

Universitätsklinikum
Hamburg-Eppendorf



Hochschule für Angewandte
Wissenschaften Hamburg
Hamburg University of Applied Sciences

**Hamburg University of Applied Sciences
Faculty Life Sciences**

Investigation of infrared laser-based sampling for the detection of
biomarkers with MALDI mass spectrometry

Master thesis

Degree course Pharmaceutical Biotechnology

submitted by

Elisabetta De Iaco



Hamburg Bergedorf

the 7th of December 2021

1st reviewer: Prof. Dr. Jörg Andrä (HAW Hamburg)

2nd reviewer: Prof. Dr. rer. nat. Hartmut Schlüter (UKE Hamburg)

This master thesis was supervised and developed in the laboratories of the Mass Spectrometric Proteomics group at the Institute of Clinical Chemistry and Laboratory Medicine UKE Hamburg.

Content table

List of abbreviations	4
List of figures	5
List of tables	8
Abstract	10
1. Theoretical background	11
1.1. Biomarkers	11
1.2. Walnuts as matrix sample	12
1.2.1. Phytochemicals in walnuts	12
1.2.2. Lipids in walnut kernel	13
1.2.2.1 Phospholipids	13
1.2.3. Proteins in walnuts	16
1.3. Biomarker identification techniques	16
1.3.1. Current methods	16
1.4. MALDI-TOF for proteomic and lipidomic analysis	18
1.5. Current sampling methods	19
1.6. Aim	20
2. Materials and methods	21
2.1. Materials	21
2.1.1. Samples	21
2.1.2. Solutions, Standards and Matrices	21
2.1.3. Equipment	22
2.1.4. Additional material	23
2.1.5. Software	23
2.2. Methods	24
2.2.1. Reference and Sample preparations	24
2.2.1.1. Extraction of lipids with the methyl-tert-butyl ether (MTBE) method	24
2.2.1.2. Extraction of proteins with the formic acid/ acetonitrile (FA/ACN) method	24
2.2.1.3. Nanosecond Infrared Laser (NIRL) ablation	24
2.2.1.4. Protein purification via C8-reversed phase chromatography	26
2.2.2. Matrix preparation	26
2.2.3. Spotting procedures	28
2.2.4. MALDI acquisition	28
2.2.5. Experimental Designs	31
2.2.5.1. Method development	31
2.2.5.2. Method optimization and validation	33
2.2.6. Sample Processing and statistical analyses	34
2.2.6.1. Sample processing	34
2.2.6.2. Statistical Analyses	37
3. Results	38
3.1. Analysis of proteins using the conventional method FA/ACN	39
3.1.1. Intra- and inter-assay for the conventional protein extraction method FA/ACN	39
3.1.2. Identified signal areas	40
3.1.3. Averaged spectra and barcode gel view for the conventional method FA/ACN	42
3.1.4. Principal Component Analysis (PCA) for the FA/ACN method according to cultivar and origin	46
3.2. Analysis of lipids using the conventional method MTBE	48
3.2.1. Phospholipid analysis with the MALDI matrix DHB	48
3.2.2. Phospholipid analysis with the MALDI matrix 9-AA	51

3.2.3.	Averaged spectra and barcode gel view in the positive and negative mode for the analysis of phospholipids extracted with the MTBE-based sampling method	59
3.2.4.	Principal Component Analysis (PCA) for samples treated with the MTBE-based method according to cultivar and origin	64
3.3.	Analysis of proteins using NIRL.....	66
3.3.1.	Direct application of the reference sample on target with HCCA	66
3.3.2.	Addition of the desalting and protein enrichment C8-reversed phase chromatography.....	67
3.3.3.	Averaged spectra and barcode gel view for the analysis of proteins extracted with NIRL.....	68
3.3.4.	Principal Component Analysis (PCA) for proteins homogenized via NIRL according to cultivar and origin	72
3.4.	Analysis of lipids using NIRL	74
3.4.1.	Direct and indirect spotting with 9-AA on NIRL ablated samples on the ground steel target	74
3.4.2.	Intra- and inter-assay for the NIRL-based method for phospholipid analysis.....	79
3.4.3.	Averaged spectra and barcode gel view for the analysis of phospholipids extracted with NIRL.....	81
3.4.4.	Principal Component Analysis (PCA) for lipids homogenized via NIRL according to cultivar	86
4.	Discussion.....	88
5.	Summary and future outlook	94
	References	95
	Declaration of Originality	100
6.	Annex	101

List of abbreviations

ACN	Acetonitrile
a.i.	normalized intensity
Chol Ester	Cholesterol ester
2,5 DHB	2,5 Dihydroxybenzoic acid
FA	Formic acid
ELSD	Evaporative Light Scattering Detector
HCCA	α cyano 4 hydroxycinnamic acid
Jug r 1	Juglans regia 1
Jug r 2	Juglans regia 2
Jug r 3	Juglans regia 3
Jug r 4	Juglans regia 4
Jug r 5	Juglans regia 5
Jug r 6	Juglans regia 6
Jzg r 7	Juglans regia 7
Jug r 8	Juglans regia 8
MALDI	Matrix assisted laser desorption ionization
MAG	Monoacylglycerol
MTBE	Methyl tert butyl ether
m/z	Mass over z ratio
PA	Phosphatidic acid
PC	Phosphatidylcholine
PCA	Principle component analysis
PE	Phosphatidylethanolamine
PI	Phosphatidylinositol
PIRL	picosecond infrared laser
PG	Phosphatidylglycerol
PS	Phosphatidylserine
NIRL	nanosecond infrared laser
r. int.	relative intensity
SM	Sphingomyelin
SNR	Signal To Noise ratio
TAG	Triacylglycerol
TAM	tumor associated macrophages
TOF	Time of flight
TFA	Trifluoroacetic acid
9 AA	9 aminoacridine

List of figures

Figure 1: Overview of all phytochemical classes present in walnuts	12
Figure 2: Overview of the phospholipid subgroups	14
Figure 3: Chemical structure of the HCCA matrix molecule	18
Figure 4: Chemical structure of the (a) 9 AA matrix molecule and (b) the 2,5 DHB matrix molecule.....	18
Figure 5: Schematic overview of the NIRL experimental setup.....	25
Figure 6: Exemplary data processing with mMass for lipid analysis	35
Figure 7: Exemplary data processing with mMass for protein analysis	36
Figure 8: Results of pre processing with Mass Up	37
Figure 9: Schematic overview of the sampling preparations used for proteins and lipids in walnut.	38
Figure 10: Intra and inter assay spectra for the conventional protein extraction method ...	39
Figure 11: Exemplary representation of the MALDI TOF spectrum obtained for proteins, using the conventional FA/ACN method.....	40
Figure 12: Representative MALDI TOF spectrum for proteins obtained, using the conventional protein extraction method FA/ACN.....	41
Figure 13: Averaged MALDI TOF spectra for all four walnut cultivars prepared with the conventional FA/ACN method in spectrum overview	42
Figure 14: Averaged MALDI TOF spectra for all four walnut cultivars prepared with the conventional FA/ACN method in barcode representation.....	45
Figure 15: 2D representation of PCA results according to cultivar for proteins prepared via the FA/ACN method	46
Figure 16: 2D representation of PCA results according to origin for proteins prepared via the FA/ACN method	47
Figure 17: Crystallization behavior of the matrix preparation 0.5 M DHB in 100% MeOH and with different sample solvents.....	48
Figure 18: Representative MALDI TOF spectra for DHB mixed with different sample solvents	49
Figure 19: Crystallization behavior of 9 AA in two different solvents and mixed with two different samples solvents	51
Figure 20: MALDI TOF spectra of the lipid standard in positive and negative ion reflector mode	52
Figure 21: Fragmentation MALDI TOF spectrum for PI 15:0 18:1 (without the NH ₄) salt.....	54
Figure 22: Comparison between the reference and the 9 AA matrix spectra (negative mode)	55
Figure 23: Spectra of the analysis of matrix/analyte ratios and dilutions on MALDI TOF signals measured in the positive ion reflector mode	56
Figure 24: Intra and inter assay spectra for the reference sample.....	58
Figure 25: Averaged MALDI TOF spectra for all four walnut cultivars for the conventional MTBE method measured in positive ion reflector mode in spectrum overview.....	59
Figure 26: Averaged MALDI TOF spectra for all four walnut cultivars for the conventional MTBE method in barcode gel overview measured with the positive polarity	60
Figure 27: Averaged MALDI TOF spectra for all four walnut cultivars for the conventional MTBE method measured in negative ion reflector mode in spectrum overview	61
Figure 28: Averaged MALDI TOF spectra for all four walnut cultivars for the conventional MTBE method in barcode gel overview measured with the negative polarity.....	63

Figure 29: 2D representation of the PCA results according to cultivar and origin for lipids measured in the positive and negative ion reflector mode prepared via MTBE based method.....	64
Figure 30: Spectrum for the direct application of the reference sample with HCCA for protein analysis.....	66
Figure 31: Intra assay spectra for proteins with the reference sample ablated with NIRL.....	67
Figure 32: Averaged MALDI TOF spectra for all four walnut cultivars for protein samples ablated with NIRL measured in linear mode in spectrum overview	68
Figure 33: Averaged MALDI TOF spectra for all four walnut cultivars for protein samples ablated with NIRL in barcode overview.....	71
Figure 34: 2D representation of the PCA results according to cultivar and origin for proteins prepared via NIRL.....	72
Figure 35: Comparison between the direct and indirect application of reference walnut sample on the MALDI TOF ground steel target in the positive ion reflector mode after NIRL ablation	75
Figure 36: Comparison between the direct and indirect application of reference walnut sample on the MALDI TOF ground steel target in the negative ion reflector mode after NIRL ablation	77
Figure 37: Intra and inter assay spectra for the ablation with NIRL in the positive ion reflector mode. The abscissa shows the m/z values in a range from 500 to 1,600	79
Figure 38: Intra and inter assay spectra for the ablation with NIRL for the negative ion reflector mode	80
Figure 39: Averaged MALDI TOF spectra for all four walnut cultivars for lipid samples ablated with NIRL measured in the positive ion mode in spectrum overview.....	81
Figure 40: Averaged MALDI TOF spectra for all four walnut cultivars for lipid samples ablated with NIRL in barcode representation	82
Figure 41: Averaged MALDI TOF spectra for all four walnut cultivars for lipid samples ablated with NIRL measured in the negative ion mode in spectrum overview	83
Figure 42: Averaged MALDI TOF spectra for all four walnut cultivars for lipid samples ablated with NIRL in barcode representation	84
Figure 43: 2D representation of the PCA according to cultivar and origin for lipids measured in the positive and negative ion reflector mode prepared via NIRL.....	86
Figure 44: Crystallization behavior of the matrix preparation 20 mg/mL DHB in 100% TA30 and with different sample solvents.....	101
Figure 45: Crystallization behavior of the matrix preparation 10 mg/mL DHB in ACN/H ₂ O 50:50 (v/v) and with different sample solvents	101
Figure 46: Crystallization behavior of the matrix preparation 10 mg/mL DHB in ACN/H ₂ O with 0.1% TFA and with different sample solvents	102
Figure 47: Crystallization behavior of the matrix preparation 30 mM DHB in MeOH/H ₂ O (90:10, v/v) and with different sample solvents	102
Figure 48: Crystallization behavior of the matrix preparation 0.5 M DHB in 100% MeOH + 0.1% TFA and with different sample solvents	103
Figure 49: Exemplary 0.5 M DHB in 100% MeOH spectrum in the range from 100 m/z to 1,600 m/z	103
Figure 50: Fragmentation spectrum for the SM phospholipid present in the lipid standard..	104
Figure 51: Fragmentation spectrum for the PC phospholipid present in the lipid standard.	104

Figure 52: Fragmentation spectrum for the Lyso PC phospholipid present in the lipid standard	105
Figure 53: Fragmentation spectrum for the Lyso PE phospholipid present in the lipid standard	106
Figure 54: Fragmentation spectrum for the PG phospholipid present in the lipid standard.	106
Figure 55: Fragmentation spectrum for the PS phospholipid present in the lipid standard.	107
Figure 56: Fragmentation spectrum of the sample specific precursor ion 666 (positive ion reflector mode). The abscissa shows the m/z values in a range from 0 to 720	108
Figure 57: Fragmentation spectrum of the sample specific precursor ion 674 (positive ion reflector mode)	108
Figure 58: Fragmentation spectrum of the sample specific precursor ion 706 (positive ion reflector mode)	109
Figure 59: Fragmentation spectrum of the sample specific precursor ion 760 (positive ion reflector mode)	109
Figure 60: Fragmentation spectrum of the sample specific precursor ion 1,516 (positive ion reflector mode)	110
Figure 61: Fragmentation spectrum of the sample specific precursor ion 1,542 (positive ion reflector mode)	110
Figure 62: Fragmentation spectrum of the sample specific precursor ion 1,566 (positive ion reflector mode)	111
Figure 63: Fragmentation spectra of the sample specific precursor ions 833 and 861 (negative ion reflector mode)	111
Figure 64: Spectrum of the condition 1:3 matrix/analyte ratio, original sample in the positive ion reflector mode	112
Figure 65: Intra and inter assay spectra for proteins with the exemplary authentic sample 33 prepared with the conventional FA/ACN method.....	112
Figure 66: Intra and inter assay spectra for lipids measured in the positive ion reflector mode with the exemplary authentic sample 33 prepared with the MTBE based method	113
Figure 67: Intra and inter assay spectra for lipids measured in the negative ion reflector mode with the exemplary authentic sample 33 prepared with the MTBE based method	114
Figure 68: Intra and inter assay spectra for proteins with the exemplary non reproducible authentic sample 33 ablated with NIRL	115
Figure 69: Inter assay spectra for proteins with the exemplary reproducible authentic sample 33 ablated with NIRL	115
Figure 70: Intra and inter assay spectra for lipids measured in the positive ion reflector mode with the exemplary authentic sample 33 ablated with NIRL	116

List of tables

Table 1: List of sample information.....	21
Table 2: List of utilized solutions	21
Table 3: List of utilized standards and matrices	22
Table 4: List of utilized equipment	22
Table 5: List of additional materials	23
Table 6: List of utilized software	23
Table 7: List of NIRL parameters used for the method development procedure	25
Table 8: MS parameters for the acquisition of the spectra with 9 AA and DHB.....	28
Table 9: List of the utilized peptide calibration standard and their respective reference m/z ratios for the positive and the negative ion mode.....	29
Table 10: Analyte ions from HCCA and their respective m/z values for the measurement of the positive and the negative ion reflector mode.....	30
Table 11: MS parameters for the acquisition of the spectra.....	30
Table 12: List of the utilized protein calibration standard and their respective reference m/z ratios	31
Table 13: List of conditions analyzed for the optimization of the MALDI acquisition procedure in the conventional method.....	32
Table 14: Parameters for statistical analysis of the MS data	37
Table 15: List of main Chandler signals, their intensity and presence for protein analysis after extraction with FA/ACN method	43
Table 16: List of main Franquette signals, their intensity and presence for protein analysis after extraction with FA/ACN method	44
Table 17: List of main Lara signals, their intensity and presence for protein analysis after extraction with FA/ACN method	44
Table 18: List of main NB signals, their intensity and presence for protein analysis after extraction with FA/ACN method	44
Table 19: List of major signals for the 0.5 M DHB in 100% MeOH mixed with different sample solvents and their m/z value and relative intensity	50
Table 20: List of lipids present in the standard, their expected m/z values and measured m/z values	53
Table 21: m/z values and their respective attributed lipid fragment for the phospholipid PI 54	
Table 22: Signals and their relative intensity and presence, respectively, for the averaged Chandler spectrum for protein analysis after NIRL ablation	70
Table 23: Signals and their relative intensity and presence, respectively, for the averaged Franquette spectrum for protein analysis after NIRL ablation.....	70
Table 24: Signals and their relative intensity and presence, respectively, for the averaged Lara spectrum for protein analysis after NIRL ablation.....	70
Table 25: Signals and their relative intensity and presence, respectively, for the averaged NB spectrum for protein analysis after NIRL ablation.....	71
Table 26: m/z values and their respective attributed lipid fragment for the phospholipid SM	104
Table 27: m/z values and their respective attributed lipid fragment for the phospholipid PC	105
Table 28: m/z values and their respective attributed lipid fragment for the phospholipid Lyso PC	105

Table 29: m/z values and their respective attributed lipid fragment for the phospholipid Lyso PE.....	106
Table 30: m/z values and their respective attributed lipid fragment for the phospholipid PG	107
Table 31: m/z values and their respective attributed lipid fragment for the phospholipid PG	107
Table 32: Complete sample list	117

Abstract

Food fraud is defined as the deliberate act of falsifying food with the intent of economic gain. In recent years, this practice has become more and more common. The increasing need of fast and reproducible methods for biomarker discovery and identification in the context of food authentication, has therefore lead to the development of novel sampling methods.

Aim of this work is the development and investigation of infrared laser based sampling method nanosecond infrared laser (NIRL) for the detection of biomarkers with matrix assisted laser desorption ionization mass spectrometry (MALDI MS). For this means, the four different cultivars Chandler, Franquette, Lara and NB from five different origins (Italy, USA, France, Hungary and Germany), harvested in the years 2017 and 2018, were used. It was shown that sample homogenization with NIRL delivers similar results to the conventional Methyl tert butyl ether (MTBE) based method. Hereby, lipidomic analysis was reproducible and allowed a volume reduction from 5,000 nL to 158 nL. However, no biomarker could be identified. No differentiation into different cultivars was possible both with NIRL nor the conventional MTBE based method.

Proteomic studies using NIRL were only partially reproducible as opposed to the conventional formic acid/acetonitrile (FA/ACN) method. The analyzed proteins of each walnut hinted at a walnut specific phenotype. Therefore, also here no potential biomarker based on the bioprofiling of the different walnut cultivars could be identified.

1. Theoretical background

1.1. Biomarkers

Food fraud is defined as the deliberate act of falsifying food and frauding consumers with the intent of economic gain.^{1]} In recent years, this practice has become more and more common. While many accidents go unnoticed, food fraud may cause devastating public health risks.^{2]}

Adulteration is commonly linked to foods and ingredients such as meat, juices, olive oil, fish, honey, dairy products, alcoholic beverages, coffee and tea as well as spices. However, all foods can be adulterated.^{2]} Therefore, to avoid potential health risks, authentication of food is of pivotal importance.

In this context, foods such as truffles or walnuts represent interesting aliments and ingredients to be analyze.^{3]} While truffles often are object of frauds due to their expensive nature^{4]}, walnuts are interesting due to their high lipid content and can serve as exemplary matrix for similar ingredients or foods.^{3]}

In order to adequately identify fraud, food analytics and the discovery of biomarkers through bioprofiling or fingerprinting can be used.^{5]}

Biomarkers are important for the elucidation of biological systems. Usually, they can be categorized in different types: molecular, histologic, radiographic and physiologic characteristics.^{5]}

Molecular biomarkers entail proteins, lipids, metabolites, mRNA molecules and genes.^{5]}

Biomarkers of histologic properties analyze and describe the morphologic appearance of e.g. cells.^{5]} Radiographic biomarkers base on pattern recognition found in medical images. These biomarkers serve the purpose of predictive identification and characterization of pathologies and the treatment response.^{6]}

Lastly, physiologic biomarkers involve a physiological state such as symptoms related to a specific pathology.^{7]}

In this work, molecular biomarkers of protein and lipid origin in different walnut cultivars were analyzed.

1.2. Walnuts as matrix sample

Juglans regia L. also known as English walnut, Persian walnut or simply walnut is a type of nut tree from the *Juglans* genus native to the region of Central Asia and the Balkans.^{8] 9]} The walnut fruit is characterized by a hard outer shell and a brain looking kernel surrounded by skin. The outer shell is surrounded by the husk.^{8]} There are about 62 different cultivars, among which the Chandler variety is the most commonly cultivated on a global level. In Europe different walnut cultivars are cultivated. Along with Chandler, also Franquette, NB, and Lara can be found. The Franquette walnut is predominantly cultivated in France, while NB and Lara are sparsed over Germany, Hungary and France.^{8]}

All walnuts are known for their high anti oxidative compounds and their bioactive properties, among which antimicrobial, anticancer, anti amnesia, anti inflammatory, and antidiabetic activity.^{10] 11]}

1.2.1. Phytochemicals in walnuts

Walnuts are rich in particular in phenolic compounds and lipids. Major differences of phytochemical components can be found between leaf, husk, the walnut fruit (kernel), and the root.^{12]}

Figure 1 summarizes all phytochemical components typically found in *Juglans regia L.*

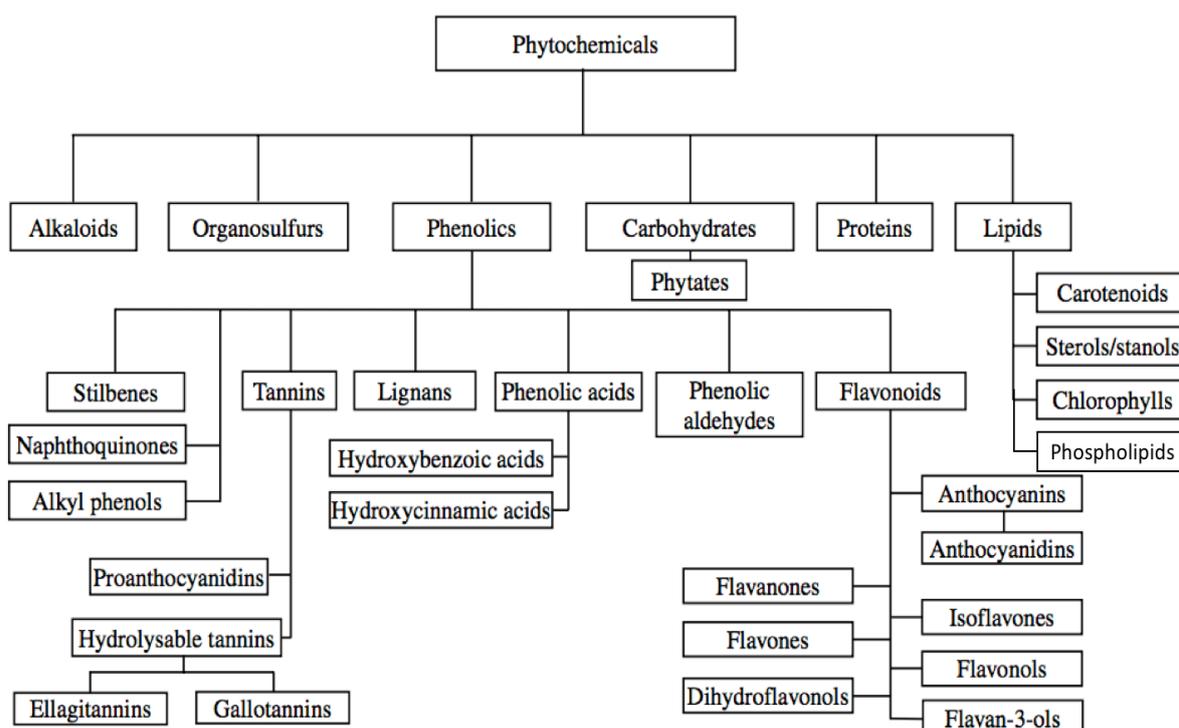


Figure 1: Overview of a phytochemical classes present in walnuts. Adapted from Robbins et al.^{13]}

The leaves are rich in phenolic acids, tannins, essential fatty acids, ascorbic acid, flavonoids, caffeic and paracomaric acid as well as juglone.^{14] 15]}

The walnut tree roots have a high phospholipid content in particular phosphatidylcholine (PC) and phosphatidic acid (PA), with 7.5 nmol/mg and 3.0 nmol/mg, respectively.^{12]}

As opposed to that, the walnut fruit (kernel) is rich in free fatty acids, tocopherols, phytosterols such as stigmasterol, tannins, glycerolipids, and phospholipids.^{14]}

1.2.2. Lipids in walnut kernel

The main constituents of walnuts are glycerolipids. In particular triacylglycerides (TAGs) of different fatty acyl chain length have been identified.^{12] 16]} The most abundant TAGs contain the fatty acids 54:3, 54:4, and 54:5.^{12]} Other identified TAG molecules contain 18:1, 18:2 and 18:3 fatty acids respectively (oleic, linoleic and linolenic acid).

TAGs have been used as biomarkers for the differentiation of origin of different cultivars.^{12] 16]} Hereby, the discrimination is possible even for cultivars of the same country, but different geographical origins.^{17]}

Free fatty acids account for only 0.20% of lipids present in walnuts. Here, the major compounds are saturated and unsaturated fatty acids with C18 chain lengths since they are relevant to phospholipid synthesis.^{18]}

The second major lipid component of walnuts is the group of phospholipids.

1.2.2.1 Phospholipids

Phospholipids are built of a hydrophilic head, containing a phosphate group, and lipophilic tail like acyl structures that can vary in length.^{19]} Based on specific head group structures attached to the free oxygen of the phosphate group, phospholipids can be differentiated into six subclasses: PA, Phosphatidylethanolamine (PE), PC, Phosphatidylserine (PS), Phosphatidylglycerol (PG) and Phosphatidylinositol (PI).^{20]} Figure 2 summarizes the structures of the different phospholipid classes.

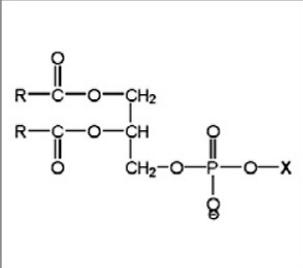
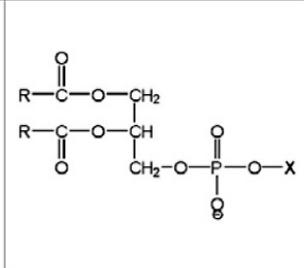
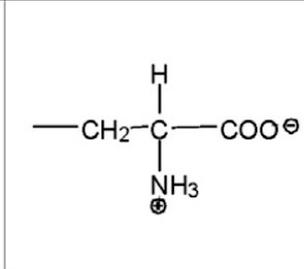
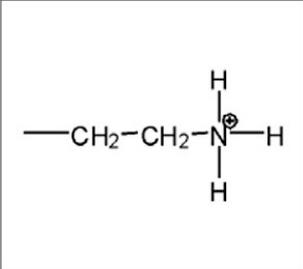
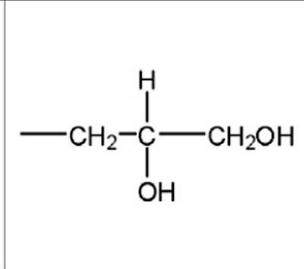
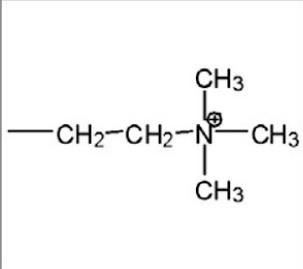
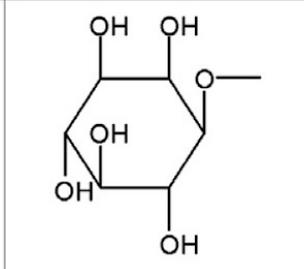
	Class		Class
H	Phosphatidic Acid (PA)		Phosphatidylserine (PS)
	Phosphatidylethanolamine (PE)		Phosphatidylglycerol (PG)
	Phosphatidylcholine (PC)		Phosphatidylinositol (PI)

Figure 2: Overview of the phospholipid subgroups.²⁰

In the walnut kernel 10–15% of lipid content are attributed to phospholipid species. It is therefore the second most abundant phytochemical in the walnut fruit after the glycerolipids.^{12]}

The simplest glycerophospholipid known as PA has hydrogen as a rest and is regarded as an anionic. Among their multiple functions, PA glycerolipids are important signaling and growth regulation molecules that can also contribute to the physical properties of membrane lipid bilayers.^{21] 22]} In walnuts, it was proven that PA concentration can vary based on different cultivars and therefore act as potential biomarker.^{17]}

The phospholipid PC is a zwitterionic molecule, containing a choline group attached to the phosphate. The molecular size of the head group is 184 Da. Yan et al. and Song et al. showed that this phospholipid is the third major component in walnuts. PC species with

saturated and unsaturated C34 fatty acid chains are the most abundant, however the presence of phospholipid species such as PC(16:0, 18:2), PC(18:0, 18:1) and Lyso phosphotidylcholine (LPC (18:1)) can also be observed.^{23]}^{17]}

The phospholipid PG contains a glycerol group (molecular weight of the head group = 170 Da). In the two Chinese cultivars Xin 2 and Qingxiang PG was not identified as a potential biomarker.^{17]}

The phospholipid PE is a zwitterionic lipid characterized by the presence of ethanolamine in the head group (molecular weight of the head group = 139.97 Da). PE is one of the least abundant phospholipid in walnuts. Contrasting works have shown that the PE content can vary based on cultivar, reaching up to 26.2% of all analyzed phospholipids.^{23]} To date, however, it was not identified as a biomarker for either cultivar or origin differentiation within walnuts.^{17]}

The anionic phospholipid PI contains an inositol group attached the oxygen of the phosphate group (molecular weight of the head group = 259 Da). It is an acidic phospholipid that can appear in both high or lower concentrations, depending on the walnut cultivars that are analyzed.^{17]}^{23]}

Lastly, the phospholipid PS is an anionic phospholipid which contains the amino acid serine. PS as shown to be a possible biomarker for the identification of different walnut types.^{17]}

Phospholipids can therefore be used for identification of cultivars since their presence may vary based on the geographical origin of a sample. This implies an adaptation of the membrane's permeability to the respective region.^{24]} These changes are triggered by aspects such as environmental stress like starvation, infections and UV light may have an impact on the phospholipid content.^{25]}^{12]}

1.2.3. Proteins in walnuts

In walnuts 15 to 20% account for proteins. The proteome and genome in walnuts has not been extensively studied. The information of the first analyzed Chandler reference is incomplete and does therefore not allow integrity of information.^{26]} However, to date, eight allergens have been identified in walnuts.^{27]} The most common allergen found in walnuts, to which both children and adults react the strongest, is the 2S albumin seed storage protein *Juglans regia 1* (Jug r 1).^{27]} Jug r 1 has a molecular weight of 15 kDa to 16 kDa. This protein belongs to the prolamin super family and is characterized by prominent alpha helical structure.^{28]}

Juglans regia 2 (Jug r 2) is a vicillin seed storage protein with a molecular weight of 44 kDa. Similarly to Jug r 1 it has been shown that Jug r 2 is a major allergen found in walnuts.^{29]}

Juglans regia 3 (Jug r 3) is the third major allergen.^{30]} This protein belongs to the non specific lipid transfer protein family. Proteins of this family are responsible for lipid transport within walnuts and other tree nuts.^{28]}

Further minor *Juglans regia* allergens are *Juglans regia 4* (molecular weight of 58.1 kDa)^{31]}^{32]}, *Juglans regia 5* (molecular weight of 20 kDa)^{33]}, *Juglans regia 6* (molecular weight of 47 kDa)^{34]}^{29]}, the profilin *Juglans regia 7* (molecular weight of 13 kDa)^{35]} and the non specific lipid transfer protein type 2 *Juglans regia 8* (molecular weight of 9 kDa).^{36]}

1.3. Biomarker identification techniques

Different methods are used in food analytics in order to identify food adulterations on both targeted and untargeted level. In all cases, fingerprints of the analyzed samples are created. These fingerprints aid in identifying additives or falsified food by comparing the spectra or chromatograms with databases.^{37]}

1.3.1. Current methods

Even if the walnut is among the most consumed nut tree fruits world wide^{38]}, the characterization of this ingredient has not been in the focus of research. In recent years, the amount of studies performed on the walnut in an effort to differentiate either by origin or species have risen significantly. For walnut analytics, targeted procedures are mostly used for lipidomic analyses.

Targeted approaches typical for authentication of plant derived species are Gas Chromatography (GC), High Performance Liquid Chromatography (HPLC), GC coupled to Electrospray Ionization tandem Mass spectrometry (ESI LC MS/MS), other mass spectrometric techniques and ELISA. In these cases, fingerprints can also give information about exact concentrations of the respective components found in the aliments. Targeted approaches are used for biomarker validation and entail a quantitative analysis.^{37]}

HPLC coupled to an evaporative light scattering detector (ELSD) was successfully implemented in the determination of triacylglycerol composition in different walnut species.^{16]} HPLC was also used for the identification of phenolic acid compounds in different walnut cultivars. Here, chlorogenic, caffeic, p coumaric, ferulic, sinapic, ellagic, and syringic acid were identified in the kernel and skin of ten different walnut cultivars.^{15]}

ESI MS without liquid chromatography or coupled to other platforms was used for walnut characterization with a successful identification of phospholipids and fatty acids differences between species.^{17]} Species characterization was also successfully carried out on truffles^{4]} and hazelnuts.^{38]}^{24]}

Gas chromatography both as only analytical method and coupled with mass spectrometry, finds application in particular for the analysis of free fatty acids and allowed the differentiation of walnuts species.^{39]}

Untargeted procedures lead to the generation of fingerprints, however no further information of the protein or lipids is given. This approach is ideal for the discovery of biomarkers and delivers the possibility of species identification. HPLC and in particular MALDI TOF procedures are used as untargeted methods.

Both HPLC and MALDI TOF as untargeted procedures have not been applied on walnut analytics. The application of MALDI TOF, however, was successfully applied on plant and animal derived foods such as truffle^{4]}, honey^{40]} and fish^{41]} for protein analysis.

1.4. MALDI-TOF for proteomic and lipidomic analysis

Matrix assisted laser desorption ionization (MALDI TOF) also known as MALDI or MALDI TOF MS is a mass spectrometric method used for identification molecules and fingerprinting.^{42] 43]} It bases on co crystallization of a matrix with the respective analyte. Upon laser irradiation, desorption takes place and leads to molecule ionization.^{44]} Both proteomics and lipidomic analyses can be performed with MALDI MS.^{45] 46] 47]} As mentioned in section 1.3.1, MALDI has found applicability in proteomic and lipidomic analysis for the differentiation and identification of species and authenticity analytics to avoid food fraud. The matrix used as golden standard for protein analysis is α cyano 4 hydroxy cinnamic acid (HCCA) a cinnamic acid derivative.^{4] 43] 47] 48]} Figure 3 shows the chemical structure of the HCCA molecule:

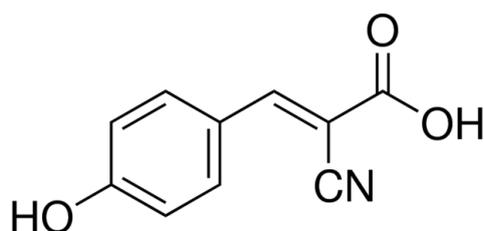


Figure 3: Chemical structure of the HCCA matrix molecule⁴⁹

For lipidomic analysis different matrices are used, depending on the lipid class of interest. Phospholipids can best be analyzed with 9 aminoacrine (9 AA) and 2,5 dihydroxybenzoic acid (2,5 DHB) a heterocycle and benzoic acid derivative respectively.^{42] 50]} Figure 4 shows the chemical structure of (a) 9 AA and (b) 2,5 DHB:

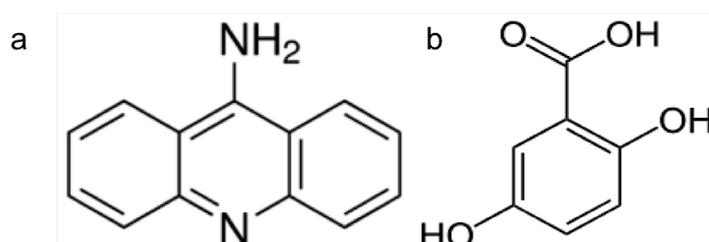


Figure 4: Chemical structure of the (a) 9-AA matrix molecule⁵¹ and (b) the 2,5-DHB matrix molecule⁵²

In this work MALDI TOF was used due its advantages. Among the main advantage of untargeted procedures, are the quick, reliable, sensitive and inexpensive measurements of a variety of products. For protein analysis, the HCCA matrix was used and for lipid analysis the 9 AA and DHB matrix.

1.5. Current sampling methods

The presence of an analyte for the acquisition via MALDI TOF is dependent on the utilized sampling procedure. To date, in literature, conventional methods based on two phase extraction procedures (solid liquid liquid extraction methods) such as tissuelyser with MTBE based extraction, Bligh and Dyer and Folch method have been utilized in order to identify biomarkers of lipid origin.^{53]}^{22]}

The Bligh and Dyer as well as Folch lipid extraction procedures base on utilization of Chloroform in order to separate lipophilic from hydrophilic substances^{22]}. Here, Chloroform represents the lower phase. The Bligh and Dyer and Folch method finds applicability for nearly all lipid classes.^{22]}

The MTBE based extraction procedure, on the other hand, uses MTBE and methanol (10:3 ratio) as extraction solvent. Since MTBE has a lower density than water, the MTBE phase is the upper phase within the two phase partitioning system.^{54]} This significantly facilitates the overall workflow and handling of each sample. However, loss of polar lipids between the two phases has been reported.^{53]} The MTBE based sampling has found applicability in particular for phospholipids and their derivatives.^{54]}

For proteins, the analysis of intact proteins has allowed the discovery and identification of several biomarkers in the context of food authentication.^{4]}^{55]}^{56]} The extraction of intact proteins is achieved via the formic acid/acetonitrile (FA/ACN) method. This method represents a significant reduction in time in comparison to, for example, digestion of proteins for the identification of biomarkers of human, plant or animal origin.^{4]}^{57]}^{58]}

While all methods allow the identification and discovery of biomarkers, time and usage of toxic reagents has to be taken into consideration as soon as a high sample amount is being analyzed. Moreover, all mentioned conventional methods were tailored for specific sample types. Therefore, optimization for each analyzed matrix is necessary.^{22]}

Infrared laser based homogenization of samples has recently been used as a means to analyze proteoms.^{59] 60]} Different laser types such as picosecond infrared laser (PIRL), and nanosecond infrared laser (NIRL) exist. Both lasers allow reproducible results in the setting of proteomics.^{59] 60]} Infrared laser based homogenization of samples allows volume and time reduction^{59]} and eliminates obstacles posed by conventional methods such as MTBE. To date, infrared laser based sampling has not yet been used for biomarker identification in the context of food authentication. In this work, a NIRL system has been used as a biomarker identification strategy for walnuts.

1.6. Aim

The aim of this work is the development and optimization of a fast, reliable and reproducible sampling procedure for proteins and lipids based on nanosecond infrared laser (NIRL) technology. Subsequent analysis should be performed via MALDI TOF mass spectrometry. Two NIRL methodologies should be established for proteins and lipids, respectively. As comparative measures for the protein (FA/ACN method) and lipid class (MTBE method) conventional sample homogenization should be used. Moreover, the study should comprise a validation for usage in routine operation to allow adequate food profiling.

2. Materials and methods

2.1. Materials

2.1.1. Samples

Walnut samples were supplied by UK Eppendorf. They were chosen based on cultivar, origin and year of harvest. The following table lists all the analyzed origin countries, the varieties and the year of harvest. A complete list with information about each sample is supplied in the annex in section 6 in the supplementary table 32.

The references for the method development were commercially available walnuts from

- Reformhaus Kontor GmbH & Co. KG: Lihn, Wallnuss Kernbruch from Chile and;
- Rapunzel Naturkost: Wallnusskernstückchen from Moldavia

Table 1: List of sample information

Data	Information
Year	2017, 2018
Cultivar	Chandler, Franquette, Lara and NB
Origin	Italy, France, Germany, Hungary and USA

2.1.2. Solutions, Standards and Matrices

Table 2: List of utilized solutions

Solutions	Manufacturer	Lot number
Acetonitril (ACN) Gradient grade for HPLC	Chemsolute Th. Geyer GmbH & Co. KG, Germany	POD097240D
Trifluoroacetic acid (TFA)	Sigma Aldrich, St. Louis, MO, USA	BCBH6349V
Formic acid (FA)	Honeywell International Inc. Fluka™ Charlotte, NC, USA	n.a.
H ₂ O for LC MS	Chemsolute Th. Geyer GmbH & Co. KG, Germany	VOL613200L
Methanol for LC MS	Chemsolute Th. Geyer GmbH & Co. KG, Germany	POD737070E
Methyl tert butyl ether	Sigma Aldrich, St. Louis, MO, USA	SZBC298MV
Isopropanol	Chemsolute Th. Geyer GmbH & Co. KG, Germany	28.1121108
Chloroform for analysis	Sigma Aldrich, St. Louis, MO, USA	K52316142 014

Table 3: List of utilized standards and matrices

Standards/ Matrices	Manufacturer	Lot number
9 Aminoacridine (9 AA)	Sigma Aldrich, MO, USA	BCCD2865
α cyano 4 hydroxycinnamic acid (HCCA)	Bruker Daltonics GmbH & Co. KG, Bremen, Germany	0000279462
2,5 Dihydroxybenzoic acid (DHB)	Bruker Daltonics GmbH & Co. KG, Bremen, Germany	BCBF5070V
Protein Calibration Standard 1 (in 0.1% TFA)	Bruker Daltonics GmbH & Co. KG, Bremen, Germany	n.a.
Lipid Calibration Standard "Light Splash"	Avanti Polar Lipids, Inc., Alabaster, AL, USA	n.a.
Peptide Calibration Standard II	Bruker Daltonics GmbH & Co. KG, Bremen, Germany	0000368297

2.1.3. Equipment

Table 4: List of utilized equipment

Equipment	Manufacturer
Vortexer REAX 2000	Heidolph Instruments GmbH & Co. KG Schwabach, Germany
Table centrifuge PicoFuge	Stratagene – Agilent Technologies Division, Santa Clara, CA, USA
Vacuum centrifuge UNIVAPO 100 H	Thermo Electron Corporation Waltham, MA, USA
Thermomixer 5436	Eppendorf AG, Hamburg, Germany
RapifleX Maldi Tissue typer	Bruker Daltonics GmbH & Co. KG, Bremen, Germany
TissueLyser II	Qiagen N.V., Hilden, Germany
Nanosecond infrared laser (NIRL) Opolette Se 2731 (tunable)	OPOTEK, Carlsbad, CA, USA
FieldMaxII: Laser Energy Meter	Coherent Inc., Santa Clara, CA, USA
2D Scanning Mirror FSM 300 Series	New Port Corporation, Irvine, CA, USA
I/O card, NI USB 6343, multifunction DAQ	National Instruments, Austin, TX, USA
2:1 Telescope 2 Ar PC 25 76	n.a.
MALDI target plate MTP 384 Ground steel target	Bruker Daltonics GmbH & Co. KG, Bremen, Germany

2.1.4. Additional material

Table 5: List of additional materials

Material	Manufacturer	Lot
Solid Phase Extraction Disk Empore™ 2214 C8 (octyl)	Bioanalytical Technologies, 3M company St. Paul, MN, USA	620071
SuperFrost/ Plus Objektträger Assistent®	Glaswarenfabrik Karl Hecht GmbH & Co. KG, Sondheim, Germany	No. 42409110

2.1.5. Software

Table 6: List of utilized software

Software	Manufacturer
OPOTEK Control Software v1.3.177	OPOTEK, Carlsbad, CA, USA
Costume control software	Jan Hahn, UKE Hamburg
OQ Labscope 2.0	Lumedita Durham, NC, USA
ITK SNAP	Open Source ^{61]}
flexControl Version 4.0 Build 46	Bruker Daltonics GmbH & Co. KG, Bremen, Germany
flexAnalysis Version 4.0 Build 40	Bruker Daltonics GmbH & Co. KG, Bremen, Germany
mMass Version 5.5.0	Martin Strohm, Germany
Mass Up Version 1.0.14	Bioscope, Lisbon, Portugal and Sing, Vigo, Spain
Prism 5.0.0.288	GraphPad Software, Inc. San Diego, CA, USA
Microsoft Excel 10, Ink	Microsoft Corporation Redmond, WA, USA

2.2. Methods

2.2.1. Reference and Sample preparations

All references and samples were either stored at 20 °C or analyzed right away by MALDI TOF mass spectrometer. The matrices and MALDI acquisition methods are described in sections 2.2.2 and 2.2.4.

2.2.1.1. Extraction of lipids with the methyl tert butyl ether (MTBE) method

For conventional lipid extraction, the methyl tert butyl ether (MTBE) two phase extraction procedure has been used. After walnut skin removal, the samples were mixed with a 10:3 (v/v) MTBE/MeOH solvent solution and lysed with the TissueLyser II for 1 min at a frequency of 25/s. Phase separation was induced via 5 min centrifugation at 16000 xg after adding water in a ratio 10:3:2.5 (v/v/v; MTBE/MeOH/H₂O). After the organic phase (upper phase) was collected in a new tube, the solvent was evaporated by a vacuum centrifuge overnight. The dried sample was resuspended in a 50:50 (v/v) ACN/H₂O solvent solution.

2.2.1.2. Extraction of proteins with the formic acid/ acetonitrile (FA/ACN) method

Protein extraction was carried out through the formic acid/acetonitrile method (FA/ACN). The method was optimized and validated by the working group of Prof. Dr. rer. nat. Schlüter at the UKE and is based on the method developed by El Karkouri et al ⁴³. Briefly, 10 to 15 mg of each walnut powder sample without skin were weighted and mixed with 400 µL extraction buffer (50/50 (v/v) ACN/ 70 % FA in H₂O). After shaking for 5 min at RT, the mixture was centrifuged at 12000 xg for 3 min. The supernatant was then collected in a fresh tube.

2.2.1.3. Nanosecond Infrared Laser (NIRL) ablation

For the development of the new sampling method, the Opolette SE 2731 Nanosecond Infrared Laser (NIRL) by OPOTEK was used. The Laser was operated with the following parameters:

Table 7: List of NIRL parameters used for the method development procedure

NIRL parameters	Value
Pulses	321 Pulses
Ablation time per sample	20 seconds
Wavelength	2950 nm
Laser energy	850 mJ
Scan range	1 mm ²
Sample N _x	9
Scan x _{min}	0.5
Scan x _{max}	0.5
Sample N _y	9
Scan y _{min}	0.5
Scan y _{max}	0.5
Sample per step	1
Waits after line	5
Repetitions	3
Layer correction	OFF

Before ablation, the samples were cut in cubes of approx. similar height to avoid laser focus adjustments after each ablated sample. The skin was peeled off. During ablation, all samples directly condensed on an aminoalkylsilane coated slide which was positioned above a remote controlled cooling stage. Figure 5 displays the experimental setup for the NIRL experiments:

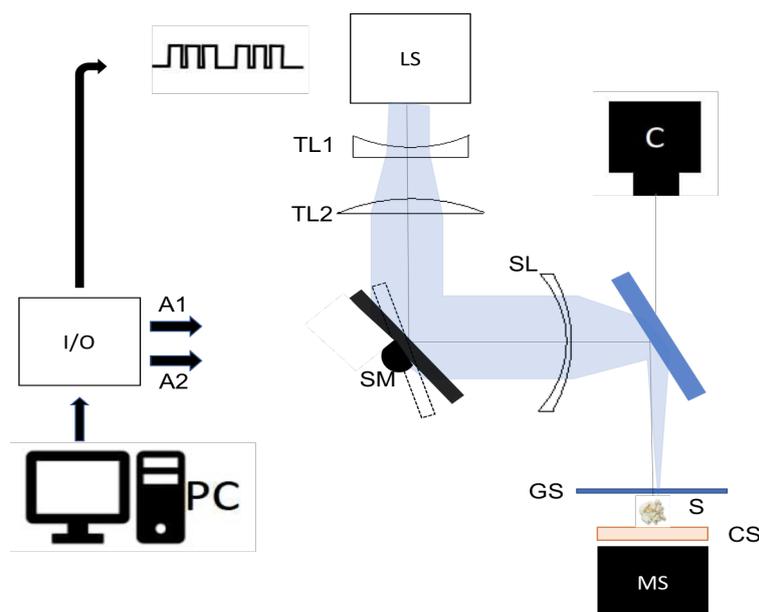


Figure 5: Schematic overview of the NIRL experimental setup. I/O = input/output card, A1/A2 = analog control signals, LS = NIRL system, TL1/ TL2 = lenses 1/2, SM = two-axis scanning mirror, GS = glass slide, SL = Scanning lenses, S = sample, CS = cooling stage, MS = manual stage, C = Camera.

For lipids, the sample was taken off the slide with 9 AA (10 mg/mL 60:40 isopropanol/ACN) matrix and directly spotted as a dried droplet on the MALDI TOF ground steel target.

Protein samples were dissolved in 0.1% TFA solution and further processed as described in the next section.

2.2.1.4. Protein purification via C8 reversed phase chromatography

Protein purification for NIRL samples was carried out using C8 chromatographic material. All steps include a subsequent centrifugation at 2500 xg for 1 min. Moreover, all steps are repeated once. The sample aspiration step entails three steps and centrifugations. Sample dissolution from slides was achieved with the help of a 0.1% TFA solution.

Prior to sample application, the chromatography material was equilibrated. The equilibration steps entailed usage of 10 μ L MeOH, followed by 10 μ L 50% ACN in water and lastly 10 μ L 0.1% TFA solvent. The different solvents were subsequently discarded.

10 μ L sample was then dissolved in 0.1% TFA solution and applied on the column and reapplied thrice after centrifugation. The column was rinsed with 0.1% TFA/5% ACN and the solvent discarded. Sample elution was achieved by application of 10 μ L 0.1% TFA in 50% ACN elution solution and one final centrifugation step into a fresh tube. Samples were analyzed right away or stored at 20 °C.

2.2.2. Matrix preparation

For the different analyses and method development three different matrices were used: 9 AA, 2,5 DHB and HCCA. The 9 AA and 2,5 DHB matrix, adequate for lipid analysis, were acquired in the positive and the negative mode (9 AA) or only the positive mode (2,5 DHB), respectively. The HCCA matrix was used for protein analysis of the samples.

9-AA:

The 9 AA matrix was prepared by dissolving 10 mg 9 AA in 1 mL 90:10 MeOH/H₂O or 60:40 isopropanol/ACN solution. To aid the dissolving process, the matrix solution was vortexed for 20 sec.

Finally, the matrix was applied on the MALDI ground steel target through different spotting procedures to allow subsequent analysis with different conditions (see section 2.2.3). The matrix solution was stored at 20 °C and freshly prepared every four weeks.

2,5-DHB:

As mentioned in section 2.2.2, the 2,5 DHB matrix can be used for the analysis of lipids and other low molecular substances. Several DHB matrix solutions were used in order to analyze the quality of crystallization and overall MALDI acquisition potential. Following 2,5 DHB preparations have been used:

- 20 mg/mL 2,5 DHB in TA30 (70:30 0.1% TFA in H₂O/ACN)
- 10 mg/mL 2,5 DHB in 50:50 ACN/H₂O
- 10 mg/mL 2,5 DHB in 50:50 ACN/H₂O with 0.1% TFA
- 30 mM 2,5 DHB in 9:1 MeOH/H₂O
- 0.5 M 2,5 DHB in 100% MeOH
- 0.5 M 2,5 DHB in 100% MeOH with 0.1% TFA

After weighing and mixing of matrix compound and solvent, the preparations were dissolved via vortexing for 20 sec. In the following sections the 2,5 DHB matrix will be referred to as DHB.

HCCA:

In order to create the HCCA matrix solution, a TA50 solvent (50:50 0.1% TFA in H₂O/ACN) was used. The HCCA matrix was prepared as a saturated solution.

In order to allow MALDI TOF analysis, 2,5 DHB and 9 AA were applied on the MALDI target through different spotting procedures with different conditions (see section 2.2.5). HCCA was applied on the MALDI target with only the thin layer procedure (see section 2.2.3). The matrix solutions were all stored at 20 °C.

2.2.3. Spotting procedures

Spotting on the target was achieved by pipetting matrix/analyte mixes with two different methods: the thin layer (HCCA) and the dried droplet (9 AA and DHB) spotting procedures.

The thin layer procedure consisted in depositing 1 part of analyte, letting it air dry for 5 min and subsequently pipetting 1 part matrix solution (drying time: 5 min) onto the MALDI ground steel target. The spots were all air dried prior to further MALDI TOF sample analysis.

For the dried droplet procedure, on the other hand, different matrix/analyte ratios were premixed and finally 1 μL of matrix/analyte mix were spotted on the MALDI ground steel target. Each spot was air dried for 5 min prior to further analysis. The different matrix/analyte ratios can be found in section 2.2.5 under the description of the preliminary experiments.

2.2.4. MALDI acquisition

Mass spectrometric analyses were carried out using the rapifleX MALDI Tissue typer by Bruker Daltonics GmbH. Hereby, the respective mass spectra were acquired by FlexControl and subsequently analyzed with FlexAnalysis 4.0 and mMass 5.5.0 by Bruker and Martin Strohm, respectively.

For Lipids:

The lipid mass spectra were recorded in the positive and negative ion reflector mode for the 9 AA matrix and only in the positive ion reflector mode for the DHB matrix. The utilized MALDI TOF parameters are listed in table 8:

Table 8: MS parameters for the acquisition of the spectra with 9-AA and DHB. Samples spotted with 9-AA were measured both in the positive and the negative ion reflector mode. Samples spotted with DHB were measured only in the positive ion reflector mode

Parameters	Value
Positive and negative ion reflector mode for the 9 AA matrix	
Ion source	20000
Detector voltage	2070
Ion extraction	2550
Shots	1000
Frequency	10000

Accumulated shots	4000
Laser energy	55% (conventional method) 30% (negative mode – NIRL) 35% (positive mode – NIRL)
Laser mode	Custom laser
Positive ion reflector mode for the DHB matrix	
Ion source	20000
Detector voltage	2075
Ion extraction	2550
Shots	1000
Frequency	10000
Accumulated shots	5000
Laser energy	50 60%
Laser mode	Custom laser

Prior to optimization a mass range from 100 m/z to 2800 m/z was analyzed. The ion accumulated laser shots were randomly distributed on the respective spots. For calibration, a peptide standard with molecules within the same range as lipids was used. The different peptides and their respective m/z values for the positive and negative ion modes are shown below, in table 9.

Table 9: List of the utilized peptide calibration standard and their respective reference m/z ratios for the positive and the negative ion mode.

Peptide calibration standard	Reference m/z
Positive ion mode	
Leucine Enkephalin [M+H] ⁺ + average	556.27657
Bradykinin [M+2H] ²⁺ + average	757.39916
Angiotension II [M+2H] ²⁺ + average	1,046.54180
Angiotension I [M+H] ⁺ + average	1,296.68480
Substance P [M+H] ⁺ + average	1,347.73540
Bombesin [M+H] ⁺ + average	1,619.82230
Negative ion mode	
Leucine Enkephalin [M+H] + average	554.26093
Bradykinin [M+2H] ² + average	755.38350
Angiotension II [M+2H] ² + average	1,044.52725
Angiotension I [M+H] + average	1,294.67025
Substance P [M+H] + average	1,345.72085
Bombesin [M+H] + average	1,617.80775

To extend the range and linearity to a m/z value of 100, two HCCA matrix ions were added. Table 10 displays the m/z values for the additional signals both in the positive and in the negative ion mode.

Table 10: Analyte ions from HCCA and their respective m/z values for the measurement of the positive and the negative reflectron mode

Matrix ion	Reference m/z
Positive ion mode	
CCA [M+H] ⁺ + average	190.04987
CCA [2M+H] ²⁺ + average	379.09246
Negative ion mode	
CCA [M+H] + average	188.03532
CCA [2M+H] ² + average	377.07791

For Proteins:

The protein mass spectra were recorded in linear mode under the conditions shown in table 11:

Table 11: MS parameters for the acquisition of the spectra

Parameters	Value
Ion source	20000
Detector voltage	448
Ion extraction	1,560
Shots	3,000 2,000
Frequency	10,000
Accumulated shots	24,000 (conventional method) 8,000 (NIRL)
Laser energy	40% (conventional method) 45% (NIRL)
Shots at raster spot	100
Limit diameter	1,200 μm
Laser mode	Thin Layer

A mass range from 2,000 m/z to 20,000 m/z was analyzed for all samples. The ion accumulated laser shots were randomly distributed on the respective spots. The utilized calibration standard by Bruker contained the different proteins shown below, in table 12.

Table 12: List of the utilized protein calibration standard and their respective reference m/z ratios

Protein calibration standard	Reference m/z
Insulin [M+H] ⁺ + average	5,734.52000
Cytochrom C [M+2H] ²⁺ + average	6,181.05000
Myoglobin [M+2H] ²⁺ + average	8,476.66000
Ubiquitin I [M+H] ⁺ + average	8,565.76000
Cytochrom C [M+H] ⁺ + average	12,360.97000
Myoglobin [M+H] ⁺ + average	16,952.31000

The protein calibration standard was chosen to allow linearity and the exact determination of reproducibility of the method. However, linearity is not present in the range 2,000 m/z to 5,000 m/z. Here, a direct comparison of the spectra was done in the mMass software to verify whether a similar signal pattern is given.

2.2.5. Experimental Designs

2.2.5.1. Method development

For the NIRL method development, the commercially available walnut (Rapunzel Naturkost) from Moldavia was used. For the conventional MTBE based method, the reference from Reformhaus Kontor GmbH & Co. KG was utilized. Samples for protein and lipid analysis were all carried out using the parameters shown in section 2.2.1.3 with the slide set up. The wavelength and the parameters used in the experiments was taken from Kwiatkowski et al^{60]}. The other parameters were based on the results obtained in the paper written by Hahn et al^{59]}.

Choice of conditions for the main experiments (preliminary experiments)

Conditions were evaluated based on several aspects thought to be influencing the results of the preliminary experiments. These parameters are listed below:

- Relative intensity of peaks above a Signal to Noise (S/N) ratio of 3,
- Range influence,
- Crystallization efficiency (also valid for choice of matrices and sample solvents), and
- Homogeneity of the spotted samples

The samples for the main experiments were subsequently only prepared with optimized conditions unless stated differently and by usage of the respective sample preparation procedure described in section 2.2.1. The following section (2.2.5.1.1) describes the conditions analyzed for method development.

2.2.5.1.1 Conventional method for lipids

After sample preparation, (see section 2.2.1) preliminary experiments for the analysis of adequate sample dilution, spotting procedure and matrix/analyte ratio were carried out. Table 13 lists all analyzed conditions for the development of the MALDI acquisition parameters for the conventional lipid method. The choice of conditions was based on the aspects listed in section 2.2.5.1

Identity of phospholipids was verified utilizing the Avanti lipid standard “Light Splash”. The standard was measured in both positive and negative ion reflector mode. Moreover, the commercially available walnut from Chile (see section 2.1.1) was used. Several negative controls were used in every preliminary and main experiment:

- 9 AA alone
- 9 AA with sample solvent ACN/H₂O and
- sample without walnut prepared according to the MTBE protocol in section 2.2.1.1

The MTBE sample without walnut was re suspended in the same sample solvent and spotted on the ground steel target with 9 AA based on the conditions mentioned in table 13.

Table 13: List of conditions analyzed for the optimization of the MALDI acquisition procedure in the conventional method.

Parameter	Condition
Matrix	9 AA (negative + positive mode) DHB (positive mode)
Sample dilutions	Original 1:10
Matrix/analyte ratio	1:1 1:2 1:3
Spotting procedure	Dried Droplet Thin Layer

2.2.5.1.2 NIRL ablation for lipids

The reference was ablated as mentioned in section 2.2.1.3. The lipid preliminary experiments subsequently involved the direct application of the condensed reference sample with 0.5 μL 9 AA on the MALDI ground steel target as well as an indirect application through 3 μL solvent and final mixing with 3 μL 9 AA and application as dried droplet procedure on the MALDI target. The spotted samples were each dried for 5 min prior to analysis. In the following the different solvents and solvents mixtures used for the preliminary experiment:

- MeOH: Chloroform 2:1, 1:1, 1:2
- Chloroform
- ACN:H₂O 1:1

2.2.5.1.3 NIRL ablation for proteins

For protein analysis, the reference sample was directly spotted on the MALDI target (here with HCCA) as described in section 2.2.5.1.2.

Moreover, the C8 reversed phase chromatography was used to desalt the samples and enrich the proteins and subsequently analyze them (see section 2.2.1.4). MALDI TOF acquisition was carried out as shown in section 2.2.4.

2.2.5.2. Method optimization and validation

Validation characteristics were defined as described in the ICH tripartite guideline for the validation of analytical procedures Q2(R1)^{62]}. All precision and robustness studies were carried out both during the preliminary and the main experiments.

Range studies

The range studies for lipid analysis were carried out during the preliminary experiments. The range was chosen based on the signal suppression caused by prominent 9 AA and DHB matrix signals.

For lipids, a final range of 500 m/z to 1,600 m/z was selected. For proteins, a range from 2,000 m/z to 20,000 m/z was chosen.

Precision studies

To guarantee an appropriate statistical analysis and precision of the developed method both inter and intra assays were carried out.

Reproducibility studies

Reproducibility studies were performed by spotting the same sample as triplicates (n=3) for the lipids and protein samples over different days, respectively. The resulting spectra were evaluated by direct comparison with the help of FlexAnalysis and mMass under the sample processing described in section 2.2.6 (see below).

Repeatability studies

Repeatability studies were carried out by spotting the same sample as triplicates (n=3) for lipids and for protein samples on the same day, respectively. The resulting spectra were evaluated by direct comparison with the help of FlexAnalysis and mMass under the sample processing described in section 2.2.6 (see below).

2.2.6. Sample Processing and statistical analyses

2.2.6.1. Sample processing

Processing by means of the software mMass 5.5.0 allowed the investigation of reproducibility and direct comparison between the samples as well as the generation of the barcode gel view for all samples.

For lipids:

The baseline was corrected with a precision and relative offset of 100 and 0, respectively. Smoothing of the spectra was obtained by Savitzky Golay algorithm, with a window size of 0.2 m/z and 3 cycles (b). Lastly, all samples were normalized to allow a comparison (c). Figure 6 summarizes the sample processing procedure.

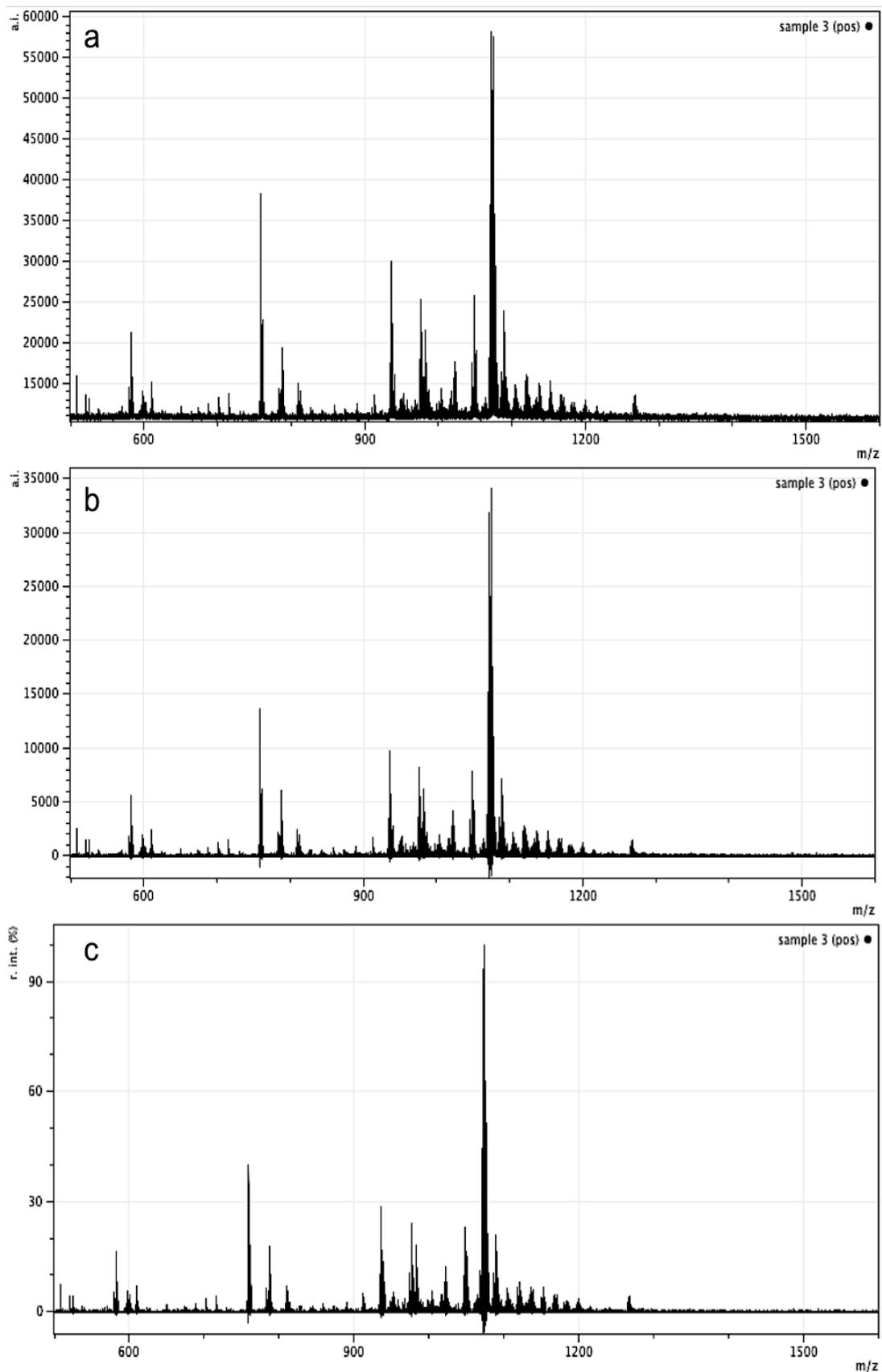


Figure 6: Exemplary data processing with mMass for peptide analysis. (a) raw data of sample 3 in the positive ion reflector mode abated with NIRL, (b) results of baseline and smoothing corrective measures of the sample. (c) resulting spectrum after relative normalization was carried out.

For proteins:

The baseline was corrected with a precision and relative offset of 95 and 70, respectively. Smoothing of the spectra was obtained by Savitzky Golay algorithm, with a window size of 5 m/z and 3 cycles (b). Lastly, all samples were normalized to allow a comparison (c). Figure 7 shows an exemplary data processing of sample 3.

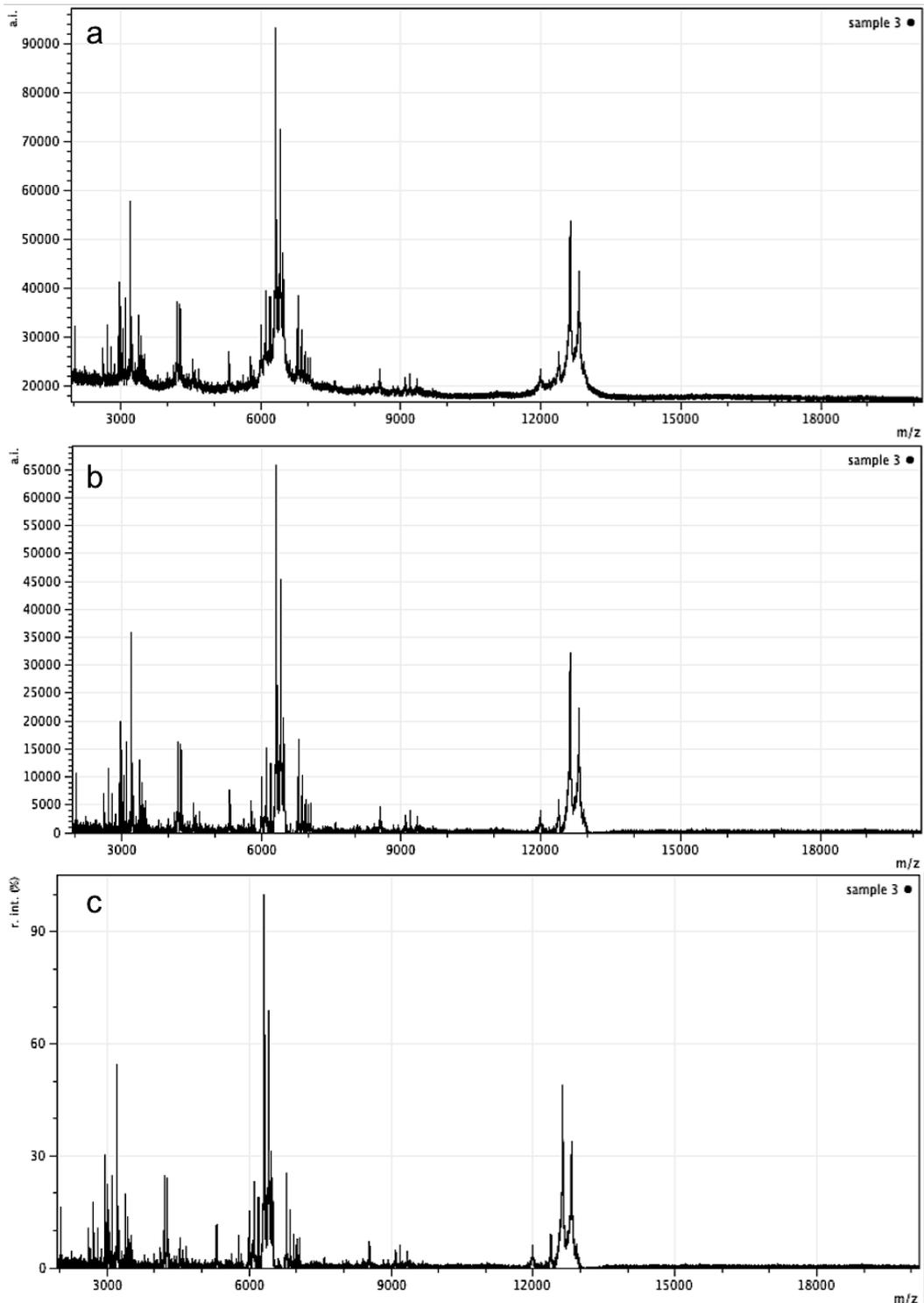


Figure 7: Exemplary data processing with mMass for protein analysis. (a) raw data of sample 3 obtained with NIRL, (b) results of baseline and smoothing correction measures on the sample. (c) resulting spectrum after relative normalization was carried out.

2.2.6.2. Statistical Analyses

Moreover, statistical analysis was carried out using Mass Up in order to create a Principle Component analysis (PCA). The raw data was pre processed using the Savitzky Golay algorithm and a TopHat baseline correction. The parameter “peak detection” was adjusted with the MALDIQuant function. The minimum peak intensity was adjusted as shown in table 14. For intra and inter peak matching the integrated MALDIQuant function was used with a tolerance of 0.002.

Table 14: Parameters for statistical analysis of the MS data. The peak detection parameters: peak scale range and amplitude threshold were the software’s default values.

Parameter	value
Peak detection (MALDIQuant function)	
SNR	3
Half/Window size	60
Minimum peak intensity	1,000 (lipids) 2,000 (proteins)
Peak matching: (MALDIQuant function)	
Tolerance	0.002

Figure 8 shows an exemplary raw data pre processing with (a) being the raw data and (b) the resulting spectrum after the parameters in table 14 have been applied. The processed spectra were then used for the PCA analysis. The PCA analysis was performed with default settings in MassUp, the Eigenvalues and single components were then exported to Excel and further imported into Prism for a 2D representation.

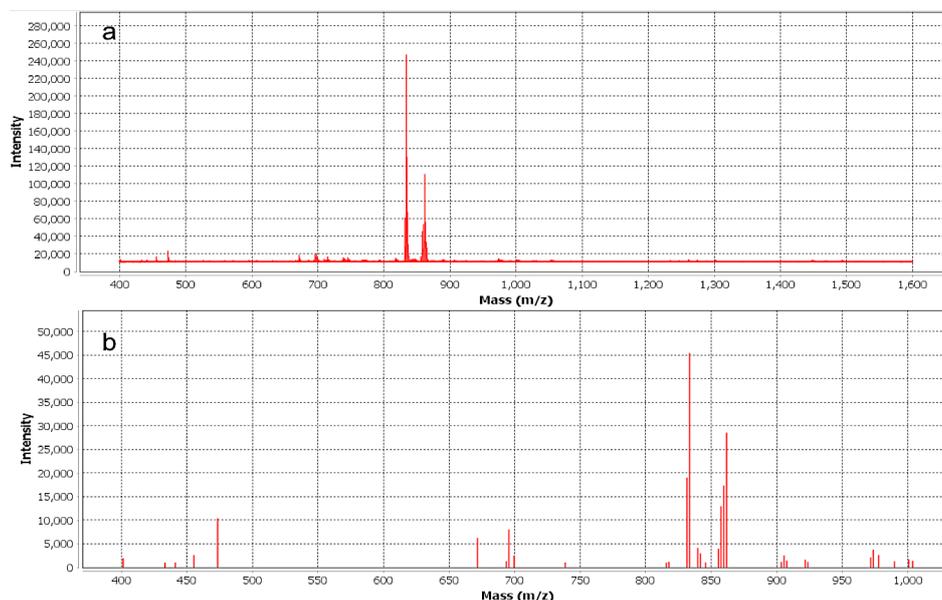


Figure 8: Results of pre-processing with Mass-Up. (a) exemplary raw data of a sample, (b) processed data after the adjustment of parameters.

3. Results

For the development of a new laser based method for biomarker identification, both conventional and NIRL sample preparations of walnuts were carried out. The conventional sample procedures entailed the sample preparation based on the FA/ACN protocol for proteins and the MTBE protocol for lipids.

The NIRL sampling method was performed for both lipids and proteins. Lipid samples were directly applied on the MALDI target, protein samples on the other hand were further purified through a C8 reversed phase chromatography procedure before being analyzed. A schematic overview is given in figure 9.

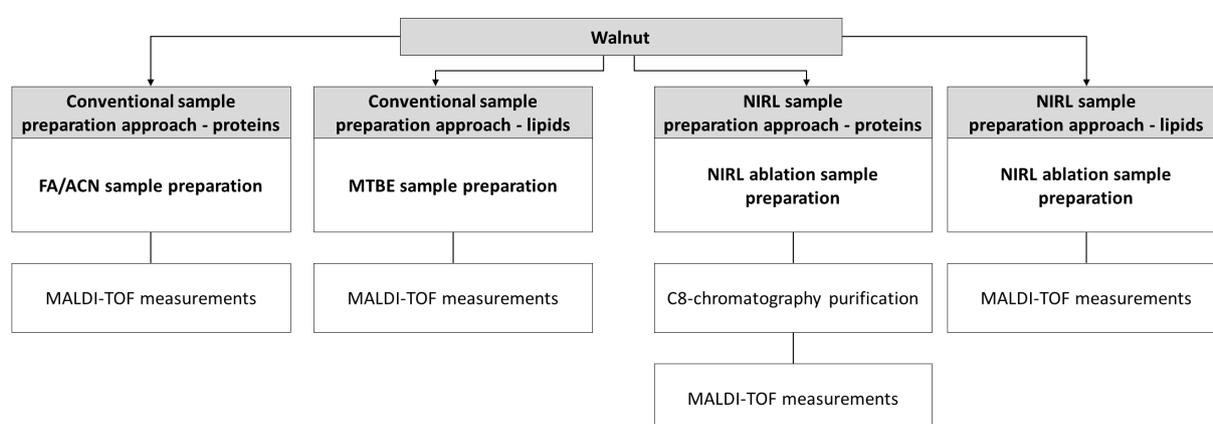


Figure 9: Schematic overview of the sample preparation used for proteins and lipids in walnut.

The new laser based method NIRL coupled to MALDI TOF mass spectrometry should serve as a fast sample profiling and first biomarker identification strategy prior to a comprehensive biomarker characterization. (The utilization of NIRL would allow a significant simplification in comparison to the conventional methods which can find application in routine laboratories.)

3.1. Analysis of proteins using the conventional method FA/ACN

One of the main aims of the thesis was to develop conventional methods as adequate comparison means for the NIRL strategy. Profiling of proteins was hereby achieved with the help of the FA/ACN method. For this work, the method developed in the working group of Prof. Dr. rer. nat Schlüter was used. The FA/ACN method was performed to determine the method's applicability on the walnut samples. The method bases on the extraction of intact proteins to allow a reliable and fast detection and identification of species.

After protein extraction, each sample was spotted with HCCA matrix (1:1) in a thin layer procedure and analyzed via MALDI TOF.

3.1.1. Intra and inter assay for the conventional protein extraction method FA/ACN

In a first step, the reproducibility of the method was tested. Both intra and inter assays were carried out on three consecutive days on sample 1 (here used as a reference). In figure 10, an exemplary result for the intra assay as well as the inter assay is shown.

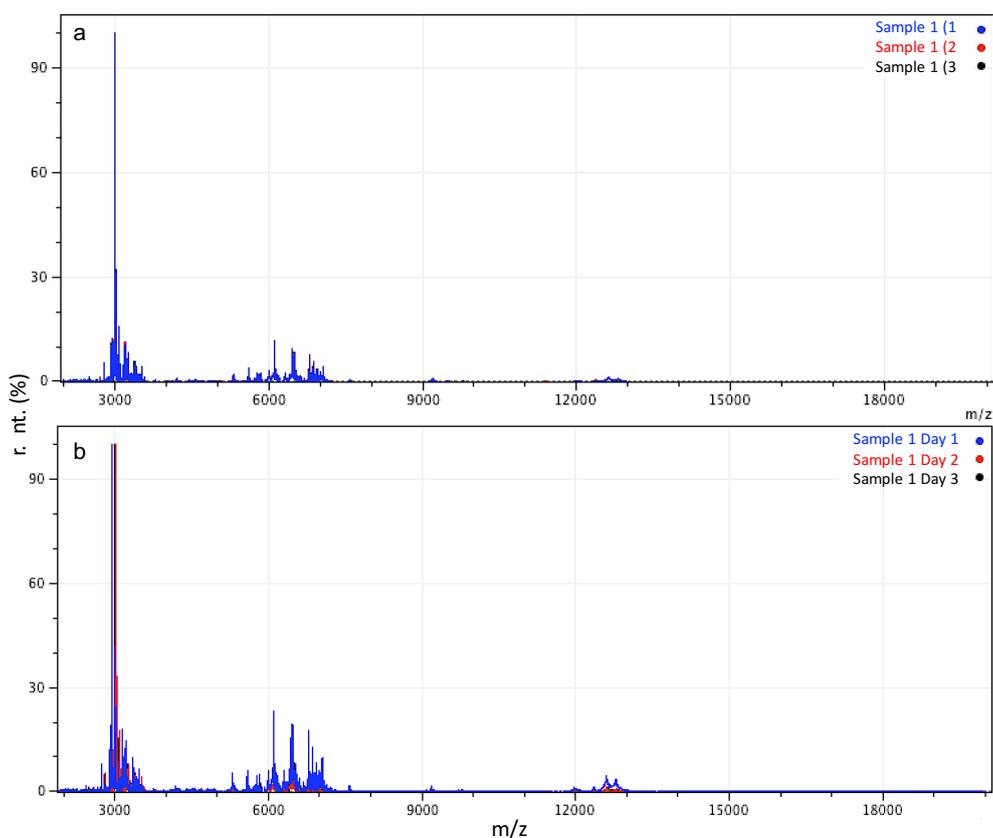


Figure 10: Intra- and inter-assay spectra for the conventional protein extraction method. (a) shows sample 1 as triplicate on day 1, (b) shows sample 1 on Day 1 to 3. The abscissa shows the m/z values in a range from 2,000 to 20,000. The ordinate represents the normalized relative intensity in %. All measurements were carried out with a laser intensity of 45% (see section 2.2.4 for further details). MALDI-TOF spectrum was processed and analyzed with mMass software as described in section 2.2.6.1.

Figures 10 (a) and (b) emphasize the reproducibility of the method. Since in figure 10 (a) the single spectra are nearly identical, only the first sample can be seen on the spectrum. In figure 10 (b) slight divergences in intensities of the overlaid spectra are visible. These are to be expected due to equipment related variances.

Overall, both reproducibility studies can be regarded as successful. The conventional method was therefore used for further analyses of all authentic samples.

3.1.2. Identified signal areas

After pre processing in mMass, analysis of all spectra revealed the presence of six distinct signal ranges within the samples. Hereby, three can be found in all spectra, while the other three may appear more prominently in only specific samples. Figure 11 shows an exemplary spectrum, representing the three main signal ranges.

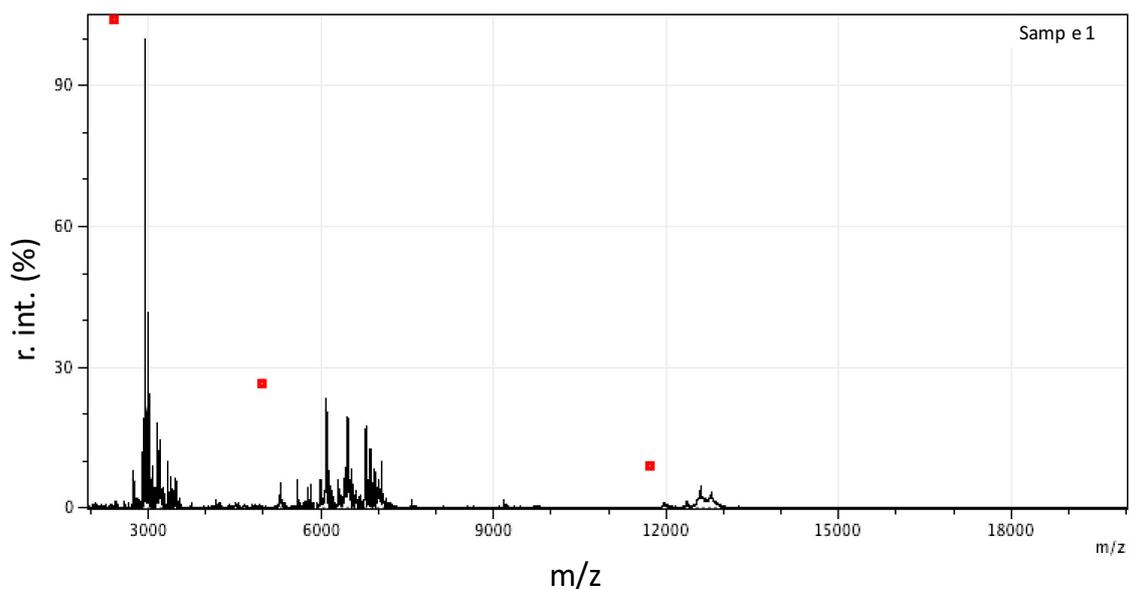


Figure 11: Exemplary representation of the MALDI-TOF spectrum obtained for proteins, using the conventional FA/ACN method. Typical signal areas are marked in red boxes with three main ranges from 3,000 m/z to 3,500 m/z, 6,000 m/z to 7,000 m/z and 12,000 m/z to 13,000 m/z. The abscissa shows the m/z values in a range from 2,000 to 20,000. The ordinate represents the normalized relative intensity in %. All measurements were carried out with a laser intensity of 45% (see section 2.2.4 for further details). MALDI-TOF spectrum was processed and analyzed with mMass software as described in section 2.2.6.1.

The main signals can be found between 3,000 m/z and 3,500 m/z, 6,000 m/z to 7,000 m/z and around 12,000 m/z. All samples showed similar behavior in regards to peak intensities and shapes (not shown) of the main signals.

For a more adequate representation of possible signals in the other ranges, spectra from different samples (sample 7 and 18) have been chosen. The signals found in between the previously mentioned ranges are sample specific.

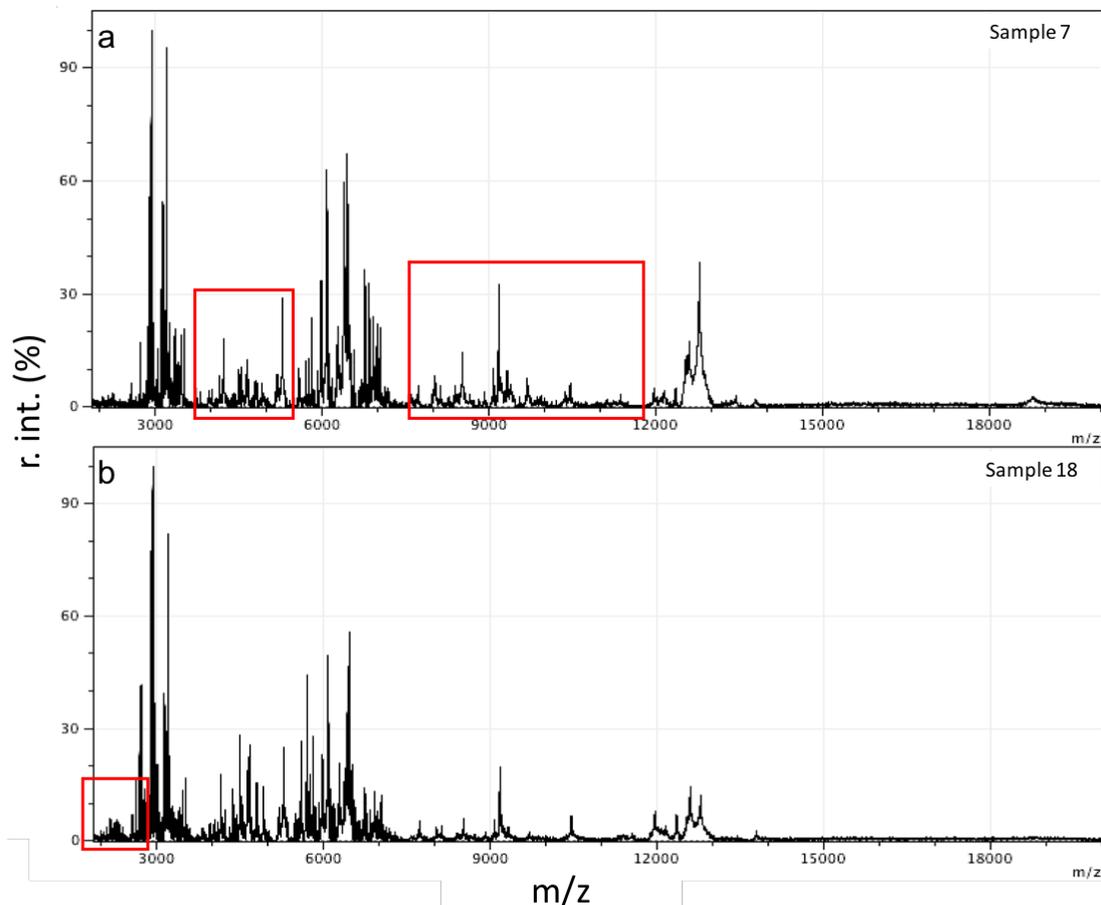


Figure 12: Representative MALDI-TOF spectrum for proteins obtained using the conventional protein extraction method FA/ACN. The additional peak signal areas found in different samples are highlighted in red boxes. Typically, additional signals are found from the ranges 2,000 m/z to 3,000 m/z (1), 4,000 m/z to 5,000 m/z (2) and 8,000 m/z to 11,000 m/z. The abscissa shows the m/z values in a range from 2,000 to 20,000. The ordinate represents the normalized relative intensity in %. All measurements were carried out with a laser intensity of 45% (see section 2.2.4 for further details). MALDI-TOF spectrum was processed and analyzed with mMass software as described in section 2.2.6.1.

Figure 12 (a) displays sample 7. Prominent additional signals are visible within the range 4,000 m/z to 5,000 m/z and 7,000 m/z to 11,000 m/z. Moreover, slight differences in the three main signal ranges in comparison to figure 11 are visible.

In subfigure (b), sample 18 also contains signals in the mentioned ranges, however additionally shows various peaks around 2,000 m/z. The samples in figure 11 and 12 demonstrate that regardless of similarities in the protein profiles, each sample has its own unique protein profile, hinting at differences on phenotype level.

3.1.3. Averaged spectra and barcode gel view for the conventional method FA/ACN

To analyze whether the samples can be discriminated based on cultivar, all samples from one species (here Chandler n= 7, Franquette n= 8, Lara n = 13 and NB n = 8) were averaged in one spectrum.

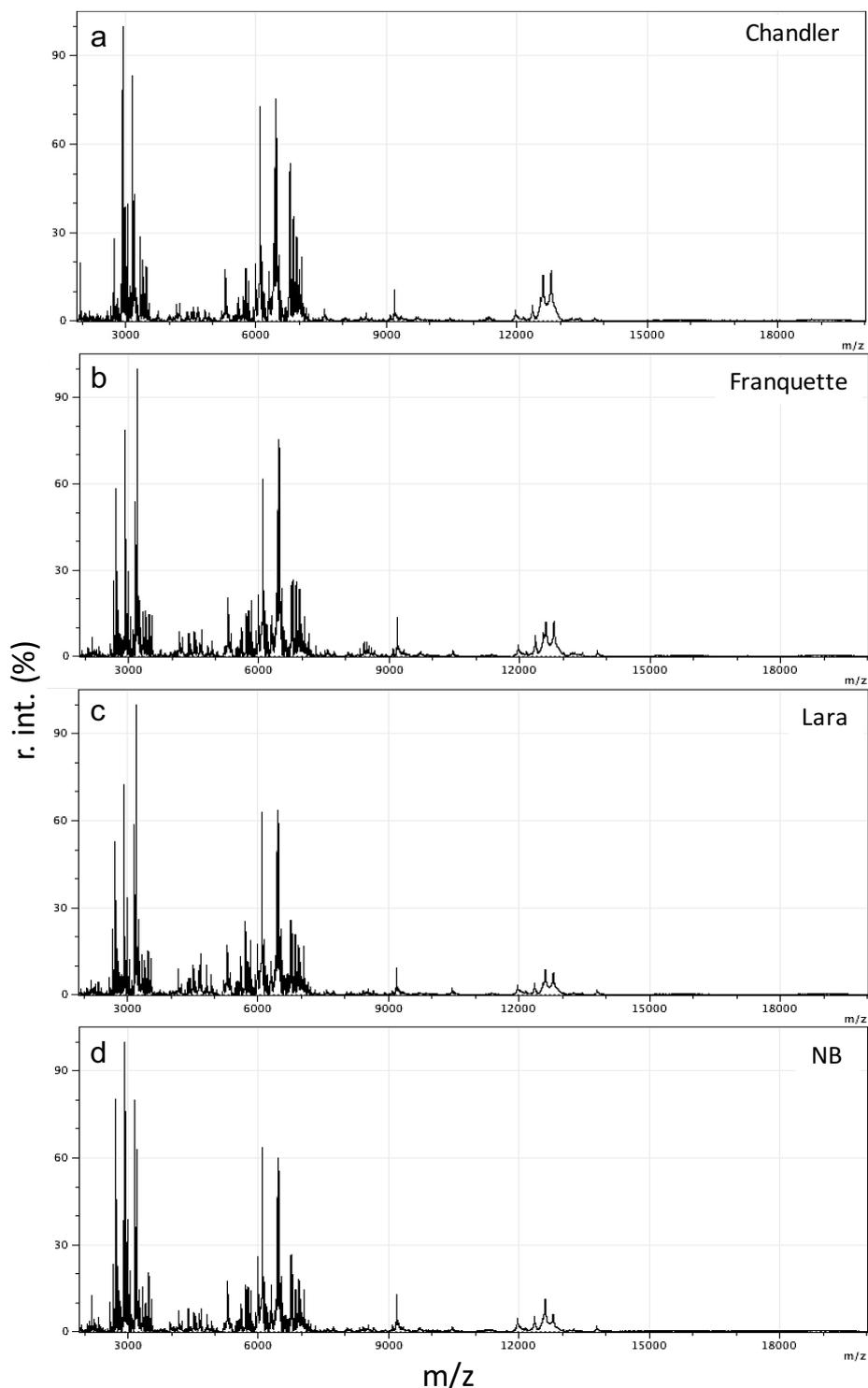


Figure 13: Averaged MALDI-TOF spectra for a four walnut cultivars prepared with the conventional FA/ACN method in one spectrum overview (n=3). (a) Chandler (n = 7), (b) Franquette (n = 8), (c) Lara (n = 13) and (d) NB (n = 8). The abscissa shows the m/z values in a range from 2,000 to 20,000. The ordinate represents the relative intensity in %. All measurements were carried out with a laser intensity of 45% (see section 2.2.4 for further details). Processing of the spectra was done with mMass software as described in section 2.2.6.1.

The spectra shown in figure 13 demonstrate differences. The major divergences can be seen in the signals within the ranges 3,000 m/z to 3,500 m/z, 4,000 m/z to 5,000 m/z and in 6,000 m/z to 7,000 m/z. In subfigure (a) the signal at 3,000 m/z shows the highest intensity (100%), followed by the peak at an m/z value of 3,100 (82% intensity). Another characteristic for Chandler is low signal intensities (4 to 5%) in the range between 4,000 m/z and 5,000 m/z as well as a prominent peak at 2,000 m/z with a normalized intensity of 20%.

In comparison, Franquette (b) has several peaks around 2,500 m/z and multiple signals with an intensity of 6% from 4,000 m/z to 5,000 m/z. Furthermore, the highest peak has an intensity of 100% and can be found at 3,100 m/z. The peaks at 2,900 m/z and 3,000 m/z are both smaller.

Lara (c) behaves similarly to Franquette with the main difference to be found in intensity discrepancies and the signals at 6,900 m/z to 7,000 m/z.

Lastly, NB (d) shows several differences in the first and second main signal ranges in comparison to all other cultivars. These divergences are noticeable due to the different signal relation to each other. Here, the highest peak can be found at 3,000 m/z and the relation of the two peaks in the range from 6,000 m/z and 6,500 m/z is inverted.

The main differences between the walnut cultivars are summarized in the tables below. The signals around approx. 12,000 m/z are not included since the signals have an identical ratio for all four cultivars.

Table 15: List of main Chandler signals, their intensity and presence for protein analysis after extraction with FA/ACN method

Signal (in m/z)	Relative intensity (in %)	Presence
2,000	20	Yes
2,900	28	Yes
3,000	100	Yes
3,100	82	Yes
6,050	73	Yes
6,500	75	Yes
6,850	52	Yes
6,900	39	Yes
6,999	20	Yes

Table 16: List of main Franquette signals, the intensity and presence for protein analysis after extraction with FA/ACN method

Signal (in m/z)	Relative intensity (in %)	Presence
2,000	n.a.	No
2,900	59	Yes
3,000	79	Yes
3,100	100	Yes
6,050	29	Yes
6,500	61	Yes
6,850	75	Yes
6,900	27	Yes
6,999	24	Yes

Table 17: List of main Lara signals, the intensity and presence for protein analysis after extraction with FA/ACN method

Signal (in m/z)	Relative intensity (in %)	Presence
2,000	n.a.	No
2,900	51	Yes
3,000	72	Yes
3,100	100	Yes
6,050	62	Yes
6,500	62	Yes
6,850	26	Yes
6,900	19	Yes
6,999	16	Yes

Table 18: List of main NB signals, the intensity and presence for protein analysis after extraction with FA/ACN method

Signal (in m/z)	Relative intensity (in %)	Presence
2,000	n.a.	No
2,900	80	Yes
3,000	100	Yes
3,100	80	Yes
6,050	62	Yes
6,500	60	Yes
6,850	27	Yes
6,900	19	Yes
6,999	16	Yes

Additionally, a barcode gel view was generated to allow quick identification of differences within the species.

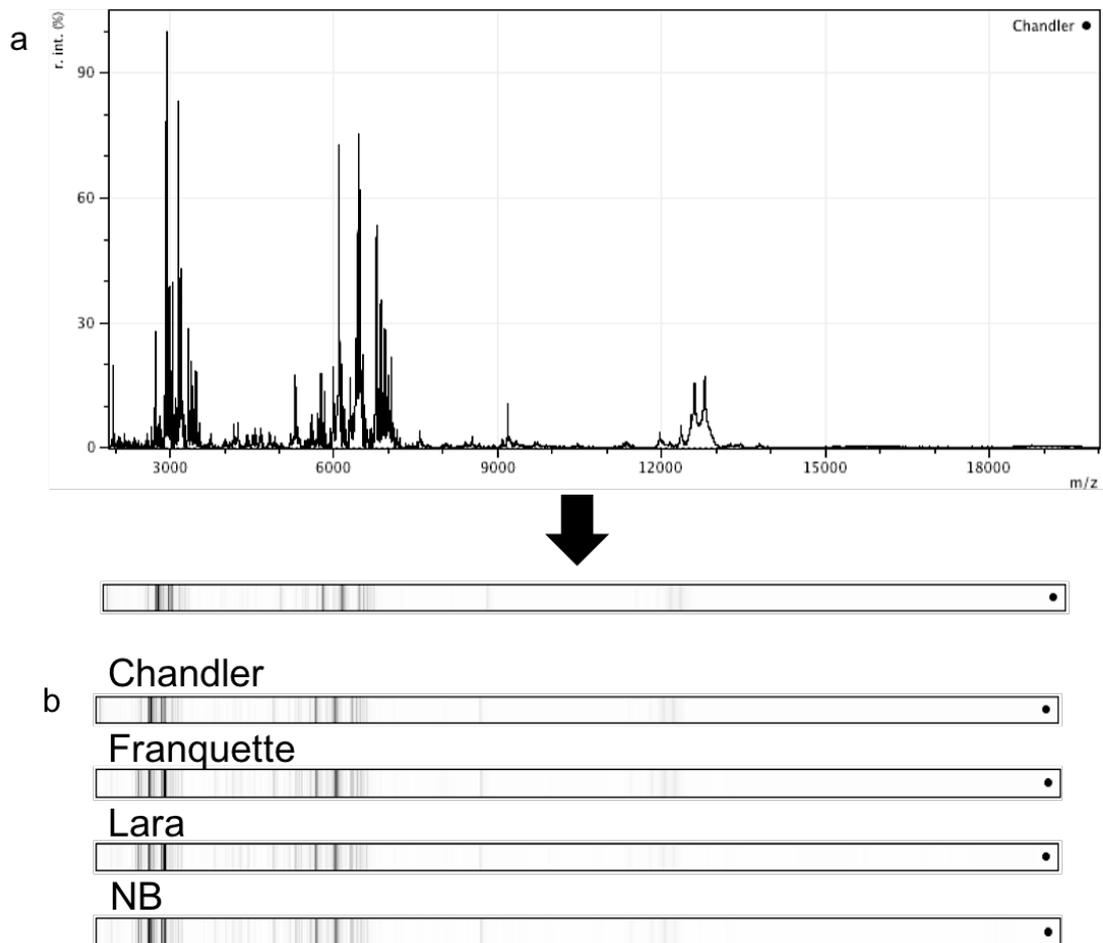


Figure 14: Averaged MALDI-TOF spectra for a four walnut cultivars prepared with the conventional FA/ACN method in barcode representation. To generate the barcode gel views signals with high intensity are translated into thick and dark bands. Low intensity signals appear as lighter and thinner bands. (a) shows a pre-processed exemplary averaged Chandler spectrum and the respective barcode gel view. (b) exhibits all cultivars. Processing of the spectra was done with mMass software as described in section 2.2.6.1.

The barcode overview (see figure 14 (b)), however, displays the differences within the range 3,000 m/z to 3,500 m/z as most prominent. Here, the divergences of the signals with the highest intensity described above are clearly visible. All other ranges show a similar behavior. Moreover, Lara and Franquette show an additional band corresponding to the signal at approx. 2,800 m/z. In both cases, this band is the relatively thin.

In summary, the different cultivars can be differentiated based on their profiles, however the major divergences are related to intensity differences rather than peak presence or absence for clear discrimination between the species.

3.1.4. Principal Component Analysis (PCA) for the FA/ACN method according to cultivar and origin

The aim of the Principle Component Analysis (PCA) was to determine whether the single samples within one cultivar cluster and to analyze how similar the single cultivars are. Figure 15 shows the results of the PCA.

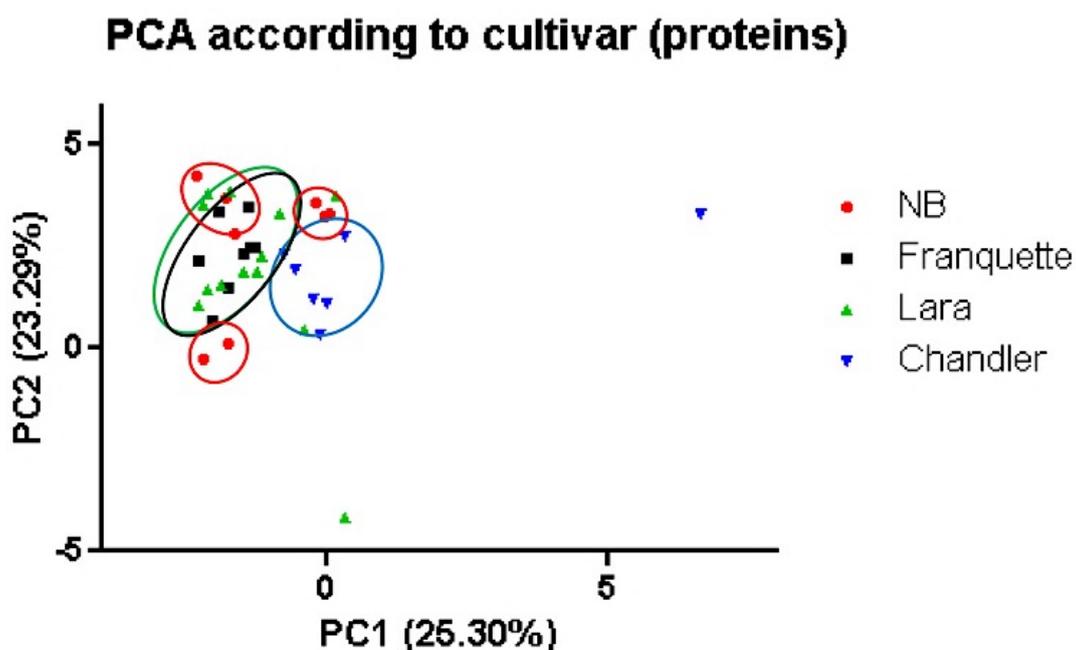


Figure 15: 2D-representation of PCA results according to cultivar for proteins prepared via the FA/ACN method. NB = red, Franquette = black, Lara = green, Chandler = blue. Samples were exported as mzXML files, loaded into MassUp and pre-processed as described in section 2.2.6.2. PC values were exported to Excel and Prism for the generation of the graphical representation.

Figure 15 represents the PCA results according to cultivar for the proteins prepared with the FA/ACN method. The values for PC2 all range between 0 and 4. The PC1 values range from 3 to 0. One Chandler and one Lara samples are outliers (PC1 = 6.5, PC2 = 4 and PC1 = 0.3 and PC2 = 4.5, respectively).

Within the Chandler cultivar all samples (outlier excluded) can be found between PC1 = 1 to 0 and PC2 = 0 to 3. Hereby, the cultivar Chandler is the only clearly distinguishable cluster. The cultivars NB, Franquette and Lara clusters are in the same range, not allowing a clear differentiation of the three different cultivars.

Overall, clusters are visible, however discrimination is not possible for three of the cultivars.

A PCA according to origin was carried out to analyze whether the samples cluster based on their respective origin (figure 16). In this work, the samples originated from Germany (n=4), the USA (n=3), Hungary (n=4), France (n=16), and Italy (n = 9).

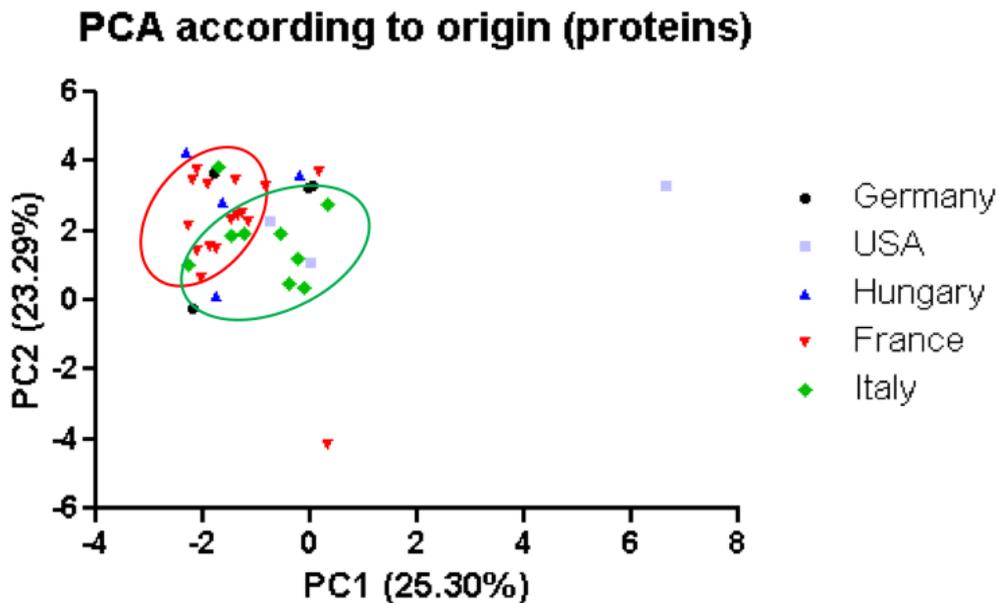


Figure 16: 2D-representation of PCA results according to origin for proteins prepared via the FA/ACN method. Germany = black, USA = light blue, Hungary = blue, France = red, and Italy = green. Samples were exported as mzXML files, loaded into MassUp and pre-processed as described in section 2.2.6.2. PC values were exported to Excel and Prism for the generation of the graphical representation.

Figure 16 shows no clear clustering for the majority of samples. The samples originating from Germany, USA and Hungary are scattered across the graph. Hereby, one USA sample shows significantly different PC values (PC1 = 6.7, PC2 = 3.2) and is therefore regarded as an outlier. As opposed to that, samples originating from France and Italy tend to cluster. In this case, an intersection at the coordinates PC1 = -2.1 and PC2 = 0.2 and PC1 = 0.6 and PC2 = 3 can be seen. Even if the sample from France and Italy slightly cluster, the sample originating from other countries are found within the respective ranges.

To sum up, no clear clustering according to origin of the respective samples is possible.

Overall, analysis of the intact proteins within the range from 2,000 m/z and 20,000 m/z does not allow a differentiation in cultivar nor origin. For the analysis of the cultivar, a slight trend is visible, however not all samples follow the same behavior.

3.2. Analysis of lipids using the conventional method MTBE

The MTBE protocol is a lipid extraction procedure based on a two phase system. Here, the molecules within a sample are separated based on solubility in the organic (MTBE) or hydrophilic phase (H_2O), respectively. Lipids will tend to accumulate in the MTBE phase. All hydrophilic molecules, on the other hand, will be found in the hydrophilic phase.

In this work, all lipid classes were extracted, however only phospholipid were further analyzed. In a first step, crystallization behavior and reproducibility of acquired spectra were tested for two matrices DHB and 9 AA.

3.2.1. Phospholipid analysis with the MALDI matrix DHB

DHB is a benzoic acid derivative which creates clearly defined crystals and is capable to visualize various lipid classes. In this work, it was used to elucidate the phospholipid composition of the four cultivars. For this purpose, the DHB matrix was prepared in different concentrations and with different solvents (see figures 44 48 in annex). Figure 17 shows the crystallization behavior of 0.5 M DHB in 100% MeOH mixed with different solvents.

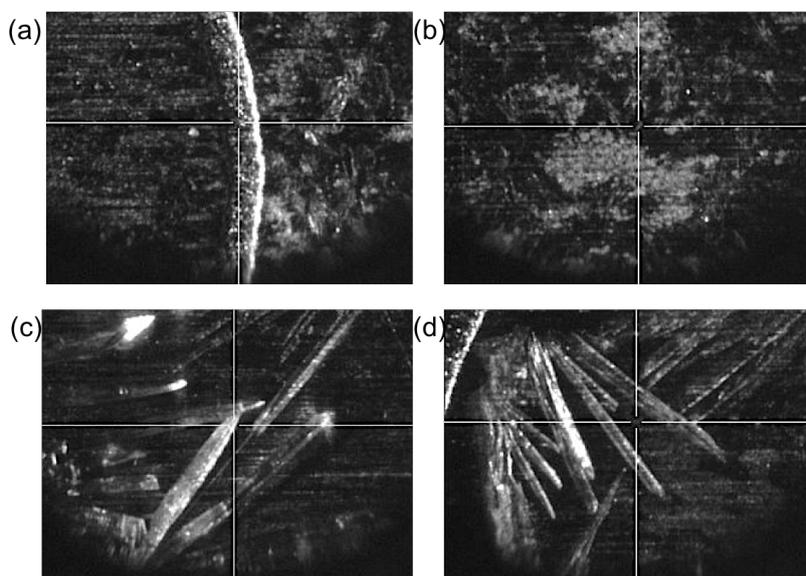


Figure 17: Crystallization behavior of the matrix preparation on 0.5 M DHB in 100% MeOH and with different sample solvents. (a) DHB, (b) DHB with 100% MeOH sample solvent, (c) DHB with MeOH/ H_2O 50:50 (v/v) sample solvent and (d) DHB with ACN/ H_2O 50:50 (v/v) sample solvent. Pictures taken with incorporated camera in MALDI-TOF equipment after spotted and air-dried for 5 min.

As shown in figure 17 (a), only small crystals can be seen. The matrix solution moved outside of the spot due to the high MeOH content. In subfigure (b) a similar behavior is visible. The crystals are small, but form patches across the spot.

In figures 17 (c) and (d) clear crystals can be seen, which developed at the borders of the

spotted matrix and mixed sample solvent. The 0.5 M DHB matrix was further tested for reproducibility with the sample solvents MeOH/H₂O 50:50 (v/v) and ACN/H₂O 50:50 (v/v), since these two preparations yielded the best crystal formation. The intra assay for both preparations is visible in figure 18.

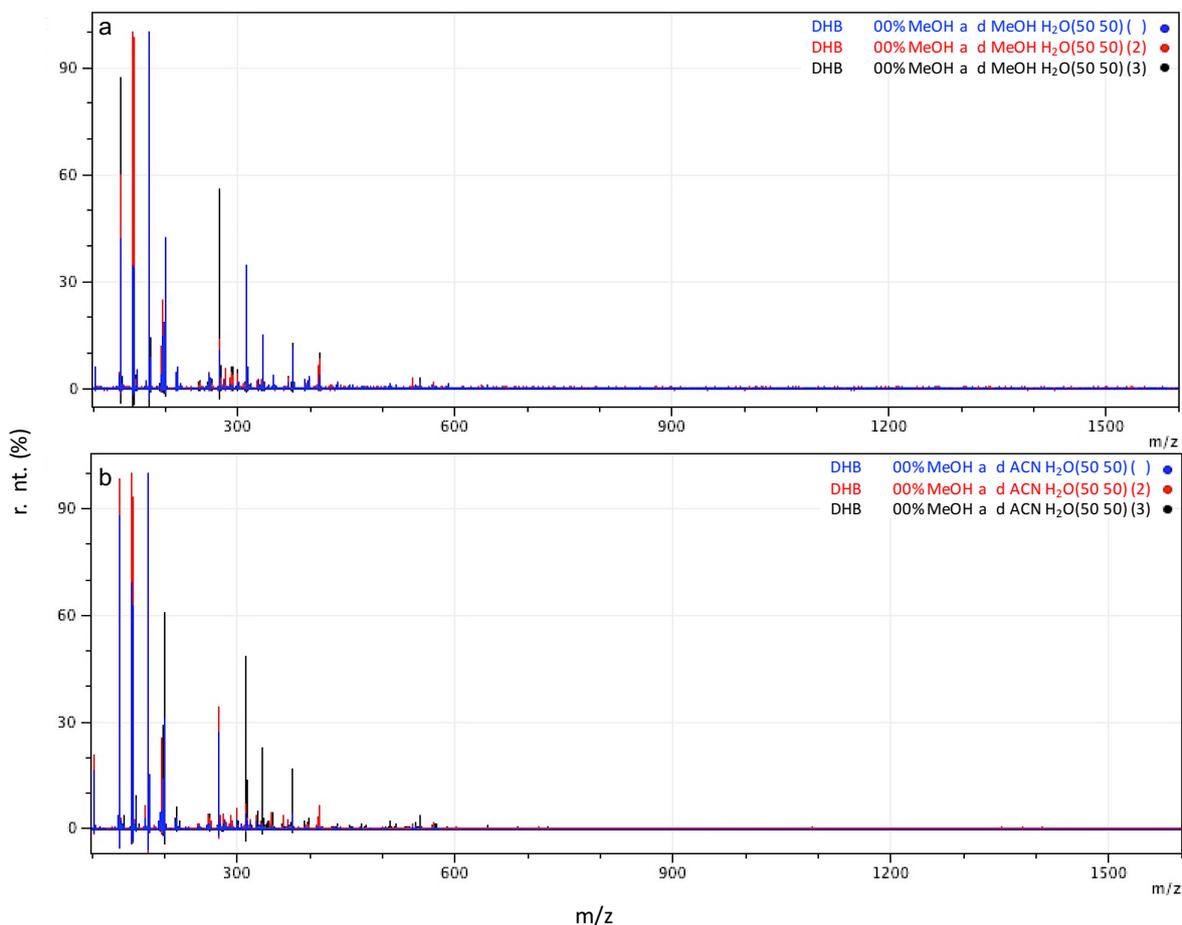


Figure 18: Representative MALDI-TOF spectra for DHB mixed with different sample solvents. (a) analysis of reproducibility of the matrix signal with MeOH/H₂O 50:50 (v/v), (b) analysis of reproducibility of the matrix signal with ACN/H₂O 50:50 (v/v). The samples were spotted on the ground steel target and air-dried for 5 min each. All measurements were carried out with a laser intensity of 35%. The abscissa shows the m/z values in a range from 100 m/z to 1500 m/z. The ordinate represents the normalized intensity of each signal. Processing of the spectra was done with mMass software as described in section 2.2.6.1

Subfigure (a) shows identical matrix signals (molecular mass DHB = 154.12, see reference spectrum in annex figure 49) in the range from 100 m/z to 210 m/z, however five prominent signals appear starting from 250 m/z to approx. 410 m/z. The signals are listed in table 19.

Table 19: List of major signals for the 0.5 M DHB in 100% MeOH mixed with different sample solvents and their m/z value and relative intensity. The relative intensities are valid for subfigure (a).

Signal	m/z value	r. int. (in %)
1	275	57
2	311	34
3	335	14
4	375	13.5
5	411	11.2

Signals with lower intensities are also visible. However, those do not show reproducibility. A similar behavior is visible in subfigure (b). Also here, only the matrix signals visible until approx. 200 m/z are reproducible. Overall, only the matrix signals until approx. 200 m/z can be regarded as reproducible.

Due to lack of crystallization, spectra from all other matrix preparations with DHB could not be recorded. The DHB matrix was therefore not used for further experiments.

3.2.2. Phospholipid analysis with the MALDI matrix 9 AA

Phospholipids were also analyzed using 9 AA as MALDI matrix. 9 AA is a heterocycle which allows the visualization of several lipid classes in both the positive and negative ion polarity. In a first step, the crystallization capability of 9 AA was determined through preliminary experiments similar to the one shown in section 3.2.1.

For this purpose, 9 AA was prepared with 90:10 MeOH/H₂O as well as 60:40 isopropanol/ACN. The matrix was spotted alone as well as mixed in matrix/sample solvent ratio of 1:1 with two possible samples solvents: MeOH/H₂O and ACN/H₂O (50:50 (v/v) each). The results of the preliminary experiment are visible in figure 19.

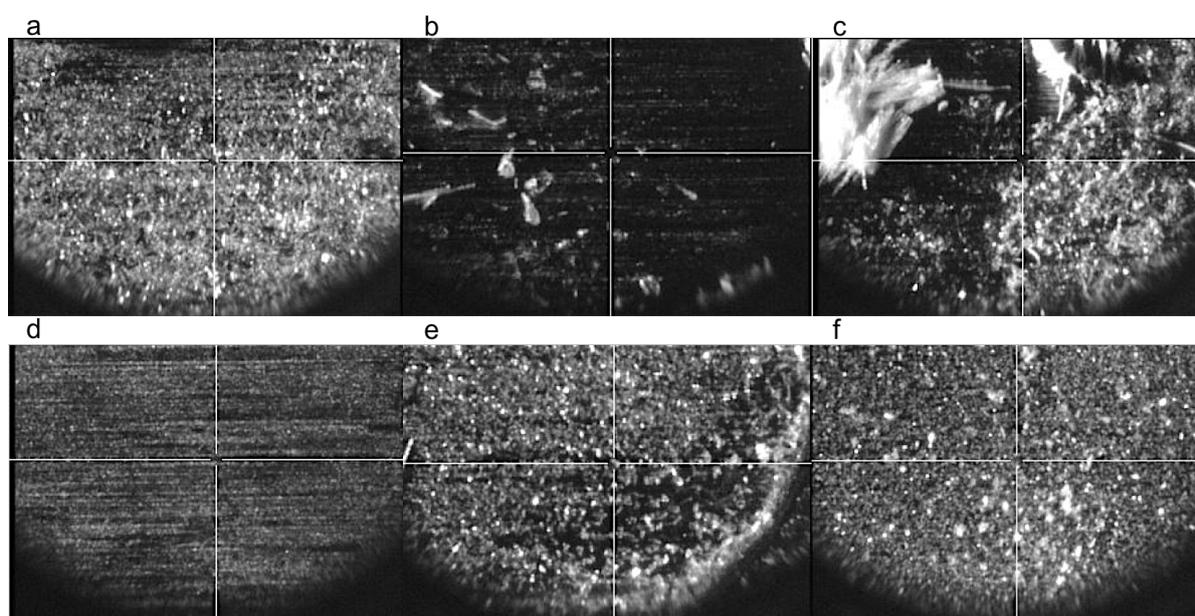


Figure 19: Crystallization behavior of 9-AA in two different solvents and mixed with two different samples solvents, respectively. In (a-c) 9-AA was prepared with 90:10 (v/v) MeOH/H₂O. Hereby, (a) 9-AA matrix in MeOH/H₂O, (b) matrix with sample solvent MeOH/H₂O 50:50 (v/v) and (c) matrix with sample solvent ACN/H₂O 50:50 (v/v). In (d-f) 9-AA was prepared with 60:40 isopropanol/ACN, (d) 9-AA matrix in isopropanol/ACN, (e) matrix with sample solvent MeOH/H₂O 50:50 (v/v) and (f) matrix with sample solvent ACN/H₂O 50:50 (v/v). Pictures taken with incorporated camera in MALDI-TOF equipment.

Subfigures 19 (a c) show the matrix prepared with 90:10 MeOH/H₂O. The crystallization pattern of 9 AA without sample solvent (a) shows relatively big crystals with single brighter spots. As opposed to that, subfigure (b) depicts little to no crystallization. The last subfigure, using the 9 AA matrix solved in MeOH/H₂O mixed with ACN/H₂O (50:50 (v/v) each) (c), shows partial crystallization. Here, the lower half of the spot resembles the crystallization of the 9 AA matrix alone. However, towards the center, aggregation of the matrix/sample solvent material can be seen.

Subfigures (d f) represent the crystallization pattern for the 9 AA matrix solved in 60:40 isopropanol/ACN. Subfigure (d) shows the typical crystallization expected by the 9 AA matrix

when spotted alone. Subfigures (e f) slightly resemble the 9 AA crystallization pattern shown in figure 19 (d). In both cases, the crystals are either only slightly bigger (f) or show uneven crystallization over the spot (e).

Due to even crystallization as well as applicability for spectra recording, the 9 AA solved in 60:40 isopropanol/ACN and the sample solvent ACN/H₂O 50:50 (v/v) were chosen.

Applicability of the 9 AA (60:40 (v/v) isopropanol/ACN) matrix has been tested in the positive and negative mode by analysis and fragmentation of the lipid standard “Light Splash” (Avanti Lipids). The spectra are visible in figure 20. Tables 20 and 21 list the identified and unidentified lipids, the expected m/z values and the exemplary identification of fragments of one of the PI phospholipid.

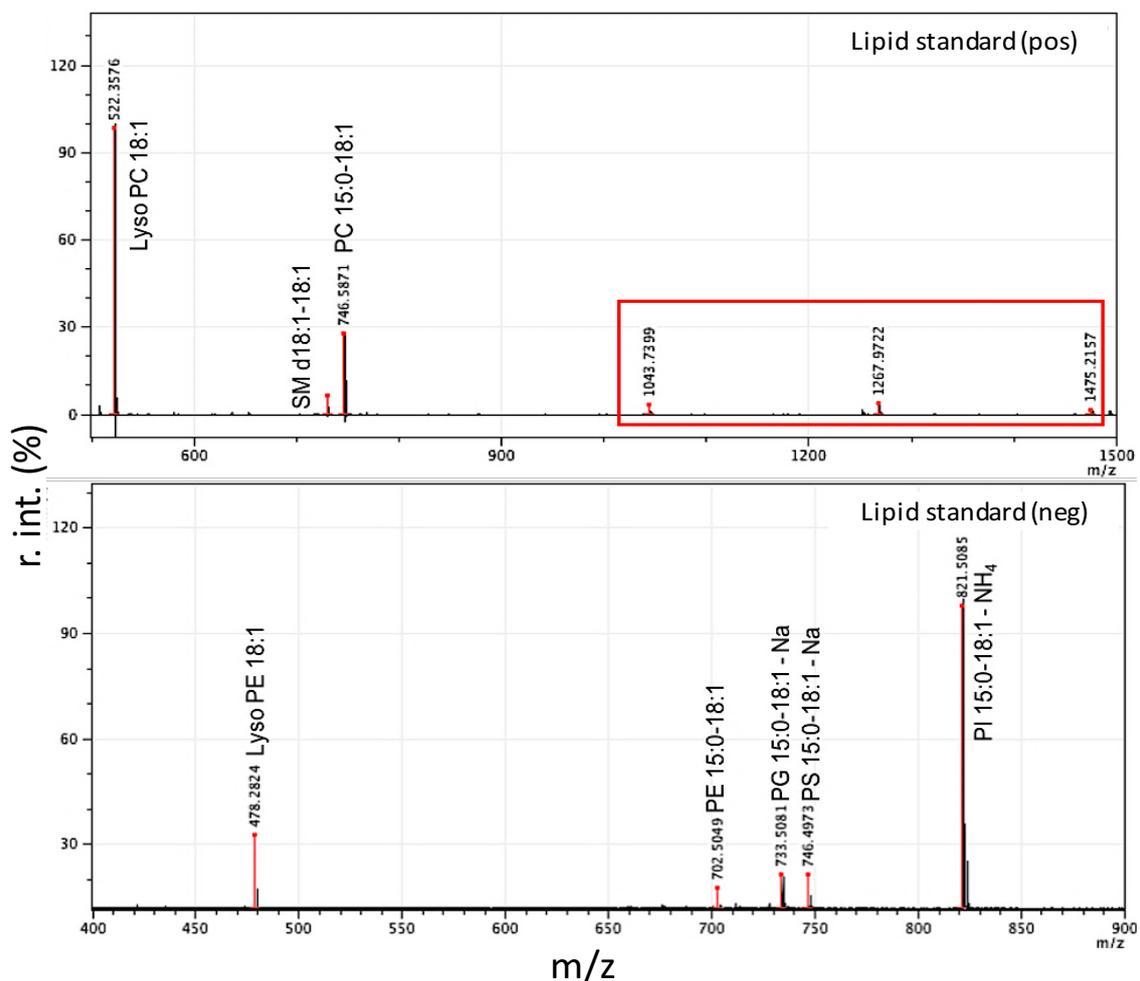


Figure 20: MALDI-TOF spectra of the lipid standard in positive and negative ion mode. Identified lipids were marked in the spectra. The abscissae show the m/z values in a range from 500 m/z to 1,500 m/z (a) and from 400 m/z to 900 m/z, respectively. The ordinate represents the normalized relative intensity (%). The spectra were recorded in positive and negative ion mode with a laser intensity of 35% (positive mode) and 30% (negative mode), respectively. The spectra were zoomed in to allow a better view of the signals. Processing of the spectra was done with mMass software as described in section 2.2.6.1. List of signals can be found in a separate table. Signals in red box were not further analyzed.

Figure 20 (a) shows the spectrum of the lipid standard in the positive ion mode. A total of

three lipids could be identified: PC (15:0 18:0) Lyso PC (18:1) and SM (d18:1 18:1). Two peaks can be seen over 1,000 m/z: 1,043.7399 m/z and 1,267.9722 m/z. These two signals could not be attributed to any lipid present in the standard. They were therefore not further analyzed.

The negative ion polarity allowed the identification of five lipids: PE (15:0 18:1), Lyso PE (18:1), PG (15:0 18:1 Na), PI (15:0 18:1 NH₄) and PS (15:0 18:1 Na). The two signals at approx. 659 m/z and approx. 675 m/z could not be attributed to any molecule in the lipid standard and were therefore not taken into consideration for further analyses.

Table 20: List of lipids present in the standard, the theoretical weight, the measured m/z values and polarity

Lipid	Molecular weight	Measured m/z value	Polarity
15:0 18:1 PC	745.56	746.59	positive
18:1 Lyso PC	521.35	522.36	positive
15:0 18:1 PE	703.52	702.52	negative
18:1 Lyso PE	479.30	478.29	negative
15:0 18:1 PG (Na salt)	756.49	733.52	negative
15:0 18:1 PI (NH ₄ salt)	839.55	821.52	negative
15:0 18:1 PS (Na salt)	769.49	746.51	negative
15:0 18:1 15:0 TG	804.72		
15:0 18:1 DG	580.51		
18:1 MG	356.29		
18:1 Chol Ester	650.60		
d18:1 18:1 SM	728.58	729.61	positive
D18:1 15:0 C15 Ceramide	523.50		

Out of the 13 lipids, 8 could be identified. Among those all phospholipids contained in the standard. The figure below (figure 21) shows the fragmentation spectrum of the phospholipid phosphoinositol (PI). All lipids were analyzed thoroughly in another project and characterized based on their fragmentation.

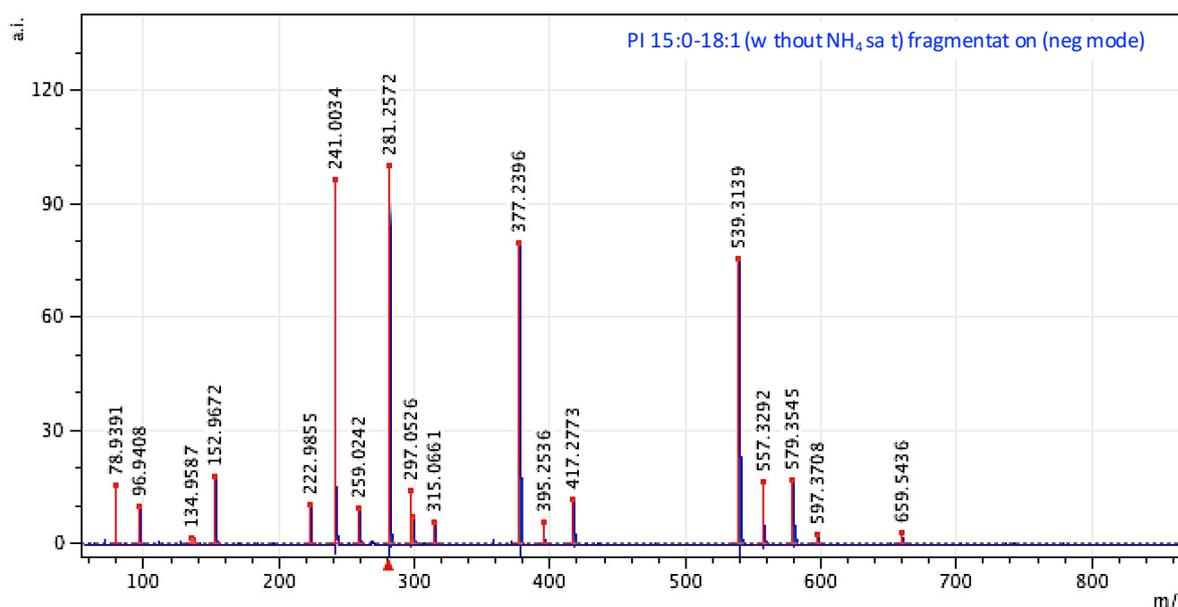


Figure 21: Fragmentation MALDI-TOF spectrum for PI 15:0-18:1 (w/out the NH₄ salt). Peak assignment based on database search LIPID MAPS [www.lipidmaps.org]. The abscissa shows the m/z values in a range from 80 to 850. The ordinate represents the normalized intensity in a.u.. Processing of the spectra was done with mMass software as described in section 2.2.6.1.

Table 21: m/z values and the respective attributed lipid fragment for the phospholipid PI. I = phosphate, FA1 = 15:1 fatty acid, FA2 = 18:1 fatty acid, P_i = phosphate group, (O) = cleavage ether bond

m/z value	Attributed lipid fragment
78.9391	P
96.9408	P + 2H ⁺
134.9587	not assigned
152.9672	FA2 olefin
222.9855	not assigned
241.0034	FA1
259.0242	HG + H ⁺
281.2572	FA2
377.2396	M I FA2
395.2536	M I FS2 (O) 2H ⁺
417.2773	M I FS1
539.3139	M FA2 2H ⁺
557.3292	M FA2 (O)
579.3545	M FA1 H ⁺
597.3708	M FA1 (O)
659.5436	M I + H ⁺

Nearly all fragments were successfully attributed to an m/z value. Hereby, fragments P, P + 2H⁺ and FA2 olefin (see figure 21) represent the basic fingerprint spectrum for the PI lipid. All other fragments are dependent on the fatty acids present in the lipid. The exemplary fragment spectrum as well as figures 50-55 in the annex were used for characterization of lipids in the reference sample.

After the analysis of the matrix applicability, lipids were extracted from a reference sample, using the MTBE protocol. In this step, the adequacy of the conventional method was tested. For an appropriate analysis, a direct comparison with the 9 AA matrix measured in the same range and under the same parameters was evaluated. The spectra are visible in figure 22.

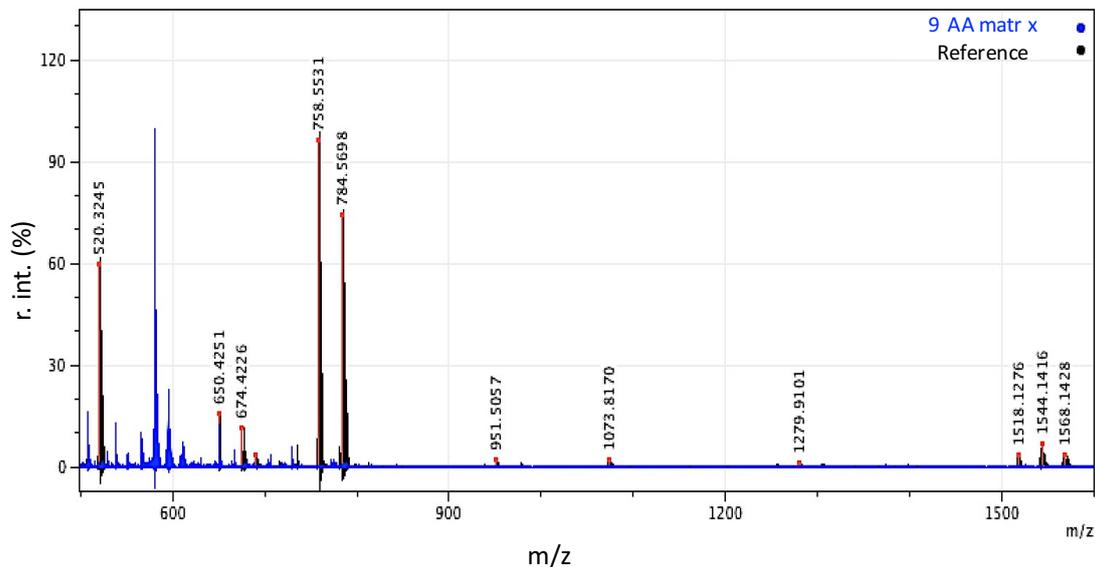


Figure 22: Comparison between the reference and the 9-AA matrix (negative sample) spectra. All signals marked with m/z values are not present and not part of the background noise caused by 9-AA. The abscissa shows the m/z values in a range from 500 to 1,600. The ordinate represents the normalized relative intensity in %. The spectra were recorded with a laser energy of 35%. The reference sample was spotted with a matrix/analyte ratio of 1:1. Both spectra were measured in the positive ion mode. Processing was done with mMass software as described in section 2.2.6.1.

Figure 22 shows the presence of specific sample related peaks. Those peaks range from approx. 666 m/z to approx. 1,550 m/z.

Characterization of reference specific signal was done with fragmentation. Signals that did not show any of the lipid specific fragments determined in the previous experiments were not further analyzed. (see figures 56-63 for reference signal fragmentation and figures 50-55 for lipid standard fragmentation and lipid fingerprint).

Additionally, figure 22 displays a strong background noise from approx. 500 m/z to 600 m/z. Only the signal at 520.32 m/z was not attributed to a matrix signal. However, signals found in this area were not further analyzed.

Next, matrix/analyte ratio and sample dilution studies were carried out. The aim of these studies was the determination of the ideal ratio and sample dilution. The experimental set was performed using the commercially available walnut from the brand Rapunzel as reference sample. The matrix/analyte ratios 1:1, 1:2 and 1:3 (see annex figure 64) as well as the original sample and a 1:10 dilution were carried out. All conditions were measured in the positive ion reflector mode. The spectra are shown in figure 23:

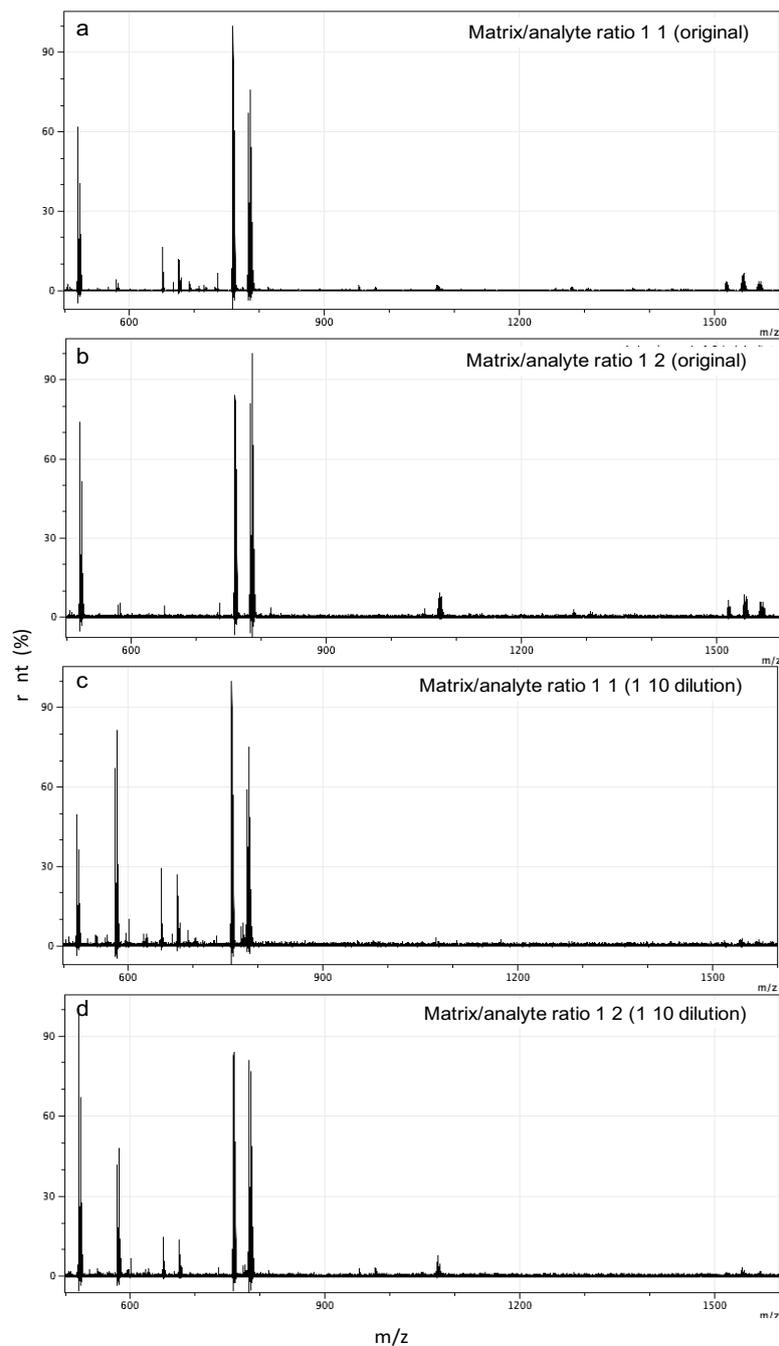


Figure 23: Spectra of the analysis of matrix/analyte ratios and dilutions on MALDI-TOF spectra measured in the positive ion reflector mode. (a) matrix/analyte ratio 1:1 non-diluted, (b) matrix/analyte ratio 1:2 non-diluted, (c) 1:10 sample dilution and 1:1 matrix/analyte ratio and (d) 1:10 dilution of the sample and 1:2 matrix/analyte ratio. All measurements were carried out with a laser intensity of 35%. The abscissa shows the m/z values in a range from 500 to 1,600. The ordinate represents the normalized relative intensity in %. Processing of the spectra was done with mMass software as described in section 2.2.6.1

Subfigure (a) shows the spectrum for the 1:1 matrix/analyte ratio and the analysis of the original sample (sample after MTBE sample preparation resuspended in 50 μ L ACN/H₂O).

The range from 500 m/z to approx. 600 m/z shows several matrix signals.

The reference specific signals are once again visible. Additionally, a signal with lower intensity can further be seen around 1,000 m/z.

Subfigure (b) represents the original sample measured in a matrix/analyte ratio of 1:2. The results are nearly identical to subfigure (a). Subfigures (c) and (d), on the other hand, represent the sample in a 1:10 dilution measured with a matrix/analyte ratio of 1:1 and 1:2, respectively. In both subfigures only the main signals at approx. 758 m/z and approx. 780 m/z are visible. The signal triplets at approx. 1,500 m/z to approx. 1,550 m/z were not detected. Subfigure (d), however, has signals at approx. 990 m/z 1,000 m/z.

Overall, the original sample lead to better spectra. Therefore, further experiments were carried out in a 1:1 matrix/analyte ratio with the original sample to minimize the sample volume as much as possible.

Before analysis of the authentic samples, the reference sample was analyzed as in intra (n=3) and interassay (n=3). For this purpose, the MTBE protocol was performed thrice on the reference sample. Figure 24 shows the respective spectra.

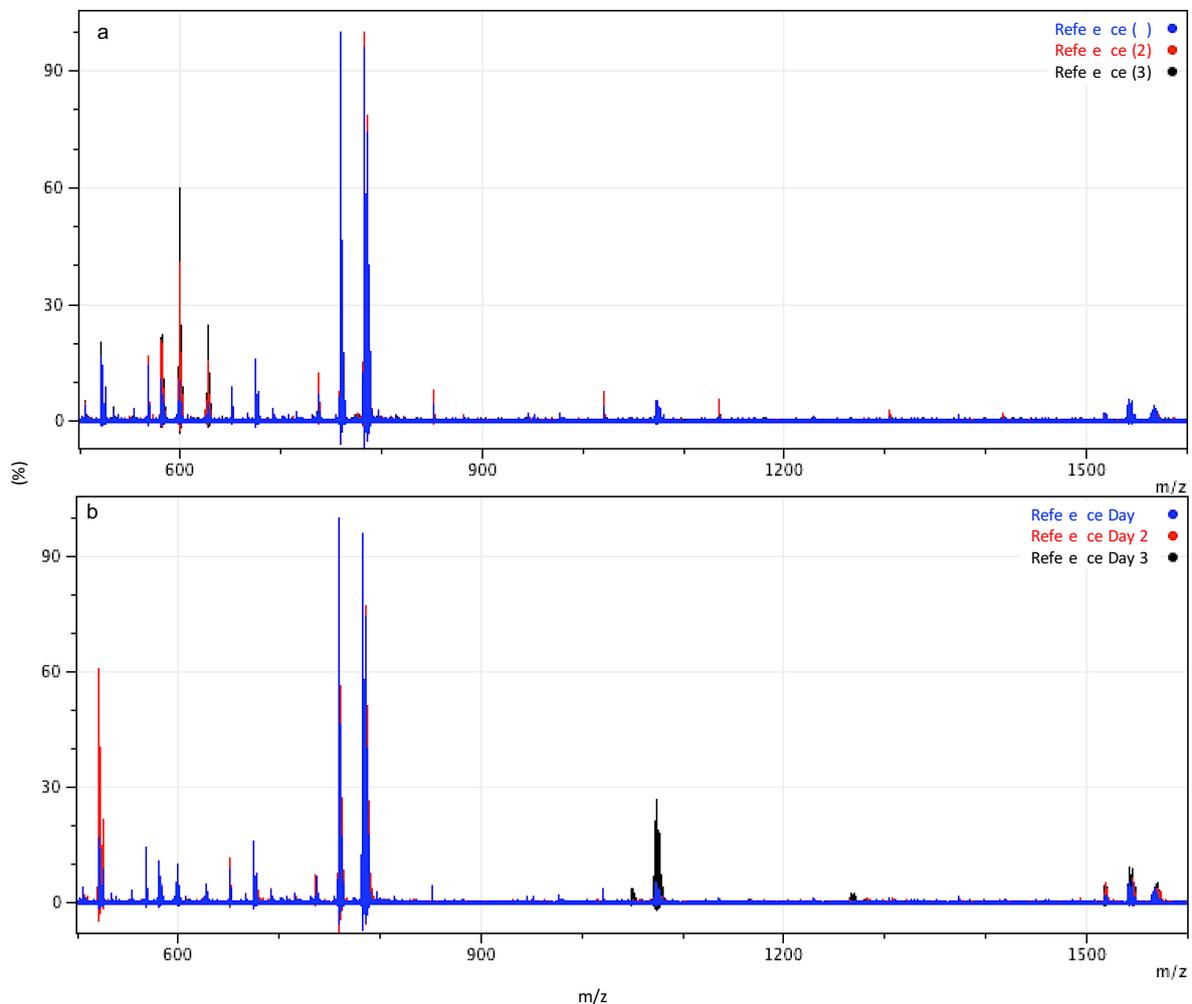


Figure 24: Intra- and inter-assay spectra for the reference sample. (a) shows the spectrum for the intra-assay results, (b) shows the inter-assay results. All samples were prepared with the MTBE protocol (sample dissolved in 50:50 (v/v) ACN/H₂O) and analyzed in a 1:1 matrix/analyte ratio in the positive ion reflector mode at a laser intensity of 35%. The abscissa shows the m/z values in a range from 500 to 1,600. The ordinate represents the normalized relative intensity in %. Processing of the spectra was done with mMass software as described in section 2.2.6.1

Subfigure (a) displays the intra assay measured in the positive ion mode. Except for reference 2 (red), all spectra (black and blue) show the same signals already detected in prior experiments. In the spectrums of reference 2 two additional signals: 1,000 m/z and approx. 1,100 m/z can be seen. Subfigure (b) shows similar results, with only the measurement on the third day exhibiting a higher intensity. The MTBE protocol was therefore further used for the analysis of the phospholipids in the authentic samples.

3.2.3. Averaged spectra and barcode gel view in the positive and negative mode for the analysis of phospholipids extracted with the MTBE based sampling method

Due to lack of surface tension of the authentic samples after MTBE extraction, the samples could not be spotted as determined in the method development. Since the problem could not be identified, the results in the following sections are all based on the MTBE extraction followed with a matrix/analyte ratio of 1:3. Figure 25 represents the averaged spectra of all four cultivars measured in the positive ion reflector mode.

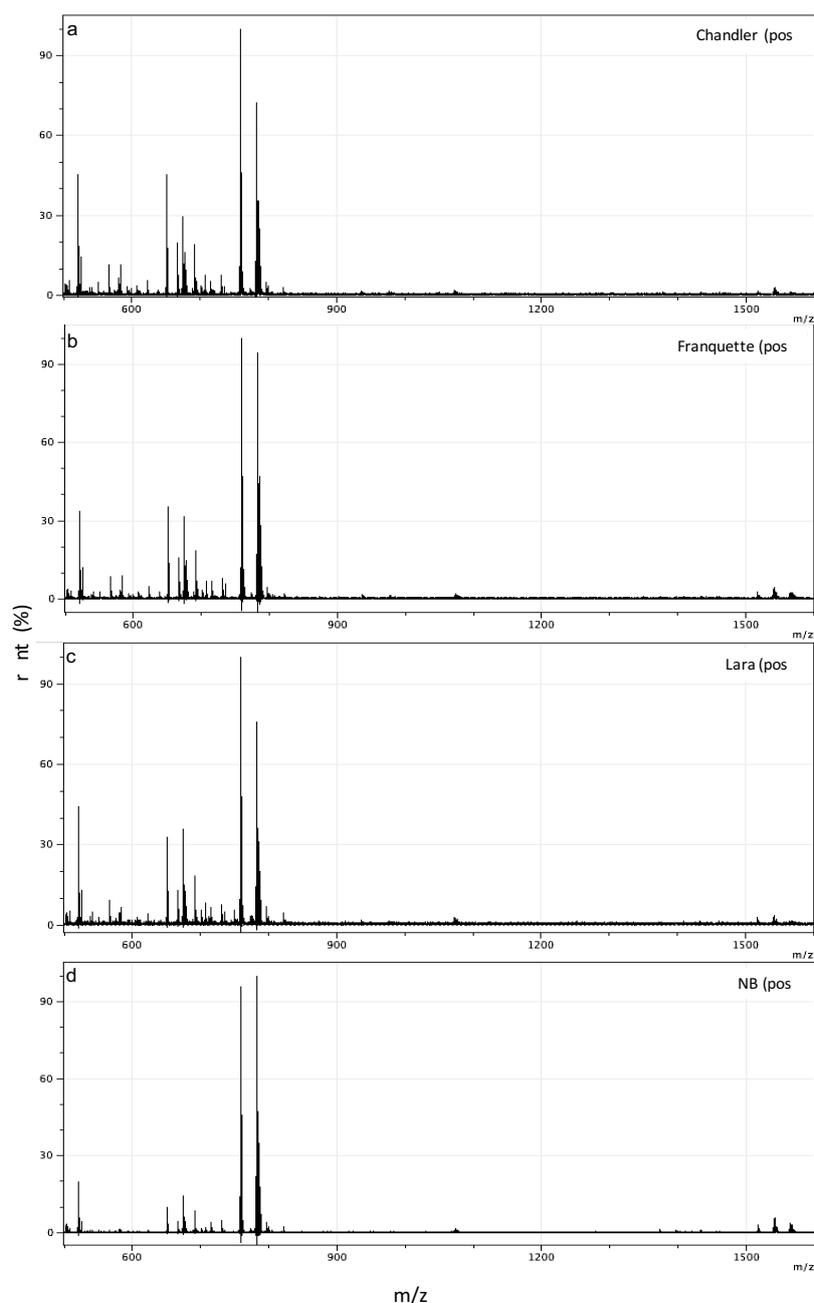


Figure 25: Averaged MALDI-TOF spectra for a four walnut cultivars for the conventional MTBE method measured in positive ion reflector mode in spectrum overview (n=3). The figure shows (a) Chandler (n=7), (b) Franquette (n=8), (c) Lara (n=13) and (d) NB (n=8). The abscissa represents the m/z values in a range from 500 to 1,600. The ordinate represents the normalized relative intensity in %. The spectra were recorded with a laser energy of 35%. Processing of the spectra was done with mMass software as described in section 2.2.6.1

In subfigure 25 (a) only the two prominent signals around approx. 758 m/z and 780 m/z could be detected. The triplet signals detected in the at 1,500 m/z are not visible, however the signal at 1,540 m/z (2% rel. int.) is hinted at. Also signals with smaller intensities can be seen between 900 m/z and 1,100 m/z (between 2% and 3%). However, similarly to the triplet signals, these peaks are only hinted at.

In comparison, subfigure 25 (b) shows clearer signal triplets at 1,500 m/z, but no peaks between 900 m/z and 1,100 m/z.

In subfigure 25 (c) the two major signals at approx. 758 m/z and 780 m/z and one signal at 1,080 m/z were detected. The signal at approx. 1,540 m/z is hinted at (rel. int. 2%).

The averaged spectrum of NB (subfigure 25 (d)) is the only one which clearly shows the triplet signals. However, also in this case the intensities are fairly low (approx. 5-7%).

All spectra show background noise. Subfigure 25 (d) is the one with less disturbances.

The averaged spectra were also represented in barcode overview (figure 26).

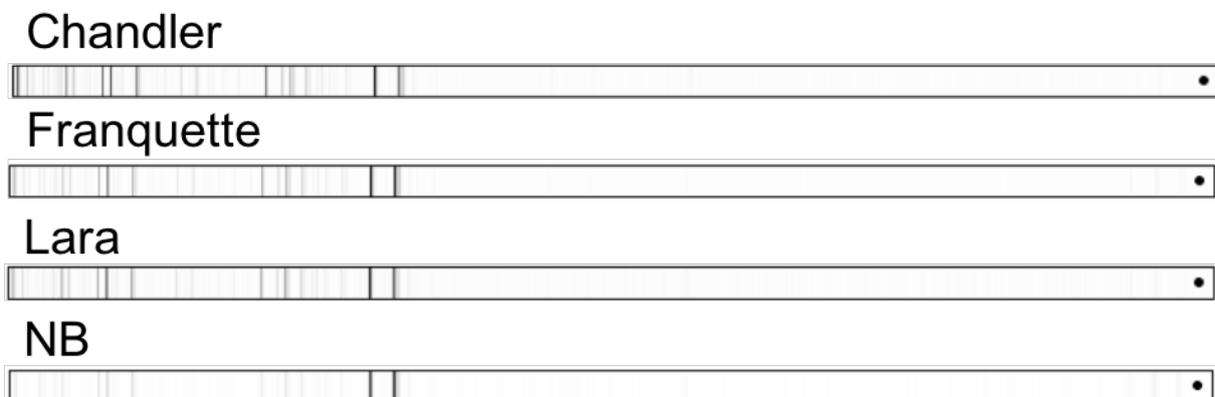


Figure 26: Averaged MALDI-TOF spectra for a four walnut cultivars for the conventional MTBE method in barcode overview measured with the positive polarity. The spectra used for the automatic generation of the barcodes were recorded with a laser energy of 35% (see section 2.2.4 for further MALDI-TOF acquisition details). Processing of the spectra was done with mMass software as described in section 2.2.6.1.

Here, the differences described previously are not visible. The signals at approx. 758 m/z and approx. 780 m/z can be clearly detected. However, signals with lower relative intensity are less visible. Only the cultivars NB and Franquette show light and thin bands in the higher m/z range. Bands corresponding range between 600 m/z and 700 m/z are visible nearly in all barcode representations except for the cultivar NB.

Phospholipid analysis was also carried in the negative ion mode. Figure 27 represents the averaged spectra for all four cultivars. The spectra for negative ion mode were also recorded in a 1:3 ratio for the same reason described above.

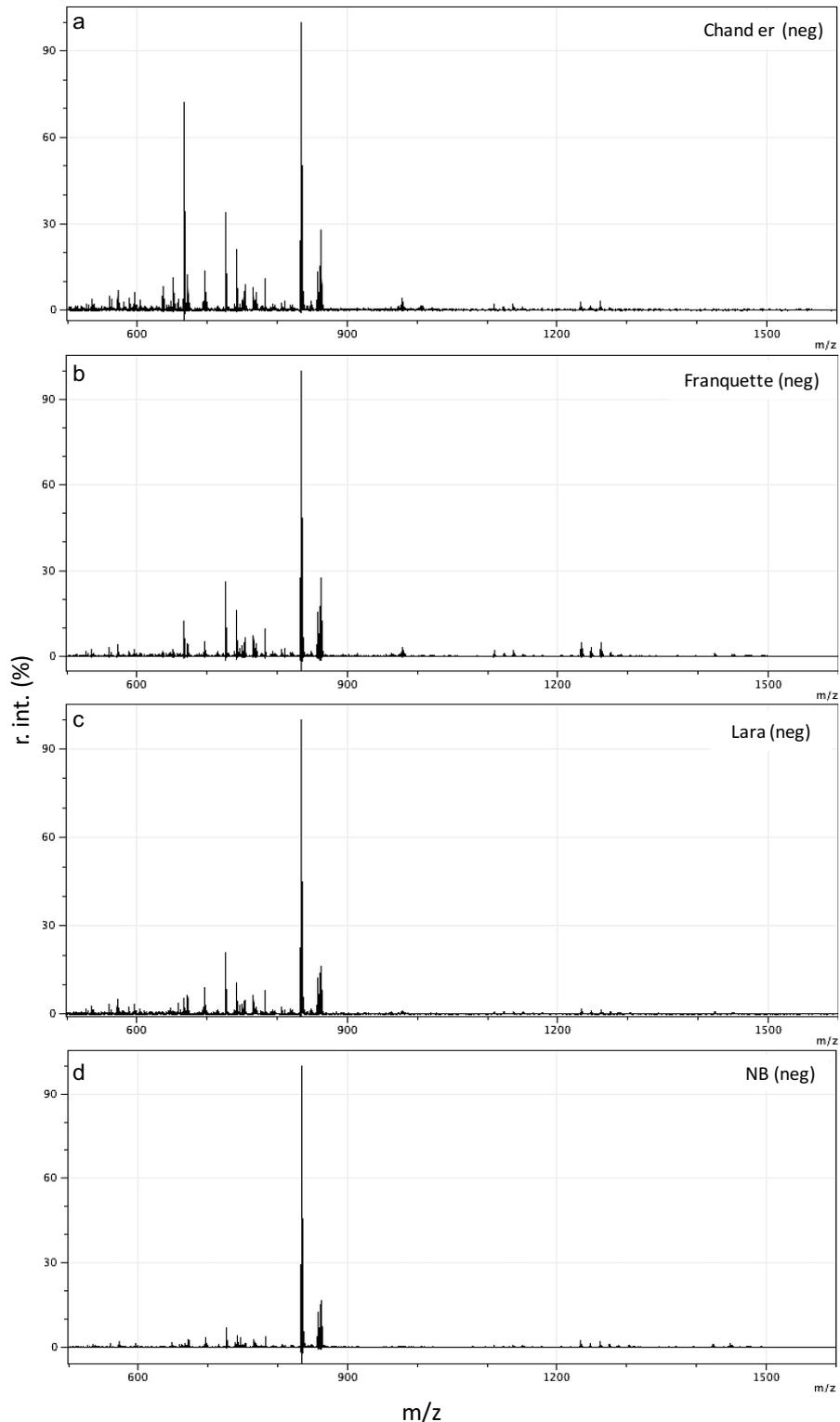


Figure 27: Averaged MALDI-TOF spectra for a four walnut cultivars for the conventional MTBE method measured in negative ion reflector mode in spectrum overview ($n=3$). (a) Chander, (b) Franquette, (c) Lara and (d) NB. The abscissa shows the m/z values in a range from 500 to 1,600. The ordinate represents the normalized relative intensity in %. The spectra were recorded with a laser energy of 35%. Processing of the spectra was done with mMass software as described in section 2.2.6.1.

Subfigure 27 (a) shows the average spectra for the cultivar Chandler. The two sample specific signals with the highest intensities are at approximately 831 m/z (relative intensity 100%) and 860 m/z (relative intensity 29%). Various signals with the relative intensity of below 5% can be seen around 1,000 m/z, 1,200 m/z, and around 1,250 m/z. In the range from 672 m/z to 800 m/z a variety of signals were detected. The signal at approx. 670 m/z shows a relative intensity of 71%. The signal at 710 m/z also shows a relatively high relative intensity of 32%.

Subfigure 27 (b) represents the averaged spectrum for the cultivar Franquette. The signals between 800 m/z and 900 m/z show the same relative intensities as Chandler. Moreover, the signals around 1,100 m/z also share the same relative intensities as the cultivar Chandler. However, the signals around approx. 1,500 m/z she was slightly higher relative intensities. In comparison to Chandler, the signals in the range from 672 m/z to approx. 800 m/z show lower relative intensities, however the same signals in a similar ratio to each other were detected.

Subfigure 27 (c) represents the averaged spectrum for the cultivar Lara. Here, only the two sample specific signals at 830 m/z and 861 m/z are visible with a relative intensity of 100% and 15% respectively. In comparison, the signals at 1,200 m/z are only hinted at. The signals within the range from 670 m/z to 800 m/z show a similar peak ratio as both Chandler and Franquette, however have a lower relative intensity.

The last subfigure 27 (d) shows the averaged spectrum measured in the negative mode for the cultivar NB. The signals in the range from 600 m/z to 800 m/z show less relative intensity in comparison to all other spectra. The sample specific signals can be seen at 830 m/z and 860 m/z around approximately 1,200 m/z and around 1,400 m/z. The signals in both high ranges (from 1,200 m/z to 1,500 m/z) however show a relative intensity below 5%.

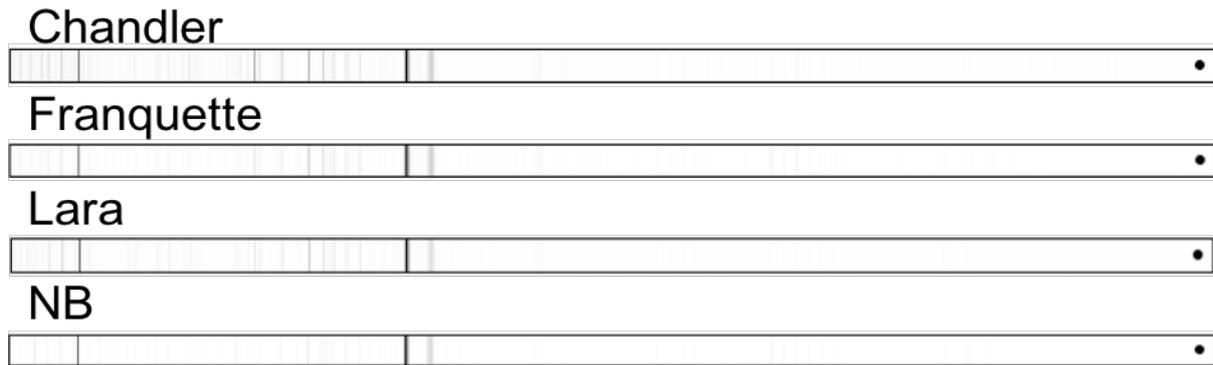


Figure 28: Averaged MALDI-TOF spectra for a four walnut cultivars for the conventional MTBE method in barcode generation measured with the negative polarity. The spectra used for the generation of the barcodes were recorded with a laser energy of 35%. Processing of the spectra was done with mMass software as described in section 2.2.6.1.

Figure 28 shows the respective barcode representations of the four averaged MALDI TOF spectra. The majority of signals were detected in the lower ranges the darkest bands are hereby the one for the signal at 830 m/z and for the signal at 600 m/z then 700 m/z. Chandler shows several thinner but slightly darker bands in the range from 700 m/z to 800 m/z. The signals corresponding to the higher m/z ranges (from 1,200 m/z to 1,500 m/z) are barely visible. Franquette and NB show faded bands corresponding to the signals around 1,200 m/z in the average spectrum.

In comparison to the barcode representation, the average spectra allow the visibility of more signals.

Overall, the MTBE protocol did not show differences between the cultivars in the positive nor the negative ion reflector mode. The usage of the 1:3 ratio is not ideal because the signals at 1,500 m/z are not always visible in the positive mode as opposed to the reference. Moreover, the different ratio leads to more background noise. The additional background noise could potentially lead to signal suppression and therefore falsify the sample specific signals visible in the spectrum. In this work, a better ratio (see figure 23) has been identified. In regards to the negative ion mode, only Franquette shows signals with relative intensities of around 5% in the ranges 1,100 m/z, and 1,200 m/z.

3.2.4. Principal Component Analysis (PCA) for samples treated with the MTBE based method according to cultivar and origin

In a last step, the results were verified with statistical analysis. The PCA analysis was performed as described in section 2.2.6.2. Figure 29 shows the results of the PCA for the lipids in the positive and negative ion reflector mode based on cultivar and origin clustering.

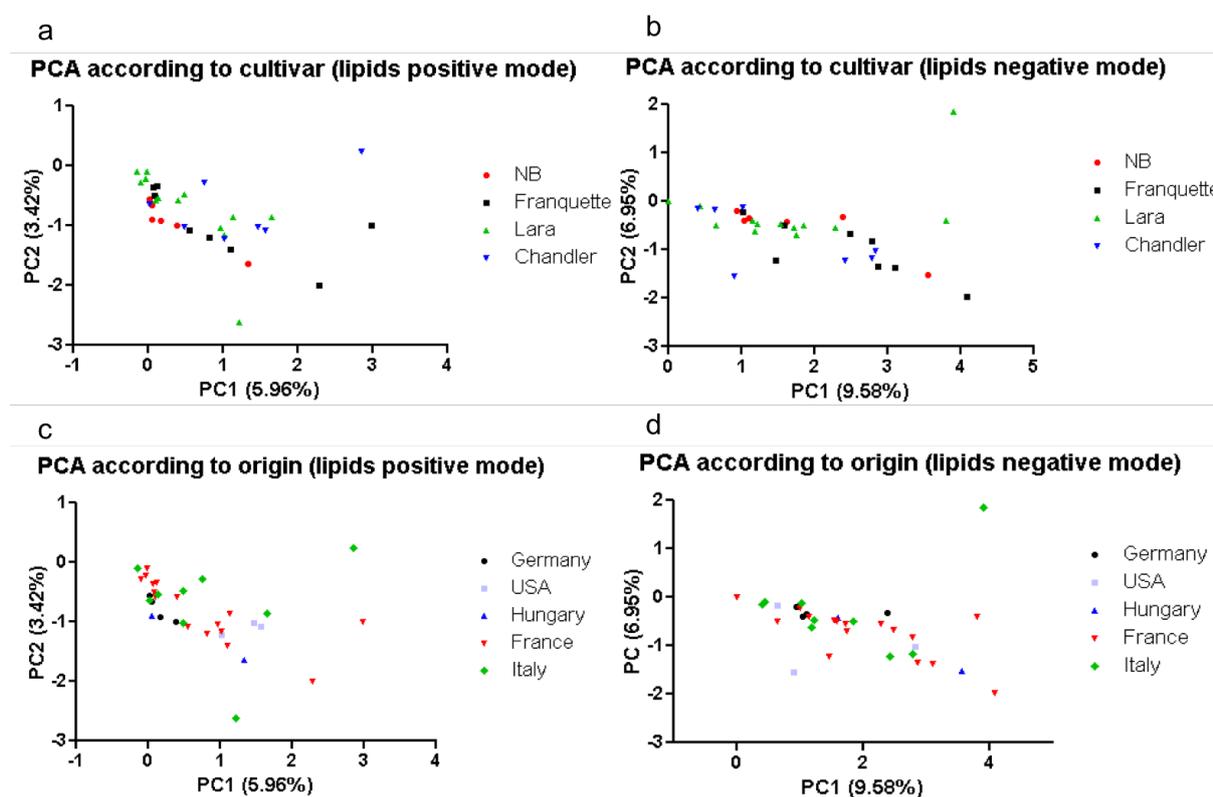


Figure 29: 2D-representation of PCA results according to cultivar and origin for lipids measured in the positive and negative ion reflector mode prepared via a MTBE-based method. (a) PCA according to cultivar for lipids measured in the positive ion reflector mode, (b) PCA according to cultivar for lipids measured in the negative ion reflector mode, (c) PCA according to origin for lipids measured in the positive ion reflector mode, (d) PCA according to origin for lipids measured in the negative ion reflector mode. Samples were exported as mzXML files, loaded into MassUp and pre-processed as described in section 2.2.6.2. PC values were exported to Excel and Prism for the generation of the graphical representation. Color coding (a)-(b): NB = red, Franquette = black, Lara = green, and Chandler = blue; (c) and (d): Germany = black, USA = light blue, Hungary = blue, France = red, and Italy = green.

Subfigure 29 (a) shows the PCA results according to cultivar for the measurements of the lipids in the positive ion mode. The cultivar NB clusters around PC1 values ranging from 0 to 0.5 and PC2 values of 0.72 to 1. An outlier is visible at PC1 = 1.3 and PC2 = 1.8. The cultivar Lara ranges from PC1 values from 0.2 to 1.6 and PC2 values from 1 to 0 with one outlier (PC1 = 1.2, PC2 = 2.89). The cultivars Franquette and Chandler, however, scatter across the graph. Several samples have similarities with Lara or NB samples. Therefore, a clear differentiation and clustering is not visible.

Subfigure 29 (b) represents the PCA results according to cultivar, resulting from the negative ion mode measurements. Here, no clustering based on cultivar is visible. Each sample varies significantly in either the PC1 or PC2 value. Aside from two outliers (PC1 = 4, PC2 = 2; PC1 = 3.8, PC2 = 0), the cultivar Lara (green) clusters better than all other cultivars. The cultivar Chandler (blue), on the other hand, shows two small clusters. One around approx. PC1 = 1 and PC2 between 0.5 and 1 and the other around approx. PC1 = 2.4 to 3 and PC2 between 1 and 0. One Chandler outlier is visible at PC1 = 1 and PC2 = 1.1. The cultivars NB and Franquette show no clustering.

A PCA according to origin was carried out to analyze whether the samples cluster based on the respective origin. In this work, the samples originated from Germany (n=4), the USA (n=3), Hungary (n=4), France (n=16), and Italy (n = 9) were used.

Subfigure 29 (c) shows the PCA results for the lipids measured in the positive ion mode prepared with the MTBE based method. Also here, no clustering is visible. The Italy samples scatter the most with three outliers (PC1 = 1.2, PC2 = 2.8; PC1 = 3, PC2 = 0.8; PC1 = 1.5, PC2 = 1). The samples from France create two smaller clusters around PC1 = 1 to 0 and PC2 = 0 and around PC1 = 0 to 1 and PC2 = 0.5 to 0. The samples from USA cluster slightly, with one sample around PC1 = 1 and PC2 = 1 as an outlier. The sample from Hungary create two smaller clusters.

The PCA results for the clustering according to origin done by using the lipids measured in the negative ion mode is shown in subfigure 29 (d). In this case, the samples from Germany show the tendency of a cluster around PC1 = 1 and PC2 = 0. However, an outlier is visible around PC1 = 2.1 and PC2 = 0.2. All other samples scatter over a PC1 value of 0 and 4 and PC2 value of 0 and 2. One significantly different outlier can be seen for the sample at PC1 = 4 and PC2 = 2, originating from Italy.

Overall, no graph shows either a clustering trend by cultivar nor by origin.

3.3. Analysis of proteins using NIRL

In this work, NIRL was used to ablate the walnuts which subsequently condensed against a slide. Prior to the analysis of the authentic samples, direct and indirect application as well as intra and inter assay were performed.

3.3.1. Direct application of the reference sample on target with HCCA

For direct application, the reference sample was spotted on the target and air dried. The dry sample was coated with 1 μ L HCCA and allowed to air dry for other 5 min. The spectra were then recorded in the linear mode in a range from 2,000 m/z to 20,000 m/z. Figure 30 shows the spectrum resulting after direct application of the reference sample mixed with HCCA.

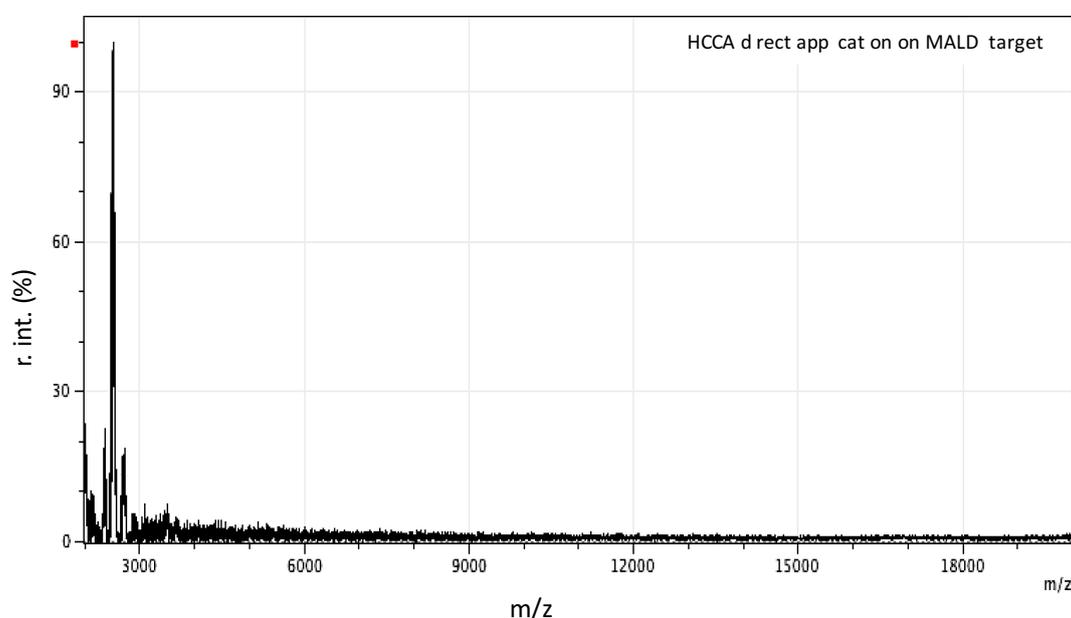


Figure 30: Spectrum for the direct application of the reference sample with HCCA for protein analysis. Sample was first spotted on the target (air-drying for 5 min) and subsequently coated with 1 μ L HCCA (air-dried for additional 5 min). The measurement was performed in the linear mode at a laser intensity of 45%. The abscissa shows the m/z values in a range from 2,000 to 20,000. The ordinate represents the normalized relative intensity in %. Processing of the spectra was done with mMass software as described in section 2.2.6.

In comparison to the proteins extracted with the conventional method here only few proteins in the range between 2,000 m/z and 3,000 m/z are visible. However, it was assumed that the condensed sample contains more proteins. Therefore, a C8 reversed phase chromatography procedure was added to the workflow after NIRL ablation. The aim of the C8 chromatography step was the purification and desalting of the sample as well as protein enrichment.

3.3.2. Addition of the desalting and protein enrichment C8 reversed phase chromatography

The results of exemplary spectra after a C8 chromatographic purification are shown in figure 31 in form of an intra assay. The sample was spotted in a thin layer procedure with HCCA (1:1, v/v).

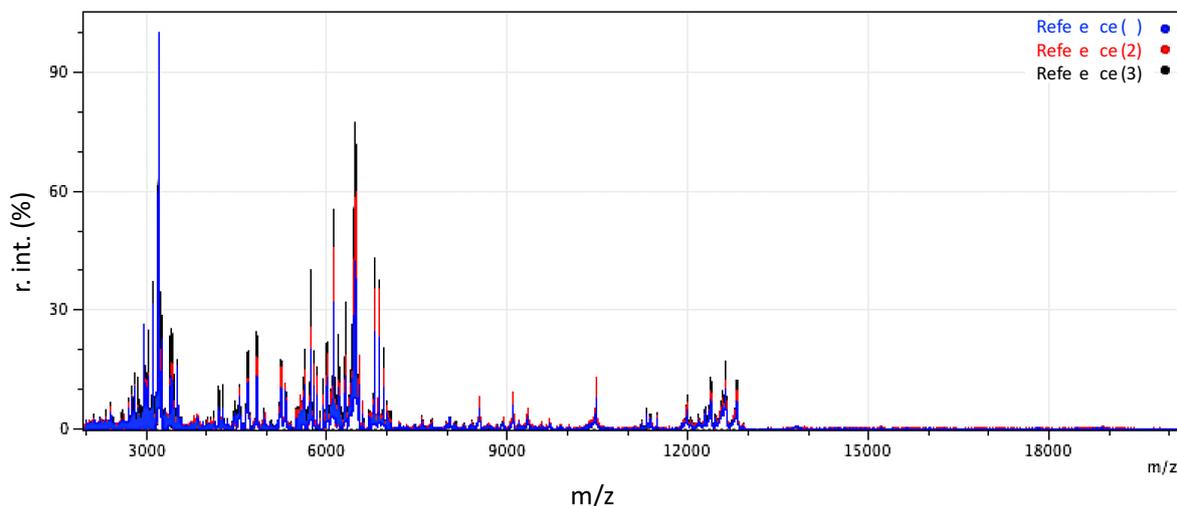


Figure 31: Intra-assay spectra for proteins with the reference sample abated with NIRL. After NIRL ablation, a sample was purified with the C8-reversed phase chromatography protocol (sample dissolved in 0.1% TFA in H₂O) and was analyzed in a 1:1 matrix/analyte ratio (application in the thin layer procedure with subsequent air-drying for 5 min for each step) in near normal mode at a laser intensity of 45%. The abscissa shows the m/z values in a range from 2,000 to 20,000. The ordinate represents the normalized relative intensity in %. Processing of the spectra was done with mMass software as described in section 2.2.6.

Figure 31 shows similar signals as the results obtained with the conventional FA/ACN method. All signals have a signal to noise (S/N) ratio above 6 and were recorded with sufficient intensity.

Additionally, the intra assay shows that the method is reproducible and can be used for further experiments. Moreover, it emphasizes the need of a purification of the condensed sample.

3.3.3. Averaged spectra and barcode gel view for the analysis of proteins extracted with NIRL

After application of the C8 chromatographic procedure on the ablated authentic samples, all averaged spectra were generated with the help of mMass. The four cultivars Chandler (n=7), Franquette (n=8), Lara (n=13) and NB (n=7) are represented. The spectra are visible in figure 32.

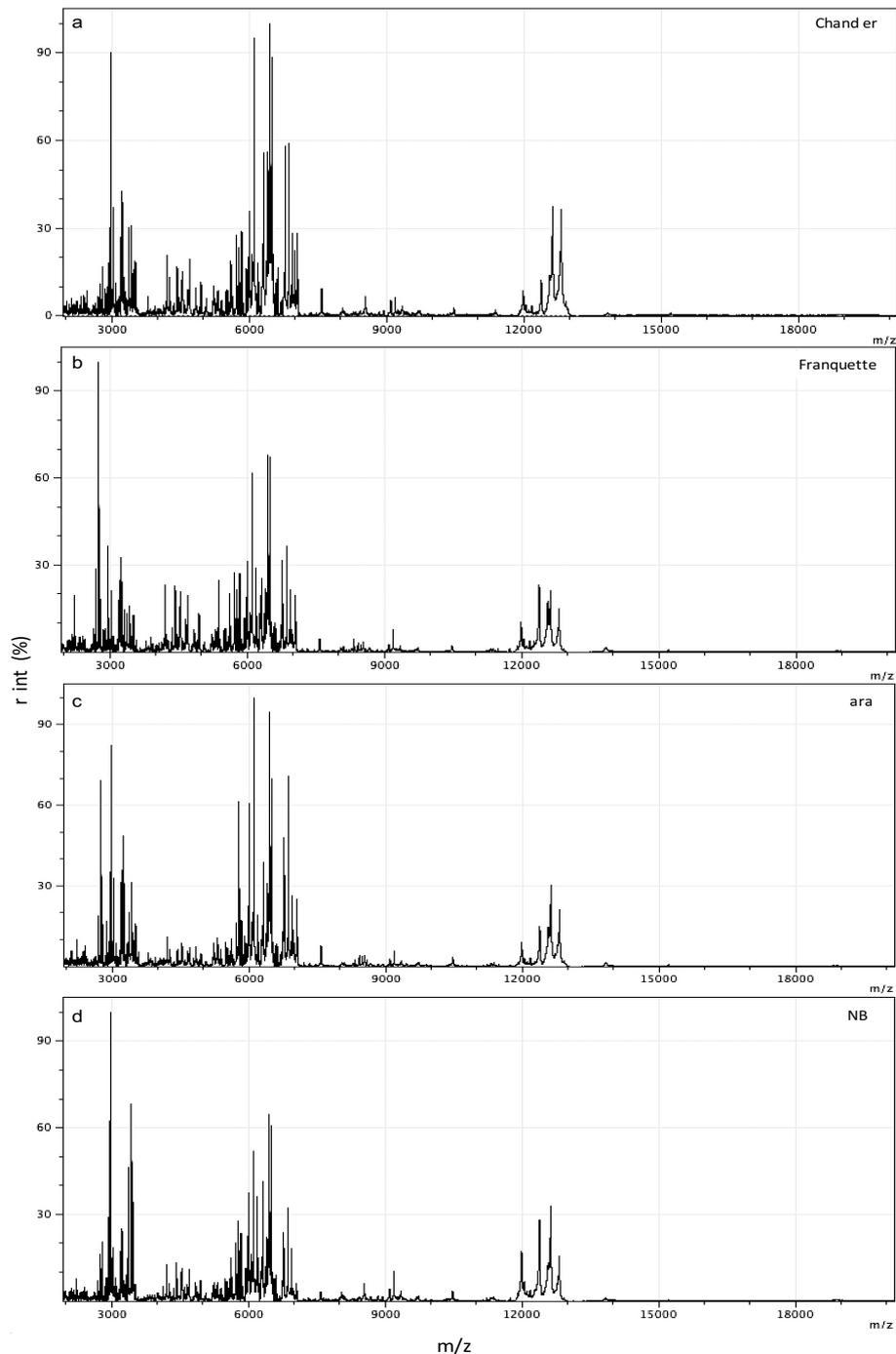


Figure 32: Averaged MALDI-TOF spectra for a four walnut cultivars for protein samples ablated with NIRL measured in near mode in spectrum overview (n=2). (a) Chandler, (b) Franquette, (c) Lara and (d) NB. The abscissa shows the m/z values in a range from 2,000 to 20,000. The ordinate represents the normalized relative intensity in %. The spectra were recorded with a laser energy of 45%. Processing of the spectra was done with mMass software.

All averaged spectra show similar signals as shown in the averaged spectra for the conventional method (figure 13), however clearer distinctions between all cultivars are visible.

Chandler (figure 32 (a)) shows few signals with low relative intensity between 2,000 m/z and 3,000 m/z. The three signals around 3,000 m/z follow a trend with relative intensities lowering from 90% for the first signal, to 42% for the second and lastly, 31% for the third signal. In the ranges between 3,500 m/z and 5,000 m/z and 7,000 m/z and 11,000 m/z several signals with lower relative intensities can be seen. Further prominent signals are in the ranges of approx. 6,000 m/z and approx. 12,000 m/z. The highest signals in the respective ranges have relative intensities of 30% and 20%, respectively.

Franquette (figure 32 (b)), on the other hand, shows several signals with high intensities. These can be seen in the range between 2,000 m/z to 3,000 m/z, between 4,000 m/z and 5,000 m/z and at approx. 9,100 m/z. Around approx. 12,000 m/z, however, the relative intensity of the signals shifted significantly. The first two signals show higher intensities in comparison to Chandler. The signals around approx. 3,000 m/z behave as described for Chandler, but vary in their intensity. A similar behavior can be seen for the signals around 6,000 m/z to 6,500 m/z with only slight intensity differences for the signals around 7,000 m/z.

Lara (figure 32 (c)) behaves similarly to Franquette, however shows lowered intensities for the ranges around 2,000 m/z, between 4,000 m/z and 5,000 m/z and between 7,000 m/z and 11,000 m/z.

NB (figure 32 (d)) shows more similarities to Chandler. Here, the range between 2,000 m/z and 3,000 m/z has few signals. The ranges between 3,500 m/z and 5,000 m/z and between 7,000 m/z 11,000 m/z show signals in similar intensities with only small differences such as the signal at approx. 9,000 m/z.

The signals within the ranges from 2,900 m/z to 3,500 m/z and approx. 6,000 m/z to 7,000 m/z also show similar behaviors, however with different signal intensities.

The signals around 12,000 m/z show high similarities with Lara, however with different relative intensities.

Tables 22 25 show the main signal and the respective signal intensities for each cultivar.

Table 22: Signals and their relative intensity and presence, respectively, for the averaged Chandler spectrum for protein analysis after NIRLab at on

Signal (in m/z)	Relative intensity (in %)	Presence
2,000	n.a.	No
2,900	90	Yes
3,000	42	Yes
3,100	31	Yes
6,050	94	Yes
6,500	100	Yes
6,850	49	Yes
6,900	29	Yes
6,999	29	Yes
12,000	10	Yes
12,400	15	Yes
12,600	40	Yes
12,800	39	Yes

Table 23: Signals and their relative intensity and presence, respectively, for the averaged Franquette spectrum for protein analysis after NIRLab at on

Signal (in m/z)	Relative intensity (in %)	Presence
2,000	n.a.	No
2,900	100	Yes
3,000	21	Yes
3,100	31	Yes
6,050	61	Yes
6,500	70	Yes
6,850	35	Yes
6,900	21	Yes
6,999	20	Yes
12,000	10	Yes
12,400	26	Yes
12,600	25	Yes
12,800	17	Yes

Table 24: Signals and their relative intensity and presence, respectively, for the averaged Lara spectrum for protein analysis after NIRLab at on

Signal (in m/z)	Relative intensity (in %)	Presence
2,000	n.a.	No
2,900	70	Yes
3,000	82	Yes
3,100	30	Yes
6,050	100	Yes
6,500	95.5	Yes

6,850	70	Yes
6,900	27	Yes
6,999	26	Yes
12,000	11	Yes
12,400	15	Yes
12,600	30	Yes
12,800	21	Yes

Table 25: Signals and their relative intensity and presence, respectively, for the averaged NB spectrum for protein analysis after NIRL ablation

Signal (in m/z)	Relative intensity (in %)	Presence
2,000	n.a.	No
2,900	100	Yes
3,000	25	Yes
3,100	69	Yes
6,050	29	Yes
6,500	52	Yes
6,850	32	Yes
6,900	20	Yes
6,999	7	Yes
12,000	18	Yes
12,400	29	Yes
12,600	32	Yes
12,800	18	Yes

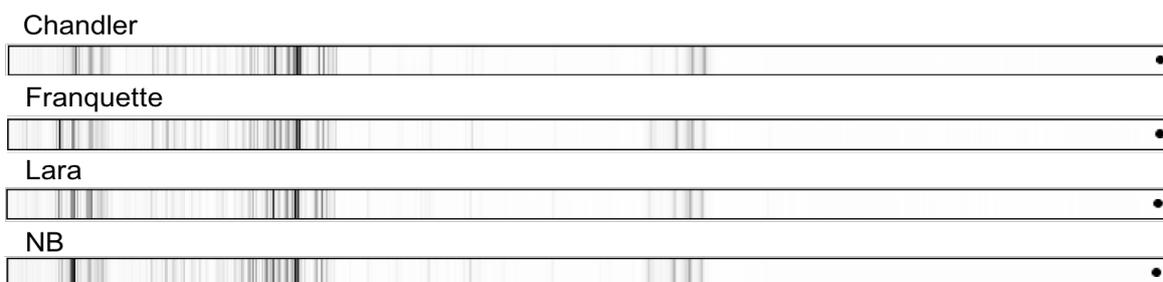


Figure 33: Averaged MALDI-TOF spectra for a four walnut cutvars for protein samples ablated with NIRL in barcode overview. The spectra used for the generation of the barcodes were recorded with a laser energy of 45%. Processing of the spectra was done with mMass software.

The differences described in the averaged spectra also visible in the generated barcode overview. Based on the signal intensities clear differentiation is possible. According to the barcode representation, Chandler and NB show a similar behavior. NB has a dark and thick band for the speak corresponding to the signal at 2,900 m/z (100%). Chandler showed only relative intensity of 90% for the same peak. This translated into a slightly thinner band. A similar behavior is also visible for the other NB bands. Franquette and Lara, on the other

hand, show a darker and thicker band around the signals corresponding to 2,800 m/z. All other bands have a similar intensity to NB and correspond to the same signals.

3.3.4. Principal Component Analysis (PCA) for proteins homogenized via NIRL according to cultivar and origin

In a last step, a PCA was performed for clustering analysis of the individual samples. The results are shown in figure 34.

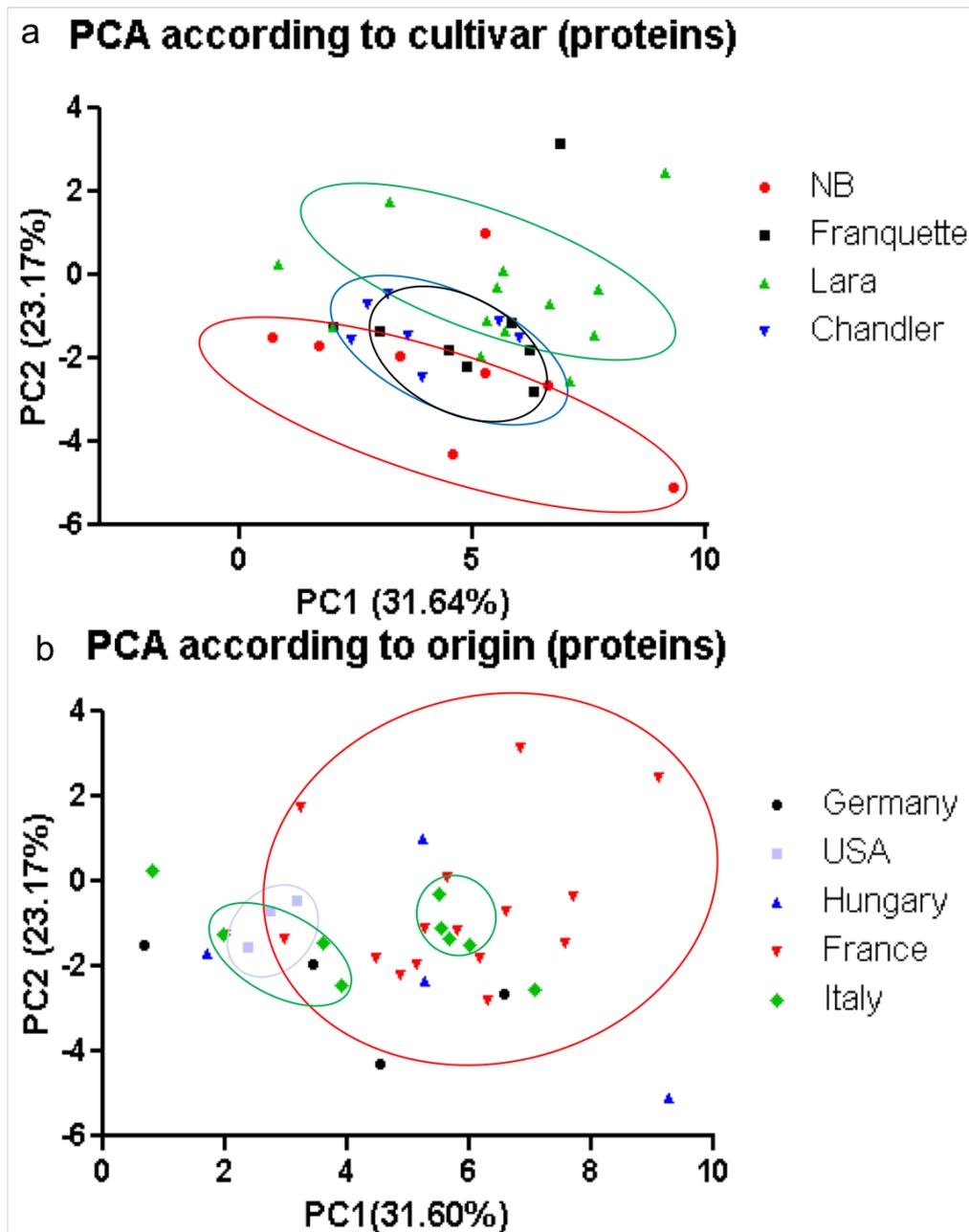


Figure 34: 2D-representation of the PCA results according to cultivar and origin for proteins prepared via NIRL. (a) PCA according to cultivar, and (b) PCA according to origin. Color code (a) NB = red, Franquette = black, Lara = green, and Chandler = blue, and (b): Germany = black, USA = light blue, Hungary = blue, France = red, and Italy = green. Samples were exported as mzXML files, loaded into MassUp and pre-processed as described in section 2.2.6.2. PC values were exported to Excel and Prism for the generation of the graphical representation.

Figure 34 represents the PCA results according to cultivar and origin for the proteins homogenized via NIRL.

No clear clustering is visible for subfigure 34 (a), however a trend is slightly recognizable. The NB samples tend to aggregate between a PC1 value of 1 to 5 and a PC2 value of 2. The Chandler samples range from PC1 values of 1 to 4 and PC2 values of 1.9 to 0. The Franquette samples show similar PC2 values as Chandler, however, the majority of samples have a PC1 value ranging between 5 and 6. And lastly the cultivar Lara, showing a PC2 range from 2 to 0 and PC1 value from approx. 5 to 7. All cultivars show several outliers, out of which Lara has the most (total of 4). These outliers are positioned at PC1 = 0.1 and PC2 = 0.1, PC1 = 4 and PC2 = 1.8, PC1 = 1 and PC2 = 1.8, and PC1 = 9.1 and PC2 = 2.

The cultivar NB also shows three significant outliers at PC1 = 9, PC2 = 5; PC1 = 4.4, PC2 = 4.1, PC1 = 5 and PC2 = 1. The cultivar Franquette has only one outlier (PC1 = 6.5 and PC2 = 3).

All Chandler samples can be found within the range mentioned before.

A PCA according to origin was carried out to analyze whether the samples cluster based on the respective origin (samples originated from Germany (n=4), the USA (n=3), Hungary (n=4), France (n=16), and Italy (n = 9)). The PCA results are visible in subfigure 34 (b). The samples coming from Germany, Hungary, Italy and France do not cluster. As opposed to that, the samples coming from USA show a cluster, ranging from PC1 values of 2.3 to 3.4 and a PC2 values from 1 to 2. The samples from Italy create two separate clusters. Even if the samples from France do not cluster, a clustering trend in the ranges from approx 4 to 8 (PC1) and 0 to 2.5 (PC2) is visible. However, overall no clustering and differentiation of the samples based on origin is possible.

To sum up, no direct differentiation into either cultivar nor origin based on PCA analysis is visible. However, a slight trend can be seen for the differentiation into cultivars.

3.4. Analysis of lipids using NIRL

Similarly to the proteins, lipids were obtained by ablation and subsequent condensation of the walnut sample on a slide.

Prior to the analysis of the authentic samples, direct and indirect application as well as intra and inter assay were performed. These preliminary experiments were used with the aim of developing the NIRL method. All MALDI TOF specific parameters that were optimized in section 3.2 were applied for the spectra acquisition of the ablated sample.

3.4.1. Direct and indirect spotting with 9 AA on NIRL ablated samples on the ground steel target

In order to identify whether a sample solvent is needed, the condensed sample was either mixed with 9 AA (isopropanol/ACN 60:40 (v/v)) on the slide and directly spotted on the MALDI target or first pre mixed with CHCl_3 /MeOH in different solvent ratios (2:1, 1:1, and 1:2), CHCl_3 or ACN/ H_2O , then mixed with 9 AA and applied on the MALDI target. For the preliminary experiments a reference was used.

The spectra are shown in figure 35 for the positive mode and figure 36 for the negative mode, respectively.

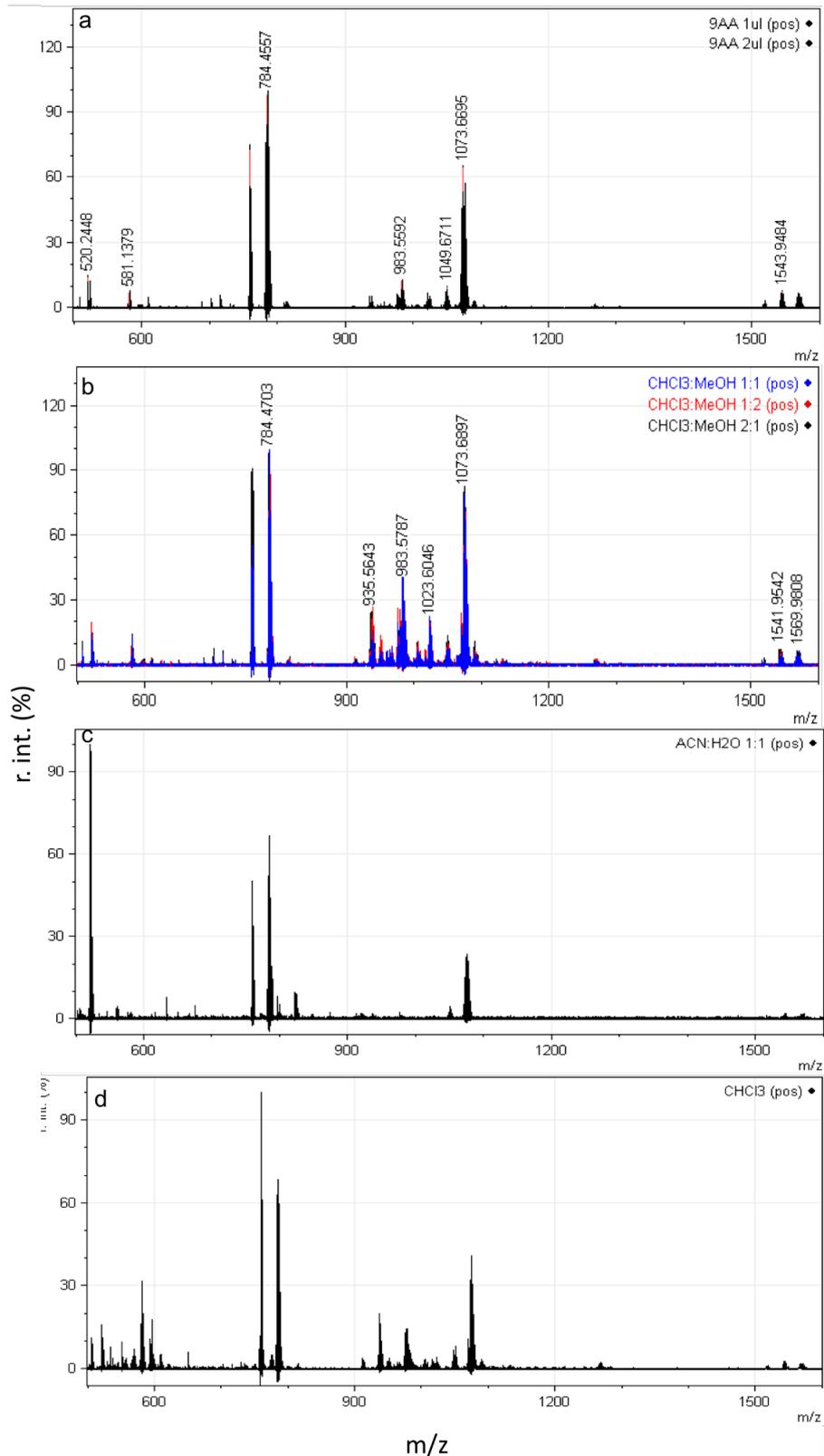


Figure 35: Comparison between the direct and indirect application of reference walnut sample on the MALDI-TOF ground steel target in the positive ion reflector mode after NIRL ablation. (a) direct application with 9-AA from the side on the MALDI-TOF ground steel target, (b) indirect application through premixing of the sample with different CHCl₃/MeOH ratios with subsequent 1:1 mixing with 9-AA matrix solution, (c) mixing with ACN/H₂O and (d) mixing with CHCl₃. The abscissa shows the m/z values in a range from 500 to 1,600 m/z. The ordinate represents the normalized relative intensity in %. The spectra were recorded with a laser energy of 35%. The signals at 520.2448 m/z and 581.1379 m/z are not relevant for product analysis. Spectra were pre-processed using the mMass software as described in section 2.2.6.1.

Subfigure 35 (a) shows the spectra of the condensed sample mixed with 9 AA. The sample was either mixed with 1 μL or 2 μL and subsequently 0.5 μL were spotted on the target. The sample peaks at 758 m/z and at 780 m/z and the triplet signals at 1,500 m/z are visible. In the range from 900 m/z to 1,000 m/z multiple signals can be seen. Hereby, the signal with the highest intensity is 1,073 m/z (approximate rel. int. 59%). In comparison to figure 25 however, the background noise is reduced.

Subfigure 35 (b) shows the spectra that were obtained if the sample was pre mixed with $\text{CHCl}_3/\text{MeOH}$ as sample solvent in different ratios (2:1, 1:1 and 1:2). The signal triplets and the two signals between 700 m/z and 800 m/z are still visible, however the range between 900 m/z and 1,000 m/z shows signals with higher intensity in comparison to Subfigure 35 (a).

Prior mixing with $\text{ACN}/\text{H}_2\text{O}$ (subfigure (c)) lead to significant increase in the background noise of 9 AA, in particular the matrix signal 510 m/z. In comparison, the sample specific signals are suppressed with mostly the two signals at 758 m/z and 780 m/z visible. Two further peaks are visible around 800 m/z to 900 m/z. The signal triplet visible in the previous two spectra are only hinted at.

Usage of Chloroform (Subfigure 35 (d)), also increases the amount of background noise created by the 9 AA matrix. However, less signal suppression is noticeable. The signals in the range from 700 m/z to 800 m/z as well as the signals from 900 m/z to 1,000 m/z were all detectable with a S/N ratio of at least over 6% relative intensity. The signal triplets around approx. 1,500 m/z were only detectable with a low S/N ratio and show low relative intensity.

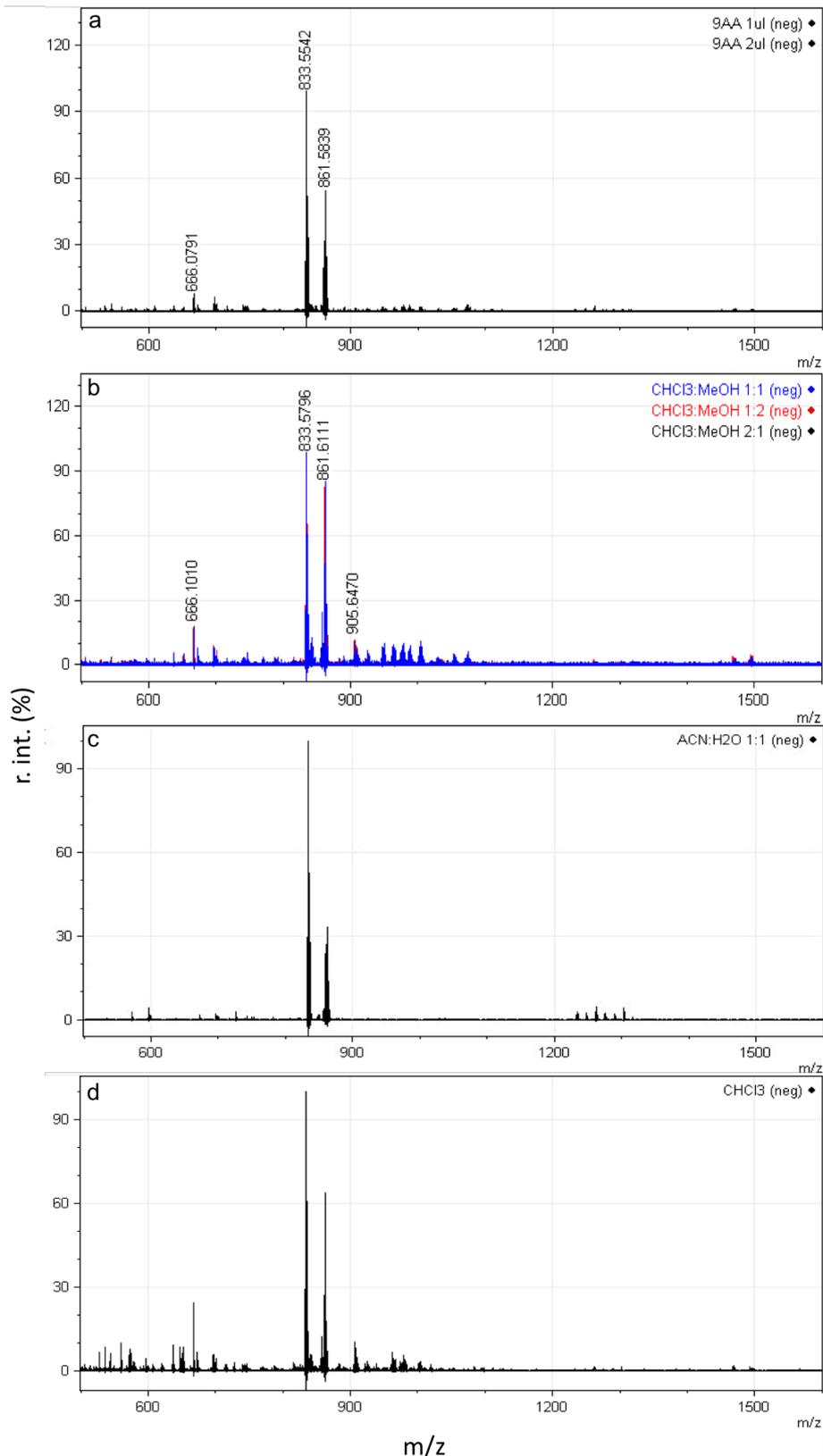


Figure 36: Comparison between the direct and indirect application of reference walnut sample on the MALDI-TOF ground steel target in the negative ion reflector mode after NIRL ablation. (a) direct application from the side on the MALDI-TOF ground steel target, (b) indirect application through premixing of the sample with different $\text{CHCl}_3/\text{MeOH}$ ratios with subsequent 1:1 mixing with 9-AA matrix solution on, (c) mixing with $\text{ACN}/\text{H}_2\text{O}$ and (d) mixing with CHCl_3 . The abscissa shows the m/z values in a range from 500 to 1,600 m/z . The ordinate represents the normalized relative intensity in %. The spectra were recorded with a laser energy of 30%. All spectra were pre-processed using the mMass software.

Subfigures 36 (a) and (b) show the negative mode for the direct and indirect application of the reference sample on the MALDI target. The indirect application of the sample yields better spectra in comparison to the direct application since more signals are visible, however, slightly more background noise was detected. Overall, the background noise is negligible. Moreover, there is no difference between any of the sample solvents used, therefore the ratios 2:1, 1:1 and 1:2 could all be used if desired.

Subfigure 36 (c) shows the spectrum of reference sample mixed with ACN/H₂O (50:50 (v/v)). Here, two prominent signals in the range from 800 m/z to 900 m/z were detected. Furthermore, different signals in the range from 1,200 m/z to 1,300 m/z can be seen. The background noise produced by 9 AA is negligible since all sample specific signals are visible with a good relative intensity and show an adequate S/N ratio. As opposed to subfigure (b), no signals in the range from 900 m/z to 1,000 m/z are detectable.

The mix with chloroform (Subfigure 36 (d)), on the other hand, shows an increase in matrix background signals in the range from 500 m/z to 600 m/z. All major sample specific signals are suppressed, however visibility of signals in the range from 900 m/z to 1,000 m/z are detectable. The signals around 1,200 m/z are hinted at, but do not show a high enough S/N ratio.

For further experiments, direct application of 9 AA was chosen because the same signals as with the chloroform sample solvent mix could be identified. Moreover, the overall handling is quicker and simpler. Pre mixing with CHCl₃, CHCl₃/MeOH or ACN/H₂O led to problems in pipetting and sample handling due to lack of surface tension.

3.4.2. Intra and inter assay for the NIRL based method for phospholipid analysis

Next, intra and inter assays were carried out. Here, after ablation the direct application with 9 AA was chosen. The spectra are visible in figure 37.

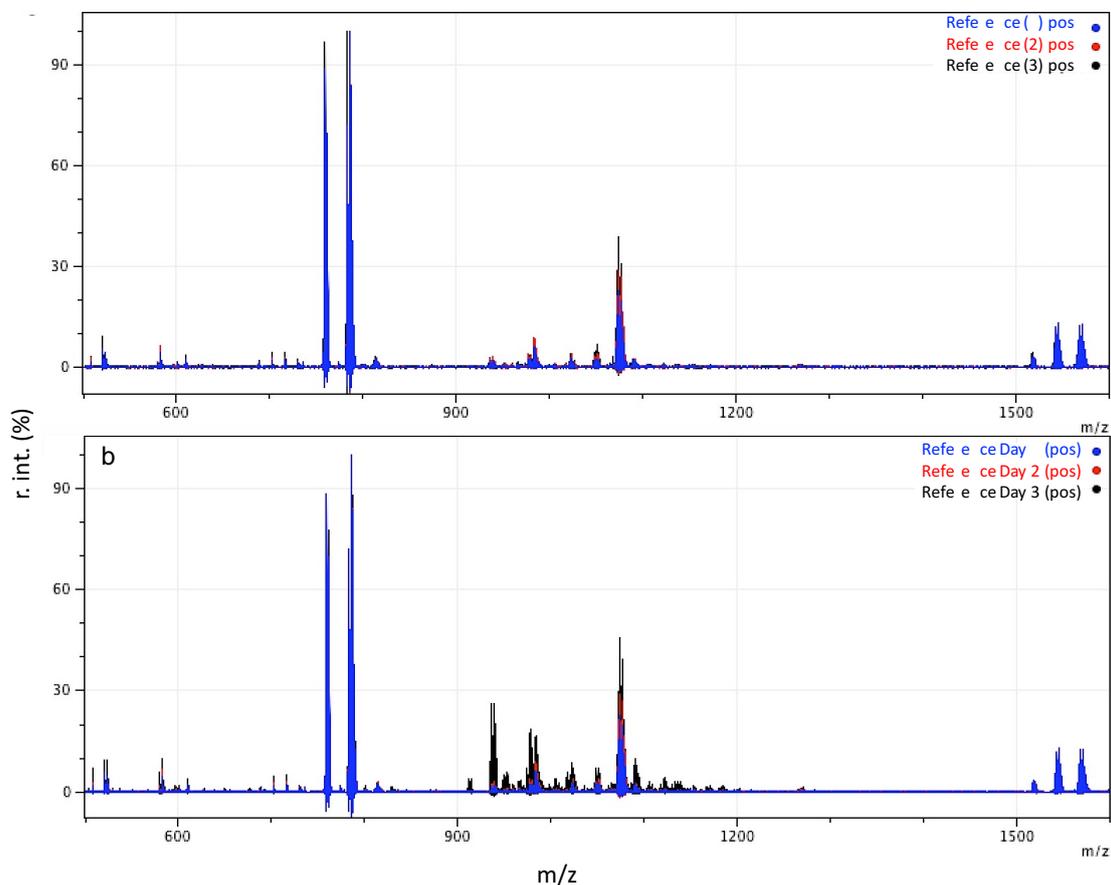


Figure 37: Intra and inter-assay spectra for the ablation on with NIRL in the positive ion reflector mode. The abscissa shows the m/z values in a range from 500 to 1,600. The ordinate represents the normalized relative intensity in %. (a) represents the intra-assay and (b) the inter-assay spectrum, respectively. The spectra were recorded with a laser energy of 35%. Processing of the spectra was done with mMass software.

Figure 37 (a) shows the intra assay on day 1 for the positive mode. As described in the previous sections, the signal triplet is visible as well as the two signals with the highest intensity and the signals, ranging from 900 m/z to 1,200 m/z . All three spectra look identical with only slight variations in the intensity of certain signals.

A similar result is visible in Subfigure 37 (b). Therefore, the intra assay and the inter assay for the positive mode were regarded as successful.

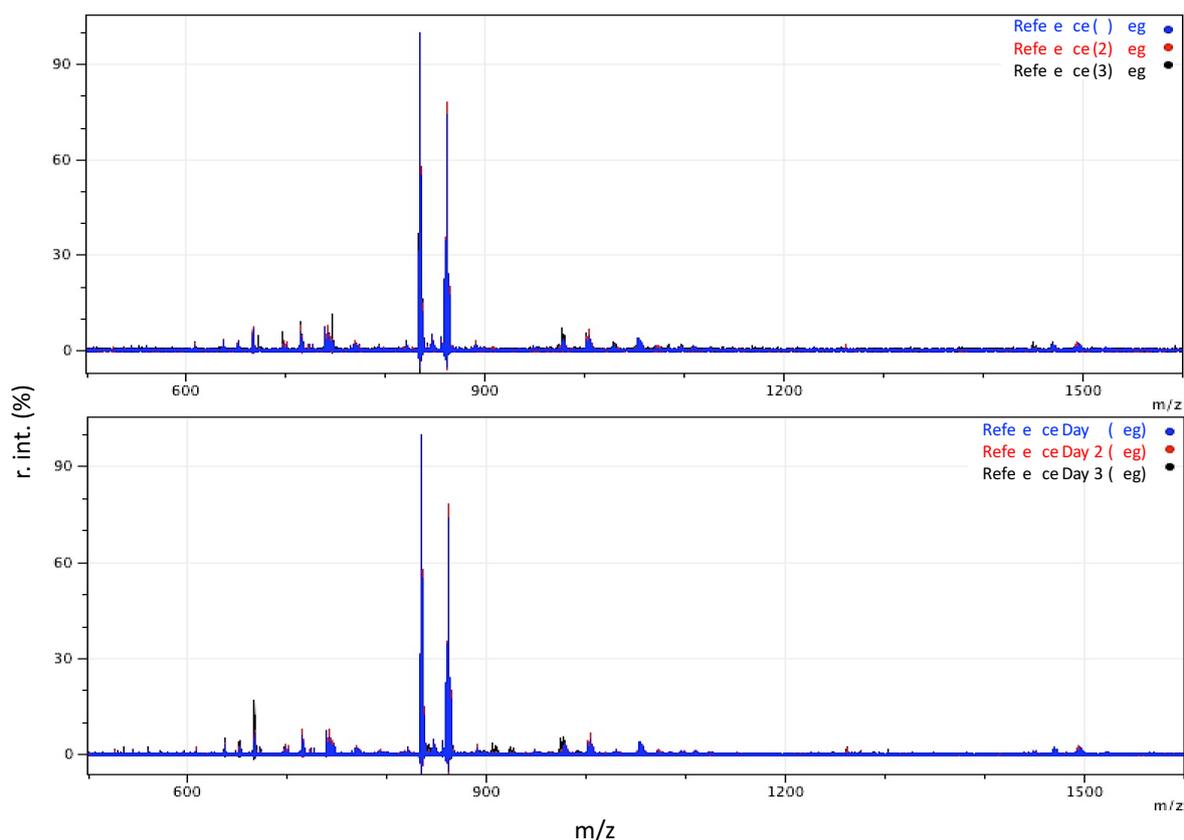


Figure 38: Intra and inter-assay spectra for the ab on with NIRL for the negative ion reflector mode. The abscissa shows the m/z values in a range from 500 to 1,600. The ordinate represents the normalized relative intensity (%). (a) represents the intra-assay and (b) the inter-assay spectrum, respectively. The spectra were recorded at a laser energy of 30%. Processing of the spectra was done with mMass software.

Figure 38 shows intra and inter assay for the negative ion polarity. All spectra show the same peaks at approx. 830 m/z and 860 m/z . Further signals are present around 1,000 m/z . Similarly to the positive ion polarity, the spectra show reproducible results. The intra and inter assays were therefore regarded as successful.

The intra and inter assay were successful for both the positive and negative polarity. The authentic samples were therefore measured with the direct application from the slide on the MALDI target via 9 AA.

3.4.3. Averaged spectra and barcode gel view for the analysis of phospholipids extracted with NIRL

As described in the previous section, the direct sample application was chosen. Figure 39 shows the four averaged spectra in the positive mode for the cultivars Chandler (n=7), Franquette (n=8), Lara (n=13), and NB (n=8), respectively.

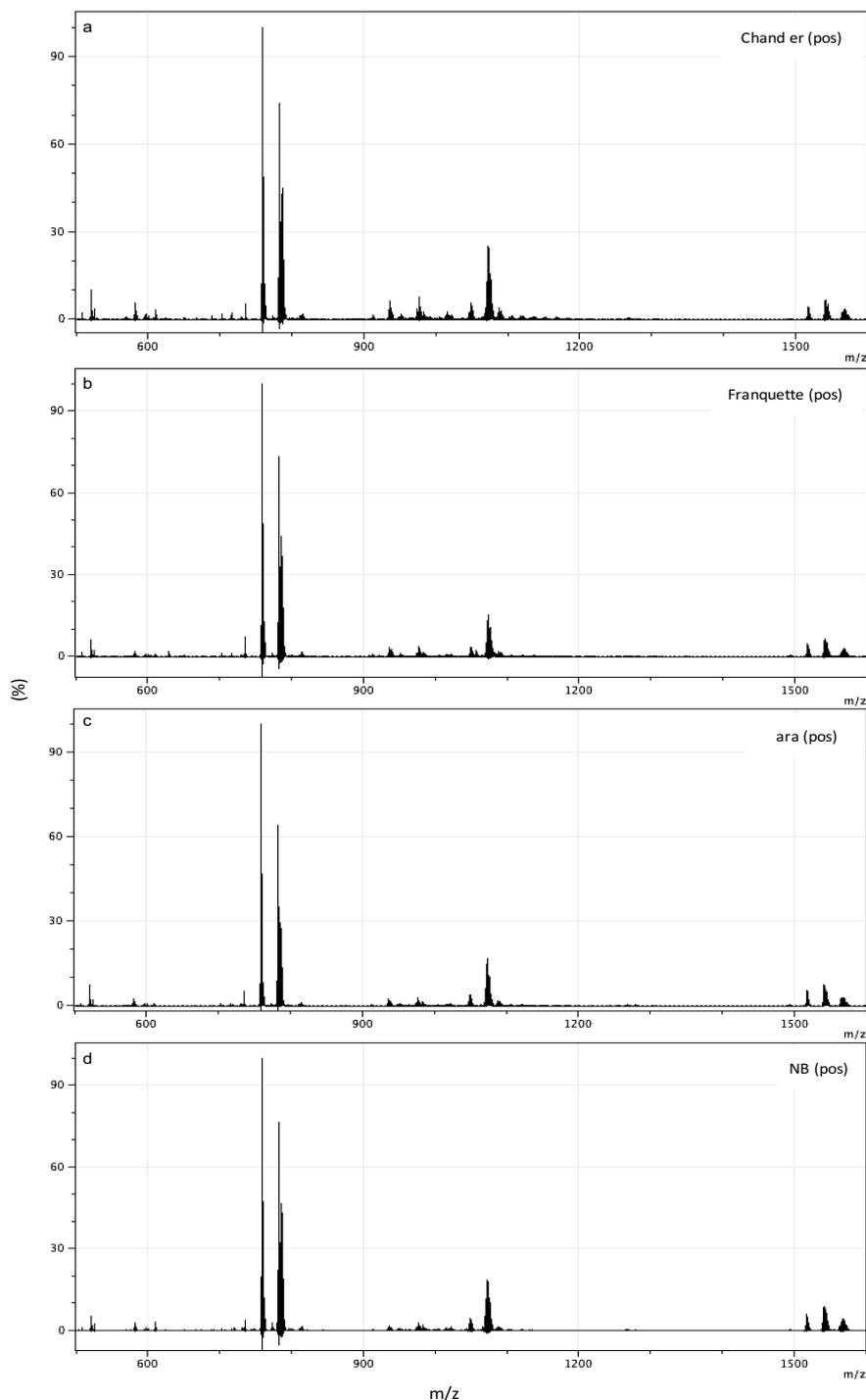


Figure 39: Averaged MALDI-TOF spectra for a four walnut cultivars for lipid samples ablated with NIRL measured in the positive ion mode in spectrum overview (n=3). The abscissa shows the m/z values in a range from 500 to 1,600. The ordinate represents the normalized relative intensity in %. The spectra were recorded with a laser energy of 35%. Processing of the spectra was done with mMass software.

All averaged spectra (figure 39) show the same signals, however, high variances in the respective intensities can be seen. Chandler has most signals with a high intensity. As opposed to that both Franquette and Lara show less relative intensity in the signals, ranging from approx. 900 m/z to 1,200 m/z. The signal triplet around 1,500 m/z is visible in all four cultivars.

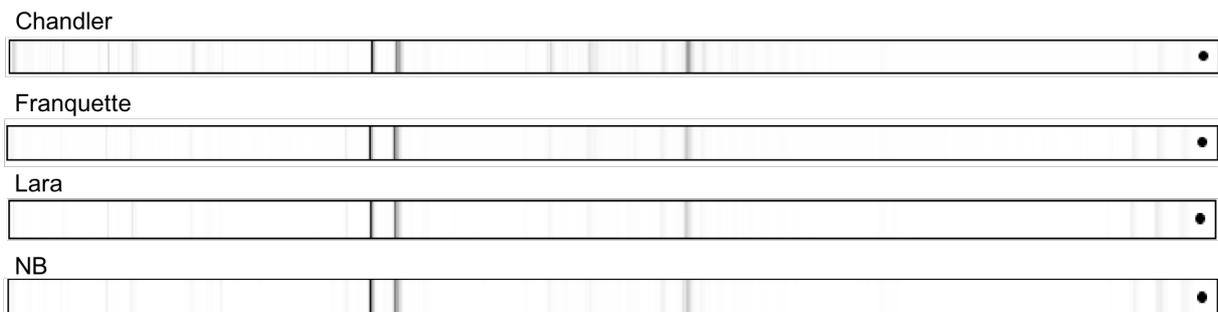


Figure 40: Averaged MALDI-TOF spectra for a four walnut cultivars for pooled samples abated with NIRL in barcode representation. The spectra used to generate the barcode overviews were recorded with a laser energy of 35% and in positive polarity. Processing of the spectra was done with mMass software.

Figure 40 shows the barcode representations of all four cultivars. The above described differences can also be seen on the following gel views. Chandler shows darker and thicker bands than the other three cultivars. This is visible in particular for the bands of the matrix signals as well as the bands around the range 900 m/z to approx. 1,100 m/z. The bands representing the signal triplet around 1,500 m/z, however, are less visible. The cultivars Franquette, Lara and NB show similar bands of nearly identical thickness and brightness. Only for Franquette, the bands for the signals triplets at 1,500 m/z are lighter.

Phospholipids were also recorded in the negative ion reflector mode. The samples were applied on the MALDI target with the direct application procedure. Following sample sizes were chosen for the cultivars: Chandler (n=7), Franquette (n=8), Lara (n=13), and NB (n=8), respectively. The averaged spectra for all cultivars is shown in figure 41.

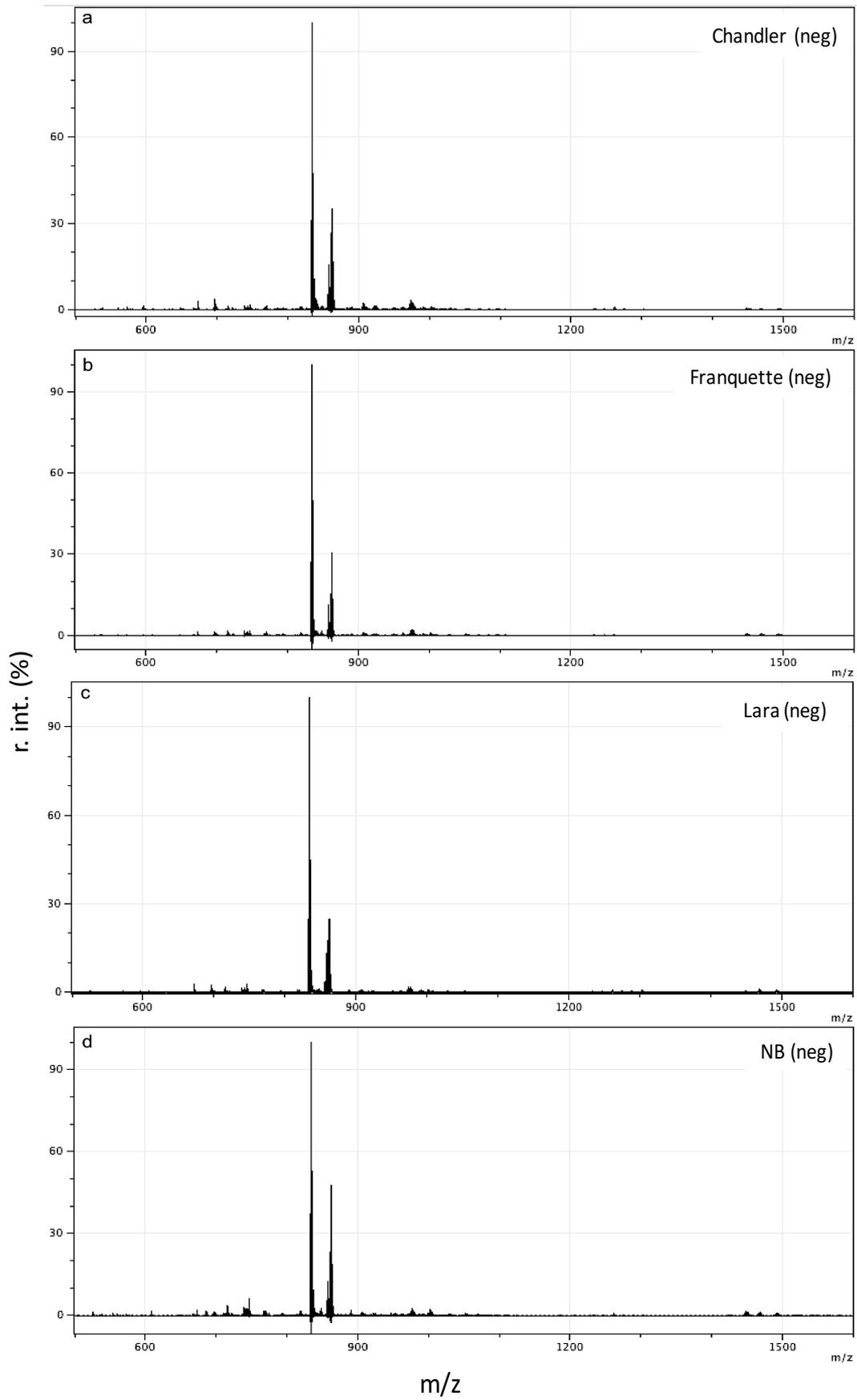


Figure 41: Averaged MALDI-TOF spectra for a four walnut cultivars for lipid samples abated with NIRL measured in the negative ion mode in spectrum overview (n=3). The abscissa shows the m/z values in a range from 500 to 1,600. The ordinate represents the normalized relative intensity in %. The spectra were recorded with a laser energy of 30%. Processing of the spectra was done with mMass software.

Subfigure 41 (a) shows the averaged spectrum of the cultivar Chandler. The two peaks with the highest relative intensity are at approx. 830 m/z and approx. 861 m/z. Several other signals between 680 m/z and 800 m/z as well as between 900 m/z and 1,000 m/z are visible, however show a relative intensity below 10%. At 1250 m/z, one signal with a relative intensity of 1% was detected. Moreover, another signal with a relative intensity of 1% was detected at approx. 1400 m/z.

In the range between 500 m/z and 600 m/z barely any matrix signal is visible. Only two signals at approx. 660 m/z and approx. 670 m/z are hinted at but show a relative intensity of below 5%.

Subfigure 41 (b) shows the average spectrum in the negative mode for the cultivar Franquette. Similarly to Chandler, two signals with high relative intensities were detected between 800 m/z and 900 m/z. Signals in the range from approx. 700 m/z to 800 m/z and in between 900 m/z and 1,000 m/z are visible, but show a small relative intensity. In addition, at approx. 1500 m/z three signal could be detected. In this range, more signals were detected for the cultivar Franquette as opposed to the cultivar Chandler.

Subfigure 41 (c) represents the averaged spectrum for the cultivar Lara. In comparison to subfigures 41 (a) and 41 (b), less signals were detected for Lara. The signals with the highest intensity are at approx. 830 m/z and 860 m/z. Several signals around 500 m/z and 800 m/z as well as 900 m/z and 1,000 m/z are visible. The signals at higher ranges are only hinted at. Lastly, subfigure 41 (d) represents the averaged spectrum of NB recorded in the negative ion mode. The behavior of all signals is similar with a cultivar Chandler. Also here no background matrix signals were detected.

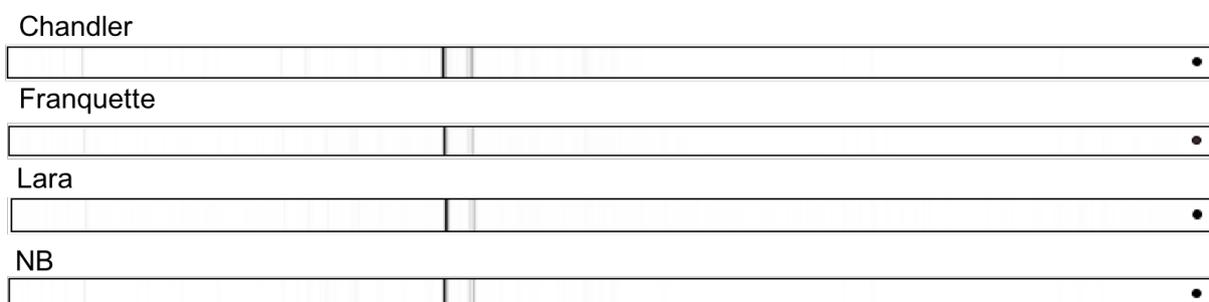


Figure 42: Averaged MALDI-TOF spectra for a four walnut cultivars for p d samples abated with NIRL n barcode representat on. The spectra used to generate the barcode overv ews were recorded with a laser energy of 35% and n negat ve po ar ty. Process ng of the spectra was done with mMass software.

Figure 42 shows the barcode representation generated for the four recorded averaged spectra. As opposed to the averaged spectra, only two bands are clearly visible. These bands correspond to the signals at 830 m/z and 860 m/z. Since the signal at 830 m/z showed a higher relative intensity in all averaged spectra, the respective bands in the barcode overview are represented thicker and darker than the signal bands corresponding to the 860 m/z signal.

Moreover, all barcode overviews show a thin band corresponding to the 580 m/z signal visible in each averaged spectrum.

Few thin and light bands are noticeable for the ranges corresponding to approx. 670 m/z to approx. 800 m/z. All barcode representations show similar bands with differences only in the brightness and thickness of the signals.

Overall, both the positive and the negative mode do not show clear differences in the phospholipids present in the four cultivars. However, all results are reproducible and deliver clear spectra without too much matrix background noise.

3.4.4. Principal Component Analysis (PCA) for lipids homogenized via NIRL according to cultivar and origin

In a last step, a PCA was performed. The results are visible in figure 43 (a) and (c) for the positive and figure 43 (b) and (d) for the negative ion mode.

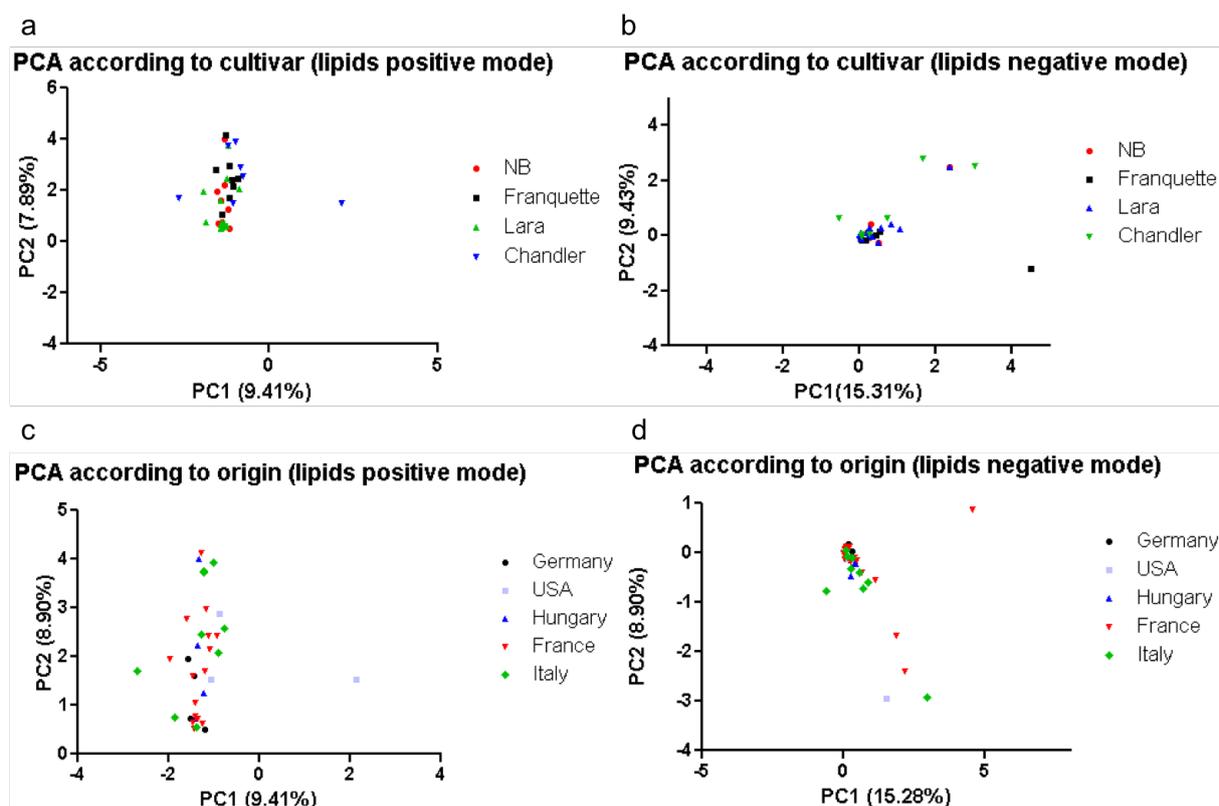


Figure 43: 2D-representation of PCA results according to cultivar and origin for lipids measured in the positive and negative ion reflector mode prepared via NIRL. (a) PCA according to cultivar for lipids measured in the positive ion reflector mode, (b) PCA according to cultivar for lipids measured in the negative ion reflector mode, (c) PCA according to origin for lipids measured in the positive ion reflector mode, (d) PCA according to origin for lipids measured in the negative ion reflector mode. Samples were exported as mzXML files, loaded into MassUp and pre-processed as described in section 2.2.6.2. PC values were exported to Excel and Prism for the generation of the graphical representation. Color coding (a)-(b): NB = red, Franquette = black, Lara = green, and Chandler = blue; (c) and (d): Germany = black, USA = light blue, Hungary = blue, France = red, and Italy = green.

Subfigure 43 (a) shows the PCA results for the lipids measured in the positive ion reflector mode and sorted according to cultivar. No clear clustering of the respective cultivars is visible, however NB, Franquette and Lara cluster closer together as opposed to Chandler. In the cultivar Chandler, a total of two samples are scattered, with a PC1 value of approx. 2 and PC2 value of 1.8 and a PC1 value of 2 and PC2 value of 1.7, respectively. For all other samples, the PC1 value is similar (approx. 1). The PC2 value, on the other hand varies between 0 and 4 and is therefore responsible for the scattering of the samples.

Subfigure 43 (b) represents the PCA results according to cultivar that resulted from the negative ion mode measurements. Here, all samples have approx. both a similar PC1 and PC2 value, hinting at identical signals. In total, 9 outliers are visible out of which two Lara and two Chandler samples show a similar PC2 value but differ in PC1 in comparison to the main cluster. The other five outliers are two Chandler, one NB, one Franquette and one Lara sample. In these cases, both the PC1 and PC2 value significantly differs.

Overall, the cultivar with most outliers is Chandler.

A PCA according to origin was carried out to analyze whether the samples cluster based on the respective origin (samples originated from Germany (n=4), the USA (n=3), Hungary (n=4), France (n=16), and Italy (n = 9)).

The PCA, using the lipid measurements in the positive ion mode (c), shows no clustering according to origin. The PC1 value is approx. identical for all samples. The PC2 value, accounting for 8,90%), however, spreads from 0.5 to approx. 4. Around 1.7 and 0.5 a small cluster of a total of 5 6 samples from France are visible. All other samples are scattered across the graph. Among the sample from the USA, one is an outlier. The PC1 value (PC1 = 2) is significantly different from all other samples.

Subfigure 43 (d) shows one major cluster around PC1 = 0 and PC2 = 0. Here, three France samples (PC1 = 4.8, PC2 = 0.99; PC1 = 2.5, PC2 = 2.3; PC1 = 2.3, PC2 = 1.8), one Italy sample (PC1 = 3, PC2 = 3) and one USA sample (PC1 = 1, PC2 = 3) show a different PC1 and PC2 value, resulting into outliers.

Overall, no graph shows either a clustering trend by cultivar nor by origin.

4. Discussion

Biomarkers of protein and lipid origin enable the authentication and the verification of foods and can help uncover food fraud. The development of simple, fast and reproducible methods is therefore of pivotal importance to ensure a rapid preliminary biomarker detection based on specific molecule profiles.

This work investigated the identification of biomarkers of both protein and lipid origin through infrared laser based sampling. The aim was to develop a rapid and reproducible sampling procedure able to overcome the restrictions faced with conventional extraction methods such as FA/ACN or MTBE.

With this scope, two nanosecond infrared laser based methods were developed. The subsequent mass spectrometric analysis was done via MALDI TOF. The results indicate that NIRL can be used as a fast method to facilitate the discovery of biomarkers in clinical settings.

Three of the sampling procedures used in this work can be utilized for the identification of the protein and phospholipid profiles of samples.

The FA/ACN method represents the conventional protein profiling analytical procedure that allows biotyping and bioprofiling based on intact proteins.^{63]} The results hint at a reproducibility of the method (annex figure 65), but the individual phenotype of each walnut does not enable clear clustering into cultivars nor origin (figures 15 16). A clear discrimination of the different cultivars is also not possible with the help of averaged spectra (figure 13). Here, the majority of signals differ in relative intensity but not in signal presence. These results stand in contrast with previous works. Hyun Seung Lee et al. and the dissertation in AG Schlüter by Dennis Krösser were able to proof that the FA/ACN method delivers reliable results for the identification of different yeast species and truffles.^{64]}^{4]} A reliable application was also found for different bacteria species.^{42]} A possible reason for the different results may reside in the kind of sample. In comparison to the above mentioned microorganisms and aliments, the walnut contains less proteins. These proteins are mostly categorized into allergens^{28]}^{65]} and proteins relevant for basic survival, accounting for only a total of 15% of molecules in walnuts.^{66]}

Moreover, it is of pivotal importance to consider which part of the walnut has been analyzed. In this work, all experiments were performed on the walnut kernel. The skin was previously removed. An identification of the protein profiles in the skin can be carried out to analyze whether additional proteins can be detected. However, the presence of several phytochemicals with antioxidant activity^{67]} may hinder adequate and reproducible results.

The conventional method MTBE has been used for the extraction of intact phospholipids. As described in the introduction (section 1.2.4), phospholipid content in different species can vary based on the origin due to different climatic conditions.^{24]} In this work, the walnuts either came from the middle European regions (Germany and France), Eastern Europe (Hungary), USA or the Mediterranean region (Italy). In a first step, the reference specific signals were fragmented to allow a classification of the molecules in the different phospholipid classes.

For the positive ion reflector mode, the signals at 758 m/z and approx. 784 m/z could both be attributed to phospholipids of the PC species. Hereby, PCs with acyl residues of different length could be identified. The signal triplet around approx. 1,500 m/z could either be a Cardiolipin or a PC molecule attached to Lyso PC. Since the fragmented spectrum (see figures 60 62 annex), showed the presence of a signal at 758 m/z for all three signals, it is assumed that a PC molecule attached to Lyso PC is more likely. For verification, further analyses are needed. The signals with lower relative intensity could not be attributed to any phospholipid. This can be explained by the limitations set by the MALDI TOF fragmentation procedure. In order to identify and subsequently fragment a signal of interest, a signal should have at least a S/N ratio of above 10 15. Signals below mentioned threshold cannot be recognized and therefore not be analyzed. In MALDI TOF, signals are not directly related to the concentration, instead to the ionization capability.^{68]} With 9 AA, certain lipids are more likely to ionize than others. Therefore, suppressive signal behavior may lead to the previously described phenomenon. However, under consideration of the lipids visibility in the lipid standard analyzed in the positive ion reflector mode (see figure 20), all other signals can be attributed to either PC or SM lipids.

In the negative ion reflector mode, two signals with high relative intensity were identified. In both cases (see figure 63 in the annex), the signals could be attributed to PI phospholipids.

All other signals could not be attributed to any phospholipid class due to the low S/N ratios. For exact identification of all lipids further analyses are required.

In regards to phospholipid profiles, the results do not allow discrimination of either cultivars nor origin as all spectra similar signals. It is assumed that the different walnuts cannot be distinguished through MALDI based on the phospholipid profiles. As opposed to that, quantitative methods, such liquid chromatography coupled to ESI or UHPLC coupled to orbitrap, have allowed the differentiation of different walnuts^{17]} and hazelnuts cultivars.^{39]} In literature, it has been shown that TAGs and fatty acids^{17]} vary depending on cultivar. The possibility of biomarker discovery among these two lipid classes using MALDI is therefore feasible. This experimental setup would allow the applicability of MALDI TOF for the discovery of biomarkers specifically for walnuts.

It has to be noted that regardless of reproducibility of the method (annex figures 66 67), crystallization was not always ideal. The crystallization pattern often resembled the crystals shown in figure 19 (a). Named issue lead to difficulties in measurement in particular of the negative reflector ion mode. The spectra often showed insufficient S/N ratios and strong background signals. This emphasized the imperative importance of crystallization for MALDI measurements due to impediment of proper ionization of the lipids. Moreover, analysis of figures 25 and 26 shows that the signal triplets at 1,500 m/z is often not visible. It was assumed that the divergences in comparison to the results obtained during method development are related to the change in matrix/analyte ratio. Since the reason for the different sample behavior could not be identified, the spectra were recorded with a 1:3 ratio.

Major drawbacks of the conventional methods are matrix background noises caused by matrices (in particular valid for 9 AA) and time intensive and laborious extraction procedures as well as use of toxic solvents for the extraction. Additionally, the usage of two phased procedures such as MTBE, may lead to the loss of polar phospholipids at the phase separation and hereby falsify the results.^{53]}

As opposed to the herein used conventional methods, NIRL allowed the development of a rapid sampling procedure which does not utilize any kind of solvent for the extraction of in

particular phospholipids. Additionally, a significant reduction of volume from around 5,000 nL for the MTBE protocol to 158 nL for NIRL has been achieved. (The respective volumes were calculated in other projects in the proteomics group AG Schlüter).

A direct application of the ablated sample on the target through 9 AA enabled the immediate identification of phospholipid profiles in walnuts. Other works have shown that a direct application of the sample with 9 AA is also possible with other types of sampling procedures not using NIRL.^{45]} These results all hint at the capability of lipids to ionize without need of desalting or purification as is often needed for proteins.^{69]} Time is therefore further reduced, allowing the applicability of this method in routine laboratories.

The phospholipid profiles that were obtained using NIRL are similar to the spectra recorded with the conventional method. The highest divergence is visible in the area between 900 m/z to 1,100 m/z (for the positive ion mode). The differences are assumed to be related to usage of ACN/H₂O as sample solvent in the conventional method. As shown in section 3.4.1, sample solvents based on chloroform such as Chloroform/MeOH mixtures (figure 35 (b)) may be more adequate. In literature, the usage of isopropanol/MeOH/Chloroform (4:2:1 (v/v/v)) enabled the recording of spectra with little background noise.^{70]} The phospholipids in the range between 900 m/z to 1,100 m/z were not fragmented, therefore no classification into lipid classes was possible.

Regardless of the simplified analysis of phospholipids, a major drawback of the NIRL method is the incapability of determining the exact matrix/analyte ratio. Even if the ablation (158 nL) and the matrix volume (0.5 µL) are known, the establishment of the matrix/analyte ratio is not possible due to the lack of complete resuspension of the sample in the 9 AA matrix. This could lead to problems in implementations of the NIRL sampling procedure on other tissues which are then analyzed via MALDI TOF. It is therefore advised to develop a method for the establishment of the phospholipid concentration in samples.

Overall, the NIRL sampling method with subsequent MALDI TOF analysis showed reproducibility both during the method development as well as the analysis of the authentic samples (annex figure 70). To date, sampling procedures based on infrared lasers were successfully applied for the analysis of proteins coupled to ESI.^{71]} Therefore, this work offers a first step towards the implementation of infrared based sampling and analysis via MALDI TOF in order to discover and identify biomarkers of lipid origin.

The method developed for the analysis of proteins does not allow the generation of reproducible spectra (annex figure 68 and 69). Since the direct application with HCCA is not possible as seen for 9 AA, a purification step using the C8 reversed phase chromatography was added to the workflow. The different sets of the authentic samples did not always show the same spectra. A possible reason for the lack of reproducibility might be the reaching of the binding capacity or secondary interactions between the chromatographic material and the proteins. An improvement of the C8 chromatographic protocol is therefore advised. In this work, three cycles for protein binding were used. The cycle number should be increased to hinder protein loss as much as possible. Overall however, loss of proteins cannot be avoided.

Similarly to lipids, no direct comparison with literature is possible since no infrared laser based sampling procedures have previously been developed for the identification of biomarkers with MALDI TOF. However, NIRL sampling was coupled to ESI. The developed method showed reproducible results for the identification of protein species.^{59]}

The PCA analyses do not allow differentiation into species since no clear clustering is visible. This is not applicable only to the different varieties, but also to the sample origin.

In literature, it was shown that walnuts can be differentiated based on their origin by analyzing the phospholipid content.^{17]} However, similarly to the averaged spectra, the PCA analyses also do not hint at a differentiation of phospholipids according to origin.

Similar results were obtained for PCA analyses of proteins. As opposed to the phospholipids, however, a slight trend is visible (figure 15 and 34) regardless of whether NIRL or the conventional method FA/ACN was used.

Analysis of the averaged spectra hints at similarities within all samples. It was therefore assumed that the PCA analysis would deliver graphs with less scattering samples. The reason for such behavior can be explained by various factors that play a role during MALDI spectra acquisition. On one hand, differences within crystallization of each single sample may lead to the recording of differing spectra. In this case, ion suppression may influence the visibility and detection of certain ions and therefore falsify the results. If spectra with such characteristics (e.g. crystallization problems for the MTBE based method) are recorded, PCA

analyses will not be able to cluster adequately, since the prime role of PCAs consists in identifying differences between samples.

In the case of the direct NIRL application through mixing with 9 AA, a further problem could consist in different molecule concentrations in the spotted matrix/analyte mix. As previously described, a major setback of the direct NIRL method is the incapability of determining the analyte concentration. Since no complete dissolution of the sample in 9 AA is present, it is possible for certain molecules to appear in different concentrations in each set. If the concentration of a molecule is too low and other e.g. phospholipids in the sample tend to ionize with more ease with 9 AA (see PI in figure 27), suppression mechanisms are likely. This would lead not only to falsification of PCA results but also to lack of reproducibility of a method.

Lastly, the presence of strong background noise could also influence PCA results. In this work, a threshold of 1,000 and 2,000 was chosen. However, it is thought that minor matrix signals (valid for 9 AA) may have influenced the calculation of the PC values. Given problem may be circumvented by usage of direct application of the sample on the target after NIRL ablation (see figures 39 and 40), matrix change^{72]} or improvement^{73]} and shift of analyzed range.

Overall, the analysis of intact proteins and phospholipids via MALDI TOF generates a fingerprint for a sample. The discovery of potential biomarkers can there therefore be facilitated. However, additional analyses are necessary such as deep characterization via ESI^{4]} in order to elucidate structure, amino acid sequence and gain additional information. An in depth analysis of the proteins was beyond the scope of this work and not necessary, since no biomarker for walnuts could be identified.

5. Summary and future outlook

In this work, it was proven that NIRL can be used as a sample preparation method in order to visualize both proteins and lipids. Hereby, a direct application of the condensed sample was achieved for lipids. For protein analysis, on the other hand, a desalting procedure was needed prior to protein visibility in the spectra. Even if in this work no biomarkers were identified, the methods find applicability within this field. It is important to further validate the NIRL method for lipids and proteins by augmenting the sample size and by improving the C8 chromatography procedure. For both lipids and proteins, studies with at least $n=4$ or higher should be performed to allow better statistical security.

A final validation of the methods can be achieved by performance of a blind test. If discrimination of the cultivars had been successful on proteomic or lipidomic level, samples of unknown origins or cultivar should be prepared according to the developed method. Comparison with averaged spectra as well as statistical tools would then allow the sorting of the analyzed sample.

A quantitative approach as described by Yan et al. (usage of UHPLC Orbitrap methods) could be used for the analysis of phospholipids in walnuts to verify whether the concentration of the respective phospholipids varies in the four analyzed cultivars.

Moreover, different lipid groups such as TAGs or free fatty acids could be analyzed. TAGs analysis can be performed with a DHB matrix. Free fatty can be analyzed with 9 AA or the MTPP matrix. Since free fatty acids are small molecules in the range between 200 m/z to 300 m/z , the background noise in the respective range of 9 AA has to be reduced. This can be achieved by mixing of 9 AA with carbon dots.

The method development for TAGs and free fatty acid should first be carried out through direct application as shown in this work. Indirect application can be performed by usage of adequate solvents or with the addition of an extraction or purification step coupled to MALDI TOF for bioprofiling and biomarker identification.

NIRL and MALDI TOF can allow a first identification of a biomarker candidate, however further analyses are needed in order to fully characterize the molecule. For this means, an ESI method should be developed. In the case of lipids signal fragmentation can be performed to enable a comparison.

References

1. https://www.bvl.bund.de/DE/Arbeitsbereiche/01_Lebensmittel/03_Verbraucher/16_Food_Fraud/02_Was%20ist%20Lebensmittelbetrug/Was%20ist%20Lebensmittelbetrug_node.html;jsessionid=EEFCAB1810B570EC8F049766B1AB799D.1_cid363 (access, 27th of November 2021, 10:50 am)
2. Johnson Renée. "Food Fraud and "Economically Motivated Adulteration" of Food and Food Ingredients" Congressional Research Service, 2014 01 10
3. Fischer Markus, Creydt Marina. "Food Profiling Analytical strategies for Food Authentication" *Journal of Agricultural and Food Chemistry*, 68, 49; 14321 14322, 2020 12 09
4. Krösner Dennis. "Mass Spectrometric based Proteomic Profiling of Truffles" Dissertation for the acquisition of the academic degree Dr. rer. nat. at the University of Hamburg, 2021 02
5. Amur Shashi. "Biomarker qualification program educational module series module 1 Biomarker terminology: speaking the same language" U.S. Food and Drug administration (access: 24th of November 2021, 10:00 am)
6. deSouza Nandita M. et al. "Validated imaging biomarkers as decision making tool in clinical trials and routine practice: current status and recommendations from the EIBALL* subcommittee of the European Society of Radiology (ESR)" *Insights into Imaging*, 10, 1; 87, 2019 08 29
7. Institute of Medicine (US) Committee on Metabolic Monitoring for Military Field. "Physiological Biomarkers for Predicting Performance" National Academics Press (US), 2004
8. Jahanban Esfahlan Ali et al. "A Comprehensive Review on the Chemical Constituents and Functional Uses of Walnut (*Juglans* spp.) Husk" *Int J Mol Sci*, 20, 16, 2019 08 12
9. Martínez García Pedro J. "The walnut (*Juglans regia*) genome sequence reveals diversity in genes coding for the biosynthesis of non structural polyphenols" *the plant journal*, 87, 5; 507 532, 2016 05 03
10. Pereira José Alberto et al. "Bioactive properties and chemical composition of six walnut (*Juglans regia* L.) cultivars" *Food and Chemical Toxicology*, 46, 6; 2103 2111, 2008 06 01
11. Ebrahimi Shrin et al. "Persian Walnut Composition and its Importance in Human Health" *Int J Enteric Pathog*, 6, 1; 3 9, 2017 12 03
12. Lu Shaoping. "An efficient and comprehensive plant glycerolipids analysis approach based on high performance liquid chromatography quadrupole time of flight mass spectrometer" *Plant Direct*, 3, 11; e00183, 2019 11
13. Robbins Katherine S. et al. „Update on the Healthful Lipid Constituents of Commercially Important Tree Nuts" *J. Agric. Food Chem.*, 59, 22; 12083 12092, 2011 11 23
14. Dlaviz Hamdollah et al. "A Review Study on Phytochemistry and Pharmacology Applications of *Juglans Regia* Plant" *Pharmacogn Rev*, 11, 22; 145 152, 2017
15. Colaric Mateja et al. "Phenolic acids, syringaldehyde, and juglone in fruits of different cultivars of *Juglans regia* L" *J Agric Food Chem*, 53, 16; 6390 6396, 2005 08 10
16. Amaral Joana S. et al. "Triacylglycerol composition of walnut (*Juglans regia* L.) cultivars: characterization by HPLC ELSD and chemometrics" *J Agric Food Chem*, 52, 26; 7964 7969, 2004 12 29

17. Yan Suxian et al. "Insights Into Walnut Lipid Metabolism From Metabolome and Transcriptome Analysis" *Front Genet*, 12; 715731, 2021
18. Bolling Jeffrey B. "The phytochemical composition and antioxidant actions of tree nuts" *Asia Pac J Clin Nutr*, 19, 1; 117 123, 2010
19. Christen Philipp, Jaussi Rolf, Benoit Roger. "Biochemie und Molekularbiologie Eine Einführung in 40 Lehreinheiten" Springer Spektrum, Chapter „Lipide und biologische Membranen“; 70 72, 2016
20. Fahy Eoin et al. "A comprehensive classification system for lipids" *Journal of Lipid Research*, 46, 5; 839 861, 2005 05 01
21. Zegarlińska Jolanta et al. "Phosphatidic acid a simple phospholipid with multiple faces" *Acta Biochim Pol*, 65, 2; 163 171, 2018
22. Ulmer Candice Z. et al. "Optimization of Folch, Bligh Dyer, and Matyash Sample to Extraction Solvent Ratios for Human Plasma Based Lipidomics Studies" *Anal Chim Acta*, 1037; 351 357, 2018 12 11
23. Song Shuang et al. "Characterization of phospholipid profiles in six kinds of nut using HILIC ESI IT TOF MS system" *Food Chem*, 240; 1171 1178, 2018 02 01
24. Klockmann Sven. "Food Fingerprinting: Metabolomic Approaches for Geographical Origin Discrimination of Hazelnuts (*Corylus avellana*) by UPLC QTOF MS" *J Agric Food Chem*, 64, 48; 9253 9262, 2016 12 07
25. Bolling Bradley W. et al. "Tree nut phytochemicals: composition, antioxidant capacity, bioactivity, impact factors. A systematic review of almonds, Brazils, cashews, hazelnuts, macadamias, pecans, pine nuts, pistachios and walnuts" *Nutrition Research Reviews*, 24, 2; 244 275, 2011 12
26. Marrano Annarita et al. "High quality chromosome scale assembly of the walnut (*Juglans regia* L.) reference genome" *Gigascience*, 9, 5; giaa050, 2020 5 20
27. Lee Jeongmin et al. "Component resolved diagnosis of walnut allergy in young children: Jug r 1 as a major walnut allergen" *Asian Pac J Allergy Immunol*, 39, 3; 190 196, 2021 09
28. Geiselhart Sabine, Hoffmann Sommergruber Karin, Bublin Merima. "Tree nut allergens" *Molecular Immunology*, 100; 71 81, 2018 08 01
29. Dubiela Pawel et al. "Jug r 6 is the allergenic vicilin present in walnut responsible for IgE cross reactivities to other tree nuts and seeds" *Sci Rep*, 8, 1; 11366, 2018 07 27
30. Dubiela Pawel et al. "Impact of lipid binding on the tertiary structure and allergenic potential of Jug r 3, the non specific lipid transfer protein from walnut" *Sci Rep*, 9, 1; 2007, 2019 02 14
31. Blankestijn M. A. "A subset of walnut allergic adults is sensitized to walnut 11S globulin Jug r 4" *Food Chem*, 248; 1206 1213, 2018 09
32. <http://www.allergen.org/viewallergen.php?aid=39>, (access: 29th of November 2021, 10:30 am)
33. <http://allergen.org/viewallergen.php?aid=876>, (access: 29th of November 2021, 11:00 am)
34. <http://allergen.org/viewallergen.php?aid=895>, (access: 29th of November 2021, 11:05 am)
35. <http://www.allergen.org/viewallergen.php?aid=912>, (access: 29th of November 2021, 11:10 am)
36. <http://www.allergen.org/viewallergen.php?aid=924>, (access: 29th of November 2021, 11:30 am)

37. Fanelli Valentina et al. "Molecular Approaches to Agri Food Traceability and Authentication: An Updated Review" *Foods*, 10, 7; 1644, 2021 07
38. Klockmann Sven et al. "Food Targeting: Geographical Origin Determination of Hazelnuts (*Corylus avellana*) by LC QqQ MS/MS Based Targeted Metabolomics Application" *J Agric Food Chem*, 65, 7; 1456 1465, 2017 02 22
39. Esteki Mahnaz et al. "Classification and authentication of Iranian walnuts according to their geographical origin based on gas chromatographic fatty acid fingerprint analysis using pattern recognition methods" *Chemometrics and Intelligent Laboratory Systems*, 171, 2017 11 01
40. Wang Jun et al. "Rapid determination of the geographical origin of honey based on protein fingerprinting and barcoding using MALDI TOF MS" *J Agric Food Chem*, 57, 21; 10081 10088, 2009 11 11
41. Stahl Antje, Schröder Uwe. "Development of a MALDI TOF MS Based Protein Fingerprint Database of Common Food Fish Allowing Fast and Reliable Identification of Fraud and Substitution" *J Agric Food Chem*, 65, 34; 7519 7527, 2017 08 30x
42. Angelini Roberto et al. "Lipidomics of intact mitochondria by MALDI TOF/MS" *Journal of Lipid Research*, 53, 7; 1417 1425, 2012 06 01
43. El Karkouri Khalid et al. „Rapid MALDI TOF MS identification of commercial truffles" *Sci Rep*, 9,1; 17686, 2019 11 27
44. Singhal Neelja et al. "MALDI TOF mass spectrometry: an emerging technology for microbial identification and diagnosis" *Frontiers in Microbiology*, 6; 791, 2015
45. Angelini Roberto et al. "MALDI TOF/MS analysis of archaeobacterial lipids in lyophilized membranes dry mixed with 9 aminoacridine" *Journal of Lipid Research*, 51, 9; 2818 2825, 2010 09 01
46. Ciarmiello Loredana F. et al. "Analysis of Different European Hazelnut (*Corylus avellana* L.) Cultivars: Authentication, Phenotypic Features, and Phenolic Profiles" *J. Agric. Food Chem.*, 62, 26; 6236 6246, 2014 07 02
47. Fuchs Beate. "An update of MALDI TOF mass spectrometry in lipid research" *Progress in Lipid Research*, 49, 4; 450 475, 2010 10 01
48. Yoon So Ra. "A novel method to rapidly distinguish the geographical origin of traditional fermented salted vegetables by mass fingerprinting" *PLoS One*, 12, 11; e0188217, 2017 11 17
49. <https://www.sigmaaldrich.com/DE/de/product/sial/39468>
(access: 29th of November 2021, 6:00 am)
50. Li Ying L., Gross Michael L., Hsu Fong Fu. "Ionic Liquid Matrices for Improved Analysis of Phospholipids by MALDI TOF Mass Spectrometry" *Journal of the American Society for Mass Spectrometry*, 16, 5; 679 682, 2005 05 01
51. <https://www.sigmaaldrich.com/DE/de/product/sial/92817>
(access: 28th of November 2021, 10:00 am)
52. <https://www.sigmaaldrich.com/DE/de/product/sigma/85707>
(access: 30th of November 2021, 7:30 am)
53. Wong Matthew Wai Kin et al. "v Comparison of Single Phase and Biphasic Extraction Protocols for Lipidomic Studies Using Human Plasma" *Frontiers in Neurology* 10; 879, 2019
54. Matyash Vitali et al. "Lipid extraction by methyl tert butyl ether for high throughput lipidomics" *J Lipid Res*, 49, 5; 1137 1146, 2008 05
55. Mazzeo Maria Fiorella et al. "Fish Authentication by MALDI TOF Mass Spectrometry" *J. Agric. Food Chem.*, 56, 23; 11071 11076, 2008 12 10

56. Agrawal Ganesh Kumar et al. „Biomarker discovery and applications for foods and beverages: proteomics to nanoproteomics” *J Proteomics*, 93; 74 92, 2013 11 20
57. Aristoteli Lina P., Molloy Mark P., Baker Mark S. “Evaluation of Endogenous Plasma Peptide Extraction Methods for Mass Spectrometric Biomarker Discovery” *J. Proteome Res.*, 6, 2; 571 581, 2007 02 01
58. Murugaiyan Jayaseelan, Roesler Uwe. “MALDI TOF MS Profiling Advances in Species Identification of Pests, Parasites, and Vectors” *Front Cell Infect Microbiol*, 7; 184, 2017 5 15
59. Hahn Jan et al. “Tissue Sampling and Homogenization in the Sub Microliter Scale with a Nanosecond Infrared Laser (NIRL) for Mass Spectrometric Proteomics” *International Journal of Molecular Sciences*, 22, 19; 10833, 2021 01
60. Kwiatkowski Marcus et al. “Homogenization of tissues via picosecond infrared laser (PIRL) ablation: Giving a closer view on the in vivo composition of protein species as compared to mechanical homogenization” *J Proteomics*, 134; 193 202, 2016 2 16
61. Yushkevich Paul A. et al. „User guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability” *Neuroimage*, 31, 3; 1116 1128, 2006 06 01
62. ICH. “ICH Harmonised tripartite guideline – Validation of analytical procedures: text and methodology Q2(R1)” Current Step 4 version, 2005
63. Kelleher Neil L. “Top Down Proteomics” *American chemical society*, 197 203, 2004 06 01
64. Lee Hyun Seung et al. “Comparison of the Bruker Biotyper and VITEK MS Matrix Assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry Systems Using a Formic Acid Extraction Method to Identify Common and Uncommon Yeast Isolates” *Ann Lab Med*, 37, 3; 223 230, 2017 05
65. Downs Melanie L. “Mass spectrometric analysis of allergens in roasted walnuts” *Journal of Proteomics*, 142; 62 69, 2016 06 16
66. Ziarati Parisa, Aryapak Sara. “Nutritive Value of Persian Walnut (*Juglans regia* L.) Orchards” *American Eurasian Journal of agricultural and environmental sciences*, 14; 1228 1235, 2014 11 18
67. Panth Nisha, Paudel Keshav Raj, Karki Rajendra. “Phytochemical profile and biological activity of *Juglans regia*” *Journal of Integrative Medicine*, 14, 5; 359 373, 2016 09 01
68. Szaéjli Emiélia, Feheér Tamaés, Medzihradsky Katalin F. “Investigating the Quantitative Nature of MALDI TOF MS” *Molecular & Cellular Proteomics*, 7, 12; 2410 2418, 2008 12 01
69. Mlynarcik Patrik et al. “Development of simple and rapid elution methods for proteins from various affinity beads for their direct MALDI TOF downstream application” *Journal of Proteomics*, 75, 14; 4529 4535, 2012 06 19
70. Cheng Hua et al. “Selective desorption/ionization of sulfatides by MALDI MS facilitated using 9 aminoacridine as matrix” *J Lipid Res*, 51, 6; 1599 1609, 2010 6
71. Wurlitzer Marcus et al. “Mass Spectrometric Lipid Profiles of Picosecond Infrared Laser Generated Tissue Aerosols Discriminate Different Brain Tissues” *Lasers in Surgery and Medicine*, 52, 3; 228 234, 2020 03
72. Ayorinde F. O. “Use of meso tetrakis(pentafluorophenyl)porphyrin as a matrix for low molecular weight alkylphenol ethoxylates in laser desorption/ ionization time of flight mass spectrometry” *Rapid Communications in Mass Spectrometry*, 13, 24; 2474 2479, 1999

73. Yongli Chen et al. "Carbon Dots and 9AA as a Binary Matrix for the Detection of Small Molecules by Matrix Assisted Laser Desorption/Ionization Mass Spectrometry" *J Am Soc Mass Spectrom*, 27; 1227-1235, 2016 04 13

Declaration of Originality

I hereby declare that the following dissertation titled “Investigation of infrared laser based sampling for the detection of biomarkers with MALDI mass spectrometry” submitted at the university of applied sciences in Hamburg Bergedorf the 07.12.2021, has been carried out by myself at the department of Mass Spectrometric Proteomics at the Institute of Clinical Chemistry and Laboratory Medicine UKE Hamburg.

All literature has been mentioned and acknowledged duly in the respective sections in the text. The appropriate list of references is provided starting from page 94 to 97.

Elisabetta De Iaco,

Hamburg, the 6th of December 2021

6. Annex

DHB Crystallization behavior studies:

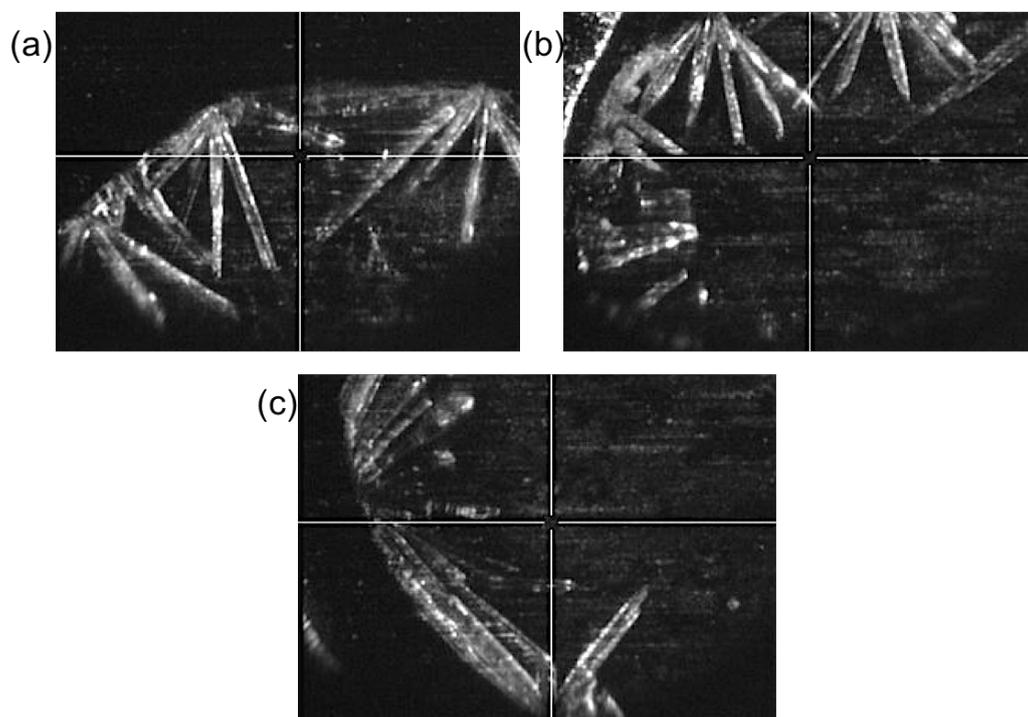


Figure 44: Crystallization behavior of the matrix prepared on 20 mg/mL DHB in 100% TA30 and with different solvent. (a) DHB, (b) DHB with MeOH/H₂O 50:50 (v/v) solvent and (c) DHB with ACN/H₂O 50:50 (v/v) solvent. Pictures taken with incorporated camera in MALDI-TOF equipment after spotted and air-dried for 5 min.

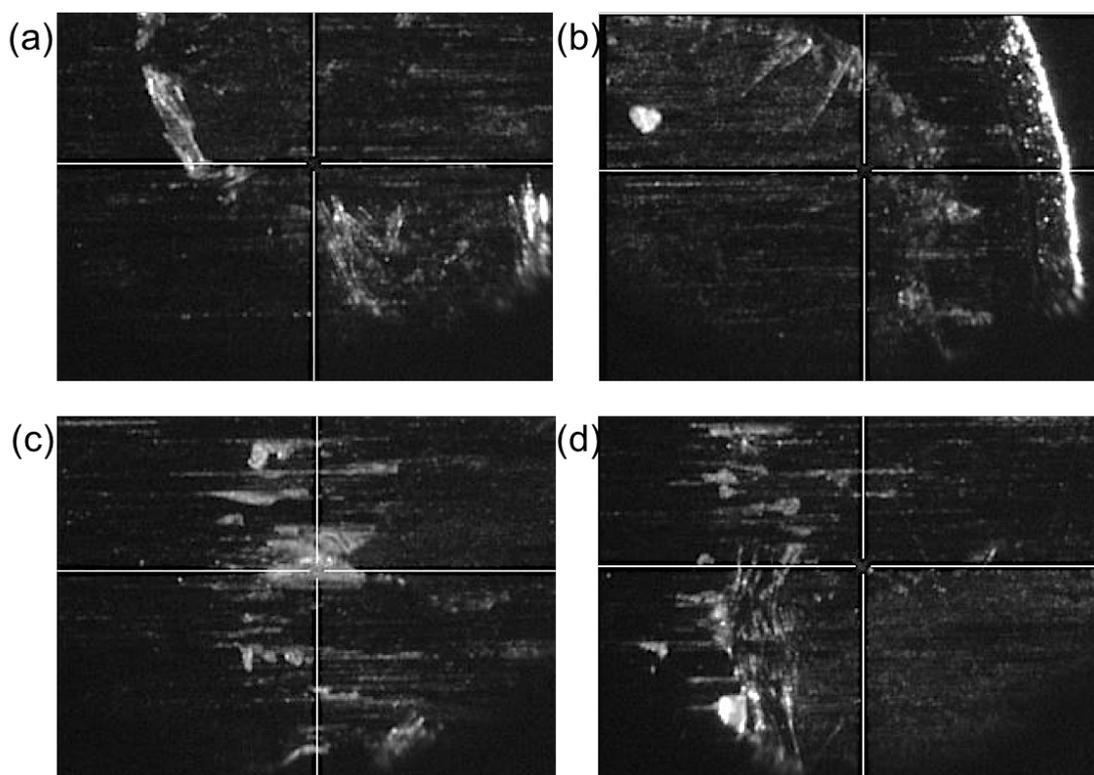


Figure 45: Crystallization behavior of the matrix prepared on 10 mg/mL DHB in ACN/H₂O 50:50 (v/v) and with different solvent. (a) DHB, (b) DHB with 100% MeOH solvent, (c) DHB with MeOH/H₂O 50:50 (v/v) solvent and (d) DHB with ACN/H₂O 50:50 (v/v) solvent. Pictures taken with incorporated camera in MALDI-TOF equipment after spotted and air-dried for 5 min.

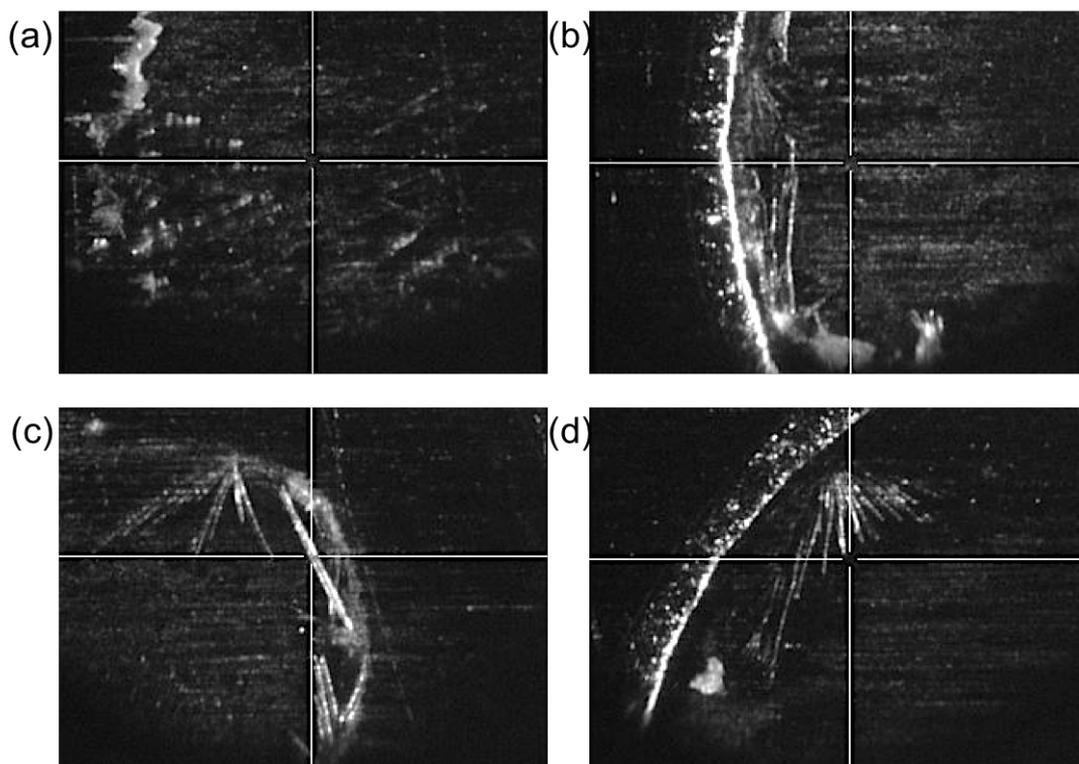


Figure 46: Crystallization behavior of the matrix preparation on 10 mg/mL DHB in ACN/H₂O with 0.1% TFA and with different sample solvents. (a) DHB, (b) DHB with 100% MeOH sample solvent, (c) DHB with MeOH/H₂O 50:5 (v/v) sample solvent and (d) DHB with ACN/H₂O 50:50 (v/v) sample solvent. Pictures taken with incorporated camera in MALDI-TOF equipment after spotted and a re-dried for 5 min.

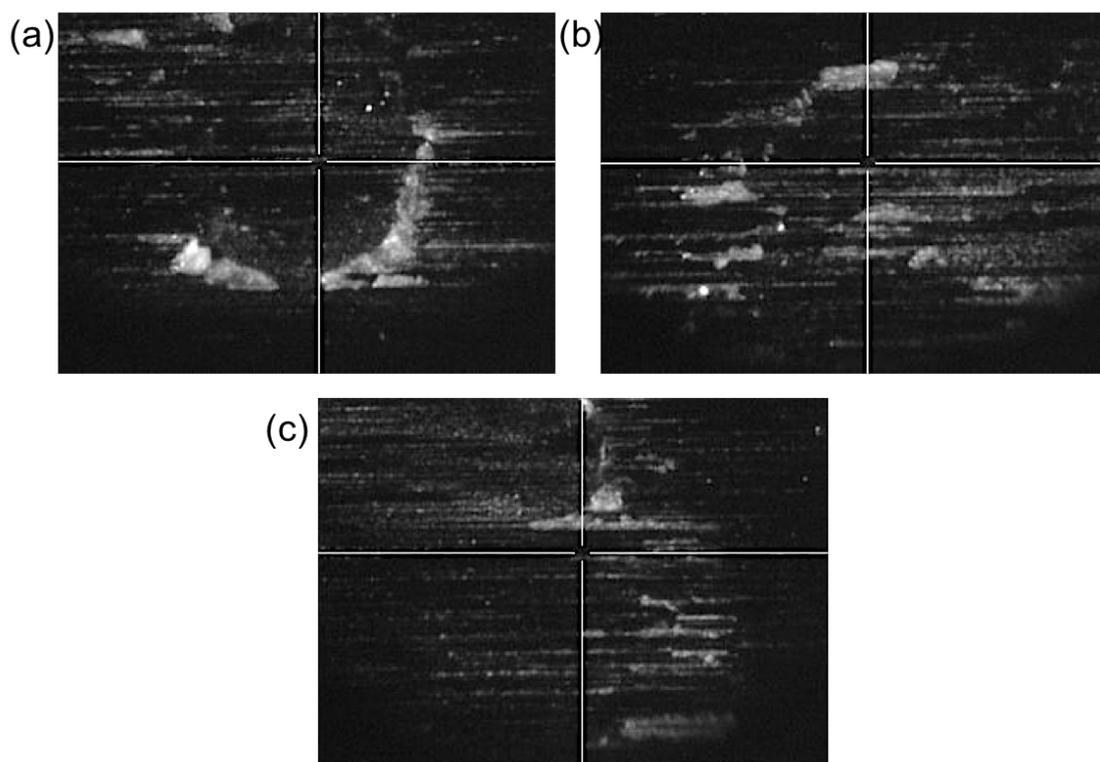


Figure 47: Crystallization behavior of the matrix preparation on 30 mM DHB in MeOH/H₂O (90:10, v/v) and with different sample solvents. (a) DHB, (b) DHB with MeOH/H₂O 50:5 (v/v) sample solvent and (c) DHB with ACN/H₂O 50:50 (v/v) sample solvent. Pictures taken with incorporated camera in MALDI-TOF equipment after spotted and a re-dried for 5 min.

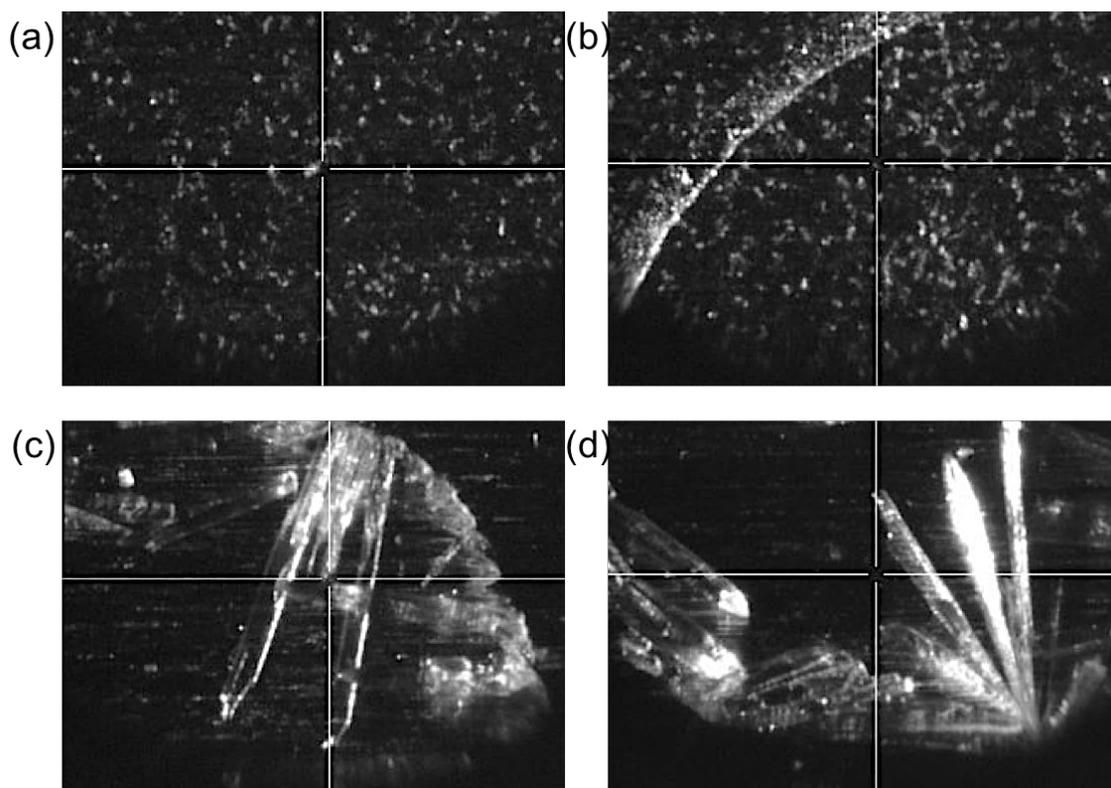


Figure 48: Crystallization behavior of the matrix preparation on 0.5 M DHB in 100% MeOH + 0.1% TFA and with different sample solvents. (a) DHB, (b) DHB with 100% MeOH sample solvent, (c) DHB with MeOH/H₂O 50:5 (v/v) sample solvent and (d) DHB with ACN/H₂O 50:50 (v/v) sample solvent. Pictures taken with incorporated camera in MALDI-TOF equipment after spotted and air-dried for 5 min.

DHB spectrum:

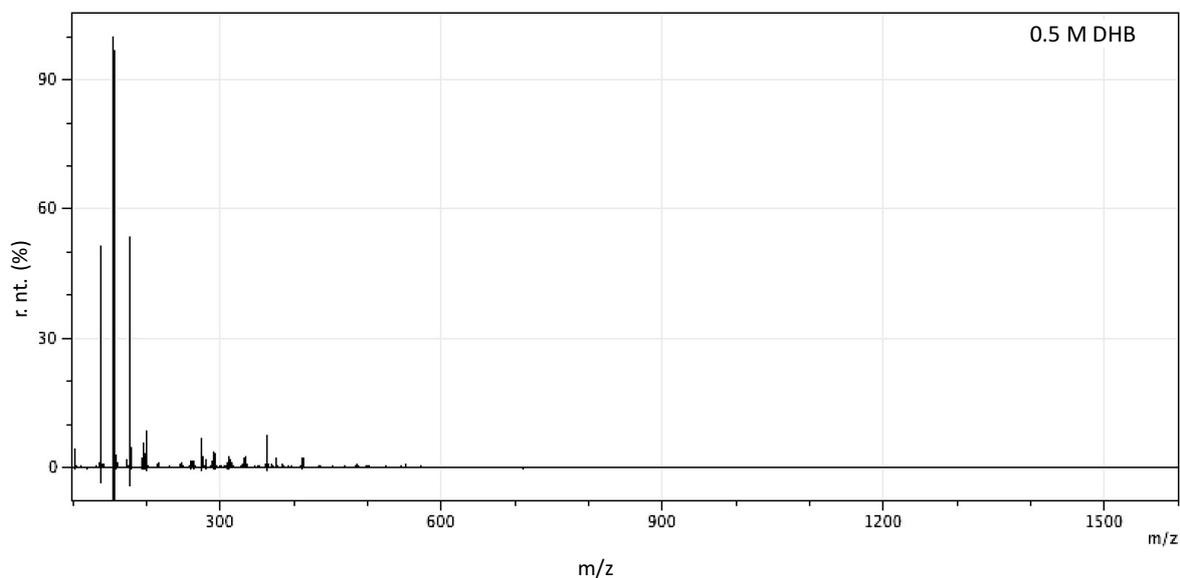


Figure 49: Exemplary 0.5 M DHB in 100% MeOH spectrum in the range from 100 m/z to 1,600 m/z. Recorded with a laser intensity of 30%. Contains no sample. The abscissa shows the m/z values in a range from 100 to 1,600. The ordinate represents the normalized relative intensity in %. Processing of the spectra was done with mMass software.

Fragmentation spectra for the lipid standard:

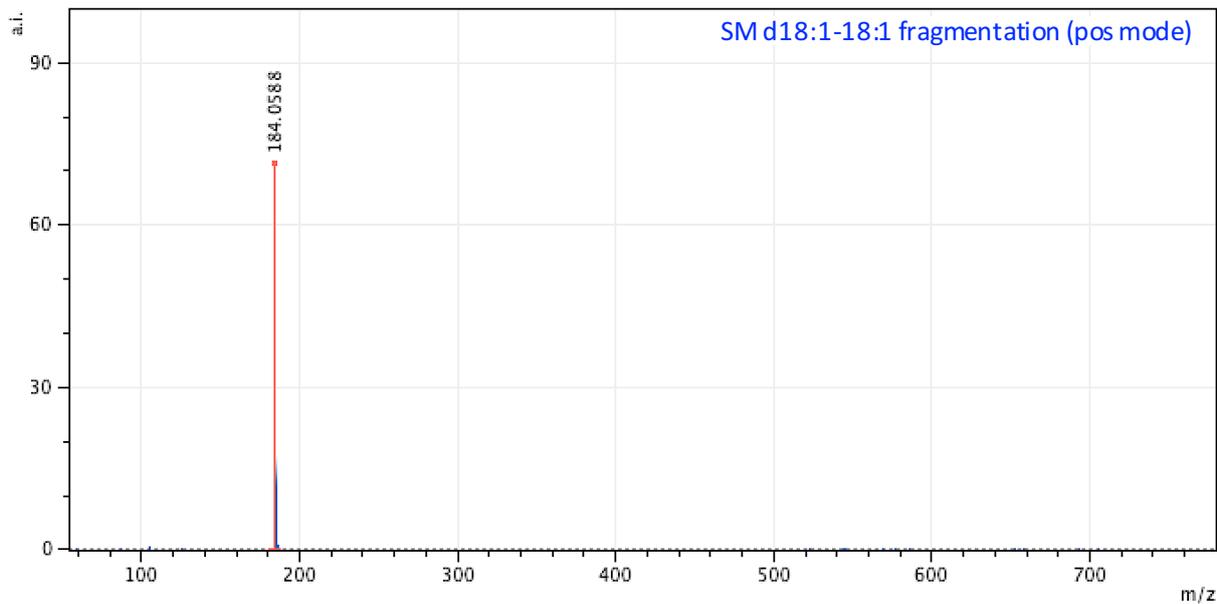


Figure 50: Fragmentation spectrum for the SM phospholipid present in the lipid standard. The abscissa shows the m/z values in a range from 80 to 780. The ordinate represents the normalized intensity in a.u. The spectra were recorded with a laser energy of 80%. Processing of the spectra was done with mMass software.

Table 26: m/z values and the respective attributed lipid fragment for the phospholipid SM. HG = head group

m/z value	attributed lipid fragment
184.0588	HG

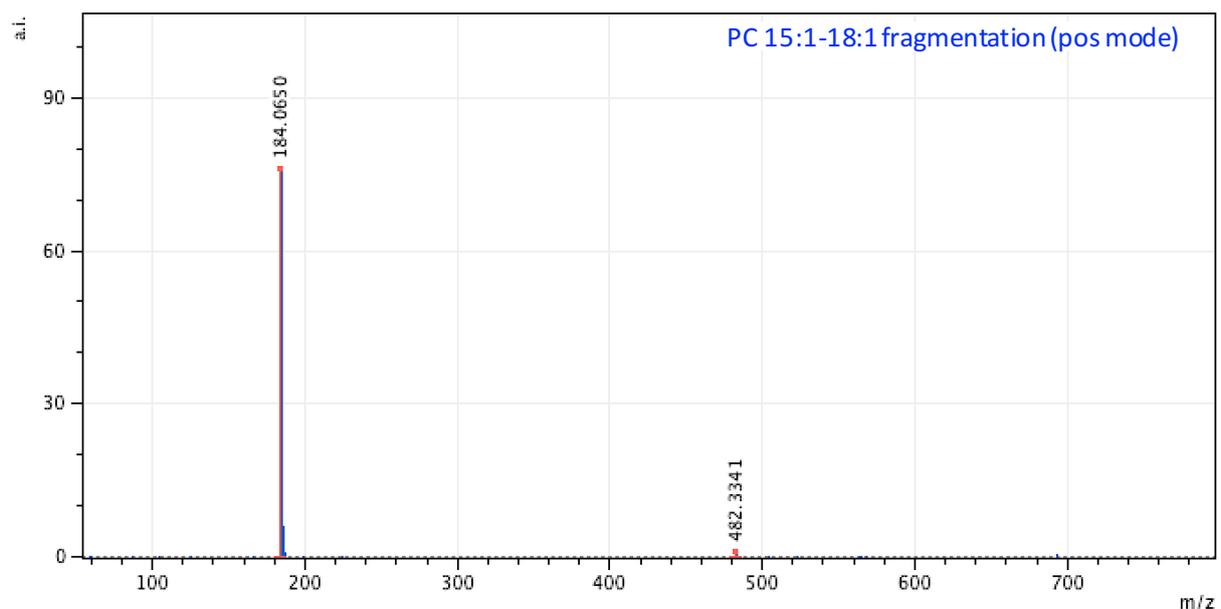


Figure 51: Fragmentation spectrum for the PC phospholipid present in the lipid standard. The abscissa shows the m/z values in a range from 80 to 780. The ordinate represents the normalized intensity in a.u. The spectra were recorded with a laser energy of 80%. Processing of the spectra was done with mMass software.

Table 27: m/z values and the respective attributed lipid fragment for the phospholipid PC. HG = head group, FA2 = 18:1 fatty acid

m/z value	attributed lipid fragment
184.0650	HG
482.3341	M FA2+H ⁺

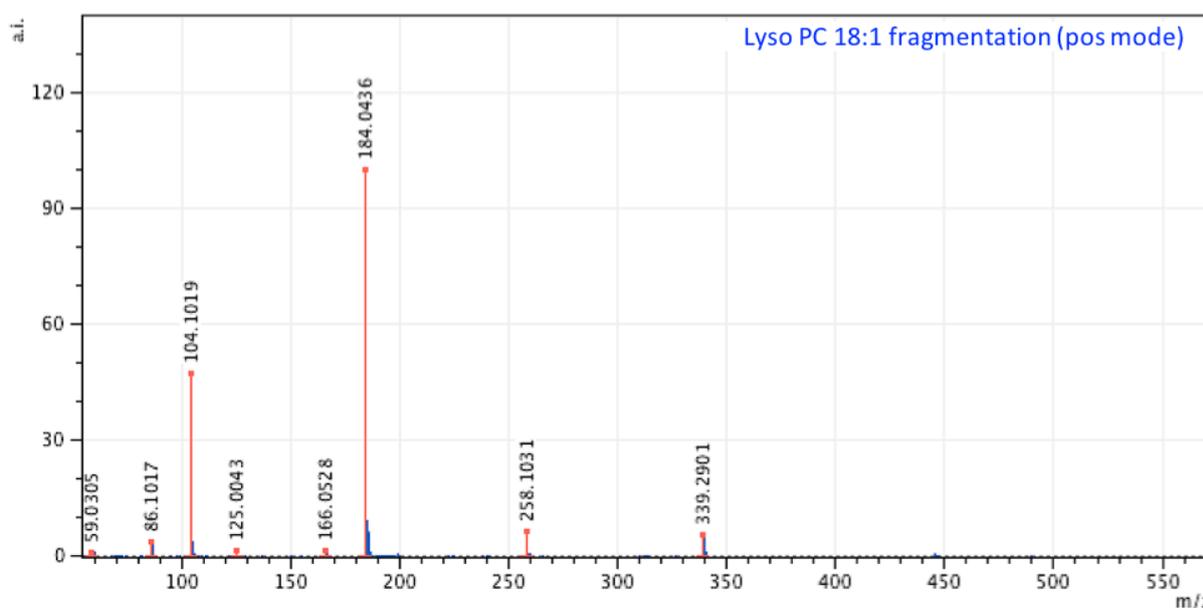


Figure 52: Fragmentation spectrum for the Lyso PC phospholipid present in the lipid standard. The abscissa shows the m/z values in a range from 0 to 570. The ordinate represents the normalized intensity in a.u. The spectra were recorded with a laser energy of 80%. Processing of the spectra was done with mMass software.

Table 28: m/z values and the respective attributed lipid fragment for the phospholipid Lyso PC. HG = head group, FA = fatty acid.

m/z value	attributed lipid fragment
59.0305	NMe ₃
86.1017	Choline H ⁺
104.1019	Choline+OH (P cleavage)
125.0043	P+Et (HG NMe ₃)
166.0528	not assigned
184.0436	HG + 2H ⁺
258.1031	M FA+H ⁺ (ester bond cleavage)
339.2901	M HG

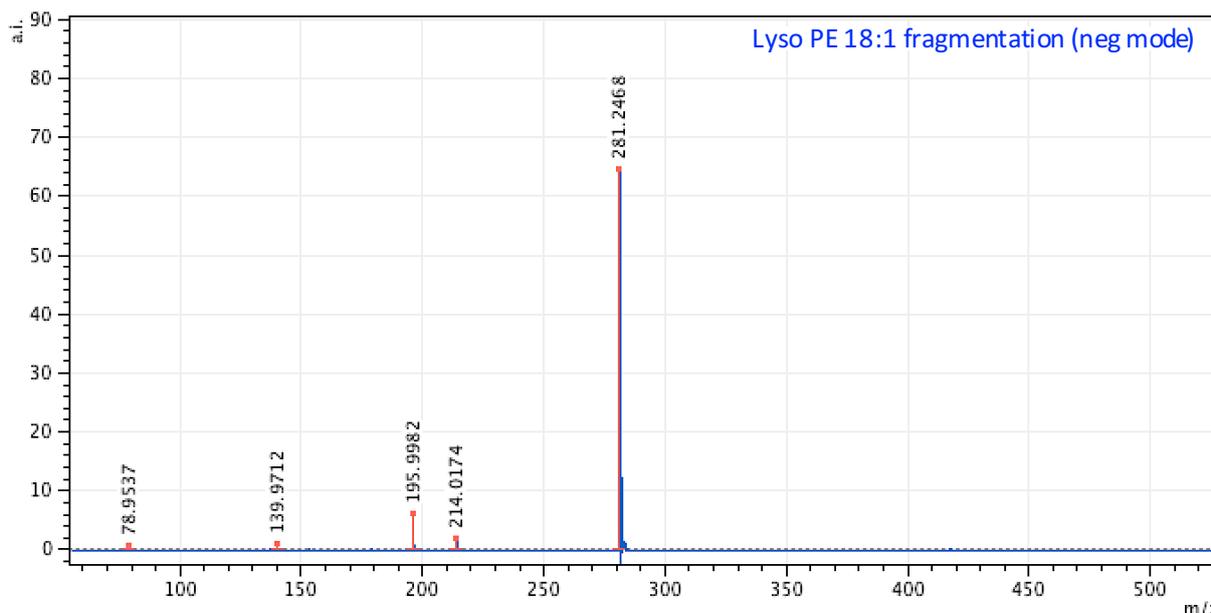


Figure 53: Fragmentation spectrum for the Lyso PE phospholipid present in the lipid standard. The abscissa shows the m/z values in a range from 0 to 560. The ordinate represents the normalized intensity in a.u. The spectra were recorded with a laser energy of 80%. Processing of the spectra was done with mMass software.

Table 29: m/z values and the respective attributed lipid fragment for the phospholipid Lyso PE. HG = head group, FA = fatty acid.

m/z value	attributed lipid fragment
78.9391	P
139.9712	HG
195.9982	M FA H ⁺
214.0174	not assigned
281.2468	FA

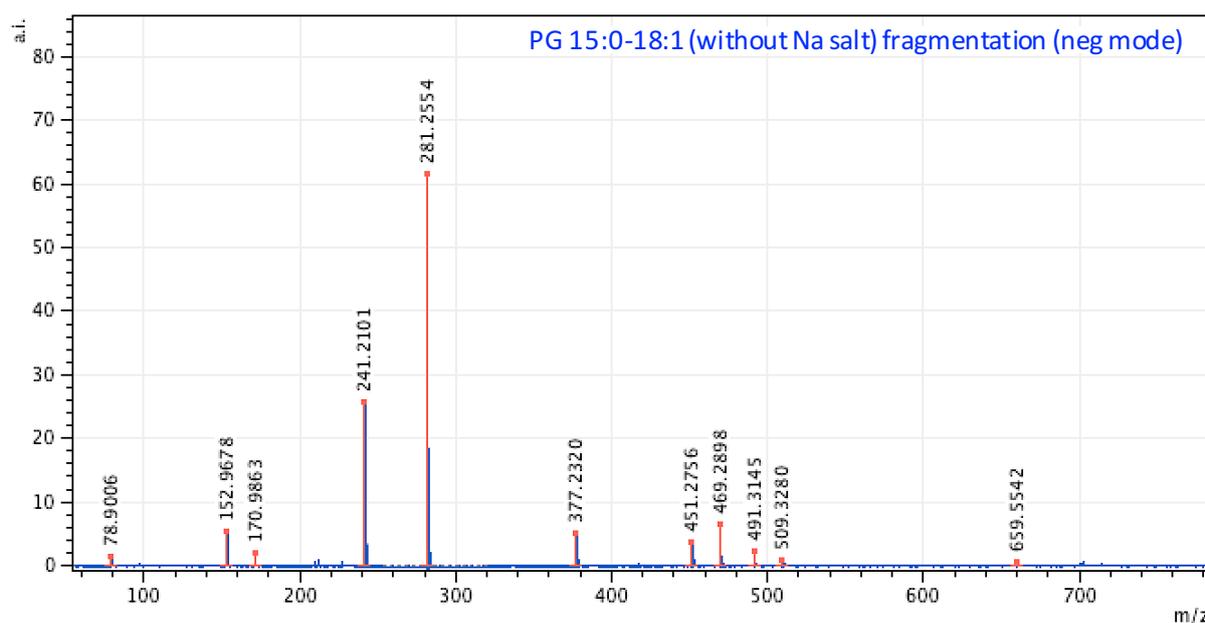


Figure 54: Fragmentation spectrum for the PG phospholipid present in the lipid standard. The abscissa shows the m/z values in a range from 80 to 780. The ordinate represents the normalized intensity in a.u. The spectra were recorded with a laser energy of 80%. Processing of the spectra was done with mMass software.

Table 30: m/z values and the respective attributed lipid fragment for the phospholipid PG. P_i = phosphate HG = head group, FA1 = 15:0 fatty acid, FA2 = 18:1 fatty acid

m/z value	attributed lipid fragment
78.9391	P
152.9672	FA2 olefin
170.9863	HG
241.2101	FA2
281.2554	FA1
377.2320	M
451.2756	M FA2 2H ⁺
469.2898	Not assigned
491.3145	M FA2 2H ⁺
509.3280	M FA1 H ⁺
659.5542	M Glycerol

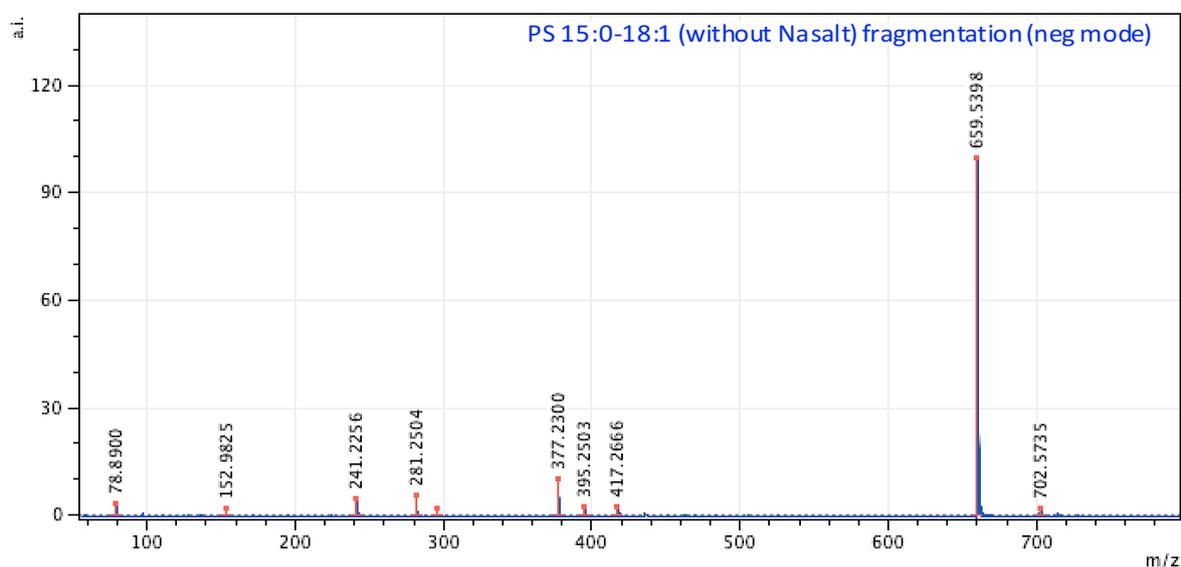


Figure 55: Fragmentation spectrum for the PS phospholipid present in the lipid standard. The abscissa shows the m/z values in a range from 80 to 780. The ordinate represents the normalized relative intensity in a.u. The spectra were recorded with a laser energy of 80%. Processing of the spectra was done with mMass software.

Table 31: m/z values and the respective attributed lipid fragment for the phospholipid PG. P_i = phosphate HG = head group, FA1 = 15:0 fatty acid, FA2 = 18:1 fatty acid

m/z value	attributed lipid fragment
78.8900	P
152.9825	FA2 olefin
241.2256	FA1
281.2504	FA2
377.2300	not assigned
395.2503	M Serine FA2
417.2666	M Serine FA1
659.5398	M Serine+H ⁺
702.5735	M CO ₂

Sample-specific fragmentation spectra:

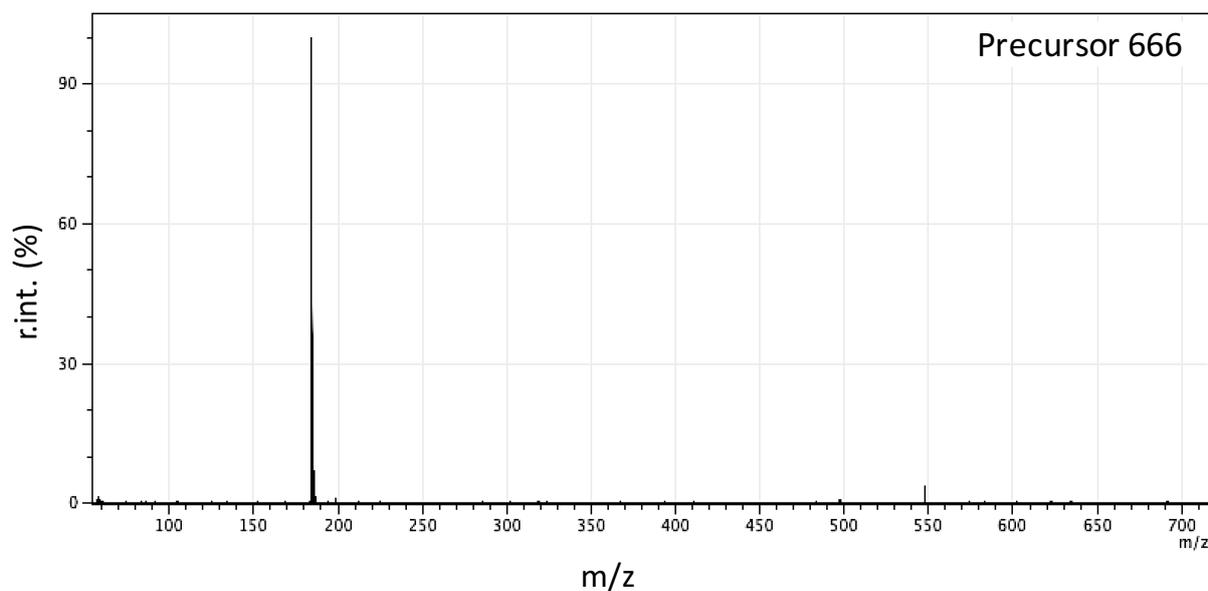


Figure 56: Fragmentation spectrum of the sample-specific precursor on 666 (positive ion reflector mode). The abscissa shows the m/z values in a range from 0 to 720. The ordinate represents the normalized relative intensity in %. The spectra were recorded with a laser energy of 80%. Processing of the spectra was done with mMass software.

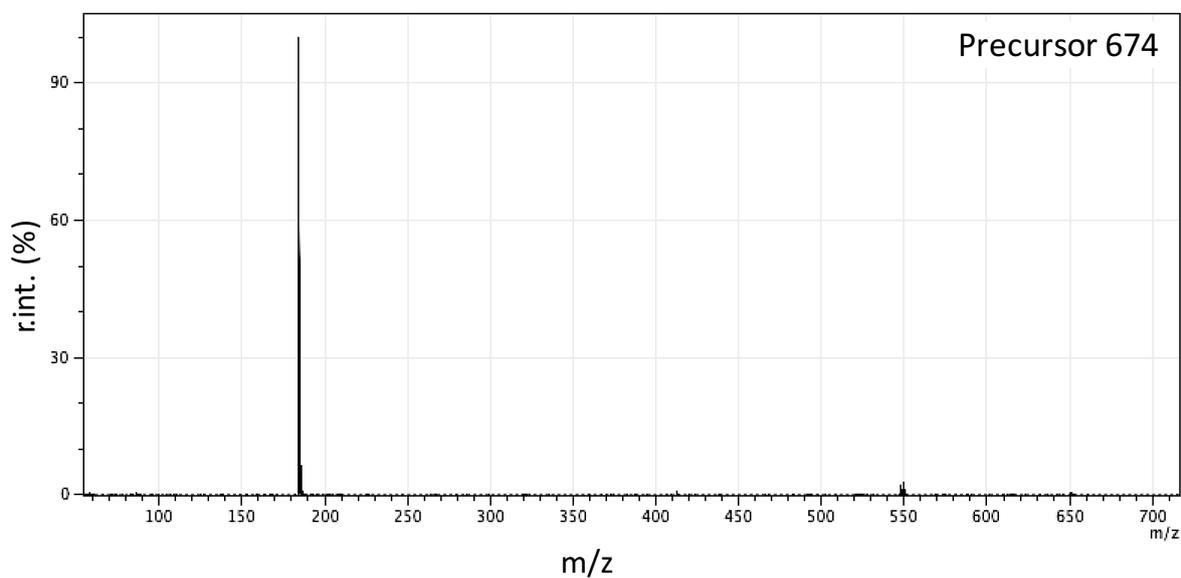


Figure 57: Fragmentation spectrum of the sample-specific precursor on 674 (positive ion reflector mode). The abscissa shows the m/z values in a range from 0 to 720. The ordinate represents the normalized relative intensity in %. The spectra were recorded with a laser energy of 80%. Processing of the spectra was done with mMass software.

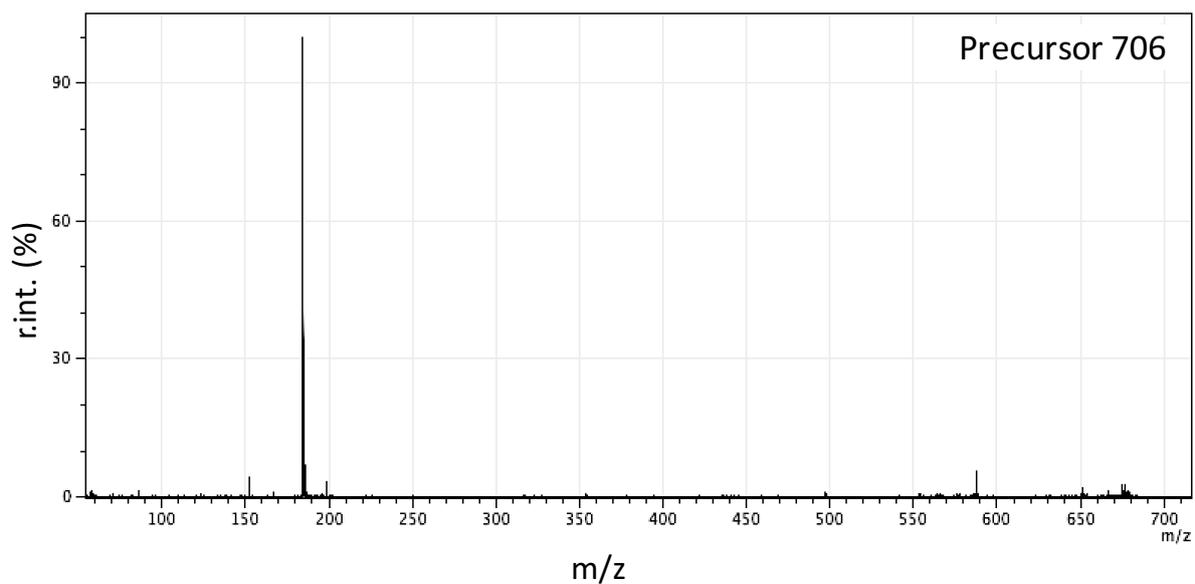


Figure 58: Fragmentation spectrum of the sample-specific precursor ion 706 (positive ion reflector mode). The abscissa shows the m/z values in a range from 0 to 720. The ordinate represents the normalized relative intensity in %. The spectra were recorded with a laser energy of 80%. Processing of the spectra was done with mMass software.

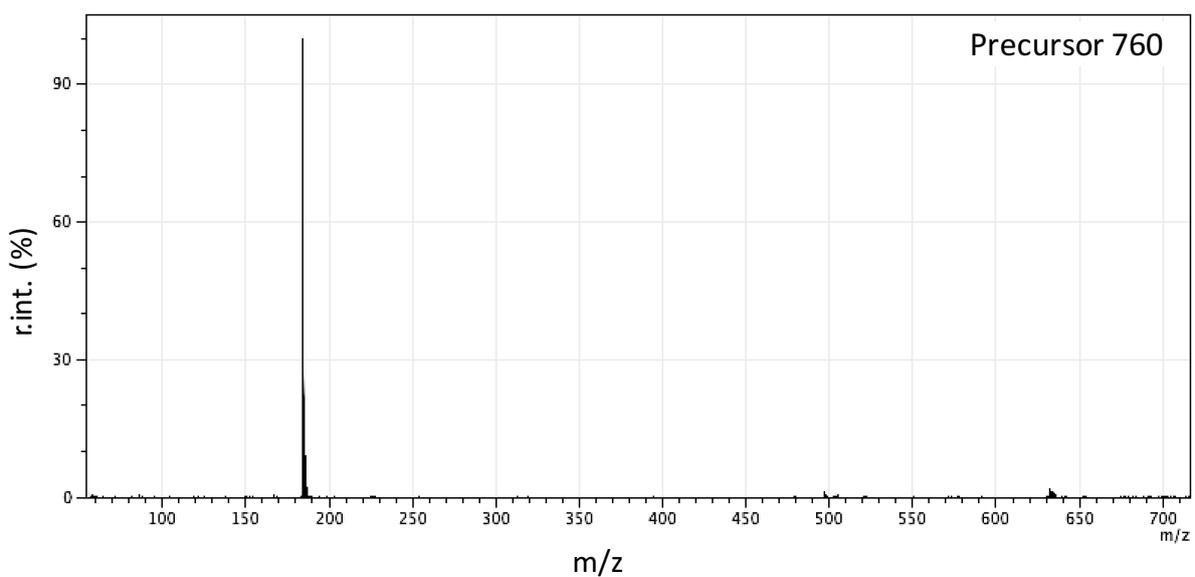


Figure 59: Fragmentation spectrum of the sample-specific precursor ion 760 (positive ion reflector mode). The abscissa shows the m/z values in a range from 0 to 720. The ordinate represents the normalized relative intensity in %. The spectra were recorded with a laser energy of 80%. Processing of the spectra was done with mMass software.

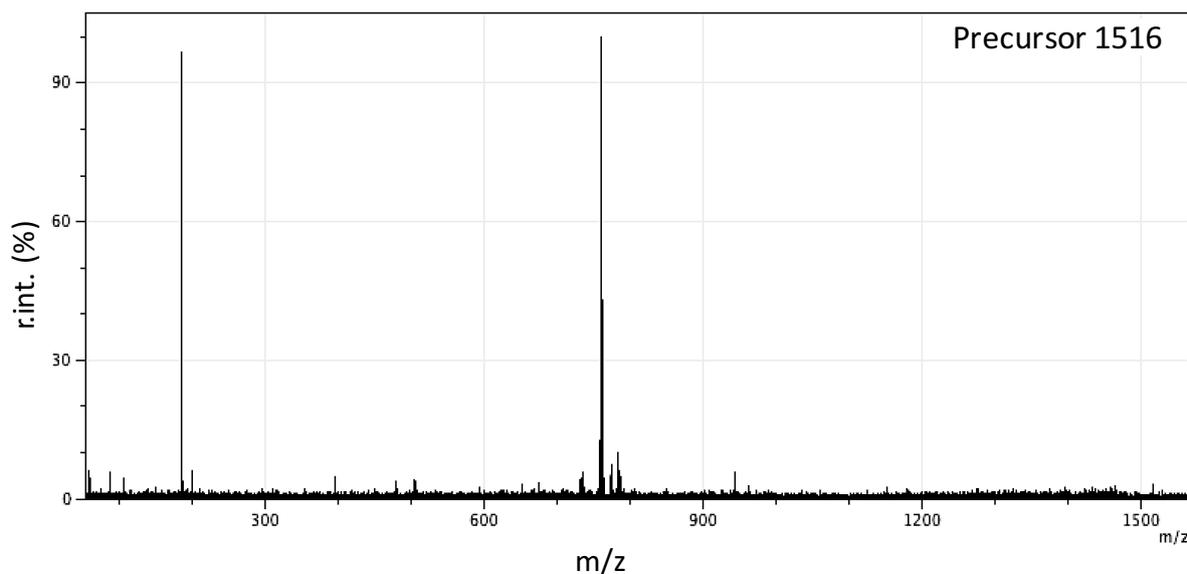


Figure 60: Fragmentation spectrum of the sample-specific precursor ion 1,516 (positive ion reflector mode). The abscissa shows the m/z values in a range from 0 to 1,500. The ordinate represents the normalized relative intensity (%). The spectra were recorded with a laser energy of 80%. Processing of the spectra was done with mMass software.

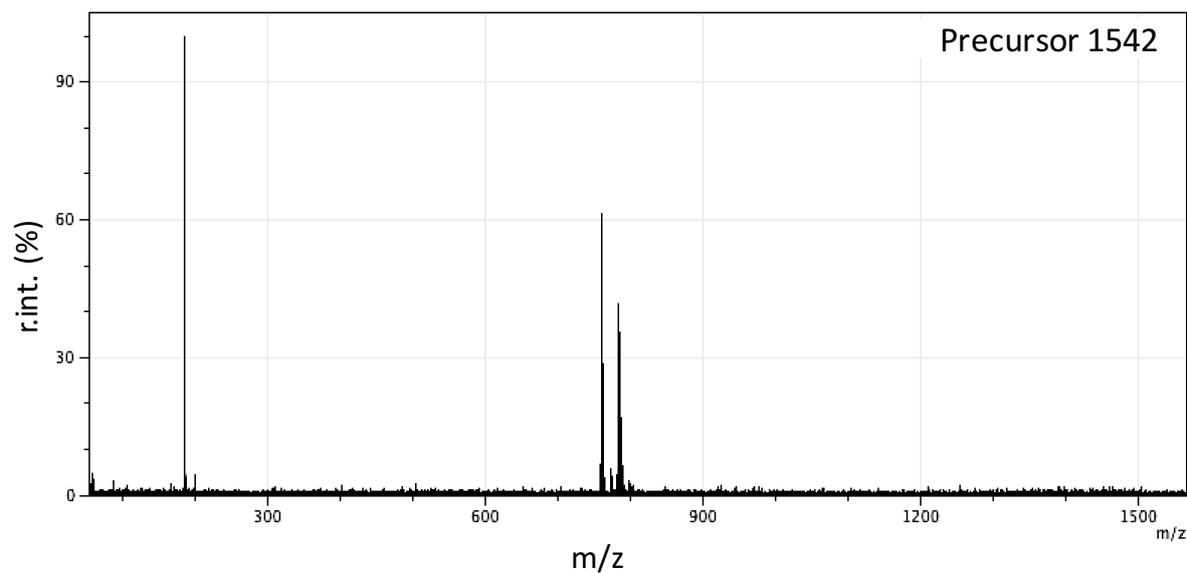


Figure 61: Fragmentation spectrum of the sample-specific precursor ion 1,542 (positive ion reflector mode). The abscissa shows the m/z values in a range from 0 to 1,500. The ordinate represents the normalized relative intensity (%). The spectra were recorded with a laser energy of 80%. Processing of the spectra was done with mMass software.

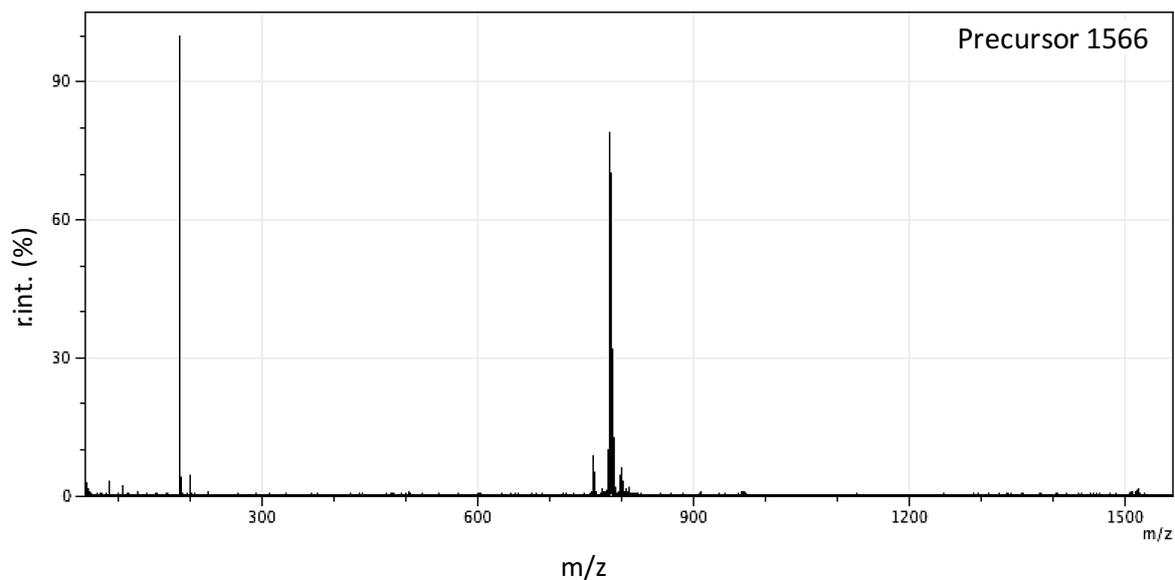


Figure 62: Fragmentation spectrum of the sample-specific precursor ion 1,566 (positive ion reflector mode). The abscissa shows the m/z values in a range from 0 to 1,500. The ordinate represents the normalized relative intensity (%). The spectra were recorded with a laser energy of 80%. Processing of the spectra was done with mMass software.

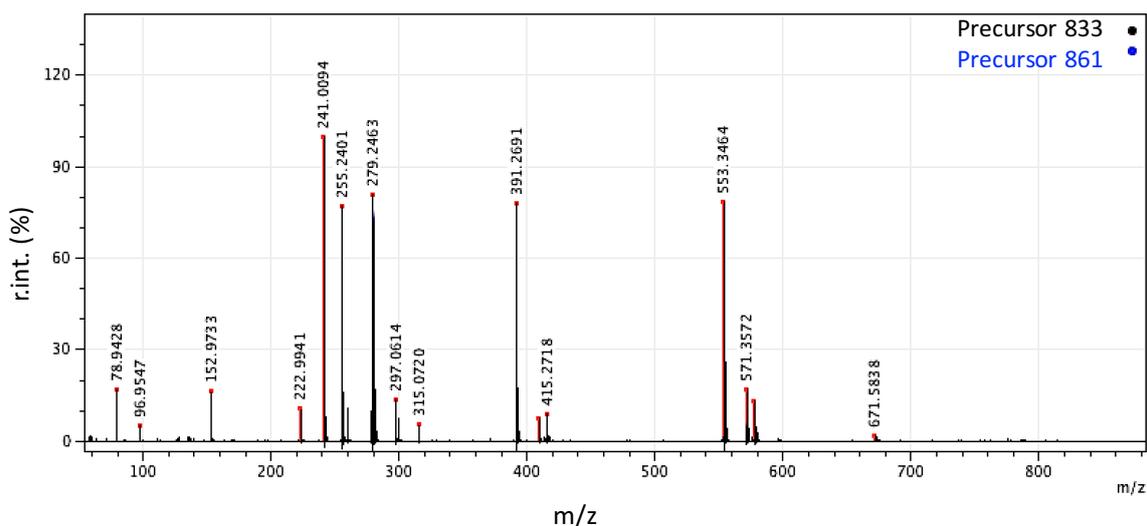


Figure 63: Fragmentation spectra of the sample-specific precursor ions 833 and 861 (negative ion reflector mode). The abscissa shows the m/z values in a range from 80 to 880. The ordinate represents the normalized relative intensity (%). The spectra were recorded with a laser energy of 80%. Processing of the spectra was done with mMass software.

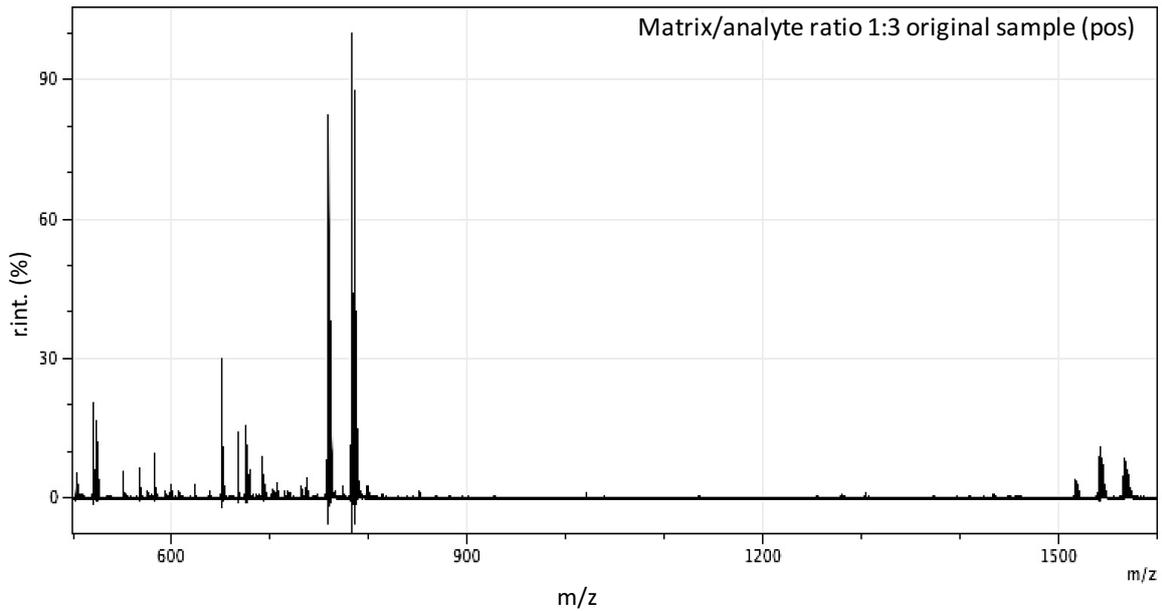


Figure 64: Spectrum of the condensation product of a 1:3 matrix/analyte ratio, original sample in the positive ion reflector mode. The measurement was carried out with a laser intensity of 35%. The abscissa shows the m/z values in a range from 500 to 1,600. The ordinate represents the normalized relative intensity in %. Processing of the spectra was done with mMass software as described in section 2.2.6.1

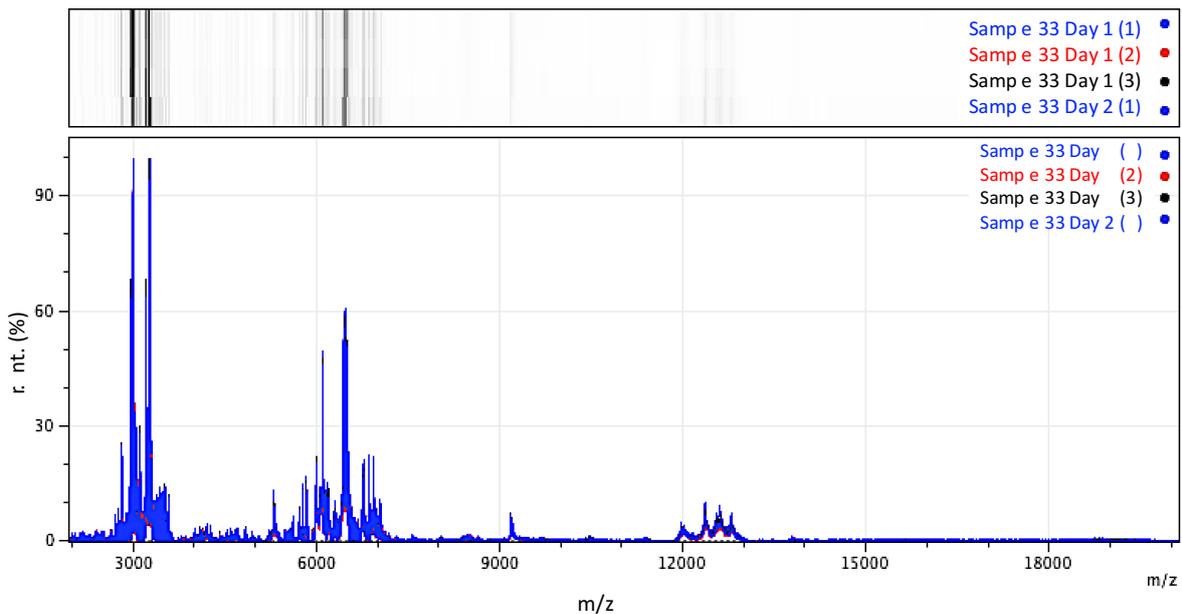


Figure 65: Intra- and inter-assay spectra for proteins with the exemplary authentic sample 33 prepared with the conventional FA/ACN method. Analysis was done by using saturated HCCA with the sample in a 1:1 matrix/analyte ratio (application of the layer procedure with subsequent air-drying for 5 min) in negative ion mode at a laser intensity of 45%. The abscissa shows the m/z values in a range from 2,000 to 20,000. The ordinate represents the normalized relative intensity in %. Processing of the spectra was done with mMass software as described in section 2.2.6.

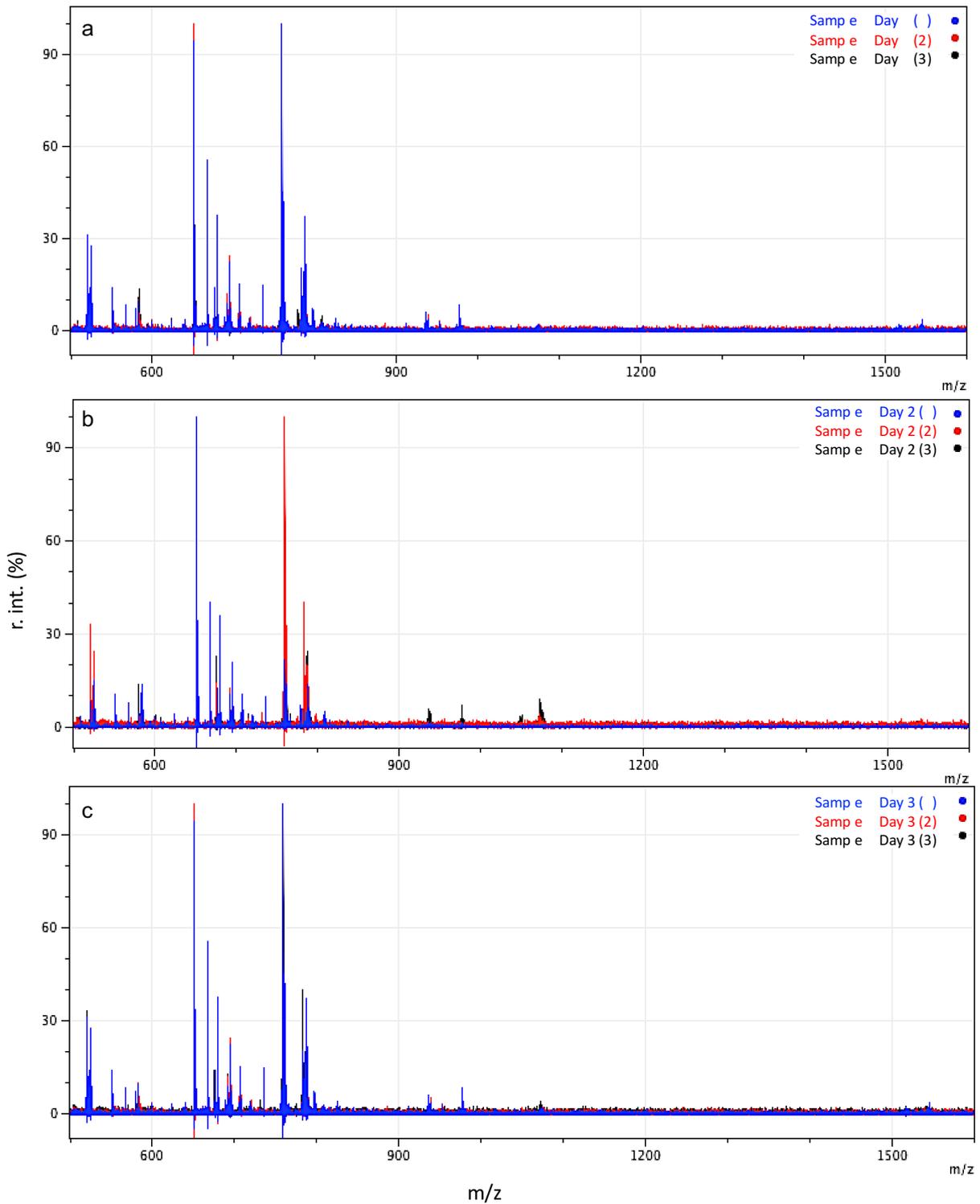


Figure 66: Intra- and inter-assay spectra for pds measured in the positive ion reflector mode with the exemplary authentic sample 1 prepared with the MTBE-based method. Analysis was done by mixing 9-AA with the sample in a 1:3 matrix/analyte ratio (application in dried droplet procedure with subsequent air-drying for 5 min) in ion reflector mode at a laser intensity of 55%. The abscissa shows the m/z values in a range from 500 to 1,600. The ordinate represents the normalized relative intensity in %. Processing of the spectra was done with mMass software as described in section 2.2.6.

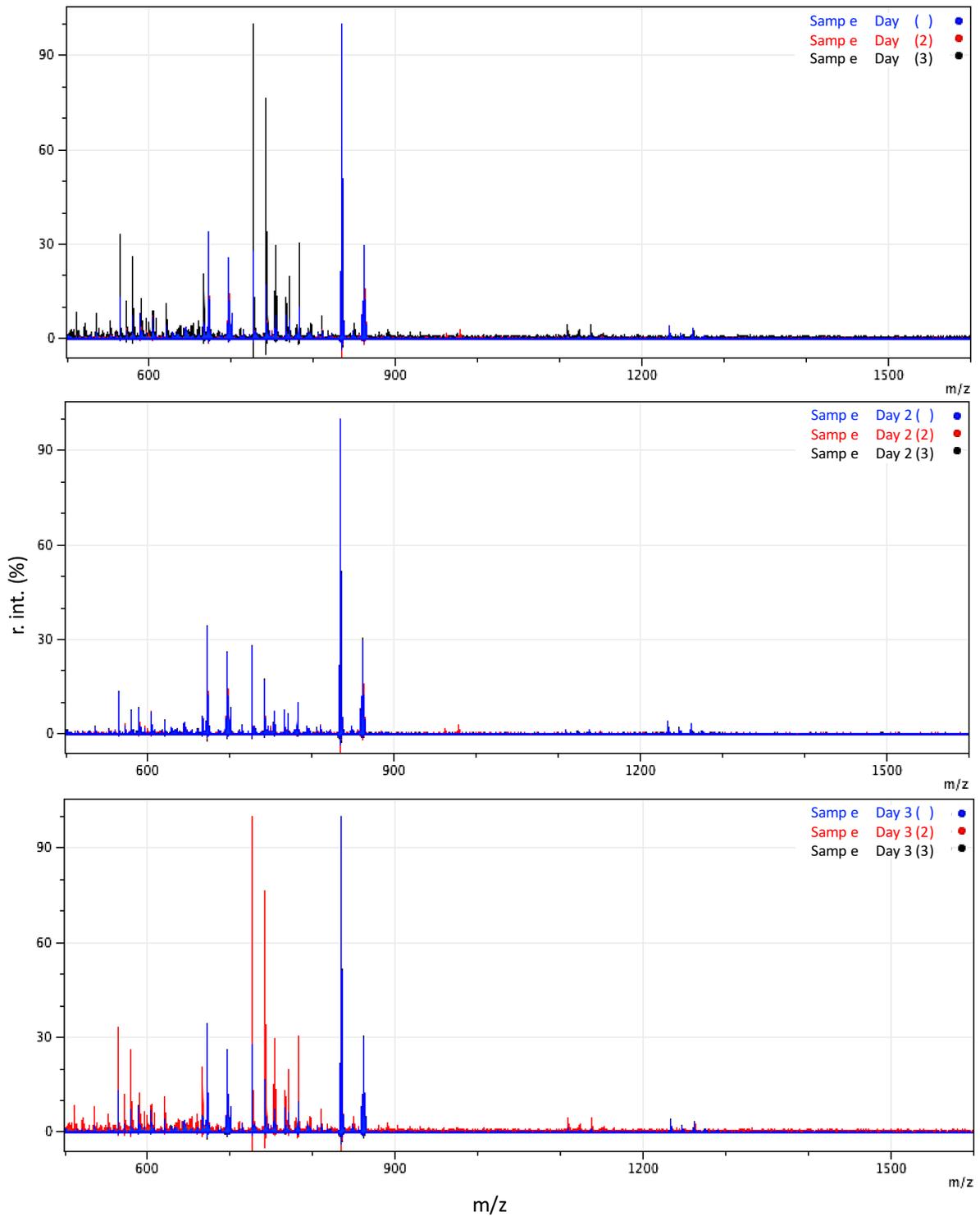


Figure 67: Intra- and inter-assay spectra for pds measured in the negative ion reflector mode with the exemplary authentic sample 1 prepared with the MTBE-based method. Analysis was done by mixing 9-AA with the sample in a 1:3 matrix/analyte ratio (application in dried drop let procedure with subsequent air-drying for 5 min) in ion reflector mode at a laser intensity of 55%. The abscissa shows the m/z values in a range from 500 to 1,600. The ordinate represents the normalized relative intensity in %. Processing of the spectra was done with mMass software as described in section 2.2.6.

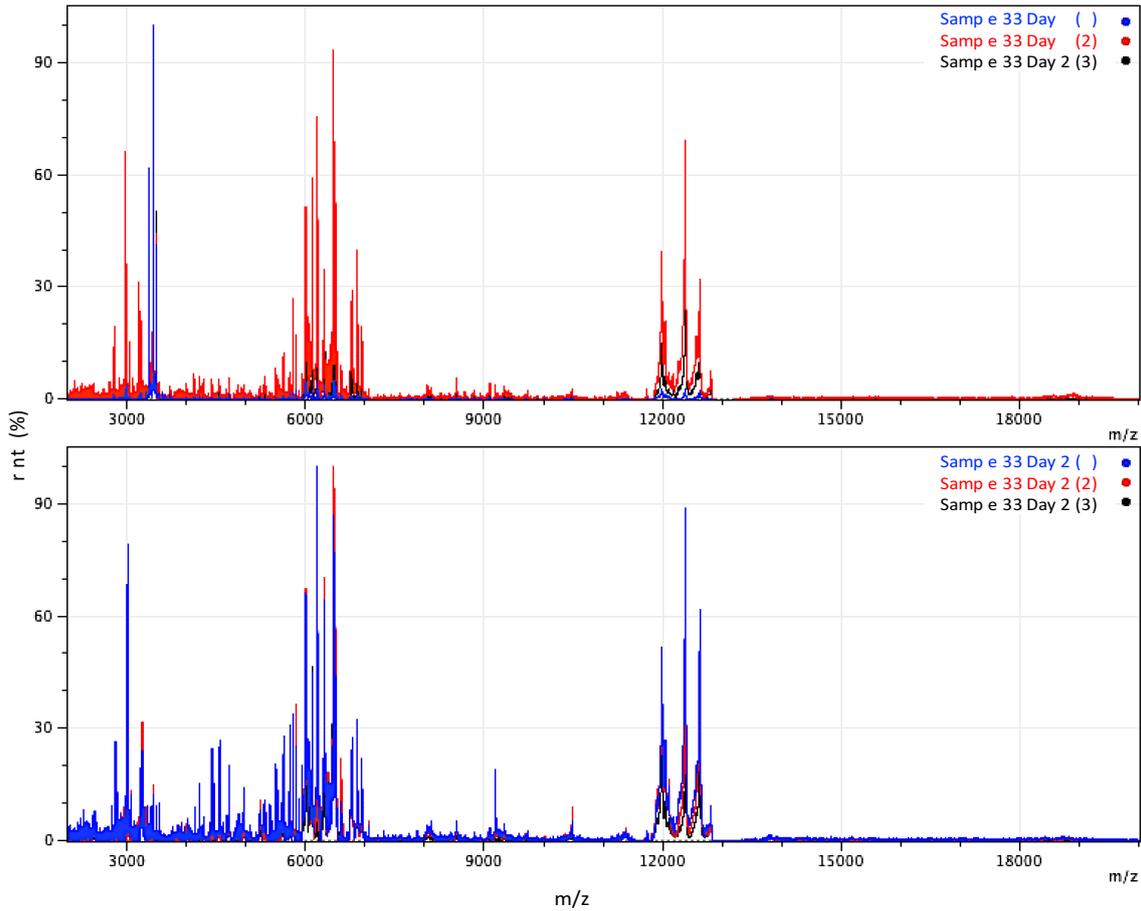


Figure 68: Intra- and inter-assay spectra for proteins with the exemplary non-reproducible authentic sample 33 abated with NIRL. Analysis was done by using saturated HCCA with the sample in a 1:1 matrix/analyte ratio (application of the layer procedure with subsequent air-drying for 5 min) in near-normal mode at a laser intensity of 45%. The abscissa shows the m/z values in a range from 2,000 to 20,000. The ordinate represents the normalized relative intensity in %. Processing of the spectra was done with mMass software as described in section 2.2.6.

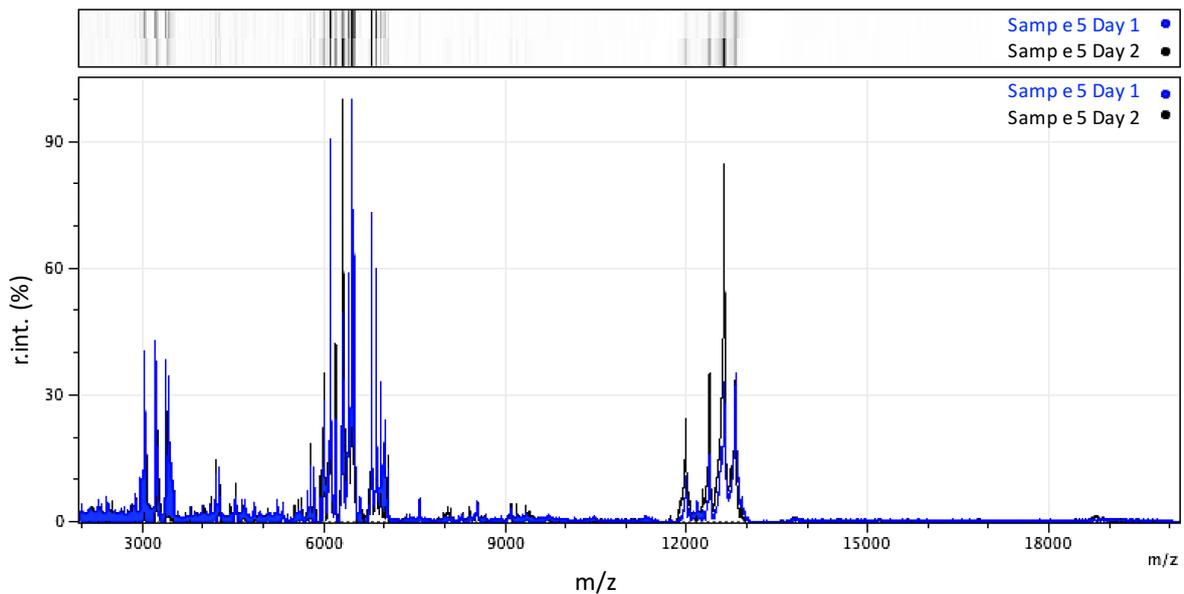


Figure 69: Inter-assay spectra for proteins with the exemplary reproducible authentic sample 5 abated with NIRL. Analysis was done by using saturated HCCA with the sample in a 1:1 matrix/analyte ratio (application of the layer procedure with subsequent air-drying for 5 min) in near-normal mode at a laser intensity of 45%. The abscissa shows the m/z values in a range from 2,000 to 20,000. The ordinate represents the normalized relative intensity in %. Processing of the spectra was done with mMass software as described in section 2.2.6.

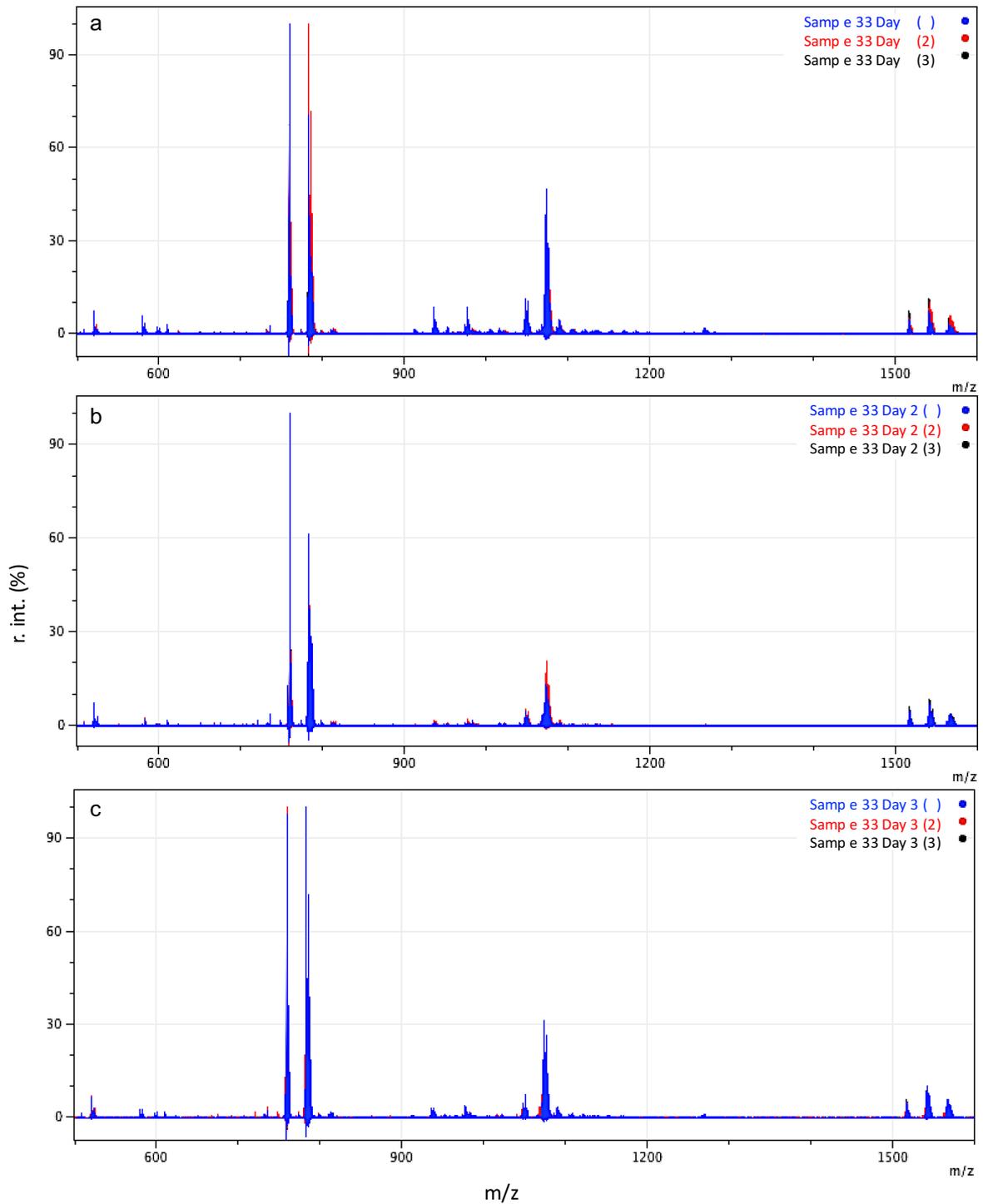


Figure 70: Intra- and inter-assay spectra for pds measured in the positive ion reflector mode with the exemplary authentic sample 33 abated with NIRL. Analysis was done by mixing 9-AA with the sample in a 1:3 matrix/analyte ratio (application in dried drop let procedure with subsequent air-drying for 5 min) in reflector ion mode at a laser intensity of 35%. The abscissa shows the m/z values in a range from 500 to 1,600. The ordinate represents the normalized relative intensity in %. Processing of the spectra was done with mMass software as described in section 2.2.6.

Table 32: Complete sample set, containing the year, the cultivar and the origin of each original analyzed sample respectively

Sam. no.	Orig. ID	cultivar	year	origin
1	159	Chandler	2018	Italy
2	171	Chandler	2018	Italy
3	23	Chandler	2017	USA
4	162	Chandler	2018	Italy
5	25	Chandler	2017	USA
6	165	Chandler	2018	Italy
7	25_2	Chandler	2017	USA
8	155	Franquette	2018	France
9	156	Franquette	2018	France
10	153	Franquette	2018	France
11	71	Franquette	2017	France
12	59	Franquette	2017	France
13	68	Franquette	2017	France
14	65	Franquette	2017	France
15	58	Franquette	2017	France
16	67	Lara	2017	France
17	148	Lara	2018	France
18	40	Lara	2017	Italy
19	72	Lara	2017	France
20	149	Lara	2018	France
21	163	Lara	2018	Italy
22	145	Lara	2018	France
23	166	Lara	2018	Italy
24	150	Lara	2018	France
25	62	Lara	2017	France
26	69	Lara	2017	France
27	172	Lara	2018	Italy
28	169	Lara	2018	Italy
29	102	NB	2018	Germany
30	103	NB	2018	Germany
31	120	NB	2018	Germany
32	130	NB	2018	Hungary

33	124	NB	2018	Hungary
34	125	NB	2018	Hungary
35	102_2	NB	2018	Germany
36	126	NB	2018	Hungary