intensity and presence of compounds.

The freezing of *Lemna minor* for short periods generally decreases the intensity of some compounds. However, freezing for long periods generates new compounds. Freezing under -80°C did not serve to arrest the activity in *Lemna minor*. Enzymes were still possibly active and degraded metabolites. Furthermore, they altered the biosynthetic pathways in different ways. Similar findings were reported for onion storage, i.e. the quercetin concentration decreased but its conjugates increased (Sharma and Lee, 2016). Bilia *et al.* reported in 2002 that the concentration of flavonoids in calendula, milk thistle, and passionflower tinctures resulted in different behaviors during storage depending on the type of flavonoids (Bilia *et al.*, 2002).

Researchers reported that the concentration of secondary metabolites was reduced due to freezing at different temperatures (Kapcum and Uriyapongson, 2018).

To avoid any false positive and false negative interpretations, it is necessary to extract *Lemna minor* freshly to determine its metabolomics fingerprint. The results can then be used in the investigation of phytochemicals in different plant extracts, which have various applications in phytomedicine research, cosmetics, food additives, and preserving agents.

The four solvents were used to extract the metabolic profile of *Phragmites australis.* The results displayed that the elimination of acidic 90% MeOH extracts will improve the statistical model as in Lemna minor. This can be described with the different occurrence of charged molecules caused by different pH values. Based on these findings the hypothesis #1 **was accepted**.

8.1.2. Analytical method development and statistical data analysis for plant metabolites using RPLC-HILIC-MS

The advances in MS are conducting untargeted metabolomics analysis through generating an accurate empirical formula, using tandem mass spectrometry with structural information observed by molecule fragmentation and normalizing retention times, and correlating with logD values. In untargeted metabolomics studies, the metabolites identification is the fundamental step to transform the analytical data into biological knowledge, which is still considered the major bottleneck. The number of identified metabolites in untargeted metabolomics studies is in general below 50% (Yang *et al.*, 2014). Thus, hypothesis #2 "Analytical workflows, applied in the determination of trace organic compounds in the aquatic environment, can be used for plant metabolites detection" was developed and tested using different databases and mass spectrometers.

To test the second hypothesis:

a- Different databases (STOFF-IDENT, Metlin, PLANT-IDENT) were used to identify *Lemna minor* metabolites.

b- The obtained *Lemna minor* metabolic profile was compared between different mass spectrometers.

The study produced an extended-polarity untargeted screening workflow that reflects the great importance of detecting (very) polar molecules separated by HILIC in the plant metabolome (in addition to the classic non-polar to polar molecules separated by RPLC using100% MeOH extracts, 50% MeOH, and 100% H_2O extracts to pave the way to new views of plant metabolomes for (very) polar molecules. Different metabolites could be identified and prioritize using different databases. Also, PLANT-IDENT enabled the identification of robinetin and norwogonin for the first time in the *Lemna minor* extracts according to our knowledge.

Moreover, the *Lemna minor* metabolic profile has small common features between the TOF and QTOF-MS. However, the different extracts could be significantly discriminated in both

MS. Regarding the solvent used for the extraction; there is no big difference because it depends on the pKa of metabolites. Further, QTOF was enabled to detect more features than single TOF-MS due to its sensitivity. The prediction model could predict significantly the corresponding MS to each data set. The feature number was the main reason for the discrimination.

The applied workflows in *Lemna minor* and *Phragmites australis* metabolites analysis will be a cornerstone for subsequent research. Despite using different mass spectrometers (systems A and B), researchers can apply this workflow with/or without reference materials to identify suspect and hidden targets. The statistical investigations of the steps of the workflow were conducted through different statistical tests to monitor the metabolite differences between different samples (workflow reliability), which were obtained with the different mass spectrometer. These steps provided additional quality control steps on the applied workflow.

Detailed information of the applied workflow including all parameter settings, and criteria should be required in all studies. Open-access tools, in addition to software (compound and spectral databases), could be a way for the enhancement of untargeted screening. Consequently, the second hypothesis **was accepted.** In the first and second hypothesis, the sample preparation and data analysis workflows parameters optimization guaranteed the untargeted metabolomics analysis of *Lemna minor* and *Phragmites australis* to detect the changes in their metabolic profile due to incubation with DCF and CBZ.

8.1.3. Study the effect of diclofenac and carbamazepine on Lemna minor and Phragmites australis metabolic profiles

Untargeted metabolomics has a fundamental function in determining the metabolic changes in plants due to xenobiotics exposure. Also, the mass spectrometric untargeted metabolomics strategy has a substantial role in investigating the biochemical changes and metabolic adaptation of plants in xenobiotics exposure cases. The changes in plant metabolic profile could be investigated through two independent workflows: (1) the revalent metabolites identification, and (2) the statistical identification of the metabolomics indicators. Thus, hypothesis# 3 “The presence of pharmaceuticals such as diclofenac and carbamazepine in the aquatic environment can affect the biosynthetic pathways of *Lemna minor* and *Phragmites australis”* was developed.

The hypothesis was tested according to the following

a- Untargeted metabolomics analysis of *Phragmites australis* incubated with 10 and 100 µM DCF and 10 & 50 µM CBZ for 96 hrs. was performed using system A, individually.

b- Further, another study of *Lemna* incubated with 10 and 100 µM DCF was performed using systems A and B. Furthermore, the *Lemna* incubation with DCF with prolonged to 8 and 12 days to investigate the successive effect of DCF on *Lemna’s* metabolic profile using system B.

The *Phragmites* study was divided into the target and untargeted analysis. The target analysis resulted in the identification of quercetin, DCF metabolites (DCF, hydroxylated DCF, DM_1, DM_2, DM_3, DM_4, DM_5, DM_6, and DM_7), and CBZ metabolites (carbamazepine-10,11-epoxide, 10,11-dihydro-10,11-dihydroxy-carbamazepine, 10,11-dihydro-10-hydroxy-carbamazepine, 9-acridine carboxaldehyde, and 2,3-dihydro-2,3-dihydroxy-carbamazepine). Carbamazepine-10,11-epoxide is also the first metabolite of carbamazepine in tomato plants and *Armoracia rusticana* root cultures (Riemenschneider *et al.*, 2017b; Sauvêtre *et al.*, 2018).

The statistical analysis revealed that the succinic acid (DM_4), propane-1,2,3-triol (DM_6), and 2- hydroxypropanoic acid (DM_7) were detected as transformation products of DCF. Hence, the increase of succinic acid (DM_4) might affect the TCA cycle (tricarboxylic acid cycle) of *P. australis*. In 2019, Sivaram and coworkers reported a high impact on the TCA cycle in maize leaves exposed to pyrene. The same result was observed in lettuce crops, which were exposed to different contaminants of emerging concern (CEC) concentrations (Hurtado *et al.*, 2017). Also, it was reported that aromatic hydrocarbons altered the osmotic balance in maize (Sivaram *et al.*, 2019).

Untargeted analysis, the multivariate statistical analysis detected the differentiating metabolic profile (DMF), which was used to investigate the *Phragmites australis* metabolism.

The enhanced pathways due to incubation with DCF and CBZ were the glycolysis biosynthesis pathway, subsequently enhanced the pyrimidine metabolism. The first is controlled with the cytochrome P450 enzyme controls glucosinolate biosynthesis, which is involved in the metabolism of DCF and CBZ during phase I (Huber *et al.*, 2012; Lee *et al.*, 2016). Therefore, DCF and CBZ exposure increased in energy consumption (Hurtado *et al.*, 2017). Pyrimidine metabolism is enhanced and leads to products that could be used in the case of salvage, that is, recovery of infections and subsequent synthesis of secondary products with specific functions in defense mechanisms. Additionally, pyrimidine metabolism provides a source of β-alanine or β-aminobutyrate, which might be an important source for the pantothenate of coenzyme A (Zrenner *et al.*, 2006).

After incubation with DCF, *P. australis* exhibited an increase in phenylalanine metabolism which aligned with the increase of quercetin content. The flavonoid 3′-monooxygenase and the flavonoid 3′,5′-hydroxylase enzymes responsible for the conversion of kaempferol to quercetin are cytochrome P450 plant types, which are triggered after incubation with DCF but not in CBZ incubation.

Further, glutathione metabolism was affected in DCF incubation but not under the influence of CBZ. glutathione is critical for the detoxification of xenobiotics, environmental stress tolerance, and, in the form of phytochelatins, also the retention of heavy metals (Gong *et al.*, 2018). The metabolism of DCF required Nicotinamide-Adenine Dinucleotide Phosphate (NADPH) as reductants (Huber *et al.*, 2012). Upon DCF incubation, succinic acid formed from the degradation of diclofenac might have fueled the TCA cycle, which has been detected significantly in extracts incubated with DCF.

The alteration of previously mentioned biosynthetic pathways enhances the *P. australis* defense mechanisms. Hence, *P. australis* responded differently to the DCF and CBZ through changing its metabolic pathway regardless of the type of drug to some extent.

The *Lemna minor* was incubated with DCF at 10 and 100 µM for 96 hr. The target and untargeted analysis showed enhancement in the glucosinolate biosynthesis, aminoacyl-tRNA biosynthesis, porphyrin and chlorophyll metabolism, tryptophan metabolism, sphingolipid metabolism, phenylalanine, tyrosine and tryptophan biosynthesis, cyanoamino acid metabolism, arachidonic acid metabolism, nicotinate and nicotinamide metabolism, phenylpropanoid biosynthesis, biosynthesis of unsaturated fatty acids, purine metabolism, pyrimidine metabolism, cysteine and methionine metabolism, valine, leucine and isoleucine degradation, valine, leucine and isoleucine biosynthesis, arginine and proline metabolism, pyruvate metabolism, monoterpenoid biosynthesis, flavonoids biosynthesis, and anthocyanin biosynthesis. The same data was investigated with MPP software to detect the altered pathways in *Lemna* incubated with DCF at two different concentrations. The results showed three changed pathways, which were glucosinolate, porphyrin and chlorophyll metabolism, and flavonoids biosynthesis. The data investigation with two software, which has different algorithms, confirmed the results, as well as indicating the accuracy of the data.

The targeted analysis showed an enhancement in the phenylpropanoid pathway (phenylalanine up-modulation), which protects it from oxidative stress against DCF. However, flavonoids were reduced in *Lemna gibba,* which was exposed to several environmental challenges (stressor); according to Akhtar *et al.*, 2010. This might be a result of the promotion of the photosynthetic electron transport chain reduction, causing flavonoid reduction. Consequently, the glucosinolate pathway was enhanced, which has a link with the phenylpropanoid pathway due to flavone 3'-*O*-methyltransferase 1. It is responsible for methylation of the 3'-hydroxy group in flavonoids, which exhibits also an activity for methylation of hydroxyl-indole glucosinolates (Chhajed *et al.*, 2020).

Further, *Lemna* preferred the formation of phenylalanine → protocatechuic acid → gallic acid through the β-oxidative pathway when incubated with 100 µM DCF (higher stress). This pathway was observed in the mutant strain of *Neurospora crassa* blocked in the conversion of 5-dehydroshikimic acid into shikimic acid (Dewick and Haslam, 1969).

At lower DCF concentrations (10 µM), *Lemna* required higher energy to tolerate the stress effect of DCF incubation through exhaustion of sugar stores. Also, it has been reported that non-soluble sugars have osmoprotectant and antioxidant activities (Sivaram *et al.*, 2019). However, in 100 µM DCF, sugar concentrations decreased which might be due to the fluctuations of sugars mechanisms that are affected by changes in the genotype (e.g. INV, SuSy, ATB2 bZIP, and α-amylase gene) and stress factors (Rosa *et al.*, 2009). The same was observed in plant response to hypoxic stress, which was accompanied by higher rates of glycolysis and ethanol fermentation causing the fast depletion of sugar stores and carbon stress (Limami *et al.*, 2008). Unsaturated fatty acids (oleic acid, undecanes, and 9,12-octadecadienoic acid) were also increased in 100 µM DCF to protect the *Lemna* against the ROS as a result of stress degradation,

which are produced from DCF and/or its transformation products under a high dose of DCF(Alkimin *et al.*, 2019). Therefore, the oleic acid, undecanes, and 9,12-octadecadienoic acid concentrations also were increased. Since, saturated and unsaturated fatty acids induce broad-spectrum resistance against infections in the plant, such as *Pseudomonas syringae* in tomatoes. Also, unsaturated fatty acids are particularly important in plant defense (Lim *et al.*, 2017). Some of them (stearic, oleic, and palmitic acids) increased in *Lemna* incubated with glyphosate, and glyphosate metribuzin mixture for 72 hrs (Kostopoulou *et al.*, 2020); this might explain the increase in reducing potential, in 100 µM DCF, as a result of unsaturated fatty acid formation/accumulation. The decrease in oleic in 10, 100 µM DCF, however, the increase of 9,12-Octadecadienoic acid (linoleic acid) at higher concentration confirmed the enhancement in arachidonic acid metabolism, which is the main constituents of *Lemna* fatty acid content (Al-Snafi, 2019). Arachidonic acid is also considered a potent elicitor of programmed cell death and defense responses, as well as induces resistance to viruses via producing plants oxylipins (Savchenko *et al.*, 2010).

Moreover, it was reported that DCF decreased the content of photosynthetic pigments, relative fluorescence decay values of chlorophyll in *Lemna* (Alkimin *et al.*, 2019). Consequently, the porphyrin and chlorophyll metabolism was enhanced in *Lemna* incubated with DCF. The pathway was identified with the two software. Finally, *Lemna minor* altered its metabolomics after incubation with DCF increased the antioxidant compounds (phenolic, unsaturated fatty acids). Those compounds could be used as a source for phytochemicals from the plants in a constructed wetland, which can be used in industrial concepts.

Lemna incubated with DCF showed an enhancement in the phenylpropanoid pathway including the flavonoids biosynthesis pathway using two different software. Therefore, the flavonoids contents in *Lemna minor* incubated with 10 and 100 µM DCF for 8 and 12 days were investigated. The results showed an increase in the treatment period enhances the flavonoids pathway; hence increase the activities of the responsible enzymes. The precursors (phenylalanine, 4-methoxycinnamic acid, *p*-coumaric acid, and naringenin) intensities were started to fade after incubation of 12 days. Thus, *Lemna* preferred to synthesize the simple phenols. Their synthesis is less complicated as it requires lower energy and less involved enzymes and cell organelle, which could be the reason. Also, *Lemna* metabolic system was exhausted to degraded and accumulate the higher concentration of DCF. Moreover, the DCF activation to the glycosyltransferases and methyltransferases seems to be decreased with increasing the incubation period. *Lemna's* metabolic system was resistant until the 12^{th} day of treatments; however, it showed a decrease in the final products.

According to the available reference standards, incubation of *Lemna* with CBZ showed an enhancement in the flavonoids pathway especially in the flavonol and flavone biosynthesis, as well as anthocyanin. Flavonoids, antioxidants, play a vital role in the defense mechanism of plants against ROS. The induction in the synthesis of B-ring-substituted seems to be a result of the CBZ effect on *Lemna* chlorophyll, which protects the plants against the light (UV) stress factor.

Concluding, the spectral and statistical results showed changes in the metabolic profile of *Lemna and Phragmites* due to DCF and CBZ incubation after 4 days. The PLS-DA and OPLS-DA analysis identified the significant differences between the controls and incubated samples. The metabolites exhibited changes in their intensities as a response to the incubations, which were related to stress defense mechanisms. Further, significant DMF was determined in *Phragmites australis* after the incubation with DCF or CBZ, individually. Different metabolic pathways were predicted from the statically identified DMF. *Phragmites australis* could use the glutathione metabolism pathway and unsaturated fatty acid pathway to protect themselves during the incubation with DCF and CBZ, respectively. Regarding *Lemna minor*, DCF and CBZ could enhance the intensities of flavonoids as a response to the stress effect directly as an antioxidant against the ROS species, as well as to activate the Phase II enzymes, or because of their effect on the chlorophyll contents of *Lemna minor* (like the response to light stress).

Also, *Lemna* could uptake and accumulate DCF for 12 days of incubation. *Lemna* showed the ability to resist the stress effect of DCF until the 12th of incubation. Furthermore, *Phragmites* and *Lemna's* glucose content was decreased to provide the required energy for both of them to adapt and degraded them.

8.1.4. Identification of changed biosynthetic pathways

The plant metabolomics approach is used for the investigation of specific metabolic pathway changes due to the pharmaceutical's stress effect. The evaluation of the different metabolites in the specific pathway indicates the changes in this pathway. The phenylalanine, tyrosine, and tryptophan biosynthesis, phenylpropanoid pathway were chosen according to the results of the untargeted metabolomics analysis in Chapters 5 and 6. However, the previously used DCF and CBZ concentrations were higher than the environmental ones. Thus, *Lemna* was incubated with SMX, TRIM, DCF, CBZ, and a mixture of them at 5 ppb, individually. *Lemna's* response is based on the drug. Regarding SMX and TRIM, they blocked the folate biosynthesis pathway in *Lemna minor* as in bacteria by the accumulation of 4-aminobenzoic acid. Interestingly, 4-aminobenzoic acid was accumulated at a higher rate incubated with CBZ than SMX and TRIM. This is might be due to the stress effect of CBZ on *Lemna*. Thus, *Lemna* required an enhancement in the carbon flux mechanism and/or the effect of CBZ on *Lemna* DNA. *Lemna* incubated with TRIM, DCF, CBZ, and MIX up-modulated the lignin pathway, which is induced due to pathogen infections. The accumulation of lignin is due to the activation of quinate/shikimate *p*-hydroxycinnamoyltransferase such as in infection in *Populous*. Cinnamate 4-hydroxylase and caffeic acid O-methyl transferase also are considered the key enzymes for the downstream metabolites. The enhancement of those three enzymes caused an accumulation in lignin. Furthermore, SMX, TRIM, and DCF were up-modulated the flavonoids pathway in incubated *Lemna*. Thus, the accumulation of phenolic compounds is due to the up-regulation of the biosynthesis of phenylpropanoid enzymes including phenylalanine ammonia-lyase, chalcone synthase, shikimate dehydrogenase, cinnamyl alcohol dehydrogenase, and polyphenol

oxidase. The pharmaceuticals were competed on the uptake by *Lemna minor*, which could be resulted in the decrease in the uptake of one drug in the presence of the other.

In conclusion, pharmaceuticals affect the metabolic pathways of plants. They change the intensities (i.e. up-modulation or down-modulation) of the plant metabolome. The thesis provides novel insights into the untargeted metabolomics analysis, which allow for comprehensive systematic analysis and essential understanding of the effect of pharmaceuticals on the plant metabolic profile, in consequence, the metabolic pathways. However, there are remaining challenges towards an application on the plants which grow in constructed wetlands.

Therefore, hypothesis # 3 **was accepted.**

8.2. Remaining challenges and suggestions for future research

The plant metabolomics approach has a fundamental function in determining the metabolic changes in plants due to xenobiotics exposure. Also, the mass spectrometric untargeted metabolomics strategy has a substantial role in investigating the biochemical changes and metabolic adaptation of plants in xenobiotics exposure cases. Nevertheless, there are still enormous challenges for future research to investigate the changes in the plant pathways. The key challenges and concepts for process development are outlined in the following.

8.2.1. The availability of the reference standards

Not all assigned pathways using the statistical analysis in this work could be investigated. In untargeted metabolomics studies, the identification of the metabolites is the fundamental step to transform the analytical data into biological knowledge, which is still considered the major bottleneck. Plants contain a wide range of compounds, over 200,000 to 1,000,000 different metabolites in the plant kingdom. Many plant secondary metabolites have been demonstrated to have “specialize” roles for adaptive significance in protection against predator and microbial infection. Therefore, the reference standards to investigate the changes in the biosynthetic pathways are needed. Further, the identification of specialized metabolites is still difficult due to the unavailability of commercial standards. The number of identified metabolites in untargeted metabolomics studies is in general below 50%. Several MS/MS databases have been established to facilitate metabolite annotation. The chemically synthesized important metabolites could be the outlook to investigate other pathways, which can be analyzed in a single run in the serial coupling of RPLC-HILIC.

8.2.2. Considering the genetic data analysis

Despite that, the plant metabolites' biosynthetic pathways and the utilized enzymes are well established. Plants are smart factors, which can exceptionally change their biosynthetic pathway. As various studies have demonstrated that, each plant may respond differently to the different stressor and the same stressor at a different dose. The genomics, transcriptomics, proteomics studies will be required to assign the changes in the DNA of the plant to investigate

the direct and indirect effects of the drug. In addition, the biomarkers for each stressor could be identified.

8.2.3. Transfer to the real samples from constructed wetland plants

The studies on the plants that have been grown in the constructed wetlands are performed mainly to investigate the uptake and accumulation of pollutants. There are few studies concerning plant metabolic pathways. The plants in constructed wetland plants are exposed to many stressors, not only in water but also in the air. The effect of different stressors on the plant metabolic pathway will be different from single drug exposure. The result of *Lemna* incubation with a mixture of SMX, TRIM, DCF, and CBZ was interesting and different from the single drug effect due to the synergistic and/or inhibition effects of the pharmaceuticals on each other.

Up to now, there are huge gaps between lab-grown plants and real samples from constructed wetlands on plant metabolomics changes research. Future research should be directed towards real samples from the constructed wetlands in the different growth stages of the plant to indicate the metabolic pathway changes are continuous during the whole life of the plant or spontaneous processes during the flowering stage.

9. Supplementary Information

9.1. Peer-reviewed journal articles and author contributions

9.1.1. Published manuscripts

1- Wahman, R., Graßmann, J., Sauvêtre, A., Schröder, P., Letzel, T. (2020). *Lemna minor* studies under various storage periods using extended-polarity extraction and metabolite non-target screening analysis. Journal of Pharmaceutical and Biomedical Analysis 188:113362.

Author contributions:

Rofida Wahman and Thomas Letzel conceptualized the research objective and designed the methodology Andrés Sauvêtre, and Peter Schröder created the test set and performed the plant growing and incubation together with Rofida Wahman. Rofida Wahman collected and analyzed the data. Rofida Wahman applied and validated the models and wrote the paper. Johanna Graßmann, Andrés Sauvêtre, Peter Schröder, and Thomas Letzel reviewed the manuscript. All authors approved the final version of the manuscript.

2-Wahman, R., Graßmann, J., Sauvêtre, A., Schröder, P., Moser, S., Letzel, T. (2021). Untargeted metabolomics studies on drug-incubated *Phragmites australis* profiles. Metabolites 11(1):2.

Author contributions:

Rofida Wahman and Thomas Letzel designed the metabolomics study. Rofida Wahman prepared the samples, performed the untargeted metabolomics analysis, and ran the analyses on the metabolomics platform. Stefan Moser created the statistical design for data evaluation and data interpretation and performed the realization together with Rofida Wahman and Thomas Letzel. Andrés Sauvêtre, Peter Schröder created the test set and performed the plant growing and incubation together with Rofida Wahman. They also supported Rofida Wahman in biological data interpretation. Rofida Wahman and Thomas Letzel conceived and drafted the manuscript. All the authors contributed with critical intellectual input. All authors have read and agreed to the published version of the manuscript.

3- Wahman, R., Cruzeiro, C., Grassmann, J., Schröder, P., Letzel, T. (2021). The changes in *Lemna minor* metabolomic profile: A response to diclofenac incubation. Chemosphere 287, 13207.

Author contributions:

Rofida Wahman, Catarina Cruzeiro, Peter Schröder and Thomas Letzel conceptualized the research objective. Rofida Wahman prepared the samples, performed the untargeted metabolomics analysis, and ran the analyses on the metabolomics platform. Rofida Wahman wrote the paper. Catarina Cruzeiro, and Peter Schröder created the test set and performed the plant growing and incubation together with Rofida Wahman. Rofida Wahman and Peter Schröder conceptualized and analyzed the biochemical part. Johanna Grassmann reviewed the

manuscript. Rofida Wahman, Catarina Cruzeiro, Peter Schröder and Thomas Letzel critically reviewed the manuscript.

9.1.2. Submitted manuscripts

1- Wahman, R., Moser, S., Bieber, S., Cruzeiro, C., Schröder, P., Gilg, A., Lesske, F., Letzel, T. (2021). Untargeted analysis of *Lemna minor* metabolites: workflow and prioritization strategy comparing highly confident features between different mass spectrometers. Submitted to Metobiltes.

Author contributions:

Rofida Wahman and Thomas Letzel designed the metabolomics study. Rofida Wahman prepared the samples, performed the untargeted metabolomics analysis, and ran the analyses on the metabolomics platform. Stefan Moser created the statistical design for data evaluation and data interpretation and performed the realization together with Rofida Wahman and Thomas Letzel. Catarina Cruzeiro and Peter Schröder created the test set and performed the plant growing together with Rofida Wahman. Rofida Wahman and Thomas Letzel conceived and drafted the manuscript. August Gilg and Frank Lesske are the IT specialists of the FOR-IDENT platform. All the authors contributed with critical intellectual input. All authors have read and agreed to the published version of the manuscript.

9.1.3. Manuscript in preparation

1-The *Lemna minor* study using system B is under-preparation for submission into Metabolomics Journal as follows:

Wahman, R., Schröder, P., Cruzeiro, C., Drewes, J., Letzel, T. (2021). The effect of prolonged diclofenac incubation on *Lemna minor* metabolomic profile: flavonoids content. Metabolomics.

Author contributions:

Rofida Wahman, Peter Schröder, Catarina Cruzeiro, and Thomas Letzel conceptualized the research objective. Rofida Wahman prepared the samples, performed the untargeted metabolomics analysis, and ran the analyses on the metabolomics platform. Rofida Wahman wrote the paper. Catarina Cruzeiro, Peter Schröder created the test set and performed the plant growing and incubation together with Rofida Wahman. Rofida Wahman and Peter Schröder conceptualized and analyzed the biochemical part. Jörg Drewes reviewed the manuscript. Rofida Wahman, Peter Schröder, Catarina Cruzeiro, and Jörg Drewes Thomas Letzel will critically review the manuscript.

9.2. Non-peer-reviewed journal articles and author contributions

1-Wahman, R., Graßmann, J., Schröder, P., Letzel, T. (2019). Plant metabolomics workflows using reversed-phase LC and HILIC with ESI-TOF-MS

Current Trends in Mass Spectrometry, LCGC. N. Am., 37 (3), 8-15.

Author contributions:

Rofida Wahman and Thomas Letzel conceptualized the research objective and designed the methodology. Rofida Wahman and Peter Schröder created the test set and performed the plant growing. Rofida Wahman collected and analyzed the data. Rofida Wahman applied and validated the models and wrote the paper. Johanna Graßmann, Peter Schröder, and Thomas Letzel reviewed the manuscript. All authors approved the final version of the manuscript.

9.3.First author contributions to national and international conferences

9.3.1. Presentations

1- Wahman, R., Graßmann, J., Schröder, P., Letzel, T. (2018). The effect of various pharmaceuticals on plant metabolites. Second Munich Metabolomics Meeting. Munich (Germany).
2- Wahman, R., Bieber, S., Letzel, T. (2020). Non-target screening workflow of plant metabolomes to reveal changes caused by an inflammatory drug incubation. The 27th ITP electrophoretic and liquid-phase separations (online).

9.3.2. Poster

1- Wahman, R., Graßmann, J., Schröder, P., and Letzel, T. (2017). Novel analytical strategies for anthropogenic compounds in plants: vegetable biomonitors for water contaminations. Langenau (Germany).
2- Wahman, R., Graßmann, J., Schröder, P., and Letzel, T. (2018). Novel analytical strategies for anthropogenic compounds in plants: vegetable biomonitors for contaminants in the environment.' SETAC (Italy).
3- Wahman, R., Minkus, S., and Letzel, T. (2019). Non-target screening workflow using the open-access compound database PLANT-IDENT for the identification of biogenic compounds in surface water. Langenau (Germany).
4- Wahman, R., and Letzel, T. (2019). The workflow of an open-access platform Plant-Ident in the identification of *Lepidium sativum* L. (garden cress) metabolites. 3rd Munich Metabolomics Meeting. Munich (Germany).
5- Wahman, R., Bieber, S., Letzel, T. (2020). Non-target screening workflow of plant metabolomes to reveal changes caused by antiepileptic drug incubation. Metabolomics (online).

References

Adamia, G., Chogovadze, M., Chokheli, L., Gigolashvili, G., Gordeziani, M., Khatisashvili, G., Kurashvili, M., Pruidze, M., Varazi, T., 2018. About the possibility of alga *Spirulina* application for phytoremediation of water polluted with 2,4,6-trinitrotoluene. Annals of Agrarian Science.

Ahammed, G.J., He, B.-B., Qian, X.-J., Zhou, Y.-H., Shi, K., Zhou, J., Yu, J.-Q., Xia, X.-J., 2017. 24-Epibrassinolide alleviates organic pollutants-retarded root elongation by promoting redox homeostasis and secondary metabolism in *Cucumis sativus* L. Environmental Pollution 229, 922-931.

Akhtar, N., Thadhani, V.M., Ul Haq, F., Khan, M.N., Ali, S., Musharraf, S.G., 2020. Rapid identification and quantification of bioactive metabolites in processed *Camellia sinensis* samples by UHPLC-ESI-MS/MS and evaluation of their antioxidant activity. Journal of Industrial and Engineering Chemistry 90, 419-426.

Akhtar, T.A., Lees, H.A., Lampi, M.A., Enstone, D., Brain, R.A., Greenberg, B.M., 2010. Photosynthetic redox imbalance influences flavonoid biosynthesis in *Lemna gibba*. Plant, Cell & Environment 33, 1205-1219.

Al-Snafi, A., 2019. *Lemna minor*: traditional uses, chemical constituents, and pharmacological effects-A review. 9, 6-11.

Aliferis, K.A., Materzok, S., Paziotou, G.N., Chrysayi-Tokousbalides, M., 2009. *Lemna minor* L. as a model organism for ecotoxicological studies performing 1H NMR fingerprinting. Chemosphere 76, 967-973.

Alkimin, G.D., Daniel, D., Dionísio, R., Soares, A., Barata, C., Nunes, B., 2019. Effects of diclofenac and salicylic acid exposure on *Lemna minor*: Is time a factor? Environmental Research 177, 108609.

Ameer, K., Shahbaz, H.M., Kwon, J.-H., 2017. Green extraction methods for polyphenols from plant matrices and their byproducts: A review. Comprehensive Reviews in Food Science and Food Safety 16, 295-315.

Appenroth, K.-J., Sree, K.S., Bog, M., Ecker, J., Seeliger, C., Böhm, V., Lorkowski, S., Sommer, K., Vetter, W., Tolzin-Banasch, K., Kirmse, R., Leiterer, M., Dawczynski, C., Liebisch, G., Jahreis, G., 2018. Nutritional value of the duckweed species of the genus *Wolffia* (Lemnaceae) as human food. Frontiers in Chemistry 6.

Araújo, W.L., Nunes-Nesi, A., Nikoloski, Z., Sweetlove, L.J., Fernie, A.R., 2012. Metabolic control and regulation of the tricarboxylic acid cycle in photosynthetic and heterotrophic plant tissues. Plant Cell Environ 35, 1-21.

Aretz, I., Meierhofer, D., 2016. Advantages and pitfalls of mass spectrometry-based metabolome profiling in systems biology. Int J Mol Sci 17.

Aronsson, P., Perttu, K., 2001. Willow vegetation filters for wastewater treatment and soil remediation combined with biomass production. The Forestry Chronicle 77, 293-299.

Azwanida, N., 2015. A review on the extraction methods uses in medicinal plants, principle, strength, and limitation. Medicinal and Aromatic plants 4, 1-6.

Bao, J., Ding, R.B., Jia, X., Liang, Y., Liu, F., Wang, K., Zhang, C., Li, P., Wang, Y., Wan, J.B., He, C., 2018. Fast identification of anticancer constituents in *Forsythiae fructus* based on metabolomics approaches. J Pharm Biomed Anal 154, 312-320.

Bartha, B., Huber, C., Schröder, P., 2014. Uptake and metabolism of diclofenac in *Typha latifolia* – How plants cope with human pharmaceutical pollution. Plant Science 227, 12-20.

Baxter, C.J., Redestig, H., Schauer, N., Repsilber, D., Patil, K.R., Nielsen, J., Selbig, J., Liu, J., Fernie, A.R., Sweetlove, L.J., 2007. The metabolic response of heterotrophic *Arabidopsis* cells to oxidative stress. Plant Physiol 143, 312-325.

Bedia, C., Cardoso, P., Dalmau, N., Garreta-Lara, E., Gómez-Canela, C., Gorrochategui, E., Navarro-Reig, M., Ortiz-Villanueva, E., Puig-Castellví, F., Tauler, R., 2018. Chapter Nineteen - Applications of metabolomics analysis in environmental research. in: Jaumot, J., Bedia, C., Tauler, R. (Eds.). Comprehensive Analytical Chemistry. Elsevier, pp. 533-582.

Bieber, S., Greco, G., Grosse, S., Letzel, T., 2017. RPLC-HILIC and SFC with mass spectrometry: polarity-extended organic molecule screening in environmental (Water) samples. Analytical chemistry 89, 7907-7914.

Bigott, Y., Chowdhury, S.P., Pérez, S., Montemurro, N., Manasfi, R., Schröder, P., 2021. Effect of the pharmaceuticals diclofenac and lamotrigine on stress responses and stress gene expression in lettuce (*Lactuca sativa*) at environmentally relevant concentrations. J. Hazard. Mater. 403.

Bilia, A.R., Bergonzi, M.C., Gallori, S., Mazzi, G., Vincieri, F.F., 2002. Stability of the constituents of *Calendula*, milk-thistle, and passionflower tinctures by LC-DAD and LC-MS. Journal of pharmaceutical and biomedical analysis 30, 613-624.

Blasco, H., Błaszczyński, J., Billaut, J.C., Nadal-Desbarats, L., Pradat, P.F., Devos, D., Moreau, C., Andres, C.R., Emond, P., Corcia, P., Słowiński, R., 2015. Comparative analysis of targeted metabolomics: Dominance-based rough set approach versus orthogonal partial least square-discriminant analysis. Journal of Biomedical Informatics 53, 291-299.

Bourioug, M., Mazzitelli, J.Y., Marty, P., Budzinsky, H., Aleya, L., Bonnafé, E., Geret, F., 2018. Assessment of *Lemna minor* (duckweed) and *Corbicula fluminea* (freshwater clam) as potential indicators of contaminated aquatic ecosystems: responses to the presence of psychoactive drug mixtures. Environmental science and pollution research international 25, 11192-11204.

Brain, R.A., Cedergreen, N., 2009. Biomarkers in aquatic plants: selection and utility. Reviews of environmental contamination and toxicology 198, 49-109.

Brain, R.A., Hanson, M.L., Solomon, K.R., Brooks, B.W., 2008a. Aquatic Plants Exposed to Pharmaceuticals: Effects and Risks. in: Whitacre, D.M. (Ed.). Reviews of environmental contamination and toxicology. Springer New York, New York, NY, pp. 67-115.

Brain, R.A., Ramirez, A.J., Fulton, B.A., Chambliss, C.K., Brooks, B.W., 2008b. Herbicidal effects of sulfamethoxazole in *Lemna gibba*: using p-aminobenzoic acid as a biomarker of effect. Environ Sci Technol 42, 8965-8970.

Brain, R.A., Solomon, K.R., Brooks, B.W., 2009. Targets, Effects and Risks in Aquatic Plants Exposed to Veterinary Antibiotics. Veterinary Pharmaceuticals in the Environment. American Chemical Society, pp. 169-189.

Bucar, F., Wube, A., Schmid, M., 2013. Natural product isolation--how to get from biological material to pure compounds. Natural product reports 30, 525-545.

Buchanan, B.B., Gruissem, W., Jones, R.L., 2015. Biochemistry and Molecular Biology of Plants. Wiley.

Capasso, C., Supuran, C.T., 2014. Sulfa and trimethoprim-like drugs – antimetabolites acting as carbonic anhydrase, dihydropteroate synthase, and dihydrofolate reductase inhibitors. Journal of Enzyme Inhibition and Medicinal Chemistry 29, 379-387.

Chakrabarti, R., Clark, W.D., Sharma, J.G., Goswami, R.K., Shrivastav, A.K., Tocher, D.R., 2018a. Mass production of *Lemna minor* and its amino acid and fatty acid profiles. Frontiers in Chemistry 6, 479.

Champagne, A., Hilbert, G., Legendre, L., Lebot, V., 2011. Diversity of anthocyanins and other phenolic compounds among tropical root crops from Vanuatu, South Pacific. Journal of Food Composition and Analysis 24, 315-325.

Chen, L., Wu, J.e., Li, Z., Liu, Q., Zhao, X., Yang, H., 2019. Metabolomic analysis of energy regulated germination and sprouting of organic mung bean (*Vigna radiata*) using NMR spectroscopy. Food Chemistry 286, 87-97.

Cheong, M.S., Seo, K.H., Chohra, H., Yoon, Y.E., Choe, H., Kantharaj, V., Lee, Y.B., 2020. Influence of sulfonamide contamination derived from veterinary antibiotics on plant growth and development. Antibiotics 9.

Chhajed, S., Mostafa, I., He, Y., Abou-Hashem, M., El-Domiaty, M., Chen, S., 2020. Glucosinolate biosynthesis and the glucosinolate-myrosinase system in plant defense. Agronomy 10.

Chong, J., Wishart, D.S., Xia, J., 2019a. Using MetaboAnalyst 4.0 for comprehensive and integrative metabolomics data analysis. Current protocols in bioinformatics 68, e86.

Chong, J., Yamamoto, M., Xia, J., 2019b. MetaboAnalystR 2.0: From raw spectra to biological insights. Metabolites 9.

Chua, M., Baldwin, T.C., Hocking, T.J., Chan, K., 2010. Traditional uses and potential health benefits of *Amorphophallus konjac* K. Koch ex N.E.Br. Journal of ethnopharmacology 128, 268-278.

Chutia, R., Abel, S., Ziegler, J., 2019. Iron and phosphate deficiency regulators concertedly control coumarin profiles in *Arabidopsis thaliana* roots during iron, phosphate, and combined deficiencies. Frontiers in Plant Science 10.

Cleuvers, M., 2003. Aquatic ecotoxicity of pharmaceuticals including the assessment of combination effects. Toxicology letters 142, 185-194.

Consonni, R., Cagliani, L.R., 2019. The potentiality of NMR-based metabolomics in food science and food authentication assessment. Magnetic Resonance in Chemistry 57, 558-578.

Crow, F.W., Tomer, K.B., Looker, J.H., Gross, M.L., 1986. Fast atom bombardment and tandem mass spectrometry for structure determination of steroid and flavonoid glycosides. Analytical Biochemistry 155, 286-307.

Cuperlovic-Culf, M., Culf, A.S., 2016. Applied metabolomics in drug discovery. Expert Opinion on Drug Discovery 11, 759-770.

De Vos, R., Schipper, B., Hall, R., 2012. High-performance liquid chromatography-mass spectrometry analysis of plant metabolites in Brassicaceae. Methods in molecular biology (Clifton, N.J.) 860, 111-128.

Delauney, A.J., Verma, D.P.S., 1993. Proline biosynthesis and osmoregulation in plants. The Plant Journal 4, 215-223.

Desbiolles, F., Moreau, X., de Jong, L., Malleret, L., Grandet-Marchant, Q., Wong-Wah-Chung, P., Laffont-Schwob, I., 2020. Advances and limits of two model species for ecotoxicological assessment of carbamazepine, two by-products, and their mixture at the environmental level in freshwater. Water Res 169, 115267.

Dewick, P.M., Haslam, E., 1969. Phenol biosynthesis in higher plants. Gallic acid. Biochemical Journal 113, 537-542.

Donnelly, P.K., Hegde, R.S., Fletcher, J.S., 1994. Growth of PCB-degrading bacteria on compounds from photosynthetic plants. Chemosphere 28, 981-988.

Doppler, M., Bueschl, C., Kluger, B., Koutnik, A., Lemmens, M., Buerstmayr, H., Rechthaler, J., Krska, R., Adam, G., Schuhmacher, R., 2019. Stable isotope-assisted plant metabolomics: combination of global and tracer-based labeling for enhanced untargeted profiling and compound annotation. Frontiers in Plant Science 10.

Drobniewska, A., Wójcik, D., Kapłan, M., Adomas, B., Piotrowicz-Cieślak, A., Nałęcz-Jawecki, G., 2017. Recovery of *Lemna minor* after exposure to sulfadimethoxine irradiated and non-irradiated in a solar simulator. Environmental science and pollution research international 24, 27642-27652.

Dzantor, E.K., Woolston, J.E., 2001. Enhancing dissipation of Aroclor 1248 (PCB) using substrate amendment in rhizosphere soil. Journal of Environmental Science and Health, Part A 36, 1861-1871.

Ekperusi, A., Sikoki, F., Nwachukwu, E., 2019. Application of common duckweed (*Lemna minor*) in phytoremediation of chemicals in the environment: State and future perspective. Chemosphere 223, 285-309.

Eloff, J.N., 1998. Which extractant should be used for the screening and isolation of antimicrobial components from plants? Journal of ethnopharmacology 60, 1-8.

Erb, M., Kliebenstein, D.J., 2020. Plant secondary metabolites as defenses, regulators, and primary metabolites: The Blurred Functional Trichotomy. Plant Physiol 184, 39-52.

Ernst, M., Silva, D.B., Silva, R.R., Vêncio, R.Z.N., Lopes, N.P., 2014. Mass spectrometry in plant metabolomics strategies: from analytical platforms to data acquisition and processing. Natural Product Reports 31, 784-806.

Ferreres, F., Gonçalves, R.F., Gil-Izquierdo, A., Valentão, P., Silva, A.M.S., Silva, J.B., Santos, D., Andrade, P.B., 2012. Further knowledge on the phenolic profile of *Colocasia esculenta* (L.) Shott. J Agric Food Chem 60, 7005-7015.

Figueiredo, A.C., Barroso, J.G., Pedro, L.G., Scheffer, J.J.C., 2008. Factors affecting secondary metabolite production in plants: volatile components and essential oils. Flavour and Fragrance Journal 23, 213-226.

Fischer, K., Sydow, S., Griebel, J., Naumov, S., Elsner, C., Thomas, I., Abdul Latif, A., Schulze, A., 2020. Enhanced removal and toxicity decline of diclofenac by combining UVA treatment and adsorption of photoproducts to polyvinylidene difluoride. Polymers 12.

Forni, C., Braglia, R., Harren, F.J., Cristescu, S.M., 2012. Stress responses of duckweed (*Lemna minor* L.) and water velvet (*Azolla filiculoides* Lam.) to anionic surfactant sodium-dodecyl-sulfate (SDS). Aquatic toxicology (Amsterdam, Netherlands) 110-111, 107-113.

Franklin, A.M., Williams, C.F., Andrews, D.M., Woodward, E.E., Watson, J.E., 2016. Uptake of three antibiotics and an antiepileptic drug by Wheat crops spray irrigated with wastewater treatment plant effluent. Journal of environmental quality 45, 546-554.

Freiwald, V., Häikiö, E., Julkunen-Tiitto, R., Holopainen, J.K., Oksanen, E., 2008. Elevated ozone modifies the feeding behavior of the common leaf weevil on hybrid aspen through shifts in developmental, chemical, and structural properties of leaves. Entomologia Experimentalis et Applicata 128, 66-72.

Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grün, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R.B., Briggs, S.P., Simcox, K., Gierl, A., 1997. Analysis of a chemical plant defense mechanism in Grasses. Science 277, 696.

Fukushima, A., Kusano, M., 2013. Recent progress in the development of metabolome databases for Plant systems biology. Frontiers in Plant Science 4.

Galili, G., 2011. The aspartate-family pathway of plants: linking production of essential amino acids with energy and stress regulation. Plant signaling & behavior 6, 192-195.

Gao, X., Sun, W., Fu, Q., Niu, X., 2014. Ultra-performance liquid chromatography coupled with electrospray ionization/quadrupole time-of-flight mass spectrometry for the rapid analysis of constituents in the traditional Chinese medical formula Danggui San. Journal of Separation Science 37, 53-60.

Giebułtowicz, J., Nałęcz-Jawecki, G., Harnisz, M., Kucharski, D., Korzeniewska, E., Płaza, G., 2020. Environmental risk and risk of resistance selection due to antimicrobials' occurrence in two polish wastewater treatment plants and receiving surface Water. Molecules (Basel, Switzerland) 25.

Gika, H.G., Theodoridis, G.A., Earll, M., Snyder, R.W., Sumner, S.J., Wilson, I.D., 2010. Does the mass spectrometer define the marker? A comparison of global metabolite profiling data generated simultaneously via UPLC-MS on two different mass spectrometers. Analytical chemistry 82, 8226-8234.

Gong, B., Sun, S., Yan, Y., Jing, X., Shi, Q., 2018. Glutathione metabolism and its function in higher plants adapting to stress. in: Gupta, D.K., Palma, J.M., Corpas, F.J. (Eds.). Antioxidants and Antioxidant Enzymes in Higher Plants. Springer International Publishing, Cham, pp. 181-205.

Gorelova, V., Ambach, L., Rébeillé, F., Stove, C., Van Der Straeten, D., 2017. Folates in plants: research advances and progress in crop biofortification. Frontiers in chemistry 5, 21-21.

Gray, K.R., Biddlestone, A.J., 1995. Engineered reed-bed systems for wastewater treatment. Trends in Biotechnology 13, 248-252.

Greco, G., Grosse, S., Letzel, T., 2013. Serial coupling of reversed-phase and zwitterionic hydrophilic interaction LC/MS for the analysis of polar and nonpolar phenols in wine. Journal of separation science 36, 1379-1388.

Gromski, P.S., Xu, Y., Kotze, H.L., Correa, E., Ellis, D.I., Armitage, E.G., Turner, M.L., Goodacre, R., 2014. Influence of missing values substitutes on multivariate analysis of metabolomics data. Metabolites 4, 433-452.

Grote, R., Niinemets, U., 2008. Modeling volatile isoprenoid emissions--a story with split ends. Plant biology (Stuttgart, Germany) 10, 8-28.

Grunwald, C., Endress, A.G., 1985. Foliar sterols in soybeans are exposed to chronic levels of ozone. Plant Physiol 77, 245-247.

Habib, M., Trajkovic, M., Fraaije, M.W., 2018. The biocatalytic synthesis of syringaresinol from 2,6-Dimethoxy-4-allylphenol in one pot using a tailored oxidase/peroxidase system. ACS Catalysis 8, 5549-5552.

Hacham, Y., Matityahu, I., Amir, R., 2017. Transgenic tobacco plants having a higher level of methionine are more sensitive to oxidative stress. Physiologia Plantarum 160, 242-252.

Haggarty, J., Burgess, K.E.V., 2017. Recent advances in liquid and gas chromatography methodology for extending coverage of the metabolome. Current Opinion in Biotechnology 43, 77-85.

Hai, F.I., Yang, S., Asif, M.B., Sencadas, V., Shawkat, S., Sanderson-Smith, M., Gorman, J., Xu, Z.-Q., Yamamoto, K., 2018. Carbamazepine as a possible anthropogenic marker in water: occurrences, toxicological effects, regulations, and removal by wastewater treatment technologies. Water 10, 107.

Hájková, M., Kummerová, M., Zezulka, Š., Babula, P., Váczi, P., 2019. Diclofenac as an environmental threat: Impact on the photosynthetic processes of *Lemna minor* chloroplasts. Chemosphere 224, 892-899.

Hammond, J.B.W., Burrell, M.M., Kruger, N.J., 1990. Effect of low temperature on the activity of phosphofructokinase from potato tubers. Planta 180, 613-616.

Harborne, J.B., Williams, C.A., 1982. Flavone and Flavonol Glycosides. in: Harborne, J.B., Mabry, T.J. (Eds.). The Flavonoids: Advances in research. Springer US, Boston, MA, pp. 261-311.

Hart, D.A., Kindel, P.K., 1970. A novel reaction involved in the degradation of apiogalacturonans from *Lemna minor* and the isolation of apibiose as a product. Biochemistry 9, 2190-2196.

Hau, J., Stadler, R., Jenny, T.A., Fay, L.B., 2001. Tandem mass spectrometric accurate mass performance of time-of-flight and Fourier transform ion cyclotron resonance mass spectrometry: a case study with pyridine derivatives. Rapid Communications in Mass Spectrometry 15, 1840-1848.

Hemmler, D., Heinzmann, S.S., Wöhr, K., Schmitt-Kopplin, P., Witting, M., 2018. Tandem HILIC-RP liquid chromatography for increased polarity coverage in food analysis. Electrophoresis 39, 1645-1653.

Hess, D., 1975. Polarity and Unequal Cell Division as Fundamentals of Differentiation. in: Hess, D. (Ed.). Plant physiology: Molecular, biochemical, and physiological fundamentals of metabolism and development. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 225-230.

Heyman, H.M., Dubery, I.A., 2016. The potential of mass spectrometry imaging in plant metabolomics: a review. Phytochemistry Reviews 15, 297-316.

Huber, C., Bartha, B., Harpaintner, R., Schröder, P., 2009. Metabolism of acetaminophen (paracetamol) in plants-two independent pathways result in the formation of glutathione and a glucose conjugate. Environmental Science and Pollution Research 16, 206.

Huber, C., Bartha, B., Schröder, P., 2012. Metabolism of diclofenac in plants-Hydroxylation is followed by glucose conjugation. Journal of Hazardous Materials 243, 250-256.

Hughes, E.H., Shanks, J.V., 2002. Metabolic engineering of plants for alkaloid production. Metabolic engineering 4, 41-48.

Hughes, R.J., Croley, T.R., Metcalfe, C.D., March, R.E., 2001. A tandem mass spectrometric study of selected characteristic flavonoids. International Journal of Mass Spectrometry 210-211, 371-385.

Hurtado, C., Parastar, H., Matamoros, V., Piña, B., Tauler, R., Bayona, J.M., 2017. Linking the morphological and metabolomic response of *Lactuca sativa* L exposed to emerging contaminants using GC × GC-MS and chemometric tools. Sci Rep 7, 6546.

Hussain, S., Naeem, M., Chaudhry, M.N., Iqbal, M.A., 2016. Accumulation of residual antibiotics in the vegetables irrigated by pharmaceutical wastewater. Exposure and Health 8, 107-115.

Igamberdiev, A.U., Kleczkowski, L.A., 2018. The glycerate and phosphorylated pathways of serine synthesis in plants: The branches of plant glycolysis linking carbon and nitrogen metabolism. Frontiers in Plant Science 9.

Iriti, M., Faoro, F., 2009. Chemical diversity and defense metabolism: how plants cope with pathogens and ozone pollution. Int J Mol Sci 10, 3371-3399.

Iwasaki, Y., Sawada, T., Hatayama, K., Ohyagi, A., Tsukuda, Y., Namekawa, K., Ito, R., Saito, K., Nakazawa, H., 2012. Separation technique for the determination of highly polar metabolites in biological samples. Metabolites 2, 496-515.

Jackson, D.M., Rufty, T.W., Heagle, A.S., Severson, R.F., Eckel, R.V.W., 2000. Survival and development of Tobacco hornworm larvae on Tobacco plants grown under elevated levels of ozone. Journal of Chemical Ecology 26, 1-19.

Jain, A., Sundriyal, M., Roshnibala, S., Kotoky, R., Kanjilal, P., Singh, H., Sundriyal, R.C., 2011. Dietary use and conservation concern of edible wetland plants at Indo-Burma hotspot: A case study from northeast India. Journal of ethnobiology and ethnomedicine 7, 29.

Jang, G.H., Kim, H.W., Lee, M.K., Jeong, S.Y., Bak, A.R., Lee, D.J., Kim, J.B., 2018. Characterization and quantification of flavonoid glycosides in the *Prunus* genus by UPLC-DAD-QTOF/MS. Saudi Journal of Biological Sciences 25, 1622-1631.

Jewell, K.S., Castronovo, S., Wick, A., Falås, P., Joss, A., Ternes, T.A., 2016. New insights into the transformation of trimethoprim during biological wastewater treatment. Water Res 88, 550-557.

Jin, P., Wang, S.Y., Wang, C.Y., Zheng, Y., 2011. Effect of cultural system and storage temperature on antioxidant capacity and phenolic compounds in strawberries. Food chemistry 124, 262-270.

Joshi, S., 2002. HPLC separation of antibiotics present in formulated and unformulated samples. J Pharm Biomed Anal 28, 795-809.

Kainulainen, P., Holopainen, J.K., Oksanen, J., 1995. Effects of SO_2 on the concentrations of carbohydrates and secondary compounds in Scots pine (*Pinus sylvestris* L.) and Norway spruce (*Picea abies* (L.) Karst.) seedlings. New Phytologist 130, 231-238.

Kairigo, P., Ngumba, E., Sundberg, L.-R., Gachanja, A., Tuhkanen, T., 2020. Contamination of surface water and river sediments by antibiotic and antiretroviral drug cocktails in low and middle-income countries: occurrence, risk and mitigation strategies. Water 12, 1376.

Kandeler, R., 2019. Entwicklungsphysiologie der Pflanzen. De Gruyter.

Kapcum, C., Uriyapongson, J., 2018. Effects of storage conditions on phytochemical and stability of purple corn cob extract powder. Food Science and Technology 38, 301-305.

Kaufmann, C.M., Grassmann, J., Letzel, T., 2016. HPLC method development for the online coupling of chromatographic *Perilla frutescens* extracts separation with xanthine oxidase enzymatic assay. J Pharm Biomed Anal 124, 347-357.

Kaufmann, C.M., Letzel, T., Grassmann, J., Pfaffl, M.W., 2017. Effect of *Perilla frutescens* extracts on porcine jejunal epithelial cells. Phytotherapy research: PTR 31, 303-311.

Kesselmeier, J., Staudt, M., 1999. Biogenic Volatile Organic Compounds (VOC): An overview on emission, physiology, and ecology. Journal of Atmospheric Chemistry 33, 23-88.

Köbbing, J., Thevs, N., Zerbe, S., 2013. The utilization of reed (*Phragmites australis*)-a review. Mires and Peat 13, 1-14.

Kohli, S.K., Handa, N., Kaur, R., Kumar, V., Khanna, K., Bakshi, P., Singh, R., Arora, S., Kaur, R., Bhardwaj, R., 2017. Role of salicylic acid in heavy metal stress tolerance: insight into the underlying mechanism. in: Nazar, R., Iqbal, N., Khan, N.A. (Eds.). Salicylic acid: A multifaceted hormone. Springer Singapore, Singapore, pp. 123-144.

Kortesmäki, E., Östman, J.R., Meierjohann, A., Brozinski, J.-M., Eklund, P., Kronberg, L., 2020. Occurrence of antibiotics in influent and effluent from 3 major wastewater-treatment plants in Finland. Environmental Toxicology and Chemistry 39, 1774-1789.

Kostopoulou, S., Ntatsi, G., Arapis, G., Aliferis, K.A., 2020. Assessment of the effects of metribuzin, glyphosate, and their mixtures on the metabolism of the model plant *Lemna minor* L. applying metabolomics. Chemosphere 239, 124582.

Lamichhane, S., Sen, P., Dickens, A.M., Hyötyläinen, T., Orešič, M., 2018. An Overview of Metabolomics data analysis: Current tools and future perspectives. in: Joaquim, J., Carmen, B., Romà, T. (Eds.). Data Analysis for Omic Sciences: Methods and Applications. Elsevier, pp. 387-413.

Langebartels, C., Wohlgemuth, H., Kschieschan, S., Grün, S., Sandermann, H., 2002. Oxidative burst and cell death in ozone-exposed plants. Plant Physiology and Biochemistry 40, 567-575.

Lavola, A., Julkunen-Tiitto, R., PÄÄKkÖNen, E., 1994. Does ozone stress change the primary or secondary metabolites of birch (*Betula pendula* Roth.)? New Phytologist 126, 637-642.

Le Gall, G., DuPont, M.S., Mellon, F.A., Davis, A.L., Collins, G.J., Verhoeyen, M.E., Colquhoun, I.J., 2003. Characterization and content of flavonoid glycosides in genetically modified Tomato (*Lycopersicon esculentum*) fruits. Journal of Agricultural and Food Chemistry 51, 2438-2446.

Le Moullec, A., Juvik, O.J., Fossen, T., 2015. First identification of natural products from the African medicinal plant *Zamioculcas zamiifolia* - A drought-resistant survivor through millions of years. Fitoterapia 106, 280-285.

Lee, D.-K., Ahn, S., Cho, H.Y., Yun, H.Y., Park, J.H., Lim, J., Lee, J., Kwon, S.W., 2016. Metabolic response induced by parasitic plant-fungus interactions hinders amino sugar and nucleotide sugar metabolism in the host. Sci Rep, p. 37434.

Lee, S.-H., Kim, H.-W., Lee, M.-K., Kim, Y.J., Asamenew, G., Chan, Y.-S., Kim, J.-B., 2018. Phenolic profiling and quantitative determination of common sage (*Salvia plebeia* R. Br.) by UPLC-DAD-QTOF/MS. European Food Research and Technology 244, 1637-1646.

Letzel, T., Bayer, A., Schulz, W., Heermann, A., Lucke, T., Greco, G., Grosse, S., Schussler, W., Sengl, M., Letzel, M., 2015. LC-MS screening techniques for wastewater analysis and analytical data handling strategies: Sartans and their transformation products as an example. Chemosphere 137, 198-206.

Letzel, T., Lucke, T., Schulz, W., Sengl, M., Letzel, M., 2014. In a class of its own-OMI (Organic Molecule Identification) in water using LC-MS(/MS): Steps from “unknown” to “identified”: a contribution to the discussion.

Lim, G.H., Singhal, R., Kachroo, A., Kachroo, P., 2017. Fatty Acid- and lipid-mediated signaling in plant defense. Annual review of phytopathology 55, 505-536.

Limami, A.M., Glévarec, G., Ricoult, C., Cliquet, J.B., Planchet, E., 2008. Concerted modulation of alanine and glutamate metabolism in young *Medicago truncatula* seedlings under hypoxic stress. Journal of experimental botany 59, 2325-2335.

Liu, F., Ying, G.-G., Tao, R., Zhao, J.-L., Yang, J.-F., Zhao, L.-F., 2009. Effects of six selected antibiotics on plant growth and soil microbial and enzymatic activities. Environmental pollution (Barking, Essex : 1987) 157, 1636-1642.

Liu, J.-H., Kitashiba, H., Wang, J., Ban, Y., Moriguchi, T., 2007. Polyamines and their ability to provide environmental stress tolerance to plants. Plant Biotechnology 24, 117-126.

Liu, L., Liu, Y.-h., Liu, C.-x., Wang, Z., Dong, J., Zhu, G.-f., Huang, X., 2013a. Potential effect and accumulation of veterinary antibiotics in *Phragmites australis* under hydroponic conditions. Ecological Engineering 53, 138-143.

Liu, M., Li, X., Liu, Y., Cao, B., 2013b. Regulation of flavanone 3-hydroxylase gene involved in the flavonoid biosynthesis pathway in response to UV-B radiation and drought stress in the

desert plant, Reaumuria soongorica. Plant physiology and biochemistry : PPB / Societe francaise de physiologie vegetale 73C, 161-167.

Liu, Z., Rochfort, S., 2014. Recent progress in polar metabolite quantification in plants using liquid chromatography-mass spectrometry. Journal of Integrative Plant Biology 56, 816-825.

Löfstedt, T., Trygg, J., 2011. OnPLS-a novel multiblock method for the modeling of predictive and orthogonal variation. Journal of Chemometrics 25, 441-455.

Ma, Y.L., Li, Q.M., Van den Heuvel, H., Claeys, M., 1997. Characterization of flavone and flavonol aglycones by collision-induced dissociation tandem mass spectrometry. Rapid Communications in Mass Spectrometry 11, 1357-1364.

Maciejewska-Potapczyk, W., Konopska, L., Olechnowicz, K., 1975. Protein in *Lemna minor* L. Biochemie und Physiologie der Pflanzen 167, 105-108.

Maeda, H., Dudareva, N., 2012. The Shikimate pathway and aromatic amino acid biosynthesis in plants. Annual Review of Plant Biology 63, 73-105.

Martínez-Piernas, A.B., Nahim-Granados, S., Polo-López, M.I., Fernández-Ibáñez, P., Murgolo, S., Mascolo, G., Agüera, A., 2019. Identification of transformation products of carbamazepine in lettuce crops irrigated with Ultraviolet-C treated water. Environ Pollut 247, 1009-1019.

Matsuo, T., Tsugawa, H., Miyagawa, H., Fukusaki, E., 2017. Integrated strategy for unknown EI-MS identification using a quality control calibration curve, multivariate analysis, EI-MS spectral database, and retention index prediction. Analytical chemistry 89, 6766-6773.

Mayo, S.J., Bogner, J., Boyce, P.C., 1998. Araceae. in: Kubitzki, K. (Ed.). Flowering plants monocotyledons: Alismatanae and Commelinanae (except Gramineae). Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 26-74.

McClure, J.W., Alston, R.E., 1966. A chemotaxonomic study of Lemnaceae. Am J Bot 53, 849-860.

Miliauskas, G., Venskutonis, P.R., van Beek, T.A., 2004. Screening of radical scavenging activity of some medicinal and aromatic plant extracts. Food Chemistry 85, 231-237.

Milke, J., Gałczyńska, M., Wróbel, J., 2020. The importance of biological and ecological properties of *Phragmites australis* (Cav.) Trin. Ex Steud., in phytoremediation of aquatic ecosystems-The review. Water 12.

Mishra, P., Gong, Z., Kelly, B.C., 2017. Assessing biological effects of fluoxetine in developing zebrafish embryos using gas chromatography-mass spectrometry-based metabolomics. Chemosphere 188, 157-167.

Moheb, A., Agharbaoui, Z., Kanapathy, F., Ibrahim, R.K., Roy, R., Sarhan, F., 2013. Tricin biosynthesis during growth of wheat under different abiotic stresses. Plant Science 201-202, 115-120.

Mohedano, R.A., Costa, R.H.R., Tavares, F.A., Belli Filho, P., 2012. High nutrient removal rate from swine wastes and protein biomass production by full-scale duckweed ponds. Bioresource Technology 112, 98-104.

Moon, Y.J., Wang, X., Morris, M.E., 2006. Dietary flavonoids: effects on xenobiotic and carcinogen metabolism. Toxicology in vitro: an international journal published in association with BIBRA 20, 187-210.

Muhit, M.A., Izumikawa, M., Umehara, K., Noguchi, H., 2016. Phenolic constituents of the Bangladeshi medicinal plant *Pothos scandens* and their anti-estrogenic, hyaluronidase inhibition, and histamine release inhibitory activities. Phytochemistry 121, 30-37.

Muñoz-Cuervo, I., Malapa, R., Michalet, S., Lebot, V., Legendre, L., 2016. Secondary metabolite diversity in taro, *Colocasia esculenta* (L.) Schott, corms. Journal of Food Composition and Analysis 52, 24-32.

Nabavi, S.M., Šamec, D., Tomczyk, M., Milella, L., Russo, D., Habtemariam, S., Suntar, I., Rastrelli, L., Daglia, M., Xiao, J., Giampieri, F., Battino, M., Sobarzo-Sanchez, E., Nabavi, S.F., Yousefi, B., Jeandet, P., Xu, S., Shirooie, S., 2020. Flavonoid biosynthetic pathways in plants: Versatile targets for metabolic engineering. Biotechnology Advances 38, 107316.

Navarro-Reig, M., Jaumot, J., Baglai, A., Vivó-Truyols, G., Schoenmakers, P.J., Tauler, R., 2017. Untargeted comprehensive two-dimensional liquid chromatography coupled with high-resolution mass spectrometry analysis of Rice metabolome using multivariate curve resolution. Analytical chemistry 89, 7675-7683.

Ndolo, V.U., Fulcher, R.G., Beta, T., 2015. Application of LC-MS-MS to identify niacin in aleurone layers of yellow corn, barley, and wheat kernels. Journal of Cereal Science 65, 88-95.

Neilson, K.A., Mariani, M., Haynes, P.A., 2011. Quantitative proteomic analysis of cold-responsive proteins in rice. Proteomics 11, 1696-1706.

Neuhaus, D., Kühl, H., Kohl, J.G., Dörfel, P., Börner, T., 1993. Investigation on the genetic diversity of *Phragmites stands* using genomic fingerprinting. Aquatic Botany 45, 357-364.

Nostro, A., Germanò, M.P., D'Angelo, V., Marino, A., Cannatelli, M.A., 2000. Extraction methods and bioautography for evaluation of medicinal plant antimicrobial activity. Letters in applied microbiology 30, 379-384.

Nürenberg, G., Schulz, M., Kunkel, U., Ternes, T.A., 2015. Development and validation of a generic nontarget method based on liquid chromatography-high-resolution mass spectrometry analysis for the evaluation of different wastewater treatment options. Journal of Chromatography A 1426, 77-90.

Obermeier, M., Schröder, C.A., Helmreich, B., Schröder, P., 2015. The enzymatic and antioxidative stress response of *Lemna minor* to copper and a chloroacetamide herbicide. Environmental science and pollution research international 22, 18495-18507.

Ohlbaum, M., Wadgaonkar, S.L., van Bruggen, J.J.A., Nancharaiah, Y.V., Lens, P.N.L., 2018. Phytoremediation of seleniferous soil leachate using the aquatic plants *Lemna minor* and Egeria densa. Ecological Engineering 120, 321-328.

Olsen, K.M., Slimestad, R., Lea, U.S., Brede, C., LØVdal, T., Ruoff, P., Verheul, M., Lillo, C., 2009. Temperature and nitrogen effects on regulators and products of the flavonoid pathway: experimental and kinetic model studies. Plant, Cell & Environment 32, 286-299.

Pal, A., Paul, A.K., 2013. Bacterial endophytes of the medicinal herb *Hygrophila spinosa* T. Anders and their antimicrobial activity. Journal of Pharmaceutical Research International 3, 795-806.

Pan, M., Wong, C.K., Chu, L.M., 2014. Distribution of antibiotics in wastewater-irrigated soils and their accumulation in vegetable crops in the Pearl River Delta, southern China. J Agric Food Chem 62, 11062-11069.

Park, M.G., Blossey, B., 2008. Importance of plant traits and herbivory for invasiveness of *Phragmites australis* (Poaceae). American Journal of Botany 95, 1557-1568.

Pereira, C.A.M., Yariwake, J.H., McCullagh, M., 2005. The distinction of the C-glycosylflavone isomer pairs orientin/isoorientin and vitexin/isovitexin using HPLC-MS exact mass measurement and in-source CID. Phytochemical Analysis 16, 295-301.

Petrussa, E., Braidot, E., Zancani, M., Peresson, C., Bertolini, A., Patui, S., Vianello, A., 2013. Plant flavonoids-biosynthesis, transport, and involvement in stress responses. Int J Mol Sci 14, 14950-14973.

Pierattini, E.C., Francini, A., Huber, C., Sebastiani, L., Schröder, P., 2018. Poplar and diclofenac pollution: A focus on physiology, oxidative stress, and uptake in plant organs. Science of The Total Environment 636, 944-952.

Pobłocka-Olech, L., van Nederkassel, A.-M., Vander Heyden, Y., Krauze-Baranowska, M., Glód, D., Baczek, T., 2007. Chromatographic analysis of salicylic compounds in different species of the genus *Salix*. Journal of Separation Science 30, 2958-2966.

Prinsloo, G., Vervoort, J., 2018. Identifying anti-HSV compounds from unrelated plants using NMR and LC-MS metabolomic analysis. Metabolomics 14, 134.

Rahman, M.E., Bin Halmi, M.I.E., Bin Abd Samad, M.Y., Uddin, M.K., Mahmud, K., Abd Shukor, M.Y., Sheikh Abdullah, S.R., Shamsuzzaman, S.M., 2020. Design, operation, and optimization of constructed wetland for removal of pollutants. Int J Environ Res Public Health 17.

Rajab, M., Greco, G., Heim, C., Helmreich, B., Letzel, T., 2013. Serial coupling of RP and zwitterionic hydrophilic interaction LC-MS: suspects screening of diclofenac transformation products by oxidation with a boron-doped diamond electrode. J Sep Sci 36, 3011-3018.

Riach, A.C., Perera, M.V.L., Florance, H.V., Penfield, S.D., Hill, J.K., 2015. Analysis of plant leaf metabolites reveals no common response to insect herbivory by *Pieris rapae* in three related host-plant species. Journal of experimental botany 66, 2547-2556.

Riemenschneider, C., Seiwert, B., Goldstein, M., Al-Raggad, M., Salameh, E., Chefetz, B., Reemtsma, T., 2017a. An LC-MS/MS method for the determination of 28 polar environmental contaminants and metabolites in vegetables irrigated with treated municipal wastewater. Analytical Methods 9, 1273-1281.

Riemenschneider, C., Seiwert, B., Moeder, M., Schwarz, D., Reemtsma, T., 2017b. Extensive transformation of the pharmaceutical carbamazepine following uptake into intact Tomato plants. Environ Sci Technol 51, 6100-6109.

Rochfort, S., 2005. Metabolomics Reviewed: A New "Omics" platform technology for systems biology and implications for Natural products research. Journal of Natural Products 68, 1813-1820.

Rosa, M., Prado, C., Podazza, G., Interdonato, R., González, J.A., Hilal, M., Prado, F.E., 2009. Soluble sugars-metabolism, sensing and abiotic stress: a complex network in the life of plants. Plant signaling & behavior 4, 388-393.

Sakalem, M.E., Negri, G., Tabach, R., 2012. Chemical composition of hydroethanolic extracts from five species of the *Passiflora* genus. Revista Brasileira de Farmacognosia 22, 1219-1232.

Salem, M., Souza, L., Serag, A., Fernie, A., Farag, M., Ezzat, S., Alseekh, S., 2020. Metabolomics in the context of plant natural products research: From sample preparation to metabolite analysis. Metabolites 10, 37.

Sánchez, V., López-Bellido, F.J., Cañizares, P., Rodríguez, L., 2018. Can electrochemistry enhance the removal of organic pollutants by phytoremediation? Journal of Environmental Management 225, 280-287.

Sanjaya, Hsiao, P.-Y., Su, R.-C., Ko, S.-S., Tong, C.-G., Yang, R.-Y., Chan, M.-T., 2008. Overexpression of *Arabidopsis thaliana* tryptophan synthase beta 1 (AtTSB1) in *Arabidopsis* and tomato confers tolerance to cadmium stress. Plant Cell Environ 31, 1074-1085.

Santos, J.L., Aparicio, I., Alonso, E., 2007. Occurrence and risk assessment of pharmaceutically active compounds in wastewater treatment plants. A case study: Seville city (Spain). Environ Int 33, 596-601.

Sarkar, U., Ravindra, K.C., Large, E., Young, C.L., Rivera-Burgos, D., Yu, J., Cirit, M., Hughes, D.J., Wishnok, J.S., Lauffenburger, D.A., Griffith, L.G., Tannenbaum, S.R., 2017. Integrated assessment of diclofenac biotransformation, pharmacokinetics, and omics-based toxicity in a three-dimensional human liver-immunocompetent coculture system. Drug Metabolism and Disposition 45, 855.

Sauvêtre, A., May, R., Harpaintner, R., Poschenrieder, C., Schröder, P., 2018. Metabolism of carbamazepine in plant roots and endophytic rhizobacteria isolated from *Phragmites australis*. Journal of hazardous materials 342, 85-95.

Sauvêtre, A., Schröder, P., 2015. Uptake of carbamazepine by rhizomes and endophytic bacteria of *Phragmites australis*. Frontiers in Plant Science 6.

Savchenko, T., Walley, J.W., Chehab, E.W., Xiao, Y., Kaspi, R., Pye, M.F., Mohamed, M.E., Lazarus, C.M., Bostock, R.M., Dehesh, K., 2010. Arachidonic acid: an evolutionarily conserved signaling molecule that modulates plant stress signaling networks. Plant Cell 22, 3193-3205.

Schäfer, M., Brütting, C., Baldwin, I.T., Kallenbach, M., 2016. High-throughput quantification of more than 100 primary- and secondary-metabolites, and phytohormones by a single solid-phase extraction-based sample preparation with analysis by UHPLC–HESI–MS/MS. Plant Methods 12, 30.

Schröder, P., Scheer, C., Belford, B.J.D., 2001. Metabolism of organic xenobiotics in plants: conjugating enzymes and metabolic endpoints. Minerva Biotecnologica 13, 85-91.

Schymanski, E.L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H.P., Hollender, J., 2014. Identifying small molecules via high-resolution mass spectrometry: communicating confidence. Environmental science & technology 48, 2097-2098.

Scossa, F., Benina, M., Alseekh, S., Zhang, Y., Fernie, A.R., 2018. The integration of metabolomics and next-generation sequencing data to elucidate the pathways of natural product metabolism in medicinal plants. Planta Medica 84, 855-873.

Sharma, A., Shahzad, B., Rehman, A., Bhardwaj, R., Landi, M., Zheng, B., 2019. Response of phenylpropanoid pathway and the role of polyphenols in plants under abiotic stress. Molecules (Basel, Switzerland) 24, 2452.

Sharma, G., Rahul, Guleria, R., Mathur, V., 2020. Differences in plant metabolites and microbes associated with Azadirachta indica with variation in air pollution. Environmental Pollution 257, 113595.

Sharma, K., Lee, Y.R., 2016. Effect of different storage temperatures on the chemical composition of onion (*Allium cepa* L.) and its enzymes. J Food Sci Technol 53, 1620-1632.

Siddiqui, M.J., Ismail, Z., Saidan, N.H., 2011. Simultaneous determination of secondary metabolites from *Vinca rosea* plant extractives by reverse-phase high-performance liquid chromatography. Pharmacognosy magazine 7, 92-96.

Simirgiotis, M., Schmeda-Hirschmann, G., Borquez, J., Kennelly, E., 2013. The *Passiflora tripartita* (banana passion) fruit: A source of bioactive flavonoid C-glycosides isolated by HSCCC and characterized by HPLC-DAD-ESI/MS/MS. Molecules (Basel, Switzerland) 18, 1672-1692.

Singer, A.C., Crowley, D.E., Thompson, I.P., 2003. Secondary plant metabolites in phytoremediation and biotransformation. Trends in Biotechnology 21, 123-130.

Sirtori, C., Agüera, A., Gernjak, W., Malato, S., 2010. Effect of water-matrix composition on Trimethoprim solar photodegradation kinetics and pathways. Water Res 44, 2735-2744.

Sivaram, A.K., Subashchandrabose, S.R., Logeshwaran, P., Lockington, R., Naidu, R., Megharaj, M., 2019. Metabolomics reveals defensive mechanisms adapted by maize on exposure to high molecular weight polycyclic aromatic hydrocarbons. Chemosphere 214, 771-780.

Smith, B.N., Meeuse, B.J., 1966. Production of volatile amines and skatole at anthesis in some arum lily species. Plant Physiol 41, 343-347.

Storozhenko, S., Navarrete, O., Ravanel, S., De Brouwer, V., Chaerle, P., Zhang, G.F., Bastien, O., Lambert, W., Rébeillé, F., Van Der Straeten, D., 2007. Cytosolic hydroxymethyl-dihydro pterin pyrophosphokinase/dihydropteroate synthase from *Arabidopsis thaliana*: a specific role in early development and stress response. The Journal of biological chemistry 282, 10749-10761.

Stringlis, I.A., de Jonge, R., Pieterse, C.M.J., 2019. The age of coumarins in plant-microbe interactions. Plant and Cell Physiology 60, 1405-1419.

Sultana, B., Anwar, F., Ashraf, M., 2009. Effect of extraction solvent/technique on the antioxidant activity of selected medicinal plant extracts. Molecules (Basel, Switzerland) 14, 2167-2180.

Sumiahadi, A., Acar, R., 2017. A review of phytoremediation technology: heavy metals uptake by plants. in: IOP Conference Series: Earth and Environmental Science (Ed.). The 4th International Conference on Sustainable Agriculture and Environment (4th ICSAE) 10–12 August 2017. IOP Publishing Ltd, Surakarta, Indonesia, p. 012023.

Sumner, L.W., Amberg, A., Barrett, D., Beale, M.H., Beger, R., Daykin, C.A., Fan, T.W.M., Fiehn, O., Goodacre, R., Griffin, J.L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A.N., Lindon, J.C., Marriott, P., Nicholls, A.W., Reily, M.D., Thaden, J.J., Viant, M.R., 2007. Proposed minimum reporting standards for chemical analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics 3, 211-221.

Sumner, L.W., Lei, Z., Nikolau, B.J., Saito, K., 2015. Modern plant metabolomics: advanced natural product gene discoveries, improved technologies, and prospects. Natural Product Reports 32, 212-229.

Sylvia Grosse, T.L., 2017. User Manual Stoff-IDENT Database. pp. 1-35.

Tandlich, R., Brežná, B., Dercová, K.n., 2001. The effect of terpenes on the biodegradation of polychlorinated biphenyls by *Pseudomonas stutzeri*. Chemosphere 44, 1547-1555.

Tantikanjana, T., Mikkelsen, M.D., Hussain, M., Halkier, B.A., Sundaresan, V., 2004. Functional analysis of the tandem-duplicated P450 genes; in glucosinolate biosynthesis and plant development by transposition-generated double mutants. Plant Physiol 135, 840.

Tehrani, R., Van Aken, B., Kaveh, R., 2012. Uptake and metabolism of pharmaceuticals and other emerging contaminants by plants. pp. 541-570.

Tohge, T., de Souza, L.P., Fernie, A.R., 2017. Current understanding of the pathways of flavonoid biosynthesis in model and crop plants. Journal of experimental botany 68, 4013-4028.

Tohge, T., Fernie, A.R., 2009. Web-based resources for mass-spectrometry-based metabolomics: A user's guide. Phytochemistry 70, 450-456.

Tolstikov, V.V., Fiehn, O., 2002. Analysis of highly polar compounds of plant origin: the combination of hydrophilic interaction chromatography and electrospray ion trap mass spectrometry. Analytical Biochemistry 301, 298-307.

Toyo'oka, T., 2008. Determination methods for biologically active compounds by ultra-performance liquid chromatography coupled with mass spectrometry: application to the analyses of pharmaceuticals, foods, plants, environments, metabonomics, and metabolomics. Journal of Chromatographic Science 46, 233-247.

Tsimogiannis, D., Samiotaki, M., Panayotou, G., Oreopoulou, V., 2007. Characterization of flavonoid subgroups and hydroxy substitution by HPLC-MS/MS. Molecules (Basel, Switzerland) 12, 593-606.

Tsolmon, B., Fang, Y., Yang, T., Guo, L., He, K., Li, G.-Y., Zhao, H., 2020. Structural identification and UPLC-ESI-QTOF-MS2 analysis of flavonoids in the aquatic plant *Landoltia punctata* and their in vitro and in vivo antioxidant activities. Food Chemistry, 128392.

Tybring, G., von Bahr, C., Bertilsson, L., Collste, H., Glaumann, H., Solbrand, M., 1981. Metabolism of carbamazepine and its epoxide metabolite in human and rat liver in vitro. Drug metabolism and disposition: the biological fate of chemicals 9, 561-564.

Tzin, V., Galili, G., 2010. The biosynthetic pathways for shikimate and aromatic amino acids in *Arabidopsis thaliana*. Arabidopsis Book 8, e0132-e0132.

Valant-Vetschera, K.M., 1985. C-Glycosylflavones as an accumulation tendency: A critical review. The Botanical Review 51, 1-52.

Van Hoeck, A., Horemans, N., Monsieurs, P., Cao, H.X., Vandenhove, H., Blust, R., 2015. The first draft genome of the aquatic model plant *Lemna minor* opens the route for future stress physiology research and biotechnological applications. Biotechnol Biofuels 8, 188-188.

Varga, M., Horvatić, J., Čelić, A., 2013. Short-term exposure of *Lemna minor* and *Lemna gibba* to mercury, cadmium, and chromium. Central European Journal of Biology 8, 1083-1093.

Vernouillet, G., Eullaffroy, P., Lajeunesse, A., Blaise, C., Gagné, F., Juneau, P., 2010. Toxic effects and bioaccumulation of carbamazepine evaluated by biomarkers measured in organisms of different trophic levels. Chemosphere 80, 1062-1068.

Vieno, N., Sillanpää, M., 2014. The fate of diclofenac in municipal wastewater treatment plant-A review. Environment International 69, 28-39.

Villette, C., Maurer, L., Wanko, A., Heintz, D., 2019. Xenobiotics metabolization in *Salix alba* leaves uncovered by mass spectrometry imaging. Metabolomics 15, 122.

Vladimirova, I.N., Georgiyants, V.A., 2014. Biologically active compounds from *Lemna minor* S. F. Gray. Pharmaceutical Chemistry Journal 47, 599-601.

Wahman, R., Graßmann, J., Sauvêtre, A., Schröder, P., Letzel, T., 2020. *Lemna minor* studies under various storage periods using extended-polarity extraction and metabolite non-target screening analysis. Journal of Pharmaceutical and Biomedical Analysis 188, 113362.

Wahman, R., Grassmann, J., Schröder, P., Letzel, T., 2019. Plant metabolomic workflows using reversed-phase LC and HILIC with ESI-TOF-MS. LC GC N. Am. 37, 8-15.

Wahman, R., Sauvêtre, A., Schröder, P., Moser, S., Letzel, T., 2021. Untargeted Metabolomics Studies on Drug-Incubated *Phragmites australis* Profiles. Metabolites 11, 2.

Wang, X., Zhong, X.-J., Zhou, N., Cai, N., Xu, J.-H., Wang, Q.-B., Li, J.-J., Liu, Q., Lin, P.-C., Shang, X.-Y., 2020. Rapid characterization of chemical constituents of the tubers of *Gymnadenia conopsea* by UPLC-Orbitrap-MS/MS analysis. Molecules (Basel, Switzerland) 25, 898.

Want, E.J., Masson, P., Michopoulos, F., Wilson, I.D., Theodoridis, G., Plumb, R.S., Shockcor, J., Loftus, N., Holmes, E., Nicholson, J.K., 2013. Global metabolic profiling of animal and human tissues via UPLC-MS. Nature Protocols 8, 17-32.

Waridel, P., Wolfender, J.-L., Ndjoko, K., Hobby, K.R., Major, H.J., Hostettmann, K., 2001. Evaluation of quadrupole time-of-flight tandem mass spectrometry and ion-trap multiple-stage mass spectrometry for the differentiation of C-glycosidic flavonoid isomers. Journal of Chromatography A 926, 29-41.

Watson, A.A., Fleet, G.W., Asano, N., Molyneux, R.J., Nash, R.J., 2001. Polyhydroxylated alkaloids-natural occurrence and therapeutic applications. Phytochemistry 56, 265-295.

Wegener, A., Gimbel, W., Werner, T., Hani, J., Ernst, D., Sandermann, H., Jr., 1997. Molecular cloning of ozone-inducible protein from *Pinus sylvestris* L. with high sequence similarity to vertebrate 3-hydroxy-3-methylglutaryl-CoA-synthase. Biochimica et Biophysica Acta 1350, 247-252.

Worley, B., Powers, R., 2013. Multivariate analysis in metabolomics. Current Metabolomics 1, 92-107.

Xi, Y., Song, Y., Johnson, D.M., Li, M., Liu, H., Huang, Y., 2018. The enhanced phytoremediation of diesel in soil by *Trifolium repens*. Ecotoxicology and Environmental Safety 154, 137-144.

Xie, M., Zhang, J., Tschaplinski, T.J., Tuskan, G.A., Chen, J.-G., Muchero, W., 2018. Regulation of lignin biosynthesis and its role in growth-defense tradeoffs. Frontiers in Plant Science 9, 1427.

Xie, X.-Y., Wang, R., Shi, Y.-P., 2013. Chemical constituents from rhizomes of *Homalomena occulta*. Zhongguo Zhong Yao Za Zhi 38, 2325-2327.

Xu, J.-J., Fang, X., Li, C.-Y., Zhao, Q., Martin, C., Chen, X.-Y., Yang, L., 2018. Characterization of *Arabidopsis thaliana* hydroxyphenylpyruvate reductases in the tyrosine conversion pathway. Frontiers in plant science 9, 1305-1305.

Xu, L.L., Xu, J.J., Zhong, K.R., Shang, Z.P., Wang, F., Wang, R.F., Zhang, L., Zhang, J.Y., Liu, B., 2017. Analysis of non-volatile chemical constituents of *Menthae haplocalycis* Herba by ultra-high performance liquid chromatography-high resolution mass spectrometry. Molecules (Basel, Switzerland) 22.

Yang, H., Ludewig, U., 2014. Lysine catabolism, amino acid transport, and systemic acquired resistance: what is the link? Plant signaling & behavior 9, e28933.

Yang, Q., Zhang, A.-h., Miao, J.-h., Sun, H., Han, Y., Yan, G.-l., Wu, F.-f., Wang, X.-j., 2019. Metabolomics biotechnology, applications, and future trends: a systematic review. RSC Advances 9, 37245-37257.

Yang, Z., Nakabayashi, R., Okazaki, Y., Mori, T., Takamatsu, S., Kitanaka, S., Kikuchi, J., Saito, K., 2014. Toward better annotation in plant metabolomics: isolation and structure elucidation of 36 specialized metabolites from *Oryza sativa* (rice) by using MS/MS and NMR analyses. Metabolomics: Official journal of the Metabolomic Society 10, 543-555.

Yonekura-Sakakibara, K., Higashi, Y., Nakabayashi, R., 2019. The Origin and Evolution of Plant Flavonoid Metabolism. Frontiers in Plant Science 10.

Zhang, H., Li, X., Yang, Q., Sun, L., Yang, X., Zhou, M., Deng, R., Bi, L., 2017. Plant growth, antibiotic uptake, and prevalence of antibiotic resistance in an endophytic system of *Pakchoi* under antibiotic exposure. Int J Environ Res Public Health 14, 1336.

Zhang, L., Hatzakis, E., Patterson, A.D., 2016. NMR-based metabolomics and its application in drug metabolism and cancer research. Current Pharmacology Reports 2, 231-240.

Zhang, X.-w., Li, Q.-h., Xu, Z.-d., Dou, J.-j., 2020. Mass spectrometry-based metabolomics in health and medical science: a systematic review. RSC Advances 10, 3092-3104.

Zhang, Z.H., Dai, Z., Hu, X.R., Lin, R.C., 2013. Isolation and structure elucidation of chemical constituents from *Pinellia ternata*. Journal of Chinese medicinal materials 36, 1620-1622.

Zhao, T., Zhang, H., Zhao, T., Zhang, X., Lu, J., Yin, T., Liang, Q., Wang, Y., Luo, G., Lan, H., Li, P., 2012. Intrarenal metabolomics reveals the association of local organic toxins with the progression of diabetic kidney disease. Journal of Pharmaceutical and Biomedical Analysis 60, 32-43.

Zhao, X., Moates, G.K., Wellner, N., Collins, S.R., Coleman, M.J., Waldron, K.W., 2014. Chemical characterization and analysis of the cell wall polysaccharides of duckweed (*Lemna minor*). Carbohydr Polym 111, 410-418.

Zhao, X., Zeng, Z., Chen, A., Lu, X., Zhao, C., Hu, C., Zhou, L., Liu, X., Wang, X., Hou, X., Ye, Y., Xu, G., 2018. A comprehensive strategy to construct an in-house database for accurate and batch identification of small molecular metabolites. Analytical Chemistry 90, 7635-7643.

Zrenner, R., Stitt, M., Sonnewald, U., Boldt, R., 2006. Pyrimidine and purine biosynthesis and degradation in plants. Annu Rev Plant Biol 57, 805-836.

Appendix

The following articles have been published

Paper-I

Wahman *et al.*, 2019

Plant Metabolomics Workflows Using Reversed-Phase LC and HILIC with ESI-TOF-MS

Current Trends in Mass Spectrometry, LCGC

Paper-II

Wahman *et al.*, 2020

Lemna minor Studies under various storage periods using extended-polarity extraction and metabolite non-target screening analysis

Journal of Pharmaceutical and Biomedical Analysis 188:113362

Paper-III

Untargeted Metabolomics Studies on Drug-Incubated *Phragmites australis* Profiles

Metabolites 11(1):2

In preparation

Paper-IV

Wahman *et al.*, 2021

Untargeted analysis of *Lemna minor* metabolites: Workflow and prioritization strategy comparing highly confident features in between extracts and different mass spectrometers. In preparation

Paper-V

Wahman *et al.*, 2021

The changes in *Lemna minor* metabolomic profile: A response to diclofenac incubation. Submitted to Chemosphere journal.

Plant Metabolomic Workflows Using Reversed-Phase LC and HILIC with ESI-TOF-MS

In the field of metabolomics, researchers seek to acquire almost complete information about the metabolic composition of a sample to provide fundamental information about the organism's cellular state. In metabolomics analysis today, typically reversed-phase liquid chromatography (LC) is coupled with specific, sensitive, and robust mass spectrometry (MS). That approach, however, misses many moderately polar, and all very polar, compounds; this situation is a problem in plant metabolomics, because plant metabolites are mainly water-soluble species and thus very polar. Here, we describe new developments in polarity-extended separations using the serial coupling of reversed-phase LC and hydrophilic interaction liquid chromatography (HILIC) separation steps, in combination with electrospray ionization-time-of-flight-mass spectrometry (ESI-TOF-MS), and the application of this approach to plant metabolomics. The resulting retention time vs. mass plots are molecular fingerprints, as well as sources of further molecular descriptors. Extraction methods, molecular analysis, and data evaluation have to be adapted to the matrix under consideration. Representative strategies using this polarity-extending approach, following so-called *suspects* and *nontargeted screening* approaches, are presented.

Rofida Wahman, Johanna Grassmann, Peter Schröder, and Thomas Letzel

Plant metabolomics, which is a comparatively young field, aims to provide almost complete molecular coverage of plant metabolites, to establish fundamental information about the plant under investigation based on changes in its metabolism (1,2). Plant metabolites can be roughly divided into primary and secondary metabolites. Secondary metabolites, although not used during plant growth and development, are important, because of their usage in plant defense, food, and medicine. Moreover, secondary metabolites play a substantial role in the adaptation of plants to environmental stress (3). Plants respond to exogenous factors through signaling pathways that induce downstream stress responses, including the modulation of gene expression and the regulation of a wide range of biochemical processes, ending with a remodeling of the metabolism (4).

Lemna minor (duckweed) is sometimes used as a model plant in plant metabolomics studies. It has also been proposed for phytoremediation of heavy metals in a glass house experiment (5). Additionally, *L. minor* has medical importance (6). Compared to other plant species, duckweed has a simple structure and rapid growth. Furthermore, *L. minor* is easy to harvest.

Targeted, suspects, and nontargeted screening strategies can be used for plant metabolite analysis (7). Targeted screening, formerly known as *quantitative analysis*, observes analytes using a reference substance. Compound identification and quantification has to be validated with isotopically labeled reference substances using mass spectrometry (MS) detection. Suspects screening typically is performed with accurate-mass and high-resolution mass spectrometry (HRMS) to observe the empirical formula of each formula present, or with tandem MS to observe specific fragment spectra. Subsequently, the empirical formula can be compared with chemical databases, such as Chemspider (http://www.chemspider.com), Chemicalize (https://chemicalize.com), or STOFF-IDENT (SI; https://www.lfu.bayern.de/stoffident), and with mass-spectrometric databases containing analytically observed mass spectra (such as MassBank, https://massbank.eu/MassBank/), or local databases in

laboratories. If these databases lead to clear hits, one then has to validate the observation by an equivalent reference substance. Provided that the identity is proven, one can use it further in targeted screening analysis applying isotopically labeled standards. As stated before, this approach may require HRMS to generate the empirical formula of expected substances.

HRMS instrumentation is strictly needed when nontargeted screening is used as an analytical strategy. This screening strategy is split into screening for so-called *hidden targets* (also referred to as *known unknowns*) and *unknown targets* (also referred to as *unknown unknowns*). The hidden-target screening approach is similar to suspect screening, but starts without a list of metabolites or analytical databases. However, nontargeted screening data frequently can be compared with data in chemical and analytical databases (especially in retrospective analyses). For this purpose, analytical platform solutions are needed to handle nontargeted screening data. FOR-IDENT (FI; https://water.for-ident.org) is an open-access example of such a platform; it enables data evaluation by retention time index, accurate mass, and other important features of the molecule analyzed. The core of FOR-IDENT is a compound database called STOFF-IDENT, which is filled with anthropogenic compounds relevant in the aqueous environment. Another compound database under development is PLANT-IDENT, which contains general plant metabolites and includes exact masses, as well as other analytical results, and makes use of a retention index tool; this database is expected to encourage nontargeted screening in plant metabolomics.

Detailed metabolomics strategies (also for so-called *unknown unknowns*) are described in a 2014 paper (8). The nontargeted screening approach is encouraged by recent advances in HRMS that provide increased mass resolution and accuracy, enabling the identification of metabolites often simply by their accurate mass determinations, and by developments in tandem MS that enable the determination of accurate fragment spectra (containing additional structural information) (9). Gas chromatography–mass spectrometry (GC–MS) and reversed-phase liquid chromatography–mass spectrometry (LC–MS) are the methods of choice for qualitative and quantitative metabolomics analysis. The ultimate aim of nontargeted screening is to be able to accurately detect, monitor, and (eventually) identify every relevant metabolite in plant extracts, which cannot be achieved by any single existing analytical method. The methodology has to be thoughtful, from initial solvent extraction through chromatography via ionization and MS to data evaluation. Recently, plant metabolites analysis has been carried out through direct examination of crude extracts or after reversed-phase LC separation coupled with quadrupole

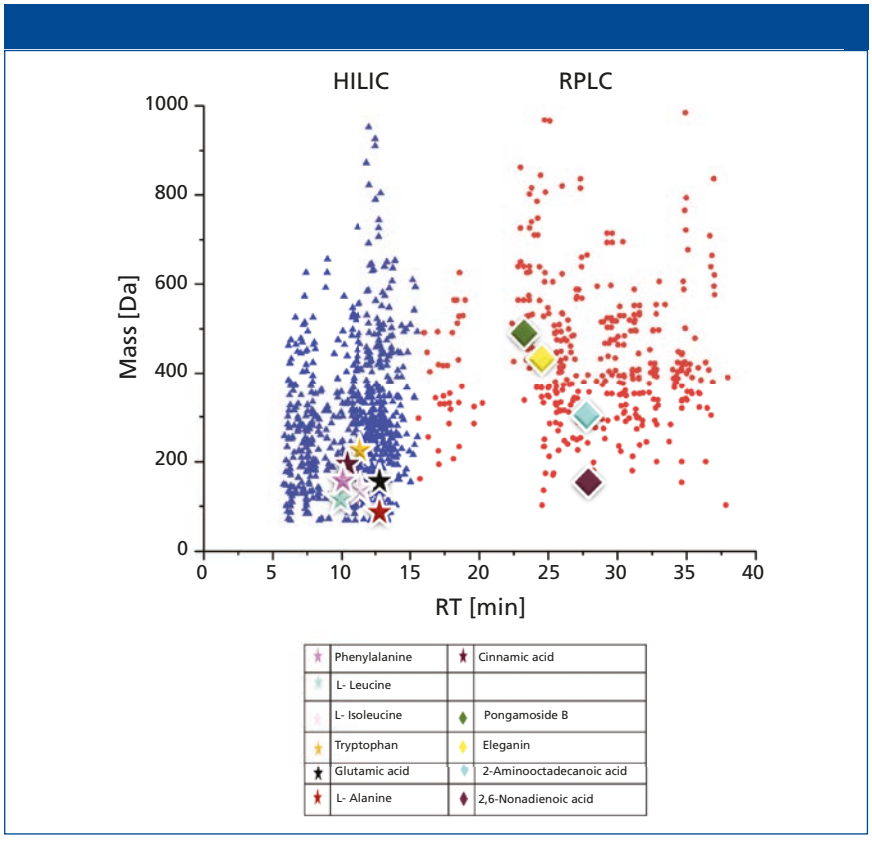


Figure 1: Retention time versus mass plot of *Lemna minor* extract (of an untreated sample). Blue triangles represent hydrophilic interaction liquid chromatography (HILIC)–retained molecules eluted from 0 to 15 min; red circles represent reversed-phase LC–retained molecules eluted from 16 to 38 min. Different colored and shaped dots represent examples of identified (stars) and expected (diamonds) compounds, as indicated in the legend.

MS, time-of-flight MS (TOF-MS), or other MS techniques, like ultrahigh-resolution Fourier transform ion cyclotron MS (FT–MS (1).

The monitoring and identification of (very) polar compounds is important, because these compounds play a vital role in plant metabolic pathway changes. The use of hydrophilic interaction liquid chromatography (HILIC) columns in plant metabolomics analysis has facilitated the retention of polar and very polar compounds in a chromatographic column prior to elution into the mass spectrometer (10). The demand for a new analytical method that can identify a plant's metabolites in a single run has been answered previously (for red wine analysis [11]). The polarity-extended serial coupling of reversed-phase LC and HILIC, in combination with HRMS (12), allows for the robust and repeatable analysis of a large variety of compounds in a single run (13,14), and has been applied in several disciplines (11,15–18). In such polarity-extended studies, a new extraction method for sample preparation was confirmed that it is applicable to polarity-extended chromatography analysis (19). The current study focuses on the use of serial coupling of reversed-phase LC and HILIC, coupled to high-resolution TOF-MS and general data evaluation strategies for plant metabolomics, applied to the analysis of *L. minor* extracts. Therefore, samples were used before and after plant incubation with an exemplary (model) pharmaceutical that can influence the metabolism of the treated plants. Accordingly, new insights are provided, using accurate-mass MS and HRMS as well as novel data evaluation workflows.

Experimental

Reagents and Chemicals

Methanol and water were obtained in LC–MS grade from VWR (Darmstadt, Germany). *L. minor* was kindly provided by the German Research Center for Environmental Health, Plant Microbe Interactions, Helmholtz Centrum of Munich. Amino acids standards were obtained from Sigma (Missouri, United States).

Sample Preparation

First, 500 mg of freeze-dried and milled *L. minor* (incubated with and without 10 µM of diclofenac, respectively) were extracted with 5 mL of methanol. Extraction solvent was sonicated at 35 KHz (Sonorex super RK 106, Bandelin, Germany) with plant material for 10 min at 4 °C. Afterwards, the samples were centrifuged at 1500 rpm for 20 min, and the supernatant was transferred to a test tube. The extraction method was triplicated under exactly the same experimental conditions. Finally, the extracts were evaporated to dryness (using SpeedVac Fischer Scientific, Sweden), and redissolved in (50:50) methanol:water. The standard solutions and the solvent extract were injected three times, respectively.

Reversed-Phase LC and HILIC combined with ESI-TOF-MS

The polarity-extended serial coupling or reversed-phase LC and HILIC was connected via a Jet Stream ESI interface to an Agilent 6230 TOF-MS instrument (both Agilent Technologies, Santa Clara, California, United States), as described recently (11). Column effluent was coupled via a T-piece to an isocratic pump (Agilent Technologies, Waldbronn, Germany), which provides a constant flow of reference solution for MS calibration prior entering the ion source of the MS. The separation system consisted of an autosampler, a column oven, two columns, two binary pumps, and a UV detector (all Agilent Technologies, Waldbronn, Germany). The initial binary pump was connected to a nonpolar 120 EC-C18 Poroshell column (Agilent Technologies, Waldbronn, Germany). The outlet of this column was connected to a ZIC-HILIC

column (Merck, Darmstadt, Germany), and to a second binary pump via a T-piece (Upchurch Scientific, IDEX, Illinois, USA). Ions were detected in positive ionization mode, with a mass range of 50–2100 Da. The instrument resolution was greater than 10,000 at *m/z* 922. The parameters were as follows: 325 °C gas temperature, 10 L/min drying gas flow, 325 °C sheath gas temperature, 7.5 L/min sheath gas flow, 45-psi nebulizer operating pressure and 100 V fragmentor voltage.

Data Processing

The data were evaluated with Agilent Mass hunter Profinder B.08.00, Mass Profiler, and Mass Profiler Professional (MPP) 13.1.1 software. ProFinder parameters applied for the automation of compound retention times (t_R), and extraction and molecular weights, were set to a peak filter of 300 counts peak height, ion species to "positive ions" with H^+, Na^+ and K^+, "charge state" to 1, and the "expected retention time" to ±3.00 min. The extracted ion chromatograms (EICs) were smoothed with a Gaussian function, using 9 data points width and 5000 points Gaussian width. Then, the "features" from the blanks dataset were deleted from sample datasets. These parameters limit the result for 2000 compound groups. The data set was exported to MPP. In MPP, the compounds were aligned. Subsequently, the retention times and masses were corrected, and C, H, O, N, and S atoms were used to compute the empirical formula of the compounds. The compounds' intensity logarithmic fold changes between the two samples are calculated and subsequently are drawn as a scatter plot. The calculated logD at pH 7.4 was based on the retention time/logD (pH 7) calibration curve of twelve different standards (18,20). The logD (pH) values were predicated from ChemAxon software (https://disco.chemaxon.com/apps/demos/logd/) and then exported to Windows Excel 2016 for further data evaluation. For hidden-target screening, the amino acids standards were organized, and injected three times. The mean of each of the standard masses (S) and t_R values (S) were calculated in daltons and in minutes, respectively. The variation between the extract and the standard mean isotopic masses (Δ ppm) was calculated according to following equation:

$$\Delta ppm \frac{(mean\ of\ Standard\ masses - mean\ of\ Sample\ masses)}{mean\ of\ sample\ masses} \times 10^6 \quad [1]$$

Moreover, the standard deviation (SD) of t_R and relative standard deviation (RSD) were calculated.

Percent of RSD= SD of compound t_R values in different injection divided by the mean of compound t_R values.

The nontargeted measurements were evaluated through FOR-IDENT (FI; https://water.for-ident.org) and Metlin (https://metlin.scripps.edu/landing_page.php?pgcontent=batch_search) databases to search for expected and hidden targets ("known unknowns"). The workflow was conducted in the steps described below.

First, the list of accurate masses was uploaded into the database. The matching features were downloaded with their physiochemical properties. For FOR-IDENT results, the feature hits below a retention time of 15 min were filtered by excluding positive logD values. The amino acids were confirmed by reference standard injection (category 1) (21,22). However, cinnamic acid and other hits from the Metlin database remained as suspects (category 2) (21,22).

The log2-normalized data were used in statistical analysis to improve normality and group 2724 metabolites. Differences between the two samples were considered significant when the *P* value (calculated using a Student's *t*-test) cut-off was 0.05. The principal component analysis (PCA) was done using the multivariate analysis of OriginPro 2017. The heatmap was developed using MPP of the same data set as used in the PCA analysis.

Results and Discussion

Formerly, the serial coupling of reversed-phase LC and HILIC with ESI-TOF-MS was established for trace organic compound analysis in wine samples (10), water samples (15), and oxidative (17) as well as enzymatic conversion screening samples (18). More recently, this coupling is also being applied in plant metabolomic studies, such as to study metabolic changes caused by stress or by external compounds. *L. minor* is a good model plant for such studies and can easily be incubated, extracted, and analyzed.

The TOF-MS technique, and its specificity, allow accurate compound detection over a broad mass range (in different matrices). The so-called *nontargeted screening strategy* was performed in this study in positive ionization mode and with a mass range of 50–2100 Da. Concerning a typical methanol extract analysis of untreated *L. minor* samples with the serial coupling of reversed-phase LC and HILIC with ESI-TOF-MS, feature extraction was possible as shown in Figure 1. For such a feature, the retention time (t_R), the accurate mass, and the signal intensity were extracted from a total ion chromatogram (TIC), each reflecting an independent (but still unknown) molecule. More than half of the isolated compounds (686) were retained and separated on the (very) polar HILIC column (blue colored triangular compounds in Figure 1; they are located with a retention time between 5 and 15 min, and have a logarithm of distribution coefficient logD [pH 7] < 0). In addition, 383 compounds were separated by reversed-phase LC (nonpolar) (red-colored circles compounds in Figure 1; they were eluted later than 16 min, and have logD values (pH 7) > 0), respectively. The empirical formula could be predicted for most features. The data sets were evaluated via nontargeted screening strategies as described above in the introduction and references therein.

First, the nontargeted screening data were evaluated by exporting results to the open-access platform FOR-IDENT. There, the features were uploaded and the normalized retention time (t_R) and accurate mass were compared with the compound database STOFF-IDENT (containing anthropogenic compounds expected in the aqueous environment). This search yielded several hits of suggested compounds with respective empirical formula and logD values. Applying this strategy resulted in hits and suggested compounds containing amino acids and organic acids (see Figure 1; specifically, molecules labeled with stars and included in the table). For example, glutamic acid and cinnamic acid were the only hits

Table I: List of example amino acids found in *Lemna minor* extract with the mean monoisotopic mass of reference standard (S), extract (M), the variation between them; mean retention time (t_R) of reference standard (S) and the extract (M) and the variation between them.

Compound Name	Mean Monoisotopic Mass (Da) (S)	Mean Monoisotopic Mass (Da) (M)	Δ ppm	Mean t_R (Min) (S)	Mean t_R (Min) (M)	Δ t_R
L-Isoleucine	131.0944	131.0947	-2.29	10.82	10.83	-0.01
L-Leucine	131.0941	131.0945	-3.05	10.79	10.83	-0.04
Phenylalanine	165.0787	165.0789	0.081	10.71	10.72	0.00
Tryptophan	204.0898	204.0899	-0.26	11.45	11.61	-0.15
Glutamic acid	147.0529	147.0525	2.49	12.63	12.71	-0.08
L-Alanine	89.0477	89.0476	1.12	12.84	12.93	-0.09

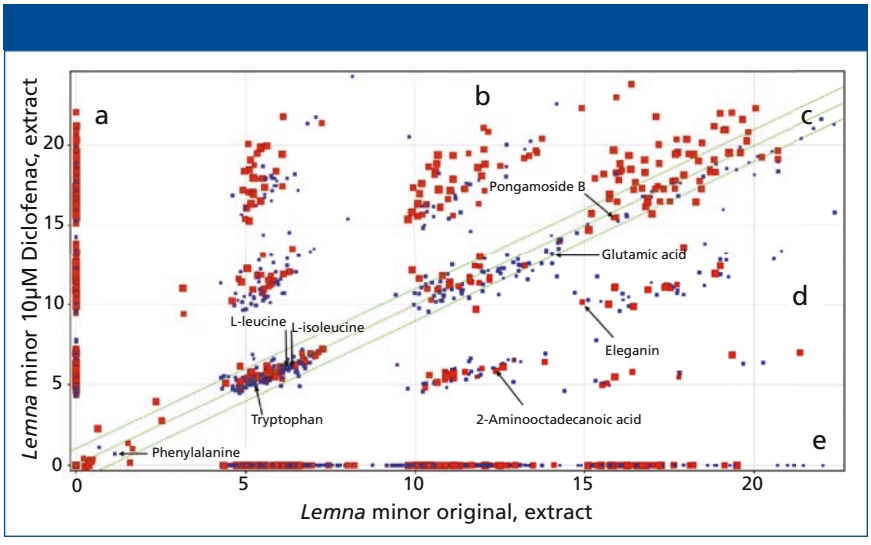


Figure 2: Feature (signal intensity) comparison plot of an extract using original untreated *Lemna minor* sample versus an extract of *L. minor* sample treated with 10 µM of diclofenac. Examples of found amino acids and expected compounds are found in both samples. Small blue squares represent retention time (t_R) (0–15 min) and big red squares represent the t_R (16–38 min) of (a) compounds only found in the *Lemna* sample treated with 10-µM diclofenac extract; (b) compounds found in both samples with increased feature intensity in treated compared to untreated *Lemna* samples; (c) compounds found in both samples; (d) compounds found in both samples with decreased feature intensity in treated compared to the untreated *Lemna* sample; and (e) compounds only found in the untreated *Lemna* sample extract.

in STOFF-IDENT for 147.0532 Da and 148.052 Da, respectively. Thus, one can confirm the identity of both very quickly and easily by measuring the respective reference substance for each compound.

However, the process is not always that simple. For example, the mass of 165.079 gives 10 matches, of which nine have a positive logD value, and one a negative value. Thus, the nine matches could be deleted, because the feature eluted in the negative logD region was retained by HILIC. The remaining hit (after deletion) was the amino acid phenylalanine and could easily be evaluated further by comparing to reference material.

Sometimes the use of compound databases does not lead to such unequivocal results. For example, a feature at 89.048 Da resulted in five hits. After excluding the two positive logD values, three hits were remaining with negative logD values. L-alanine was confirmed by a reference standard also eluting in the HILIC region. In general, this strategy leads to explicit and filtered suggestions for compound identities.

Furthermore, tryptophan, leucine, and isoleucine (as shown in Figure 1) were identified similarly. After excluding the positive logD results, the remaining hits with negative logD values were either the corresponding amino acid, or various synthetic compounds. Later hits are typically anthropogenic (brought into the environment), and cannot be originated in the plant, thus they were neglected in this study. Reference standards for the remaining amino acids were injected to prove the identity and presence in the *L. minor* extracts, which also were reported previously in the literature (24). Finally, for the successful application of the hidden-target screening strategy (7) using the compound database STOFF-IDENT, amino acid standards were available, directly confirming the hits of expected amino acids in *L. minor* methanol extracts. The amino acids identified are phenylalanine, L-leucine, L-isoleucine, tryptophan, glutamic acid, and L-alanine, as labeled in Figure 1. The amino acid standards were injected in triplicate, and the mean of each standard mass (S) and retention time were calculated. The identification was done through comparing the differences between the mean retention times of the standard and extract, which was in the range of 0 min (phenylalanine) to 0.15 min (tryptophan). Moreover, the mass variation between the standard mass (S) and the extract mass (M) is less than 5 ppm. In Table I, the six identified amino acids were listed with their mean isotopic mass and mean retetention times. In addition, standards mean isotopic mass and retention times were calculated for the three injections.

The identification was done using the following workflow. The differences

Table II: Examples of compounds found in *Lemna minor* extract using a hidden-target screening strategy, showing the monoisotopic mass in literature (L), mean monoisotopic in extract (M), the variation between them, the mean retention time (t_R) of the extract (M), standard deviation, relative standard deviation, logD in predicted (P), and logD experimental (E).

Compound Name	Monoisotopic Mass (Da) (L)	Mean Monoisotopic Mass (Da) (M)	Δ ppm	Mean t_R (Min) (M)	SD of t_R (Min)	RSD	Log D (P) pH 7.4	Log D (E) pH 7
Cinnamic acid	148.0524	148.052	2	10.72	0.033	0.003	-0.81	-0.32
Pongamoside B	470.1213	470.122	1	23.29	0.084	0.004	0.40	0.97
Eleganin	434.1577	434.159	3	24.34	0.029	0.001	0.76	1.3
2-Aminoocta-decanoic acid	299.2824	299.282	1	27.61	0.034	0.001	3.91	2.31
2,6-Nonadienoic acid	154.0994	154.0995	2	27.00	0.016	0.001	1.3	2.12

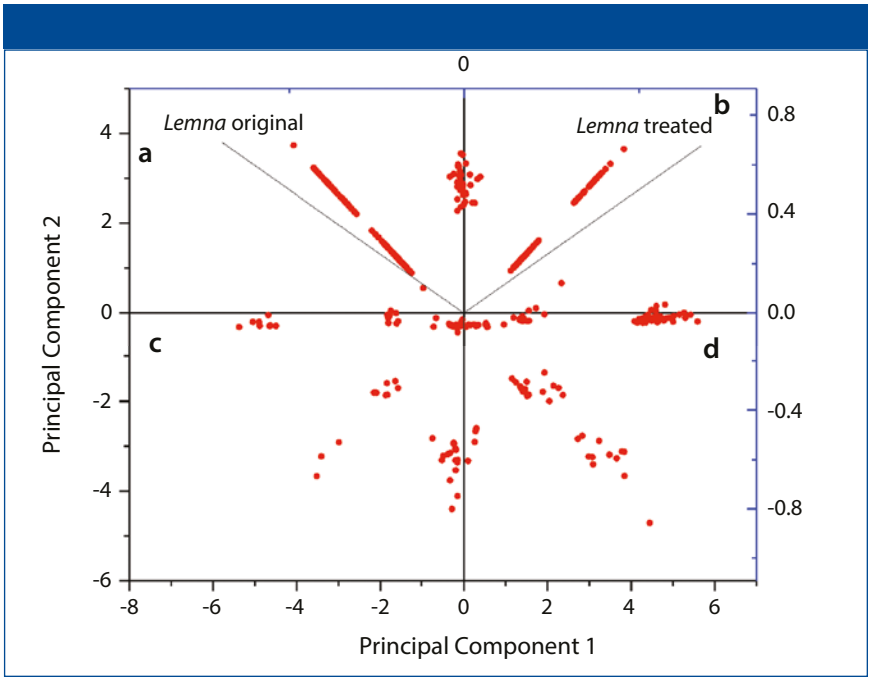


Figure 3: Principal component analysis (PCA) plot of features occurring in untreated *Lemna minor* sample extract versus *L. minor* sample extract treated with 10-µM diclofenac: (a) Dots represent the compounds related positively to untreated sample and negatively to treated sample; (b) dots represent compounds that are positively related to both samples; (c) dots represent compounds that are negatively related to both samples; (d) dots represent compounds related positively to treated sample and negatively to untreated sample.

in t_R (Δt_R) between the standard and the extract were in the range of ≈0–0.15 min. The variations between the extract and the standard mean isotopic masses (Δ ppm) were calculated in the six amino acids. The Δ ppm values were in the range of –3.05 for L-leucine to +0.08 for phenylalanine. All the amino acids were in the same retention time and mass range as formerly reported for the serial coupling of HILIC and reversed-phase LC in an aqueous environment (15). Also, without using tandem MS (and its significant fragment spectra) the hits resulted in category 1 of the identification scheme published earlier (21,22).

Generally, nontargeted screening data from plant extracts can be compared with an available metabolomics compound database like Metlin (https://metlin.scripps.edu/landing_page.php?pgcontent=batch_search) to find expected molecular hits using the hidden-target strategy described previously. A disadvantage compared to the FOR-IDENT platform is that other databases like Metlin have no automated comparison functionality. In applying the Metlin database, four compounds were identified by manually comparing the t_R and logD of the compounds with the literature; these compounds were pongamoside B, eleganin, 2-aminooctadeconic acid, and 2,6-nonadienoic acid (Figure 1). These compounds and cinnamic acid (additionally observed in FOR-IDENT) were not yet confirmed by standards; thus, these compounds remain in category 2 as suggested by references (21,22). However, the four compounds presumably identified through Metlin and cinnamic acid by the same workflow are listed in Table II. The variation between measured monoisotopic mass and the mass found in the literature is less than ±4 ppm. Moreover, the standard deviation (SD) of t_R and relative standard deviation (RSD) were calculated. The all-compounds RSD values were <0.1% for the three injections, which indicates the reproducibility of the LC method, and as formerly reported for the serial coupling of HILIC and reversed-phase LC (15). Consequently, the t_R for each compound could be used in the standards t_R indices calibration curve to calculate the logD (pH 7) (20,25). The lower polarity limit of polar compounds was set to a logD (pH 7) value of zero because of the reversed-phase LC column used; therefore, it can likely retain compounds above this polarity. The polarity region below logD zero in this study is restricted to HILIC (15). Pongamoside B, eleganin, 2-aminooctadeconic acid, and 2,6-nonadienoic acid were retained by reversed-phase LC, thus their t_R values were longer than 15 min. The experimental logD values of suspect compounds were calculated from the 12 standards calibration curve of logD (pH7) and retention time index. The difference between experimental and predicated logD (pH 7) values was

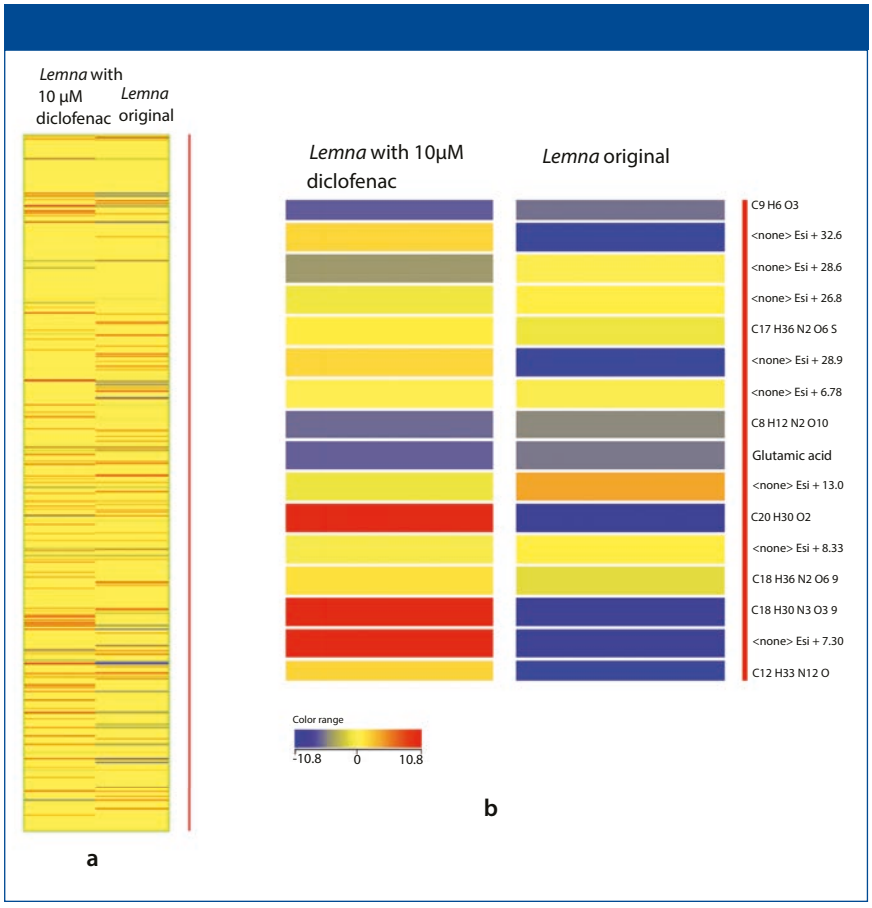


Figure 4: (a) Heatmap plot of different compound intensities in untreated *Lemna minor* sample extract versus *Lemna minor* sample extract treated with 10 µM of diclofenac (with glutamic acid as an concrete example). (b) Magnified view of a part of the heatmap. Red color indicates higher intensity of compound in the sample. The compound intensity is ≤10.8. However, the blue color indicates lower intensity, which means intensity is ≥ −10.8. The yellow color indicates the absence of compound in the sample, which means the intensity is equal to zero.

in the range of 0.5 (cinnamic acid) to 1.6 (2-aminooctadecanoic acid). Subsequently, the calculated logD values for all compounds are compatible with the literature logD values. In consideration of supporting parameters, such as variation between masses, t_R, and logD values of cinnamic acid, pongamoside B, eleganin, 2-aminooctadeconic acid, and 2,6-nonadienoic acid were presumably identified through the Metlin database using the hidden-target screening strategy.

However, identifying molecules and "calling them by name" often is not the goal of nontargeted screening measurements or workflows. Moreover, nontargeted screening of the total sample may inform researchers about recent changes in the metabolome in combination with statistical tools and graphical visualization. For instance, direct feature comparison (19), PCA (26), heat plots (27), and clustering (28) were used to reduce and visualize the complex metabolomics datasets (23).

A study in the laboratory using the model plant *L. minor* was performed by incubation with 10 µM of the pharmaceutical diclofenac, which may be enriched in the plant, metabolized in the plant, or change the plant's metabolic pathways. The latter can be monitored in principle by nontargeted screening measurement and subsequent application of statistical tools like the above stated direct sample comparisons or PCA.

In sample comparisons, features were found reflecting *L. minor* samples with and without pharmaceutical incubation, which were located on the *y*-axis in Figure 2a, and on the *x*-axis in Figure 2e, respectively. Other compounds were not affected by incubation, and had the same signal intensity in the untreated as well as treated samples, thus located in the middle part of the comparison plot (see Figure 2c). However, some compound intensities were increased due to incubation, and are located in the upper part in Figure 2b. Others were decreased due to incubation, and are found in the lower part of Figure 2d.

The category 1 identified phenylalanine, tryptophan, L-leucine, glutamic acid, L-isoleucine, and category 2 identified pongamoside B were found in the two samples with the same intensity (see Figure 2c). Consequently, incubation with diclofenac or its degradation products might not affect their biosynthetic pathways. However, the incubation of *L. minor* with diclofenac caused decreases in the intensity of some compounds. Category 2 identified eleganin and 2-aminooctadececanoic acid have a higher intensity in the original samples, indicating that diclofenac or its degradation products affected their biosynthetic pathways. Accordingly, incubation of *L. minor* with 10-µM diclofenac causes changes in its metabolome by altering the intensity or disappearance of the compounds.

PCA of the samples showed that samples were grouped into two groups (Figure 3). *L. minor* with 10-µM diclofenac was related to principal component 1 (PC1), whereas the original untreated sample was related to principal component 2 (PC2). The dots (features) located in the upper left part (the positive part of principal component 2) represent the compounds related to the untreated sample (see Figure 3a). Also, the compounds related positively to both samples are located in the upper right part of Figure 3b. Furthermore, the compounds related negatively to both samples are located in lower left part of Figure 3c. However, the compounds related to treated samples are drawn in the positive region of PC2 (see Figure 3d).

In addition, the heat map of the *L. minor* with 10-µM of diclofenac and the original samples presented the compounds intensities in the two samples. The blue color indicates that the compounds have low intensity and the red color indicates that the compounds have higher intensity. As an example,

glutamic acid was found with low intensity –5.81 and –6.62 in *L. minor* original untreated and treated with 10 µM of diclofenac, respectively (Figures 4a and 4b).

Conclusion

Advances in MS are driving nontargeted screening by generating an accurate empirical formula, using tandem MS with structural information observed by molecule fragmentation (not in this study), and normalizing retention times and correlating the latter with logD values. Thus, the application of data evaluation platforms and compound databases can be helpful in identifying compounds that lack reference standards, by searching for the exact masses in a hidden-target screening approach, and to analyze suspect metabolites in nontargeted screening. A new database resembling FOR-IDENT, that will focus on plant metabolites, will be launched. The new database, PLANT-IDENT, will have the same concept and advantages of the FOR-IDENT platform. Moreover, the new database will decrease the number of hits compared to chemical databases like Chemspider.com, and thus will avoid many false positive results. The launch of this database will help researchers who use highly sensitive mass spectrometers in identification of plant metabolomics through nontargeted screening.

In this study, extracts containing untreated or pharmaceutical-incubated *L. minor* plants were investigated with nontargeted screening strategies like hidden-target screening. The applied workflow in *L. minor* metabolite analysis will be a touchstone for subsequent research. Even when using different mass spectrometers, researchers can apply this workflow with or without reference materials to identify suspect and hidden targets. The statistical nontargeted screening workflow was conducted through different statistical tests to monitor metabolite differences between different samples. In addition, it allows the monitoring of metabolites changes. Studies that are more complex might require more complex or combined statistical tools.

Acknowledgments

This study was partially supported by the Bavarian State Ministry of the Environment and Consumer Protection. Dr. Andrés Sauvêtre is thanked for the cultivation, incubation and harvesting of the applied *L. minor* samples.

References

(1) R. Hall, M. Beale, O. Fiehn, N. Hardy, L. Sumner, and R. Bino, *Plant Cell* **14**(7), 1437 LP-1440 (2002).
(2) F.T. Jorge, T.A. Mata, and A. Carla, *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **374**(2079), 20150370 (2016).
3. J.N. Kabera, E. Semana, A.R. Mussa, and X. He, J. Pharm. Pharmacol. 2, 377–392 (2014). doi:10.1016/0300-9084(96)82199-7.
4. N. Shitan, *Biosci. Biotechnol. Biochem.* **80**(7), 1283–1293 (2016).
5. S.H. Bokhari, I. Ahmad, M. Mahmood-Ul-Hassan, and A. Mohammad, *Int. J. Phytoremediation* **18**(1), 25-32 (2016). doi:10.1080/15226514.2015.1058331.
6. X. Zhao, G.K. Moates, N. Wellner, S.R.A. Collins, M.J. Coleman, and K.W. Waldron, *Carbohydr. Polym.* **111**, 410–418 (2014). doi:10.1016/j.carbpol.2014.04.079.
7. T. Letzel, A. Bayer, W. Schulz, A. Heermann, T. Lucke, G. Grecoa, S. Grosse, W. Schüssler, M. Sengl, and M. Letzel, *Chemosphere* **37**, 198-206 (2015). doi:10.1016/j.chemosphere.2015.06.083.
8. T. Letzel, *Lab More Int.* 14–18 (2014). at http://www.int.laborundmore.com/archive/555305/Nontargeted-screening,-suspected-target-screening-and-target-screening-–-of-technologies-and-philosophies,-databases-and-crafts.html.
9. T. Letzel and J. E. Drewes, *ACS Symp. Ser.* pp. 175–181 (2016). doi:10.1021/bk-2016-1242.ch010.
10. T. Letzel and G. Greco, *J. Chromatogr. Sci.* **51**(7), 684–693 (2013).
11. G. Greco, S. Grosse, and T. Letzel, *J. Sep. Sci.* **36**(8), 1379–1388 (2013).
12. G. Greco and T. Letzel, *LCGC North Amer.* **10**(2s), 40–44 (2012).
13. G. Greco, S. Grosse, and T. Letzel, *J. Sep. Sci.* **37**(6), 630–634 (2014).
14. G. Greco, A. Boltner, and T. Letzel, *Am. J. Mod. Chromatogr.* **1**(1), 12–25 (2014).
15. S. Bieber, G. Greco, S. Grosse, and T. Letzel, *Anal. Chem.* **89**(15), 7907-7914 (2017). doi:10.1021/acs.analchem.7b00859.
16. S. Bieber and T. Letzel, *LCGC Europe* **31**(11), 602–608 (2018).
17. M. Rajab, G. Greco, C. Heim, B. Helmreich, and T. Letzel, *J. Sep. Sci.* **36**(18), 3011–3018 (2013).
18. L.F. Stadlmair, S. Grosse, T. Letzel, J.E. Drewes, and J. Grassmann, *Anal. Bioanal. Chem.* **411**(2), 339-351 (2018). doi:10.1007/s00216-018-1442-7.
19. R. Wahman, J. Grassmann, P. Schröder, and T. Letzel, *Prep.*
20. S. Grosse, and T. Letzel, *User Man. Stoff-IDENT Database* **4.1,** 1–33 (2016).
21. E.L. Schymanski, J. Jeon, R. Gulde, K. Fenner, M. Ruff, H.P. Singer, and J. Hollender, *Environ. Sci. Technol.* **48**(4), 2097–2098 (2014).
22. T. Letzel, T. Lucke, W. Schulz, M. Sengl, and M. Letzel, *Lab More Int.* 24–28 (2014).
23. J.E. Schollée, E.L. Schymanski, and J. Hollender, *Transform. Prod. Chem. by Nontargeted Suspect Screen.– Strateg. Work. Vol. 1* **1241,** 4–45 (American Chemical Society, 2016).
24. W. Maciejewska-Potapczyk, L. Konopska, and K. Olechnowicz, *Biochem. und Physiol. der Pflanz.* **167**(1), 105–108 (1975).
25. FOR-IDENT (2019). available at https://www.for-ident.org/.
26. Y. Wang, L. Xu, H. Shen, J. Wang, W. Liu, X. Zhu, R. Wang, X. Sun, and L. Liu, *Sci. Rep.* **5,** 18296 (2015).
27. Q. Zhang, Y. Shi, L. Ma, X. Yi, and J. Ruan, *PLoS One* **9**(11), e112572 (2014).
28. P.H. Benton, J. Ivanisevic, D. Rinehart, A. Epstein, M.E. Kurczy, M.D. Boska, H.E. Gendelman, and G. Siuzdak, *Metabolomics* **11**(4), 1029–1034 (2015).

Rofida Wahman, Johanna Grassmann and **Thomas Letzel** are with the Technical University of Munich in Munich, Germany. **Peter Schröder** is with the German Research Center for Environmental Health in Munich, Germany. Direct correspondence to T.Letzel@tum.de.

Journal of Pharmaceutical and Biomedical Analysis 188 (2020) 113362

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba

Lemna minor studies under various storage periods using extended-polarity extraction and metabolite non-target screening analysis

Rofida Wahman[a], Johanna Graßmann[a], Andrés Sauvêtre[b], Peter Schröder[b], Thomas Letzel[a,*]

[a] Chair of Urban Water Systems Engineering, Technical University of Munich, Am Coulombwall 3, 85748, Garching, Germany
[b] German Research Center for Environmental Health, Research Unit Comparative Microbiome Analysis, Helmholtz Centrum Munich, Ingolstädter Street 1, 85764, Neuherberg, Germany

ARTICLE INFO

Article history:
Received 27 November 2019
Received in revised form 4 May 2020
Accepted 7 May 2020
Available online 28 May 2020

Keywords:
Lemna minor
Extended polarity extraction method
Extended polarity chromatographic separation
Metabolomics
Non-target screening
RPLC-HILIC-ESI-TOF-MS analysis
Storage effect

ABSTRACT

Plant metabolomic studies cover a broad band of compounds, including various functional groups with different polarities and other physiochemical properties. For this reason, specific optimized methods are needed in order to enable efficient and non-destructive extraction of molecules over a large range of LogD values. This study presents a simple and efficient extraction procedure for *Lemna minor* samples demonstrating polarity extension of the molecular range. The *Lemna* samples chosen were kept under the following storage conditions: 1) fresh, 2) stored for a few days at −80 °C, and 3) stored for 6 months at −80 °C. The samples were extracted using five specifically chosen solvents: 100 % ethanol, 100 % methanol (MeOH), acidic 90 % MeOH (MeOH-water-formic acid (FAC) (90:9.5:0.5, v/v/v), MeOH-water (50:50, v/v), and 100 % water. The final extraction procedure was conducted subject to three solvent conditions, and the subsequent polarity-extended analysis was applied for *Lemna minor* samples using RPLC-HILIC-ESI-TOF-MS. The extraction yield is in descending order (acidic 90 % MeOH), 50 % MeOH, 100 % water and 100 % MeOH. The results displayed significant molecular differences, both in the extracts investigated and in the fresh *Lemna* samples, compared to stored samples, in terms of the extraction yield and reducing contents as well as the number of features. The storage of *Lemna minor* resulted in changes to the fingerprint of its metabolites as the reducing contents increased. The comparisons enable a direct view of molecule characterizations, in terms of their polarity, molecular mass, and signal intensity. This parametric information would appear ideal for further statistical data analysis. Consequently, the extraction procedure and the analysis/data evaluation are highly suitable for the so-called extended-polarity non-target screening procedure.

1. Introduction

The field of metabolomics represents a relatively new approach to the systematic study of metabolites and enables their presence and content to be determined in such biological samples as plants. The information gained from metabolomics can make a profound contribution to understanding the interactive nature of secondary metabolic networks in plants and their responses to environmental and genetic changes. It can also provide unique insights into the fundamental nature of plant phenotypes with regard to development, physiology, tissue identity, resistance, and biodiversity. Metabolomics fingerprinting can therefore be beneficial in drug discovery, gene-function analysis, and multiple diagnostic methods in phytomedicines [1]. *Lemna minor* (duckweed) is an aquatic plant that belongs to the Lemnaceae family. Thanks to its high reproducibility rate and great simplicity, it is often used in experimental studies. In 2010, Zhang et al. reported that *Lemna minor* could be used as a fast, inexpensive and reproducible model plant system for the study of host-pathogen interactions [2]. Moreover, it is used as food in the treatment of flu-like illnesses and for reducing body responses to allergies [3]. The plant extract should be obtained with minimum pretreatment, in order to prohibit the loss of metabolites. Consequently, the extract can be analyzed on the basis of a non-target screening strategy. However, challenges such

* Corresponding author.
E-mail address: T.Letzel@tum.de (T. Letzel).

https://doi.org/10.1016/j.jpba.2020.113362
0731-7085/

as massive differences in abundance and polarities limit the analytical power of plant metabolites [4]. Thus, sample preparation (extraction method) is one of the most important challenges in the analysis of plant metabolites. In recent years, several extraction methods of plant materials have been presented in the literature. The amount of extractable components is mainly influenced by the strength of the extraction method as well as the efficiency of the solvent in dissolving these compounds. Azwanida, enumerated the limitations of both conventional and modern techniques of plant extraction in 2015. The conventional methods of extraction, such as maceration, infusion, diffusion and decoction, waste many solvents, which is a serious environmental issue [5]. Also, the Soxhlet extraction method causes environmental pollution and is not suitable for the extraction of thermolabile compounds. In 2009, Sultana et al. revealed that the reflux extraction method reduces the scavenging activity of extracts, due to the degenerative reactions of antioxidant compounds. The heat accelerates the oxidation of antioxidant compounds in the determination of extraction yield and antioxidant activity in different parts of a plant [6]. There is a consensus among scientists that alongside their advantages, modern techniques such as solid-phase extraction, microwave-assisted extraction, accelerated solvent extraction, ultrasound-assisted extraction and supercritical fluid extraction also display weaknesses [7]. What is required today is a type of polarity-extended extraction that can be utilized in new polarity-extended chromatographic separation systems.

By employing various separation methods along with mass spectrometric detection, metabolomics has enabled the modification of the chemistry in phytomedicines as well as nutrition and toxicology fields. It comprises two classes: targeted and non-targeted metabolomics [8]. Targeted screening, formerly known as 'quantitative analysis', observes analytes using isotopically labeled reference substances. Non-targeted metabolomics aims to assign as many compounds as possible in a plant sample [9]. The analytical tools employed in non-target screening of plant metabolomes with LC-ESI-MS/MS techniques can yield 'big data'. However, these metabolite classes are difficult to retard and separate analytically (both from each other and from the matrix components) using the traditional and most commonly used method of reversed-phase liquid chromatography (RPLC), due to the great polarity variation [10]. Furthermore, the plant's polar metabolites have low adsorption on non-polar surfaces such as C18 RPLC material. They also display a low efficiency of electrospray ionization. Such molecules are often involved in a multitude of metabolic pathways and play an important role in the recognition of crucial metabolic changes [11]. Liu and Rochfort (2014), discussed the application of different stationary phases to the separation of polar compounds [12], such as HILIC, which is employed to characterize and separate (very) polar compounds in biological samples [13]. Therefore, it is essential to develop a separation technique with an extended polarity range that allows simultaneous monitoring of non-polar, polar, and (very) polar plant metabolites. Serial coupling of RPLC and HILIC is used to separate compounds with extended-polarity chromatography in a single run. Moreover, serial RPLC-HILIC can be coupled with accurate high-resolution time-of-flight mass spectrometers (TOF-MS), which provide accurate detection of non-polar to (very) polar compounds, including the empirical formula for each one [14]. Systems with parallel coupling are also being developed for extended-polarity separation [15]. The present paper explores an optimized extended polarity extraction method along with the effects of storage and the subsequent analysis of different *Lemna minor* samples. In this way, the study makes a great contribution to the analysis of plant metabolites using RPLC-HILIC-ESI-TOF-MS.

2. Materials and methods

2.1. Reagents and chemicals

LC–MS grade methanol and water were obtained from VWR, Darmstadt, Germany. Formic acid and amino acids standards were obtained from Sigma-Aldrich, Steinheim, Germany. Ethanol was purchased from AppliChem, Darmstadt, Germany. Gallic acid and sodium carbonate were obtained from Acros Organics, Niederau, Germany for reducing content determination. Folin-Ciocalteau reagent was purchased from Merck Chemicals, Darmstadt, Germany. *Lemna minor* was kindly provided by the German Research Center for Environmental Health, Plant-Microbe Interactions, Helmholtz Centrum of Munich, Germany.

2.2. Plant sample

Lemna minor L. was cultivated in aquaria according to Obermeier et al. [16] with minor modifications. Plants were grown at 23 °C with a photoperiod of 16–8 h and an average light intensity of 43 μmol m^{-2} s^{-1}. *Lemna* fronds were subcultured every two weeks in 24 L of Steinberg medium made up of (in mg/L) 85 $NaNO_3$, 13.4 KH_2PO_4, 75 $MgSO_4{\cdot}7\ H_2O$, 36 $CaCl_2{\cdot}2\ H_2O$, 20 Na_2CO_3, 1 H_3BO_3, 0.2 $MnCl_2{\cdot}4\ H_2O$, 0.01 $Na_2MoO_4{\cdot}2\ H_2O$, 0.05 $ZnSO_4{\cdot}7\ H_2O$, 0.005 $CuSo_4{\cdot}5\ H_2O$, 0.01 $Co(NO_3)_2{\cdot}6\ H_2O$, 0.84 $FeCl_3{\cdot}6\ H_2O$ and 1.4 Na_2-EDTA·2 H_2O. The plants were then harvested, rinsed briefly in tap water, dried with lint-free tissue paper, and finally frozen in liquid nitrogen. Samples were kept at –80 °C until further processing. The fresh samples that were frozen for several days were from the same plant generation.

2.3. Extraction

Three types of *Lemna minor* sample were extracted: 1) fresh plants LF, 2) plants frozen for several days at −80 °C, LFD and 3) plants frozen for 6 months at -80 °C (frozen months; LFM). For each sample, 500 mg of freeze-dried and milled powder was extracted with a) 100 % methanol (MeOH), b) acidic 90 % MeOH (MeOH-water-formic acid (FAC) (90:9.5:0.5, v/v/v)), c) MeOH-water (50:50, v/v), and d) 100 % water (H_2O), respectively. Additionally, the frozen plant material LFM was extracted with 100 % ethanol. Extracts with plant powder were sonicated (Sonorex super RK 106, Bandelin, Berlin, Germany) for 10 min at 4 °C. Although this extraction method uses an ultrasonic bath, extracts with plant powder were sonicated for 10 min at 4 °C and a frequency of 35 kHz to protect the samples from heating and loss of thermolabile compounds. The samples were then centrifuged (Z 200 A Universal Compact Centrifuge, Hermle LaborTechnik GmbH, Germany) at 1500 rpm for 20 min and the supernatants transferred to clean glass test tubes. The extraction process was triplicated in identical experimental conditions. Finally, the extracts were evaporated to dryness (using a SpeedVac, Fischer Scientific, Göteborg, Sweden) and dissolved in MeOH-H_2O (50:50, v/v) [17,18], as shown in (Fig. S1).

2.4. Reducing contents

100 μL of Folin-Ciocalteau reagent-H_2O (1:9, v/v) added to the 20 μL sample (Gallic acid at different concentrations (to plot the calibration curve) or plant extracts dissolved in MeOH-H_2O (50:50, v/v)) with and mixed well. 100 μL of sodium carbonate-H_2O (7.5 g: 100, w/v) solution were then added and the mixture incubated for 60 min in darkness at 25 °C. The absorbance was measured in triplicate at 765 nm. The reducing content was calculated from the

calibration curve as gallic acid equivalent (GAE) [18].

2.5. Instruments

The Agilent 1260 Infinity LC-systems comprised an autosampler, column oven, two columns and two binary pumps. An online degasser, a mixing chamber, and a UV detector were used to perform reversed-phase (RPLC) and zwitterion hydrophilic interaction liquid chromatography (HILIC) in its serial coupling. The first binary pump and the autosampler were connected to the reversed-phase separation column, a Poroshell 120 EC-C18 (50.0 × 3.0 mm, 2.7 μm; Agilent Technologies). The outlet of this column was connected to the HILIC column, which was a ZIC-HILIC column (150 × 2.1 mm, 5 μm, 200 Å; Merck Sequent, Umea, Sweden). The columns were coupled through a T-piece (Upchurch, IDEX Europe GmbH, Erlangen, Germany). The third port of the T-piece was connected to the HILIC flow pump, (Fig. S2). The injection volume was 10 μL. The mobile phase of the serial coupling (RPLC-HILIC-ToF-MS) was employed as follows: Solvent A: 10 mM ammonium acetate in 90:10 (v/v) water–acetonitrile; solvent B: 10 mM ammonium acetate in 10:90 (v/v) water–acetonitrile; solvent C: acetonitrile; solvent D: water. Further mobile phase conditions are summarized in Table S1 [14,19]. Samples were detected with a 'time-of-flight' mass spectrometer (Agilent Technologies, Waldbronn, Germany), equipped with the Jet Stream ESI interface. Ions were detected in positive ionization mode with a mass range of 50–2100 Dalton. The resolution of the instrument was better than 10,000 at *m/z* 922. The parameters were as follows: 325 °C gas temperature, 10 L/min drying gas flow, 325 °C sheath gas temperature, 7.5 L/min sheath gas flow, 45-psi nebulizer operating pressure and 100 V fragmentor voltage.

2.6. Data evaluation

2.6.1. Non-spectroscopic data evaluation

The extraction yield was calculated according to the equation:

$$\text{Extractionyield} = \text{evaporatedextractablematter(g)}/0.5\text{g}\times 100, \quad (1)$$

The residue of the extracts was weighed and divided by the weight of the milled powder.

2.6.2. Mass spectrometric data evaluation

The data were acquired with MassHunter Workstation LC/MS Data Acquisition software B 05.00, (Agilent Technologies, Waldbronn, Germany) and processed with Agilent Profinder B.06.00 (Agilent Technologies) to extract the RT and determine the exact mass of various *Lemna minor* extracts for the triplicate injections of each sample, after detaching the features found in the blank samples. The parameters were set to a peak filter of 1000 counts peak height, ion species to 'positive ions' with H^+, Na^+, K^+, and NH_4^+, 'charge state' to 1, 'expected RT' to ±3.00 min, and mass to ±5 ppm. The extracted EICs were smoothed with a Gaussian function using a function width of 9 points and a Gaussian width of 5000 points. This limits the result to 2000 compound groups, which were subsequently exported as. cef files to Mass Profiler Professional (MPP) 13.1.1 for further data evaluation. In MPP, the retention time correction was done subsequently by the regression curve without standard delta RT corresponding to each compound's RT. The compounds were then aligned according to the following parameters: RT window of 0.15 min and mass window 5 ppm. The cut-off was set at twofold and all entities with fold change values larger or equal to two were displayed. The different extracts are represented on the scatter diagrams in accordance with the fold change (FC), which is the absolute ratio between the normalized averaged intensities of the compound between two extracts. The fold change was calculated according to the following equation, in which the data is considered without the logarithm:

$$y = (FC)x,\ y = x,\ y = (1 \div FC)x \quad (2)$$

$$y = x + \log(FC),\ y = x,\ y = x - \log(FC) \quad (3)$$

Data and statistical analysis were conducted in Origin 2017 (Origin Lab Corporation, Academic) with SIMCA statistical software (Malmö, Sweden), respectively. Further analysis and data evaluation was performed with Microsoft Excel 2016 (Washington, USA).

3. Results

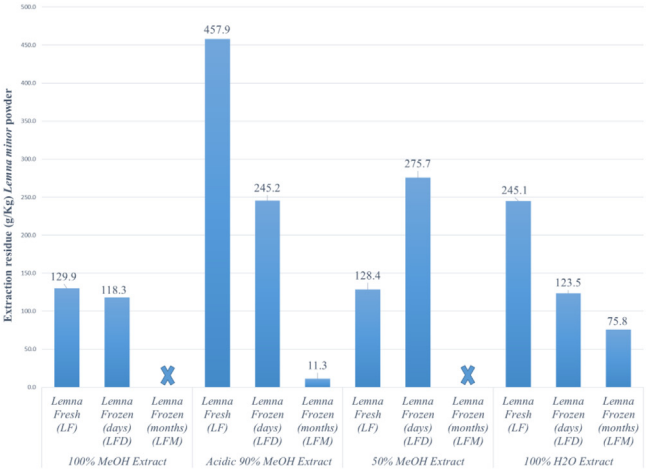
There are many factors that can influence the analysis of plant metabolites. This study investigates the effects of extraction and storage procedures on the (secondary) metabolite content of exemplary plant samples using a non-target screening strategy. The extended-polarity extraction method using *Lemna minor* enabled analysis of its metabolites with RPLC-HILIC-ESI-TOF-MS. Initially, five different solvents were used to ensure the effective extraction of molecules with a broad polarity variation. It was found that increasing the aqueous content of the solvent increased the content of extracted (very) polar compounds. On the other hand, it decreased the solubility of moderated polar to non-polar compounds. The converse behavior could be observed using less polar solvents such as ethanol. The non-target screening strategy reflects the differentiation between various dissolved metabolites, in addition to, the behavior of *Lemna minor* metabolites under different storage periods.

3.1. The extraction method and yields of Lemna minor

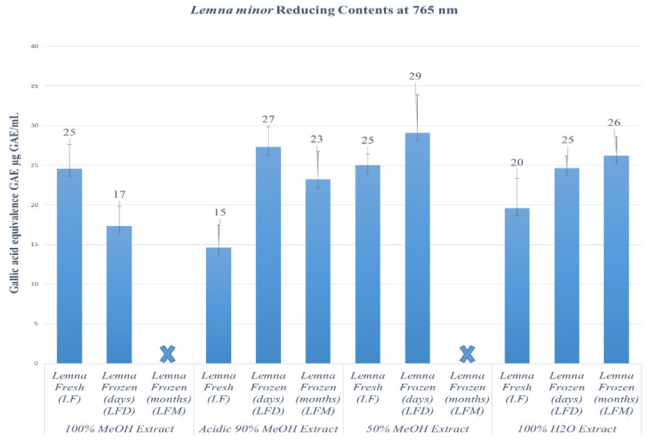
This study applied the extraction method previously approved by Kaufmann et al. [17,18]. The method provides a high extraction yield with minimal pretreatment. As stated in the literature, the extraction method was modified to suit the requirements of extended-polarity chromatography in non-target screening analysis. Fig. 1A shows the extraction yield plot of *Lemna minor* expressed as g/kg *Lemna* powder). Different trends in extraction yields were obtained with the different solvents, although the weight of the *Lemna minor* powder and the extraction process remained the same for all extraction solvents. A comparison of the extraction yields of *Lemna* Fresh (LF), frozen for days (LFD) and frozen for months (LFM) shows that the different *Lemna* samples have a range of between 0 and 457.9 g/Kg of LF in acidic 90 % MeOH extract. LF has, in general, the high extraction yields 129.9, 457.9 and 245.1 g/Kg in100 % MeOH, acidic 90 % MeOH and 100 % H_2O extracts, respectively, but LFD also gives significantly high yields, whereas the measured yields of LFM seem to be rather low. However, as reported in [20], acidified MeOH was applied as a solvent to improve the extraction of the plant metabolites. In general, good extraction efficiencies for *Lemna minor* were obtained with acidic 90 % MeOH, but also in 50 % MeOH and 100 % H_2O solvents.

3.2. Reducing contents of Lemna minor

The reducing contents were expressed as gallic acid equivalents (GAE), measured by molecular absorbance at 765 nm; these are presented in (Fig. 1B). According to the results, the extracts contained a mixture of reducing compounds with different polarities. The reducing contents for LF, LFD, and LFM ranged from 15 μg GAE/mL in an acidic 90 % MeOH extract to LFD with 29 μg GAE/mL in a 50 % MeOH extract of LFD. In general, LFD in the extracts often creates a significantly higher reducing content compared to the same LF extract. The freezing of *Lemna minor* clearly increases the reducing contents in the extracts (excluding the 100 % MeOH extract).



(A)



(B)

Fig. 1. (A) The extraction yield is expressed in g/Kg of *Lemna minor* Fresh (LF), Frozen for days (LFD), and Frozen for months (LFM). In 100 % MeOH, acidic 90 % MeOH, 50 % MeOH, and 100 % H2O extracts, respectively. (B) The reducing contents of *Lemna minor* Fresh, Frozen for days, and Frozen for months in 100 % MeOH, acidic 90 % MeOH, 50 % MeOH, and 100 % H2O extracts, respectively. The reducing contents are expressed as Gallic acid equivalent measured at the absorbance of 765 nm.

Lemna minor contains different compounds of various natures, such as fatty acids, saturated carboxylic acids, aldehydes, ketones, aromatic alcohols and phytosterols [21]. Moreover, duckweed contains protocatechuic aldehyde, p-hydroxyl benzaldehyde, truxillic acid and p-coumaric acid [3].

In conclusion, there is no direct relation between the extraction yield and the reducing contents of different *Lemna minor* samples in the different extracts.

3.3. *RPLC-HILIC-MS Analysis of different extracts of* Lemna minor *samples*

Fresh *Lemna minor* samples were extracted with the four solvents 100 % MeOH, acidic 90 % MeOH, 50 % MeOH and 100 % H_2O, producing the results presented above. The extraction method is illustrated in detail in (Fig. S1). The four fresh *Lemna minor* LF extracts were injected into a RPLC-HILIC-ESI-TOF-MS system. The mass spectrometric total ion chromatogram (TIC) was analyzed to determine the feature (i.e. molecule) numbers after excluding the experimental background. The TIC was interpreted to obtain extracted ion chromatograms (EICs), which enabled the extraction of the retention time/mass plot for each extract (see Fig. 2). In the RT/mass plot, the RT of features are plotted on the x-axis in minutes

Table 1
The number of compounds in the four extracts of *Lemna minor* fresh (LF) separated by HILIC with RT (0-15 min) and RPLC columns with RT (16-33 min).

Sample	*Lemna minor* fresh (LF)		
	HILIC	RPLC	Total
100 % H_2O	806	292	1098
50 % MeOH	701	410	1111
Acidic 90 % MeOH	595	502	1097
100 % MeOH	686	383	1069

while molecular masses are plotted on the y-axis in Dalton units. The polar to (very) polar molecules separated by HILIC are marked by blue triangles and the polar to non-polar molecules separated by RPLC are marked by red circles. The reason why the retention time of 15 min can be translated into a logD value of zero is that the break between the molecules eluting from HILIC and RPLC was in a retention window of 16 and 21 min. A hard limit was defined at the RT for metformin of 15.0 ± 0.3 min, as it was the last very polar standard compound eluting from the HILIC column (Table S2). A detailed definition of the polarity classification used in this study was published in Bieber et al. [14]. In (Table 1), the feature values for the four extracts of LF are ordered in decreasing solvent polarity: 100 % H_2O > 50 % MeOH > 90 % MeOH (+ 0.5 % FAC) > 100 % MeOH. The 0.5 % formic acid lowered the pH of the solvent, resulting in the lowest number of compounds in the HILIC component (see Table 1). Increasing the water content to 50 % led to a minor increase in the compounds found in the HILIC and RPLC components (see Table 1). Although 100 % aqueous extract has a similar quantity of total compounds, the highest significant compound number of polar and (very) polar molecules appeared in the HILIC component. On the other hand, the molecule number in the RPLC component decreased significantly.

So far, the comparison of the four extracts has been presented in terms of general values, such as extraction yield, reducing content and number of compounds. However, it is possible that the molecules in various extracts may be different or the same. For this reason, the features from the retention time/mass plot are evaluated with regard to their precise mass (in other words, their empirical formula, i.e. their identical atomic composition). The empirical formula for most of the compounds was determined using the mass spectrometric supplier software Mass Profiler Professional (MPP). This was the basis on which the comparison between the compounds of the different extraction samples was performed. Due to the significantly different pH values and resulting other charging properties, the acidic 90 % MeOH extracts will be neglected from this point on in the study. The results for the 100 % MeOH, 50 % MeOH and 100 % H_2O extracts of *Lemna minor* fresh LF are summarized in the Venn diagrams in Fig. 3. The Venn diagrams show the unique, overlapping compounds between the extracts (inside the cycles) as well as the total number of compounds in each extract (outside of the cycles). The blue numbers again represent the HILIC-separated compounds, while the red numbers represent the RPLC-separated compounds. As can be seen in Table 1 and Fig. 3, the total amount of extracted molecules remains similar, and the absolute number of detected molecules is always in a HILIC / RPLC ratio larger than 1. Moreover, it is obvious that the 100 % H_2O extract has the highest number of HILIC-separated molecules.

Although the 50 % MeOH extract has a large quantity of different molecules in the HILIC component, it has a less unique number compared to 100 % MeOH and 100 % H_2O extracts (see Fig. 3A). On the other hand, 50 % MeOH extract overlapped in a large number of compounds with 100 % MeOH and 100 % H_2O extracts, as a consequence of the polarity median of both solvents. However, the latter two extracts – with a wide polarity discrepancy – contain 36 common compounds, and all three extracts contain at least 35 common

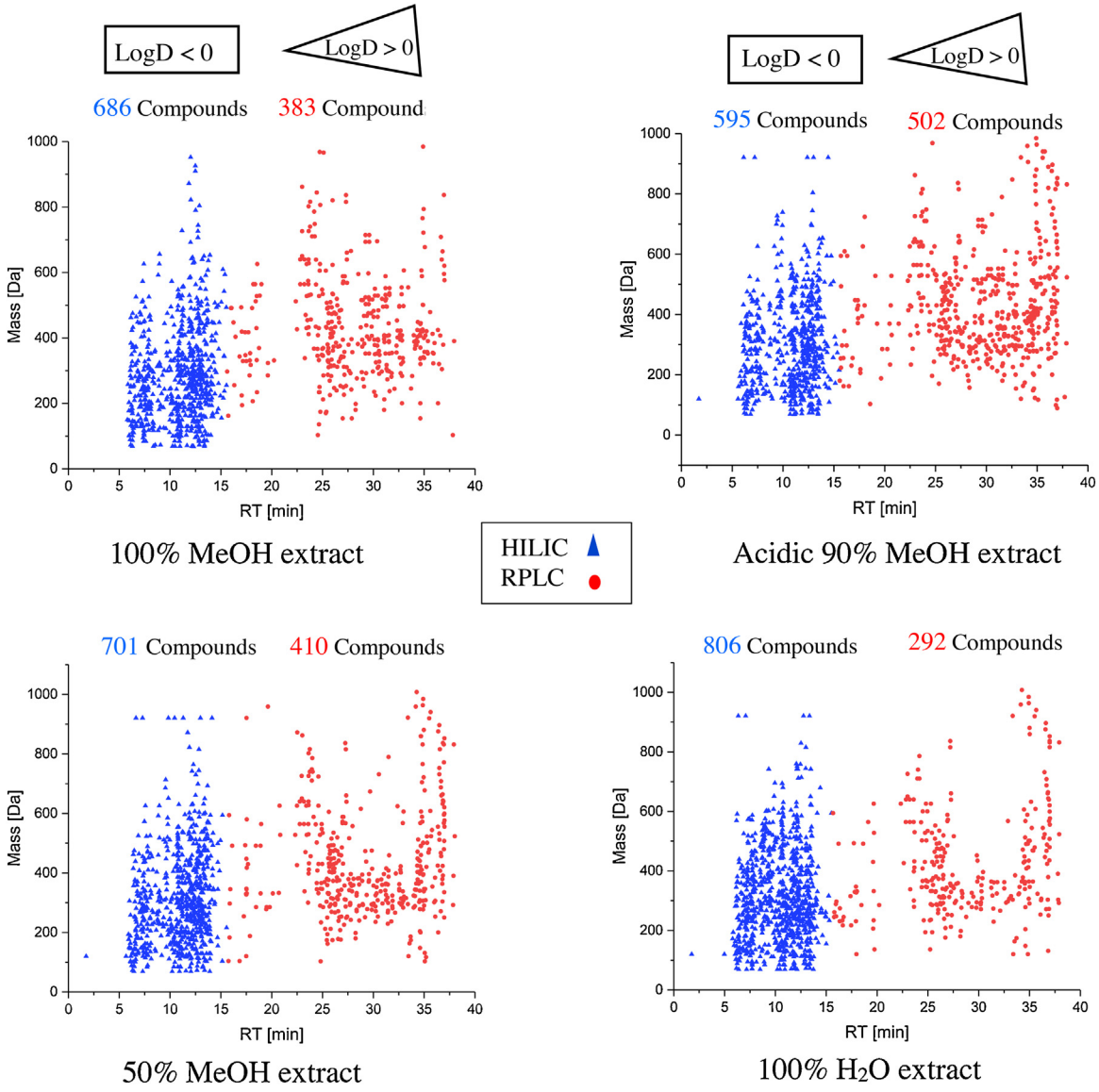


Fig. 2. Retention time / mass plot of *Lemna minor* fresh sample (LF) 100 % MeOH, acidic 90 % MeOH, 50 % MeOH, and 100 % H2O extracts, respectively. The numbers above each plot represent the total number of molecules separated with HILIC in blue and RPLC in red. Blue, triangles represent HILIC retarded molecules eluting (0-15) min, which have LogD values below zero and red, circles represent RPLC retarded molecules eluting (16–38) min, which have LogD values above zero. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

molecules. The same is almost true for the RPLC component (in an inverse manner).

Although the 50 % MeOH extract has a large quantity of different molecules in the RPLC component, it has a less unique number compared to 100 % MeOH (see Fig. 3B). On the other hand, 50 % MeOH extract overlapped in a large number of compounds with 100 % MeOH and 100 % H_2O extracts because of the polarity median of both solvents. However, the latter two extracts – with a wide polarity discrepancy – contain nine common compounds, and all three extracts contain at least 26 common molecules. The small quantity of non-polar molecules in the MeOH extraction – observed with RPLC – confirms the previous polarity argumentation.

3.4. *Comparison of* Lemna minor *samples: fresh, frozen (days) and frozen (months)*

The chemical and bioactive nature of plant metabolites necessitates the use of methodologies that do not alter their concentration or structure. Previous studies have shown that the metabolite yield could be affected by changing the solvent used for extraction, the duration of extraction, and the storage periods of the plant material or extracts, as well as the biochemical activity, heat, light, vacuum, and drying procedures [22]. It is therefore important, when investigating molecular fingerprints using a non-target screening strategy, to study the effect of storage on plant metabolites. It has recently been shown that storage can alter the plant's metabolites [23]. Three samples of *Lemna minor* were investigated in this study: fresh (LF), stored for days at −80 °C (LFD), and stored for months at −80 °C (LFM). The extraction yield and the reducing contents of the samples are described above and shown in (Fig. 1A and B) respectively. The acidic 90 % MeOH has a considerable reducing content, although, it has a lower feature number in LFM compared to the 100 % H_2O extract, (Fig. S2). The three 100 % H_2O extract samples (analyzed with RPLC-HILIC-ESI-TOF-MS) were shown as examples in this study and the results discussed in terms of their presence in the samples. An orthogonal partial least squares discriminant analysis (OPLS-DA) model was drawn up with two components. The horizontal component distinguishes between groups, while,

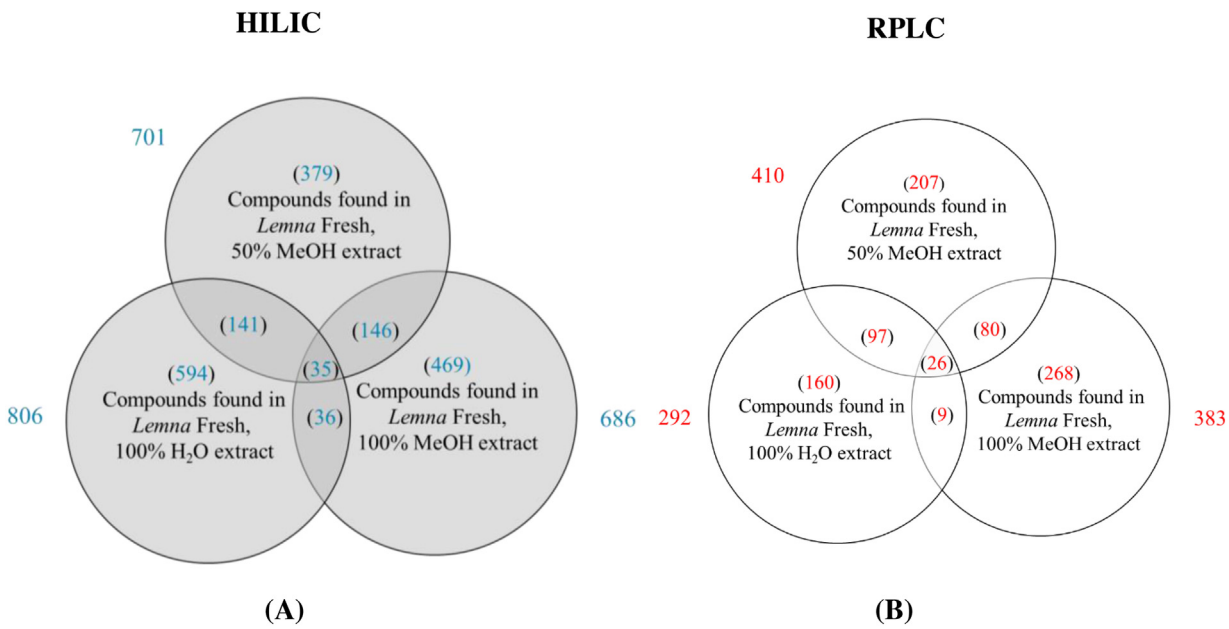


Fig. 3. Venn diagrams summarizing the number of shared and unique compounds in three extracts of *Lemna minor* Fresh sample. (A) Compounds measured in the HILIC column (with a RT ≤ 15) and are blue colored. (B) Compounds measured in the RPLC column (with a RT > 15) and are red-colored. The number of the total compounds found outside each circle, unique compounds of the extracts in the inside and the shared compounds. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

the vertical one differentiates within a group. Although it resembles PCA, the data is guided by known class information taken from the prediction model. Consequently, OPLS-DA is a supervised modeling approach. In Fig. 4, the OPLS-DA model is at significance level α = 0.05. The OPLS-DA model characterized 86 % of the variation in X (R^2X (cum) = 0.86); 100 % in response Y (R^2Y(cum) = 1). The Y was predicted by the 7-fold cross-validation (Q^2(cum) = 1). The high-value parameters illustrate the good classification and prediction ability of the OPLS-DA model. Furthermore, the model demonstrates the significant strength and reproducibility of the data. The model indicates that the LFD and LFM are clustered separately from the LF. The vertical component distinguishes between LFD and LFM. Moreover, it shows that LF and LFD are more correlated. This supports the argument that the features in LF and LFD are highly correlated.

The features were extracted in Venn diagrams containing the LF, LFD, and LFM 100 % H_2O extracts, as presented in (Fig. 5). The unique, overlapped and total compound numbers are presented as described above. The blue color (on the left-hand side) represents compounds separated with the HILIC column, whereas the red color (on the right-hand side) represents compounds separated with the RPLC column. The direct feature comparison between the LF, LFD, and LFM 100 % H_2O extracts shows that LFM has a significantly lower amount of molecules in the HILIC component, whereas there is a significantly higher amount of molecules in LFM that cannot be found in LF and LFD. Accordingly, (very) polar compounds might be degraded into less polar compounds by freezing, these having been retarded on the RPLC. Furthermore, these compound structures may be alternated into a less polar compound. In addition, LFM has the highest number of unique compounds in the RPLC component, which confirms the polarity argumentation.

The number of overlapping compounds in the HILIC and RPLC components between LFD and LF is significantly higher than between LFM and LF. However, the total amount of overlapping compounds in the three extracts is very small in comparison with the numbers of unique compounds. In particular, the LFM sample differs significantly (in the overall polarity region) from the other two samples.

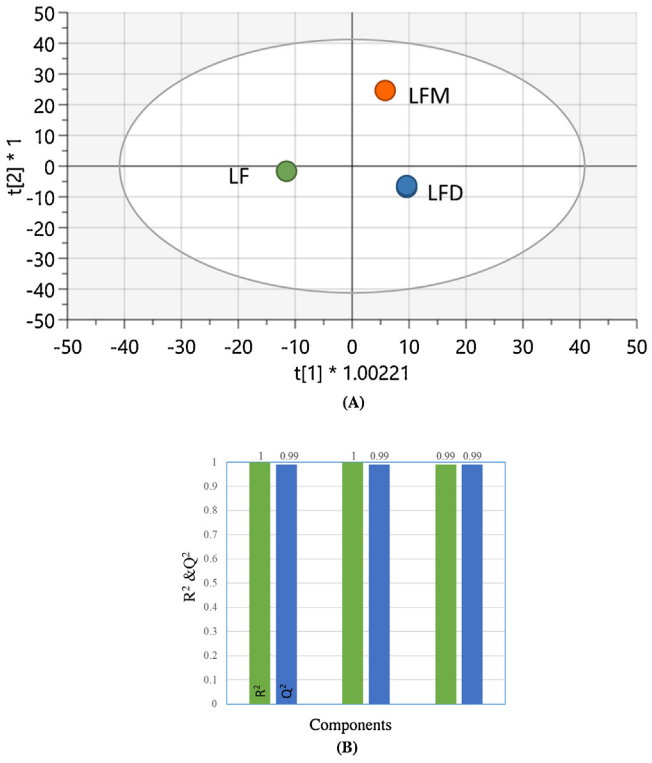


Fig. 4. (A) The Orthogonal partial least squares discriminant analysis (OPLS-DA) model of *Lemna minor* fresh sample (LF), frozen for days (LFD) and frozen for months (LFM) 100 % H2O extracts. Scaled proportionally to R2X, R2X[1] = 0.238, R2X[2] = 0.243, Ellipse: Hotelling's T2 (95 %). (B) X/Y overview plot displays the individual cumulative R2 (green columns) and Q2 (blue columns) for the goodness of fits and cross-validation parameters. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Therefore, the data was considered without the logarithm, according to the fold change. The data was plotted between two samples, as shown in (Fig. 6

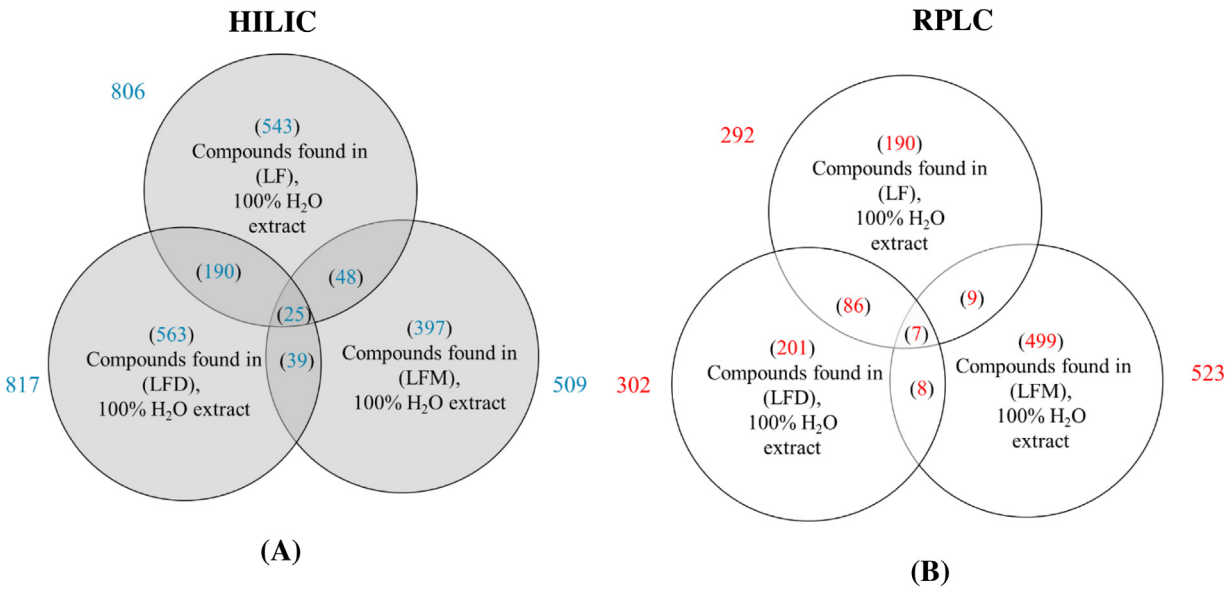


Fig. 5. Venn diagrams summarizing the number of shared and unique compounds in 100 % H2O extracts of *Lemna minor* fresh sample (LF), frozen for days (LFD) and frozen for months (LFM). (A) Compounds measured in the HILIC column (with a RT ≤ 15) and are blue colored. (B) Compounds measured in the RPLC column (with a RT > 15) and are red-colored. The number of the total compounds found outside each circle, unique compounds of the extracts in the inside and the shared compounds. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

). Fig. 6A shows a bilateral comparison plot of an LFD fingerprint versus an LF fingerprint. Compounds exclusively found in an LFD 100 % H_2O extract were located on the y-axis (Fig. 6A) (i.e. plot region I), whereas LF 100 % H_2O extract compounds are exclusively located on the x-axis (Fig. 6A), (i.e. plot region V). The compounds presented in region II (Fig. 6A) are found in both extracts with higher signal intensities in LFD 100 % H_2O extract. Compounds close to the axis halving line (region III) of (Fig. 6A) were found in both extracts obtained with similar signal intensities. However, the compounds in region IV (Fig. 6A) have higher signal intensities in the LF 100 % H_2O extract. The results reflect that *Lemna minor* might be degrading its original metabolites due to the longer storage period at −80 °C. In Fig. 6B, the unique compounds of LF 100 % H_2O extract were located on the x-axis and the compounds unique to the LFM 100 % H_2O extract were located on the y-axis. However, Fig. 6C presented unique compounds in the LFD 100 % H_2O extract on the x-axis versus unique LFM 100 % H_2O extract compounds on the y-axis. The compounds on the y-axis could not be found in LF and LFD, which may be due to degradation of the *Lemna minor* metabolites. Moreover, in a direct bilateral comparison between LFD and LFM, the compounds of both samples have higher signal intensities in the LFM, as shown in (Fig. 6C) region II).

There are fourteen compounds in Table 2 that could be suggested in the LF 100 % H_2O extract by comparing measured masses with molecular masses of compounds found in the Lemnaceae and/or Araceae family. On the one hand, there are the four amino acids L-histidine, L-leucine, L-isoleucine and L-tyrosine, which were previously identified in *Lemna minor*, and, on the other hand, the three flavonoids luteolin 7-rutinoside, heptamethoxyflavone, and kaempferol-3-O-rutinoside. Furthermore, dimethyl malate, nicotinic acid, adenine, 2,5 dihydroxymethyl-3,4-dihydroxypyrrolidin (DMDP), uracil-1-beta-D-arabinofuranoside, 3,4-dihydroxycinnamic acid methyl ester, and 3,4,5-trimethoxyallylbenzene were also characterized.

L-histidine, L-leucine, L-isoleucine were also identified in varying intensities in the LFD sample (see Table 3).

Moreover, in the LFD 100 % H_2O extract, a small amount of L-methionine was identified in the *Lemna minor,* (Table 2). Also, the L-phenylalanine signal intensity increased, although the L-tyrosine signal intensity decreased below the measured limit. Consequently, the signal intensity of heptamethoxyflavone increased (Table 3). Isorhamnetin-3-galactoside was also characterized. However, there was a slight decrease in schaftoside, isoschaftoside, and isovitexin-7-O-xyloside signal intensities.

In addition, 2-aminohexanedioic acid, a metabolite of lysine metabolism, was suggested in addition to this, hexose and apiose were also characterized. Consequently, 2,6-dideoxy-2,6-iminoheptitol, which is a potent α-glucosidase inhibitor, could also be characterized.

In the LFM 100 % H2O extract, there was a significant increase in the expected L-methionine signals. However, the signal intensities of other amino acids decreased below the detection limit. A flavanoid scoparin 2′′-xyloside was potentially characterized, (Table 2). Furthermore, schaftoside, isoschaftoside and Isovitexin-7-O-xyloside signals increased significantly in LFM (Table 3). Signal intensities also increased for 2,5 dihydroxymethyl-3,4-dihydroxypyrrolidin (DMDP) and 2,6-dideoxy-2,6-iminoheptitol. The sugars fell below the detection limit.

4. Discussion

Developments in analytical technologies have resulted in a substantial expansion to the approach used in metabolomic research [1]. Consequently, the investigation focused on the changes in the metabolomics of *Lemna minor* due to differences in storage conditions, using a non-target screening strategy. The non-target screening required an extraction method with minimal pretreatment. In general, LF produces higher extraction yields than LFD and LFM in 100 % MeOH, acidic 90 % MeOH and 100 % H2O extracts, respectively. The disparity in the extraction yield of different extracts may be due to the different solubilities of extractable components resulting from their varied chemical compositions. Previous authors were generally more interested in questions concerning 100 % MeOH and 100 % EtOH extraction power. Sultana et al. (2009), concluded that the 100 % MeOH extract produced higher extraction yields of total phenolic and flavonoids from various organs of different plants than the 100 % EtOH extract [6]. Furthermore, Eloff, 1998, pointed out that 100 % MeOH in extrac-

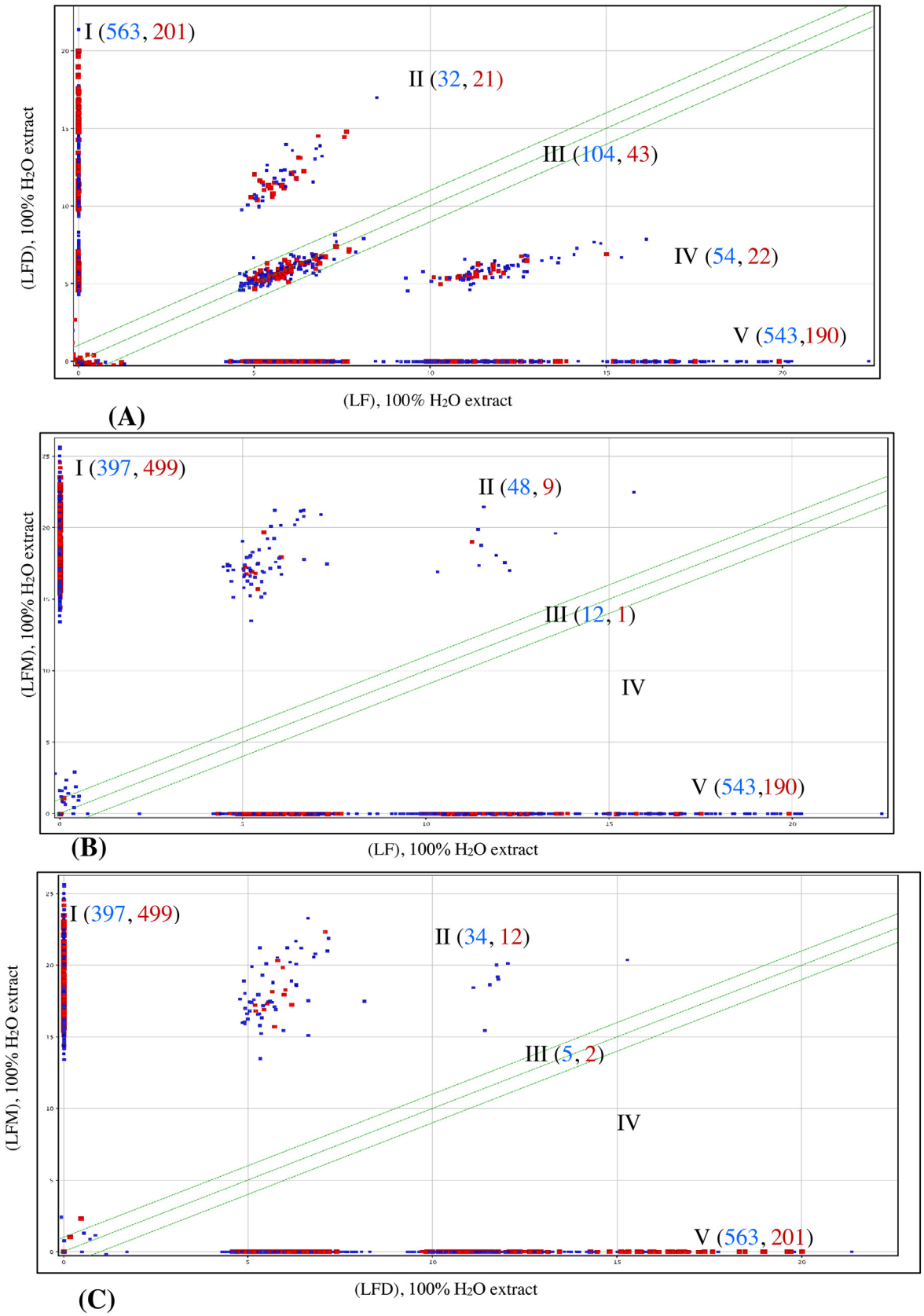


Fig. 6. (A) Feature (signal intensity) comparison plot of *Lemna minor* Fresh (LF) 100 % H2O extract versus *Lemna minor* Frozen for days (LFD) 100 % H2O extract. (B) *Lemna minor* Fresh (LF) 100 % H2O extract versus *Lemna minor* Frozen for months (LFM) 100 % H2O extract. (C) *Lemna minor* Frozen for days (LFD) 100 % aqueous extract versus *Lemna minor* for months (LFM) 100 % aqueous extract. Blue, small squares represent RT (0-15 min) and red, big squares represent the RT (16-33 min). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).
I. Compounds exclusively found on y-axes.

Table 2
List of molecular names characterized in *Lemna minor* Fresh (LF), Frozen for days (LFD), and Frozen for months (LFM) 100 % H2O extracts with a ToF-MS along with the monoisotopic masses values with respective deviations in ppm, logD values at pH 7, respective the intensities and Retention times of compounds.

Name	Measured Mass (Da)	Mono-isotopic mass (Da)	Δ ppm	RT (Min)	LogD (pH7)	Intensity*
LF						
Dimethyl Malate [31]^{X}	162.0533	162.0530	−1.85	7.72	−0.82	6.91
L-histidine [32]*	155.0689	155.0690	−6.45	14.98	−3.70	12.99
Nictonic acid [33]	123.0314	123.0320	3.25	9.24	−2.76	12.56
Adenin [34]	135.0544	135.0540	−2.96	6.81	−0.58	16.42
2,5-Dihydroxymethyl-3,4-dihydroxypyrrolidin (DMDP) [35]^{X}	163.0835	163.0840	3.077	6.37	−4.76	6.48
Luteolin 7-rutinoside [36]^{X}	594.1601	594.1580	−3.53	15.40	−0.77	6.02
1-beta-D-ribofuranosyluracil [37]*	244.0701	244.0695	2.46	8.97	−2.42	6.80
3,4-dihydroxycinnamic acid methyl ester [38]^{X}	194.0584	194.0580	−2.06	25.41	1.91	5.60
L-leucine L-isoleucine [32]*	131.0952	131.0950	−1.53	10.88	−1.59	7.20
Heptamethoxyflavone [39]*	432.1413	432.1420	−1.62	23.74	1.89	5.79
Kaempferol-3-O-rutinoside [40]$^{+}$	594.1601	594.1585	2.69	15.4	−1.72	6.02
3,4,5-Trimethoxyallylbenzene [41]^{X}	208.11	208.1099	0.48	8.80	2.60	5.49
L-tyrosin [32]*	181.074	181.0740	9.39	11.82	−1.49	5.95
LFD						
Isorhamnetin-3-galactoside [42]^{X}	478.1109	478.1110	0.21	13.48	−0.74	5.20
D-hexose [3]*	180.0632	180.0633	−0.56	7.07	−2.93	7.65
Apiose [43]*	150.0526	150.0530	2.67	6.29	−2.44	5.60
2,6-Dideoxy-2,6-iminoheptitol [35]^{X}	193.0945	193.095	2.60	6.46	−4.21	5.55
L-methionine [32]*	149.0505	149.0510	3.35	9.99	−2.19	6.63
2-Aminohexanedioic acid [44]^{X}	161.0702	161.0690	−7.45	13.23	−5.29	5.18
L-phenylalanine [32]*	165.0791	165.0790	0.61	10.58	−1.19	7.46
LFM						
L-methionine [32]*	149.0505	149.0510	3.35	9.99	−2.19	19.68
Scoparin 2''-xyloside [47]^{X}	594.1594	594.1580	−2.36	16.82	−2.25	20.80
Found in all samples						
Isovitexin-7-O-xyloside[48]^{X}	564.1536	564.1480	−9.93	23.60	−2.06	–
Isoschaftoside [45]* Schaftoside [46]^{X}	564.1492	564.1479	2.30	19.15	−3.91	–

References are marked as (*) for literature dealing with *Lemna minor*, (x) for literature discussing compounds from Lemnaceae family and/or Araceae family, and (+) for literature in which the core-flavonoid was found without glycosylation or glucuronisation.The references can be found in the supplementary part.

* It is the absolute ratio between the normalized averaged intensities of the compound between two extracts.

tion was more efficient than 100 % EtOH in extraction for eluting antimicrobial compounds in different plants based on different parameters [24]. A survey conducted by Miliauskas et al. (2004) showed that the 100 % MeOH extract of twelve different plants had a high level of radical scavenging activity. This strongly suggested that 100 % MeOH extracts contained the highest number of antioxidant compounds attributed to the defense mechanism of the plant that protects it against stress factors [25]. In the present study, 100 % EtOH as an extraction solvent has a significantly lower efficient extraction yield than 100 % MeOH as an extraction solvent, as a result of which EtOH was not applied any further in this study. It can be concluded from the above findings that different solvent systems dissolve components differently, and that a variety of solutions with a wide range of polarities may be used for maximum extraction. However, it should be borne in mind that a higher extraction yield does not necessarily result in a high reducing content or a higher number of different compounds, which is indicated by the intrinsic nature of the components (such as the acidic 90 % MeOH LFM shown in Fig. S3). On the other hand, the reducing contents of the *Lemna minor* increased with freezing, which can change their phenolic behavior during the freezing period by way of molecular transformation, such as oxidation.

The higher yield of LF extracts was analyzed using a RPLC-HILIC-ESI-TOF-MS system. 100 % H_2O displays the highest number of features, while 100 % MeOH has the lowest number. Although the 100 % MeOH extract is the most non-polar solvent in this study, it generally has a higher number of compounds in the HILIC component compared to the RPLC component. The same is true of extractions containing MeOH. MeOH is an amphiphilic compound with a polarity index of 5.1. This means that MeOH is able to dissolve non-polar and (very) polar compounds. The 0.5 % formic acid reduced the number of compounds in the HILIC component. This may be due to the neutralization of acidic compounds that elute later on, resulting in the highest number of observed molecules (in the RPLC section). Based on this argumentation, it can be expected that acidic extraction will produce a small quantity of molecules with basic functional groups. Moreover, the presence of formic acid in the acidic 90 % MeOH extract may cleave the glucoside bonds and release the phenolic compounds [20]. The project supported the non-target screening (NTS) strategy, which aims to assign as many compounds as possible. Thus, the combined results from the various extracts may be a precise reflection of the NTS strategy. Subsequently, the acidic 90 % MeOH extract was eliminated due to the differences in pH values. The different compounds of 100 % MeOH, 50 % MeOH, and 100 % H_2O extracts of *Lemna minor* fresh LF were compared in Venn diagrams, as shown in Fig. 3. The 50 % MeOH extract overlapped in a large number of compounds with 100 % MeOH and 100 % H_2O extracts in the HILIC and RPLC components, as a consequence of the polarity median of both solvents. The combination of 100 % MeOH, 50 % MeOH and 100 % H_2O extracts (with similar pH-values) is appropriate for *Lemna minor* extraction and determining its chromatographic fingerprint with RPLC-HILIC-ESI-TOF-MS. With regard to the three solvents, a very large number of *Lemna minor* metabo-

II. Compounds found in both samples with increased feature intensity on the y-axis compared to the x-axis.
III. Compounds found in both samples.
IV. Compounds found in both samples with decreased feature intensity on the y-axis compared to the x-axis.
V. Compounds found exclusively on x-axes.

Table 3
List of molecular names characterized in *Lemna minor* Fresh (LF), Frozen for days (LFD), and Frozen for months (LFM) 100 % H_2O extracts with a ToF-MS along with respective intensities.

Name	Intensities*		
	LF	LFD	LFM
L-histidine	12.99	6.27	–
L-leucine L-isoleucine	7.20	15	–
Heptamethoxyflavone	5.79	6.07	–
2,5-Dihydroxymethyl-3,4-dihydroxypyrrolidin (DMDP)	6.48	–	18.1
L-methionine	–	6.63	19.68
2,6-Dideoxy-2,6-iminoheptitol	–	5.55	16.73
Isovitexin-7-O-xyloside	6.04	5.97	17.92
Isoschaftoside Schaftoside	6.05	6.04	19.45

* It is the absolute ratio between the normalized averaged intensities of the compound between two extracts.

lites were extracted. An overall picture of the metabolites was obtained.

It is convenient to use a non-target screening strategy as a basis for the extended-polarity chromatographic fingerprint workflow with plant extracts. In particular, the large number of (very) polar molecules dissolved in 100 % H_2O extract and detected in the HILIC component highlights its applicability for future measurements in the (very) polar region.

The comparison of LF, LFD, and LFM metabolites using the OPLS-DA model showed that the data was significantly separated, also LFD and LFD was more correlated than LFM. In a bilateral comparison, the metabolites of *Lemna minor* were located differently according to the storage periods. Some metabolites were degraded due to having been stored for a long period at −80 °C. In addition, the relative intensities of the metabolites changed due to storage.

Various compounds were detected and identified. Fourteen compounds could be suggested in LF 100 % H2O extract and three amino acids (L-histidine, L-leucine, and L-isoleucine) were identified by reference material. In the LFD sample, the L-phenylalanine signal intensity increased, which is a precursor of the flavonoids' biosynthesis pathway in the plant. Hence, the signal intensity of heptamethoxyflavone was increased and isorhamnetin-3-galactoside could be characterized. On the other hand, there was a slight decrease in the signal intensities for schaftoside, isoschaftoside and isovitexin-7-O-xyloside. New flavonoids were characterized in the LFM sample. Moreover, there was a significant increase in schaftoside, isoschaftoside, and Isovitexin-7-O-xyloside signals. Also, there was a significant increase in L-methionine. It increases especially in plants suffering from stress conditions [26].

The intensities of polar carbohydrates (such as hexose) fell below the measured detection limit. Thus, LFM had the lowest number in the HILIC component, which may be due to the increase in reducing sugar ('low-temperature sweetening') [27]. The same effect was detected with exposure of Scots pine to SO_2 whereby the small concentration increased the production of carbohydrates and the larger one reduced it [28]. Hence, the freezing of *Lemna minor* for short periods generally decreases the intensity of some compounds. However, freezing for long periods clearly generates new compounds. Freezing under −80 °C did not serve to arrest the activity in *Lemna minor*. It is possible that some enzymes were still active and degraded metabolites. Furthermore, they altered the biosynthetic pathways in different ways. Similar findings were reported for onion storage, i.e. the quercetin concentration decreased but its conjugates increased [23]. Bilia et al. reported in 2002 that the concentration of flavonoids in calendula, milk-thistle and passion flower tinctures resulted in different behaviors during storage depending on the type of flavonoids [29]. Researchers reported that the concentration of secondary metabolites was reduced due to freezing at different temperatures [30].

In conclusion, it is essential to use fresh samples to perform correct analytical measurements in (plant) metabolomic (i.e. non-target screening) studies, as plant freezing changes the molecular content.

5. Conclusions

Three types of *Lemna minor* samples were extracted using different solvents. The extractions were then dried and redissolved residues analyzed using extended-polarity RPLC-HILIC-TOF-MS to determine their chromatographic/mass spectrometric fingerprint. This study produced an extended-polarity non-target screening workflow that clearly reflects the great importance of detecting (very) polar molecules separated by HILIC in the plant metabolome (in addition to the classic non-polar to polar molecules separated by RPLC). The ultimate decision to use 100 % MeOH extracts, 50 % MeOH and 100 % H_2O extracts paves the way to new views of plant metabolomes for (very) polar molecules. Non-target screening was applied to study the aging behavior of *Lemna minor* under different freezing periods. The comparison of compounds in *Lemna minor* fresh and frozen samples showed that freezing under −80 °C of *Lemna* changes the intensity and presence of compounds. In order to avoid any false positive and false negative interpretations, it is necessary to extract *Lemna minor* freshly in order to determine its metabolomic fingerprint. The results can then be used in phytomedicine research. Furthermore, the features of unknown identity in this study can be subjected to further investigation in the form of various univariate and/or multivariate statistical analyses. This will then lead to mutual correlations of features. Where features relevant to a process are characterized by this analytical strategy, compound and/or analytical libraries can be used to identify the compounds. A new and exemplarly novel compound library containing plant metabolites (like 'PLANT-IDENT on the open-access platform FOR- IDENT (https://www.for-ident.org)) will help to identify positive results. This and other data evaluation will be published soon.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was partially funded by the Bavarian State Ministry of the Environment and Consumer Protection as well as the Cultural Affairs and Mission Sector of the Egyptian Ministry of Higher Education. Dr. Stefan Moser is thanked for his support in applying the OPLS-DA model.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jpba.2020.113362.

References

[1] A. Fukushima, M. Kusano, Recent progress in the development of metabolome databases for plant systems biology, Front. Plant Sci. 4 (73) (2013).
[2] Y. Zhang, Y. Hu, B. Yang, F. Ma, P. Lu, L. Li, C. Wan, S. Rayner, S. Chen, Duckweed (Lemna minor) as a model plant system for the study of human microbial pathogenesis, PLoS One 5 (10) (2010) e13527.
[3] X. Zhao, G.K. Moates, N. Wellner, S.R. Collins, M.J. Coleman, K.W. Waldron, Chemical characterisation and analysis of the cell wall polysaccharides of duckweed (Lemna minor), Carbohydr. Polym. 111 (2014) 410–418.

[4] M. Schäfer, C. Brütting, I.T. Baldwin, M. Kallenbach, High-throughput quantification of more than 100 primary- and secondary-metabolites, and phytohormones by a single solid-phase extraction based sample preparation with analysis by UHPLC–HESI–MS/MS, Plant Methods 12 (1) (2016) 30.
[5] P.N.N. Azwanida, A review on the extraction methods use in medicinal plants, principle, strength and limitation, Med. Aromat. Plants (Los Angel) 4 (3) (2015) 1–6.
[6] B. Sultana, F. Anwar, M. Ashraf, Effect of extraction solvent/technique on the antioxidant activity of selected medicinal plant extracts, Molecules 14 (6) (2009) 2167–2180.
[7] K. Ameer, H.M. Shahbaz, J.-H. Kwon, Green extraction methods for polyphenols from plant matrices and their byproducts: a review, Compr. Rev. Food Sci. Food Saf. 16 (2) (2017) 295–315.
[8] J. Bao, R.B. Ding, X. Jia, Y. Liang, F. Liu, K. Wang, C. Zhang, P. Li, Y. Wang, J.B. Wan, C. He, Fast identification of anticancer constituents in Forsythiae Fructus based on metabolomics approaches, J. Pharm. Biomed. Anal. 154 (2018) 312–320.
[9] T. Letzel, Non-target Screening, Suspected-target Screening and Target Screening – of Technologies and Philosophies, Databases and Crafts, Lab&more, 2014.
[10] S. Joshi, HPLC separation of antibiotics present in formulated and unformulated samples, J. Pharm. Biomed. Anal. 28 (5) (2002) 795–809.
[11] Y. Iwasaki, T. Sawada, K. Hatayama, A. Ohyagi, Y. Tsukuda, K. Namekawa, R. Ito, K. Saito, H. Nakazawa, Separation technique for the determination of highly polar metabolites in biological samples, Metabolites 2 (3) (2012) 496–515.
[12] Z. Liu, S. Rochfort, Recent progress in polar metabolite quantification in plants using liquid chromatography–mass spectrometry, J. Integr. Plant Biol. 56 (9) (2014) 816–825.
[13] F. Bucar, A. Wube, M. Schmid, Natural product isolation – how to get from biological material to pure compounds, Nat. Prod. Rep. 30 (4) (2013) 525–545.
[14] S. Bieber, G. Greco, S. Grosse, T. Letzel, RPLC-HILIC and SFC with mass spectrometry: polarity-extended organic molecule screening in environmental (Water) samples, Anal. Chem. 89 (15) (2017) 7907–7914.
[15] D. Hemmler, S.S. Heinzmann, K. Wohr, P. Schmitt-Kopplin, M. Witting, Tandem HILIC-RP liquid chromatography for increased polarity coverage in food analysis, Electrophoresis 39 (13) (2018) 1645–1653.
[16] M. Obermeier, C.A. Schroder, B. Helmreich, P. Schroder, The enzymatic and antioxidative stress response of Lemna minor to copper and a chloroacetamide herbicide, Environ. Sci. Pollut. Res. Int. 22 (23) (2015) 18495–18507.
[17] C.M. Kaufmann, J. Grassmann, T. Letzel, HPLC method development for the online-coupling of chromatographic Perilla frutescens extract separation with xanthine oxidase enzymatic assay, J. Pharm. Biomed. Anal. 124 (2016) 347–357.
[18] C.M. Kaufmann, T. Letzel, J. Grassmann, M.W. Pfaffl, Effect of Perilla frutescens extracts on porcine jejunal epithelial cells, Phytother. Res. 31 (2) (2017) 303–311.
[19] G. Greco, S. Grosse, T. Letzel, Serial coupling of reversed-phase and zwitterionic hydrophilic interaction LC/MS for the analysis of polar and nonpolar phenols in wine, J. Sep. Sci. 36 (8) (2013) 1379–1388.
[20] A. Nostro, M.P. Germano, V. D'Angelo, A. Marino, M.A. Cannatelli, Extraction methods and bioautography for evaluation of medicinal plant antimicrobial activity, Lett. Appl. Microbiol. 30 (5) (2000) 379–384.
[21] I.N. Vladimirova, V.A. Georgiyants, Biologically Active Compounds from Lemna Minor S. F. Gray, Pharm. Chem. J. 47 (11) (2014) 599–601.
[22] P. Jin, S.Y. Wang, C.Y. Wang, Y. Zheng, Effect of cultural system and storage temperature on antioxidant capacity and phenolic compounds in strawberries, Food Chem. 124 (1) (2011) 262–270.
[23] K. Sharma, Y.R. Lee, Effect of different storage temperature on chemical composition of onion (Allium cepa L.) and its enzymes, J. Food Sci. Technol. 53 (3) (2016) 1620–1632.
[24] J.N. Eloff, Which extractant should be used for the screening and isolation of antimicrobial components from plants? J. Ethnopharmacol. 60 (1) (1998) 1–8.
[25] G. Miliauskas, P.R. Venskutonis, T.A. van Beek, Screening of radical scavenging activity of some medicinal and aromatic plant extracts, Food Chem. 85 (2) (2004) 231–237.
[26] Y. Hacham, I. Matityahu, R. Amir, Transgenic tobacco plants having a higher level of methionine are more sensitive to oxidative stress, Physiol. Plant. 160 (3) (2017) 242–252.
[27] J.B.W. Hammond, M.M. Burrell, N.J. Kruger, Effect of low temperature on the activity of phosphofructokinase from potato tubers, Planta 180 (4) (1990) 613–616.
[28] P. Kainulainen, J.K. Holopainen, J. Oksanen, Effects of SO2 on the concentrations of carbohydrates and secondary compounds in Scots pine (Pinus sylvestris L.) and Norway spruce (Picea abies (L.) Karst.) seedlings, New Phytol. 130 (2) (1995) 231–238.
[29] A.R. Bilia, M.C. Bergonzi, S. Gallori, G. Mazzi, F.F. Vincieri, Stability of the constituents of Calendula, milk-thistle and passionflower tinctures by LC-DAD and LC-MS, J. Pharm. Biomed. Anal. 30 (3) (2002) 613–624.
[30] C. Kapcum, J. Uriyapongson, Effects of storage conditions on phytochemical and stability of purple corn cob extract powder, Food Sci. Technol. (2018).

Supplementary information to the manuscript

Lemna minor Studies under Various Storage Periods using Polarity-Extended Extraction and Metabolite Non-Target Screening Analysis

Rofida Wahman[1], Johanna Graßmann[1], Andrés Sauvêtre[2], Peter Schröder[2], and Thomas Letzel[1,*]

[1] Chair of Urban Water Systems Engineering, Technical University of Munich, Am Coulombwall 3, 85748 Garching

[2] German Research Center for Environmental Health, Research Unit Comparative Microbiome Analysis, Helmholtz Centrum Munich, Ingolstädter Street 1, 85764 Neuherberg

***** Correspondence: T.Letzel@tum.de;

2. Materials and Methods

2.1. Reagents and chemicals

Acetonitrile LC-MS grade was purchased from VWR (Darmstadt, Germany). Ammonium acetate, gabapentin, monuron, chloridazon, carbetamide, metobromuron, sotalol, quinoxyfen, and metconazol, were obtained from Sigma, Darmstadt, Germany. Further, metformin was obtained from Fluka, Buchs, Switzerland. Furthermore, chlorbromuron and diazinon were obtained from Dr. Ehrenstorfer, Augsburg, Germany

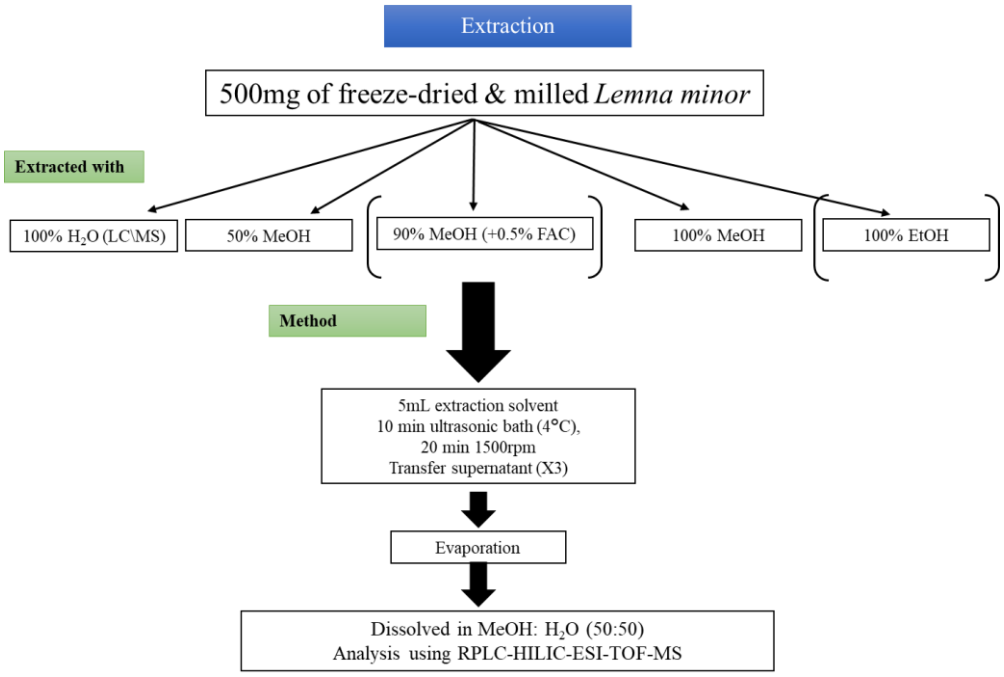


Fig. S1. The extraction method of *Lemna minor* Kaufmann *et al*., [33, 34]. Extraction solvents in brackets were used for special experiments. However, other solvents were used for the total study.

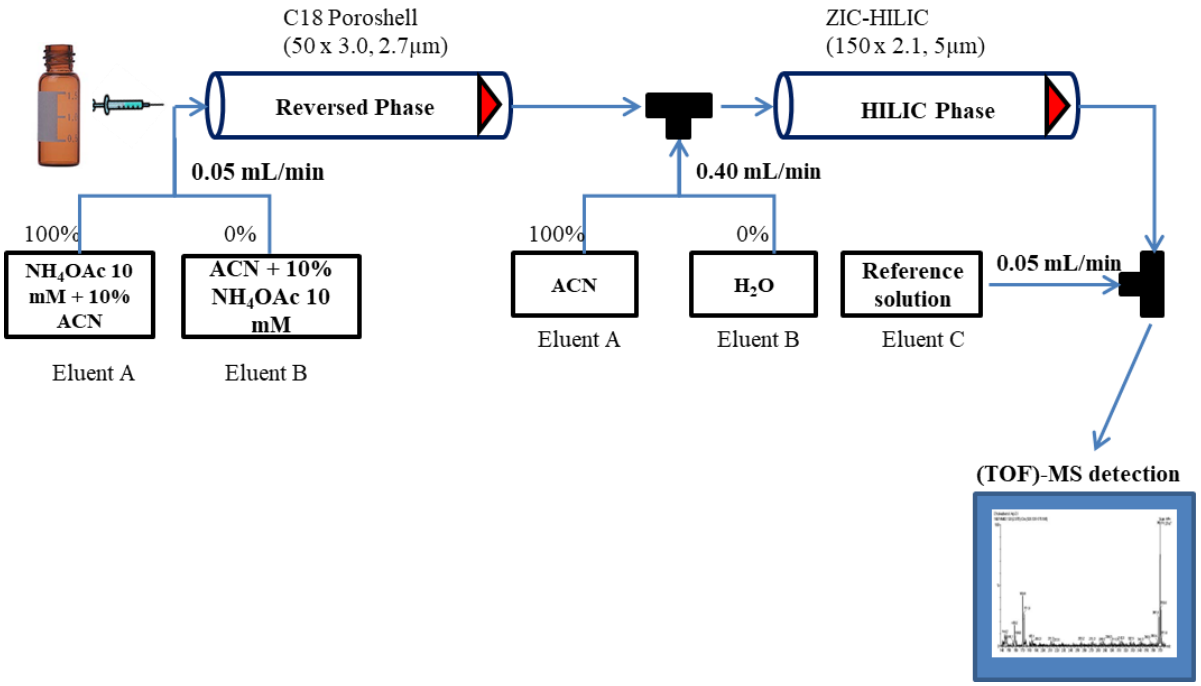


Fig. S2. RPLC-HILIC-TOF-MS scheme with two binary pumps and isocratic pump to deliver an electrospray ionization enhancing solvent flow and mass correction standards (Greco *et al*, J. Sep. Sci., 36, 1379-1388 (2013).

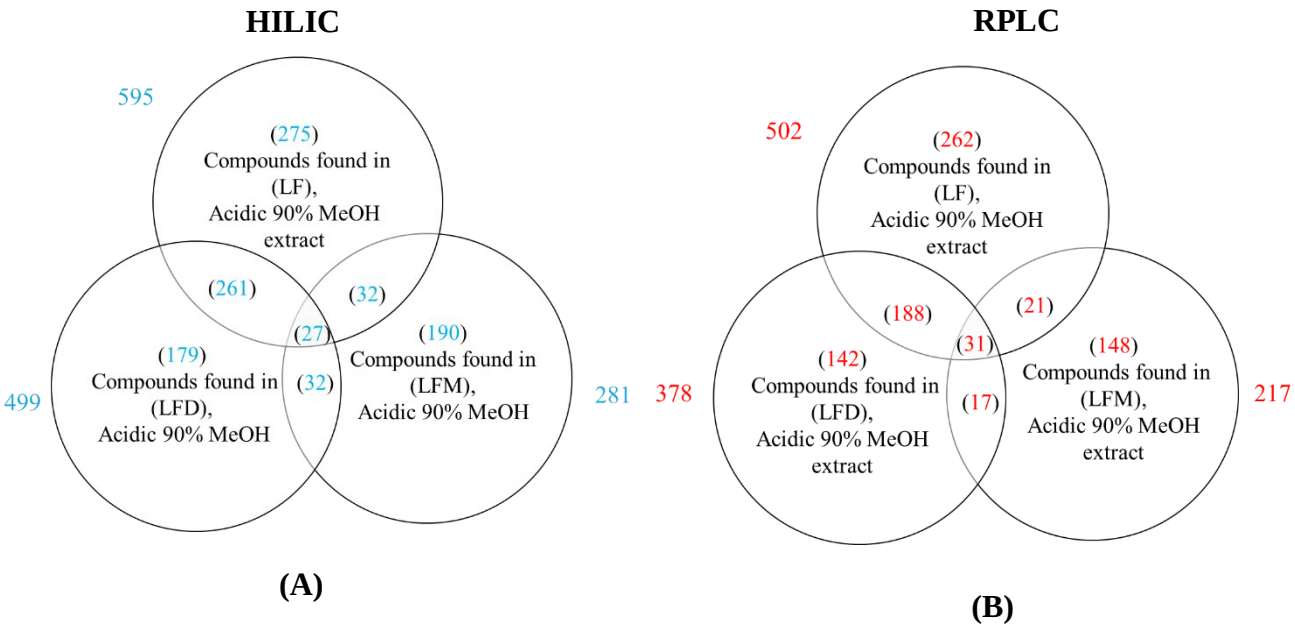


Fig. S3. Venn diagrams summarizing the number of shared and unique compounds in acidic 90% MeOH extracts of *Lemna minor* fresh sample (LF), frozen for days (LFD) and frozen for months (LFM). (A) Compounds measured in the HILIC column (with a RT ≤ 15) and are blue colored. (B) Compounds measured in the RPLC column (with a RT > 15) and are red-colored. The number of the total compounds found outside each circle, unique compounds of the extracts in the inside and the shared compounds.

Table (S1). Mobile phase condition of RPLC-HILIC-TOF-MS.

Binary Pump 1				**Binary Pump 1**			
Time (min)	**Flow rate (mL/min)**	**A%**	**B%**	**Time (min)**	**Flow rate (mL/min)**	**C%**	**D%**
0	0.05	100	0	0	0.4	100	0
7	0.05	100	0	6	0.4	100	0
12	0.05	50	50	13	0.4	60	40
13	0.1	50	50	32	0.4	60	40
22	0.1	0	100	33	0.8	100	0
32	0.1	0	100	53	0.8	100	0
33	0.1	100	0	54	0.4	100	0
53	0.1	100	0	58	0.4	100	0
54	0.05	100	0				
58	0.05	100	0				

Table (S2). The standards mixture used to determine the transition between RPLC and HILIC columns. Monoisotopic mass in the literature (L), means monoisotopic in different injection and the variation between them, also, mean RT of the different injections, standard deviation (SD), and relative standard deviation (RSD) were listed.

Name	Mono isotopic Mass (Da) (L)	Mean Mono isotopic Mass (Da) (M)	Δppm	Mean RT (min)	SD	%RSD
Metformin	129.1014	129.1020	-4.45	14.87	0.08	0.51
Gabapentin	171.1259	171.1256	2.05	9.46	0.07	0.71
Monuron	198.0560	198.0552	4.04	26.46	0.04	0.15
Chloridazon	221.0356	221.0353	1.36	24.51	0.03	0.12
Carbetamid	236.1161	236.1164	-1.38	25.97	0.05	0.21
Metobromuron	258.0004	258.0000	1.65	29.94	0.04	0.13
Sotalol	272.1195	272.1196	-0.18	11.16	0.03	0.24
Chlorbromuron	291.9615	291.9603	4.28	29.94	0.03	0.11
Diazinon	304.1010	304.1011	-0.44	35.00	0.03	0.07
Quinoxyfen	306.9967	306.9984	-5.46	34.77	0.01	0.02
Metconazol	319.1451	319.1449	0.63	28.39	0.08	0.28

References

[31] X.Y. Xie, R. Wang, Y.P. Shi, [Chemical constituents from rizomes of Homalomena occulta], Zhongguo Zhong yao za zhi = Zhongguo zhongyao zazhi = China journal of Chinese materia medica 38(14) (2013) 2325-7.
[32] R. Chakrabarti, W.D. Clark, J.G. Sharma, R.K. Goswami, A.K. Shrivastav, D.R. Tocher, Mass Production of Lemna minor and Its Amino Acid and Fatty Acid Profiles, Front Chem 6 (2018) 479-479.
[33] M. Chua, T.C. Baldwin, T.J. Hocking, K. Chan, Traditional uses and potential health benefits of Amorphophallus konjac K. Koch ex N.E.Br, J Ethnopharmacol 128(2) (2010) 268-78.
[34] Z.H. Zhang, Z. Dai, X.R. Hu, R.C. Lin, [Isolation and structure elucidation of chemical constituents from Pinellia ternata], Zhong yao cai = Zhongyaocai = Journal of Chinese medicinal materials 36(10) (2013) 1620-2.
[35] A.A. Watson, G.W.J. Fleet, N. Asano, R.J. Molyneux, R.J. Nash, Polyhydroxylated alkaloids — natural occurrence and therapeutic applications, Phytochemistry 56(3) (2001) 265-295.
[36] F. Ferreres, R.F. Goncalves, A. Gil-Izquierdo, P. Valentao, A.M. Silva, J.B. Silva, D. Santos, P.B. Andrade, Further knowledge on the phenolic profile of Colocasia esculenta (L.) Shott, J Agric Food Chem 60(28) (2012) 7005-15.
[37] R. Kandeler, Entwicklungsphysiologie der Pflanzen, 2019.
[38] A. Le Moullec, O.J. Juvik, T. Fossen, First identification of natural products from the African medicinal plant Zamioculcas zamiifolia - A drought resistant survivor through millions of years, Fitoterapia 106 (2015) 280-5.
[39] J.B. Harborne, T.J. Mabry, H. Mabry The Flavonoids, 1975.
[40] J.W. McClure, R.E. Alston, A CHEMOTAXONOMIC STUDY OF LEMNACEAE, American Journal of Botany 53(9) (1966) 849-860.
[41] K.-J. Appenroth, K.S. Sree, M. Bog, J. Ecker, C. Seeliger, V. Böhm, S. Lorkowski, K. Sommer, W. Vetter, K. Tolzin-Banasch, R. Kirmse, M. Leiterer, C. Dawczynski, G. Liebisch, G. Jahreis, Nutritional Value of the Duckweed Species of the Genus Wolffia (Lemnaceae) as Human Food, Front Chem 6 (2018) 483-483.
[42] A. Champagne, G. Hilbert, L. Legendre, V. Lebot, Diversity of anthocyanins and other phenolic compounds among tropical root crops from Vanuatu, South Pacific, Journal of Food Composition and Analysis 24(3) (2011) 315-325.
[43] D.A. Hart, P.K. Kindel, A novel reaction involved in the degradation of apiogalacturonans from Lemna minor and the isolation of apibiose as a product, Biochemistry 9(10) (1970) 2190-6.
[44] B.N. Smith, B.J. Meeuse, Production of volatile amines and skatole at anthesis in some arum lily species, Plant physiology 41(2) (1966) 343-347.
[45] V.-V. Karin Maria, C-Glycosylflavones as an Accumulation Tendency: A Critical Review, Botanical Review 51(1) (1985) 1-52.
[46] I. Muñoz-Cuervo, R. Malapa, S. Michalet, V. Lebot, L. Legendre, Secondary metabolite diversity in taro, #Colocasia esculenta# (L.) Schott, corms, Journal of Food Composition and Analysis 52 (2016) 24-32.
[47] M.A. Muhit, M. Izumikawa, K. Umehara, H. Noguchi, Phenolic constituents of the Bangladeshi medicinal plant Pothos scandens and their anti-estrogenic, hyaluronidase inhibition, and histamine release inhibitory activities, Phytochemistry 121 (2016) 30-37.
[48] C.A. Williams, J.B. Harborne, S.J. Mayo, Anthocyanin pigments and leaf flavonoids in the family araceae, Phytochemistry 20(2) (1981) 217-234.

metabolites

MDPI

Article

Untargeted Metabolomics Studies on Drug-Incubated *Phragmites australis* Profiles

Rofida Wahman [1], Andres Sauvêtre [2,†], Peter Schröder [2], Stefan Moser [3] and Thomas Letzel [1,4,*]

[1] Chair of Urban Water Systems Engineering, Technical University of Munich, Am Coulombwall 3, 85748 Garching, Germany; rofida.wahman@tum.de
[2] German Research Center for Environmental Health, Research Unit Comparative Microbiome Analysis, Helmholtz Centrum Munich, Ingolstadt Street 1, 85764 Neuherberg, Germany; andre.sauvetre@umontpellier.fr (A.S.); peter.schroeder@helmholtz-muenchen.de (P.S.)
[3] Stefan Moser Process Optimization, Weberweg 3, D-83131 Nußdorf am Inn, Germany; stefan_moser@web.de
[4] Analytisches Forschungsinstitut für Non-Target Screening GmbH (AFIN-TS GmbH), Am Mittleren Moos 48, D-86167 Augsburg, Germany
* Correspondence: T.letzel@tum.de
† Current address: HydroSciences Montpellier, UMR 5569, Faculté de Pharmacie, University of Montpellier, Avenue Charles Flahault 15, 34000 Montpellier, France.

Abstract: Plants produce a huge number of functionally and chemically different natural products that play an important role in linking the plant with the adjacent environment. Plants can also absorb and transform external organic compounds (xenobiotics). Currently there are only a few studies concerning the effects of xenobiotics and their transformation products on plant metabolites using a mass spectrometric untargeted screening strategy. This study was designed to investigate the changes of the *Phragmites australis* metabolome following/after diclofenac or carbamazepine incubation, using a serial coupling of reversed-phase liquid chromatography (RPLC) and hydrophilic interaction liquid chromatography (HILIC) combined with accurate high-resolution time-of-flight mass spectrometer (TOF-MS). An untargeted screening strategy of metabolic fingerprints was developed to purposefully compare samples from differently treated *P. australis* plants, revealing that *P. australis* responded to each drug differently. When solvents with significantly different polarities were used, the metabolic profiles of *P. australis* were found to change significantly. For instance, the production of polyphenols (such as quercetin) in the plant increased after diclofenac incubation. Moreover, the pathway of unsaturated organic acids became more prominent, eventually as a reaction to protect the cells against reactive oxygen species (ROS). Hence, *P. australis* exhibited an adaptive mechanism to cope with each drug. Consequently, the untargeted screening approach is essential for understanding the complex response of plants to xenobiotics.

Keywords: *P. australis* metabolic profile; untargeted metabolomics; diclofenac; carbamazepine; orthogonal partial least square-discriminant analysis (OPLS-DA)

Citation: Wahman, R.; Sauvêtre, A.; Schröder, P.; Moser, S.; Letzel, T. Untargeted Metabolomics Studies on Drug-Incubated *Phragmites australis* Profiles. *Metabolites* **2021**, *11*, 2. https://dx.doi.org/10.3390/metabo11010002

Received: 12 November 2020
Accepted: 21 December 2020
Published: 22 December 2020

1. Introduction

Throughout human history, plants have been the most important source for pioneering medicines, flavors and industrial materials. Plants produce up to 200,000 natural products with a vast chemical diversity using a wide range of enzymes and substrates [1]. Among them, primary metabolites are relatively few and are defined as fundamental to plant physiology. In contrast, secondary metabolites (specialized metabolites) are essentially diverse. The high diversity of specialized metabolites leads to functional variation. Specialized metabolites play important roles in the interactions between the plant and its environment since they are involved in protection against environmental stress, competition or pollinator attraction and some are involved in vegetative or floral development. Furthermore, they play a vital role in the conception of environmental signals and their translation in

biography traits. Therefore, they may play major roles in allowing the organism to sustain environmental constraints.

Among many other uses, plants have been used to restore environmental quality via adsorption and accumulation of organic xenobiotics (such as pesticides, dyes, drugs, etc.) inside their tissues. In the process of phytoremediation, plants can degrade these external compounds through a specific enzymatic scaffold consisting of three phases [2,3]. In this process, they utilize enzymes like cytochromes P450 or peroxidases to activate the xenobiotics in phase I and conjugating enzymes as glucosyl-transferases, malonyl-transferases or glutathione S-transferases in phase II for detoxification [4,5].

Diclofenac (DCF) and carbamazepine (CBZ) belong to the problematic xenobiotics in European wastewater treatment plants since they can harm the ecotoxicological equilibrium. They have not only been detected in wastewater but also in biosolids and wastewater effluents [6–8]. Recently, the maximal measured concentrations of DCF and CBZ in municipal wastewater has ranged between 440 and 7100 $ng \cdot L^{-1}$ and 1075–6300 $ng \cdot L^{-1}$ of, respectively [9,10]. DCF is one of the most commonly used non-steroidal anti-inflammatory drugs, while CBZ has been used widely as an antiepileptic and mood stabilizer since the 1970s. Both DCF and CBZ are taken up and translocated into the aerial parts of plants, where they can be accumulated or metabolized into more or less toxic products [7,11].

One of these species is *P. australis*, known as "common reed". *P. australis* belongs to the Poaceae family and is an invasive plant that spreads worldwide [12]. It has been used for a long time in wetlands to remove pollutants, reduce nitrogen loads and provide oxygen to the rhizosphere [5,13]. Furthermore, it is used as a resource for traditional crafts and fodder. In some regions like Northern China, it is grown as a crop and its leaves are used in the treatment of bronchitis and cholera [12].

The chemical diversity (such as polarity, stability, reactivity or ionization) of plant molecules prohibits investigation of the metabolites´ full picture and understanding changes in their occurrence. The analysis purpose and chemical nature of plant metabolites determine the analytical technique. Also, the untargeted concept demands an analytical device, which can separate different classes of metabolites with a wide range of polarities. For untargeted metabolite analysis, several analytical techniques are available. These approaches apply liquid chromatography-mass spectrometry (LC-MS) [14], gas chromatography-mass spectrometry (GC-MS) [15] or nuclear magnetic resonance (NMR) [16] to analyze a large number of different chemical metabolites classes within one single analysis. NMR has analytical reproducibility and a non-destructive nature; however, it has relatively low sensitivity compared to MS. Besides, a mass spectrometric untargeted screening strategy can identify a large number of molecules independently to provide a preferably holistic picture of the plant's metabolome. Also, the thermal stability of the stationary phase, metabolites and their derivatives, which might introduce artifacts, limit the metabolome coverage derived by GC-MS. Thus, the usage of LC-MS has expanded rapidly over the past ten years in untargeted metabolomics analysis [17].

The serial coupling of reversed-phase liquid chromatography (RPLC) and hydrophilic interaction liquid chromatography (HILIC) is often used to separate compounds with differing polarities in a single run [18]. Further, connecting the serial RPLC-HILIC coupling with an accurate high-resolution time-of-flight mass spectrometer (HRMS) provides the detection of a wide range of metabolites.

Recently, the awareness of untargeted metabolomics analysis has increased due to its capabilities in the assessment of xenobiotics exposure/specific biomarkers and the risk of contaminants to living organisms. For example, the metabolic fingerprints of *Plantago lanceolata* showed various chemical changes as a response to different stresses [19].

Metabolic fingerprinting experiments aim to determine relative differences between two or more systems elucidating a biological relationship. Therefore, statistical strategies are typically used in a chemometrics style. Univariate and multivariate statistics can be used as standard approaches to extract relevant information from complex datasets [20–23].

This study aims to investigate the effect of DCF, CBZ and their transformation products on *P. australis* metabolites using mass spectrometric untargeted screening analysis. By studying the fingerprint of *P. australis* leaves, rhizomes and roots with/without incubation at two different concentrations, 10 and 100 µM and 10 and 50 µM of DCF, as well as CBZ, respectively. Also, a statistical workflow was used to discriminate against the changes in *P. australis* metabolites and detect the differentiating metabolic profiles.

2. Results

Diclofenac and carbamazepine are considered problematic environmental pollutants because they can be found in surface waters (mainly in wastewater treatment plant effluents) in high concentrations. Several plant species have been found capable of absorbing and detoxifying them. Despite this, the effects of diclofenac and carbamazepine and their transformation products on plant metabolic profiles need more investigation. Consequently, *P. australis* was incubated with 10 and 100 µM DCF and with 10 and 50 µM CBZ, respectively. Each plant (i.e., leaves, rhizomes and roots, respectively) was extracted with four different solvents. The different samples were injected into a robust and reproducible serial coupling RPLC-HILIC in hyphenation with electron spray ionization- time of flight- mass spectrometer (ESI-TOF-MS). Both robustness and reproducibility of the analytical system was checked and proven with a mixture of 13 reference standards (Tables S1 and S2). The statistical evaluation discriminated between the different extraction solvents, plant parts and incubation. Moreover, the statistical analysis was enabled to assign the changes in *P. australis* metabolic profiles and the differentiating metabolic profiles (DMF) between the control and incubated samples. Furthermore, different DCF and CBZ transformation products could be identified or suspected. Thus, changes in *P. australis* biosynthetic pathways could be predicted.

2.1. Metabolic Profiling Elucidation in Phragmites australis Extracts with RPLC-HILIC-ESI-TOF-MS

A mixture of 13 reference standards was injected at the beginning/end and fixed intervals during the experimental sets (i.e., after each extract batch). The results of the standard mixture proved the accuracy, repeatability and reproducibility of the analytical system. Mass and RT of the standards during the experiment showed an acceptable deviation of less than 8 ppm (with a ToF system from the year 2012) and % RSD less than 2 except for gabapentin, carbetamid and sotalol. For more information, readers are referred to Wahman et al., 2019 [18].

The four different extracts 100% MeOH, acidic 90% MeOH, 50% MeOH and 100% H_2O of *P. australis* leaf, rhizome and root were analyzed similarly with RPLC-HILIC-ESI-TOF-MS coupling as previously described [24–26]. The obtained mass spectrometric total ion chromatograms (TICs) were interpreted to extract the feature (extracted masses, RT and signal abundance) according to the parameters mentioned in Section 4.6.1. Background signals (i.e., all the peaks detected in the corresponding blank) were deleted to avoid false positives. The Retention Time (RT)/Mass plots of the different background were presented in (Figure S1). Lastly, the features found in the triplicate injections were considered for further analysis. Exemplarily, Figure 1 represents (RT)/Mass plots for 100% MeOH extracts of *P. australis* leaf, rhizome and root in positive ion mode. The highly polar to polar compounds eluted at RT < 15 min, with logD values below zero (HILIC part). The nonpolar compounds were eluted at RT > 15 min, with logD values above zero (RPLC part). The differences in chromatographic fingerprints reflected the variability in metabolite profiles (and composition) in the leaf, rhizome and root samples. Detailed information and a description of data analysis have already been discussed in previous publications [18,26].

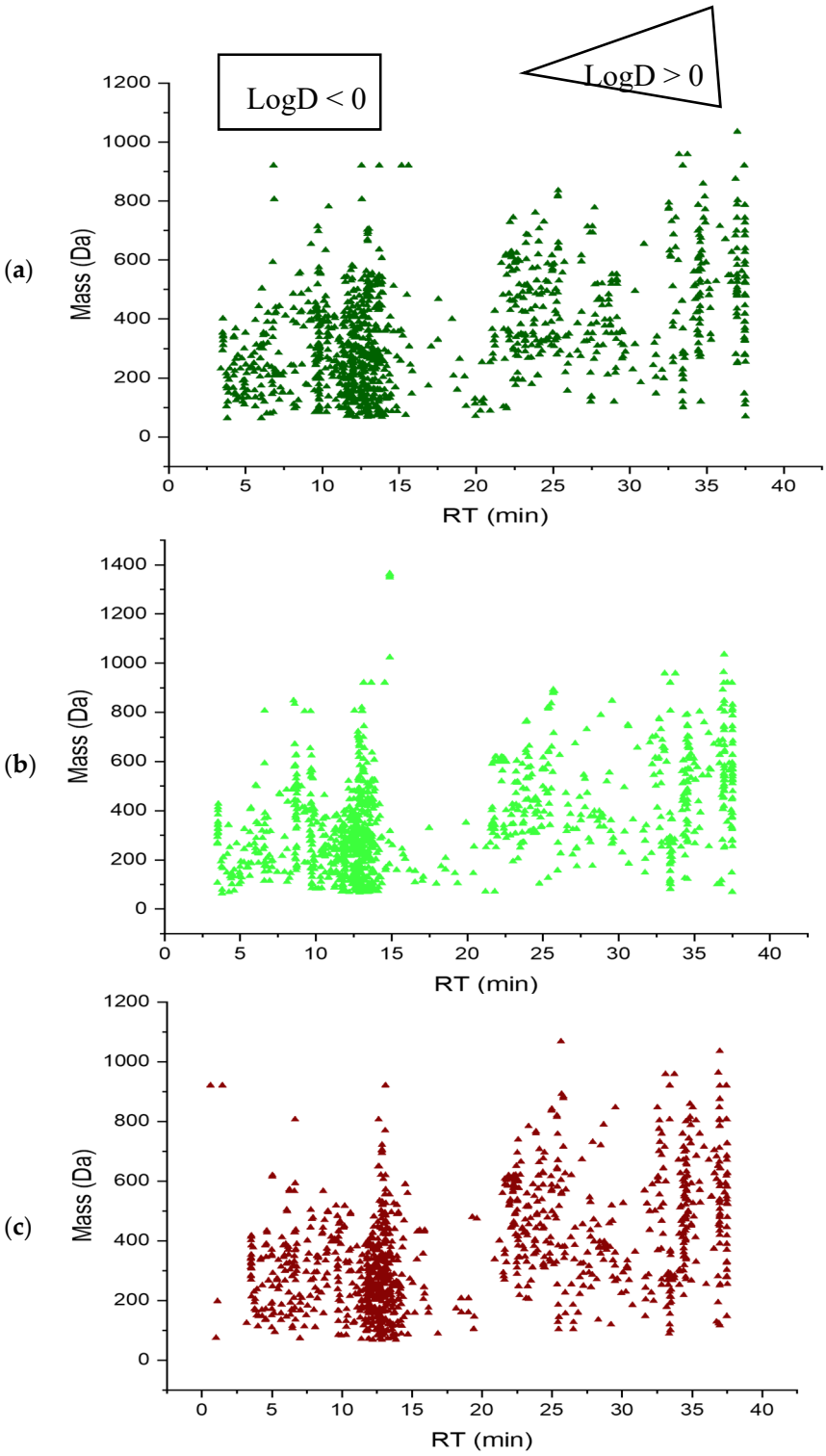
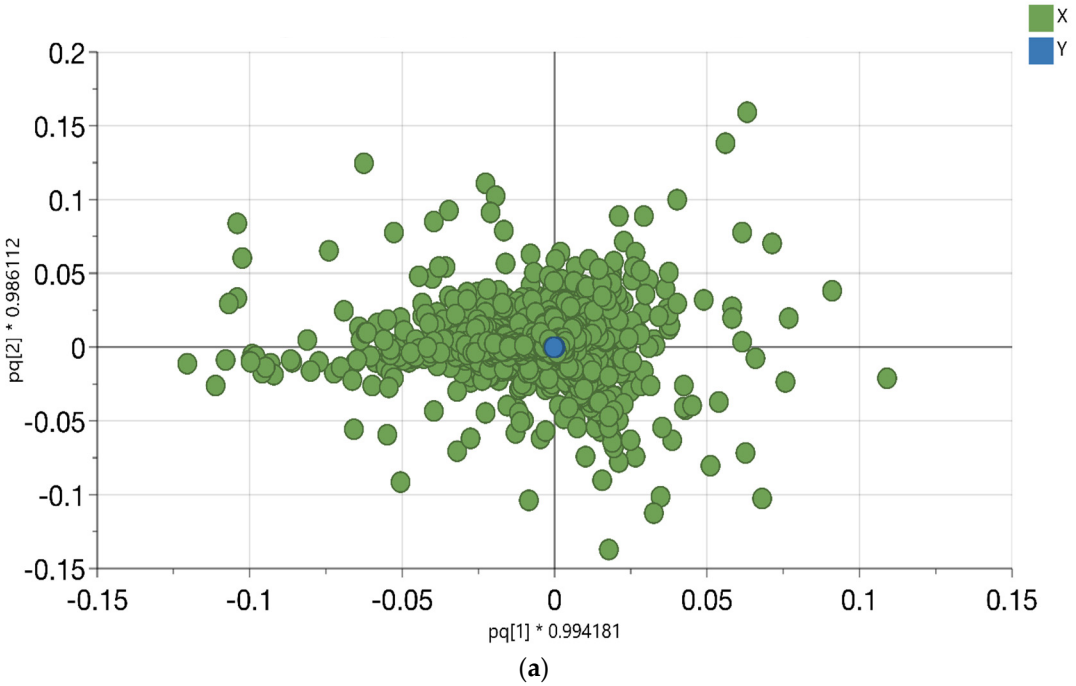


Figure 1. Retention time (RT)/Mass plot of *Phragmites australis* 100% Methanol extracts analyzed by RPLC-HILIC-ESI-TOF-MS in positive electrospray ionization mode. (**a**) Leaf; (**b**) Rhizome; (**c**) Root, which showed the features' separation according to their polarity and detected according to their *m/z*.

2.2. Different Extracts of Phragmites australis's Metabolic Fingerprints

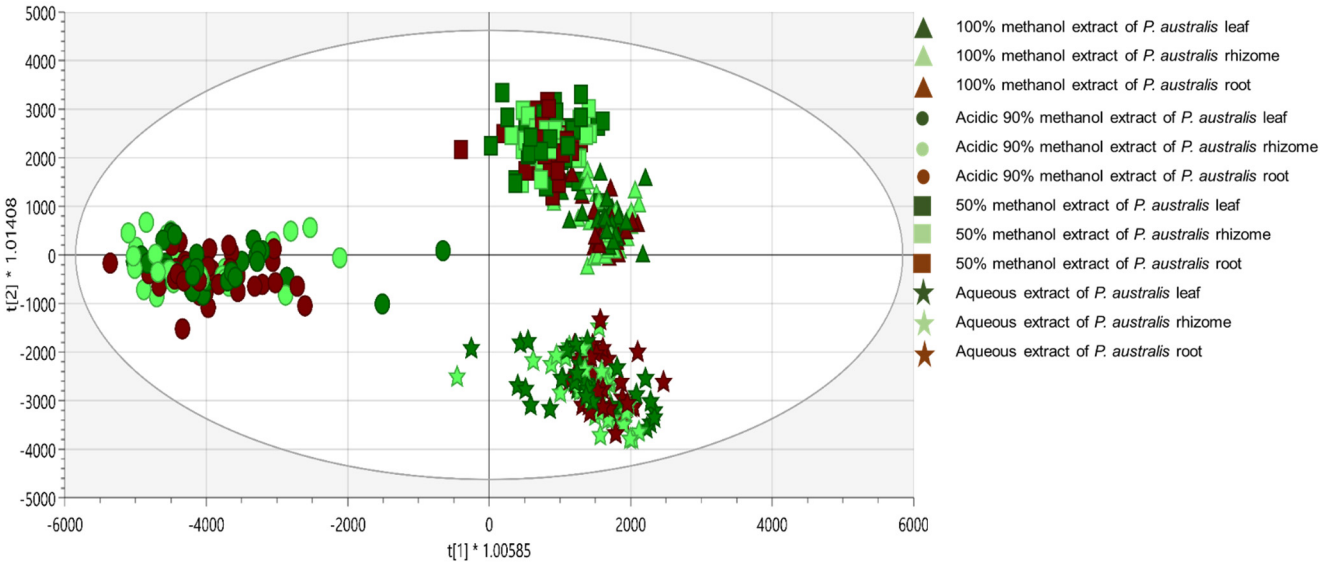
Untargeted metabolomics analysis demands minimal pretreatment methods to allow the detection of almost all the sample metabolites. Different leaf, rhizome and root samples were extracted with the described four solvents (i.e., 100% MeOH, acidic 90% MeOH, 50% MeOH and 100% H_2O), individually, which ensured the extraction of a wide range

of metabolites from nonpolar to highly polar metabolites relative to the solvents, which were used in the extraction process. All the extracts were analyzed along a metabolites' fingerprint strategy. In orthogonal partial least square-discriminant analysis (OPLS-DA), the variables were *P. australis* metabolites, which were plotted in the loading score plot (as in Figure 2a). The OPLS-DA model described the variables according to the solvents of the extraction class. Of the data variations, 18.3% (R^2X(cum)) are responsible to distinguish between the classes that were previously established based on solvents. The rest of the variation (orthogonal components) describes the variation within the solvent classes. The high value of those parameters indicates that the OPLS-DA model had a good classification and prediction efficiency to distinguish between different extracts, even though it described one variation (i.e., solvent type). This is owing to the accuracy of variables, which were separated with a robust and reproducible LC-system. In (Figure 2b), the samples are distributed according to t1 (predicative component) and t2 (orthogonal component). The predictive component (t1) separated the samples into two groups, the first group (negative side) contained the acidic 90% MeOH extracts and the second group (positive side) consisted of 100% MeOH, 50% MeOH and 100% H_2O extracts (Figure 2b). Moreover, the orthogonal component (t2) described the differences within the group. Consequently, it separated the second group into 100%MeOH and 50%MeOH in the positive part and the 100% H_2O extracts in the negative part (Figure 2b). Further, in (Figure 2c), good separation was reached according to the cumulative goodness of fit and the cross-validation parameters of each variable R^2 and Q^2, respectively. Therefore, the model had no risk of overfitting.

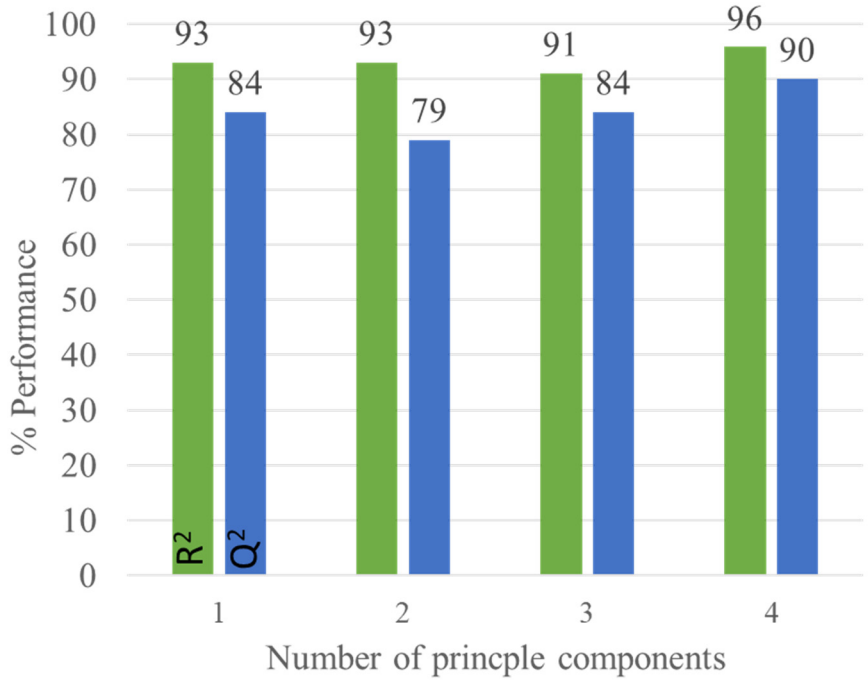


(**a**)

Figure 2. *Cont.*

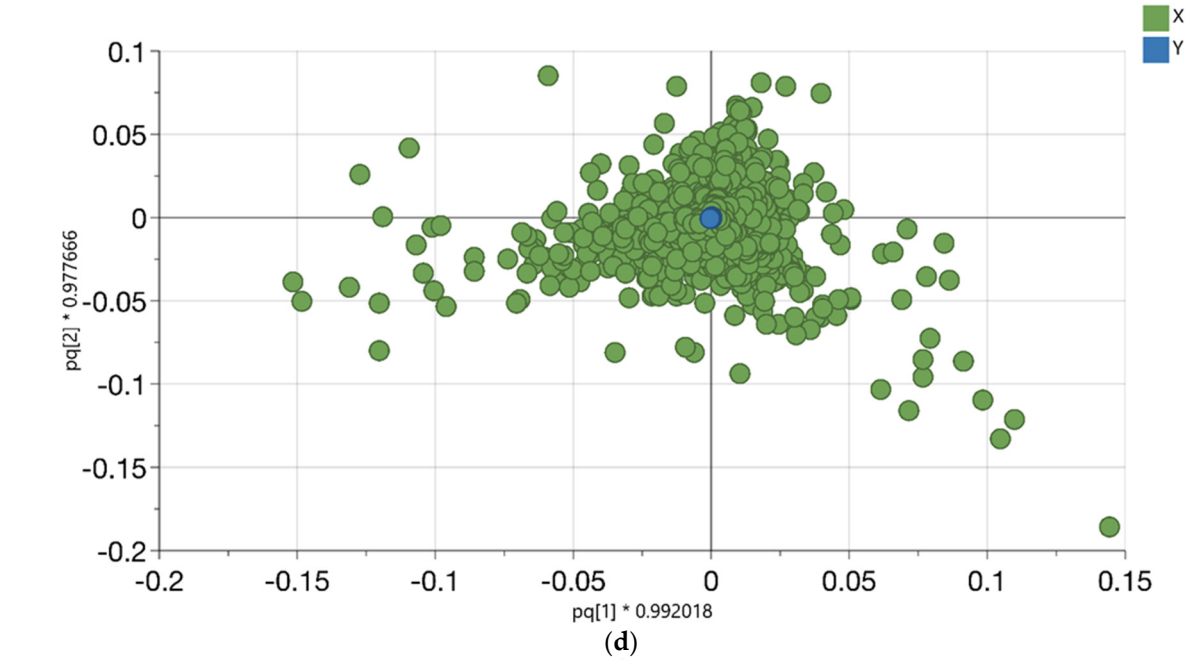


(**b**)

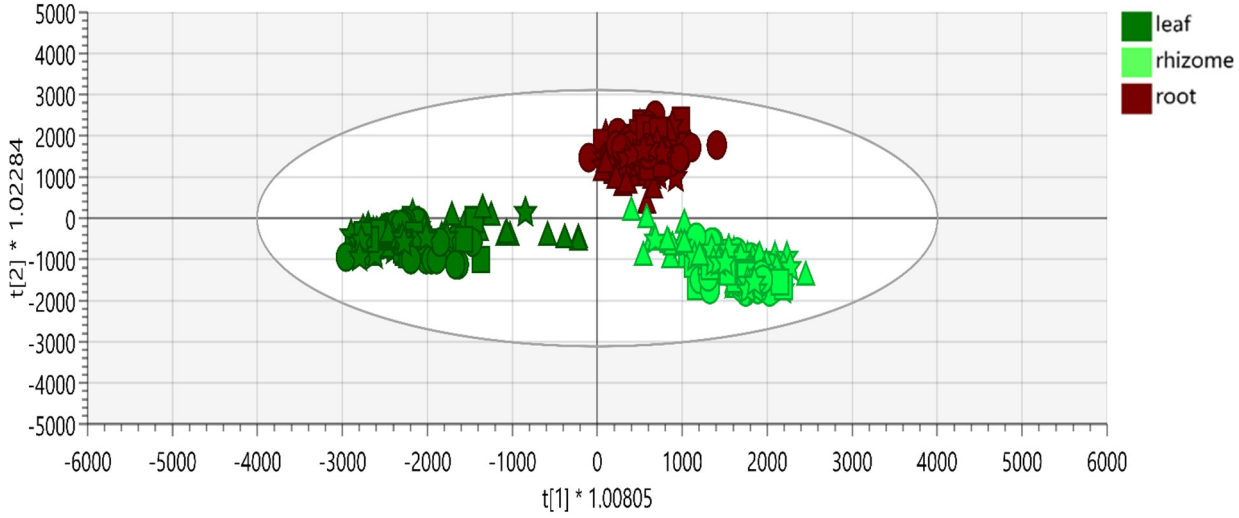


(**c**)

Figure 2. *Cont.*



(**d**)



(**e**)

Figure 2. (**a**) The OPLS-DA score plot of different *Phragmites australis* extracts with a confidence limit of 95% discriminating according to the solvent used in the extraction. The variables were plotted according to the first principal component (t1) and the orthogonal component (t2). The triangles represent 100% methanol extracts, circles represent acidic 90% methanol extracts, the squares represent 50% methanol extracts and the stars represent aqueous extracts, respectively. The green color represents leaf samples, the light green color represents rhizome samples and the brown color represents root samples, respectively. Each symbol represents one observation of *P. australis* leaf, rhizome and root plant part; (**b**) loading scatter plot for the selected principal components; (**c**) the Q^2/R^2 Overview plot displays the individual cumulative R^2 (green columns) and Q^2 (blue columns) for the goodness of fits and cross-validation parameters; (**d**) the loading scatter plot for the selected principal components displays the relation between the different *Phragmites australis* samples and the chosen metabolites; (**e**) the OPLS-DA score plot of the different parts of *Phragmites australis* with a confidence limit of 95%, discriminating according to the plant part.

2.3. Phragmites australis Leaf, Rhizome and Root Metabolic Fingerprints

In this section, the 11,442 variables were differentiated according to the plant part that they originated from. The variables were in the loading score plot (Figure 2d). This step was performed to improve the statistical significance of the dataset through additional

cross-validation. Also, the OPLS-DA analysis was conducted to investigate the structure of the data. Another 18% of the data was investigated through the variation between X (metabolites) and Y (plant parts) given by $R^2X(cum)$ and explained approximately 94% of the variations in the various samples ($R^2Y(cum)$). We found the predictive value of the model was ($Q^2(cum)$ = 74%), which was created by OPLS-DA. The cross-validation performance was confirmed by analysis of variance (ANOVA). OPLS-DA discriminates the different samples of *P. australis* leaf, rhizome and root regardless of incubation, with or without DCF or CBZ and regardless of the extraction solvent composition according to the plant part (Figure 2e).

The predictive component t1 differentiated between leaf extracts in one group (negative part) and the root and rhizomes extracts in the second group (positive part), as shown in (Figure 2e). However, the t2 (orthogonal component) differentiated rhizomes sample in the negative part and the roots in the positive one.

The quality of the OPLS-DA module was expressed by the cumulative value of the goodness of fit and the cross-validation for each value R^2 and Q^2, respectively, as shown in the Q^2/R^2 Overview plot (Figure S2A). The high values of the previous parameters indicated a good classification and prediction efficiency to distinguish between different plant parts.

2.4. Untargeted Metabolomics Analysis of Phragmites australis Incubated with DCF or CBZ

The untargeted metabolomics analysis of *P. australis* incubated with/without DCF or CBZ, respectively, was performed using OPLS-DA to assign the changes in its metabolic fingerprint. The *P. australis* metabolic fingerprints of different extraction solvents with various plant parts were investigated in Sections 2.2 and 2.3, respectively, to test the organization and reliability of the data. Then, the large data set was used to perform the untargeted analysis and assign a list of metabolites that determined the distance between different groups. Also, the metabolic markers of *P. australis* incubated with/without DCF or CBZ were plotted each by the OPLS-DA, which represented the variability in metabolic patterns due to the different incubation (Figure 3, that is, the loading plot in Figure 3a and the OPLS-DA score plot in 3B). OPLS-DA analysis showed the identified and unidentified metabolites, which distinguished the different clusters according to the characteristic change of control or incubated samples metabolite profiles. It was used to enhance the quality of pairwise classification analysis. The relative high goodness of fit indicated the good separation of different incubation groups of *P. australis* R^2 and Q^2, respectively, (Figure S2B). To facilitate the interpretation and visualization of OPLS-DA, the S-plot was drawn to illustrate the model's influence with accuracy in the search for differentiating metabolic profile (DMF). S-plot analysis represented the highest contributing signals for the control and incubation of *P. australis* with DCF or CBZ (Figure 4a,b, respectively). The DMF was extracted from S-plots, which were marked in red color. They were selected based on their contribution to the variation and correlation within the data set between the X-variables and the predictive component t1 (p (corr) vector). Hence, they were considered functionally in the combined form of a metabolic profile, which was distinguished between the control and the incubated *P. australis*. In Figure 4c,d, the contribution plots extend the data to better visualization and indicate regions responsible for sample clustering [22]. The contribution plot summarized the changing trends in the metabolites in the pairwise groups (i.e., control and incubated samples) through expressing the fold change of the metabolites between the control and incubated samples. As examples, succinic acid (DM_4), propane-1,2,3-triol (DM_6) and 2- hydroxypropanoic acid (DM_7) were detected as transformation products of DCF (DMs, see Table 1), thus they appeared significant in the positive part of the plot, (Figure 4c). Also, quercetin (compound 1 in Figure 4c,d) was found in elevated signal heights after *P. australis* exposure to DCF. Quercetin is a flavonoid that protects the plant against reactive oxygen species (ROS) during exposure to pharmaceuticals.

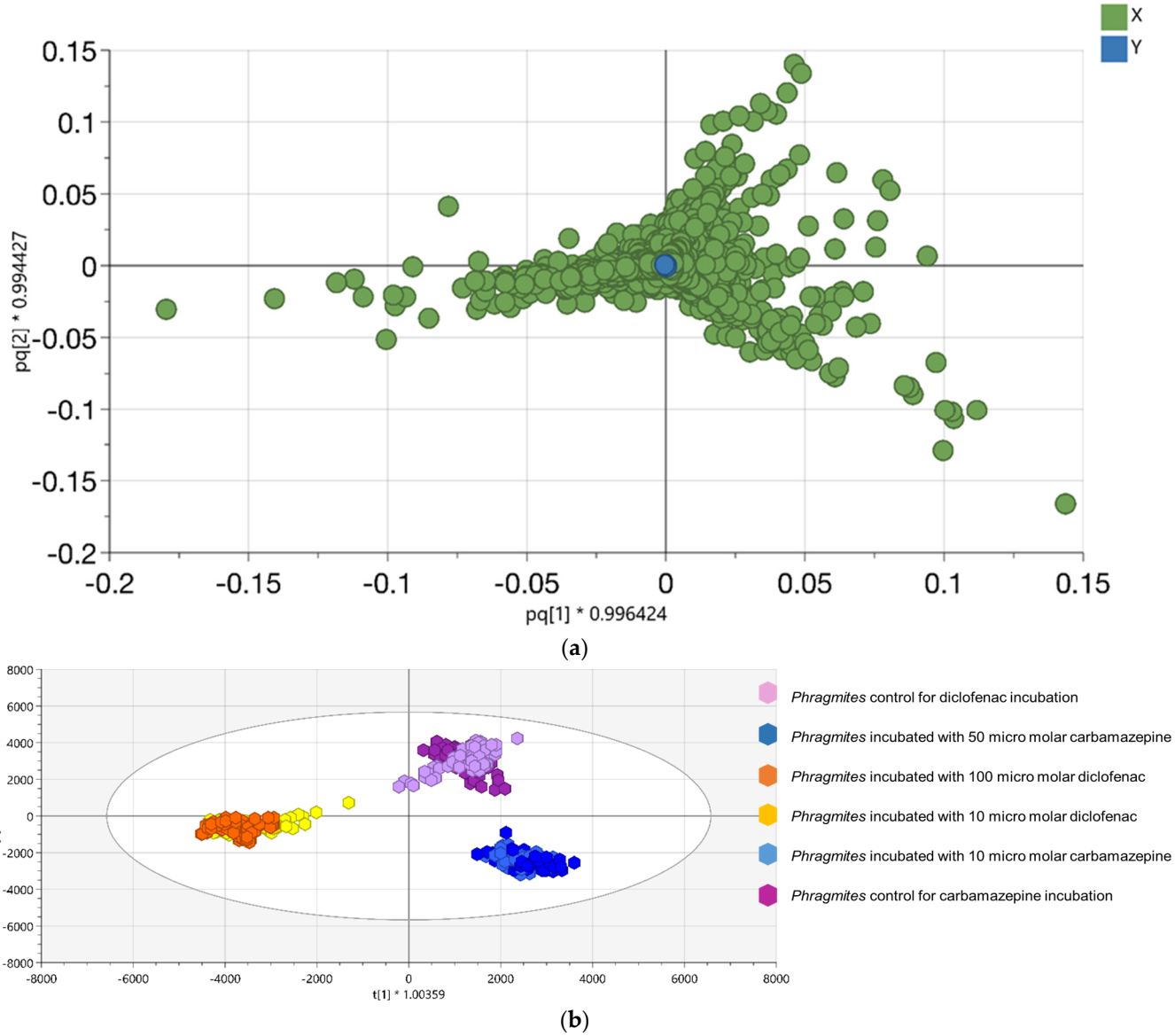
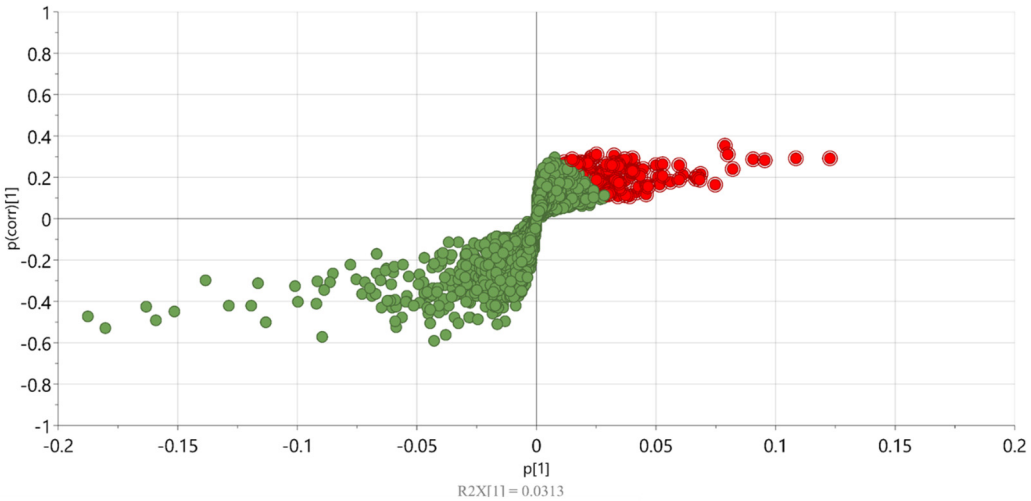
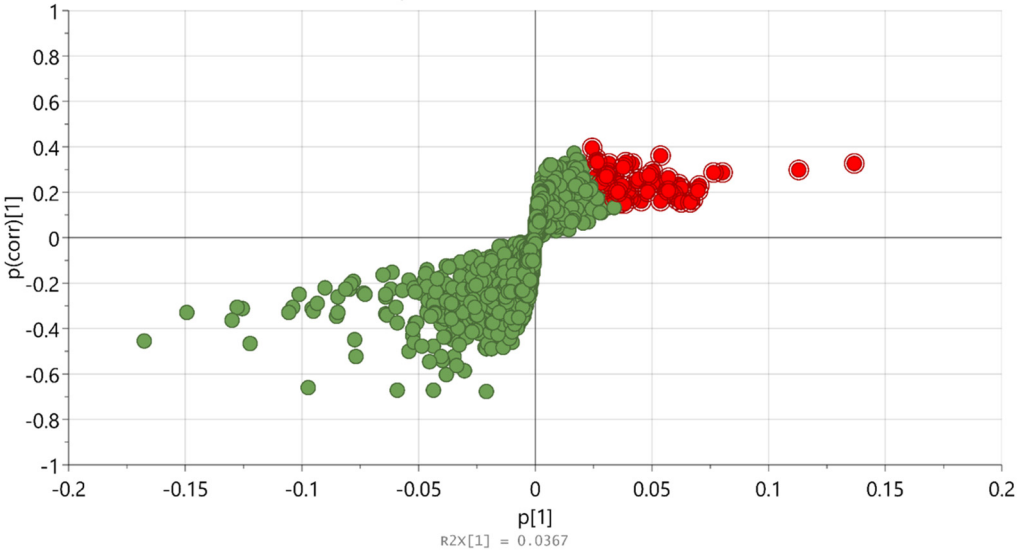


Figure 3. (**a**) The loading plot displays the relation between the different *Phragmites australis* samples and the chosen metabolite; (**b**) the OPLS-DA score plot of different *Phragmites australis* samples incubated with 10 and 50 µM carbamazepine, 10 and 100 µM diclofenac, individually. The confidence limit is 95%. For carbamazepine incubation the purple color represents the control group, the light blue represents a sample incubated with 10 µM carbamazepine and the blue color represents samples incubated with 50 µM carbamazepine. For diclofenac incubation the light purple color represents the control group, the yellow color represents samples incubated with 10 µM diclofenac and the orange color represents samples incubated with 100 µM diclofenac.

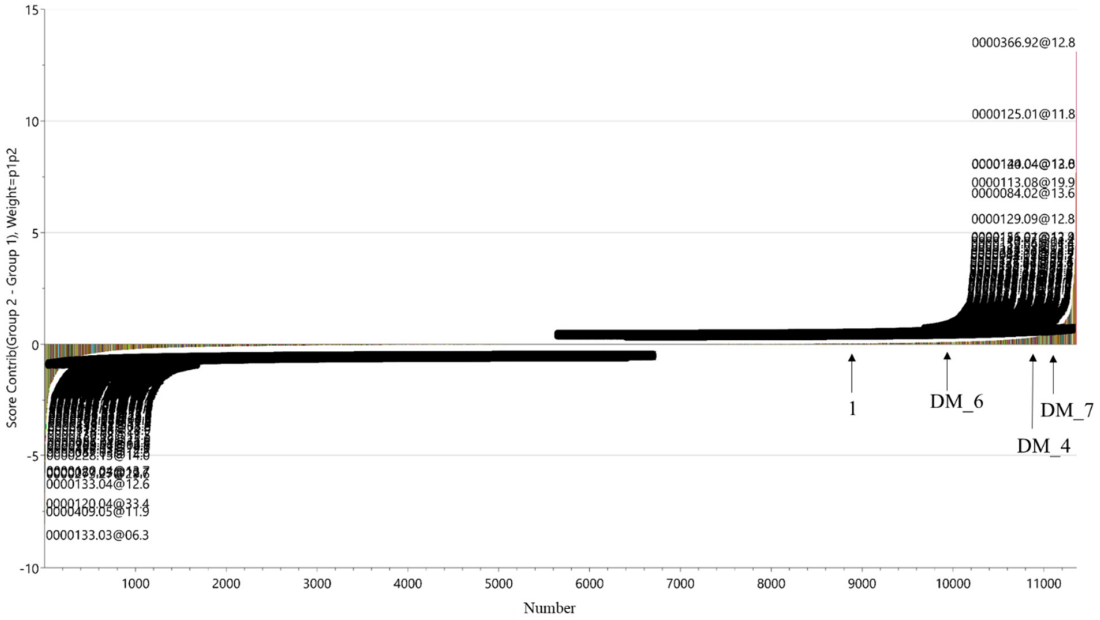


(**a**)

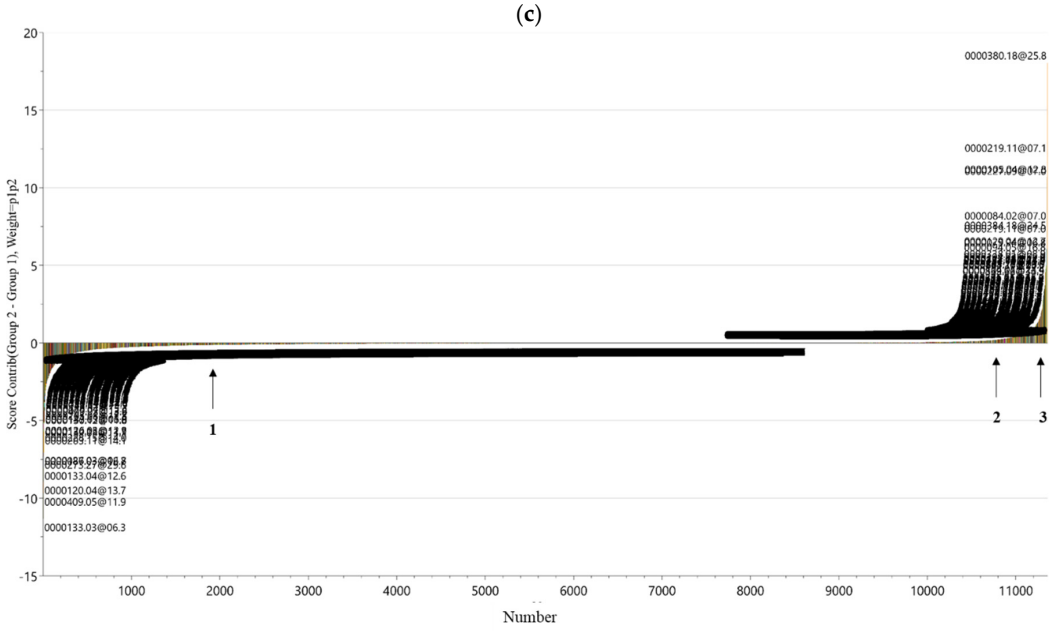


(**b**)

Figure 4. *Cont.*



(**c**)



(**d**)

Figure 4. (**a**) S-plot of *Phragmites australis* control and incubated with 10 and 100 µM diclofenac samples; (**b**) S-plot of *Phragmites australis* control and incubated with 10 and 50 µM carbamazepine samples. The S-plot provides the visualization of the loading components (green color) of OPLS-DA to enable the interpretation of the data. The red-labeled compounds represent the differentiating metabolic profile (DMF) of each incubation; (**c**) the contribution plot shows the up- and down-regulated compounds due to the incubation of *Phragmites australis* with 10 and 100 µM diclofenac, individually. Down-regulated compounds have negative values, while up-regulated compounds have positive values. (1) Quercetin, DM_6, DM_4 and DM_7 have up-regulated in *Phragmites australis* due to incubation with diclofenac at (8820, 0.0385581), (9968, 0101411), (10,817, 0.279365) and (10,985, 0.389), respectively; (**d**) the contribution plot of *Phragmites australis* with 10 and 50 µM carbamazepine, individually. Down-regulated compounds have negative values as compound 1 which was quercetin at (2033, −0.118706). Up-regulated compounds have positive values as compounds 2 and 3 which were 2,3-dihydro-2,3-dihydroxycarbamazepine at (10,921, 0.288274) and carbamazepine-10,11-epoxide at (11,270, 1.27941), respectively.

Table 1. List of diclofenac (DCF) transformation products detected in different *Phragmites australis* leaf, rhizome and roots incubated with 10 and 100 µM diclofenac samples with the monoisotopic mass in the literature (L), the mean monoisotopic mass of *Phragmites australis* (Ph), the variation between them, mean RT of *Phragmites australis* (Ph), standard deviation, relative standard deviation and LogD (pH = 7.4) were listed. The logD values were predicted from ChemAxon software (https://disco.chemaxon.com/apps/demos/logd/).

DCF Transformed Products	Name	Mono Isotopic Mass (Da) (L) (Rajab, Greco et al. 2013)	Mean Mono Isotopic Mass (Da) (Ph)	Δppm	Mean RT (Min) (Ph)	SD of RT (Min)	RSD	LogD (pH = 7.4)	Leaf	Rhizome	Root
DM_1	2-Hydroxypropanoic acid	152.0473	152.0475	−0.99	8.1	0.05	0.67	−1.86	√	√	√
DM_2	2-(Hydroxymethyl)benzene-1,4-diol	140.0473	140.0472	0.71	12.5	0.06	0.51	0.60	√	√	√
DM_3	2-Hydroxysuccinic acid	134.0215	134.0215	0.07	12.6	0.03	0.23	−6.81	√	√	√
DM_4	Succinic acid	118.0266	118.0271	−3.95	6.8	0.04	0.55	−1.99	√	√	√
DM_5	Fumaric acid	116.0101	116.01	0.49	12.6	0.05	0.42	−2.00	√	√	√
DM_6	Propane-1,2,3-triol	92.0473	92.04703	2.9	7.2	0.07	0.99	−1.84	√	×	×
DM_7	2-Hydroxypropanoic acid	90.0317	90.03207	−4.07	12.1	0.04	0.32	−1.00	√	×	×

In CBZ incubation, the contribution plot showed that in incubated samples 2,3-dihydro-2,3-dihydroxy-carbamazepine (compound **2** in Figure 4d) and carbamazepine-10,11-epoxide (compound **3** in Figure 4d) could be identified as metabolites (evaluated and proven with reference standards) and appeared with significant intensity. The carbamazepine-10,11-epoxide was statistically significant variable important (VIP) > 1 and $p < 0.05$. However, in this case quercetin (compound **1** in Figure 4d) intensity was found to decrease significantly.

2.5. Metabolism of Diclofenac in Phragmites australis

P. australis was exposed for 96 h to diclofenac in the two concentrations of 10 and 100 µM, respectively. Thereafter, the data was processed with MassHunter Workstation Software Profinder B.06.00 (Agilent Technologies, Waldbronn, Germany) to detect the DCF molecule and its metabolites (including background subtraction). Diclofenac and its hydroxylated metabolites were detected in the roots and rhizomes of *P. australis*. Diclofenac was identified at 295.0173 Da with 2 ppm deviation from the monoisotopic mass, which eluted at 26 min (logD (pH7) > 0). The EICs of diclofenac in extracts and standard are shown in Figure S3. Further, hydroxylated metabolites were suspected according to the previously mentioned criteria at 311.0114 Da with 0.6 ppm deviation from the monoisotopic mass, which eluted at 24 min (logD (pH7) > 0) [27]. Moreover, the analysis of different extracts of *P australis* samples based on the mass, RT and LogD (pH7) revealed seven proposed metabolites of diclofenac, as summarized in Table 1 and Figure S3. The comparison of mass spectra from leaf, rhizome and root resulted in evidence for five metabolites of diclofenac in all parts of the treated samples of *P. australis*. However, DM_6 and DM_7 were detected in leaf extracts (Table 1).

2.6. Metabolism of Carbamazepine in Phragmites australis

P. australis was exposed for 96 h to carbamazepine in two CBZ concentrations of 10 and 50 µM, respectively. Thereafter, the data was processed with MassHunter Workstation Software Profinder B.06.00 to detect the compound and its metabolites (including background subtraction). According to Sauvêtre and co-workers (2018), four different pathways for carbamazepine metabolism (PCM) were investigated in the plant [28].

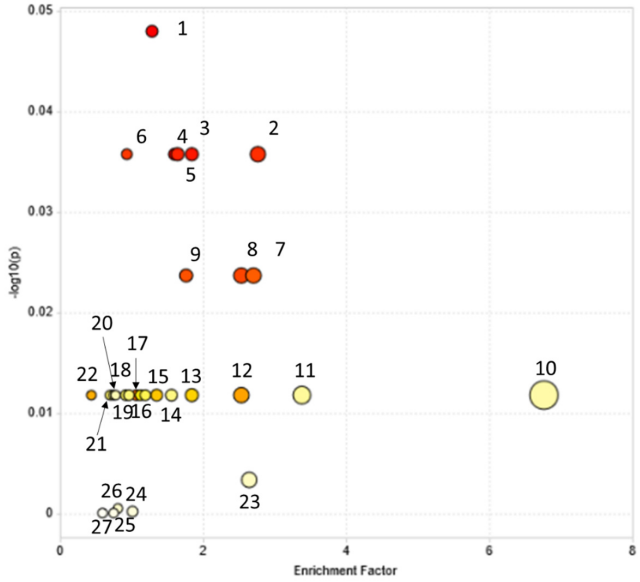
The different transformation products were identified in different *P. australis* samples using reference standards. The mean monoisotopic mass of standards and different samples and their absolute variation were tabulated, as well as the retention time in both standards and samples and the variation between them (Table S3, Figure S4). Carbamazepine-10,11-epoxide, 10,11-dihydro-10,11-dihydroxy-carbamazepine, 10,11-dihydro-10-hydroxy-carbamazepine, 9-acridine carboxaldehyde and 2,3-dihydro-2,3-dihydroxy-carbamazepine were identified in *P. australis* incubated samples. The mass deviation was less than 5 ppm and the deviation in the RT was less than 0.3 min. Carbamazepine-10,11-epoxide and 9-acridine carboxaldehyde have been identified in leaf, rhizome and root in both incubation levels 10 and 50 µM. 10,11-dihydro-10,11-dihydroxy-carbamazepine and 2,3-dihydro-2,3-dihydroxy-carbamazepine were identified in root and rhizome extracts in both incubation concentrations. Moreover, the first was identified in leaf extracts of 50 µM carbamazepine; however, the latter was not detectable in leaf extracts, respectively.

2.7. Impacts of DCF and CBZ on Phragmites australis Metabolic Pathways

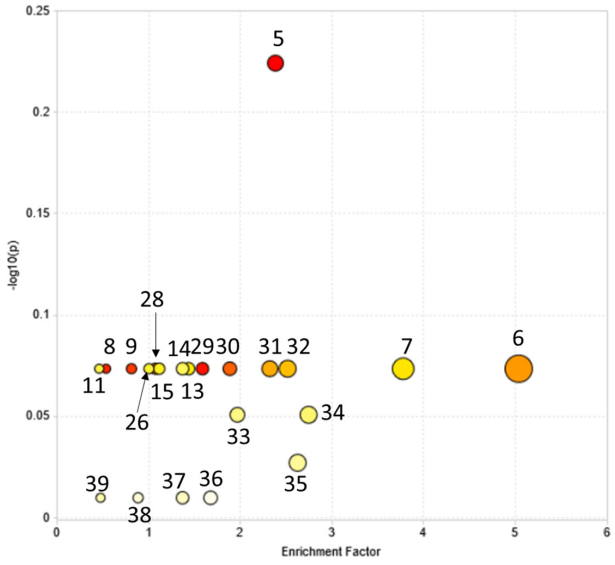
Metabolic pathway analysis was performed to identify the pathways that were induced upon the incubation of *P. australis* with DCF or CBZ via MetaboAnalyst 4.0 software (Montreal, QC, Canada) based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The results are established on OPLS-DA and S-plot analysis of different metabolites with/without DCF or CBZ incubation depending on the p (corr) vector in the OPLS-DA module. Consequently, these metabolites represented the metabolic profile (DMF) which differentiate between control and incubated samples.

For the investigation, the mummichog algorithm was the first implementation of this concept to infer pathway activities from a ranked list of MS peaks identified by untargeted metabolomics. The DMF was applied to test their participation in pathways using a combination of network analysis and functional enrichment analysis and listed in Tables S4 and S5. The pathways exhibiting $p < 0.05$ were considered a statistically significant metabolic pathway, meaning that they were affected with DCF or CBZ or their transformation products via MetaboAnalyst 4.0 software based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.

KEGG pathway analysis of the analyzed metabolites using the pathway data set of *A. thaliana* matched 27 pathways in *P. australis* incubated with DCF. Further, 11 were significantly biologically active and are listed in the table in (Figure 5a). They were identified as glycolysis/gluconeogenesis, ascorbate and aldarate metabolism, fructose and mannose metabolism, galactose metabolism, the pentose phosphate pathway, arginine biosynthesis, alanine, aspartate and glutamate metabolism, purine metabolism, pyrimidine metabolism, glutathione metabolism and phenylalanine metabolism.



(**a**)



(**b**)

Figure 5. *Cont.*

1	Glycolysis / Gluconeogenesis	20	Glyoxylate and dicarboxylate metabolism
2	Ascorbate and aldarate metabolism	21	Cyanoamino acid metabolism
3	Fructose and mannose metabolism	22	Starch and sucrose metabolism
4	Galactose metabolism	23	Amino sugar and nucleotide sugar metabolism
5	Pentose phosphate pathway	24	Inositol phosphate metabolism
6	Arginine biosynthesis	25	Sphingolipid metabolism
7	Alanine, aspartate, and glutamate metabolism	26	Pyruvate metabolism
8	Purine metabolism	27	Nicotinate and nicotinamide metabolism
9	Pyrimidine metabolism	28	One carbon pool by folate
10	Phosphatidylinositol signaling system	29	Arachidonic acid metabolism
11	Aminoacyl-tRNA biosynthesis	30	Fatty acid biosynthesis
12	Phenylalanine metabolism	31	Tyrosine metabolism
13	Glycine, serine, and threonine metabolism	32	Tryptophan metabolism
14	Valine, leucine, and isoleucine degradation	33	Pantothenate and COA biosynthesis
15	Valine, leucine, and isoleucine biosynthesis	34	Carbon fixation in photosynthetic organisms
16	Lysine degradation	35	Folates biosynthesis
17	Glutathione metabolism	36	Biosynthesis of unsaturated fatty acids
18	Histidine metabolism	37	Taurine hypotaurine metabolism
19	Phenylalanine, tyrosine, and tryptophan biosynthesis	38	Terpenoid backbone biosynthesis
		39	β-alanine metabolism

Figure 5. Overview of pathway analysis using MetaboAnalyst 4.0. For metabolite set enrichment analysis of *Phragmites australis* differentiating metabolic profile (DMF) after incubation with (**a**) 10 and 100 µM diclofenac; (**b**) 10 and 50 µM carbamazepine. The overview displays all matched pathways as circles. The color and size of each circle are based on the *p* value and the impact of the pathway value, respectively. The names of the metabolic pathways are listed in the table.

After CBZ incubation, 22 pathways were significantly altered. Twelve pathways have a more significant *p* value, relating them to the effect of CBZ and/or its transformation products (Figure 5b). They were the pentose phosphate pathway, purine metabolism, pyrimidine metabolism, fatty acid biosynthesis, arachidonic acid metabolism, tyrosine metabolism, tryptophan metabolism, β-Alanine metabolism, arginine and proline metabolism, pantothenate and CoA biosynthesis, carbon fixation in photosynthetic organisms and folate biosynthesis. The previous pathways were induced after incubation with DCF or CBZ.

3. Discussion

P. australis metabolic profiles were investigated in different samples with RPLC-HILIC-ESI-TOF-MS. The system accuracy and reproducibility were checked using a mixture of 13 reference standards. The results of the quality control mixture revealed that the serial coupling produces robust and reproducible data, which was used for further statistical analysis. *P.australis* metabolomics' statistical analysis of the four different extracts reveals that the 90% acidic MeOH extracts are very different from the other extracts of various *P. australis* leaf, rhizome and root samples. However, the 100% MeOH, 50% MeOH and 100% H_2O extracts are more related to each other. 100% MeOH was generally located in the middle between 50% MeOH and 100% H_2O extracts, however, representing a similar solution behavior to all three extract types. A large t2 range value was observed mostly in acidic 90% MeOH extracts, that is, a value far above the critical limits in the score space. Hence, this is likely to be an outlying observation. The elimination of acidic 90% MeOH extracts might improve the model. Also, most of the metabolites extracted with acidic 90% MeOH have been detected in three other solvents. The same observation was made for *Lemna minor* extracts and can be described with the different occurrence of charged molecules caused by different pH values [26].

The statistical software allowed cross-validation between the datasets. Furthermore, an additional step of cross-validation of the data was done through the differentiation between the x (metabolites) and the plant parts. The results reveal the significance of the data organization. Hence, the compounds extracted from roots and rhizomes are more correlated than compounds in leaves, which supports the physiological similarities between them [29]. After the additional cross-validation step, the OPLS-DA model was built to assign the variables, which differentiate between control and incubated samples. OPLS-DA is a supervised approach that tends to improve the separation between (two or more) groups of samples. For this reason, it is widely used for classification purposes and biomarker identification and/or differentiating metabolic profiles (DMF) in metabolomics studies [21].

The DMF of *P. australis* was extracted. Further, the fold change of the different variables in the control and incubated samples show an increase in quercetin in DCF incubation; however, in this case of CBZ, it was found to decrease significantly. *P. australis* responds to the incubation with different classes of compounds, with the potential to protect it against ROS, such as polyphenols and flavonoids. Also, it might be that CBZ does not cause oxidative stress to the same extent as DCF. Furthermore, succinic acid (DM_4), propane-1,2,3-triol (DM_6) and 2-hydroxypropanoic acid (DM_7) appeared significant in the positive part of the plot with VIP > 1 and $p < 0.05$ (Figure 4c). They were shifted toward the middle of the figure more than CBZ transformation products, indicating that they originated partially from diclofenac. The same was in CBZ transformation products, which were not produced by *P. australis.* The carbamazepine-10,11-epoxide was statistically significant: VIP > 1 and $p < 0.05$. Carbamazepine-10,11-epoxide is considered a main transformed product of carbamazepine in plants (compound **3** in Figure 4d). It is also the first metabolite of carbamazepine in tomato plants and *Armoracia rusticana* root cultures [28,30]. Hence, the increase of succinic acid (DM_4) might affect the TCA cycle (tricarboxylic acid cycle) of *P. australis.*

Also, diclofenac and its hydroxylated metabolites were identified in *P. australis* extracts. It has been reported that formation of hydroxylated metabolites is the first step in the detoxification of diclofenac in the plant. These results reveal that P450 monooxygenases or peroxidases were involved to detoxify diclofenac [11].

In CBZ incubation, different transformation products were identified, which originated from different metabolism pathways (PCM). The 10,11-diOH pathway (PCM) has been investigated comprehensively in plants like cucumber, tomato, sweet potato, lettuce, carrot and horseradish [28,31]. The main transformed product is carbamazepine-10,11-epoxide (Martínez-Piernas et al., 2019), which has been identified in leaf, rhizome and root in both incubations levels 10 and 50 µM, as shown in Table S3 [32]. This is the first oxidation step, which was observed in different organisms from bacteria and fungi to mammals. It was conducted by cytochrome P450 and/or peroxidases, leading to different sub-pathways (PCM).

The first sub-pathway (PCM) started with cleavage of the epoxide bond and the hydroxylation to form 10,11-dihydro-10,11-dihydroxy-carbamazepine. This step is catalyzed by epoxide hydrolase enzymes [33]. A recent study reported further metabolism of 10,11-dihydro-10,11-dihydroxy-carbamazepine through hydroxylation of the benzene ring [32]. 10,11-dihydro-10,11-dihydroxy-carbamazepine was identified in root and rhizome extracts in both incubation concentrations, as shown in Table S3. However, 10,11-dihydro-10-hydroxy-carbamazepine was identified in leaf extracts in the 50 µM carbamazepine. It seems that the 10,11-dihydro-10,11-dihydroxy-carbamazepine was transferred to rhizome and leaf for further metabolism. This explains why it reaches the leaves in high concentration from the highly loaded rhizome tissue. Thus, *P. australis* can degrade carbamazepine through a 10,11-diOH pathway (PCM), initiated in the root and completed in leaf and rhizome.

Further, *P. australis* metabolized the carbamazepine-10,11-epoxide into 9-acridine carboxaldehyde, which is analogous to other studies [28,31,32]. The 9-acridine carboxaldehyde

is a reactive compound that was identified in leaf, rhizome and root extracts of plants exposed to 10 and 50 µM carbamazepine. Further metabolites downstream from 9-acridine carboxaldehyde could not be detected. 9-acridine carboxaldehyde is typically transformed into acridone, which is a non-toxic compound with the formation of the two intermediates acridine and 9-hydroxy acridine [28]. In our conditions, they were below the detectable limit or did not arise. Because 9-acridine carboxaldehyde was identified in leaf, rhizome and root extracts in both concentrations, it may be assumed that there are further steps in the acridine metabolic pathway (PCM) that could be detected, like in lettuce, with longer incubation periods [32].

The last sub-pathway (PCM) is similar to the 10, 11-diOH pathway in the first steps. It includes a consecutive oxidation reaction on the aromatic group. Carbamazepine was metabolized to 2,3-dihydro-2,3-dihydroxy-carbamazepine which comprises two steps with the participation of a cytochrome P450 or a peroxidase and an epoxide hydrolase [33]. 2,3-dihydro-2,3-dihydroxy-carbamazepine was identified in the *P. australis* rhizome and root in both incubation levels. However, 2,3-dihydro-2,3-dihydroxy-carbamazepine was not detectable in leaf extracts. It is moreover possible that the leaf metabolizes CBZ to 4-hydroxy-carbamazepine as in the leaves of the tomato plant. It was identified formerly in lettuce after 4 days of exposure; however, the researchers spiked the lettuce with 1 mg L^{-1} carbamazepine [30,32].

Thus, it is possible that *P. australis* takes up and metabolizes carbamazepine in different mechanisms. Still, all the metabolites resemble phase I metabolism in plant metabolism and not the glycosylated transformation products (phase II) [3].

The mummichog algorithm of the DMF between control and incubated samples shows an induction in the eleven and twelve pathways due to incubation with DCF or CBZ, respectively. They were identified as glycolysis/gluconeogenesis, ascorbate and aldarate metabolism, fructose and mannose metabolism, galactose metabolism, the pentose phosphate pathway, arginine biosynthesis, alanine, aspartate and glutamate metabolism, purine metabolism, pyrimidine metabolism, glutathione metabolism and phenylalanine metabolism in DCF incubation.

After CBZ incubation, the pentose phosphate pathway, purine metabolism, pyrimidine metabolism, fatty acid biosynthesis, arachidonic acid metabolism, tyrosine metabolism, tryptophan metabolism, β-Alanine metabolism, arginine and proline metabolism, pantothenate and CoA biosynthesis, carbon fixation in photosynthetic organisms and folate biosynthesis pathways were induced after the incubation of CBZ.

The glucosinolate biosynthesis pathway provides the plant with defense compounds against biting insects and worms. Furthermore, the cytochrome P450 enzyme controls glucosinolate biosynthesis, which is involved in the metabolism of DCF and CBZ during phase I [11,34]. The decline in sugars after plant exposure to DCF and CBZ indicates an increase in energy consumption [35].

Pyrimidine metabolism could be proposed to be a result of glucosinolate biosynthesis pathway alteration, which up-regulates the biosynthesis of other stress-induced pathways [36]. Pyrimidine metabolism is enhanced and leads to products that could be used in the case of salvage, that is, recovery of infections and subsequent synthesis of secondary products with specific functions in defense mechanisms. Additionally, pyrimidine metabolism provides a source of β-alanine or β-aminobutyrate, which might be an important source for the pantothenate of coenzyme A [37].

Furthermore, fatty acid biosynthesis and the related pyruvate metabolism were induced after incubation of *P. australis* with DCF and CBZ. After incubation with DCF, *P. australis* exhibits some responses differently from incubation with CBZ, since it showed an increase in phenylalanine metabolism which aligned with the increase of quercetin content. The flavonoid 3′-monooxygenase and the flavonoid 3′,5′-hydroxylase enzymes responsible for the conversion of kaempferol to quercetin are cytochrome P450 plant types, which are triggered after incubation with DCF. However, the quercetin levels were lower after CBZ incubation (Figure 4d).

Glutathione metabolism was affected in DCF incubation but not under the influence of CBZ. Glutathione metabolism is a central part of the antioxidative ascorbate–glutathione cycle. Further, glutathione is critical for the detoxification of xenobiotics, environmental stress tolerance and, in the form of phytochelatins, also the retention of heavy metals [38]. The metabolism of DCF required Nicotinamide-Adenine Dinucleotide Phosphate (NADPH) as reductants [11]. Upon DCF incubation, succinic acid formed from the degradation of diclofenac might have fueled the TCA cycle, which has been detected significantly in extracts incubated with DCF (Figure 4c).

In 2019, Sivaram and coworkers reported a high impact on the TCA cycle in maize leaves exposed to pyrene [39]. Consequently, the γ-Aminobutyric acid (GABA) shunt bypasses two steps of the TCA cycle. Moreover, it has an important pathway under stress conditions and it is associated with numerous physiological responses, including the regulation of cytosolic pH, carbon fluxes into the TCA cycle, nitrogen metabolism, osmoregulation and plant-pathogen interaction. Increased GABA levels also occur in response to changing environmental conditions and represent another adaptive mechanism in the attempt to maintain the rate of respiration under certain harmful conditions (Araújo et al., 2012) [40].

The same result was observed in lettuce crops, which were exposed to different contaminants of emerging concern (CEC) concentrations [35]. Also, it was reported that aromatic hydrocarbons altered the osmotic balance in maize [39]. The alteration of previously mentioned biosynthetic pathways enhances the *P. australis* defense mechanisms. They also seem to be involved in the transformation of DCF and CBZ.

Finally, *P. australis* used the glutathione metabolism pathway to defend itself against the DCF and seemingly, used the unsaturated fatty acid pathway to protect itself during the incubation with CBZ. Hence, *P. australis* responded differently to the DCF and CBZ through changing its metabolic pathway regardless of the type of drug to some extent. Consequently, specific changes in several common metabolic pathways can be considered as a marker for pollutant exposure in *P. australis*. However, each drug has fingerprints of the alteration of distinct metabolic pathways, which might be connected to its metabolites and the enzymes involved in metabolism. Therefore, the induced or changed pathways could be used as indicators for the exposure of the plant to DCF or CBZ.

4. Materials and Methods

4.1. Reagents and Chemicals

LC-MS grade methanol and water were obtained from VWR, Darmstadt, Germany. Quercetin dehydrate (≥95%, high performance liquid chromatography (HPLC)) was purchased from Enzo Life Sciences GmbH, Lörrach, Germany. Diclofenac (>99%) was obtained from Cayman Chemical Company, Ann Arbor, Michigan, MI, USA. Glyphosate (100 quality level, HPLC), gabapentin (200 quality level, HPLC), monuron (100 quality level, HPLC), chloridazon (100 quality level, HPLC), carbetamide (100 quality level, HPLC), metobromuron (100 quality level, HPLC), sotalol (≥98%), quinoxyfen (100 quality level, HPLC), metconazol (100 quality level, HPLC) and fenofibrate (≥99%) were obtained from Sigma, Darmstadt, Germany. Metformin (300 quality level, HPLC) was obtained from Fluka, Buchs, Switzerland. Furthermore, chlorbromuron (99.24%) and diazinon (99.53%) were obtained from Dr. Ehrenstorfer, Augsburg, Germany. Carbamazepine, 2,3-dihydro-2,3-dihydroxycarbamazepine, 10,11-dihydro-10,11-dihydroxy-carbamazepine, 10,11-dihydro-10-hydroxy-carbamazepine, 9-acridine carboxaldehyde and carbamazepine-10,11-epoxide were kindly provided by the German Research Center for Environmental Health, Comparative Microbiome Analysis (COMI), Helmholtz Centrum of Munich, Munich, Germany.

4.2. Plant Samples

Twelve *P. australis* plants were grown in semi-hydroponic conditions in the greenhouse as described by Sauvêtre and Schröder earlier, in 2015 [7]. Plants were grown in Hoagland solution made of (in mg/L) 472.30 $Ca(NO_3)_2 \cdot 4H_2O$, 202.22 KNO_3, 492.96 $MgSO_4 \cdot 7H_2O$, 68.04 KH_2PO_4, 80.04 NH_4NO_3, H_3BO_3, 1.8 $MnCl_2 \cdot 4H_2O$, 0.2 $ZnSO_4 \cdot 7H_2O$,

0.1 $CuSO_4 \cdot 5H_2O$, 0.025 $NaMoO_4$ and 3.67 FeNa-Ethylenediaminetetraacetic acid. Plants (approximately 0.8 m in height) of uniform size were selected and placed into individual pots containing 2 L of spiked Hoagland solution Each pot contained one plant and was arranged in the greenhouse following a completely randomized design. The nutrient medium was spiked with a stock solution to reach the desired final concentration (diclofenac, 10 and 100 µM and carbamazepine, 10 and 50 µM, respectively). Control plants growing in Hoagland solution (spiked with the same amount of solvent as incubated plants) were used to obtain a reference plant matrix. Two pots were set up for each of the four exposure concentrations. Each assay consisted of duplicates arranged in the greenhouse following a randomized design. To compensate for water losses by evapo-transpiration, distilled water was added daily to the pots for a final volume of 2 L. Plants were exposed for 4 days before they were harvested. Harvested material was divided into roots, rhizomes and leaves, frozen in liquid nitrogen and stored at −81 °C until further processing. *P. australis* was kindly provided by the German Research Center for Environmental Health, Comparative Microbiome Analysis (COMI), Helmholtz Center Munich, Munich, Germany.

4.3. Extraction

Twelve samples of *P. australis* leaves, rhizomes and roots, respectively, were collected, frozen under liquid N2. Then, the samples were freeze-dried and milled (Retsch S1 planetary ball mill, Retsch GmbH, Haan, Germany). Duplicates of each plant part were extracted with (a) 100% methanol (MeOH), (b) 90% MeOH (MeOH-water-formic acid (FAC) (90:9.5:0.5, $v/v/v$), (c) MeOH-water (50:50, v/v) and (d) 100% water (H_2O), respectively. The solvents containing 500 mg plant powder were sonicated (Sonorex super RK 106, Bandelin, Germany) for 10 min at 4 °C with 35 kHz frequency. Then, samples were centrifuged (Z 200 A Universal Compact Centrifuge, Hermle LaborTechnik GmbH, Wehingen, Germany) at 1500 rpm/261.6× g for 20 min and the supernatants were transferred to clean glass test tubes. The extraction process was triplicated in identical experimental conditions. Finally, the extracts were evaporated to dryness (using a SpeedVac, Fischer Scientific, Göteborg, Sweden) and dissolved in (50:50 (v/v %)) MeOH: H_2O [18,26].

4.4. Instruments

Filtered samples and standards (with 22 µm filter, Analytics Shops, Munich, Germany) were separated by LC (Agilent 1260 Infinity) consisting of an autosampler, two columns, two binary pumps, an online degasser, a mixing chamber and a UV detector. The LC-system was used to perform reversed-phase and zwitterion hydrophilic interaction liquid chromatography (HILIC) in its serial coupling. The reversed-phase separation column was a Poroshell 120 EC-C18 (50.0 × 3.0 mm, 2.7 µm, Agilent Technologies Waldbronn, Germany). The HILIC column was a ZIC-HILIC column (150 × 2.1 mm, 5 µm, 200 Å, Merck Sequant, Umea, Sweden). Columns were coupled through a T-piece (Upchurch, IDEX Europe GmbH, Erlangen, Germany). The third port of the T-piece was connected to the HILIC flow pump. The injection volume was 10 µL. Further details, like the mobile phase of the RPLC-HILIC serial coupling and other settings, are described in References [24–26].

Samples and reference mixes were analyzed with a HRMS "time-of-flight" (TOF) mass spectrometer equipped with a Jet Stream ESI interface (Agilent Technologies, Waldbronn, Germany) and hyphenated with RPLC-HILIC chromatography. The parameters were as follows: 325 °C gas temperature, 10 L/min drying gas flow, 325 °C sheath gas temperature, 7.5 L/min sheath gas flow, 45-psi nebulizer operating pressure and 100 V fragmentor voltage. Ions were detected in positive ionization mode with a mass range of 50–2100 Daltons. The resolution of the instrument was better than 10,000 at m/z 922.

4.5. Quality Control of the RPLC-HILIC-ESI-TOF-MS System

The robustness and reproducibility of the RPLC-HILIC-ESI-TOF-MS system were tested with a standard mixture containing 13 different reference standards with a 20 µM final concentration. The mixture consisted of metformin, glyphosate, gabapentin, monuron,

chloridazon, carbetamide, metobromuron, sotalol, chlorbromuron, diazinon, quinoxyfen, metconazol and fenofibrate. The mixture was injected at the beginning/end of the experiment series and at fixed intervals during the experiment. It was injected after each extraction batch.

The absolute variation between the literature monoisotopic mass and the mean of measured isotopic masses (Δppm) was computed according to the following equation:

$$\Delta\text{ppm} = (\text{monoisotopic mass of standard} - \text{mean of standard masses})/\text{monoisotopic mass of standard} \times 10^6 \quad (1)$$

Moreover, the standard deviations (SD) of RT and Relative standard deviations (RSD) were calculated:

$$\%\text{ of RSD} = \text{SD of compound RTs in different injection}/\text{Mean of compound RTs} \quad (2)$$

The results are summarized in (Tables S1 and S2). The absolute mass deviation ranged from 0.2 (Da) to 7 (Da). The RT standard deviation was less than 1%. Moreover, the relative standard deviation (%RSD) ranged from 0.5% to 3.6%. Thus, the results indicate the accuracy, repeatability and reproducibility of the LC system as previously reported in the investigation of plant metabolites in *Lemna minor* samples [18].

4.6. Data Evaluation

4.6.1. Spectrometric Data Evaluation

Data acquired with MassHunter Workstation LC/MS Data Acquisition software B 05.00, (Agilent Technologies, Waldbronn, Germany) was subsequently analyzed with Profinder B.06.00 (Agilent Technologies, Waldbronn, Germany) to extract the so-called "features" by their retention times (RT), molecular mass and their peak intensity in various *P. australis* extracts. This was performed in a combination of the 3-fold injections of each sample after removing the features found in the corresponding blank samples. The parameters are set to a peak filter of 1000 counts peak height, ion species to "positive ions" with H^+, Na^+, K^+ and $NH4^+$, "charge state" to 1, the "expected RT" to ±3.00 min and the mass to ±10 ppm. The extracted ion chromatograms (EICs) were smoothed with a Gaussian function using 9 points function and 5000 points Gaussian width. This limits the result finally to 2000 compound groups.

4.6.2. DCF and CBZ Transformation Products Detection

DCF and CBZ transformation product standards were analyzed using Agilent Profinder B.06.00 (Agilent Technologies). They were detected in the *P. australis* extracts, with masses estimated at ±10 ppm and RT ±0.3 min of the exact mass and RT of the standards, respectively. The metabolites were identified and suspected (when analytical standards were not available) after RPLC-HILIC-ESI-TOF-MS separation in the suspect analysis. (Suspects screening typically is performed with accurate and high-resolution mass spectrometers to observe the empirical formula of each molecule present and/or with tandem-mass spectrometry to observe specific fragment spectra). A local database was built using MassHunter PCDL Manager B.04.00 (Agilent Technologies, Waldbronn, Germany). Further, the logD (pH7) was the third parameter used to certify the identity of metabolites. The highly polar to polar compounds eluted from the HILIC column at RT < 15 min, with logD values below zero. The nonpolar compounds were eluted from the RP column at RT > 15 min, with logD values above zero. Metabolites within the criteria of mass, RT and logD (pH7) in the suspect analysis were considered.

4.6.3. Statistical Data Analysis

Data statistical analyses were conducted with SIMCA 16 software (Malmö, Sweden). Further analysis and data evaluation were performed with Microsoft Access and Excel 2016 (Redmond, Washington, WA, USA) and OriginPro 2019, Origin Lab cooperation, Northampton, MA, USA.

The preprocessed data is a matrix. The rows are the exact masses, retention times (RTs) and abundances of each sample, which were listed in a Microsoft Excel Sheet. For statistical analysis, it is common to handle the data matrices with rows as observations and columns as compounds [36]. Therefore, the data was organized in the Microsoft Access Database file (DBF), which was exclusively built to be suitable for SIMCA 16. The main advantage of the data matrix is the inherent support to align quantitative data (plant part, plant number, extraction solvent and drug incubation) along with related metadata (i.e., feature annotations/abundance as columns and sample annotations as rows). In DBF, the RTs, masses and abundances were connected to the corresponding plant part (i.e., leaf, rhizome and root), plant number (i.e., plants 1 and 2), extraction solvent and drug incubation. Once the matrix was built, comprehensive statistical analyses could be performed by using the vast range of functions provided by the software. The matrix consisted of 432 observations (i.e., the incubation with/without DCF or CBZ) and 11,442 variables (features). Furthermore, the plant part, plant number and extraction solvent were used as secondary observations. The data was not transformed and centered; however, it was scaled. The data was analyzed according to the following two strategies, considering the statistical analyses in untargeted metabolomics. Further information about the data setup can be found in the supplementary material. By default, SIMCA provides an algorithm called "cross-validation" to get the most valid model by calculating the adequate number of principal components to prevent overfitting of the data in the model. Furthermore, the software provides a large number of visual diagrams (score plots (with Hotelling's ellipse), DModX (Distant to Model) and statistic tables to assess the quality of the model in addition to the R^2 and Q^2.

1. Metabolite fingerprinting was used to capture metabolite patterns across metabolite profiles. They are characterized without further identification steps (i.e., without need for standard reference material). Partial Least Squares (PLS) and Orthogonal Partial Least Squares regression-Discriminant Analysis (OPLS-DA) were used to relate sets of X-variables (such as plant part, plant number, extraction solvent and drug incubation) to the metabolites matrix. SIMCA 16 has a tool called Multiblock Orthogonal Component Analysis (MOCA). MOCA's concept is used to accomplish a fast and accurate analysis of multiple blocks of data (variables) registered for the same set of observations. MOCA aims at extracting the information in complex multi-block data analytics. Furthermore, it will extract two sets of components: the joint and the unique components. The quality of the models is described by R^2 and Q^2 values, where R^2 is the proportion of variance in the data explained by the models and indicates the goodness of fit and Q^2 is the proportion of variance in the data predictable by the model and expresses predictability [41].
2. Metabolite profiling which uses sets of predefined metabolites were studied in different samples of *P. australis* and differences in metabolites were usually related to the incubation with DCF or CBZ. Metabolite/variable selection was conducted to observe only the most significant metabolite candidates that explain the differences between the samples using S- and contribution-plots. The statistical models were built with confidence limits at 95%. Also, the differentiating metabolic profile (DMF) was chosen based on their contribution to the variation and correlation within the data sets. The related metabolic pathways were analyzed using MetaboAnalyst 4.0. Moreover, their contributions and biological clarifications were described based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The KEGG pathway analysis tool was used by the Arabidopsis thaliana database. The pathway analysis module combines the enrichment analysis and topology analysis based on KEGG. Fisher's test was used to generate p values. The p value was equal to 0.05, which indicates the fundamental connection of the identified metabolite with their respective metabolite and not due to the random chance [42,43].

5. Conclusions

Metabolites of *P. australis* influenced by pharmaceuticals were investigated using RPLC-HILIC-ESI-TOF-MS. The experimental data and the statistical analysis revealed a change in the metabolites' fingerprint between the different extracts, different plant parts and upon incubation with DCF and CBZ. The PLS and OPLS-DA identified the statistically significant clusters between the different groups. Further, significant DMF was determined in *P. australis* after the incubation with DCF or CBZ, individually. Different metabolic pathways were predicted from the statically identified DMF. These pathways were related mainly to the defense of the plant against stressful environmental conditions. *P. australis* adapted to each drug differently. *P. australis* could putatively use the glutathione metabolism pathway and unsaturated fatty acid pathway to protect itself during the incubation with DCF and CBZ, respectively. The results reveal insights into the metabolic profile of the species´ adaptation to different pollutants. This study may set a cornerstone for understanding the changes in the plant metabolism after incubation with DCF and CBZ. Also, the mass spectrometric untargeted metabolomics strategy has a substantial role in investigating the biochemical changes and metabolic adaptation of plants in xenobiotics exposure cases.

Supplementary Materials: The supplementary materials are available online at https://www.mdpi.com/2218-1989/11/1/2/s1, Figure S1: Retention time (RT)/Mass plot of the background was analyzed by RPLC-HILIC-ESI-TOF-MS in positive electrospray ionization mode, Figure S2: The Q^2/ R^2 Overview plot displays the individual cumulative R^2 (green columns) and Q^2 (blue columns) and Q^2 for the goodness of fits and cross-validation parameters (a) *P. australis* different parts. (b) *P. australis* different incubation, Figure S3: EICs were corresponding to measured diclofenac (right) and the reference standard (left), which were identified in the extracts of *Phragmites australis* leaf, rhizome, and roots incubated with 10 and 100 µM diclofenac. Also, EICs relative to transformed products are suspected in the extracts of *Phragmites australis* leaf, rhizome, and roots incubated with 10 and 100 µM diclofenac, Figure S4: EICs were corresponding to carbamazepine (CBZ) and its transformed product standards (left), which were identified in the extracts of *Phragmites australis* leaf, rhizome, and roots incubated with 10 and 50µM carbamazepine (measured right), Table S1: The standards compounds of the quality control external calibration mixture, monoisotopic mass in the literature (L), monoisotopic in different injection and the mean of them, the variation between monoisotopic mass in the literature (L), and the mean of measured monoisotopic mass, and mean mass standard deviation (SD) are listed, Table S2: The standards compounds of the quality control external calibration mixture, the single RT, mean RT of the different injections, mean RT standard deviation (SD), and relative standard deviation (RSD) of the standards are listed, Table S3: List of carbamazepine (CBZ) transformed product identified in *Phragmites australis* different samples with the mean monoisotopic mass in the standards (S), the mean monoisotopic mass of *Phragmites australis* (Ph), the variation between them, mean RT of standards (S), mean RT of *Phragmites australis* (Ph), and the variation between them were listed, Table S4: The differentiating metabolic profile (DMF) metabolites of *Phragmites australis* due to incubation with 10 & 100 µM diclofenac were extracted from the S-plot, monoisotopic mass, *p*-value, and t. score are listed, Table S5. The differentiating metabolic profile (DMF) metabolites of *Phragmites australis* due to incubation with 10 & 50 µM carbamazepine were extracted from the S-plot, monoisotopic mass, *p*-value, and t.score are listed.

Author Contributions: R.W. and T.L. designed the metabolomics study. R.W. prepared the samples, performed the untargeted metabolomics analysis and ran the analyses in the metabolomics platform. S.M. created the statistical design for data evaluation and data interpretation and performed the realization together with R.W. and T.L., A.S. and P.S. created the test set and performed the plant growing and incubation together with R.W. They also supported R.W. in biological data interpretation. R.W. and T.L. conceived and drafted the manuscript. All the authors contributed with critical intellectual input. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the Bavarian State Ministry of the Environment and Consumer Protection as well as the Cultural Affairs and Mission Sector of the Egyptian Ministry of Higher Education.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data is available at https://www.ebi.ac.uk/metabolights/ with ID—sMTBLS2321.

Acknowledgments: This work was assisted by the German Research Foundation (DFG) and the Technical University of Munich (TUM) in the scope of the Open Access Publishing Program.

Conflicts of Interest: The authors declare that they have no conflict of interest. There is no connection between AFIN-TS GMbH and the subject of this manuscript.

References

1. Tohge, T.; Nishiyama, Y.; Hirai, M.Y.; Yano, M.; Nakajima, J.-I.; Awazuhara, M.; Inoue, E.; Takahashi, H.; Goodenowe, D.B.; Kitayama, M.; et al. *Identification of Genes Involved in Anthocyanin Accumulation by Integrated Analysis of Metabolome and Transcriptome in Pap1-Overexpressing Arabidopsis Plants*; Springer: Dordrecht, The Netherlands, 2007; pp. 159–168.
2. Sandermann, H.; Diesperger, H.; Scheel, D. *Metabolism of Xenobiotics by Plant Cell Cultures*; Springer: Berlin/Heidelberg, Germany, 1977; pp. 178–196.
3. Schröder, P.; Scheer, P.; Belford, E.J.D. Metabolism of organic xenobiotics in plants: Conjugating enzymes and metabolic endpoints. *Minerva Biotecnol.* **2002**, *13*, 85–91.
4. Schröder, P.; Maier, H.; Debus, R. Detoxification of Herbicides in *Phragmites australis*. *Z. Nat. C* **2005**, *60*, 317. [CrossRef]
5. Villette, C.; Maurer, L.; Wanko, A.; Heintz, D. Xenobiotics metabolization in Salix alba leaves uncovered by mass spectrometry imaging. *Metabolomics* **2019**, *15*, 122. [CrossRef]
6. Letzel, M.; Metzner, G.; Letzel, T. Exposure assessment of the pharmaceutical diclofenac based on long-term measurements of the aquatic input. *Environ. Int.* **2009**, *35*, 363–368. [CrossRef] [PubMed]
7. Sauvêtre, A.; Schröder, P. Uptake of carbamazepine by rhizomes and endophytic bacteria of *Phragmites australis*. *Front. Plant Sci.* **2015**, *6*, 83. [CrossRef] [PubMed]
8. Thelusmond, J.-R.; Kawka, E.; Strathmann, T.J.; Cupples, A.M. Diclofenac, carbamazepine and triclocarban biodegradation in agricultural soils and the microorganisms and metabolic pathways affected. *Sci. Total Environ.* **2018**, *640–641*, 1393–1410. [CrossRef] [PubMed]
9. Vieno, N.; Sillanpää, M. Fate of diclofenac in municipal wastewater treatment plant—A review. *Environ. Int.* **2014**, *69*, 28–39. [CrossRef] [PubMed]
10. Hai, F.I.; Yang, S.; Asif, M.B.; Sencadas, V.; Shawkat, S.; Sanderson-Smith, M.; Gorman, J.; Xu, Z.-Q.; Yamamoto, K. Carbamazepine as a Possible Anthropogenic Marker in Water: Occurrences, Toxicological Effects, Regulations and Removal by Wastewater Treatment Technologies. *Water* **2018**, *10*, 107. [CrossRef]
11. Huber, C.; Bartha, B.; Schröder, P. Metabolism of diclofenac in plants—hydroxylation is followed by glucose conjugation. *J. Hazard. Mater.* **2012**, *243*, 250–256. [CrossRef]
12. Park, M.G.; Blossey, B. Importance of plant traits, and herbivory for invasiveness of *Phragmites australis* (Poaceae). *Am. J. Bot.* **2008**, *95*, 1557–1568. [CrossRef]
13. Gray, K.R.; Biddlestone, A.J. Engineered reed-bed systems for wastewater treatment. *Trends Biotechnol.* **1995**, *13*, 248–252. [CrossRef]
14. Zhou, B.; Xiao, J.F.; Tuli, L.; Ressom, H.W. LC-MS-based metabolomics. *Mol. Biosyst.* **2012**, *8*, 470–481. [CrossRef] [PubMed]
15. Beale, D.; Pinu, F.; Kouremenos, K.; Poojary, M.; Narayana, V.; Boughton, B.; Kanojia, K.; Dayalan, S.; Jones, O.; Dias, D. Review of recent developments in GC–MS approaches to metabolomics-based research. *Metabolomics* **2018**, *14*, 152. [CrossRef]
16. Emwas, A.-H.; Roy, R.; McKay, R.T.; Tenori, L.; Saccenti, E.; Gowda, G.A.N.; Raftery, D.; Alahmari, F.; Jaremko, L.; Jaremko, M.; et al. NMR Spectroscopy for Metabolomics Research. *Metabolites* **2019**, *9*, 123. [CrossRef] [PubMed]
17. Aretz, I.; Meierhofer, D. Advantages, and pitfalls of mass spectrometry-based metabolome profiling in systems biology. *Int. J. Mol. Sci.* **2016**, *17*, 632. [CrossRef]
18. Wahman, R.; Grassmann, J.; Schröder, P.; Letzel, T. Plant metabolomic workflows using reversed-phase LC and HILIC with ESI-TOF-MS. *LCGC N. Am.* **2019**, *37*, 8–15.
19. Riach, A.C.; Perera, M.V.L.; Florance, H.V.; Penfield, S.D.; Hill, J.K. Analysis of plant leaf metabolites reveals no common response to insect herbivory by *Pieris rapae* in three related host-plant species. *J. Exp. Bot.* **2015**, *66*, 2547–2556. [CrossRef]
20. Gromski, P.S.; Xu, Y.; Kotze, H.L.; Correa, E.; Ellis, D.I.; Armitage, E.G.; Turner, M.L.; Goodacre, R. Influence of missing values substitutes on multivariate analysis of metabolomics data. *Metabolites* **2014**, *4*, 433–452. [CrossRef]
21. Lamichhane, S.; Sen, P.; Dickens, A.M.; Hyötyläinen, T.; Orešič, M. Chapter Fourteen—An Overview of metabolomics data analysis: Current tools and future perspectives. *Compr. Anal. Chem.* **2018**, *82*, 387–413.
22. Prinsloo, G.; Vervoort, J. Identifying anti-HSV compounds from unrelated plants using NMR and LC-MS metabolomic analysis. *Metabolomics* **2018**, *14*, 134. [CrossRef]
23. Worley, B.; Powers, R. Multivariate Analysis in Metabolomics. *Curr. Metab.* **2013**, *1*, 92–107.
24. Bieber, S.; Greco, G.; Grosse, S.; Letzel, T. RPLC-HILIC, and SFC with mass spectrometry: Polarity-extended organic molecule screening in environmental (Water) samples. *Anal. Chem.* **2017**, *89*, 7907–7914. [CrossRef] [PubMed]

25. Greco, G.; Grosse, S.; Letzel, T. Serial coupling of reversed-phase and zwitterionic hydrophilic interaction LC/MS for the analysis of polar and nonpolar phenols in wine. *J. Sep. Sci.* **2013**, *36*, 1379–1388. [CrossRef] [PubMed]
26. Wahman, R.; Graßmann, J.; Sauvêtre, A.; Schröder, P.; Letzel, T. *Lemna minor* studies under various storage periods using extended-polarity extraction and metabolite non-target screening analysis. *J. Pharm. Biomed. Anal.* **2020**, *188*, 113362. [CrossRef]
27. Rajab, M.; Greco, G.; Heim, C.; Helmreich, B.; Letzel, T. Serial coupling of RP and zwitterionic hydrophilic interaction LC-MS: Suspects screening of diclofenac transformation products by oxidation with a boron-doped diamond electrode. *J. Sep. Sci.* **2013**, *36*, 3011–3018. [CrossRef]
28. Sauvetre, A.; May, R.; Harpaintner, R.; Poschenrieder, C.; Schröder, P. Metabolism of carbamazepine in plant roots and endophytic rhizobacteria isolated from *Phragmites australis*. *J. Hazard. Mater.* **2018**, *342*, 85–95. [CrossRef]
29. Hess, D. *Plant Physiology*; Springer: Berlin/Heidelberg, Germany, 1975; p. 334.
30. Riemenschneider, C.; Seiwert, B.; Schwarz, D.; Reemtsma, T. Extensive Transformation of the pharmaceutical carbamazepine following uptake into intact tomato plants. *Environ. Sci. Technol.* **2017**, *51*, 6100–6109. [CrossRef]
31. Riemenschneider, C.; Seiwert, B.; Goldstein, M.; Al-Raggad, M.; Salameh, E.; Chefetz, B.; Reemtsma, T. An LC-MS/MS method for the determination of 28 polar environmental contaminants and metabolites in vegetables irrigated with treated municipal wastewater. *Anal. Methods* **2017**, *9*, 1273–1281. [CrossRef]
32. Martínez-Piernas, A.B.; Nahim-Granados, S.; Polo-López, M.I.; Fernández-Ibáñez, P.; Murgolo, S.; Mascolo, G.; Agüera, A. Identification of transformation products of carbamazepine in lettuce crops irrigated with Ultraviolet-C treated water. *Environ. Pollut.* **2019**, *247*, 1009–1019. [CrossRef]
33. Tybring, G.; von Bahr, C.; Bertilsson, L.; Collste, H.; Glaumann, H.; Solbrand, M. Metabolism of carbamazepine and its epoxide metabolite in human and rat liver in vitro. *Drug Metab. Dispos.* **1981**, *9*, 561–564.
34. Lee, D.-K.; Ahn, S.; Cho, H.Y.; Yun, H.Y.; Park, J.H.; Lim, J.; Lee, J.; Kwon, S.W. Metabolic response induced by parasitic plant-fungus interactions hinder amino sugar and nucleotide sugar metabolism in the host. *Sci. Rep.* **2016**, *6*, 1–11. [CrossRef] [PubMed]
35. Hurtado, C.; Parastar, H.; Matamoros, V.; Piña, B.; Tauler, R.; Bayona, J.M. Linking the morphological and metabolomic response of *Lactuca sativa* L exposed to emerging contaminants using GC × GC-MS and chemometric tools. *Sci. Rep.* **2017**, *7*, 6546. [CrossRef] [PubMed]
36. Tantikanjana, T.; Mikkelsen, M.D.; Hussain, M.; Halkier, B.A.; Sundaresan, V. Functional Analysis of the tandem-duplicated P450 genes SPS/BUS/CYP79F1 and CYP79F2 in glucosinolate biosynthesis and plant development by Ds transposition-generated double mutants. *Plant Physiol.* **2004**, *135*, 840–848. [CrossRef]
37. Zrenner, R.; Stitt, M.; Sonnewald, U.; Boldt, R. Pyrimidine and purine biosynthesis and degradation in plants. *Annu. Rev. Plant Biol.* **2006**, *57*, 805–836. [CrossRef] [PubMed]
38. Gong, B.; Sun, S.; Yan, Y.; Jing, X.; Shi, Q. Glutathione Metabolism and Its Function in Higher Plants Adapting to Stress. In *Antioxidants and Antioxidant Enzymes in Higher Plants*; Springer International Publishing: Heidlberg, Germany, 2018; pp. 181–205.
39. Sivaram, A.K.; Subashchandrabose, S.R.; Logeshwaran, P.; Lockington, R.; Naidu, R.; Megharaj, M. Metabolomics reveals defensive mechanisms adapted by maize on exposure to high molecular weight polycyclic aromatic hydrocarbons. *Chemosphere* **2019**, *214*, 771–780. [CrossRef] [PubMed]
40. Araújo, W.L.; Nunes-Nesi, A.; Nikoloski, Z.; Sweetlove, L.J.; Fernie, A.R. Metabolic control and regulation of the tricarboxylic acid cycle in photosynthetic and heterotrophic plant tissues. *Plant Cell Environ.* **2012**, *35*, 1–21. [CrossRef]
41. Löfstedt, T.; Trygg, J. OnPLS—A novel multiblock method for the modeling of predictive and orthogonal variation. *J. Chemom.* **2011**, *25*, 441–455.
42. Chong, J.; Wishart, D.S.; Xia, J. Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. *Curr. Protoc. Bioinform.* **2019**, *68*, e86. [CrossRef]
43. Chong, J.; Yamamoto, M.; Xia, J. MetaboAnalystR 2.0: From Raw Spectra to Biological Insights. *Metabolites* **2019**, *9*, 57. [CrossRef]

Untargeted metabolomics studies on drug-incubated Phragmites australis profiles

Rofida Wahman[1], Andres Sauvêtre [2+], Peter Schröder [2], Stefan Moser [3], Thomas Letzel[1,4,*]

[1] Chair of Urban Water Systems Engineering, Technical University of Munich, Am Coulombwall 3, 85748 Garching, Germany.

[2] German Research Center for Environmental Health, Research Unit Comparative Microbiome Analysis, Helmholtz Centrum Munich, Ingolstadt street 1, 85764 Neuherberg, Germany

[3] Stefan Moser Process Optimization, Weberweg 3, D-83131 Nußdorf am Inn, Germany

[4] Analytisches Forschungsinstitut für Untarget Screening GmbH (AFIN-TS), Am Mittleren Moos 48, D - 86167 Augsburg, Germany

*Correspondence: T.letzel@tum.de

+ Current address: HydroSciences Montpellier, UMR 5569, Faculté de Pharmacie, University of Montpellier, Avenue Charles Flahault 15, 34000, Montpellier, France

2. Materials and methods:

2.1. Data Processing:

2.1.1. Access data file Creation

The experiment setup was inserted into the access database. Each run has a unique numerical ID from 1 to 432. Masses, RTs, and an abundance of compounds from each run were merged to the ID by inserting them into the access database. The internal Access programming helped to visualization and check that all the data was correctly inserted. This was done by reviewing the graphs and the corresponding pre documented additional data like (Solvent, part of the plant …..). After this from the dataset, a pivot matrix data table was calculated.

The pivot data table was arranged that masses@Rt are used as rows while the columns documented the accompanied abundance. The algorithm to attach the abundance to mass@Rt used the first value of abundance. Different approaches of mapping were tested such as (Mean, Average, First Value, Last Value…) but turned out to process a similar outcome in further analysis. Therefore, combination mass@Rt has occurred only in a unique combination with abundance.

Because of the limited possible number of columns (IDs) in Access the data table needed to be split into single excel files via a script before reunion it in SIMCA.

After the Excel, files with all the runs were merged in SIMCA the data table needed to be transposed to treat the different mass@RT as variables (Columns) and the different runs as observations (Rows).

After the union and transposing the data, the additional documentation like (Solvent, part of the plant …..) was pasted into SIMCA as well.

In the statistical software (SIMCA), the additional information (Solvent, part of the plant …..) were defined as secondary observations. This means that this information is not used within the developed models as variables.

As expected, some variables Mass@RT weren't found in all observation runs. Accordingly, the pivot table used to merge the data did document this with missing data. This is not very helpful in analyzing the data because the statistical software would see this data as "missing". Instead of "no occurrence". To put this right the empty cells were replaced with zeros

After this, the data was used to build the first model. In this starting PCA analysis, the untreated data was stored as a reference and to start the basic analysis with further models.

This is an especially important step to get an overview of the overall pattern in the data. The most important tools are a.) the score scatters plot which presents the consistency of the data using the uses the hosteling ellipse. This ellipse represents a 95% confidence interval in the multidimensional space. Observation outside this ellipse is remarkably interesting and needs to be investigated. Sometimes these Observations could also be identified as outliers.

Also the DModX "Distance to the Model in X" could give insights about the portion of the Variance (Predicted – Observed) which couldn't be described by the Model, to get a better understanding of what the model is capable of and what might be very unlikely and need more detailed investigation.

In the data, no anomalies or outliers have been found.

Within the first analysis, the underlying correlation pattern is represented with clusters in the score scattered plot, who summarized the information of all investigation runs in each one data point. With the help of the 2nd observations, the data set can be colored accordingly to check if the for-instance solvent or part of the plant does have significant uniqueness to expose the observation in one of the clusters. This is an easy way to analyze the clusters using the secondary information of documentation without considering the extra information to build the model. Then, the OPLS-DA model was built as illustrated in the results part.

2.1.2. Metabolomics data analysis

The DMF of *Phragmites australis* assigned with the OPLS-DA and S-plots were extracted. The extracted data were returned to the original data. It is impossible to identify a pathway depend on just a mass. To get around this issue, a key concept is to shift the unit of analysis from individual compounds to individual pathways or a group of functionally related compounds. The mummichog algorithm is the first implementation of this concept to infer pathway activities from a ranked list of MS peaks identified by untarget metabolomics. The original algorithm implements an over-representation analysis (ORA) method to evaluate pathway-level enrichment based on significant features assigned with the statistical analysis. Users need to specify a pre-defined cutoff based on *p*-values. For further details about the original implementation, please refer to Li *et al*. 2013. The mass accuracy was set to 5 ppm on the positive mode. The *p*-value cutoff was assigned to 0.05 to delineate between significantly enriched and non-significantly enriched features.

Table S1. The standards compounds of the quality control external calibration mixture, monoisotopic mass in the literature (L), monoisotopic in different injection and the mean of them, the variation between monoisotopic mass in the literature (L), and the mean of measured monoisotopic mass, and mean mass standard deviation (SD) are listed.

Name	Mono isotopic Mass (Da) (L)	Mono isotopic Mass (Da) (1st inj.)	Mono isotopic Mass (Da) (2nd inj.)	Mono isotopic Mass (Da) (3rd inj.)	Mono isotopic Mass (Da) (4th inj.)	Mono isotopic Mass (Da) (5th inj.)	Mono isotopic Mass (Da) (6th inj.)	Mean Mono isotopic Mass (Da)	Δ ppm	SD
Metformin	129.1014	129.1006	129.1001	129.1000	129.1005	129.1015	129.1014	129.1007	5.78	0.001
Glyphosat	169.0140	169.0132	169.0134	169.0138	169.0123	169.0120	169.0136	169.0131	5.62	0.001
Gabapentin	171.1259	171.1252	171.1254	171.1254	171.1231	171.1231	171.1253	171.1246	7.81	0.001
Monuron	198.0560	198.0566	198.0568	198.0569	198.0563	198.0564	198.0564	198.0566	-2.96	0
Chloridazon	221.0356	221.0359	221.0362	221.0361	221.0355	221.0359	221.0355	221.0359	-1.22	0
Carbetamid	236.1161	236.1176	236.1179	236.1180	236.1164	236.1169	236.1169	236.1173	-5.1	0.001
Metobromuron	258.0004	258.0012	258.0014	258.0010	258.0006	258.0025	258.0060	258.0021	-6.54	0.002
Sotalol	272.1195	272.1193	272.1193	272.1192	272.1192	272.1211	272.1195	272.1196	-0.55	0.001
Chlorbromuron	291.9615	291.9616	291.9614	291.9613	291.9613	291.9614	291.9614	291.9614	0.21	0
Diazinon	304.1010	304.1010	304.1019	304.1014	304.1004	304.1001	304.1022	304.1012	-0.42	0.001
Quinoxyfen	306.9967	306.9961	306.9961	306.9976	306.9948	306.9964	306.9967	306.9963	1.32	0.001
Metconazol	319.1451	319.1456	319.1452	319.1443	319.1452	319.1462	319.1451	319.1453	-0.46	0.001
Fenofibrat	360.1128	360.1135	360.1141	360.1137	360.1132	360.1141	360.1139	360.1138	-2.58	0

Table S2. The standards compounds of the quality control external calibration mixture, the single RT, mean RT of the different injections, mean RT standard deviation (SD), and relative standard deviation (RSD) of the standards are listed.

Name	RT (Min) (1st inj.)	RT (Min) (2nd inj.)	RT (Min) (3rd inj.)	RT (Min) (4th inj.)	RT (Min) (5th inj.)	RT (Min) (6th inj.)	Mean RT (Min)	SD	% RSD
Metformin	14.33	14.27	14.07	14.30	14.30	14.27	14.26	0.09	0.7
Glyphosat	13.73	13.72	14.39	13.85	14.03	13.81	13.92	0.25	1.8
Gabapentin	7.64	7.52	7.48	7.28	7.85	7.58	7.56	0.19	2.5
Monuron	24.32	24.22	24.25	24.59	24.59	24.23	24.37	0.17	0.7
Chloridazon	22.00	21.93	21.99	21.45	21.36	21.85	21.76	0.28	1.3
Carbetamid	23.90	23.84	23.84	22.66	22.65	23.85	23.46	0.62	2.6
Metobromuron	26.27	26.16	26.21	25.67	25.92	26.05	26.04	0.22	0.9
Sotalol	9.60	9.23	9.11	9.19	9.80	9.90	9.47	0.34	3.6
Chlorbromuron	26.92	27.61	27.66	27.41	27.72	28.00	27.55	0.37	1.3
Diazinon	33.45	33.49	33.44	33.07	33.38	34.33	33.53	0.42	1.3
Quinoxyfen	35.17	35.54	35.20	35.17	34.52	34.52	35.02	0.41	1.2
Metconazol	28.64	28.59	28.57	28.65	28.49	28.28	28.54	0.14	0.5
Fenofibrat	32.90	33.11	32.87	33.16	33.19	33.38	33.10	0.19	0.6

Table S3. List of carbamazepine (CBZ) transformed product identified in *Phragmites australis* different samples with the mean monoisotopic mass in the standards (S), the mean monoisotopic mass of *Phragmites australis* (Ph), the variation between them, mean RT of standards (S), mean RT of *Phragmites australis* (Ph), and the variation between them were listed. The logD values were predicated from ChemAxon software (https://disco.chemaxon.com/apps/demos/logd/)

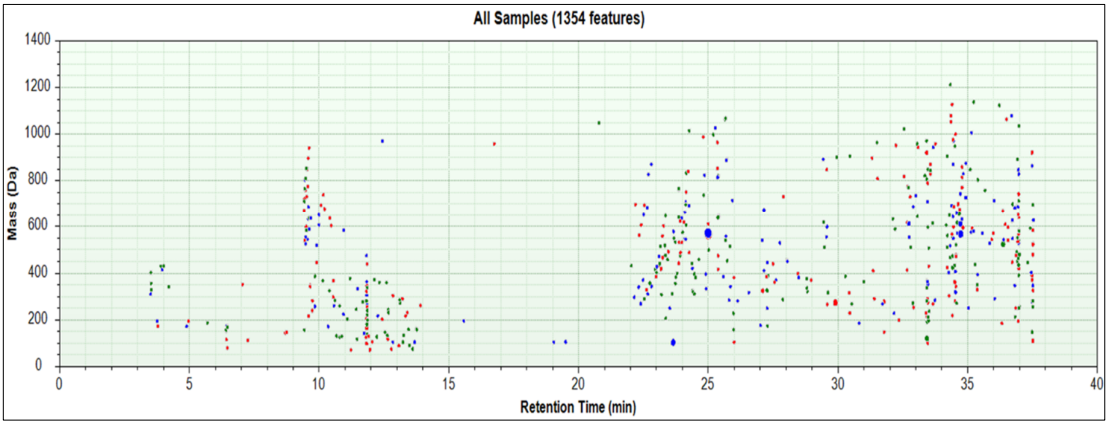
CBZ Transformed Products	Mean Mono isotopic Mass (Da) (S)	Mean Mono isotopic Mass (Da) (Ph)	Δ ppm	Mean RT (Min) (S)	Mean RT (Min) (Ph)	Δ RT	LogD (pH=7.4)
Carbamazepine-10,11-epoxide	252.0903	252.0891	4.81	23.57	23.39	0.18	1.97
10,11-Dihydro-10,11-dihydroxy-carbamazepine	270.10	270.0994	2.16	22.34	22.12	0.22	0.81
10,11-Dihydro-10-hydroxy-carbamazepine	254.1055	254.1044	4.33	23.06	22.93	0.13	1.73
9-Acridine carboxaldehyde	207.0684	207.0682	1.10	33.46	33.27	-0.19	2.98
2,3-Dihydro-2,3-dihydroxy-carbamazepine	270.1	270.0994	2.16	5.37	5.20	0.17	-0.13

Table S4. The differentiating metabolic profile (DMF) metabolites of *Phragmites australis* due to incubation with 10 & 100 µM diclofenac were extracted from the S-plot, monoisotopic mass, *p*-value, and t. score are listed

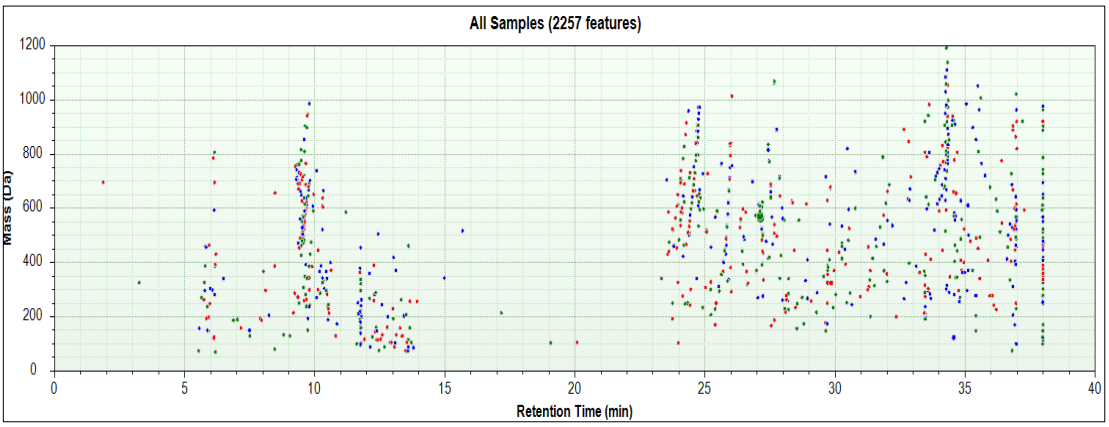
Monoisotopic mass	*p*-value	t.score	Monoisotopic mass	*p*-value	t.score
68.0234	0.030106	0.23683	155.0692	0.015114	0.237591
68.0264	0.022125	0.166365	156.006	0.068255	0.18542
69.0214	0.039998	0.289065	159.0894	0.021152	0.171257
69.0577	0.019497	0.164378	162.0533	0.021843	0.153318
69.0586	0.024319	0.219215	170.0186	0.024278	0.200248
74.034	0.017752	0.260474	180.0193	0.015685	0.262302
78.9832	0.032256	0.307147	180.021	0.029023	0.236323
84.0216	0.041884	0.125878	180.021	0.022126	0.224066
84.0216	0.068541	0.213997	180.0215	0.027404	0.222718
84.0216	0.021316	0.091268	180.0215	0.01918	0.170941
84.0219	0.030039	0.250664	194.115	0.021004	0.155939
87.23	0.095546	0.279915	197.0906	0.032444	0.17429
87.23	0.028594	0.151951	197.0906	0.029217	0.160687
89.0477	0.030512	0.191343	202.1285	0.023499	0.194255
89.0477	0.022961	0.212673	208.9804	0.033769	0.18385
97.9683	0.027681	0.239861	208.9804	0.026954	0.166529
101.05	0.020528	0.258464	214.0096	0.015891	0.198022
103.9845	0.037352	0.18452	214.0096	0.017781	0.175736
103.9848	0.021889	0.127428	219.1034	0.016703	0.183509
104.0351	0.022192	0.200123	240.1469	0.022109	0.147296
105.0393	0.025857	0.213747	241.1054	0.018804	0.27248
105.0429	0.05	0.259898	255.0755	0.029164	0.126878
105.0429	0.019939	0.121926	255.1678	0.020379	0.150431
105.0776	0.022807	0.233001	255.1678	0.040065	0.227302
105.0788	0.027732	0.20118	270.0712	0.059532	0.185823
110.0845	0.026338	0.168961	270.0723	0.056079	0.178862
115.0287	0.022279	0.189553	270.9991	0.019459	0.2047
115.0635	0.050546	0.199762	273.1539	0.052786	0.207747
117.0782	0.042451	0.154643	273.156	0.030913	0.140906
117.0782	0.035544	0.124902	273.156	0.037379	0.114409
117.0782	0.06642	0.192607	273.156	0.022518	0.080971
117.0792	0.034949	0.149076	273.1788	0.015009	0.205866
117.1149	0.020489	0.152183	276.0773	0.030484	0.140702
118.0413	0.022657	0.106052	282.1656	0.016142	0.259492
120.0435	0.039739	0.206169	303.1539	0.032277	0.224316
120.0435	0.078967	0.351932	303.1546	0.044533	0.23808
120.0436	0.021475	0.203871	308.1591	0.032854	0.187574
120.0436	0.016406	0.226417	320.0991	0.016778	0.204282
126.0411	0.023179	0.132794	320.0991	0.013498	0.240874
129.0783	0.018318	0.188127	320.0991	0.016778	0.204282
131.0947	0.027175	0.102577	320.0991	0.013498	0.240874
131.0951	0.045885	0.118508	368.1104	0.030716	0.121481
131.0951	0.022295	0.075155	368.1104	0.030716	0.121481
137.0814	0.025135	0.148162	373.1014	0.021469	0.268151
144.0367	0.079941	0.309844	373.1014	0.021469	0.268151
144.0379	0.018853	0.13714	532.6444	0.034617	0.174168
144.0391	0.023973	0.101013	532.6444	0.034617	0.174168
147.9867	0.025314	0.205985	607.2324	0.043548	0.207635
155.0686	0.016005	0.182058	607.2324	0.043548	0.207635

Table S5. The differentiating metabolic profile (DMF) metabolites of *Phragmites australis* due to incubation with 10 & 50 µM carbamazepine were extracted from the S-plot, monoisotopic mass, *p*-value, and t.score are listed

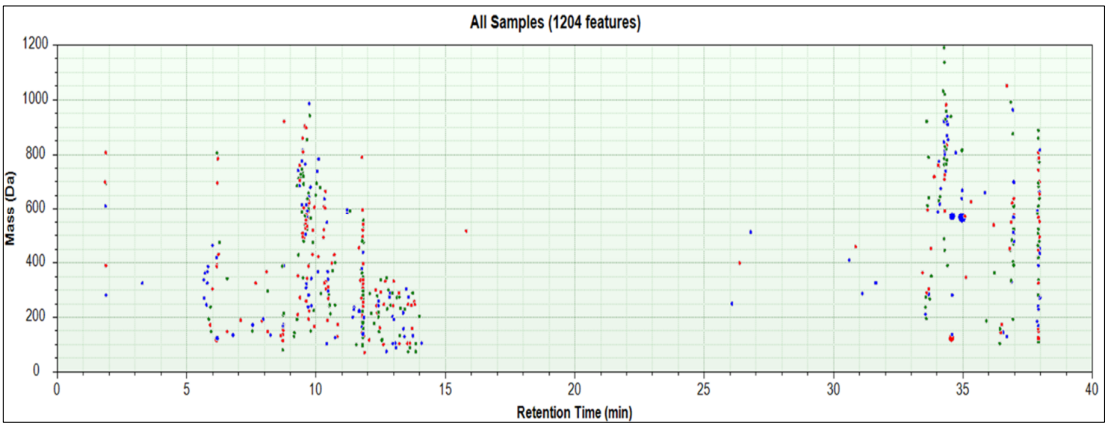
Monoisotopic mass	*p*-value	t.score	Monoisotopic mass	*p*-value	t.score
70.0536	0.013284	0.194421	186.1002	0.023673	0.249651
79.9409	0.061161	0.164513	190.1096	0.025954	0.214713
83.0374	0.010699	0.146955	201.1732	0.015503	0.10034
84.0215	0.061734	0.23701	206.1515	0.006323	0.136052
87.0323	0.059627	0.204729	215.1882	0.004611	0.212767
89.0479	0.01228	0.145846	216.0968	0.00502	0.173375
89.048	0.00947	0.189974	217.9675	0.002192	0.167609
89.0482	0.050044	0.23896	219.1104	0.067305	0.167342
103.0993	0.023768	0.123503	227.0906	0.012969	0.208512
103.9847	0.017263	0.148329	227.0907	0.076145	0.287427
103.9849	0.045547	0.162259	227.1889	0.004641	0.197283
103.985	0.006327	0.238841	241.2396	0.029797	0.211693
103.9851	0.056276	0.212792	241.2767	0.050552	0.292318
105.0423	0.08038	0.287973	244.1779	0.019101	0.336869
105.0427	0.08038	0.287973	249.1216	0.026224	0.27147
105.0787	0.007499	0.275313	251.0576	0.023621	0.238833
111.0433	0.029081	0.144429	255.0234	0.020722	0.247003
113.0855	0.035166	0.161382	287.2807	0.026274	0.346846
114.9497	0.019858	0.235099	289.0658	0.038021	0.151319
115.0266	0.03549	0.149447	294.0162	0.017712	0.346654
115.0266	0.038415	0.171592	295.2506	0.020323	0.310185
115.0633	0.022501	0.18596	301.1896	0.048709	0.278219
117.0785	0.041413	0.328885	302.1923	0.02723	0.325911
119.9588	0.033128	0.180568	303.2487	0.024313	0.397787
119.9589	0.007138	0.132	330.075	0.020419	0.266759
120.0436	0.037921	0.199761	330.1892	0.037219	0.230546
120.0436	0.03734	0.156266	343.2363	0.023681	0.241985
120.0436	0.034371	0.19167	344.1459	0.017875	0.301713
120.0436	0.037921	0.199761	352.0677	0.031794	0.327461
125.0143	0.0077	0.214742	364.0989	0.035662	0.201401
126.0654	0.008069	0.204016	382.1076	0.048453	0.203364
129.0425	0.053529	0.362021	382.147	0.059465	0.191069
129.0426	0.053529	0.362021	384.1757	0.053612	0.161827
137.0783	0.019625	0.203304	384.1787	0.062787	0.154608
138.9917	0.008553	0.133195	384.1787	0.066648	0.157788
147.9796	0.024155	0.218384	416.2029	0.023989	0.224429
147.9796	0.026522	0.232374	419.2163	0.030838	0.288205
147.9854	0.024155	0.218384	479.4187	0.020506	0.260596
157.1463	0.008491	0.21146	481.4001	0.026541	0.331124
167.9698	0.009256	0.132112	512.1178	0.030593	0.267982
171.1086	0.007084	0.196078	606.9501	0.05694	0.262249
180.0167	0.021764	0.191732	869.6639	0.069832	0.207301
186.099	0.023673	0.249651			


All Samples (1354 features)
Mass (Da)
Retention Time (min)

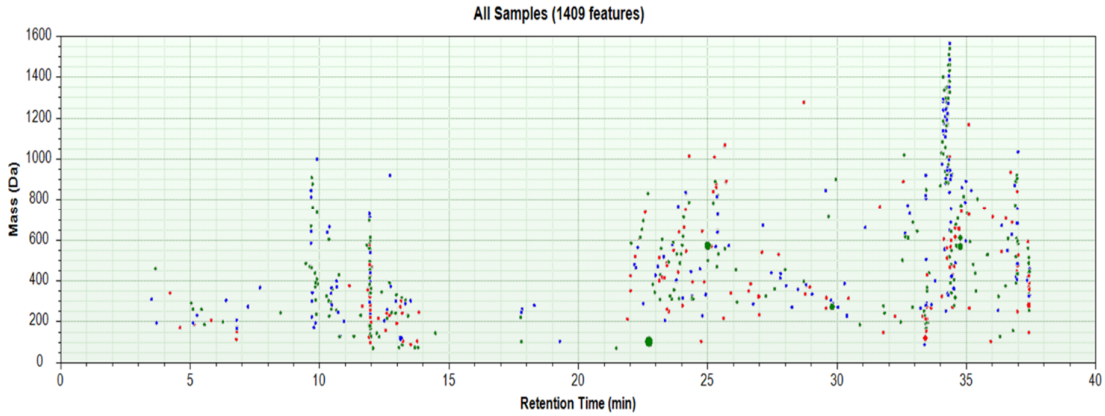
(a)


All Samples (2257 features)
Mass (Da)
Retention Time (min)

(b)


All Samples (1204 features)
Mass (Da)
Retention Time (min)

(c)



(d)

Figure S1. Retention time (RT)/Mass plot of the background was analyzed by RPLC-HILIC-ESI-TOF-MS in positive electrospray ionization mode. (a) 100% Methanol, (b)Acidic 90% methanol, (c) 50% Methanol, and (d) 100% Aqueous.

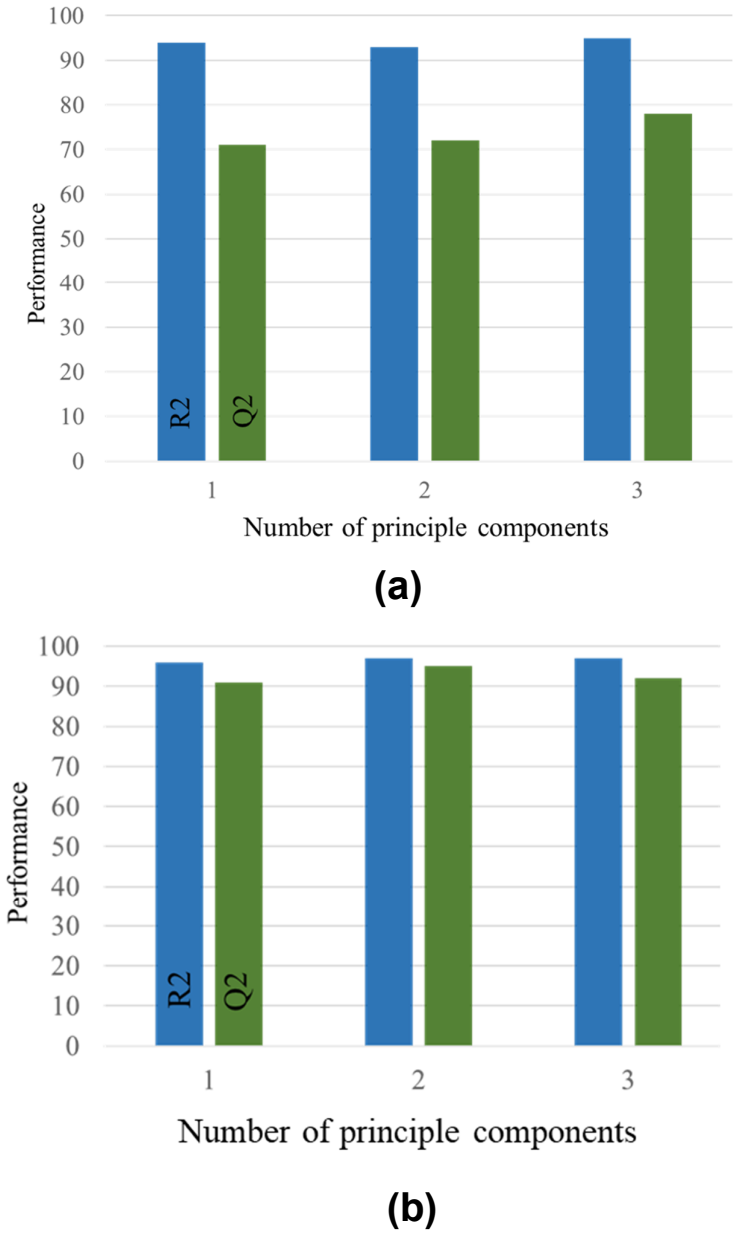
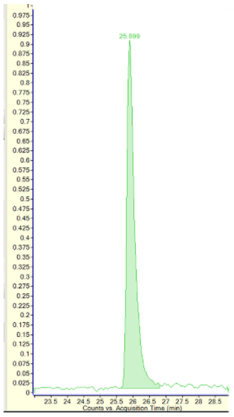
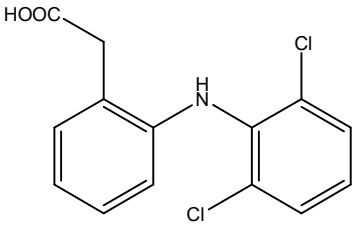


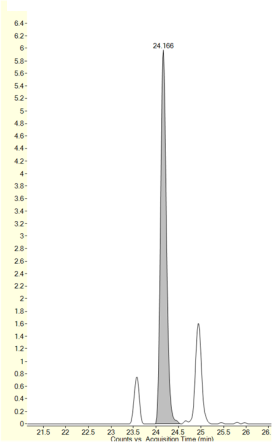
Figure S2. The Q^2/ R^2 Overview plot displays the individual cumulative R^2 (green columns) and Q^2 (blue columns) and Q^2 for the goodness of fits and cross-validation parameters (a) *P. australis* different parts. (b) *P. australis* different incubation

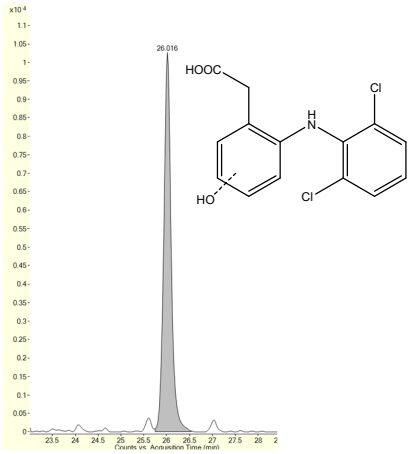
Diclofenac reference standards

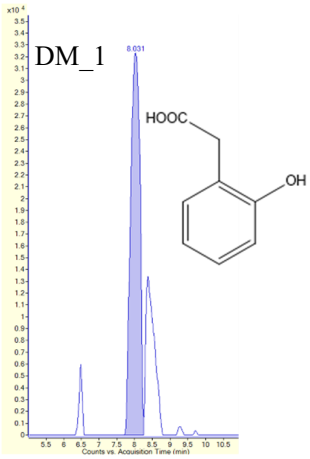


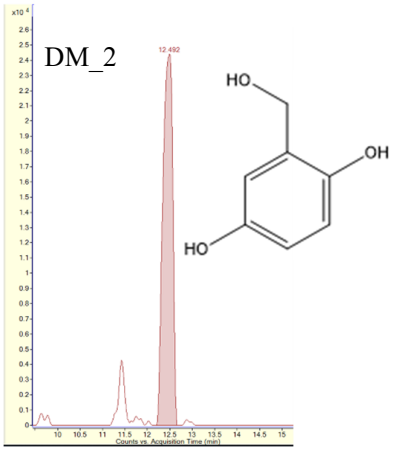


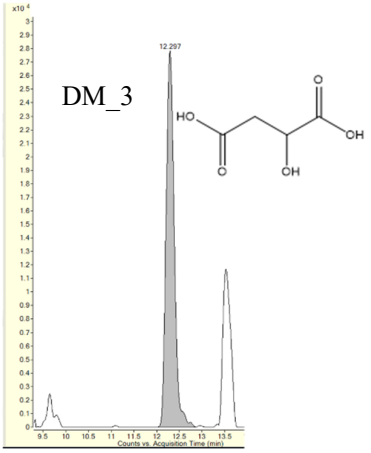
Diclofenac

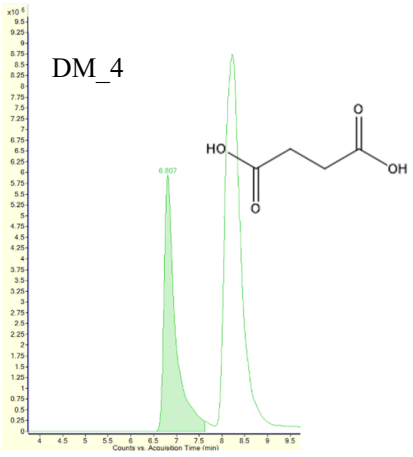


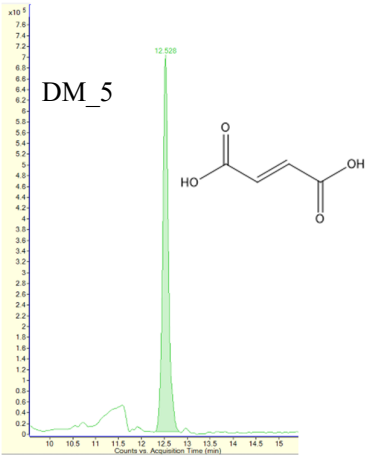












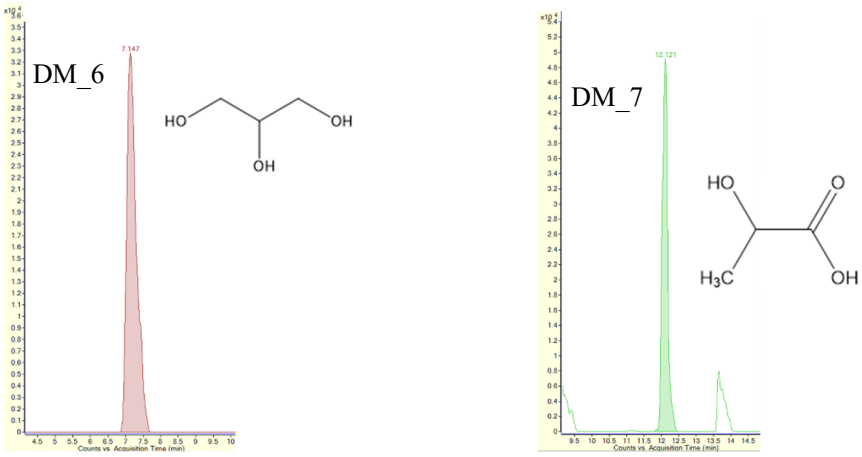
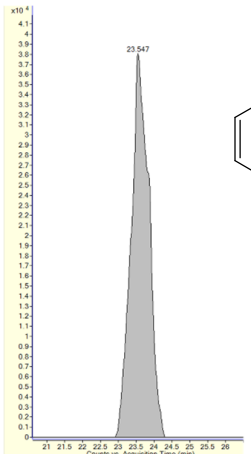
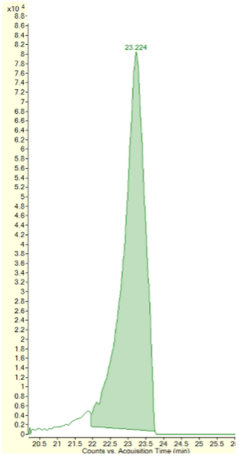
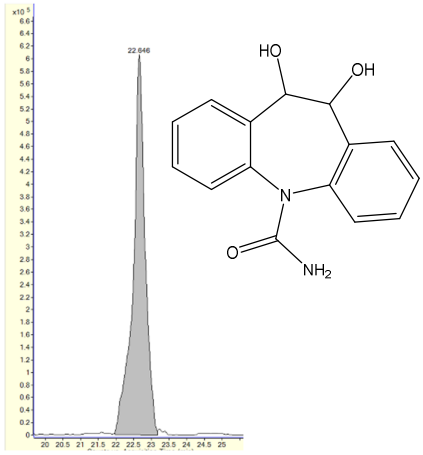
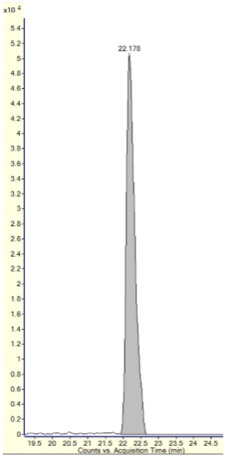
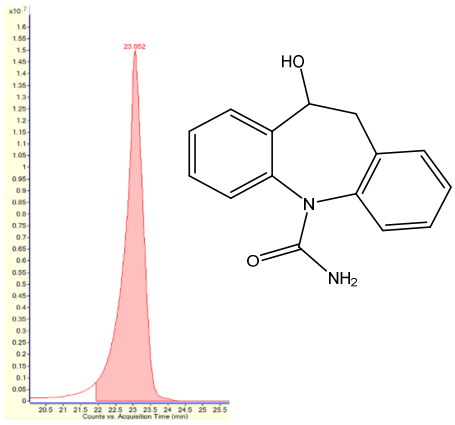
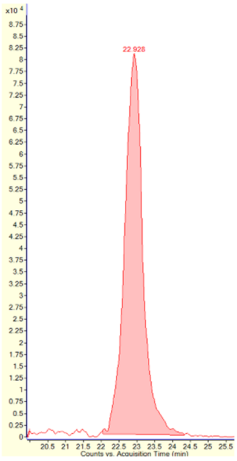


Figure S3. EICs were corresponding to measured diclofenac (right) and the reference standard (left), which were identified in the extracts of *Phragmites australis* leaf, rhizome, and roots incubated with 10 and 100 µM diclofenac. Also, EICs relative to transformed products are suspected in the extracts of *Phragmites australis* leaf, rhizome, and roots incubated with 10 and 100 µM diclofenac.

	Standard reference EICs	*Phragmites* EICs
Carbamazepine-10,11-epoxide	Mass= 252.0903 Da	Mass= 252.0891 Da
		
10,11-Dihydro-10,11-dihydroxy-carbamazepine	Mass=270.10 Da	Mass=270.0994 Da
		
10,11-Dihydro-10-hydroxy-carbamazepine	Mass= 254.1055 Da	Mass= 254.1044 Da
		

9-Acridine carboxaldehyde	Mass= 207.0684 Da	Mass=207.0682
2,3-Dihydro-2,3-dihydroxy-carbamazepine	Mass= 270.1 Da	Mass= 270.0994 Da

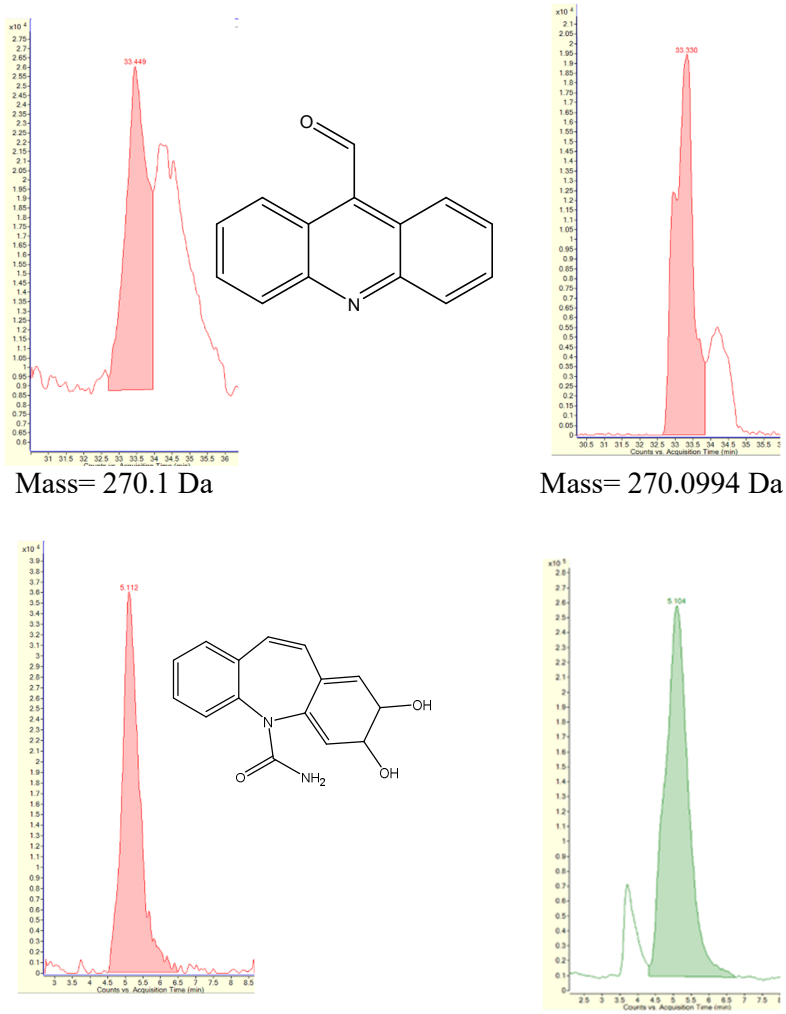


Figure S4. EICs were corresponding to carbamazepine (CBZ) and its transformed product standards (left), which were identified in the extracts of *Phragmites australis* leaf, rhizome, and roots incubated with 10 and 50µM carbamazepine (measured right).

metabolites

MDPI

(Article)

Untargeted analysis of *Lemna minor* metabolites: Workflow and prioritization strategy comparing highly confident features between different mass spectrometers

Rofida Wahman[1, 2], Stefan Moser [3], Stefan Bieber [4], Catarina Cruzeiro [5], Peter Schröder [5], August Gilg[6], Frank Lesske[6], and Thomas Letzel [1, 4*]

[1] Chair of Urban Water Systems Engineering, Technical University of Munich, Am Coulombwall 3, 85748 Garching, Germany.
[2] Pharmacognosy Department, Faculty of Pharmacy, Assiut University, Assiut, Egypt.
[3] Stefan Moser Process Optimization, Weberweg 3, D-83131 Nußdorf am Inn, Germany
[4] Analytisches Forschungsinstitut für Non-Target Screening (AFIN-TS GmbH), Am Mittleren Moos 48, D - 86167 Augsburg, Germany.
[5] German Research Center for Environmental Health, Research Unit Comparative Microbiome Analysis, Helmholtz Centrum Munich, Ingolstadt street 1, 85764 Neuherberg, Germany
[6] University of Applied Sciences, Am Hofgarten 4, 85354 Freising, Weihenstephan, Germany
* Correspondence: t.letzel@afin-ts.de

Abstract: The metabolomics approach provides a vast array of analytical data sets, which require a comprehensive analytical, statistical, and biochemical workflow revealing changes in metabolic profiles. The biological interpretation of mass spectrometric metabolomics results is still obstructed by the reliable identification of the metabolites as well as annotation and/or classification. *Lemna minor* was extracted using various solvents and analyzed utilizing polarity-extended liquid chromatography (RPLC-HILIC) connected to two mass spectrometers, individually. This paper introduces three topics for a novel untargeted workflow:

A prioritization tool using the significant features in the FOR-IDENT platform with the compound database PLANT-IDENT (filled with relevant compounds of Lemnaceae, Poaceae, Brassicaceae, Nymphaceae families) according to different criteria such as accurate mass (i.e. empirical formula, retention time (i.e. polarity and LogD (pH 7)), and mass spectrometer signal intensity. Finally, untargeted analysis was performed using the compound database PLANT-IDENT in the FOR-IDENT platform as a prioritization and identification source. Consequently, forty-four (amino acids, vitamins, flavonoids, and flavnoidal glycosides) compounds could be validated in *Lemna* metabolic profile using reference standards. The flavonoids and flavonoidal glycosides, which have an essential antioxidant activity for plants and humans. Also, robinetin and norwogonin were for the first time identified in the *Lemna minor* extracts according to our knowledge.

- A statistical procedure to prioritize the relevant features (dependent and independent of the mass spectrometer using predictive methodology Orthogonal Partial Least Squares-Discriminant Analysis (OPLS-DA)).
- A comparison study of metabolome samples of the same plant material origin in different mass spectrometers.

Keywords: metabolomics; extended-polarity chromatography; *Lemna minor*; annotation; PLANT-IDENT, FOR-IDENT, untargeted screening workflow.

1. Introduction

The metabolomics approach aims to contact the identification of metabolites and changes in their concentrations. The modern development of analytical techniques expanded the use of metabolomics in biological systems investigations. This approach permits remarkable insights into regulation mechanisms as well as studying responses to different perturbations. Over the centuries, target metabolomics analysis has become a routine application in different files. In untargeted metabolomics studies, the identification of the metabolites is the fundamental step to transform the analytical data into biological knowledge, which is still considered the major bottleneck. The number of identified metabolites in untargeted metabolomics studies is in general below 50%. In LC-HRMS-based approaches, metabolites can be identified if the retention time (RT), the ion mass, MS/MS fragment spectra are successfully matched with an authentic reference standard measured at the same instrument due to the presence of over 200,000 to 1,000,000 different metabolites in the plant kingdom. Thus, the identification of specialized metabolites is still difficult due to the unavailability of the reference standard, which has been demonstrated to have 'specialized' roles for adaptive significance in protection against predator and microbial infection [1, 2]. Several MS/MS databases have been established to facilitate metabolite annotation, such as MassBank, METLIN (http://metlin.scripps.edu/index.php), LipidBlast (https://fiehnlab.ucdavis.edu/projects/lipidblast), and ReSpect (http://spectra.psc.riken.jp). Moreover, plant metabolomics databases have been developed such as KNApSAcK (http://kanaya.naist.jp/KNApSAcK/), MetaCyc (http://metacyc.org), PlantMetabolomics.org (http://www.plantmetabolomics.org), KEGG (http://www.genome.jp/kegg/), and PRIMe (http://prime.psc.riken.jp/). The choice of databases can have an essential influence on the study's interpretation. The small, well-defined database can suggest more meaningful hits compared to searches in huge ones, which is containing several (hundred-) thousand compounds from many different fields that are not related to the study scientific and biological question [1-3]. PLANT-IDENT(https://water.for-ident.org/#!home) is a specific database for the metabolites of the Lemnaceae family in addition to Poaceae, Brassicaceae, Nymphaeaceae, and, compounds from different families, which is organized for rapid search retrieval by a computer in FOR-IDENT platform. The FOR-IDENT platform contains different parameters retention time (RT), mass, and MS/MS in silico fragmentation tool, in addition to retention time index (RTI), logD, and LogP values. Hence, a large number of the peak were annotated, which could enhance the biomarkers and indicators identification in untargeted metabolomics analysis. Usually, the first step toward metabolites identification is the separation technique, which provides a reproducible, precise, and wide range of polarity separation. A serial coupling of two columns with different polarities secured the separation of different compounds from the complex matrix (e.g., unfractionated plant extracts). Serial coupling of hydrophilic interaction liquid chromatography (HILIC) coupled to reversed-phase liquid chromatography (RPLC) was hyphenated with a highly accurate high-resolution mass spectrometer (HRMS). This serial coupling-QTOF provides a large amount of data, consequently adequate workflow to process and analyze it. QTOF-mass spectrometers are the most widely used mass spectrometer in metabolomics because they offer the highest scan rates, high efficiency, resolution, and mass accuracy [4]. Although metabolomics data processing is a straightforward process, the metabolite identification process continues to be problematic. From peak picking to identify the metabolites, the workflow contains numerous steps, which require to be precise, compatible, and with less effort and time-consuming. All the steps required quality control procedures for trustworthy metabolomics analysis out-

comes. The workflow has several preprocessing steps, starting from peak picking and signal to noise threshold detection. Following by deleting the background and alignments of different data sets. The workflow ends with the automated metabolite identification algorithms and the biological interpretation of the data [5, 6].

The most common proposed workflow applies to the food and nutrition sciences. These fields adopt metabolomics- especially the untargeted strategy -as an analytical tool for decades. Also, medical research [7] sciences metabolomics reflecting new insights that are correlated with other clinical variables [8]. New techniques were revealed on the molecular level between drugs and the human body. Besides, metabolomics was considered in environmental sciences as a technique to evaluate the physiological responses of the drug action, toxic effects, and metabolic disorders on the organisms from the initial chemical interaction to the final adverse outcome (adverse outcome pathways (AOP)) [8]. *Lemna minor* is known as duckweed, which is tiny, free-floating, aquatic green. Fast growth, microbial reduction, and high nutrient and metal accumulation potential are the factors that candidate *Lemna* for phytoremediation research. It belongs to the family Lemnaceae with other 4 aquatic genera *[9]*. The family members contain various chemical constituents such as lipid, protein, amino acids, organic acids, sterol, terpenes, and flavonoids [10, 11].

Also, the second potential problem of metabolomics is the difficulty of comparing the results obtained in different laboratories, generated on different mass spectrometers, and with different methods of data processing. Also, the second potential problem of metabolomics is the difficulty of comparing the results obtained in different laboratories, generated on different mass spectrometers, and with different methods of data processing. Therefore, the marker identified in a plant in one laboratory cannot be replicated, on the same and/or a different plant due to the differences in the instrumentation used rather than in the samples themselves [4].

This study divide into three parts, which aimed to

a)-prove the suspects and untargeted RPLC-HILIC-MS/MS screening strategic relevance of plant samples using the PLANT IDENT database in the FOR-IDENT platform.

b)- Further, the plant metabolites identification workflow was described from the theoretical predictions to the final analyses in *Lemna* metabolic profile using reference materials. The usage of the open-source PLANT-IDENT database to identify the different metabolites represents an addition to plant metabolomics. A PLANT-IDENT database contains plant metabolites that will encourage untargeted screening in plant metabolomics.

c)- Furthermore, the metabolomics data was analyzed with predictive methodology Orthogonal Partial Least Squares-Discriminant Analysis (OPLS-DA) between different mass spectrometers. The discrimination method was adopted to validate the features, that were extracted with the workflow. Consequently, the standard statistical methods of metabolomics data investigation were used in the identification of relevant variables (i.e. conditional attributes to each solvent), which related to the discrimination analysis. Three different extracts were systematically analyzed with the workflow.

2. Results and Discussion

Lemna was extracted with different solvents (100% MeOH, 50% MeOH, 90% MeOH (+0.5% formic acid) and 100% H_2O). Each solvent was injected into RPLC-HILIC-ESI-single TOF-MS (system A) in triplicate to extract a wide polarity range of *Lemna* metabolites, which finally resulted in final data analysis and comparison of the three solvents (100% MeOH, 50% MeOH, and 100% H_2O) due to the reasons as shown in the recent publication Wahman et al [12]. Subsequently, the extracts were analyzed using a RPLC-HILIC-ESI-QTOF-MS/MS (system B) obtaining additional MSMS data. The raw data was checked for analytical quality by internal standards detection and the results are listed in (Table S1).

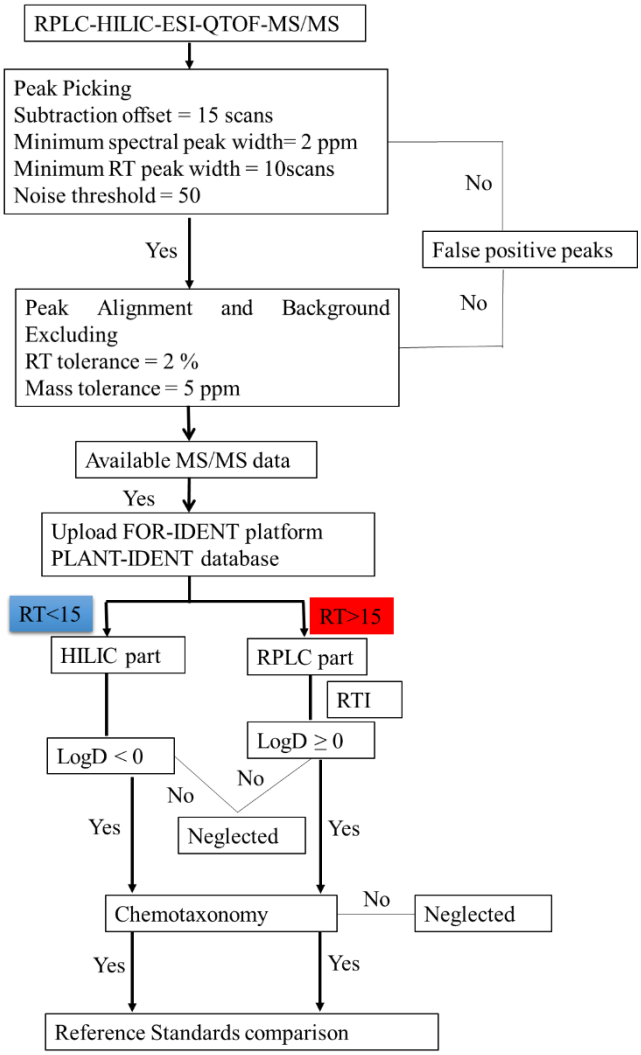
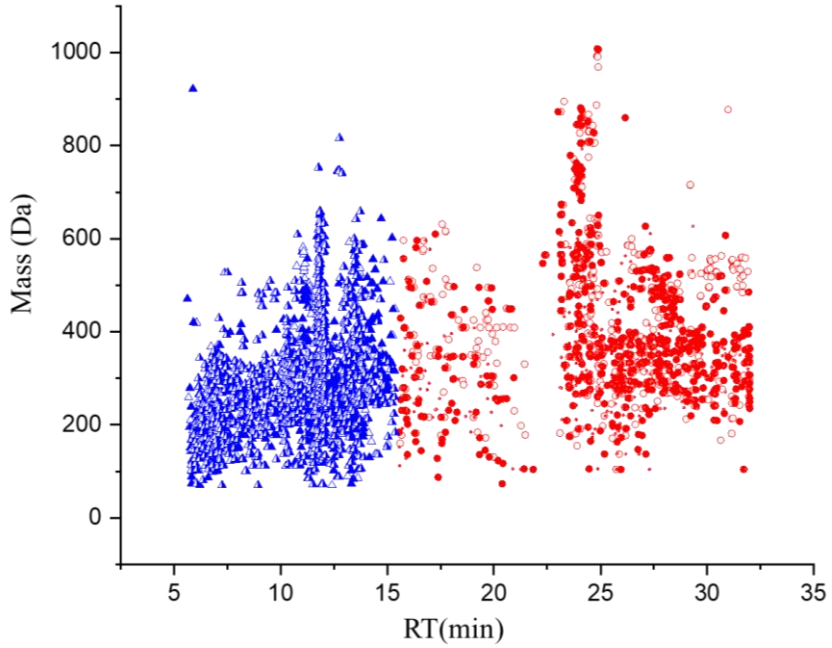


Figure 1: The non-target workflow flowchart (from precursor ion mass into names) to evaluate the MSMS data obtained with RPLC-HILIC-ESI-QTOF-MS/MS (system B). The parameters of peak picking and alignments were set according to the internal standards. The features were prioritized using the PLANT-IDENT database in the FOR-IDENT platform after different filtration steps.

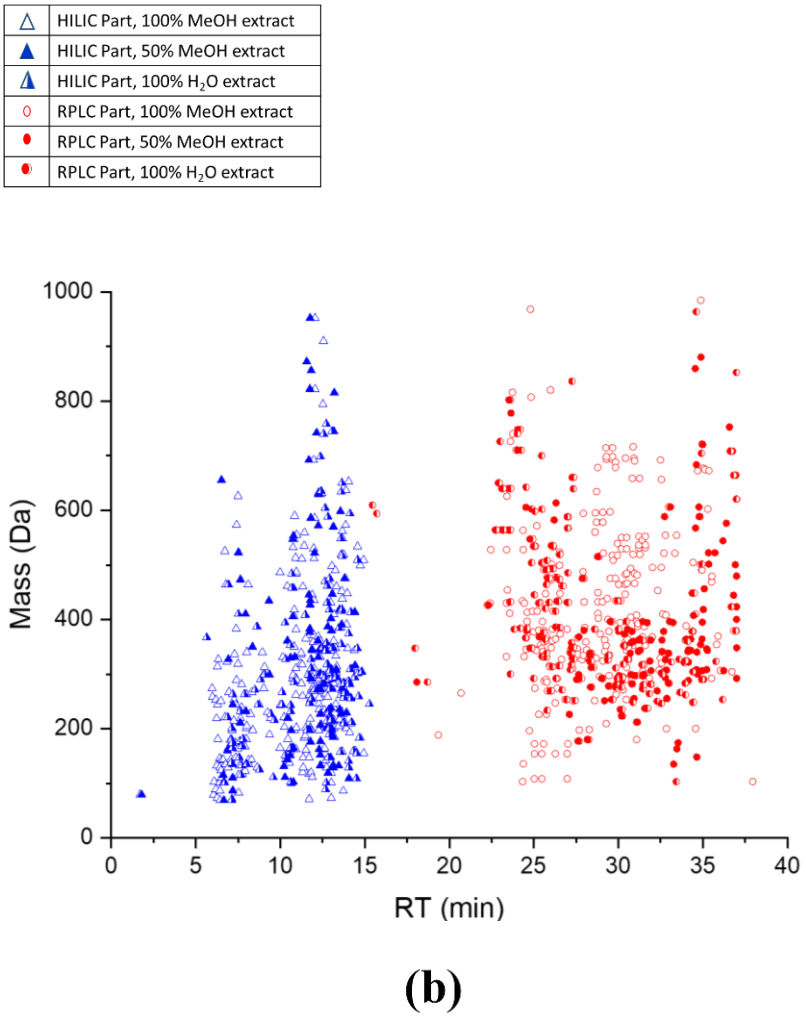
2.1. Lemna minor analysis using system A and B

In the current study, each extract was injected into the system in triplicate. The obtained data was preprocessing according to the workflow mentioned and as shown in (Figure 1). A comprehensive strategy was performed to identify metabolites in untargeted metabolomics. The extracts were analyzed with RPLC-HILIC-MS-based untargeted metabolomics analysis. Then, peak picking was performed via Sciex OS software achieving precursor ions and retention time. As described in the experimental part. Database searching and scoring were performed. Consequently, the mentioned parameters led to 1121, 1287, and 1498 features in 100% MeOH, 50% MeOH, and 100% H_2O extracts, respectively. The RT-mass plot of the obtained features in three solvents (100% MeOH, 50%MeOH, and 100% H_2O) was plotted in (Figure 2a) according to their RT (min.) and *m*/*z* (Da). The data set was separated at 15 min. The features of the *Lemna minor* metabolic profile were separated into two parts: first (5-15 min.) contained the highly polar to moderate polar features, which were eluted from the HILIC column. The second part is from 15 min. until the end of the run, which contained the nonpolar features.

Moreover, *Lemna minor* metabolic profile was obtained with the time of flight mass spectrometer (single TOF-MS) (system A) with ESI positive ionization mode (Figure 2b). The workflow of the data was mentioned in [12]. Interestingly, the data showed significant differences between the two MS (1069, 1111, 1098 features in 100% MeOH, 50% MeOH, and 100% H_2O extracts, respectively (Figure 2), which was expected due to the high sensitivity of system A compared to system B. Also, the intermediate polar part showed a fundamental difference between both systems (i.e. number of features) because this part was characterized by the start of RPLC elution and change in the gradient of the mobile phase see which were mentioned in Wahman *et al*., 2020 [12].



(a)



(b)

Figure 2: Mass-retention time (RT) plot of *Lemna minor* metabolic profile which was extracted with 100% MeOH, 50% MeOH, and 100% H_2O extracts (a) system A (QTOF); (b) system B (single TOF). The blue color represents the HILIC eluted metabolites. Also, the red color represents the RPLC eluted metabolites.

2.2. The strategy of Lemna minor metabolites identification based on the PLANT-IDENT database

Metabolites identification from known databases depend substantiality on the following: (1) comprehensive data integration and weighting of retention time, mass, and MSMS fragments, and (2) employment or ignoring of MSMS intensity information [13]. For the PLANT-IDENT database, the uploaded analytical data could be compared and transferred into physicochemical parameters like polarity, empirical formula via the FOR-IDENT platform. The prioritization depends on this transformation and the comparison with compounds and their physio-chemical properties. RT and mass are applied simultaneously as a prerequisite with MSMS data for scoring evaluation. The

determination scoring of metabolites is calculated based on the accurate finding of the features RT, mass, MSMS data extracted from samples, and PLANT-IDENT, which is achieved by using the predetermined differences values Δppm. Then, to reduce the false fragments determination, the intensity threshold is defined. The threshold differentiates the line between the signals and noise level. The stationary phase (reversed-phase (RP) and HILIC) searching is simultaneously employed for feature searching.

2.2.1. PLANT-IDENT batch searching and scoring.

a) First, the features eluted from the HILIC column with RT< 15 minutes. Features were uploaded into the FOR- IDENT platform using the PLANT-IDENT database (Figure S1). The parameters set according to the experiment at pH=7, the absolute mass deviation of the precursor ion and fragments was 5 ppm. The intensity threshold was 5 with a positive ion (H^+). The search was scored according to the mass screening, and MS/MS. Then, the results were downloaded. The PLANT-IDENT database suggested 239 candidates matching features in *Lemna's* metabolomics. All the results with LogD $\geqslant$ 0 were eliminated. In the end, the results have priority "look at" was considered and standards reference injected for validation (Figure S1). Subsequently, the filtration parameters decreased the number of the results from 239 to 41 candidates. Those metabolites were annotated and classified in the second level. They could be classified to level one by confirmation with standards references injections

b) Secondly, the RP part with RT > 15 minutes, the second part of each data set was uploaded. The retention time of the standards mixture was also uploaded to normalize the RT (Table S2). Here the scoring of suspected compounds depends on the same parameters in addition to RTI screening. Each is 25% of the total score. The 188 features were suggested as a matching candidate. After, the adjusted logD $\geqslant$ 0 candidates only were considered. This led to forty-two candidates. Those metabolites were annotated and classified in the second level. They could be classified to level one by confirmation with standards references injections

2.2.2. Identification of *Lemna minor* metabolites

Lemna's metabolic profile was interpreted according to the following workflow (from precursor ion mass into names) from system A. The annotated metabolites (level 2) were classified to level 1 as proven by standard reference injection. Each identified compound has a mass division of less than 5 ppm, and LogD < 0 in the HILIC part or Log D > 0 in the RP column part. 44 compounds were identified with the reference standards and classified into level 1 listed in Table 1. The 44 compounds have RT deviation of less than 1.0 min. (therein peonidin below 0.6 min. and 36 compounds below 0.3 min), and a mass deviation of less than 5ppm between the reference standard mass and the *Lemna minor* candidate feature mass. 16 compounds were eluted with the HILIC column: vitexin, niacin, nicotinamide, phenylalanine, Leucine/isoleucine, tryptophan, valine, tyrosine, proline, glutamic acid, aspartic acid, di-L-alanine, 4-methoxy cinnamic acid, alanine, threonine, and serine. They have LogD < 0 at (pH=7). In our previous studies, the amino acids (phenylalanine, proline, tryptophan, alanine, tyrosine, aspartic acid, isoleucine, serine,

and valine) were identified in *L. minor* using system A. The RT and mass deviations were also less than 0.3 min and 5 ppm, respectively, which were accepted [12, 14]. Thus, the results from both systems a similar in RT and mass deviations, which were accepted.

The extracted ion chromatogram, *m/z* spectrum, and MS/MS spectrum of niacin system A were shown as an example in (Figure S2). The suggested 28 compounds were separated from the RP column: apigenin-6,8-diglucopyranoside, apigenin 7-glucoside, apigenin-5-glucoside, apigenin-6-arabinopyranoside-8-glucopyranose, apigenin, robinetin, quercetin, luteolin, kaempferol, acacetin, orientin, Isoorientin, peonidin, 6-methoxy-flavon, flavon, naringenin-7-O-glucoside, quercetin-3-glucoside, saponarin, 5-hydroxy-6-methoxy-flavon, luteolin-3',7-di-O-glucoside, apiin, chrysoeriol, umbelliferone, norwogonin, isovitexin, tricin, galangin, and myricetin. Further, the MSMS fragments of the reference standard and the identified peaks in the *Lemna* metabolic profile were compared with the literature, as shown in (Table 1).

The extracted ion chromatogram, mass spectrum, and MS/MS spectrum of apiin were shown as an example in (Figure S3).

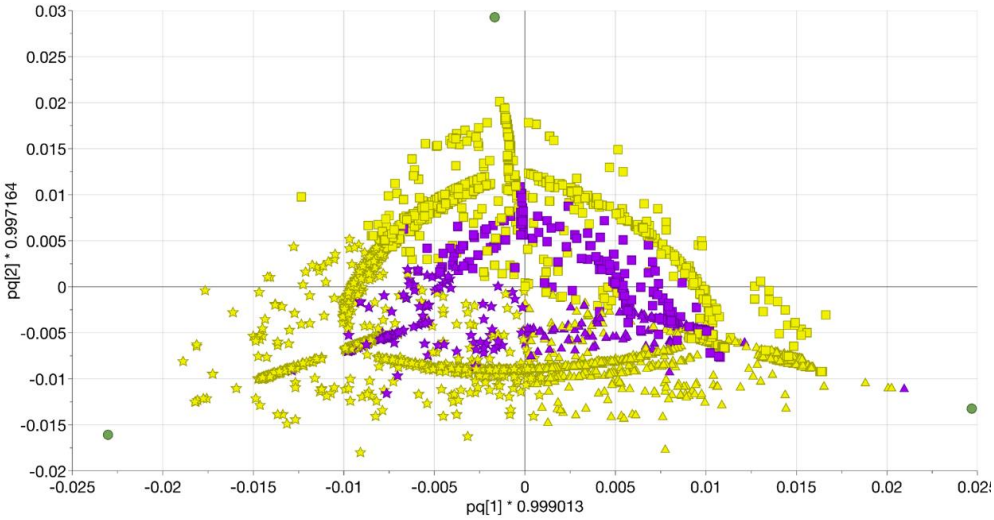
According to chemotaxonomy criteria, results, 42 compounds were separated previously from Lemnaceae or Araceae family. Further, robinetin and norwogonin were not reported from Lemnaceae or Araceae family according to the available literature. Those two compounds were not filtered with the last filter of the chemotaxonomy because they are a pentahydroxy-flavone and 5,7,8-trihydroxy-flavone, respectively. Also, luteolin, quercetin, kaempferol, apigenin, and myricetin and their glucosides have been identified in the family before [15, 16]. Besides, robinetin is known as 5-deoxy-myricetin. Therefore, robinetin and norwogonin reference standards were injected. Robinetin and norwogonin were for the first time identified in the *Lemna minor* extracts according to our knowledge.

2.5. OPLS-DA analysis of Lemna minor metabolic profile obtained with system A and B

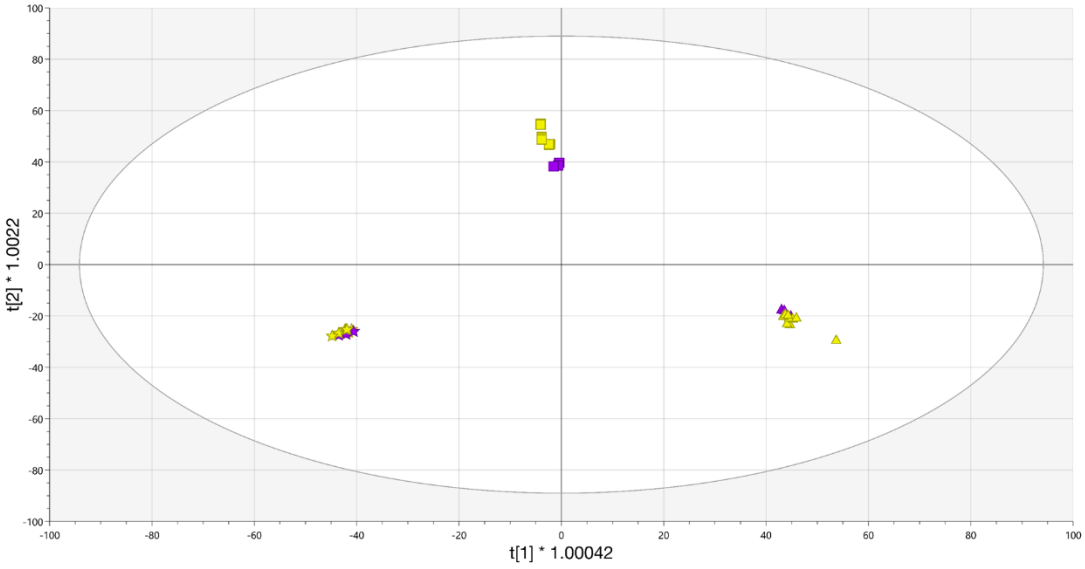
Regarding the untargeted analysis, it is fundamental to assign the common metabolic profile between the different data sets of the same organism and/or different organisms obtained from the same and/or different laboratories. The untargeted analysis concept depends mainly on the databases. Thus, the datasets were investigated statistically to show the differences and the commons between the two MS, besides, identification of *L. minor* common metabolites via the PLANT-IDENT database.

The processed data was statistically investigated to test the data quality or (workflow reliability) and information transfer between different mass spectrometric systems and/or laboratories. The OPLS-DA was built to the three extracts of three *Lemna* samples, which were analyzed with single TOF-MS (system A), and QTOF-MS/MS (system B) respectively. Each extract with three injections was plotted individually. The Score plot of OPLS-DA explained 99.2% of the variations in the various extracts (R^2Y (cum)) with a higher predictive value (Q^2 (cum) = 0.877) (Figure 3). The first component (t1) separated the 100% H_2O extracts in the negative part and the 100%MeOH and 50%MeOH extracts in the positive part. The orthogonal component (t2) distinguish between the 100%MeOH and 50%MeOH in the positive and negative part, respectively. The loading scatters plot

showed which variables (features) were expressed differently between the different extracts. The variable is responsible for the discrimination analysis. Hence, coloring the score plot according to the variable corresponding to the extracts emphasize the separation between the different extracts (Figure 3a). In the loading score, the features were colored according to their main solvent, extracting them from the *L. minor.* The features present in the main solvent, which could dissolve them in higher intensity. the main solvent for each feature is the best choice to extract them from the plant with high intensity. Furthermore, the mass spectrometers could also be distinguished. Concluding, from the statistical analysis, the *L. minor* metabolic profile has small common features between the single TOF and QTOF-MS. However, the different extracts could be significantly discriminated against in both MS, which could be used for further generations of *Lemna* metabolite measurements. Regarding the solvent used for the extraction, there is no big difference because it depends on the pKa of metabolites. Further, QTOF was enabled to detect more features than single TOF-MS due to its sensitivity.



(a)



(b)

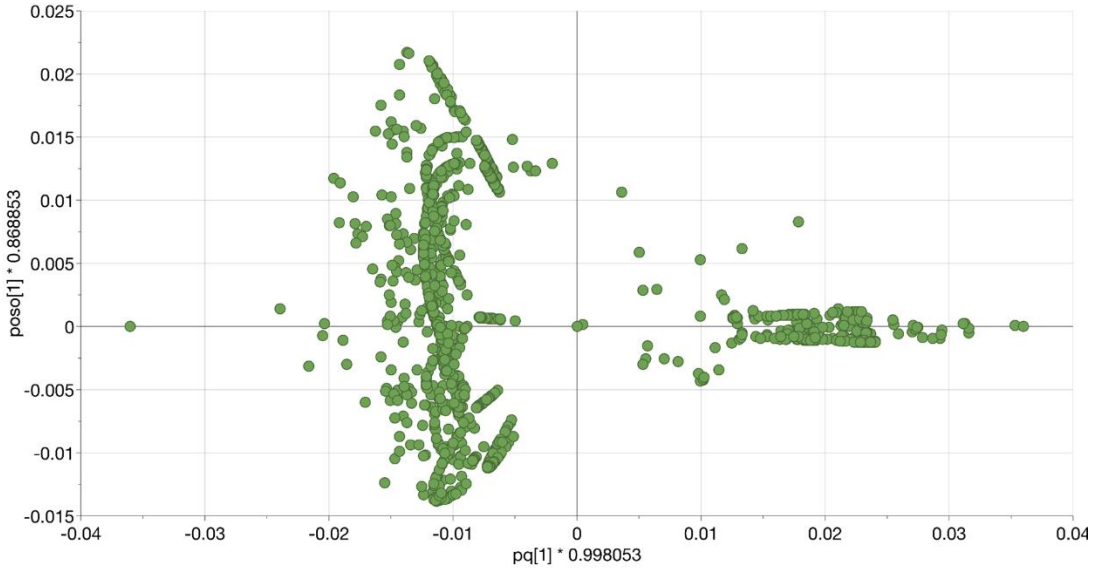
Figure 3: (a) The loading plot displays the relation between the different *Lemna minor* samples 100% MeOH (triangular), 50% MeOH (stars), and 100% H_2O (squares) extracts analyzed with TOF (violet) and QTOF (yellow); (b) OPLS-DA score scatter plot of *Lemna minor* samples. The confidence limit is 95%.

2.6. OPLS-DA analysis of Lemna minor100% MeOH and 100%H2O extracts

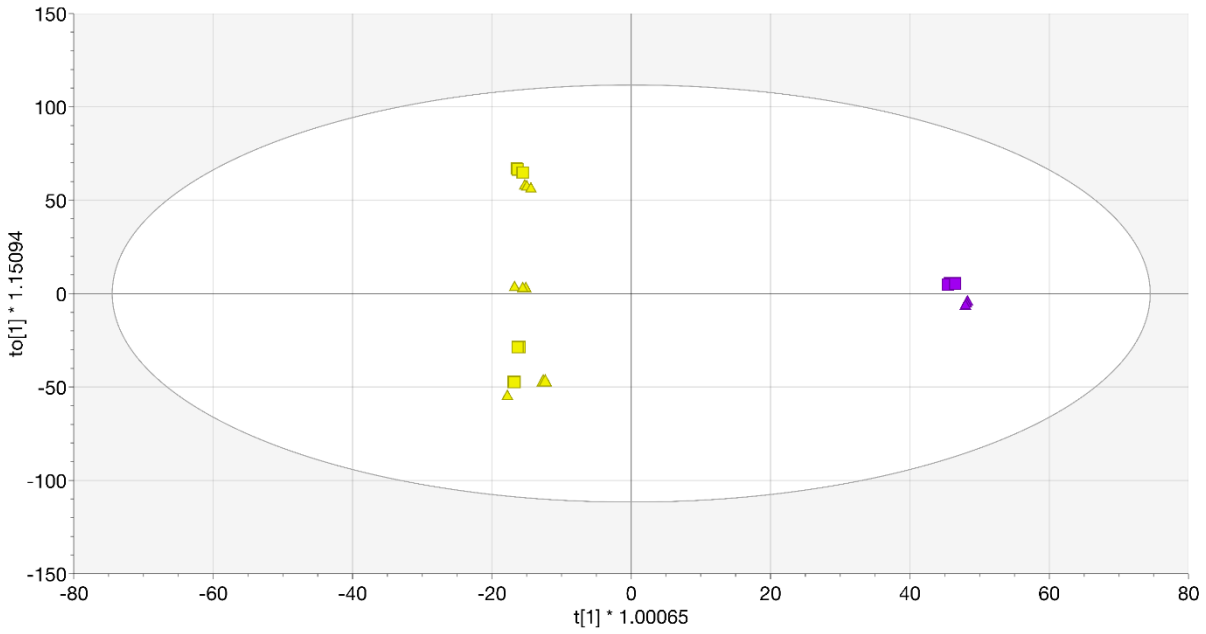
In this chapter, the data of *Lemna* (system B) was compared with obtained *Lemna* metabolic profile using single TOF-MS (system A) to validate the workflow using the PLANT-IDENT database and information transfer in between different mass spectrometric systems and/or laboratories.

The statistical comparison of the data, as a part of the untargeted screening strategy, was revealed that the two sets are significantly different. The OPLS-DA was built to the 100% MeOH and 100%H_2O extracts of *Lemna* samples, which were analyzed with a single TOF and QTOF. Each extract with three injections was plotted individually. The Score plot of OPLS-DA explained 99.8% of the variations in the various extracts (R^2Y (cum)) with a higher predictive value (Q^2 (cum) = 0.891) (Figure 4). The first components separated the different extracts according to the mass spectrometers. Further, the s-plot was built from the OPLS-DA model to identify the common metabolites, which were identified by the two machines between the two extracts. The features were with higher p(corr) values in 100% MeOH extract means that they have a higher intensity on 100% MeOH comparing to 100% H_2O extract and vice versa. Hence, the common compounds were chosen according to p(corr) values which were small approximately equal, but not equal to zero (which mean not significant to both extracts). Further, the negative values were avoided because they have less intensity in one or both extracts. The compounds were marked in red, which represented the common metabolic profile of *Lemna minor* in both solvents (systems A and B) (Figure 4c) and listed in Table S3. The OPLS-DA loading score showed that QTOF-MS features number was higher than the single TOF-MS feature due to the high

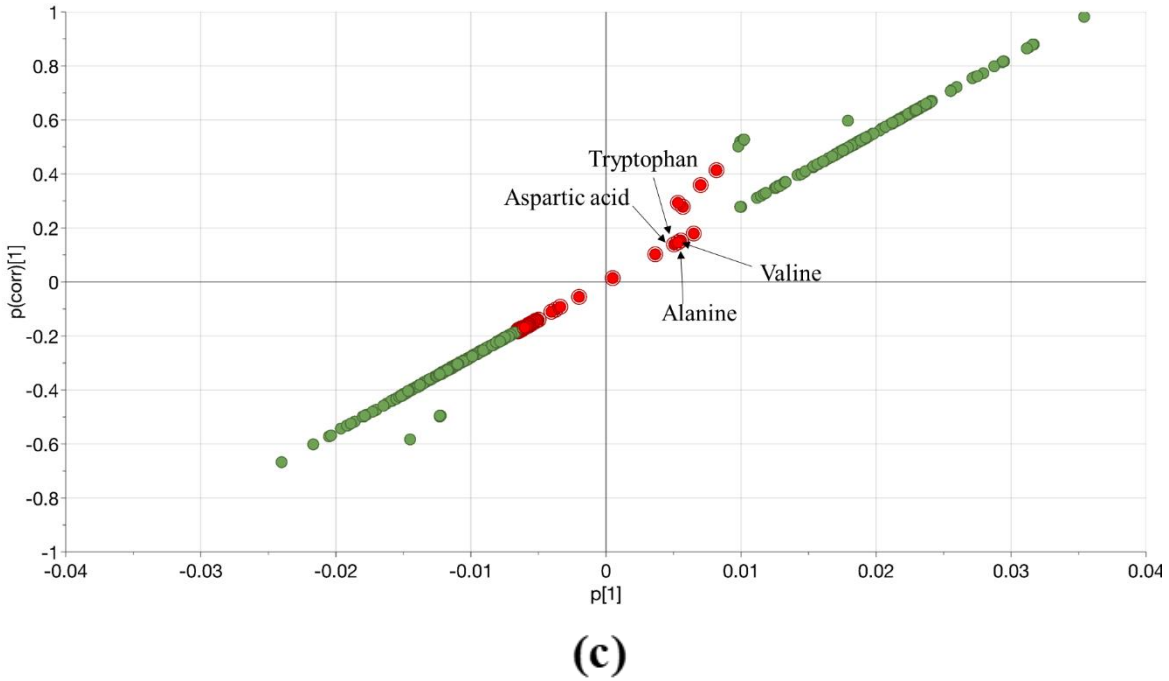
sensitivity of the first. Furthermore, each dataset was obtained with the different workflow of different software.



(a)



(b)



(c)

Figure 4: (a) The loading plot displays the relation between the different *Lemna minor* samples 100% MeOH and 100% H_2O extracts analyzed with TOF and QTOF; (b) OPLS-DA score scatter plot of *Lemna minor* samples 100% MeOH (triangular) and 100% H_2O (squares) extracts analyzed with TOF (violet) and QTOF (yellow). The confidence limit is 95%; (c) The S-plot of *Lemna minor* different samples, the marked molecules represent the common one between the two machines. Tryptophan, aspartic acid, alanine, and valine were prioritized by the PLANT-IDENT database and identified using reference standards after

2. 7. Common metabolic profile of Lemna minor metabolites identification

The obtained features were classified according to RT, and, prioritized using the PLANT-IDENT database in the FOR-IDENT platform. The search was scored according to the mass screening, and MS/MS. The PLANT-IDENT database suggested pegamine, tryptophan, aspartic acid, alanine, valine, and betaine (trimethylglycine) candidates matching features in both 100% MeOH, and 100% H_2O extracts. They have with LogD ≥ 0 at (pH= 7) except pergamine, which was eliminated. They also have priority "look at" and standards reference injected for validation except for betaine (Figure S1). Subsequently, the filtration parameters decreased the number of the results into 5 metabolites, listed in (Table S4).

Those metabolites were annotated and classified in the second level. They could be classified into level one by confirmation with standards reference injections.

In our previous study, the amino acids (phenylalanine, proline, tryptophan, alanine, tyrosine, aspartic acid, isoleucine, serine, and valine) were identified in *L. minor* using system B [12]. Despite being also identified in system A (Table S5), they were not detected with the common metabolic profile between both systems A& B. This is due to differences in the intensities of the amino acids between the two extracts.

Secondly, the RP part with RT > 15 minutes, the second part of each data set was uploaded into the PLANT-IDENT database in the FOR-IDENT platform, individually. Two features were suggested as a matching candidate in 100% MeOH, and 100%H_2O extracts. They did not have any matches in the PLANT-IDENT database. Those metabolites were annotated and classified in the third level.

3. Materials and Methods

3.1. Reagents and chemicals

LC-MS grade methanol and water were obtained from VWR, Darmstadt, Germany. 6-Amino-1,3-dimethyl-5-(formylamino)uracil (certified reference material), chlortoluron (100 quality level, HPLC), famotidine (100 quality level, HPLC), vidarabine (certified reference material), etilefrine (pharmaceutical primary standard), and 2,4-diamino-6-(hydroxymethyl)pteridine hydrochloride (100 quality level, HPLC), and apigenin (≥ 97%), monuron (100 quality level, HPLC), chloridazon (100 quality level, HPLC), carbetamide (100 quality level, HPLC), metobromuron (100 quality level, HPLC), sotalol (≥ 98%), metconazol (100 quality level, HPLC), vitexin (≥ 95%), flavone (≥ 99%), DL-Ala-DL-Ala (98%), nicotinic acid (≥ 99.5%), nicotinamide (≥ 99%), and galangin (≥ 97%)were obtained from Sigma, Darmstadt, Germany. Chlorbromuron (99.24%) was obtained from Dr. Ehrenstorfer, Augsburg, Germany. Acacetin was purchased from Fluka, Buchs, Switzerland. Apigenin-7-glucoside, Apigenin-5-glucoside, apigenin-6-arabinoside-8-glucoside, Apigenin-6,8-di-glucoside, apiin, orientin, kaempferol-7-rhamnoside, peonidin, norwogonin, luteolin-3,7-di-glucoside, 6-methoxy flavone, 4-methoxy cinnamic acid, naringenin-7-glucoside, isovitexin, tricin, quercetin-3-glucoside, saponarin, myricetin, 5-hydroxy-6-methoxy flavone, chrysoeriol, isovitexin, robinetin, umbelliferone were kindly provided by Center of Life and Food Science Weihenstephan, Biotechnology of Natural Products, Technical University of Munich, Germany.

3.2. Plant Sample

Lemna minor L. was cultivated in aquaria according to Obermeier *et al*. [16] with small modifications. Plants were grown at 23 °C with a photoperiod of 16-8 h and an average light intensity of 43 μmol m^{-2} s^{-1}. *Lemna* fronds were subcultured every two weeks in 24 L of Steinberg medium made of (in mg/L) 85 $NaNO_3$, 13.4 KH_2PO_4, 75 $MgSO_4 \cdot 7\ H_2O$, 36 $CaCl_2 \cdot 2\ H_2O$, 20 Na_2CO_3, 1 H_3BO_3, 0.2 $MnCl_2 \cdot 4\ H_2O$, 0.01 $Na_2MoO_4 \cdot 2\ H_2O$, 0.05 $ZnSO_4 \cdot 7\ H_2O$, 0.005 $CuSo_4 \cdot 5\ H_2O$, 0.01 $Co(NO_3)_2 \cdot 6\ H_2O$, 0.84 $FeCl_3 \cdot 6\ H_2O$ and 1.4 Na_2-EDTA$\cdot 2\ H_2O$. After the plants were harvested, shortly rinsed with tap water, dried with lint-free tissue paper, and were finally frozen in liquid nitrogen. Samples were kept at -80°C until further processing. *Lemna* was kindly provided by the German Research Center for

Environmental Health, Plant-Microbe Interactions, Helmholtz Centrum of Munich, Munich, Germany.

3.3. Extraction

500 mg of *Lemna* freeze-dried and milled powder was extracted with a) 100% methanol (MeOH), b) MeOH-water (50:50, v/v), and c) 100% water (H_2O), respectively. Extracts with plant powder were sonicated (Sonorex super RK 106, Bandelin, Berlin, Germany) for 10 minutes at 4ºC. Despite using the ultrasonic bath in this extraction method, extracts with plant powder were sonicated for 10 minutes at 4ºC and with a 35 kHz frequency. Then, samples were centrifuged (NuWind Multi-Application Bench Top Refrigerated Centrifuge 2, NuAire, USA) at 1500 rpm for 20 minutes at 4ºC and the supernatants were transferred to clean glass test tubes. The extraction process was triplicated in identical experimental conditions. Finally, the extracts were evaporated to dryness (using a SpeedVac, Fischer Scientific, Göteborg, Sweden) and dissolved in (MeOH-H_2O (50:50, v/v) [12, 17].

3.4. Instruments

a) System A

Lemna metabolomics and reference standards were separated with LC (Agilent 1260 Infinity) (system A) consisting of an autosampler, two columns, two binary pumps, an online degasser, a mixing chamber, and a UV detector. The LC-system was used to perform reversed-phase and zwitterion hydrophilic interaction liquid chromatography (HILIC) in its serial coupling. The reversed-phase separation column was a Poroshell 120 EC-C18 (50.0 × 3.0 mm, 2.7 μm; Agilent Technologies). The HILIC column was a ZIC-HILIC column (150 × 2.1 mm, 5 μm, 200 Å; Merck Sequant, Umea, Sweden). Columns were coupled through a T-piece (Upchurch, IDEX Europe GmbH, Erlangen, Germany). The third port of the T-piece was connected to the HILIC flow pump. The injection volume was 10 μL. Further details, like the mobile phase of the RPLC-HILIC serial coupling and other settings, are described in [12, 18, 19].

b) System B

Lemna metabolic profile mass spectrums were performed on a Triple TOF 5600system (AB SCIEX triple TOF 5600, Darmstadt, Germany) (system B) with a Duospray ion source and a Turbolonspray ESI probe in positive ion mode. The MSMS data were collected from the eight parallel experiments at a collision energy of 40 ± 20 eV in the full cycle scan.

The mass spectrum parameters were set as the following parameters: ISVF (ion spray voltage floating) 2000 kV and turbo spray temperature, 650°C. DP clustering potential (DP), and collision energy (CE) were set to 46, and 40 V, respectively. The nebulizer and the auxiliary gas were both nitrogen. Also, the nebulizer gas (gas 1), the heater gas (gas 2), and the curtain gas were set to 44, 50, and 29 psi, respectively.

3.5. Internal standards

Each sample and blanks were spiked with a standards mixture of 12 substances. The mixture consists of 6-amino-1,3-dimethyl-5-(formylamino) uracil, chlortoluron, famotidine, vidarabine, etilefrine, monuron, carbetamide, metobromuron, sotalol, chlorbromuron, metconazol, and 2,4-diamino-6-(hydroxymethyl) pteridine hydrochloride to obtain a final concentration of 20µM each. The absolute variation between literature monoisotopic mass and the mean of measured isotopic masses (Δ ppm). The results are summarized in (Table S1). The absolute mass deviation ranged from 0.09 (ppm) to 2.81 (ppm). The RT standard deviation was less than 0.3. Moreover, the relative standard deviation (%RSD) was less than 2 % as shown in (Table S1). These parameters were used to correlate the features in different samples.

3.6. *Data collection and preprocessing*

a) - System A

Data acquired with MassHunter Workstation LC/MS Data Acquisition software B 05.00, (Agilent Technologies, Waldbronn, Germany) was subsequently analyzed with Profinder B.06.00 (Agilent Technologies) to extract the so-called 'features' by their retention times (RT), molecular mass, and their peak intensity in various extracts. This was performed in a combination of the 3-fold injections of each sample after removing the features found in the corresponding blank samples. The parameters are set to a peak filter of 1000 counts peak height, ion species to "positive ions" with H+, Na+, K+, and NH4+, "charge state" to 1, the "expected RT" to ±3.00 min, and the mass to ±10 ppm. The extracted ion chromatograms (EICs) were smoothed with a Gaussian function using 9 points function width and 5000 points Gaussian width. This limits the result finally to 2000 compound groups.

b)- System B

3.6.1. Peak picking

Data collection and preprocessing are the first and the important step in metabolomics analysis because they determine the quality of the data for the next steps. The data was obtained using the Analyst Software version (TF 1.7.1). Then, it was preprocessed with MarkerView Software (version 1.3.1). The parameter was optimized according to the target analysis of the internal standards in all samples and blanks. Also, transformation and/or normalization of the data acquired in this step, which is required for univariate and multivariate statistical analysis. The minimum and maximum retention times were 5 and 34 min respectively. The subtraction offset was 15 scans. The noise threshold and subtraction multiplication factors were 50 and 1, respectively. Further, the minimum spectral peak width was 2 ppm (Figure 1). The noise threshold and subtraction multiplication factors have a fundamental impact on peak picking especially for the low abundance metabolites due to the matrix effect of unfractionated extracts. This was achieved using internal standards.

3.6.2. Alignment and filtering

The three injections of each extract feature list were compared and aligned. The variation between different injections was determined according to tolerance in RT and mass. The retention time tolerance is 2%, which is assigned according to the target analysis of the internal standards. Further, mass tolerance was 5 ppm based on QTOF specification [20]. Then, the background was deleted (i.e. the features found in the blank corresponding to each solvent were deleted from the same extract) (Figure 1). *Lemna minor* metabolic profile investigation demands alignments of features, which were found in different extracts and the different injections of similar extracts. Further, the features were deleted, which were found in the corresponding blank. This step was fundamental in the untargeted metabolomics analysis to avoid misleading statements.

Also, the retention time index and mass tolerance were determined according to the results of the internal standard, which led to a decrease in the number of false-positive, as well as, negative peaks, and reduce the number of features of the same metabolite. The minimum variation in the noise threshold affected the peak extraction by increasing the number of false-positive peaks or subtract real peaks apart. The isotopic, adduct (e.g., $[M+Na]^+$, $[M+K]^+$ or $[M+NH4]^+$), and fragments ions were removed or combined to produce one metabolite corresponding to one ion [13].

3.6.3. PLANT-IDENT batch searching and scoring

After feature extractions for systems a,) and b), respectively (as shown in 3.6.2.), the data was organized in an excel file, which was uploaded into the FOR-IDENT platform. The molecular masses of detected metabolites were compared with compounds stored in the database, resulting in suggestions for their identification. The house-made compound database PLANT-IDENT containing various plant metabolites was created and implemented in the FOR-IDENT platform to prioritize and annotate plant metabolites. The PLANT-IDENT supports the untargeted metabolomics screening analysis through the identification of plant metabolites. The PLANT-IDENT is an open-access compound data. PLANT-IDENT contains up to 3019 plant metabolites from Lemnaceae, Poaceae, Brassicaceae, Nymphaceae families, in addition to, flavonoids, and nitrogen-containing plant metabolites. The compound's name and plant source were collected from the available literature (articles and/or books) which was cited in the database. The chemical identifiers (smiles, LogD, international chemical identifier (InChi Key)) were gathered with a research crawler from the PubChem and Chemspider databases. Then, the physicochemical properties and spectral data were crawled through the PubChem database. The LogD was used to support the identification of compounds through hydrophilicity for the HILIC eluted compounds. Further, the retention time index (RTI) was used to support the identification of RPLC eluted compounds via hydrophobicity. The retention time of the standards mixture (Table S2) was also uploaded to normalize the RT. The normalization was performed according to the calibration curve (between the RT and Log D at pH=7) of the target analysis of the standard mixture, for more details

readers referred to [14]. The mixture consists of compounds, which have affixed logD at different pHs (i.e. 'compounds with stable logP values) (Table S2).

The fragmentations were compared through *In-Silico fragmentation* MetFrag [21]. Furthermore, the results have priority "look at" was considered and standards reference injected for validation. Moreover, chemotaxonomy was applied. PLANT-IDENT contained the plant name and/or Family name, which was reported to contain the corresponding metabolites. The processed data from the system (A or B), respectively, was uploaded into the FOR-IDENT platform (https://water.for-ident.org/#!home). The search process was performed according to the following parameters: pH= 7, precursor ion mass deviation 5 ppm, the intensity threshold was 5 counts per second (CPS), ion species was positive. The metabolites were considered according to their score, which depends on mass screening and MS/MS data (later in system B)). Then, the results were considered when they have negative LogD at pH= 7. In contrast, the metabolites were suspected in the RP column eluted compounds according to the score of the same parameter in addition to the retention time index (RTI) [22]. The metabolites were considered when they have positive LogD. The last step was the chemotaxonomy criteria. Each metabolite has a plant and/or family name reference according to the available literature (Figure 1).

3.6.4. Classification scheme

The different extracts were analyzed according to the previously mentioned workflow. Then, the plant metabolites classification was performed according to the scheme of [23]. The scheme has 5 levels. The first one is the identification by the reference standard. In the second level, the identification was performed by various criteria such as (retention time behavior, accurate mass (i.e. empirical formula), fragmentation, and chemotaxonomical criteria). In the third level, the identification was performed by comparison of accurate mass and fragments from different laboratories. In the fourth level, the identification is done by molecular formula or fragments comparison the last level is the mass identification without further information. This classification scheme enhances the identification of plant metabolites in untargeted metabolomics analysis.

3.6.5. Orthogonal partial least square-discriminant analysis

The quality of features was statistically investigated from both systems. The experiment was repeated twice. In each experiment, *Lemna* was grown in the same condition and extracted with 100% MeOH, 50% MeOH, and 100% H_2O extracts. OPLS is the orthogonal modification of PLS regression analysis methods, which both are supervised statistical models on the contrary of PCA [24]. The OPLS fits well for metabolomics analysis because it can analyze a large number of variables (8940) for a small sample size (3 extracts). The OPLS separates the variables (X) into two directions linear and orthogonal to Y. The OPLS model can be visualized with the score plot and the loading plot of the variables, which describe the contribution of variables. The OPLS quality is described with the cumulative variation in the matrix (features) or $R^2X(cum)$,

the cumulative variation in the Y matrix (extracts) or R^2-Y(cum), and the cross-validated predictive ability or Q^2(cum) values. R^2 is defined as a fraction of the variance explained by a component. Cross-validation of R^2 gives Q^2, which represents the proportion of total variation predicted by a component. Thus, the R^2 indicates how well the variation of a variable is explained, and Q^2 how well a variable could be predicted and estimated by cross-validation [25]. The dataset organization and different statistical parameters were mentioned in detail in [26]. Additionally, the detector parameter was included as a secondary observation.

4. Conclusions

Lemna's metabolic profile was investigated using an extended-polarity LC system. The untargeted workflow was applied using an open access PLANT-IDENT database. PLANT-IDENT database contains up to 3019 naturally occurring compounds with their physicochemical properties, in addition to chemotaxonomic data. The workflow considered the features represented in triplicates to minimize the false-positive results. The identification of *Lemna* metabolites proceeded according to different filters: LogD, mass deviation, MSMS fragment comparison, and chemotaxonomy filter. Moreover, compounds were identified using reference standards. Further, the workflow was investigated with statistical analysis to validate the reliability and information transfer between different mass spectrometric systems and/or laboratories. Furthermore, PLANT-IDENT could be used in the identification of important variable(s). PLANT-IDENT can transform the spectral data (mass, RT) into names. Thus, the untargeted plant metabolomics research will be enhanced via utilizing the workflow combined with the PLANT-IDENT database. The PLANT-IDENT database is under work to include more plant members.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S1: title, Table S1: title, Video S1: title.

Author Contributions R.W. and T.L. designed the metabolomics study. R.W. prepared the samples, performed the untargeted metabolomics analysis, and ran the analyses in the metabolomics platform. S.M. created the statistical design for data evaluation and data interpretation and performed the realization together with R.W. and T.L., C.C. and P.S. created the test set and performed the plant growing and incubation together with R.W. R.W. and T.L. conceived and drafted the manuscript. All the authors contributed with critical intellectual input. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the Bavarian State Ministry of the Environment and Consumer Protection as well as the Cultural Affairs and Mission Sector of the Egyptian Ministry of Higher Education.

Acknowledgments: This work was assisted by the German Research Foundation (DFG) and the Technical University of Munich (TUM) in the scope of the Open Access Publishing Program). We thank Dr. Axel Besa and Sciex for providing the TripleToF mass spectrometer and valuable support

Conflicts of Interest: The authors declare that they have no conflict of interest. There is no connection between AFIN-TS GMbH and the subject of this manuscript.

References

1. Doppler, M.; Bueschl, C.; Kluger, B.; Koutnik, A.; Lemmens, M.; Buerstmayr, H.; Rechthaler, J.; Krska, R.; Adam, G.; Schuhmacher, R., Stable Isotope–Assisted Plant Metabolomics: Combination of Global and Tracer-Based Labeling for Enhanced Untargeted Profiling and Compound Annotation. *Frontiers in Plant Science* **2019,** 10, (1366).
2. Yang, Z.; Nakabayashi, R.; Okazaki, Y.; Mori, T.; Takamatsu, S.; Kitanaka, S.; Kikuchi, J.; Saito, K., Toward better annotation in plant metabolomics: isolation and structure elucidation of 36 specialized metabolites from Oryza sativa (rice) by using MS/MS and NMR analyses. *Metabolomics* **2014,** 10, (4), 543-555.
3. Sorokina, M.; Steinbeck, C., Review on natural products databases: where to find data in 2020. *Journal of Cheminformatics* **2020,** 12, (1), 20.
4. Gika, H. G.; Theodoridis, G. A.; Earll, M.; Snyder, R. W.; Sumner, S. J.; Wilson, I. D., Does the Mass Spectrometer Define the Marker? A Comparison of Global Metabolite Profiling Data Generated Simultaneously via UPLC-MS on Two Different Mass Spectrometers. *Analytical chemistry* **2010,** 82, (19), 8226-8234.
5. Want, E. J.; Masson, P.; Michopoulos, F.; Wilson, I. D.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Loftus, N.; Holmes, E.; Nicholson, J. K., Global metabolic profiling of animal and human tissues via UPLC-MS. *Nature Protocols* **2013,** 8, (1), 17-32.
6. Martins, M. C. M.; Caldana, C.; Wolf, L. D.; de Abreu, L. G. F., The Importance of Experimental Design, Quality Assurance, and Control in Plant Metabolomics Experiments. *Methods in molecular biology (Clifton, N.J.)* **2018,** 1778, 3-17.
7. Zhang, X.-w.; Li, Q.-h.; Xu, Z.-d.; Dou, J.-j., Mass spectrometry-based metabolomics in health and medical science: a systematic review. *RSC Advances* **2020,** 10, (6), 3092-3104.
8. Bedia, C.; Cardoso, P.; Dalmau, N.; Garreta-Lara, E.; Gómez-Canela, C.; Gorrochategui, E.; Navarro-Reig, M.; Ortiz-Villanueva, E.; Puig-Castellví, F.; Tauler, R., Chapter Nineteen - Applications of Metabolomics Analysis in Environmental Research. In *Comprehensive Analytical Chemistry*, Jaumot, J.; Bedia, C.; Tauler, R., Eds. Elsevier: 2018; Vol. 82, pp 533-582.
9. Ali, Z.; Waheed, H.; Kazi, A. G.; Hayat, A.; Ahmad, M., Chapter 16 - Duckweed: An Efficient Hyperaccumulator of Heavy Metals in Water Bodies. In-*Plant Metal Interaction*, Ahmad, P., Ed. Elsevier: 2016; pp 411-429.
10. Kotowska, U.; Piotrowska, A.; Isidorova, A. G.; Bajguz, A.; Isidorov, V. A., Gas chromatographic-mass spectrometric investigation of the chemical composition of the aquatic plant Wolffia arrhiza (Lemnaceae). *Oceanological and Hydrobiological Studies* **2013,** 42, (2), 181-187.
11. Jensen, S. R.; Nielsen, B. J., Chemical Characters. In *The Families of the Monocotyledons: Structure, Evolution, and Taxonomy*, Springer Berlin Heidelberg: Berlin, Heidelberg, 1985; pp 17-22.
12. Wahman, R.; Graßmann, J.; Sauvêtre, A.; Schröder, P.; Letzel, T., Lemna minor studies under various storage periods using extended-polarity extraction and metabolite non-target screening analysis. *Journal of Pharmaceutical and Biomedical Analysis* **2020,** 188, 113362.
13. Zhao, X.; Zeng, Z.; Chen, A.; Lu, X.; Zhao, C.; Hu, C.; Zhou, L.; Liu, X.; Wang, X.; Hou, X.; Ye, Y.; Xu, G., Comprehensive Strategy to Construct In-House Database for Accurate and Batch Identification of Small Molecular Metabolites. *Analytical Chemistry* **2018,** 90, (12), 7635-7643.
14. Wahman, R.; Grassmann, J.; Schröder, P.; Letzel, T., Plant metabolomic workflows using reversed-phase LC and HILIC with ESI-TOF-MS. *LC GC N. Am.* **2019,** 37, (3), 8-15.
15. AKHTAR, T. A.; LEES, H. A.; LAMPI, M. A.; ENSTONE, D.; BRAIN, R. A.; GREENBERG, B. M., Photosynthetic redox imbalance influences flavonoid biosynthesis in Lemna gibba. *Plant, Cell & Environment* **2010,** 33, (7), 1205-1219.
16. Tsolmon, B.; Fang, Y.; Yang, T.; Guo, L.; He, K.; Li, G.-Y.; Zhao, H., Structural identification and UPLC-ESI-QTOF-MS2 analysis of flavonoids in the aquatic plant Landoltia punctata and their in vitro and in vivo antioxidant activities. *Food Chemistry* **2020**, 128392.
17. Wahman, R.; Grassmann, J.; Schröder, P.; Letzel, T., Plant Metabolomic Workflows Using Reversed-Phase LC and HILIC with ESI-TOF-MS. *LCGC North America* **2019,** 37, (3), 8-15.

18. Bieber, S.; Greco, G.; Grosse, S.; Letzel, T., RPLC-HILIC, and SFC with Mass Spectrometry: Polarity-Extended Organic Molecule Screening in Environmental (Water) Samples. *Analytical chemistry* **2017,** 89, (15), 7907-7914.
19. Greco, G.; Grosse, S.; Letzel, T., Serial coupling of reversed-phase and zwitterionic hydrophilic interaction LC/MS for the analysis of polar and nonpolar phenols in wine. *Journal of Separation Science* **2013,** 36, (8), 1379-1388.
20. Nürenberg, G.; Schulz, M.; Kunkel, U.; Ternes, T. A., Development and validation of a generic nontarget method based on liquid chromatography – high-resolution mass spectrometry analysis for the evaluation of different wastewater treatment options. *Journal of Chromatography A* **2015,** 1426, 77-90.
21. Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S., MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *Journal of Cheminformatics* **2016,** 8, (1), 3.
22. Sylvia Grosse, T. L., User Manual Stoff-IDENT Database. 2017; Vol. 4.1, pp 1-35.
23. Letzel, T.; Lucke, T.; Schulz, W.; Sengl, M.; Letzel, M., In a class of its own – OMI (Organic Molecule Identification) in water using LC-MS(/MS): Steps from "unknown" to "identified": a contribution to the discussion. *Lab & More International* **2014**, 24-28.
24. Blasco, H.; Błaszczyński, J.; Billaut, J. C.; Nadal-Desbarats, L.; Pradat, P. F.; Devos, D.; Moreau, C.; Andres, C. R.; Emond, P.; Corcia, P.; Słowiński, R., Comparative analysis of targeted metabolomics: Dominance-based rough set approach versus orthogonal partial least square-discriminant analysis. *Journal of Biomedical Informatics* **2015,** 53, 291-299.
25. Löfstedt, T.; Trygg, J., OnPLS—a novel multiblock method for the modeling of predictive and orthogonal variation. *Journal of Chemometrics* **2011,** 25, (8), 441-455.
26. Wahman, R.; Sauvêtre, A.; Schröder, P.; Moser, S.; Letzel, T., Untargeted Metabolomics Studies on Drug-Incubated Phragmites australis Profiles. *Metabolites* **2021,** 11, (1), 2.
27. Waridel, P.; Wolfender, J.-L.; Ndjoko, K.; Hobby, K. R.; Major, H. J.; Hostettmann, K., Evaluation of quadrupole time-of-flight tandem mass spectrometry and ion-trap multiple-stage mass spectrometry for the differentiation of C-glycosidic flavonoid isomers. *Journal of Chromatography A* **2001,** 926, (1), 29-41.
28. Ndolo, V. U.; Fulcher, R. G.; Beta, T., Application of LC-MS-MS to identify niacin in aleurone layers of yellow corn, barley, and wheat kernels. *Journal of Cereal Science* **2015,** 65, 88-95.
29. Hau, J.; Stadler, R.; Jenny, T. A.; Fay, L. B., Tandem mass spectrometric accurate mass performance of time-of-flight and Fourier transform ion cyclotron resonance mass spectrometry: a case study with pyridine derivatives. *Rapid Communications in Mass Spectrometry* **2001,** 15, (19), 1840-1848.
30. Wang, X.; Zhong, X.-J.; Zhou, N.; Cai, N.; Xu, J.-H.; Wang, Q.-B.; Li, J.-J.; Liu, Q.; Lin, P.-C.; Shang, X.-Y., Rapid Characterization of Chemical Constituents of the Tubers of Gymnadenia conopsea by UPLC–Orbitrap–MS/MS Analysis. *Molecules (Basel, Switzerland)* **2020,** 25, (4), 898.
31. Sakalem, M. E.; Negri, G.; Tabach, R., Chemical composition of hydroethanolic extracts from five species of the Passiflora genus. *Revista Brasileira de Farmacognosia* **2012,** 22, 1219-1232.
32. Tsimogiannis, D.; Samiotaki, M.; Panayotou, G.; Oreopoulou, V., Characterization of flavonoid subgroups and hydroxy substitution by HPLC-MS/MS. *Molecules (Basel, Switzerland)* **2007,** 12, (3), 593-606.
33. Xu, L. L.; Xu, J. J.; Zhong, K. R.; Shang, Z. P.; Wang, F.; Wang, R. F.; Zhang, L.; Zhang, J. Y.; Liu, B., Analysis of Non-Volatile Chemical Constituents of Menthae Haplocalycis Herba by Ultra-High Performance Liquid Chromatography-High Resolution Mass Spectrometry. *Molecules (Basel, Switzerland)* **2017,** 22, (10).
34. Lee, S.-H.; Kim, H.-W.; Lee, M.-K.; Kim, Y. J.; Asamenew, G.; Cha, Y.-S.; Kim, J.-B., Phenolic profiling and quantitative determination of common sage (Salvia plebeia R. Br.) by UPLC-DAD-QTOF/MS. *European Food Research and Technology* **2018,** 244, (9), 1637-1646.

35. Le Gall, G.; DuPont, M. S.; Mellon, F. A.; Davis, A. L.; Collins, G. J.; Verhoeyen, M. E.; Colquhoun, I. J., Characterization and Content of Flavonoid Glycosides in Genetically Modified Tomato (Lycopersicon esculentum) Fruits. *Journal of Agricultural and Food Chemistry* **2003,** 51, (9), 2438-2446.
36. Jang, G. H.; Kim, H. W.; Lee, M. K.; Jeong, S. Y.; Bak, A. R.; Lee, D. J.; Kim, J. B., Characterization and quantification of flavonoid glycosides in the Prunus genus by UPLC-DAD-QTOF/MS. *Saudi Journal of Biological Sciences* **2018,** 25, (8), 1622-1631.
37. Akhtar, N.; Thadhani, V. M.; Ul Haq, F.; Khan, M. N.; Ali, S.; Musharraf, S. G., Rapid identification and quantification of bioactive metabolites in processed Camellia sinensis samples by UHPLC-ESI-MS/MS and evaluation of their antioxidant activity. *Journal of Industrial and Engineering Chemistry* **2020,** 90, 419-426.
38. Crow, F. W.; Tomer, K. B.; Looker, J. H.; Gross, M. L., Fast atom bombardment and tandem mass spectrometry for structure determination of steroid and flavonoid glycosides. *Analytical Biochemistry* **1986,** 155, (2), 286-307.
39. Pereira, C. A. M.; Yariwake, J. H.; McCullagh, M., Distinction of the C-glycosylflavone isomer pairs orientin/isoorientin and vitexin/isovitexin using HPLC-MS exact mass measurement and in-source CID. *Phytochemical Analysis* **2005,** 16, (5), 295-301.
40. Simirgiotis, M.; Schmeda-Hirschmann, G.; Borquez, J.; Kennelly, E., The Passiflora tripartita (banana passion) fruit: A source of bioactive flavonoid C-glycosides isolated by HSCCC and characterized by HPLC-DAD-ESI/MS/MS. *Molecules (Basel, Switzerland)* **2013,** 18, 1672-92.
41. Gao, X.; Sun, W.; Fu, Q.; Niu, X., Ultra-performance liquid chromatography coupled with electrospray ionization/quadrupole time-of-flight mass spectrometry for the rapid analysis of constituents in the traditional Chinese medical formula Danggui San. *Journal of Separation Science* **2014,** 37, (1-2), 53-60.
42. Yang, Z.; Nakabayashi, R.; Okazaki, Y.; Mori, T.; Takamatsu, S.; Kitanaka, S.; Kikuchi, J.; Saito, K., Toward better annotation in plant metabolomics: isolation and structure elucidation of 36 specialized metabolites from Oryza sativa (rice) by using MS/MS and NMR analyses. *Metabolomics: Official journal of the Metabolomic Society* **2014,** 10, (4), 543-555.
43. Hughes, R. J.; Croley, T. R.; Metcalfe, C. D.; March, R. E., A tandem mass spectrometric study of selected characteristic flavonoids22Dedicated to Professor N.M.M. Nibbering for his many contributions to mass spectrometry. *International Journal of Mass Spectrometry* **2001,** 210-211, 371-385.
44. Ma, Y. L.; Li, Q. M.; Van den Heuvel, H.; Claeys, M., Characterization of flavone and flavonol aglycones by collision-induced dissociation tandem mass spectrometry. *Rapid Communications in Mass Spectrometry* **1997,** 11, (12), 1357-1364.

Table1. The compounds were identified in *Lemna minor* metabolic profile identified by PLANT-IDENT. Retention time (RT) means of standards (S), measured (M), and the RT deviation. Also, the mass means of standards (S), measured (M), and the deviation between them. The mean fragments of standards and measured were compared with the literature and listed with the references. Compounds could be detected in system B (single TOF-MS) marked with *

Compound Name	RT (S) [Min]	RT (M) [Min]	ΔRT [Min]	Mass (S) [Da]	Mass (M) [Da]	Δppm	MSMS fragments	References
Vitexin	7.5	7.3	0.2	433.1133	433.1129	0.8	433;415;397;379;337;313;283	[27]
Niacin*	7.6	7.8	-0.2	124.0394	124.0393	0.7	124;96;80;78	[28]
Nicotinamide	7.8	7.6	0.1	123.0554	123.0553	0.8	123;106;80;78	[29]
Phenylalanine*	11.0	11.1	-0.1	166.0866	166.08627	2.0	120;103;77	MassBank of North America (MoNA)
Leucine/Isoleucine*	11.2	11.2	-0.1	132.1018	132.1020	-1.4	86;69;44;30	(MoNA)
Tryptophan*	11.7	11.7	0.0	205.0973	205.0970	1.8	188;146;144	(MoNA)
Valine*	12.1	11.9	0.1	118.0863	118.0862	0.8	72;71;55	(MoNA)
Tyrosine*	12.3	12.2	0.1	182.0811	182.0810	1.9	136;123;119	(MoNA)
Proline*	12.4	12.4	0.0	116.0705	116.0707	0.3	70;68;43	(MoNA)
Glutamic acid*	12.5	12.6	-0.2	147.0434	147.0430	3.0	130;102;84	(MoNA)
Aspartic acid*	12.7	12.7	0.0	134.0447	134.0447	-0.2	134;115	(MoNA)
Di-L-Alanine	12.7	12.8	-0.1	161.0928	161.0920	4.9	161;115;90	(MoNA)
4-Methoxy cinnamic acid	13.4	13.1	0.3	179.0706	179.0708	-0.9	147;137	[30]
Alanine*	13.4	13.2	0.2	90.0550	90.0548	2.1	44;28	(MoNA)
Threonine*	13.6	13.4	0.2	120.0656	120.0653	2.8	73;56	(MoNA)
Serine*	14.0	13.8	0.2	106.0500	116.0499	0.9	60;42;43	(MoNA)
Apigenin-6,8-diglucopyranoside*	15.8	15.7	0.1	595.1659	595.1658	0.2	595; 383	[31]

Compound Name	**RT (S) [Min]**	**RT (M) [Min]**	**ΔRT [Min]**	**Mass (S) [Da]**	**Mass (M) [Da]**	**Δppm**	**MSMS fragments**	**References**
Robinetin	15.8	15.9	-0.1	303.0494	303.0493	0.3	285;267;147	(MoNA)
Quercetin	24.8	24.9	-0.1	303.0549	303.0544	1.7	303;285;257;229;165	[32]
Luteolin	24.8	24.6	0.2	287.0562	287.0557	1.6	287;269;241;153	[32]
Acacetin	28.8	29.1	-0.3	285.0759	285.0760	-0.4	285;242;153	[33]
Apigenin 7-glucoside	24.8	24.7	0.1	433.1130	433.1132	-0.4	433;271	[34]
Orientin	25.1	25.2	-0.1	449.1123	449.1134	-2.6	449; 329	[31]
Peonidin	25.6	25.2	0.4	302.0785	302.0792	-2.4	302;283;197	(MoNA)
6-Methoxyflavone	30.4	30.6	-0.2	253.0879	253.0881	-0.7	253; 238; 210	NIST
Luteolin-3',7-di-O-glucoside	23.8	23.6	0.3	611.1640	611.1622	2.8	611;449;287	(MoNA)
Kaempferol	29.0	29.1	-0.1	287.0531	287.0540	-3.1	287;269;231;165;153;133	[32]
Apigenin	26.8	26.7	0.1	271.0603	271.0604	-0.6	271;253;153	[32]
Flavon	29.9	29.6	0.2	223.0756	223.0748	3.6	223;178;152;121	(MoNA)
Naringenin-7-O-glucoside	25.0	24.1	0.9	435.1298	435.1285	2.9	435;273	[35]
Quercetin-3-glucoside	24.2	24.3	-0.1	465.1018	465.1022	-0.7	465; 303	[36]
Saponarin	23.8	24.0	-0.2	595.1638	595.1663	-4.2	433;415;397;367;337;283; 271	[37]
5-Hydroxy-6-Methoxyflavon	31.3	31.1	0.1	269.0823	269.0819	1.3	269;254;104	(MoNA)
Apiin	24.6	23.8	0.9	565.1566	565.1559	1.3	433;313	[38]
Chrysoeriol	26.9	26.8	0.1	301.0731	301.0722	2.9	286;121	[33]

746 **Compound Name**	**RT (S) [Min]**	**RT (M) [Min]**	**ΔRT [Min]**	**Mass (S) [Da]**	**Mass (M) [Da]**	**Δppm**	**MSMS fragments**	**References**
Isoorientin	23.8	23.6	0.2	449.1085	449.1095	-2.1	499;329;299;165	[39]
Umbelliferone	24.7	24.4	0.2	163.0396	163.0391	2.9	135;107	[30]
Apigenin-5-glucoside	24.2	23.9	0.3	433.1127	433.1127	0.1	433;271	[34]
Apigenin-6-arabopyranoside-8-glucopyranose	23.4	23.3	0.2	565.1550	565.1557	-1.2	565;547;379;337;325;295;121	[40]
Norwogonin	24.2	24.0	0.2	271.0604	271.0599	1.8	271;253;241;225	[41]
Isovitexin	24.1	23.9	0.2	433.1125	433.1134	-2.0	313;295;284;283;267	[27]
Tricin	26.8	26.3	0.6	331.0811	331.0796	4.7	331;315	[42]
Galangin	29.4	29.4	-0.1	271.0602	271.0608	2.3	271;253	[43]
Myricetin	25.1	25.1	0.0	319.0440	319.0453	4.0	301;283;265;111	[44]

The changes in *Lemna minor* metabolomic profile: a response to diclofenac incubation

Rofida Wahman[1,2], Catarina Cruzeiro[3,*], Johanna Graßmann[1], Peter Schröder[3] and Thomas Letzel[1,4]

[1]Chair of Urban Water Systems Engineering, Technical University of Munich, Am Coulombwall 3, 85748 Garching

[2]Pharmacognosy Department, Faculty of Pharmacy, Assiut University, Assiut, Arab Republic of Egypt/Egypt

[3]German Research Center for Environmental Health, Research Unit Comparative Microbiome Analysis, Helmholtz Zentrum München, Ingolstädter Street 1, 85764 Neuherberg

[4]Analytisches Forschungsinstitut für Non-Target Screening GmbH (AFIN-TS), Am Mittleren Moos 48, 86167 Augsburg, Germany

***Corresponding author**: catarina.cruzeiro@helmholtz-muenchen.de; Phone: +49 89 3187 49254

Abstract

Metabolomics is an emerging approach that investigates the changes in the metabolome profile. In the present study, *Lemna minor* -considered as an experimental aquatic plant model— was incubated with 10 and 100 µM diclofenac (DCF) for 96 hours, respectively. Knowing that DCF is internationally often problematic in wastewater effluents and that it might affect particularly the metabolic profiles in aquatic plants, mainly the oxidoreductase, dehydrogenase, peroxidase, and glutathione reductase activities, here it was hypothesized (H) that in the common duckweed, DCF might increase the phenolic and flavonoids pathways, as an antioxidant response to this stress (H1). Also, it was expected DCF to alternate the physiological characteristics, especially the molecular interaction and biochemical properties, of *Lemna* (H2). Metabolic changes were investigated with target and untargeted screening analysis using RPLC-HILIC-ESI-TOF-MS. Twelve amino acids were identified in all treatments, together with three organic acids (*p*-coumaric, cinnamic, and sinapic acids). In untargeted screening, the important metabolites to discriminate between different treatments were assigned to *Lemna* such as organic acids, lignin, sugars, amino acids, dipeptides, flavonoids, biflavonoids, fatty acids, among others. In resume, *Lemna* responded to both DCF concentrations, showing different stress patterns. A similar metabolic response had already been identified in other studies in exposing *Lemna* to other anthropogenic stressors (like pesticides).

Keywords: untargeted screening analysis, aquatic plants, anti-inflammatory compounds, RPLC-HILIC-ESI-TOF-MS, organic and amino acids, reducing potential

Highlights: 85 characters per sentence with spaces

1. The reducing contents increased due to diclofenac (DCF) incubation;
2. *Lemna´s* amino acid profile changed after DCF incubation;
3. *Lemna´s* metabolic profile (by untargeted screening analysis) changed upon DCF incubation.

Introduction

Metabolomics is an approach for the overall investigation of metabolite variations in biological systems and is essential to characterize these profiles as the metabolites change significantly during biochemical reactions. Consequently, metabolic profiling can be used as a robust tool to discuss the metabolic response of plants regarding environmental disorders, such as xenobiotics, nutrient deficiency, high salinity, and temperature stress (Kralova et al., 2012). As known, primary (e.g. amino acids (AA)) and specialized (e.g. fatty acids, and flavonoids) metabolites reflect the plant's functional and physiological states of the cell and organism, respectively (Wu et al., 2020).

Lemna minor (commonly named duckweed) is the largest genus of the family *Lemna*ceae and can be found in tropical and subtropical countries. *Lemna* has an ecologically important role in the absorption of excess nutrients, heavy metals, and other contaminants (Chakrabarti et al., 2018). It is a free-floating aquatic plant that is mainly used as fish and livestock feed as a source of various essential and non-essential AA, polyunsaturated fatty acids, β-carotene, and xanthophylls (Cao et al., 2018).

Plant metabolomics investigations have been formerly conducted with liquid chromatography (LC) and gas chromatography (GC) connected to the mass spectrometer (MS), respectively (Tugizimana et al., 2015; Kim J-Y et al., 2017). Moreover, a serial coupling of reversed-phase liquid chromatography (RPLC) and hydrophilic interaction liquid chromatography (HILIC) connected to a high-resolution mass spectrometer (HRMS) has been used to detect a wide range of polarities (from non-polar to high polar compounds; Bieber et al., 2017) in plant extracts in a single run (Greco et al., 2013; Wahman et al., 2019). However, metabolomics data require an analysis strategy that could recognize the changes between datasets. Thus, an untargeted screening strategy has been developed to assign such changes in metabolomics studies (Wahman et al., 2019). Using this strategy, the changes in plant metabolic profile could be investigated through two independent workflows: (1) the identification of the relevant metabolites, and (2) the statistical identification of the ‘unknown’ metabolomics indicators.

One prominent environmental problem is water pollution, especially with slow and/or non/transformed pollutants. Diclofenac (DCF) is a widely distributed non-steroidal anti-inflammatory drug that can be found in surface waters and is considered an environmental risk factor (Huber et al., 2012). Further, in 37 countries, including Germany, the average concentration of DCF is up to 1µg L^{-1} in surface water (Letzel et al., 2009; Fischer et al., 2020). The estimated removal efficacy for DCF using conventional wastewater treatment plants varies from 39% to 70% (Fischer et al., 2020). Furthermore, DCF's photo-transformation products have high potential toxicity at concentration levels close to environmental concentrations (Schmitt et al., 2007). In 2019, Alkimin and co-authors (2019) reported that DCF incubation in *Lemna* caused a decrease in the content of photosynthetic pigments, relative fluorescence decay values of chlorophyll, and oxidoreductase and dehydrogenase activities. However, it led to increases in non-photochemical quenching, amount of reactive nitrogen and oxygen species in roots, lipid peroxidation, oxidized ascorbate and thiols, and glutathione-reductase activity (Alkimin et al., 2019). These findings induce concerns regarding the chronic exposure of plants in constructed wetlands.

Knowing that DCF particularly affects the *Lemna* metabolic profiles (mainly the oxidoreductase, dehydrogenase, peroxidase, and glutathione reductase activities; Alkimin et al., 2019), here it was hypothesized (H1) that an increase of the reducing potential can be expected, as an antioxidant response to this stress. The reducing potential involves all compounds, which can reduce the Folin-Ciocalteu agent, such as phenolic compounds, ascorbic acid, dehydroascorbic acid, and reducing sugars (e.g. glucose and fructose; Rangel et al., 2013). They act as primary antioxidants or free radical scavengers. For this reason, it is reasonable to determine their total amount in the *Lemna* extracts.

Using duckweed as a potential phytoremediator (i.e., to clean the environment in long-term exposure), the plant's survival mechanisms might be affected when continuously exposed to DCF. Therefore, an alteration in its physiological characteristics, especially the molecular interaction and biochemical properties, during exposure to the recalcitrant DCF is expected (H2).

This study uses a (target and untargeted) analytical workflow and to evaluate the changes in the metabolic profile of *Lemna* when incubated with two concentrations of DCF 10 and 00 µM. Moreover, an untargeted metabolomics analysis strategy was conducted to characterize the metabolic profile of *Lemna* and reveal the changes due to DCF incubation.

2. Material and Methods

2.1. Experimental trial and sample extraction

The experiments were performed with fronds of *Lemna minor* L., grown in aquaria under controlled conditions. Each, control, 10, and 100 µM DCF treatment was maintained for four days in Steinberg medium at 23 °C with a photoperiod of 8-16 h and an average light intensity of 43 µmol m^{-2} s^{-1}, as described by Obermeier *et al*., (2015). After freeze-drying and grinding the *Lemna*, samples were extracted with 100% MeOH and 100% H_2O, separately as mentioned in (Wahman et al., 2020). Detailed experimental setup and extraction method are provided in the supplementary material.

2.2. Total reducing potential

The total reducing potential was estimated by the Folin-Ciocalteu method (Singleton *et al*., 1999) as described in Wahman et al., 2020. The quantification of phenolics was based on the standard curve (with seven nominal concentrations) generated with the use of gallic acid, and expressed as gallic acid equivalent; details are mentioned in the supplementary material as well.

2.3. Instrumental analysis and quality controls

Lemna metabolic profile was obtained using a reversed-phase column Poroshell 120 EC-C18 (50.0 × 3.0 mm, 2.7 µm; Agilent Technologies, Waldbronn, Germany) coupled to a ZIC-HILIC column (150 × 2.1 mm, 5 µm, 200 Å; Merck Sequant, Umea, Sweden) with a T-piece (Upchurch, IDEX Europe GmbH, Erlangen, Germany), which was connected to the HILIC flow pump. The reversed-phase liquid chromatography (RPLC) mobile phase was 10 mM ammonium acetate in water: acetonitrile (90:10, v/v) (A) and 10 mM ammonium acetate in water: acetonitrile (10:90, v/v) (B). For the HILIC, acetonitrile (C) and water (D) were used as a mobile phase. The pH value of both was about 7. The gradient data was done as described by (Wahman et al., 2020). The injection volume was 10 µL injected in triplicates. The ion masses were detected with a 'time-of-flight' mass spectrometer (6230 TOF-MS; Agilent Technologies, Santa Clara, CA, USA), equipped with Jet Stream ESI interface in positive electrospray ionization mode with the following parameters: 325 ºC gas temperature, 10 L min^{-1} drying gas flow, 325 ºC sheath gas temperature, 7.5 L min^{-1} sheath gas flow, 45-psi nebulizer operating pressure, and 100 V fragmentor voltage. Ions were detected in positive ionization mode with a mass range of 50-2100 Dalton. Mass accuracy calibration was performed with a reference solution that consisted of 125 nM purine and 6.25 nM HP-921 MS tuning mix (Agilent Technologies, Waldbronn, Germany) in methanol/water (90/0, v/v). The resolution of the instrument was better than 10,000 at m/z 922 (using HP-921).

Three different standard mixtures were injected in triplicate during sample analysis at regular intervals (at the beginning, middle, and end of each batch) to confirm the analytical system's

robustness. The mixtures consisted of M1 (kaempferol, rutin, and taxifolin), M2 (apigenin, resveratrol), and M3 (galangin, diclofenac, flavone, vitexin, and quercetin). All compounds were injected at a final concentration of 20 µM. The mass deviation, retention time (RT), standard deviation (SD), and relative standard deviation (RSD) were calculated and used to determine the stability, reproducibility, and accuracy of the LC system.

2.4. Metabolites identification

AA and organic acids standards were injected at a final concentration of 20 µM into RPLC-HILIC-ESI-TOF-MS. The mean monoisotopic mass and RT for each standard were calculated in Daltons and minutes, respectively. The compounds were identified when the absolute deviation in masses and RT were $\Delta ppm \leq 10$ Dalton and $\Delta RT \leq 0.3$ min., respectively. The AA in *Lemna* extracts were later confirmed using more differentiated QTOF analysis, see details in the supplementary material (Table S2).

2.5. Untargeted screening workflow

All samples were analyzed with RPLC-HILIC-ESI-TOF-MS. The data was processed with Agilent Profinder B.06.00 Software (Agilent Technologies, Santa Clara, CA, USA) to perform the peak picking. The main parameters were: ppm, which represented mass deviation (set for ±10 ppm), minimum/maximum chromatographic peak height (set with a filter of 1000 counts), and chromatographic signal-to-noise threshold (set to 3X the threshold setting of the MassHunter Data Acquisition Software). The complete workflow of data analysis with different parameters is provided in the supplementary material.

2.6. Data and Statistical analyses

The extracted data was then submitted to statistical analysis to investigate the changes in the *Lemna* metabolites profile after exposure to 10 and 100 µM DCF.

Statistical analyses were performed to assess differences in the total reducing potential, for the extracts (100% MeOH and 100% H_2O) between the untreated control and the 10 and 100 µM DCF treatments, using one-way ANOVA. The total reducing potential data were initially checked for normality (Shapiro-Wilk's test) and homogeneity of variances (Brown-Forsythe's test). The same data was transformed to try to fit the assumptions for the analysis. When not possible, a Kruskal-Wallis test was applied. Additionally, QQplots were also executed to verify the distribution functions of the statistical variables; data were also checked for outliers (ROUT alpha = 0.05). Dunn's comparison test (non-parametric) and Tukey's multiple comparison test (parametric) were applied as post-hoc tests to assess differences between treatments for the 100% MeOH and 100% H_2O extracts, individually. All statistical analyses were done using GraphPad Prism version 6.00.

The extracted compound groups were imported to Mass profile and Mass Profiler Professional Software (MPP, v.13.1.1) to start the untargeted data analysis workflow and the statistical analysis (Wahman et al., 2020); details are mentioned in the supplementary material. The obtained data set (from 100% MeOH and 100% H_2O extracts) were used to assess the *Lemna* metabolic profile for all the treatments (untreated control and incubated with 10 and 100 µM DCF, individually).

The MetaboAnalyst 4.0 software was used for the generation of Partial Least Square-Discriminant Analysis (PLS-DA) score plot, and the dendrogram using the metabolic profile of the treatments: untreated control, incubated with 10 and 100 µM DCF.

The PLS-DA was used to discriminate the similarities and differences in the metabolites profile among the different treatments. The PLS-DA module was calculated between the metabolites data (X: variables) and the permuted treatments (Y: class labels) using the optimal number of components, which were determined by cross-validation for the model based on the original class assignment (Chong et al., 2019). The most important metabolites are based on the weighted coefficients of the PLS-DA model, which were used to discriminate between different treatment metabolic profiles. Pre-selection criteria of 50 were established for the coefficients (is a relative measure of variability that indicates the size of a standard deviation to its mean) for each treatment (untreated control, 10, and 100 µM DCF). About 76 compounds could not be identified via the in-house database due to the limited size of the database. These compounds without identify suggestions are listed in the supplementary material, Table S4.

3. Results

3.1 RPLC-HILIC-ESI-TOF-MS analysis

A wide polarity range of metabolites was extracted from *Lemna* samples by using two different solvents (100% MeOH and 100% H_2O, individually), allowing untargeted metabolomics analysis. The %RSD of the RT of the quality control ranged between 0.1 and 0.3%, showing a very high chromatographic robustness. Further, the mass deviation was less than 5 ppm. The results are summarized in Table S1. Subsequently, the RTs and masses were assigned to perform the untargeted analysis of the different treatments *Lemna* underwent in this study. The *Lemna* metabolic fingerprint showed low %RSD values (< 2%) indicating the robustness and the reproducibility of the method.

The RT-mass plots displayed different patterns (i.e., dispersion of features according to polarity, Fig. S1). For both extracts, plots are divided into two parts: 1) the polar part (from 0-15 min. that is eluted from the HILIC phase material), and 2) the non-polar fraction (from 15

min until the end of the run, that is eluted from the RP column). Using the 100% H_2O extract, significantly higher amount molecules (2232) could be separated in the HILIC phase than in the 100% MeOH extract (1965). However, in the RPLC phase, more molecules were separated and detected in the 100% MeOH (1704) than in the 100% H_2O extract (986).

For the comprehensible investigation of different treatments, a scatter plot between untreated control and incubated samples (10 and 100 µM DCF) was drawn according to the average intensity of different molecules, respectively (see Fig. S2).

The scatter plot shows a change in duckweed metabolites intensities and metabolomics upon 10 and 100 µM DCF exposure (Fig. S2). For example, looking into the part I and V of this scatter plot, some compounds are absent in the untreated control samples and other new ones are present in the treated ones; detailed information can be seen in Table S3.

3.2. Total reducing potential

Considering the extracts independently (Fig. 1), the increasing pattern for the 100% MeOH extracts was observed, as untreated control < 10 µM DCF< 100 µM DCF while for the 100% H_2O extracts, *Lemna* untreated control, and the one incubated with 10 µM DCF samples presented approximately the same gallic acid equivalent content (ca. 730 µg GAE/g DW). Furthermore, the reducing potential increased significantly (about 1.8-fold) in samples incubated with 100 µM DCF (Fig. 1A).

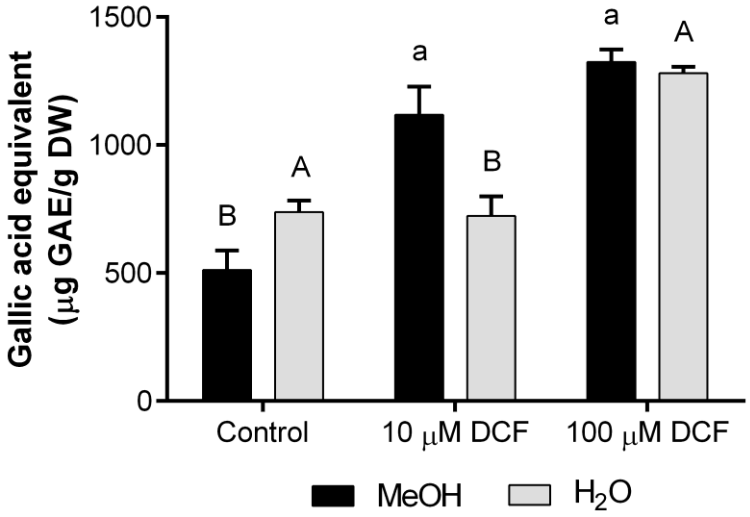


Fig. 1: Total reducing potential expressed as gallic acid equivalent (µg GAE/ g of dry weight (DW)) extracted from 100% MeOH (black) and 100% H_2O (grey) and represented by treatments (untreated control, 10 and 100 µM DCF); lower-case (100% MeOH) and upper-case (100% H_2O) letters represent statistical differences between DCF treatments and untreated control group (one-way ANOVA followed by the Dunn's multiple comparisons test); (n=9) ± SD.

232
233
234 3.3 Metabolites identification
235 The changes in *Lemna* metabolic profile were investigated with reference standards to follow
236 the changes in the intensities after 10 and 100 µM DCF. Several AA (serine, tyrosine, aspartic
237 acid, glutamic acid, proline, histidine, phenylalanine, tryptophan, valine, alanine, isoleucine,
238 leucine, and proline) were identified in duckweed samples extracted with 100% MeOH. The
239 pattern of each AA was compared in the different samples incubated with 10 and 100 µM DCF
240 against untreated control and expressed as a relative percentage against untreated control in
241 100% MeOH extracts (Table 1 and Fig. 2A).

Table 1: List of amino acids (AA) and organic acids detected in the different *Lemna* treatments (untreated control, 10 and 100 µM DCF) from 100% MeOH extract; details such as name, the monoisotopic mean mass, and the corresponding mean RT, the absolute deviation in mass (Δppm), and RT (ΔRT), the intensities (inten.) of the standards and the treated samples are represented.

Name	Standard		Control					10 µM DCF					100 µM DCF				
	Mean Mono isotopic Mass (Da)	Mean RT (min)	Mean Mono isotopic Mass (Da)	Δppm	Mean RT (min)	ΔRT	Inten.	Mean Mono isotopic Mass (Da)	Δppm	Mean RT (min)	ΔRT	Inten.	Mean Mono isotopic Mass (Da)	Δppm	Mean RT (min)	ΔRT	Inten.
Serine	105.0428	13.44	105.0433	4.13	13.45	0.02	13.74	105.0433	4.76	13.49	0.05	1.45	—	—	—	—	—
Tyrosine	181.0730	11.96	181.0726	2.49	11.94	0.02	5.67	181.0725	1.75	11.96	0.001	8.54	—	—	—	—	—
Aspartic acid	133.0375	12.71	133.0369	4.76	12.88	0.17	5.41	133.0375	0	12.83	0.12	12.93	133.0376	0.5	12.93	0.04	3.16
Glutamic acid	147.0529	12.63	147.0522	2.04	12.75	0.04	3.00	147.0520	1.81	12.73	0.02	25.90	147.0529	0	12.77	0.04	2.60
Proline	115.0636	12.17	115.0633	2.35	12.03	0.14	1.76	115.0632	3.48	12.40	-0.24	3.33	115.0634	-0.32	12.36	-0.19	31.75
Histidine	155.0687	15.35	155.0696	6.02	14.65	0.7	9.42	155.0692	3.44	15.05	0.3	12.83	155.0704	3.44	15.23	0.11	3.11
Phenylalanine	165.0787	10.71	165.0786	2.30	10.70	0.01	0.16	165.0783	4.32	10.70	0.01	3.68	165.0799	5.77	10.67	0.04	7.27
Tryptophan	204.0899	11.45	204.0900	0.82	11.43	0.03	0.84	204.0898	0.33	11.43	0.02	0	204.0908	4.25	11.46	0.01	3.46
Valine	117.079	11.64	117.0790	0.28	11.60	0.03	3.52	117.0791	0.85	11.62	0.02	4.60	117.0796	5.12	11.71	0.07	3.68
Alanine	89.0477	12.84	89.0478	1.12	13.11	0.27	0.89	89.0475	2.25	12.99	0.15	12.46	89.0479	1.87	12.89	0.06	1.823
Isoleucine	131.0944	10.82	131.0949	4.07	10.87	0.04	12	131.0947	2.8	10.83	0.01	13.33	131.0947	2.8	10.67	0.16	13.12
Leucine	131.0941	10.79	131.0943	1.78	10.98	0.19	3.07	131.0947	4.83	10.68	0.11	9.94	131.0947	4.83	10.67	0.12	10.43
Cinnamic acid*	148.0519	10.61	148.0521	1.35	10.6	0.01	4.55	148.0520	0.90	10.52	0.09	5.15	148.0523	2.70	10.50	0.11	3.08
***p*-coumaric acid**	164.0468	7.72	164.0479	6.50	7.75	0.03	2.00	164.0472	2.64	7.80	0.05	1.00	—	—	—	—	—
Sinapic acid	224.0683	11.95	224.0677	2.83	12.14	0.19	2.69	224.0679	1.79	12.17	0.22	6.87	—	—	—	—	—

The symbol (—) means below the LOD
**Identified only in the 100% H_2O extract*

The relative amount of tyrosine increased in *Lemna* 100% MeOH incubated with 10 µM DCF, while it decreased under the measurable limit in 100 µM DCF incubation. The relative amount of serine decreased in *Lemna* incubated with 10 µM DCF and fell below the LOD for the 100 µM DCF.

Further, aspartic acid, glutamic acid, histidine, valine, and alanine patterns increased in *Lemna* 100% MeOH incubated with 10 µM DCF and decreased in *Lemna* incubated with 100 µM DCF. For phenylalanine, proline, leucine, and isoleucine levels increased in both incubated samples. They were directly proportional to DCF concentration. However, tryptophan decreased in *Lemna* incubated with 10 µM DCF and increased in the 100 µM incubated sample.

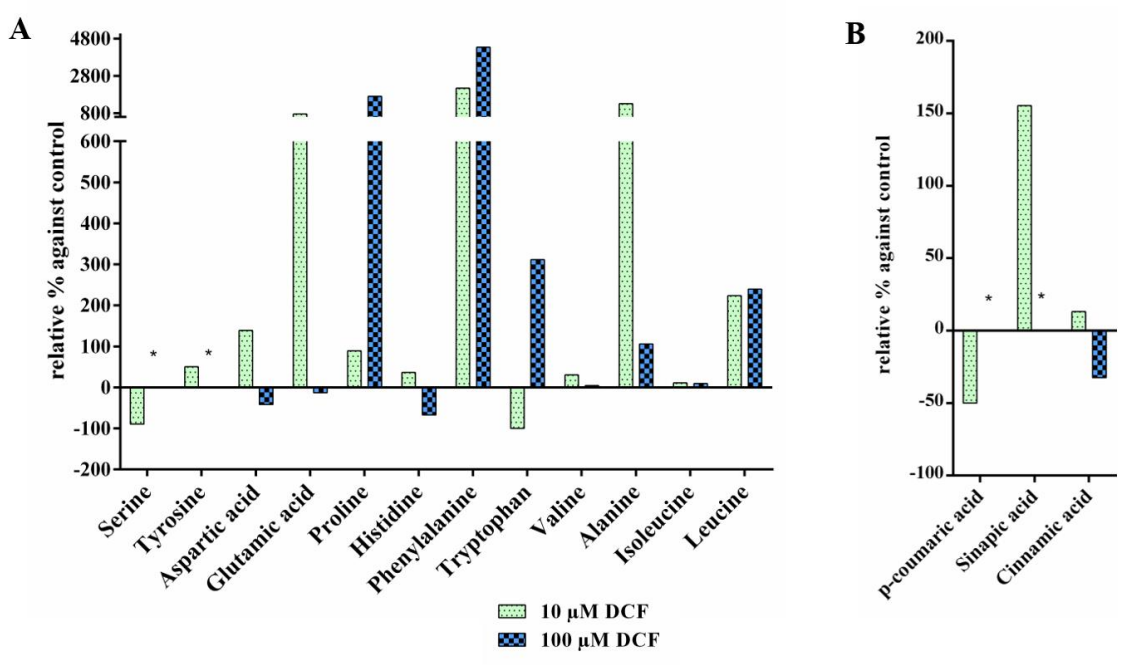


Fig. 2: Relative intensities (%) of different metabolites identified in 10 and 100 µM DCF treatments against the untreated control; A) amino acids (AA); B) organic acids; the asterisk symbol (*) means intensities below the LOD.

Regarding organic acids, *p*-coumaric and sinapic acids were identified in all the samples extracted in 100% MeOH. In 10 µM DCF, these compounds presented opposite pattern; i.e., *p*-coumaric acid decreased while the sinapic acid increased. For 100 µM DCF, both were below the LOD. For cinnamic acid, it was increased and decreased in the 10 µM DCF and 100 µM DCF in 100% H_2O extract, respectively, as shown in Table 1 and displayed in Fig. 2B.

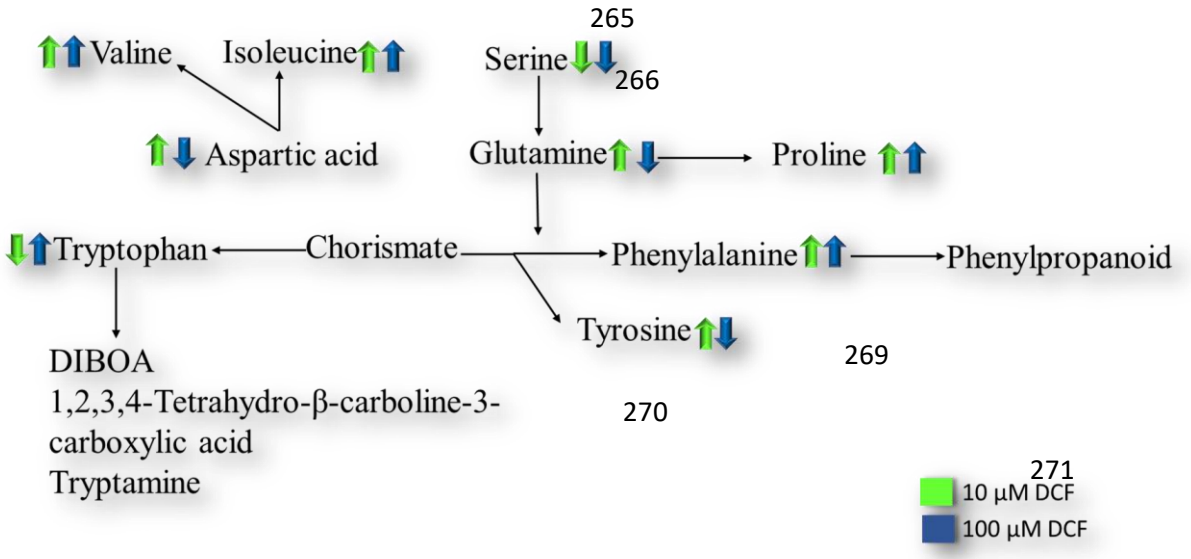


Fig.3: The schematic of the investigated amino acids biosynthesis pathways in *Lemna* exposed to DCF; the green and blue colors represent the *Lemna* incubated with 10 µM DCF and 100 µM DCF, respectively.

3.4 Alterations in *Lemna* metabolic profile

The present study was built on the LC-MS metabolomics analysis to investigate the metabolic profile of *Lemna* changes due to incubation with 10 and 100 µM DCF. After assigning an accurate mass, 1177 metabolites were selected and used to differentiate between untreated control, and 10 and 100 µM DCF treatments, which are represented in separated clusters in the PLS-DA plot (Fig. 4).

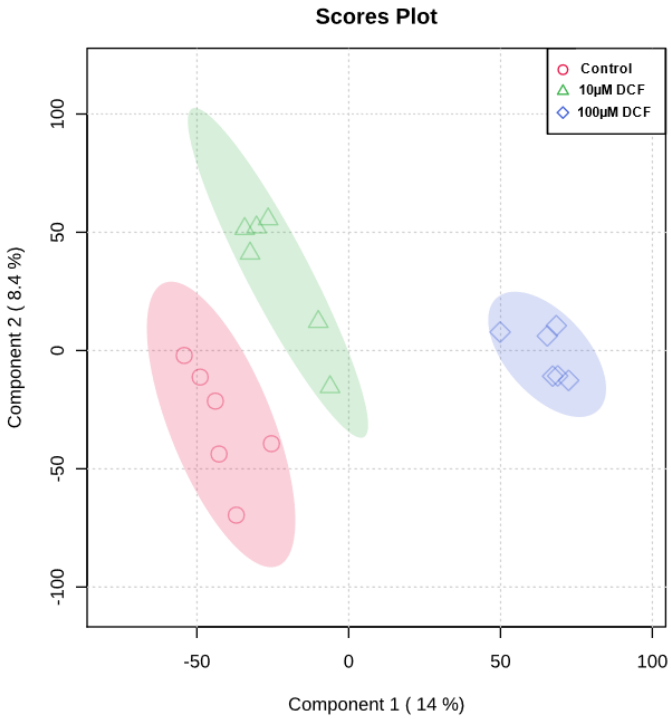


Fig. 4 PLS-DA Scores plot of *Lemna* untreated control (red) and 10 and 100 µM DCF treatments (green and blue, respectively) between the first and second components using both 100% MeOH and 100% H_2O extracts. The explained variances are shown in brackets (95% confidence level); circles represent untreated control, triangles represent 10 µM DCF, and diamonds represent 100 µM DCF treatment samples (n=6).

The first and second significant PLS components explained 14.0% and 8.4% of the total variance, respectively. The separation was observed clearly between treatments being more evident for the 100 µM DCF incubation. Within treatments, a high variation was observed among replicates due to the two different types of extraction (100% H_2O and 100% MeOH) and biological variability.

The same data was plotted in a dendrogram of hierarchical cluster analysis for detailed visualization to illustrate the correlation between different extracts and treatments (Fig. S3). The hierarchical cluster analysis demonstrates the individual extracts behind the differentiation between the treatments.

The contribution of metabolites to the discrimination between the *Lemna* untreated control and incubated with DCF metabolic profiles is based on the coefficients of the PLS-DA, which were given in (Table 2), for more information readers referred to supplementary material. A total of 108 compounds were selected, through the coefficients of PLS-DA, as important metabolites in the discrimination between the different treatments. 30% of them were detected as organic acids, lignin, sugars, AA, dipeptide, flavonoids, bioflavonoids, fatty acids, and miscellaneous (Table 2).

The concentrations (intensities) according to the Metaboanalyst software varied among treatments:

- gallic acid and truxillic acid, N-alanyl-alanine (dipeptide), Luteolin-8-C-beta-D-glucopyranoside (flavonoid), propelargonidin and n-pentadecanoic acid (fatty acids), 1,2,3,4-tetrahydro-β-carboline-3-carboxylic acid, decreased due to incubation with DCF;
- syringaresinol (lignan), D-hexose and ribose (sugars), myricetin, vicenin-1 and pentamethoxyflavone (flavonoids), 1-pentadecane carboxylic acid and stearamide (fatty acids), and 2,4-dihydroxy-1,4-benzoxazine-one (DIBOA) increased in the 10 µM DCF treatment;
- 3-heptadecyl-5-methoxy phenol, 3-ketosphingosine, 5,8,11,14-eicosatetraenoic acid, oleamide, C14 fatty acids, and undecenoic acid derivatives (all fatty acids) increased in 100 µM DCF mainly to untreated control.

Table 2. The relevant coefficients of compounds, produced by *Lemna*, indicate the effect of DCF incubation. The range of concentrations (intensities) (low to high) is represented by a designated color of the bar.

Compound name	mean	Coefficients		
		untreated control	10 µM DCF	100 µM DCF
Organic acids				
Gallic acid	34.44	50.18	41.02	12.13
Hydroxymethyl furfural	55.57	43.52	40.95	82.25
Nicotinic acid	80.85	98.57	80.18	63.81
Truxillic acid	68.88	100.00	76.91	29.72
Lignan				
Syringaresinol	26.99	12.03	16.01	52.95
Symplocosin	45.59	39.18	34.92	62.66
Miscellaneous				
1,2,3,4-Tetrahydro-β-carboline-3-carboxylic acid	52.53	73.70	58.07	25.83
N,N-Dimethyltryptamine	47.08	47.80	41.05	52.39
5-Hydroxytryptamine	25.27	8.70	13.06	54.04
2,4-Dihydroxy-1,4-benzoxanzin-34H-one (DIBOA)	49.94	46.62	42.31	60.90
Sugars				
D-Hexose	47.52	36.57	35.42	70.58
Ribose	53.29	47.10	42.01	70.75
Amino acids				
Glutamine	30.10	8.59	13.22	68.48
Histidine	78.13	74.23	65.63	94.53
Dipeptide				
N-Alanyl-alanine	41.34	69.82	52.59	1.62
Flavanoids				
Luteolin-8-C-beta-D-glucopyranoside	50.64	73.15	57.61	21.15
Myricetin	40.05	65.73	47.56	6.86
Vicenin 1	35.54	56.51	42.76	7.34
Pentamethoxyflavone	38.40	19.09	21.61	74.51
Biflavanoids				
Propelargonidin	28.27	14.80	17.97	52.04
Sciadopitysin	26.78	12.71	16.31	51.32
Fatty acids				
Propylamylcarbinol	31.03	51.78	39.72	1.60
1-Monopalmitoylglycerol	30.40	14.68	18.90	57.63
1-Pentadecanecarboxylic acid	50.77	74.25	55.27	22.79
3-Heptadecyl-5-methoxyphenol	33.27	14.95	18.71	66.15

(continued)

Compound name	mean	Coefficient		
		untreated control	10 µM DCF	100 µM DCF
3-ketosphingosine*	41.76	50.91	39.88	34.47
5,8,11,14-Eicosatetraenoicacid	46.44	45.32	40.31	53.69
Oleamide*	53.18	75.28	56.70	27.55
n-Pentadecanoic acid	42.19	53.87	43.11	29.59
Stearamide*	34.71	50.99	38.11	15.03
Undecenoic acid derviatives*	42.84	57.21	46.14	25.18
C14 fatty acid	51.26	71.98	54.30	27.49

4. Discussion

In the last decades, metabolomics has been expanded especially in plant science. This approach has been used to illustrate the stress response, characterization of biomonitors, and identifying the influential metabolites (Kumar et al., 2017). The present study investigated the potential alteration in *Lemna* metabolic profile after incubation with DCF (10 and 100 µM) over four days based on the fact that a lower concentration of DCF (20 µg L^{-1}) can already act as a signal, and might affect the circadian expression of reactive oxygen species response selected genes in lettuce (Bigott et al., 2021).

The extraction was performed with 100% MeOH and 100% H_2O to study the metabolic profile due to incubation. The large difference in the polarity ensured the extraction of a wide range of metabolites. Further, the LC-system consisted of a two-column setup in serial coupling which allowed separation of a wide range of different metabolites, from very-polar to non-polar ones (Wahman et al., 2019; Wahman et al., 2020) in positive ionization mode, which is sufficient for untargeted screening strategy to obtain a global overview of the differences and similarities between samples, as suggested by De Vos and coworkers (2012).

The scatter plot (Fig S2) showed an alteration in the metabolic profile as expected; hence, some compounds disappeared and others appeared after DCF incubation; these results show that plants were capable of dealing with DCF, even at higher concentrations. Duckweed increased the production of stress-defensive compounds such as fatty acids, saturated carboxylic acids, and flavonoids (Vladimirova and Georgiyants, 2014), as we expected (H1), which was already mentioned in other studies using duckweed. For example, in 2012, Forni and coworkers revealed that *Lemna* has higher production of phenols during the first three days of treatment with sodium dodecyl sulfate (an anionic surfactant) with a higher concentration (up to 50 mg L^{-1}) than when exposed for seven days to a lower dosage (25 mg L^{-1}), even if the phenol content was lower than the untreated control for the last one. Further, in 2013, Varga

and co-authors showed that the total phenolic content of *Lemna* changed significantly when exposed, for 24 hours, to Hg in comparison to other metals (Cd or Cr) at concentrations ranging from 0.02 to 20 mg L^{-1}. Furthermore, in 2020, Kostopoulou and coworkers reported that *Lemna* increases the production of aromatic amino acids (AAAs) after incubation with drugs such as glyphosate, metribuzin, and their mixture.

In the same way, the AA profile also changed. In plants, serine biosynthesis proceeds by different pathways. As known, plants use the phosphorylated pathway in response to an infection, and to environmental and abiotic stresses (Igamberdiev and Kleczkowski, 2018). In this pathway, plants produce glutamate, which, in turn, synthesizes proline. In plants, intracellular proline levels have been found to increase by more than 100-fold during stress, as observed in 100 µM DCF treatment. Rhodes and co-authors (1986), suggested that proline increased after exposure of *Lemna* to methionine sulfoximine. Later, it was been proven that the concentration of proline increased because the glutamate pool served as its precursor (Delauney and Verma, 1993). Also, the glutamic acid concentration decreased in *Lemna* incubated with metribuzin, glyphosate, and their mixtures for 72 hours (Kostopoulou et al., 2020); this may explain the decrease in the concentration of glutamic acid in our data for 100 µM DCF treatments.

Tryptophan, tyrosine, and phenylalanine are the AAAs, which are required for protein biosynthesis in all living cells. In plants, AAAs serve as precursors of a wide variety of plant natural products that play crucial roles in plant growth, development, reproduction, defense, and environmental responses such as alkaloids, phytoalexins, and indole glucosinolates as well as numerous phenolic compounds (Maeda and Dudareva, 2012). The increase in phenylalanine pattern was observed when incubated with DCF, which was accompanied by an increase of cinnamic acid, supporting once again H1. Cinnamic acid is the first compound in the phenylpropanoid pathway that begins with the deamination of phenylalanine (Maeda and Dudareva, 2012). In light of the current data, it seems that the phenylpropanoid pathway is induced, which protects *Lemna* from oxidative stress against DCF, increasing the patterns of myricetin, syringaresinol, pentamethoxyflavone, and vicenin 1. Phenylpropanoids also increased to protect *Lemna* against reactive oxygen species (ROS) (Buchanan et al., 2015). The biflavonoids pattern decreased in both treatments reaching lower levels for the highest concentration of treatment (100 µM DCF). The flavonoids content also decreased, when *Lemna gibba* was exposed to several environmental challenges (stressors); according to Akhtar and co-authors (2010), this might be a result of the promotion of the photosynthetic electron transport chain reduction, causing flavonoid reduction. Moreover, *p*-coumaric and sinapic acid decreased

for the 100 µM treatment, which can be related to the alternative usage of NADPH to degrade DCF instead of producing these organic acids (Huber et al., 2012); the same profile was also reported by Kostopoulou et al., (2020).

It has been known that the utilization of glutamate decreases the intensity of tryptophan since it uses glutamine in its biosynthesis (Delauney and Verma, 1993). The same was observed in the current study with *Lemna* incubated with 10 µM DCF. Interestingly, the opposite behavior was observed for the plants incubated with 100 µM DCF, which might be due to over-expression of tryptophan synthase β gene-like like it happened for plants treated with Cd (Sanjaya et al., 2008). Higher DIBOA levels were observed for 10 µM DCF incubation. As we know, DIBOA is synthesized by tryptophan synthase α that catalyzes its formation from indole (as a precursor) instead of tryptophan (Frey et al., 1997; Buchanan et al., 2015).

Aspartic acid increased in *Lemna* incubated with10 µM DCF, which might be due to an increase in isoleucine and valine synthesis. Rhodes and coworkers (1986), expected an increase of isoleucine due to protein degradation, in *Lemna* exposed to methionine sulfoximine, for 24 hours. However, a recent study indicated that the increase is due to aspartic acid catabolism, such as in a study where *Arabidopsis* was subjected to bacterial infection (Yang and Ludewig, 2014). However, for the 100 µM DCF incubation, the opposite was observed for aspartic acid. This might be a consequence of the highest concentration of DCF due to the usage of aspartic acid to provide energy to the tricarboxylic acid cycle (TCA), as a response to abiotic stress (Galili, 2011).

The alanine pattern fluctuated significantly between *Lemna* incubated with 10 µM and 100 µM DCF. The same was observed in plant response to hypoxic stress, which was accompanied by higher rates of glycolysis and ethanol fermentation causing fast depletion of sugar stores and carbon stress (Limami et al., 2008). Consequently, in 10 µM DCF incubation, hexoses such as glucose, galactose, mannose, raffinose, and ribose, potentially increased due to higher energy demand to tolerate the stress effect of DCF incubation through exhaustion of sugar stores. Also, it has been reported that non-soluble sugars have osmoprotectant and antioxidant activities (Sivaram et al., 2019). However, in 100 µM DCF incubation, their concentrations decreased which might be due to the fluctuation of sugars due to mechanisms that are affected by changes in the genotype (e.g., INV, SuSy, ATB2 bZIP, and α-amylase gene) and stress factors (Rosa et al., 2009).

Similar patterns were also reported for *Lemna* exposed to pesticides (Kostopoulou et al., 2020). In that experiment, aspartic acid, isoleucine, and valine concentrations increased after 72 hours

of exposure, while for alanine, the scenario was not consistent, changing among pesticide treatments.

The untargeted screening can assign important variables without reference standards. In the current study, in untreated control samples, gallic, nicotinic, and n-pentadecanoic acids have a higher concentration of incubation than 100 µM DCF followed by 10 µM DCF.

Gallic acid plays an important role in plant defense against stress conditions. This finding is illustrated by the increase in phenylalanine and the concomitant decrease in *p*-coumaric, sinapic, and truxillic acids in DCF incubated samples. Gallic acid has two pathways either through phenylalanine and 5-dehydroshikimic acid. It seems that *Lemna* preferred the formation of phenylalanine → protocatechuic acid → gallic acid through the β-oxidative pathway when incubated with DCF. This pathway was observed in the mutant strain of *Neurospora crassa* that blocks the conversion of 5-dehydroshikimic acid into shikimic acid leading to a strict production of gallic acid from protocatechuic acid (Dewick and Haslam, 1969).

The syringaresinol that is biosynthesized from sinapic alcohol by peroxidase enzymes (Habib et al., 2018), increased in 10 µM DCF treatment; this is an expected result of activation of peroxidase and oxidase enzymes in *Lemna* by DCF (Alkimin et al., 2019).

Moreover, the sinapic acid intensity increased together with syringaresinol, showing that 10 µM DCF evoked the phenylpropanoid pathway as expected (H1 and H2). Regarding 100 µM DCF treatment, the saturated and unsaturated fatty acids, and sphingosine concentrations increased to protect the *Lemna* against the ROS because of stress degradation, which are produced in the presence of DCF and/or its transformation products under a high dose of DCF (Alkimin et al., 2019). Therefore, the n-pentadecanoic acid concentration also increased.

Saturated and unsaturated fatty acids induce broad-spectrum resistance against infections in the plant, such as *Pseudomonas syringae* in tomatoes (Lim et al., 2017). Some of them (stearic, oleic, and palmitic acids) increased in *Lemna* incubated with glyphosate, and glyphosate metribuzin mixture for 72 hrs (Kostopoulou et al., 2020); this might explain the increase in reducing potential in the 100 µM DCF treatment, because of unsaturated fatty acid formation/accumulation.

5. Conclusions

Metabolomics analysis of *Lemna* incubated with 10 and 100 µM DCF was performed using RPLC-HILIC-ESI-TOF-MS. The spectral and statistical results showed changes in the metabolic profile of *Lemna* due to DCF incubation after 4 days. The PLS-DA analysis identified

the significant differences between the untreated control, samples treated with 10 and 100 µM DCF, resepectively. The AA and organic acids were exhibited changes in their intensities as a response to DCF incubation, where were most of them occurred for the highest concentration, being related to stress defense mechanisms. However, it could not explain the causality, in some points, as for the organic acids. The untargeted strategy enabled the investigation of changes in *Lemna* metabolic profile such as organic acids, lignin, sugars, AA, dipeptides, flavonoids, bioflavonoids, fatty acids, and some others more. Hence, untargeted metabolomics has a fundamental function in determining the metabolic changes in plants due to xenobiotics exposure. These results provided insights into untargeted metabolomics as they serve as a workflow to monitor the changes in *Lemna* metabolic profile that can be extended to target quantitative studies In conclusion, *Lemna* responded differently to both treatments, showing that concentrations have a great impact on the metabolic profile of this aquatic plant, as it was also observed with other anthropogenic stressors (like pesticides). Such a difference in reactions might influence the efficiency of phytoremediation or productivity of aquatic species.

6. CRediT author statement

R. Wahman: conceptualization, methodology, validation, formal analysis, investigation, writing-original draft, visualization; **C. Cruzeiro:** conceptualization, formal analysis, writing-review and editing, visualization, supervision; **J. Grassmann:** writing-review and editing, supervision; **P. Schröder:** conceptualization, methodology, resources, writing-review, and editing, supervision; **T. Letzel:** conceptualization, methodology, resources, writing-review and editing, funding acquisition, supervision.

7. Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

8. Acknowledgments

This research was partially funded by the project “Novel analysis strategies for pollutants in plants” (TLK01U-69448), financed by the Bavarian State Ministry of the Environment and Consumer Protection (BstMUG).

R Wahman was granted by the Cultural Affairs and Mission Sector of the Egyptian Ministry of Higher Education and C Cruzeiro was funded by the Water Joint Programming Initiative

(WATER 21015 JPI) through the European research project IDOUM - Innovative Decentralized and low-cost treatment systems for Optimal Urban wastewater Management.

7. References

Akhtar, T.A., Lees, H.A., Lampi, M.A., Enstone, D., Brain, R.A., Greenberg, B.M., 2010. Photosynthetic redox imbalance influences flavonoid biosynthesis in *Lemna gibba*. Plant Cell Environ 33, 1205-1219.

Alkimin, G.D., Daniel, D., Dionisio, R., Soares, A., Barata, C., Nunes, B., 2019. Effects of diclofenac and salicylic acid exposure on *Lemna minor*: Is time a factor? Environ Res 177, 108609.

Bieber, S., Greco, G., Grosse, S., Letzel, T., 2017. RPLC-HILIC and SFC with Mass Spectrometry: Polarity-Extended Organic Molecule Screening in Environmental (Water) Samples. Anal Chem 89, 7907-7914.

Bigott, Y., Chowdhury, S.P., Pérez, S., Montemurro, N., Manasfi, R., Schröder, P., 2021. Effect of the pharmaceuticals diclofenac and lamotrigine on stress responses and stress gene expression in lettuce (*Lactuca sativa*) at environmentally relevant concentrations. Journal of hazardous materials 403, 123881.

Buchanan, B.B., Gruissem, W., Jones, R.L., 2015. Biochemistry and Molecular Biology of Plants. Wiley.

Cao, H.X., Fourounjian, P., Wang, W., 2018. The importance and potential of duckweeds as a model and crop plant for biomass-based applications and beyond. in: Hussain, C.M. (Ed.). Handbook of Environmental Materials Management. Springer International Publishing, Cham, pp. 1-16.

Chakrabarti, R., Clark, W.D., Sharma, J.G., Goswami, R.K., Shrivastav, A.K., Tocher, D.R., 2018. Mass production of *Lemna minor* and its amino acid and fatty acid profiles. Front Chem 6.

Chong, J., Wishart, D.S., Xia, J., 2019. Using MetaboAnalyst 4.0 for comprehensive and integrative metabolomics data analysis. Current Protocols in Bioinformatics 68, e86.

De Vos, R., Schipper, B., Hall, R., 2012. High-Performance Liquid Chromatography-Mass Spectrometry Analysis of Plant Metabolites in Brassicaceae. Methods in molecular biology (Clifton, N.J.) 860, 111-128.

Delauney, A.J., Verma, D.P.S., 1993. Proline biosynthesis and osmoregulation in plants. The Plant Journal 4, 215-223.

Dewick, P.M., Haslam, E., 1969. Phenol biosynthesis in higher plants. Gallic acid. Biochem J 113, 537-542.

Fischer, K., Sydow, S., Griebel, J., Naumov, S., Elsner, C., Thomas, I., Abdul Latif, A., Schulze, A., 2020. Enhanced removal and toxicity decline of diclofenac by combining UVA treatment and adsorption of photoproducts to polyvinylidene difluoride. Polymers 12, 2340. https://doi.org/10.3390/polym12102340

Forni, C., Braglia, R., Harren, F.J.M., Cristescu, S.M., 2012. Stress responses of duckweed (*Lemna minor* L.) and water velvet (*Azolla filiculoides* Lam.) to anionic surfactant sodium-dodecyl-sulfate (SDS). Aquatic Toxicology 110-111, 107-113.

Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grün, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R.B., Briggs, S.P., Simcox, K., Gierl, A., 1997. Analysis of a chemical plant defense mechanism in grasses. Science (New York, N.Y.) 277, 696-699.

Galili, G., 2011. The aspartate-family pathway of plants: linking production of essential amino acids with energy and stress regulation. Plant signaling and behavior 6, 192-195.

Greco, G., Grosse, S., Letzel, T., 2013. Serial coupling of reversed-phase and zwitterionic hydrophilic interaction LC/MS for the analysis of polar and nonpolar phenols in wine. Journal of Separation Science 36, 1379-1388.

Letzel, M., Metzner, G., Letzel, T., 2009. Exposure assessment of the pharmaceutical diclofenac based on long-term measurements of the aquatic input. Environment International 35, 363-368.

Limami, A.M., Glévarec, G., Ricoult, C., Cliquet, J.B., Planchet, E., 2008. Concerted modulation of alanine and glutamate metabolism in young *Medicago truncatula* seedlings under hypoxic stress. J Exp Bot 59, 2325-2335.

Habib, M., Trajkovic, M., Fraaije, M.W., 2018. The biocatalytic synthesis of syringaresinol from 2,6-dimethoxy-4-allylphenol in one pot using a tailored oxidase/peroxidase system. ACS Catalysis 8, 5549-5552.

Huber, C., Bartha, B., Schröder, P., 2012. Metabolism of diclofenac in plants-hydroxylation is followed by glucose conjugation. Journal of Hazardous Materials 243, 250-256.

Igamberdiev, A., Kleczkowski, L., 2018. The glycerate and phosphorylated pathways of serine synthesis in plants: The branches of plant glycolysis linking carbon and nitrogen metabolism. Frontiers in Plant Science 9: 318.

Kim J-Y, Kim H-Y, Jeon J-Y, Kim D-M, Zhou Y, Lee JS, 2017. Effects of coronatine elicitation on growth and metabolic profiles of *Lemna paucicostata* culture. PloS one 12.

Kostopoulou, S., Ntatsi, G., Arapis, G., Aliferis, K.A., 2020. Assessment of the effects of metribuzin, glyphosate, and their mixtures on the metabolism of the model plant *Lemna minor* L. applying metabolomics. Chemosphere 239, 124582.

Kralova, K., Jampilek, J., Ostrovsky, I., 2012. Metabolomics - Useful tool for the study of plant responses to abiotic stresses. Ecological Chemistry and Engineering S 19, 133-161.

Kumar, R., Bohra, A., Pandey, A.K., Pandey, M.K., Kumar, A., 2017. Metabolomics for plant improvement: Status and prospects. Frontiers in Plant Science 8. 1302. DOI: 10.3389/fpls.2017.01302

Lim, G.H., Singhal, R., Kachroo, A., Kachroo, P., 2017. Fatty acid- and lipid-mediated signaling in plant defense. Annual Review of Phytopathology 55, 505-536.

Maeda, H., Dudareva, N., 2012. The shikimate pathway and aromatic amino acid biosynthesis in plants. Annual Review of Plant Biology 63, 73-105.

Obermeier, M., Schröder, C.A., Helmreich, B., Schröder, P., 2015. The enzymatic and antioxidative stress response of *Lemna minor* to copper and a chloroacetamide herbicide. Environmental Science and Pollution Research 22, 18495-18507.

Rangel, J.C., Benavides Lozano, J., Heredia, J., Cisneros-Zevallos, L., Jacobo-Velázquez, D., 2013. The Folin-Ciocalteu assay revisited: Improvement of its specificity for total phenolic content determination. Analytical Methods 5, 5990.

574 Rhodes, D., Deal, L., Haworth, P., Jamieson, G.C., Reuter, C.C., Ericson, M.C., 1986. Amino
575 acid metabolism of *Lemna minor* L.: I. Responses to methionine sulfoximine. Plant
576 Physiology 82, 1057-1062.

577 Rosa, M., Prado, C., Podazza, G., Interdonato, R., González, J.A., Hilal, M., Prado, F.E., 2009.
578 Soluble sugars--metabolism, sensing and abiotic stress: a complex network in the life of
579 plants. Plant signaling and behavior 4, 388-393.
580

581 Sanjaya, Hsiao, P.Y., Su, R.C., Ko, S.S., Tong, C.G., Yang, R.Y., Chan, M.T., 2008.
582 Overexpression of *Arabidopsis thaliana* tryptophan synthase beta 1 (AtTSB1) in Arabidopsis
583 and tomato confers tolerance to cadmium stress. Plant Cell Environ 31, 1074-1085.

584 Schmitt, M., Bartels, P., Adler, N., Altenburger, R., 2007. Phytotoxicity assessment of
585 diclofenac and its phototransformation products. Analytical and Bioanalytical Chemistry 387,
586 1389-1396.

587 Singleton, V.L., Orthofer, R., Lamuela-Raventós, R.M., 1999. [14] Analysis of total phenols
588 and other oxidation substrates and antioxidants by means of folin-ciocalteu reagent. Methods
589 in Enzymology. Academic Press, pp. 152-178.Sivaram, A.K., Subashchandrabose, S.R.,
590 Logeshwaran, P., Lockington, R., Naidu, R., Megharaj, M., 2019. Metabolomics reveals
591 defensive mechanisms adapted by maize on exposure to high molecular weight polycyclic
592 aromatic hydrocarbons. Chemosphere 214, 771-780.

593 Thelusmond, J.-R., Kawka, E., Strathmann, T.J., Cupples, A.M., 2018. Diclofenac,
594 carbamazepine, and triclocarban biodegradation in agricultural soils and the microorganisms
595 and metabolic pathways affected. Science of the Total Environment 640-641, 1393-1410.

596 Tugizimana, F., Ncube, E.N., Steenkamp, P.A., Dubery, I.A., 2015. Metabolomics-derived
597 insights into the manipulation of terpenoid synthesis in *Centella asiatica* cells by methyl
598 jasmonate. Plant Biotechnology Reports 9, 125-136.

599 Varga, M., Horvatić, J., Čelić, A., 2013. Short-term exposure of *Lemna minor* and *Lemna gibba*
600 to mercury, cadmium, and chromium. Central European Journal of Biology 8, 1083-1093.

601 Wahman, R., Graßmann, J., Sauvêtre, A., Schröder, P., Letzel, T., 2020. *Lemna minor* studies
602 under various storage periods using extended-polarity extraction and metabolite non-target
603 screening analysis. Journal of Pharmaceutical and Biomedical Analysis 188, 113362.

604 Wahman, R., Grassmann, J., Schröder, P., Letzel, T., 2019. Plant metabolomic workflows using
605 reversed-phase LC and HILIC with ESI-TOF-MS. LCGC North America 37, 8-15.Wu, Q.,
606 Zhao, X., Chen, C., Zhang, Z., Yu, F., 2020. Metabolite profiling and classification of
607 developing *Styrax tonkinensis* kernels. Metabolites 10, 21.
608

609 Yang, H., Ludewig, U., 2014. Lysine catabolism, amino acid transport, and systemic acquired
610 resistance: what is the link? Plant signaling and behavior 9, e28933-e28933.
611

612 Zhang, Y., Hu, Y., Yang, B., Ma, F., Lu, P., Li, L., Wan, C., Rayner, S., Chen, S., 2010.
613 Duckweed (*Lemna minor*) as a model plant system for the study of human microbial
614 pathogenesis. PloS one 5, e13527-e13527.

615