

TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Analytische Lebensmittelchemie

Mass-spectrometric characterization of stress-induced metabolite profiles in poplar

Moritz Kaling

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. E. Grill

Prüfer der Dissertation:

1. apl. Prof. Dr. Ph. Schmitt-Kopplin
2. Univ.-Prof. Dr. M. Rychlik
3. apl. Prof. Dr. J. P. Schnitzler

Die Dissertation wurde am 27.05.2016 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 26.09.2016 angenommen.

To my family

Summary

As one of the fastest growing tree species with good regeneration from rootstocks, poplars are the perfect fit for biofuel, wood and fibre production. Poplar plantations are, however, subjected to a variety of environmental stresses, which can drastically alter plant fitness and biomass production. Additionally, poplars emit the volatile organic compound isoprene, which is a precursor of ozone and secondary aerosols in the troposphere. Genetically modified non-isoprene emitting poplars could be a well-fitting alternative to isoprene emitting poplars normally used for biomass production. However, isoprene is an important molecule in the stress response of poplars and how the lack of isoprene emission affects the performance of poplars, is functionally and biochemically not entirely understood.

Metabolites are a direct result of changes in gene, transcript and protein expressions. Thus they represent the perfect read-out for studying the impact of environmental factors as well as genetic modifications on plant performance. The scientific discipline, which aims at measuring the complete set of metabolites within an organism, the metabolome, is non-targeted metabolomics.

This thesis aims at elucidating the influence of environmental stress factors on the primary and secondary metabolism of wild type *Populus x canescens* (grey poplar), a naturally occurring hybrid, and its non-isoprene emitting mutants generated by RNA interference. Non-targeted metabolomics using high resolution mass-spectrometers, such as the ICR-FT/MS and the UPLC-qToF-MS, have been utilized in the present thesis studying the metabolic response of grey poplars to (i) high ultraviolet-B irradiation, (ii) knock-down of isoprene synthase, and (iii) to leaf herbivory with and without ectomycorrhiza. Multivariate statistics and mass-difference analysis enabled the extraction and visualization of specific and comprehensive metabolic perturbations for an enhanced phenotypic description.

In summary, this thesis provides a comprehensive overview how high resolution non-targeted metabolomics is efficiently used to deepen our understanding of plant metabolic adaption to environmental stimuli.

Zusammenfassung

Die Pappel zählt zu den schnellwachsenden Baumarten und ist daher perfekt geeignet für die Produktion von Biotreibstoff, Holz und Fasern. Jedoch wird die Ausbeute von Pappelplantagen von negativen Umwelteinflüssen stark beeinträchtigt. Ein weiterer Nachteil ist, dass Pappeln starke Isopren Emitter sind. Isopren ist eine flüchtige organische Verbindung, welche die Bildung von Ozon und sekundären Aerosolen in der Troposphäre begünstigt. Genetisch veränderte nicht Isopren emittierende Pappeln wären eine perfekte Alternative für die Produktion von Biomasse. Allerdings erfüllt Isopren wichtige Funktionen unter Stressbedingungen und es ist daher unklar wie sich der Verlust von Isopren auf die Fitness und Wuchsleistung der Pappel auswirkt.

Metabolite befinden sich am Nächsten zum zellulären Phänotyp und sind das direkte Resultat von Änderungen in der Gen-, Transkript- und Proteinexpression. Die wissenschaftliche Disziplin welche sich mit der Messung des gesamten Metaboloms in einem Organismus beschäftigt ist die ungerichtete Metabolomik.

Das Ziel dieser Doktorarbeit ist es den Einfluss von verschiedenen Umwelteinflüssen auf den Primär- und Sekundärstoffwechsel der Graupappel (*Populus x canescens*), einem natürlich vorkommendem Hybrid, und deren nicht Isopren emittierenden Mutanten, die über RNA Interferenz erzeugt wurden zu untersuchen. Dabei werden ungerichtete Metabolom Analysen mit hochauflösenden Massenspektrometern, wie dem ICR-FT/MS und dem UPLC-qToF-MS, verwendet, um die metabolische Antwort der Graupappel (i) zu hoher ultravioletter Strahlung (UV-B, 290-315 nm), (ii) dem Fehlen der Isopren Synthase, (iii) dem Einfluß von Blattherbivoren mit und ohne Mykorrhizierung der Wurzel zu untersuchen. Zur Identifikation von spezifischen und globalen Stoffwechseländerungen wird die Multivariate Statistik zusammen mit der Massendifferenz Analyse angewandt.

Diese Dissertation zeigt wie man die ungerichtete Metabolomik mit hochauflösenden Massenspektrometern effizient für die Untersuchung des Pflanzenstoffwechsels verwenden kann, um unser Wissen über die metabolische Stressantwort zu erweitern.

List of Abbreviations

AMF	Arbuscular mycorrhizal fungus
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
CoA	Coenzyme A
cLSM	Confocal laser scanning microscopy
DI	Direct injection
DMAPP	Dimethylallylpyrophosphate
DNA	Deoxyribonucleic acid
EMF	Ectomycorrhizal fungi
ESI	Electrospray Ionization
EV	Empty vector
FA	Fatty acid
FPP	Farnesyl pyrophosphate
gdh2	Glutamate dehydrogenase 2
GLV	Green leaf volatile
GPP	Geranyl pyrophosphate
GUS	β -glucuronidase /green fluorescent vector controls
HMBPP	(E)-4-Hydroxy-3-methyl-but-2-enyl pyrophosphate
HTG	Hemiterpene glucoside
ICR-FT/MS	Fourier Transform Ion Cyclotron Resonance Mass Spectrometry
IE	Isoprene emitting poplar
IPP	Isopentenyl pyrophosphate
ISPS	Isoprene synthase
LC	Liquid chromatography
MB	Mycorrhized with beetle
MC	Mycorrhized control plants
MDA	Malondialdehyde

MDEA	Mass-Difference enrichment analysis
MDiA	Mass-Difference analysis
MDiN	Mass-Difference network
MEP	Non-mevalonate pathway
MS	Mass spectrometry/Mass spectrometer
MS/MS	Tandem mass spectrometry
MT	Monoterpene
m/z	Mass to charge ratio
NB	Non-mycorrhized with beetle
NC	Non-mycorrhized control plants
NE	Non-isoprene emitting poplar
nia	Nitrite reductase
nrt1	Nitrate transporter
OPLS-DA	Orthogonal Partial Least Squares Discriminant Analysis
PCA	Principal Component Analysis
PEP	Phospho <i>eno</i> pyruvate
Phe	Phenylalanine
PLS	Partial Least Square Regression Analysis
REMD	Reaction Equivalent Mass-Difference
RNA	Ribonucleic acid
RNAi	RNA interference
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
RT	Retention time
SIM	Multiple adjacent selected ion monitoring
SQT	Sesquiterpene
TCA	Tricarboxylic acid cycle
TIC	Total ion current

ToF-MS	Time of Flight Mass spectrometer
Tyr	Tyrosine
UPLC	Ultra-performance liquid chromatography
UV-B	Ultraviolet-B radiation
VIP	Variable influence of projection score
VOC	Volatile organic compounds
WT	Wild type

Table of content

Summary	i
Zusammenfassung	ii
List of Abbreviations.....	iii
Table of content	vi
1 Introduction	1
1.1 Motivation and Overview.....	1
1.2 Metabolomics.....	4
1.2.1 The importance of metabolism in life.....	4
1.2.2 Metabolomics.....	5
1.3 Plant metabolomics: a broad scientific field.....	5
1.4 The importance of poplar as a model system.....	7
1.4.1 Poplars secondary metabolism	7
1.4.2 Non-isoprene emitting poplars	9
1.5 Metabolomics as a tool to study environmental effects on poplar	10
1.5.1 Background on UV irradiation	10
1.5.2 The influence of leaf herbivores on metabolism	11
1.5.3 The symbiosis of mycorrhizal fungi and plant roots	12
1.6 Analytical techniques	13
1.6.1 Mass Spectrometry	14
1.6.2 Electrospray ionization.....	15
1.6.3 Fourier transform ion cyclotron resonance mass spectrometry	15
1.6.4 Time of flight mass spectrometers	17
1.6.5 Liquid Chromatography.....	18
1.7 Data Analytical Techniques used for metabolomics	19
1.7.1 Principal component analysis.....	19
1.7.2 Partial least squares	20
1.7.3 Mass-difference network analysis	21
2 Non-targeted metabolomic comparison of isoprene emitting and non-emitting poplars grown under high UV irradiation	22
2.1 Background.....	22
2.2 Experimental setup	22
2.2.1 Physiological measurements	23
2.3 Results.....	24

2.3.1	Normalization and statistical analysis of the ICR-FT/MS data	24
2.3.2	Non-targeted metabolomics reveals extensive metabolic rearrangements related to UV-B treatment and genetic background.....	26
2.3.3	Investigation of line dependent metabolic changes under UV exposure.....	28
2.3.4	UV-B exposure causes line dependent responses in the biosynthesis of phenolic compounds	32
2.3.5	Linking isoprene emission capacity and the diminished production of phenolic compounds	35
3	Mass-difference network analysis of non-isoprene and isoprene emitting poplars.....	37
3.1	Background.....	37
3.2	Development of an ideal workflow for mass-difference analysis on plant extracts.	38
3.2.1	Creation of a plant based reaction equivalent mass-difference list	38
3.2.2	Sum formula annotation and statistical analysis	41
3.2.3	Reaction equivalent mass-differences based statistics	43
3.3	Results and Discussion.....	45
3.3.1	Interpretation of KEGG annotations	45
3.3.2	Prenylation-REMDs directly link metabolism to genetic modification.....	46
3.3.3	Manually curated dicarboxylic acid REMDs explain new biochemical relationships.....	49
3.3.4	Under-representation of nitrogen containing REMDs in non-emitting poplars	50
3.3.5	Oxidative stress related REMDs highlight the non-emitting poplars	51
3.3.6	Over-representation of phosphoenolpyruvate and 2-oxo acid REMDs in non-emitting poplars.....	52
3.3.7	Over-representation of phenolic REMDs in isoprene-emitting poplars	53
3.3.8	Mass difference analysis as a data-driven approach for the location of metabolic pathways	53
3.3.9	The mass-difference network incidence matrix describes the origin of the unknown compositional space	54
3.3.10	Network modules contain compound class information.....	57
3.3.11	Tandem mass spectrometric measurements validate MDEA findings in the non-isoprene emitting poplars	58
3.3.1	MS/MS experiments validate and extend the conjugation of phenolics in the IE genotype.....	60
4	Investigation of the metabolic response of non-isoprene and isoprene emitting poplars to the poplar leaf beetle <i>C. populi</i>	62
4.1	Background.....	62
4.2	Experimental Setup.....	63
4.3	Results and Discussion.....	64

4.3.1	High leaf development and herbivore dependent metabolic variability	64
4.3.2	Identification and extraction of discriminant mass-features	65
4.3.3	Annotation and comparison of discriminant mass-features	65
4.3.4	The metabolic response of poplars to <i>C. populi</i>	67
5	Impact of root colonization on the leaf metabolome of poplars and the behavior of <i>C. populi</i>	70
5.1	Background.....	70
5.2	Experimental setup	70
5.3	Results and Discussion.....	71
5.3.1	Summary of physiological results.....	71
5.3.2	Statistical analysis of GC-MS data	71
5.3.3	The development of a workflow for REMD-transcript mapping	73
5.3.4	Root mycorrhization and beetle feeding causes metabolic alterations in poplar leaves	75
5.3.5	Inoculation and herbivory impacts the CHO and CHNO compositional space of leaves	77
5.3.6	Mycorrhization by EMF impacts the phenolic metabolism of poplar leaves	79
5.3.7	Mapping of transcripts on metabolite data via REMDs	80
5.3.8	Over-represented REMDs show an induction of aldoxim condensations and validate prior knowledge about herbivore induced pathways.....	83
5.3.9	Combination of (-/+)-LC-MS, GC-MS and transcript data	86
5.3.10	Proposed mechanism of induced defensive system caused by EMF inoculation	88
6	Conclusions and outlook	92
Appendix	94
Methods	94
Instruments	98
Chemicals	99
Software	101
List of figures	107
List of tables	110
List of equations	111
Literature	112
Acknowledgements	131
List of Scientific Communication	132
Curriculum Vitae	133
Eidesstattliche Erklärung	134

1 Introduction

1.1 Motivation and Overview

Fast resolution and sensitivity improvements in mass spectrometry enable the detection and quantification of thousands of mass features with one single analysis. For that reason the newest member of the 'omics' sciences, namely metabolomics, receives a lot of attention from scientists. Metabolomics aims at describing the whole set of small molecules, called metabolites, within an organism under defined environmental conditions and resolve it in space and time (Fiehn, 2002; Goodacre et al., 2004; Goodacre, 2005). Metabolites are the organisms' final chemotype/phenotype resulting from changes in gene/transcript and protein expressions (Sumner et al., 2003; Saito and Matsuda, 2010). The analysis of metabolomes can either be done in a targeted way, which aims at quantifying a set of metabolites with similar physiochemical properties (Fiehn, 2002), or in a non-targeted approach with the aim to detect and semi-quantify as many metabolites as possible in order to obtain a global view upon changes in the metabolic composition of an organism (Goodacre et al., 2004).

Plants are part of an ever-changing environment, which means that they have to quickly adjust to environmental stresses in order to maintain their fitness. For that reason phytometabolomes (plant metabolomes) belong to the most complex metabolomes and fulfill a variety of functions.

Poplars belong to those trees, which are best characterized on the genomic, transcriptomic and proteomic level (Tuskan et al., 2006, Jansson and Douglas, 2007). Additionally they are of high economic importance since they are used for biofuel, wood and fibre production (Sannigrahi et al., 2010). The disadvantage of poplar plantations, however, is the huge emission of the volatile organic compound (VOC) isoprene (2-methyl-1,3-butadiene) (Behnke et al., 2012), which acts as an indirect greenhouse gas by extending the half-life of methane in the atmosphere (Lelieveld et al., 1998; Poisson et al., 2000) and takes part in the formation of secondary organic aerosols (Claeys et al., 2004; Carslaw et al., 2010) as well as ozone (Chameides et al., 1988; Kiendler-Scharr et al., 2012) in anthropogenically polluted atmosphere. Non-isoprene emitting (NE) *Populus x canescens* (Behnke et al., 2007) could be a perfect alternative for isoprene emitting wild type (IE) poplars since they would decrease their negative impact on atmospheric chemistry (Behnke et al., 2012). Unfortunately isoprene is an important molecule in the stress response to sub-optimal environmental conditions in poplar (Behnke et al., 2007; Behnke et al., 2009; Behnke et al., 2010b; Way et al., 2013). Whether NE poplar could replace wild-type poplars under stress conditions is still not entirely clear and needs to be studied in more detail.

This thesis aims to compare the influence of biotic and abiotic stress factors on the metabolic composition of IE *Populus x canescens* (Aiton. Smith), a natural hybrid of *P. tremula* x *P. alba*, and its non-isoprene emitting mutants (Behnke et al., 2007) generated by RNA interference suppressing isoprene synthase (ISPS). The idea was to combine the expertise on poplar physiology, metabolism and adaptation from the research group Environmental Simulation (EUS), directed by Professor Dr. Jörg-Peter Schnitzler, with the expertise on high resolution analytical chemistry, non-targeted metabolomics and data handling from the research group Analytical BioGeoChemistry (BGC), directed by Professor Dr. Philippe Schmitt-Kopplin. Large-scale poplar experiments were either conducted in the sun simulators at EUS or under free-air conditions in cooperation with colleagues (Dr. Anna Müller and Professor Dr. Andrea Polle) from the department “Forstbotanik und Baumphysiologie” at the Georg-August University in Göttingen. Non-targeted metabolomic analyses of plant samples, which were collected during these experiments, were performed on the comprehensive metabolomics platform at BGC. During the thesis three different experiments were analyzed, which are subdivided into four main chapters (Fig. 1.1).

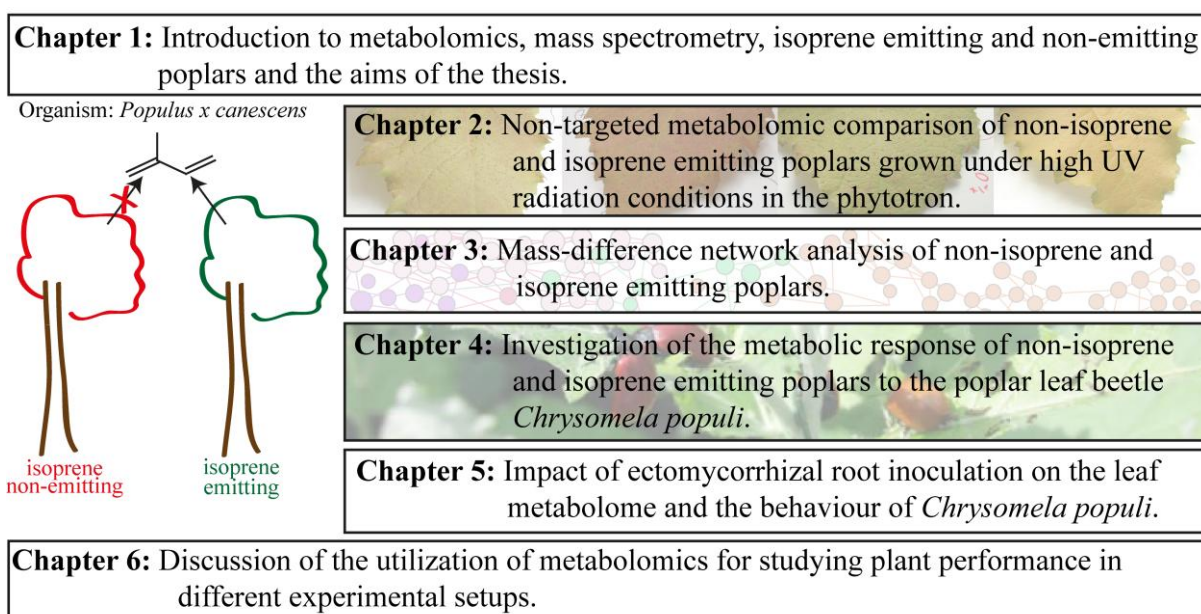


Figure 1.1 Chapter overview of the thesis.

The first experiment (Chapter 2, Fig. 1.1) was performed in the sun simulators (phytotron at EUS) and addressed the question whether isoprene emitting (IE) and non-emitting (NE) poplars differ in their physiological and metabolic adjustments to first time UV-B (290-315 nm) exposure. Therefore DI-ICR-FT/MS non-targeted metabolomic experiments were combined with gas exchange analysis and confocal microscopy. Analyzing the metabolomic data revealed prominent differences in the metabolic compositions of NE and IE poplars.

Since the NE and IE poplar lines became well characterized in previous works (Behnke et al., 2007; Behnke et al., 2010a; Behnke et al., 2010b; Way et al., 2013; Ghirardo et al., 2014), a proof-of-principle analysis of a novel data analytical strategy called Mass-Difference enrichment analysis (MDEA), invented by Dr. Franco Moritz (Moritz et al., 2015), was performed with the newly obtained data sets. MDEA aims at overcoming common limitations in metabolomic experiments such as database coverage and compound identifications. The basic idea is to extend mass spectra with a mass-difference domain pertaining to the remaining –omics disciplines, which can be used for the pheno/genotype linkage. Chapter 3 provides in depth explanations and results.

The second experiment was conducted at a free-air side in Göttingen, enabling the exposure of transgenic plants under almost natural environmental conditions, and dealt with the question whether specialized poplar leaf beetles (*Chroymela populi* L.) prefer feeding on non-isoprene or isoprene emitting poplars. It revealed that *C. populi* had a slight preference towards IE poplars. Since NE and IE poplar drastically differ in their metabolic profiles (see Chapter 2 & 3) a non-targeted metabolomic analyses with the ICR-FT/MS was performed to unravel whether the metabolic composition of the poplar genotypes affect leaf palatability. The results of this experiment are shown in Chapter 4.

Chapter 5 deals with the third experiment, in which the tripartite interaction between the ectomycorrhizal fungus (EMF) *Laccaria bicolor* (Maire) P.D. Orton), IE poplars and *C. populi* was investigated, was also run in Göttingen (similar experimental setup as in Chapter 4). The significant changes on the behavior of *C. populi* on inoculated plants lead to a systems biology investigation in which comprehensive cross-platform (UPLC-qToF-MS & ICR-FT/MS) non-targeted metabolomic data is merged with transcriptomics data (measured and analyzed in Göttingen by Dr. Anna Müller and Professor Dr. Andrea Polle) via the additional mass-difference domain compiled in Chapter 3. An overview of the thesis structure is provided in Fig. 1.1.

1.2 Metabolomics

1.2.1 The importance of metabolism in life

Organisms are highly ordered structures in which information needs to be transported and processed fast in order to quickly respond to stimuli to maintain their fitness. The cellular information flows from genes (DNA) via transcripts (RNA) to proteins. Proteins then can catalyze chemical reactions, which generate of low molecular weight molecules (<1500 Da), called metabolites (Fig. 1.2). This flow of information used to be the central dogma in molecular biology (Crick, 1956; Crick, 1970). These days, it is known that this information flow is tightly regulated in order to orchestrate its' complex interacting set of reactions which has the crucial role to supply energy under all environmental conditions.

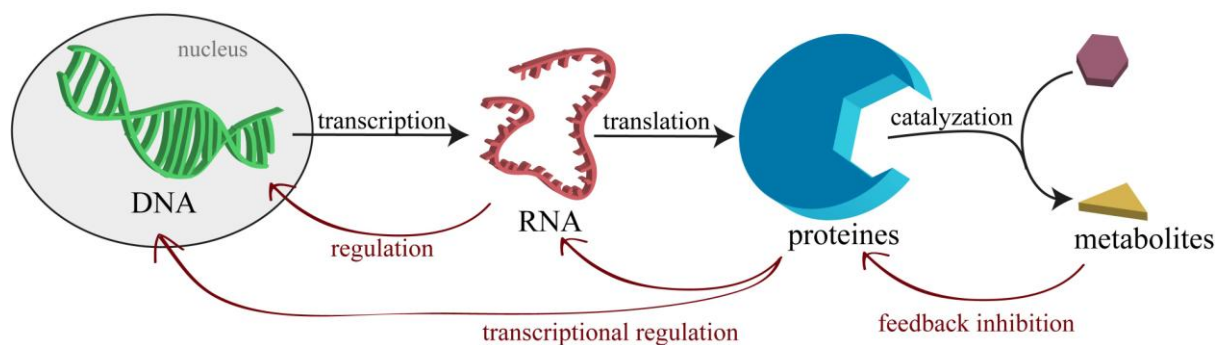


Figure 1.2. Cellular flow of information.

To achieve this, the balance between the energy demanding construction of molecules and their precursors (anabolism) and the energy releasing break-down of molecules (catabolism) must be adapted (Wilson, 2013). The efficiency of metabolic adaption is a central part in the organisms' fitness and consequently drives evolution on a molecular level.

In the last decade scientists, which before were more focused on analyzing genes and proteins, became more aware of the importance of overall metabolomic pattern because many questions remained unaddressed. The reason for the neglect of comprehensively analyzing metabolomic pattern simply was the technological limitation. A single organism produces thousands of different metabolites, e.g. a single *Arabidopsis* plant produces $\approx 5,000$ metabolites (Saito and Matsuda, 2010). Additionally, the high variability in cellular concentrations of metabolites is in the 10^6 range (Saito and Matsuda, 2010). These facts address major challenges to the analytical techniques in order to measure a representative metabolic snapshot. So how can we measure metabolites?

1.2.2 Metabolomics

The analysis of the complete set of metabolites under defined environmental conditions with the aim to provide a complete picture of an organism's physiological status at a certain point in time is called metabolomics (Fiehn, 2002). Compared to the other "omics" disciplines, such as genomics, transcriptomics and proteomics, metabolomics is still in its infancy. However, the rapidly growing number of publications illustrates the immense scientific potential (Fig. 1.3). A major reason for this is the broad interdisciplinary applicability of metabolomics. Thus far it has been successfully used in medical, environmental and biotechnology research fields to get a better understanding of gene function (Raamsdonk et al., 2001; Patil et al., 2005), diseases (Kaddurah-Daouk et al., 2008), drugs (Kim et al., 2010; Lan and Jia, 2010), host-pathogen interactions (Müller et al., 2013b; Witting et al., 2015), bioengineering and abiotic (Kaplan et al., 2004; Bundy et al., 2009; Kusano et al., 2011; Way et al., 2013) and biotic stresses (Marti et al., 2013).

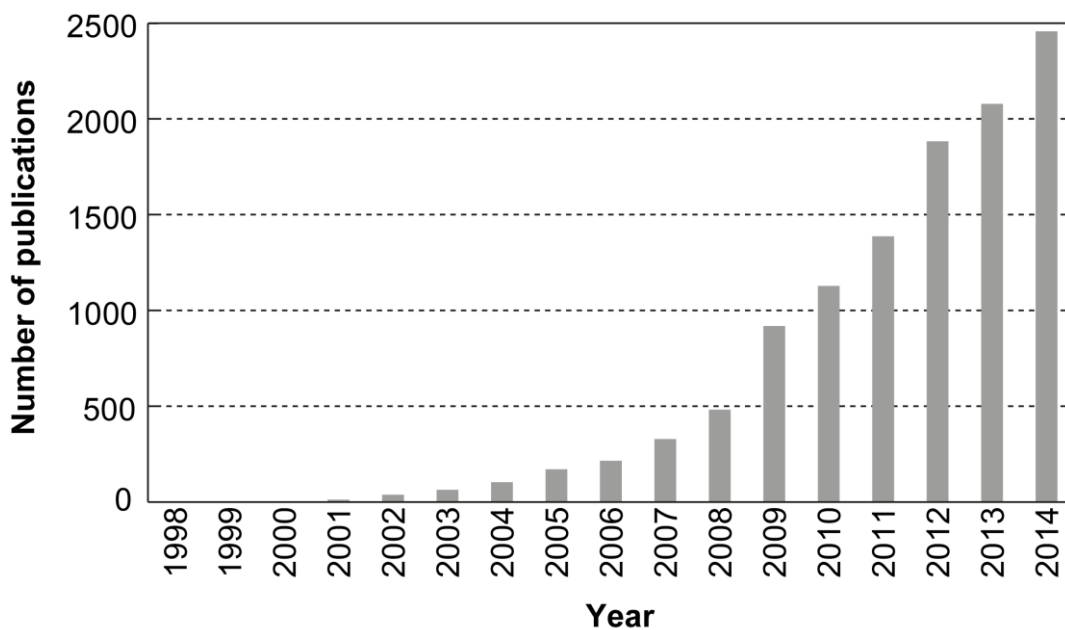


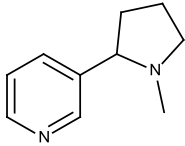
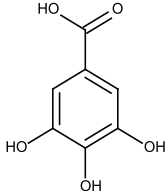
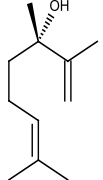
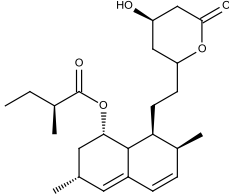
Figure 1.3. Pubmed search results for the term meabolomics (<http://www.ncbi.nlm.nih.gov/pubmed>).

1.3 Plant metabolomics: a broad scientific field

What makes plants so important? Plants are able to convert sunlight into chemical energy (photosynthesis), which then is stored as carbohydrates. This important ability puts plants at the basis of ecosystems by serving as the primary producers of organic matter. Other organisms consume plants in order to produce energy and survive.

As plants generally are sessile organisms, they have to cope with many biotic and abiotic stresses such as extreme temperature, nutrient limitation, drought, UV-B (290-315 nm) irradiation and exposure to pathogens or herbivores (Mittler, 2006; Dorantes-Acosta et al., 2012). This extreme evolutionary pressure resulted in the ability to synthesize a large variety of secondary metabolites. Those metabolites are not obligatory for the survival of plants/organisms as they are not directly linked to their growth, development or reproduction. Yet, they drastically improve their survivability and fitness by functioning as structure stabilizing, defense and signal compounds (Wink, 2003). More than 200,000 different secondary metabolites are present in the plant kingdom (Hartmann, 2007). Secondary metabolites are classified according to their biosynthetic origin (Table 1.1).

Table 1.1: Classes of secondary metabolites.

Compound class	Biosynthetic origin	Example structure	Estimated number
Alkaloids	Amino acids	 Nicotine	>12 000 (Croteau et al., 2000)
Phenolics	Shikimic acid	 Gallic acid	>8000 (Croteau et al., 2000)
Terpenes	IPP and DMAPP	 Linalool	>25 000 (Croteau et al., 2000)
Polyketides and fatty acids	Acetate	 Lovastatin	No data

As secondary metabolites improve the fitness of plants, they can also improve the fitness/life quality of humans. For centuries humans have used plants as medicine and nowadays 25% of drugs are derived from plants (Rates, 2001; Amin et al., 2009). In other

words, plants represent a huge source for the discovery of novel bioactive compounds. High throughput screening of plant extracts in order to find and identify unknown metabolites with useful properties can be achieved with metabolomics. However, there are many more applications and scientific questions for metabolomic analysis of plants such as phenotyping, gene characterization, plant quality improvement, chemical ecology, signaling networks, systemic acquired resistance and plant adaption to biotic and abiotic stresses.

1.4 The importance of poplar as a model system

Poplar is an important model system in plant biology since (i) it has a relatively close phylogenetic relationship to *Arabidopsis thaliana*, (ii) it allows studying processes which are absent in other annual model plants, such as wood-development and (iii) poplar exhibits a high amount of natural variation (Jansson and Douglas, 2007). Yet, the importance of poplar goes far beyond the model system approach. Their rapid growth rate and good rootstock regeneration make them the perfect tree species for bio-energy and wood production (Sannigrahi et al., 2010). Indeed poplar trees can be grown in short rotation coppices, which can be harvested in intervals of 3 to 5 years (Aylott et al., 2008) for biomass production. For that reasons scientists are eager to understand and improve wood formation, stress tolerance and herbivore resistance of poplars.

1.4.1 Poplars secondary metabolism

Poplar trees are known for their high biosynthetic production of shikimate/phenylpropanoid-derived phenolics such as phenolic glycosides, hydroxycinnamic acids and flavonoids (Boeckler et al., 2011). Although these three phenolic classes are structurally different, their cellular functions partially overlap. They all protect against high light intensities (Schaberg et al., 1998), protect cellular structures and are involved in signaling (Mandal et al., 2010). However, phenolic glycosides and hydroxycinnamic acids are primarily used to defend against herbivores and pathogens (Nicholson and Hammerschmidt, 1992; Boeckler et al., 2011), whereas flavonoids are involved in flower pigmentation (Ferreyra et al., 2012), symbiotic nitrogen fixation (Mandal et al., 2010; Abdel-Lateif et al., 2012) and seed development (Lepiniec et al., 2006; Doughty et al., 2014).

The biosynthesis of aromatic compounds channels through the plastidic shikimate pathway, to which >30% of photosynthetically fixed carbon is directed (Maeda and Dudareva, 2012). The end product of the shikimate pathway is chorismate, which either can be used for the biosynthesis of phenolic glycosides or it is transformed via phenylalanine (aromatic amino acids) to the phenylpropanoid cinnamate. Cinnamate represents an

important cross-road in the biosynthesis of phenolic compounds since it is used for the production of hydroxycinnamic acids, flavonoids (via phenylpropanoids and chalcones), coumarins and benzyl derivatives such as salicylate.

Another important class of secondary metabolites synthesized by poplars are terpenes (or terpenoids/isoprenoids), which are involved in a variety of essential physiological processes such as photosynthesis (Singsaas, 2000) and plant growth and development (Tholl, 2015). *In planta* terpenes can be produced by two spatially separated pathways, namely the cytosolic mevalonate pathway and the plastidic 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway (Cheng et al., 2007). Both yield the two isomeric C5 molecules isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). Head-tail conjugation of IPP and DMAPP results in the formation of the monoterpene precursor geranyl pyrophosphate (GPP, C10). Follow up condensations with IPP yields, farnesyl pyrophosphate (FPP, C15) and geranylgeranyl pyrophosphate (GGPP, C20), the precursors of sesqui- and diterpenes (Cheng et al., 2007). The utilized pathway thereby dictates the metabolic fate of DMAPP and IPP. The cytosolic mevalonate pathway is mainly used for the production of steroids, terpenoid quinones, sesqui- and triterpenes (Tholl and Lee, 2011), whereas the plastidic MEP pathway mainly produces isoprene (2-methyl-1,3-butadiene, C5), monoterpenes, diterpenes and higher terpenes such as carotenoids (Dubey et al., 2003; Tholl and Lee, 2011).

Mono- and sesquiterpenes represent an important group of volatile organic compounds (VOC) emitted by plants. They are key components in the intra- and inter-species communication by acting as repellents against herbivores or as attractants for natural herbivore enemies and possible pollinators (Holopainen, 2004).

The hemiterpene isoprene, however, is the VOC that is emitted the most by poplar and other tree species. The global amount of emitted isoprene is similar to that of methane (Guenther et al., 1995). Contributing to methane oxidation and by acting as a major precursor of ozone and secondary organic aerosols in the troposphere, isoprene significantly impacts atmospheric chemistry (Fuentes et al., 2000). Recently it was shown that isoprene represents the main carbon sink of MEP pathway intermediates in poplar leaves to which 99% of the MEP carbon flux is directed (Ghirardo et al., 2014). Although the amount of photosynthetically fixed carbon used for isoprene biosynthesis is high, the physiological function of isoprene is still under debate. Many works showed that isoprene improves thermotolerance (Singsaas and Sharkey, 1997; Behnke et al., 2007; Sasaki et al., 2007), protects against oxidative stress (Loreto et al., 2001; Velikova et al., 2008; Vickers et al., 2009), increases thylakoid membrane stability (Velikova et al., 2011) and attracts predators of plant enemies (Loivamäki et al., 2008). Other hypothesized that isoprene functions as a metabolic 'safety valve' (Rosenstiel et al., 2004).

1.4.2 Non-isoprene emitting poplars

The emission of isoprene represents a major disadvantage of poplar plantations (Behnke et al., 2012). To reduce the impact of poplar plantations on the atmosphere and to study the physiological function of isoprene in more detail, transgenic non-emitting grey poplars (NE), where the isoprene synthase (*PcISPS*) is knocked-down by RNA interference (RNAi), can be used (Behnke et al., 2007; Behnke et al., 2012). A number of studies have shown that the loss of isoprene emission in grey poplar results in physiological (Behnke et al., 2007; Behnke et al., 2010b), transcriptomic (Janz et al., 2010; Behnke et al., 2010b), proteomic (Velikova et al., 2014) and metabolic changes (Behnke et al., 2010a; Janz et al., 2010; Way et al., 2013). Interestingly, loss of isoprene emission results in a reduced production of phenolic compounds (Behnke et al., 2010a). The MEP and shikimate pathways are linked through phosphoenolpyruvate, which is consumed by both pathways. Moreover, blocking ISPS results in accumulation of DMAPP in the chloroplast (Behnke et al., 2007; Behnke et al., 2010a; Ghirardo et al., 2014). High DMAPP levels inhibit the MEP pathway entry enzyme desoxy xylulose 5-phosphate synthase (DXS) consequently reducing the carbon flux (Ghirardo et al., 2014).

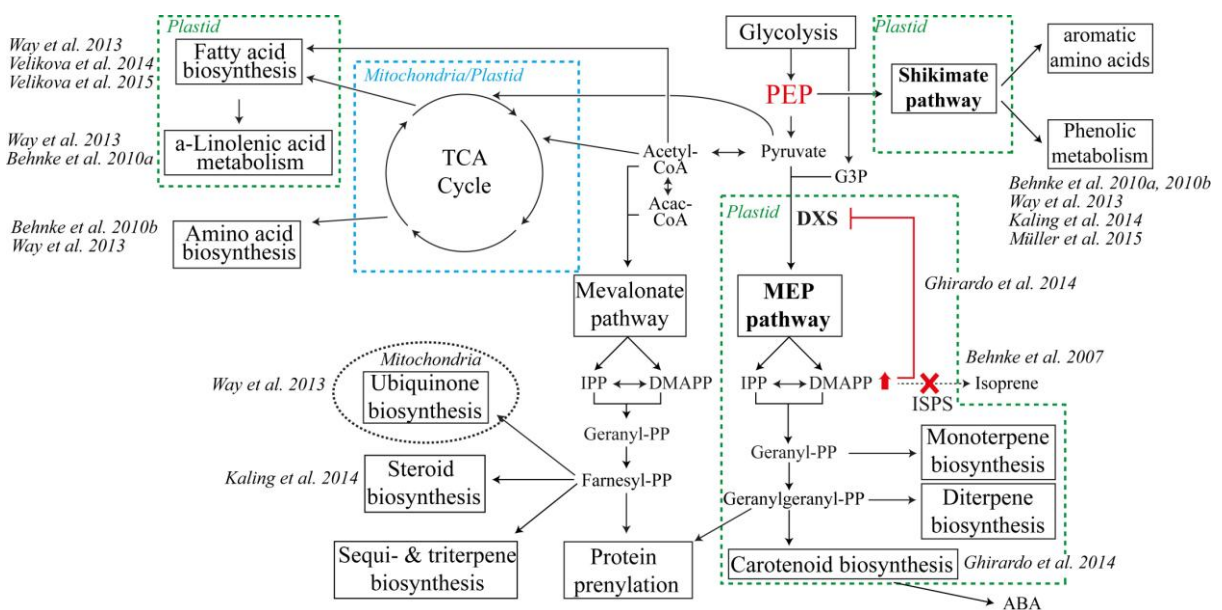


Figure 1.4. Overview of changes in the metabolism of *P. x canescens* resulting from *PcISPS* knockdown by RNAi.

Another metabolic pathway affected by the ISPS knockdown is the fatty acid metabolism (Way et al., 2013). Here, leaves of NE poplars possess higher amounts of unsaturated and saturated fatty acids as well as oxylipins. Quite contrary to that a recent work investigated the lipid composition of thylakoid membranes revealing a lower amount of lipids and fatty

acids within NE chloroplastic membranes (Velikova et al., 2015). Consequently, the loss of isoprene emission alters the cellular distribution of lipids, leading to lower amounts in thylakoid membranes but an increase in total cellular lipid concentration.

Proteome analysis of NE chloroplast also unraveled different protein levels between NE and IE poplars (Velikova et al., 2014), whereby NE chloroplasts had higher amounts of proteins involved in photosynthesis, including proteins involved in lipid metabolism, and lower amounts of ribosomal and histone modulating proteins (Velikova et al., 2014). An overview of metabolic changes resulting from ISPS knockdown is given in Figure 1.4. Physiological differences resulting from ISPS knockdown include lower net CO₂ assimilation rates, higher growing rates and an increased susceptibility for fungal infections as well as for herbivores (Behnke et al., 2012).

1.5 Metabolomics as a tool to study environmental effects on poplar

As already mentioned in the previous section, plants have to cope with a variety of environmental stimuli, which can either be harmful or beneficial. Generally harmful conditions are called stresses and can be classified into abiotic and biotic stresses. Abiotic stresses are defined as non-living stress factors such as nutrient and space limitations, drought-, light-, heat- and cold stress and many more (Nakashima et al., 2009). In contrast to that, biotic stresses are defined as the negative impact of other organisms such as bacteria, fungi or herbivores (Fujita et al., 2006).

But plants can also experience beneficial stimuli such as growth promoting bacteria and mycorrhizal fungi, which enhance the growth and nutrient supply of plants. Explaining all environmental factors would go beyond the aim of this thesis. Because of that the environmental conditions investigated in this thesis are shortly introduced in the following paragraphs.

1.5.1 Background on UV irradiation

A part of natural sunlight is ultraviolet A (UV-A, 315-400 nm) and ultraviolet B (UV-B, 290-315 nm) radiation. Since atmospheric ozone blocks most of UV-B, the amount of radiation that reaches plant species depends on latitude, season, time of day, cloud cover and altitude (Seckmeyer et al., 2008).

Natural levels of UV-B radiation act as an environmental regulator for plants (Jenkins, 2009; Hideg et al., 2013), but high UV-B doses impairs cellular and macromolecular

structures in the vegetative plant tissue. To decide whether a certain UV-B dosage is “natural” or “harmful”, plants have to sense and quantify UV-B radiation. In fact the UV-B photoreceptor UV RESISTANCE LOCUS 8 (UVR8) not only enables plants to sense UV-B radiation, but also triggers the metabolic plant defense mechanisms such as antioxidant levels, photo-repair or accumulation of UV-B screening pigments if the UV-B dose is too high (Rizzini et al., 2011; Heijde and Ulm, 2012).

Well-known metabolic pathways responding to UV-B are the phenolic pathways (Li et al., 1993; Lavola, 1998; Warren et al., 2003; Kaspar et al., 2010; Kusano et al., 2011; Tossi et al., 2011), the ascorbate (Costa et al., 2002; Lidon et al., 2012) and the tocopherol metabolism (Carletti et al., 2003). Compounds synthesized by these pathways either function as UV-B absorber in the epidermal cell layers in order to protect the photosynthetic apparatus (Burchard et al., 2000), or they quench reactive oxygen and reactive nitrogen species (ROS/RNS) that are produced during the stress response (Lidon et al., 2012; Hideg et al., 2013). In this context isoprene plays a role as a central molecule in ROS/RNS signaling/quenching (Vanzo, 2014). It already has been shown that increased UV-B radiation influences isoprene emission (Harley et al., 1996; Tiiva et al., 2007). Based on this, it is obvious to study the effects of UV-B irradiation on IE and NE poplars in detail. Especially, the absence of isoprene emission and the reduced production of phenolic compounds in NE poplar could pave the way to unravel the function of isoprene under UV stress.

1.5.2 The influence of leaf herbivores on metabolism

Long living plant species such as poplars have to cope year after year with an ever-changing community of antagonistic species (Whitham et al., 1996; Jansson and Douglas, 2007; Philippe and Bohlmann, 2007). To better adapt to these community changes, different direct and indirect defensive systems have evolved. Poplars (and other *Salicacea* species) are best known for their high production of phenolic compounds such as phenolic glycosides, hydroxycinnamic acids and condensed tannins (Boeckler et al., 2011; see chapter 1.6.1.). These compounds possess a high bioactivity and hinder the digestion of consumed plant proteins (Philippe and Bohlmann, 2007; Boeckler et al., 2011).

This defensive system is very efficient against generalist insect herbivores. Specialized herbivores, however, are less sensitive to phenolic glycosides. The larvae of *Chrysomela populi* for example are even capable to sequester phenolics for their own safety (Deroe and Pasteels, 1982).

C. populi, the poplar leaf beetle, is a specialized herbivore which causes severe damage not only on natural growing poplars, but also on poplar plantations (Gruppe et al., 1999; Brilli et al., 2009). However, plant phenolics are just one aspect of the comprehensive response

upon herbivore attack (Philippe and Bohlmann, 2007). Damaging the leaf surface triggers the emission of many, structurally different, VOCs (Pare and Tumlinson, 1999). Within the first few minutes small membrane breakdown products, also called green leaf volatiles, are released (Hassan et al., 2015). These compounds function as important signaling molecules and can be sensed by other plants priming their defense. Additionally the emission of monoterpenes, sesquiterpenes and phenolics is induced and can last for days after the initial herbivore attack (Brilli et al., 2009). This emission profile results in a specific blend that can either function as a direct insect repellent or as an attractant for insect predators.

Whether isoprene plays a role in the poplar-insect communication is still unclear. In transgenic *Arabidopsis* (Loivamäki et al., 2008) and tobacco plants overexpressing ISPS it has been shown that isoprene affects insect behavior (Laothawornkitkul et al., 2008). NE poplars would be the perfect system to study the effect of isoprene emission on the poplar-insect communication network.

1.5.3 The symbiosis of mycorrhizal fungi and plant roots

In recent years the attention of plant scientists is rapidly shifting from above- to belowground, simply because many interactions taking place on the root level are poorly understood. One example is the ancient interaction between plant roots and mycorrhizal fungi, which is found on more than 90% of all plant species and in almost any environment (Behie and Bidochka, 2014). The term mycorrhizal fungus refers to a heterogeneous group of many different fungal taxa, which colonize plant roots. The fungus, thereby, transfers nutrients, such as ammonia and phosphate, to the plant in exchange for plant derived carbohydrates (Behie and Bidochka, 2014). Estimates indicated that approximately 20% of plant photo assimilates are transferred to the fungus (Jakobsen and Rosendahl, 1990). However, the plant benefits with an increased tolerance to biotic and abiotic stresses (Feng et al., 2002)

Based on the anatomical traits of the fungus, mycorrhizas are sub-divided into two main groups, specified as (i) endomycorrhiza fungi, which penetrate the cell walls of epidermal and cortex root cells, and (ii) ectomycorrhiza fungi (EMF), who typically colonize the root tips and form an intercellular Hartig net surrounding the epidermal cells of plant roots (Kottke and Oberwinkler, 1987).

A typical EMF species, which colonizes natural growing poplar as well as poplar grown under lab conditions is *Laccaria bicolor* (Müller et al., 2013a). Just recently a study demonstrated that the inoculation with *L. bicolor* alters the root auxin metabolism of poplars (Vayssières et al., 2015). Another group did even report changes in leaves secondary metabolism when poplars were inoculated with the EMF *Hebeloma mesophaeum* (Pfabel et

al., 2012). Changes in the leaf metabolic composition after inoculation with mycorrhizal fungi have been shown to be species-specific (Schweiger et al., 2014). Whether these alterations in the phytometabolom, however, cause changes in the behavior of above ground herbivores still needs to be elucidated.

1.6 Analytical techniques

Metabolomic experiments can be separated into two approaches. First, the targeted approach that aims to quantify and identify selected sets of metabolites/compound classes which possess similar physical and chemical properties. It therefore represents the hypotheses driven or confirming approach. For example if a protein is knocked-down, metabolites whose positions are in the affected pathway would be chosen for metabolic analysis. The major disadvantage, however, is that possible metabolic ‘side-effects’ of the genetic modification are not included. To avoid this, one needs to use the second approach, which is called the non-targeted or “curiosity driven approach”. Its purpose is to obtain a global view upon changes in metabolism by measuring and quantifying as many metabolites as possible (Naz et al., 2014). Non-targeted measurements create hypotheses, which in turn can be validated again by targeted methods. The typical workflow of metabolomic experiments is exemplified in Figure 1.5. Generally, it can be divided into three steps being (i) sample preparation, (ii) measurement and (iii) data evaluation.

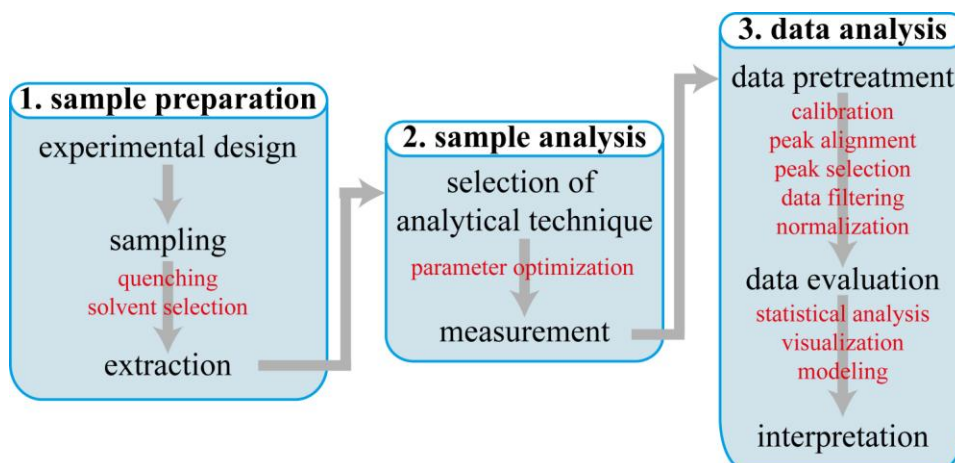


Figure 1.5. The general workflow of metabolomic experiments can be separated into three steps.

The most important aspect that needs to be taken care of during sample preparation is the immense speed of metabolic adaption to external stimuli. To counteract this, metabolism needs to be quenched rapidly when samples are taken, which is usually achieved with liquid

nitrogen (Faijes et al., 2007). Moreover, since metabolite extraction is depending on the extraction solvent, the aim of the study needs to be clarified prior to the extraction procedure. For example an interest in cellular lipid concentrations needs other solvent properties than a method that aims measuring amino acids. The methods used for sample and data analysis in this work will be introduced in the following sections.

1.6.1 Mass Spectrometry

One way of puzzling together the molecular ingredients of plant extracts, is to measure their masses. Having exact masses, molecular compositions can be calculated. Modern mass spectrometers (MS) are able to measure thousands of masses simultaneously, which makes them the perfect tool for metabolomic analysis. However, MS instruments are not able to measure exact masses of neutral, uncharged, molecules. For that reason, molecules first need to be ionized and transferred into the gas phase (Kearle and Tang, 1993). In consequence the MS instruments measure the mass and the charge of molecules, called mass-to-charge ratio (m/z). In principle every MS instrument consists of the following three parts; (i) an ion source, (ii) a mass analyzer and (iii) a detector. Whereby, the choice of ion source and mass analyzer depends on the analytic question that needs to be addressed. The strength and quality of mass spectrometers and their spectra can be determined with following parameters:

The first one is the mass resolution R , which is defined as the ability to distinguish peaks with similar m/z values in a mass spectrum (Gross, 2011; Murray et al., 2013). This can be formulated as follows (where m_{50} is the peak width at half of the peak height (FWHM)):

$$R = \frac{m}{\Delta m_{50}} \quad \text{Eq. 1}$$

The second parameter is the mass accuracy (Murray et al., 2013), which is defined to be the difference of measured and theoretical mass. In most cases the relative mass error (mass dependent) is calculated as follows:

$$\text{relative mass error [ppm]} = \frac{(m_{\text{measured}} - m_{\text{theoretical}})}{m_{\text{theoretical}}} \cdot 10^6 \quad \text{Eq. 2}$$

Another important instrumental parameter is the sensitivity of the mass spectrometer. It is defined as the minimum analyte concentration that is needed to detect a signal at a specified signal-to-noise ratio (S/N) (Forcisi et al., 2013). Consequently the sensitivity is compound

dependent. In the following sections the MS techniques, which were used in this thesis will briefly be introduced.

1.6.2 Electro spray ionization

Compound ionization with consecutive transfer into the gas phase represents the initial step of MS analysis (Kearle and Tang, 1993). In order to perform untargeted metabolome analysis, the ionization efficiency should cover a broad range of different metabolite classes. A well-fitting ionization technique, commonly used for metabolomic experiments, is electrospray ionization (ESI) (Forcisi et al., 2013). The reason for that is that on the one hand ESI represents a 'soft' ionization technique with almost no molecule fragmentation and on the other hand it fulfills the previously stated broadband ionization efficiency (Theodoridis et al., 2012; Forcisi et al., 2013).

The sample is sprayed through a thin metal needle (also called spray needle) at the tip of which a very high voltage is applied into the ESI source, usually 3-4.5 kV for small molecules (Kearle and Verkerk, 2009). At a distance of 1-3 cm a counter electrode is placed to which the much lower voltage of 0.5 kV is applied. The high voltage causes the sample to disperse into an aerosol of charged analyte-solvent droplets (Kearle and Verkerk, 2009). Through a second inlet a hot flow (200°C) of nitrogen enters the ion source, which causes solvent evaporation (dry gas). Consequently, the droplet size shrinks and the charge density of the droplet surface increases. When the repelling ion forces become too strong, the droplets explode. This process repeats itself until dissolved ions arise (Kearle and Verkerk, 2009).

Based on the voltage polarity, which is applied in the ESI source, positive or negative ions are produced. In the negative ionization mode, the ESI source functions as a strong base which ionizes acidic molecules, whereas in the positive ionization the ESI source represents a strong acid ionizing compounds with basic moieties such as amines. In addition to the spray conditions, the ionization efficiency also depends on molecular properties such as pKa value, hydrophobicity, surface activity, LogP and polarizability (Oss et al., 2010). However, ESI ionization is suitable for a large number of analytes and is therefore broadly used for mass spectrometric applications (Kostiainen and Kauppila, 2009; Oss et al., 2010).

1.6.3 Fourier transform ion cyclotron resonance mass spectrometry

Due to the high resolution and mass accuracy, Fourier transform ion cyclotron resonance mass spectrometry (ICR-FT/MS) has unique advantages as an analytical tool for metabolomic experiments (Brown et al., 2005). Its high mass accuracy enables the assignment of sum formulas to thousands of mass features. ICR-FT/MS spectrometers consist of the following parts: (i) the ionization unit, (ii) ion optics (ion funnels and focusing

lenses), (iii) the quadrupole, octapole and hexapole units, (iv) the ICR cell and, (v) the magnet. Ion detection in ICR-FT/MS takes place in the ICR cell and is based on the ion motion in high homogeneous magnetic fields (Marshall et al., 1998). The ICR cell consists of two detection, two excitation and two trapping plates. Once exposed to high magnetic fields, ions start to oscillate circularly within the ICR cell (Marshall et al., 1998). The reason for that is the Lorentz force (Eq. 3), which is defined as:

$$F = m \frac{dv}{dt} = zv \times B \quad \text{Eq. 3}$$

where m is the ion mass, z is the ion charge, v is the ion velocity and B represents the strength of the magnetic field. The Lorentz force (Eq. 3) can be converted to the cyclotron frequency w_c :

$$w_c = \frac{zB}{m} \quad \text{Eq. 4}$$

Equation 2 can then be rearranged to the mass to charge ratio:

$$\frac{m}{z} = \frac{B}{w_c} \quad \text{Eq. 5}$$

To enable ion detection, a radio frequency (RF) is applied through two excitation plates. If the RF is in resonance with a certain ion present in the ICR cell, the ion oscillation radius increases. Consequently, the ion movements induce a current, which corresponds to the cyclotron frequency, in the detector plates (Marshall et al., 1998; Barrow et al., 2005). The frequencies and their respective intensities are then translated into a mass spectrum (m/z over intensity) through Fourier transformation.

As the per-scan sensitivity of ICR-FT/MS measurements is low, high mass accuracy and resolution can only be achieved with long scanning times (Chernushevich et al., 2001; Forcisi et al., 2015). Thus, to use the full potential of ICR-FT/MS spectrometers, the samples must be measured through direct injection (DI).

A disadvantage of DI measurements, however, is that it is not possible to distinguish isomeric molecules. To overcome this problem, two different MS/MS experiments can be performed on ICR-FT/MS spectrometers: (i) the collision induced dissociation (CID) fragmentations in which all ions entering the hexapole are fragmented and (ii) sustained off-resonance irradiation (SORI) experiments in which isolated ions are fragmented within the ICR cell. However, due to their high price and high lab space requirements, ICR-FT/MS spectrometers are not widely distributed.

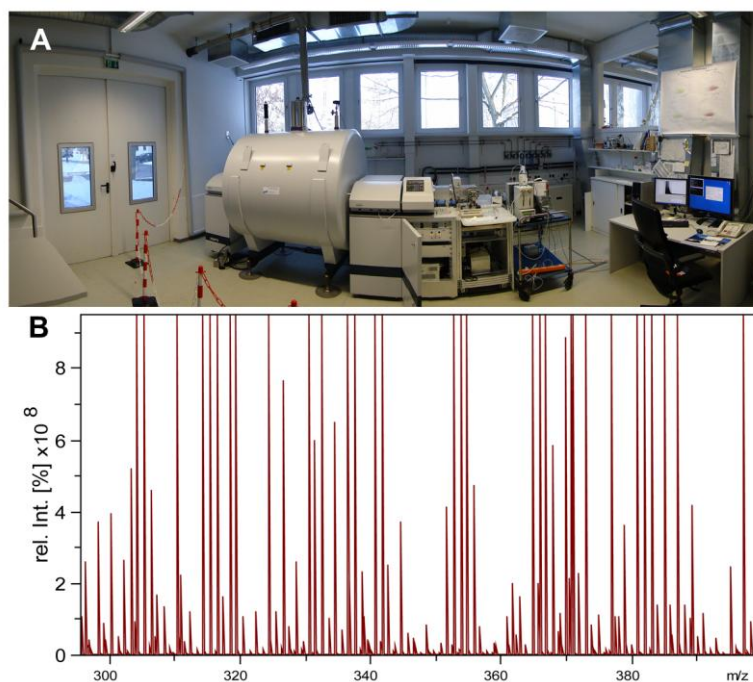


Figure 1.6. Representation of an ICR-FT/MS. (A) Picture of the 12 Tesla ICR-FT/MS spectrometer used in this thesis. (B) ICR-FT/MS spectrum of a grey poplar leaf extract acquired at 2 Megawords.

1.6.4 Time of flight mass spectrometers

The perfect mass spectrometric match for LC coupling, are time of flight (TOF) instruments. They provide high scanning rates, are able to measure very small m/z features and give sufficient resolution (Forcisi et al., 2013). In TOF instruments, as already indicated by its name, the mass to charge ratio is determined by exact flight time measurements. To achieve this, ions are accelerated simultaneously by an electric field in which ions with the same charge receive the same kinetic energy (Wiley and McLaren, 1955). Thereby, the ions enter a flight tube having a detector placed at its end, which records the exact time of ion impact (Cotter, 1992). Having the same kinetic energy, the flight time of ions is mass dependent in that light ions travel faster than heavy ones (Wiley and McLaren, 1955). The whole relationship can be expressed as follows:

$$TOF = L \sqrt{\frac{m}{2zeV}} \quad Eq. 6$$

where L is the length of the flight tube and eV is the electric field strength. Consequently, the resolution of TOF instruments depends on the length of the flight tube. For that reason, ToF development focuses on the increase of flight paths mainly through the improvement of ion reflection techniques within the flight tube. A picture of the ToF instrument, used in this thesis, is shown in Figure 1.7.

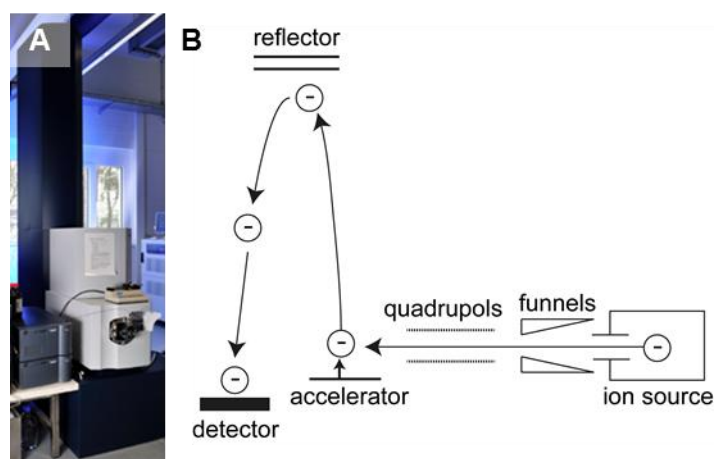


Figure 1.7. Representation of a ToF-MS. (A) Picture of the qToF-MS system used in this thesis. (B) Illustration of ion guidance through the qToF-MS.

1.6.5 Liquid Chromatography

Complex mixtures such as plant extracts contain thousands of molecules with unique physical and chemical properties. These characteristic properties can be used to separate such mixtures prior to their mass spectrometric detection. A well-suited separation technique, which is widely used for plant metabolomics, is liquid chromatography (LC) (De Vos et al., 2007).

The principle of LC separation is based on the distribution of molecules within a two-phased system consisting of one mobile (fluid) and one stationary phase (Dorsey and Dill, 1989). The sample is dissolved in the mobile phase and then flushed through/over the stationary phase. Thereby, the speed of molecules is dependent on their interactions with the surface of the stationary phase. The time a particular molecule needs to pass through the stationary phase is called retention time (RT) (Dorsey and Dill, 1989). In modern chromatography the stationary phase is packed in metal columns with different length, particle shape, inner diameters and particle sizes. These parameters, together with the physiochemical properties of the stationary phase, determine the separation efficiency, which is expressed as the number of theoretical plates normalized by the column length (Eq. 5, van Deemter equation, Van Deemter et al., 1956):

$$HEPT = A + \frac{B}{u} + (C_s + C_m) \cdot u \quad \text{Eq. 7}$$

where A is the Eddy diffusion, B represents the longitudinal diffusion, C_s is the mass transfer of the stationary phase and C_m of the mobile phase. A theoretical plate represents a fictional

space within the chromatographic column in which the two phases are in equilibrium with each other. Notably, a theoretical plate does not exist, it just serves to better understand the separation process (Giddings, 1969). Nevertheless, the higher the number of theoretical plates, the higher the chromatographic separation efficiency.

Thereby, the physical and chemical properties of the stationary as well as of the mobile phase are of major importance. The separation method most commonly used is reversed phase chromatography (RP), in which a mobile phase gradient of water to organic solvent is flushed through a non-polar stationary phase (Rainville et al., 2007). Standard stationary column phases consist of C18 or C8 capped silica particles. Typical mobile phase gradients, which are suitable for MS coupling, are e.g., water/acetonitrile or water/methanol gradients.

Coupling LC systems, such as high performance liquid chromatography (HPLC) or ultra-performance liquid chromatography (UPLC), to MS instruments has several advantages: (i) the reduction of matrix effects, (ii) separation of isomers and (iii) the additional retention time information enabling a better characterization of the molecule of interest.

1.7 Data Analytical Techniques used for metabolomics

Using modern high-resolution mass spectrometers for metabolomic experiments typically results in a large amount of data, which needs to be put in context to the experimental background. In other words: thousands of mass features must be explained with a handful of experimental variables. Thereby, the overall aim is to identify biomarkers and metabolic patterns that directly respond to the experimental question.

To achieve this, chemometric methods, which are defined as ‘the art of extracting chemically relevant information from data produced in chemical experiments...’ (Wold, 1995), are applied. The different mathematical and statistical procedures grouped in the term chemometrics can be subdivided into two classes: (i) unsupervised and (ii) supervised methods (Lucio, 2009). Unsupervised methods are used to capture the overall data pattern and reduce data dimensionality, whereas supervised methods are used to extract important variables to a predefined grouping. Preferably used for metabolomic data analysis are multivariate techniques such as principal component analysis (PCA) and partial least squares (PLS) regression analysis (Trygg et al., 2007), which are explained in the following sections.

1.7.1 Principal component analysis

Principal component analysis (PCA) is an unsupervised multivariate statistical method, which aims to explain as much variation as possible. The data-matrix containing sample

(called observations) and peak information (variables) is thereby reduced through orthogonal transformation into a few linearly uncorrelated latent variables called principle components (PC). Notably, the PCs represent the Eigen vectors of the covariance matrix and are ranked from large to small amounts of explained variance. It is important to know that each PC consists of two vectors, namely the score and the loading vector, whereby the score equals the latent variable, which is linked to the original variable by the loading vector. Now, it is possible to visualize the data matrix by plotting either two score vectors or two loading vectors against each other (Fig. 1.8). Combination of these two plots enables finding variables positively or negatively correlating to the observations (Boccard et al., 2010; Marti et al., 2013).

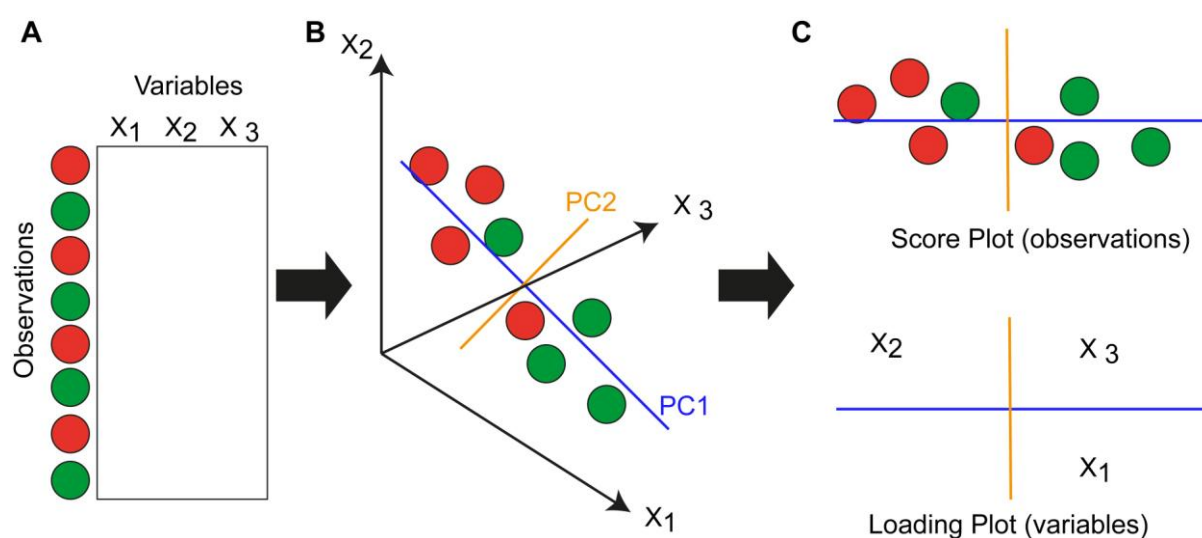


Figure 1.8. Schematic representation of Principal Component Analysis (PCA). Starting with (A) the data-matrix and (B) its projection into principle components (PC) as well as (C) the resulting score and loading plots. (modified according to Boccard et al., 2010)

In summary, PCA reduces data dimensionality, enables to easily check the data quality and visualizes the correlation between observations and their variables (Lucio, 2008).

1.7.2 Partial least squares

To extract important mass features and classify the experimental context based on metabolic pattern, supervised methods, like partial least squares discriminant analysis (PLS-DA), are needed. In principle, as the wording “supervised” already implies, PLS-DA looks for variables correlating to a user made sample grouping/classification by maximizing their covariance (Wiklund et al., 2008). The output is a so-called PLS model. Like other statistical and machine learning models, PLS-DA is in danger of explaining data noise instead of group-classification, a phenomenon called over fitting (Trygg and Wold, 2002). To avoid this

effect, different techniques such as cross-validation or permutation tests must be used. The model validity is expressed as the goodness of prediction (Q^2). Discriminant variables can now be selected by using the variable influence on projection (VIP) score, whereby a VIP score bigger than 1 is considered as being discriminant (Galindo-Prieto et al., 2014).

1.7.3 Mass-difference network analysis

A bottleneck in metabolomics is the poor database coverage. In most cases only up to 20% of mass features can be annotated by databases, emphasizing the huge gap of knowledge, which is present in metabolomic studies. So how can we use and explain the whole set of mass features? Although, high-resolution mass spectrometers have the advantage of measuring exact masses with very high accuracies, their full power is not fully utilized in metabolomic applications. By using specific rules, exact masses can be directly converted to the molecule's sum formula. Of course, fragmentation (MS/MS) measurements are necessary for further structural evaluations. However, an insufficiency in ion abundances often hinders MS/MS fragmentations. Yet, it is possible to use the MS/MS data evaluation procedure as a template to improve full scan MS data interpretability.

Similar to the calculation of neutral losses in MS/MS data evaluation, the subtraction of annotated sum formulas does also yield a kind of "neutral loss", called mass-difference (Breitling et al., 2006; Tziotis et al., 2011). Based on this feature, a whole MS spectrum can be converted and visualized in a network where each node represents a sum formula and each edge represents a mass difference (Tziotis et al., 2011). This concept has been introduced in 2006 by Breitling et al. to predict metabolic pathways from conventional full scan spectra. To achieve this, a predefined list of mass-differences, mimicking biochemical reactions, was used. In this initial work, however, the method was only applied in the known compositional space of KEGG, meaning that unknown, not annotated mass features were left out of the data evaluation procedure.

In 2011 Tziotis et al. adopted this idea and extended it to create a mass-difference based network algorithm, called NetCalc, for the entirely data-driven assignment of sum formulas to m/z features. Since then this algorithm was used in a variety of studies to visualize the complexity of natural organic matter (NOM) and metabolism of different species (Müller et al., 2013b; Walker et al., 2014; Zhang et al., 2014; Jeandet et al., 2015; Moritz et al., 2015; Witting et al., 2015). The advantage of such network based sum formula assignments is that on the one hand the full set of mass features is available for the data interpretation and on the other hand that the mass-difference dimension can be used.

2 Non-targeted metabolomic comparison of isoprene emitting and non-emitting poplars grown under high UV irradiation

Parts of this chapter were published in:

Kaling M., Kanawati B., Ghirardo A., Albert A., Winkler J.B., Heller W., Barta C., Loreto F., Schmitt-Kopplin P. and Schnitzler J.P. (2015) UV-B mediated metabolic rearrangements in poplar revealed by non-targeted metabolomics. *Plant, Cell & Environment*, 38, 892-904.

2.1 Background

High irradiation intensities containing the UV-B part (290-315 nm) of the solar spectrum typically result in high production and accumulation of ROS/RNS species, especially of superoxide radicals in the exposed plant tissue. An important molecule involved in the protection of leaves against ROS species, is isoprene. Non-isoprene emitting grey poplars represented therefore, an excellent system to study the role of isoprene emission in the adaptation of leaves to UV-B exposure.

2.2 Experimental setup

To study the interplay between UV-B radiation and isoprene emission in more detail, the transgenic NE poplar lines Ra2 (Behnke et al., 2007) and Rb7 (Kaling et al., 2015) were grown together with IE wild type and empty vector control plants to a 20 leaf stem stage under greenhouse conditions in the absence of UV-B radiation. Then plants were transferred into the sun simulators of EUS (Thiel et al., 1996) for an acclimation phase of 10 days maintaining the relative humidity (50-70%) and temperature profiles (26-30°C) as in the greenhouse, while photosynthetic active radiation (PAR) was slowly increased to a maximum of $700 \mu\text{mol m}^{-2} \text{s}^{-1}$. After acclimation, plants were divided into two batches. Half of each group was transferred into the simulator where UV stress was generated by providing a high UV-B irradiation of 2.8 W m^{-2} and 19 W m^{-2} of UV-A over 13 consecutive days (Fig. 2.1A). The other half stayed in the initial simulator as control plants. Other environmental conditions were identical in the two simulators. The three leaves above and below leaf #9, plus leaf number nine, were randomly harvested after plants went through the entire daily

length of maximum UV-irradiation at 0, 12, 36, 60, 108, 209 and 300 hours, day 0, 1, 2, 3, 5, 9 and 13 respectively, of UV-B exposure (Fig. 2.1B).

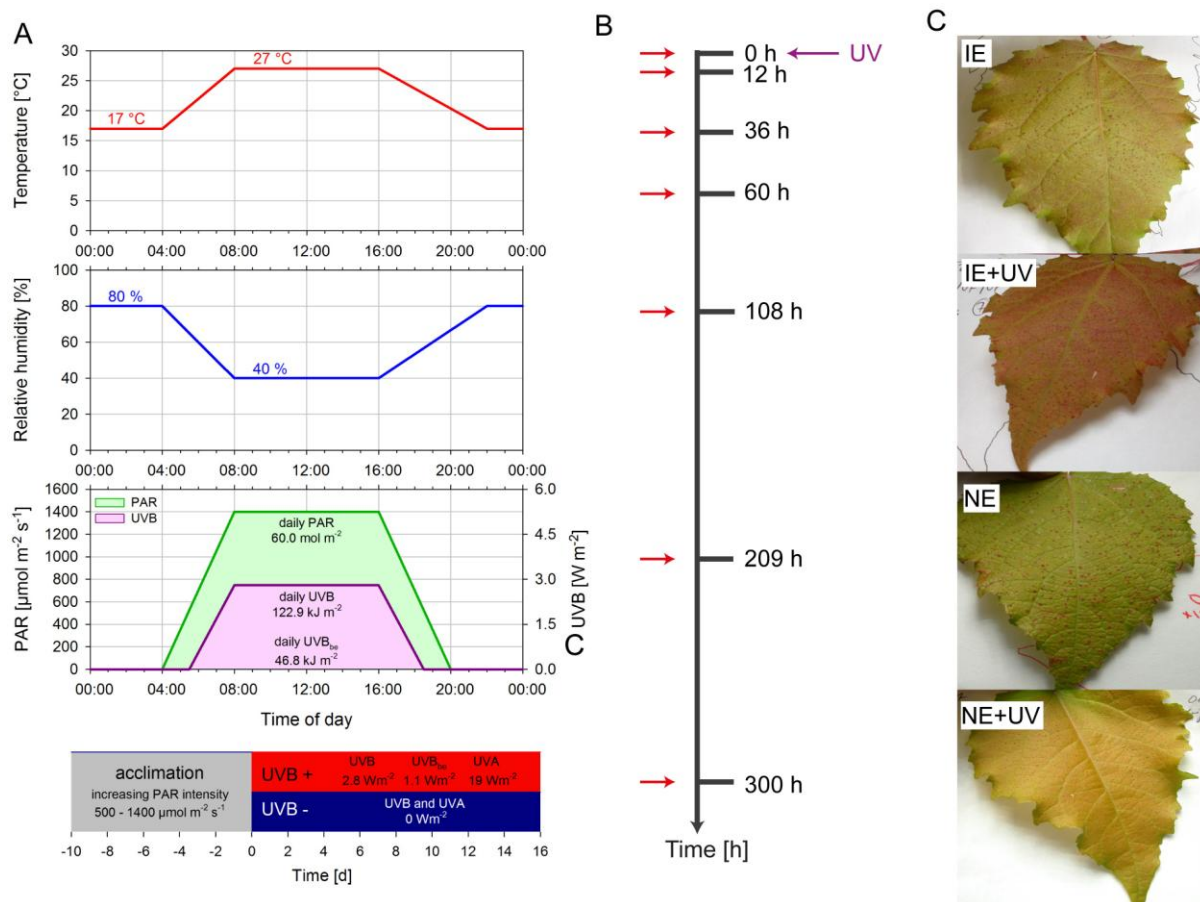


Figure 2.1. Experimental setup of the phytotron experiment. (A) Temperature, relative humidity, photosynthetic active radiation (PAR) and UV-B conditions, (B) leaf harvest time points are indicated by red arrows and (C) emerging leaf phenotypes after 300 hours of UV exposure (modified according to Kaling et al., 2015).

2.2.1 Physiological measurements

After 300 hours (13 days) of UV-B treatment, IE and NE plants developed distinct visible leaf color phenotypes (Fig. 2.1C). The alteration of leaf coloration is a strong indication of changes in pigment compositions. Indeed, HPLC measurements of phenolic compounds in NE and IE leaves revealed the different accumulation of UV-absorbing pigments (Kaling et al., 2015). To investigate whether these changes lead to differences in light penetration into the leaves, confocal laser scanning microscopy (cLSM) of z-sections of leaf surfaces was performed.

The cLSM results supported the initial leaf color observations in that the exposure to UV-B radiation is line and UV-treatment dependent. Statistical analysis of leaf grey scale intensities showed significant differences in penetration depth and light attenuation between

UV-treated and control IE leaves ($P < 0.001$), and between UV-treated IE and NE leaves, but no difference between UV-treated and control NE leaves was discovered (Kaling et al., 2015).

Measurements of net CO_2 assimilation showed a decrease of 22% to 37% under UV-B exposure in the NE lines from day one to nine. Until day 13, however, NE poplars completely recovered their assimilation rates. Additionally the isoprene emission rates decreased when plants were exposed to UV-B radiation over the course of the experiment.

2.3 Results

2.3.1 Normalization and statistical analysis of the ICR-FT/MS data

The leaf extracts were measured over three consecutive days in three batches. The ICR-FT/MS spectra were internally calibrated and exported as ascii files to peak height lists at a signal to noise ratio of 2 using the Data Analysis 4.0 software package (Bruker, Bremen, Germany). The peak lists were combined to a peak matrix with an error of 1.5 ppm using an in-house written tool (Lucio et al., 2011). In order to have a first overview of the ICR-FT/MS raw data, unsupervised PCA analysis was performed (Fig. 2.2A). Unfortunately the PCA revealed a strong batch effect in that batch one clustered together with the majority of batch two, with the exception of the last eight measurements of batch two, which formed a separate cluster together with batch three. The plot of the total ion current (TIC), defined as the sum of all intensities, as a function of sample position showed a significant increase of the TIC in batches two and three, respectively (Fig. 2.2B). Most probably the sensitivity of the ICR-FT/MS spectrometer changed, resulting in higher amplitudes of electrical noise. The batch effect was eliminated by data filtration and normalization as follows: (i) removal of single mass events, (ii) deletion of peaks which were detected in less than 50% of all biological replicates ($n=4$), (iii) removal of peaks with no corresponding ^{13}C isotopic peak to avoid signals generated by electrical noise and (iv) TIC normalization was performed according to equation 8:

$$I_j = \left(\frac{I_i}{\text{TIC}_n} \right) \cdot \text{TIC}_{\text{av}} \quad \text{Eq. 8}$$

where I_j is the normalized intensity, I_i is the peak intensity, TIC_n is the TIC of the current spectrum and TIC_{av} is the average TIC over all spectra.

The remaining 4412 masses were subjected to PCA analysis (Fig. 2.2C). The PCA explains a distinct clustering of NE and IE plants (line effect) in the PC1xPC2 score plot (Fig. 2.2D).

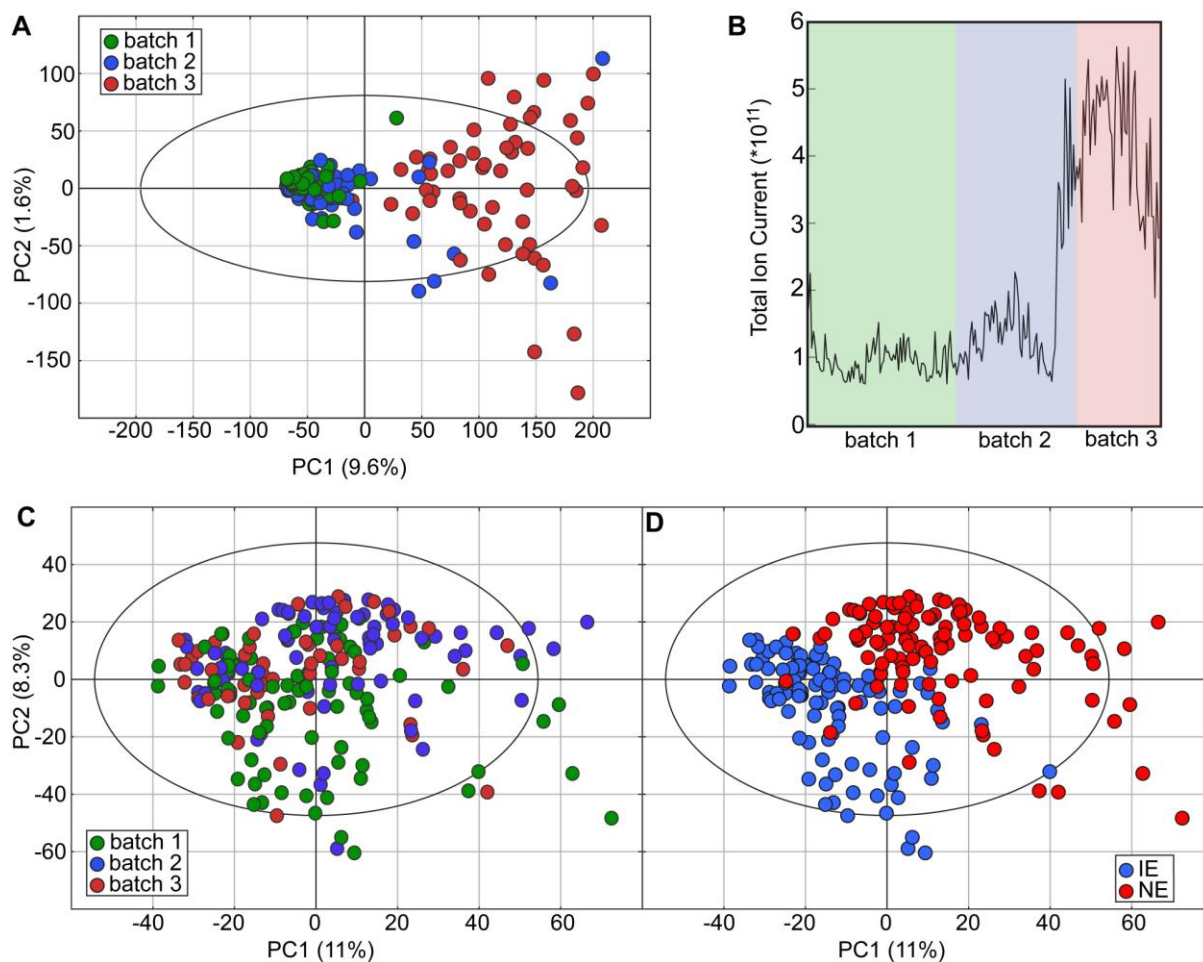


Figure 2.2. Results obtained by ICR-FT/MS measurements. (A) PCA of raw ICR-FT/MS data colored according to batches, (B) total ion current (TIC) plotted as a function of sample order, (C) PCA scatter plot of TIC-normalized and filtered ICR-FT/MS data colored according to batched and (D) according to isoprene emission capacity.

To identify and extract discriminant masses, the average peak intensity of each mass calculated over all four biological replicates ($n=4$) was utilized as variables “X” for PCA and Partial Least Squares (PLS) Regression, respectively, using the software package The Unscrambler 8.0 (CAMO Software AS, Oslo, Norway). The X-variables were centered and scaled to unit variance. Both PCA and PLS models were validated using full cross-validation. Additionally, to test for overfitting, the stability of PLS models was assessed using the jack-knifing based approach of Martens Uncertainty Test (Martens and Martens, 2000).

Separate PLS models were calculated to differentiate samples according to their isoprene emission ability (i.e. NE/IE), time behavior and UV-B exposure. Discriminant masses were

discovered by means of Martens Uncertainty Test from PLS models type 2 calculated with two Y-variables, the first describing NE/IE ($Y=0/1$), the second describing the different time points ($Y=0-6$). Similarly, a second PLS model type 1 was calculated for the extraction of discriminant masses describing the applied UV treatment with one Y-variable UV-B-/UV-B+ ($Y=0/1$). The fitness of a PLS model was measured by a regression between the measured- vs predicted- Y scores. Slopes and coefficient of determinations (R^2) of those regressions were always > 0.91 and > 0.96 for calibration models, and > 0.82 and > 0.92 for the validation models, respectively.

2.3.2 Non-targeted metabolomics reveals extensive metabolic rearrangements related to UV-B treatment and genetic background

A more detailed analysis of the PCA score plot revealed a time-dependent sub-clustering of UV-exposed from control plants within the NE and IE clusters (Fig. 2.3A). Thereby, the longer plants were exposed to UV-B radiation, the more separation exists between control and UV-treated plants in the PCA score plot. The metabolome-based clustering corresponds to the initial leaf-phenotype observation, indicating an alteration of leaf metabolism.

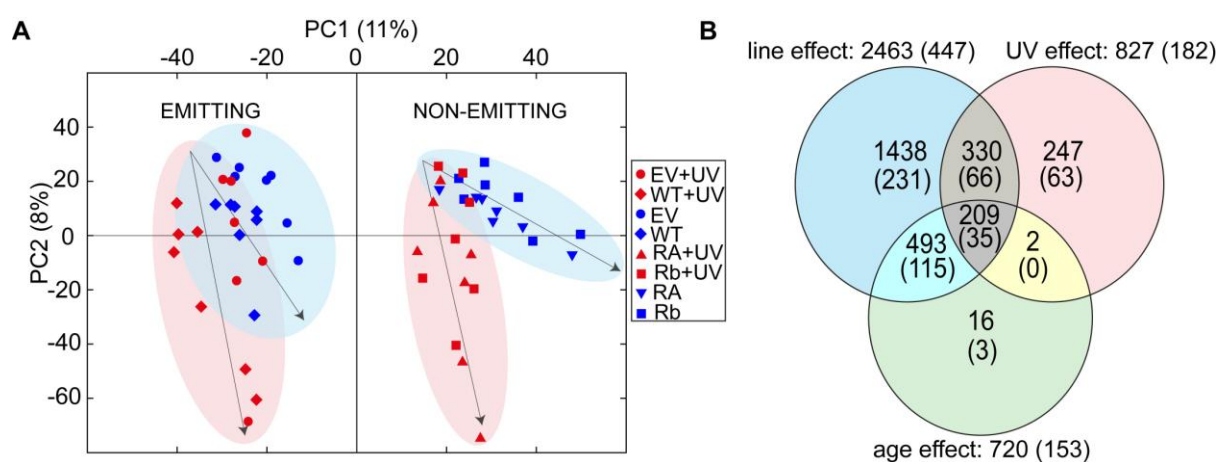


Figure 2.3. Statistical analysis of normalized ICR-FT/MS data. (A) PCA score plot of peak intensities the arrows indicate the time trend of the data and (B) Venn-plot of masses discriminant for the different experimental variables i.e. line, UV-exposure and ageing. The number of annotated metabolites is written in brackets (modified according to Kaling et al., 2015).

From each PCA cluster the discriminant variables were extracted via separate PLS models. Out of a total of 2735 discriminant masses 2463 masses were found to be discriminant for the separation of IE from NE plants (line effect), 827 masses were discriminant for the UV-B treatment and 720 masses were responsible for the separation of time points (Fig. 2.3B). To exclude leaf ageing effects, an UV and line independent PLS

model in which all discriminant masses were subjected as X-variables, was calculated. It explaining a total of 84% Y-variance within the first 2 components and had a slope of 0.89 and correlation coefficient of 0.98 between measured and predicted Y of the validation set of samples. Only 16 masses were found to be discriminant for ageing (Fig. 2.3B).

To check which masses were discriminant for more than one experimental variable, they all were compared in a Venn-plot (Fig. 2.3B). A striking number of 539 masses were significant for both the line- and UV-differentiation.

For metabolite identification the mass-features were uploaded to the MassTriX 3 server selecting *Populus trichocarpa* as organism (Suhre and Schmitt-Kopplin, 2008), PoplarCyc database (Zhang et al., 2010), Knapsack database (Oikawa et al., 2006) and Metlin database (Smith et al., 2005) with an maximal error acceptance of 3 ppm. After removing compound classes that were not present in poplar, such as alkaloids and pesticides, a total of 1012 (22.9%) annotations remained of which 513 were discriminant (Fig. 2.3B).

All metabolites were classified according to their respective compound classes and KEGG metabolic pathways, irrespective of the database in which they were annotated (Fig. 2.4). This revealed line- and UV-B-dependent metabolic perturbations of the primary and secondary metabolism.

Pathways that were strongly affected by the ISPS silencing included carbohydrate, lipid, phenolic and terpene metabolism. Previous work on NE poplar already showed the suppressed production of phenolic compounds on the metabolic level as well as on the transcript level (Behnke et al., 2010a). Additionally, the fatty acid metabolism is known to be affected by the loss of isoprene emission. Interestingly, some masses were annotated as iridoids, which represent cyclic monoterpene glucosides and are synthesized through the plastidic MEP pathway. Until now, this compound class has not been verified to be present in poplar. However, the respective P450 enzymes, which are needed for their biosynthesis are putatively predicted to be present in the poplar genome (Tuskan et al., 2006). Unfortunately, due to low ion abundances, no MS/MS spectra could be measured.

Similar pathways showed also a strong response to UV-B exposure, the majority (70.6 %) of annotated metabolites being phenolic compounds or steroids. Only few metabolites belonged to carbohydrate, amino acid, energy or terpenoid metabolism. However, terpenoid metabolism such as the biosynthesis of diterpenes and carotenoids seemed to be unaffected by the UV-B exposure.

Non-targeted metabolomic comparison of isoprene emitting and non-emitting poplars grown under high UV irradiation

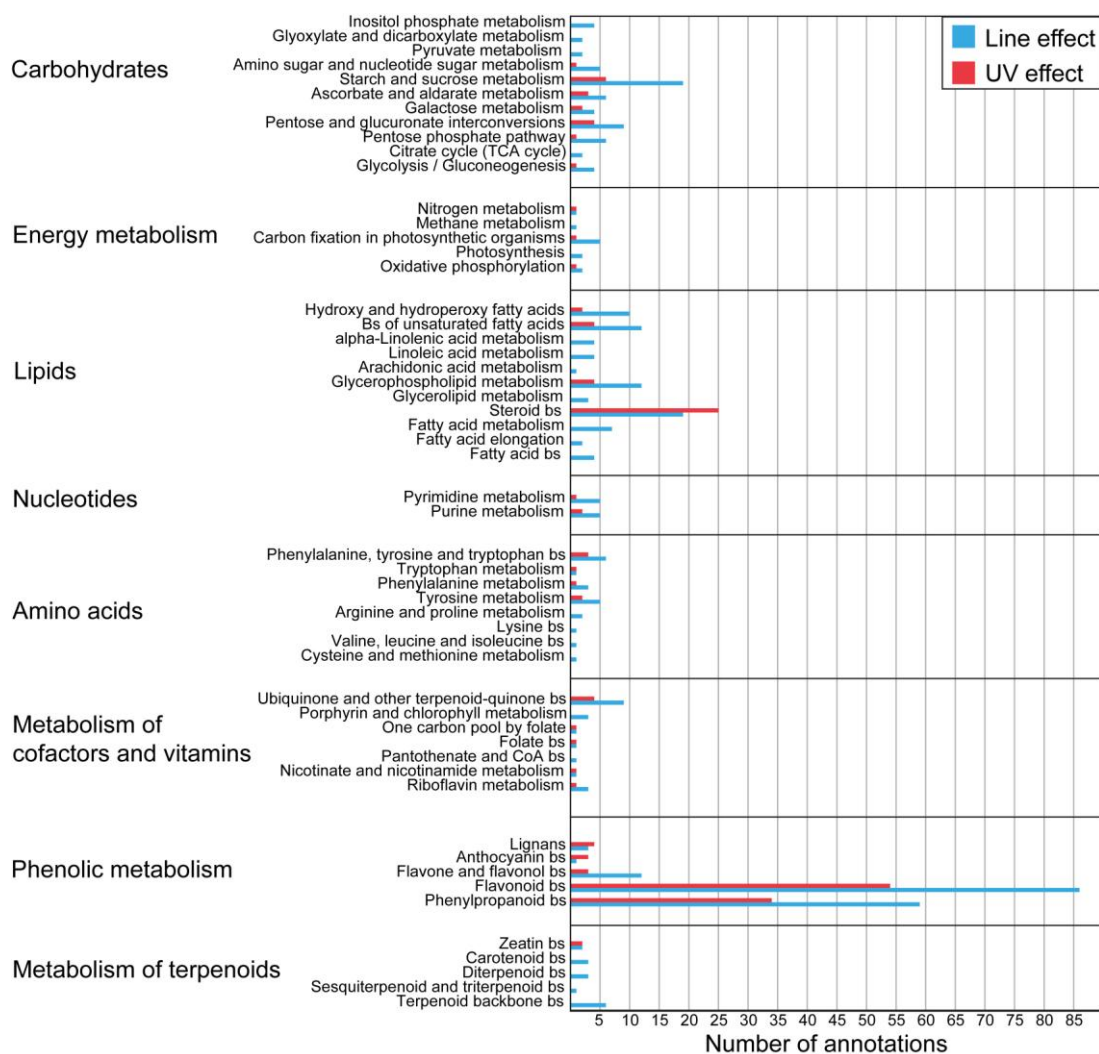


Figure 2.4. Database annotations for masses discriminant for either line- or the UV-B effect. The annotations were arranged according to their respective compound classes and KEGG metabolic pathways. Annotations were not corrected for multiple pathway entries (modified according to Kaling et al., 2015).

2.3.3 Investigation of line dependent metabolic changes under UV exposure

To take a closer inspection on the 539 mass features discriminant for isoprene emission capacity and UV-exposure (Fig. 2.3), the respective peaks and their intensities were subjected to hierarchical clustering analysis (Fig. 2.5). The dendrogram revealed the presence of three major classes consisting of (i) UV-B-exposed NE leaves, (ii) UV-B-exposed IE leaves and (iii) leaves sampled after 209 and 300 h of UV-B exposure. This complements and validates the results obtained by the PCA.

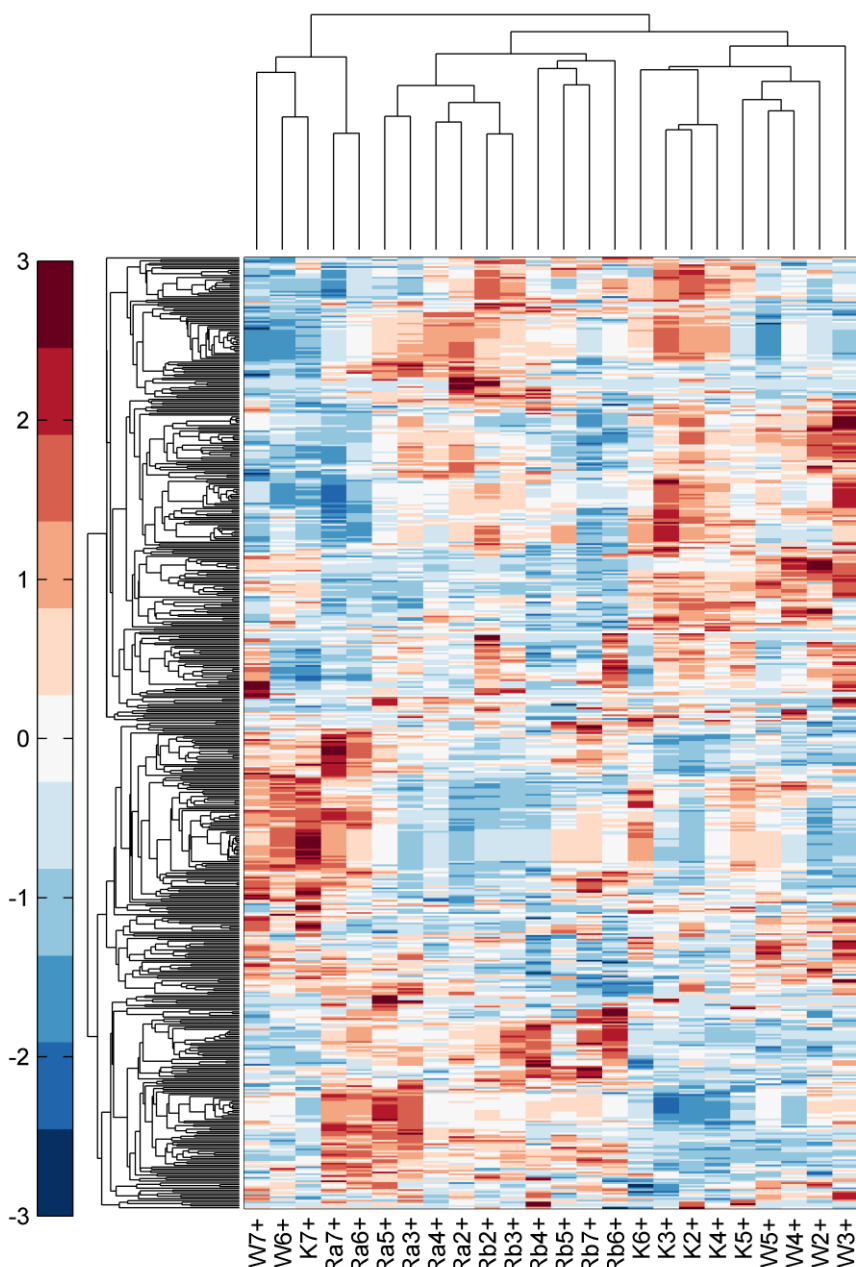


Figure 2.5. Clustermap of masses discriminant for both, isoprene emission capacity and UV-B exposure. Numbers indicate the sampling order as follows: 2 = 12h, 3 = 36, 4 = 60 h, 5 = 108 h, 6 = 209 h and 7 = 300 h of UV-B exposure.

The 2735 discriminant mass features were subjected to the Hierarchical Clustering Explorer (HCE), which allows the supervised study of time trends. The focus was on three major time profiles: (i) line-dependent metabolic differences which are independent of UV-B exposure (Fig. 2.6 A&B), (ii) a constant increase of metabolite concentrations over the course of the experiment (Fig. 2.6C), and (iii) a fast decrease of metabolite concentrations within the first 12 hours of UV-B exposure (Fig. 2.6D). The latter two were already observed by HPLC analysis of phenolic compounds (see Fig. 1 in Kaling et al., 2015). Indeed, 47 %

(1287 out of 2735) of the discriminant mass features gathered in one of these four different time profiles (Fig. 2.6). Overall, 972 mass features matched the time profile of the UV-independent line effect (Fig. 2.6A-B), of which 604 were up-regulated in NE plants (Fig. 2.6A) and 368 mass features were down-regulated in NE poplars (Fig. 2.6B).

Database annotations of both clusters showed that the silencing of ISPS causes an alteration of plant phenolic and lipid concentrations that is conserved upon UV-B exposure. Additionally, NE poplar had higher contents of carbohydrates and terpenes compared to IE poplar. Especially products of the MEP pathway, such as DMAPP, FPP and farnesyl monophosphate, were up-regulated in the NE genotypes. Entry reactions of the shikimate and MEP pathway consume intermediates and products of glycolysis and of the pentosephosphate pathway, i.e. pyruvate, glyceraldehyde 3-phosphate, PEP and E4P (Maeda and Dudareva, 2012; Banerjee and Sharkey, 2014). Consequently, a decreased production of phenolic compounds with a subsequent decrease in MEP metabolic fluxes (Ghirardo et al., 2014) may lead to increasing carbohydrate contents, by way of feedback inhibition of phosphofructokinase (PFK). PFK catalyzes the rate-determining step of glycolysis and is inhibited by high cellular ATP concentrations (Passonneau and Lowry, 1964). Important steps of the MEP and shikimate pathway consume ATP. A flux reduction in both pathways therefore saves ATP molecules, leading to higher ATP/AMP ratios, which might reduce PFK activity.

Three gibberellin derivatives, i.e. gibberellin A8-catabolite, gibberellin A28 and gibberellin glucoside, which are diterpene phytohormones that regulate plant growth (Brian, 1959), were also found to be up-regulated in NE leaves. A two-year field study, in which the performance of NE and IE plants was compared, showed an increase of 6.9% net growth of NE poplars (Behnke et al., 2012). An increase of gibberellin concentrations partially explains the slight increase of NE net growth.

87 % of annotated UV-induced metabolites either were flavonoids, phenylpropanoids or phenolic glucosides (Fig. 2.6C). This finding complements the cLSM measurements, showing the reduced light transmission into UV-B exposed leaf z-sections (Kaling et al., 2015), and the time dependent accumulation of anthocyanins and flavonoids determined by HPLC (Kaling et al., 2015). It is known that under UV-B exposure, anthocyanins and flavonoids accumulate in vacuoles and cell nuclei of leaf's epidermal tissues for photo-protecting and antioxidant purposes (Schmelzer et al., 1988; Yamasaki et al., 1997; Burchard et al., 2000; Agati et al., 2012). The cellular redox-potential is of major importance under high UV-B doses because cellular ROS concentration rise, leading to lipid and protein oxidation and finally to cytochrome c release (Nawkar et al., 2013).

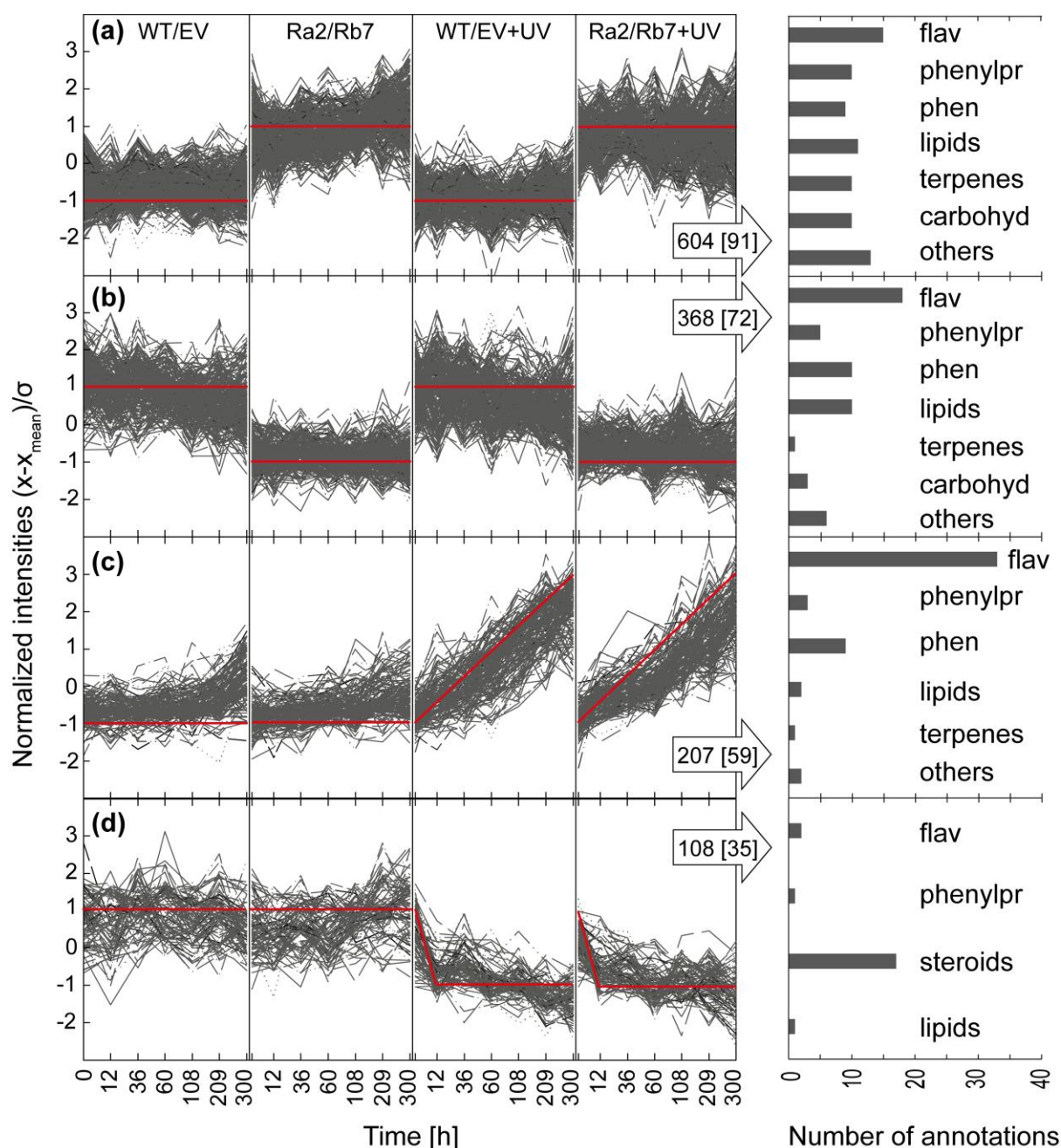


Figure 2.6. Time behavior of normalized signal intensities of all discriminant masses. The masses cluster in four major time profiles: up-regulated metabolites in (A) NE (Ra2/Rb7), (B) in IE (WT/EV), (C) metabolites induced under UV-B irradiation and (D) metabolites decreasing within the first 12 hours of UV-B exposure. Red lines indicate the time profile entered into the profile search function of HCE 3.5 and grey lines masses, which matched the entered profile. The number of masses within a cluster is given in the arrows. Numbers in brackets stand for the amount of annotations. The annotations were grouped into compound classes and visualized in bar plots. Flav: flavonoids; phenylpr: phenylpropanoids; phen: phenolic; carbohyd: carbohydrates (reprinted with permission from John Wiley and Sons).

After 12 hours of UV-B exposure, a total of 108 mass features, of which 35 (32%) could be annotated as metabolites, decreased drastically. Interestingly, 16 of those annotations

were steroids, such as vitamine D2 derivatives. Only a few masses were annotated as phenolics or lipids. The HPLC analysis of plant phenolics revealed a similar time trend for hydroxycinnamic acid derivatives (Kaling et al., 2015). However, the strong decrease of steroid concentrations has not been described yet. Moreover, the down-regulation of steroid biosynthesis contradicts earlier work on *Vitis vinifera L.* (Gil et al., 2012), where an up-regulation of membrane-related steroids was reported. Steroid biosynthesis channels through the MEP antagonistic cytosolic mevalonate pathway (Bishop and Yokota, 2001).

Both pathways are linked through pyruvate, which on the one hand represents an important substrate for DXS that catalyzes the entry reaction of the MEP pathway, and on the other hand, it can be converted to acetyl-CoA, of which three molecules are needed to form 3-hydroxy-3-methylglutaryl-CoA, an important molecule in the mevalonate pathway (Banerjee and Sharkey, 2014). Consequently, the four pathways affected by UV-B exposure and by the genetic modification, namely MEP pathway, mevalonate pathway, fatty acid biosynthesis and the shikimate pathway, are all connected through the important building blocks PEP, pyruvate and acetyl-CoA.

The drop in steroid contents with a subsequent reduction of isoprene emissions rates under UV-B radiation (Kaling et al., 2015) might illustrate a step towards adjusting metabolic fluxes in the direction of phenolic biosynthesis, facilitating UV-B adaption in grey poplar. During the process of metabolic adaption, the activity and regulation of pyruvate kinase, which determines the PEP/pyruvate balance, may represent the key process of UV-B adaption. The absence of isoprene emission is accompanied by a decrease in C-flux through the MEP pathway (Ghirardo et al 2014) and therefore may alter the PEP/pyruvate balance which leads to changes in the biosynthesis of phenolic compounds under control and UV conditions in NE poplars.

2.3.4 UV-B exposure causes line dependent responses in the biosynthesis of phenolic compounds

A detailed analysis of the flavonoid and phenylpropanoid metabolism, revealed line dependent regulatory processes (Fig. 2.7). The biosynthesis of phenylpropanoids starts with either phenylpyruvate, a precursor of Phe, or with 4-hydroxyphenylpyruvate, a precursor of Tyr. Interestingly, IE poplars had relative higher amounts of phenylpyruvate, whereas NE plants had higher amounts of 4-hydroxyphenylpyruvate. Over the course of the experiment, the average \log_2 -ratio of peak intensities of phenylpyruvate divided by 4-hydroxyphenylpyruvate in UV unexposed IE leaves was 1.07 compared to 0.18 in UV unexposed NE leaves ($P = 5.24 \cdot 10^{-7}$). When UV-B radiation was applied, the concentration of both compounds decreased, which implies an increase of turnover of these compounds in both phenotypes. Still, the line dependent \log_2 ratios are significantly different ($P = 8.8 \cdot 10^{-4}$).

These line and UV-B-dependent changes could either indicate an alteration of 4-hydroxyphenylpyruvate aminotransferase and phenylpyruvate aminotransferase activities, respectively or a preference for the biosynthesis of Phe and Tyr via arogonate. Both scenarios result in increasing amounts of Phe and Tyr, the key substrates for phenylpropanoid biosynthesis. Under UV-exposure, changes in phenylpropanoids were on biosynthetic end products, such as sinapoyl malate, chlorogenate and 4-hydroxycinnamyl alcohol 4-D-glucose (Fig. 2.7). Especially, sinapate esters are well known for their UV-B protecting properties (Li et al., 1993; Landry et al., 1995). In UV-B exposed flavonoid deficient *Arabidopsis* mutants, an accumulation of sinapoyl malate has been observed to compensate the lack of flavonoids (Landry et al., 1995; Kusano et al., 2011).

Similar line dependent trends were found for the phenolic glycosides coniferin and syringin (Fig. 2.7), which both are precursors for lignin biosynthesis (Boerjan et al., 2003). NE poplar had higher amounts of coniferin whereas IE poplars had higher amounts of syringin. This is also reflected in fold changes ($\log_2[\text{NE}_{\text{syringing}}/\text{NE}_{\text{coniferin}}]$), where control NE poplars had a ratio of -0.43 and control IE plants of 0.22 ($P = 4.7 \cdot 10^{-6}$). Exposure to UV caused an increase of syringin concentration, with similar trends in log ratios (-0.21 in NE and 0.37 in IE, $P = 8.3 \cdot 10^{-6}$), in both phenotypes. Consequently, it might be speculated whether these differences are also present in the lignin composition of NE and IE plants. Strikingly, a previously conducted wood composition analysis showed lower concentration of syringyl lignin in NE plants compared to IE plants (Behnke et al., 2012).

Comparable trends are also found in the biosynthesis of flavonoids. IE genotypes had higher contents of metabolites leading to hesperidin, while NE plants preferred the biosynthetic route to naringin. Interestingly, flavonoids used for anthocyanine biosynthesis were up regulated in IE leaves, while the precursors for proanthocyanidin biosynthesis i.e. galocatechin and catechin were down regulated.

Accumulation of anthocyanines in leaf vacuoles typically results in the development of red leaf colors (Manetas, 2006), which was observed in IE poplars under UVB exposure (Fig. 2.1C), during leaf senescence. In contrast to that, proanthocyanidines, whose relative concentrations are higher in NE plants, are usually colorless or brown (Koes et al., 2005). Determination of anthocyanine concentrations by HPLC revealed a faster production of anthocyanines under UV-B exposure in IE poplars (Kaling et al., 2015). The cellular pH value sets the mesomerism of the anthocyanine ion and therefore its color. *In vitro* the red flavylium cation is the predominant species at a pH value of 2. At higher pH values ($\text{pH} > 4$) the colorless hemi-acetal and chalcone structures dominate, resulting in bleaching (Lima et al., 2002). However, the information regarding the ion mesomerism of anthocyanins might be lost upon extraction and the acidic workup procedure that is needed for anthocyanine determination by HPLC-UV-VIS analysis (Kaling et al., 2015).

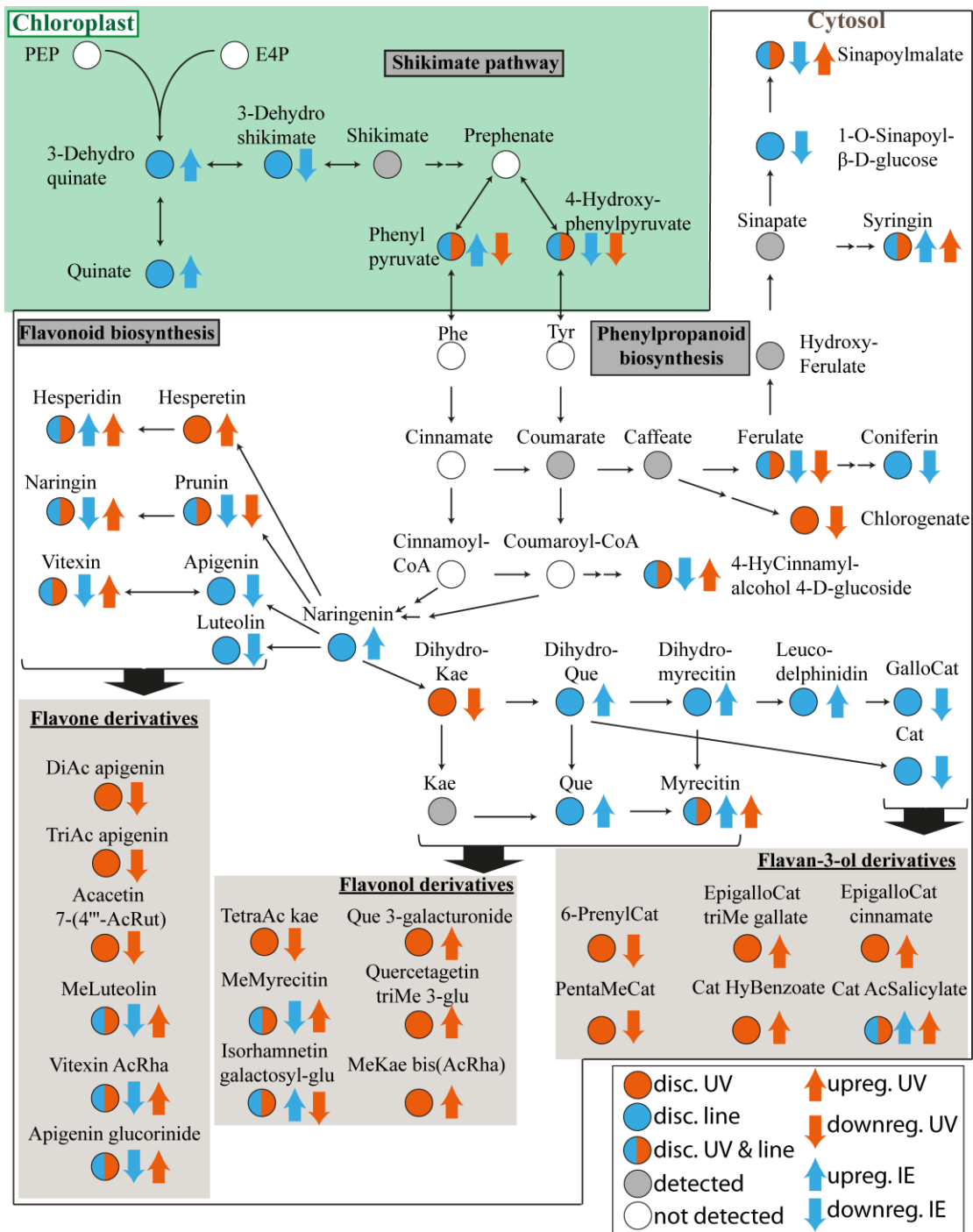


Figure 2.7. Map of the shikimate, phenylpropanoid and flavonoid pathways/biosynthesis and their respective compartmentalization. Ac, acetyl; Cat, catechin; downreg, downregulated; glu, glucose; Hy, hydroxyl; Kae, kaempferol; Me, methoxy; Que, quercetin; Rha, rhamnoside; Rut, rutenoside; upreg, upregulated (modified and reprinted with permission from John Wiley and Sons).

Additionally, it has been reported that anthocyanins form charge-transfer complexes with hydroxycinnamic acids and hydroxyflavones resulting in co-pigmentation and also in stabilization of the red colored flavylum cation (Ferreira Da Silva et al., 2005). Indeed, the cLSM measurements showed differences in the accumulation of hydroxy cinnamic acids and

flavonoids between NE and IE leaves. Previous work on red-leaved plants revealed the accumulation of colorless phenolic compound that contribute more than anthocyanins to the overall absorbance of UV radiation (Woodall and Stewart, 1998).

Under UV-B NE lines also induced flavone biosynthesis as seen by higher content of apigenin, luteolin, vitexin and their biosynthetically modified derivatives (Fig. 2.7). Rather than on pathway intermediates UV-B radiation caused a change on biosynthetic end products and polyphenols, such as methylated and acetylated flavonoids. O-Methylations and acetylations increase lipophilicity (Ibrahim et al., 1987; Harborne and Williams, 2000; Zhou et al., 2006), which facilitates the embedment into biomembranes (Brisson et al., 1986; Charest et al., 1986). Additionally, O-methylation blue shifts the absorption maxima, typically to the 250-320 nm region (Harborne and Williams, 2000). An increase in flavonoid lipophilicity might be correlated with transportation processes responsible for the integration of flavonoids into cell walls of the epidermal tissue or chloroplasts, where they function as UV-absorbers and quench singlet oxygen (Agati et al., 2007).

Catechin and epicatechin based polyphenolic compounds connected to a galloyl residue, also known as theaflavins (Quideau et al., 2011), were also induced upon UV-B exposure. The hydroxybenzoic acid moieties of theaflavins hold an absorption maxima within the UV-B range (290-315 nm), making them perfect UV-absorbing pigments (Pietta, 2000; Quideau et al., 2011). On top of that, a key characteristic of polyphenols is their ability to directly quench ROS species (Pietta, 2000; Zhou et al., 2005; Quideau et al., 2011).

2.3.5 Linking isoprene emission capacity and the diminished production of phenolic compounds

Although the preference for biosynthetic routes of NE and IE genotypes varied, the overall production of phenolic compounds is reduced and delayed in NE poplars. Besides changes in the PEP/pyruvate ratio, another reason might be the successive reorganization of signaling processes involving isoprene. Characteristic abilities of isoprene include the protection against the damaging properties of ROS species (Loreto and Schnitzler, 2010) as well as to modulate ROS/RNS signals (Vickers et al., 2009; Loreto and Schnitzler, 2010). This can include H₂O₂-mediated signaling cascades associated with flavonoid and phenylpropanoid biosynthesis. Earlier work shows that NE poplar leaves accumulate *in vivo* higher contents of H₂O₂ upon illumination with PAR (Behnke et al., 2010a).

It is known that certain H₂O₂ concentrations and changes in cellular redox potential (Heine et al., 2004; Taylor and Grotewold, 2005; Dubos et al., 2010) activate transcription factors triggering anthocyanine biosynthesis under high light intensities (Biosynthesis et al., 2005). Parts of this transcriptional regulation are MYB (myeloblastosis) transcription factors, which are UV-B inducible and lead to the activation of proanthocyanidin biosynthesis in

poplar (Mellway et al., 2009). In *Arabidopsis* certain MYB factors control flavonoid accumulation (Stracke et al., 2007) and are activated together with genes responsible for leaf pigment production under UV-B exposure (Zhao et al., 2007).

Previous transcriptional analysis of NE poplars demonstrated a down-regulation of MYB or MYP-like transcription factors (orthologs of AtTTG1, coding for a WD40 protein, AtMYB4, AtMYB58) that coincided with reduced proanthocyanidin accumulation (Behnke et al., 2010). In poplar (*P. trichocarpa*), a total of 192 genes are coding for R2R3-MYB transcription factors (Wilkins et al., 2009) of which MYB134 is induced by UV-B leading to the activation of proanthocyanidin biosynthesis in poplar (Mellway et al., 2009). Whether some of these transcription factors are responsible for the line dependent behavior under UV-B exposure is a logical question, which should be addressed in future experiments.

3 Mass-difference network analysis of non-isoprene and isoprene emitting poplars

Parts of this chapter were published in:

Kaling M.*, Moritz F.*, Schnitzler J.P. and Schmitt-Kopplin P. (2015) Mass difference analysis as a generic framework for data-driven plant systems chemical biology. (submitted)

*both authors contributed equally

3.1 Background

The non-targeted metabolomics data convincingly showed the metabolic perturbations of NE and IE plants under UV-B exposure. A common problem of metabolomic experiments, however, is that only 10 to 30 % of mass features are directly tractable to databases. Combining a variety of metabolic databases, such as Popcyc, Masstrix, Knapsack, Metlin, KEGG and LipidMaps (through Masstrix), 22.9% of mass features could be annotated. For example the KEGG database coverage of the UV poplar dataset is only 3%. As already described in chapter 1.4 one possibility of including the whole set of mass features into the data evaluation process is mass difference network (MDiN) analysis and mass difference enrichment analysis (MDEA). These concepts have already been introduced on natural organic matter (Zhang et al., 2014) and on the analysis of human exhaled breath condensate (Moritz et al., 2015). However, a proof-of-principle experiment that convincingly validates the theory behind this powerful method is still needed. Isoprene emitting grey poplars and their well-characterized non-isoprene emitting mutants are the perfect fit for the validation of the mass-difference method. The following sections explain and describe how mass-differences can be used to completely explain metabolic perturbations resulting from gene knockout in a data-driven manner.

3.2 Development of an ideal workflow for mass-difference analysis on plant extracts

3.2.1 Creation of a plant based reaction equivalent mass-difference list

The theory behind mass-differences is that all complex (bio)chemical reactions can be explained by stoichiometry because they are subjected to the law of mass conservation. Converting enzyme catalyzed and spontaneous reactions into sum formulas inherits three important molecular masses: (i) product masses, (ii) educt masses and (iii) the mass of the leaving group. Within a mass spectrum the information of the reaction partner or leaving group is not directly accessible, but it can be extracted by subtracting educt from product masses (Fig. 3.1).

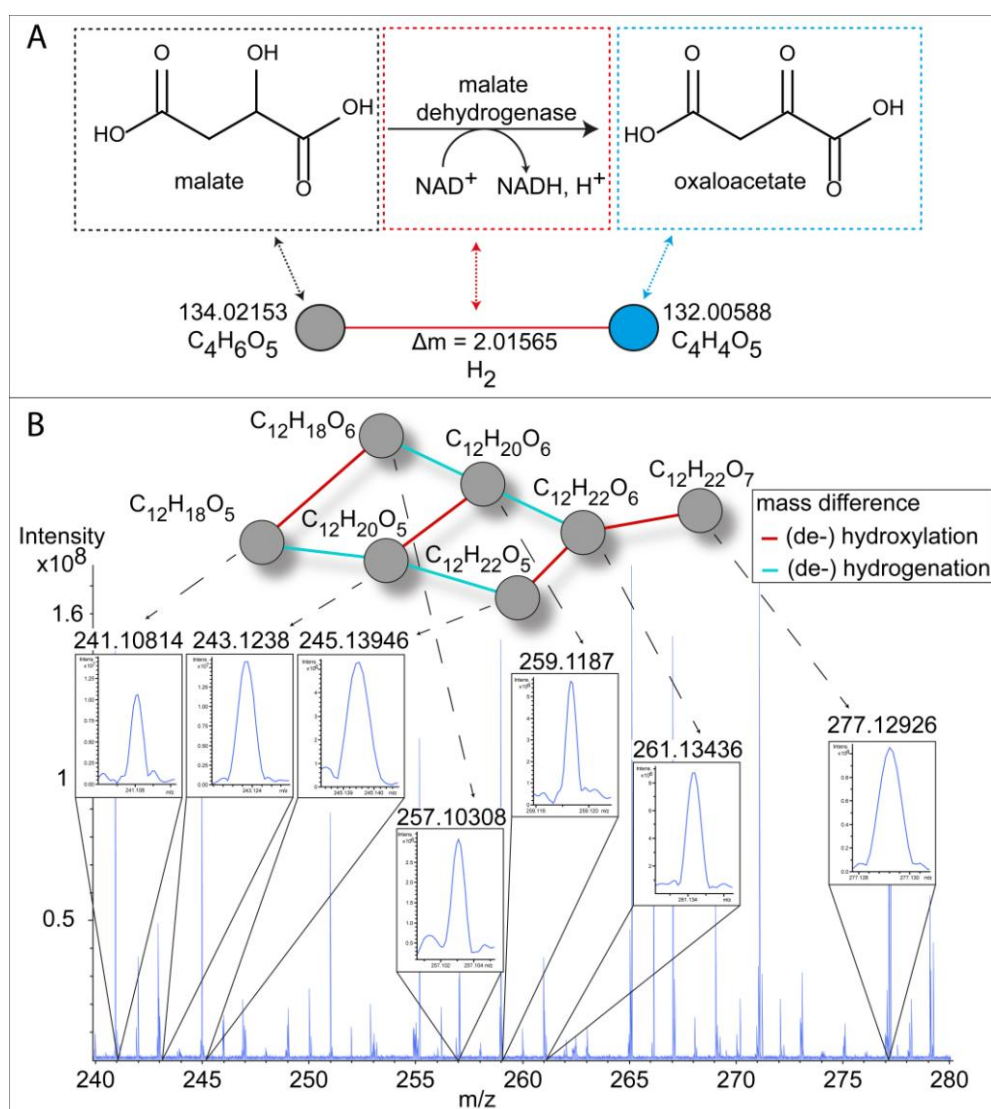
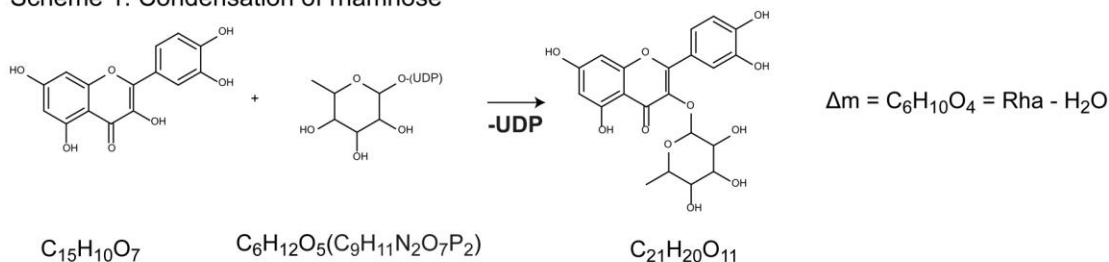


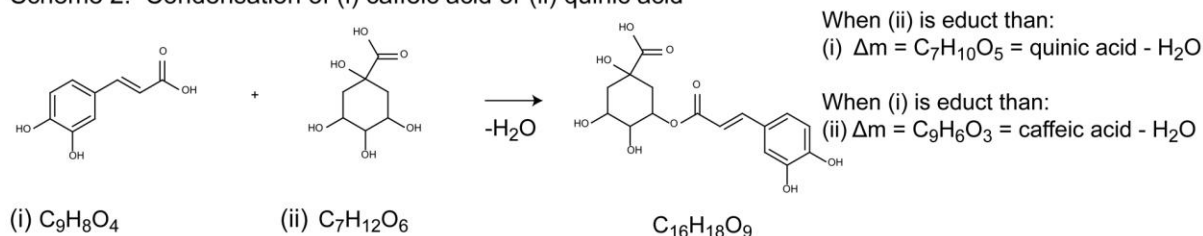
Figure 3.1. The theory of mass difference networks. (A) Conversion of a biochemical reaction into network nodes and edges. (B) Conversion of a mass spectrum into a network.

The REMD set, which was used for the characterization of the grey poplar metabolism, contained the functional set presented in Tziotis et al., (2011) with the extensions made by Zhang et al., (2014) and Moritz et al., (2015). Additionally, plant/poplar specific reactions, such as rhamnose, glucose, quinate, phenolic and PEP related REMDs, were added (Fig. 3.2, Table 3.1).

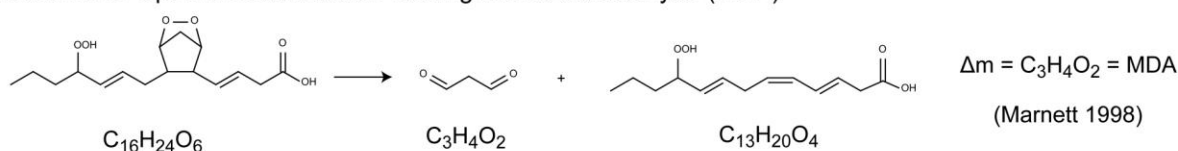
Scheme 1: Condensation of rhamnose



Scheme 2: Condensation of (i) caffeic acid or (ii) quinic acid



Scheme 3: Spontaneous reaction leading to malondialdehyde (MDA)



Scheme 4: Spontaneous reaction of malondialdehyde with DNA bases

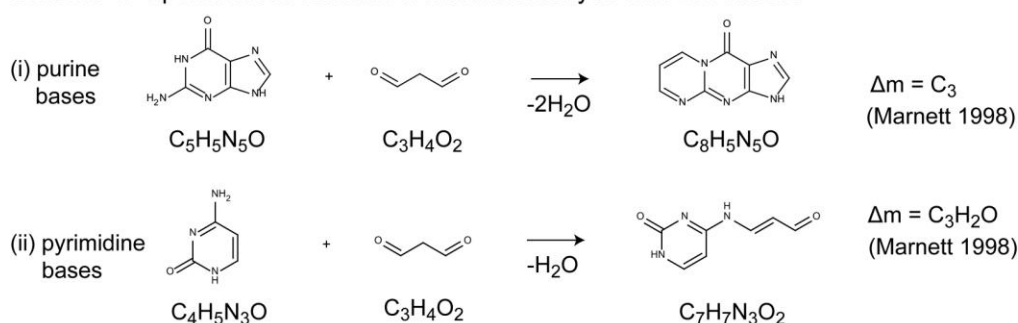


Figure 3.2. Examples of biochemical reactions mimicked by REMDs.

Another important set of reactions pertained to malondialdehyde (MDA), which is a marker for the cellular lipid peroxidation (Moore and Roberts, 1998). MDA is formed via a spontaneous radical reaction (Fig. 3.2; Pryor et al., 1976). Since MDA is a highly reactive molecule it undergoes several spontaneous reactions, e.g. addition to DNA nucleobases, which were also added to the plant REMD list. It is to be noted that co-substrates, e.g.

NADPH or ATP, and reaction activating intermediates, e.g. CoA and uridine diphosphate adducts, are explicitly considered to be part of the mass difference reactions. Figure 3.2 exemplifies the conversion of plant reactions into their respective REMDs.

For this purpose, a total of 248 reactions were translated into mass-differences for which the term reaction equivalent mass-difference (REMD) was coined. This concept extends the approach of Breitling et al., (2006) whose idea was to convert textbook reactions such as hydrogenation, oxygenation, transamination, hydroxylation, etc., into mass-differences for the prediction of metabolic pathways.

Using REMDs together with their respective sum formulas (calculated from mass features) it is possible to create mass-difference networks (MDiN), in which REMDs are represented by the network edges and the mass signal information is given by the network nodes (Fig. 3.1B). In network theory edges represent an interaction between different nodes, which in the case of MDiN are compositional differences between mass features.

In total 19 important plant reactions were translated into REMDs (Table 3.1). The 248 REMDs were classified according to their functionality, such as amino acids, 2-oxo-acids, phenolics, fatty acids, dicarboxylic acids, terpenes etc., and according to their reaction class, representing the mechanistic principle behind the mimicked reaction (see Table 3.1).

Another approach to create a REMD set was introduced by Weber and Viant (2011) who used the KEGG reaction-pairs (rpairs) downloaded from KEGGAPI (Weber and Viant approach, (<http://www.kegg.jp/kegg/rest/keggapi.html>)). The advantage of this approach is that rpairs directly pertain to enzymatic reactions, meaning no manual creation is needed. The disadvantage, however, is that rpairs are limited to known metabolic transformations lacking spontaneous reactions such as the MDA formation. The REMD-rpair list contained 301 REMDs sharing 99 entries with the curated REMD list.

Both approaches were used on the poplar ICR-FT/MS dataset, whereby the curated REMDs were used to annotate sum formulas to MS signals according to Tziotis et al., (2011).

Table 3.1. Plant based reaction equivalent mass-differences (REMDs). The reaction class represents the mechanistic principle behind the mimicked reaction: A = condensation/hydrolyses and F = others are shown in the table. Other mechanisms which are not included in the table are: B = decarboxylative condensation, C = condensation on hydrogenated carbonyls, D = decarboxylative addition and E = THF related reactions.

Reaction partner	Reaction class	Functionality	Mass-difference	Sum formula
Benzylaldehyde	A	Phenolic	88.0313	C ₇ H ₄
Benzyl alcohol	A	Phenolic	90.04695	C ₇ H ₆
Salicyl aldehyde	A	Phenolic	104.02622	C ₇ H ₄ O
Salicyl alcohol	A	Phenolic	106.04187	C ₇ H ₆ O
Salicylic acid	A	Phenolic	120.02113	C ₇ H ₄ O ₂
Cinnamic acid	A	Phenolic	130.04187	C ₉ H ₆ O
Shikimic acid	A	Phenolic	156.04226	C ₇ H ₈ O ₄
Ferulic acid	A	Phenolic	176.04735	C ₁₀ H ₈ O ₃
Erythrose EC 2.5.1.54	F	Tetrose/PEP related	120.04226	C ₄ H ₈ O ₄
MDA-DNA Purine	F	Malondialdehyde	36	C ₃
MDA-DNA Pyrimidine	F	Malondialdehyde	54.01057	C ₃ H ₂ O
MDA formation	F	Malondialdehyde	72.02113	C ₃ H ₄ O ₂
Rhamnose	A	Hexose	146.05791	C ₆ H ₁₀ O ₄
Glucose	A	Hexose	162.05283	C ₆ H ₁₀ O ₅
Quinate	A	Cyclohexane	174.05283	C ₇ H ₁₀ O ₅
+ Phosphate, -CO ₂	F	Transphosphorylation/ PEP related	35.9765	CH ₁ OP
Pyrophosphate	A	Pyrophosphate cleavage	159.93266	H ₂ P ₂ O ₆
Phenylpyruvic acid/Coumaric acid	A	2-oxo-acid/Phenolic	146.03678	C ₉ H ₆ O ₂
Pentose condensation	F	Pentose	132.04226	C ₅ H ₈ O ₄

3.2.2 Sum formula annotation and statistical analysis

The NetCalc based sum formula annotation resulted in 2276 ± 886 non-randomly occurring mass features in each spectrum, surmounting to a total of 4335 monoisotopic sum formula annotations. The standard deviation of annotation error at each m/z is smaller than 0.1 ppm even though the plot is generated over 211 spectra. Overall, the elemental sum formula compositions of NE and IE poplars were CHO (4335), N (2205), S (1938) and P (892). Now, the MDiN is reconstructed based on the theoretical masses of sum formula annotations. This process resulted in the formation of x edges. The decisive point is to

extract the biological important REMDs out of the edge haystack. In order to achieve this, a node-based statistical measure is needed.

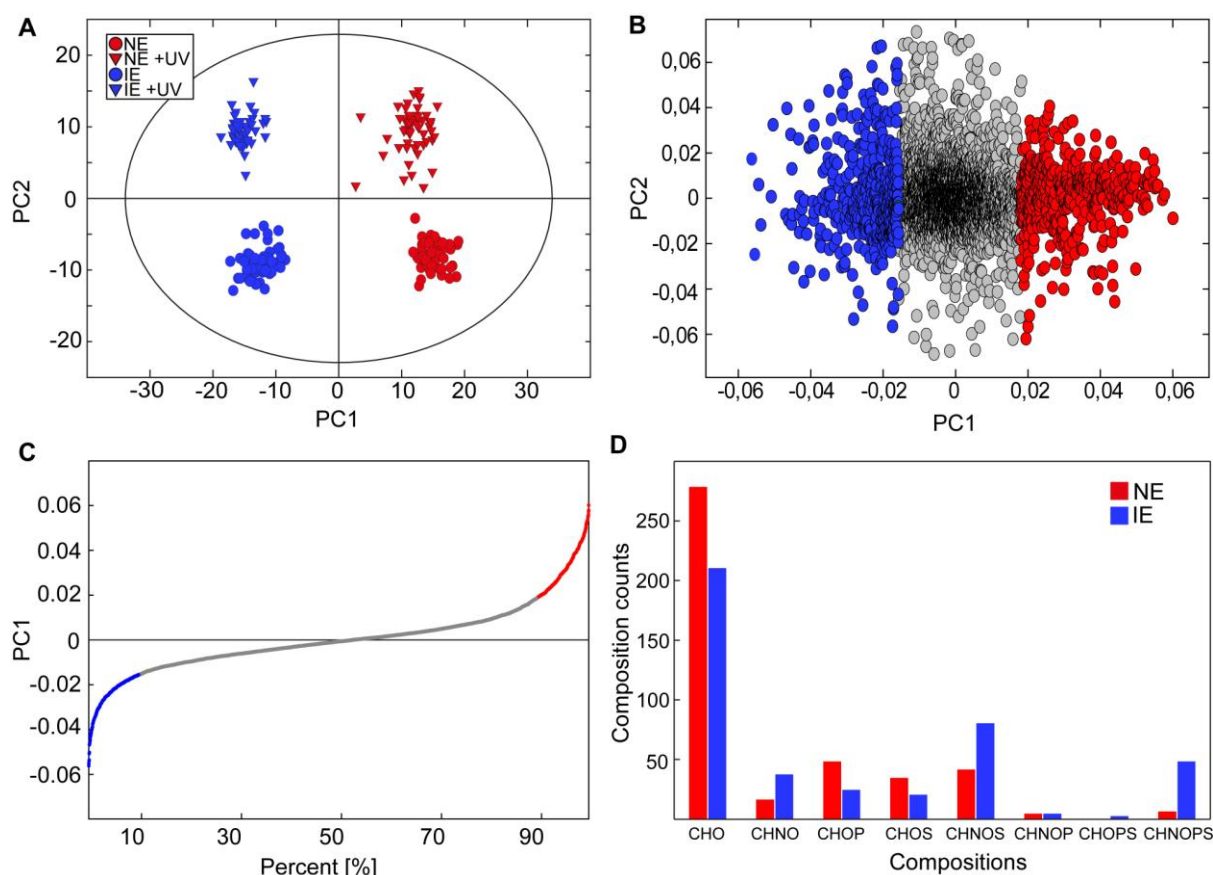


Figure 3.3. Statistical analyses of sum formula annotations from NetCalc. (A) Score plot of OPLS-DA model, (B) loading plot showing the selection of discriminant sum formulas based on (C) the top 10% of NE and IE respectively and (D) elemental compositions of discriminant sum formulas.

Since the NetCalc peak matrix differs from the one used in Chapter 2, the mass feature based statistics had to be re-done. Therefore, orthogonal Partial least squares discriminant analysis (OPLS-DA) was used to select sum formulas discriminant for the knockdown of ISPS. The multivariate model ($R^2Y(\text{cum}) = 0.957$, $Q^2(\text{cum}) = 0.767$) showed the distinct clustering of IE from NE plants in principal component 1 and the clustering of UV-exposed from control plants in principle component 2 (Fig. 3.3A). For the selection of discriminant sum formulas the X-variable loadings were used, where the top 10% of NE up-regulated and top 10% of NE down-regulated sum formulas were extracted (Fig. 3.3B-C).

Comparing the elemental compositions of sum formulas discriminant for NE and IE revealed that CHO, CHOP and CHOS compositions were up regulated by trend in NE plants (Fig. 3.3D). In contrast to that discriminant sum formulas of IE plants nitrogen containing compositions were up regulated (Fig. 3.3D).

Having the molecular formula it is possible to calculate the cyclomatic number (Senior, 1951), which is a measure for the molecules unsaturation. Interestingly, the degree of unsaturation was significantly lower in NE plants ($u_{NE} = 6.6$, $u_{IE} = 9.1$, $p = 2.2 \cdot 10^{-19}$), a fact that coincides with previous publications on NE poplar stating the reduced production of phenolic compounds (Behnke et al., 2010a; Kaling et al., 2015). Looking at the significantly lower H/C ratios and H/(C+N) ratios in IE ($p_{H/C} = 1.4 \cdot 10^{-12}$, $p_{H/(C+N)} = 2.2 \cdot 10^{-23}$) strengthens this finding.

3.2.3 Reaction equivalent mass-differences based statistics

Up to this point the workflow can be summed up in the following steps:

- (i) Peak alignment and calibration of mass spectra,
- (ii) Sum formula annotation of mass features,
- (iii) Network reconstruction and REMD assignment on theoretical monoisotopic exact masses using both REMD-sets, to get rid of post-calibration random errors.
- (iv) Statistical evaluation of annotated sum formulas (nodes) in order to define a node-set of interest.

The third step resulted in the assignment of 63,608 edges when using the curated REMD list and 65,180 edges when using the rpairs. At first glance this process led to an immense increase of data, however, using the nodal statistical information the edges can be divided into three sets. The first edge set consists of all edges E , e.g. 63,608 for the curated REMD list, which are connected to the whole node population. The second edge set includes edges connected to nodes discriminant for NE, and the third edge set contains edges connected to discriminant IE nodes. Now it is possible to perform enrichment analysis similar to Moritz et al., (2015) and test for over/under-representation of REMDs, using the Fisher exact test as follows: (1) reconstruct an MDiN from a population of nodes P and a set of edges E , (2) define a sample set S that contains nodes of interest, (3) count the set of edges, E_S , that are connected to S , (4) count the number of edges of one REMD type R in E ; and (5) count the corresponding number of edges, R_S , that are connected to S (for more details see Moritz et al., 2015).

As a result of this procedure each REMD yields two z-scores that correspond to their involvement in the group classifications (IE and NE). Consequently REMD-based enrichment analysis reduced the immense amount of data from 63,608 edges back to 248 REMDs (curated REMD-list, 65,180 to 301 for rpairs). It has to be noted, that this concept does not differentiate whether compounds/metabolites are known to metabolic databases or not; it is entirely data driven. MDiA includes all the measured data and REMD-based database

knowledge can be incorporated in into metabolomic MDiN interpretation. The whole workflow of MDEA and MDiA is shown in Figure 3.4.

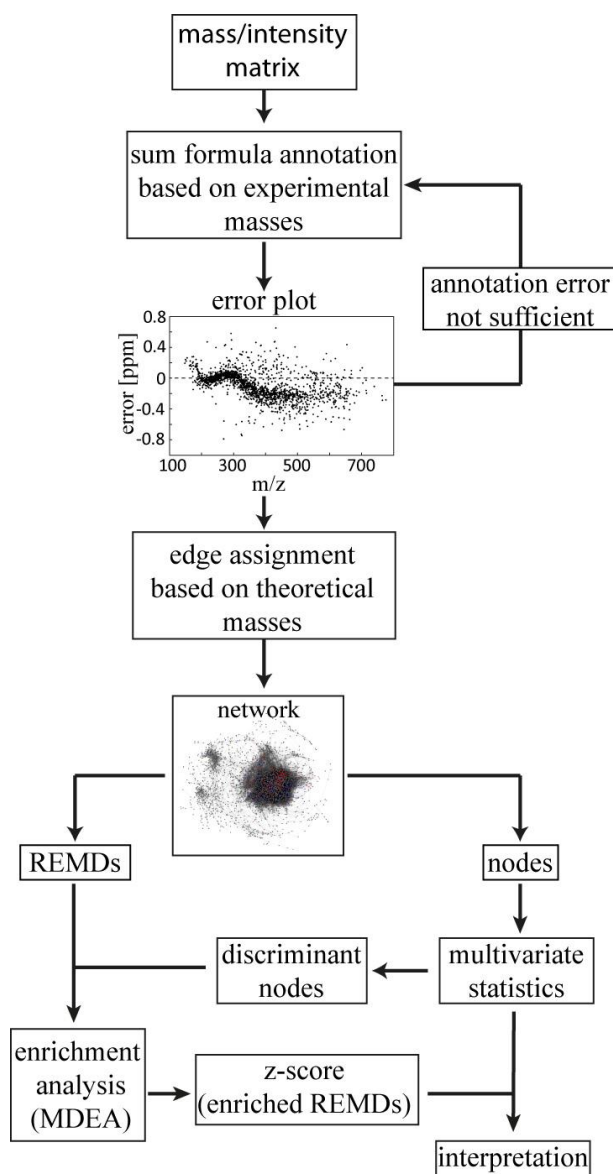


Figure 3.4. Scheme of the mass difference enrichment analysis workflow.

3.3 Results and Discussion

3.3.1 Interpretation of KEGG annotations

For metabolite annotations of sum formulas the widely, maybe most accepted, database KEGG was used, allowing only compounds present on *Populus trichocarpa* pathway maps. A total of 242 KEGG compounds, including isomers, were found. Removing isomers, only 129 KEGG compounds (3% of annotated sum formulas) remained.

Among them, 17 sum formulas (55 KEGG compounds) corresponded to discriminant NE NetCalc-annotations, while 22 sum formulas (47 KEGG compounds) belonged to discriminant IE NetCalc-annotations. As partly shown in chapter 2, the conventional way to describe metabolomes is based on pathway map participation of discriminant metabolites (Kankainen et al., 2011). In this case, however, biological statements that are based on 3% of a metabolome cannot be representative for a phenotypic description.

Among the 17 discriminant sum formulas found in NE genotypes, only two, namely DMAPP/IPP and GPP, were related to terpenoid biosynthesis. Five sum formulas belonged to the common carbohydrates, namely $C_6H_{12}O_6$ for hexoses, $C_{12}H_{22}O_{11}$ for disaccharides, $C_6H_{10}O_7$ for hexuronic acids, $C_6H_{13}O_9P$ for hexose-phosphates and $C_7H_{15}O_{10}P$ for heptose-phosphates. The remaining KEGG annotations for NE consisted of malate, two fatty acids, three flavonoids, three phenolics and one phytosterol.

22 KEGG annotations (47 KEGG compounds) were up regulated in IE. They consisted of seven flavonoids, four phenylpropanoid derivatives, three unsaturated fatty acids (including linoleic acid), four carbohydrate derivatives, two shikimate pathway intermediates as well as one phytosterol and dihydroxyacetone phosphate.

The small sample size of KEGG annotations only allows limited interpretations. The first eye-catching difference concerns the up-regulation of DMAPP as well as of GPP. Since the genetic modification involves the MEP pathway, which is used for the production of DMAPP and GPP, metabolic alterations happening down-stream of ISPS are somehow expected.

The fact that hexose and hexose-phosphate sum formulas are enriched in NE, while hexose-bisphosphate is depleted, indicates feedback inhibition of phosphofruktokinase, which can be inhibited either by high ATP/AMP ratios or by PEP (Kelly and Latzko, 1977; Stitt, 1990). The depletion of dihydroxyacetone phosphate in NE genotypes fits that interpretation. As already mentioned in the Chapters 1 and 2, both PEP and dihydroxyacetone phosphate are the substrates for DXS, which catalyzes the entry reaction of the MEP pathway. Increasing amounts of hexose would indicate a subsequent increase of PEP concentration. However, PEP could not be annotated. The decrease in dihydroxyacetone phosphate concurred with a decrease in pentose-phosphate, which would indicate a limited energy metabolism in NE plants.

Malate was the only KEGG annotation pertaining to the TCA cycle, preventing TCA based interpretations. Information regarding the up-regulation of linoleic acid derivatives in IE is misleading: the linoleic acid pathway has an association to oxidative stress (op den Camp et al., 2003; Moller et al., 2007). However, previous publications convincingly showed higher oxidative stress in NE rather than in IE (Behnke et al., 2010a).

The observed alterations in the flavonoid, phenylpropanoid and phenolic metabolism have already been published (Behnke et al., 2010a; Behnke et al., 2010b), but the reason behind their response to the ISPS knock down cannot be explained.

In summary, the limiting amount of KEGG annotations leads to an inflation of interpretations, where correct interpretations cannot be distinguished from incorrect ones. If the KEGG annotations are visualized and explained by KEGG pathway participations, the picture gets even worse.

For NE poplars, 56 pathway entries, instead of 17, and for IE plants 50 pathway entries instead of 22 are obtained. This is due to the overlap of KEGG pathway maps, which generally leads to double and triple entries. An example is the biosynthesis of amino acids, which in this case is indicated as being of importance even if no sum formula contains nitrogen. Similar interpretations can be drawn looking at the terpenoid pathways. It is known that leaves of NE plants have a much larger plastidic DMAPP pool as IE plants (Ghirardo et al., 2014), yet, only the cytosolic ubiquinone- and terpenoid-quinone biosynthesis is listed among the addressed KEGG maps. Information regarding plastidic terpenoid biosynthesis is not found.

Taken together, both the data interpretation by KEGG sum formula matching and also by KEGG pathway participation results in interpretations, which cannot be backed up with a sufficient amount of data. In the following sections it is shown, how mass difference analysis together with mass-difference enrichment analysis (MDEA) establishes to explain metabolic perturbations even if 97% of mass features are unknown.

3.3.2 Prenylation-REMDs directly link metabolism to genetic modification

For a general overview of the REMD patterns, the two z-scores of each REMD can be visualized in a scatter-plot (Fig. 3.5). In total 40 out of 248 curated REMDs (16%) were over-represented in NE poplars (Fig. 3.5A). Strikingly, three of those describe the mass-differences of mono-, di- and tri-prenylation reactions. In addition the mono-prenylation yielded the highest z-score ($z = 4.54$) of the entire dataset. Similar results were also obtained using the KEGG rpairs in which 37 out of 301 REMDs (12.3%) were enriched in NE poplars (Fig. 3.5B). Seven of those described terpene based reactions, whereby the mono-prenylation KEGG-rpair (e.g. RP01673) yielded the highest z-score ($z=4.08$) of this list. Parts

of the over-represented terpene pairs were also the di- and tri-prenylation. Consequently the pair findings are in agreement with the terpene REMMD findings obtained with the curated REMD list.

These findings represent the major metabolic phenotype as a result of the knockdown of ISPS. As mentioned previously, among the few KEGG annotations DMAPP (substrate of ISPS) and GPP, which are both substrates for di- and tri prenylation reactions, were upregulated in NE. Additionally DMAPP and IPP are the building blocks of GPP, F-PP and other terpenes. More importantly, DMAPP, GPP and F-PP are used for the prenylation of phenolics (Shen et al., 2012), zeatin biosynthesis (Mok et al., 2000) and for the posttranslational modification of proteins (Zhang and Casey, 1996).

¹³C labeling techniques revealed that isoprene emission is the major C-sink of the MEP pathway (Ghirardo et al., 2014). In addition the knockdown of ISPS not only results in the accumulation of its substrate DMAPP in chloroplasts, but also the C-flux towards pigment and chlorophyll biosynthesis increases. However, the increase in C-flux does not explain the complete metabolic fate of DMAPP (Ghirardo et al., 2014).

Interestingly other authors have identified other hemiterpene glucosides (HTG), namely 4-hydroxy-2-methyl-2-buten-1-yl-O-glucopyranoside (4-HMBG, C₁₁H₂₀O₇, Ward et al., 2011) and 2-C-methyl-D-erythritol-O-4-β-D-glucopyranoside (2-MEG, Gonzalez-Cabanelas et al., 2015), in *Arabidopsis*. Both HTG sum formulas were upregulated in NE and had a high connectivity with the prenylation REMDs (Fig. 3C). The HTGs are synthesized via MEP pathway intermediates located up-stream of ISPS and DMAPP/IPP. 4-HMBG was identified by Ward et al., (2011) under nitrogen starvation. The authors showed that 4-HMBG is synthesized via (E)-4-Hydroxy-3-methyl-but-2-enyl pyrophosphate (HMBPP), which is the direct precursor of DMAPP/IPP. Interestingly they argue that 4-HMBG fulfills isoprene like functions under stress conditions in *Arabidopsis* (Ward et al., 2011).

2-MEG was identified in hydroxy-methyl-butenyl-diphosphat-synthase (HDS) deficient *Arabidopsis* mutants (Gonzales-Cabanelas et al 2015). HDS catalyzes the enzymatic step leading to HMBPP. The lack of HDS results in the accumulation of its substrate 2-C-Methyl-D-erythritol-2,4-cyclodiphosphate and 2-C-methyl-D-erythritol. To synthesize 2-MEG, 2-C-methyl-D-erythritol is exported out of the chloroplast and glycosylated in the cytosol (Gonzalez-Cabanelas et al., 2015).

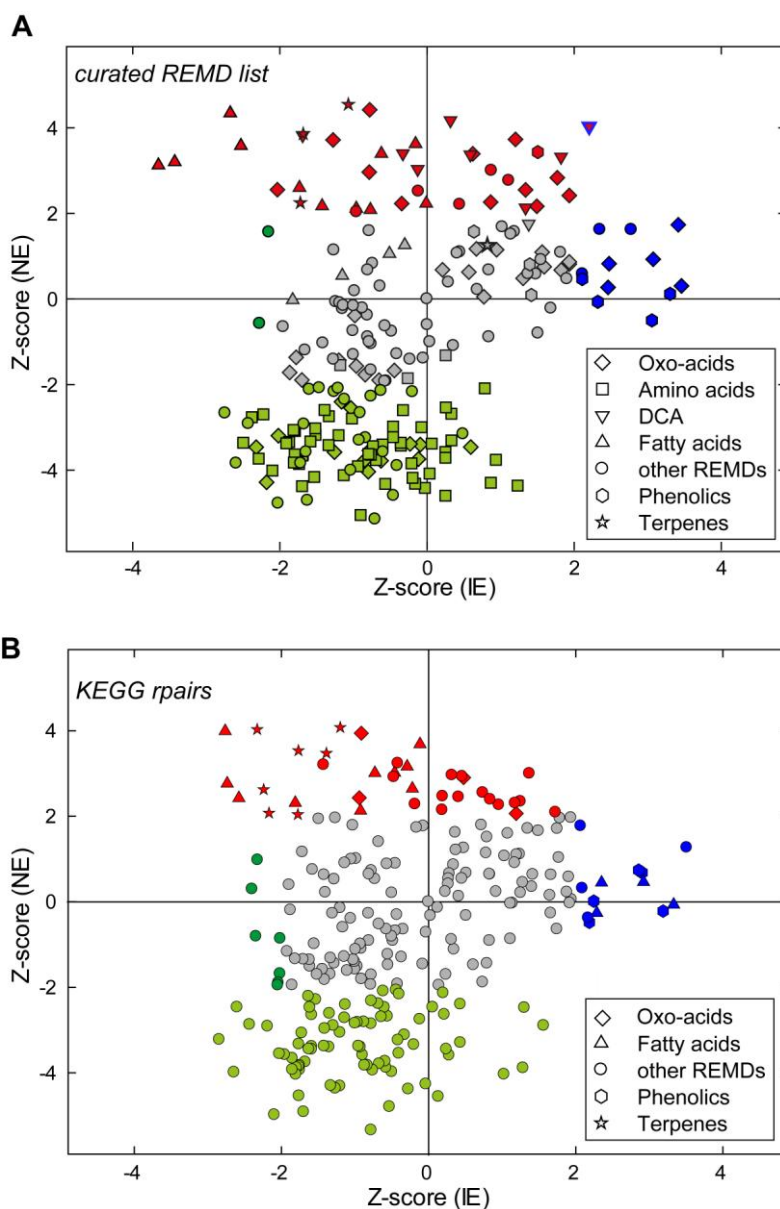


Figure 3.5. Z-score scatter plots of NE and IE plants. (A) Z-score scatter plot of the curated REMD list and (B) z-score scatter plot of KEGG rpairs. Red: enriched in NE genotypes, blue: enriched in IE genotypes, grey: non-significant REMDs, light and dark green: depleted in NE or IE.

These findings are in accordance to Weise et al., (2013), who showed that acid hydrolysis, a commonly used technique for the quantification of DMAPP available for isoprene synthesis, does result in higher amounts than quantified by LC/MS, indicating the presence of yet unknown prenylated compounds that also release isoprene when being hydrolyzed (Weise et al., 2013). Strikingly, we found that hexose and other NE-up-regulated sugars frequently use prenylation REMDs to connect to nodes of unknown identity; reasonably, those sum formulas fit to prenylated glycosides (Fig. 3.5).

The upregulation of terpene REMDs directly links the phenotypic/genotypic background to the metabotype of NE.

3.3.3 Manually curated dicarboxylic acid REMDs explain new biochemical relationships

An REMD-class exclusively present in the curated REMD-list are the dicarboxylic acids (DCA). Interestingly, eight out of nine (89%) DCA-REMDs were enriched in NE poplars (Fig. 3.5A). It's known that DCAs azelaic acid ($z = 2.13$) and pimelic acid ($z = 3.37$) are produced either through lipoxigenases (LOX) or via spontaneous fragmentation of oxidized lipids (Zoeller et al., 2012). Furthermore, azelaic acid is an important signaling molecule, priming immune responses *in planta* (Jung et al., 2009), and pimelic acid is a precursor in biotin biosynthesis (Marquet et al., 2001). A cellular function regarding the remaining DCA is not described. However, by looking at the REMD describing the decarboxylative condensation of adipic acid one immediately finds similarities with prenylation REMDs in that its mass-differences of $\Delta m = C_5H_8O$ equals a prenylation REMD ($\Delta m = C_5H_8$) followed by an O-insertion ($\Delta m = O$).

Further investigation of REMDs enriched in NE revealed similar relationships such as that (i) the condensation of decanoic acid ($z = 4.34$), which describes the transfer of $C_{10}H_{18}O$, can alternatively be described by a di-prenylation ($\Delta m = C_{10}H_{16}$) followed by hydrolysis (H_2O) and (ii) the decarboxylative condensation of ketohexanoic acid ($\Delta m = C_5H_{10}O$, $z = 4.42$) is similar to a two-step reaction involving prenylation and hydrolysis ($\Delta m = C_5H_8 + H_2O$).

In the next step, a subgraph containing these REMDs, which are connected to at least one upregulated NE node, was extracted (Fig. 3.6A). Thereby, the three REMD groups formed triangle motifs, which are important motifs in network theory and are described by the clustering coefficient (Fig. 3.6B-D; Barabási and Oltvai, 2004). The above mentioned 4-HMBG (upregulated in NE), identified by Ward et al., (2011), was part of the prenylation-hydroxylation-adipic acid triangle together with the glucose node (upregulated in NE) and an unknown sum formula $C_{11}H_{20}O_6$ (Fig. 3.6B and 3.3E). Consequently, this triangle motif explains the biological relationship found by Ward et al., (2011) in that the adipic acid REMD mimics the addition of HMBPP to a glucose moiety with a subsequent pyrophosphate cleavage.

The upregulation of the HTG sum formula plus the enrichment of its REMD might be indicative for an alternative route of prenylation to, on the one hand, reduce the amount of plastidic DMAPP and on the other, to compensate the loss of isoprene emission. Those two metabolic adjustments seem reasonable since Ward et al., (2011) argue that the formation of HTG under stress is similar to isoprene emission in trees. In consequence this would indicate the presence of an alternative route of prenylation through the direct precursor of DMAPP, namely HMBPP. Additionally this REMD context explains that the triangle nodes ($C_6H_{12}O_6$, $C_{11}H_{20}O_6$ and $C_{11}H_{20}O_7$, Fig. 3.3B and Fig. 3.3E) all share at least one

characteristic moiety, which in this case is glucose. Thus, triangles might also improve compound identification.

A function regarding the remaining DCAs is not yet described, but it might be hypothesized that their corresponding REMDs are markers for variants of the oxilipin pathway that do not start from linoleic acid but rather from other unsaturated FAs; probably with alternative double bond positions (Zoeller et al., 2012). Furthermore, it might be speculated, whether the DCA-REMDs represent instable metabolic intermediates, similar to pimeloyl-CoA (Streit and Entcheva, 2003), which are hardly detectable with MS techniques. If so, they might actually be used as building blocks by a yet unknown pathway. The DCA REMDs are a perfect example how MDEA can be used in a targeted manner for example if the task is to evaluate the degradation, transfer or the possible targets of xenobiotics.

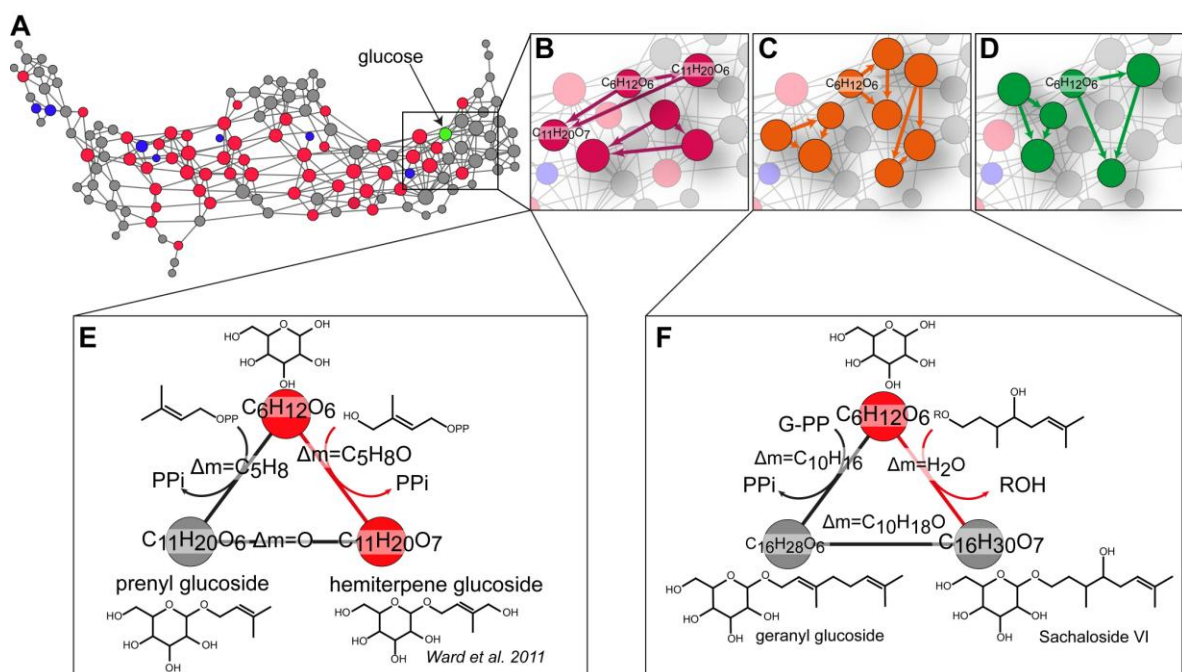


Figure 3.6. Triangle motifs show novel biochemical relationships. (A) Graph laid out using prenylation, geranylation, 2-ketohexanoic acid type b, adipic acid type b, decanoic acid type a, condensation and hydroxylation. Triangle motifs containing: (B) prenylation, hydroxylation and adipic acid (C) decanoic acid, geranylation and condensation, (D) 2-hexanoic acid, prenylation and condensation. (E) Detailed explanation of triangle shown in Fig. 3.6B and, (F) detailed explanation of triangle motif shown in Fig. 3.6C.

3.3.4 Under-representation of nitrogen containing REMDs in non-emitting poplars

MDEA yielded a total of 96 under-represented REMDs in the manual list and 82 under-represented REMDs in the rpair-list. In both lists, only a single REMD, which belonged to the KEGG rpairs, did not contain nitrogen. Since Ward et al., (2011) identified the HTG under

nitrogen limitation, showing the tight regulation of the N to C balance, it can be stated that the under-representation of N-containing REMDs is indicative for a direct involvement with the NE phenotype. In addition, the compositional evaluations already indicated a loss of CHNO, CHNOS and CHNOP compounds in NE.

Within manual list 56 out of the 96 under-represented REMDs described amino acid reactions (56 out of 59 amino acid REMDs). One- and two-step transamination REMDs that are closely related to amino acids were also depleted in NE. The major source for N-metabolism is ammonia, which enters the primary metabolism in the form of carbamoyl-phosphate (Masclaux-Daubresse et al., 2010). The corresponding REMD is strongly underrepresented in NE ($z = -4.71$) and phosphorylation is enriched in NE ($z = 2.51$). It is to be noted, that it is likely that a major proportion of normally available phosphate is not released, as DMAPP is accumulated. It is possible that a consequent cytosolic and mitochondrial lack of P causes impaired N-incorporation.

The other entry point of ammonia into the metabolism is glutamine and glutamate (Bernard and Habash, 2009; Chellamuthu et al., 2014). With a z-score of -5.04, glutamine is the second most underrepresented NE-REMD; directly after formimine transfer ($z = -5.15$), which is THF dependent. The under-representation, however, shows that nitrogen REMDs are not responsible for the observed metabolic changes in NE.

3.3.5 Oxidative stress related REMDs highlight the non-emitting poplars

Given that the knockdown of ISPS is accompanied by an alteration in the fatty acid (FA) biosynthesis typically leading to higher amounts of unsaturated and saturated FA within NE leaves (Way et al., 2013), it sounds reasonable to take a closer look at FA REMDs. The curated list of REMDs included 15 FA reactions, 11 of which (73%) were over-represented in NE, whereas four were depleted in IE (Fig. 3.5). Notably the majority of those FA REMDs were acids with less than ten carbon atoms. Evaluation of the rpairs gave similar results in that ten FA and oxylipin mimicking REMDs were over-represented in NE.

The highest FA z-score of the curated list was obtained by the condensation of decanoic acid ($z(\text{NE}) = 4.34$ and $z(\text{IE}) = -2.67$), which forms together with the di-prenylation and hydrolysis REMDs, a triangle motif (Fig. 3.6C and Fig. 3.6F). This triangle was connected to sum formula annotations fitting to prenyl glycosides, monoterpenes and carbohydrates (Fig. 3.6F). The high z-score of decanoic acid is partly due to its' close similarity with the geranylation REMD. The importance of isoprene for membrane stability and fluidity has been well described in literature, but it also quenches ROS/RNS, protecting plant cells against oxidative stress (Affek and Yakir, 2003). High oxidative stress leads to increased concentrations of malondialdehyde (MDA), a marker for lipid hydroperoxidation (Moore and

Roberts, 1998). MDA is formed through a spontaneous radical reaction, meaning no enzyme activity can be measured (Pryor et al., 1976). Consequently, cellular concentrations of MDA are determined by the thiobarbituric acid reactive substances (TBARS) assay (Hodges et al., 1999). Strikingly, the REMD describing the cellular formation of MDA was over-represented ($z = 2.77$) in NE lines (Fig. 3B). On top of that, MDEA performed with the rpairs yielded the over-representation of two different types of H_2O_2 reactions such as (i) $\Delta m = H_2O_2$ and (ii) $\Delta m = O_2$. The latter REMD would describe the formation of hydroperoxide functionality (maybe resulting from a reaction with a superoxide radical), which, if it is transferred to an unsaturated FA, represents a possible precursor for the spontaneous formation of MDA. The connection between the over-representation of FA hydroperoxidation with the subsequent over-representation of the MDA formation perfectly matches previously published work that showed the accumulation of MDA and of H_2O_2 in NE leaves (Behnke et al., 2010a).

Oxidative stress related REMDs also dominate the Kegg-based MDEA results and therewith confirm both, our “curated REMD”-based interpretations and literature-based knowledge/hypotheses. REMDs that pertained to α -linoleic acid residuals (3/4) and oxilipin metabolism (3/3); one of three MDA reactions and one of two peroxidation reactions were significantly enriched in association to NE nodes.

3.3.6 Over-representation of phosphoenolpyruvate and 2-oxo acid REMDs in non-emitting poplars

Twelve aliphatic or acidic 2-oxo-acid REMDs were enriched in NE (Fig. 3B). The highest (2nd overall) z-score ($z=4.42$) of this REMD group was obtained by the decarboxylative addition of ketohexanoic acid. As stated previously this mass difference ($C_5H_{10}O$) forms a triangle with the prenylation REMD (Fig. 3.6D).

Four 2OA REMDs, namely pyruvic acid, hydroxypyruvic acid and two ketosuccinic acid REMDs, were related to phosphoenolpyruvate (PEP) and pyruvate. Together with the enrichment of the erythrose and transphosphorylation REMDs in NE, a total of six significant REMDs are correlated to reactions involving PEP and pyruvate.

PEP represents an important metabolic intersection, as it is central for carbon allocation (Davies, 1979; Dizengremel et al., 2012). PEP is consumed by four different metabolic pathways, one being the genetically modified MEP pathway (Dizengremel et al., 2012). The silencing of ISPS results in a lower flux through the MEP pathway and therefore might induce the re-navigation of carbon flux towards other pathways through PEP and pyruvate. The mitochondrial TCA cycle is also PEP dependent. With glutarate and 2-ketoglutarate, two important TCA intermediate REMDs, were enriched in NE poplars. Commonly, it is improbable to observe significant changes in TCA metabolite levels even if the TCA carbon flow is altered. TCA cycle intermediates are under strong anapleurotic/homeostatic control

and TCA cycle aberrations can only be observed with labeling experiments (Boyle and Morgan, 2009). Using MDEA, statements on TCA cycle carbon flow can still be made, because the metabolic/anabolic fate of these compounds is elucidated. Furthermore, the enrichment of the condensation and decarboxylative addition of ketoisovaleric acid in NE, which is the biosynthetic precursor of valine and is synthesized via pyruvate, was found.

3.3.7 Over-representation of phenolic REMDs in isoprene-emitting poplars

The PEP and pyruvate involvement stretches even further, as seven out of 13 IE over-represented REMDs from the curated list were either aromatic or characteristic for the shikimate pathway (Fig. 3.5B). The same pattern was also observed in the KEGG rpair list, where also seven out of 13 over-represented REMDs mimicked phenolic and shikimate pathway REMDs. Biosynthesis of aromatic compounds passes through the shikimate pathway, which also requires PEP (Herrmann and Weaver, 1999). Strikingly, earlier metabolomic and transcriptomic experiments on NE poplars showed the diminished production of phenolics in NE compared to IE poplars (Behnke et al., 2010a). The authors observed a down-regulation of transcript levels coding for shikimate pathway enzymes and more importantly of 3-deoxy-7-phosphoheptulonate synthase (DAHP, EC 2.5.1.54), which catalyzes the starting reaction of the shikimate pathway, and of phenylalanine ammonia-lyase (PAL, EC 4.3.1.24). Both of these enzymes catalyze the entry reactions for the biosynthesis of phenolic pathways. In addition, also the nodes pertaining to dehydroquinate and quinate, two intermediates of the shikimate pathway, were upregulated in IE (Fig. 3.7B).

3.3.8 Mass difference analysis as a data-driven approach for the location of metabolic pathways

Its biological interpretability aside, the MDEA results can also be used to select REMDs for network visualization purposes (Fig. 3.7). Here the hydroxylation and hydration REMDs together with the top six over-represented REMDs in NE was used in order to extract a sub-graph enriched with up-regulated NE nodes (Fig. 3.7A). Thereby the hydration and hydroxylation REMDs are used to have a sufficient REMD (edge) backbone to reduce the number of small sub-networks. Typically, the more REMDs are used for the visualization process, the more complex gets the network topology, which might hinder the biological interpretability.

Interestingly the edge based selection of nodes resulted in five sub-graphs explaining the different possible elemental compositions affected by the genetic modification such as CHO, CHOP, CHNO, CHOS and CHNOS (see also Fig. 3.3 and Chapter 3.1.2), whereby, annotations of CHO, CHOS and CHOP exhibited many discriminant nodes upregulated in

NE. Breitling et al. (2006) previously stated that using MDiNs, it is possible to find pathways within the known metabolic realm. The data driven MDiN approach enables to locate a real “pathway motif” representing the shikimate pathway within the extracted CHO sub-graph (Fig. 3G). However, a criterion directly identifying possible metabolic pathways must still be developed.

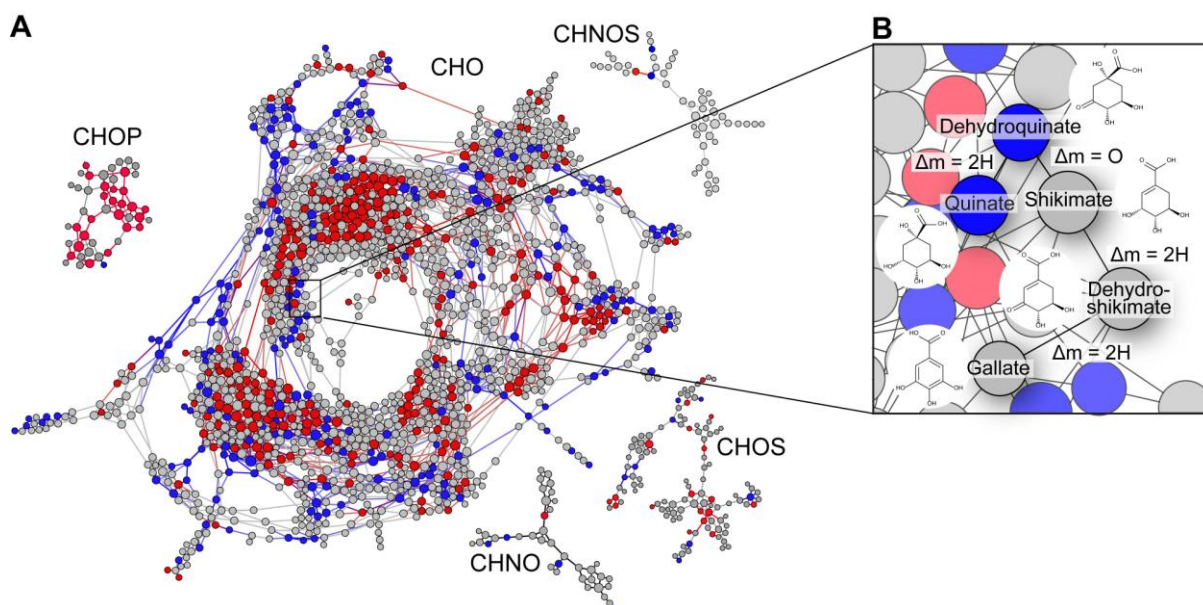


Figure 3.7. Mass difference networks inherit pathway information. (A) Sub-graph laid out using hydration, hydroxylation and the top six over-represented REMDs in NE illustrating the formation of compositional networks, namely CHO, CHOP, CHOS, CHNO and CHNOS; red: upregulated NE nodes, blue: downregulated NE nodes, grey nodes: not significant. (B) Magnified view of the shikimate pathway.

3.3.9 The mass-difference network incidence matrix describes the origin of the unknown compositional space

An incidence matrix is a network representation in which edges are mapped on their respective nodes (Fulkerson and Gross, 1965). Consequently each edge (REMD) is treated as a distinct variable. The aim of using the incidence matrix is to describe how unknown IE and NE nodes are formed from the known KEGG space. To achieve this, all KEGG compound entries pertaining to the organism *Populus trichocarpa* were downloaded via KEGGAPI. In total 854 single KEGG compounds were downloaded. These KEGG nodes were co-networked with the 4,206 nodes of the poplar dataset (POP). Of course the 129 annotated sum formulas of the POP dataset were excluded. Additionally, both REMD lists were combined which resulted in a total of 450 different REMDs.

The co-networking resulted in that 492 KEGG sum formulas were directly connected to 2,316 different POP sum formulas. Commonly KEGG compounds are ordered and classified

according to their location on KEGG metabolic pathway maps. This often results in double and triple entries, as many metabolites are located on several pathway maps. Another problem is that multiple compound classes can be located on a single KEGG pathway map, such as in the phenylpropanoid or flavonoid metabolism where amino acids, phenolic glycosides, phenylpropanoids and flavonoids are found. To improve the interpretability of the incidence matrix results, the KEGG compounds were classified according to the HMDB compound classification (HMDB compound classes). If the KEGG sum formula was not found in the HMDB database, it was classified manually. In total 42 HMDB compound classes were assigned to 486 KEGG compounds.

Now, the edges are replaced by REMDs and the nodes by the HMDB compound classes within the KEGG-POP incidence matrix, whereby non-zero entries represent the total number of connections to the respective compound class with one particular REMD.

The over-represented REMDs and their respective compound classes are visualized in a clustergram (Fig. 3.8). According to the IM clustergram, carbohydrate-containing compounds are incorporated into POP at a very high frequency. More importantly, how the carbohydrates are built into POP varies between IE and NE poplars. Biochemistry in IE leaves preferentially conjugates aromatic moieties with carbohydrates that likely result in (poly)phenolic glycosides and prefers to perform hydroxyl-methylation reactions. In contrast to that, NE prefers to perform farnesylation and geranylation reactions with carbohydrates and *vice versa* (carbohydrates on terpenes). The NE KEGG terpene nodes were frequently converted with saccharides such as pentose condensation, keto-octose transfer, two disaccharide condensations and the condensation of ribose 5-phosphate. Together with the previous findings, which were obtained by MDEA, it substantiates the presence of yet unknown terpene glucosides in poplar and highlights the fact that NE poplars have a much more pronounced terpene metabolism.

Additionally, the incidence matrix also shows the impact on fatty- and dicarboxylic acids in that it visualizes their structural relationship via a clustering of FA KEGG nodes and the NE REMDs succinate transfer to aldehydes and pimelate condensation. Furthermore, REMDs pertaining to fatty acids such as decanoic acid condensation, hexanoic acid condensation, 13-HPOT-cleavage, dodecanoic acid condensation and the decarboxylative condensation of 3PG were NE specific. Substrates for these REMDs were either carbohydrate containing compounds or phosphorus containing compounds. Strikingly, the initial sum formula annotations already indicated differences in phosphorous-containing compounds. The incidence matrix confirmed this observation and additionally revealed different usage patterns. Despite FA condensations, NE plants preferably use geranylation on phosphorus compounds.

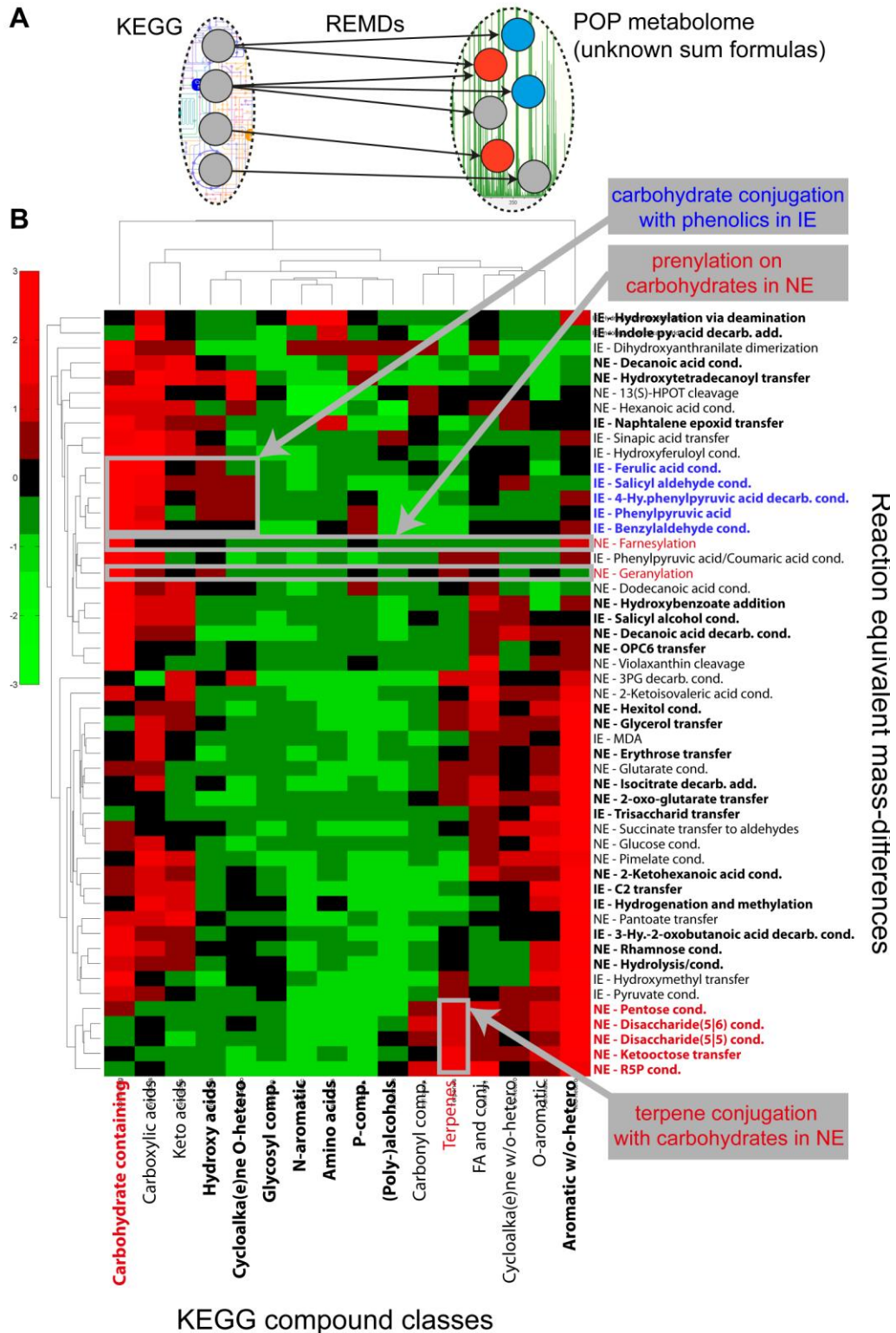


Figure 3.8. Clustermap of the incidence matrix. (A) Exemplifies the formation of discriminant unknown sum formulas of NE and IE from known KEGG metabolites. (B) The clustermap, in which the most important findings are highlighted (red font color: significant for NE; blue font color: significant for IE)

On the contrary, IE plants preferably conjugated aromatic metabolites. Additionally, IE metabolites are significantly generated by modification of KEGG metabolites with

phenylpropanoids. 10 of 21 enriched REMDs were associated to the shikimate pathway. Besides using phenylpropanoids, IE plants used N-aromatics and amino acids more frequently, which corresponds to the MDEA results that were mined in POP.

The introduced methodology facilitates the definition of targeted strategies to investigate very specific aspects of a largely unknown metabolism. Ultimately, it allows the mining of candidate enzymes by querying databases for the combined information of source compound class and acting REMD.

3.3.10 Network modules contain compound class information

The previous sections convincingly showed how MDEA and MDiNs improve data interpretability when almost all mass-features are unknown. If the annotations of Chapter 2 are matched to the annotated sum formulas, the database coverage is increased to 22.2%. Now, the aim is to elucidate whether graph theory can be used to improve compound identification. Therefore, eight REMDs were chosen for the sub-graph extraction (two backbone, two over-represented and four under-represented REMDs). This resulted in 12 sub-networks (Fig. 3.9A and Fig. 3.9B) that were laid-out with the open-source software gephi (<http://gephi.github.io>). Gephi allows calculating network parameters such as the network modularity, which describes the network structure (Barabási and Oltvai, 2004). Modularity searches for sub-clusters in any given network. In these sub-clusters nodes of the same module have a high connectivity, but have sparse connections with nodes of other modules.

Calculating the network modularity revealed a total of 19 in the main network (Fig. 3.9B). Now each module can be visualized in a pie plot, which shows the compound class annotations found in the respective module (Fig. 3.9C). This revealed that many modules consisted of only up to three compound classes. Hence there is a high probability that unknown nodes of the same module belong to the same compound class. For example the nodes of the shikimate pathway were located in the same module (Fig. 3.9D).

If an unknown node of interest would be present in one of these modules, it can already be assigned to a compound class. Furthermore, this information would help developing a targeted LC-MS method with the aim of structural identification.

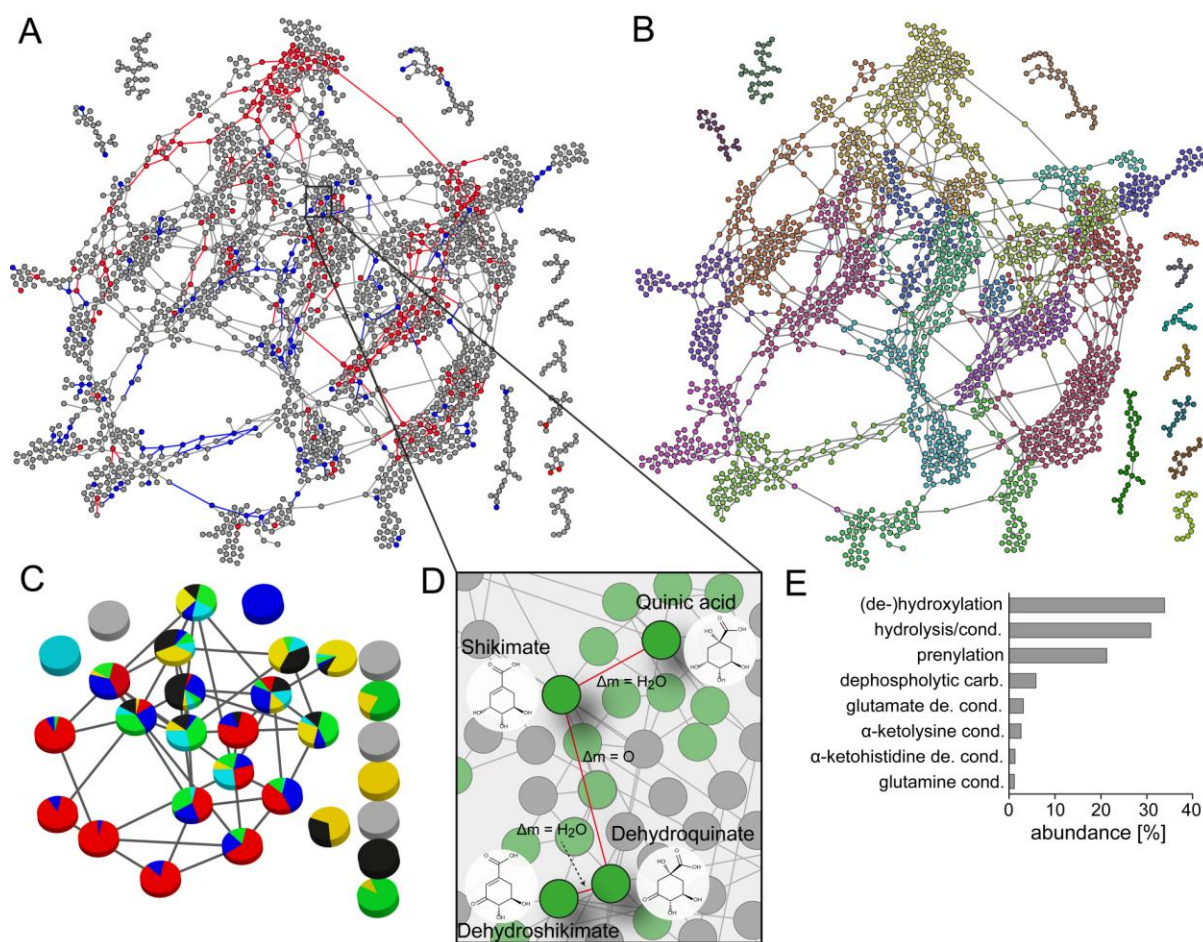


Figure 3.9. Modules of mass difference networks contain compound information. (A) Extracted sub-graph obtained by using REMDs shown in Fig. 3.9E; red: upregulated NE nodes, blue: downregulated NE nodes. (B) The same graph colored according to the different modularity classes. (C) Pie plot representation of compound classes present in each modularity class; red, flavonoids; blue, phenolics; green, terpenes; yellow, lipids; teal, carbohydrates; black, other primary metabolites; grey, unknown. (D) Shikimate pathway found in Fig. 3.9A. (E) Relative abundance of each REMD present in the sub-graph.

3.3.11 Tandem mass spectrometric measurements validate MDEA findings in the non-isoprene emitting poplars

As described previously, the main focus of this technique is to demonstrate how the entirety of acquired m/z data, be it of high or low ion abundance, can be integrated into biochemical contextualization. As single feature identification is a tedious work and in many cases not possible, multiple adjacent selected ion monitoring (SIM) MS/MS (Southam et al., 2007) on leaflets of both, the NE and IE genotypes, was performed.

Markers of interest suffer from less penalization using the SIM-MS/MS approach as compared to smaller SIM windows conventionally used for MS/MS. The larger SIM window allows for marker features of low abundance to contribute to the daughter-ion space. The mass range of the network shown in Figure 3.6 ranging from 245 to 455 Da was chosen for

analyses. MDEA was used to mine neutral losses that were significantly associated to the NE and IE markers as compared to all parent-daughter ion pairs of a given SIM window.

The MDEA results of the SIM-MS/MS experiments yielded an overrepresentation of 49 REMDs in the NE genotype using the curated REMD-list (Fig. 3.10A). The correlation coefficient between all full-scan and SIM-MS/MS REMD z-scores from the NE samples was 0.76 (Fig. 3.10E).

25 REMDs were over-represented in both, the full scan and the SIM-MS/MS results (Fig. 3.10A). Strikingly, the three REMDs, which were part of the triangle motifs in Figure 3.6, namely the prenylation, the decarboxylative condensation of adipic acid and of ketohexanoic acid, were also over-represented in the SIM-MS/MS data (Fig. 3.10B). This observation directly validates the full-scan MDEA data in which it was shown that MDEA establishes a direct linkage between the metabotype and the genotype. Additionally it substantiates the presence of unknown prenylated compounds in NE poplars (Fig. 3.6).

Close similarities between conventional full-scan and SIM-MS/MS results were observed in the DCA REMDs, where seven out of the eight over-represented full-scan REMDs were also associated to the NE poplars in the SIM-MS/MS data (Fig. 3.10B).

Additionally, 17 2-oxo acid REMDs, of which 11 described the cleavage of aliphatic 2-oxo acids, were over-represented in the SIM-MS/MS spectra of NE plant extracts. This is in agreement with the observation that ten out of those 17 2-oxo acid REMDs were also over-represented in the full-scan REMD results of this genotype (Fig. 3.10B).

Interestingly, the REMDs that only were over-represented in the SIM-MS/MS results pertained to four two carbon 2-oxo acid building blocks (pyruvate-related), two hydroxyphenylpyruvate REMDs and one ketoglutarate REMD. Six FA related REMDs were enriched in the NE SIM-MS/MS experiments. Four of them overlapped with REMDs that were enriched in the full-scan MDEA results of the NE genotype (Fig. 3.10B). Two of those REMDs describe butanoic acid reactions (type A and B), one the hexanoic type A reaction and one dodecanoic acid type A reaction (Fig. 3.10B).

Another finding confirming the results obtained by the incidence matrix is the overrepresentation of six carbohydrate REMDs in the SIM-MS/SM NE poplar data. The results of the incidence matrix show that those REMDs were used for terpene modifications to form unknown terpene glycosides in NE poplar (Fig. 3.10B). In plants, many secondary metabolites, especially phenolic metabolites, are glycosylated. Poplar trees are known for their high production of phenolics, such as flavonoids, phenolic glycosides and phenylpropanoids (Babst et al., 2010; Boeckler et al., 2011). These compounds differ drastically in their glycosylation patterns and because of that neutral losses of carbohydrates are often observed in MS/MS experiments of plant extracts (Kachlicki et al., 2008; Abreu et

al., 2011). Overall, a high correlation between the z-score profiles of the NE incidence matrix and of the NE SIM-MS/MS was found (Fig. 3.10E).

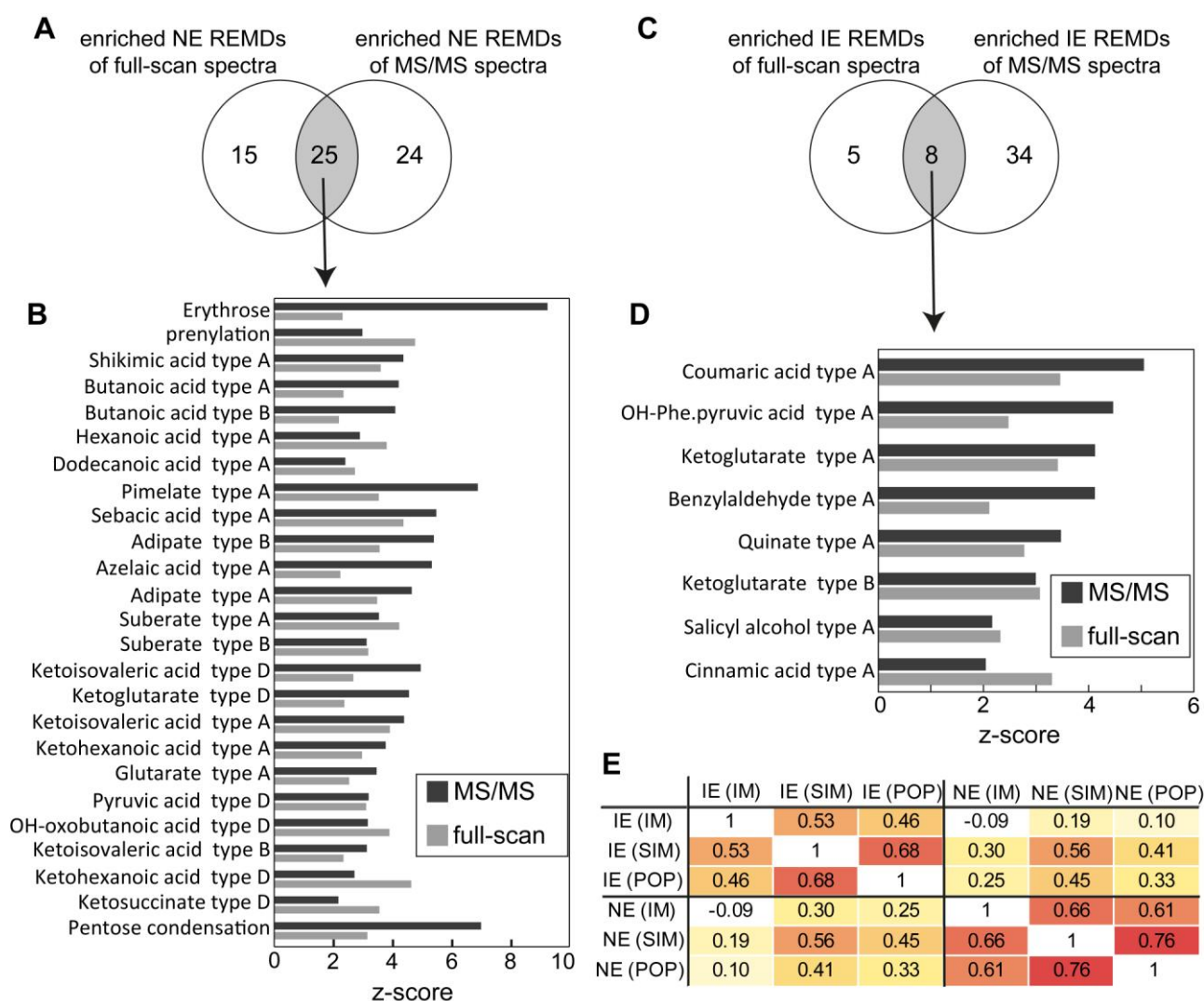


Figure 3.10. Comparison of SIM-MS/MS MDEA with full-scan MDEA results. (A) Venn-plot illustration of enriched NE full scan and SIM-MS/MS REMDs. (B) The overlapping NE REMDs are illustrated in z-score bar charts. (C) Venn-plot illustration of enriched IE full scan and SIM-MS/MS REMDs. (D) The overlapping IE REMDs are illustrated in z-score bar-charts. (E) Heatmap showing the correlation coefficients of REMD z-score vectors obtained by full-scan MDEA (POP), incidence matrix MDEA (IM) and SIM-MS/MS MDEA (SIM).

3.3.1 MS/MS experiments validate and extend the conjugation of phenolics in the IE genotype

42 REMDs were over-represented in the SIM-MS/MS data set of IE poplars. The correlation coefficient between the REMD z-scores of full-scan and SIM-MS/MS MDEA analysis in the IE poplars was 0.68 (Fig. 3.10E). Eight REMDs yielded significant z-scores in both measurement types (Fig. 3.10C). Five of those REMDs described phenolic reactions (including two aromatic 2-oxo-acids), one characterizes the condensation of quinate and the

remaining two described ketoglutarate reactions (Fig. 3.10D). These results do not only validate the full-scan REMD results (Fig. 3.6), but they also complement the known literature showing higher contents of phenolics in leaves of IE compared to leaves of the NE genotypes (Behnke et al., 2010a; Way et al., 2013; Kaling et al., 2015).

Additionally, four phenolic REMDs and three aromatic 2-oxo acid REMDs were over-represented in the SIM-MS/MS results, thus further complementing the full-scan MDEA results in IE poplars.

4 Investigation of the metabolic response of non-isoprene and isoprene emitting poplars to the poplar leaf beetle *C. populi*

Parts of this chapter were published in:

Müller A.*, Kaling M.*, Faubert P., Gort G., Smid Hans M., Van Loon Joop J.A., Dicke M., Kanawati B., Schmitt-Kopplin P., Polle A., Schnitzler J.P., Rosenkranz M. (2015) Isoprene emission in poplar is not important for the feeding behaviour of poplar leaf beetles. *BMC Plant Biology*, 15:165.

*Both authors contributed equally to this work.

4.1 Background

Non-isoprene emitting poplars showed a visual distinguishable leaf-color phenotype resulting from variations in the biosynthesis of phenolic compounds (see above). These compounds do also fulfill important functions in the protection against specialized leaf herbivores. Because of that it is reasonable to address the question whether typical poplar herbivores, such as *C. populi*, prefer to consume either NE or IE poplars under close to natural conditions.

Electroantennographic experiments revealed that *C. populi* detects higher terpenes but is unable to sense isoprene (Müller et al., 2015). Additionally, GC-MS analysis of VOC emission profiles showed line- and herbivore-dependent differences (Müller et al., 2015), which have no significant effect on the preference of *C. populi* for either IE or NE poplars under greenhouse conditions (Müller et al., 2015). Since no-differences were found using bioassays, the remaining question is: How does *C. populi* behave under field conditions, when they are given the free choice between IE and NE plants? Therefore, a two-week field experiment was performed in free-air cages located in Göttingen was performed. The biological data from the 14-day outdoor experiment revealed that *C. populi* fed significantly more on IE than on NE poplar. Additionally, the beetles exclusively fed on young poplar leaves (for more details see Müller et al., 2015).

To clarify whether the changes in poplar leaf beetle behavior are caused by dissimilarities in the metabolic composition of NE and IE as well as between young and old poplar leaves, non-targeted metabolomics with the ICR-FT/MS was performed.

4.2 Experimental Setup

Non-isoprene emitting (NE) poplars (genotypes RA1 & RA22), isoprene emitting (IE) wild type and IE transgenic (β -glucuronidase (GUS)/green fluorescent (GFP) vector controls) were used (for more details on the RA and GUS(GFP) lines see Behnke et al., 2007). Trees used for the field bioassay experiment had reached the 20 leaves stem-stage. *Chrysomela populi* were collected in poplar plantations in Southern Germany near Freising, Scheyern and Sigmaringen.

Fourteen-day bioassays under close-to-natural conditions were conducted in eight small cages covered with mesh screen, which were located at Göttingen in a caged area with permission to work with transgenic plants. Three weeks before the beetles were placed into the treatment cages, four grey poplar trees per genotype (RA1, RA22, GUS and WT) were put into each cage (16 plants per cage).

At day zero (2 pm), 60 *C. populi* adult individuals were released in the middle of each of the four cages. Leaves were sampled and immediately frozen in liquid nitrogen on days 4, 8 and 14 (at 2 pm). At each time point one “young” leaf (number 5 from apex) and one “old” leaf (leaf between leaf number 10 and 15 from apex) was sampled. Leaf samples of each genotype in one cage, with the upper and lower leaves separated, were pooled for ICR-FT/MS measurements.

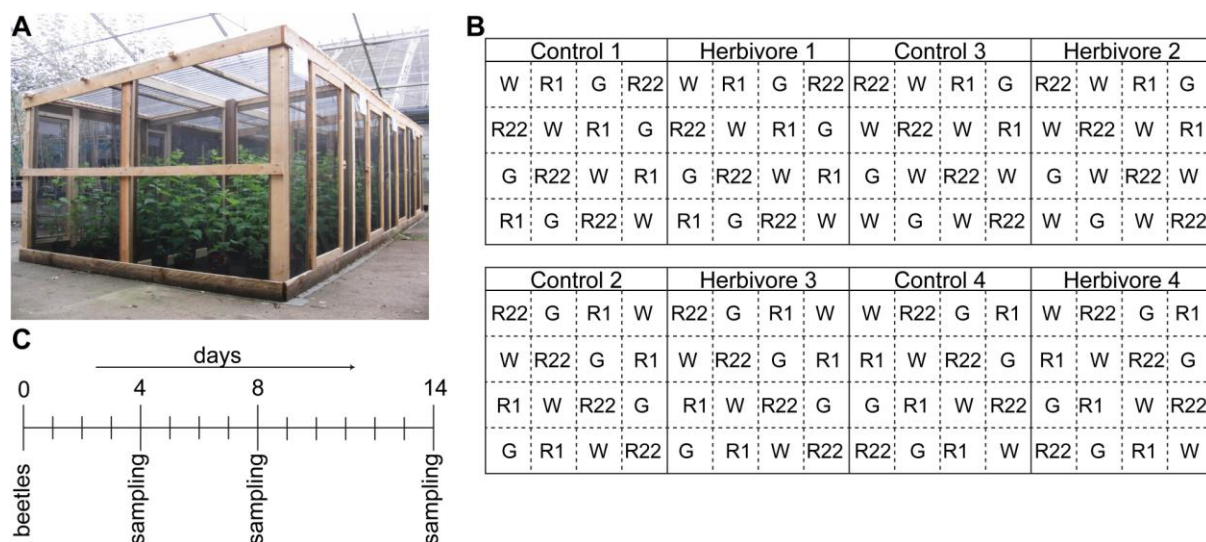


Figure 4.1. Experimental setup of the *C. populi* bioassay under field conditions. (A) Photo of free-air cages in Göttingen, (B) the arrangement of IE (Wild type: W; GUS: G) and NE (RA1: R1; RA22; R22) poplars in each cage and, (C) experimental timeline highlighting the sampling days for ICR-FT/MS measurements (modified according to Müller et al., 2015).

4.3 Results and Discussion

4.3.1 High leaf development and herbivore dependent metabolic variability

Unsupervised evaluation of the ICR-FT/MS data (size of intensity matrix: 24x8923) via PCA revealed a distinct clustering of young and old leaves in the PC1xPC2 score plot. The PCA also details a time-dependent increase in metabolic variation between young and old leaves because leaf samples taken at day 8 and 14 were even stronger separated in PC1 than samples taken at day 4 (Fig. 4.2A). This substantial metabolic difference between young and old leaves might be one reason why the beetles exclusively fed on young poplar leaves.

Analyzing the PC2xPC3 score plot showed the clear presence of time-dependent metabolic alterations caused by the herbivore treatment (Fig. 4.2B). Similar to the time trend found in PC1xPC2, infested leaves, which were harvested at day 14, formed a distinct cluster in the PCA model. Unfortunately, line-dependent differences were not explained by the PCA model. The reason for that might be that those metabolic adjustments are very small compared to dominant changes in the metabolic composition caused by leaf-ageing and herbivory. Nevertheless, the PCA model described leaf-age, beetle feeding and sampling time and explained 58% of the total data variance.

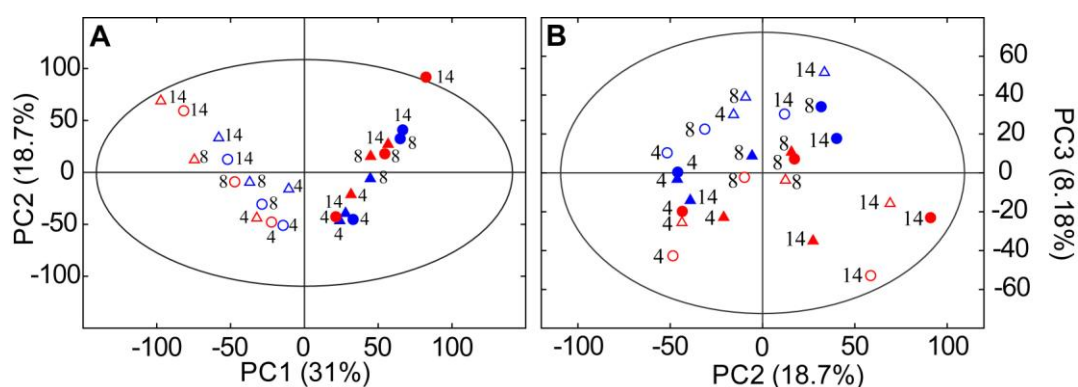


Figure 4.2. Principal Component Analysis of ICR-FT/MS data. (A) PC1xPC2 score scatter plot showing the separation of young and old leaves and (B) PC2xPC3 score scatter plot visualizes the distinct clustering of fed and unfed plants (modified according to Müller et al., 2015). Blue symbols, control plants; red symbols, infested plants; triangles, NE poplars; circles, IE poplars; filled symbols, young leaves; no-fill symbols, old leaves.

4.3.2 Identification and extraction of discriminant mass-features

To extract masses responsible for the observed PCA separation and to clarify whether NE and IE poplars respond differently to herbivore attack, separate OPLS-DA models for IE and NE poplars were calculated, in which leaf-age and herbivory were subjected as Y-variables (e.g. for leaf development: $Y(\text{young}) = 1$, $Y(\text{old}) = 0$). Masses with a variable influence of projection score (VIP) > 1 and a cross-validation standard error smaller than the VIP value (cvSE) were extracted and further tested with a Wilcoxon-Mann-Whitney-U Test. Masses with a VIP score >1 and a P-value <0.05 were considered as discriminant.

In IE poplars 2076 (23% of detected mass features) and 2333 (26%) masses were discriminant for leaf-age at control and herbivore conditions, respectively. A similar pattern was found in NE poplars were 2274 masses were discriminant for leaf development at control and 2505 at herbivore conditions. Consequently, 23%-28% of detected mass-features are significantly changed between young and old leaves.

1211 metabolites were found to be discriminant for herbivory in young IE leaves and 328 in old IE leaves. Even though the beetles almost exclusively fed on young leaves, metabolic changes in the systemic old leaves are expected and are caused by systemic signaling networks between leaves in order to prime the defensive systems of the undamaged adjacent leaves (Jung et al., 2009). In NE poplars, 1544 masses were discriminant for herbivory in young and 1283 in old leaves. Interestingly, the systemic response to herbivory is more pronounced in NE than in IE poplars.

4.3.3 Annotation and comparison of discriminant mass-features

To unravel the metabolic differences between NE and IE leaves, the discriminant masses were compared in Venn-Plots (Fig. 4.3). In both experimental variables (leaf development and herbivory) many masses did not overlap between NE and IE, which indicate the presence of genotype-specific processes in leaf ageing and herbivore attack.

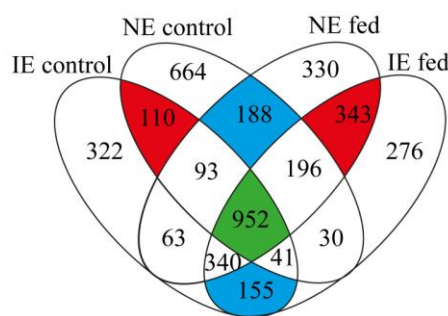
Despite the treatment specific metabolic responses between young and old leaves, a total number of 952 masses overlapped in all four scenarios (Fig. 4.3A). These masses represent the fundamental differences in the molecular composition that exists between old and young leaves. Since these compositional differences might be one of the reasons why *C. populi* avoids feeding on old leaves, the 952 masses were uploaded for metabolite annotation to the massrix and metlin server. In total 118 masses (12.4%) could be annotated as metabolites and assigned to compound classes (Fig. 4.3B).

It was found that young and old leaves differed in their composition of phenolics, lipids, carbohydrates, steroids and terpenes (Fig. 4.3B). Because of their high biological activity especially, flavonoids and phenolic glycosides are very important compounds in plant-insect

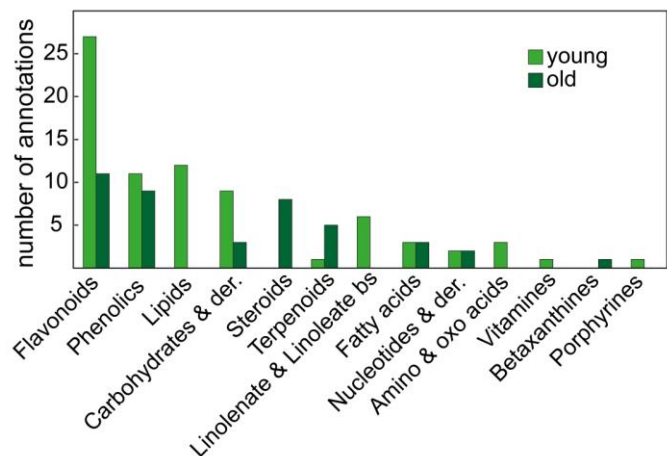
interactions (Simmonds, 2003; Treutter, 2005; Lattanzio et al., 2006). By binding to plant proteins in the gut of leaf consuming herbivores, phenolics complicate and even prevent the digestion (Boeckler et al., 2011, Philippe and Bohlmann 2007).

Young leaves exhibited higher amounts of carbohydrates and lipids, especially metabolites of the linoleic acid metabolism. Since young leaves are not fully developed, the biosynthesis of lipids and carbohydrates should run in full gear in order to meet on the one hand the high energy requirements and on the other hand provide enough building blocks for the assembly of biomembranes that are needed for leaf growth.

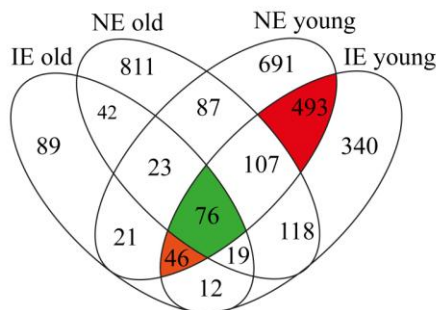
A leaf-age effect



B



C beetle effect



D linolenic acid metabolism

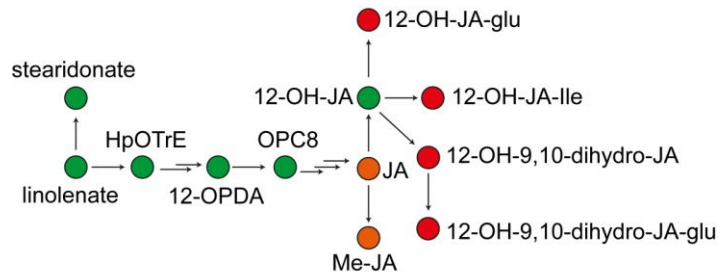


Figure 4.3. Comparison of discriminant masses. (A) Venn-plot comparing the leaf-age discriminant masses at the different experimental variables (herbivory and line). (B) Annotations of masses that are discriminant for leaf-age and are independent of other experimental variables (herbivory and line effect). (C) Venn-plot comparing the masses discriminant for herbivory. (D) Detailed view of the α -linolenic acid metabolism, colors indicate their location in the Venn-diagram shown in C (modified according to Müller et al., 2015).

Old leaves had higher amounts of steroids (phytosterols) and terpenes such as gibberellins. Additionally, higher amounts of the MEP intermediate 2-C-Methyl-D-erythritol 4-phosphate were found in old leaves. Emission of isoprene starts at a certain environmental stage because very young leaves do not express ISPS. Moreover there is a direct competition between the biosynthesis of isoprene and the synthesis of photosynthetic

pigments as well as of cytokines (Rasulov et al., 2014). In other words, while the biosynthesis of terpenes in young leaves did just start, old leaves already produce isoprene.

The next question was whether herbivory caused different metabolic adjustments in IE versus NE poplars (Fig. 4.3C), which might explain the herbivore preference towards IE. Once again, genotype dependent metabolic deviations in local (young) and adjacent undamaged (old) leaves were found (Fig. 4.3C). A more pronounced response in local and adjacent undamaged leaves in NE (811 adjacent undamaged and 691 local masses) than in the IE genotypes (89 adjacent undamaged and 340 local masses) was found.

But also a large portion of discriminant masses overlapped, which points out a conserved metabolic response to herbivory in local (493 masses) as well as in adjacent undamaged (42 masses) leaves.

Additionally, 76 masses were discriminant for herbivory in both leaf types for IE and NE (Fig. 4.3C). Database annotations revealed that the majority of those metabolites were fatty acids and phospholipids. One of the main pathways induced upon herbivore attack, is the linolenic acid metabolism, which generates important signaling molecules, better known as jasmonates (Wasternack, 2007). These are involved in leaf priming and in the synthesis of green leaf volatiles (Wasternack and Parthier, 1997; Weber, 2002; Matsui, 2006). For that reason a more detailed look upon the changes of metabolites derived by the α -linolenic acid pathway (also known as: octadecanoic acid pathway) was taken (Fig. 4.3D). Many intermediates of this pathway were found to be upregulated under herbivory in both leaf developmental stages as well as in both genotypes (Fig. 4.3D), whereas glycosylation, hydroxylation and amino acid modifications of jasmonic acid were induced exclusively in young leaves (Fig. 4.3D). Old leaves exhibited higher amounts of jasmonate and methyl-jasmonate. Arachidonic acid and azelaic acid are also signaling molecules that are important in orchestrating stress-induced signaling networks (Jung et al., 2009; Savchenko et al., 2010). Both were differentially regulated between young and old leaves.

4.3.4 The metabolic response of poplars to *C. populi*

The 2956 masses discriminant for herbivory (Fig. 4.3C) were subjected to hierarchical clustering analysis (HCA) using the Pearson correlation coefficient (Fig. 4.4A). Thereby, young leaves clustered separately from old leaves and also sub-clustered in control and infested leaves (Fig. 4.4A). In total 361 (12%) discriminant masses were annotated as metabolites using databases. To further elucidate the poplar response to herbivory, the annotated metabolites were examined in more detail.

It was found that 20 and 147 metabolites appeared after herbivory in old and young IE leaves (see Venn-Plot Fig. 4.4B), respectively, whereas 121 and 103 metabolites were found

in old and young NE leaves (see Venn-Plot Fig. 4.4C), respectively. In total, 40 metabolites changed in both young and old IE leaves and 45 in young and old NE leaves (Fig. 4.4B-C).

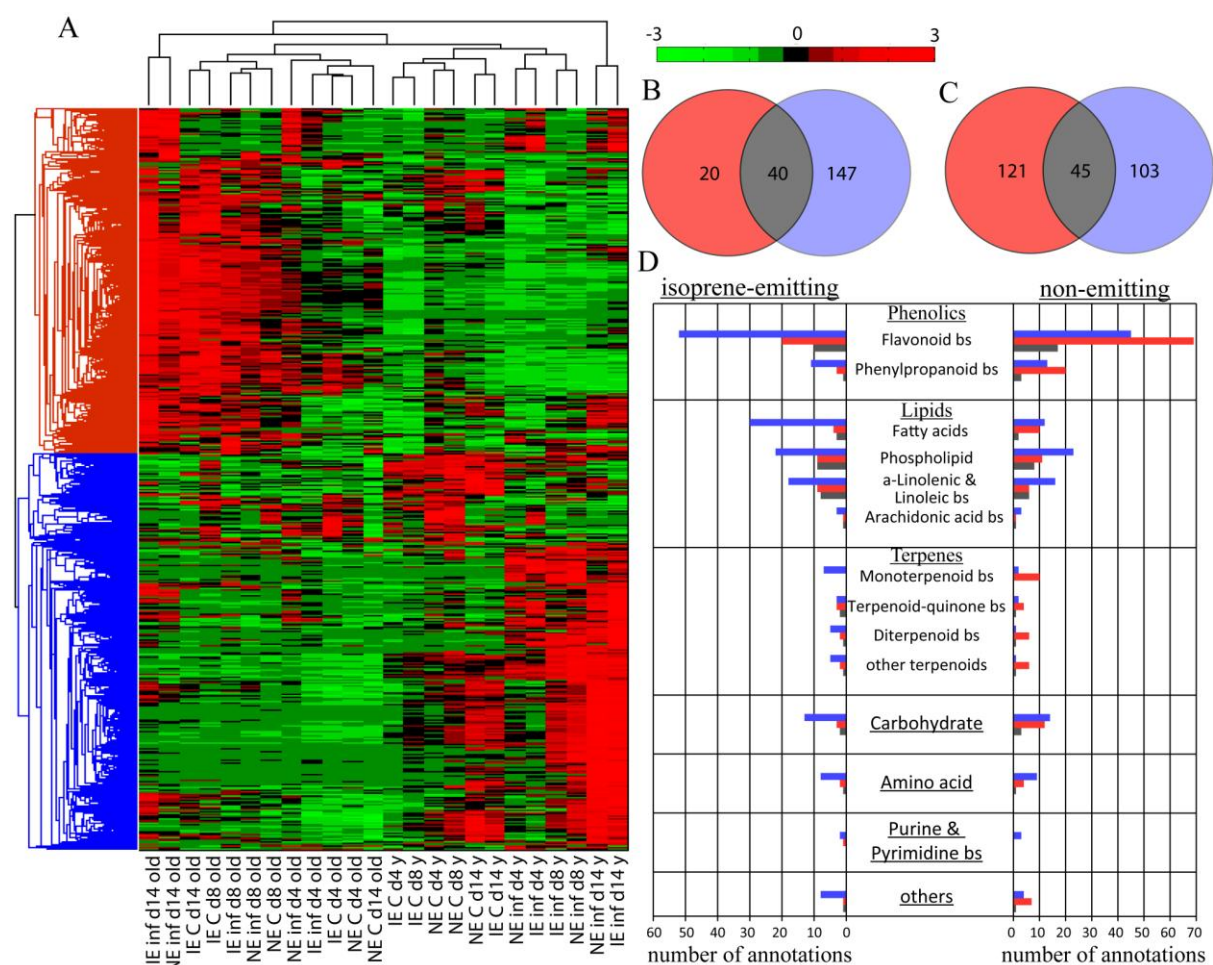


Figure 4.4. Clustermap and database annotations of ICR-FT/MS intensity profiles of discriminant masses. (A) Clustermap of discriminant masses. Venn plots of annotated discriminant metabolites of infested young (red color) and old leaves (blue color) in (B) IE and, (C) NE plants. All database annotations are shown in bar charts ordered according to compound classes and metabolic pathways (red: young leaves, blue: old leaves, grey: discriminant for both leaf types; inf: infested; d4: day 4; d8: day 8; d14: day 14; bs: biosynthesis; C: control; y: young, reprinted from Müller et al., 2015; publisher: BioMed Central).

The metabolites were grouped according to their respective compound classes and biosynthetic routes (Fig. 4.4D). This revealed an influence of herbivory in both leaf age classes over various pathways. The results show the differential regulation of the phenolics metabolism: in NE poplars, more phenylpropanoids were found in old leaves than in young leaves, whereas the opposite was observed for the IE poplars. Flavonoids, on the other hand, were upregulated in both leaf ages in infested plants. Furthermore, 3-dehydroquinate and shikimate, which are important precursors of plant phenolics, were upregulated in infested young old leaves and showed no change in adjacent undamaged leaves.

As already mentioned in the previous section, especially the changes in biosynthesis of phenolics, may alter the leaf attractiveness for herbivores and could be responsible for the slight infestation preference of *C. populi* towards IE poplars. For example it was shown that alterations in the leaves phenolic composition resulting from UV-B exposure in *Populus trichocarpa* influenced the leaf consumption from *Chrysomela scripta* (Warren et al., 2002). Furthermore, Boeckler et al., (2014) showed that poplars genetically modified in the tannin pathway (enhanced tannin content) increased the palatability of leaves for herbivores induced by the subsequent down-regulation of phenolic glycosides.

Moreover, infested young leaves showed a drastic increase in phospholipids and in metabolites of the α -linolenic, linoleic and arachidonic acid biosynthetic pathways. A few of these metabolites, particularly those of the α -linolenic acid biosynthetic pathways, were also found to be discriminant in old leaves. Several of the identified fatty acids are directly involved in plant signaling (Fig. 4.3D). Alterations in the levels of important signaling molecules maybe one reason for the different phenolic profiles observed between IE and NE poplars, therefore, influencing the defense chemistry of NE plants. Additionally, NE poplars need to replace the function of isoprene with other metabolites to maintain the efficiency of the photosynthetic apparatus.

However, it has to be noted that compared to the metabolomic differences between the infested young and adjacent undamaged old leaves, the genotype-dependent differences were minor. But in this experimental setup 62% of detected mass-features were discriminant for at least one of the four experimental variables (time, line, herbivory or leaf development).

In addition, NE poplars seem to behave differently under natural conditions than under controlled conditions that were studied in Chapter 2 and 3.

5 Impact of root colonization on the leaf metabolome of poplars and the behavior of *C. populi*

5.1 Background

Up to now only negative environmental factors (UV-B radiation and herbivory) were studied. In nature there are, however, many positive interactions that can enhance the overall performance of plants. One of those interactions is the symbiosis between plant roots and ectomycorrhizal fungi (EMF), which enhances nutrient uptake as well as stress tolerance. Such beneficial intercommunications can affect above ground plant-insect interactions. To study the reasons for these behavioral changes, free-air experiments with wildtype IE poplars inoculated with the EMF *Laccaria bicolor* were performed in Göttingen using the same experimental equipment as described in Chapter 4. Non-targeted metabolomics with the UPLC-qToF-MS and the ICR-FT/MS was performed of leaf samples that were harvested during the experiment. Thereby the main task was to create a workflow for the mapping of RNA sequencing data (analyzed by Anna Müller; Müller, 2014) on the REMDs of the MDiNs (see Chapter 3).

5.2 Experimental setup

Sixty-four *Populus x canescens* plants were inoculated with the EMF *Laccaria bicolor* according to Müller et al. (2013). As control plants, 64 non-inoculated plants of the same developmental stage were used. Both plant types (inoculated and non-inoculated poplars) were grown free-air conditions until three weeks before the start of the experiment.

Then the plants were transferred into the cages (see Fig. 5.1A) such as that each inoculated plant was surrounded by four non-inoculated poplars and *vice versa* (Fig. 5.1B). Sixteen plants were placed in each cage, half of which were inoculated with EMF (Fig. 5.1B). After acclimation (ten weeks after the inoculation), in four cages a total of 320 *C. populi* beetles (80 per cage) were released. The remaining four cages were used as a control. The duration of the experiment was set for 14 days. For metabolomic analysis, leaf number five from the apex was harvested after eight days of beetle exposure. Leaf samples of the four mycorrhized with beetle (MB), non-mycorrhized with beetle (NB), mycorrhized control (MC) or non-mycorrhized control (NC) poplars that were present in each cage were pooled (for more detail see Müller, 2014).

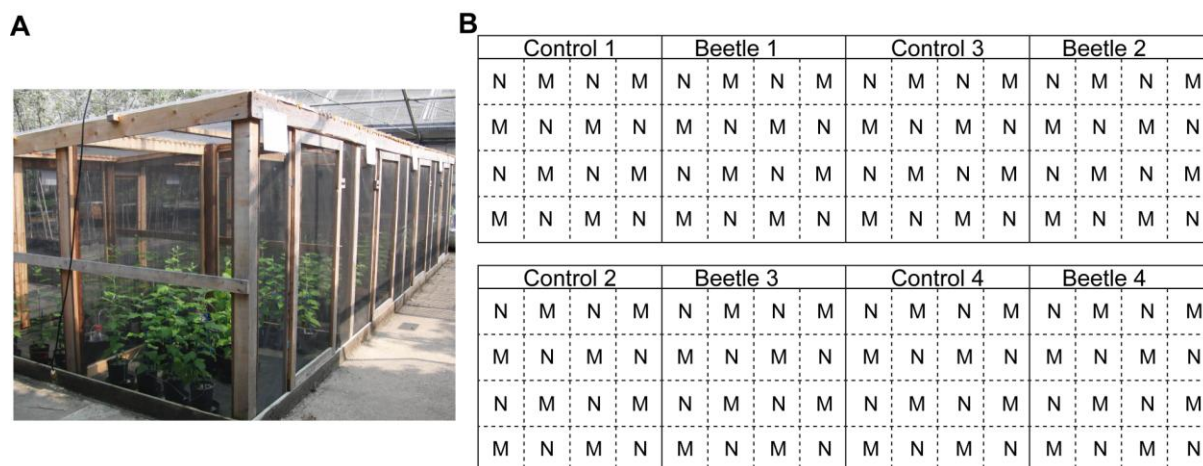


Figure 5.1. Experimental setup of the *C. populi* behavior study with EMF inoculated (M) and non-inoculated (N) poplars under field conditions. (A) Photo of the free-air cages in Göttingen and (B) the arrangement of poplars in each cage. (modified according to Müller 2014)

5.3 Results and Discussion

5.3.1 Summary of physiological results

After 14 days the percent of root tip mycorrhization of inoculated poplars was determined as being 9.5 ± 0.6 %. On non-inoculated plants no mycorrhizas were found. Throughout the experiment *C. populi* preferred feeding on non-inoculated plants as seen by the significantly higher leaf damage of infested non-inoculated poplars (NB) compared to infested inoculated poplars (MB; for more details see Müller, 2014). Additionally, a significantly higher number of eggs were found on leaves of NB than on MB ones (Müller, 2014).

In addition to the metabolomics experiments presented in the present thesis, transcriptomics was performed in Göttingen by Dr. Anna Müller (Müller, 2014). Transcriptional changes were found in leaves of inoculated poplars when compared to non-inoculated poplars (for more detail see Müller, 2014). The aim of this Chapter is to merge the transcriptional data with metabolite and VOC data via mass difference networks (MDiN) and REMDs.

5.3.2 Statistical analysis of GC-MS data

Plant volatiles were collected with twisters, analyzed with the GC-MS (personal communication Maaria Rosenkranz and Kerstin Koch, both EUS) and made available for analysis in the present thesis. In total 42 volatiles were detected and grouped to five compound classes as follows; (i) monoterpenes (MT), sesquiterpenes (SQT), (ii) phenolics, (iv) furans and, (v) green leaf volatiles together with fatty acids and aldehydes (GLV&FA&Alde).

For statistical analysis, the VOC concentrations were utilized as variable “X” to OPLS-DA analysis. The X-variables were pareto-scaled and the four experimental parameters (MC/NC, MB/NB, MB/MC, NB/NC) were assigned as Y-variables. Unfortunately, PLS and OPLS-DA models explaining the mycorrhization (MB/NB & MC/NC) and the beetle effect (NB/NC & MB/MC) separately, had a negative Q^2 . For that reason the beetle treatments were combined in that NB and MB (n=6) were tested against MC and NC (n=6). Discriminant volatiles were extracted from the S-plot, which was constructed from the OPLS-DA model, based on their correlation value ($p(\text{corr}) > 0.8$, Fig. 5.2A).

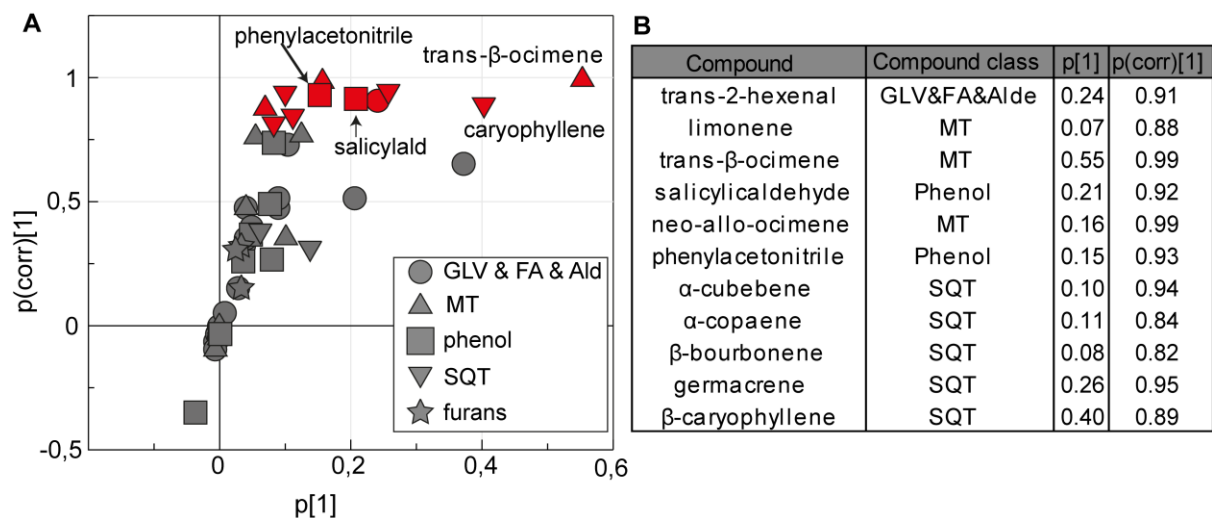


Figure 5.2. Statistical analysis of the GC-MS data. (A) S-plot showing the contributions of volatiles ($p(\text{corr})[1]$ = correlation coefficient between X and the score vector 1; $p[1]$ = score vector 1) for the comparison of fed (n=6) versus unfed (n=6) poplars, significant volatiles are colored in red and shown in table (B). Legend: GLV: green leaf volatile, FA: fatty acid, Alde: aldehyde, MT: monoterpene, SQT: sesquiterpene.

Eleven volatiles were discriminant for the beetle treatment, of which eight belonged to either mono- or sesquiterpenes (Fig. 5.2B). The MT beta-ocimene and the SQT caryophyllene had the highest contribution to the OPLS-DA model. Both compounds are well known volatiles that are induced in herbivore-infested plants to attract parasitoids (De Moraes et al., 1998). *C. populi* uses MT and SQT to search for high quality young poplar leaves (Brilli et al., 2009). Other induced volatiles consisted of GLV and phenolic compounds such as salicylaldehyde and phenylacetoneitrile. The latter compound belongs to the group of phytoaldoximes which are synthesized from amino acids and function either as attractants for natural herbivore enemies or, due to their high biological activity, as direct herbivore repellents (Irmisch et al., 2013; Irmisch et al., 2014).

Generally, aldoximes are minor components of herbivore-induced volatile blends (Irmisch et al., 2014). Recently it has been shown that aldoximes are released shortly after herbivore

attack and that their emission stops when the herbivore is removed from the damaged leaf (Clavijo McCormick et al., 2014). Consequently the emission of aldoximes depends on the number of insect currently feeding on the plant, which explains the high variability in phenylacetonitrile concentration between the biological replicates. Phenylacetonitrile was, however, exclusively detected in fed poplars, validating the transcriptomic results in which transcripts coding for enzymes of the aldoxime biosynthesis were found to be upregulated under beetle infestation (Müller, 2014).

5.3.3 The development of a workflow for REMD-transcript mapping

Very recently it was shown that the network based sum formula assignment of Tziotis et al. (2011), which is usually used on high-resolution MS instruments, is also applicable on mass spectrometers with lower resolution such as qToF systems (Forcisi et al., 2015). Furthermore, Chapter 3 successfully showed how mass spectrometric data can be used linking metabo- to genotypes via REMDs pertaining to the macromolecular -omes. Hence, the logical consequence would be the development of a workflow, which aims at mapping the results obtained by mass-difference analysis on genomic, transcriptomic or proteomic data. Since in this case next generation sequencing data were made available (personal communication Anna Müller; Müller, 2014), the REMDs could be mapped on enzyme coding transcripts. A scheme of the data analytical workflow is given in Figure 5.3.

It starts with the assignment of sum formulas to high resolution direct injection (-)ICR-FT/MS m/z features. To achieve this, the filtered ICR-FT/MS m/z -list was subjected to NetCalc in order to perform 10 individual annotation runs using a REMD list consisting of both the manually curated REMD and the REMD-rpair list. In total 2789 sum formulas could be assigned.

First it is necessary to annotate the low-resolution ToF data for both the negative and positive ionization modes. To ensure maximum accuracy of the annotation procedure, the determination of starter masses is crucial. Therefore, some typical plant phenolics were measured and their respective retention time (RT) and m/z value was compared with the peak-list of the non-targeted LC measurements. The concurring m/z features were used together with m/z features, whose sum formulas could be determined by MS/MS experiments, as starting masses for the NetCalc annotation procedure on (-)LC-MS data. 17 starting masses were used for 100 consecutive NetCalc annotation runs. Sum formulas with an annotation frequency of 100% were used for further analysis (Forcisi et al., 2015).

Secondly, to further improve the accuracy of the (-)LC-MS sum formula annotations, they were matched on the ICR-FT/MS annotations. The 432 sum formulas (without isomers) which were assigned in both the (-)LC-MS and the (-)ICR-FT/MS measurements were

utilized as new starting masses for a follow-up round of 100 consecutive (-)LC-MS annotation runs. Sum formulas with an annotation frequency better than 90% were used to reconstruct a theoretical (-)LC-MS MDiN. Finally, 2422 stable sum formulas were obtained.

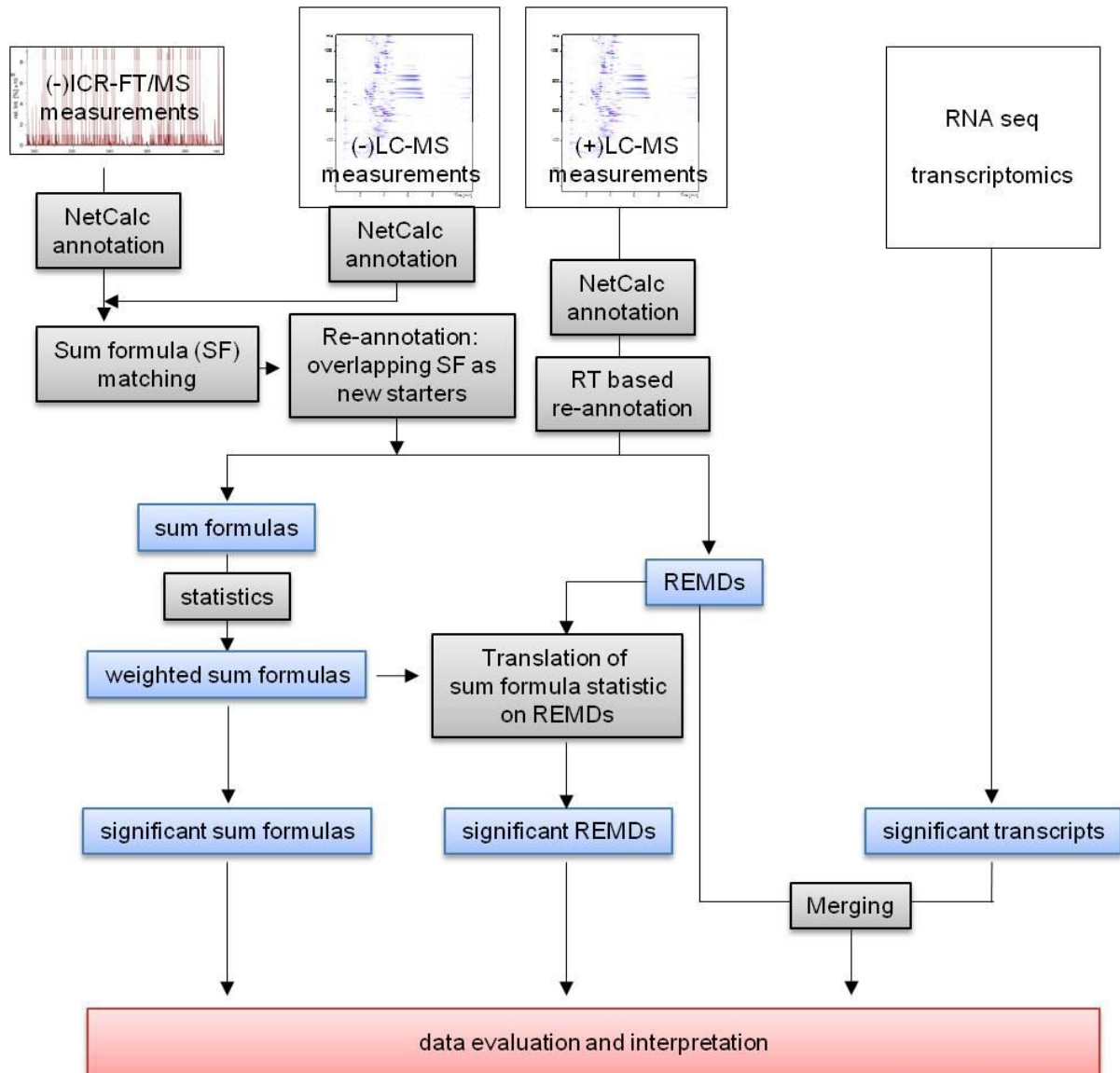


Figure 5.3. The extension of the mass difference analysis workflow which aims at mapping the transcriptomic data on their respective REMDs (compare to Figure 3.4).

Unfortunately, due to several ion-suppressing compounds in the extracts, no representative positive ICR-FT/MS spectra could be measured, thus preventing the matching on high-resolution data. In order to still assign sum formulas to (+)LC-MS m/z features, the workflow needed to be adjusted. Similar to the (-)LC-MS annotation procedure, the initial step implies the selection of trustworthy starting masses for an initial run of 100 NetCalc annotations (combined REMD list). However, that is where the similarity between

the (-)LC-MS and the (+)LC-MS procedure ends (Fig. 5.3). To annotate (+)LC-MS m/z features, sum formulas possessing an annotation frequency larger than 90% were used for mass-difference mining and the creation of mass-difference RT rules (for detailed information see Forcisi et al., 2015). The resulting RT mass-differences were used for a second round of 100 NetCalc annotations in order to delete wrong sum formula annotations. After that, the 2644 sum formulas with an annotation frequency >90% were used for the reconstruction of a theoretical MDiN by using the combined REMD list. The last step resulted in 2644 sum formulas that were used for further data analysis (see Materials and Methods).

5.3.4 Root mycorrhization and beetle feeding causes metabolic alterations in poplar leaves

The NetCalc annotations were subjected to PCA analysis (Fig. 5.4A-B). Within the (-)LC-MS PCA model, each experimental condition - namely NC, MC, NB and MB - was positioned separately in one PCA quadrant, thus illustrating the presence of both a mycorrhization and a herbivore leaf-metabotype (Fig. 5.4A).

The (+)LC-MS PCA model did, however, only show a separation between infested and non-infested plants in the PC1xPC2 score plot (Fig. 5.4B). The mycorrhization effect (EMF effect) was found in PC3 and accounted for 10.5% of explained data variance (Supporting Fig. S1). To extract discriminant sum formulas, for each experimental condition and ionization source polarity, separate OPLS-DA models were calculated.

Within the (-)LC-MS measurements, a total of 39 sum formulas were found to be discriminant for EMF inoculation under control (MC/NC) and 86 under beetle conditions (MB/NB), respectively (Fig. 5.4C). A major part of these sum formulas, namely 25 in MC/NC and 42 in MB/NB (14 sum formulas overlapped), were downregulated. The remaining 14 and 44 sum formulas were upregulated in control conditions and under herbivory, respectively, but here only two sum formulas possess the same behavior in both EMF scenarios (Fig. 5.4C). Statistical analysis of the (+)LC-MS data yielded 146 and 167 sum formulas discriminant for the EMF effect under control and beetle conditions, respectively (Fig. 5.4D). Here, 56 sum formulas were downregulated under control and 76 under beetle conditions, 12 being downregulated simultaneously (Fig. 5.4D). Additionally, 90 and 91 sum formulas were upregulated in control and beetle conditions, respectively (Fig. 5.4D).

Beetle attack resulted in a much stronger metabolic response in the leaves compared to the changes caused by the EMF. Statistical analysis of the (-)LC-MS data, resulted in a total of 220 and 228 sum formulas discriminant for the beetle effect in inoculated and non-inoculated poplars, respectively (Fig. 5.4E). The calculation of log ratios showed that 76 of these sum formulas were downregulated under beetle attack in non-inoculated and 100 in EMF poplars. Consequently, 144 and 128 sum formulas were upregulated in both non-

inoculated and EMF poplars, respectively (Fig. 5.4E). Similar results were obtained by analyzing the (+)LC-MS measurements, in which 518 sum formulas were discriminant for the NB/NC comparison and 452 sum formulas were found to be discriminant for the MB/MC comparison (Fig. 5.4F). Thereby, the majority of sum formulas (59% in NB/NC and 65% in MB/MC) were upregulated in both beetle effect comparisons (Fig. 5.4F). The remaining sum formulas were downregulated.

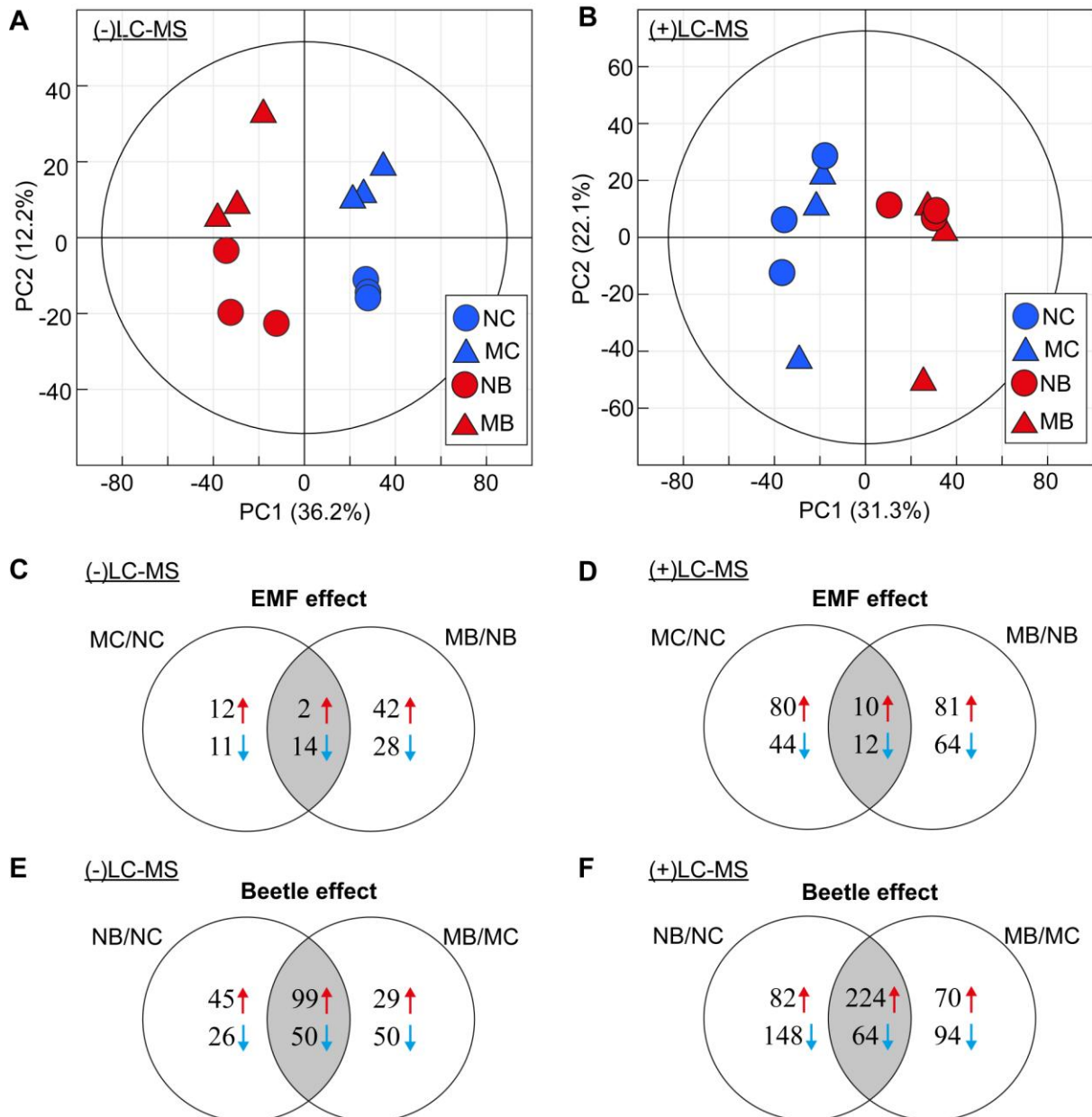


Figure 5.4. Statistical results of the LC-MS measurements. (A) PCA score plot of (-)LC-MS data and (B) PCA score plot of (+)LC-MS data. Venn-plots of sum formulas discriminant for the EMF effect in (C) (-)LC-MS- and (D) (+)LC-MS measurements. Venn-plots comparing sum formulas discriminant for the beetle effect obtained by (E) (-)LC-MS and (F) (+)LC-MS measurements. Red arrows: upregulated sum formulas; blue arrows: downregulated sum formulas.

5.3.5 Inoculation and herbivory impacts the CHO and CHNO compositional space of leaves

Comparing the elemental composition of sum formulas discriminant for mycorrhization and herbivory, revealed that the majority belonged either to CHO or to CHNO sum formulas (Fig. 5.5A).

This result can also be seen by analyzing the molecular compositions of discriminant sum formulas. In the (-)LC-MS measurements a total of 108 sum formulas were discriminant for the EMF inoculation (MC/NC and MB/NB, Fig. 5.4C), of which 101 (93.5%) belonged to CHO, two to CHNO, four to CHOS and one to CHS containing sum formulas (Fig. 5.5A).

Different patterns were found by analyzing the sum formula compositions that were upregulated under EMF inoculation. Here, the beetle attack caused the upregulation of 43 CHO and one CHS sum formula, whereas under control conditions only 12 CHO and two CHNO sum formulas were upregulated (Fig. 5.5A). This means that inoculation induces a stronger response of certain CHO containing compounds when the leaves are attacked by herbivores. These sum formulas possessed a very high degree of saturation as seen by the high cyclomatic number $u = 18.18$ (Table 5.1), which is a measure for the molecules' saturation (Senior, 1951), thus hinting towards a stronger induction of specific aromatic compounds in MB compared to NB. Additionally, a significantly lower degree of saturation was found in EMF downregulated than in EMF upregulated sum formulas (Table 5.1; $u_{\text{down}}=13.17$, $u_{\text{up}}= 17.38$, $P=0.0025$).

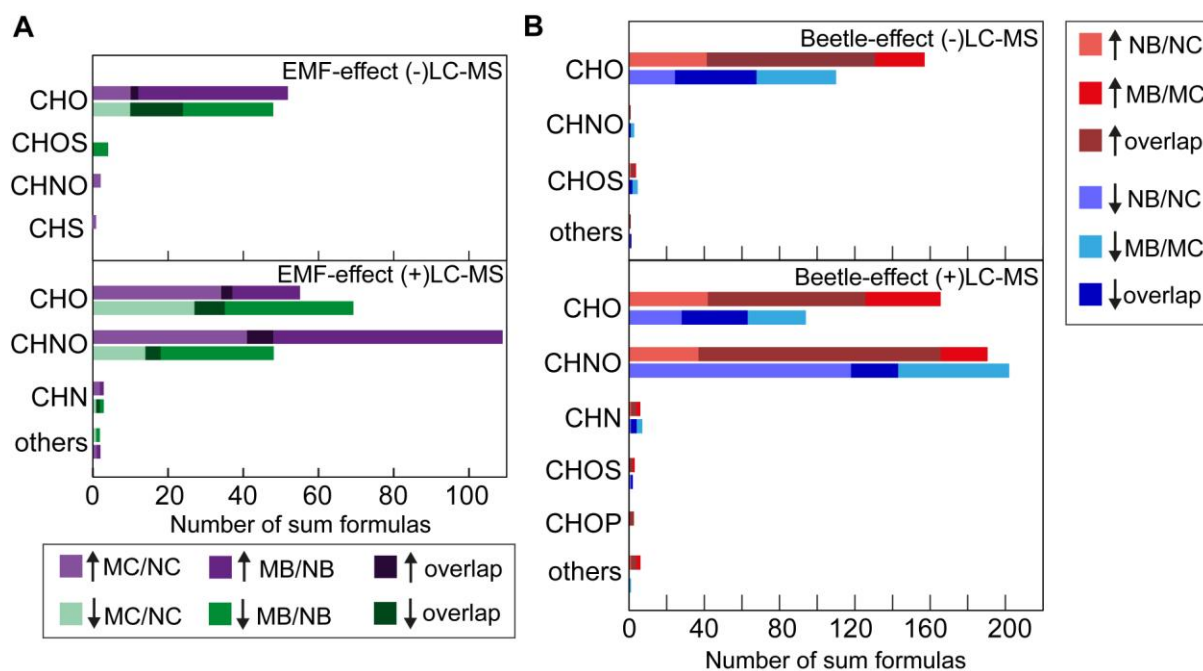


Figure 5.5. The compositional space of discriminant sum formulas. Sum formulas discriminant for (A) EMF inoculation and (B) beetle infestation.

The majority, namely 158 (54%), of the EMF compositional space from the (+)LC-MS sum formula annotations belonged to CHNO compositions (Fig. 5.4A). 110 of those were upregulated in EMF plants. As already mentioned in the introduction (Paragraph 1.5.3), the formation of a mycorrhiza improves the plants' nitrogen uptake from the soil (Boxman and Roelofs, 1988; Parniske, 2008). Consequently the finding of enhanced CHNO compounds in systemic leaves of inoculated poplars is a direct result of increased nitrogen availability. A significantly lower cyclomatic number, i.e. $u=9.33$ (Table 5.1), was found for sum formulas discriminant for the EMF in the positive than in the negative mode.

Statistical analysis of the (-)LC-MS data led to the identification of 299 sum formulas discriminant for beetle infestation, with 285 of those belonging to CHO compositions (Fig. 5.4B). The cyclomatic number between up- and downregulated sum formulas was identical ($u=13.67$; Table 5.1) and no differences between non-inoculated (NB/NC) and inoculated (MB/MC) grey poplars were found.

Table 5.1. Cyclomatic number u of discriminant sum formulas.

	EMF-effect			Beetle-effect		
	$u(\text{MC/NC})$	$u(\text{MB/NB})$	$u(\text{average})$	$u(\text{NB/NC})$	$u(\text{MB/MC})$	$u(\text{average})$
(-)LC-MS: upreg.	16.35	18.18	17.38	13.42	13.53	13.67
downreg.	13.5	12.98	13.17	12.66	14.09	13.67
(+)LC-MS: upreg.	9.07	9.89	9.55	10.87	10.86	10.86
downreg.	9.78	8.48	9.01	11.29	9.81	10.45

393 of the 681 sum formulas (57.71%), which were found to be discriminant for beetle infestation, belonged to CHNO compositions (Fig. 5.5B). Notably, the absence of the EMF under beetle attack resulted in a higher number of downregulated CHNO compounds than in the presence of the EMF (Fig. 5.5B). This finding also demonstrates the better nitrogen supply in leaves of mycorrhized plants.

5.3.6 Mycorrhization by EMF impacts the phenolic metabolism of poplar leaves

First indications for metabolic changes pertaining to the nitrogen and to phenolic metabolism were found by analyzing sum formulas discriminant for either EMF inoculation or beetle infestation. To further analyze this, all discriminant sum formulas were uploaded to the Masstrix 3 server to annotate them as metabolites (Suhre and Schmitt-Kopplin, 2008). Overall 69 metabolites (30%) were annotated in the negative and 158 (20%) in the positive mode. All annotations were grouped into their respective compound classes and assigned to the experimental condition for which they were discriminant (Figure 5.6A-B).

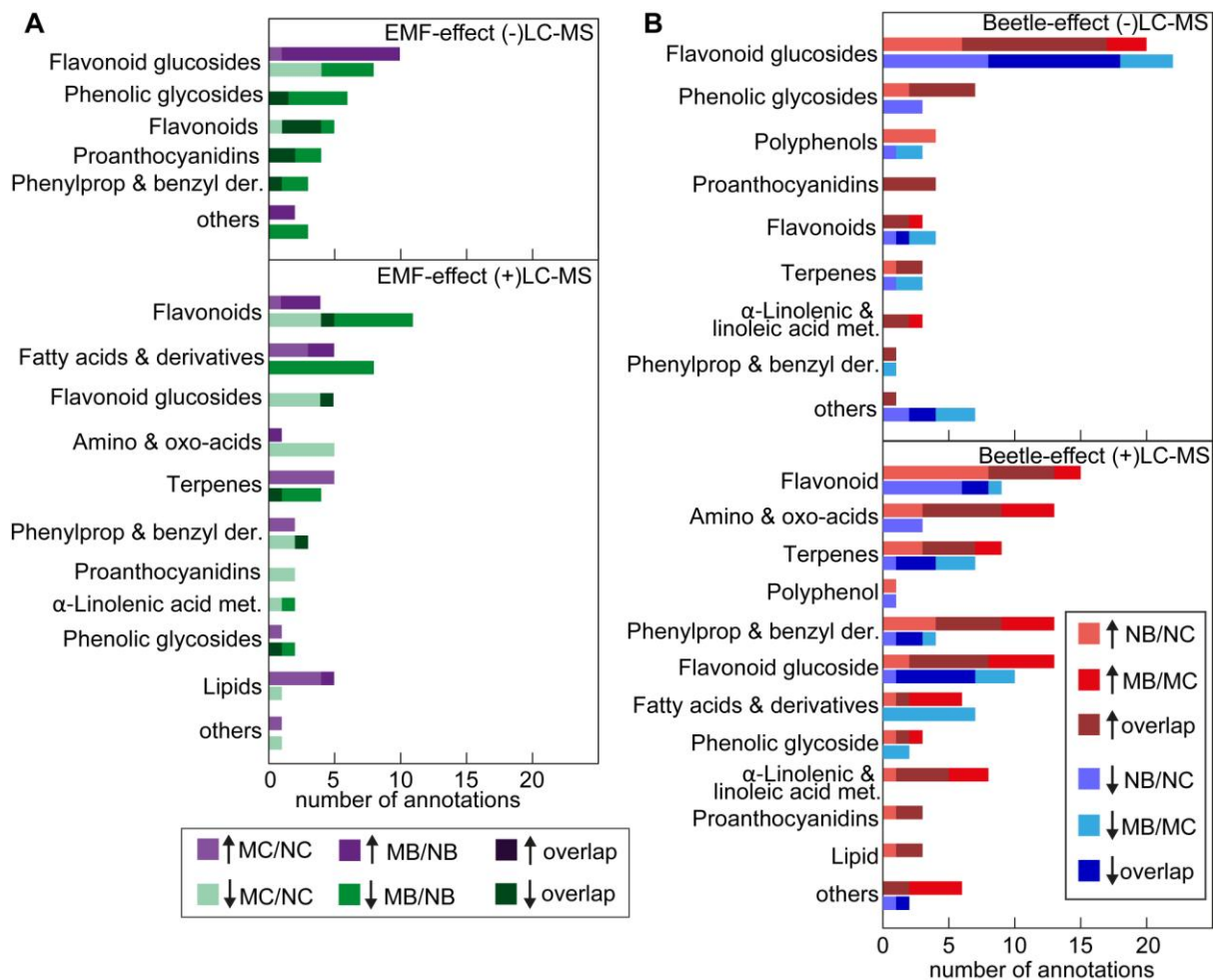


Figure 5.6. Illustration of database annotations obtained by uploading the discriminant sum formulas. Metabolites discriminant for (A) EMF inoculation and (B) beetle infestation.

This grouping revealed a clear connection between EMF inoculation of roots and the phenolic metabolism of leaves (Fig. 5.6). A total of 75 annotated metabolites in leaves that were discriminant for EMF inoculation were phenolics. Their majority, namely 68% (51 metabolites), were downregulated and 17 out of the 24 upregulated phenolic compounds

were only induced when beetles fed on the foliage of EMF poplars (MB/NB). Other compound classes changing in the leaves under EMF inoculation included lipids, terpenes and fatty acids (Fig. 5.6).

Beetle infestation on the one hand, lead to an upregulation of a variety of bioactive compounds in leaves, such as flavonoids and phenolic glycosides, and on the other, it induced the upregulation of lipids and fatty acids, especially of those belonging to the α -linolenic acid metabolism (Fig. 5.6B). The compounds of this metabolic pathway are involved in leaf priming and in the synthesis of green leaf volatiles via lipoxygenases (Wasternack and Parthier, 1997; Weber, 2002; Matsui, 2006).

Interestingly, both the EMF inoculation and the beetle infestation strongly affected leaf metabolites belonging to phenolic pathways, such as the biosynthesis of phenylalanine, phenylpropanoids, flavonoids and proanthocyanidins (Fig. 5.6). However, the way in which these pathways respond to EMF inoculation differed substantially from the way they react to beetle infestation. While mycorrhization caused a downregulation of flavonoid precursors in leaves - such as quercetin, dihydroquercetin, kaempferol, dihydrokaempferol and catechine – as well as of proanthocyanidins, namely two proanthocyanidine dimers and two trimers (identified by MS/MS, Supplementary Fig. S2), the beetle infestation induced their upregulation. Strikingly, these patterns are in accordance to the transcriptomic results in which transcripts coding for chalcone synthase 1 (CHS1), flavanone 3-hydroxylase (F3H), flavonoid 3'-monooxygenase (tt7), dihydroflavonol 4-reductase (DFR), anthocyanidin reductase (BAN) and leucoanthocyanidin dioxygenase (LDOX) were downregulated under EMF inoculation and upregulated under beetle infestation (Müller, 2014; Fig. 5.6A).

5.3.7 Mapping of transcripts on metabolite data via REMDs

One classical way of cross-omics data visualization is the integration of both transcriptomic and metabolomic log-fold changes into KEGG pathway maps (Fig. 5.7A). This method was chosen to layout the flavonoid and proanthocyanidin pathways (Fig. 5.7A). As seen in Figure 5.7, it shows an almost perfect accordance between both the metabolite and transcript abundances as the biosynthesis of flavonoids and proanthocyanidins was downregulated under EMF inoculation, but upregulated under beetle infestation.

Previous works showed that the downregulation of phenolic pathways serves as a good indicator for the cellular supply of nitrogen (Muzika, 1993; Keski-Saari and Julkunen-Tiitto, 2003). Generally, a high availability of nitrogen results in the downregulation of secondary

metabolic processes in order to provide a sufficient amount of carbon skeletons for nitrogen fixation and amino acid biosynthesis (Chellamuthu et al., 2014).

As already described in Chapter 3, the MDiN based annotation procedure has the advantage of having access to an additional domain for data interpretation, namely the REMDs. Since they are derived from biochemical reactions and therefore pertain to the macromolecular omes, they are well suited for systems biology integration by means of transcriptomic/metabolomic data fusion. In addition, on one hand, spontaneous reactions, such as the malondialdehyde formation, can also be grasped by REMDs therefore complementing the transcriptomic results. On the other hand, transcripts can also complement metabolomic/REMD results since they carry regulatory information, e.g. the myb transcription factor family, explaining metabolic induction/inhibiting processes.

To achieve cross-omics data fusion by means of REMD-transcript matching, first all transcript coding for metabolic active enzymes were assigned to their respective EC-numbers using R. Secondly, the EC numbers were matched on their respective KEGG rpairs (downloaded from KEGGAPI) and finally to the REMDs of the MDiN. For the comparison of non-infested EMF plants (MC/NC) 30 REMDs described reactions catalyzed by one of the significant metabolic active enzymes found for MC/NC. Next a sub-graph (Fig. 5.7C) containing only sum formulas connected by at least one of the 30 REMDs was extracted (negative LC-MS data was used). It was constructed from 650 edges (REMDs/transcripts) and 522 nodes (metabolites). The graphical visualization showed a high modularity (0.76) and clustered into 14 communities (Fig. 5.7C). Strikingly, it contained the entire flavonoid pathway relationship, whereby for the first time, the transcripts coding for the FLS, TT7, DFR, LAR and LAR flavonoid pathway enzymes were directly mapped on their respective REMDs which in turn were connected to their specific flavonoid educt-product pairs (Fig. 5.7B).

Furthermore, the flavonoid monomers, namely naringenin, dihydrokaempferol, quercetin, dihydroquercetin, leucocyanidin and catechin, were located in one single network community (Fig. 5.7B-C). In accordance to that, both of the proanthocyanidin dimers were found in a separate network community and on top of that the proanthocyanidin trimers clustered in a separated sub-graph. Given these facts, it is obvious that structurally similar metabolites arrange in network communities. This complements and validates the results obtained in Chapter 3 in which a compound class-based community clustering was observed (Chapter 3.3.10).

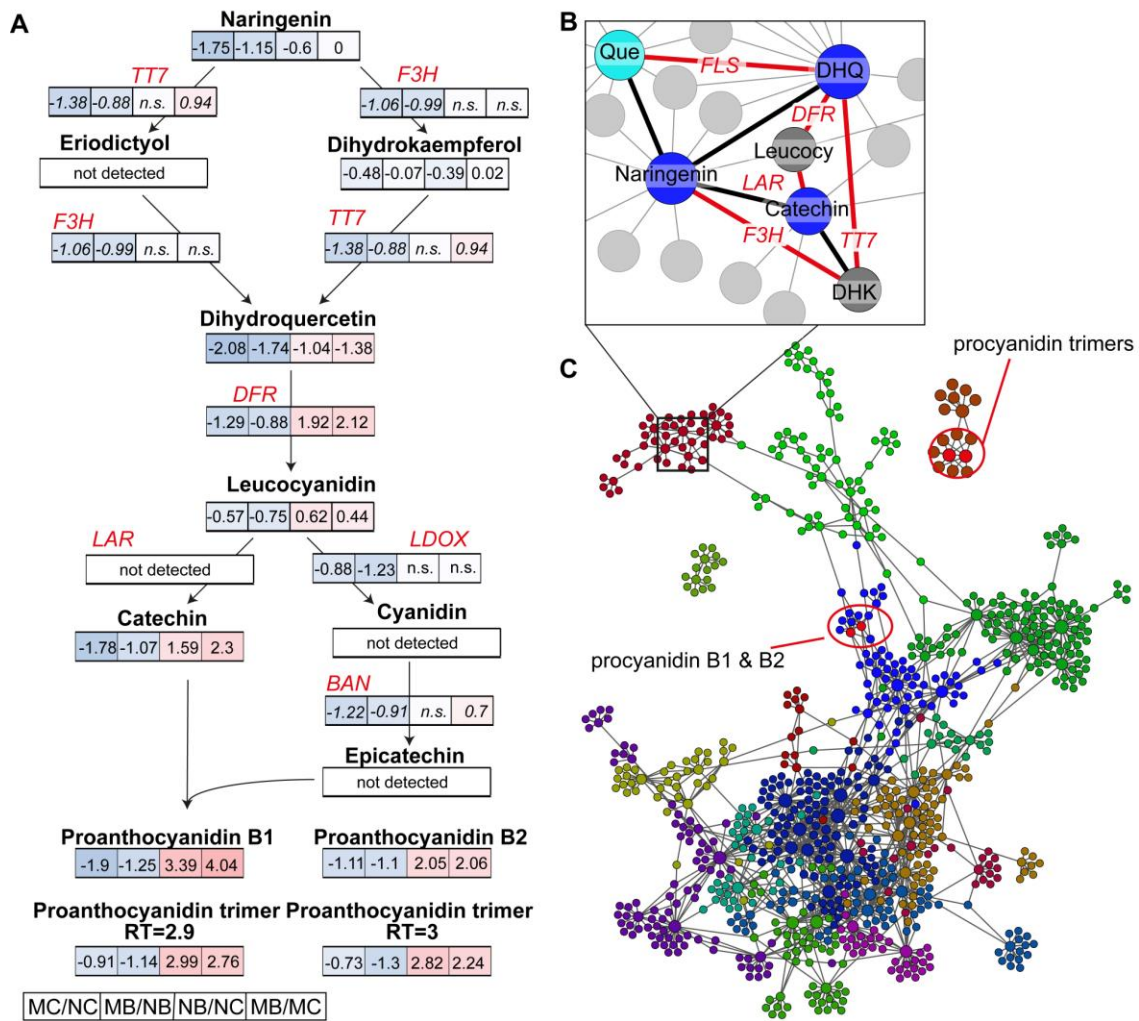


Figure 5.7. Flavonoid biosynthesis exemplifies the transcriptomic-metabolomic data matching via REMDs. (A) Log₂ fold changes of metabolite and transcript levels of the proanthocyanidin pathway. (B) (-)UPLC-qToF-MS mass-difference network of the flavonoid/proanthocyanidin biosynthesis with transcripts matched on the REMDs.

Flavonoids and hydrolysable tannins, such as proanthocyanidins, are important compounds in the chemical defense of poplars against leaf herbivores (Philippe and Bohlmann, 2007). With this in mind it is somehow astonishing that EMF inoculation induces the downregulation of both metabolite and transcript levels in leaves pertaining to this important pathway under beetle infestation. Still the analysis of plant performance showed that *C. populi* preferred to feed and ovipose on non-inoculated poplar leaves (Müller, 2014). Consequently, the mycorrhization of roots with *L. bicolor* must induce systemic defensive systems that are in leaves more effective against the specialist herbivore *C. populi* than flavonoids and their derivatives.

5.3.8 Over-represented REMDs show an induction of aldoxime condensations and validate prior knowledge about herbivore induced pathways

As shown in Chapter 3, mass-difference enrichment analysis (MDEA) complements and extends knowledge bases. Here, MDEA was implemented on the (-)LC-MS data to further explain the metabolic and transcriptomic findings and indeed, its results verified and deepened many of them as follows:

First, under EMF control conditions (MC/NC) 13 REMDs were over-represented in MC poplars of which one described hydrolysis/condensation reactions ($\Delta m = H_2O$). The use of water molecules for metabolite modifications is among the most frequently used reactions within living organisms (Breitling et al., 2006). More importantly, it directly refers to the flavonoid metabolism, which before was found to be of major importance for in mycorrhized plants.

Secondly, another important finding was the higher content of transcripts coding for enzymes of the aldoxime biosynthesis in mycorrhized and beetle infested (MB) poplars compared to EMF free beetle treated (NB) poplars (Müller, 2014). MDEA yielded the overrepresentation of three nitrile REMDs in MC, namely (i) the propionitrile transfer, (ii) the 2-hydroxy-2-methylbutanenitrile and (iii) the hydroxymandelonitrile condensation (Fig. 5.8A-B). Additionally, in MB the highest z-score was obtained by the prunasin condensation REMD ($z=3.65$; Fig. 5.8B). This REMD is the glycosylated form of hydroxymandelonitrile, whose REMD, was also overrepresented in MB/NB as well as in MC/NC. At the same time the condensation of phenylalanine was also overrepresented in MB poplars. Strikingly, all of these REMDs correspond to KEGG rpaths found in the cyanoamino acid metabolism in which the biosynthesis of the volatile aldoxime phenylacetonitrile is located. This VOC was detected in the GC/MS measurements and was significantly upregulated in beetle infested poplars (Fig. 5.2). Additionally, three transcripts coding for P450 monooxygenases, of which two were recently characterized to catalyze the formation of volatile nitriles (Irmisch et al., 2013), which are derived from aldoximes were upregulated under herbivore attack. Two of those P450 enzymes, namely CYP79D6 and CYP79D7, possess broad substrate specificity using at least five different amino acids as substrates (Irmisch et al., 2013). Among them is isoleucine, the biosynthetic precursor of 2-hydroxy-2-methyl-butanenitrile whose REMD was associated to the EMF effect (Fig. 5.8D).

Taken together, these findings substantiate the higher production of volatile nitriles as defensive compounds in leaves of mycorrhized poplars, which is most probably due to a better nitrogen allocation under symbiotic conditions.

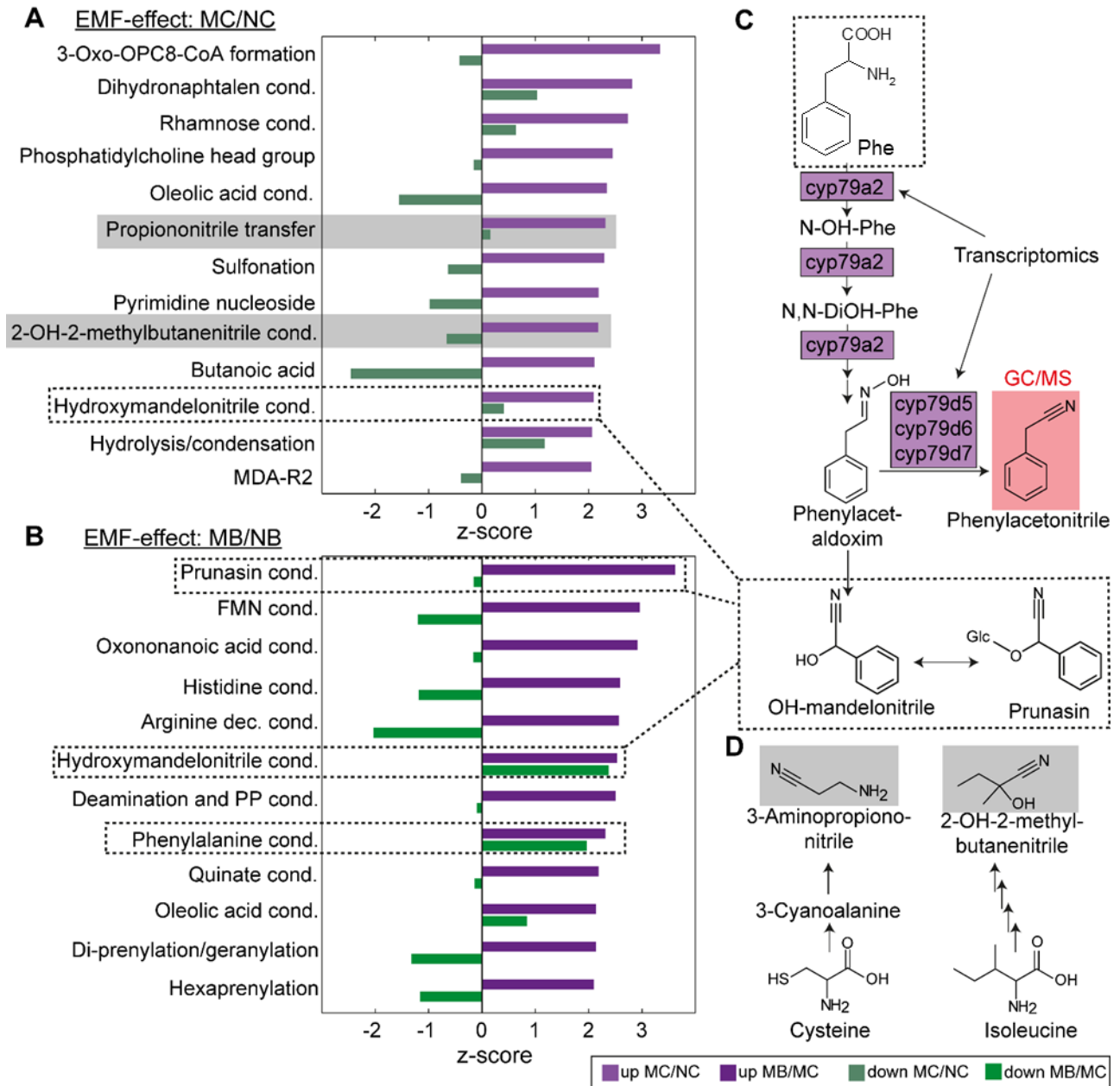


Figure 5.8. Z-scores obtained by MDEA analysis of the EMF-effect in the (-)LC-MS samples. Overrepresented REMDs in (A) the MC/NC comparison and in (B) the MB/NB comparison. REMDs highlighted with dashed lines pertain to the cyanoamino acid metabolism KEGG pathway map starting from (C) phenylalanine (Phe), and REMDs highlighted in grey start from either (D) cysteine or isoleucine. Red background: detected with GC/MS; purple background: detected via transcriptomics; cond.: condensation; DiOH: dihydroxy; FMN: flavinmononucleotide; Glc: glucose; MDA: malondialdehyde; OH: hydroxyl.

Third, the REMD class representing fatty acid reactions was overrepresented in leaves of mycorrhized poplars. In the MC/NC comparison the alpha-linoleic acid REMD explaining the condensation/hydroxylation of a C16 fatty acid yielded the highest z-score. Additionally, the oleolic acid condensation (C18 fatty acid) REMD was upregulated in both EMF conditions (MC/NC & MB/NB). Plus, the REMD of the important signaling compound 9-oxononanoic

acid was overrepresented in mycorrhized and beetle treated (MB) poplars (Fig. 5.8). This fatty acid plays an important part in systemic acquired resistance by priming undamaged leaves (Wittek et al., 2014).

Table 5.1. REMDs overrepresented for beetle feeding in (A) the NB/NC and (B) the MB/MC comparison.

A			
NB/NC			
REMD name	Reaction Class	Mass difference	Z-score (N B)
Linoleic acid condensation	Kegg	260.214015	5.22
Double Hydrogenation	Kegg	2.99314	3.42
Prenylation	F/Kegg	288.06339	3.19
Vicinal dienol to nitroso functionality	Kegg	9.98435	2.69
Aminobenzoate condensation	Kegg	118.02661	2.63
Addition of Phosphate	Kegg	95.037114	2.55
Aminoethylphosphonic acid condensation	Kegg	99.047284	2.55
Tyrosine	A	162.052825	2.43
5-Amino-2-oxopentanoic acid	A	113.011294	2.42
Cyanoalanine	Kegg	95.961247	2.42
3-Hydroxy-2-oxobutanoic acid	B/Kegg	56.026215	2.36
Benzylaldehyde	A	88.016045	2.29
Asparagine	A	114.031695	2.23
DNA-R1	F	54.010565	2.21
Loss of oxygenated monoterpene moiety	Kegg	74.000395	2.19
Arginine	A	156.078645	2.19
2-Ketoisovaleric acid	D/Kegg	72.02113	2.18

B			
MB/MC			
REMD name	Reaction Class	Mass difference	Z-score (MB)
Linoleic acid condensation	Kegg	262.229665	6.48
Vicinal dienol to nitroso functionality	Kegg	12.995249	3.25
2-Oxoarginine	A	155.069477	3.09
Oleic acid condensation	Kegg	264.245315	3.02
3-Hydroxy-2-oxobutanoic acid	B/Kegg	56.026215	2.66
Phosphatidylcholine head group	A	239.092261	2.55
2-Ketoisovaleric acid	D/Kegg	72.057515	2.30
Prenylation	F/Kegg	68.0626	2.15
Arginine	A	156.101111	2.05

Forth, the MDEA results verified the induction of the α -linoleic acid metabolism. Independently from the mycorrhization status under herbivory the condensation of linoleic acid yielded the highest z-score (Table 5.1). When plants suffer leaf damage the α -linoleic acid metabolism is induced. The enzymatic cleavage of α -linoleic acid through lipoxygenases creates C6 volatiles, better known as GLV which function as important

signaling molecules priming defense mechanisms of undamaged leaves (Arimura et al., 2009). Additionally, the α -linoleic acid pathway is used for the biosynthesis of jasmonic acids, an important signaling molecule in the plant response to biotic stresses (Delker et al., 2006).

Futhermore, the prenylation was overrepresented under herbivory (Table 5.1). As beetle feeding caused an induction of MT and SQT emission and the fact that some cellular terpenes, such as steroids, were upregulated in infested plants, validate this finding.

5.3.9 Combination of (-/+)-LC-MS, GC-MS and transcript data

Not only do REMDs improve the data interpretation but also they allow a more simplified cross-platform MS data matching and visualization. In the here presented case the combination of three different MS measurements, namely (+)LC-MS, (-)LC-MS and GC/MS, is necessary by means of MDiN and REMDs. Therefore, the theoretical masses of all three MS methods were cross-networked using NetCalc. On this cross-platform MDiN consisting of 4525 nodes and 114,428 edges, MDEA was performed for each experimental condition.

To enable a global view of all metabolic alterations induced by mycorrhization and herbivory in poplar leaves, all discriminant sum formulas were extracted. For the graphical visualization the 68 REMDs that were overrepresented in mycorrhized leaves were chosen (Fig. 5.9). Strikingly, the volatiles were not randomly cross-networked into the MDiN. They were directly associated to sum formulas pertaining to their respective intracellular biosynthetic pathway (Fig. 5.9B) was generated. For example sesquiterpenes were connected to farnesoic acid, the carboxylic acid of farnesyl/farnesylpyrophosphate, the C15 precursor of sesquiterpene synthases. Similar findings were obtained for phenylacetonitrile and salicylaldehyde, which both were connected to phenolic precursors of their respective biosynthetic pathway.

Additionally, this network representation visualizes the effect of mycorrhization on the nitrogen metabolism of leaves namely that the CHNO compositional space (Fig. 5.9A) was upregulated under EMF inoculation but downregulated under beetle attack.

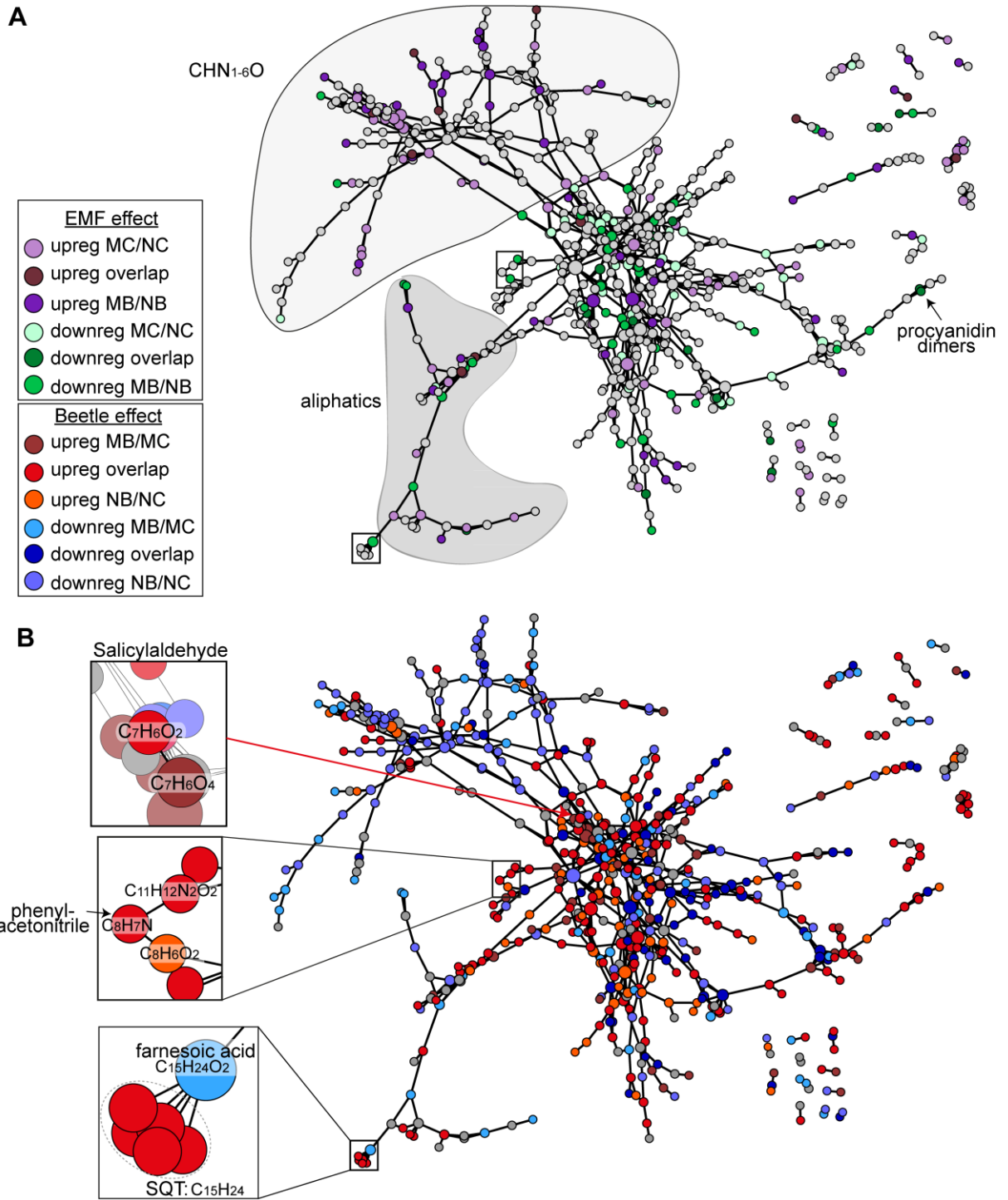


Figure 5.9. Cross-platform mass difference network. (A) Discriminant sum formulas colored according to the EMF effect and (B) to the beetle-effect.

5.3.10 Proposed mechanism of induced defensive system caused by EMF inoculation

Beneficial mycorrhizal associations provide plants with a better resistance against a variety of biotic and abiotic stresses (Nehls et al., 2010). This is also the case for the here investigated symbiosis between *L. bicolor* and *Populus x canescens* in which the mycorrhizal formation decreased the attractiveness of poplar leaves for the specialist herbivore *C. populi*, caused by compositional changes in the plant leaves (Fig. 5.10). Just recently a work reported that arbuscular mycorrhizal fungus (AMF) formation induces species-specific changes in the chemical composition in leaves of five different plant species (Schweiger et al., 2014). While other works report that AMF-inoculated *Plantago lanceolata* were equipped with higher amounts of bioactive metabolites in their leaves (Gange and West, 1994; Fontana et al., 2009), the EMF inoculation of grey poplars induced a downregulation of bioactive phenolic compounds, namely tannins, flavonoids and phenolic glycosides, and of their respective enzymes (Fig. 5.7; Fig. 5.10). Especially the proanthocyanidins behaved differently, as two proanthocyanidin dimers and trimers were downregulated in leaves when roots were mycorrhized but became highly upregulated under herbivory (Fig. 5.7A). Nevertheless, the global phenolic composition of poplar leaves don't seem to be driving the *C. populi* decision as the beetles prefer the phenolic-rich leaves of non-mycorrhized plants that contained higher amounts of phenolic compounds. This is in accordance with another work on transgenic poplars with increased production of condensed tannins which showed that two generalist herbivores preferred tannin rich leaves over moderate-phenolic wild type poplar leaves (Boeckler et al., 2014). Consequently, the reason behind the downregulation of phenolic pathways in leaves is most probably the better nitrogen allocation caused by the EMF inoculation. Various studies showed that high nitrogen availability leads to a significant decrease in total phenolics as plants are in urgent need of carbon skeletons for nitrogen fixation (Muzika, 1993; Coviella et al., 2002; Keski-Saari and Julkunen-Tiitto, 2003; Ruan et al., 2010). Indeed, higher amounts of CHNO compounds were found in leaves of mycorrhized poplars (Fig. 5.4). Moreover, beetle attack on EMF inoculated plants facilitated a transcriptional induction of nitrate transporter (*nrt1*), nitrite reductase (*nia*) and of glutamate dehydrogenase 2 (*gdh2*) which are important enzymes in nitrogen allocation (Fig. 5.10; Sakakibara et al., 2006).

Additionally, transcript levels of amino acid transporters and nicotianamine synthase 3 (*nas3*) were upregulated under herbivory-free conditions (Fig. 5.10). Nicotianamine is an important metal chelating compound transporting Fe, Zn and Mg needed as Co-factors for protein biosynthesis (von Wirén et al., 1999; Takahashi et al., 2003; Zheng et al., 2010). Excelling metal-ion transportation, nitrogen allocation and amino acid biosynthesis most probably facilitates an improved induction of protein biosynthesis under herbivory as seen by

the higher amounts of transcripts coding for defensive-related proteins such as kunitz protease inhibitors (KPIs) and chitinases, in mycorrhized plants (Fig. 5.10). Protease inhibitors complicate the digestion of plant-derived proteins in the herbivore's gut by inhibiting their proteases (Philippe and Bohlmann, 2007). Previous publications showed the wound-induced accumulation of chitinases in poplar (Collinge et al., 1993; Clarke et al., 1998). These enzymes hydrolyze the glycosidic bonds of chitin, the building block of an insect derived biopolymer, and therefore function as defensive proteins. For example, the expression of poplar chitinase in tomato inhibited the development of potato beetles (Lawrence and Novak, 2006)

Leaves of mycorrhized poplars possessed also higher transcript abundances of glutathione S-transferases (GST; Fig. 5.10). Those enzymes are involved in detoxification processes introducing a glutathione moiety to xenobiotics facilitating their degradation (Edwards et al., 2000).

Improved nitrogen allocation in leaves of EMF inoculated plants might also be the reason for the higher transcript abundance of three P450 monooxygenases, which catalyze the formation of volatile nitriles from aldoximes, when herbivory took place. VOC analysis revealed the induction of the volatile aromatic nitrile phenylacetoneitrile during herbivory. These findings were supported by the MDEA results, which displayed an overrepresentation of four different nitrile REMDs pertaining to the KEGG cyanoamino acid metabolism. Generally nitriles are minor components of the herbivore-induced volatile blend in grey poplar. Moreover, grey poplars are also very weak nitrile/aldoxime emitters compared to black poplar (Irmisch et al., 2013). However, nitriles and aldoximes are very effective as direct herbivore repellents and as attractants of natural herbivore enemies (Irmisch et al., 2013; Irmisch et al., 2014).

Monoterpenes, sesquiterpenes, 2-hexenal and salicylaldehyde were amongst the other herbivore-induced plant volatiles (Fig. 5.10). Those compounds are usually always emitted upon herbivore infestation to prime other leaves' defense systems and attract natural herbivore enemies (Arimura et al., 2005; Clavijo McCormick et al., 2012).

Additionally increased transcript abundances of isoprene synthase (*isps*) were found in beetle infested and mycorrhized plants. Isoprene has previously been shown to interfere tri-trophic plant insect interactions repelling parasitic wasps searching for their insect host (Loivamäki et al., 2008). Chapter 4 already dealt with the question whether isoprene is involved in poplar-insect communication (Müller et al., 2015). An important finding of this work was that *C. populi* is unable to sense isoprene (Müller et al., 2015). Consequently, it is questionable whether the transcriptional induction of *isps* is crucial for direct insect repelling purposes. As described in the introduction the function of isoprene is still under debate. In the present context two functions of isoprene are possible.

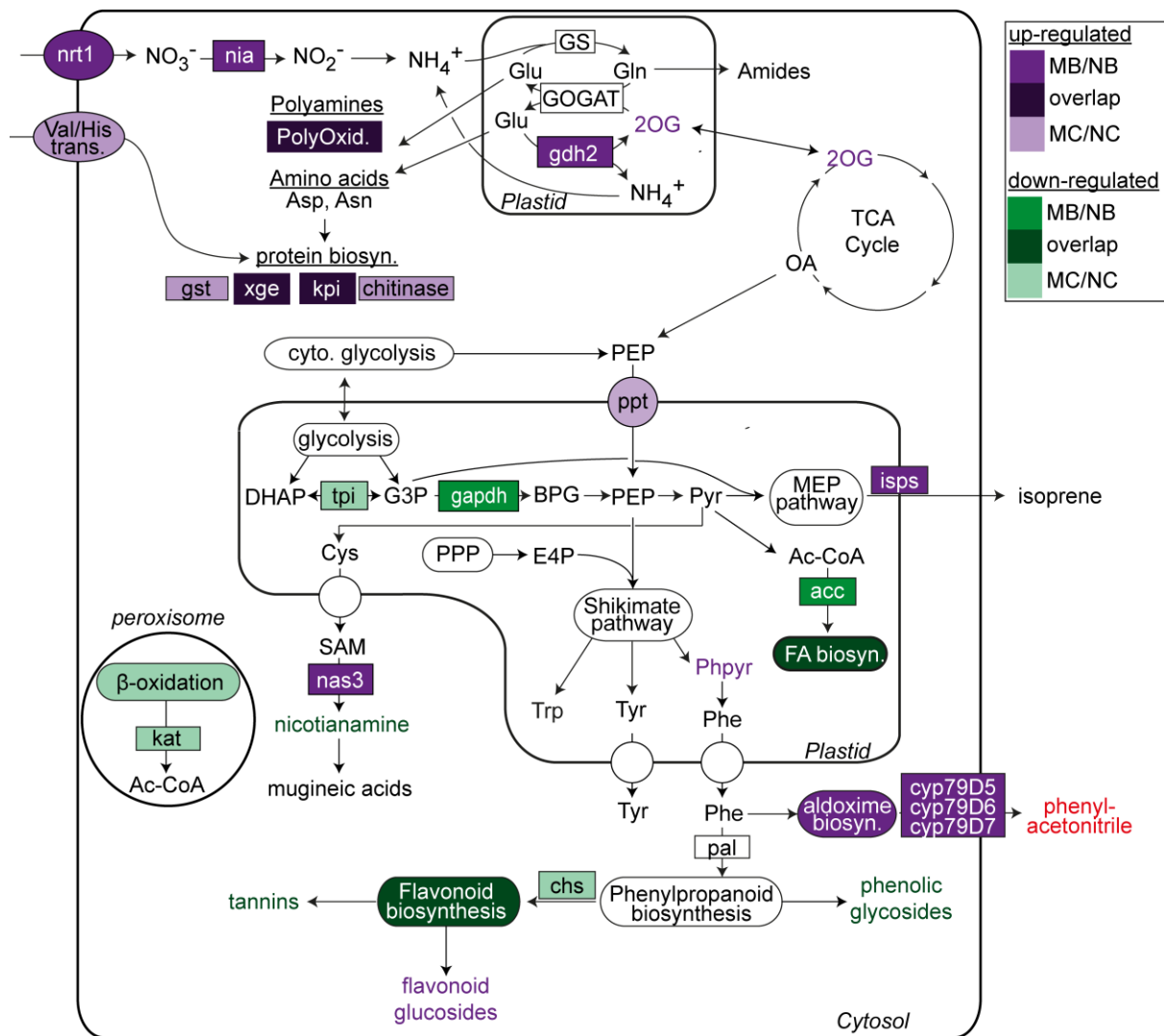


Figure 5.10. Illustration of systemic changes resulting from EMF inoculation within a poplar cell. Transcriptional changes of enzymes are represented in colored boxes and of transmembrane transporters in colored circles, biosynthetic routes are illustrated in rounded rectangles that are colored according to enzymatic and metabolic regulation patterns found within them. Metabolites and compound classes are also colored according to their respective regulation pattern. Metabolites which are colored in red were detected by GC/MS measurements. 2OG, 2-oxoglutarate; acc, acetyl-CoA carboxylase; Ac-CoA, acetyl-CoA; Asn, asparagine; Asp, aspartic acid; chs, chalcone synthase; Cys, cysteine; cyto, cytosol; DHAP, dihydroxyacetone phosphate; E4P, erythrose 4-phosphate; FA, fatty acid; G3P, glyceraldehyde 3-phosphate; gapdh, glyceraldehyde-3-phosphate dehydrogenase; gdh2, glutamate dehydrogenase 2; GOGAT, glutamine oxoglutarate aminotransferase; GS, glutamate synthase; gst, glutathione S-transferase; Gln, glutamine; Glu, glutamic acid; isps, isoprene synthase; kat, 3-ketoacyl-CoA thiolase; kpi, kunitz protease inhibitor; nas3, nicotianamine synthase; nia, nitrite reductase; nrt1, nitrate transporter; OA, oxaloacetate; pal, phenylalanine ammonia-lyase; PEP, phospho*enol*pyruvate; PPP, pentose phosphate pathway; PPT, phospho*enol*pyruvate/phosphate translocator; Phpyr, phenylpyruvate; PolyOxid, polyamine oxidase; SAM: S-adenoyl-L-methionine; tpi, triosephosphate isomerase; Trp, tryptophane; Tyr, tyrosine; XGE, xyloglucan endotransglucosylase; biosyn, biosynthesis; trans, transporter.

First, is the ability of isoprene to influence the NO-signaling cascade of poplar by impairing the S-nitrosylation of proteins (Vanzo, 2015). The post-translational modification of cysteine moieties is a central component that triggers plant responses to oxidative stress (Moreau et al., 2010). Second, isoprene emission regulates the C-flux through the MEP pathway. To synthesize insect repellants such as SQTs and MTs an increase in MEP-C-flux is needed. Higher isoprene emission might be crucial in order to prevent plastidic DMAPP accumulation that would inhibit DXS through feedback regulation (Ghirardo et al., 2014). An improved herbivore defense through mycorrhization without any trade-offs in plant performance was not found as young mycorrhized poplars showed lower biomass production (Müller, 2014). A reduction in plant growth is a common trait in AMF inoculated plants (Schweiger et al., 2014).

In summary, mycorrhization of poplar roots with EMF resulted in a transcriptional and metabolic adjustment of defense mechanisms consisting of volatiles, protease inhibitors and plant phenolics in the plants' foliage making leaves more effective in their defense against leaf beetles. The boost in nitrogen allocation enables poplars to increase the production of protease inhibitors and volatile nitriles and lower the production of plant phenolics in leaves in order to better fight specialist herbivores such as *C. populi*.

As described in Chapter 3, REMDs were introduced to provide a framework for a better integration of mass-spectrometry based metabolomic data into systems biology. MDEA analysis enabled the selection of important REMDs linking the metabo- to the phenotype. The EMF context pushes mass-difference network analysis by means of REMDs and MDEA one step further, as on the one hand the technique was used on MS data with lower resolution (Forcisi et al., 2015) and on the other hand the overrepresented REMDs were directly mapped on significant transcripts. This does not only validate the connection of REMDs with other omics domains but also solidifies the improved integration of metabolites into systems biology. The improvement does also consist in a complementation of regulatory transcriptional/metabolic process, e.g. transcription factors.

Alterations happening in the flavonoid biosynthesis could be integrated by MDiN as the hydroxylation and hydrolysis REMDs were mapped on transcripts coding for *ban*, *f3h*, *fls*, *dfr* and *tt7* which were connected to the flavonoids quercetin, naringenin, dihydrokaempferol, dihydroquercetin, catechin and leucocyanidin (Fig. 5.7). This might facilitate the assignment of gene functions as MDiN patterns of node and edge significant might ensure the functional embedment of significant genes or transcripts.

The applicability of MDiNs for cross-MS data matching was also shown to be a valuable tool. Moritz et al. (2015) already dealt with the assignment of intracellular precursors of volatiles detected in human breath condensate. Here, the matching of GC/MS and LC/MS data via MDiNs resulted in that the majority of VOCs was in direct connection to their cellular

biosynthetic precursor allowing a global systemic view on metabolism. This comprehensive MDiN is analyzable by means of MDEA in which overrepresented REMDs account for metabolic alterations and establish a direct relation with the remaining –omic disciplines.

6 Conclusions and outlook

Future climate scenarios predict a higher frequency of drought and heat periods that are projected to impact tree mortality (Allen et al., 2010). Consequently, to cover the increasing demand of renewable energies in the future, a higher production of biomass with a simultaneously increased resistance against environmental stresses is needed. Here, one of the most important trees used for the production of bioenergy, namely grey poplar (*Populus x canescense*), was studied by means of non-targeted metabolomics.

It turned out that non-targeted metabolomics using modern mass spectrometers that are suitable for high-throughput analysis, namely the ICR-FT/MS and the UPLC-qToF-MS, have an enormous potential to unravel molecular mechanisms responsible for plants' environmental adaptation. The metabolic alterations of IE and NE poplar genotypes caused by UV-B exposure (Chapter 2) as well as by beetle infestation (Chapter 4) could be explained at hand of non-targeted ICR-FT/MS measurements. Additionally, it was found that metabolomic experiments significantly improved the knowledge about tripartite interactions between poplar, *L. bicolor* and *C. populi*. The mycorrhization of roots with EMF increased the defense mechanisms of poplars as seen by the lower leaf damage caused by *C. populi*. and reduced oviposition. However, young EMF-inoculated poplars did produce less biomass as compared to non-mycorrhized plants. For that reason, field studies needs to be performed in order to study whether the strategic inoculation of poplars in plantations increases biomass production as compared to non-inoculated poplar plantations.

The common limitation in plant metabolomic experiments is the identification of compounds. In particular the low database coverage is accountable for the incompleteness of metabolic descriptions as only the small fraction of m/z features that is found within databases is used for its explanation. This is a major drawback because plant metabolomes belong to the most complex and diverse chemical systems in biology (Grotewold, 2005). As illustrated herein, mass difference analysis overcomes this typical limitation as the statistical information of each m/z feature is directly translated to its corresponding REMDs, which likewise are representative for reactions as well as their involved genes and enzymes.

This approach was validated in Chapter 3, where the most significant REMDs could directly link the metabotype to the PcISPS knockdown, as well as in Chapter 4, in which

significant transcripts could directly be mapped on their respective REMDs. Consequently, MDiNs pave the way for an improved integration of metabolomics into plant systems biology.

Appendix

Methods

Metabolite extraction procedures

The leaf material was grinded under liquid nitrogen to a fine powder and stored at -80°C. From each sample 5 mg (Chapter 2) or 20 mg (Chapter 4) of powdered leaf material was extracted with 1 ml -20°C extraction solvent consisting of methanol/isopropanol/water with a ratio of 1:1:1 (v/v/v) in an ultrasonic bath for 15 min. Subsequently the solution was centrifuged at 10,000 g for 10 min at 4°C. The supernatant was removed and diluted with extraction solvent in a ratio of 1:25 (v/v, Chapter 2) 1:12.5 (v/v, Chapter 4) and directly injected into the ICR-FT/MS.

For UPLC-qToF-MS measurements, 50 mg of powdered leaf material was extracted two consecutive times with one ml of -20°C extraction solvent consisting of methanol/water (8:2 [v/v]) in an ultrasonic bath filled with ice water for 15 min (the decision to change the extraction solvent was made because it lowered the drying time in the speed-vac). Subsequently, the solution was centrifuged at 10,000 g for 10 min at 4°C. A total of 1.5 ml of supernatant was removed and divided into two aliquotes of 750 µl samples. For quality control (QC) 20 µl of each extract were taken and combined. One aliquote was used for direct injection DI-ICR-FT/MS measurements. From the second aliquote the extraction solvent was removed in a speed-vac and stored at -80°C for further analysis. Prior to the LC-MS measurements, the dried samples were resolved in 500 µl of 20% acetonitrile in water and centrifuged at 14,000 rpm at 4°C for 10 mins.

ICR-FT/MS measurements

Chapter 2 & 4

For non-targeted metabolome analysis, ultra-high-resolution mass spectra were acquired using a ICR-FT/MS (APEX Qe, Bruker, Bremen, Germany) equipped with a 12-Tesla superconducting magnet and an APOLLO II electrospray (ESI) source. Samples were introduced into the microelectrospray source (Agilent sprayer, Waldbronn, Germany) at a flow rate of 120 µl h⁻¹ with a nebulizer gas pressure of 20 psi and a drying gas pressure of 15 psi at 180°C using a Gilson autosampler (Sample changer 223, Gilson Inc., Middleton, USA). All measurements were performed in the negative ionization mode over a mass range of m/z 100 – 1000 in Chapter 2 or 100 – 1200 in Chapter 4. The spectra were acquired over

a time domain transient of 2 Megawords with 438 (Chapter 2) or 488 (Chapter 4) accumulated scans for each sample. To exclude memory effects in the electrospray source, samples were randomized, and the mass spectrometer was cleaned with 50% methanol solution between each run for seven minutes at a high flow rate of 300 $\mu\text{L}/\text{min}$.

Chapter 3

One IE leaf and one NE leaf were investigated by fragmentation experiments that were performed using the multiple adjacent selected ion monitoring (SIM) method (Southam et al., 2007). ESI parameters were set as described in Chapter 2&4. The spectra were acquired over a time domain transient of 4 Megawords and an Ion-accumulation time of 1.3 s. The SIM window size was set to 30 Da. The SIM window was first centered around 260 m/z and was then shifted towards 440 m/z values in 15 Da steps (13 windows). Fragmentation of each SIM window was performed with four different fragmentation energies: (i) 0 eV, (ii) 5 eV, (iii) 10 eV and (iv) 15 eV. Each spectrum was acquired for 56 scans.

Chapter 5

Ultra-high resolution mass spectra were acquired using a Fourier transform ion cyclotron resonance mass spectrometer (ICR-FT/MS, APEX Qe, Bruker, Bremen, Germany) equipped with a 12-Tesla superconducting magnet and an APOLLO II electrospray (ESI) source. Samples were introduced into the microelectrospray source (Agilent sprayer, Waldbronn, Germany) at a flow rate $120 \mu\text{L h}^{-1}$ with a nebulizer gas pressure of 20 psi and a drying gas pressure of 15 psi at 200 °C by a Gilson autosampler (Sample changer 223, Gilson Inc., Middleton, USA). All measurements were performed in the negative ionization mode over a mass range of m/z 100 – 1100. The spectra were acquired over a time domain transient of 4 Megawords with 400 accumulated scans for each sample. To exclude memory effects in the ICR cell, samples were randomized and the mass spectrometer was cleaned with Methanol between each run for 7 consecutive minutes. The mass spectrometer was calibrated on arginine clusters.

ICR-FT/MS data analysis

Chapter 2

The mass spectra were internally calibrated and exported to peak height lists as ascii files at a signal to noise ratio of 2 using the Data Analysis 4.0 software package (Bruker, Bremen, Germany). The peak lists were combined to a peak matrix with an error of 1.5 ppm using an in-house written tool (Lucio et al., 2011). Peaks with just 1 non-zero intensity (single mass events) were removed from the matrix as well as peaks that were detected in less than 50%

of all biological replicates. After that, a ^{13}C isotopic peak filter was applied, deleting peaks with no corresponding ^{13}C isotopic peak to avoid signals generated by electrical noise. Overall 4412 masses remained after all filtration processes (Kaling et al., 2015).

Chapter 3

The SIM-MS/MS data was internally calibrated, aligned and exported to peak height lists as ascii files at a S/N-ratio of 4 and a relative threshold intensity of 0 using the Data Analysis 4.3 software (Bruker, Bremen, Germany). Annotation was performed as described in chapter 3. Each spectrum was divided into parent and daughter sections. MDiNs were created using KEGG rpairs and the manually curated REMD list. For MDEA, only valid parent→daughter (P→D) pairs per spectrum were considered. MDEA variables were generated as follows: E: The sum of all REMDs that were P→D pairs was calculated, both for NE and IE. R: The sum of all REMD that were P→D pairs with P being a marker for either NE or IE was calculated. ES: The frequency of P→D pairs for each REMD was calculated. RS: The frequency of P→D pairs with P being a marker was calculated for each REMD.

Chapter 4

For post-processing the measured mass spectra were aligned, internally calibrated and exported to peak height lists as ascii files at a signal to noise ratio of 4 and a relative threshold intensity of $1 \cdot 10^{-8}\%$ using the Data Analysis 4.1 software package (Bruker, Bremen, Germany). Peak lists were combined to a peak matrix with an error of 1 ppm using an in-house written tool (Lucio et al., 2011). Single mass events, isotopic peaks (^{13}C , ^{15}N , ^{34}S and ^{18}O), side-peaks, peaks with an unusual absolute mass-defect and peaks detected in less than 50% of the biological samples were removed from the matrix.

The zero intensities present in the peak matrix were replaced with the ICR-FT/MS threshold intensity of $1 \cdot 10^6$ counts. Prior to the statistical analysis, peak intensities were \log_2 transformed and the poplar plants were grouped to their respective isoprene-emission capacity in IE (WT/GUS) and NE (RA1/RA22) (Müller et al., 2015).

Chapter 5

The ICR-FT/MS mass spectra were aligned, internally calibrated and exported to peak height lists in the ascii format with a signal to noise ratio of 4 using the Data Analysis 4.2 software package (Bruker, Bremen, Germany). The mass lists were combined to a mass/intensity matrix with an error window of 1 ppm using an in-house written tool (Lucio et al., 2011). Single mass events and peaks that were present only once in each of the four different experimental conditions were removed from the matrix. In the next step, unusual mass-defects and isotopic peaks (^{13}C , ^{15}N , ^{34}S and ^{18}O) were removed from the matrix.

For sum formula annotation to the unknown m/z-features, the filtered mass-list was subjected to the mass-difference based NetCalc algorithm of Tziotis et al., (2011). The NetCalc annotation procedure was repeated 10 times. Sum formulas which were annotated in each individual run were used for LC-MS matching.

UPLC-qToF-MS measurements

LC-MS measurements were performed on a Waters Acquity UPLC System (Waters GmbH, Eschborn, Germany) coupled to a Bruker maXis ToF-MS (Bruker Daltonic, Bremen, Germany). The chromatographic separation was achieved on a Grace Vision HT C18-HL column 150 mm x 2 mm inner diameter with 1.5 μm particles (W.R. Grace and Company, Maryland, USA).

Eluent A was 5% acetonitrile in water with 0.1% of formic acid and eluent B was acetonitrile with 0.1% of formic acid. The gradient elution started with an initial isocratic hold of 0.5% B for 1 minute, followed by a linear increase to 99.5% B in 5.4 minutes and a further isocratic step of 99.5% B for 3.6 minutes. In 0.5 min the initial conditions of 0.5% B were restored. To equilibrate the initial column conditions, 0.5% B was hold for 5 minutes. The flowrate was 400 $\mu\text{l}/\text{min}$ and the column temperature was set at 40°C. The auto-sampler was set to 4°C. From each sample two technical replicates were measured in both the positive and the negative ionization mode. Prior to sample analysis 10 QC samples were injected for column conditioning. Every 10th measurement QC and blank samples were injected. Mass calibration was achieved with low concentration ESI Tuning Mix (Agilent, Waldbronn, Germany).

The MS was operated as follows: nebulizer pressure was set to 2 bar, dry gas flow was 8 l/min, dry gas temperature was 200°C, capillary voltage was set to 4000 V and the end plate offset was -500 V. Mass spectra were acquired in a mass range of 50-1100 m/z.

UPLC-qToF-MS data analysis

The LC-MS spectra were internally calibrated with ESI tune mix. Each Bruker spectrum file was separately imported into the GeneData Refiner MS software. First, a chemical noise reduction was applied. Secondly, a retention time (RT) alignment was performed. Thirdly, m/z features were identified using the summed-peak-detection feature implemented in the GeneData software. Fourthly, only peaks which were present in at least 10% of mass spectra were used for isotope clustering. The resulting peak matrix was exported and used for further processing steps.

Unfortunately the statistical analysis of raw MS data (NetCalc sum formula annotations, the workflow and results are explained in Chapter 5) revealed severe batch effects as well as variations between the technical replicates. To adjust for these deviations, cyclic loess normalization was applied. Therefore, peak intensities were log₂ transformed and imported into R to perform cyclic loess normalization according to Ejigu et al., (2013). Therefore the following R packages are needed:

- Bioconductor package (<http://www.bioconductor.org/>)
- Library(affy)

Following parameters were used for the normalization (Ejigu et al., 2013):

- loess.data = normalize.loess(data, epsilon = 0.05, maxit = 2, log.it = FALSE, verbose = TRUE, span = 0.75, family.loess = "gaussian")

Two normalization cycles (maxit=2) were needed in order to eliminate the batch effects of the negative mass spectra and one cycle (maxit=1) was needed to eliminate the batch effect of the positive measurements.

Afterwards the average peak intensities of the technical replicates were calculated. Subsequently the averaged matrix was imported into SIMCA to perform multivariate statistical analysis. Discriminant features were determined by OPLS-DA analysis. For each experimental condition, namely MC/NC, MB/NB, NB/NC and MB/MC, separate OPLS-DA models were calculated, in which either mycorrhization or beetle infestation leaf-age was subjected as Y-variables (e.g. for mycorrhization: Y(inoculated) = 1, Y(non-inoculated) = 0). Mass features with a variable influence of projection score (VIP) > 1 and an acceptable cross-validation standard error (cvSE) were extracted. When these features possessed a log₂-fold change >1 or <-1 they were considered as discriminant.

Instruments

ICR-FT/MS

- 12 Tesla solariX - Bruker, Bremen, Germany
- Gilson autosampler – Gilson Inc., Middleton, USA

UPLC-qToF-MS

- Aquity UPLC system – Waters, Milford, USA
- maXis 3G UHR-ToF-MS - Bruker, Bremen, Germany

Speed-Vac

- Savant SPD121P SpeedVac - Thermo Scientific

Centrifuges

- Benchtop centrifuge 5804R - Eppendorf
- Benchtop centrifuge Z 233 MK - Hermle

Ultrasonic baths

- Sonorex Super, 10P
- Sonorex Super, RK 255 H

Balances

- Sartorius analytic, A 120 S
- Kern, ABT120-4M

Vortex

- Vortex Genie 2 - Bender and Hobein AG

Chemicals

Formic acid	Fluka
ESI Tune-mix	Agilent

Metabolite standards

<u>Name</u>	<u>Company</u>
2',4',6',3,4-Pentahydroxychalcone	Sigma
7,4-Dihydroxyflavanon	Roth
Apigenin	Extra
Apigenin-7-glucoside	Roth
Apiin	Roth
Auxin	Sigma
Caffeic acid	Roth
Catechine	Sigma
Dihydrokaempferol (DHK)	synthetisch
Dihydroquercetin (DHQ)/Taxifolin	Roth

Gallocatechine-gallate	Sigma
Hesperidin	Sigma
Indole 2-carboxylic acid	Sigma
Isoquercitrin	Roth
Isorhamnetin	Roth
Kaempferol	Roth
Kaempferide	Roth
Luteolin	Roth
Luteolin-7-glucoside	Roth
Myricetin	Roth
Naringenin	Serva
Naringenin-7-glucoside	Roth
Orientin	Extra
p-Coumaric acid	Sigma
Procyanidine B1	Sigma
Procyanidine B2	Sigma
Quercetin	Fluka
Que-3-galactoside	Roth
Quercetagetin	Roth
Quercetin-3-arabinoside	Roth
Quercetin-3-rhamnoside	Sigma
Rhamnetin	Roth
Rutin	Serva
Salicylic acid	Sigma
Shikimic acid	Roth

Solvents

<u>Solvent name</u>	<u>Description</u>
Methanol	LC-MS CHROMASOLV, Fluka
Isopropanol	LC-MS CHROMASOLV, Fluka
Acetonitrile	LC-MS CHROMASOLV, Fluka
Water	Merck Milli-Q® Integral water purification system

Software

DataAnalysis v4.0-v4.2	Bruker, Bremen, Germany
Gephi 0.8.2beta	Open Source software: http://gephi.github.io/
HCE 3.5	Open Source software: http://www.cs.umd.edu/hcil/hce/
Matlab R2011b	MathWorks, Ismaning, Germany
Matrix Generator	In-house written tool (Lucio et al., 2011)
Excel 2010	Microsoft, Redmond WA, USA
NetCalc	In-house written tool (Tziotis et al., 2011)
R 3.2	Open Source software: https://www.r-project.org/
SigmaPlot 12.0	Systat Software Inc., San Jose, CA USA
SIMCA-P 13.0	Umetrics, Umeå, Sweden
Venny 2.0.2	http://bioinfogp.cnb.csic.es/tools/venny/

Supplemental Material

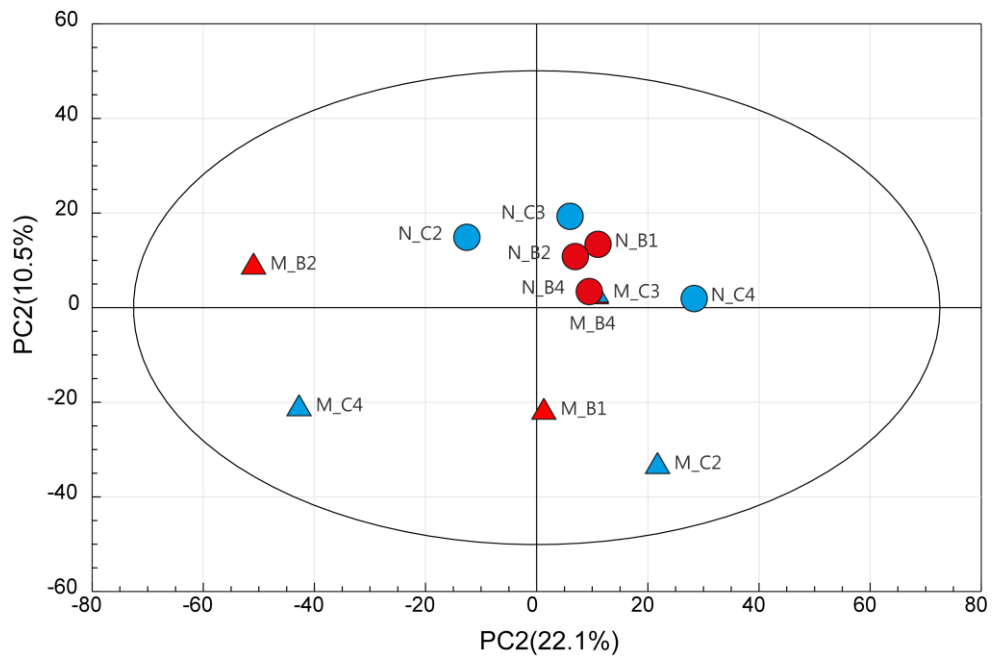


Figure S1. PC2xPC3 score plot of (+)LC-MS data analyzed in Chapter 5.

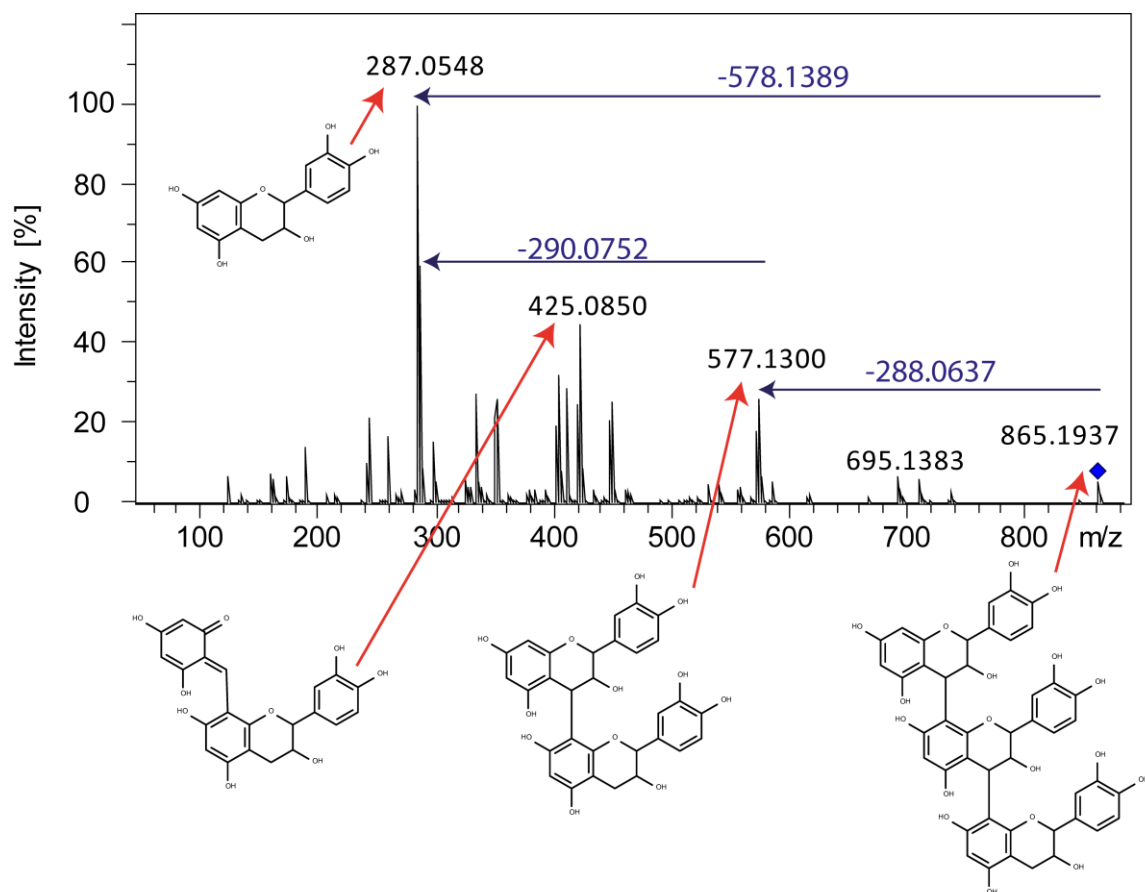


Figure S2. MS/MS spectrum of proanthocyanidin trimers (fragmentation mechanisms was proposed by Callemien and Collin, 2008).

Table S1. Measured LC-MS standards.

Name	Sum Formula	Monoisotopic mass	[M-H]-	[M+H]+	RT [min]
2',4',6',3,4-OH-chalcone	C15H12O6	288.06339	287.056114	289.070666	4
7,4-Dihydroxyflavanon	C15H10O4	254.05791	253.050634	255.065186	3.9
Apigenin	C15H10O5	270.052825	269.045549	271.060101	4.5
Apigenin-7-glucoside	C21H20O10	432.10565	431.098374	433.112926	3.6
Apiin	C26H28O14	564.14791	563.140634	565.155186	3.7
Auxin	C10H9NO2	175.063329	174.056053	176.070605	4
Caffeic acid	C9H8O4	180.04226	179.034984	181.049536	3.1
Catechine	C15H14O6	290.07904	289.071764	291.086316	3.5
Dihydrokaempferol	C15H12O6	288.0634	287.0561	289.0707	3.9
Dihydroquercetin	C15H12O7	304.0583	303.0510	305.0656	3.6
Gallocatechine-gallate	C22H18O10	442.09	441.082724	443.097276	3.3
Hesperidin	C28H34O15	610.189775	609.182499	611.197051	3.7
Indole 2-carboxylic acid	C9H7NO2	161.047679	160.040403	162.054955	4.2
Isoquercitrin	C21H20O12	464.09548	463.088204	465.102756	4
Isorhamnetin	C16H12O7	316.058305	315.051029	317.065581	4.6
Kaempferol	C15H10O6	286.04774	285.040464	287.055016	4.5
Kaempferide	C16H12O6	300.0634	299.0561	301.0707	5.1
Luteolin	C15H10O6	286.04774	285.040464	287.055016	4.1
Luteolin-7-glucoside	C21H20O11	448.100565	447.093289	449.107841	3.5
Myricetin	C15H10O8	318.03757	317.030294	319.044846	3.4
Naringenin	C15H12O5	272.068475	271.061199	273.075751	3.6
Naringenin-7-glucoside	C21H22O10	434.1213	433.114024	435.128576	3.7
Orientin	C21H20O11	448.100565	447.093289	449.107841	3.2
p-Coumaric acid	C9H8O3	164.047345	163.040069	165.054621	3.5
Procyanidine B1	C30H26O12	578.14243	577.135154	579.149706	2.9
Procyanidine B2	C30H26O12	578.14243	577.135154	579.149706	3.1
Que-3-galactoside	C21H20O12	464.09548	463.088204	465.102756	3.5
Quercetagetin	C15H10O8	318.03757	317.030294	319.044846	3.3
Quercetin-3-arabinoside	C20H18O11	434.084915	433.077639	435.092191	3.6
Quercetin-3-rhamnoside	C21H20O11	448.100565	447.093289	449.107841	3.6
Rhamnetin	C16H12O7	316.058305	315.051029	317.065581	4.7
Rutin	C27H30O16	610.15339	609.146114	611.160666	3.3
Salicylic acid	C7H6O3	138.031695	137.024419	139.038971	4
Shikimic acid	C7H10O5	174.052825	173.045549	175.060101	1

Table S2. REMDs overrepresented in inoculated poplars in the MC/NC comparison.

Description	Reaction Class	Mass Difference	Z-Score(MC)
Serine	C	89.047679	7.15
Pteridine reaction	Kegg	195.075625	6.24
Amination	F/Kegg	15.010899	4.73
Two step Deamination/Transamination	F/Kegg	0.984016	4.71
Oxolysis of a C2-moiety yielding Glyoxylate and Water	Kegg	220.2191	4.67
Pteridine reaction	Kegg	177.06506	4.63
Adenosine condensation	F/Kegg	249.08619	4.23
Deamination followed by hydroxylation	Kegg	16.978931	3.99
Leucine/Isoleucine	A	113.084064	3.98
Choline condensation	Kegg	85.089149	3.95
Asparagine	A	114.042928	3.81
Acetone Cyanohydrin attachment	Kegg	67.042199	3.80
Hexadecanoic acid	A/Kegg	238.229665	3.65
5-Amino-2-oxopentanoic acid	B	69.057849	3.58
Indole pyruvic acid	B	141.057849	3.48
Adenine condensation	F/Kegg	117.04393	3.46
Glutamic acid	C/Kegg	131.058244	3.44
Indole pyruvic acid	D	159.068414	3.36
Aspartic acid	C/Kegg	117.042594	3.35
Valine	B	55.078584	3.23
Alanine	B	27.047284	3.22
Propionitrile transfer	Kegg	52.042533	3.19
Hydroxymandelonitrile condensation	Kegg	131.037114	3.18
Lysine	A	128.094963	3.14
Leucine/Isoleucine	B	69.094234	3.14
Histidine	A	137.058912	3.08
2-Amino adipate condensation	Kegg	143.058244	3.05
Pyridoxal	F	149.047679	3.03
Addition of 2-Amino adipate	Kegg	129.078979	3.00
(9Z,12Z,15Z)-Octadecatrienoic acid condensation	Kegg	260.214015	2.99
5-Amino-2-oxopentanoic acid	D	87.068414	2.97
2-Oxosuccinamic acid	B	69.021464	2.95
Glutamic acid	B	85.052764	2.93
Alanine	C	73.052764	2.91
Proline	B	53.062934	2.90
NO• - H +neutral (nitrosylation)	F/Kegg	28.990164	2.89
Indole, Pyruvate, Ammonia	F/Kegg	73.016379	2.86
Hydro-peroxidation	F/Kegg	31.98983	2.79
Arginine	B	112.111281	2.76
Pyridoxamine	F	150.079313	2.74
Hexadecanoic acid	B/Kegg	194.239835	2.72
Pyridoxine	F	151.063329	2.72
Dodecanoic acid	B	138.177235	2.72
Arachidonic acid condensation	Kegg	286.229665	2.62

Niacinamide	F/Kegg	104.037448	2.55
6-Amino-2-oxohecanoic acid	B	83.073499	2.52
Guanosine condensation	F	265.081105	2.50
Guanine condensation	F	133.038845	2.50
Aminobutyraldehyde condensation	Kegg	71.073499	2.49
Proline	A	97.052764	2.44
Niacine	F/Kegg	105.021464	2.44
Threonine	A/Kegg	101.047679	2.39
Imidazole pyruvic acid	A	136.027278	2.38
Corresponds the disproportionation of a carotenoid	Kegg	218.167065	2.33
Aspartate-4-semialdehyde condensation	Kegg	99.032029	2.31
Indole pyruvic acid	A	185.047679	2.30
Glutamine	B	84.068748	2.30
Tetradecanoic acid	A/Kegg	210.198365	2.28
Amination of Enol	Kegg	17.026549	2.26
Tryptophan	B	142.089483	2.20
Serine	A/Kegg	87.032029	2.19
Transsulfonierung (OH replaced by HSO₃)	Kegg	63.961901	2.19
Aminobenzoate condensation	Kegg	119.037114	2.18
Imidazole pyruvic acid	D	110.048013	2.08
Valine or proline	C	99.068414	2.08
2-Keto-glutaramic acid	B	83.037114	2.05
Eicosatrienoic acid condensation	Kegg	288.245315	2.04
Glycine	A/Kegg	57.021464	2.02

List of figures

Figure 1.1 Chapter overview of the thesis.	2
Figure 1.2. Cellular flow of information.	4
Figure 1.3. Pubmed search results for the term meabolomics (http://www.ncbi.nlm.nih.gov/pubmed).	5
Figure 1.4. Overview of changes in the metabolism of <i>P. x canescens</i> resulting from PclSPS knockdown by RNAi.	9
Figure 1.5. The general workflow of metabolomic experiments can be separated into three steps.	13
Figure 1.6. Representation of an ICR-FT/MS. (A) Picture of the 12 Tesla ICR-FT/MS spectrometer used in this thesis. (B) ICR-FT/MS spectrum of a grey poplar leaf extract acquired at 2 Megawatts.	17
Figure 1.7. Representation of a ToF-MS. (A) Picture of the qToF-MS system used in this thesis. (B) Illustration of ion guidance through the qToF-MS.	18
Figure 1.8. Schematic representation of Principal Component Analysis (PCA). Starting with (A) the data-matrix and (B) its projection into principle components (PC) as well as (C) the resulting score and loading plots. (modified according to Boccard et al., 2010).....	20
Figure 2.1. Experimental setup of the phytotron experiment. (A) Temperature, relative humidity, photosynthetic active radiation (PAR) and UV-B conditions, (B) leaf harvest time points are indicated by red arrows and (C) emerging leaf phenotypes after 300 hours of UV exposure (modified according to Kaling et al., 2015).....	23
Figure 2.2. Results obtained by ICR-FT/MS measurements. (A) PCA of raw ICR-FT/MS data colored according to batches, (B) total ion current (TIC) plotted as a function of sample order, (C) PCA scatter plot of TIC-normalized and filtered ICR-FT/MS data colored according to batched and (D) according to isoprene emission capacity.....	25
Figure 2.3. Statistical analysis of normalized ICR-FT/MS data. (A) PCA score plot of peak intensities the arrows indicate the time trend of the data and (B) Venn-plot of masses discriminant for the different experimental variables i.e. line, UV-exposure and ageing. The number of annotated metabolites is written in brackets (modified according to Kaling et al., 2015).	26
Figure 2.4. Database annotations for masses discriminant for either line- or the UV-B effect. The annotations were arranged according to their respective compound classes and KEGG metabolic pathways. Annotations were not corrected for multiple pathway entries (modified according to Kaling et al., 2015).....	28
Figure 2.5. Clustermap of masses discriminant for both, isoprene emission capacity and UV-B exposure. Numbers indicate the sampling order as follows: 2 = 12h, 3 = 36, 4 = 60 h, 5 = 108 h, 6 = 209 h and 7 = 300 h of UV-B exposure.	29
Figure 2.6. Time behavior of normalized signal intensities of all discriminant masses. The masses cluster in four major time profiles: up-regulated metabolites in (A) NE (Ra2/Rb7), (B) in IE (WT/EV), (C) metabolites induced under UV-B irradiation and (D) metabolites decreasing within the first 12 hours of UV-B exposure. Red lines indicate the time profile entered into the profile search function of HCE 3.5 and grey lines masses, which matched the entered profile. The number of masses within a cluster is given in the arrows. Numbers in brackets stand for the amount of annotations. The annotations were grouped into compound classes and visualized in bar plots. Flav: flavonoids; phenylpr: phenylpropanoids;	

phen: phenolic; carbohyd: carbohydrates (reprinted with permission from John Wiley and Sons).	31
Figure 2.7. Map of the shikimate, phenylpropanoid and flavonoid pathways/biosynthesis and their respective compartmentalization. Ac, acetyl; Cat, catechin; downreg, downregulated; glu, glucose; Hy, hydroxyl; Kae, kaempferol; Me, methoxy; Que, quercetin; Rha, rhamnoside; Rut, rutenoside; upreg, upregulated (modified and reprinted with permission from John Wiley and Sons).	34
Figure 3.1. The theory of mass difference networks. (A) Conversion of a biochemical reaction into network nodes and edges. (B) Conversion of a mass spectrum into a network.	38
Figure 3.2. Examples of biochemical reactions mimicked by REMDs.....	39
Figure 3.3. Statistical analyses of sum formula annotations from NetCalc. (A) Score plot of OPLS-DA model, (B) loading plot showing the selection of discriminant sum formulas based on (C) the top 10% of NE and IE respectively and (D) elemental compositions of discriminant sum formulas.	42
Figure 3.4. Scheme of the mass difference enrichment analysis workflow.	44
Figure 3.5. Z-score scatter plots of NE and IE plants. (A) Z-score scatter plot of the curated REMD list and (B) z-score scatter plot of KEGG rpairs. Red: enriched in NE genotypes, blue: enriched in IE genotypes, grey: non-significant REMDs, light and dark green: depleted in NE or IE.	48
Figure 3.6. Triangle motifs show novel biochemical relationships. (A) Graph laid out using prenylation, geranylation, 2-ketohexanoic acid type b, adipic acid type b, decanoic acid type a, condensation and hydroxylation. Triangle motifs containing: (B) prenylation, hydroxylation and adipic acid (C) decanoic acid, geranylation and condensation, (D) 2-hexanoic acid, prenylation and condensation. (E) Detailed explanation of triangle shown in Fig. 3.6B and, (F) detailed explanation of triangle motif shown in Fig. 3.6C.	50
Figure 3.7. Mass difference networks inherit pathway information. (A) Sub-graph laid out using hydration, hydroxylation and the top six over-represented REMDs in NE illustrating the formation of compositional networks, namely CHO, CHOP, CHOS, CHNO and CHNOS; red: upregulated NE nodes, blue: downregulated NE nodes, grey nodes: not significant. (B) Magnified view of the shikimate pathway.	54
Figure 3.8. Clustermap of the incidence matrix. (A) Exemplifies the formation of discriminant unknown sum formulas of NE and IE from known KEGG metabolites. (B) The clustermap, in which the most important findings are highlighted (red font color: significant for NE; blue font color: significant for IE).....	56
Figure 3.9. Modules of mass difference networks contain compound information. (A) Extracted sub-graph obtained by using REMDs shown in Fig. 3.9E; red: upregulated NE nodes, blue: downregulated NE nodes. (B) The same graph colored according to the different modularity classes. (C) Pie plot representation of compound classes present in each modularity class; red, flavonoids; blue, phenolics; green, terpenes; yellow, lipids; teal, carbohydrates; black, other primary metabolites; grey, unknown. (D) Shikimate pathway found in Fig. 3.9A. (E) Relative abundance of each REMD present in the sub-graph.	58
Figure 3.10. Comparison of SIM-MS/MS MDEA with full-scan MDEA results. (A) Venn-plot illustration of enriched NE full scan and SIM-MS/MS REMDs. (B) The overlapping NE REMDs are illustrated in z-score bar charts. (C) Venn-plot illustration of enriched IE full scan and SIM-MS/MS REMDs. (D) The overlapping IE REMDs are illustrated in z-score bar-charts. (E) Heatmap showing the correlation coefficients of REMD z-score vectors obtained by full-scan MDEA (POP), incidence matrix MDEA (IM) and SIM-MS/MS MDEA (SIM).	60

Figure 4.1. Experimental setup of the <i>C. populi</i> bioassay under field conditions. (A) Photo of free-air cages in Göttingen, (B) the arrangement of IE (Wild type: W; GUS: G) and NE (RA1: R1; RA22; R22) poplars in each cage and, (C) experimental timeline highlighting the sampling days for ICR-FT/MS measurements (modified according to Müller et al., 2015)...	63
Figure 4.2. Principal Component Analysis of ICR-FT/MS data. (A) PC1xPC2 score scatter plot showing the separation of young and old leaves and (B) PC2xPC3 score scatter plot visualizes the distinct clustering of fed and unfed plants (modified according to Müller et al., 2015). Blue symbols, control plants; red symbols, infested plants; triangles, NE poplars; circles, IE poplars; filled symbols, young leaves; no-fill symbols, old leaves.....	64
Figure 4.3. Comparison of discriminant masses. (A) Venn-plot comparing the leaf-age discriminant masses at the different experimental variables (herbivory and line). (B) Annotations of masses that are discriminant for leaf-age and are independent of other experimental variables (herbivory and line effect). (C) Venn-plot comparing the masses discriminant for herbivory. (D) Detailed view of the α -linolenic acid metabolism, colors indicate their location in the Venn-diagram shown in C (modified according to Müller et al., 2015).	66
Figure 4.4. Clustermap and database annotations of ICR-FT/MS intensity profiles of discriminant masses. (A) Clustermap of discriminant masses. Venn plots of annotated discriminant metabolites of infested young (red color) and old leaves (blue color) in (B) IE and, (C) NE plants. All database annotations are shown in bar charts ordered according to compound classes and metabolic pathways (red: young leaves, blue: old leaves, grey: discriminant for both leaf types; inf: infested; d4: day 4; d8: day 8; d14: day 14; bs: biosynthesis: C: control; y: young, reprinted from Müller et al., 2015; publisher: BioMed Central).....	68
Figure 5.1. Experimental setup of the <i>C. populi</i> behavior study with EMF inoculated (M) and non-inoculated (n) poplars under field conditions. (A) Photo of the free-air cages in Göttingen and (B) the arrangement of poplars in each cage. (modified according to Müller 2014)	71
Figure 5.2. Statistical analysis of the GC-MS data. (A) S-plot showing the contributions of volatiles ($p(\text{corr})[1]$ = correlation coefficient between X and the score vector 1; $p[1]$ = score vector 1) for the comparison of fed (n=6) versus unfed (n=6) poplars, significant volatiles are colored in red and shown in table (B). Legend: GLV: green leaf volatile, FA: fatty acid, Alde: aldehyde, MT: monoterpene, SQT: sesquiterpene.	72
Figure 5.3. The extension of the mass difference analysis workflow which aims at mapping the transcriptomic data on their respective REMDs (compare to Figure 3.4).	74
Figure 5.4. Statistical results of the LC-MS measurements. (A) PCA score plot of (-)LC-MS data and (B) PCA score plot of (+)LC-MS data. Venn-plots of sum formulas discriminant for the EMF effect in (C) (-)LC-MS- and (D) (+)LC-MS measurements. Venn-plots comparing sum formulas discriminant for the beetle effect obtained by (E) (-)LC-MS and (F) (+)LC-MS measurements. Red arrows: upregulated sum formulas; blue arrows: downregulated sum formulas.	76
Figure 5.5. The compositional space of discriminant sum formulas. Sum formulas discriminant for (A) EMF inoculation and (B) beetle infestation.	77
Figure 5.6. Illustration of database annotations obtained by uploading the discriminant sum formulas. Metabolites discriminant for (A) EMF inoculation and (B) beetle infestation.	79
Figure 5.7. Flavonoid biosynthesis exemplifies the transcriptomic-metabolomic data matching via REMDs. (A) Log2 fold changes of metabolite and transcript levels of the proanthocyanidin pathway. (B) (-)UPLC-qToF-MS mass-difference network of the flavonoid/proanthocyanidin biosynthesis with transcripts matched on the REMDs.	82

Figure 5.8. Z-scores obtained by MDEA analysis of the EMF-effect in the (-)LC-MS samples. Overrepresented REMDs in (A) the MC/NC comparison and in (B) the MB/NB comparison. REMDs highlighted with dashed lines pertain to the cyanoamino acid metabolism KEGG pathway map starting from (C) phenylalanine (Phe), and REMDs highlighted in grey start from either (D) cysteine or isoleucine. Red background: detected with GC/MS; purple background: detected via transcriptomics; cond.: condensation; DiOH: dihydroxy; FMN: flavinmononucleotide; Glc: glucose; MDA: malondialdehyde; OH: hydroxyl.	84
Figure 5.9. Cross-platform mass difference network. (A) Discriminant sum formulas colored according to the EMF effect and (B) to the beetle-effect.....	87
Figure 5.10. Illustration of systemic changes resulting from EMF inoculation within a poplar cell. Transcriptional changes of enzymes are represented in colored boxes and of transmembrane transporters in colored circles, biosynthetic routes are illustrated in rounded rectangles that are colored according to enzymatic and metabolic regulation patterns found within them. Metabolites and compound classes are also colored according to their respective regulation pattern. Metabolites which are colored in red were detected by GC/MS measurements. 2OG, 2-oxoglutarate; acc, acetyl-CoA carboxylase; Ac-CoA, acetyl-CoA; Asn, asparagine; Asp, aspartic acid; chs, chalcone synthase; Cys, cysteine; cyto, cytosol; DHAP, dihydroxyacetone phosphate; E4P, erythrose 4-phosphate; FA, fatty acid; G3P, glyceraldehyde 3-phosphate; gapdh, glyceraldehyde-3-phosphate dehydrogenase; gdh2, glutamate dehydrogenase 2; GOGAT, glutamine oxoglutarate aminotransferase; GS, glutamate synthase; gst, glutathione S-transferase; Gln, glutamine; Glu, glutamic acid; isps, isoprene synthase; kat, 3-ketoacyl-CoA thiolase; kpi, kunitz protease inhibitor; nas3, nicotianamine synthase; nia, nitrite reductase; nrt1, nitrate transporter; OA, oxaloacetate; pal, phenylalanine ammonia-lyase; PEP, phosphoenolpyruvate; PPP, pentose phosphate pathway; PPT, phosphoenolpyruvate/phosphate translocator; Phpyr, phenylpyruvate; PolyOxid, polyamine oxidase; SAM: S-adenosyl-L-methionine; tpi, triosephosphate isomerase; Trp, tryptophane; Tyr, tyrosine; XGE, xyloglucan endotransglucosylase; biosyn, biosynthesis; trans, transporter.	90

List of tables

Table 1.1: Classes of secondary metabolites.....	5
Table 3.1. Plant based reaction equivalent mass-differences (REMDs). The reaction class represents the mechanistic principle behind the mimicked reaction: A = condensation/hydrolyses and F = others are shown in the table. Other mechanisms which are not included in the table are: B = decarboxylative condensation, C = condensation on hydrogenated carbonyls, D = decarboxylative addition and E = THF related reactions.	41
Table 5.1. Cyclomatic number u of discriminant sum formulas.	78
Table S1. Measured LC-MS standards.	104
Table S2. REMDs overrepresented in inoculated poplars in the MC/NC comparison. ...	105

List of equations

Equation 1: Mass Resolution.....	14
Equation 2: Relative mass error in ppm.....	14
Equation 3: Lorentz force.....	16
Equation 4: Cyclotron frequency.....	16
Equation 5: Mass to charge ratio.....	16
Equation 6: Time-of-flight definition.....	17
Equation 7: Van Deemter equation.....	18
Equation 8: Total ion current normalization.....	24

Literature

- Abdel-Lateif K, Bogusz D, Hocher V** (2012) The role of flavonoids in the establishment of plant roots endosymbioses with arbuscular mycorrhiza fungi, rhizobia and Frankia bacteria. *Plant Signal Behav* **7**: 636–641
- Abreu IN, Ahnlund M, Moritz T, Albrechtsen BR** (2011) UHPLC-ESI/TOFMS determination of salicylate-like phenolic glycosides in *Populus tremula* leaves. *J Chem Ecol* **37**: 857–870
- Affek HP, Yakir D** (2003) Natural abundance carbon isotope composition of isoprene reflects incomplete coupling between isoprene synthesis and photosynthetic carbon flow. *Plant Physiol* **131**: 1727–1736
- Agati G, Azzarello E, Pollastri S, Tattini M** (2012) Flavonoids as antioxidants in plants: Location and functional significance. *Plant Sci* **196**: 67–76
- Agati G, Matteini P, Goti A, Tattini M** (2007) Chloroplast-located flavonoids can scavenge singlet oxygen. *New Phytol* **174**: 77–89
- Allen CD, Macalady AK, Chenchouni H, Bachelet D, McDowell N, Vennetier M, Kitzberger T, Rigling A, Breshears DD, Hogg EHT** (2010) A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *For Ecol Manage* **259**: 660–684
- Amin A, Gali-Muhtasib H, Ocker M, Schneider-Stock R** (2009) Overview of Major Classes of Plant-Derived Anticancer Drugs. *Int J Biomed Sci* **5**: 1–11
- Arimura G -i., Matsui K, Takabayashi J** (2009) Chemical and Molecular Ecology of Herbivore-Induced Plant Volatiles: Proximate Factors and Their Ultimate Functions. *Plant Cell Physiol* **50**: 911–923
- Arimura G, Kost C, Boland W** (2005) Herbivore-induced, indirect plant defences. *Biochim Biophys Acta - Mol Cell Biol Lipids* **1734**: 91–111
- Aylott MJ, Casella E, Tubby I, Street NR, Smith P, Taylor G** (2008) Yield and spatial supply of bioenergy poplar and willow short-rotation coppice in the UK. *New Phytol* **178**: 358–370
- Babst B a., Harding S a., Tsai CJ** (2010) Biosynthesis of phenolic glycosides from phenylpropanoid and benzenoid precursors in populus. *J Chem Ecol* **36**: 286–297
- Banerjee a, Sharkey TD** (2014) Methylerythritol 4-phosphate (MEP) pathway metabolic regulation. *Nat Prod Rep* **31**: 1043–55
- Barabási A-L, Oltvai ZN** (2004) Network biology: understanding the cell's functional organization. *Nat Rev Genet* **5**: 101–113

- Barrow MP, Burkitt WI, Derrick PJ** (2005) Principles of Fourier transform ion cyclotron resonance mass spectrometry and its application in structural biology. *Analyst* **130**: 18–28
- Behie SW, Bidochka MJ** (2014) Nutrient transfer in plant–fungal symbioses. *Trends Plant Sci* **19**: 734–740
- Behnke K, Ehltng B, Teuber M, Bauerfeind M, Louis S, Hansch R, Polle A, Bohlmann J, Schnitzler J-P** (2007) Transgenic, non-isoprene emitting poplars don't like it hot. *Plant J* **51**: 485–499
- Behnke K, Grote R, Bruggemann N, Zimmer I, Zhou G, Elobeid M, Janz D, Polle A, Schnitzler J-P** (2012) Isoprene emission-free poplars--a chance to reduce the impact from poplar plantations on the atmosphere. *New Phytol* **194**: 70–82
- Behnke K, Kaiser A, Zimmer I, Bruggemann N, Janz D, Polle A, Hampp R, Hansch R, Popko J, Schmitt-Kopplin P, et al** (2010a) RNAi-mediated suppression of isoprene emission in poplar transiently impacts phenolic metabolism under high temperature and high light intensities: a transcriptomic and metabolomic analysis. *Plant Mol Biol* **74**: 61–75
- Behnke K, Kleist E, Uerlings R, Wildt J, Rennenberg H, Schnitzler J-P** (2009) RNAi-mediated suppression of isoprene biosynthesis in hybrid poplar impacts ozone tolerance. *Tree Physiol* **29**: 725–736
- Behnke K, Loivamäki M, Zimmer I, Rennenberg H, Schnitzler JP, Louis S** (2010b) Isoprene emission protects photosynthesis in sunfleck exposed Grey poplar. *Photosynth Res* **104**: 5–17
- Bernard SM, Habash DZ** (2009) The importance of cytosolic glutamine synthetase in nitrogen assimilation and recycling. *New Phytol* **182**: 608–620
- Bishop GJ, Yokota T** (2001) Plants Steroid Hormones, Brassinosteroids: Current Highlights of Molecular Aspects on their Synthesis/Metabolism, Transport, Perception and Response. *Plant Cell Physiol* **42**: 114–120
- Boccard J, Veuthey J-L, Rudaz S** (2010) Knowledge discovery in metabolomics: an overview of MS data handling. *J Sep Sci* **33**: 290–304
- Boeckler GA, Gershenzon J, Unsicker SB** (2011) Phenolic glycosides of the Salicaceae and their role as anti-herbivore defenses. *Phytochemistry* **72**: 1497–1509
- Boeckler GA, Towns M, Unsicker SB, Mellway RD, Yip L, Hilke I, Gershenzon J, Constabel CP** (2014) Transgenic upregulation of the condensed tannin pathway in poplar leads to a dramatic shift in leaf palatability for two tree-feeding Lepidoptera. *J Chem Ecol* **40**: 150–158
- Boerjan W, Ralph J, Baucher M** (2003) Lignin biosynthesis. *Annu Rev Plant Biol* **54**: 519–546

- Boxman AW, Roelofs JGM** (1988) Some effects of nitrate versus ammonium nutrition on the nutrient fluxes in *Pinus sylvestris* seedlings. Effects of mycorrhizal infection. *Can J Bot* **66**: 1091–1097
- Boyle NR, Morgan JA** (2009) Flux balance analysis of primary metabolism in *Chlamydomonas reinhardtii*. *BMC Syst Biol* **3**: 4
- Breitling R, Ritchie S, Goodenowe D, Stewart ML, Barrett MP** (2006) Ab initio prediction of metabolic networks using Fourier transform mass spectrometry data. *Metabolomics* **2**: 155–164
- Brian PW** (1959) Effects of gibberellins on plant growth and development. *Biol Rev* **34**: 37–77
- Brilli F, Ciccio P, Frattoni M, Prestininzi M, Spanedda AF, Loreto F** (2009) Constitutive and herbivore-induced monoterpenes emitted by *Populus x euroamericana* leaves are key volatiles that orient *Chrysomela populi* beetles. *Plant Cell Environ* **32**: 542–552
- Brisson L, Vacha WEK, Ibrahim RK** (1986) Localization of partially methylated flavonol glucosides in *Chrysosplenium americanum*. II. Immunofluorescence. *Plant Sci* **44**: 175–181
- Brown SC, Kruppa G, Dasseux J-L** (2005) Metabolomics applications of FT-ICR mass spectrometry. *Mass Spectrom Rev* **24**: 223–231
- Bundy JG, Davey MP, Viant MR** (2009) Environmental metabolomics: a critical review and future perspectives. *Metabolomics* **5**: 3–21
- Burchard P, Bilger W, Weissenböck G** (2000) Contribution of hydroxycinnamates and flavonoids to, epidermal shielding of UV-A and UV-B radiation in developing rye primary leaves as assessed by ultraviolet-induced chlorophyll fluorescence measurements. *Plant, Cell Environ* **23**: 1373–1380
- Callemien D, Collin S** (2008) Use of RP-HPLC-ESI (-)-MS/MS to differentiate various proanthocyanidin isomers in lager beer extracts. *J Am Soc Brew Chem*: **66**, 109-115
- Carletti P, Masi A, Wonisch A, Grill D, Tausz M, Ferretti M** (2003) Changes in antioxidant and pigment pool dimensions in UV-B irradiated maize seedlings. *Environ Exp Bot* **50**: 149–157
- Carslaw KS, Boucher O, Spracklen D V, Mann GW, Rae JGL, Woodward S, Kulmala M** (2010) A review of natural aerosol interactions and feedbacks within the Earth system. *Atmos Chem Phys* **10**: 1701–1737
- Chameides WL, Lindsay RW, Richardson J, Kiang CS** (1988) The role of biogenic hydrocarbons in urban photochemical smog: Atlanta as a case study. *Sci* **241** : 1473–1475
- Charest PM, Brisson L, Ibrahim RK** (1986) Ultrastructural features of flavonoid accumulation in leaf cells of *Chrysosplenium americanum*. *Protoplasma* **134**: 95–101

- Chellamuthu V-R, Ermilova E, Lapina T, Lüddecke J, Minaeva E, Herrmann C, Hartmann MD, Forchhammer K** (2014) A Widespread Glutamine-Sensing Mechanism in the Plant Kingdom. *Cell* **159**: 1188–1199
- Cheng A-X, Lou Y-G, Mao Y-B, Lu S, Wang L-J, Chen X-Y** (2007) Plant Terpenoids: Biosynthesis and ecological functions. *J Integr Plant Biol* **49**: 179–186
- Chernushevich I V, Loboda A V, Thomson BA** (2001) An introduction to quadrupole-time-of-flight mass spectrometry. *J Mass Spectrom* **36**: 849–865
- Claeys M, Graham B, Vas G, Wang W, Vermeylen R, Pashynska V, Cafmeyer J, Guyon P, Andreae MO, Artaxo P, et al** (2004) Formation of Secondary Organic Aerosols Through Photooxidation of Isoprene. *Sci* **303** : 1173–1176
- Clarke HRG, Lawrence SD, Flaskerud J, Korhnek TE, Gordon MP, Davis JM** (1998) Chitinase accumulates systemically in wounded poplar trees. *Physiol Plant* **103**: 154–161
- Clavijo McCormick A, Irmisch S, Reinecke A, Boeckler GA, Veit D, REICHEL T M, HANSSON BS, GERSHENZON J, KÖLLNER TG, UNSICKER SB** (2014) Herbivore-induced volatile emission in black poplar: regulation and role in attracting herbivore enemies. *Plant Cell Environ* **37**: 1909–1923
- Clavijo McCormick A, Unsicker SB, Gershenzon J** (2012) The specificity of herbivore-induced plant volatiles in attracting herbivore enemies. *Trends Plant Sci* **17**: 303–310
- Collinge DB, Kragh KM, Mikkelsen JD, Nielsen KK, Rasmussen U, Vad K** (1993) Plant chitinases. *Plant J* **3**: 31–40
- Costa H, Gallego SM, Tomaro ML** (2002) Effect of UV-B radiation on antioxidant defense system in sunflower cotyledons. *Plant Sci* **162**: 939–945
- Cotter RJ** (1992) Time-of-flight mass spectrometry for the structural analysis of biological molecules. *Anal Chem* **64**: 1027A–1039A
- Coviella CE, Stipanovic RD, Trumble JT** (2002) Plant allocation to defensive compounds: interactions between elevated CO₂ and nitrogen in transgenic cotton plants. *J Exp Bot* **53**: 323–331
- Crick F** (1970) Central dogma of molecular biology. *Nature* **227**: 561–563
- Crick FHC** (1956) On protein synthesis. *Symp Soc Exp Biol* **XII**: 139–163
- Croteau R, Kutchan TM, Lewis NG** (2000) Secondary Metabolites. *Biochem Mol Biol Plants* **7**: 1250–1318
- Davies DD** (1979) The central role of phosphoenolpyruvate in plant metabolism. *Annu Rev Plant Physiol* **30**: 131–158
- Delker C, Stenzel I, Hause B, Miersch O, Feussner I, Wasternack C** (2006) Jasmonate Biosynthesis in *Arabidopsis thaliana*- Enzymes, Products, Regulation. *Plant Biol* **8**: 297–306

- Deroe C, Pasteels JM** (1982) Distribution of adult defense glands in chrysomelids (Coleoptera: Chrysomelidae) and its significance in the evolution of defense mechanisms within the family. *J Chem Ecol* **8**: 67–82
- Dizengremel P, Vaultier MN, Le Thiec D, Cabane M, Bagard M, Gerant D, Gerard J, Dghim AA, Richet N, Afif D, et al** (2012) Phosphoenolpyruvate is at the crossroads of leaf metabolic responses to ozone stress. *New Phytol* **195**: 512–517
- Dorantes-Acosta AE, Sánchez-Hernández CV, Arteaga-Vazquez MA** (2012) Biotic stress in plants: life lessons from your parents and grandparents . *Front. Genet.* **3** :
- Dorsey JG, Dill K a.** (1989) The molecular mechanism of retention in reversed-phase liquid chromatography. *Chem Rev* **89**: 331–346
- Doughty J, Aljabri M, Scott RJ** (2014) Flavonoids and the regulation of seed size in *Arabidopsis*. *Biochem Soc Trans* **42**: 364–369
- Dubey VS, Bhalla R, Luthra R** (2003) An overview of the non-mevalonate pathway for terpenoid biosynthesis in plants. *J Biosci* **28**: 637–646
- Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L** (2010) MYB transcription factors in *Arabidopsis*. *Trends Plant Sci* **15**: 573–581
- Edwards R, Dixon DP, Walbot V** (2000) Plant glutathione S-transferases: enzymes with multiple functions in sickness and in health. *Trends Plant Sci* **5**: 193–198
- Ejigu BA, Valkenburg D, Baggerman G, Vanaerschot M, Witters E, Dujardin J-C, Burzykowski T, Berg M** (2013) Evaluation of Normalization Methods to Pave the Way Towards Large-Scale LC-MS-Based Metabolomics Profiling Experiments. *OMICS: A Journal of Integrative Biology* **17**: 473-485
- Faijes M, Mars A, Smid E** (2007) Comparison of quenching and extraction methodologies for metabolome analysis of *Lactobacillus plantarum*. *Microb Cell Fact* **6**: 27
- Feng G, Zhang F, Li X, Tian C, Tang C, Rengel Z** (2002) Improved tolerance of maize plants to salt stress by arbuscular mycorrhiza is related to higher accumulation of soluble sugars in roots. *Mycorrhiza* **12**: 185-190
- Ferreira Da Silva P, Lima JC, Freitas A a., Shimizu K, Maçanita AL, Quina FH** (2005) Charge-transfer complexation as a general phenomenon in the copigmentation of anthocyanins. *J Phys Chem A* **109**: 7329–7338
- Fiehn O** (2002) Metabolomics--the link between genotypes and phenotypes. *Plant Mol Biol* **48**: 155–171
- Fontana A, Reichelt M, Hempel S, Gershenzon J, Unsicker SB** (2009) The effects of arbuscular mycorrhizal fungi on direct and indirect defense metabolites of *Plantago lanceolata* L. *J Chem Ecol* **35**: 833–843
- Forcisi S, Moritz F, Kanawati B, Tziotis D, Lehmann R, Schmitt-Kopplin P** (2013) Liquid chromatography–mass spectrometry in metabolomics research: Mass analyzers in ultra

- high pressure liquid chromatography coupling. *J Chromatogr A* **1292**: 51–65
- Forcisi S, Moritz F, Lucio M, Lehmann R, Stefan N, Schmitt-Kopplin P** (2015) Solutions for low and high accuracy mass spectrometric data matching: a data-driven annotation strategy in nontargeted metabolomics. *Anal Chem* **87**: 8917–8924
- Fuentes JD, Gu L, Lerdau M, Atkinson R, Baldocchi D, Bottenheim JW, Ciccioli P, Lamb B, Geron C, Guenther A, Sharkey TD, Stockwell W** (2000) Biogenic hydrocarbons in the atmospheric boundary layer: a review. *Bull Amer Meteor Soc* **81**: 1537–1575
- Fujita M, Fujita Y, Noutoshi Y, Takahashi F, Narusaka Y, Yamaguchi-Shinozaki K, Shinozaki K** (2006) Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. *Curr Opin Plant Biol* **9**: 436–442
- Fulkerson DR, Gross O** (1965) Incidence matrices and interval graphs. *Pacific J Math* **15**: 835–855
- Galindo-Prieto B, Eriksson L, Trygg J** (2014) Variable influence on projection (VIP) for orthogonal projections to latent structures (OPLS). *J Chemom* **28**: 623–632
- Gange AC, West HM** (1994) Interactions between arbuscular mycorrhizal fungi and foliar-feeding insects in *Plantago lanceolata* L. *New Phytol* **128**: 79–87
- Ghirardo A, Wright LP, Bi Z, Rosenkranz M, Pulido P, Rodríguez-Concepción M, Niinemets Ü, Brüggemann N, Gershenzon J, Schnitzler J-P** (2014) Metabolic flux analysis of plastidic isoprenoid biosynthesis in poplar leaves emitting and non-emitting isoprene. *Plant Physiol* **165**: 37–51
- Giddings JC** (1969) Generation of variance, “theoretical plates,” resolution, and peak capacity in electrophoresis and sedimentation. *Sep Sci Technol* **4**: 181–189
- Gil M, Pontin M, Berli F, Bottini R, Piccoli P** (2012) Metabolism of terpenes in the response of grape (*Vitis vinifera* L.) leaf tissues to UV-B radiation. *Phytochemistry* **77**: 89–98
- Gonzalez-Cabanelas D, Wright LP, Paetz C, Onkokesung N, Gershenzon J, Rodriguez-Concepcion M, Phillips MA** (2015) The diversion of 2-C-methyl-d-erythritol-2,4-cyclodiphosphate from the 2-C-methyl-d-erythritol 4-phosphate pathway to hemiterpene glycosides mediates stress responses in *Arabidopsis thaliana*. *Plant J* **82**: 122–137
- Goodacre R** (2005) Metabolomics – the way forward. *Metabolomics* **1**: 1–2
- Goodacre R, Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB** (2004) Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends Biotechnol* **22**: 245–252
- Guenther A, Hewitt CN, Erickson D, Fall R, Geron C, Graedel T, Harley P, Klinger L, Lerdau M, Mckay WA, et al** (1995) A global model of natural volatile organic

- compound emissions. *J Geophys Res Atmos* **100**: 8873-8892
- Gross JH** (2011) Mass Spectrometry - A textbook. 2nd Edition.
- Grotewold E** (2005) Plant metabolic diversity: a regulatory perspective. *Trends Plant Sci* **10**: 57–62
- Gruppe A, Fusseder M, Schopf R** (1999) Short rotation plantations of aspen and balsam poplar on former arable land in Germany: defoliating insects and leaf constituents. *For Ecol Manage* **121**: 113–122
- Harborne JB, Williams CA** (2000) Advances in flavonoid research since 1992. *Phytochemistry* **55**: 481–504
- Harley P, Deem G, Flint S, Caldwell M** (1996) Effects of growth under elevated UV-B on photosynthesis and isoprene emission in *Quercus gambelii* and *Mucuna pruriens*. *Glob Chang Biol* **2**: 149–154
- Hartmann T** (2007) From waste products to ecochemicals: fifty years research of plant secondary metabolism. *Phytochemistry* **68**: 2831–2846
- Hassan MN, Zainal Z, Ismail I** (2015) Green leaf volatiles: biosynthesis, biological functions and their applications in biotechnology. *Plant Biotechnol. J.*
- Heijde M, Ulm R** (2012) UV-B photoreceptor-mediated signalling in plants. *Trends Plant Sci* **17**: 230–237
- Heine GF, Hernandez JM, Grotewold E** (2004) Two cysteines in plant R2R3 MYB domains participate in REDOX-dependent DNA binding. *J Biol Chem* **279**: 37878–37885
- Herrmann KM, Weaver LM** (1999) The shikimate pathway. *Annu Rev Plant Biol* **50**: 473–503
- Hideg É, Jansen MAK, Strid Å** (2013) UV-B exposure, ROS, and stress: inseparable companions or loosely linked associates? *Trends Plant Sci* **18**: 107–115
- Hodges DM, DeLong JM, Forney CF, Prange RK** (1999) Improving the thiobarbituric acid-reactive-substances assay for estimating lipid peroxidation in plant tissues containing anthocyanin and other interfering compounds. *Planta* **207**: 604–611
- Holopainen JK** (2004) Multiple functions of inducible plant volatiles. *Trends Plant Sci* **9**: 529–533
- Ibrahim RK, De Luca V, Khouri H, Latchinian L, Brisson L, Charest PM** (1987) Enzymology and compartmentation of polymethylated flavonol glucosides in *chrysosplenium americanum*. *Phytochemistry* **26**: 1237–1245
- Irmisch S, Clavijo McCormick A, Boeckler GA, Schmidt A, Reichelt M, Schneider B, Block K, Schnitzler J-P, Gershenzon J, Unsicker SB, et al** (2013) Two Herbivore-Induced Cytochrome P450 Enzymes CYP79D6 and CYP79D7 Catalyze the Formation of Volatile Aldoximes Involved in Poplar Defense. *Plant Cell* **25**: 4737–4754
- Irmisch S, Clavijo McCormick A, Günther J, Schmidt A, Boeckler GA, Gershenzon J,**

- Unsicker SB, Köllner TG** (2014) Herbivore-induced poplar cytochrome P450 enzymes of the CYP71 family convert aldoximes to nitriles which repel a generalist caterpillar. *Plant J* **80**: 1095–1107
- Jakobsen I, Rosendahl L** (1990) Carbon flow into soil and external hyphae from roots of mycorrhizal cucumber plants. *New Phytol* **115**: 77–83
- Jansson S, Douglas CJ** (2007) *Populus*: A Model System for Plant Biology. *Annu Rev Plant Biol* **58**: 435–458
- Janz D, Behnke K, Schnitzler JP, Kanawati B, Schmitt-Kopplin P, Polle A** (2010) Pathway analysis of the transcriptome and metabolome of salt sensitive and tolerant poplar species reveals evolutionary adaptation of stress tolerance mechanisms. *BMC Plant Biol.* 10:
- Jeandet P, Heinzmann SS, Roullier-Gall C, Cilindre C, Aron A, Deville MA, Moritz F, Karbowskiak T, Demarville D, Brun C, et al** (2015) Chemical messages in 170-year-old champagne bottles from the Baltic Sea: Revealing tastes from the past. *Proc Natl Acad Sci, U S A.* doi: D - NLM: PMC4434772 OTO - NOTNLM
- Jenkins GI** (2009) Signal transduction in responses to UV-B radiation. *Annu Rev Plant Biol* **60**: 407–431
- Jung HW, Tschaplinski TJ, Wang L, Glazebrook J, Greenberg JT** (2009) Priming in systemic plant immunity. *Science (80-)* **324**: 89–91
- Kachlicki P, Einhorn J, Muth D, Kerhoas L, Stobiecki M** (2008) Evaluation of glycosylation and malonylation patterns in flavonoid glycosides during LC/MS/MS metabolite profiling. *J Mass Spectrom* **43**: 572–586
- Kaddurah-Daouk R, Kristal BS, Weinshilboum RM** (2008) Metabolomics: A Global Biochemical Approach to Drug Response and Disease. *Annu Rev Pharmacol Toxicol* **48**: 653–683
- Kaling M, Kanawati B, Ghirardo A, Albert A, Winkler JB, Heller W, Barta C, Loreto F, Schmitt-Kopplin P, Schnitzler Jör-P** (2015) UV-B mediated metabolic rearrangements in poplar revealed by non-targeted metabolomics. *Plant Cell Environ* **38**: 892–904
- Kankainen M, Gopalacharyulu P, Holm L, Oresic M** (2011) MPEA--metabolite pathway enrichment analysis. *Bioinformatics* **27**: 1878–1879
- Kaplan F, Kopka J, Haskell DW, Zhao W, Schiller KC, Gatzke N, Sung DY, Guy CL, Molecular P, Program CB, et al** (2004) Exploring the Temperature-Stress Metabolome. *Plant Physiol* **136**: 4159–4168
- Kaspar S, Matros A, Mock H** (2010) Proteome and Flavonoid Analysis Reveals Distinct Responses of Epidermal Tissue and Whole Leaves upon UV - B Radiation of Barley (*Hordeum vulgare L.*) Seedlings research articles. 2402–2411

- Kebarle P, Tang L** (1993) From ions in solution to ions in the gas phase. *Society* 65:
- Kebarle P, Verkerk UH** (2009) Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass Spectrom Rev* **28(6)**: 898–917
- Kelly GJ, Latzko E** (1977) Chloroplast phosphofructokinase II. Partial purification, kinetic and regulatory properties. *Plant Physiol* **60**: 295–299
- Keski-Saari S, Julkunen-Tiitto R** (2003) Resource allocation in different parts of juvenile mountain birch plants: effect of nitrogen supply on seedling phenolics and growth. *Physiol Plant* **118**: 114–126
- Kiendler-Scharr A, Andres S, Bachner M, Behnke K, Broch S, Hofzumahaus A, Holland F, Kleist E, Mentel TF, Rubach F, et al** (2012) Isoprene in poplar emissions: effects on new particle formation and OH concentrations. *Atmos Chem Phys* **12**: 1021–1030
- Kim HK, Wilson EG, Choi YH, Verpoorte R** (2010) Metabolomics: a tool for anticancer lead-finding from natural products. *Planta Med* **76**: 1094–1102
- Koes R, Verweij W, Quattrocchio F** (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends Plant Sci* **10**: 236–242
- Kostiainen R, Kauppila TJ** (2009) Effect of eluent on the ionization process in liquid chromatography–mass spectrometry. *J Chromatogr A* **1216**: 685–699
- Kottke I, Oberwinkler F** (1987) The cellular structure of the Hartig net: coenocytic and transfer cell-like organization. *Nord J Bot* **7**: 85–95
- Kusano M, Tohge T, Fukushima A, Kobayashi M, Hayashi N, Otsuki H, Kondou Y, Goto H, Kawashima M, Matsuda F, et al** (2011) Metabolomics reveals comprehensive reprogramming involving two independent metabolic responses of *Arabidopsis* to UV-B light. *Plant J* **67**: 354–69
- Lan K, Jia W** (2010) An integrated metabolomics and pharmacokinetics strategy for multi-component drugs evaluation. *Curr Drug Metab* **11**: 105–114
- Landry LG, Chapple CC, Last RL** (1995) *Arabidopsis* mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. *Plant Physiol* **109**: 1159–1166
- Laothawornkitkul J, Paul ND, Vickers CE, Possell M, Taylor JE, Mullineaux PM, Hewitt CN** (2008) Isoprene emissions influence herbivore feeding decisions. *Plant Cell Environ* **31**: 1410–1415
- Lattanzio V., Lattanzio VMT., Cardinali A** (2006) Role of phenolics in the resistance mechanisms of plants against fungal pathogens and insects. *Phytochemistry* 23–67
- Lavola A** (1998) Accumulation of flavonoids and related compounds in birch induced by UV-B irradiance. *Tree Physiol* **18**: 53–58
- Lawrence SD, Novak NG** (2006) Expression of poplar chitinase in tomato leads to inhibition of development in Colorado potato beetle. *Biotechnol Lett* **28**: 593–599

- Lelieveld JOS, Crutzen PJ, Denthner F** (1998) Changing concentration, lifetime and climate forcing of atmospheric methane. *Tellus B* **50**: 128–150
- Lepiniec L, Debeaujon I, Routaboul JM, Baudry A, Pourcel L, Nesi N, Caboche M** (2006) Genetics and biochemistry of seed flavonoids. *Annu Rev Plant Biol* **57**: 405–430
- Li J, Ou-Lee T, Raba R, Amundson R, Last R** (1993) Arabidopsis Flavonoid Mutants Are Hypersensitive to UV-B Irradiation. *Plant Cell* **5**: 171–179
- Lidon FJC, Teixeira M, Ramalho JC** (2012) Decay of the Chloroplast Pool of Ascorbate Switches on the Oxidative Burst in UV-B-Irradiated Rice. *J Agron Crop Sci* **198**: 130–144
- Lima JC, Vautier-Giongo C, Lopes A, Melo E, Quina FH, Macanita AL** (2002) Color stabilization of anthocyanins: Effect of SDS micelles on the acid-base and hydration kinetics of malvidin 3-glucoside (Oenin). *J Phys Chem A* **106**: 5851–5859
- Loivamäki M, Mumm R, Dicke M, Schnitzler J-P** (2008) Isoprene interferes with the attraction of bodyguards by herbaceous plants. *Proc Natl Acad Sci U S A* **105**: 17430–17435
- Loreto F, Mannozi M, Maris C, Nascetti P, Ferranti F, Pasqualini S** (2001) Ozone quenching properties of isoprene and its antioxidant role in leaves. *Plant Physiol* **126**: 993–1000
- Loreto F, Schnitzler JP** (2010) Abiotic stresses and induced BVOCs. *Trends Plant Sci* **15**: 154–166
- Lucio M** (2009) Datamining metabolomics: the convergence point of non-target approach and statistical investigation. Dissertation, Technische Universität München
- Lucio M, Fekete A, Frommberger M, Schmitt-Kopplin P** (2011) Metabolomics: High-Resolution Tools Offer to Follow Bacterial Growth on a Molecular Level. *Handb. Mol. Microb. Ecol. I*. John Wiley & Sons, Inc., pp 683–695
- Maeda H, Dudareva N** (2012) The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants. *Annu Rev Plant Biol* **63**: 73–105
- Mandal SM, Chakraborty D, Dey S** (2010) Phenolic acids act as signaling molecules in plant-microbe symbioses. *Plant Signal Behav* **5**: 359–368
- Manetas Y** (2006) Why some leaves are anthocyanic and why most anthocyanic leaves are red? *Flora - Morphol Distrib Funct Ecol Plants* **201**: 163–177
- Marquet A, Bui BTS, Florentin D** (2001) Biosynthesis of biotin and lipoic acid. *Vitam Horm* **61**: 51–101
- Marshall AG, Hendrickson CL, Jackson GS** (1998) Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrom Rev* **17**: 1–35
- Martens H, Martens M** (2000) Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Qual Prefer* **11**: 5–

- Marti G, Erb M, Boccard J, Glauser G, Doyen GR, Villard N, Robert C a M, Turlings TCJ, Rudaz S, Wolfender JL** (2013) Metabolomics reveals herbivore-induced metabolites of resistance and susceptibility in maize leaves and roots. *Plant, Cell Environ* **36**: 621–639
- Masclaux-Daubresse C, Daniel-Vedele F, Dechorgnat J, Chardon F, Gaufichon L, Suzuki A** (2010) Nitrogen uptake, assimilation and remobilization in plants: challenges for sustainable and productive agriculture. *Ann Bot* **105**: 1141–1157
- Matsui K** (2006) Green leaf volatiles: hydroperoxide lyase pathway of oxylipin metabolism. *Curr Opin Plant Biol* **9**: 274–280
- Mellway RD, Tran LT, Prouse MB, Campbell MM, Constabel CP** (2009) The wound-, pathogen-, and ultraviolet B-responsive MY134 gene encodes an R2R3 MYB transcription factor that regulates proanthocyanidin synthesis in poplar. *Plant Physiol* **150**: 924–941
- Mittler R** (2006) Abiotic stress, the field environment and stress combination. *Trends Plant Sci* **11**: 15–19
- Mok MC, Martin RC, Mok DWS** (2000) Cytokinins: Biosynthesis, metabolism and perception. *Vitr Cell Dev Biol* **36**: 102–107
- Moller IM, Jensen PE, Hansson A** (2007) Oxidative modifications to cellular components in plants. *Annu Rev Plant Biol* **58**: 459–481
- Moore K, Roberts LJ** (1998) Measurement of lipid peroxidation. *Free Radic Res* **28**: 659–671
- De Moraes CM, Lewis WJ, Pare PW, Alborn HT, Tumlinson JH** (1998) Herbivore-infested plants selectively attract parasitoids. *Nature* **393**: 570–573
- Moritz F, Janicka M, Zygler A, Forcisi S, Kot-Wasik A, Kot J, Gebefügi I, Namiesnik J, Schmitt-Kopplin P** (2015) The compositional space of exhaled breath condensate and its link to the human breath volatilome. *J Breath Res* **9**: 27105
- Müller A** (2014) Communication between mycorrhizal fungi and poplar. Dissertation, Georg-August-Universität Göttingen
- Müller A, Kaling M, Faubert P, Gort G, Smid H, Van Loon J, Dicke M, Kanawati B, Schmitt-Kopplin P, Polle A, et al** (2015) Isoprene emission by poplar is not important for the feeding behaviour of poplar leaf beetles. *BMC Plant Biol* **15**: 165
- Müller A, Volmer K, Mishra-Knyrim M, Polle A** (2013a) Growing poplars for research with and without mycorrhizas. *Front Plant Sci* **4**: 332
- Müller C, Dietz I, Tziotis D, Moritz F, Rupp J, Schmitt-Kopplin P** (2013b) Molecular cartography in acute *Chlamydia pneumoniae* infections—a non-targeted metabolomics approach. *Anal Bioanal Chem* **405**: 5119–5131

- Murray KK, Boyd RK, Eberlin MN, Langley GJ, Li L, Naito Y** (2013) Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013). *Pure Appl Chem* **85**: 1515–1609
- Muzika R-M** (1993) Terpenes and phenolics in response to nitrogen fertilization: a test of the carbon/nutrient balance hypothesis. *Chemoecology* **4**: 3–7
- Nakashima K, Ito Y, Yamaguchi-Shinozaki K** (2009) Transcriptional regulatory networks in response to abiotic stresses in *Arabidopsis* and grasses. *Plant Physiol* **149**: 88–95
- Nawkar GM, Maibam P, Park JH, Sahi VP, Lee SY, Kang CH** (2013) UV-induced cell death in plants. *Int J Mol Sci* **14**: 1608–1628
- Naz S, Vallejo M, García A, Barbas C** (2014) Method validation strategies involved in non-targeted metabolomics. *J Chromatogr A* **1353**: 99–105
- Nehls U, Göhringer F, Wittulsky S, Dietz S** (2010) Fungal carbohydrate support in the ectomycorrhizal symbiosis: a review. *Plant Biol* **12**: 292–301
- Nicholson RL, Hammerschmidt R** (1992) Phenolic Compounds and Their Role in Disease Resistance. *Annu Rev Phytopathol* **30**: 369–389
- Oikawa A, Nakamura Y, Ogura T, Kimura A, Suzuki H, Sakurai N, Shinbo Y, Shibata D, Kanaya S, Ohta D** (2006) Clarification of Pathway-Specific Inhibition by Fourier Transform Ion Cyclotron Resonance/Mass Spectrometry-Based Metabolic Phenotyping Studies. *Plant Physiol* **142**: 398–413
- op den Camp RG, Przybyla D, Ochsenbein C, Laloi C, Kim C, Danon A, Wagner D, Hideg E, Gobel C, Feussner I, et al** (2003) Rapid induction of distinct stress responses after the release of singlet oxygen in *Arabidopsis*. *Plant Cell* **15**: 2320–2332
- Oss M, Krueve A Fau - Herodes K, Herodes K Fau - Leito I, Leito I** (2010) Electrospray ionization efficiency scale of organic compounds. *Anal Chem* **82**: 2865–72
- Pare PW, Tumlinson JH** (1999) Update on Plant-Insect Interactions Plant Volatiles as a Defense against Insect Herbivores BY RELEASING GREATER AMOUNTS OF A VARIETY. *Plant Physiol* **121**: 325–331
- Parniske M** (2008) Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nat Rev Microbiol* **6**: 763–775
- Passonneau J V, Lowry OH** (1964) The role of phosphofructokinase in metabolic regulation. *Adv Enzyme Regul* **2**: 265–274
- Patil K, Rocha I, Forster J, Nielsen J** (2005) Evolutionary programming as a platform for in silico metabolic engineering. *BMC Bioinformatics* **6**: 308
- Pfabel C, Eckhardt KU, Baum C, Struck C, Frey P, Weih M** (2012) Impact of ectomycorrhizal colonization and rust infection on the secondary metabolism of poplar (*Populus trichocarpa* x *deltoides*). *Tree Physiol* **32**: 1357–1364
- Philippe RN, Bohlmann J** (2007) Poplar defense against insect herbivores. *Can J Bot Can*

Bot **85**: 1111–1126

- Pietta PG** (2000) Flavonoids as antioxidants. *J Nat Prod* **63**: 1035–1042
- Poisson N, Kanakidou M, Crutzen P** (2000) Impact of Non-Methane Hydrocarbons on Tropospheric Chemistry and the Oxidizing Power of the Global Troposphere: 3-Dimensional Modelling Results. *J Atmos Chem* **36**: 157–230
- Pryor W, Stanley JP, Blair E** (1976) Autoxidation of polyunsaturated fatty acids: II. A suggested mechanism for the formation of TBA-reactive materials from prostaglandin-like endoperoxides. *Lipids* **11**: 370–379
- Quideau S, Deffieux D, Douat-Casassus C, Pouységou L** (2011) Plant Polyphenols: Chemical Properties, Biological Activities, and Synthesis. *Angew Chemie Int Ed* **50**: 586–621
- Raamsdonk LM, Teusink B, Broadhurst D, Zhang N, Hayes A, Walsh MC, Berden JA, Brindle KM, Kell DB, Rowland JJ, et al** (2001) A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat Biotech* **19**: 45–50
- Rainville PD, Stumpf CL, Shockcor JP, Plumb RS, Nicholson JK** (2007) Novel application of reversed-phase UPLC-oeTOF-MS for lipid analysis in complex biological mixtures: a new tool for lipidomics. *J Proteome Res* **6**: 552–558
- Rasulov B, Bichele I, Laisk AGU, Niinemets Ü** (2014) Competition between isoprene emission and pigment synthesis during leaf development in aspen. *Plant Cell Environ* **37**: 724–741
- Rates SM.** (2001) Plants as source of drugs. *Toxicol* **39**: 603–613
- Rizzini L, Favory J-J, Cloix C, Faggionato D, O'Hara A, Kaiserli E, Baumeister R, Schäfer E, Nagy F, Jenkins GI, et al** (2011) Perception of UV-B by the Arabidopsis UVR8 protein. *Science* **332**: 103–106
- Rosenstiel TN, Ebbets AL, Khatri WC, Fall R, Monson RK** (2004) Induction of Poplar Leaf Nitrate Reductase: A Test of Extrachloroplastic Control of Isoprene Emission Rate. *Plant Biol* **6**: 12–21
- Ruan J, Haerdter R, Gerendás J** (2010) Impact of nitrogen supply on carbon/nitrogen allocation: a case study on amino acids and catechins in green tea [*Camellia sinensis* (L.) O. Kuntze] plants*. *Plant Biol* **12**: 724–734
- Saito K, Matsuda F** (2010) Metabolomics for Functional Genomics, Systems Biology, and Biotechnology. *Annu Rev Plant Biol* **61**: 463–489
- Sakakibara H, Takei K, Hirose N** (2006) Interactions between nitrogen and cytokinin in the regulation of metabolism and development. *Trends Plant Sci* **11**: 440–448
- Sannigrahi P, Ragauskas AJ, Tuskan GA** (2010) Poplar as a feedstock for biofuels: A review of compositional characteristics. *Biofuels, Bioprod Biorefining* **4**: 209–226

- Sasaki K, Saito T, Lamsa M, Oksman-Caldentey KM, Suzuki M, Ohyama K, Muranaka T, Ohara K, Yazaki K** (2007) Plants utilize isoprene emission as a thermotolerance mechanism. *Plant Cell Physiol* **48**: 1254–1262
- Savchenko T, Walley JW, Chehab EW, Xiao Y, Kaspi R, Pye MF, Mohamed ME, Lazarus CM, Bostock RM, Dehesh K** (2010) Arachidonic acid: an evolutionarily conserved signaling molecule modulates plant stress signaling network. *Plant Cell* **22**: 3193–3205
- Schaberg PG, Shane JB, Cali PF, Donnelly JR, Strimbeck GR** (1998) Photosynthetic capacity of red spruce during winter. *Tree Physiol* **18**: 271–276
- Schmelzer E, Jahn W, Hahlbrock K** (1988) In situ localization of light-induced chalcone synthase mRNA, chalcone synthase, and flavonoid end products in epidermal cells of parsley leaves. *Proc Natl Acad Sci U S A* **85**: 2989–2993
- Schweiger R, Baier MC, Persicke M, Müller C** (2014) High specificity in plant leaf metabolic responses to arbuscular mycorrhiza. *Nat Commun* **5**: 1–11
- Seckmeyer G, Pissulla D, Glandorf M, Henriques D, Johnsen B, Webb A, Siani A-M, Bais A, Kjeldstad B, Brogniez C, et al** (2008) Variability of UV Irradiance in Europe. *Photochem Photobiol* **84**: 172–179
- Senior JK** (1951) Partitions and their representative graphs. *Am J Math* 663–689
- Shen G, Huhman D, Lei Z, Snyder J, Sumner LW, Dixon RA** (2012) Characterization of an isoflavonoid-specific prenyltransferase from *Lupinus albus*. *Plant Physiol* **159**: 70–80
- Simmonds MSJ** (2003) Flavonoid–insect interactions: recent advances in our knowledge. *Phytochemistry* **64**: 21–30
- Singsaas E** (2000) Terpenes and the thermotolerance of photosynthesis. *New Phytol* **146**: 1–4
- Singsaas EL, Sharkey TD** (1997) Effects of transient changes in PAR and leaf temperature on isoprene emission. *Plant Physiol* **114**: 374
- Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, Custodio DE, Abagyan R, Siuzdak G** (2005) METLIN – A metabolite mass spectral database. *Ther Drug Monit* **27**: 747–751
- Southam AD, Payne TG, Cooper HJ, Arvanitis TN, Viant MR** (2007) Dynamic range and mass accuracy of wide-scan direct infusion nanoelectrospray fourier transform ion cyclotron resonance mass spectrometry-based metabolomics increased by the spectral stitching method. *Anal Chem* **79**: 4595–4602
- Stitt M** (1990) Fructose-2, 6-bisphosphate as a regulatory molecule in plants. *Annu Rev Plant Biol* **41**: 153–185
- Stracke R, Ishihara H, Huep G, Barsch A, Mehrtens F, Niehaus K, Weisshaar B** (2007) Differential regulation of closely related R2R3-MYB transcription factors controls

- flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant J* **50**: 660–677
- Streit WR, Entcheva P** (2003) Biotin in microbes, the genes involved in its biosynthesis, its biochemical role and perspectives for biotechnological production. *Appl Microbiol Biotechnol* **61**: 21–31
- Suhre K, Schmitt-Kopplin P** (2008) MassTRIX: mass translator into pathways. *Nucleic Acids Res* **36**: W481–W484
- Sumner LW, Mendes P, Dixon RA** (2003) Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochemistry* **62**: 817–836
- Takahashi M, Terada Y, Nakai I, Nakanishi H, Yoshimura E, Mori S, Nishizawa NK** (2003) Role of nicotianamine in the intracellular delivery of metals and plant reproductive development. *Plant Cell* **15**: 1263–1280
- Taylor LP, Grotewold E** (2005) Flavonoids as developmental regulators. *Curr Opin Plant Biol* **8**: 317–23
- Theodoridis GA, Gika HG, Want EJ, Wilson ID** (2012) Liquid chromatography-mass spectrometry based global metabolite profiling: a review. *Anal Chim Acta* **711**: 7–16
- Thiel S, Döhning T, Köfferlein M, Kosak A, Martin P, Seidlitz HK** (1996) A Phytotron for Plant Stress Research: How Far Can Artificial Lighting Compare to Natural Sunlight? *J Plant Physiol* **148**: 456–463
- Tholl D** (2015) Biosynthesis and biological functions of terpenoids in plants. *Adv Biochem Eng Biotechnol* **148**: 63–106
- Tholl D, Lee S** (2011) Terpene Specialized Metabolism in *Arabidopsis thaliana*. *Arabidopsis Book* **9**: e0143
- Tiiva P, Rinnan R, Faubert P, Räsänen J, Holopainen T, Kyrö E, Holopainen JK** (2007) Isoprene emission from a subarctic peatland under enhanced UV-B radiation. *New Phytol* **176**: 346–355
- Tossi V, Amenta M, Lamattina L, Cassia R** (2011) Nitric oxide enhances plant ultraviolet-B protection up-regulating gene expression of the phenylpropanoid biosynthetic pathway. *Plant, Cell Environ* **34**: 909–921
- Treutter D** (2005) Significance of Flavonoids in Plant Resistance and Enhancement of Their Biosynthesis. *Plant Biol* **7**: 581–591
- Trygg J, Holmes E, Lundstedt T** (2007) Chemometrics in Metabonomics. *J Proteome Res* **6**: 469–479
- Trygg J, Wold S** (2002) Orthogonal projections to latent structures (O-PLS). *J Chemom* **16**: 119–128
- Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, et al** (2006) The genome of black cottonwood,

- Populus trichocarpa* (Torr. & Gray). *Science* (80-) **313**: 1596–1604
- Tziotis D, Hertkorn N, Schmitt-Kopplin P, Eur JMS** (2011) Kendrick-analogous network visualisation of ion cyclotron resonance Fourier transform mass spectra: improved options for the assignment of elemental compositions and the classification of organic molecular complexity. *Eur J Mass Spectrom* 415–421
- Van Deemter JJ, Zuiderweg FJ, Klinkenberg A van** (1956) Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography. *Chem Eng Sci* **5**: 271–289
- Vanderauwera S, Vandenameele S, Zimmermann P, Breusegem F Van, Langebartels C, Gruissem W, Inze D** (2005) Genome-wide analysis of hydrogen peroxide-regulated gene expression in *Arabidopsis* reveals a high light-induced transcriptional cluster involved in anthocyanin biosynthesis. *Plant Physiol* **139**: 806–821
- Vanzo E** (2014) The impact of isoprene emission on stress-induced signaling pathways in Grey poplar. Dissertation, Albert-Ludwigs-Universität Freiburg im Breisgau
- Vayssières A, Pěňčík A, Felten J, Kohler A, Ljung K, Martin FM, Legue V** (2015) Development of the *Populus-Laccaria bicolor* ectomycorrhiza modifies root auxin metabolism, signalling and response. *Plant Physiol* **169**: pp.114.255620
- Velikova V, Fares S, Loreto F** (2008) Isoprene and nitric oxide reduce damages in leaves exposed to oxidative stress. *Plant Cell Environ* **31**: 1882–1894
- Velikova V, Ghirardo A, Vanzo E, Merl J, Hauck SM, Schnitzler JP** (2014) Genetic manipulation of isoprene emissions in poplar plants remodels the chloroplast proteome. *J Proteome Res* **13**: 2005–2018
- Velikova V, Muller C, Ghirardo A, Rock TM, Aichler M, Walch A, Schmitt-Kopplin P, Schnitzler JP** (2015) Knocking Down of Isoprene Emission Modifies the Lipid Matrix of Thylakoid Membranes and Influences the Chloroplast Ultrastructure in Poplar. *Plant Physiol* **168**: 859–870
- Velikova V, Sharkey TD, Loreto F** (2012) Stabilization of thylakoid membranes in isoprene-emitting plants reduces formation of reactive oxygen species. *Plant Signal Behav* **7**: 139–141
- Velikova V, Varkonyi Z, Szabo M, Maslenkova L, Nogues I, Kovacs L, Peeva V, Busheva M, Garab G, Sharkey TD, et al** (2011) Increased thermostability of thylakoid membranes in isoprene-emitting leaves probed with three biophysical techniques. *Plant Physiol* **157**: 905–916
- Vickers CE, Possell M, Cojocariu CI, Velikova VB, Laothawornkitkul J, Ryan A, Mullineaux PM, Nicholas Hewitt C** (2009) Isoprene synthesis protects transgenic tobacco plants from oxidative stress. *Plant Cell Env* **32**: 520–531
- De Vos RC, Moco S, Lommen A, Keurentjes JJ, Bino RJ, Hall RD** (2007) Untargeted

- large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nat Protoc* **2**: 778–791
- Walker A, Pfitzner B, Neschen S, Kahle M, Harir M, Lucio M, Moritz F, Tziotis D, Witting M, Rothballer M, et al** (2014) Distinct signatures of host-microbial meta-metabolome and gut microbiome in two C57BL/6 strains under high-fat diet. *ISME J* **8**: 2380–2396
- Ward JL, Baker JM, Llewellyn AM, Hawkins ND, Beale MH** (2011) Metabolomic analysis of *Arabidopsis* reveals hemiterpenoid glycosides as products of a nitrate ion-regulated, carbon flux overflow. *Proc Natl Acad Sci USA* **108**: 10762–10767
- Warren JM, Bassman JH, Eigenbrode S** (2002) Leaf chemical changes induced in *Populus trichocarpa* by enhanced UV-B radiation and concomitant effects on herbivory by *Chrysomela scripta* (Coleoptera: *Chrysomelidae*). *Tree Physiol* **22**: 1137–1146
- Warren JM, Bassman JH, Fellman JK, Mattinson DS, Eigenbrode S** (2003) Ultraviolet-B radiation alters phenolic salicylate and flavonoid composition of *Populus trichocarpa* leaves. *Tree Physiol* **23**: 527–535
- Wasternack C** (2007) Jasmonates: An update on biosynthesis, signal transduction and action in plant stress response, growth and development. *Ann Bot* **100**: 681–697
- Wasternack C, Parthier B** (1997) Jasmonate-signalled plant gene expression. *Trends Plant Sci* **2**: 302–307
- Way D a., Ghirardo A, Kanawati B, Esperschütz J, Monson RK, Jackson RB, Schmitt-Kopplin P, Schnitzler JP** (2013) Increasing atmospheric CO₂ reduces metabolic and physiological differences between isoprene- and non-isoprene-emitting poplars. *New Phytol* **200**: 534–546
- Weber H** (2002) Fatty acid-derived signals in plants. *Trends Plant Sci* **7**: 217–224
- Weise SE, Li Z, Sutter AE, Corrion A, Banerjee A, Sharkey TD** (2013) Measuring dimethylallyl diphosphate available for isoprene synthesis. *Anal Biochem* **435**: 27–34
- Whitham TG, Floate KD, Martinssen GD, Driebe EM, Keim P** (1996) Ecological and evolutionary implications of hybridization: *Populus*-herbivore interactions. *Biol Popul its Implic Manag Conserv* **40337**: 247
- Wiklund S, Johansson E, Sjöström L, Mellerowicz EJ, Edlund U, Shockcor JP, Gottfries J, Moritz T, Trygg J** (2008) Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Anal Chem* **80**: 115–122
- Wiley WC, McLaren IH** (1955) Time-of-Flight mass spectrometer with improved resolution. *Rev Sci Instrum* **26**: 1150
- Wilkins O, Nahal H, Foong J, Provart NJ, Campbell MM** (2009) Expansion and Diversification of the *Populus* R2R3-MYB Family of Transcription Factors. *Plant Physiol* **149**: 981–993

- Wilson DF** (2013) Regulation of cellular metabolism: programming and maintaining metabolic homeostasis. *J Appl Physiol* **115**: 1583–1588
- Wink M** (2003) Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry* **64**: 3–19
- von Wirén N, Klair S, Bansal S, Briat J-F, Khodr H, Shioiri T, Leigh RA, Hider RC** (1999) Nicotianamine chelates both FeIII and FeII. Implications for metal transport in plants. *Plant Physiol* **119**: 1107–1114
- Wittek F, Hoffmann T, Kanawati B, Bichlmeier M, Knappe C, Wenig M, Schmitt-Kopplin P, Parker JE, Schwab W, Vlot AC** (2014) *Arabidopsis* ENHANCED DISEASE SUSCEPTIBILITY1 promotes systemic acquired resistance via azelaic acid and its precursor 9-oxo nonanoic acid. *J Exp Bot* **65**: 5919–5931
- Witting M, Lucio M, Tziotis D, Wagele B, Suhre K, Voulhoux R, Garvis S, Schmitt-Kopplin P** (2015) DI-ICR-FT-MS-based high-throughput deep metabolotyping: a case study of the *Caenorhabditis elegans*-*Pseudomonas aeruginosa* infection model. *Anal Bioanal Chem* **407**: 1059–1073
- Wold S** (1995) Chemometrics; what do we mean with it, and what do we want from it? *Chemom Intell Lab Syst* **30**: 109–115
- Woodall GS, Stewart GR** (1998) Do anthocyanins play a role in UV protection of the red juvenile leaves of *Syzygium*? *J Exp Bot* **49**: 1447–1450
- Yamasaki H, Sakihama Y, Ikehara N** (1997) Flavonoid-Peroxidase Reaction as a Detoxification Mechanism of Plant Cells against H₂O₂. *Plant Physiol* **115**: 1405–1412
- Zhang F, Harir M, Moritz F, Zhang J, Witting M, Wu Y, Schmitt-Kopplin P, Fekete A, Gaspar A, Hertkorn N** (2014) Molecular and structural characterization of dissolved organic matter during and post cyanobacterial bloom in Taihu by combination of NMR spectroscopy and FTICR mass spectrometry. *Water Res* **57**: 280–294
- Zhang FL, Casey PJ** (1996) Protein prenylation: molecular mechanisms and functional consequences. *Annu Rev Biochem* **65**: 241–269
- Zhang P, Dreher K, Karthikeyan A, Chi A, Pujar A, Caspi R, Karp P, Kirkup V, Latendresse M, Lee C, et al** (2010) Creation of a genome-wide metabolic pathway database for *Populus trichocarpa* using a new approach for reconstruction and curation of metabolic pathways for plants. *Plant Physiol* **153**: 1479–1491
- Zhao J, Zhang W, Zhao Y, Gong X, Guo L, Zhu G, Wang X, Gong Z, Schumaker KS, Guo Y** (2007) SAD2, an importin -like protein, is required for UV-B response in *Arabidopsis* by mediating myb4 nuclear trafficking. *Plant Cell* **19**: 3805–3818
- Zheng L, Cheng Z, Ai C, Jiang X, Bei X, Zheng Y, Glahn RP, Welch RM, Miller DD, Lei XG** (2010) Nicotianamine, a novel enhancer of rice iron bioavailability to humans. *PLoS One* **5**: e10190

- Zhou B, Miao Q, Yang L, Liu Z-L** (2005) Antioxidative effects of flavonols and their glycosides against the free-radical-induced peroxidation of linoleic acid in solution and in micelles. *Chem – A Eur J* **11**: 680–691
- Zhou J-M, Gold ND, Martin VJJ, Wollenweber E, Ibrahim RK** (2006) Sequential O-methylation of tricetin by a single gene product in wheat. *Biochim Biophys Acta* **1760**: 1115–24
- Zoeller M, Stingl N, Krischke M, Fekete A, Waller F, Berger S, Mueller MJ** (2012) Lipid profiling of the *Arabidopsis* hypersensitive response reveals specific lipid peroxidation and fragmentation processes: biogenesis of pimelic and azelaic acid. *Plant Physiol* **160**: 365–378

Acknowledgements

This work would not have been possible without the help and guidance of many people.

I would like to thank my two supervisors Prof. Dr Philippe Schmitt-Kopplin and Prof. Dr. Jörg-Peter Schnitzler for giving me the opportunity to work on this very interesting topic and for their guidance and enthusiasm about the work as well as for letting me be part of their very nice research groups.

I would like to thank Dr. Basem Kanawati for introducing me into FT-ICR/MS spectrometry and Matlab programming.

I would like to thank Dr. Andrea Ghirardo for the introduction into multivariate data analysis, the very good coffee and the many discussions as well as the basketball sessions.

I would like to thank Dr. Franco Moritz for the knowledge I gained about MS data analysis, the hours we worked together and of course for the friendship.

I would like to thank Dr. Michael Witting for introducing me into LC-MS.

I would like to thank Dr. Norbert Hertkorn for the guidance in NMR analysis.

I would like to thank Prof. Dr. Andrea Polle and Dr. Anna Müller for the nice collaboration.

Many thanks to all of my colleagues, especially to my office, it was a nice time.

Very special thanks go to my family! My wife Tamara for always being there and supporting me through the ups and downs, the time we shared and the many special moments we had together, I love you! My children Emilian and Amelie for always reminding me what life is all about, I love you! My parents Michael and Rosemarie for always supporting me in life and for always being there, this would not have been possible without you, love you!

My parents-in-law Harald and Brigitte for their support, especially during the last three years, and for being there for us, love you. Many thanks to Benjamin and Christina for their support, love you.

List of Scientific Communication

Publications

Kaling M., Kanawati B., Ghirardo A., Albert A., Winkler J.B., Heller W., Barta C., Loreto F., Schmitt-Kopplin P. and Schnitzler J.P. (2015) UV-B mediated metabolic rearrangements in poplar revealed by non-targeted metabolomics. *Plant, Cell & Environment*, 38, 892-904.

Müller A.*, Kaling M.*, Faubert P., Gort G., Smid Hans M., Van Loon Joop J.A., Dicke M., Kanawati B., Schmitt-Kopplin P., Polle A., Schnitzler J.P., Rosenkranz M. (2015) Isoprene emission in poplar is not important for the feeding behaviour of poplar leaf beetles. *BMC Plant Biology*, 15:165.

*Both authors contributed equally to this work

Kaling M.*, Moritz F.*, Schnitzler J.P. and Schmitt-Kopplin P. (2015) Mass difference analysis as a generic framework for data-driven plant systems chemical biology. (submitted)

*both authors contributed equally to this work

Posters

Phytotron Conference 2014, München (Germany), UV-B mediated metabolic rearrangements in poplar revealed by non-targeted metabolomics, poster, M Kaling, B Kanawati, A Ghirardo, A Albert, JB Winkler, W Heller, C Barta, F Loreto, P Schmitt-Kopplin and JP Schnitzler

Curriculum Vitae

Persönliche Daten

Name	Moritz Kaling
Geburtsdatum	04.03.1986
Geburtsort	Tucson, USA
Staatsangehörigkeit	Deutsch

Doktorarbeit

03/2012 – 12/2015	Promotion am Helmholtz-Zentrum-München in den Arbeitsgruppen Analytische BioGeoChemie und Experimentelle Umweltsimulation Titel: Massenspektrometrische Bestimmung von Stress-induzierten Metabolitprofilen in Pappeln.
-------------------	--

Studium

10/2009 – 09/2011	Chemiestudium an der Technischen Universität München, Abschluss: Master of Science Thema der Masterarbeit: Charakterisierung der Metabolit- und Stoffwechselprofile in <i>Saccharomyces cerevisiae</i> Wildtyp und deren Δ Hsp12 Mutante.
10/2006 – 09/2009	Chemiestudium an der Technischen Universität München, Abschluss: Bachelor of Science

Zivildienst

06/2005 – 02/2006	Zivildienstleistender, Hausmeistertätigkeiten im Altenheim der Bürgerhospitalstiftung in Speyer
-------------------	---

Schulbildung

09/1996 – 04/2005	Friedrich-Magnus-Schwerd-Gymnasium in Speyer, Abschluss: Abitur
08/1992 – 08/1996	Grundschule

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die bei der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der TUM zur Promotionsprüfung vorgelegte Arbeit mit dem Titel: „Mass-spectrometric characterization of stress-induced metabolite profiles in poplar“ unter der Anleitung und Betreuung durch apl. Prof. Dr. Philippe Schmitt-Kopplin ohne sonstige Hilfe erstellt und bei der Abfassung nur die gemäß § 6 Abs. 6 und 7 Satz 2 angegebenen Hilfsmittel benutzt habe.

(x) Ich habe keine Organisation eingeschaltet, die gegen Entgelt Betreuerinnen und Betreuer für die Anfertigung von Dissertationen sucht, oder die mir obliegenden Pflichten hinsichtlich der Prüfungsleistungen für mich ganz oder teilweise erledigt.

(x) Ich habe die Dissertation in dieser oder ähnlicher Form in keinem anderen Prüfungsverfahren als Prüfungsleistung vorgelegt.

(x) Ich habe den angestrebten Doktorgrad noch nicht erworben und bin nicht in einem früheren Promotionsverfahren für den angestrebten Doktorgrad endgültig gescheitert.

Die öffentlich zugängliche Promotionsordnung der TUM ist mir bekannt, insbesondere habe ich die Bedeutung von § 28 (Nichtigkeit der Promotion) und § 29 (Entzug des Doktorgrades) zur Kenntnis genommen. Ich bin mir der Konsequenzen einer falschen Eidesstattlichen Erklärung bewusst.

Neuherberg, den 29.12.2015

Moritz Kaling