

PhD THESIS IN PHARMACY

NIKOLAOS LEMONAKIS

*State of the art analytical
techniques in the
phytochemical and biological
evaluation of bioactive
substances from *Olea europea*.
Targeted quantification and
metabolomics in different case
studies*



National and Kapodistrian
University Of Athens

*State of the art analytical techniques in
the phytochemical and biological
evaluation of bioactive substances from
Olea europea. Targeted quantification
and metabolomics in different case studies*

NIKOLAOS LEMONAKIS

*State of the art analytical techniques in
the phytochemical and biological
evaluation of bioactive substances from
Olea europea. Targeted quantification
and metabolomics in different case studies*

Department of Pharmacognosy and Natural Products Chemistry
School of Pharmacy
National and Kapodistrian University of Athens
Greece
2014

EXAMINING COMMITTEE for the PhD Thesis

Prof. A.L. Skaltsounis¹

Assistant Prof. E. Gikas²

Researcher B' C. Tamvakopoulos³

Prof. E. Mikros²

Associate Prof. A. Tsarbopoulos⁴

Associate Prof N. Thomaidis⁵

Prof. J.L. Wolfender⁶

¹Department of Pharmacognosy and Natural Products Chemistry, Faculty of Pharmacy, National and Kapodistrian University of Athens

²Department of Pharmaceutical Chemistry, Faculty of Pharmacy, National and Kapodistrian University of Athens

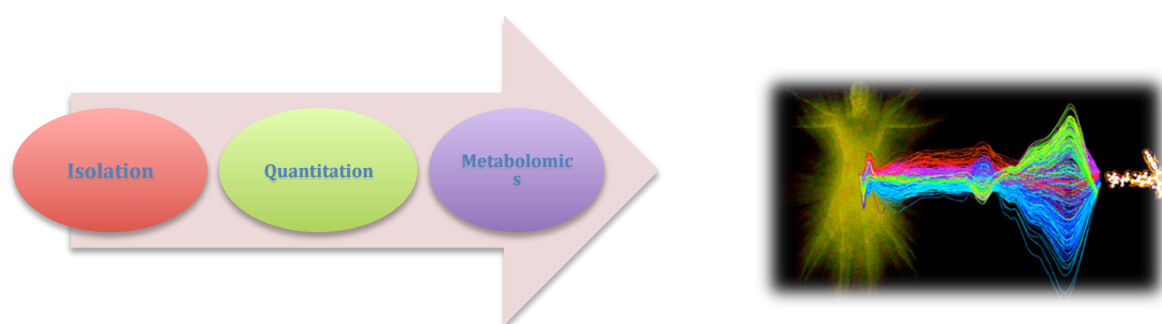
³Laboratory of Pharmacology-Pharmacotechnology, Biomedical Research Foundation Academy of Athens

⁴Department of pharmacology, Medical School, National and Kapodistrian University of Athens

⁵Laboratory of Analytical Chemistry, Department of Chemistry, National and Kapodistrian University of Athens

⁶School of Pharmaceutical Sciences, EPGL, University of Geneva, University of Lausanne

NIKOLAOS LEMONAKIS



The research reported in this thesis was carried out at the Department of Pharmacognosy and Natural Products Chemistry, Faculty of Pharmacy, National and Kapodistrian University of Athens

Table of Contents

Summary		i
Chapter 1	Optimization of parameters affecting signal intensity to an LTQ-Orbitrap: A DOE approach using oleuropein and hydroxytyrosol as model compounds	1
Chapter 2	Development and validation of a combined methodology for assessing the total quality control of herbal medicinal products – Application to oleuropein preparations	44
Chapter 3	Validation of a bioanalytical methodology based on UHPLC-HRMS/MS for the quantification of oleuropein in human serum. Application to a pharmacokinetic study	85
Chapter 4	Effect of supplementation with olive leaf extract enriched in oleuropein (Benolea® EFLA®943), on the metabolome and on the redox status of athletes' blood and urine. A metabolomic approach	111
Chapter 5	Metabolomics study of hydroxytyrosol's administration for the treatment of metabolic syndrome in a rat model	234
Chapter 6	Metabolism of hydroxytyrosol after its administration to a metabolic syndrome rat model	312
Chapter 7	Metabolomics-based time series analysis of hydroxytyrosol in rat plasma	332
Chapter 8	Impact of a functionalized olive oil extract on the uterus and the bone in a model of postmenopausal osteoporosis	346
Acknowledgments		370
Publications		375
CV		

Δε μπορείς να ανακαλύψεις νέους ωκεανούς αν δεν έχεις
το κουράγιο να χάσεις την ακτή από τα μάτια σου

Πλάτωνας

Dedicated to the memory of Tasos

S

Summary

Nikolaos Lemonakis

Olive derived substances occupy a central role in the diet of the Mediterranean basin, with low occurrence of cardiovascular diseases and cancer incidents. It is evidenced that such molecules (especially the secoiridoid oleuropein and the polyphenol hydroxytyrosol) could play a significant role in the prevention of the aforementioned as well as various other diseases, and could ameliorate the human well being in terms of health status. Thus, their study is of paramount importance.

The current scientific trends in the field of analytical chemistry enable the in-depth exploration of their phytochemical properties, the characterization of their fate after administration, as well as their impact on the systems biology of the organism. In this context, integrated methodologies, covering all the aforementioned aspects, need to be developed.

In the current PhD Thesis, as the sensitive determination of the examined substances is highly desirable, a design of experiments approach has been developed in order to optimize the analytical conditions for all the methodologies that were developed thereafter. As the health effects of oleuropein, as a herb medicinal product, consists a major concern, the total quality of these formulations was explored, employing both targeted and untargeted (multivariate) approaches. The administration of the aforementioned oleuropein-based herb medical products in athletes, aiming to verify their effect on the human redox status after exercise, have been studied using a pharmacokinetic and a MS-based metabolomic approach. In the same context, the administration of hydroxytyrosol in an induced metabolic syndrome rat model has been performed. Furthermore, the metabolic fate (metabolism) of hydroxytyrosol has been explored employing an MS-based approach, whereas a

metabolomic-based time series study has been used to gain insight of the substance's fate and its impact on the organism. In a parallel mode, administration of the total polyphenolic fraction derived from olive oil, has been realized in order to understand its impact on osteoporosis. The isolation of oleuropein has been attempted using olive leaves as the source, whereas hydroxytyrosol was isolated from olive oil mill wastes. Finally, the total polyphenolic fraction was obtained from pure extra virgin olive oil.

Overall, in the current PhD Thesis, state of the art analytical techniques along with cutting edge biological experiments have offered new knowledge on the analytical and biochemical impact of the promising bioactive substances from *Olea europea*.

1

Optimization of parameters affecting signal intensity to an LTQ-Orbitrap: A DOE approach using oleuropein and hydroxytyrosol as model compounds

This chapter was submitted for publication in the Journal of the American Society for Mass Spectrometry

ABSTRACT

INTRODUCTION: A multistage optimization of all the parameters affecting detection/response in an LTQ-orbitrap analyzer was performed, using a design of experiments methodology. The signal intensity, a critical issue for mass analysis, was investigated for the first time and the optimization process was completed in three successive steps, taking into account the three main regions of an orbitrap, the ion generation, the ion transmission and the ion detection regions. Oleuropein and hydroxytyrosol were selected as the model compounds.

METHODS: A high-resolution LTQ Orbitrap Discovery mass spectrometer was used for the determination of the analytes of interest. Thus, oleuropein and hydroxytyrosol were infused via the embedded to the instrument syringe pump and they were analyzed employing electrospray ionization (ESI) in the negative high-resolution full scan ion mode.

RESULTS: The involved parameters of the three main regions of the LTQ-orbitrap were independently optimized in terms of maximum sensitivity. In this context, factorial design, response surface model and Plackett Burman experiments were performed and analysis of variance was carried out to evaluate the validity of the statistical model and to determine the most significant parameters for signal intensity.

CONCLUSIONS: The optimum MS conditions for each analyte were summarized and the method optimum condition was achieved by maximizing the desirability function. Our observation showed good agreement between the statistically predicted optimum response and the responses collected at the predicted optimum condition.

KEYWORDS

Design of Experiments, DOE, LTQ-Orbitrap, oleuropein, hydroxytyrosol, signal intensity

INTRODUCTION

Orbitrap is a novel Fourier-transform (FT) mass spectrometer that is used for high-resolution/high-mass accuracy measurements. Nowadays there are different commercial models of orbitrap instruments i.e. LTQ-orbitrap Discovery (LTQ stands for “linear trap quadrupole”), LTQ-orbitrap XL, Velos, Elite, Exactive, Q Exactive and Fusion. Orbitrap is widespread in the fields of proteomics¹, metabolomics² and generally in the omics technologies as well as in various other analytical applications, such as structure elucidation³ and dereplication⁴. These applications are widely performed in the research field as well as in the pharmaceutical industry.

A crucial step before performing any of the aforementioned applications is the signal generation via MS. This step is carried out by infusing the individual compounds under study in the mass spectrometer and measuring their signal intensity. Understanding of the signal generation is essential for increasing the sensitivity of the instrument, which can dramatically improve the overall performance of the instrument.

The LTQ-Orbitrap is a hybrid mass spectrometer, which consists from two main components, the linear ion trap and the orbitrap, as illustrated in Figure 1. Entering into details, initially, the electrospray source creates ions employing a high voltage field with the aid of two gas flows, i.e. the sheath and the auxiliary gas. Multiple parameters also contribute to the efficient ion generation such as the capillary voltage and the capillary temperature or to the initial ion transmission from the source, such as the tube lens voltage. In a next step, the generated ions are transferred by radio frequency-only (RF) multipoles to the linear ion trap (LIT)^{5, 6} where they are trapped employing a combination of RF, direct current (DC) and fringing potentials.

Consequently, the ions can be radially ejected from the LIT into a pair of secondary electron multipliers⁷ as in a typical LTQ-only experiment, or axially ejected and then transferred through an RF-only octopole and the gate electrode into the C-trap as in typical LTQ-orbitrap experiment.

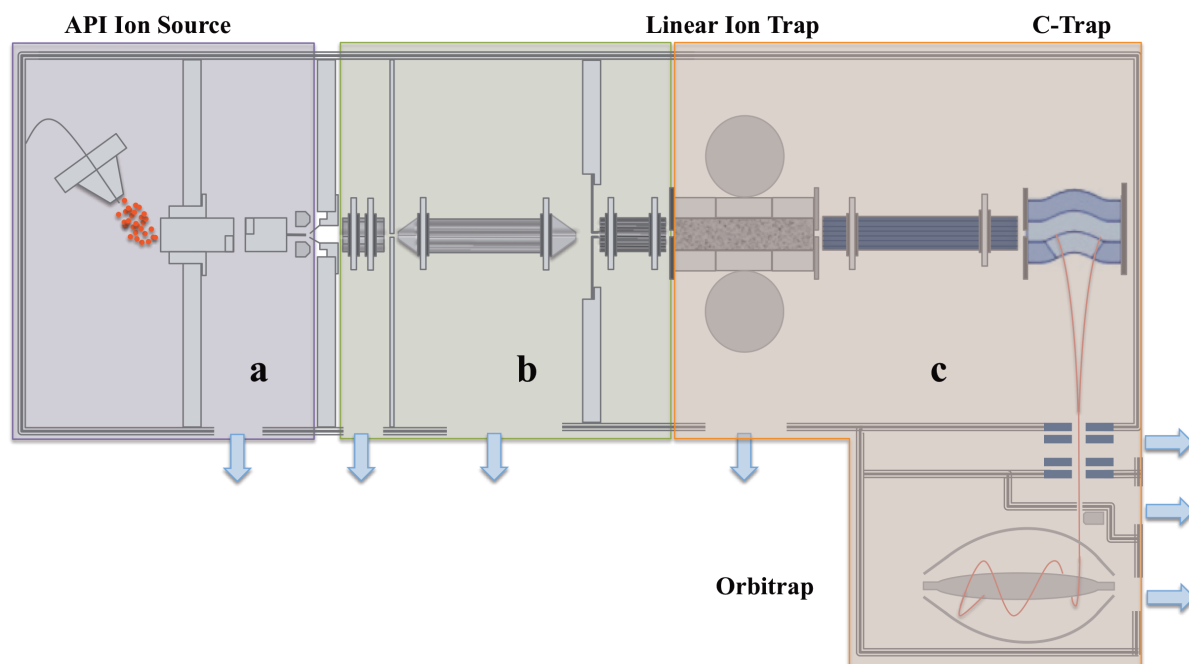


Figure 1. Schematic of the commercial LTQ-Orbitrap Discovery mass spectrometer with the API ion source, the linear ion trap the c-trap and the orbitrap: **(a)** ion generation region (purple box), **(b)** ion transmission region (green box) and **(c)** ion detection region (orange box).

At the beginning of an orbitrap experiment, the ions form a long arc along the curved axis of the C-trap whereas they are axially confined by applying a potential to both the gate and the trap electrodes. During the C-trapping process, the entering ions begin to lose energy by collisions with N₂ whereas the curve shape of the C-trap forces the ions to undergo multiple oscillations and thus undergo collisional cooling. After ejection from the C-trap, the ions are accelerated through an S-shaped

trajectory⁸ and converge into a tight ion cloud, which then passes through the entrance aperture of the orbitrap where they oscillate with characteristic frequencies according to their m/z values generating an electric signal to the orbitrap electrodes. Finally, the produced multiplex signal from the ion movement detected is subjected to Fourier-transformation producing the spectra. The sensitivity of the LIT- and FT-trapping is regulated by the automatic gain control algorithm (AGC) embedded to the instrument software.

Generally, the basic procedure followed for the optimization of MS parameters in order to increase the signal intensity, is auto-tuning, which is usually adequate and it is executed by performing one factor at a time (OFAT) experiments⁹. A major weakness of this route is that the possible interactions among examined factors are not taken into account and this leads to a lack of information about the interaction of variables and sometimes results in the wrong conclusions. Therefore, application of optimal conditions, estimated via the OFAT strategy, can result in reduction of the method performance.

This gap can be filled employing the design of experiments (DOE) theory that allows studying the influence of several process parameters on one or more responses of the design. Furthermore, DOE can create mathematical models for the responses optimizing the selected factors that affect the system^{10, 11}. DOE is able to determine the optimal experimental conditions using a minimum number of experiments, whereas an additional major advantage is its inherent ability to predict interactions between the various experimental factors. Generally, the DOE strategy comprises of performing a series of experiments in accordance to a predefined plan, revealing the significant information about the interaction between the examined parameters and

the response; in the current case the signal intensity of selected compounds to an LTQ-orbitrap Discovery mass spectrometer.

There are different types of DOE depending on the study to be performed. The most commonly used types are the screening designs (SD) and the response surface methodologies (RSM). SDs are the designs that are used to uncover the most influential factors in a high-throughput manner. A subcategory of SD, namely the fractional-factorial designs (FFD)¹², represent an efficient method for investigating the effects of several parameters performing only a fraction of a full-factorial design.

Central composite designs (CCD), which belong to the RSM, are designs used for modeling a response surface and subsequently optimizing the levels of the parameters involved, in order to obtain a desired response. Briefly they are created by the fusion of three experimental run levels executed in random: a factorial design augmented by star points and also center values of all factor ranges¹³. In order to assure the rotatability of the design, the star points are spaced at a distance α from the center, which is based on the desired properties of the design.

Among their various applications¹⁴⁻¹⁸, the DOE experiments have been applied in the development of a variety of mass spectrometry-based analytical methodologies. Briefly, DOE has been implemented in the optimization of the ion source interface parameters¹⁹⁻²³ as well as to the optimization of an LC-MS method for both the LC and MS performance simultaneously^{10, 19} or independently²⁴. DOE has also been used for optimization of the ion trap²⁵ or triple quadrupole factors²⁶ whereas recent publications are focused on the parameter optimization of an LTQ-orbitrap in order to improve the proteome coverage^{27, 28}. In addition, factors affecting the accuracy of the mass measurements²⁹ were also optimized following a DOE strategy.

Nowadays, there is a need for the ultra sensitive determination of *Olea europea* bioactive compounds, such as oleuropein (OE) and hydroxytyrosol (HT), in various matrices e.g. biological specimens, extracts etc. OE is a natural secoiridoid glycoside which has been evidenced to exhibit various biological properties such as cardioprotective³⁰, anti-ischemic³¹, antioxidant³², etc. Its main degradation product, HT, is also a naturally occurring compound, which can fortify health through its powerful antioxidant capacity³³, as well as its anti-inflammatory³⁴ and antiplatelet aggregation action³⁵, etc.

Many analytical methodologies have been developed and applied for the determination of OE and HT in various matrices such as olive oil³⁶, table olives³⁷, herbal medicinal products³⁸, byproducts³⁹ and biological fluids⁴⁰. Thus, OE has been analysed employing UV⁴¹, fluorescence⁴² and MS/MS³⁸. It is apparent from the literature that the method of choice for the ultra sensitive determination of the aforementioned compounds is mass spectrometry (MS).

The current paper presents a three-step optimization strategy, aiming to improve the signal intensity of specific bioactive compounds from *Olea europaea* (OE and HT) serving as model compounds in an ESI-LTQ orbitrap Discovery. Overall, a DOE strategy has been employed for the design of a structured, systematic method for determining the factors that affect the sensitivity in an ESI-LTQ orbitrap in terms of signal intensity.

EXPERIMENTAL

Chemical and Reagents

OE and HT (Figure 2) were isolated from olive leaves and olive mill wastewater respectively based on previously described procedures^{31, 43}. Their purity was found to be more than 98% (HPLC-UV) and their structures were identified by 1 & 2D NMR and High Resolution Mass Spectrometry (HRMS). No specific permissions were required as the examined plant is an endemic species widely spread in the Mediterranean region, as well as in Australia, the American continent, Africa and Asia. Furthermore *Olea europea* is not included in any list of endangered or protected species. All solvents used in this study were of LC-MS grade. Acetonitrile (ACN), methanol (MeOH) and water were purchased from Fluka/Riedel-de Haën (Switzerland) and they were MS grade.

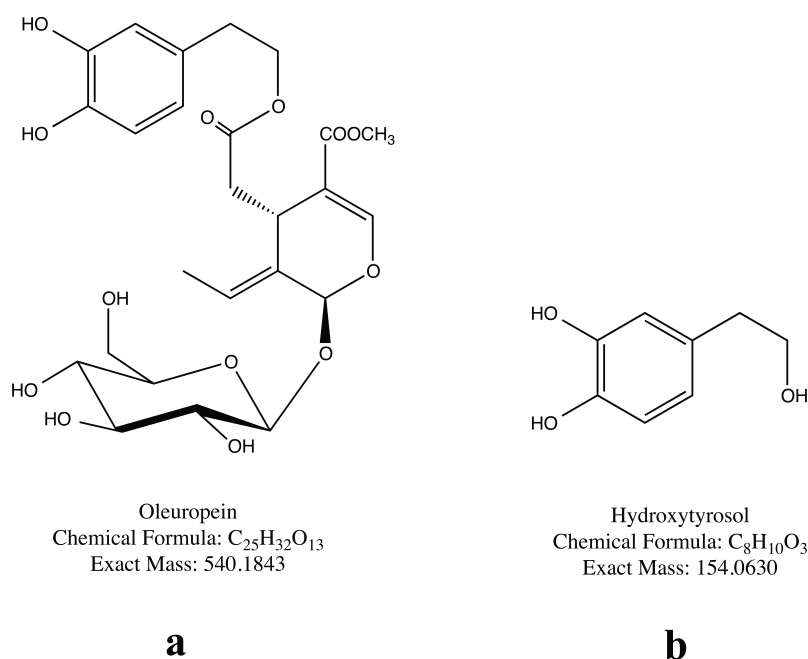


Figure 2. Structures of OE (**a**) and HT (**b**).

Analytes

Stock solutions of OE and HT were prepared at a concentration of 1.0 mg/mL in MeOH and dilutions were made up at 1 ug/mL in ACN/water (8:2, v/v). The above solution, (experimental solution -EXS), was used for the completion of the experiments.

Mass spectrometry

Mass spectrometry experiments were carried out on a hybrid LTQ Orbitrap Discovery mass spectrometer (Thermo Scientific, Germany), which was equipped with an electrospray ionisation (ESI) MAX2 source. Mass analysis was performed in the full scan mode over a mass range of 50–1000 m/z, using a resolution of 30000. EXS (OE: $[M-H]^+ = 539.1756$ and HT: $[M-H]^+ = 153.0557$) was used to perform a set of auto-tuning procedures but also to perform the rest of the experiments for the DOE approach, as described below. The EXS infused at 5 uL/min via the embedded to the instrument syringe pump, provided the signal intensity of each analyte, expressed in counts per second (cps). Mass spectra were collected in the negative-ion mode over 30 scans. The mean intensity values for both analytes, used for all experiments throughout, were obtained by averaging the acquired spectra.

In order to gain insight about the optimal instrumental conditions, the auto-tuning procedure (essentially an OFAT procedure) has been employed, as proposed by the instrument vendor. The results have been used as part of a scoping experiment²⁹. In detail, the EXS was initially infused at 5 uL/min with the automatic tuning routine focusing at m/z 539.1 and 153.1 (for OE and HT respectively, $[M-H]^-$). The

parameters optimized via the auto-tuning procedure can be found in Table 1. These values were examined to compare the optimum values for the examined parameters after auto-tuning with these that were resulted from the DOE study.

Software

The Design-ExpertW 7.0.0 software (Stat-Ease Inc., Minneapolis, USA) was used for performing the experimental plans, data interpretation and for determining which factors affect the signal intensity of the tested compounds after infusion in the orbitrap. For data acquisition, the Xcalibur 2.07 software (Thermo Fisher Scientific Inc., San Jose, USA) was used.

Optimization regions	Parameters	Process			
		Auto-tuning Values	Optimal Values after DOE	Auto-tuning Values	Optimal Values after DOE
		OE		HT	
Ion generation (ESI Source)	Sheath gas	76	30	78	30
	Aux Gas	3	3	6	6
	Sweep gas	3	3	1	1
	Spray voltage	3.1	3.1	3.1	3.1
	Capillary voltage	-8	11	-18	11
	Capillary Temperature	356	356	356	356
	Tube Lens	-78	-48	-38	-48
Ion transmittion (Ion Optics)	Multipole 00	2.5	2.5	2.5	2.5
	Lens 0	1	1	1.5	1.5
	Multipole 0	4	4	4.5	4.5
	Lens 1	31	31	39	39
	Gate Lens	78	78	8	8
	Multipole 1	13.5	13.5	18.5	18.5
	Multipole RF	850	850	850	850
Ion detection (Injection Control)	Front Lens	3.25	4	3.5	4
	AGC Target Value (IT)	30000	30000	30000	30000
	Maximum Injection Time (IT)	10	10	10	10
	AGC Target Value (FT)	200000	45500	200000	45500
	Maximum Injection Time (FT)	500	317	500	317

Table 1. Optimized parameters via the auto-tuning procedure and via the DOE approach for OE and HT.

		Factorial 2 Level Factorial Design 2^{7-1}			Central Composite Response Surface Design					Factorial Plackett Burman Design			Factorial 2 Level Factorial Design 2^{8-2}			Central Composite Response Surface Design				
		Low actual (-1)	Mean (0)	High actual (+1)	$-\alpha$ (-1.68 \times low actual)	Low actual (-1)	Mean (0)	High actual (+1)	$+\alpha$ (1.68 \times low actual)	Low actual (-1)	Mean (0)	High actual (+1)	Low actual (-1)	Mean (0)	High actual (+1)	$-\alpha$ (-1.41 \times low actual)	Low actual (-1)	Mean (0)	High actual (+1)	$+\alpha$ (+1.41 \times low actual)
Optimization regions	Parameters																			
Ion generation (ESI Source)																				
	Sheath gas	13	31.5	50	70	30	45	60	20											
	Aux Gas	0	8	16																
	Sweep gas																			
	Spray voltage	3	3.25	3.50																
	Capillary voltage	-50	-15	20	-50	-30	0	30	50											
	Capillary Temp.	300	350	400																
	Tube Lens	-50	-17.5	15	-72	-60	42.5	-25	-13											
Ion transmission (Ion Optics)																				
	Multipole 00									0	1.5	3	0	1.5	3					
	Lens 0									2	4.5	7	2	4.5	7					
	Multipole 0									4.5	5.75	7	4.5	5.75	7					
	Lens 1									4.5	7.25	10	4.5	7.25	10					
	Gate Lens									4.5	69.75	135	4.5	69.75	135					
	Multipole 1									4.7	69.85	135	4.7	69.85	135					
	Multipole RF									200	525	850	200	525	850					
	Front Lens									4	7	10	4	7	10					
Ion detection (Injection Control)																				
	AGC Target Value (IT)															5000	10125	22500	34874	40000
	Max. Injection Time (IT)															50	57	75	93	100

Chapter 1

	AGC Target Value (FT)															10000	45147	130000	214900	250000
	Max. Injection Time (FT)															200	317	600	883	1000

Table 2. Overview of the 3-step optimization strategy. Parameters investigated and their corresponding levels.

		Factorial 2 Level Factorial Design 2^{7-1}			Central Composite Response Surface Design					Factorial Plackett Burman Design			Factorial 2 Level Factorial Design 2^{8-2}			Central Composite Response Surface Design				
		Low actual (-1)	Mean (0)	High actual (+1)	$-\alpha$ (-1.68 \times low actual)	Low actual (-1)	Mean (0)	High actual (+1)	$+\alpha$ (1.68 \times low actual)	Low actual (-1)	Mean (0)	High actual (+1)	Low actual (-1)	Mean (0)	High actual (+1)	$-\alpha$ (-1.41 \times low actual)	Low actual (-1)	Mean (0)	High actual (+1)	$+\alpha$ (+1.41 \times low actual)
Optimization regions	Parameters																			
Ion generation (ESI Source)																				
	Sheath gas	13	31.5	50	70	30	45	60	20											
	Aux Gas	0	8	16																
	Sweep gas																			
	Spray voltage	3	3.25	3.50																
	Capillary voltage	-50	-15	20	-50	-30	0	30	50											
	Capillary Temp.	300	350	400																
	Tube Lens	-50	-17.5	15	-72	-60	42.5	-25	-13											
Ion transmission (Ion Optics)																				
	Multipole 00									0	1.5	3	0	1.5	3					
	Lens 0									2	4.5	7	2	4.5	7					
	Multipole 0									4.5	5.75	7	4.5	5.75	7					
	Lens 1									4.5	7.25	10	4.5	7.25	10					
	Gate Lens									4.5	69.75	135	4.5	69.75	135					
	Multipole 1									4.7	69.85	135	4.7	69.85	135					
	Multipole RF									200	525	850	200	525	850					
	Front Lens									4	7	10	4	7	10					
Ion detection (Injection Control)																				
	AGC Target Value (IT)															5000	10125	22500	34874	40000
	Max. Injection Time (IT)															50	57	75	93	100

Chapter 1

	AGC Target Value (FT)															10000	45147	130000	214900	250000
	Max. Injection Time (FT)															200	317	600	883	1000

Table 2. Overview of the 3-step optimization strategy. Parameters investigated and their corresponding levels.

Optimization regarding the ion detection region

The final step of the proposed optimization strategy was the investigation of the possible parameters affecting the ion trapping / detection region to the signal intensity of OE and HT. In this context, the AGC target and the maximum injection time values of the linear ion trap (IT) and the orbitrap, i.e. the maximum injection time FT (A), maximum injection time IT (B), AGC target value FT (C) and the AGC target value IT (D) were optimized applying the RSM methodology. No screening experiment was performed in this step as only a limited number of parameters is involved. This CCD plan ($\alpha = 1.41$) was composed of 30 experiments and the factors investigated, as well as their levels, are presented in Table 2. The generated results were again converted into response contour plots, which were used to locate the optimal operating conditions.

RESULTS AND DISCUSSION

As mentioned above, a three-step optimization study was performed. The ion generation region (ESI source), where 7 factors were investigated, the ion transmission region (ion optics) where 8 factors were investigated and finally the optimisation were applied to the ion detection region (injection control) where 4 factors were investigated. Thus the optimum values calculated after the ion generation step, were used as input for the ion transmission region optimization, and following, the optimized values were employed for the estimation of the optimal ion trapping / detection values. Finally, the overall improvement of the signal intensity of OE and HT was studied employing the whole set of optimum conditions.

Optimization regarding the ion generation region

In general, factorial designs are used to estimate the effects of the main factors and their interactions thereof and they are useful for screening a large number of factors in order to identify the most significant ones²⁹. Due to the fact that seven parameters affect the signal intensity to a MAX2-ESI source, a screening experiment was initially performed in order to determine the most influential ones. A full 2 level factorial design (all combinations of the seven factors were tested at the high and low level) would correspond to 128 experiments and would permit estimation of the individual factors main effects and all their interactions. Nevertheless, performing 128 experiments would not be practical because of the large number of experiments to be performed. Thus, the aim was to reduce the number of experiments needed to optimize the response variable under evaluation. For this purpose a 2⁷⁻¹ two level fractional factorial design was applied, in order to save 64 experiments out of 128 without losing important information. This design is known as a resolution VII design, where there is not confounding between main factors or between two-factor interactions. Since the main factors and the two factor interactions are not aliased and furthermore can be distinguished from higher-order interactions, this design represents a practical choice for a screening study. The alias list corresponding to the aforementioned design is shown in Table S1. Thus, 69 experiments (64 experiments plus 5 center points) were performed in a randomized order in order to avoid any systematic trend. Both responses have been analyzed using a natural logarithmic transformation due to the fact that the responses order of magnitude was larger than 10.

The abovementioned screening study showed that the most significant parameters affecting both OE and HT signal intensity are the tube lens voltage (F), the sheath gas flow (A), the capillary voltage (E) and the interaction EF (capillary voltage value \times tube lens

voltage value). The half normal probability plots (Figure 4a and Figure 4b) have been used for the identification of the effects that deviate from normal distribution and thus affect both the OE and HT signal.

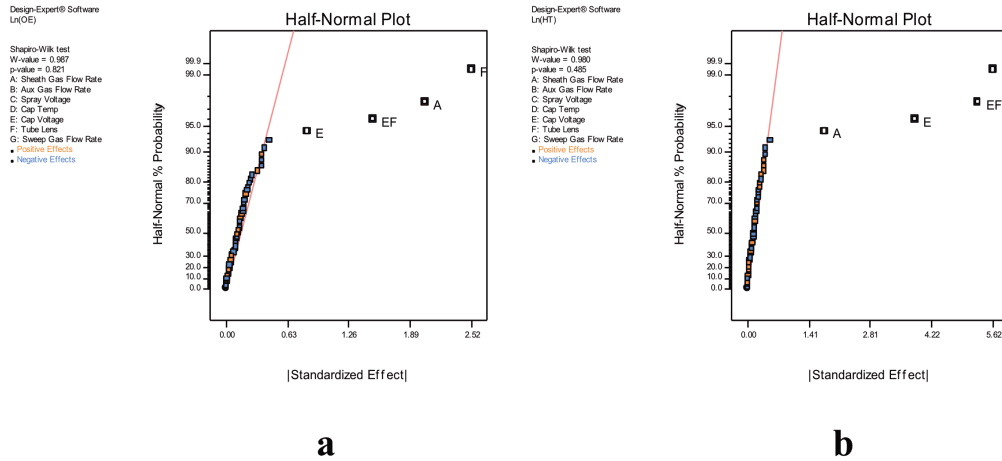


Figure 4. Half normal probability plots of the parameters used in the ion generation region: **(a)** OE and **(b)** HT. The points that deviate from the orange line are the significant parameters.

If there was no real influence of the factors, the estimated effects are expected to follow normal distribution regarding the signal, whereas effects having an influence would deviate from the normal distribution trend. In the half normal plots of OE and HT, the straight line represents the normal distribution whereas effects that are located out of the line are the significant effects because as they do not follow a normal distribution. In addition, the most significant effects are located to the far right of the graph and thus they can be considered as factors that are not part of the normal distribution and thus they influence the response. The selection of the important effects stops when all non-normal distributed factors have been recognized (points that are on or close to the line). It should be also noted that the half normal plot x-axis shows the standardized effect size. Taking the above information into account and also trying to locate the most important effects, the selection stopped at E for OE and at A for HT (Figure 4). It is evident that factor F (tube lens voltage) exhibits the most significant

effect on the signal of both analytes, whereas factor E (capillary voltage) is the less important for OE and factor A (sheath gas) for HT. Furthermore, the analysis of variance (ANOVA) exhibited that the 3 main effects [tube lens voltage (F), sheath gas flow (A), capillary voltage (E)] and the interaction EF are statistically significant at the 95% confidence limit with the p-values of the F-test lower than 0.05 in both OE and HT cases. All three and higher-level interactions were pooled into error in order to estimate the noise of the experiment. Moreover, the adequate precision value that has been calculated, is 23 for OE and 41 for HT, which is in large excess of the proposed value of 4, which shows that the signal to noise ratio is much smaller than the actual effect size. The final equations in terms of actual and coded factors for OE and HT are tabulated in Table 3. A series of diagnostic tests performed (e.g. residual normal probability test, residuals vs. run test) do not indicate the presence of any outliers for both OE and HT.

Compounds under study	Final equations in terms of coded factors	Final equations in terms of actual factors
OE	$\ln(\text{OE}) = +11.17 + 1.02 * A - 0.42 * E - 1.26 * F - 0.75 * E * F$	$\ln(\text{OE}) = +8.40 + 0.06 * \text{Sheath Gas Flow Rate} - 0.02 * \text{Cap Voltage} - 0.05 * \text{Tube Lens Voltage} - 6.61\text{E-}04 * \text{Cap Voltage} * \text{Tube Lens Voltage}$
HT	$\ln(\text{HT}) = +7.68 + 0.88 * A - 1.92 * E - 2.81 * F - 2.63 * E * F$	$\ln(\text{HT}) = +3.24 + 0.05 * \text{Sheath Gas Flow Rate} - 0.10 * \text{Cap Voltage} - 0.12 * \text{Tube Lens Voltage} - 2.31\text{E-}03 * \text{Cap Voltage} * \text{Tube Lens Voltage}$

Table 3. Final equations in terms of actual and coded factors for OE and HT, derived from the 2^{7-1} fractional factorial design performed in the ion generation region.

As mentioned above, an RSM model for the factors F, A and E was designed and applied in order to discover the optimum settings for the variables and to explore their performance over the whole experimental domain, including any interactions in respect to the signal intensity. The CCD approach was used for modeling the responses discovered by the previous 2 level factorial design. Such an experiment employs the standard 2k factorial points originating from the apexes of a cube, along with 2k axially-spaced points and it is

characterized by orthogonally and rotatability. In this case, three numerical factors [tube lens voltage (F), sheath gas flow (A) and capillary voltage (E)] were considered, resulting in the design of 19 experiments i.e. 8 cube, 6 star and 5 center points. The rest of the ESI source variables in the ion generation region namely the aux gas flow (B), spray voltage (C), capillary temperature (D) and sweep gas flow (G), that were deemed as non-significant during the factorial design, were set at their optimum levels (Table 1), auto-tuning values obtained by the embedded auto- tune and calibration routines of the instrument. Essentially, the signal does not change in a statistically significant manner by altering the instrument parameter values in the range studied. Two responses were assessed i.e. the OE and HT signal intensity. In order to select the appropriate model for OE and HT, ANOVA has been performed comparing models of different complexity e.g. linear vs. 2 factor interactions, 2 factor interactions vs. quadratic etc. and in all cases the quadratic model has found to be the most appropriate one. Furthermore a natural logarithmic transformation of the responses was applied, as their order of magnitude was more than 10. The response surface model showed that the A, E, F and a quadratic terms E^2 and F^2 were the most significant ones ($p < 0.05$) for OE while the A, E, F, the 2 level interaction EF and the quadratic terms E^2 and F^2 in the case of HT ($p < 0.05$). The final equations in terms of actual and coded factors for OE and HT are tabulated in Table 4.

Compounds under study	Final equations in terms of coded factors	Final equations in terms of actual factors
OE	$\ln(OE) = +13.45 - 0.09 * A + 0.10 * B - 0.22 * C - 2.08E-03 * A * B + 0.02 * A * C - 0.03 * B * C - 0.05 * A^2 - 0.11 * B^2 - 0.09 * C^2$	$\ln(OE) = +11.99 + 0.02 * \text{Sheath Gas Flow Rate} + 1.57E-03 * \text{Capillary Voltage} - 0.04 * \text{Tube Lens Voltage} - 4.62E-06 * \text{Sheath Gas Flow Rate} * \text{Capillary Voltage} + 7.66E-05 * \text{Sheath Gas Flow Rate} * \text{Tube Lens Voltage} - 4.85E-05 * \text{Capillary Voltage} * \text{Tube Lens Voltage} - 2.38E-04 * \text{Sheath Gas Flow Rate}^2 - 1.28E-04 * \text{Capillary Voltage}^2 - 3.08E-04 * \text{Tube Lens Voltage}^2$
HT	$\ln(HT) = +11.94 - 0.07 * A + 0.23 * B + 0.18 * C + 0.02 * A * B + 2.57E-03 * A * C$	$\ln(HT) = +10.93 + 8.41E-04 * \text{Sheath Gas Flow Rate} - 6.31E-03 * \text{Capillary Voltage} - 0.06 * \text{Tube Lens Voltage} + 3.33E-05 * \text{Sheath Gas Flow Rate} *$

$- 0.15 * B * C - 0.01 * A^2 - 0.37 * B^2 - 0.26 * C^2$	$\text{Capillary Voltage} + 9.78\text{E-}06 * \text{Sheath Gas Flow Rate} * \text{Tube Lens Voltage} - 2.92\text{E-}04 * \text{Capillary Voltage} * \text{Tube Lens Voltage} - 5.84\text{E-}05 * \text{Sheath Gas Flow Rate}^2 - 4.12\text{E-}04 * \text{Capillary Voltage}^2 - 8.53\text{E-}04 * \text{Tube Lens Voltage}^2$
---	---

Table 4. Final equations in terms of actual and coded factors for OE and HT, resulted from the RSM design performed in the ion generation region.

For the visual representation of the results, a series of different plots was constructed showing the impact of the proposed factors on the signal intensity. Figure 5 shows the one-factor plots of the \ln (OE signal intensity) (Figure 5a) and \ln (HT signal intensity) (Figure 5b) vs. the significant terms identified. These plots describe the actual changes in the response when varying the factor of interest, while retaining all other factors at their average values. Figure 5, shows the influence of the parameter F for OE and E for HT intensities respectively with all the other factors set at their mean values. It is apparent that factor F for OE and factor E for HT, are curved as expected due to the quadratic terms fitted (F^2 and E^2 respectively). The optimum OE intensity was obtained for a tube lens value (F) of -60 V whereas in the case of HT intensity, the optimum value was obtained for a tube capillary voltage (E) value of 30 V.

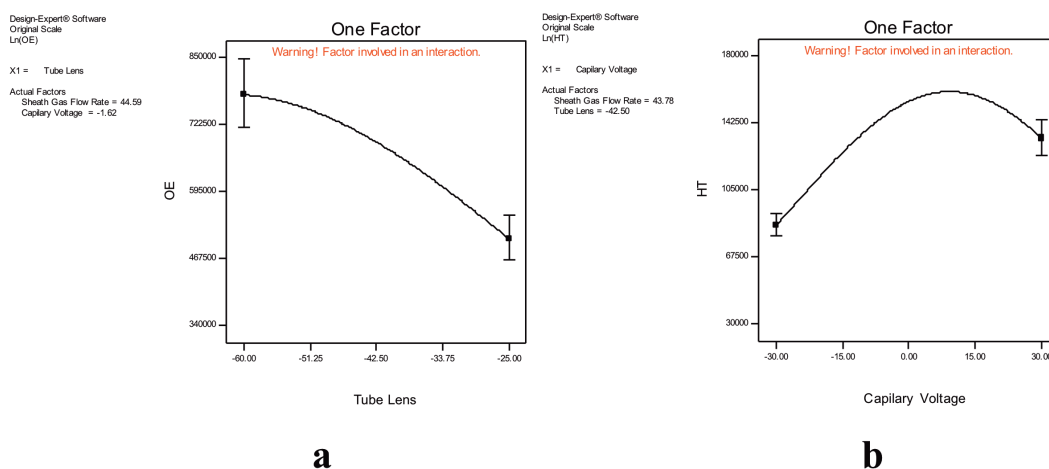


Figure 5. One-factor plots of the OE signal intensity vs. the tube lens **(a)** and the HT signal intensity vs. the capillary voltage **(b)** in the ion generation region.

In order to locate the optimal values of the above-examined parameters, numerical optimization was implemented for all the responses. The limits used for the sheath gas flow variable varied from 30 to 60 arb units, the capillary voltage ranged from -30 to 30 V, while the tube lens voltage limits were set between -60 and -25. The simplex algorithm has been applied in order to obtain the maximum of each 3-D space response. A set of solutions has been found affording for each analyte individually affording their optimal response, while the desirability function (the combined optimized response for both substances simultaneously with an optimal value and the value of combined desirability function for the simultaneous analysis of the two variables (OE and HT intensity), has reaches to 0.93. The obtained equation roots of the i.e. the values for which the OE and HT signal is maximized simultaneously are: sheath gas flow rate 30 arbitrary units, capillary voltage 11 V and tube lens voltage 46 V. All desirability surface contour plots, which are combinations of the tube lens voltage, the sheath gas flow rate and the capillary voltage, are shown in Figure 6.

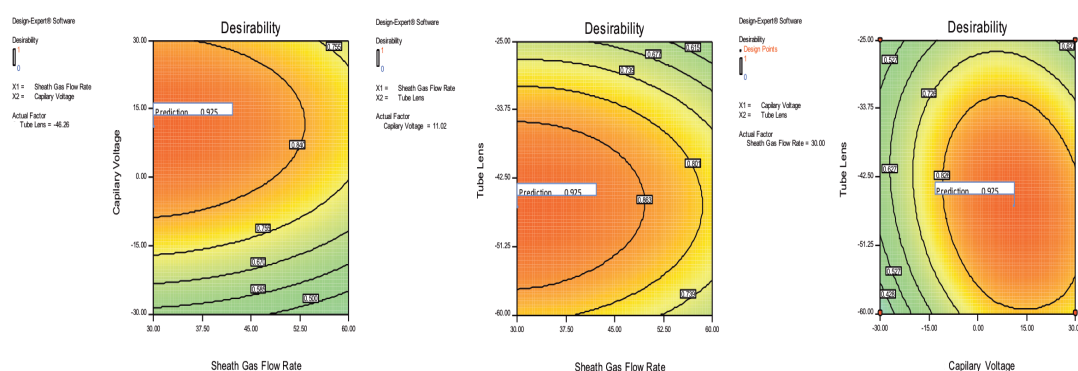


Figure 6. Desirability contour plots of sheath gas vs. capillary voltage, sheath gas vs. tube lens and capillary voltage vs. tube lens in the ion generation region.

Ion transmission region optimization

The second step was focused on the investigation of the possible ion transmission region effects to the OE and HT signal intensity. In this case also, a screening experiment was initially performed in order to determine the most influential parameters, due to large number (eight) of factors that can be altered in the ion transmission region. More specifically, the effects of the multipole 00 (A), lens 0 (B), multipole 0 (C), lens 1 (D), gate lens (E), multipole 1 (F), multipole RF (G) and front lens (H) were investigated applying a Plackett Burman factorial design in order to explore a n-dimensional experimental space using n+1 experiments. Such experiments are regularly used when the parameters that potentially contribute to the model are more than seven. This design was implemented via 17 experiments in total, encompassing five replications in the central point. The parameters used are the ones determined through the auto-tune procedure as set by the manufacturer. Each factor's values (low and high) are equally distributed throughout all the experiments. Generally the Plackett-Burman screening experiments indicate the main factors influencing the measured responses. As mentioned above, eleven parameters have been considered in the current screening design, eight of them being actual (A, B, C, D, E, F, G and H) whereas the rest three of them being “dummy” variables, which can either be as -1 or +1.

The experimental conditions for the Plackett-Burman screening design along with the values for the responses are summarized in Table 2. Statistical analysis was performed by ANOVA. The half normal probability plot, the Pareto chart as well as the ANOVA tables (Table S2 for OE and Table S3 for HT) for both analytes of interest, indicate that only one parameter i.e. the front lens (H), has been found to be as statistically significant with the p-value of the F-test being lower than 0.05 for the both examined analytes. Furthermore, the overall models have also found to be significant for both OE and HT respectively. The

adequate precision value was found to be 6 for OE and 5 for HT, which is again in excess of the proposed value of 4 which again showing as before that the signal to noise ratio is much smaller than the actual effect size. A series of diagnostic tests performed (residual normal probability test, residuals vs. run test, residuals vs. predicted test and predicted vs. actual test) do not indicate the presence of any outliers for both OE and HT.

The results from the Plackett Burman design reveal the existence of only one significant parameter in the ion transmission region that affects the signal intensity. The theoretical basis of the Plackett Burman experiments suggests that the main factors are potentially aliased with two way or higher interactions. Thus, in order to exclude the possibility of the front lens factor to be just an interaction, a 2 Level Factorial Design 2^{8-2} was applied subsequently, in order to recalculate the significant parameters in the ion transmission region. Such a design, which is a resolution V experiment, excludes the possibility of a main factor to be confounded with two level or higher interactions, as can be verified by its alias structure. Thus, 69 experiments (64 experiments plus 5 center points) were performed in random order. A square root transformation was again applied to the OE and HT intensity. The study verified that the front lens voltage (H) was the only significant parameter of the OE and HT models as can be shown to the ANOVA tables (Table S4 for OE and Table S5 for HT).

It is well known, that the correct ion kinetic energy is causing the ions to be efficiently trapped in the linear ion trap thus playing a crucial role to the signal intensity. Since the front lens voltage was found to be the only significant parameter in the ion transmission region, it is evident that this is the parameter that actually influences the ion kinetic energy and essentially the sensitivity of a hybrid LTQ orbitrap instrument.

Ion detection region optimization

The third step of the three-step optimization strategy is referred to the discovery of the possible effect of the ion detection region to the signal intensity of the tested compounds. For this purpose, an CCD based RSM design has been chosen, evaluating the following parameters: the maximum injection time FT (A), maximum injection time IT (B), AGC target value FT (C) and the AGC target value IT (D). Overall the design comprised of 30 experiments. The rest of the variables (too the ion generation and transmission regions) that were optimized in the previous steps, were set at their optimized values for the completion of these experiments. The ANOVA demonstrated that the linear model fits adequately for both responses, with the corresponding p-values < 0.0001. In these models the A and C were found to be the significant terms (p-value < 0.0001). A series of diagnostic tests performed (e.g. residual normal probability test, residuals vs. run test) do not indicate the presence of any outliers for both OE and HT. The final equations for the OE and HT signal intensities accordingly in terms of actual and coded factors are shown in Table 5.

Compounds under study	Final equations in terms of coded factors	Final equations in terms of actual factors
OE	OE = +17748.33 - 1955.75 * A - 294.12 * B - 2002.65 * C - 127.77 * D	OE = + 26445.44 - 6.91 * Max. Inject. Time (FT) - 16.64 * Max. Inject. Time (IT) - 0.02 * AGC Target value (FT) - 0.01 * AGC Target value (IT)
HT	HT = + 23476.20 - 2863.71 * A - 274.61 * B - 2972.78 * C - 132.18 * D	HT = + 35510.94 - 10.12 * Max. Inject. Time (FT) - 15.53 * Max. Inject. Time (IT) - 0.04 * AGC Target value (FT) - 0.01 * AGC Target value (IT)

Table 5. Final equations in terms of actual and coded factors for OE and HT, derived from the RSM design performed in the ion detection region.

In order to locate the optimum solutions for the equations found, numerical optimization has been performed with the goal of maximizing the desirability function the signal intensity, i.e. to maximize simultaneously the signal intensity of both OE and HT. The value of

combined desirability function for the simultaneous analysis of the two variables (OE and HT intensity) was found to be 0.96, which indicates that the signal intensity is actually maximized for both substances. Furthermore, the final optimal conditions for the ion detection were defined as follows: maximum injection time FT 317 ms and AGC target value FT 45000, as a compromise of its contradictory effects on signal intensity of OE and HT. A desirability surface plot vs. maximum injection time FT and AGC target value FT is shown in Figure 7. The respective recommended by the vendor target values were 500 ms for the FT's maximum injection time and 500000 for the FT's AGC target value FT⁴⁴. These results also show that the IT does not play any significant role in the signal intensity of the tested compounds.

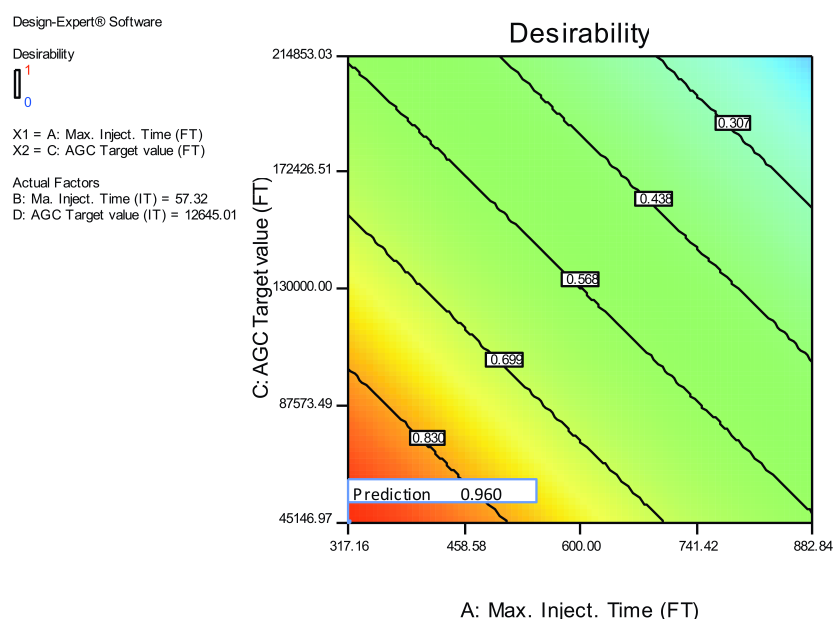


Figure 7. Desirability contour plot of the maximum injection time FT vs. the AGC target value FT in the ion detection region.

Overall, reducing the FT's AGC target value by a factor of 4.5 and the FT's maximum injection time by a factor of 1.6, the signal intensity increases by 24% in the case of OE and

46% in the case of HT. These results exhibit that the less the ions stored in the FT trap, the higher the signal intensity, a result that was not expected. A possible explanation could be a space charging effect that actually takes place in the orbitrap, leading to a significant loss of sensitivity. In addition, this effect seems to be mass independent. Furthermore, the elapsed scan time was nearly reduced to its half (1.02 msec vs. 0.48 msec). This significant improvement of the elapsed scan time is of high importance for various orbitrap-based applications, such as metabolomics, where high scan rate along with high sensitivity are required. Finally, after re-evaluation of the data in terms of mass accuracy for both OE and HT, it is evident that mass accuracy did not change during the aforementioned optimization (sub-ppm alteration).

Concluding the actual values in the ion generation are considered as substance specific, they depend on their structure and need to be optimized independently for every substance in order to gain the maximum sensitivity. In other words three parameters to the ion generation and one to the ion transmission region should be optimized. On the other hand the ion detection region parameter could be considered independent of the substances' structure thus the proposed values could be used as such. It should be noted that the proposed methodology has been applied to a metabolomics study with substantial increase of the sensitivity³⁸.

CONCLUSION

In order to investigate the parameters affecting the signal intensity on the LTQ-orbitrap Discovery, 19 parameters were examined taking into account the three involved regions, i.e. the ion generation, the ion transmission and the ion detection region. Using a three-step optimization strategy and employing various experimental designs, it was determined that the

tube lens voltage, the sheath gas flow and the capillary voltage parameters were the significant ones in the ion generation region (ESI source), whereas the other parameters of the ESI source and their interactions were not. Furthermore, the front lens was found to be the only parameter in the ion transmission region (ion optics) that has a key role. A last and very important observation concerning the ion detection region (injection control), is that the FT's AGC target value and the FT's maximum injection time were found to be the significant parameters in the examined region, and more specifically that the less ions stored in the FT trap, the higher the intensity reached. OE and HT were used as representative model compound for the completion of the experiments.

ACKNOWLEDGEMENTS

The present work was funded by SYNERGASIA 2009 PROGRAMME, which is co-funded by the European Regional Development Fund and National Resources (Project code: OSTEOPRO-1076). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No additional external funding received for this study.

REFERENCES

1. Hebert, A.S., Richards, A.L., Bailey, D.J., Ulbrich, A., Coughlin, E.E., Westphall, M.S., Coon, J.J.: The One Hour Yeast Proteome. *Mol. Cell. Proteomics* 13(1), 339–347 (2013).
2. Liu, S.Y., Zhang, R.L., Kang, H., Fan, Z.J., Du, Z.: Human liver tissue metabolic profiling research on hepatitis B virus-related hepatocellular carcinoma. *World J. Gastroenterol.* 19, 3423–3432 (2013).
3. Regal, P., Seijas, J.A., Cepeda, A., Fente, C.: Structure elucidation and HPLC-MS/MS determination of a potential biomarker for estradiol administration in cattle. *Anal. Bioanal. Chem.* 405, 9537–9546 (2013).
4. Tchoumtchoua, J., Njamen, D., Mbanya, J.C., Skaltsounis, A.L., Halabalaki, M.: Structure-oriented UHPLC-LTQ Orbitrap-based approach as a dereplication strategy for the identification of isoflavonoids from *Amphimas pterocarpoides* crude extract. *J. Mass Spectrom.* 48, 561–575 (2013).
5. Perry, R.H., Cooks, R.G., Noll, R.J.: Orbitrap mass spectrometry: instrumentation, ion motion and applications. *Mass Spectrom. Rev.* 27, 661–699 (2008).
6. Makarov, A., Denisov, E., Kholomeev, A., Balschun, W., Lange, O., Strupat, K., Horning, S.: Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Anal. Chem.* 78, 2113–2120 (2006).
7. Schwartz, J.C., Senko, M.W., Syka, J.E.P.: A two-dimensional quadrupole ion trap mass spectrometer. *J. Am. Soc. Mass Spectrom.* 13, 659–669 (2002).
8. Hardman, M., Makarov, A.A.: Interfacing the orbitrap mass analyzer to an electrospray ion source. *Anal. Chem.* 75, 1699–1705 (2003).
9. Soler, C., Hamilton, B., Furey, A., James, K.J., Mañes, J., Picó, Y.: Optimization of LC-MS/MS using triple quadrupole mass analyzer for the simultaneous analysis of carbosulfan and its main metabolites in oranges. *Anal. Chim. Acta.* 571, 1–11 (2006).
10. Székely, G., Henriques, B., Gil, M.: of experiments as a tool for LC-MS/MS method development for the trace analysis of the potentially genotoxic 4 dimethylaminopyridine impurity in glucocorticoids. *Biomed. Anal.* 70, 251–258 (2012).
11. Dejaegher, B., Heyden, Y. Vander: Experimental designs and their recent advances in set-up, data interpretation, and analytical applications. *J. Pharm. Biomed. Anal.* 56, 141–158 (2011).
12. Riter, L., Vitek, O.: Statistical design of experiments as a tool in mass spectrometry. *J. Mass Spectrom.* 40(5), 565–579 (2005).
13. Hibbert, D.B.: Experimental design in chromatography: a tutorial review. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* 910, 2–13 (2012).
14. Ha, P.K., Tong, B.C., Westra, W.H., Sanchez-Cespedes, M., Parrella, P., Zahurak, M., Sidransky, D., Califano, J.A.: Mitochondrial C-tract alteration in premalignant lesions of the head and neck: a marker for progression and clonal proliferation. *Clin. Cancer Res.* 8, 2260–2265 (2002).

15. Taamalli, A., Arráez-Román, D., Ibañez, E., Zarrouk, M., Segura-Carretero, A., Fernández-Gutiérrez, A.: Optimization of microwave-assisted extraction for the characterization of olive leaf phenolic compounds by using HPLC-ESI-TOF-MS/IT-MS(2). *J. Agric. Food Chem.* 60, 791–798 (2012).
16. Annadurai, G., Juang, R.S., Lee, D.J.: Factor optimization for phenol removal using activated carbon immobilized with *Pseudomonas putida*. *J. Environ. Sci. Health. A. Tox. Hazard. Subst. Environ. Eng.* 37, 149–161 (2002).
17. Sarkar, J., Khan, G.G., Basumallick, A.: The microscopic origin of self-organized nanostripe pattern formation on an electropolished aluminium surface. *Nanotechnology.* 20, 095604 (2009).
18. Xynos, N., Papaefstathiou, G., Gikas, E., Argyropoulou, A., Aligiannis, N., Skaltsounis, A.L.: Design optimization study of the extraction of olive leaves performed with pressurized liquid extraction using response surface methodology. *Sep. Purif. Technol.* 122, 323–330 (2014).
19. Maragou, N.C., Thomaidis, N.S., Koupparis, M.: Optimization and comparison of ESI and APCI LC-MS/MS methods: a case study of Irgarol 1051, Diuron, and their degradation products in environmental samples. *J. Am. Soc. Mass Spectrom.* 22, 1826–1838 (2011).
20. Kostić, N., Dotsikas, Y., Malenović, A., Stojanović, B.J., Rakić, T., Ivanović, D., Medenica, M.: Stepwise optimization approach for improving LC-MS/MS analysis of zwitterionic antiepileptic drugs with implementation of experimental design. *J. Mass Spectrom.* 48, 875–884 (2013).
21. Dillon, L.A., Stone, V.N., Croasdell, L.A., Fielden, P.R., Goddard, N.J., Thomas, C.L.P.: Optimisation of secondary electrospray ionisation (SESI) for the trace determination of gas-phase volatile organic compounds. *Analyst.* 135, 306–314 (2010).
22. Gruending, T., Guilhaus, M., Barner-Kowollik, C.: Design of Experiment (DoE) as a Tool for the Optimization of Source Conditions in SEC-ESI-MS of Functional Synthetic Polymers Synthesized via ATRP. *Macromol. Rapid Commun.* 30, 589–597 (2009).
23. Bylund, D., Danielsson, R., Markides, K.: Optimization strategy for liquid chromatography–electrospray ionization mass spectrometry methods. *Analyst.* 125, 1970–1976 (2000).
24. Zachariadis, G., Rosenberg, E.: Doehlert-type experimental design in optimization of a hybrid electrospray ionization ion trap time-of-flight mass spectrometry technique for glutathione determination. *J. Mass Spectrom.* 27, 489–499 (2013).
25. Moberg, M., Markides, K.E., Bylund, D.: Multi-parameter investigation of tandem mass spectrometry in a linear ion trap using response surface modelling. *J. Mass Spectrom.* 40, 317–324 (2005).
26. Kiyonami, R., Schoen, A., Prakash, A., Peterman, S., Zabrouskov, V., Picotti, P., Aebersold, R., Huhmer, A., Domon, B.: Increased selectivity, analytical precision, and throughput in targeted proteomics. *Mol. Cell. Proteomics.* 10(2), M110.002931 (2011).
27. Andrews, G.L., Dean, R. a, Hawkrige, A.M., Muddiman, D.C.: Improving proteome coverage on a LTQ-Orbitrap using design of experiments. *J. Am. Soc. Mass Spectrom.* 22, 773–783 (2011).
28. Randall, S.M., Cardasis, H.L., Muddiman, D.C.: Factorial experimental designs elucidate significant variables affecting data acquisition on a quadrupole orbitrap mass spectrometer. *J. Am. Soc. Mass Spectrom.* 24, 1501–1512 (2013).

29. Laures, A., Wolff, J.C.: Investigation into the factors affecting accuracy of mass measurements on a time-of-flight mass spectrometer using Design of Experiment. *Rapid Commun. Mass Spectrom.* 21(4), 529–535 (2007).
30. Omar, S.H.: Cardioprotective and neuroprotective roles of oleuropein in olive. *Saudi Pharm. J.* 18(3), 111–121 (2010).
31. Andreadou, I., Iliodromitis, E.K., Mikros, E., Constantinou, M., Agalias, A., Magiatis, P., Skaltsounis, A.L., Kamber, E., Tsantili-Kakoulidou, A., Kremastinos, D.T.: The Olive Constituent Oleuropein Exhibits Anti-Ischemic, Antioxidative, and Hypolipidemic Effects in Anesthetized Rabbits. *J. Nutr.* 136, 2213–2219 (2006).
32. Kontogianni, V.G., Gerothanassis, I.P.: Phenolic compounds and antioxidant activity of olive leaf extracts. *Nat. Prod. Res.* 26, 186–189 (2012).
33. Quirantes-Piné, R., Herranz-López, M., Funes, L., Borrás-Linares, I., Micol, V., Segura-Carretero, A., Fernández-Gutiérrez, A.: Phenylpropanoids and their metabolites are the major compounds responsible for blood-cell protection against oxidative stress after administration of *Lippia citriodora* in rats. *Phytomedicine*. 20, 1112–1118 (2013).
34. Tutino, V., Caruso, M.G., Messa, C., Perri, E., Notarnicola, M.: Antiproliferative, antioxidant and anti-inflammatory effects of hydroxytyrosol on human hepatoma HepG2 and Hep3B cell lines. *Anticancer Res.* 32, 5371–5377 (2012).
35. Reyes, J.J., De La Cruz, J.P., Muñoz-Marin, J., Guerrero, A., Lopez-Villodres, J.A., Madrona, A., Espartero, J.L., Gonzalez-Correa, J.A.: Antiplatelet effect of new lipophilic hydroxytyrosol alkyl ether derivatives in human blood. *Eur. J. Nutr.* 52, 591–599 (2013).
36. Dierkes, G., Krieger, S., Du, R., Bongartz, A., Schmitz, O.J., Hayen, H., Dück, R.: High-performance liquid chromatography-mass spectrometry profiling of phenolic compounds for evaluation of olive oil bitterness and pungency. *J. Agric. Food Chem.* 60, 7597–7606 (2012).
37. Zoidou, E., Melliou, E.: Identification of Throuba Thassos, a traditional Greek table olive variety, as a nutritional rich source of oleuropein. *J. Agric. Food Chem.* 58, 46–50 (2010).
38. Lemonakis, N., Gikas, E., Halabalaki, M., Skaltsounis, A.L.: Development and Validation of a Combined Methodology for Assessing the Total Quality Control of Herbal Medicinal Products – Application to Oleuropein Preparations. *PLoS One*. 21 8(10), e78277 (2013).
39. Lozano-Sánchez, J.: Wastes generated during the storage of extra virgin olive oil as a natural source of phenolic compounds. *J. Agric. Food. Chem.* 59(21), 11491–11500 (2011).
40. Bazoti, F.N., Gikas, E., Tsaibopoulos, A.: Simultaneous quantification of oleuropein and its metabolites in rat plasma by liquid chromatography electrospray ionization tandem mass spectrometry. *Biomed. Chromatogr.* 24, 506–515 (2010).
41. Ortega-García, F., Peragón, J.: HPLC analysis of oleuropein, hydroxytyrosol, and tyrosol in stems and roots of *Olea europaea* L. cv. Picual during ripening. *J. Sci. Food Agric.* 90, 2295–2300 (2010).
42. Tan, H.W., Tuck, K.L., Stupans, I., Hayball, P.J.: Simultaneous determination of oleuropein and hydroxytyrosol in rat plasma using liquid chromatography with fluorescence detection. *J. Chromatogr. B.* 785, 187–191 (2003).
43. Agalias, A., Magiatis, P., Skaltsounis, A.L., Mikros, E., Tsaibopoulos, A., Gikas, E., Spanos, I., Manios, T.: A new process for the management of olive oil mill waste water and recovery of natural antioxidants. *J. Agric. Food Chem.* 55, 2671–2676 (2007).

44. LTQ Orbitrap Discovery TM. Getting Started. Tune Plus 2.4. Revision A - 122 5860.
(http://www.unifr.ch/inph/vclab/assets/files/Software/LTQ-Orbitrap-Discovery_Getting_Started.pdf).

SUPPORTING INFORMATION

Estimated Terms	Aliased Terms
Intercept	Intercept + EFG
[A]	A - EFG
[B]	B - CFG
[C]	C - CFG
[D]	D + EFG
[E]	E - EFG
[F]	F + EFG
[G]	G + EFG
[AB]	AB + CFG
[AC]	AC + CFG
[AD]	AD - EFG
[AE]	AE + EFG
[AF]	AF - EFG
[AG]	AG - EFG
[BC]	BC + EFG
[BD]	BD - CFG
[BE]	BE + CFG
[BF]	BF - CFG
[BG]	BG - CFG
[CD]	CD - CFG
[CE]	CE + CFG
[CF]	CF - CFG
[CG]	CG - CFG
[DE]	DE - EFG
[DF]	DF + EFG
[DG]	DG + EFG
[EF]	EF - EFG
[EG]	EG - EFG
[FG]	FG + EFG
[ABC]	ABC - EFG
[ABD]	ABD + CFG
[ABE]	ABE - CFG
[ABF]	ABF + CFG
[ABG]	ABG + CFG
[ACD]	ACD + CFG
[ACE]	ACE - CFG
[ACF]	ACF + CFG
[ACG]	ACG + CFG
[ADE]	ADE + EFG
[ADF]	ADF - EFG
[ADG]	ADG - EFG

Estimated Terms	Aliased Terms
Intercept	Intercept + EFG
[AEF]	AEF + EFG
[AEG]	AEG + EFG
[AFG]	AFG - EFG
[BCD]	BCD + EFG
[BCE]	BCE - EFG
[BCF]	BCF + EFG
[BCG]	BCG + EFG
[BDE]	BDE + CFG
[BDF]	BDF - CFG
[BDG]	BDG - CFG
[BEF]	BEF + CFG
[BEG]	BEG + CFG
[BFG]	BFG - CFG
[CDE]	CDE + CFG
[CDF]	CDF - CFG
[CDG]	CDG - CFG
[CEF]	CEF + CFG
[CEG]	CEG + CFG
[DEF]	DEF - EFG
[DEG]	DEG - EFG
[DFG]	DFG + EFG

Table S1. Alias list of the 2^{7-1} fractional factorial design performed in the ion generation region.

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob>F	Observation
Model	10.76	1	10.76	14.12	0.0021	<i>significant</i>
H-Front Lens	10.76	1	10.76	14.12	0.0021	
Curvature	1.36	1	1.36	1.78	0.2031	<i>not significant</i>
Residual	10.67	14	0.76			
Lack of Fit	10.67	10	1.07	4340.37	< 0.0001	<i>significant</i>
Pure Error	9,84E-01	4	2,46E-01			
Cor Total	22.79	16				

Transform: Natural log, Constant: 0

Table S2. ANOVA table of OE Plackett Burman design in the ion transmission region.

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob>F	Observation
Model	4.26	1	4.26	8.57	0.0110	<i>significant</i>
H-Front Lens	4.26	1	4.26	8.57	0.0110	
Curvature	1.58	1	1.58	3.17	0.0965	<i>not significant</i>
Residual	6.96	14	0.50			
Lack of Fit	6.95	10	0.70	1910.48	< 0.0001	<i>significant</i>
Pure Error	1,46E+00	4	3,64E-01			
Cor Total	12.79	16				

Transform: Natural log, Constant: 0

Table S3. ANOVA table of HT Plackett Burman design in the ion transmission region.

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob>F	Observation
Model	1,88E+09	1	1,88E+09	79.13	< 0.0001	<i>significant</i>
H-Front Lens	1,88E+09	1	1,88E+09	79.13	< 0.0001	
Curvature	9688.09	1	9688.09	0.41	0.5258	<i>not significant</i>
Residual	1,57E+09	66	23814.97			
Lack of Fit	1,57E+09	62	25348.52	563.37	< 0.0001	<i>significant</i>
Pure Error	179.98	4	44.99			
Cor Total	3,47E+09	68				

Transform: Natural log, Constant: 0

Table S4. ANOVA table of OE 2⁸⁻² 2-Level Factorial Design in the ion transmission region.

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob>F	Observation
Model	1,16E+08	1	1,16E+08	40.63	< 0.0001	<i>significant</i>
H-Front Lens	1,16E+08	1	1,16E+08	40.63	< 0.0001	
Curvature	7874.49	1	7874.49	2.75	0.1019	<i>not significant</i>
Residual	1,89E+08	66	2862.27			
Lack of Fit	1,89E+08	62	3045.28	119.03	0.0001	<i>significant</i>
Pure Error	102.33	4	25.58			
Cor Total	3,13E+08	68				

Transform: Natural log, Constant: 0

Table S5. ANOVA table of HT 2⁸⁻² 2-Level Factorial Design in the ion transmission region.

ANNEX

Optimization regarding the ion generation region

A series of diagnostic tests performed in the screening design do not indicate the presence of any outliers for both OE and HT. Thus, residual normal probability test (Figure AN1), residuals vs. run test (Figure AN2), box-cox for power transforms test (Figure AN3) for OE and the corresponding ones for HT (Figure AN4, Figure AN5 and Figure AN6)

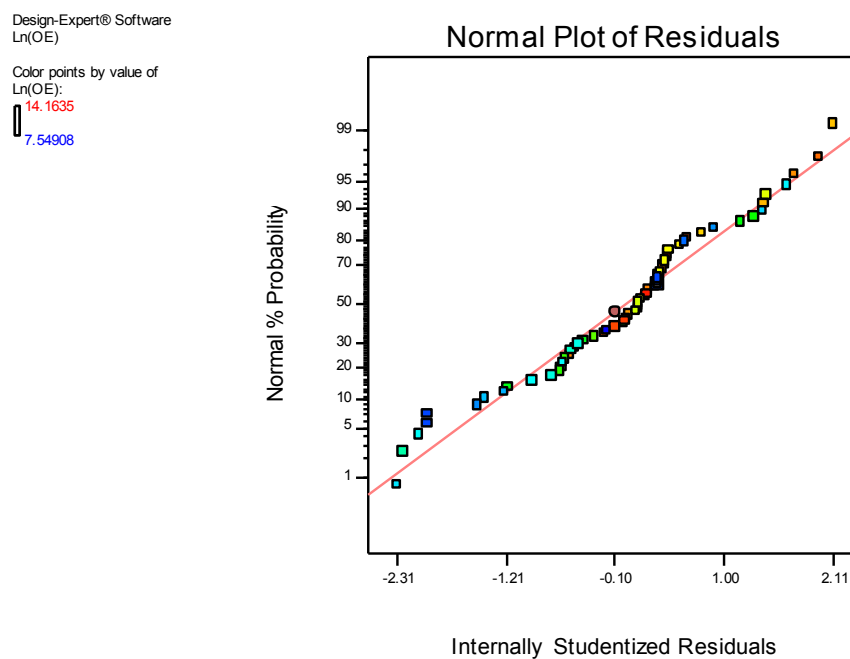


Figure AN1. Residual normal probability plot of OE

Design-Expert® Software
Ln(OE)

Color points by value of
Ln(OE):
14.1635
7.54908

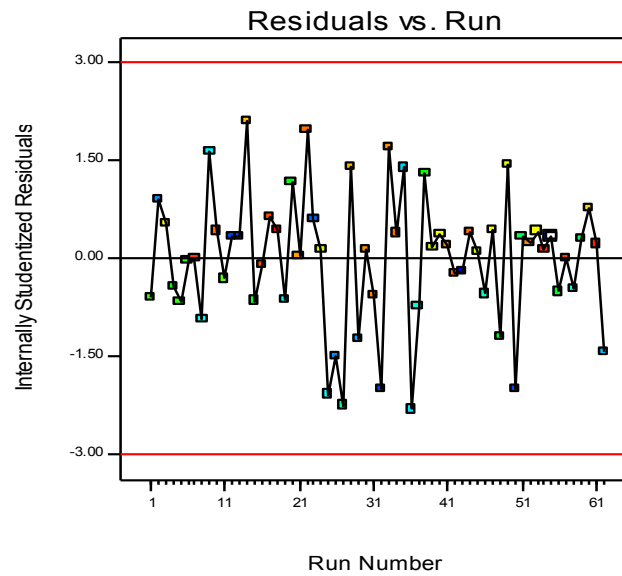


Figure AN2. Residuals vs. run plot of OE

Design-Expert® Software
Ln(OE)

Lambda
Current = 0
Best = 0.05
Low C.I. = -0.08
High C.I. = 0.18

Recommend transform:
Log
(Lambda = 0)

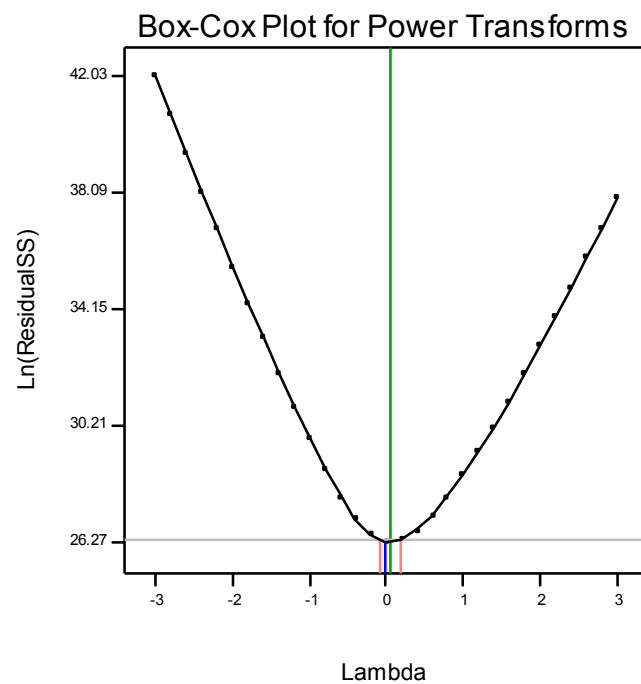


Figure AN3. Box-cox plot of OE

Design-Expert® Software
Ln(HT)

Color points by value of
Ln(HT):

12.7358
0

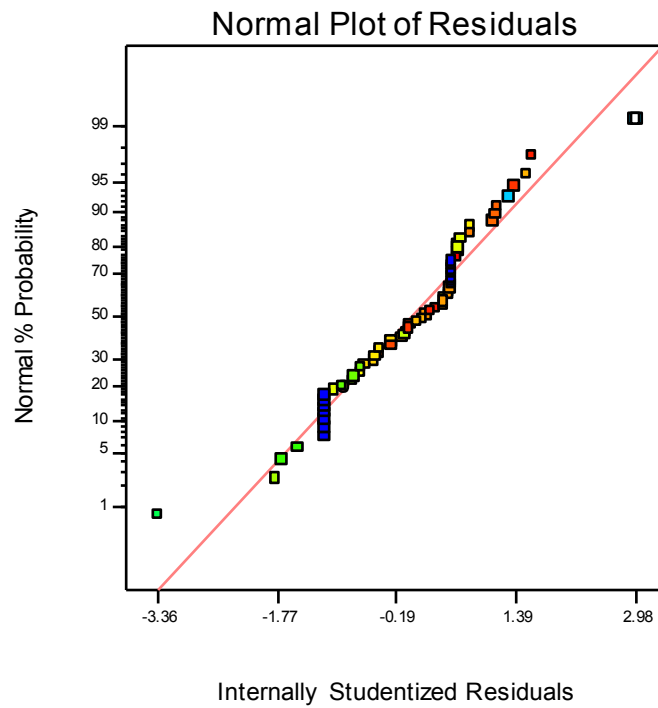


Figure AN4. Residual normal probability plot of HT

Design-Expert® Software
Ln(HT)

Color points by value of
Ln(HT):

12.7358
0

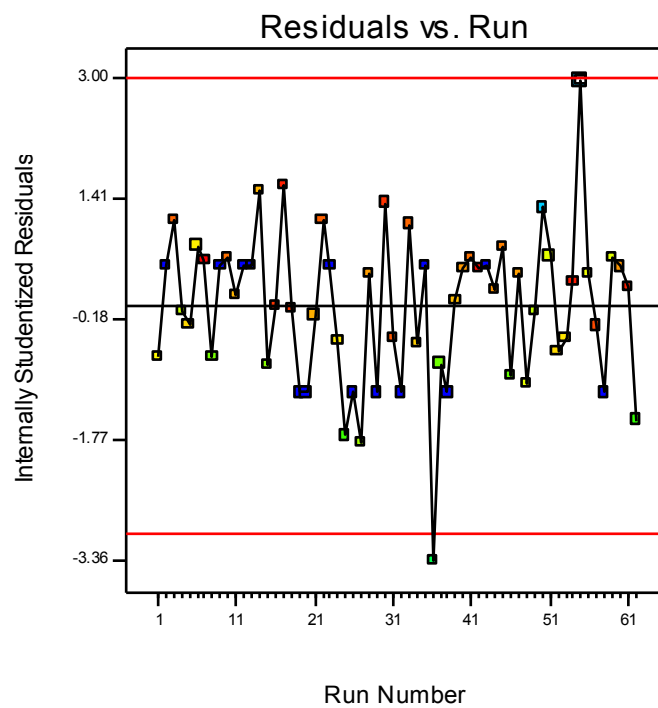


Figure AN5. Residuals vs. run plot of HT

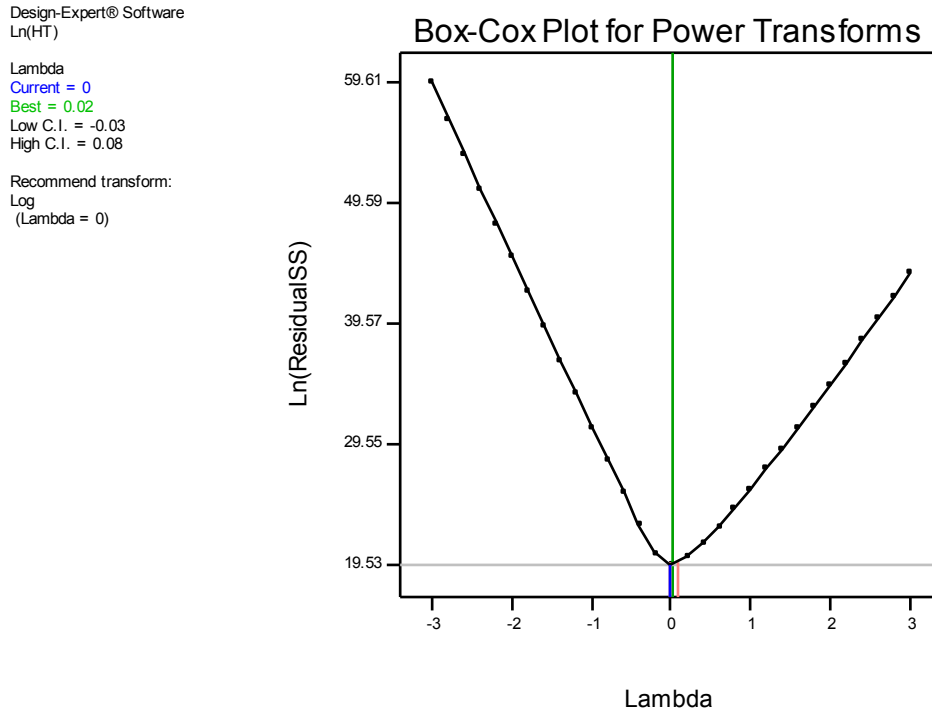


Figure AN6. Box-cox plot of HT

The response surface plots in Figure AN7 summarize the influence of two factors, the capillary voltage and the tube lens over the ln (OE signal intensity) and the ln (HT signal intensity) respectively, with the third factor, namely the sheath gas flow, set at its mean value.

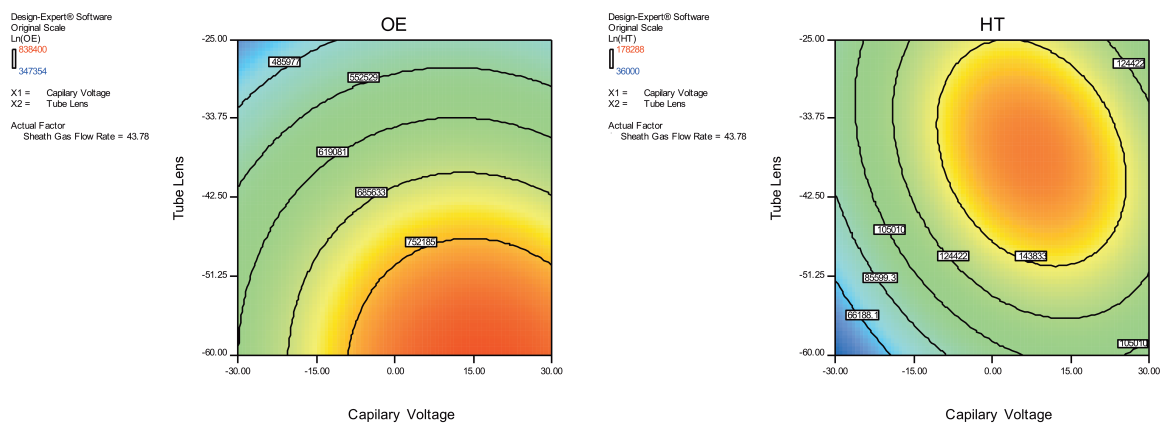


Figure AN7. Response surface plots of OE and HT. The influence of the capillary voltage and tube lens over the ln (OE signal intensity) and the ln (HT signal intensity) respectively, with sheath gas flow at its mean value, is shown.

2

Development and validation of a combined methodology for assessing the total quality control of herbal medicinal products – Application to oleuropein preparations

This chapter is published as: Lemonakis N, Gikas E, Halabalaki M, Skaltsounis AL

PLoS One. 2013 Oct 21;8(10):e78277. doi: 10.1371/journal.pone.0078277.

ABSTRACT

Oleuropein (OE) is a secoiridoid glycoside, which occurs mostly in the Oleaceae family presenting several pharmacological properties, including antioxidant, cardio-protective, anti-atherogenic effects etc. Based on these findings OE is commercially available, as Herbal Medicinal Product (HMP), claimed for its antioxidant effects. As there are general provisions of the medicine regulating bodies e.g. European Medicines Agency, the quality of the HMP's must always be demonstrated. Therefore, a novel LC-MS methodology was developed and validated for the simultaneous quantification of OE and its main degradation product, hydroxytyrosol (HT), for the relevant OE claimed HMP's. The internal standard (IS) methodology was employed and separation of OE, HT and IS was achieved on a C18 Fused Core column with 3.1 min overall run time employing the SIM method for the analytical signal acquisition. The method was validated according to the International Conference on Harmonisation requirements and the results show adequate linearity ($r^2 > 0.99$) over a wide concentration range [0.1–15 $\mu\text{g/mL}$ ($n=12$)] and a LLOQ value of 0.1 $\mu\text{g/mL}$, for both OE and HT. Furthermore, as it would be beneficial to control the quality taking into account all the substances of the OE claimed HMP's; a metabolomics-like approach has been developed and applied for the total quality control of the different preparations employing UHPLC-HRMS-multivariate analysis (MVA). Four OE-claimed commercial HMP's have been randomly selected and MVA similarity-based measurements were performed. The results showed that the examined samples could also be differentiated as evidenced according to their scores plot. Batch to batch reproducibility between the samples of the same brand has also been determined and found to be acceptable. Overall, the developed combined

methodology has been found to be an efficient tool for the monitoring of the HMP's total quality. Only one OE HMP has been found to be consistent to its label claim.

INTRODUCTION

Oleuropein (OE) is a natural secoiridoid glycoside, occurring mainly in the *Olea* genus of the Oleaceae family and it is the most well studied phenolic compound in olive cultivars¹⁻³. OE is one of the many individual components of the Mediterranean diet and it has been proved to be exhibiting protective activity against an array of common chronic pathological conditions⁴⁻⁵. The molecule has been shown to exert several pharmacological properties, including antioxidant⁶, antimicrobial⁷, cardioprotective⁸⁻⁹, anti-ischemic¹⁰, antiatherogenic¹¹, anti-inflammatory¹²⁻¹³, antidiabetic¹⁴, antiproliferative in prostate cell lines¹⁵ etc. Thus, several formulations of OE are commercially available in many countries (approximately 23), as food supplements or herbal medicinal products (HMP)¹⁶. Its main degradation product, which is also a natural occurring substance in olive products, is hydroxytyrosol (HT), which also exhibits several interesting biological properties¹⁷.

As HMP is defined “any medicinal product, exclusively containing as active ingredients one or more herbal substances or one or more herbal preparations, or one or more such herbal substances in combination with one or more such herbal preparations”¹⁸. Furthermore, as stated by the European Medicines Agency (EMA), “irrespective of the regulatory pathway to access the market, the quality of the herbal medicinal product must always be demonstrated”¹⁹. HMP's are products of high commercial interest as proved by the wide increase of their use worldwide but due to

their high profitability a series of concerns are rising regarding their quality. Thus new reliable and fast quality control methodologies are required for the assessment of HMP's quality, in order to meet the appropriate standards required for the protection and promotion of public health. It should be noted that the quality control of HMP's represents a demanding task, as they are usually complex mixtures, extracts or enriched extracts, often containing several hundreds of minor constituents that are impossible to be accurately quantified one by one, taking into account the current status of instrumental capabilities. It is also widely known that synergistic interactions are of vital importance in phytomedicines²⁰, plant extracts and HMP's, with one or a few ingredients determining their therapeutic effects but the synergy of all the substances yielding the optimum therapeutic efficacy or possible toxicity on the other hand. In other words, the beneficial effect of a plant extract or HMP cannot be attributed to one single substance, but is the result of the synergistic interplay between many extract's constituents, usually called the multicomponent system. The aforementioned biological activities represent the health claims registered to the corresponding regulatory bodies. Thus it is evident that the quality control of HMP's as a whole represents an extremely demanding task, as according to the aforementioned concepts, the accurate quantification of each one of several hundreds of minor constituents cannot be performed. On the other hand the measurement of the quality of HMP's taking into account as many constituents as possible is crucial and at least the batch to batch reproducibility of such products should be demonstrated, as it ensures the same biological activity (due to both the main constituent and the synergistic effects) for the patients receiving the selected commercial formulation. Overall the quality control of the HMP's as a whole should also be controlled and reported. In the current work the term "total quality" reflects not only the accurate

quantitation of the main bioactive substance (OE) but also the simultaneous evaluation of the total phytochemical profile of the investigated OE claimed HMP's in each formulation.

Many analytical methods have been developed for the determination of OE in various matrices such as olive oil²¹, table olives²², byproducts²³ and biological fluids²⁴. Thus, OE has been analysed employing HPLC-UV²⁵, HPLC-fluorescence²⁶ and HPLC-MS/MS^{23,27}. Other methods besides liquid chromatography have also been employed such as GC-MS²⁸, capillary electrophoresis²⁹ and FT-IR³⁰. However no analytical methodologies have been developed and validated for the simultaneous quantification of OE and its main degradation product HT in HMP's. In order such methodologies applied in industrial environments should exhibit high turnaround time, with minimal sample preparation and fast analysis duration. For this purpose, a novel, rapid and high throughput LC-MS analytical methodology was developed and validated according to the International Conference on Harmonisation (ICH) requirements³¹, for the simultaneous quantification of OE and HT in a selected range of OE commercial HMP's. This methodology allows for the targeted determination of the two analytes in complex mixtures of HMPS's. As the modern chromatographic trend goes towards fast analyses, the Fused Core column technology³² has been employed. Such columns can achieve ultra fast separations even under HPLC conditions³³. In addition, emphasis was given to the isolation of the matrix of the HMP's, as it was necessary for the correct validation of the analytical methodology. As already referred, HMP's are highly complex mixtures and thus it is difficult to simulate them artificially, in order to accurately estimate various analytical parameters, such as the matrix effect and the selectivity. For this purpose isolation of the matrix of the HMP's was performed, using an adsorbent XAD – 7HP resin and the

former has been used throughout the validation procedure, e.g. in the same way plasma^{34,35} or urine³⁴ are used in the bioanalytical method validation. The first aim of the current work was to develop and validate a rapid, sensitive and accurate methodology, to quantify simultaneously OE and HT in commercial HMP's and apply this methodology to selected OE containing HMP's in order to assess their conformity according to their label claim.

Although LC-MS based approaches are widely employed for the targeted quantification of the major bioactive components of the corresponding HMP's, similar approaches for the evaluation of HMP's regarding their profiling as well as their total quality control are less common, partly because they are more demanding in terms of instrumentation and data evaluation. It should be noted that hyphenated high resolution techniques such as UHPLC-HRMS (orbitrap or qQToF) are commonly recommended for the profiling of extracts and therefore for the assessment of HMP's quality control, because they offer high separation ability and simultaneously the ability to record accurate mass data of all the contained HMP's substances³⁶. Such analytical methodologies in combination with Multivariate analysis (MVA), such as principal component analysis (PCA) or partial least square analysis (PLS), enable the comprehensive assessment of complex HMP's and sample classification of diverse status, origin or quality in samples. This approach uses the same principles as metabolomics, which is a branch of science that concerns the total metabolome of integrated biological systems as well as the dynamic responses to the alterations of endogenous and/or exogenous factors³⁷. A second aim of the present study is to develop a metabolomics-like approach, combining UHPLC-(-ESI)-HRMS and MVA for studying the "total quality" characteristics of the selected HMP's in

terms of their minor component uniformity with the aid to assess their batch to batch reproducibility.

Overall, in the current study, an HPLC-(-ESI)-MS and an UHPLC-(-ESI)-HRMS approach were developed in order to evaluate the holistic qualities of commercial OE claimed HMP's, exploring both the quantity of the bioactive compound (OE) in different commercial OE claimed product and the similarity of their phytochemical profiles.

MATERIALS AND METHODS

Chemicals and Reagents

The analytical reference standards, OE and HT (Figure 1), were isolated from olive leaves and olive mill wastewater respectively based on previously described procedures^{10, 38}. No specific permissions were required as the examined plant is an endemic species widely spread in the Mediterranean region, as well as in Australia, the American continent, Africa and Asia. Furthermore *Olea europea* is not included in any list of endangered or protected species.

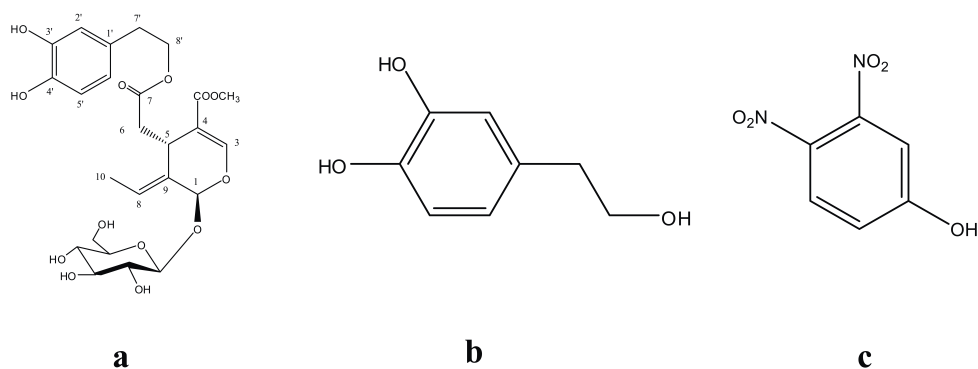


Figure 1. Structures of OE (a), HT (b) and the IS (2,4-dinitrophenol) (c)

Their purity was found to be more than 98% (HPLC-UV) and their structures were identified by High Resolution Mass Spectrometry (HRMS) and 1 & 2D NMR. The internal standard (IS), 2,4-dinitrophenol (Figure 1), was purchased from Sigma – Aldrich (Greece). All solvents used in this study were of LC-MS grade. Acetonitrile (ACN), methanol (MeOH), water and formic acid, were purchased from Fluka/Riedel-de Haën (Switzerland). 2,2-Diphenyl-1-picrylhydrazyl (DPPH) for DPPH experiments was purchased from Sigma – Aldrich (Greece). Absolute ethanol for DPPH dissolving was purchased from Sigma – Aldrich (Greece). The Amberlite XAD-7HP resin used in the absorption experiment was supplied by Sigma – Aldrich (Greece).

Samples

Five samples (a, b, c, d, e) from eight different production batches (A, B, C, D, E, F, G, H) over a year of Brand1 (B1) were analyzed. B1A batch, which was a special preparation for further biological experiments, was containing a higher amount of OE. All the other samples have been provided by the local market. Thus five samples (a, b, c, d, e) from each of three different vendors (two European and one Chinese) of OE claimed HMP's have been acquired from the local market. [Brand2 (B2), Brand3 (B3), Brand4 (B4)]. A list of the samples used for the analysis is tabulated in Table 1.

Brands (B)	Brand1 (B1)	Brand2 (B2)	Brand3 (B3)	Brand4 (B4)
Batches (n=8)	B1Aa...e*	B2Aa...e	B3Aa...e	B4Aa...e
	B1Ba...e	-	-	-
	B1Ca...e	-	-	-
	B1Da...e	-	-	-
	B1Ea...e	-	-	-
	B1Fa...e	-	-	-
	B1Ga...e	-	-	-
	B1Ha...e	-	-	-

*The following coding has been used. BxYz, x= brand, Y= batch, z= sample number

Table 1. A summary of tested samples.

Equipment

Nuclear Magnetic Resonance (NMR) spectra were recorded with the aid of a Bruker Avance III spectrometer operating at 600.11 MHz (Bruker Biospin GmbH, Reinsteten, Germany). HPLC analyses for the quantification study were performed employing a Surveyor system (Thermo Scientific, Germany) equipped with a binary pump, an autosampler, an online vacuum degasser and a temperature-controlled column compartment. Mass spectrometry for the OE and HT quantitation was performed on a single quadrupole MSQ mass spectrometer (Thermo Scientific, Germany) connected to the HPLC system. UHPLC for the quality control - multivariate analysis study was performed employing an Accela system (Thermo Scientific, Germany) equipped with a binary pump, an autosampler, an online vacuum degasser and a temperature-controlled column compartment. High-resolution mass spectrometry was performed on a hybrid LTQ Orbitrap Discovery XL mass spectrometer (Thermo Scientific, Germany). Centrifuging was performed by a Mikro 200R centrifuge (Hettich Lab Technology, Germany). Evaporation of the samples was performed with the aid of a Buchi evaporator system consisting of a rotavapor R-210, a vacuum pump V-700 and a vacuum controller V-850 (Buchi, Switzerland). A micro plate Reader, Infinite M200[®] PRO (Tecan, Austria) was used for the evaluation of the DPPH activity of the capsules.

Simultaneous quantification of OE and HT in HMP's – Validation of the analytical methodology

LC-MS Conditions. An Ascentis Express RP-C18 Fused Core (100 x 2.1 mm, 2.7 μ m) column (Supelco Analytical, Germany) was used at a flow rate of 0.3 mL/min for the chromatographic separation. The mobile phase consisted of: solvents A - water, 0.1% formic acid and B - ACN. A gradient methodology (total run time of 3.1 min) was used for the separation of OE, HT and the IS as follows: 0 to 0.5 min: 90% A: 10% B, 0.5 to 0.6 min: 70% A: 30% B, 0.6 to 1.5 min: 70% A: 30% B, 1.5 to 2.5 min: 30% A: 70% B, 2.5 to 3.1 min: 90% A: 10% B. The column temperature was kept at 30 °C throughout all experiments and the injection volume was 10 μ L.

The mass spectrometer parameters (source temperature, dwell time, cone voltage) were optimized for the selected ion monitoring (SIM) mode in order to achieve optimal sensitivity and selectivity (Table 2). Thus the mass spectrometer (MSQ) was operated in the negative ion mode under the following conditions: ESI voltage, -4500 V; source temperature, 450°C. All compounds were analyzed as their respective deprotonated ions $[M-H]^-$.

MS	Name	Mass	Time (min)	Dwell Time	Cone (V)
SIM	HT	153	0-3.1	0.07	65
SIM	IS	183	0-3.1	0.07	65
SIM	OE	539	0-3.1	0.07	75

Table 2. Mass spectrometry parameters - SIM conditions.

Preparation of the matrix. Initially, 2.1 g of capsules content were dissolved in 100 mL of HPLC grade water and sonicated for 5 min. The preparation of the resin was as follows: 150 g of XAD – 7HP resin were placed in a separation funnel, washed by 100 mL HPLC grade water, in order to remove any suspended solids and subsequently the resin was regenerated by concurrent feeding of 100 mL MeOH. In a final step, the resin was washed with 150 mL HPLC grade water prior to its use. The aforementioned capsule's content solution (100 mL) was loaded onto the XAD – 7HP resin. The eluate was collected and the solvent was evaporated using a rotary evaporator, affording a dry residue of 1.5 g. The absence of OE and HT was confirmed by LC-MS analysis employing the developed quantitation methodology. Thus, this eluate has been used as matrix throughout the validation procedure.

Preparation of standard solutions, calibration curve and quality control (QC) samples. 14.57 mg of matrix was suspended in 2.43 mL MeOH, followed by brief vortexing affording a suspension of 6 mg/mL. Consequently the suspension was sonicated for 5 min, centrifuged (12,000 rpm) for 3 min, and the supernatant was transferred into a glass tube and diluted 1/100 (v/v %) in aq. 0.1% formic acid/ACN 90/10 (v/v %) giving a solution 60 µg/mL of the matrix (Matrix-Matched Solution, MMS). An aliquot was transferred to an appropriate vial (1.8 mL) for LC-MS analysis. The MMS was used throughout all the validation procedure. Stock solutions of analytes, OE, HT and IS, were prepared at concentration of 1 mg/mL in MeOH and stored in the dark at 4 °C. Working solutions were daily prepared by diluting appropriate volumes of stock solutions in MMS to obtain 100 µg/mL solutions for each analyte of interest. The IS working solution was prepared by diluting an appropriate volume of the corresponding stock solution in MMS, in order to obtain a

20 µg/mL concentration level. The calibration curves covered the range 0.1-15 µg/mL for both analytes, whereas the concentration of IS was 2 µg/mL. Different sets of solutions were prepared specifically, in order to be used as QC samples. QC samples were prepared in the same way at three concentration levels, a low (LQC 4.0 µg/mL), a medium (MQC 9.0 µg/mL) and a high (HQC 12.5 µg/mL).

Sample Preparation-Extraction Protocol. A total of fifty-five capsules, 5 from each brand and/or batch were processed (see Table 1). Samples were prepared for LC-MS analysis according to the following protocol: The content of each capsule was suspended in MeOH followed by brief vortexing, affording a suspension of 6 mg/mL. The samples were sonicated for 5 min, subsequently centrifuged (12000 rpm) for 3 min, and the supernatant was transferred into glass tubes and diluted 1/100 (v/v) in aq. 0.1% formic acid/ACN 90/10 (v/v). An aliquot of the samples at the expected concentration level of 60 µg/mL was transferred to appropriate vials for analysis by LC-MS.

Quantification. Quantification was performed employing the SIM negative ion mode for OE (MW: 540 g/mol), HT (MW: 154 g/mol) and IS (MW: 184 g/mol). The ion beams (m/z 539 for OE, m/z 153 for HT and m/z 183 for IS) were optimized by adjustment of the cone voltage, the probe temperature and dwell time for each compound. An FWHM value of 0.7 was used throughout all quantification experiments. The IS methodology has been employed for the quantification of both analytes based in the calculation of the peak – area ratios of each analyte to that of IS.

Validation of the analytical methodology

The method validation was performed in accordance to the International Conference on Harmonisation (ICH) Q2 (R1) guidelines by evaluating stability, specificity, linearity, precision, accuracy, quantitation and detection limit, robustness and system suitability parameters³⁰.

Specificity-Recovery-Matrix Effect. Specificity has been examined by the injection of six individually prepared MMS solutions and the presence of possible interferences at the corresponding t_R of the analytes and the IS was investigated. The matrix effect was evaluated comparing the slope of two calibration curves, one prepared in aq. 0.1% formic acid/ACN 90/10 (v/v) and the other prepared in MMS, as described above. Recovery was determined at the three different QC levels (LQC, MQC and HQC), by comparing the concentration of analytes determined for matrix samples (n=5) spiked with the analytes before the extraction (samples prepared by spiking the MMS in the dry state with the appropriate volumes from the standard stock solutions of analytes and the IS and air dried in room temperature in order to remove the solvent), with the corresponding areas of MMS samples spiked with the target analytes at the same concentration post-extraction. The same procedure was employed for the determination of the IS recovery.

Regression model. Calibration curves ranging from 0.1 to 15 $\mu\text{g/mL}$ for both analytes of interest were constructed and analyzed on three separate days and fitted using linear regression analysis employing $1/x$ weighting. The 0.0 point was neither included as a point nor the calibration curve was forced to pass through it.

Quantitation and detection limit. The LOQ was calculated comparing measured analyte signals with those of blank samples (samples not spiked with analytes) and establishing the minimum concentration at which the analytes can be reliably quantified leading to a typical signal-to-noise ratio of 10:1. The LOD was calculated in the same manner by comparing measured analyte signals from samples with known low concentrations to those of blank samples and establishing the minimum concentration at which the analytes can be reliably detected with a signal-to-noise ratio of 3:1.

Repeatability, intermediate precision and accuracy. Five replicates of each QC sample at the three QC concentration levels (LQC, MQC, HQC) as well as at the limit of quantitation (LOQ) and upper limit of quantitation (ULOQ), were determined in order to evaluate the repeatability and the intermediate precision of the assay. The results were expressed as the standard deviation (SD) and the relative standard deviation (%RSD) for all QC, LOQ and ULOQ samples. Accuracy was determined for the same samples as the % standard error (% Er) between the mean concentrations calculated and the nominal concentrations.

Robustness. For the establishment of the method's robustness, experimental conditions were deliberately altered and several parameters were calculated. This was achieved by altering three conditions by $\pm 5\%$: setting the flow rate of the mobile phase at 0.285 mL/min and 0.315 mL/min, the column temperature at 28.5 °C and 31.5 °C and the capillary temperature at 427.5 °C and 472.5 °C. All experiments were performed at the MQC level (n=5) calculating the chromatographic peak area of each

analyte, the retention time (t_R), the 10% peak asymmetry, the 10% peak width and the USP theoretical plates number. The results were expressed as %RSD.

Total quality control of HMP's

The same samples (Table 1) have also been used for the total quality control of HMP's.

UHPLC-HRMS Analysis. A Hypersil Gold C18 50 x 2.1, 1.9 μm reversed phase column (Thermo Scientific, Germany) was used at a flow rate of 0.35 mL/min for the chromatographic separation. The mobile phase consisted of solvents A: water, 0.1% formic acid and B: ACN. A gradient method (total run time of 11 min) was used for the profiling of the samples as follows: 0 to 8 min: 95% A: 5% B, 8 to 8.5 min: 60% A: 40% B, 8.5 to 9 min: 5% A: 95% B, 9 to 9.5 min: 95% A: 5% B, 9.5 to 11 min: 95% A: 5% B. Column temperature was kept at 30 °C throughout all experiments and the injection volume was 4 μL .

A hybrid high resolution mass spectrometer (LTQ-Orbitrap Discovery) operating in negative ion mode under the following conditions: capillary temperature, 430°C; capillary voltage, -60 V; tube lens, -90 V; source voltage, 3.30 kV; sheath gas flow, 50 arb. units; aux gas flow, 30 arb. units have been employed. Analysis was performed in the full scan ion mode, using a resolution of 30000.

DPPH radical scavenging assay. The antioxidant capacity of the OE claimed HMP's was measured employing the DPPH radical scavenging method³⁹. Briefly, a DPPH solution was prepared daily dissolving 12.5 mg in 100 mL of absolute ethanol and stored at 4 °C in the dark until use. All tested HMP samples were diluted in

MeOH up to a concentration level of 6 mg/mL. A 7 μ L aliquot of each HMP sample solution was added to 193 μ L of the DPPH solution into a 96-wellplate to complete the final reaction. After 30 min at room temperature in the dark, the absorbance was measured at 517 nm in a micro plate spectrophotometer reader. The DPPH–ethanol solution with the addition of 7 μ L MeOH was used as a control sample. Gallic acid was used as a reference for the DPPH radical scavenging method. Measurements were performed in triplicate using the Magellan software (Tecan, Austria) and the results were expressed as mean \pm %RSD.

Data Analysis. Data pre-processing for the multivariate statistical analysis was performed using the MZmine 2.2 software (<http://mzmine.sourceforge.net>). The procedures of peak detection, deconvolution, normalization, deisotoping and alignment have been applied to data directly imported as raw files from the mass spectrometer. The generated peak list (accurate mass - retention time vs intensity) has been exported and manipulated to Excel using the CONCATANATE, ROUND and TRANSPOSE commands. The generated .xls file has been exported to SIMCA P+ 10.5 (Umetrics, Umea, Sweden) for the multivariate analysis.

Principal Component Analysis. PCA has been performed using both unit variance and Pareto scaling; confidence level on parameters was set at 95% whereas 200 maximum iterations have been employed. The scores values have been used for the evaluation of the results. The R^2 value has been used in order to evaluate the optimal number of contributing parameters. A cut-off value of 10% explained variance has been used for the optimal number of pc's kept for the construction of the PCA model.

RESULTS AND DISCUSSION

An outline of the described methodology is depicted in Figure 2.

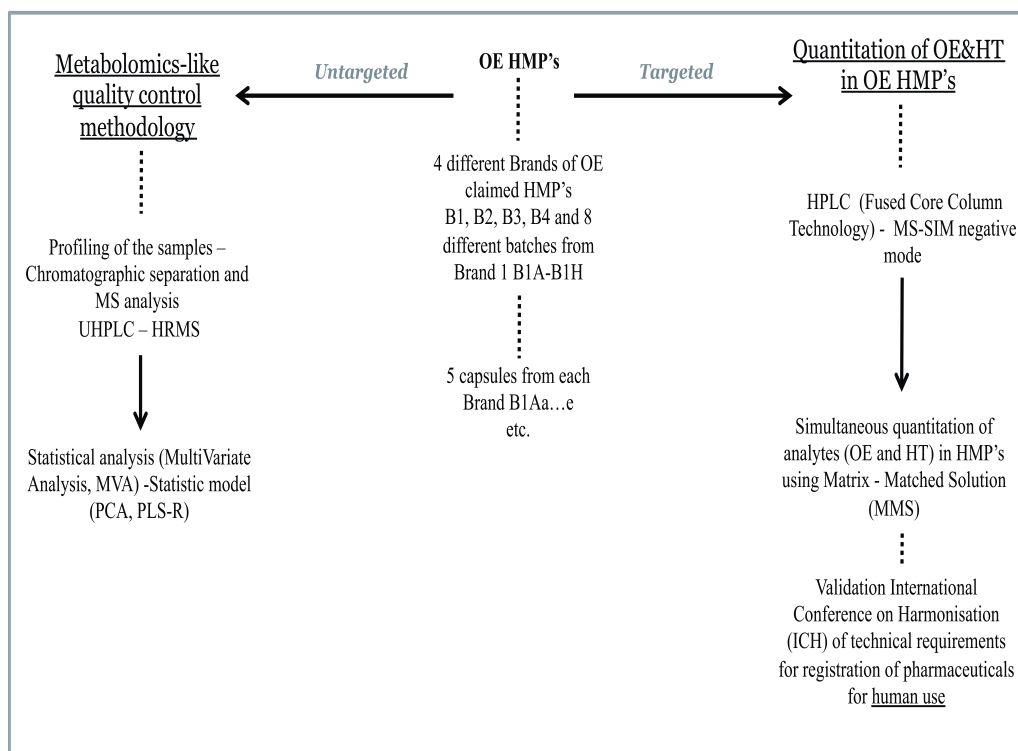


Figure 2. Workflow of the developed total quality control methodology. The right part of the figure corresponds to quantification of OE and HT in selected OE claimed HMP's and the left part corresponds to the metabolomics-like quality control methodology of the selected OE claimed HMP's.

Simultaneous quantification of OE and HT in HMP's – Validation of the analytical methodology

LC-MS conditions. Chromatographic conditions were analyzed by varying parameters such as mobile phase, gradient elution type, flow rate, type of column chemistry (C18 & C18-NPS, C8, and Phenyl) and column geometries (250 x 4.6 mm, 5µm; 100 x 2.1 mm, 5µm; 50 x 2.1 mm, 5µm, 100 x 2.1 mm, 2.7 µm). Employing the described methodology (section 2.4.1) on an Ascentis Express C18 (Fused Core) 100

x 2.1 mm, 2.7 μ m reversed phase column, sharp peaks were achieved leading to enhanced signal intensity and adequate resolution of the analytes from possible interfering components. Furthermore the total analysis time did not exceeded 3.1 min affording a rapid chromatographic methodology. Resolution between all the chromatographic peaks was greater than 2, which ascertains that all analytes are baseline separated.

Mass spectroscopic parameters (Table 2) were optimized in order to achieve the highest signal for all compounds of interest. In addition, optimization of scan and dwell time afforded well described peaks and highly stable signals.

Matrix Preparation. Isolation of the matrix was achieved following the described methodology (Preparation of the matrix section) and the absence of OE and HT was confirmed using the developed LC-MS method. A typical LC-SIM chromatogram of matrix used as the blank solution is shown in Figure 3a. The adsorbent XAD – 7HP resin can selectively retain the polyphenolic content of the extract (capsules) whereas the remaining compounds (matrix) are eluted during the procedure. As the OE and HT represent the main polyphenolic compounds of this extract (olive leaf extract), it is anticipated that the resulting eluate i.e. the matrix used, closely resembles the initial extract (capsule).

Sample Preparation-Extraction Protocol. For the efficient cleanup of the capsule samples and the isolation of the analytes, many different protocols were evaluated for the sample preparation, such as solid phase extraction and solid - liquid extraction with various solvents such as MeOH, ethanol, propanol and water. Solid - liquid extraction with MeOH was chosen as it ensured good extraction efficiency and

high reproducibility for all analytes (OE and HT) and the IS and no interference from many co-eluted components of the capsules, whereas the analysis cost and the corresponding time remained at a minimum in terms of sample preparation.

Quantification. Quantification was performed in the SIM negative ion mode for OE, HT and IS. In this study, 2,4-dinitrophenol was chosen as the IS for the LC-MS method since it exhibited high and repeatable recovery and chromatographic behavior with the analytes, it has been shown to be stable under the described experimental conditions (chemical and signal stability), and it did not interfere with the analysis of the compounds of interest. Figure 3b shows chromatograms for the MMS spiked with OE, HT and IS at the 2 µg/mL level. Blank extracts (MMS) that were processed using the described methodology, produced no discernible peaks for OE, HT and IS (shown in Figure 3a).

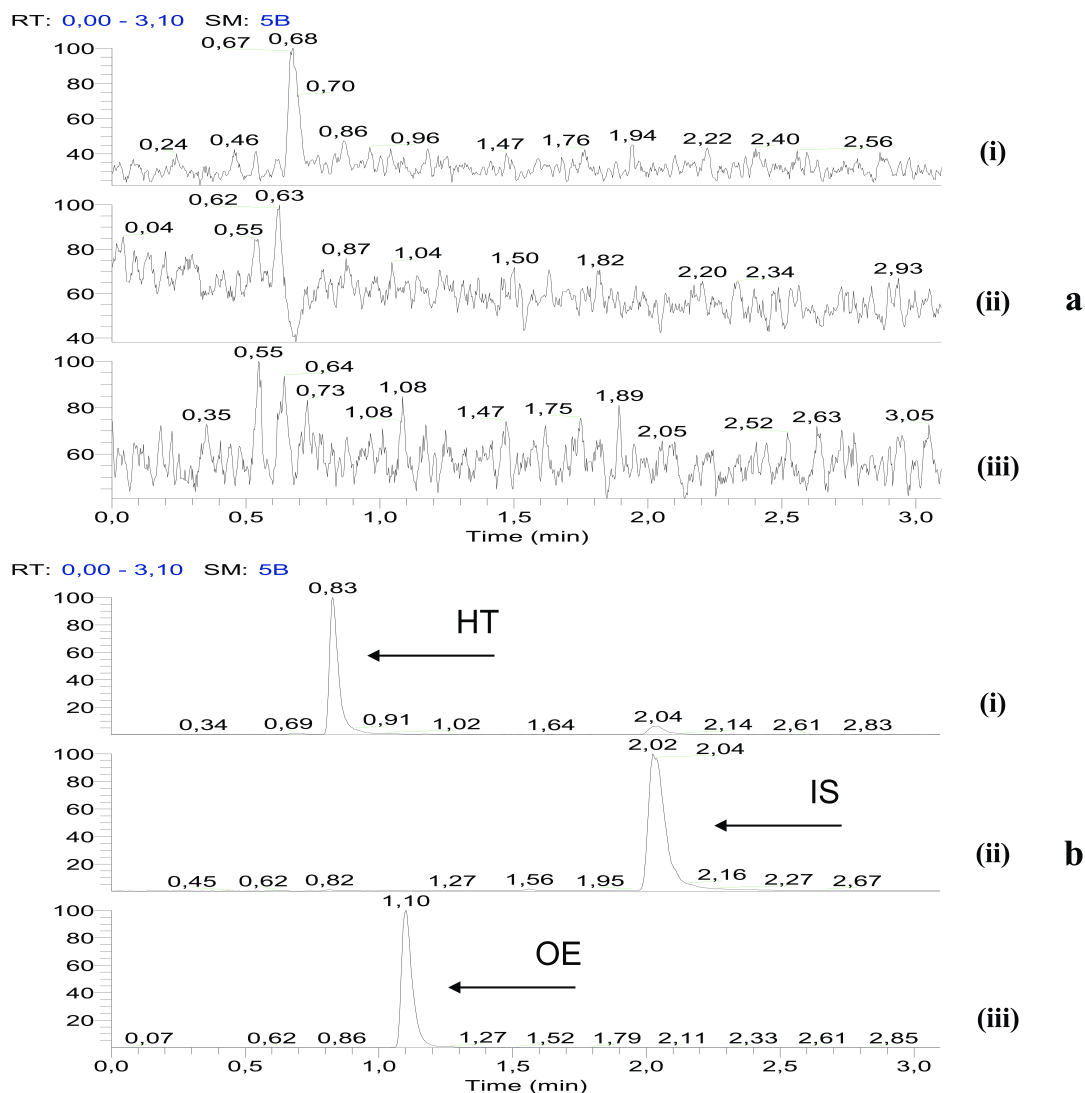


Figure 3. Typical chromatograms of the targeted quantitation of OE and HT using the proposed methodology. **a.** LC-SIM chromatogram of matrix used as the blank solution (MMS), **b.** LC-SIM chromatogram of the MMS spiked with OE, HT and IS at 2 µg/mL. The ions monitored were m/z 152.5-153.5 for HT (i), m/z 182.5-183.5 for the IS (ii) and m/z 538.5-539.5 for OE (iii) respectively.

Specificity-Recovery-Matrix Effect. The specificity was evaluated and no interference was found to the corresponding t_R of the analytes and the IS. The matrix effect was evaluated by comparing the slopes of two calibration curves, one prepared in aq. 0.1% formic acid/ACN 90/10 (v/v) and the other prepared in MMS, as described above. The slope ratios were found to be 0.85 and 0.94 for OE and HT

respectively. The data exhibit that the analyte signal does not essentially alter when the MMS was used.

Recovery was determined at the three different QC levels (LQC, MQC and HQC) and calculated as $\%R = \frac{\text{pre spiked concentration of analyte}}{\text{post spiked concentration of analyte}} \times 100$. High recovery rates render the method suitable for the analysis of OE and HT in the selected preparations (B1, B2, B3 and B4).

Regression model. Calibration curves of OE and HT were linear over the range 0.1-15.0 $\mu\text{g/mL}$, with a correlation coefficient (r^2) > 0.99. Typical equations for the calibration curves are shown in Table 3 and were used for the quantitation of OE and HT in OE claimed HMP's. Performing a t-test it has been also shown that the intercept was not significantly different from 0, at the 95% confidence level.

Analytes	Equations	R^2	Weighting
OE	$C_{OE} = 2.3 \cdot 10^{-1} (\pm 4.3 \cdot 10^{-4}) \text{ Area Ratio} - 5.3 \cdot 10^{-2} (\pm 3.0 \cdot 10^{-2})$ $R^2 = 0.997$ W: 1/X	0.997	1/X
HT	$C_{HT} = 3.1 \cdot 10^{-2} (\pm 4.5 \cdot 10^{-4}) \text{ Area Ratio} - 3.0 \cdot 10^{-4} (\pm 3.1 \cdot 10^{-4})$ $R^2 = 0.9993$ W: 1/X	0.9993	1/X

Table 3. Typical equations for the calibration curves of OE and HT.

Quantitation and detection limit. The LOQ of both for OE and HT was found to be 0.1 $\mu\text{g/mL}$ and is determined with acceptable accuracy and precision at this concentration level according to the ICH regulations. The LOQ for both analytes defined as the concentration that yielded a peak with a signal/noise ratio of 10 and at least 3 times the response compared to that blank extracts (samples not spiked with analytes). Furthermore the accuracy, expressed as the relative percentage error %Er was found to be lower than 3.8% whereas the precision expressed as the %relative

standard deviation %RSD was found to be less than 14.9% for both analytes at the LOQ level. The data are shown in Table 4. The LOD was found to be 0.03 µg/mL for both analytes. The data shows that the described method is of acceptable sensitivity for the analysis of OE and HT in formulations.

(n=5)	Repeatability		Intermediate Precision		Accuracy	
Nominal Value ($\mu\text{g/mL}$)	OE [Mean \pm SD ^a , (%RSD ^b)]	HT [Mean \pm SD ^a , (%RSD ^b)]	OE [Mean \pm SD ^a , (%RSD ^b)]	HT [Mean \pm SD ^a , (%RSD ^b)]	OE % Std Error ^c	HT % Std Error ^c
ULOQ	14.6 \pm 0.7 (4.0)	14.9 \pm 0.2 (1.0)	14.7 \pm 0.6 (4.1)	16.1 \pm 1.4 (9.3)	15.0	10.1
HQC	13.0 \pm 0.8 (5.2)	12.0 \pm 0.3 (2.3)	12.0 \pm 0.4 (3.7)	12.6 \pm 0.7 (5.2)	12.5	1.7
MQC	9.1 \pm 1.4 (14.9)	9.0 \pm 0.3 (3.1)	9.2 \pm 0.4 (4.0)	8.9 \pm 0.5 (5.3)	9.0	-1.3
LQC	4.1 \pm 0.1 (3.7)	3.9 \pm 0.1 (2.3)	4.3 \pm 0.2 (3.8)	4.2 \pm 0.3 (7.9)	4.0	0.3
LOQ	0.10 \pm 0.01 (8.4)	0.11 \pm 0.02 (14.4)	0.1 \pm 0.01 (14.9)	0.11 \pm 0.02 (18.9)	0.1	3.8

^aStandard deviation, ^b% Relative standard deviation, ^c% Standard error

Table 4. Validation results of the proposed methodology.

Repeatability, intermediate precision and accuracy. Repeatability, intermediate precision and accuracy were determined by analyzing replicates (n=5) at the three QC concentration levels (LQC, MQC and HQC), as well as at the LOQ and ULOQ level. The repeatability derived from all QC levels as well as LOQ and ULOQ, did not exhibit values more than 14.4% for OE and more than 14.9% for HT respectively. The intermediate precision did not exhibit values more than 14.9% for OE and 18.9% for HT respectively for all QC, LOQ and ULOQ levels. The %Er derived was found to be within $\pm 10.1\%$ for OE and $\pm 2.7\%$ for HT. The results of accuracy and precision of the method for the two analytes are summarized in Table 4 and proved that the described method can reliably use for the quantitation of OE and HT in OE claimed HMP's.

Robustness. To ensure the insensitivity of the HPLC method to minor changes in the experimental conditions, it is important to demonstrate the robustness of the analytical method. None of the alterations caused a significant change more than 2% for all the parameters examined, proving that the method is sufficiently robust to be used for quality control measurements. Thus, %RSD of the resolution between OE and HT was found to be less than 2% when the conditions were deliberately altered according to the robustness scheme. Changing the flow rate of the mobile phase from 0.300 mL/min to 0.285 mL/min and 0.315 mL/min the %RSD of the resolution between peaks did not exceed 3.4%. It was also noticed that changing the column temperature from 30 °C to 28.5 °C and 31.5 °C, the %RSD of the resolution between peaks was found to be less than 1.5%. Furthermore, by altering the capillary temperature from 450 °C to 427.5 °C and 472.5 °C the peak area for all peaks was not significant affected.

Measurement of OE and HT in HMP's. Quantification of OE and HT in HMP's was based on standard curves that were prepared daily. From the analysis of the results it has been found that only one brand, i.e. B1, was precise concerning the content of OE in accordance to its label claim. In one brand, B4, neither OE nor HT were found, whereas two brands, B2 and B3, were found to be out of their label claim. The content of OE and HT (mg/Cap) in the examined OE claimed capsules from different brands are shown in Table 5 and in Table S1.

n=5	OE		HT	
	Mean (mg/cap)	%RSD ^a	Mean (mg/cap)	%RSD ^a
Brand 1 (300 mg cap)	51,1	2,5	1,8	0,6
Brand 2 (150 mg cap)	8,9	0,6	NF	-
Brand 3 (450 mg cap)	26,1	1,0	NT	-
Brand 4 (200 mg cap)	NF	-	NF	-

a% Relative standard deviation

NF Not Found

Table 5. Assayed OE and HT capsule content from the different HMP's analyzed.

Total quality control of HMP's

Validation of UHPLC-HRMS analysis. In order to assess the reproducibility of the method, ten QC identical samples were analysed at the beginning and at the end of the procedure. The variation of t_R was less than 0.5% RSD and the variation of accurate mass was less than 3 ppm, demonstrating the excellent reproducibility of the procedure. Peak area RSD was found to be lower than 6.2% for all the analytes.

Antioxidant activity. In order to gain insight to the possible antioxidant activity of the different HMP's, the DPPH radical scavenging method was performed to the list of the examined samples. As it shown in Figure 4 the inhibition caused by B1 samples

was found to be higher than the inhibition caused by B2 and B3 samples whereas B4 samples did not cause any inhibition. Specifically, B1 antioxidant capacity was found to be 70.3 ± 1.7 , for B2 antioxidant capacity was found to be 16.8 ± 2.6 , 21.3 ± 2.1 for B3 and no inhibition was noticed for B4. The results show that the OE content (Figure 4) shows a high linear correlation ($\text{DPPH Activity \%} = 4.1 \times \% \text{ OE/Cap} - 2.2$, $r^2 = 0.996$) to the total antioxidant capacity of the examined samples.

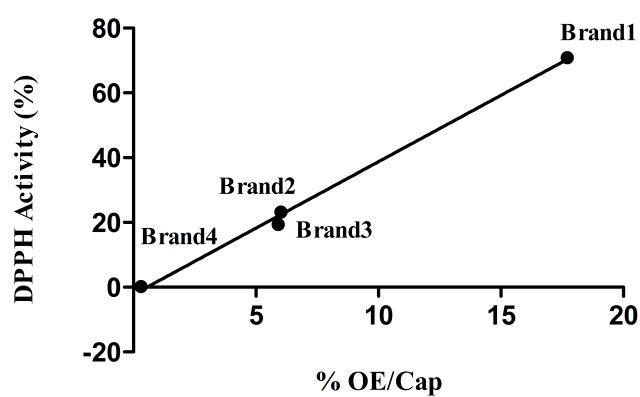


Figure 4. Linear regression of the % OE content in the capsules vs DPPH activity %, (the correlation coefficient was found to be $r^2 = 0.996$)

MVA Analysis. MVA has been employed using the PCA statistical model. Various parameters have been used, such as different scaling e.g. UV or Pareto, and different combinations of Q^2 and R^2 in order to assess the number of factors that described as much variance as possible without increasing the complexity of the model excessively. The pareto scaling method has been finally selected as it afforded tighter groups with larger separation. An arbitrary cutoff value of 10% for the explained variance for the pc's has been employed, in an effort to keep the simplest form for the PCA model. This has been verified by the scree plot, which showed rapid inflection after the second pc. Incorporation of more pc's led to an insignificant

improvement of the model. Thus, only the two first pc's have been taken into account. In a first approach all samples were included to the dataset, i.e. all samples from all the four Brands i.e five random samples from eight different batches from B1 (B1A – B1H), i.e. seven production batches and one special preparation batch, and in addition five random samples from one batch from the rest of the brands (B2, B3 and B4). The PCA analysis exhibited clear grouping of the different brand samples examined. Thus the scores plot shown in Figure 5a demonstrates the separation of the four brands in different groups applying pareto scaling, which proves the ability of the method to discriminate OE claimed HMP's from different manufacturers based not only on the target component quantity but also on the overall profile of the extract. The most important metabolites responsible for this grouping (loadings) are summarized in Table S2 as marker ions, corresponding retention times, possible identifications and their peak areas and the relative peak areas. It should be noted that features with identical m/z but different retention times possibly correspond to isomers. Furthermore, an analysis for possible adducts (M-H₂O-H, M-H, M+Na-2H, M+Cl, M+K-2H, M+FA-H, M+Br, 2M-H, 2M+FA-H) has been performed for all the features in order to verify the possibility of occurrence of chromatographic peaks due to adduct formation. The first pc holds for the largest part of variation (30.8%) between the different brands, while the second pc holds for the 11.1% of the variation.

As described before the second major target of the current work was to examine the quality of the production procedure of HMP's, based on the uniformity of the contained minor substances profile and not only taking into account the content of the active label substance OE for establishing the batch-to-batch reproducibility of the finished products. Such an approach will ensure the proper, within specifications, pharmacological activity of the HMP's respecting possible synergism concerns. This

can be assured not only residing in the content uniformity concerning the concentration of major compound, but also controlling the amounts of the minor constituents between batches. Hence, concerning the Brand B1, eight different batches were sampled and analyzed by the described UHPLC-HRMS methodology in order to assess the batch-to-batch reproducibility based on their total chemical content.

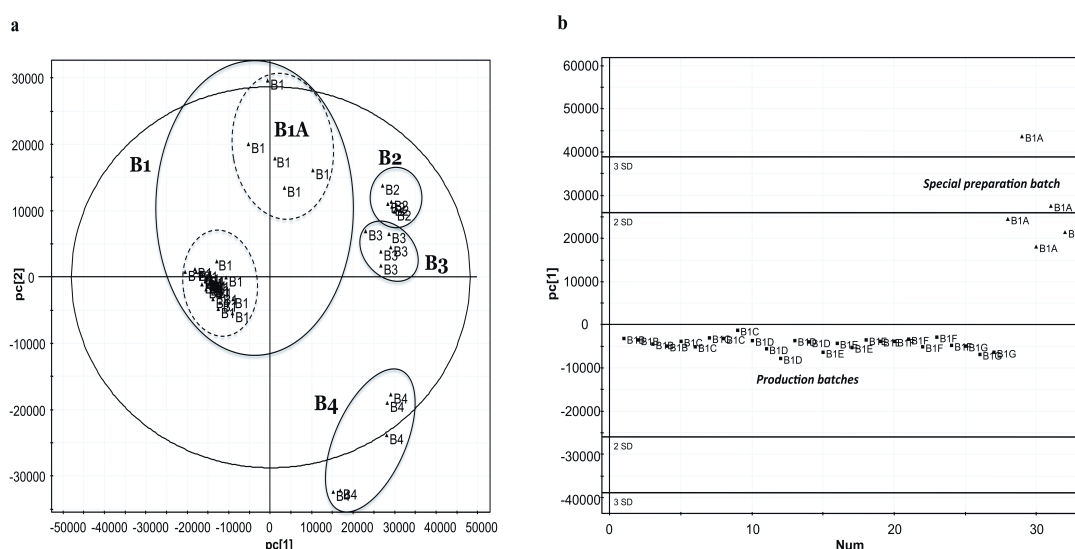


Figure 5. Multivariate analysis of the data obtained from the untargeted quality control approach. **a.** PCA scores plot of the four brands (B1-B4) ($n=5$ capsules) and the different preparations of B1 (B1A-B1H) ($n=5$ capsules) obtained after pareto scaling. PCA analysis shows clear grouping of the different examined samples and indicates the excellent batch-to-batch reproducibility obtained from B1, **b.** PCA based batch to batch reproducibility of the total content of Brand 1. The production batches exhibit low batch to batch variability (<1 SD values) based on their total chemical content. The special preparation batch is separately grouped.

Figure 5b shows only the B1 samples (production batches and the special preparation batch) for reasons of clarity, i.e. the MVA is not confounded by the presence of the other brands, which could potentially lead to tighter grouping of the production batches due to the large separation distance from the other batches in pc1 as can be seen in Figure 5a. It is clear that all data points from the production batches

lie within one SD from the mean, whereas the special preparation batch lies near or above the 2 SD limit. This reflects the excellent batch-to-batch reproducibility obtained from the production batches of B1 and furthermore it proves that the even a slightly different preparation (B1A) from the same production line (B1B-B1H) could be distinguished. Thus it seems that any “out-of-specification” batches could be easily detected. It is also evident from Figure 5a that the different preparations of B1 (B1A vs. B1B-B1H) can separate according to the second principal component (pc2) whereas in the first principal component (pc1) they seem indistinguishable, proving that slight differences play a critical role in the grouping of them. Overall it could be reasonably proposed that the quality of HMP's could be divided in two parts the first one concerning the active ingredient content whereas the second part should involve the reproducibility of the overall production reflecting the total quality of the extract used as well as the production procedures. Thus extracts from different sources or undergone different production procedures (e.g. longer storage, different storage temperature, differentiated production procedures etc.) could potentially lead to essentially different products from the pharmacological point of view. In the current work a workflow is proposed in order to overcome such issues.

CONCLUSION

Overall an LC-(-ESI)-MS method in the SIM mode for the quantification of OE and its main degradation product, HT, in different selected commercial OE claimed HMP's was developed and validated according to the ICH requirements. Despite the fact that SRM mode in MS is considered as more sensitive and selective, such

instrumentation is rather expensive and complicated. Thus, an SRM based methodology is not fit of purpose for the current analysis, considering also the fact that the targeted compounds are present in large amounts to the HMP's under investigation. Results regarding specificity, recovery, matrix effect, regression model, quantitation and detection limit, repeatability, intermediate precision, accuracy and robustness are satisfactory, and the developed method is applicable for its intended purpose. Furthermore, a simple and inexpensive sample preparation procedure followed by a fast analytical run of 3.1 min (UHPLC-like analysis on a Fused Core column) to analyze and quantify all analytes as well as the IS has been employed. Unlike the common practice, the matrix of the OE claimed HMP has been isolated and used in order to assess critical parameters such as the matrix effect; a crucial parameter that should be always considered when dealing with MS based detection.

The measured content of OE in the preparation B1 was in agreement with the concentration reported on label, and confirmed the validity of the developed method. In contrast, preparation B2 and B3 were out of specification whereas in B4 OE was found to be completely absent. In fact three out of four preparations were not compliant to the specifications, thus verifying the need of new validated methodologies for the quality control of HMP's.

Additionally a UHPLC-(-ESI)-HRMS-based metabolomics-like approach was developed in order to evaluate the holistic quality of the four different brands of commercial OE claimed HMP's. By this approach, certain quality issues of commercial OE claimed HMP's were revealed. The data obtained by the present investigation led to a satisfactory and quick characterization methodology of complex mixtures of HMP's. Nevertheless, it should be noted that UHPLC-(-ESI)-HRMS-based profiling is more sensitive but less reproducible than its NMR counterpart,

which is less sensitive but more reproducible. Thus, NMR based profiling could be an attractive alternative in this issue.

This work could potentially consist the basis for further investigation on the quality control and the stability of commercial HMP's. Concluding the proposed methodology can be used for routine quantitative analysis, quality control as well as for studies on the production process of HMP's. Overall, the proposed methodology should be incorporated to the production process of HMP's or even pharmaceuticals in a long-term manner. Such an approach will enhance the confidence and will reassure the high quality of the final products.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Frutarom Switzerland Ltd., Rütawisstrasse 7, CH-8820 Wädenswil, Switzerland, for providing the different production batches of one of the OE claimed HMP's using in this study.

REFERENCES

1. Agalias A, Melliou E, Magiatis P, Tsarbopoulos A, Mitaku S (2005) Quantitation of polyphenols and secoiridoids in decoctions of *Olea europaea* leaves from ten Greek cultivated varieties. J Liq Chromatogr 28: 1557–1571.
2. Rmani N, Mulinacci P, Pinelli F, Vincieri F, Cimato A (1999) Polyphenolic content in five tuscan cultivars of *Olea europaea* L. J Agric Food Chem 47: 964-967.
3. Japon-Lujan R, Luque de Castro MD (2006) Superheated liquid extraction of oleuropein and related biophenols from olive leaves. J Chromatogr A 1136: 185–191.
4. Renaud S, De Lorgeril M, Delaye J, Guidollet J, Jacquard F, et al. (1995) Cretan Mediterranean diet for prevention of coronary heart disease Am J Clin Nutr 6: 1360–1367.
5. Galli C, Visioli F (1999) Antioxidant and other activities of phenolics in olives/olive oil, typical components of the mediterranean diet. Lipids 34: 23-26.
6. Kontogianni VG, Gerothanassis IP (2012) Phenolic compounds and antioxidant activity of olive leaf extracts. Nat Prod Res 26: 186–189.
7. Pereira AP, Ferreira ICFR, Marcelino F, Valentao P, Andrade PB, et al. (2007) Phenolic compounds and antimicrobial activity of olive (*Olea europaea* L. Cv. Cobrancosa) leaves. Molecules 12 (5): 1153–1162.
8. Omar SH (2010) Cardioprotective and neuroprotective roles of oleuropein in olive. Saudi Pharmaceutical Journal 18: 111–121.
9. Keys A (1997) Coronary heart disease in seven countries. 1970. Nutrition 13(3): 249-253.
10. Andreadou I, Iliodromitis EK, Mikros E, Constantinou M, Agalias A, et al. (2006) The olive constituent oleuropein exhibits anti-ischemic, antioxidative, and hypolipidemic effects in anesthetized rabbits. J Nutr 136: 2213–2219.
11. Visioli F, Galli C (2001) Antiatherogenic components of olive oil. Curr Atheroscler Rep 3: 64–67.
12. Scoditti E, Calabriso N, Massaro M, Pellegrino M, Storelli C, et al. (2012) Mediterranean diet polyphenols reduce inflammatory angiogenesis through MMP-9 and COX-2 inhibition in human vascular endothelial cells: A potentially protective mechanism in atherosclerotic vascular disease and cancer. Arch Biochem Biophys 527: 81-89.
13. Poudyal H, Campbell F, Brown L (2010) Olive Leaf Extract Attenuates Cardiac, Hepatic, and Metabolic Changes in High Carbohydrate High Fat Fed Rats. J Nutr 140: 946–953.
14. Jemai H, El Feki A, Sayadi S (2009) Antidiabetic and Antioxidant Effects of Hydroxytyrosol and Oleuropein from Olive Leaves in Alloxan Diabetic Rats. J Agric Food Chem 57(19): 8798-8804.

15. Acquaviva R, Di Giacomo C, Sorrenti V, Galvano F, Santangelo R, et al. (2012) Antiproliferative Effect Of Oleuropein In Prostate Cell Lines. *Int J Oncol* 41(1): 31-38.
16. Assessment report on *Olea europaea* L., folium. Based on Article 16d(1), Article 16f and Article 16h of Directive 2001/83/EC as amended (traditional use), Assessment report on *Olea europaea* L., folium, EMA/HMPC/430506/2009 (EMA website). Available: http://www.ema.europa.eu/docs/en_GB/document_library/Herbal_-HMPC_assessment_report/2012/04/WC500125459.pdf. Accessed 2013 Oct 03).
17. Granados-Principal S, Quiles JL, Ramirez-Tortosa CL, Sanchez-Rovira P, Ramirez-Tortosa MC (2010) Hydroxytyrosol: from laboratory investigations to future clinical trials. *Nutr Rev* 68(4): 191-206.
18. Directive 2004/24/EC of the European Parliament and of the Council, amending, as regards traditional herbal medicinal products, Directive 2001/83/EC on the Community code relating to medicinal products for human use, Official Journal of the European Union (2004) L 136-85 (Available: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:136:0085:0090:en:PDF>).
19. EMA website. Available: http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000208.jsp&mid=WC0b01ac05800240cf. Accessed 2013 Oct 03.
20. Williamson EM et al. (2001) Synergy and other interactions in phytomedicines. *Phytomedicine* 5: 401-409.
21. Dierkes G, Krieger S, Duck R, Bongartz A, Schmitz OJ, et al. (2012) High-Performance Liquid Chromatography–Mass Spectrometry Profiling of Phenolic Compounds for Evaluation of Olive Oil Bitterness and Pungency. *J Agric Food Chem* 60: 7597–7606.
22. Zoidou E, Melliou E, Gikas E, Tsarbopoulos A, Magiatis P, et al. (2010) Identification of Throuba Thassos, a Traditional Greek Table Olive Variety, as a Nutritional Rich Source of Oleuropein. *J Agric Food Chem* 58: 46–50.
23. Sanchez JL, Giambanelli E, Quirantes Pine R, Cerretani L, et al. (2011) Wastes Generated during the Storage of Extra Virgin Olive Oil as a Natural Source of Phenolic Compounds. *J Agric Food Chem* 59: 11491–11500.
24. Bazoti FN, Gikas E, Tsarbopoulos A (2010) Simultaneous quantification of oleuropein and its metabolites in rat plasma by liquid chromatography electrospray ionization tandem mass spectrometry. *Biomed Chromatogr* 24: 506–515.
25. Ortega Garcia F, Peragon J (2010) HPLC analysis of oleuropein, hydroxytyrosol, and tyrosol in stems and roots of *Olea europaea* L. cv. Picual during ripening. *J Sci Food Agric* 90: 2295–2300.
26. Tan HW, Tuck KL, Stupans I, Hayball PJ (2003) Simultaneous determination of oleuropein and hydroxytyrosol in rat plasma using liquid chromatography with fluorescence detection. *J Chromatogr B* 785: 187–191.
27. Zhoua T, Qiana T, Wanga X, Lia X, Caoa L, et al. (2011) Application of LC-MS/MS method for the in vivo metabolite determination of oleuropein after intravenous administration to rat. *Biomed Chromatogr* 25(12): 1360-1363.

28. Briante R, Patumi M, Terenziani S, Bismuto E, Febbraio F, et al. (2002) *Olea Europaea* L. Leaf Extract And Derivatives: Antioxidant Properties. *J Agric Food Chem* 50: 4934–4940.
29. Aturki Z, Fanali S, D’Orazio G, Rocco A, Rosati C (2008) Analysis of phenolic compounds in extra virgin olive oil by using reversed-phase capillary electrochromatography. *Electrophoresis* 29: 1643–1650.
30. Aouidi F, Dupuy N, Artaud J, Roussos S, Msallem M, et al. (2012) Rapid quantitative determination of oleuropein in olive leaves (*Olea europaea*) using mid-infrared spectroscopy combined with chemometric analyses. *Industrial Crops and Products* 37: 292–297.
31. International Conference On Harmonisation Of Technical Requirements For Registration Of Pharmaceuticals For Human Use ICH Harmonised Tripartite Guideline Validation Of Analytical Procedures: Text And Methodology Q2 (R1)
Current Step 4 version Parent Guideline dated 27 October 1994 (Complementary Guideline on Methodology dated 6 November 1996 incorporated in November 2005) (ICH website. Available: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q2_R1/Step4/Q2_R1_Guideline.pdf. Accessed 2013 Oct 03).
32. Yang P, Litwinski G.R, Pursch M, McCabe T, Kuppannan K (2009) Separation of natural product using columns packed with Fused-Core particles. *J Sep Sci* 32: 1816–1822.
33. Manchón N, D’Arrigo M, García-Lafuente A, Guillamón E, Villares A, et al. (2010) Fast analysis of isoflavones by high-performance liquid chromatography using a column packed with fused-core particles. *Talanta* 82(5): 1986–1994.
34. Orozco-Solano MI, Ferreiro-Veraa C, Priego-Capotea F, Luque de Castro MD (2012) Automated method for determination of olive oil phenols and metabolites in human plasma and application in intervention studies. *J Chromatogr A* 1258: 108–116.
35. Hadad GM, Abdel Salam RA, Emara S (2012) Validated and optimized high-performance liquid chromatographic determination of tizoxanide, the main active metabolite of nitazoxanide in human urine, plasma and breast milk. *J Chromatogr Sci* 50(6): 509–515.
36. Lan K, Zhang Y, Yang J, Xu L (2010) Simple quality assessment approach for herbal extracts using high performance liquid chromatography-UV based metabolomics platform. *J Chromatogr A* 1217: 1414–1418.
37. Zhang HM, Lib SL, Zhang H, Wang Y, Zhao ZL, et al. (2012) Holistic quality evaluation of commercial white and red ginseng using a UPLC-QTOF-MS/MS-based metabolomics approach. *J Pharm Biomed Anal* 62: 258–273.
38. Agalias A, Magiatis P, Skaltsounis AL, Mikros E, Tsaibopoulos A, et al. (2007) A New Process For The Management Of Olive Oil Mill Waste Water And Recovery Of Natural Antioxidants. *J Agric Food Chem* 55: 2671–2676.
39. Stagos D, Portesis N, Spanou C, Mossialos D, Aligiannis N, et al. (2012) Correlation of total polyphenolic content with antioxidant and antibacterial activity of 24 extracts from Greek domestic Lamiaceae species. *Food Chem Toxicol* 50(11): 4115–4124.

SUPPORTING INFORMATION

Filename	Area	ISTD Area	Area Ratio	Amount	Units	RT	Weight of Cap content (mg)	OE/Cap (mg/Cap)	Mean	Stdv	%
B4	NF	2716458	NF	NF	ug/mL	NF	350	NF			
B4	NF	2582360	NF	NF	ug/mL	NF	350	NF			
B4	NF	2651643	NF	NF	ug/mL	NF	350	NF			
B4	NF	2571472	NF	NF	ug/mL	NF	350	NF			
B4	NF	2569926	NF	NF	ug/mL	NF	350	NF	NF	-	-
B1	6573460	2681149	2,452	9,789	ug/mL	1,19	300	48,94			
B1	6816836	2638898	2,583	10,419	ug/mL	1,19	300	52,1			
B1	6501454	2703677	2,405	9,567	ug/mL	1,18	300	47,83			
B1	6897847	2626305	2,626	10,63	ug/mL	1,18	300	53,15			
B1	6914625	2626011	2,633	10,663	ug/mL	1,18	300	53,31	51,07	2,52	17,02
B2	2756145	2653020	1,039	3,788	ug/mL	1,18	150	9,47			

Filename	Area	ISTD Area	Area Ratio	Amount	Units	RT	Weight of Cap content (mg)	OE/Cap (mg/Cap)	Mean	Stdv	%
B2	2252813	2595122	0,868	3,137	ug/mL	1,17	150	7,84			
B2	2602319	2655608	0,98	3,562	ug/mL	1,18	150	8,9			
B2	2702387	2698448	1,001	3,644	ug/mL	1,17	150	9,11			
B2	2672726	2627282	1,017	3,705	ug/mL	1,17	150	9,26	8,92	0,64	5,94
B3	2376414	2626136	0,905	3,276	ug/mL	1,16	450	24,57			
B3	2579109	2674541	0,964	3,502	ug/mL	1,17	450	26,26			
B3	2638579	2706454	0,975	3,542	ug/mL	1,16	450	26,57			
B3	2456625	2584207	0,951	3,45	ug/mL	1,14	450	25,87			
B3	2598111	2598054	1	3,638	ug/mL	1,13	450	27,29	26,11	1,01	5,8

Table S1A. Targeted analysis data. Peak areas of OE and IS in the examined HMP's and quantification thereof by the proposed LC-SIM methodology.

File Name	Area	ISTD Area	Area Ratio	Amount	Units	t _R	Weight of Cap content (mg)	HT/Cap (mg/Cap)	Mean	Stdv	%
B4	NF	2716457,611	NF	NF	ug/mL	NF	350	NF			
B4	NF	2582359,539	NF	NF	ug/mL	NF	350	NF			
B4	NF	2651642,788	NF	NF	ug/mL	NF	350	NF			
B4	NF	2571471,61	NF	NF	ug/mL	NF	350	NF			
B4	NF	2569925,957	NF	NF	ug/mL	NF	350	NF	NF	-	-
B1	27187,88957	2681148,813	0,010140388	0,244784894	ug/mL	0,875920951	300	1,22			
B1	38351,46856	2638897,574	0,01453314	0,384194004	ug/mL	0,872701347	300	1,92			
B1	25141,48941	2703676,971	0,009298999	0,218082429	ug/mL	0,872700274	300	1,09			
B1	46309,89342	2626305,062	0,017633098	0,482574764	ug/mL	0,879141867	300	2,41			
B1	44763,90828	2626010,933	0,017046353	0,463953728	ug/mL	0,875921607	300	2,32	1,79	0,61	0,6

File Name	Area	ISTD Area	Area Ratio	Amount	Units	t _R	Weight of Cap content (mg)	HT/Cap (mg/Cap)	Mean	Stdv	%
B2	6956,8249 76	2653020,0 26	0,0026222 29	NF	ug/mL	0,8694814 44	150	NF			
B2	4652,8894 34	2595121,9 54	0,0017929 37	NF	ug/mL	0,8727008 7	150	NF			
B2	6936,8285 19	2655608,0 45	0,0026121 43	NF	ug/mL	0,8759183 88	150	NF			
B2	6483,7181 23	2698447,6 61	0,0024027 59	NF	ug/mL	0,8662602 9	150	NF			
B2	6271,6305 51	2627282,0 99	0,0023871 17	NF	ug/mL	0,8759212 49	150	NF	NF	-	-
B3	8027,9360 47	2626135,5 33	0,0030569 39	NT	ug/mL	0,8662610 65	450	NT			
B3	15943,440 58	2674540,6 15	0,0059611 88	0,1121531 19	ug/mL	0,8598199 49	450	0,84			
B3	10139,904 48	2706453,8 5	0,0037465 65	NT	ug/mL	0,8598200 08	450	NT			
B3	9689,3896 79	2584207,4 78	0,0037494 63	NT	ug/mL	0,8404986 86	450	NT			
B3	9569,4364	2598053,8	0,0036833	NT	ug/mL	0,8501600	450	NT	NT	-	-

File Name	Area	ISTD Area	Area Ratio	Amount	Units	t _R	Weight of Cap content (mg)	HT/Cap (mg/Cap)	Mean	Stdv	%
	27	59	09			03					

Table S1B. Targeted analysis data. Peak areas of HT and IS in the examined HMP's and quantification thereof by the proposed LC-SIM methodology.

Excel File (available upon request)

Table S2. The most important loadings (metabolites) i.e. marker ions, corresponding retention times, possible identifications (as determined from an in-house substance library) and their peak areas and relative peak areas (%), responsible for the grouping of the four different brands of HMP's.

ANNEX

Besides the PCA model (Figure 1a), a PLS model was constructed also in order to correlate the antioxidant activity with the corresponding preparations (brands). Figure 1b shows that all data points were clustered separately in four classes in the PLS scores plot exhibiting the differences in DPPH activity of the four examined brands. From the corresponding PLS loadings plot (Figure 1c) it is evident that despite the fact that OE was the main discriminating loading of B1, there were few other loadings that also were of relatively of high importance.

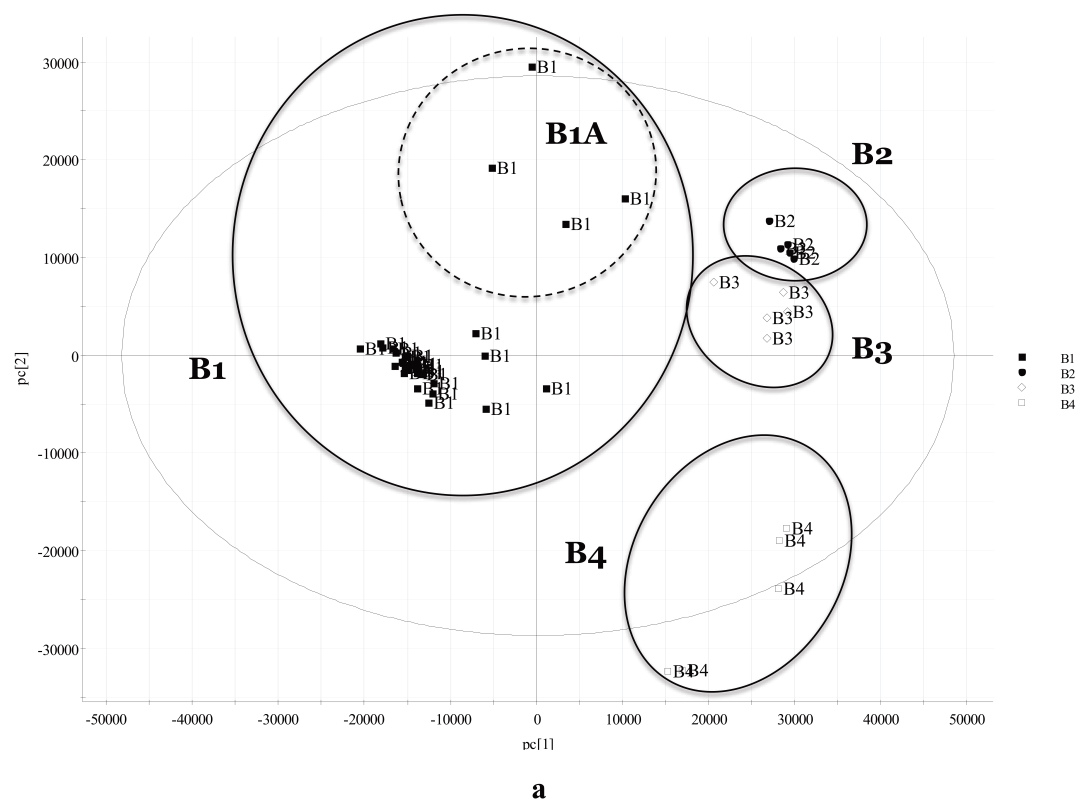


Figure 5a. PCA scores plot of the four brands (B1-B4) (n=5 capsules) and the different preparations of B1 (B1A-B1H) (n=5 capsules). PCA analysis shows clear clustering of the different examined samples and excellent batch-to-batch reproducibility obtained from B1.

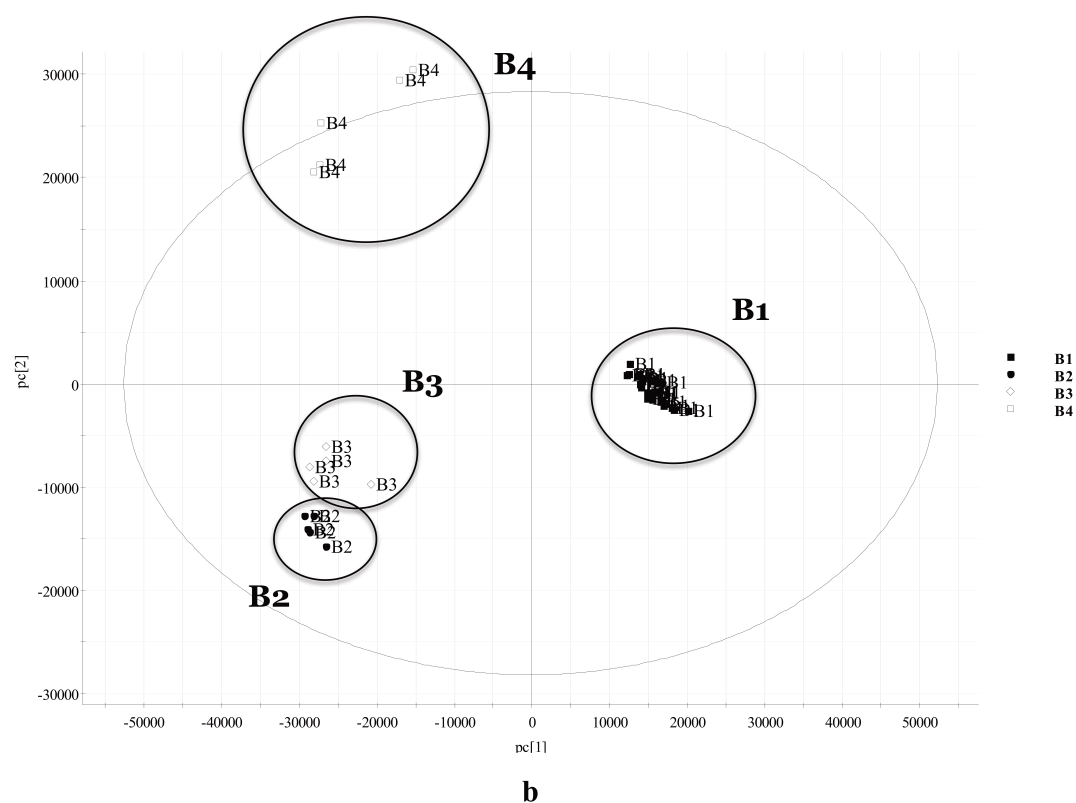


Figure 5b. PLS-DA scores plot of the four brands (B1-B4). All points are clustered separately in four classes.

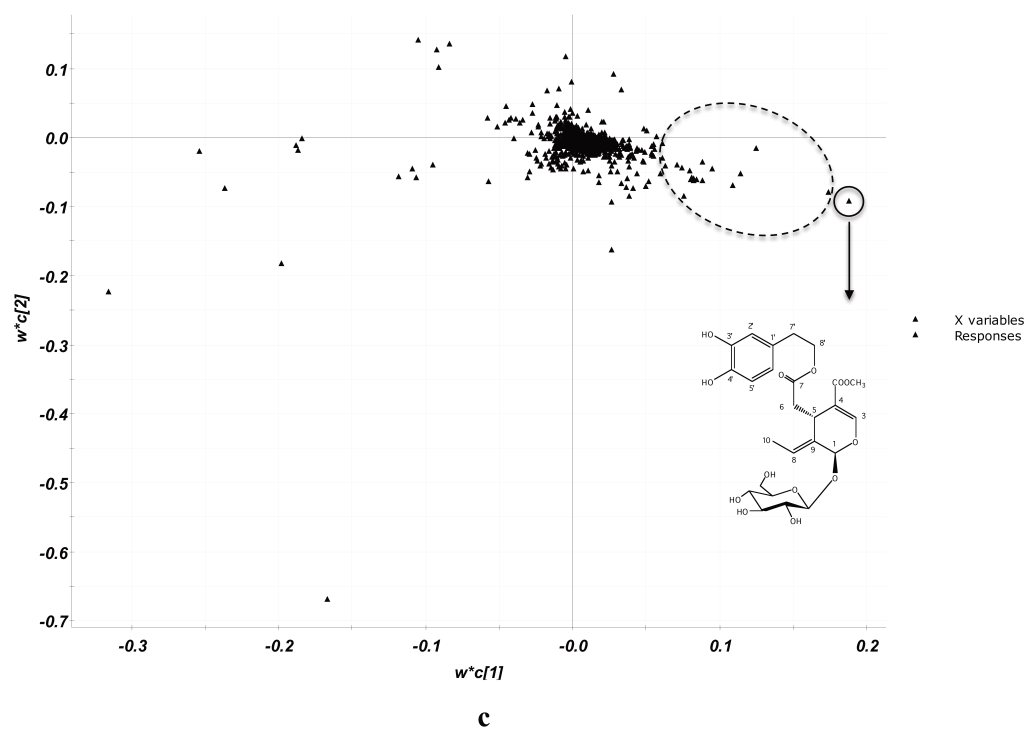


Figure 5c. PLS-DA loadings plot. OE is the main discriminating loading of B1. The cycled loadings are also of high importance

3

Validation of a bioanalytical methodology based on UHPLC-HRMS/MS for the quantification of oleuropein in human serum. Application to a pharmacokinetic study

This chapter will be submitted for publication

ABSTRACT

A highly sensitive, rapid and specific ultra performance liquid chromatography coupled to electrospray ionization high resolution tandem mass spectrometry [UHPLC-(-)ESI-HRMS/MS (orbitrap)] methodology, was developed and validated in order investigate the absorption of oleuropein (OE) in human subjects. Serum samples were collected in predefined time points, after oral administration of an olive leaf extract enriched in OE (204.4 mg OE) in two subjects. Subsequently, samples were analyzed by the developed methodology after a simple solid phase extraction step. Chromatographic separation was operated with aq. formic acid, 0.1% (v/v) and ACN following a gradient program at a flow rate of 0.45 mL min^{-1} on an RP-C18 (50 x 2.1 mm, 1.9 μm) column with a total run time of 2.7 min. The method was validated and successfully applied to the determination of OE in human serum, with the pharmacokinetic analysis of the data revealing a double peak phenomenon.

INTRODUCTION

One of the most important and characteristic components of the Mediterranean diet is the olive oil, while *Olea europea L.* (Oleace) cultivation represents one of the most important and traditional agricultural activities in the Mediterranean countries. It is well known that olive oil promotes health and its attribution to the Mediterranean diet is very high. The beneficial effects of olive oil consumption appear to be due in part to the its phenolic fraction¹. While the olive fruits are typically used for human consumption, olive leaves are also a rich source of the same type of phenolic compounds: oleuropeosides, flavones, flavonols, flavan-3-ols and substituted phenolics such as hydroxytyrosol and tyrosol². Among them, oleuropein (OE) and hydroxytyrosol are the most abundant phenolic compounds found in olive leaves³.

OE is a nontoxic natural secoiridoid glycoside, occurring mainly in the *Olea* genus of the Oleaceae family and it is present in large amounts in olive leaves but in smaller quantities in olive oil and table fruits. It is the most well studied compound in olive cultivars^{4,5} and as one of the most prominent compounds of the Mediterranean diet; it has been proved to exhibit protective activity against an array of common chronic pathological conditions^{6,7}, e.g. antioxidant⁸, antimicrobial⁹, cardioprotective^{10,11}, anti-ischemic¹², antiatherogenic¹³, anti-inflammatory^{14,15}, antidiabetic¹⁶, anti-atherosclerotic^{17,18} etc.

The biological effects of a compound are related, to the time course of its concentration or its active metabolites in the blood stream¹⁹. Therefore, bioavailability studies are key points in clinical studies and are considered critical to establish the effect of a compound in the body²⁰. Drug bioavailability is influenced by several factors such as the administration route, the genetic phenotype, the age, the gender,

the food interaction and the feeding condition. Studies concerning the influence of these factors on the phenolics bioavailability are scarce. Significant number of studies indicate an influence of food matrices, excipients and carriers, age, sex and hormonal status²¹⁻²³. Recently, a human absorption and metabolism study of OE and its degradation product, hydroxytyrosol, administered as olive leaf extract, was carried out²⁴. The form of OE administration and the delivery method, as well as the gender of the subjects, were found to affect its absorption and metabolism. However, studies focused on the bioavailability of OE, administered as an olive leaf extract, in young healthy male athletes have not been conducted so far. This is of particular interest, as the antioxidant effect of OE is considered to inhibit the imposed alteration of the redox status of the athletes due to exercise.

Many analytical methods have been developed for the determination of OE in various matrices such as olive oil²⁵, table olives²⁶, byproducts²⁷ and biological fluids²⁸. Thus, OE has been analysed employing high performance liquid chromatography coupled with ultra violet (HPLC-UV)²⁹, HPLC-fluorescence³⁰ and HPLC coupled with tandem mass spectrometry (HPLC-MS/MS)³¹. Other methods besides liquid chromatography have also been employed such as gas chromatography MS/MS (GC-MS)³², capillary electrophoresis³³ and Fourier transform infra red (FT-IR)³⁴. However no analytical methodologies have been developed and validated for the quantification of OE in human serum, using state of the art chromatography techniques hyphenated to a high resolution tandem mass spectrometry (orbitrap) (UHPLC-HRMS/MS). This could be considered of high importance, as the new chromatographic techniques enhance chromatographic resolution and decrease analysis time, whereas high resolution mass spectrometry technologies enhance the specificity and specificity of the employed methodologies. For this purpose, a novel,

rapid and high throughput UHPLC-HRMS/MS analytical methodology, employing an orbitrap analyzer, was developed and validated according to the International Conference on Harmonisation (ICH) Q2 (R1)³⁵ requirements, the U.S. Department of Health and Human Services Food and Drug Administration (FDA)³⁶ and the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP). This methodology allows for the targeted quantification of OE in human serum after administration of an olive leaf extract enriched in OE, in young healthy male athletes.

MATERIALS AND METHODS

Pharmacokinetic study design

The pharmacokinetic study (PK) was conducted in a clinical environment under medical supervision, in the Aristoteleio University of Thessaloniki, Greece. In this study, two young healthy physical education male students, namely PK_A and PK_B, in the age range of 20-22 years, received an olive leaf extract enriched in OE. Thus 1200 mg of Benolea[®] EFLA[®]943 (in the form of four 300 mg capsules), containing 204.4 mg OE (see Chapter “Development and validation of a combined methodology for assessing the total quality control of herbal medicinal products - Application to oleuropein preparations”) was orally administered and subsequently they provided 10 ml of venous blood every 15 min between 0 and 1 hour, every 30 min between 1 and 3 hours, and every 1 hour between 3 and 6 hours through a venous catheter. Baseline sampling has also been performed before administration of Benolea[®]. The subjects were required to avoid intense physical activity, intake of alcohol, pain relief tablets,

or nutritional supplements the day before the administration and blood sampling. They were also instructed not to eat or drink after 10 pm the day before.

The protocol was approved by the medical ethical committee of the Aristotle University of Thessaloniki and has been conducted in accordance with the ICH-GCP guidelines (ICH GCP, 1996). All participants gave written informed consent in order to participate in the current study.

Biofluid collection

Blood samples were collected in appropriate tubes. Serum was separated immediately after the collection by centrifugation of the blood samples at 1800 g for 10 min at room temperature and 500 μ L aliquots were stored at -80°C until subsequent analysis. Serum samples were used for the targeted quantification of OE described below.

Samples

Twenty-four serum samples (12 from each subject) were collected and analyzed for the PK study. A list of samples along with their corresponding labels is tabulated in Table 1.

Time points (h)/Subjects	PK _A	PK _B
0	PK _A 0	PK _B 0
0,25	PK _A 0.25	PK _B 0.25
0,5	PK _A 0.5	PK _B 0.5
0,75	PK _A 0.75	PK _B 0.75
1	PK _A 1	PK _B 1
1,5	PK _A 1.5	PK _B 1.5
2	PK _A 2	PK _B 2
2,5	PK _A 2.5	PK _B 2.5
3	PK _A 3	PK _B 3
4	PK _A 4	PK _B 4
5	PK _A 5	PK _B 5
6	PK _A 6	PK _B 6

Table 1. A list of samples from the two individuals PK_A and PK_B at the different time points along with their corresponding labels.

Chemicals and reagents

The analytical reference standard of OE for the targeted quantification study was isolated from olive leaves following a previously described procedure³⁷ and its molecular structure is depicted in Figure 1. No specific permissions were required as the examined plant is an endemic species widely spread in the Mediterranean region, as well as in Australia, the American continent, Africa and Asia. Furthermore *Olea europea* is not included in any list of endangered or protected species. Its purity was found to be more than 98% (HPLC-UV) and its structure was identified by High Resolution Mass Spectrometry (HRMS) and 1 & 2D NMR. The internal standard (IS), 2,4-dinitrophenol (Figure 1), was purchased from Sigma-Aldrich (Darmstadt, Germany). All solvents used in this study were of LC-MS grade. Acetonitrile (ACN), methanol (MeOH), water and formic acid, were purchased from Fluka/Riedel-de Haën (Switzerland). Human serum for the validation of the PK study was obtained from healthy volunteers under written consent.

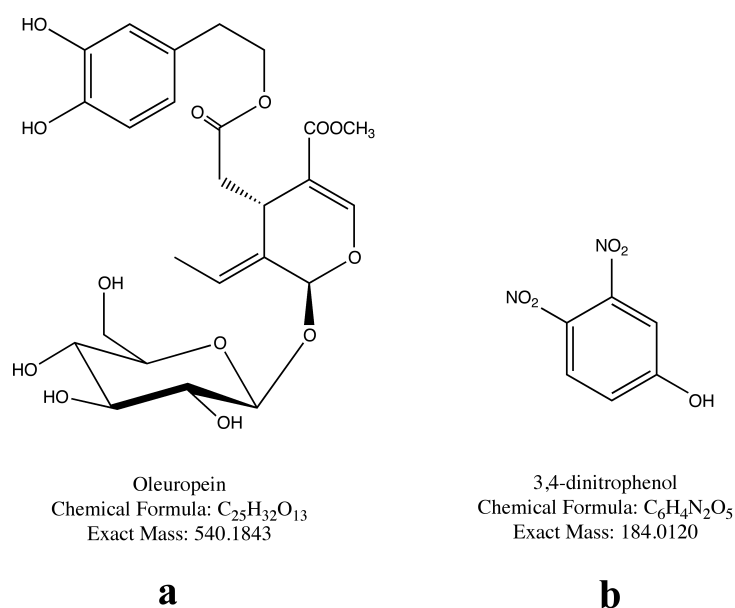


Figure 1. Structures of OE (a) and IS (b).

Instrumentation

UHPLC-ESI-HRMS/MS analysis for the PK study of OE (targeted quantification of OE) was performed on a hybrid LTQ Orbitrap Discovery XL mass spectrometer (Thermo Scientific, Germany) coupled to an Accela HPLC system (Thermo Scientific, Germany) equipped with a binary pump, an autosampler, an online vacuum degasser and a temperature-controlled column compartment and by applying the selected reaction monitoring (SRM) methodology. Centrifuging of the serum samples was performed in a Mikro 200R centrifuge (Hettich Lab Technology, Germany). Evaporation of the serum samples was performed with the aid of a GeneVac HT-4X EZ-2 series evaporator Lyospeed ENABLED (Genevac Ltd, UK).

Sample pretreatment

In the current PK study, a total of 12 serum samples (PK0, PK0.25, PK0.5, PK0.75, PK1, PK1.5, PK2, PK2.5, PK3, PK4, PK5 and PK6) from each subject (PK_A and PK_B) at the corresponding time points 0.0, 0.25, 0.5, 0.75, 1.0, 1.5, 2.0, 2.5, 3.0, 4.0, 5.0, 6.0 h, were prepared for UHPLC-HRMS/MS analysis by solid phase extraction (SPE). Thus, the cleanup of the serum samples was performed using Oasis HLB[®] Cartridges (30 mg, 1 cm³) obtained from Waters (Milford, MA, USA). Each cartridge was conditioned with 500 µL MeOH and equilibrated with 1000 µL of water. A 65 µL aliquot of serum sample spiked with 10 µL of an 1 µg/mL IS solution was carefully loaded onto the cartridge. A washing step with 200 µL of aq. formic acid 0.1% (v/v) solution followed, and the SPE cartridge was dried under vacuum for 1 min. The analytes were eluted with 1000 µL of ACN:MeOH (50:50) and the eluate was evaporated to dryness by means of the GeneVac apparatus. The residue was

reconstituted of 65 μ L initial mobile phase [aq. 0.1% formic acid:ACN, 9:1 (v/v)] and was analyzed with the developed UHPLC-ESI-HRMS/MS method described below.

Quantification of OE in serum

UHPLC-ESI-HRMS/MS analysis. An RP-C18 Hypersil Gold column (50 x 2.1 mm, 1.9 μ m) (Supelco Analytical, Germany) was used at a flow rate of 0.45 mL/min for the chromatographic separation of OE and IS. The mobile phase consisted of solvents A: aq. formic acid 0.1% (v/v) and solvent B: ACN. A gradient method (total run time of 2.7 min including the equilibration time) was used as follows: 0 to 0.4 min: 90% A: 10% B, 0.4 to 0.7 min: 60% A: 40% B, 0.7 to 1.9 min: 60% A: 40% B, 1.9 to 2.5 min: 32% A: 68% B, 2.5 to 2.7 min: 90% A: 10% B. The column temperature was kept at 40 °C throughout all experiments while the autosampler tray temperature was set at 4 °C. The injection volume was 5 μ L.

The ESI-HRMS/MS parameters for ion generation, ion transmission and ion detection procedures (capillary temperature, capillary voltage, spray voltage, sheath gas and auxiliary gas) were optimized for the SRM mode in order to achieve optimal sensitivity and selectivity. Thus the ESI-HRMS/MS was operated in the negative ion mode under the following conditions: capillary temperature, 356°C; capillary voltage, 11 V; tube lens, -49 V; source voltage, 3.10 kV; sheath gas flow, 30 arb. units; aux gas flow, 10 arb. units. Analysis was performed in the SRM mode, using a resolution of 30000 whereas the isolation width was set at 2 amu. All compounds were analyzed as their respective deprotonated molecular ions $[M-H]^-$. The UHPLC-HRMS parameters used in the PK study of OE are summarized in Table 2.

LC parameters

Column: RP-C18 Hypersil Gold column (50 x 2.1 mm, 1.9 μm) reversed phase

Mobile phase: aq. formic acid, 0.1% (v/v) (A) and ACN (B)

Flow rate: 0.45 mL min⁻¹

Column temperature: 40 °C

Tray temperature: 4 °C

Injection volume: 5 μL .

Gradient Program

Time (min)	A%	B%
0	90	10
0.4	60	40
0.7	60	40
1.9	32	68
2.5	90	10
2.7	90	10

MS parameters

ESI negative mode

30000 resolution

Centroid mode

Capillary temperature (°C): 356

Capillary voltage (V): 11

Tube lens (V): -49

Spray voltage (kV): 3.10

Sheath gas flow (arb. units): 30

Aux gas flow (arb. units): 10

Table 2: UHPLC-HRMS (orbitrap) parameters used in PK study of OE. Gradient elution program and mass spectrometry parameters applied for UHPLC-HRMS (orbitrap) analysis in (-) ESI mode.

Preparation of standard solutions, calibration curve and quality control (QC)

samples. Stock solutions of the analyte (OE) and the IS, were prepared at a concentration level of 1 mg/mL in MeOH and stored in the dark at 4 °C. Working solutions were daily prepared by diluting appropriate volumes of stock solutions in the initial mobile phase ratio to obtain 100, 10 and 1 $\mu\text{g/mL}$ solutions for the analyte and the IS. The IS working solution was prepared by diluting an appropriate volume of the corresponding stock solution in the initial mobile phase ratio, in order to obtain a 10 $\mu\text{g/mL}$ concentration level. The dynamic range of the OE calibration curve was 0.001 - 5 $\mu\text{g/mL}$ (0.001, 0.005, 0.01, 0.05, 0.075, 0.1, 0.4, 0.5, 1, 2, 3 and 5), whereas

the concentration of IS was kept at 1 µg/mL. Different sets of solutions were prepared individually, in order to be used as QC samples which were prepared at four concentration levels, a low (LQC 0.013 µg/mL), a medium1 (MQC1 0.2 µg/mL), a medium2 (MQC2 0.75 µg/mL) and a high (HQC 2.5 µg/mL).

Quantification. Quantification was performed employing the SRM negative ion mode for OE (MW: 540.1837 g/mol) and IS (MW: 184.0115 g/mol). The ion beams (m/z 539.1759 for OE and m/z 183.0036 for IS) were optimized by adjustment of the ion generation, ion transmission and ion detection region parameters for each compound while the transitions of m/z 539.1 (isolation width 2 amu) \rightarrow 377.1234 \pm 0.0005 (for OE) and m/z 183.1 (isolation width 2 amu) \rightarrow 153.0068 \pm 0.0005 (for IS) were optimized by adjustment of the normalized collision energy using the collision induced dissociation (CID) option. The IS methodology has been employed for the quantification of OE, based to the calculation of the peak-area ratio of the analyte vs. that of the IS.

Validation of the bioanalytical methodology

The method validation was performed in accordance to ICH Q2 R1³⁵, the FDA³⁶ and the EMA CHMP guidelines for bioanalytical methodologies³⁸ by evaluating the stability, specificity, linearity, precision, accuracy, lower limit of quantification, the robustness and the system suitability parameters.

Calibration curves construction. Three kinds of calibration curves were constructed, i.e. in solvent, in pre- and in post-spiked plasma in order to estimate validation parameters.

Regression model. Calibration curves ranging from 0.001 to 5 $\mu\text{g/mL}$ were analyzed on three separate days and fitted using quadratic regression analysis employing $1/x$ weighting. The methodology exhibited acceptable accuracy for the concentration range studied (back calculated values were within 15% of nominal values). Pre- and post-spiked serum samples were prepared for UHPLC-(-)ESI-HRMS/MS analysis by SPE as described above.

Specificity-Recovery-Matrix Effect. Specificity has been assessed by the injection of five individually prepared blank serum samples and the presence of possible interferences at the corresponding t_R of the analyte and the IS. The blank serum samples were prepared using the described SPE methodology. The matrix effect was evaluated comparing the slope of two calibration curves, one prepared in aq. 0.1% formic acid/ACN 90/10 (v/v) and the other prepared in serum, as described above. Recovery was determined at the four different QC levels (LQC, MQC1, MQC2 and HQC), comparing the serum concentration of OE and IS for pre-spiked samples ($n=5$) with that of post-spiked samples, using the aforementioned sample preparation methodology.

Lower limit of quantification. The LLOQ was calculated as the concentration affording an OE signal with $\%RSD < 20$ and $\%Error < 20$ ($n=5$). In addition the signal to noise ratio (S/N ratio) criterion was set as $S/N > 5$.

Repeatability, intermediate precision and accuracy. Five replicates of each QC sample at the four QC concentration levels (LQC, MQC1, MQC2 and HQC) were determined in order to evaluate the repeatability (within-run) and the intermediate

precision (between-run) of the assay. The results were expressed as the standard deviation (SD) and the relative standard deviation (%RSD) for all QC samples. The rejection criterion was set as %RSD<20 for the LLOQ and <15 for the QC samples. Accuracy was determined for the same samples as the % standard error (% Er) between the mean concentrations calculated and the nominal concentrations. The rejection criterion was set as %Er<20 for the LLOQ and <15 for the QC samples.

Robustness. For the establishment of the method's robustness, experimental conditions were deliberately slightly altered and several parameters were calculated. This was achieved by altering three conditions by $\pm 5\%$: setting the flow rate of the mobile phase at 0.4275 mL/min and 0.4725 mL/min, the column temperature at 38 °C and 42 °C and the capillary temperature at 373.8 °C and 338.2 °C. All experiments were performed at the MQC2 level (n=5) calculating the chromatographic peak area of OE, the retention time (t_R), the 10% peak asymmetry, the 10% peak width and the USP theoretical plates number. The results were expressed as %RSD.

RESULTS AND DISCUSSION

UHPLC-(-)ESI-SRM conditions

Chromatographic conditions were examined by varying parameters such as mobile phase (organic modifiers and buffers), gradient elution type, flow rate, type of column chemistry (C18 & C18-NPS, C8, and Phenyl) and column geometries (100 x 2.1 mm, 1.9 μ m; 50 x 2.1 mm, 1.9 μ m, 100 x 2.1 mm, 1.7 μ m). Employing the described methodology on a Hypersil Gold C18 50 x 2.1 mm, 1.9 μ m reversed phase column,

sharp peaks were achieved leading to enhanced signal intensity and adequate resolution of the analyte from possible interfering endogenous serum components. Furthermore the total analysis time did not exceeded 2.7 min, a fact that makes the analysis very fast, affording a rapid chromatographic methodology. Resolution between all the chromatographic peaks was greater than 2, which ascertains that the OE and IS are baseline separated.

Mass spectroscopic parameters (Table 2) were optimized in order to achieve the highest signal for the OE and IS. In addition, optimization of the collision energy (Table 3) of the tested compounds afforded well-described transitions and highly stable SRM signal for OE.

HRMS/MS	Name	Parent mass	Product mass (center mass \pm mass range)	Isolation width (amu)	CID (V)
SRM	OE	539.1	377.1234 \pm 0.0005	2	21
SRM	IS	183.1	153.0068 \pm 0.0005	2	36

Table 3. Mass spectrometry parameters - SRM conditions.

Sample Preparation - Extraction Protocol

For the efficient cleanup of the PK serum samples and the isolation of the analyte, many protocols were evaluated for the serum sample preparation, such as protein precipitation, SPE and liquid - liquid extraction. Various solvents such as MeOH, ethanol, acetone, propanol, ethyl acetate and chloroform were tested for the liquid – liquid extraction and protein precipitation applying different protocols. In SPE methodologies, many sorbents such as C18, C8 and hydrophilic-lipophilic balance

(HLB) with different particle sizes (60, 30, 25, 15 and 5 μm) and quantities (100, 50, 30 and 10 mg/cartridge) were evaluated applying various elution (MeOH, ethanol, acetone, ACN or mixtures of them) and wash solvents (water or aqueous buffers) and various volumes as well as. SPE using Oasis HLB[®] Cartridges (30 mg, 1 cm^3) was chosen as preparation/extraction method of the pharmacokinetic serum samples, as it ensured good extraction efficiency and high reproducibility for the OE and IS and no interference from many co-eluted endogenous serum components was detected, whereas the corresponding time remained at a minimum in terms of sample preparation.

Generally, HLB modifiers are styrene-based polymers developed for the SPE of a wide range of acidic, basic, and neutral compounds from various aqueous matrices samples, such as serum, plasma and urine. They have become the phase of choice for the development of new extraction methods that recover, concentrate, and cleanup small molecules (e.g., OE and HT) from complex aqueous sample matrixes (e.g., biological fluids, water, etc.). HLB's are typically comprised of a hydrophobic component and a hydrophilic component and for that reason they offer distinct advantages over traditional silica-based reversed phase SPE chemistries. The compounds of interest are retained in the cartridge and consequently they are eluted with an aqueous miscible solvent such as methanol and/or acetonitrile. In addition, because hydrophilic polymer phases contain polar functionalities, they maintain their capability for higher retention and excellent recoveries even if the sorbent runs dry, which means there is no need to take extraordinary steps to keep the sorbent beds from drying out during the critical steps prior to sample loading and this fact often makes associated extraction methods more reproducible and robust.

Quantification

Quantification was performed in the SRM negative ion mode for the OE and IS. In this study, 2,4-dinitrophenol was chosen as the IS for the LC-SRM method since it exhibited high and repeatable recovery and chromatographic behavior with the analyte, it has been shown to be stable under the described experimental conditions (chemical and signal stability), it has not been reported up to date as an endogenous serum metabolite, and it did not interfere with the analysis of the compound of interest. Figure 2a shows a typical UHPLC-SRM chromatogram of a spiked with OE and IS serum sample, at a concentration of the LQC (0.013 $\mu\text{g/mL}$), while Figure 2b shows a typical UHPLC-SRM chromatogram of a spiked with OE and IS serum sample, at a concentration of the 1 $\mu\text{g/mL}$. Blank extracts (non spike serum sample) that were processed using the described methodology, produced no discernible peaks for OE and IS. A representative blank serum extract is shown in Figure 2c.

Following the optimization of precursor ions (m/z 539.1759 for OE and m/z 183.0036 for IS), the product ion spectra were obtained for the specific precursor ions that yielded ions m/z 377.1234 for OE and m/z 153.0068 for IS. Thus, the transitions of m/z 539.1 \rightarrow 377.1234 \pm 0.0005 for OE and m/z 183.1 \rightarrow 153.0068 \pm 0.0005 for IS were chosen for the quantification studies using an isolation width of 2 amu, since upon optimization of parameters collision energy, they yielded the highest sensitivity. Optimized values for the collision energy of the tested compounds are shown in Table 3.

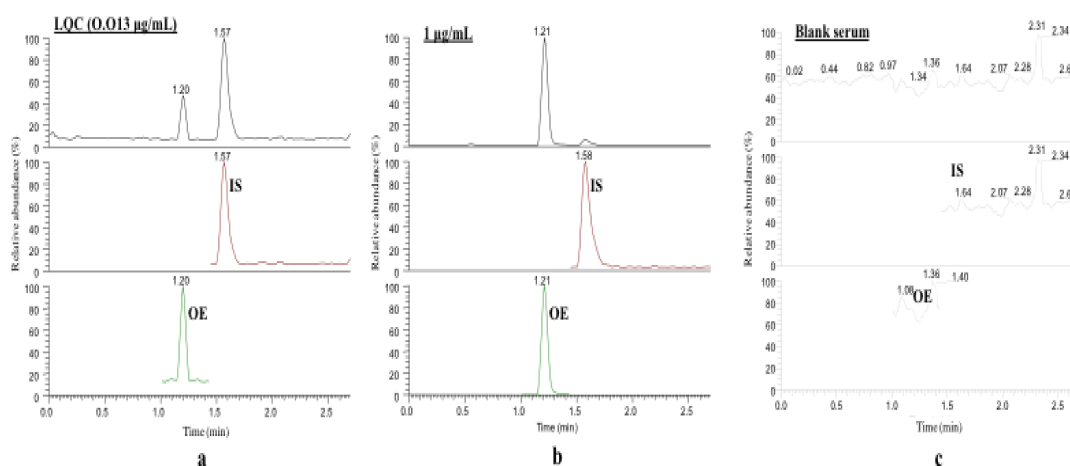


Figure 3. Typical chromatograms of the targeted quantitation of OE using the proposed methodology. **a.** UHPL-SRM chromatogram of a spiked with OE and IS serum sample, at a concentration of the LQC (0.013 $\mu\text{g/mL}$), **b.** UHPL-SRM chromatogram of a spiked with OE and IS serum sample, at a concentration of 1 $\mu\text{g/mL}$, **c.** UHPL-SRM chromatogram of a blank serum sample (non spike sample).

Validation of the bioanalytical methodology

Specificity - Recovery - Matrix Effect. The specificity was evaluated and no interference from the endogenous serum metabolites was found to the corresponding t_R of the analyte and IS. The matrix effect was evaluated by comparing the slopes of two calibration curves, one prepared in aq. 0.1% formic acid/ACN 90/10 (v/v) and the other prepared in serum as described above. The slope ratios were found to be 0.95 for OE. The data exhibit that the analyte signal does not essentially alter when the calibration curves was prepared in serum.

Recovery was determined at the four different QC levels (LQC, MQC1, MQC2 and HQC) ($n=5$) and calculated as $\%R = \frac{\text{pre spiked concentration of analyte}}{\text{post spiked concentration of analyte}} \times 100$. The calculated recoveries of OE at the four different QC levels (0.013, 0.2, 0.75 and 2.5 $\mu\text{g/mL}$) exhibited high recovery rates ($>95\%$), render the method suitable for the analysis of OE in human serum.

Regression model. A quadratic model was fitted to the OE calibration curves over the range 0.001 to 5 µg/mL, with a correlation coefficient (R^2)>0.9995. Typical equation for the calibration curve used for the quantitation of OE in the pharmacokinetic serum samples is shown in Table 5.

Analyte	Equation	R^2	Weighting
OE	$C_{OE} = 2.1 \cdot 10^{-2} (\pm 1.1 \cdot 10^{-4}) \text{ Area Ratio} - 3.3 \cdot 10^{-6} (\pm 4.6 \cdot 10^{-7}) \text{ Area Ratio}^2 - 0.03 (\pm 0.03)$	0.9995	$1/X$

Table 5. Typical equation for the calibration curve of OE.

Quantitation and detection limit. The LOQ of OE was found to be 0.001 µg/mL and is determined with acceptable accuracy and precision at this concentration level according to the ICH regulations. The LOQ for the analyte defined as the concentration that yielded a peak with a signal/noise ratio of 10 and at least 3 times the response compared to that blank extracts (samples not spiked with the analyte). Furthermore the accuracy, expressed as the relative percentage error %Er was found to be lower than 3.8% whereas the precision expressed as the %relative standard deviation %RSD was found to be less than 14.9% at the LOQ level. The data are shown in Table 6. Furthermore, the LOD was found to be 0.0003 µg/mL. The data shows that the described method is of acceptable sensitivity for the analysis of OE in human serum.

(n=5)	Repeatability	Intermediate Precision	Intra Accuracy Assay
Nominal Value (µg/mL)	OE [Mean ± SD ^a , (%RSD ^b)]	OE [Mean ± SD ^a , (%RSD ^b)]	OE %Error ^c
LQC (0.013)	0.0114 ± 0.0022 (19.3)	0.0157 ± 0.0008 (0.48)	11.98
MQC1 (0.2)	0.197 ± 0.025 (13.1)	0.2720 ± 0.0129 (4.75)	1.54
MQC2 (0.75)	0.813 ± 0.082 (10.2)	1.020 ± 0.101 (10.12)	-8.50
HQC (2.5)	2.28 ± 0.21 (9.2)	1.948 ± 0.058 (3.00)	8.72

^aStandard deviation, ^b% Relative standard deviation, ^c% Standard error

Table 6. Validation results of the proposed bioanalytical methodology.

Repeatability, intermediate precision and accuracy. Repeatability, intermediate precision and accuracy were determined by analyzing replicates (n=5) at the four QC concentration levels (0.013, 0.2, 0.75 and 2.5 µg/mL). The repeatability derived from all QC levels did not exhibit values more than 19.3% whereas the intermediate precision did not exhibit values more than 10.1%. The %Er derived was found to be within ±12%. The results of accuracy and precision of the method are summarized in Table 6 and proved that the described method can reliably use for the quantitation of OE in the human serum samples.

Robustness. To ensure the insensitivity of the HPLC method to minor changes in the experimental conditions, it is important to demonstrate the robustness of the analytical method. None of the alterations caused a significant change more than 2% for all the parameters examined, proving that the method is sufficiently robust to be used for quality control measurements. Thus, %RSD of the resolution between OE and IS was found to be less than 2% when the conditions were deliberately altered

according to the robustness scheme. Changing the flow rate of the mobile phase from 0.45 mL/min to 0.4275 mL/min and 0.4725 mL/min the %RSD of the resolution between peaks did not exceed 3.6%. It was also noticed that changing the column temperature from 40 °C to 38 °C and 42 °C, the %RSD of the resolution between peaks was found to be less than 1.8%. Furthermore, by altering the capillary temperature from 356 °C to 338.2 °C and 373.8 °C the peak area for all peaks was not significant affected.

Measurement of OE in serum samples

Human serum samples from the PK study were prepared and measured by the developed methodology. Quantification of OE in the serum samples of the two subjects (PK_A and PK_B) at the different time points (0.0, 0.25, 0.5, 0.75, 1.0, 1.5, 2.0, 2.5, 3.0, 4.0, 5.0, 6.0 h) was based on standard curve that was prepared on the day of analysis. Kinetic parameters, i.e. T_{\max} (h) and C_{\max} (ng/mL) are presented in Table 7. The areas under the concentration-time curves (AUCs) values were calculated using the trapezoidal method of integration and are also presented in Table 7.

Kinetic Parameters/Subjects	PK_A	PK_B
$t_{\max 1}$ (h)	0.5	0.5
$C_{\max 1}$ (ng/mL)	118.6	33.6
$t_{\max 2}$ (h)	2	4
$C_{\max 2}$ (ng/mL)	146.9	109.5
AUC_{0-6h} (h ng/mL)	145.9	164.9

Table 7. Kinetic parameters following oral administration of Benolea® (204.4 mg OE) in human serum. t_{\max} , C_{\max} and AUC values, for the two cohorts PK_A and PK_B are given.

Graphical representations of the OE concentration circulating in human serum of the two cohorts (PK_A and PK_B) can be found in Figures 4 and 5, respectively. From the aforementioned results, it is evident that the PK of OE follows a multiple peak phenomenon (Figures 4 and 5).

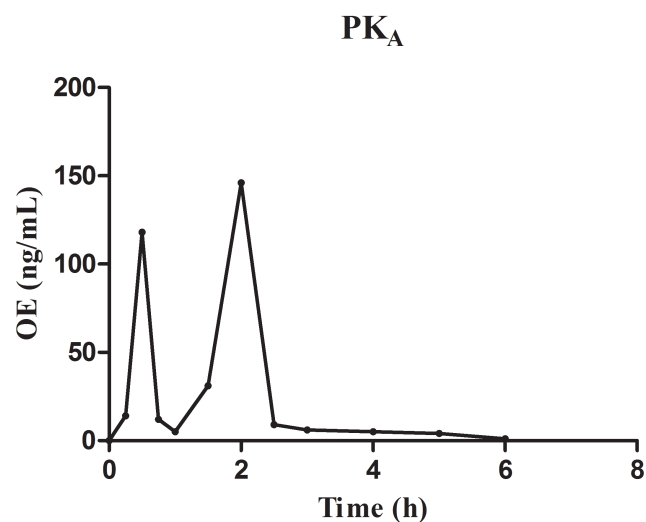


Figure 4. Serum concentration–time profile (pharmacokinetic profile) of OE in subject PK_A after oral administration of Benolea® (204.4 mg OE).

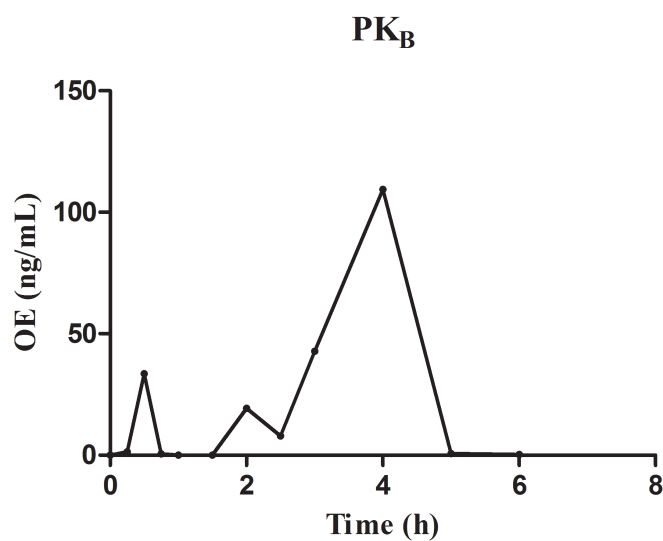


Figure 5. Serum concentration–time profile (pharmacokinetic profile) of OE in subject PK_B after oral administration of Benolea® (204.4 mg OE).

This effect randomly appears in PK studies and could be attributed to a variety of reasons. Thus, it can be attributed to enterohepatic recirculation, delayed gastric emptying, and variability of absorption³⁹. Furthermore, other factors related to the formulation, be it the drug chemical entity itself or other formulation-related factors such as the excipients incorporated into the product design, could also contribute to this phenomenon. Finally, other factors such as hormonal and dietary factors and biochemical differences in the regional areas of the gastrointestinal tract, could also play a key role⁴⁰.

In the current case, it has to be noted that the AUCs of the two subjects PK_A and PK_B are similar (Table 7), implying similar degree of absorption. The first absorption peak (C_{max1}, t_{max1}) seems to exist at the same time, 0.5 h, for both subjects, which could be contributed to the first stage of absorption in the stomach. The second absorption peak (T_{max} = 2 h and C_{max} = 146.9 ng/mL for PK_A and T_{max} = 4 h and C_{max} = 109.5 ng/mL for PK_B) could presumably be attributed to the enterohepatic reabsorption of OE and the differences in C_{max} and T_{max} in the two subjects, reflects the differences to the physiology of the two subjects. Another concern is the possible differentiation in the PH conditions in the stomach and the large intestine, which could alter the absorption rate extensively. It is also essential to recognize the complex nature of the administered formulation, which could presumably impose some bioavailability/PK concerns. Concluding, more experiments should be performed with the contribution of a larger number of participants, in order to gain deeper insight regarding the *per os* absorption of OE.

CONCLUSION

To our knowledge this is the first time that an UHPLC-HRMS/MS (orbitrap) methodology for the quantification of OE was developed and validated in human serum. This method was successfully applied to bio-analysis involving the absorption/kinetic study of the molecule in two human subjects after oral administration of Benolea® (204.4 mg OE). The results of absorption study indicated that OE was quickly absorbed showing a multiple peak phenomenon, slightly different in the two subjects. Nevertheless, the AUCs of the two subjects, PK_A and PK_B , were found to be similar. It is evident that more experiments are needed and ideally with a larger number of participants.

REFERENCES

1. Cicerale, S., Conlan, X. A., Sinclair, A. J. & Keast, R. S. J. Chemistry and health of olive oil phenolics. *Crit. Rev. Food Sci. Nutr.* **49**, 218–36 (2009).
2. García-Villalba, R., Larrosa, M., Possemiers, S., Tomás-Barberán, F. a & Espín, J. C. Bioavailability of phenolics from an oleuropein-rich olive (*Olea europaea*) leaf extract and its acute effect on plasma antioxidant status: comparison between pre- and postmenopausal women. *Eur. J. Nutr.* (2013). doi:10.1007/s00394-013-0604-9
3. El, S. N. & Karakaya, S. Olive tree (*Olea europaea*) leaves: potential beneficial effects on human health. *Nutr. Rev.* **67**, 632–8 (2009).
4. Japón-Luján, R. & Luque de Castro, M. D. Superheated liquid extraction of oleuropein and related biophenols from olive leaves. *J. Chromatogr. A* **1136**, 185–91 (2006).
5. Romani, A., Mulinacci, N., Pinelli, P., Vincieri, F. F. & Cimato, A. Polyphenolic content in five tuscan cultivars of *Olea europaea* L. *J. Agric. Food Chem.* **47**, 964–7 (1999).
6. Renaud, S. *et al.* Cretan Mediterranean diet for prevention of coronary heart disease. *Am. J. Clin. Nutr.* **61**, 1360S–1367S (1995).
7. Galli, C. & Visioli, F. Antioxidant and other activities of phenolics in olives/olive oil, typical components of the Mediterranean diet. *Lipids* **34 Suppl**, S23–6 (1999).
8. Kontogianni, V. G. & Gerothanassis, I. P. Phenolic compounds and antioxidant activity of olive leaf extracts. *Nat. Prod. Res.* **26**, 186–9 (2012).
9. Pereira, A. P. *et al.* Phenolic compounds and antimicrobial activity of olive (*Olea europaea* L. Cv. Cobrançosa) leaves. *Molecules* **12**, 1153–62 (2007).
10. Omar, S. H. Cardioprotective and neuroprotective roles of oleuropein in olive. *Saudi Pharm. J. SPJ Off. Publ. Saudi Pharm. Soc.* **18**, 111–121 (2010).
11. Keys, A. *et al.* The diet and 15-year death rate in the seven countries study. *Am. J. Epidemiol.* **124**, 903–15 (1986).
12. Andreadou, I. *et al.* The olive constituent oleuropein exhibits anti-ischemic, antioxidative, and hypolipidemic effects in anesthetized rabbits. *J. Nutr.* **136**, 2213–9 (2006).
13. Visioli, F. & Galli, C. Antiatherogenic components of olive oil. *Curr. Atheroscler. Rep.* **3**, 64–67 (2001).
14. Scoditti, E. *et al.* Mediterranean diet polyphenols reduce inflammatory angiogenesis through MMP-9 and COX-2 inhibition in human vascular endothelial cells: a potentially protective mechanism in atherosclerotic vascular disease and cancer. *Arch. Biochem. Biophys.* **527**, 81–9 (2012).

15. Poudyal, H., Campbell, F. & Brown, L. Olive leaf extract attenuates cardiac, hepatic, and metabolic changes in high carbohydrate-, high fat-fed rats. *J. Nutr.* **140**, 946–53 (2010).
16. Jemai, H., El Feki, A. & Sayadi, S. Antidiabetic and antioxidant effects of hydroxytyrosol and oleuropein from olive leaves in alloxan-diabetic rats. *J. Agric. Food Chem.* **57**, 8798–804 (2009).
17. Impellizzeri, D. *et al.* The effects of oleuropein aglycone, an olive oil compound, in a mouse model of carrageenan-induced pleurisy. *Clin. Nutr.* **30**, 533–40 (2011).
18. Impellizzeri, D. *et al.* The effects of a polyphenol present in olive oil, oleuropein aglycone, in an experimental model of spinal cord injury in mice. *Biochem. Pharmacol.* **83**, 1413–26 (2012).
19. Tian, F. *et al.* Comparative study on pharmacokinetics of a series of anticholinergics, atropine, anisodamine, anisodine, scopolamine and tiotropium in rats. *Eur. J. Drug Metab. Pharmacokinet.* (2014). doi:10.1007/s13318-014-0192-y
20. Chen, M. L. *et al.* Bioavailability and bioequivalence: an FDA regulatory overview. *Pharm. Res.* **18**, 1645–50 (2001).
21. Zoidou, E. *et al.* Oleuropein as a bioactive constituent added in milk and yogurt. *Food Chem.* **158**, 319–324 (2014).
22. Cassidy, A. *et al.* Factors affecting the bioavailability of soy isoflavones in humans after ingestion of physiologically relevant levels from different soy foods. *J. Nutr.* **136**, 45–51 (2006).
23. Roowi, S., Mullen, W., Edwards, C. A. & Crozier, A. Yoghurt impacts on the excretion of phenolic acids derived from colonic breakdown of orange juice flavanones in humans. *Mol. Nutr. Food Res.* **53 Suppl 1**, S68–75 (2009).
24. De Bock, M. *et al.* Human absorption and metabolism of oleuropein and hydroxytyrosol ingested as olive (*Olea europaea* L.) leaf extract. *Mol. Nutr. Food Res.* **57**, 2079–85 (2013).
25. Dierkes, G. *et al.* High-performance liquid chromatography-mass spectrometry profiling of phenolic compounds for evaluation of olive oil bitterness and pungency. *J. Agric. Food Chem.* **60**, 7597–606 (2012).
26. Zoidou, E. *et al.* Identification of Throuba Thassos, a traditional Greek table olive variety, as a nutritional rich source of oleuropein. *J. Agric. Food Chem.* **58**, 46–50 (2010).
27. Lozano-Sánchez, J. *et al.* Wastes generated during the storage of extra virgin olive oil as a natural source of phenolic compounds. *J. Agric. Food Chem.* **59**, 11491–500 (2011).
28. Bazoti, F. N., Gikas, E. & Tsiaropoulos, A. Simultaneous quantification of oleuropein and its metabolites in rat plasma by liquid chromatography electrospray ionization tandem mass spectrometry. *Biomed. Chromatogr.* **24**, 506–15 (2010).

29. Ortega-García, F. & Peragón, J. HPLC analysis of oleuropein, hydroxytyrosol, and tyrosol in stems and roots of *Olea europaea* L. cv. Picual during ripening. *J. Sci. Food Agric.* **90**, 2295–300 (2010).
30. Tan, H.-W., Tuck, K. L., Stupans, I. & Hayball, P. J. *Simultaneous determination of oleuropein and hydroxytyrosol in rat plasma using liquid chromatography with fluorescence detection.* *J. Chromatogr. B* **785**, 187–191 (2003).
31. Zhou, T. *et al.* Application of LC-MS/MS method for the in vivo metabolite determination of oleuropein after intravenous administration to rat. *Biomed. Chromatogr.* **25**, 1360–3 (2011).
32. Zafra-Gómez, a, Luzón-Toro, B., Jiménez-Díaz, I., Ballesteros, O. & Navalón, a. Quantification of phenolic antioxidants in rat cerebrospinal fluid by GC-MS after oral administration of compounds. *J. Pharm. Biomed. Anal.* **53**, 103–8 (2010).
33. Aturki, Z., Fanali, S., D’Orazio, G., Rocco, A. & Rosati, C. Analysis of phenolic compounds in extra virgin olive oil by using reversed-phase capillary electrochromatography. *Electrophoresis* **29**, 1643–50 (2008).
34. Aouidi, F., Dupuy, N. & Artaud, J. Rapid quantitative determination of oleuropein in olive leaves (*Olea europaea*) using mid-infrared spectroscopy combined with chemometric analyses. *Ind. Crop. ...* **37**, 292–297 (2012).
35. Version, C. S. 4 & 2005), P. G. dated 27 O. 1994 (Complementary G. on M. dated 6 N. 1996 incorporated in N. INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE. ICH HARMONISED TRIPARTITE GUIDELINE. VALIDATION OF ANALYTICAL PROCEDURES: TEXT AND METHODOLOGY Q2(R1). (2005).
36. Validation, B. M. *Guidance for Industry Bioanalytical Method Validation Guidance for Industry Bioanalytical Method Validation.* (2013).
37. Andreadou, I. *et al.* The Olive Constituent Oleuropein Exhibits Anti-Ischemic, Antioxidative, and Hypolipidemic Effects in Anesthetized Rabbits. *J. Nutr.* **136**, 2213–2219 (2006).
38. Wharf, C. & Kingdom, U. Guideline on bioanalytical method validation Guideline on bioanalytical method validation Table of contents. **44**, (2012).
39. Godfrey, K. R., Arundel, P. A., Dong, Z. & Bryant, R. Modelling the Double Peak Phenomenon in pharmacokinetics. *Comput. Methods Programs Biomed.* **104**, 62–9 (2011).
40. Davies, N. M., Takemoto, J. K., Brocks, D. R. & Yáñez, J. A. Multiple peaking phenomena in pharmacokinetic disposition. *Clin. Pharmacokinet.* **49**, 351–77 (2010).

4

Effect of supplementation with olive leaf extract enriched in oleuropein (Benolea® EFLA®943), on the metabolome and on the redox status of athletes' blood and urine. A metabolomic approach

This chapter will be submitted for publication

ABSTRACT

Oleuropein (OE) is a secoiridoid glycoside, which occurs mostly in the Oleaceae family presenting several pharmacological properties. *In vitro* assessments, as well as studies on experimental animal models, have shown that OE possesses among others, hypolipidemic and antioxidant properties. Based on these findings, several dietary supplements that contain olive leaf extracts enriched in OE, are commercially available in many countries. The aim of the current study was to examine the effect of the supplementation with such an extract (Benolea[®] EFLA[®]943), on the serum and urine metabolome of young healthy male athletes. For this purpose, applying a randomized, balanced, double-blind manner study, 9 young healthy males (physical education students) received either Benolea[®] or placebo for one week, followed by a two week washout period and subsequent dosing by the alternate scheme (crossover design). Urine and serum samples have been analyzed by UHPLC-HRMS, followed by evaluation of several multivariate methods for the data analysis. Finally, the data interpretation was completed by a multilevel metabolomic approach, as it found to be more efficient for such designs. Subsequent metabolic pathway analysis of the significant features discovered has revealed that the tryptophan and the acylcarnitines biochemistry have been mostly influenced. In addition, possible changes in the lipidemic profile and the blood and urine redox status were investigated.

INTRODUCTION

It is evident, that the nutraceutical market exploring the potential health benefits of olive products is expanding rapidly. As the concentration of the most well-known health promoting olive bioactives e.g. oleuropein, hydroxytyrosol etc., is generally higher in the leaves than in the industrial processed olive oil or in the table fruits, this oriented the interest in leaves' exploitation. Thus, olive leaves that were once discarded as by-products of tree pruning are now considered as a valuable source of high added value substances.

Oleuropein (OE) is a nontoxic natural secoiridoid glycoside, occurring mainly in the *Olea* genus of the Oleaceae family and it is present in large amounts in olive leaves but in smaller quantities in olive oil and table fruits. It is the most well studied compound in olive cultivars^{1,2} and as one of the most prominent compounds of the Mediterranean diet; it has been proved to exhibit protective activity against an array of common chronic pathological conditions^{3,4}, e.g. antioxidant⁵, antimicrobial⁶, cardioprotective^{7,8}, anti-ischemic⁹, antiatherogenic¹⁰, anti-inflammatory^{11,12}, antidiabetic¹³, anti-atherosclerotic^{14,15} etc. Especially concerning its antioxidant activity, several *in vitro* studies have demonstrated that OE has a high antioxidant activity comparable to a hydrosoluble analog of tocopherol⁹.

Recently, the European Food Safety Authority (EFSA) has endorsed the health claim that “the consumption of olive oil polyphenols contributes to the protection of blood lipids to oxidative damage”¹⁶. Thus, several formulations of olive leaf extracts enriched in OE are becoming commercially available in many countries (approximately 23), as food supplements or herbal medicinal products¹⁷. However, while the antioxidant health benefits of a Mediterranean diet rich in olive oil

polyphenols is well established, more clinical studies examining the effects of olive polyphenols supplementation on the redox status of athletes are needed.

Oxidative stress can be defined as the increase in intracellular steady state concentration of oxidants over physiological values^{18–21}. As physical exercise may enhance the oxidative stress of an organism due to the elevated catabolism that may produce large amounts of oxidized substances, athletes can be considered as an ideal study group. Following a feedback regulation mechanism athletes following a regular training program may exhibit a significant increase in antioxidant status comparing to healthy sedentary people²², due to an adaptative response to controlled physical activity. During a period of overload training, this adaptative response cannot control the free radical production, which could prove detrimental to the body²³. At the start of an exercise program, muscle damage in athletes has often occurred²⁴. Exercise-induced muscle damage is attributed to an exercise intensity-dependent increase in oxidative stress produced by mitochondrial process²⁵. Although small increases in exercise induced reactive oxygen species (ROS) are important for stimulating cellular growth and maximizing muscular force production, excessive accumulation leads to a pro-oxidant environment, which can damage DNA. Rapid recovery from the prior exercise is important for both beginners performing regular exercise to improve their health and athletes preparing for competition. Thus, antioxidant supplementation to attenuate exercise-induced muscle injury, inflammation, and pain may facilitate success in both health promotion and sport competition.

The World Health Organization (WHO) has introduced the Global Recommendations on Physical Activity for Health²⁶, which encourage people of all ages to begin moderately and gradually progress to higher levels of physical activity.

Antioxidants may be administrated before any exercise especially before high level training and results in the annihilation of the free radicals generated by the exercise that potentially overwhelm the defensive mechanisms, causing oxidative stress^{27,28}. Different types of supplementation such as vitamin E²⁹, vitamin C³⁰ or polyphenols revealed that it was possible to enhance the adapted antioxidant status. However in the same context studies that evaluated the impact of the antioxidant effect on the performance or endurance in athletes exhibited results that differed largely depending on the type of supplementation, subjects and protocol³¹.

A few studies have been performed in human subjects concerning the antioxidant effects of OE. Thus, olive (*olea europaea l.*) leaf polyphenols were found to improve insulin sensitivity in middle-aged overweight men in a randomized, placebo-controlled, crossover trial³², whereas a possible association between the consumption of a rich-polyphenol extra virgin olive oil and an enhancement of the endogen antioxidant system in healthy elderly people has been postulated³³. Furthermore, it has been demonstrated that young and middle-aged men and women that were consuming a moderate amount of EVOO regularly have increased their antioxidant status^{34,35}. The aforementioned positive effects of EVOO have been probably attributed to its high polyphenolic content and especially its high content in secoiridoid derivatives^{36,37}. Interestingly however, the effect of supplementation with olive leaf extract enriched in OE in relation with the blood and urine metabolome and the redox status of athletes have not been reported up to date. Thus up-to-date very limited data exists, concerning the effects of olive polyphenols, and especially of olive leaf extract enriched in OE, on the lipidemic profile, the blood and urine redox status and the serum and urine metabolome of young healthy male athletes. In this context, a

randomized, double-blinded, placebo-controlled, crossover trial encompassing supplementation with olive leaf extract enriched in OE has been designed.

Metabolomics can be defined as the study of the complete set of a biological entities' metabolites. In turn metabolites can be considered as the low-molecular-weight intermediates, which are context dependent, varying according to the physiology, developmental or pathological state of the cell, tissue, organ or organism. In other words metabolomics is the recording of all the possible circulating (blood), excreted (urine) or localized (tissues) metabolites, at least in two different states, one being the control, (e.g. dosed vs. non-dosed subjects) and the statistical evaluation of their differences. This approach aims towards revealing the metabolites that are up- or down- regulated due to the action of some kind of stress (physiological, environmental or even chemical/pharmaceutical). The main analytical tools used in metabolomics are liquid chromatography coupled to mass spectrometry (LC-MS) and NMR whereas the evaluation of the results is performed in the majority of instances by multivariate statistical analysis (MVA), e.g. PCA, PLS-DA etc.

The application of liquid chromatography coupled to high resolution mass spectrometry (LC-HRMS) offers significant advantages for metabolomics studies and is frequently considered as one of the gold standard techniques in metabolomics era. Fast analysis of a large variety of compounds using a small sample volume, little to minimal sample treatment, high reproducibility and high sensitivity jointly, and along with the capability of obtaining unique structural information even for unknown substances, render the technique indispensable for modern metabolomics.

Nowadays, crossover designs have started to be used in combination with 'omics' techniques, resulting in paired multi- variate data sets³⁸⁻⁴⁰. Metabolomic time series experiments with different subjects have also the same paired data structure^{41,42}, and

the analysis of such data sets can possibly be improved when exploiting the design underlying the study. However, in the analysis of these multivariate-paired data, the study design is not always considered.

One of the major challenges in crossover supplementation/nutrition metabolomic studies is the detection and identification of metabolites in different biofluids, such as serum and urine, which can be linked to possible changes in human metabolome after the supplementation period^{43–45}. In such studies (as in the current one), the metabolic responses observed are usually small, since the volunteers are in metabolic homeostasis whereas they are generally in a health state⁴⁶. Thus, considering the fact that the biological variation between the individuals is usually much larger than the effects caused by the administration, the evaluation of the results is generally very demanding.

As mentioned above metabolomics data from human studies are usually characterized by a large degree of variation between the subjects. This consists a distinct difference from the majority of metabolomics animal-based studies where the metabolic variation between the test animals is usually less abundant^{45,47}. The large variation between human subjects can cause to two significant problems in the analysis. The first is that any small and subtle treatment effects (e.g. dietary responses) can be easily overlooked, especially when the effect is smaller than the intrinsic variation between the subjects. The second problem is that the response and the impact of the treatment effect may differ significantly between the subjects. This implies that the average treatment effect may not be the most relevant after the data analysis procedure, in studies where subsets of subjects respond differently upon a dietary intervention. Therefore a frequently used solution, followed in clinical or

nutritional intervention studies, is the use of a crossover design. In a crossover study all subjects acts as their own control.

In order to gain insight of the results of a crossover nutrition study it is usually helpful to decompose the total into separate variances (e.g. between subjects variation, within subjects values variation etc.) in order to uncover minor treatment effects. Since most traditional multivariate data analysis techniques such as Principal Component Analysis (PCA), Partial Least Squares (PLS) and Partial-Least Squares-Discriminant Analysis (PLS-DA) do not handle the decomposition of the total variance into account, the power of these MVA methods is not fully employed, thus they cannot be the method of choice in such cases. Therefore the main limitation in employing PCA or PLS (-DA) in crossover experiments is that the net treatment effect is not separated from the biological variation between the subjects and the variation caused by the different conditions (e.g. variation caused by the week of administration or by the day of administration). As a result, subtle treatment effects within the subjects are often largely overwhelmed by the strong biological variation between subjects⁴⁸. On the other hand, a great advantage of crossover designs is that treatment (intervention) are evaluated on the same subjects, allowing comparison at the individual level rather than on the group level.

According to the previous remarks, in the present study, the chemometric procedure employed, decomposed the total variation in individual ones. Such an approach was imperative as the analysis was concerning a crossover-designed human nutritional study, in which the metabolic impact of an olive OE-enriched leaf extract consumption on the urine and serum is evaluated. To find systematic differences among the different intervention groups, a chemometric multilevel approach was performed on the data, taking into account the total variation of the design. The

combination of multilevel data analysis and PLS-DA is introduced as a new MVA approach to investigate treatment effects in crossover designed experiments⁴⁰.

Thus, the aim of the randomized, double-blind, crossover clinical study presented in the current study was to evaluate the effects of the supplementation with an OE-enriched olive leaf extract, (Benolea[®] EFLA[®]943), on the serum and urine metabolome of young healthy male athletes. In addition, possible changes in the lipidemic profile and the blood and urine redox status were investigated. In this context it is assumed that supplementation with OE would improve the antioxidant status of the athletes and would have positive antioxidant impact on their recovery after a training period, preventing the occurrence of the DNA damage. The biochemical parameters of the human redox status are also evaluated through a metabolomics approach aiming to highlight the changes in the biochemistry of the athletes after the completion of the protocol.

MATERIALS AND METHODS

Crossover Study Design

The study was carried out in the Artistoteleio University of Thessaloniki, Greece. Nine young healthy males (A, B, C, D, E, F, G, H and I), physical education students in the age range of 20-22 years, participated in a double-blind crossover design. The participants received 1200 mg Benolea[®] EFLA[®]943 (olive leaf extract enriched in OE) or placebo, as nontransparent capsules, for one week in a randomized, balanced, and double-blind manner (Figure 1). The participants provided blood and urine samples every morning before the supplementation. The advantage of performing a

double-blind study is the fortification against voluntary or subconscious changes in diet or physical activity that might affect the results. During the first week of supplementation, the participants were requested to adhere according to a given dietary plan and record their actual diets and physical activity. After the first week of the supplementation, two weeks of washout have followed, after which each participant received the alternate supplement for one week (crossover design). In this way each one has served as the control for himself. During the second week of supplementation, each participant repeated the diet and physical activity that he recorded during the first week of supplementation, as depicted in Figure 1. Furthermore, each of the volunteers was requested to follow a similar lifestyle pattern for the duration of the study. During the two first days of each week the subjects were receiving no supplementation in order to provide samples as baseline measurements.

The protocol was approved by the medical ethical committee of the Aristotle University of Thessaloniki and has been conducted in accordance with the ICH-GCP guidelines (ICH GCP, 1996). All participants gave written informed consent in order to participate in the current study. The quality of the aforementioned administered extract was validated in chapter “Development and validation of a combined methodology for assessing the total quality control of herbal medicinal products - Application to oleuropein preparations”

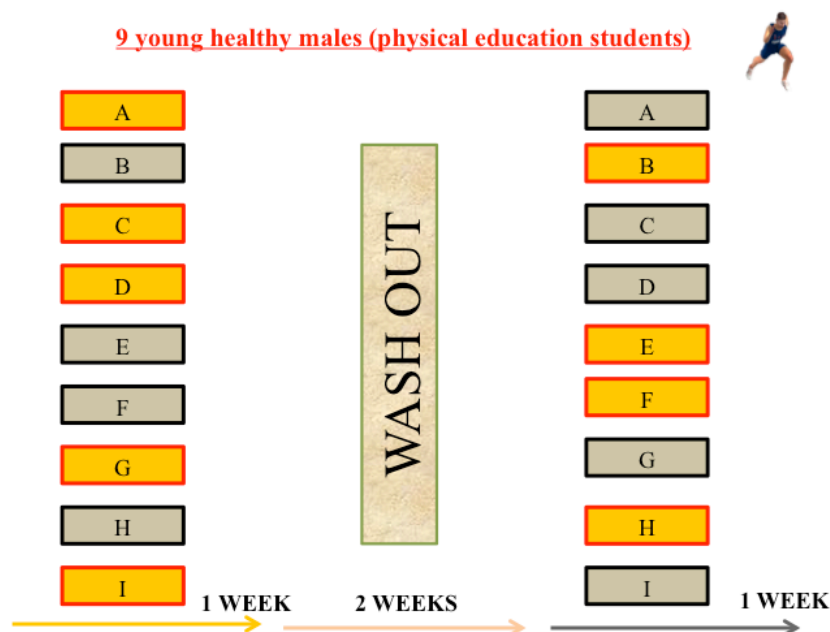


Figure 1. Randomized, balanced, double-blind manner crossover design. The individuals (A-I) received either the oleuropein supplement (orange boxes) or the placebo (grey boxes) in the first week of the protocol. After two weeks of washout each individual received the alternate supplement for one week.

Biofluid collection

Urine samples for the measurements of 8-hydroxy-2'-deoxyguanosine (serving as an index of DNA damage) and the metabolomic analysis were collected in appropriate urine collection vessels over a time span of 24 h after the supplementation, including a 12-h overnight fast. The vessels were supplied with a gelatinized layer of metaphosphoric acid to inhibit bacterial growth. Urine samples were centrifuged and aliquots were stored at -80°C before use.

Blood samples were collected in appropriate tubes. Serum and erythrocytes were separated immediately after the collection by centrifugation of the blood samples at 1800 g for 10 min at room temperature and 500 μL aliquots were stored at -80°C until subsequent analysis. Serum samples were used for the analysis of the lipidemic profile (triglycerides, total cholesterol, HDL cholesterol, and LDL cholesterol) and for

the metabolomic analysis whereas erythrocytes for the measurement of glutathione, the main antioxidant in blood.

Samples for the metabolomic study

The participants received Benolea[®] or placebo capsules, and the corresponding samples (serum and urine) were marked either as oleuropein treated (O) or placebo (P). The corresponding samples from the baseline period (the two first days of each week) were marked as baseline samples (B). Therefore, 126 serum and 126 urine samples were collected and analyzed for the metabolomic study. The list of samples along with their corresponding labels is tabulated in Table 1.

Weeks		1 st week							2 weeks wash out	2 nd week						
Subjects	Days	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
A		1	11	21	31	41	51	61		1B	11B	21B	31B	41B	51B	61B
B		3	13	23	33	43	53	63		3B	13B	23B	33B	43B	53B	63B
C		4	14	24	34	44	54	64		4B	14B	24B	34B	44B	54B	64B
D		5	15	25	35	45	55	65		5B	15B	25B	35B	45B	55B	65B
E		6	16	26	36	46	56	66		6B	16B	26B	36B	46B	56B	66B
F		7	17	27	37	47	57	67		7B	17B	27B	37B	47B	57B	67B
G		8	18	28	38	48	58	68		8B	18B	28B	38B	48B	58B	68B
H		9	19	29	39	49	59	69		9B	19B	29B	39B	49B	59B	69B
I		10	20	30	40	50	60	70		10B	20B	30B	40B	50B	60B	70B

Table 1. A list of the tested samples from the crossover study along with their corresponding labels. Baseline samples (B) are in green, oleuropein samples (O) are in black and placebo (P) samples are in red.

Chemicals, reagents and instrumentation

All solvents used in this study were of LC-MS grade. Acetonitrile (ACN), methanol (MeOH), water and formic acid, were purchased from Fluka/Riedel-de Haën (Switzerland). Ultra HPLC (UHPLC) for the MVA analysis of the crossover study was performed employing an Accela system (Thermo Scientific, Germany) equipped with a binary pump, an autosampler, an online vacuum degasser and a temperature-controlled column compartment. HRMS was performed on a hybrid LTQ Orbitrap Discovery XL mass spectrometer (Thermo Scientific, Germany). Centrifuging of the serum and urine samples for the different sample pretreatment protocols was performed by a Mikro 200R centrifuge (Hettich Lab Technology, Germany). Evaporation of the serum samples was performed with the aid of a GeneVac HT-4X EZ-2 series evaporator Lyospeed ENABLED (Genevac Ltd, UK).

Sample pretreatment

14 serum samples from each subject (7 samples from the first week and 7 samples from the second week), 126 serum samples in total, were prepared following a sample preparation protocol, part of a described methodology for large-scale metabolic profiling of serum samples⁴⁹. Hence, the serum samples were allowed initially to thaw on ice at 4 °C for 30–60 min. Following, a 200 µL aliquot from each serum sample were placed into a labeled 1.5 mL eppendorf tube and 600 µL of MeOH were added followed by vortexing for 15 s. The samples were subsequently centrifuged at 15800 g for 15 min at room temperature to pellet the protein precipitate, and 185 µL aliquots from the supernatant were transferred into two separate labeled 1.5 mL eppendorf tubes and finally evaporated in the GeneVac for 90 min at 50 °C. Consequently, 100 µL of water was added to each dry sample, vortexed for 15 s, centrifuged at 15800 g for 15

min and transferred to 200 μ L inserts placed in appropriate screw capped autosampler vials.

Urine samples from the crossover study were prepared for the UHPLC-ESI-HRMS analysis following also a sample preparation protocol, appropriate for global metabolic profiling for urine using UPLC-MS⁵⁰. Briefly, 14 urine samples from each subject (7 samples from the first week and 7 samples from the second week), 126 urine samples in total, were prepared as follow: 60 μ L of each urine sample were initially centrifuged at 10000 g for 10 min at 4 °C to remove particulates and finally 50 μ L from each centrifuged urine sample were transferred to 200 μ L inserts placed in appropriate screw capped autosampler vials with the addition of 100 μ L water. The mixture was briefly vortexed for 30 s.

Biochemical analyses

A series of biochemical analyses has also been conducted in order to discern the effect of OE supplementation on the redox status of the subjects. Thus measurements of glutathione (the main antioxidant in blood) were carried out using erythrocytes, followed by a series of blood analyses, i.e. analysis of the lipidemic profile (triglycerides, total cholesterol, HDL cholesterol, and LDL cholesterol) by photometric methods, and measurement of parameters of the redox status (urate, bilirubin, malondialdehyde, and protein carbonyls). Finally, measurements of 8-hydroxy-2'-deoxyguanosine (an index of DNA damage) have been performed in the corresponding athletes urine samples. Data have been analyzed by two-way (supplement-by-time) ANOVA with repeated measures on both factors. All biochemical analyses have been conducted at the School of Physical Education and

Sports Science at Aristotle University of Thessaloniki (Greece), by the team of Professor V. Mougios.

UHPLC-HRMS Analysis

Serum samples acquisition. An ACQUITY UPLC BEH C18 (2.1 × 100 mm, 1.7 μm) reversed phase column (Waters Corp., Milford, MA, USA) preceded by a precolumn (Waters VanGuard 5 x 2.1 mm, 1.7 μm) was used for the chromatographic separation in the serum samples. The mobile phase consisted of solvents A: aq. formic acid, 0.1% (v/v) and B: methanol formic acid, 0.1% (v/v). Different gradient elutions⁴⁹ were performed for positive and negative ion mode detection, as described in Table 2, with flow rates of 0.36 and 0.40 mL/min for positive and negative ion mode detection, respectively.

In detail, a gradient method (total run time of 24 min) was used for the profiling of the serum samples in positive ion mode as follows: 0 to 1 min: 100% A: 0% B, 1 to 16 min: from 100% A: 0% B to 0%A: 100%B, 16 to 20 min: 0% A: 100% B, 20 to 22 min: from 0% A: 100% B to 100%A: 0% B, 22 to 24 min: 100% A: 0% B. For the profiling of the serum samples in the negative ion mode, a slightly different gradient method (total run time of 26 min) was used as follows: 0 to 2 min: 100% A: 0% B, 2 to 17 min: from 100% A: 0% B to 0%A: 100%B, 17 to 22 min: 0% A: 100% B, 22 to 24 min: from 0% A: 100% B to 100%A: 0% B, 24 to 26 min: 100% A: 0% B. Column temperature was kept at 50 °C throughout all experiments while the autosampler tray temperature was set at 4 °C and the injection volume was 10 μL.

UHPLC parameters

Column: ACQUITY UPLC BEH C18 (2.1 × 100 mm, 1.7 μm) reversed phase

Mobile phase: aq. formic acid, 0.1% (v/v) (A) and methanol formic acid, 0.1% (v/v) (B)

Flow rate: 0.36 mL min⁻¹ ESI (+) / 0.40 mL min⁻¹ ESI (-)

Column temperature: 50 °C

Tray temperature: 4 °C

Injection volume: 10 μL

ESI (+)	Gradient Program		
	Time (min)	A%	B%
	0	100	0
	1	100	0
	16	0	100
	20	0	100
	22	100	0
	24	100	0
ESI (-)	Gradient Program		
	Time (min)	A%	B%
	0	100	0
	2	100	0
	17	0	100
	22	0	100
	24	100	0
	26	100	0

HRMS parameters

30000 resolution

Centroid mode

ESI (+)	Capillary temperature (°C): 356
	Capillary voltage (V): -60
	Tube lens (V): 110
	Spray voltage (kV): 3.50
	Sheath gas flow (arb. units): 30
	Aux gas flow (arb. units): 10
ESI (-)	Capillary temperature (°C): 356
	Capillary voltage (V): 20
	Tube lens (V): -49
	Spray voltage (kV): 3.10
	Sheath gas flow (arb. units): 30
	Aux gas flow (arb. units): 10

Table 2. Gradient elution programs and MS parameters applied for the UHPLC-HRMS analysis of the serum samples in (+) and (-) ESI modes.

A hybrid HRMS (LTQ-Orbitrap Discovery) instrument, operated separately in ESI positive and negative ion mode has been used. The conditions for the HRMS in each ionization mode were set as follow: for the ESI (+) mode, capillary temperature,

356°C; capillary voltage, -60 V; tube lens, 110 V; source voltage, 3.50 kV; sheath gas flow, 30 arb. units; aux gas flow, 10 arb. units, while for the ESI (-) mode, capillary temperature, 356°C; capillary voltage, 20 V; tube lens, -49 V; source voltage, 3.10 kV; sheath gas flow, 30 arb. units; aux gas flow, 10 arb. units (Table 2). Analysis was performed using the Fourier transform mass spectrometry (FTMS) full scan ion mode, applying a resolution of 30000, while acquisition of the mass spectra was performed in every case using the centroid mode.

Urine samples acquisition. An ACQUITY UPLC BEH C18 (2.1 × 100 mm, 1.7 µm) reversed phase column (Waters Corp., Milford, MA, USA) preceded by a precolumn (Waters VanGuard 5 x 2.1 mm, 1.7 µm) was also used for the chromatographic separation in the urine samples. The mobile phase consisted of solvents A: aq. formic acid, 0.1% (v/v) and B: ACN formic acid, 0.1% (v/v). A gradient elution was performed for positive and negative ESI mode detection, as described in Table 3, using a flow rate of 0.5 mL/min. In detail, a gradient method (total run time of 12 min) was used for the profiling of the urine samples as follows: 0 to 1 min: 99% A: 1% B, 1 to 3 min: from 99% A: 1% B to 85%A: 15%B, 3 to 6 min: from 85% A: 15% B to 50% A: 50% B, 6 to 9 min: from 50% A: 50% B to 5%A: 95% B, 9 to 10 min: 5% A: 95% B, 10 to 10.1 min: from 5% A: 95% B to 99%A: 1% B, 10.1 to 12 min: 99% A: 1% B. Column temperature was kept at 40 °C throughout all experiments while the tray temperature was set at 4 °C and the injection volume was 5 µL.

UHPLC parameters

Column: ACQUITY UPLC BEH C18 (2.1 × 100 mm, 1.7 μm) reversed phase
 Mobile phase: aq. formic acid, 0.1% (v/v) (A) and ACN formic acid, 0.1% (v/v) (B)
 Flow rate: 0.5 mL min⁻¹
 Column temperature: 50 °C
 Tray temperature: 4 °C
 Injection volume: 10 μL

Gradient Program		
Time (min)	A%	B%
0	99	1
1	99	1
3	85	15
6	50	50
9	5	95
10	5	95
10.1	99	1
12	99	1

HRMS parameters

30000 resolution

Centroid mode

ESI (+)	Capillary temperature (°C): 356
	Capillary voltage (V): -60
	Tube lens (V): 110
	Spray voltage (kV): 3.50
	Sheath gas flow (arb. units): 30
	Aux gas flow (arb. units): 10
ESI (-)	Capillary temperature (°C): 356
	Capillary voltage (V): 20
	Tube lens (V): -49
	Spray voltage (kV): 3.10
	Sheath gas flow (arb. units): 30
	Aux gas flow (arb. units): 10

Table 3. Gradient elution program and MS parameters applied for the UHPLC-HRMS analysis of the urine samples in (+) and (-) ESI modes.

A hybrid HRMS (LTQ-Orbitrap Discovery) instrument, operated separately in ESI positive and negative ion mode has been used. The conditions for the HRMS in each ionization mode were set as follow: for the ESI (+) mode, capillary temperature, 356°C; capillary voltage, -60 V; tube lens, 110 V; source voltage, 3.50 kV; sheath gas flow, 30 arb. units; aux gas flow, 10 arb. units, while for the ESI (-) mode, capillary temperature, 356°C; capillary voltage, 20 V; tube lens, -49 V; source voltage, 3.10

kV; sheath gas flow, 30 arb. units; aux gas flow, 10 arb. units (Table 3). Analysis was performed using the Fourier transform mass spectrometry (FTMS) full scan ion mode, applying a resolution of 30000, while acquisition of the mass spectra was performed in every case using the centroid mode.

Quality control samples

In the current study, for both serum and urine, pooled samples were used as quality control (QC) samples. Generally, a QC sample is usually prepared by thoroughly mixing aliquots from all the unknowns of the experimental sample set and is subsequently processed following the developed sample preparation protocol. Thus, for the serum metabolomic study, equal aliquots from the 124 samples were thoroughly mixed and the mixture was subsequently processed in the same manner as the serum samples (sample pretreatment section). The same procedure was followed also for the urine metabolomic study. Therefore, the urine QC samples was prepared by mixing equal quantity of the 124 urine samples and the pretreatment of the QC samples was carried out as described above (sample pretreatment section).

For both serum and urine metabolomic studies, five QC samples were injected at the beginning of each analytical batch, one QC every ten samples throughout the run and five QC samples at the end of the batch (Figure 2). In order to eliminate the instrumental concerns, the ion transfer tube has been removed and cleaned every 50 injections. Furthermore, software-based correction has been applied to the data applying data normalization algorithms. Following this approach, the obtained data sets could be easily evaluated in terms of repeatability; possible outliers can be detected, whereas deterioration of the instrumental performance could be recorded.

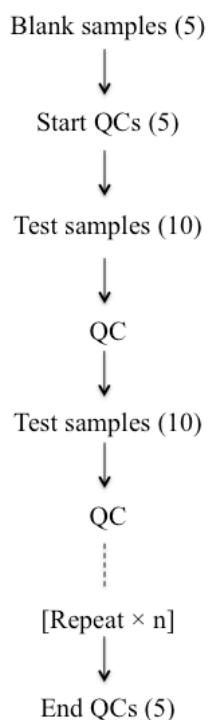


Figure 2. Schematic representation of a typical run in UHPLC-HRMS-based serum and urine metabolomics studies.

Data preprocessing

The data preprocessing procedures can be subdivided in two distinct processes; the *peak picking* procedure and the *matrix preparation* step. The software used for the data preprocessing procedures in order to prepare them for the MVA statistical analysis was: Xcalibur[®] (Thermo Fisher Scientific), MZmine 2.10 (<http://mzmine.sourceforge.net>), Microsoft Excel 2012, the web-based MetaboAnalyst software (<http://www.metaboanalyst.ca>) and the xMSanalyzer package (<http://userwww.service.emory.edu/~kuppall2/xMSanalyzer/>) implemented to the freely available R statistical language 3.0.2 (<http://cran.r-project.org>).

Peak picking procedures

Initially, Xcalibur[®] was used to convert the instrument native Xcalibur data files (*.raw) to the general and more exchangeable cdf data format (*.cdf) with the

embedded to the software Xconvert program. Data were then imported to MZmine 2.10 and processed applying the peak detection, deconvolution, normalization, deisotoping, alignment and gap filling procedures.

MZmine process. Initially, the serum data set from ESI (+) acquisition was imported as centroid LC–HRMS individual files (*.cdf) in MZmine and appropriately filtered to generate corresponding new data files with corrected baseline intensities. In detail, the baseline correction was performed using the base peak intensity chromatogram type and applying a value of 500 for the smoothing, 0.001 for the asymmetry and a value of 1 for the m/z bin width. Then, a list of ions for each scan in the baseline-corrected data file was generated for every data file using the centroid mass detection module and applying the appropriate level for the elimination of noise, in this case $1E4$ counts per second (cps). The centroid module is suitable for already centroided data, while the noise level is considered as the minimum intensity of a data point to be considered as part of a chromatogram. All data points below this intensity level were ignored. Subsequently, the chromatogram builder algorithm has been applied, where the mass lists generated for each MS scan in the previous procedure were used in order to build a chromatogram for each mass that could be detected continuously over a predetermined number of scans (defined in terms of peak width). This step was carried out applying the following parameters: 0.05 min minimum time span, $1E4$ minimum height and 5 ppm m/z tolerance. Following, the generated chromatograms were deconvoluted into individual peaks by the Wavelets (XCMS) deconvolution algorithm setting the signal/noise threshold to 5, the wavelet scales from 0.1 to 5 min and the peak duration range from 0.1 to 1 min. To avoid duplicating the peaks due to the presence of their isotopic pattern, the deconvoluted data were

processed using the isotopic peak grouper algorithm, which removed the additional isotopic peaks from the peak list. To accomplish this operation, the m/z tolerance was set to 5 ppm, the t_R tolerance to 0.05 min whereas only single charged ions were considered. After this step, possible adducts or peak complexes were recognized using the identification of adducts and peak complexes module and consequently unwanted systematic bias between measurements was removed applying the normalization module. The corresponding peaks from the B, P and O serum samples, which were analyzed in (+) ESI mode, were compared using the random sample consensus (RANSAC) aligner algorithm, since it provides the best results among all the tested algorithms⁵¹. The purpose of peak list alignment was to match relevant peaks in terms of t_R and m/z values across multiple samples and to generate a new aligned peak list. The parameters used in this step were 5 ppm m/z tolerance, 0.03 min t_R tolerance and 0.05 min t_R tolerance after correction, 15000 RANSAC iterations, 20% minimum number of points, a threshold value of 4 employing a non-linear model. Following alignment, the resulting peak list may contain missing peaks as a product of deficient peak detection or a mistake in the alignment of different peak lists. The fact that one peak is missing after the alignment does not imply that the peak does not exist but rather in most cases it is present but was undetected by the previous algorithms. Thus, the peak finder algorithm were used to fill the gaps in the peak list when it is possible according with the following parameters: 80% intensity tolerance, 5 ppm m/z tolerance and 0.05 min t_R tolerance, where the m/z and the t_R tolerance define the window where the algorithm should find the new peak. At the end of the procedure, a peak list (accurate mass - t_R vs. intensity) was generated and exported as .csv file. All the parameters used for the aforementioned processes in MZmine, are summarized in Table 4.

The serum data set from (-) ESI acquisition was treated in the same manner described above, using different parameter values, specific and related to the data, and a peak list (accurate mass - t_R vs. intensity) was again generated and exported as .csv file. All the MZmine parameters used in the serum data set from ESI (-) acquisition e.g. baseline correction; mass detection, chromatogram deconvolution, chromatogram builder, deisotoping, normalization and alignment are tabulated in Table 4.

Similarly, the urine data sets from (+) and (-) ESI acquisitions were processed through the MZmine toolbox, concluding to the generation of the corresponding peak lists (accurate mass - t_R vs. intensity). The parameters used are summarized in Table 4.

This method resulted in 3155 aligned features in serum (+) ESI and 2547 in serum (-) ESI data sets, while 2076 aligned features were detected in urine (+) ESI and 3046 in urine (-) ESI data sets.

	Serum		Urine	
	(+) ESI	(-) ESI	(+) ESI	(-) ESI
Baseline correction				
Type	base peak chromatogram	base peak chromatogram	base peak chromatogram	base peak chromatogram
Smoothing	500	500	500	500
Asymmetry	0.001	0.001	0.001	0.001
m/z bin width	1	1	1	1
Mass detection				
Algorithm	centroid	centroid	centroid	centroid
Noise level (cps)	1E4	3E4	8E4	1E5
Chromatogram builder				
Minimum time spam (min)	0.05	0.05	0.05	0.05
Minimum height (cps)	1E4	3E4	8E4	1E5
m/z tolerance (ppm)	5	5	5	5
Chromatogram deconvolution				
Algorithm	Wavelets (XCMS)	Wavelets (XCMS)	Wavelets (XCMS)	Wavelets (XCMS)
S/N threshold	5	5	5	5
Wavelet scales (min)	0.1-5	0.1-6	0.1-1	0.1-4

Peak duration range (min)	0.1-1	0.1-1	0.1-1	0.1-1
Deisotoping				
Algorithm	Isotopic peaks grouper	Isotopic peaks grouper	Isotopic peaks grouper	Isotopic peaks grouper
m/z tolerance (ppm)	5	5	5	5
t_R tolerance (min)	0.05	0.05	0.05	0.05
Identification of adducts	√	√	√	√
Identification of peak complexes	√	√	√	√
Alignment				
Algorithm	RANSAC	RANSAC	RANSAC	RANSAC
m/z tolerance (ppm)	5	5	5	5
t_R tolerance (min)	0.03	0.03	0.03	0.03
t_R tolerance (min) after correction	0.05	0.05	0.05	0.05
RANSAC iterations	15000	15000	15000	15000
Minimum number of points (%)	20	20	20	20
Threshold	4	4	4	4
Model	non linear model	non linear model	non linear model	non linear model
Gap filling				
Algorithm	Peak finder	Peak finder	Peak finder	Peak finder
Intensity tolerance (%)	80	80	80	80
m/z tolerance (ppm)	5	5	5	5
t_R tolerance (min)	0.05	0.05	0.05	0.05

Table 4. MZmine parameters for the data preprocessing of serum and urine datasets from both (+) and (-) ESI-UHPLC-HRMS analyses.

The same MZmine procedure has been also performed using the “baseline cut-off” algorithm for the data deconvolution. The resulted peak lists were compared via the xMSanalyzer package in R (see section).

Preparation of the data matrix

Microsoft Excel process. The generated peak lists (accurate mass - t_R vs. intensity) have been exported to Microsoft Excel 2012 and manipulated using the CONCATANATE, ROUND and TRANSPOSE commands. The generated .xls file has been converted to .csv file or .txt files and was saved for further preprocessing.

xMSanalyzer. The generated peak lists obtained from the different algorithms in the previous step, were imported to xMSanalyzer to compare the extracted features in terms of quality and similarity. Initially, following the appropriate commands, the two peak lists were compared via a Heatmap and the correlation of the common features were revealed. In a next step, a Venn diagram was used in order to determine the number of the unique and common features from the two selected algorithms and the aforementioned information was extracted in a .csv file for further process. Threshold values of 1 ppm for the m/z and 1 sec for the t_R , have been employed. The appropriate commands used in the xMSanalyzer package implemented to the freely available R statistical language 3.0.2 for the comparison of the two deconvolution algorithms are tabulated in Supplementary method 1.

MetaboAnalyst process. The generated peak list (accurate mass - t_R vs. intensity) from the previous step was imported as .csv file, formatted according to the instructions⁵², to the Metaboanalyst online metabolomics platform, to estimate and impute possible missing values in the sample peak list. Initially, rejection of features containing >20% zeros in both groups was performed. Then, the probabilistic principal component analysis (PPCA) was chosen in order to impute the possible missing values and applied to the initial .csv file. Therefore, a new peak list was generated, in which the missing values had been replaced by non-zero values. This new generated peak list (*.csv file) was used for the data analysis process.

In a next step, intensity normalization was performed to the peak list in order to adjust the differences among the samples. The normalization was achieved by choosing the option “normalization by a reference sample” and more specifically by averaging all the samples (pseudo-reference sample) in the control group. Following

this approach, a normalized peak list (*.csv file) was generated, wherein the intensities of the different sample features were normalized. This normalized peak list was further used for the data analysis process. The missing value estimation and the normalization process were also applied to the generated peak lists from all the data sets.

This method reduced the number of features to 1248 aligned features in serum (+) ESI and 1296 in serum (-) ESI data set, while 644 aligned features were detected in urine (+) ESI and 1228 in urine (-) ESI data set.

In a next step, after the missing value estimation process, the data were normalized applying the raw-wise normalization method, in order to reduce any systematic bias during the sample collection. More specifically, normalization was performed to the generated peak list (accurate mass - t_R vs. intensity) resulted from the different data sets, both serum and urine, in both (+) and (-) ESI modes, by a control sample each time, also known as the probabilistic quotient normalization. This method is considered as robust enough, to account for different dilution effects during sample preparation.

Data analysis

MVA data analysis, including PCA, PLS-DA and orthogonal projection to latent structures–discriminant analysis (OPLS-DA), were performed using SIMCA P+ 10.5 (Umetrics, Umea, Sweden) and EZinfo (Umetrics, Umea, Sweden). PCA and PLS-DA were also performed using the MetaboAnalyst online platform being preceded by the normalization and data transformation steps. Kernel based MVA methods have also been employed such as kPCA using a variety of kernels such as polynomial Gaussian etc. employing the kernlab R package. Sparse principal component analysis

(sPCA), sparse independent principal component analysis (siPCA) sparse partial least squares–discriminant analysis (sPLS-DA) and Multilevel sparse partial least squares–discriminant analysis (ML-sPLS-DA) was performed using mixOmics package (<http://cran.r-project.org/web/packages/mixOmics/index.html>) implemented with the freely available R statistical language 3.0.2 (<http://cran.r-project.org>).

Univariate data analysis was performed using analysis of variance (ANOVA), which does not take covariance between the features into account. ANOVA was performed on the important metabolites and their intensity values selected as biomarkers in the ML-sPLS-DA models.

Principal component analysis. PCA, being an unsupervised method, has been initially used to investigate the clustering of the QC samples in the corresponding analytical runs and to explain the percentage of variation. In addition, PCA has been also used to explore any trends or outliers in the data and to obtain an overview of variation among the groups. PCA was performed using both unit variance (UV) and Pareto scaling; confidence level on parameters was set at the 95% level whereas 200 maximum iterations have been employed. The scores values have been used for the evaluation of the results. The combination of R^2 and Q^2 values has been used in order to evaluate the optimal number of contributing parameters, as well as the scree plot. A cut-off value of 10% explained variance has been used for the optimal number of pc's kept for the construction of the PCA model.

Sparse principal component analysis. sPCA, a variant of PCA, was performed initially for estimating a limited number of the most influential loadings (sparse loadings), allowing the efficient variable selection. sPCA was based on singular value

decomposition (SVD) and sparsity is achieved via LASSO^{53,54} penalization. The value “TRUE” was set in the scale parameter, in order to obtain orthogonal sparse loading vectors, since in sPCA methods the orthogonality between the pcs and the loading vectors is lost. 100 variables were selected on each PCA dimension using the command KeepX. Cross-validation was performed to compute the R^2 and Q^2 values. The appropriate commands used in the mixOmics package implemented to the freely available R statistical language 3.0.2 for the competition of the sPCA analyses are tabulated in Supplementary method 2.

Partial least squares–discriminant analysis. PLS-DA has been used in a supervised manner for further exploration of the difference between biomarker clusters and to display the maximum covariance of metabolic data with a defined Y variable (class) in the data set. The cross-validated cumulative Q^2 (Q^2_{cum}) value was used as a degree of the predictive value of the PLS-DA model. A Q^2 value of 1 indicates maximum predictive power, whereas Q^2 values close to or below 0 indicate a lack of predictive power. As a rule of thumb, models with $Q^2_{\text{cum}} > 50\%$ are considered to be of good predictive power in this context.

Orthogonal partial least squares–discriminant analysis. OPLS-DA, a supervised pattern recognition approach, was used as a predictive model to discriminate the groups (classes) and to identify the differential metabolites in the different groups. The OPLS-DA model maximizes the covariance between the measured data of X variable (peak intensities) and the discriminant response of Y variable (class) within the groups. The quality of the model was determined by the goodness of fit and variation in the X (R^2X) and Y (R^2Y) variables and the

predictability based on the fraction correctly predicted in a 1/7 cross-validation (Q^2). Seven-round cross-validation was performed to eliminate the risk of overfit with only 1 predictive component for 2 classes. OPLS-DA shows more subtle changes in the occurrence and concentration of specific metabolites by focusing on compounds responsible for the discrimination between 2 groups.

Sparse partial least squares–discriminant analysis. sPLS-DA was used as a sparse variant of the PLS-DA method, to allow for the selection of the most influential variables in accordance to the PLS-DA. Thus, 3 components (dimensions) were chosen and the number of selected variables was set to 100. Variable selection was achieved using the LASSO penalization on the pair of loading vectors. Cross-validation was performed to compute the R^2 and Q^2 values. sPLS-DA is based on the same concept as sPLS to allow variable selection, except that the variables are only selected in the X data set and in a supervised framework using non continuous variables, i.e. the X-variables with respect to different categories of the samples.

The parameters to choose by the user are the number of components “ncomp” and the number of variables to choose in the X data set “keepX”. The appropriate commands used in the mixOmics package implemented to the freely available R statistical language 3.0.2 for the competition of the sPLS-DA analyses are tabulated in Supplementary method 2.

Multilevel sparse partial least squares–discriminant analysis. ML-sPLS-DA is an MVA method used for the pair-wise comparison of different groups in crossover studies, where each subject undergoes a control measurement and a treatment in a random order⁴⁰. ML-sPLS-DA is the method of choice in human studies since, the

inter-individual variability caused by genetic, dietary, lifestyle and environmental factors can obscure nutrition-related metabolic effects. ML-sPLS-DA was performed using the mixOmics package implemented with the freely available R statistical language 3.0.2 by applying the appropriate commands (Supplementary method 3).

The quality of the ML-PLS-DA models was judged by the goodness-of-fit parameter (R^2) and the predictive ability parameter (Q^2), which is calculated by an internal cross-validation of the data and the predictability calculated on a leave-out basis. All the aforementioned procedures have been performed using the “tune” command, as implemented in the ML-sPLS-DA sub-routine in the mixOmics package of R language (Supplementary method 3).

Variable selection - variable importance in the projection. The variable importance in the projection (VIP) values of each variable (accurate mass - t_R vs. intensity) in the different non-sparse MVA models was calculated to indicate the contribution to the classification of samples. Variables with a VIP value >1.5 , were considered important in discriminating between groups.

Sparse loadings. The first 50 sparse loadings as obtained from the ML-sPLS-DA models were considered important in discriminating between groups and were selected for further evaluation. A plot containing the selected sparse variables was constructed, using the “PlotVar” command and by applying a “rad.in” value of 0.9 (Supplementary method 3).

Scaling. The use of an appropriate scaling technique is high of importance, since PCA and PLS-DA are scale-dependent methods. Using the appropriate scaling

technique, medium and small features in the data sets could be considered as important, as the large features do. Among the different scaling techniques, Pareto-scaling and Unit Variance (UV) scaling are the most frequently used ones to LC-MS data. An important disadvantage with UV-scaling has been concerned the increase of spectral noise, which may lead to severe overfitting problems in PCA and PLS-DA. Pareto scaling compromises between the extremes of no-scaling and UV scaling because the square root of the standard deviation is used as the scaling factor in contrast to the standard deviation value, used in the UV scaling.

Identification and structural elucidation of potential target metabolites

The important features, which correspond to accurate mass - t_R , selected using the aforementioned MVA approaches were putatively assigned to specific metabolites by match the selected features in terms of mass accuracy in online databases. More specifically, three publicly available databases of MS spectra of metabolites: (a) Human Metabolome Data Base (HMDB) (<http://www.hmdb.ca/>), (b) METLIN Metabolomics Database (<http://metlin.scripps.edu/index.php>), (c) Kyoto Encyclopedia of Genes and Genomes (KEGG) (<http://www.genome.jp/kegg/>), (d) ChemSpider free chemical structure database (<http://www.chemspider.com>), (e) MassBank database (<http://www.massbank.jp>) and (f) LIPID MAPS (<http://www.lipidmaps.org>) lipidomics gateway were used and furthermore the data were compared with those from previously published literature. Furthermore, additional data were used to confirm the proposed structures, such as the isotopic pattern and the RDB information. The matching criterion for the comparison of the expected isotopic pattern with the measured one was kept to 5% of matching, whereas the RDB, which provides the degree of molecule's unsaturation, was considered. Additionally, a

deconvolution step has been performed to the corresponding spectra, using the appropriate software i.e. the Mass Frontier 5.0.1 and its freely available counterpart AMDIS (<http://chemdata.nist.gov/dokuwiki/doku.php?id=chemdata:amdis>), in order to gain information about a possible fragmentation that takes place in the source. All the putatively identified metabolites were further estimated through a t-test involving the abundance of the signal in the O and P groups, with the aim to reveal any statistically significant differences concerning their means. The significant level has been set to 95%.

Metabolic pathway analysis

Metabolic pathway analysis (MPA) was performed with MetPA (<http://metpa.metabolomics.ca/MetPA/faces/Home.jsp>) accessing the (a) Human Metabolome Data Base (HMDB) (<http://www.hmdb.ca/>), (b) METLIN Metabolomics Database (<http://metlin.scripps.edu/index.php>), (c) Kyoto Encyclopedia of Genes and Genomes (KEGG) (<http://www.genome.jp/kegg/>), (d) ChemSpider free chemical structure database (<http://www.chemspider.com>), (e) MassBank database (<http://www.massbank.jp>) and (f) LIPID MAPS (<http://www.lipidmaps.org>), to identify the affected metabolic pathways analysis and facilitate the visualization of the results. The significant identified features were imported as compound names or HMDB/KEGG ID's and possible associated biological pathways were investigated. Three parameters needed to be specified for pathway analysis: the specific pathway library, the algorithm for pathway enrichment analysis described above, and the algorithm for topological analysis⁵². From the list of pathway libraries, *Homo sapiens* (human), which contains 80 pathways, was selected for MPA. The p-value (from pathway enrichment analysis) indicates the statistical significance of association of

the altered metabolites with the pathway and the pathway impact value (from pathway topological analysis) is calculated as the sum of the importance measures of the matched metabolites normalized by the sum of the importance measures of all metabolites in the pathway^{55,56}. Those pathways with $p < 0.05$ or an impact value > 0.1 were considered and filtered out as potential target pathways.

RESULTS AND DISCUSSION

Crossover study - Metabolomic approach

In total, 9 males, aged between 20 and 22 years, were participated in the randomized, balanced and double-blind manner crossover study. One subject (subject A) took antibiotic during the study, as he reported it, and he was separated into outlier in the urine PCA model. In the serum PCA model he was not considered as an outlier, since antibiotics and their metabolites are excreted mostly in urine. Other patients were not considered as outliers in the initial serum and urine PCA models, a fact that is consistent with their daily reports for dietary plan, actual diets and physical activities. In addition, there was no significant difference in the demographic characteristics (age, height, weight) of the participants.

Biochemical analyses

From the biochemical analyses it was found that olive leaf extract systematically changed the levels of triglycerides in plasma but only sporadically the levels of total cholesterol, HDL cholesterol and LDL cholesterol. These results need to be verified with a larger cohort of participants. Data not shown.

UHPLC-HRMS analysis

Most applications in LC-HRMS metabolomics era use reverse phase column chemistries, although other chemistries (e.g., hydrophilic interaction chromatography, HILIC etc.) are also used. The use of reversed-phase chemistries provides efficient retention and separation of relatively nonpolar metabolites across a large molecular weight range (50 to >1.500) and includes high-molecular-weight lipid species (e.g., phospholipids and triglycerides) and nonpolar amino acids (e.g., tryptophan). Polar metabolites elute in the column void volume or early in the chromatographic run wherein efficient retention and separation is not achieved. For these reasons, a reverse phase column was selected for the completion of the experiments of the current study.

The organic solvent used in the gradient elution is typically either methanol or acetonitrile. Acetonitrile is more nonpolar (or lipophilic and hence ‘stronger’) organic solvent and operates at lower backpressures, although in the current experience it could present problems with elevated background and reduced sensitivity at high acetonitrile concentrations. For this reason, methanol was chosen in the instrument optimization procedure, as it provided appropriate chromatographic separation. The optimization process defined two separate optimal sets of conditions for (+) ESI and (-) ESI ion modes because of differences in the metabolites detected and the metabolic profiles. The application of the same gradient elution program in both (+) ESI and (-) ESI offers advantages in metabolite identification, as a single metabolite will be recorded with the same t_R in both ion modes, and, when feasible, source polarity switching can be carried out.

Quality control samples

A major challenge regarding LC-MS-based metabolomic methodologies is the risk concerning the degradation of the analytical instruments performance (in this case an LTQ-Orbitrap Discovery) with time. Thus, it is of high importance for the analyst to ensure the quality of the data. A variety of approaches have been used for certifying that the results obtained from global metabolomic studies are valid, such as the addition of internal standards in the tested samples or the use of test mixtures or QC samples⁴⁹.

Typical QC's should be theoretically identical biological samples, with a metabolic and sample matrix composition similar to those of the unknowns. An ideal QC sample for a small sample set would be either a pooled sample derived from the mixture of the unknowns or a commercially available biofluid. Pooled QC samples present a number of advantages, including being as close as possible to the composition of the unknown samples, whereas they are well suited for small studies in which all of the samples are available in a sufficient quantity (e.g., small clinical trials or animal studies). However, pooled QC samples are not always applicable in large-scale studies in which the sample number is extremely high and the processing of these studies is often in progress while the metabolomic analysis has already started.

Data preprocessing

Data preprocessing is a key procedure for every aspect of metabolomic data analysis, especially for data acquired via untargeted metabolomic approaches.

Peak picking procedure. A crucial step before proceeding to data analysis was to optimize the deconvolution process by applying different algorithms to the data sets

and evaluating the extracted features (m/z) from each algorithm and comparing them by finding the unique or overlapping features. It should be noted that there is no universally accepted algorithm for the data peak picking, so the evaluation of the existing algorithms is mandatory. Thus, two algorithms, namely the “baseline cut-off” and the “Wavelets” algorithm, have been evaluated as stated in aforementioned workflow, for facilitating metabolite detection by MZmine. For this purpose, the xMSanalyzer R package has been used to compare the data resulted from the two procedures. Initially, a Heatmap (Figure 3) comparing the two peak lists has been constructed. The x-axis from the Heatmap represents the data from the “Wavelets” method, whereas the y-axis the corresponding data from the “baseline cut-off” method. The matrix depicts the covariance values with red color coding, denoting the high correlation features (common features), whereas the blue color stands for the unique features. It is evident that the two matrices exhibited high correlation (large red areas in the Heatmap), which indicates that a second step of common features identification should follow.

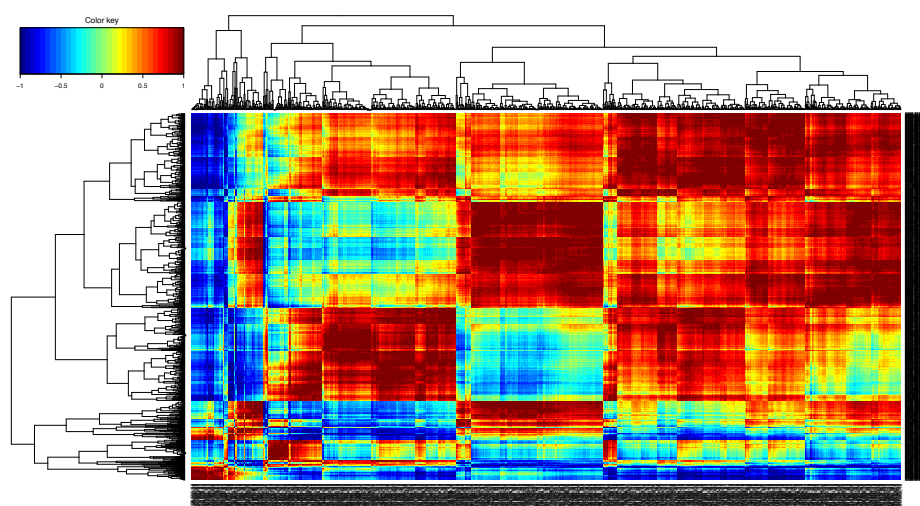


Figure 3. Heatmap visualization of the features detected by the “baseline cut-off” and “Wavelets” algorithms. X-axis: features from the “Wavelets” algorithm; Y-axis: features from the “baseline cut-off” algorithms. Color key indicates covariance features, blue: low, red: high.

Therefore, using the appropriate commands (Supplementary Method 1), a Venn diagram has been constructed, showing both common and unique features (Figure 4). The data shows that from a total number of 1492 features discovered, 887 (60%) were common, 203 (13%) were unique for the “baseline cut-off” algorithm and 402 (27%) were unique for the “Wavelets” algorithm. The common features were assembled in a peak list and submitted to further evaluation.

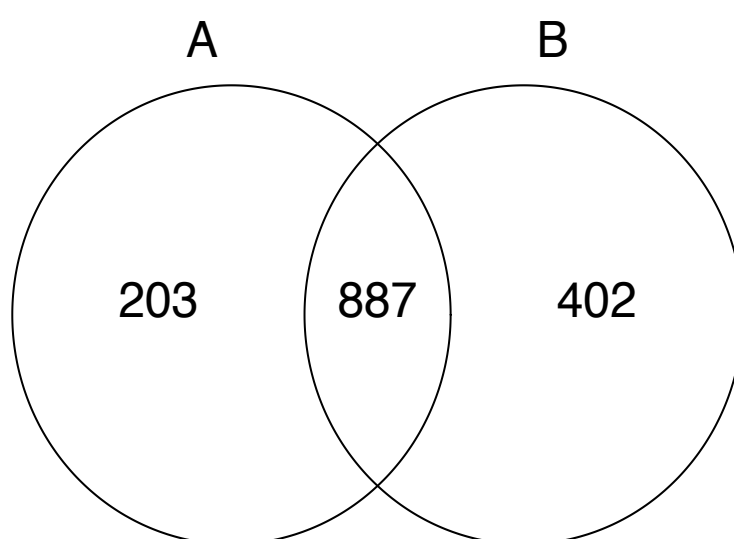


Figure 4. Venn diagram from the two tested algorithms; “baseline cut-off” (A) and “Wavelets” (B). 887 common features are located in the common section.

MetaboAnalyst process. For this purpose, many functions are available in the MetaboAnalyst process, including missing value imputation and normalization procedures. Depending on the type of metabolomics experiment being analyzed, there may be a substantial number of missing values present in the data set. A variety of methods have been implemented to deal with the missing values problem. By default, MetaboAnalyst treats missing values as being present but assigning them a low signal intensity (even below the detection limit). Consequently, they are replaced by half of

the minimum positive value detected in the data, whereas samples with too many missing values (>50% by default) were excluded either manually or automatically^{52,57}. Alternately, there are several computational methods, such as imputation of the missing values by mean/median, PPCA, Bayesian PCA (BPCA) or singular value decomposition (SVD)⁵⁸. PPCA combines an expectation maximization approach with a probabilistic model. The expectation maximization approach is based on the assumption that the latent variables (scores) as well as the noise come from normal distributions. In standard PCA, data points far from the training set but close to the subspace defined by the principal components, fit the model equally well. PPCA, on the other hand, defines a density model such that the likelihood for data points far from the training set is much lower, even if they are close to the principal subspace⁵⁸. Overall, PPCA is the most suitable method for large sets of centroid data, provides good results, also overcoming the performance problems of the other methods, thus was chosen for the missing value estimation step.

After completion of the missing value estimation step, a normalization process was followed. This procedure aims to reduce systematic bias during sample collection. The normalization by the sum method is often used for binned spectra data in which the total spectral area is assumed to be constant; the normalization by a feature can be used to adjust the feature values (i.e. concentrations) of each sample against a spike-in, an internal standard or a physiological constant (i.e., urinary creatinine); the normalization by a sample, also known as probabilistic quotient normalization⁵⁹. In the current normalization process the option “normalization by a reference sample” and more specifically by averaging all the samples (pseudo-reference sample) in the control group, was chosen for the normalization of the data. This method rescales each sample by the most probable dilution factor, which is calculated as the median of

the quotients between all corresponding features of the sample and the reference. Furthermore, as QC samples have been inserted between the samples throughout all the analytical batch, it is expected that the last methodology could also compensate for the batch effect.

MVA Analysis

In the serum and urine data sets of the 9 individual human subjects, from both (+) and (-) ESI acquisitions, a wide range of low-molecular weight metabolites can be identified. Representative base peak intensity (BPI) UHPLC-(+)ESI-HRMS serum chromatograms from individual B in the first week of administration (Benolea[®] administration) in juxtaposition with the second week of administration (placebo administration) are shown in Figure 5. Differences in metabolic profiles due to the administration (O vs. P) could even be observed by visual inspection.

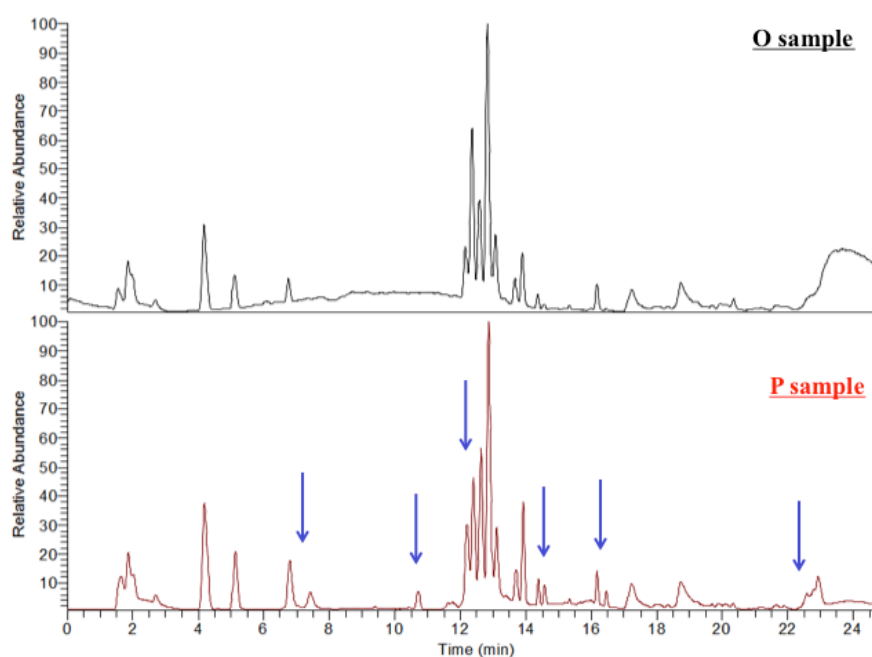


Figure 5. BPI UHPLC-(+)ESI-HRMS serum chromatograms from individual B in the first week of administration (Benolea® administration) (top) in juxtaposition with the second week of administration (placebo administration) (below). Differences could be observed (blue arrows) by visual inspection and can offer complementary information regarding the metabolite profiles.

Furthermore, differences in metabolic profiles were observed between different individuals due to large biological variation. Representative base peak intensity (BPI) UHPLC-(+)ESI-HRMS serum chromatograms of the D and I individuals in the first week of administration (placebo) are shown in Figure 6.

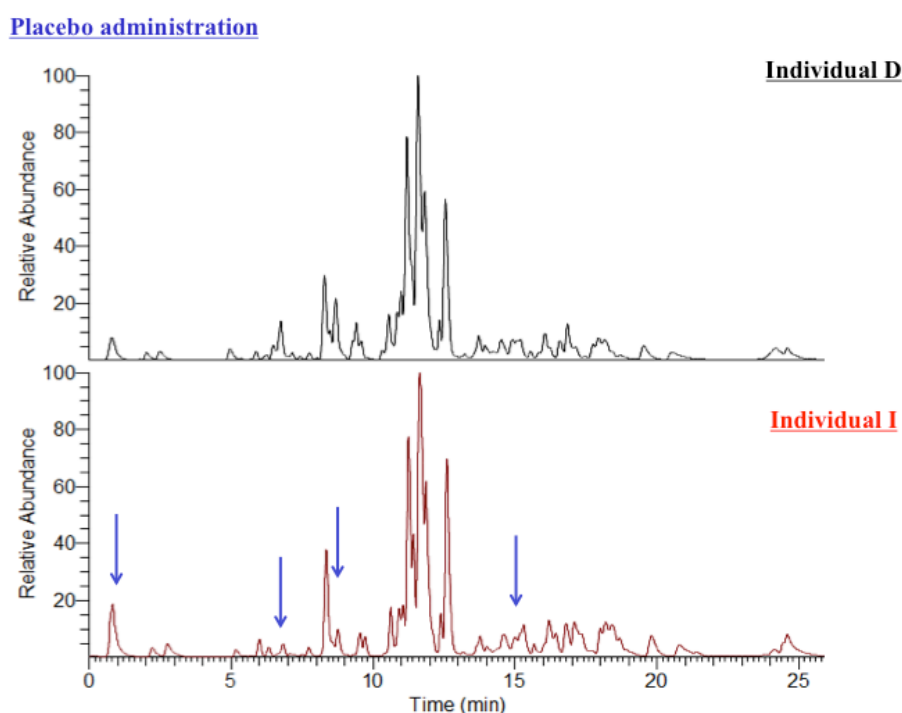


Figure 6. BPI UHPLC-(+)ESI-HRMS serum chromatograms from D (top) and I (below) individuals in the first week of administration (placebo). Differences could be observed (blue arrows) by visual inspection and can offer complementary information regarding the metabolite profiles of the individuals.

MVA has been employed using various statistical models, such as PCA, siPCA, sPCA, PLS-DA, OPLS-DA, sPLS-DA and ML-sPLS-DA (k-PCA, SVM Random Forests, Hierarchical/KNN clustering). Data were processed for MVA implementing

several combinations concerning the group/class in which they are belong to (e.g. O vs. P or O vs. B, etc.) or the days of administration (e.g. all the days of administration, first day vs. last day, two last days, etc.) or even the individuality (e.g. all the individual subjects simultaneously, one individual at a time or random combination). Various parameters have been used, such as different scaling e.g. UV or Pareto, and different combinations of Q^2 and R^2 in order to assess the number of factors that described as much variance as possible without increasing the complexity of the models excessively.

Data Quality Assurance. In order to assess the repeatability of the method, twenty-two QC samples were analysed during each analytical set. The quality of the obtained data was evaluated using two criteria; the tightness of the resulting QC's clusters and the variation of selected peak's characteristics, e.g. the t_R (min), the peak area and the mass accuracy (ppm). For this purpose three peaks, found in all chromatogram, have been selected, at the begging, middle and end of the chromatogram. The rejection criteria established were $>1\%$ RSD for the t_R , $>10\%$ RSD for the peak area and >5 ppm for the mass accuracy. The coefficient of variation of t_R was less than 0.6% RSD and the variation of accurate mass was less than 3 ppm for the selected peaks (i.e. hippuric acid in urine samples analyzed by negative ESI mode, L-tryptophan in urine samples analyzed by positive ESI mode, uric acid in serum samples analyzed by negative ESI and tiglylglycine in serum samples analyzed by positive ESI), demonstrating the excellent repeatability of the proposed methodology. Peak area RSD was found to be lower than 5% for all the selected analytes. This result also provides assurance that there are no major run-order-related changes occurring as the analysis proceeds. A representative trend plot of the selected

peak #228 ($m/z = 355.2817$ and $t_R = 8.0$ min) from (+) ESI urine dataset is depicted in Figure 7, showing the excellent area repeatability of the QC samples.

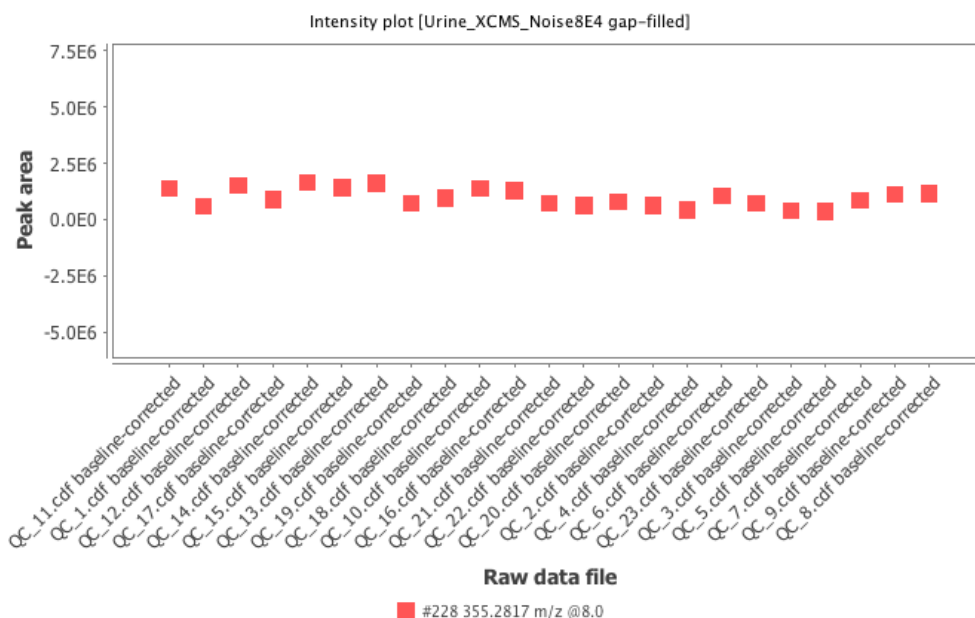


Figure 7. Representative trend plot of the selected peak #228 obtained from (+) ESI urine dataset. Red squares represent the QC samples. The excellent area repeatability of the QC samples is shown.

Furthermore, PCA has been used to obtain an overview of the clustering of the QC samples and ascertaining the analytical data quality, by visual inspection. Various parameters have been used, such as different scaling e.g. UV or Pareto, and different combinations of Q2 and R2 in order to assess the number of factors that described as much variance as possible without increasing the complexity of the model excessively. An arbitrary cutoff value of 10% for the explained variance for the pc's has been employed, in an effort to keep the simplest form for the PCA model. This has been verified by the scree plot, which showed rapid inflection after the second pc. Incorporation of more pc's led to an insignificant improvement of the model and only the two first pc's have been taken into account thereof. Hence, PCA models were created applying the UV scaling, as it afforded tighter groups with larger separation,

analyzing all the samples (B, P and O) including the QC samples. The samples that were used in this step of the MVA can be found in Table 1.

Generally, for a robust method, the QC samples should form a tight cluster, denoting that the behavior of the analytical system does not change appreciably during the course of the analysis.

As an example, a two-dimensional PCA scores plot of urine samples (black) and QCs (red) obtained by UHPLC-(+)ESI-HRMS is presented in Figure 8. It should be noted that features with identical m/z but different t_R possibly correspond to isomers. Furthermore, an analysis for possible adducts (M-H₂O-H, M-H, M+Na-2H, M+Cl, M+K-2H, M+FA-H, M+Br, 2M-H, 2M+FA-H) has been performed for all the features in order to verify the possibility of occurrence of chromatographic peaks due to adduct formation. The quality of the QC data was estimated by inspecting the tightness of the clustering among the QC samples. This clearly shows the excellent stability of the analysis.

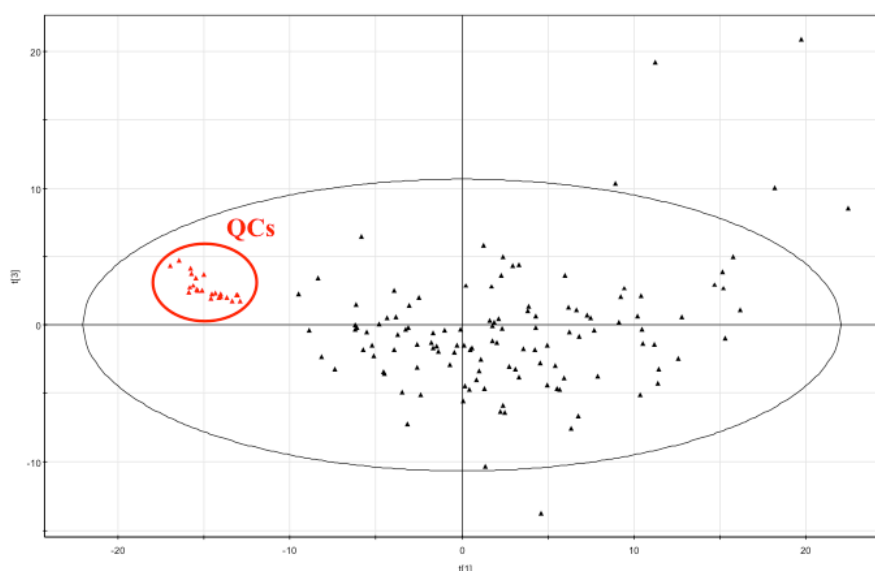


Figure 8. PCA scores plot of the urine samples (B, P and O) after UV scaling obtained from UHPLC-(+)ESI-HRMS analysis. Black triangles represent the tested samples. Red triangles (QC samples) are clustered tightly together in the dataset, proving the validity of the analysis.

The aforementioned procedure was applied also to the other datasets in which the essential stability of the QC samples through the analytical runs were found to be adequate. No instrument failure indicating a decline in sensitivity, t_R shifts, or changes in mass accuracy was observed. Therefore, the data gave confidence both about the robustness of the UHPLC-ESI-orbitrap system operating conditions and the reliability of all the datasets for further statistical analysis.

Detection of trends and possible outliers. Furthermore, PCA has been used also to explore any possible trends or to detect putative outliers in the serum and urine data sets. Thus, PCA models were created applying the UV scaling, taking into account all the samples (B, P, and O). The samples that were used in this step of the MVA can be found in Table 1.

After careful evaluation of the PCA models created, no trend towards the differentiation between B (green), P (red), and O (black) samples was found. As an example, a representative PCA scores plot of the (+) ESI urine data set (Figure 9), clearly shows the inability of the PCA model to highlight the differences between these samples and to cluster them according to their status (B, P and O).

Overall, the PCA analysis of the serum and urine data sets (positive and negative acquisition) did not exhibit any clustering of the different samples examined, related to either the administration (B, P or O grouping) or to the individuality of the participants or even to the days of administration. This can be attributed to the presence of numerous metabolites detected in the samples and caused by the great heterogeneity of the subjects or from the different nutritional habits of the individuals. Thus, it seems that PCA cannot be effectively used for human cross over nutritional intervention studies, as a predictive tool.

For the outlier detection issue, an example of the data analysis workflow followed can be demonstrated using the UHPLC-(+)ESI-HRMS urine data set. The PCA model was created applying the UV scaling, taking into account all the samples (B, P, and O).

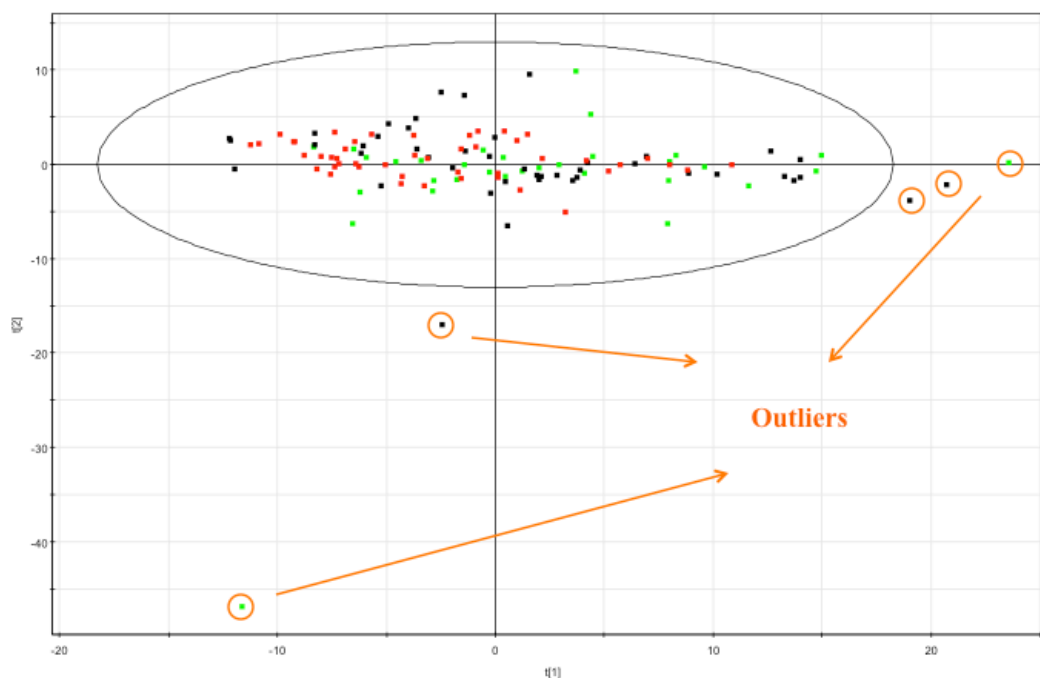


Figure 9: PCA scores plot of the urine samples (B, P and O) after UV scaling obtained from UHPLC-(+)ESI-HRMS analysis. No trend towards the differentiation between B (green points), P (red points), and O (black points) samples was observed. Five outlier samples in the PCA model are marked in orange circles.

Therefore, 5 samples were detected as outliers in the PCA model and were removed from subsequent analysis. It is noteworthy that these urine samples belong to one athlete, i.e. the subject A, who was following an antibiotic treatment during the second week of the administration, as reported in his daily reports. The LC-MS chromatograms of these outlying samples were inspected to confirm the presence of peaks belonging to the antibiotic itself or its metabolites. Figure 9 shows a PCA

scores plot of all the samples B (green), P (red) and O (black), displaying the presence of the outliers from the subject A. These data suggests that medicine-induced variations in the urine metabolome can be revealed using metabolomics. Interestingly, no subject A related outliers were detected in the serum metabolome, which indicates the homeostatic ability of the blood.

Examination of each individual separately. Due to the large biological variation between the individuals, which is reflected to their metabolome, PCA models were created for each subject separately (8 PCA models for 8 subjects for each dataset). In other words the clustering of the individual subjects between the two weeks of the crossover administration (P *vs.* O) has been examined. The samples that were used in this step of the MVA can be found in Table 5.

Indicatively, the trend of representative individuals concerning the week of O dosing *vs.* P dosing is clearly depicted in Figure 10, where the O and P samples from the (+) ESI urine data set were analyzed. Thus, it is clearly shown, that the P urine samples (red) are clustered separately from the O urine samples (black) in each subject respectively. This trend was repeated also in all the remaining urine and serum datasets. The ability of the PCA method to distinguish the administration effect to the individual level seems to be adequate, but on the other hand it seems that the large biological variation supersedes the respective variation caused by the administration effect and hinders the PCA clustering capacity (Figure 9). Therefore, a simple PCA analysis is not enough to afford a global model in this case.

Weeks		1 st week								2 nd week						
Subjects	Days	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
A		1	11	21	31	41	51	61		1B	11B	21B	31B	41B	51B	61B

Weeks		1 st week								2 nd week						
Subjects	Days	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
B		3	13	23	33	43	53	63		3B	13B	23B	33B	43B	53B	63B

Weeks		1 st week								2 nd week						
Subjects	Days	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
C		4	14	24	34	44	54	64		4B	14B	24B	34B	44B	54B	64B

Weeks		1 st week								2 nd week						
Subjects	Days	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
D		5	15	25	35	45	55	65		5B	15B	25B	35B	45B	55B	65B

Weeks		1 st week								2 nd week						
Subjects	Days	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
E		6	16	26	36	46	56	66		6B	16B	26B	36B	46B	56B	66B

Weeks		1 st week							+	2 nd week						
Subjects	Days	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
F		7	17	27	37	47	57	67		7B	17B	27B	37B	47B	57B	67B

Weeks		1 st week								2 nd week						
Subjects	Days	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
G		8	18	28	38	48	58	68		8B	18B	28B	38B	48B	58B	68B

Weeks		1 st week								2 nd week						
Subjects	Days	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
H		9	19	29	39	49	59	69		9B	19B	29B	39B	49B	59B	69B

Weeks		1 st week								2 nd week						
Subjects	Days	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
I		10	20	30	40	50	60	70		10B	20B	30B	40B	50B	60B	70B

Table 5. A list of the tested samples along with their corresponding labels, used for the examination of each individual separately. Oleuropein samples (O) are in black and placebo (P) samples are in red. Baseline samples (light grey) were not considered in this step of MVA.

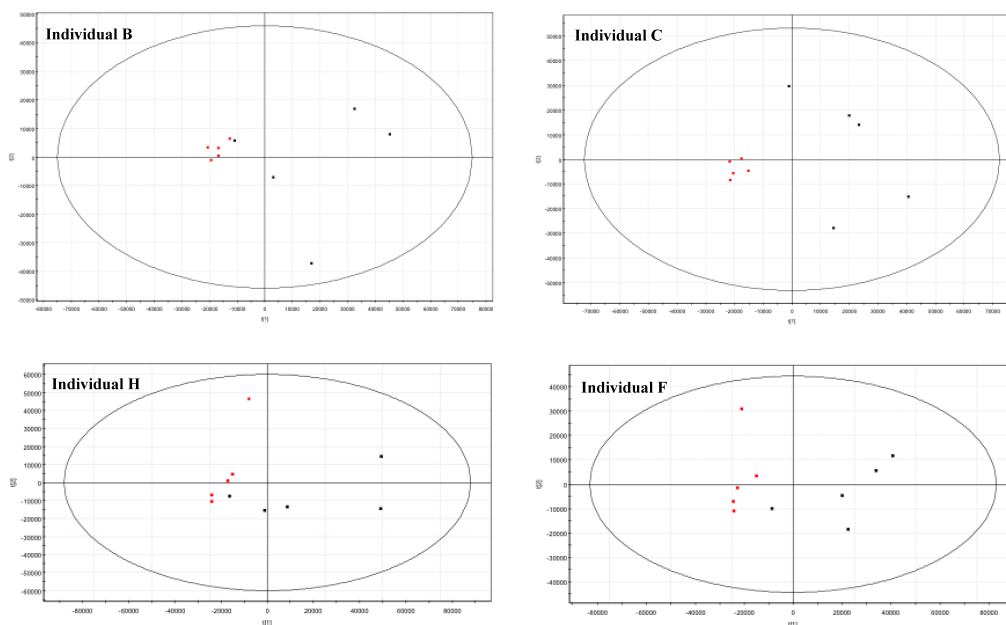


Figure 9: PCA scores plots of representative individuals' (B, C, H and F) urine samples (P and O) after UV scaling, obtained from UHPLC-(+)ESI-HRMS analysis. A clear trend towards the differentiation between P (red points) and O (black points) samples was observed in each individual separately.

The respective loadings plot of the aforementioned analyses were evaluated and no OE or OE metabolite related peaks have been detected, which indicated that the individual subject's clustering could be attributed to the alteration of the native metabolome. No effort has been made to evaluate the loadings plots further due to their high complexity.

Application of various PCA variants. Besides PCA, other unsupervised methods were employed in the manner of unsupervised methods, such as sPCA and siPCA, in order to estimate the pc's with sparse loadings and to allow variable selection. Furthermore, k-PCA has also been employed using the following functions:

- bfgdot Radial Basis kernel function "Gaussian",
- polydot Polynomial kernel function,
- vanilladot Linear kernel function.

- tanhdot Hyperbolic tangent kernel function,
- laplacedot Laplacian kernel function,
- besseldot Bessel kernel function,
- anovadot ANOVA RBF kernel function and,
- splinedot Spline kernel

Similar to the PCA results have been obtained using abovementioned MVA methodologies, underpinning the inability of unsupervised methodologies to obtain a global model of the OE administration to athletes.

PLS-DA and OPLS-DA models. After outliers have been excluded and the trends of the administration in each subject have been discovered, PLS-DA and OPLS-DA models were generated to maximize the differentiation between P dosed subjects and O dosed ones or between the subjects before administration (B samples) and after administration (P or O samples). Other combinations were also examined using these supervised models, such as B *vs.* O or P *vs.* O samples in selected days of administration (first *vs.* last day of administration). These supervised MVA models did not led to any significant clustering and hence to identification of metabolites (either serum or urine) responsible for this differentiation.

More specifically, PLS-DA analyses were carried out also to maximize the covariance between the measured data of X variable (peak intensities) and the response of Y variable (B, P or O) within the groups and to determine whether the different samples (B, P and O) could be systematically distinguish. All data were normalized, as discussed above. In addition, various data scalings were investigated (including UV and Pareto). An arbitrary cutoff value of 10% for the explained variance for the components has been employed, in an effort to keep the simplest

form for the PLS-DA models. Incorporation of more components led to an insignificant improvement of the model. Thus, only the two first components have been taken into account. Figure 10 shows a PLS-DA scores plot (UV scaled) of B, P and O urine samples from all the days of the administration obtained by UHPLC-(+)ESI-HRMS. The samples that were used in this step of the MVA can be found in Table 1. It is evident that no significant clustering was observed between the B samples (green points) the O samples (black points) and the P samples (red points) in this PLS-DA scores plot. The same results were observed, when the other data sets (urine and serum in both positive and negative ion modes) were also analysed by PLS-DA either using UV or Pareto scaling.

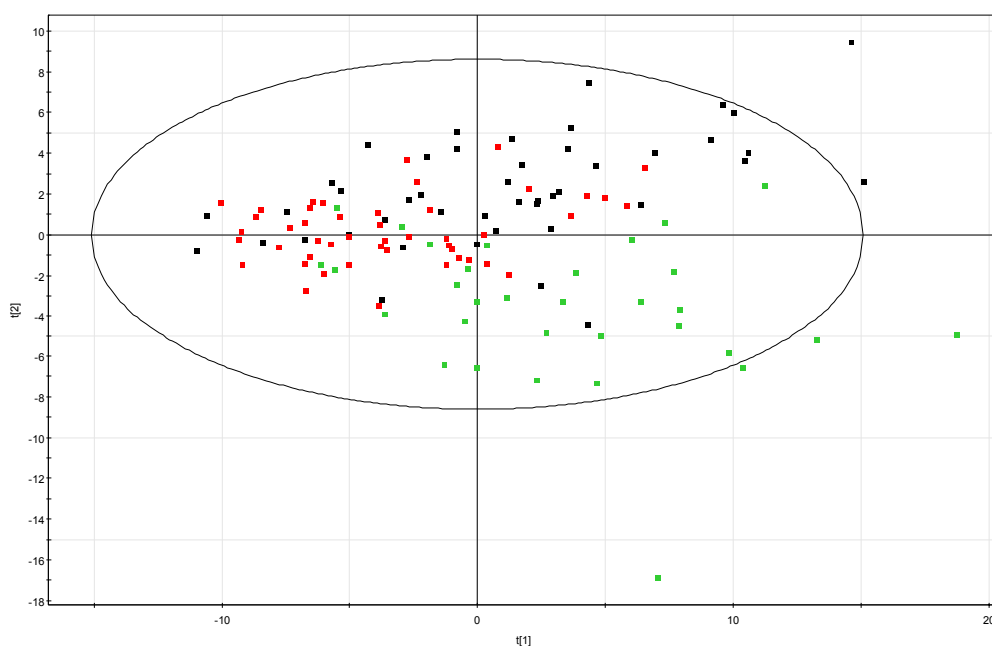


Figure 10: PLS-DA scores plot (UV scaled) of B, P and O urine samples from all the days of the administration, obtained by UHPLC-(+)ESI-HRMS. No significant clustering was observed between the B samples (green points) the O samples (black points) and the P samples (red points).

In order to highlight the administration effect and reveal the differences between the crossover administrations, the baseline B samples were excluded from the analysis

and PLS-DA models were created including only the O and P samples. The samples that were used in this step of the MVA can be found in Table 1. Thus, as an example, the PLS-DA scores plot of P and O urine samples from all the days of the administration, obtained by UHPLC-(+)ESI-HRMS is depicted in Figure 11.

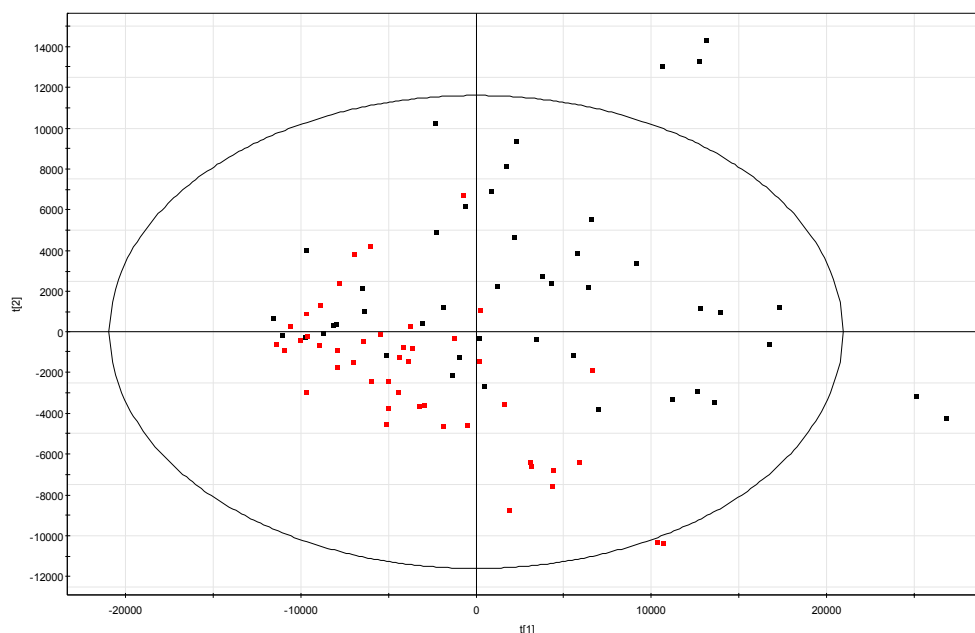


Figure 11: PLS-DA scores plot (UV scaled) of P and O urine samples from all the days of the administration, obtained by UHPLC-(+)ESI-HRMS. No significant clustering was observed between the O samples (black points) and the P samples (red points).

It is clear that the O samples (black points) cannot be clearly separated from the P samples (red points) in the PLS-DA scores plot and none of the PLS-DA dimensions are sufficient to distinguish the two groups. It should be noted also, that no clear clustering between the O and P samples was noted even when the data were observed in the dimensions of the first and the third or of the second and the third PLS-DA components. The same results were observed, when the other data sets (urine and serum in both positive and negative ion modes) were analysed using the PLS-DA method. The respective loadings scatter plots and bi-plots of the aforementioned PLS-

DA analyses were not evaluated further since the effect caused by the individuality of the humans (large biological variation) and the “day-based effect” as well, are still present.

In a next step, OPLS-DA analysis was performed in the data sets as it could be performed better than PLS-DA because there is a significant divergence in the within-class variation between the two classes (P and O). In this case also, as it can be shown in the OPLS-DA scores plot (Figure 12), it is evident that the OPLS-DA model was not able to provide a reasonable class estimation of the examined data. Even though the distribution of the scores of the subjects in the P group is slightly different from the O group and the observations obtained in the P samples (red) and the O samples (black) have the tendency to form two separate classes/groups, still we can observe a large dispersion across the two groups. Similar observations were resulted when the other data sets were analysed and thus no effort has been made to evaluate the loadings scatter plots, bi-plots and s-plots further due to their high complexity.

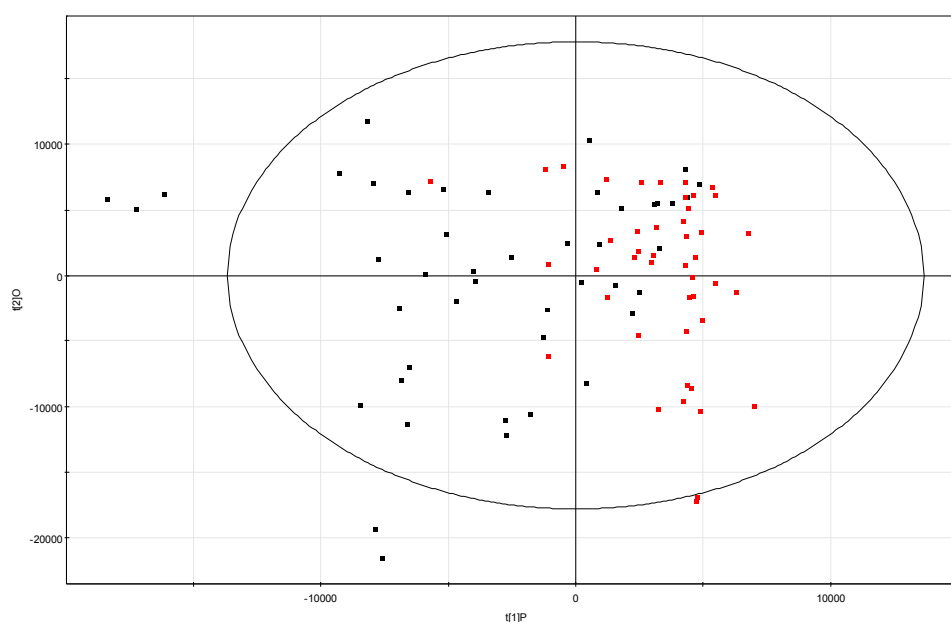


Figure 12: OPLS-DA scores plot (UV scaled) of P and O urine samples from all the days of the administration, obtained by UHPLC-(+)ESI-HRMS. No significant clustering was observed between

the O samples (black points) and the P samples (red points), even if they have the tendency to form two separate classes/groups.

Multilevel sPLS-DA

In the crossover studies, two main types of variation can be distinguished, the biological (between the subjects) and the variation induced by the administration of OE (within the subjects). Other types of variation could also be recognized in the current study (presumably of secondary importance); the variation caused by the week of administration (before or after the wash out period) - the week-based effect, the variation caused by the day of administration (possible cumulative effect) - the day-based effect, and the analytical variation – the analytical error (Figure 13). The total variation is resulting from the summary of the aforementioned individual variations, as follows:

$$\text{Var} = \Sigma(\text{Var}_{\text{within}} + \text{Var}_{\text{between}} + \text{Var}_{\text{week}} + \text{Var}_{\text{day}} + \text{Var}_{\text{analytical error}})$$

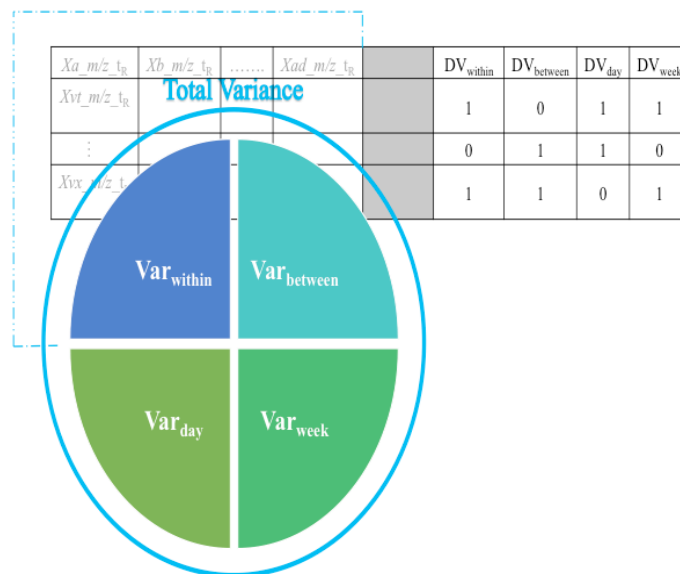


Figure 13. Total variance in a crossover study can be represented from the sum of individual variances. The white space can be thought as the analytical error. *Var* variation, *DV* discriminant variable.

Due to the fact that the fundamental variations existing between the subjects usually outweigh the administration's effects³⁹, which are typically small, and the classical MVA approaches (PCA, PLS-DA, OPLS-DA etc.) did not lead to any enlightening result, a ML-PLS-DA was performed on the serum and urine data sets. The latter approach can decompose the total variance, into singular variations, enabling the identification of the individual effects (e.g. the between subject variation, the week-based effect etc.) on the response. In other words, following the ML-PLS-DA approach, it is possible to minimize the aliasing effect of the other parameters except the administration, allowing to explore the hypothesis that the OE administration led to systematic metabolic changes. This will also facilitate the identification of possible metabolites as the most differentiating biomarkers caused by the administration (P or O). Overall such an approach was advantageous because it could separate the between subject from the within subject variation, prior to the MVA, taking into account other secondary types of variation in order to achieve better and clear clustering of the different groups.

After completing the classical MVA using different statistical approaches, such as PCA, PLS-DA, OPLS-DA etc., resulting to the inability of these processes to lead in safe conclusions, the ML-sPLS-DA analysis was performed on the serum and urine data sets. More specifically, the data sets resulted after missing value estimation and normalization, were initially used, setting the between ($\text{Var}_{\text{between}}$) and the within ($\text{Var}_{\text{within}}$) subjects variations, the week-based (Var_{week}) and the day-based (Var_{day}) effect as discriminate variables. Thus the discriminate variables used were **1.** the between subject variation (from A to I for the 9 individual subject respectively), **2.** the within subject variation (P or O for placebo or oleuropein administration respectively), **3.** the week-based effect (1 or 2 for the first or the second week of

administration respectively), and **4.** the day-based effect (from 1 to 5 for the different days of P or O administration in the two weeks of the crossover design).

In order to achieve the optimal results two parameters need to be considered: the number of discriminant vectors and the number of variables to select on each dimension (PLS component)⁶⁰. According to the Le Cao proposal⁶¹ due to the large sample number (>15) the k-fold cross-validation approach has been adopted. In the 1-level ML-PLS-DA approach, the classification error has been computed and effort has been made in every case to find the appropriate parameters that would afford small values⁶² (i.e. small classification values correspond to a robust model). Alternatively, when applying the 2-level ML-PLS-DA approach, the correlation between latent variables has been used as the criterion to assess the robustness of the proposed model (i.e. large correlation values in that case afford a robust model)⁶². After this step, the full data were analyzed given the tuned parameters.

Initially, an 1-level ML-sPLS-DA was applied to the datasets, keeping one variation at a time as discriminate variable (i.e. the within subject variation, $\text{Var}_{\text{within}}$) and subtracting another one simultaneously (i.e. the between subject variation, $\text{Var}_{\text{between}}$), in order to find out which variation has the highest impact to the data sets. The clustering, examined by visual inspection, and the tune parameter values resulting from the tuning procedure, have been used as the selection criteria for evaluating the importance of each singular variance. Furthermore the sparse variant of ML-PLS-DA has been employed in order to facilitate the variable selection. As an example, the 1-level ML-sPLS-DA scores plot of P and O urine samples from all the days of the administration, obtained by UHPLC-(+)ESI-HRMS is depicted in Figure 14. The samples that have been used in this step of the MVA can be found in Table 1.

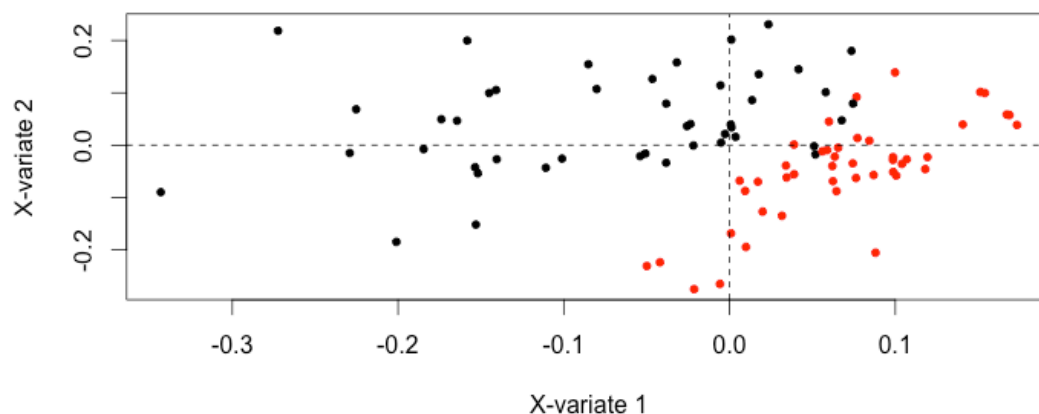


Figure 14. 1-level ML-PLS-DA analysis of the P and O urine samples from all the days of administration, obtained by UHPLC-(+)ESI-HRMS. The multilevel PLS-DA scores plot (X-variate 1, X-variate 2) of the administration effect $\text{Var}_{\text{within}}$ with the biological variation ($\text{Var}_{\text{between}}$) being subtracted is shown. Red circles represent the P samples. Black circles represent the O samples. A clear clustering mainly due to the first component, between P and O samples, is shown.

The cross-validation tuning criterion afforded for each sPLS-DA dimension a classification error of (0.12, 0.18, 0.21) for variable selection sizes of 50, 60, 70 metabolites on each dimension. It is clearly shown that, by subtracting the biological variation of individuals from the total variation, the administration effect is clearly emerging and the P (red) and O (black) samples can now be well separated according to the first component (x-variate 1). However in the second component (x-variate 2) they seem indistinguishable, proving that slight differences play a critical role in their grouping and the administration effect was not equal for all the subjects. The current observations leads to the conclusion that all subjects do not respond equally to the treatment, and demonstrates that subtracting the correct variation that obscures the classification according to the desired variable (e.g. the administration of OE in the current case), a single component could be sufficient to assess the different treatment effects within the subjects.

As a subsequent step of the previous multilevel analysis procedure, a 1-level ML-sPLS-DA was again applied to the same data sets, leaving as before the within subjects variation $\text{Var}_{\text{within}}$ (P or O administration) as the discriminate variable, but this time subtracting the variation caused by the week of administration Var_{week} (week-based effect). The cross-validation tuning criterion afforded an acceptable classification error for each sPLS-DA dimension (<0.22), for variable selection sizes of 50, 60, 70 metabolites correspondingly. The 1-level ML-sPLS-DA scores plot obtained is shown in Figure 15. It is clearly shown that the P (red) and O (black) samples are clustered separately, which leads to the conclusion that the week-based effect is playing also a critical role in the classification (subtracting the week-based effect from the total variation, the administration effect is highlighted). That could be possibly attributed to environmental changes during the weeks of the administration, to differences during the sample handling, to storage instability, or even to the wash out period between the two weeks of the administration.

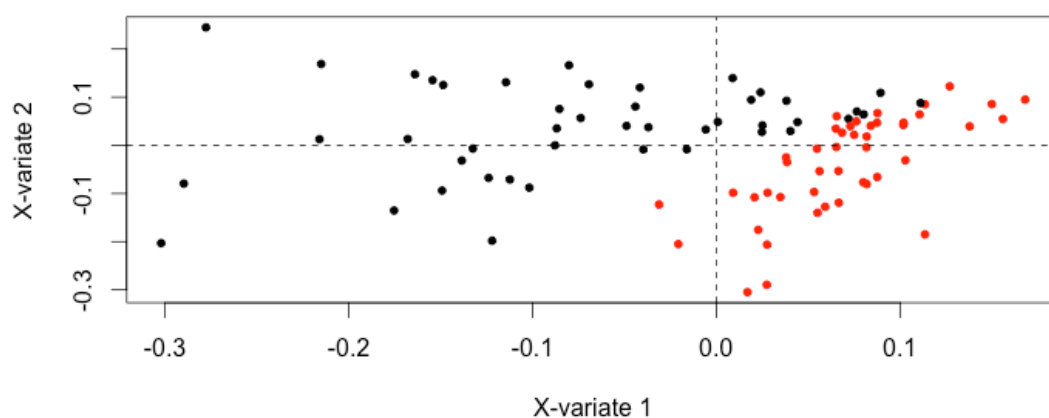


Figure 15. 1-level ML-PLS-DA analysis of the P and O urine samples from all the days of administration, obtained by UHPLC-(+)ESI-HRMS. The multilevel PLS-DA scores plot (X-variate 1, X-variate 2) of the administration effect $\text{Var}_{\text{within}}$ with the week-based effect (Var_{week}) being subtracted is shown. Red circles represent the P samples. Black circles represent the O samples. A clear clustering mainly due to the first component, between P and O samples, is shown.

Furthermore in order to investigate the day-based effect on the within subject variation, an 1-level ML-sPLS-DA was applied to the same data sets, leaving as before the within subjects variation $\text{Var}_{\text{within}}$ (different administration P or O) as the discriminate variable, but this time subtracting the variation caused by the day of the administration Var_{day} (day-based effect). As depicted in the 1-level ML-sPLS-DA scores plot obtained (Figure 16), the P (red) and O (black) samples are clustered separately whereas the classification error that calculated from the cross-validation tuning procedure was found to range in acceptable levels (<0.22). Two probabilities can be formulated regarding the repeated measurements, either that there is a cumulative effect which would lead the metabolome to move further away from the P samples with the day of administration in the metabolome space, i.e. the samples from the first days of O administration would be closer to the P samples, or in the opposite case to retract to the P samples location due to the induction of the corresponding metabolism enzymes, a “resistance-like” effect.

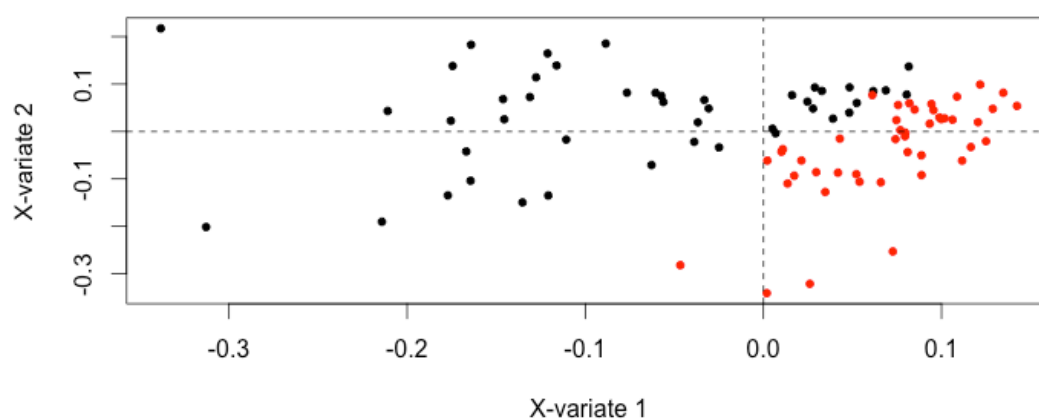


Figure 16. 1-level ML-PLS-DA analysis of the P and O urine samples from all the days of administration, obtained by UHPLC-(+)ESI-HRMS. The multilevel PLS-DA scores plot (X-variate 1, X-variate 2) of the administration effect $\text{Var}_{\text{within}}$ with the day-based effect (Var_{day}) being subtracted is shown. Red circles represent the P samples. Black circles represent the O samples. A clear clustering mainly due to the first component, between P and O samples, is shown.

As expected, two clusters can be observed in all the above case. It is evident from the scores plot (Figure 17) that the samples follow the pattern described in the second case i.e. the samples from the fifth day are located closer to the P samples compared to the ones from the initial administration. Thus it would be helpful to highlight this effect by examining the two extremes of the administration days i.e. the first vs. the last administration day (section).

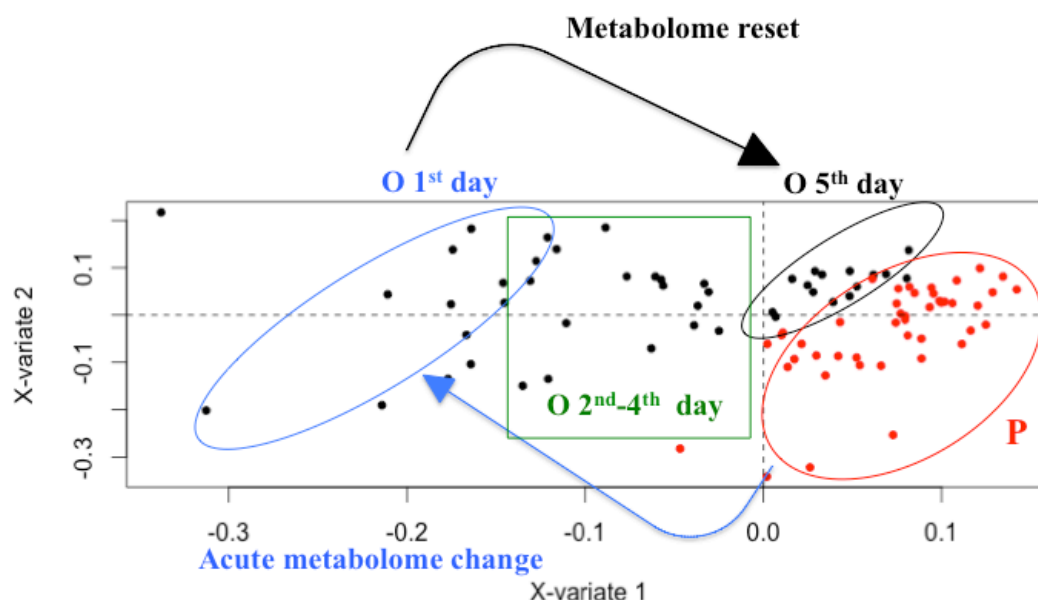


Figure 17. 1-level ML-PLS-DA analysis of the P and O urine samples from all the days of administration, obtained by UHPLC-(+)ESI-HRMS - **Figure 16** annotated. Red circles represent the P samples. Black circle represent the O samples from the 5th day of the administration. Blue circle represents the O samples from the 1st day of administration. Green box represents O samples from 2nd-4th day of administration. An acute metabolome change between P and O 1st day is evident (blue arrow), whereas a gradual metabolome reset is readily observed (black arrow).

The current findings do not support the expected theoretical model (Figure 17A) where the repeated administration should produce a different metabolic profile that presumably continuous to differentiate after each administration from the control

group (an accumulation effect) and ultimately should stabilize on a different metabolic profile. In contrast, the observed phenomenon is consistent with a reset of the metabolome to its initial state (Figure 17A).

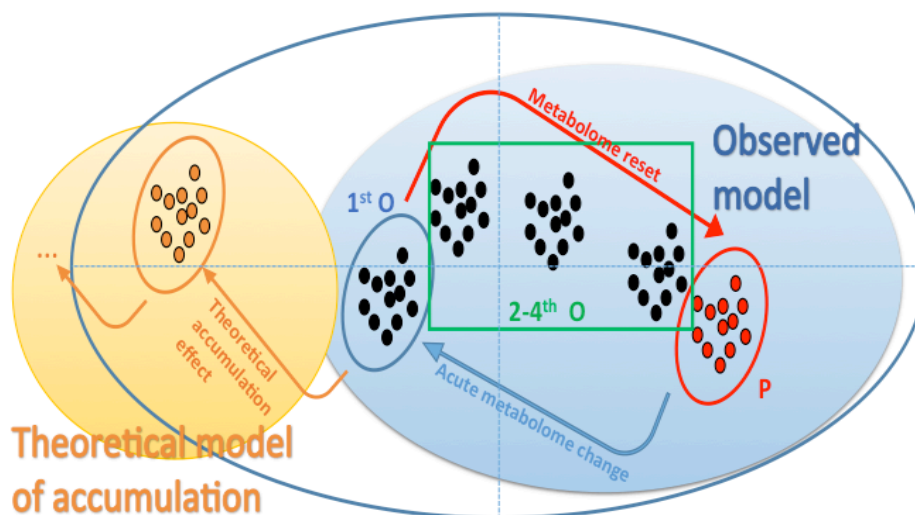


Figure 17A. Theoretical model of accumulation vs. observed in respect of PLS-DA approach.

Overall, taking into account the above observations from the 1-level ML-sPLS-DA approaches, it could be reasonably deduced that the between subject variation along with the week-based and the day-based effects outweighed the effect caused by the administration. Hence, the subtraction of their variation from the data sets is of high importance, in order to highlight the administration effect alone, and thus proceed to the evaluation of the loadings that are linked only to the administration effect (possible biomarkers).

After completion of the 1-level ML-sPLS-DA approaches, a 2-level ML-sPLS-DA approach was applied to the same data sets, keeping two variations at a time as the discriminate variables and subtracting one simultaneously. In other words the grouping algorithm will use two discriminant values (e.g. week-based effect and administration effect) and it will subtract a third effect (e.g. biological variation

effect) in the same calculation. The clustering, examined by visual inspection, and the tune parameter values resulting from the tuning procedure (the correlation between latent variables), have been used as the selection criteria for evaluating the importance of each singular variance.

Thus, a 2-level ML-sPLS-DA analysis was performed with the same data sets, subtracting initially the between subject variation $\text{Var}_{\text{between}}$ from the total variance and keeping the combination of the within subject variation and the day-based effect ($\text{Var}_{\text{within}} + \text{Var}_{\text{day}}$) as the discriminant variables. As an example, the 2-level ML-sPLS-DA 3-D scores plot of P and O urine samples from all the administration days (UHPLC-(+)ESI-HRMS) is depicted in Figure 18. The samples used can be found in Table 1. The sequential (one component at a time) validation tuning criterion afforded an acceptable maximum correlation of (0.85, 0.83, 0.85) for each sPLS-DA dimension using variable selection sizes of 50, 60, 70 metabolites on each dimension respectively. The value of 0.53 for the correlation in the fourth dimension indicated that only 3 sPLS-DA dimensions should be chosen for this analysis. As it demonstrated in Figure 18, the P (red) and O (black) samples are separated in two clusters exhibiting a strong effect of the administration on the metabolism of the subjects.

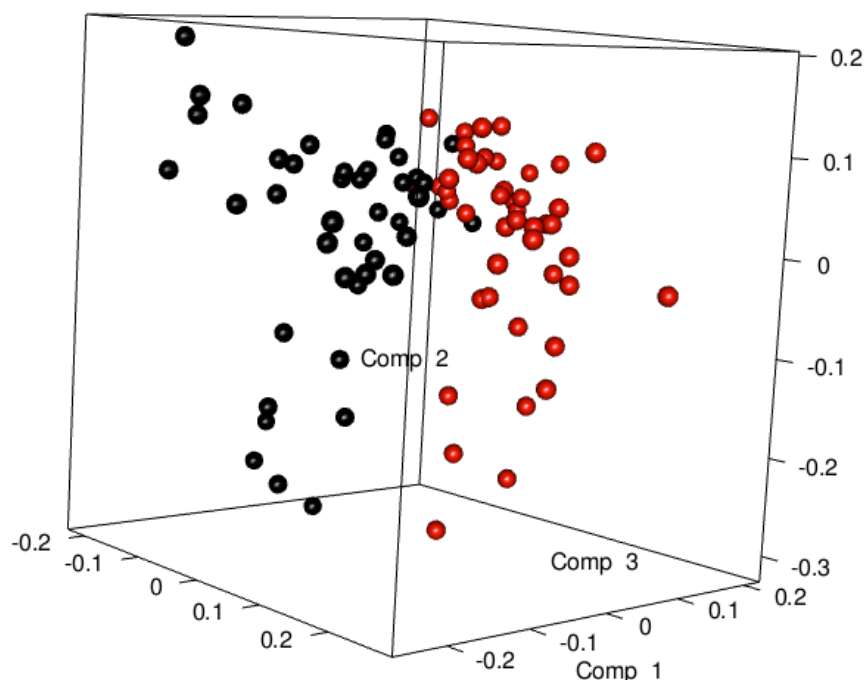


Figure 18. 2-level ML-PLS-DA analysis of the P and O urine samples from all the days of administration, obtained by UHPLC-(+)ESI-HRMS. The 3-D multilevel PLS-DA scores plot (Comp 1, Comp 2, Comp 3) of the combination of the within subject variation and the day-based effect ($\text{Var}_{\text{within}} + \text{Var}_{\text{day}}$) with the biological variation ($\text{Var}_{\text{between}}$) being subtracted is shown. Red circles represent the P samples. Black circles represent the O samples. A clear clustering between P and O samples, is shown.

Nevertheless it seems that the repetitive dosing does not create any accumulation effect, as no trend regarding the administration days is evident. In that case one would expect that the metabolome should migrate further away from the placebo samples in a day dependent fashion (the 2nd day further than the 1st one, the 3rd further than the second etc.). On the contrary the samples of the 4th and the 5th day seem to retract to the placebo samples as shown in Figure 17. A possible explanation could be that some kind of resistance due to enzyme induction is created (as O could be considered as an xenobiotic). Thus a different administration scheme should be considered either encompassing lower dosing or diminishing the frequency of the administration.

As a subsequent step of the previous analysis, a 2-level ML-sPLS-DA was again applied to the same data sets, subtracting this time the day-based effect Var_{day} from

the total variance and keeping as discriminant variable the combination of the within subject variation and the week-based effect ($\text{Var}_{\text{within}} + \text{Var}_{\text{week}}$). Figure 19 shows the 2-level ML-sPLS-DA scores plot (P and O urine samples - all administration days - UHPLC-(+)ESI-HRMS). The samples that have been used in this step of the MVA can be found in Table 1.

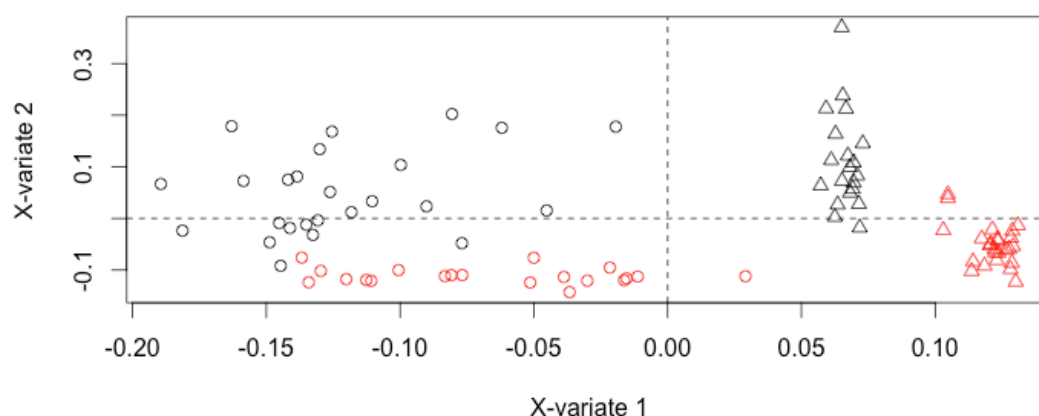


Figure 19. 2-level ML-PLS-DA analysis of the P and O urine samples from all the days of administration, obtained by UHPLC-(+)ESI-HRMS. The multilevel PLS-DA scores plot (X-variate 1, X-variate 2) of the combination of the within subject variation and the week-based effect ($\text{Var}_{\text{within}} + \text{Var}_{\text{week}}$) as the discriminant values and with the day-based effect (Var_{day}) being subtracted is shown. Circles represent the samples from the first week of administration and triangles represent the samples from the second week of administration. Black color stands for the O samples and red color stands for the P samples. A clear clustering between the first and the second week due to the first component and also between the P and O samples due to the second component is shown.

The validation tuning criterion presented maximum correlation of 0.96, 0.96, 0.96 for each sPLS-DA dimension, employing variable selection sizes of 50, 60, 70 metabolites respectively on each dimension. It is clearly shown that, subtracting the day-based effect Var_{day} and keeping the combination $\text{Var}_{\text{within}} - \text{Var}_{\text{week}}$ as the discriminant variables, a strong week-based effect is highlighted where the samples from the first week of the administration (circles) are clustered separately from those of the second week (triangles) according to the first component (x-variate 1).

Furthermore, in the second component (x-variate 2), strong clustering is again observed, which is resulting from the different administration scheme ($\text{Var}_{\text{within}}$). Thus the P (red) and O (black) samples are well separated, as it demonstrated in Figure 19. The current observations leads to the conclusion that along with the administration scheme, the week of intervention also adds a significant amount of variation. The reason of the observation is not determined but it can be speculated that such a difference could be attributed to variation in sample handling or/and storage, in differences to the environmental – physiological situation of the subjects as well as to a possible induced instability of the biomaterial used.

In the same manner as before the individuality of the subjects has been taken into account thus it has been used as the second discriminant variable along with the administration. The subtracted variance was the day-based effect in this instance. The results show that a strong discrimination can be observed in this case, showing that when taking the biological variation into account it could be considered as the most influential factor of all. This could be clearly anticipated in the representative Figure 9 by the clustering of the subjects mainly due to their biological individuality whereas the administration seems to contribute minimally. In other words 9 clusters formed, encompassing mainly each individual and in a much lesser degree the administration information.

As mentioned above, it is essential to highlight the day-based effect by examining the two extremes of the administration days i.e. the first (1st) vs. the last (5th), subtracting the between subjects variation $\text{Var}_{\text{between}}$ from the total variance and keeping as discriminant variables the within subject variation and the day-based effect ($\text{Var}_{\text{within}}$ and Var_{day}). As an example, the chemometric analysis of the urine UHPLC-(+)ESI-HRMS data set, is presented. Figure 20 shows the scores plot of the 2-level

ML-sPLS-DA analysis of P and O urine samples from the 1st and 5th day of administration. The samples that were used can be found in Table 6.

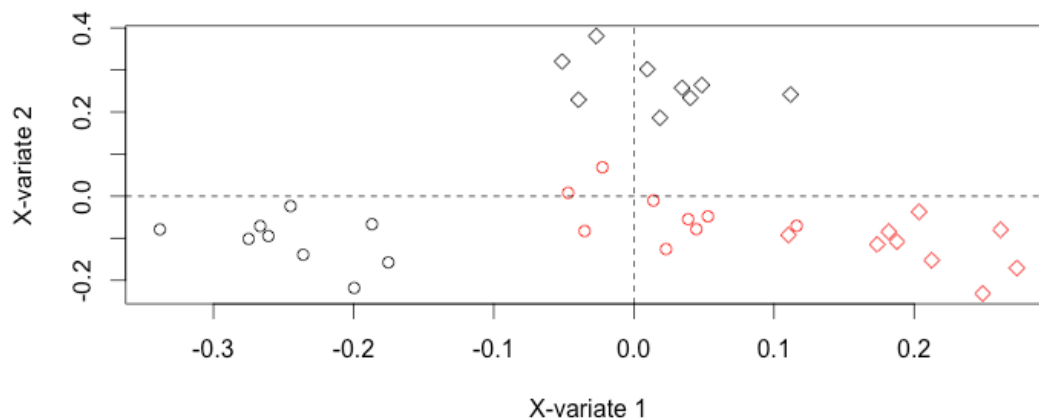


Figure 20. 2-level ML-PLS-DA analysis of the P and O urine samples from the two extremes of the administration days i.e. the first (1st) vs. the last (5th), obtained by UHPLC-(+)ESI-HRMS. The multilevel PLS-DA scores plot (X-variate 1, X-variate 2) of the combination of the within subject variation and the day-based effect ($\text{Var}_{\text{within}} + \text{Var}_{\text{day}}$) with the biological variation ($\text{Var}_{\text{between}}$) being subtracted is shown. Circles represent the samples from the 1st day of administration and diamonds represent the samples from the 5th day of administration. Black color stands for the O samples and red color stands for the P samples. Clear clustering between the 1st and the 5th day and also between the P and O samples is shown.

A maximum correlation of 0.98, 0.99, 0.99 for variable selection sizes of 50, 60, 70 metabolites respectively for each sPLS-DA dimension, was obtained applying the validation tuning criterion. As demonstrated in Figure 20, by subtracting the biological variation of individuals ($\text{Var}_{\text{between}}$) from the total variation and keeping the combination of the within subject variation and the day-based effect as discriminant variables in the decomposed data set, two main clusters can now be observed exhibiting a systematic difference between the P (red) and O (black) administration scheme thereof. Besides that, clustering can be also observed concerning the day of administration. Figure 20, clearly shows the difference between the 1st and the 5th day of the administration where the O samples from the 1st day of the administration

(black circles) are well separated from the O samples from the 5th day of the administration (black diamonds). Similar clustering was also observed in the case of P samples, where the P samples from the first day of the administration (red circles) are well separated from the P samples from the fifth day of the administration (red diamonds). Another representation of Figure 20 can be seen in Figure 21.

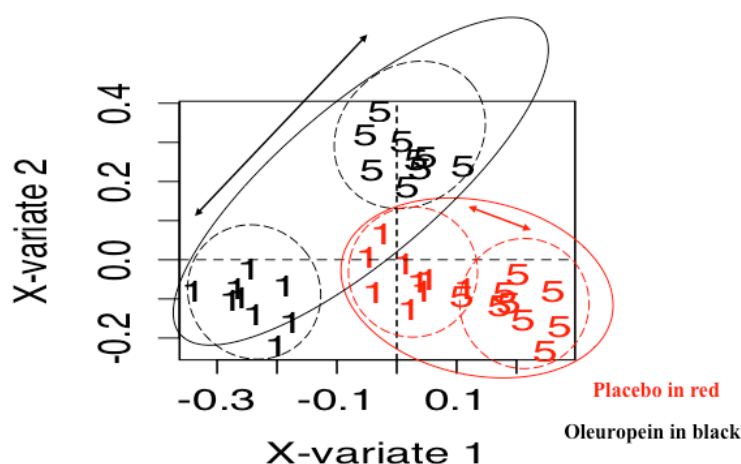


Figure 21. 2-level ML-PLS-DA analysis of the P and O urine samples from the two extremes of the administration days i.e. the first (1st) vs. the last (5th), obtained by UHPLC-(+)ESI-HRMS – **Figure 20** annotated. Circles represent the samples from the 1st day of administration and diamonds represent the samples from the 5th day of administration. Black color stands for the O samples and red color stands for the P samples whereas “1” stands for the 1st day and “5” for the 5th day. The distance between the 1st and 5th day metabolome is larger for the O samples, which shows the effect of OE administration. Furthermore, O samples from 5th day merges with the P samples in a common cluster – a metabolome reset effect.

It is evident that there is a day effect, which is notably smaller to the P group compare to the O group (the length of the samples between the centroids of the clusters is smaller in the case of P samples). In other words, there is a marked effect of OE on the metabolome. Furthermore, according to the first component the 5th day of OE administration is indistinguishable from the P group, a fact that further supports the assumption of the metabolome reset. Samples from the 5th day of OE

administration, seems to have a trend to merge in a common cluster with the P samples.

Similar observations were obtained from the MVA of the remaining serum and urine data sets. Therefore it has been considered necessary to proceed with the evaluation of the loadings plots, in order to discover possible biomarkers related to OE administration.

Overall it is evident that simple MVA analysis is not sufficient for the full evaluation of such complex experiments due to the fact that “sub-variances” contribute to the total variance, so the last should be decomposed mathematically. Thus multilevel approaches should be employed in order to discern safe conclusions. In the current case seems that the rest parts of variation play a near equally important role as the administration scheme, a fact that should be also studied.

Weeks		1 st week							2 weeks wash out	2 nd week						
Subjects	Days	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
A		1	11	21	31	41	51	61		1B	11B	21B	31B	41B	51B	61B
B		3	13	23	33	43	53	63		3B	13B	23B	33B	43B	53B	63B
C		4	14	24	34	44	54	64		4B	14B	24B	34B	44B	54B	64B
D		5	15	25	35	45	55	65		5B	15B	25B	35B	45B	55B	65B
E		6	16	26	36	46	56	66		6B	16B	26B	36B	46B	56B	66B
F		7	17	27	37	47	57	67		7B	17B	27B	37B	47B	57B	67B
G		8	18	28	38	48	58	68		8B	18B	28B	38B	48B	58B	68B
H		9	19	29	39	49	59	69		9B	19B	29B	39B	49B	59B	69B
I		10	20	30	40	50	60	70		10B	20B	30B	40B	50B	60B	70B

Table 6. A list of the tested samples along with their corresponding labels, used for the completion of the 2-level ML-sPLS-DA analysis of P and O urine samples from the 1st and 5th day of administration. Oleuropein samples (O) are in black and placebo (P) samples are in red. Samples in light grey color were not considered in this step of MVA

Metabolites

The aim of an ML-PLS-DA approach is not only to establish if the metabolite profiles of the different groups could be systematically distinguished, but also to identify the contributing variables (potential biomarkers). The loadings plots provide this information and are particularly useful in complex models such as in the current case. In order to investigate a limited number of putative metabolites that were the most important ones, the sparse variant of ML-PLS-DA namely the ML-sPLS-DA has been used. Two strategies have been employed in order to extract the most influential loadings from the statistical analysis for all the datasets. According to the first strategy, the ten most important loadings (those with the largest absolute value) of the X-variate 1 (corresponding to component 1) have been considered. These can be considered as the metabolites that contribute the most to the largest latent variable (the X-axis) in the data set. In a second step the ten most important loadings (those with the largest absolute value) on the X-variate 2 (corresponding to component 2) have been considered and the same procedure has been followed for the selection of metabolites in X-variate 3. Following, the next 10 metabolites in the X-variate have been considered and the procedure stops when all sparse components have been selected and evaluated (Figure 22a, b and c). The second strategy involves the selection of metabolites using adjustable-radius Euclidean distances to the ML-sPLS-DA variable plot. The different radii that have been employed started from the 0.9 value and were decreasing using a 0.1 step. In other words ever-decreasing ellipses have been used and the variables between the Hotelling ellipse and the adjustable-radius one have been selected Figure 22d. A representative variables plot (loadings plot) of the P and O urine samples from the 1st and the 5th day of the administration, obtained by UHPLC-(+)ESI-HRMS is depicted in Figure 23.

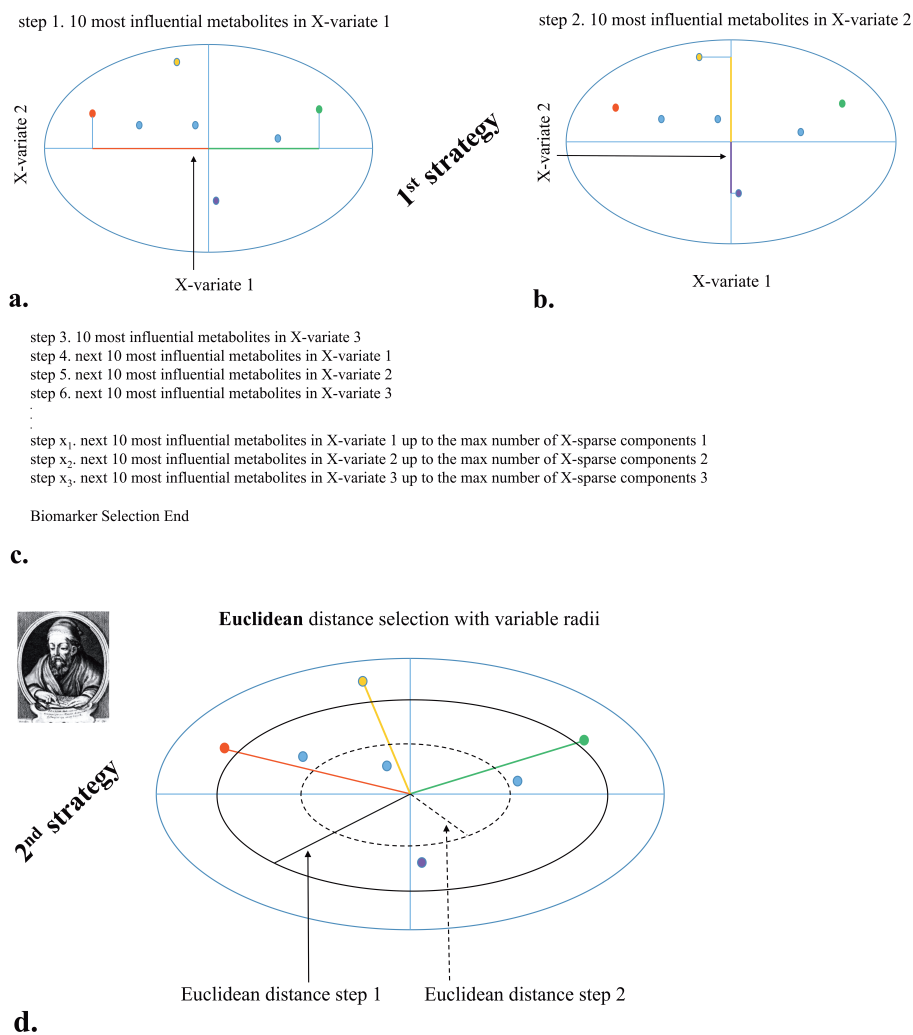


Figure 22. Strategies for metabolites' selection. The first one (**a**, **b** and **c**) is based in the distance from the origin in respect to an individual principal component. The second one (**d**) is based on an adjustable Euclidean distance from the origin.

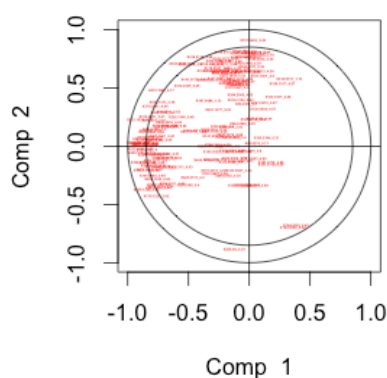


Figure 23. A representative variables plot (loadings plot) of the P and O urine samples from the 1st and the 5th day of the administration, obtained by UHPLC-(+)ESI-HRMS. A radius value of 0.85 has been applied for the selection of the metabolites.

The sparse loadings selected by either strategy, are summarized in Table 7 as a. metabolite id., b. m/z feature, c. t_R (min), d. complex/dimer/product/precursor, e. experimental deconvoluted m/z (precursor), f. theoretical m/z (precursor), g. Δm (ppm), h. corresponding RDB value, i. deconvoluted pseudo MS/MS ions, j. possible identification name, k. possible molecular formula, l. corresponding monoisotopic exact mass, m. change trend in O group compared with P group (up or down regulation) and n. corresponding dataset (serum or urine, positive or negative ESI dataset). It should be noted that features with identical m/z but different t_R possibly correspond to isomers. Furthermore, an analysis for possible adducts (M-H₂O-H, M-H, M+Na-2H, M+Cl, M+K-2H, M+FA-H, M+Br, 2M-H, 2M+FA-H) has been performed for all the features in order to verify the possibility of occurrence of chromatographic peaks due to adduct formation. These features can also be identified as present or not in the raw data.

The next step was to identify the selected features, or at least, a subset of the most dominant ones, as these are clearly important metabolites for distinguishing the P and O groups. The classical approach is to match the selected features in terms of mass accuracy in online databases in order to recognize possible metabolites. Nevertheless, careful selection of the respective online databases should be made in order to reassure as much as possible the correct identification of the features found, e.g. the Human Metabolome Data Base (HMDB) can provide reliable information for the human metabolome, whereas the use of an alternative non-specific database (ChemSpider database) could potentially obscure the procedure. Furthermore, additional data can be used to confirm the proposed structures, such as the isotopic pattern and the RDB information. More specifically, the comparison of the expected isotopic pattern with the measured one should fulfill the criterion of 5% of matching,

whereas the RDB, which provides the degree of molecule's unsaturation, should correspond to the proposed structure.

Id.	<i>m/z</i>	<i>t_R</i> (min)	Complex/ Dimer/ Product/ Precursor*	Experimental Deconvoluted <i>m/z</i> (Precursor)	Theoretical <i>m/z</i> (Precursor)	Δm (ppm)	RDB	Pseudo MS/MS ions	Name	Molecular Formula	Exact Mass	Trend	Dataset
M1	188,0704	1,02	Product	205,0970	205,0972	-0,9	6,5	-	L-Tryptophan	C ₁₁ H ₁₂ N ₂ O ₂	204,0899	↑	Urine ESI(+)
M2	188,0704	1,46	Precursor	188,0704	188,0703	-1,0	7,5	-	Indoleacrylic acid	C ₁₁ H ₉ NO ₂	187,1947	↑	Urine ESI(+)
M3	161,1070	3,96	Precursor	161,1070	161,1073	-1,9	5,5	-	Tryptamine	C ₁₀ H ₁₂ NO ₂	160,2157	↑	Urine ESI(+)
M4	206,0810	4,88	Precursor	206,0810	206,0812	0,2	6,5	-	Indolelactic acid	C ₁₁ H ₁₁ NO ₃	205,0738	↑	Urine ESI(+)
M5	221,1534	4,77	Product	286,2013	286,2013	0	2,5	-	2-Octenoylcarnitine	C ₁₅ H ₂₇ NO ₄	285,3792	↑	Urine ESI(+)
M6	279,1338	3,02	Precursor	279,1338	279,1339	-0,5	6,5	-	N1-(alpha-D- ribosyl)-5,6- dimethyl- benzimidazole/L- phenylalanyl-L- hydroxyproline	C ₁₄ H ₁₈ N ₂ O ₄	278,3037	↑	Urine ESI(+)
M7	316,2481	5,67	Precursor	316,2481	316,2482	-0,4	1,5	-	Decanoylcarnitine	C ₁₇ H ₃₃ NO ₄	315,4482	↑	Urine ESI(+)
M8	251,1277	5,80	Precursor	251,1277	251,1278	-0,3	5,5	-	Ubiquinone-1	C ₁₄ H ₁₈ O ₄	250,2903	↑	Urine ESI(+)
M9	314,2325	5,30	Precursor	314,2325	314,2326	-0,4	2,5	-	9-Decenoylcarnitine	C ₁₇ H ₃₁ NO ₄	313,4323	↑	Urine ESI(+)
M10	302,2325	5,23	Precursor	302,2325	302,2326	-0,6	1,5	-	2,6 Dimethylheptanoyl carnitine	C ₁₆ H ₃₁ NO ₄	301,4216	↑	Urine ESI(+)
M11	328,2481	5,92	Precursor	328,2481	328,2482	-0,2	2,5	-	Carnitine scaffold	C ₁₈ H ₃₄ NO ₄	327,2404	↑	Urine ESI(+)
M12	130,0494	4,30	Product	147,0760	147,0764	-2,7	1,5	-	L-Glutamine	C ₅ H ₁₀ N ₂ O ₃	146,1445	↓	Urine ESI(+)
M13	293,1471	4,30	Dimer (2M+H)**					-					
M14	147,0760	4,30	Precursor					-					
M15	286,2012	4,55	Precursor	286,2012	286,2013	-0,5	2,5	-	2-Octenoylcarnitine	C ₁₅ H ₂₇ NO ₄	285,3792	↓	Urine ESI(+)
M16	330,2274	4,24	Precursor	330,2274	330,2275	-0,6	2,5	-	6-Keto- decanoylcarnitine	C ₁₇ H ₃₁ NO ₅	329,4317	↑	Urine ESI(+)

Id.	m/z	t_R (min)	Complex/ Dimer/ Product/ Precursor*	Experimental Deconvoluted m/z (Precursor)	Theoretical m/z (Precursor)	Δm (ppm)	RDB	Pseudo MS/MS ions	Name	Molecular Formula	Exact Mass	Trend	Dataset
M17	105,0329	1,71	Product	118,0646	118,0651	-4,6	5,5	-	Indole	C ₈ H ₈ N	117.1479	↓	Urine ESI(+)
M18	310,2012	4,86	Precursor	310,2012	310,2013	-0,2	4,5	-	Carnitine scaffold	C ₁₇ H ₂₇ NO ₄	309,1935	↑	Urine ESI(+)
M19	137,0593	4,49	Precursor	137,0593	137,0597	-3	4,5		Phenylacetic acid	C ₈ H ₈ O ₂	136.0524	↑	Urine ESI(+)
M20	349,1121	2,96	Precursor	349,1121	349,1104	4,8	8,5	215/183/171/157/133	Estrone-sulfate	C ₁₈ H ₂₂ O ₅ S	350,4290	↑	Urine ESI(-)
M21	377,0854	0,30	Complex of 189,0406 and 187,0381***	UNKNOWN								↑	Urine ESI(-)
M22	201,1134	3,58	Precursor	201,1134	201,1121	6,2	2,5	183/157	Sebacic acid	C ₁₀ H ₁₈ O ₄	202,2475	↓	Urine ESI(-)
M23	389,0989	0,90	Dimer (2M+H)**	194,0452	194,0448	2,1	6,5	178/96	Hydroxyhippuric acid	C ₉ H ₉ NO ₄	195,1721	↓	Urine ESI(-)
M24	357,1092	1,71	Dimer (2M+H)**	178,0501	178,0499	1,3	6,5	134/96	Hippuric acid	C ₉ H ₉ NO ₃	179,1727	↑	Urine ESI(-)
M25	239,1610	16,0	Precursor	UNKNOWN								↑	Serum ESI(+)
M26	758,5666	20,1	Precursor	758,5666	758,5694	-3,7	3,5	-	Glycerophospholipid Skeleton	C ₄₂ H ₈₀ NO ₈ P	757,5621	↑	Serum ESI(+)
M27	174,0132	2,4	Precursor	UNKNOWN								↓	Serum ESI(+)
M28	158,0806	10,4	Precursor	158,0806	158,0812	-3,1	2,5	79/80/97/98/99	Tiglylglycine	C ₇ H ₁₁ NO ₃	157,1671	↓	Serum ESI(+)
M29	414,2993	10,4	Precursor	414,2993	414,3003	-2,3	7,5	-	N-docosahexaenoyl GABA	C ₂₆ H ₃₉ NO ₃	413,2924	↓	Serum ESI(+)
M30	252,9552	12,4	Precursor	UNKNOWN								↓	Serum ESI(+)
M31	97,9681	12,9	Precursor	UNKNOWN								↑	Serum ESI(+)
M32	154,9896	13,2	Precursor	UNKNOWN								↓	Serum ESI(+)
M33	670,2881	11,6	Precursor	UNKNOWN								↑	Serum

Id.	<i>m/z</i>	<i>t_R</i> (min)	Complex/ Dimer/ Product/ Precursor*	Experimental Deconvoluted <i>m/z</i> (Precursor)	Theoretical <i>m/z</i> (Precursor)	Δm (ppm)	RDB	Pseudo MS/MS ions	Name	Molecular Formula	Exact Mass	Trend	Dataset
													ESI(-)
M34	465,3042	24,1	Product	465,3042	465,3033	1,9	5,5	97/385	Cholesterol sulfate	C ₂₇ H ₄₆ O ₄ S	466,7170	↓	Serum ESI(-)
M35	167,0214	0,9	Product	167,0214	167,0200	6,2	6,5	96/124	Uric acid	C ₅ H ₅ O ₃ N ₄	168,1103	↓	Serum ESI(-)
M36	391,2853	9,9	Product	391,2853	391,2843	1,2	5,5	-	Bile acid	C ₂₄ H ₄₀ O ₄	392,2926	↓	Serum ESI(-)
M37	103,0399	1,1	Product	103,0399	103,0390	3,7	1,5	-	Hydroxybutyric acid	C ₄ H ₈ O ₃	104,1045	↓	Serum ESI(-)
M38	135,0302	1,0	Product	135,0302	135,0288	6,1	1,5	-	Erythronic acid/Threonic acid	C ₄ H ₈ O ₅	136,1033	↓	Serum ESI(-)

^a (↑): up-regulated in O group. (↓): down-regulated in O group.

Table 7: Candidate metabolites identified in O and P groups from serum- and urine-based metabolomic approaches obtained from both (+) and (-) ESI analyses. The features are presented as a. metabolite id., b. *m/z* feature, c. *t_R* (min), d. complex/dimer/product/precursor, e. experimental deconvoluted *m/z* (precursor), f. theoretical *m/z* (precursor), g. Δm (ppm), h. corresponding RDB value, i. deconvoluted pseudo MS/MS ions, j. possible identification name, k. possible molecular formula, l. corresponding monoisotopic exact mass, m. change trend in O group compared with P group (up or down regulation) and n. corresponding dataset (serum or urine, positive or negative ESI dataset).

Additionally, there is a possibility of induced fragmentation, which would lead to other ions besides the molecular (pseudo molecular or deprotonated). In order to elaborate on this possibility, a deconvolution step has been performed to the corresponding spectra, using the appropriate software i.e. the Mass Frontier 5.0.1 and its freely available counterpart AMDIS. The major disadvantage of the AMDIS software is that the embedded to the software algorithms could be only used for unit resolution spectra. In contrary Mass Frontier is using proprietary algorithms that take the mass accuracy into advantage. Thus, in order to identify any possible fragments that are connected to a molecular ion, a spectral deconvolution procedure has also been included to the workflow. In other words, a deconvolution step has been employed in order to produce possible pseudo-fragmentation spectra due to procedures such as in source-fragmentation. Such algorithms operate by monitoring the ions that rise and fall simultaneously meaning that all the ions that follow the same pattern under a selected chromatographic peak originate from the same substance. This way mass spectra that include possible fragments along with the molecular ion can be revealed. These pseudo MS/MS fragments (product ions) and their possible parent ions were submitted again to online database searching exploiting also the isotopic pattern and RDB information. An illustrative example follows. The m/z feature 188.0704 with $t_R=1.02$ min has been selected as one of the first loadings in X-variate 1 from the urine (+) ESI dataset. On line database searching resulted to 3 hits only, respecting the posed 3 ppm mass accuracy limits (Figure 24). The isotopic pattern one of the candidate metabolites, the deethylatrazine, did not fulfill the posed criteria thus the hit has been disqualified (Figure 25). Furthermore the remaining hits are not associated to any meaningful human metabolite.

Total: 3 Metabolites

METLIN ID	MASS	Δ ppm	NAME	MS/MS	STRUCTURE
34509	[M+H] ⁺ m/z 188.0706 M 187.0633	0	3-amino-2-naphthoic acid Formula: C ₁₁ H ₉ NO ₂ CAS: 5959-52-4	View	
5702	[M+H] ⁺ m/z 188.0706 M 187.0633	1	Indoleacrylic acid Formula: C ₁₁ H ₉ NO ₂ CAS: 1204-06-4	View	
66434	[M+H] ⁺ m/z 188.0697 M 187.0625	3	Deethylatrazine Formula: C ₆ H ₁₀ ClN ₅ CAS: 6190-65-4	NO	

Figure 24. The tryptophan example. The m/z feature 188.0704 with $tR=1.02$ min has been selected as one of the first loadings in X-variate 1 from the urine (+) ESI dataset. On line database searching (METLIN example) resulted to 3 hits only, respecting the posed 3 ppm mass accuracy limits. The isotopic pattern one of the candidate metabolites did not fulfill the posed criteria and thus the hit has been disqualified. Furthermore the remaining hits are not associated to any meaningful human metabolite.

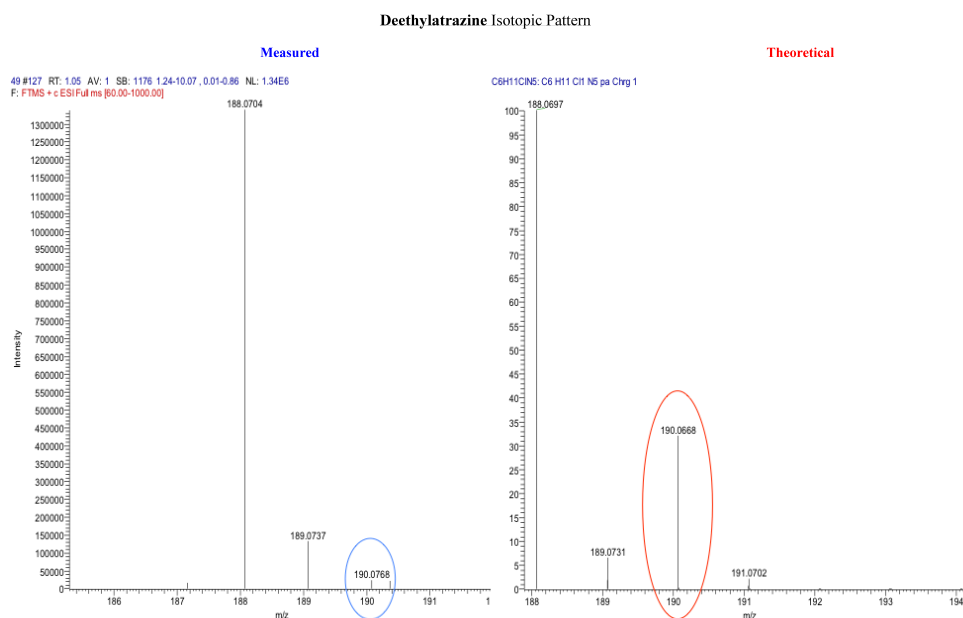


Figure 25. Comparison of deethylatrazine's measured and theoretical isotopic pattern. The theoretical isotope 190.07 due to the presence of Cl in the molecule (red ellipse), is not present in the measured isotopic pattern (blue) and thus the hit has been disqualified.

Application of the AMDIS algorithm to the chromatographic peak associated mass spectra of the selected feature, revealed two more “co-eluting” ions, i.e. 205 and 146, sharing a common pattern (the “model 1” equation in AMDIS). Thus, it has been speculated that the m/z 205 ion corresponded to the pseudo-molecular ion whereas the 188 and the 146 corresponded to fragments. Database searching employing both HMDB and METLIN, indicated that L-tryptophan exhibits the correct accurate molecular mass and shares the correct fragmentation pattern. In addition L-tryptophan has the correct isotopic pattern and RDB, and thus the combination of all the aforementioned information allows for the unequivocal identification of the hit.

The same identification has been achieved by the MassFrontier software, with the 205.0972 and 146.0597 ions being recognized (Figure 26). However in other cases AMDIS failed to lead to the correct assignment of the model presumably due to the low-resolution restriction of the algorithm. Thus Mass Frontier has been used as the method of choice for evaluating the spectra. Retrospection of the data (XIC extraction for each sample) obtained, it has been found that L-tryptophan is up regulated in the O group (statistically significant difference by ANOVA).

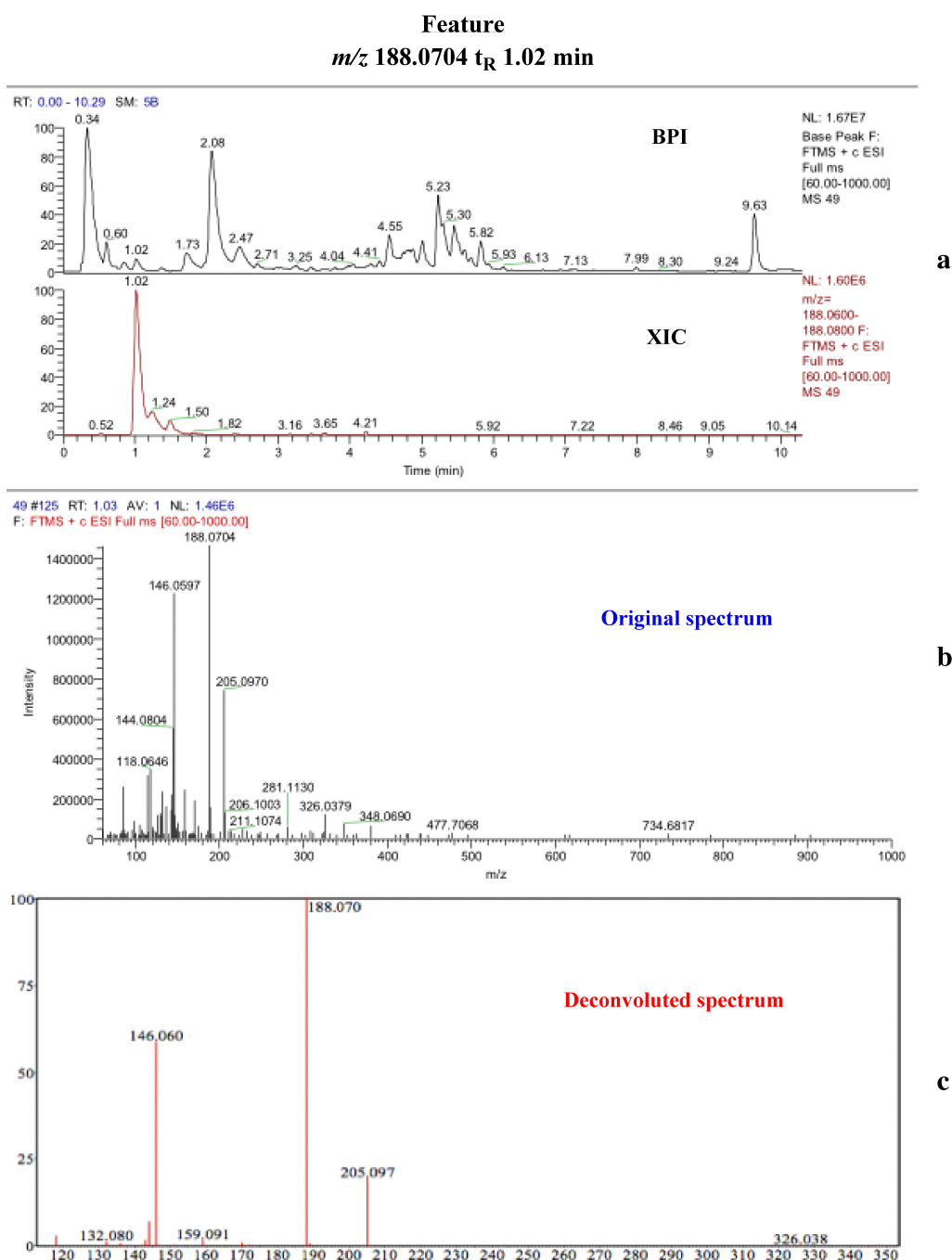


Figure 26. **a.** BPI chromatogram of a representative O sample and XIC of the selected feature m/z 188.0704 at 1.02 min. **b.** Original spectrum of the peak at 1.02 min. **c.** Deconvoluted spectrum of the selected peak at 1.02 min after the MassFrontier approach. Three ions under the selected peak exhibited the same covariance vs. time; the 205.097, 188.070 and 146.060.

The “carnitine story”. Another example is the m/z 302.2325 feature at t_R =5.23 min. It has also been selected as one of the first loadings in X-variate 1 from the urine (+) ESI dataset. Following the previous described methodology, a mass spectrum

model has been created, which has exhibited a major pseudo-fragment at m/z 85.0278. According to the described criteria and the database search, the corresponding substance has been identified as 2,6 dimethylheptanoyl carnitine. Applying an XIC filter based on that mass, a series of peaks became eminent, all of them encompassing the aforementioned peak, but presenting different pseudo-molecular ions. The database searching showed that they all belong to the acylcarnitine class with different fatty acids esterified. Thus 5 carnitines have been identified as up- or down- regulated in response to the OE administration. Furthermore in order to identify more carnitine related molecules but also to differentiate between them, the multiple mass defect filter technique has been employed using the MetWorksTM 1.3 software (Thermo Fisher Scientific, Inc., San Jose, CA, USA). Five new filters have been constructed with one, two, three, four and five methylene moieties and their combination with one and two methyl groups and with the introduction of one and two unsaturations. E.g. the first group has $1CH_2$, $1CH_2+1CH_3$, $1CH_2+2CH_3$, $1CH_2-2H$, and $1CH_2-4H$. Thus a series of carnitines have been identified and their deconvoluted spectra have been produced.

Furthermore, other metabolites were recognized as dimers ($2M+H$), i.e. M13, M23 and M24, or even as non-covalent complexes i.e. M21. Finally, after evaluation of the candidates using the above-described procedure, a few metabolites have not been possible to get recognized and probably being unknown unknowns⁶³, i.e. M21, M25, M27 and M30-M33. In a final step, in order to verify whether the metabolites (Table 7) were statistically up or down regulated in the examined samples, a t-test has been performed. The M1, M5 and M19 metabolites were found as statistically significant metabolites (at the 95% level). ANOVA results of the aforementioned significant metabolites can be found in Table S1, S2 and S3. All the rest metabolites were found

to be very close to significance levels. In column “trend”, the up/down regulation ratio can be found. All datasets have been processed with the proposed methodology in terms of metabolite identification and the results can be found in Table 7.

A comparative study of the significant metabolites in the two different sample groups (P vs. O) is depicted in Figure 27. It should be noted, that the heights of the bars correspond to the area of the corresponding metabolites found in the different groups, whereas the error bars signify the uncertainty of the measurements in terms of the mean of the standard error. Notably, all the significantly differentially secreted metabolites between the two groups were higher in the O group compared to the P group. A possible explanation could be either due to the OE administration.

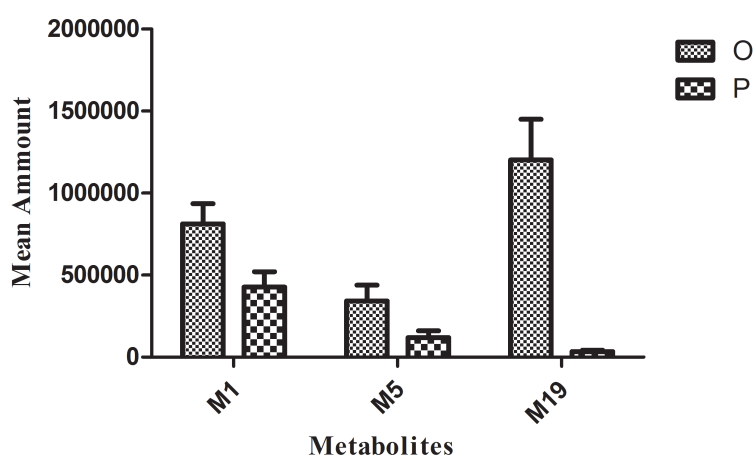


Figure 27: Mean amount of the significant metabolites in the two different groups O and P. The heights of the bars correspond to the area of the corresponding metabolites found in the different groups. The error bars signify the uncertainty of the measurements in terms of mean of the standard error.

Metabolic Pathway Analysis

After evaluation of the loadings and identification of the putative metabolites (Table 7), effort has been spend to explore possible biological pathways in which they

are involved. Thus, the metabolites expressed as a list of compound names have been uploaded to the Metaboanalyst online software for pathway analysis. After auto-correction to standardize compound names, selection of the *Homo sapiens* (human) pathway library and with the aid of Hypergeometric test for the over representation analysis, a set of results has been generated. Thus, respecting the double criterion $p < 0.5$ or impact value > 0.1 , revealed that the metabolic pathways involved are: tryptophan metabolism, phenylalanine, tyrosine and tryptophan biosynthesis, nitrogen metabolism and alanine, aspartate and glutamate metabolism (Table 8).

	Total	Expected	Hits	Raw p	Holm p	FDR	Impact
Tryptophan metabolism	79	0.69	5	4.52E-04	3.61E-02	3.61E-02	0.15
Phenylalanine, tyrosine and tryptophan biosynthesis	27	0.24	2	2.23E-02	0	8.93E-01	0
Nitrogen metabolism	39	0.34	2	4.43E-02	0	0	0
Phenylalanine metabolism	45	0.39	2	5.73E-02	0	0	0.09
D-Glutamine and D-glutamate metabolism	11	0.10	1	9.21E-02	0	0	0.03
Aminoacyl-tRNA biosynthesis	75	0.65	2	1.38E-01	0	0	0
Riboflavin metabolism	21	0.18	1	1.69E-01	0	0	0
Purine metabolism	92	0.80	2	1.90E-01	0	0	0.01
Alanine, aspartate and glutamate metabolism	24	0.21	1	1.90E-01	0	0	0.21
Steroid hormone biosynthesis	99	0.86	2	2.13E-01	0	0	0
Butanoate metabolism	40	0.35	1	2.98E-01	0	0	0.03
Ascorbate and aldarate metabolism	45	0.39	1	3.28E-01	0	0	0.01

	Total	Expected	Hits	Raw p	Holm p	FDR	Impact
Glycine, serine and threonine metabolism	48	0.42	1	3.46E-01	0	0	0
Fatty acid metabolism	50	0.44	1	3.58E-01	0	0	0
Pyrimidine metabolism	60	0.52	1	4.13E-01	0	0	0
Arginine and proline metabolism	77	0.67	1	4.96E-01	0	0	0

Table 8: Metabolic pathways, obtained by the MPA. Total is the total number of compounds in the pathway; the Hits is the actually matched number from the user uploaded data; the Raw p is the original p value calculated from the enrichment analysis; the FDR is the false discovery rate; the Impact is the pathway impact value calculated from pathway topology analysis. The significantly altered pathways by the treatment according to the double criterion ($p < 0.05$ or $\text{impact} > 0.1$) are in red.

The summary of the pathway analysis is also represented as an illustration in Figure 28, in which the color and size of each circle is based on the p values and pathway impact values, respectively.

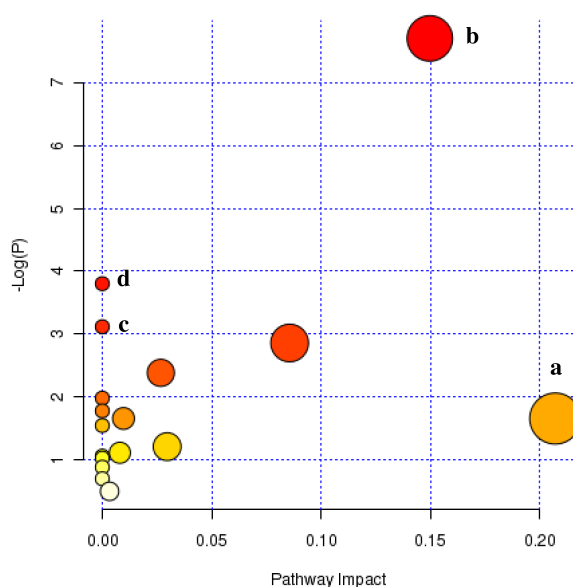


Figure 28: Summary of the pathway analysis for the OE administration to young healthy athletes. All the matched pathways are represented as circles. The color intensity and the radius of the circle

indicate the p value and pathway impact value respectively. **a.** alanine, aspartate and glutamate metabolism, **b.** tryptophan metabolism, **c.** nitrogen metabolism and **d.** phenylalanine, tyrosine and tryptophan biosynthesis.

Biological evaluation

After the identification of the differentially expressed metabolites, it is evident that OE administration increases the circulating acylcarnitines in serum. This denotes the elevated fatty acids catabolism, as it is well known that acylcarnitines mediate their transfer in the mitochondria. The elevated levels of fatty acids in plasma result to their enhanced binding on the serum albumin and thus the release of tryptophan from the protein. This could explain the finding of up-regulated tryptophan metabolism found in O samples (Figure 29), where five metabolites out of thirty-five in total (M1-M5) contributing in this metabolic pathway could be identified.

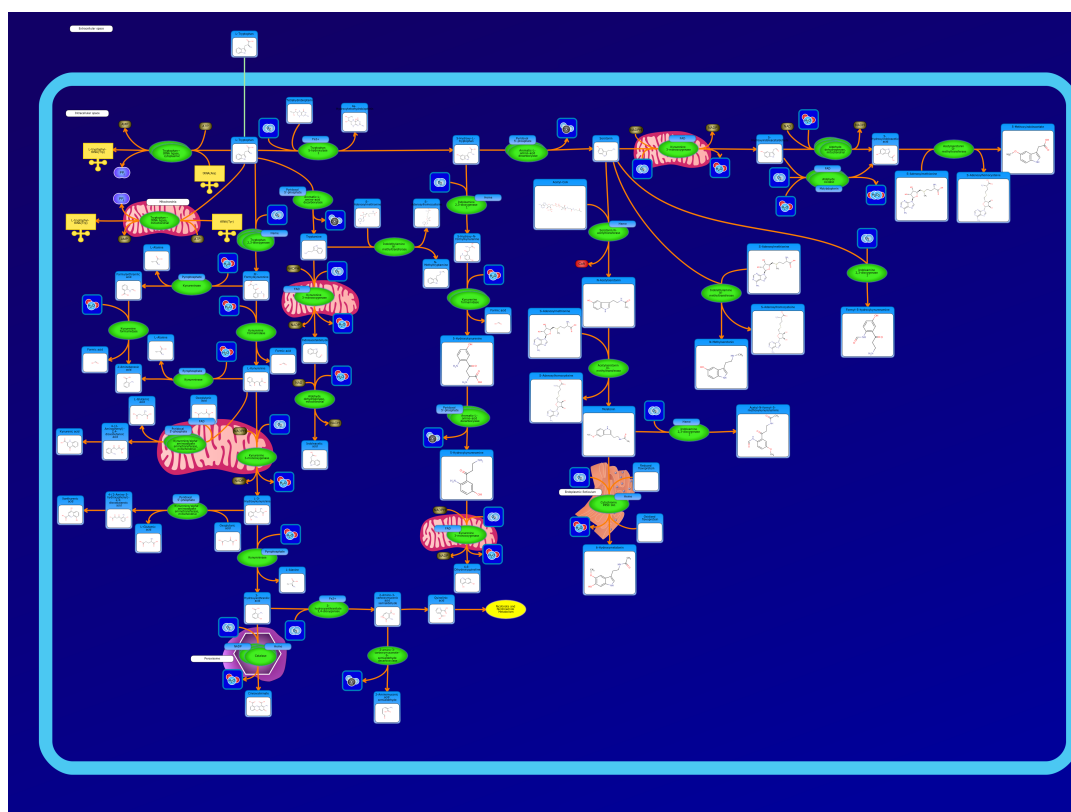


Figure 29. Tryptophan metabolism.

Such an effect has been observed in trained rats, where the nutritional supplementation with fatty acids⁶⁴, results in increased concentration of the so-called “free tryptophan” portion. This could in turn lead to elevated levels of the amino acid (tryptophan) in the brain, given a fact that the blood brain barrier is easily permeable for tryptophan. As tryptophan is an immediate precursor of serotonin, it can act positively on endurance and sensation of effort⁶⁵. Thus, it has been observed that the total exercise time was increased 49.4% after receiving L-tryptophan than after receiving the placebo in healthy sportsmen, as described in the literature⁶⁶. The central fatigue phenomenon that has been correlated with elevated L-tryptophan levels is difficult to be explained and it could possibly be attributed to the enhanced training period that is induced by the activity of the amino acid. Overall, consumption of OE-based herb medicinal products could potentially lead to boost training capability.

CONCLUSION

In the present study, an integrated mass spectrometry-based metabolomic approach consists from two steps; an ML and a MVA approach, has been utilized in order to analyze serum and urine samples from a double-blind crossover design, encompassing the administration an olive leaf extract enriched in OE) to young healthy males. In crossover studies, it is evident that when the total variance of metabolomics data is taken into account during the MVA, the power and the interpretability of the results greatly improves. ML approach provides information about the diversity and abundance of the treatment effects across the study population, which is usually

small. It allows investigation of the between subject variation which is completely separated from the within subject variation, taking into account also other types of variation, such as variation caused by the week of administration or from the day of the administration.

Different sPLS-DA models were created and applied to the datasets. The ML-sPLS-DA approach selects features separating the groups of subjects on a single data set, demonstrated the good performance of the model. After data processing and careful validation of the results, significant and clear clustering between P and O group in ML-sPLS-DA models has been found.

Furthermore, it is evident that MPA of the serum and urine metabolites found in the samples from the crossover study is highly important for assessing inherent network properties in reconstructed biochemical reaction networks. In the present study, metabolomic analysis with the aid of MPA, regarding OE's administration, was inferred from changes in the individuals during the crossover study. Thus, 38 differentially expressed metabolites were identified revealing pathway-specific expression profiles.

Overall, supplementation with an olive leaf extract enriched in OE has been found to alter the serum and urine metabolomes of the athletes compared to the placebo administration, demonstrating that OE could have some impact in the training. Importantly, the findings support that OE administration altered an array of factors concerning crucial biochemical pathways that are implicated in the physical activity. Thus, the metabolism of tryptophan was found to be up-regulated after OE supplementation and increases the circulating acylcarnitines in serum and urine. However, studies with the participation of a larger number of subjects and also targeted MS/MS experiments are required.

REFERENCES

1. Japón-Luján, R. & Luque de Castro, M. D. Superheated liquid extraction of oleuropein and related biophenols from olive leaves. *J. Chromatogr. A* **1136**, 185–91 (2006).
2. Romani, A., Mulinacci, N., Pinelli, P., Vincieri, F. F. & Cimato, A. Polyphenolic content in five tuscan cultivars of *Olea europaea* L. *J. Agric. Food Chem.* **47**, 964–7 (1999).
3. Renaud, S. *et al.* Cretan Mediterranean diet for prevention of coronary heart disease. *Am. J. Clin. Nutr.* **61**, 1360S–1367S (1995).
4. Galli, C. & Visioli, F. Antioxidant and other activities of phenolics in olives/olive oil, typical components of the Mediterranean diet. *Lipids* **34 Suppl**, S23–6 (1999).
5. Kontogianni, V. G. & Gerothanassis, I. P. Phenolic compounds and antioxidant activity of olive leaf extracts. *Nat. Prod. Res.* **26**, 186–9 (2012).
6. Pereira, A. P. *et al.* Phenolic compounds and antimicrobial activity of olive (*Olea europaea* L. Cv. Cobrançosa) leaves. *Molecules* **12**, 1153–62 (2007).
7. Omar, S. H. Cardioprotective and neuroprotective roles of oleuropein in olive. *Saudi Pharm. J. SPJ Off. Publ. Saudi Pharm. Soc.* **18**, 111–121 (2010).
8. Keys, A. *et al.* The diet and 15-year death rate in the seven countries study. *Am. J. Epidemiol.* **124**, 903–15 (1986).
9. Andreadou, I. *et al.* The olive constituent oleuropein exhibits anti-ischemic, antioxidative, and hypolipidemic effects in anesthetized rabbits. *J. Nutr.* **136**, 2213–9 (2006).
10. Visioli, F. & Galli, C. Antiatherogenic components of olive oil. *Curr. Atheroscler. Rep.* **3**, 64–67 (2001).
11. Scoditti, E. *et al.* Mediterranean diet polyphenols reduce inflammatory angiogenesis through MMP-9 and COX-2 inhibition in human vascular endothelial cells: a potentially protective mechanism in atherosclerotic vascular disease and cancer. *Arch. Biochem. Biophys.* **527**, 81–9 (2012).
12. Poudyal, H., Campbell, F. & Brown, L. Olive leaf extract attenuates cardiac, hepatic, and metabolic changes in high carbohydrate-, high fat-fed rats. *J. Nutr.* **140**, 946–53 (2010).
13. Jemai, H., El Feki, A. & Sayadi, S. Antidiabetic and antioxidant effects of hydroxytyrosol and oleuropein from olive leaves in alloxan-diabetic rats. *J. Agric. Food Chem.* **57**, 8798–804 (2009).
14. Impellizzeri, D. *et al.* The effects of oleuropein aglycone, an olive oil compound, in a mouse model of carrageenan-induced pleurisy. *Clin. Nutr.* **30**, 533–40 (2011).

15. Impellizzeri, D. *et al.* The effects of a polyphenol present in olive oil, oleuropein aglycone, in an experimental model of spinal cord injury in mice. *Biochem. Pharmacol.* **83**, 1413–26 (2012).
16. AGOSTONI, C. V. Scientific Opinion on the substantiation of health claims related to polyphenols in olive and protection of LDL particles from oxidative damage (ID 1333, 1638, 1639, 1696, 2865), maintenance of normal blood HDL-cholesterol concentrations (ID 1639), mainte. *EFSA J.* **9**, 2033.1–2033.25 (2012).
17. Assessment Report on *Olea europaea* L, folium. Based on Article 16d(1), Article 16f and Article 16h of Directive 2001/83/EC as amended (traditional use). Assessment report on *Olea europaea* L., folium, EMA/HMPC/430506/2009. (2011). at <http://www.ema.europa.eu/docs/en_GB/document_library/Herbal_-HMPC_assessment_report/2012/04/WC500125459.pdf>
18. Cazzola, R., Russo-Volpe, S., Cervato, G. & Cestaro, B. Biochemical assessments of oxidative stress, erythrocyte membrane fluidity and antioxidant status in professional soccer players and sedentary controls. *Eur. J. Clin. Invest.* **33**, 924–30 (2003).
19. Tasat, D. R. *et al.* Radiation effects on oxidative metabolism in young and aged rat alveolar macrophages. *Cell. Mol. Biol. (Noisy-le-grand)*. **48**, 529–35 (2002).
20. Radak, Z., Chung, H. Y., Koltai, E., Taylor, A. W. & Goto, S. Exercise, oxidative stress and hormesis. *Ageing Res. Rev.* **7**, 34–42 (2008).
21. Radak, Z., Chung, H. Y. & Goto, S. Systemic adaptation to oxidative challenge induced by regular exercise. *Free Radic. Biol. Med.* **44**, 153–9 (2008).
22. Brites, F. D. *et al.* Soccer players under regular training show oxidative stress but an improved plasma antioxidant status. *Clin. Sci. (Lond)*. **96**, 381–5 (1999).
23. Kłapcińska, B. *et al.* Evaluation of autoantibodies against oxidized LDL (oLAB) and blood antioxidant status in professional soccer players. *Int. J. Sports Med.* **26**, 71–8
24. Teixeira, V. H., Valente, H. F., Casal, S. I., Marques, F. P. & Moreira, P. A. Blood antioxidant and oxidative stress biomarkers acute responses to a 1000-m kayak sprint in elite male kayakers. *J. Sports Med. Phys. Fitness* **53**, 71–9 (2013).
25. Seifi-Skishahr, F., Siahkohian, M. & Nakhostin-Roohi, B. Influence of aerobic exercise at high and moderate intensities on lipid peroxidation in untrained men. *J. Sports Med. Phys. Fitness* **48**, 515–21 (2008).
26. Geneva. World Health Organization:Global recommendations on physical activity for health. ISBN 9789241599979 (2010). at <<http://www.who.int/dietphysicalactivity/publications/9789241599979/en/>>
27. Gomez-Cabrera, M.-C., Domenech, E. & Viña, J. Moderate exercise is an antioxidant: upregulation of antioxidant genes by training. *Free Radic. Biol. Med.* **44**, 126–31 (2008).
28. Gomez-Cabrera, M.-C. *et al.* Oral administration of vitamin C decreases muscle mitochondrial biogenesis and hampers training-induced adaptations in endurance performance. *Am. J. Clin. Nutr.* **87**, 142–9 (2008).

29. Margaritis, I., Palazzetti, S., Rousseau, A.-S., Richard, M.-J. & Favier, A. Antioxidant supplementation and tapering exercise improve exercise-induced antioxidant response. *J. Am. Coll. Nutr.* **22**, 147–56 (2003).
30. Rokitzki, L., Hinkel, S., Klemp, C., Cufi, D. & Keul, J. Dietary, serum and urine ascorbic acid status in male athletes. *Int. J. Sports Med.* **15**, 435–40 (1994).
31. Clarkson, P. M. & Thompson, H. S. Antioxidants: what role do they play in physical activity and health? *Am. J. Clin. Nutr.* **72**, 637S–46S (2000).
32. De Bock, M. *et al.* Olive (*Olea europaea* L.) leaf polyphenols improve insulin sensitivity in middle-aged overweight men: a randomized, placebo-controlled, crossover trial. *PLoS One* **8**, e57622 (2013).
33. Oliveras-López, M.-J. *et al.* Extra virgin olive oil (EVOO) consumption and antioxidant status in healthy institutionalized elderly humans. *Arch. Gerontol. Geriatr.* **57**, 234–42
34. Ruano, J. *et al.* Intake of phenol-rich virgin olive oil improves the postprandial prothrombotic profile in hypercholesterolemic patients. *Am. J. Clin. Nutr.* **86**, 341–6 (2007).
35. Salvini, S. *et al.* Daily consumption of a high-phenol extra-virgin olive oil reduces oxidative DNA damage in postmenopausal women. *Br. J. Nutr.* **95**, 742–51 (2006).
36. Franconi, F. *et al.* Antioxidant effect of two virgin olive oils depends on the concentration and composition of minor polar compounds. *J. Agric. Food Chem.* **54**, 3121–5 (2006).
37. Jacomelli, M. *et al.* Dietary extra-virgin olive oil rich in phenolic antioxidants and the aging process: long-term effects in the rat. *J. Nutr. Biochem.* **21**, 290–6 (2010).
38. Bertram, H. C. *et al.* NMR-based metabonomic studies reveal changes in the biochemical profile of plasma and urine from pigs fed high-fibre rye bread. *Br. J. Nutr.* **95**, 955–62 (2006).
39. Pohjanen, E. *et al.* A multivariate screening strategy for investigating metabolic effects of strenuous physical exercise in human serum. *J. Proteome Res.* **6**, 2113–20 (2007).
40. Van Velzen, E. J. J. *et al.* Multilevel data analysis of a crossover designed human nutritional intervention study. *J. Proteome Res.* **7**, 4483–91 (2008).
41. Jansen, T. *et al.* Long-term retrospective analysis of mackerel spawning in the North Sea: a new time series and modeling approach to CPR data. *PLoS One* **7**, e38758 (2012).
42. Rantalainen, M. *et al.* Piecewise multivariate modelling of sequential metabolic profiling data. *BMC Bioinformatics* **9**, 105 (2008).
43. Gibney, M. J. *et al.* Metabolomics in human nutrition: opportunities and challenges. *Am. J. Clin. Nutr.* **82**, 497–503 (2005).

44. Solanky, K. S. *et al.* Biofluid ¹H NMR-based metabonomic techniques in nutrition research - metabolic effects of dietary isoflavones in humans. *J. Nutr. Biochem.* **16**, 236–44 (2005).
45. Rezzi, S., Ramadan, Z., Fay, L. B. & Kochhar, S. Nutritional metabonomics: applications and perspectives. *J. Proteome Res.* **6**, 513–25 (2007).
46. Van der Greef, J. *et al.* The art and practice of systems biology in medicine: mapping patterns of relationships. *J. Proteome Res.* **6**, 1540–59 (2007).
47. Westerhuis, J. a, van Velzen, E. J. J., Hoefsloot, H. C. J. & Smilde, A. K. Multivariate paired data analysis: multilevel PLSDA versus OPLSDA. *Metabolomics* **6**, 119–128 (2010).
48. Lenz, E. M. *et al.* Metabonomics, dietary influences and cultural differences: a ¹H NMR-based study of urine samples obtained from healthy British and Swedish subjects. *J. Pharm. Biomed. Anal.* **36**, 841–9 (2004).
49. Dunn, W. B. *et al.* Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* **6**, 1060–83 (2011).
50. Want, E. J. *et al.* Global metabolic profiling procedures for urine using UPLC-MS. *Nat. Protoc.* **5**, 1005–18 (2010).
51. Pluskal, T., Castillo, S., Villar-Briones, A. & Oresic, M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11**, 395 (2010).
52. Xia, J. & Wishart, D. S. Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nat. Protoc.* **6**, 743–60 (2011).
53. Shen, R., Wang, S. & Mo, Q. SPARSE INTEGRATIVE CLUSTERING OF MULTIPLE OMICS DATA SETS. *Ann. Appl. Stat.* **7**, 269–294 (2013).
54. Zhu, D. *et al.* Exploring high-order functional interactions via structurally-weighted LASSO models. *Inf. Process. Med. Imaging* **23**, 13–24 (2013).
55. Xia, J. & Wishart, D. S. MetPA: a web-based metabolomics tool for pathway analysis and visualization. *Bioinformatics* **26**, 2342–4 (2010).
56. Xia, J. & Wishart, D. S. Metabolomic data processing, analysis, and interpretation using MetaboAnalyst. *Curr. Protoc. Bioinformatics* **Chapter 14**, Unit 14.10 (2011).
57. Xia, J., Mandal, R., Sinelnikov, I. V, Broadhurst, D. & Wishart, D. S. MetaboAnalyst 2.0--a comprehensive server for metabolomic data analysis. *Nucleic Acids Res.* **40**, W127–33 (2012).
58. Stacklies, W., Redestig, H., Scholz, M., Walther, D. & Selbig, J. pcaMethods--a bioconductor package providing PCA methods for incomplete data. *Bioinformatics* **23**, 1164–7 (2007).

59. Dieterle, F., Ross, A., Schlotterbeck, G. & Senn, H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in ¹H NMR metabonomics. *Anal. Chem.* **78**, 4281–90 (2006).
60. Lê Cao, K.-A., Boitard, S. & Besse, P. Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics* **12**, 253 (2011).
61. Ambroise, C. & McLachlan, G. J. Selection bias in gene extraction on the basis of microarray gene-expression data. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 6562–6 (2002).
62. Lique, B., Lê Cao, K.-A., Hocini, H. & Thiébaud, R. A novel approach for biomarker selection and the integration of repeated measures experiments from two assays. *BMC Bioinformatics* **13**, 325 (2012).
63. Romo, T. D. & Grossfield, A. Unknown Unknowns: the Challenge of Systematic and Statistical Error in Molecular Dynamics Simulations. *Biophys. J.* **106**, 1553–1554 (2014).
64. Chaouloff, F. Effects of acute physical exercise on central serotonergic systems. *Med. Sci. Sports Exerc.* **29**, 58–62 (1997).
65. Badawy, A. Novel nutritional treatment for manic and psychotic disorders: a review of tryptophan and tyrosine depletion studies and the potential of protein-based formulations using glycomacropeptide. *Psychopharmacology (Berl)*. **228**, 347–58 (2013).
66. Segura, R. & Ventura, J. L. Effect of L-tryptophan supplementation on exercise performance. *Int. J. Sports Med.* **9**, 301–5 (1988).

SUPPORTING INFORMATION

Analysis of Variance (One-Way) Metabolite M1						
Summary						
<i>Groups</i>	<i>Sample size</i>	<i>Sum</i>	<i>Mean</i>	<i>Variance</i>		
<i>O</i>	18	15.558.222,91	864.345,71722	3,73796E+11		
<i>P</i>	18	7.690.860,74	427.270,04111	1,58347E+11		
ANOVA						
<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p-level</i>	<i>F crit</i>
Between Groups	1,71932E+12	1	1,71932E+12	6,46186	0,01575	4,13002
Within Groups	9,04643E+12	34	2,66071E+11			
<i>Total</i>	1,07657E+13	35				

Table S1. ANOVA results for the significant metabolite M1.

Analysis of Variance (One-Way) Metabolite M5						
Summary						
<i>Groups</i>	<i>Sample size</i>	<i>Sum</i>	<i>Mean</i>	<i>Variance</i>		
<i>O</i>	18	6.153.587,153	341.865,95294	1,68643E+11		
<i>P</i>	18	2.147.063,846	119.281,32478	3,20009E+10		
ANOVA						
<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p-level</i>	<i>F crit</i>
Between Groups	4,45895E+11	1	4,45895E+11	4,44463	0,04246	4,13002
Within Groups	3,41095E+12	34	1,00322E+11			
<i>Total</i>	3,85685E+12	35				

Table S1. ANOVA results for the significant metabolite M5.

Analysis of Variance (One-Way) Metabolite M19						
Summary						
<i>Groups</i>	<i>Sample size</i>	<i>Sum</i>	<i>Mean</i>	<i>Variance</i>		
<i>O</i>	18	33.925.758,693	1.884.764,37183	1,02931E+13		
<i>P</i>	18	623.969,718	34.664,98433	1.016.438.443,15771		
ANOVA						
<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p-level</i>	<i>F crit</i>
Between Groups	3,08058E+13	1	3,08058E+13	5,98514	0,01975	4,13002
Within Groups	1,75E+14	34	5,14705E+12			
<i>Total</i>	2,05805E+14	35				

Table S3. ANOVA results for the significant metabolite M19.

Supplementary Method 1.

Find overlapping

`library(xMSanalyzer)`

#Please refer to the manual for parameter description:

#<http://userwww.service.emory.edu/~kuppall2/xMSanalyzer/xMSanalyzer-manual.pdf>

#####Parameters to
change#####

#1) Full path of file A, first column should be the list of m/zs, second column should

#the retention times

```
input_file_A<="/Users/nicklem26/Desktop/Algorithm 1.txt"
```

#2) Full path of file B, first column should be the list of m/zs, second column should

#the retention times

```
input_file_B<="/Users/nicklem26/Desktop/Algorithm 2.txt"
```

#3) Name of dataset A

```
name_a="A"
```

#4) Name of dataset B

```
name_b="B"
```

#5) +/- ppm threshold for matching m/zs

```
mz.thresh=0
```

#6) +/- secs threshold for comparing retention times

```
time.thresh=5
```

#7) Output location for the Venn diagram

```
outloc="/Users/nicklem26/Desktop"
```



```
#####
```

```
dataA<-read.table(input_file_A,sep="\t",header=TRUE)
```

```
dataB<-read.table(input_file_B,sep="\t",header=TRUE)
```

```
overlapping<- find.Overlapping.mzs(dataA, dataB, mz.thresh  
= 0, time.thresh = 5,
```

```
alignment.tool=NA)
```

```
print(overlapping)
```

Get overlapping and unique features – Venn diagrams

```
library(xMSanalyzer)
```

```
#Please refer to the manual for parameter description:
```

```
#http://userwww.service.emory.edu/~kuppall2/xMSanalyzer/xM  
Sanalyzer-manual.pdf
```

```
#####Parameters to
```

```
change#####
```

```
#1) Full path of file A, first column should be the list  
of m/zs, second column should
```

```
#the retention times
```

```
input_file_A<-"/Users/nicklem26/Scientific/PhD/My  
Projects/Oleuropein/Oleuropein_Mougios/Metabolomics/Serum
```

```
/Positive/R/xMSanalyzer/Text files to be  
imported/SerumPos_Cutoff_FTMS Filt_gapped_07Nov13.txt"
```

#2) Full path of file B, first column should be the list
of m/zs, second column should

#the retention times

```
input_file_B<-" /Users/nicklem26/Scientific/PhD/My  
Projects/Oleuropein/Oleuropein_Mougios/Metabolomics/Serum  
/Positive/R/xMSanalyzer/Text files to be  
imported/SerumPos_Cutoff_gapped_07Nov13.txt"
```

#3) Name of dataset A

```
name_a="A"
```

#4) Name of dataset B

```
name_b="B"
```

#5) +/- ppm threshold for matching m/zs

```
mz.thresh=5
```

#6) +/- secs threshold for comparing retention times

```
time.thresh=5
```

#7) Output location for the Venn diagram

```
outloc=" /Users/nicklem26/Scientific/PhD/My  
Projects/Oleuropein/Oleuropein_Mougios/Metabolomics/Serum  
/Positive/R/xMSanalyzer/Exported figures/Venn Diagramms"
```

```
#####

dataA<-read.table(input_file_A,sep="\t",header=TRUE)

dataB<-read.table(input_file_B,sep="\t",header=TRUE)

g1<-
  getVenn(dataA=dataA,name_a=name_a,name_b=name_b,dataB=dataB,
    mz.thresh=mz.thresh,time.thresh=time.thresh,xMSanalyzer.outloc=outloc,alignment.tool=NA)

names(g1)

head(g1$common)

head(g1$uniqueA)
```

Load overlapping data in a text

```
library(xMSanalyzer)

#Please refer to the manual for parameter description:
#http://userwww.service.emory.edu/~kuppal2/xMSanalyzer/xMSanalyzer-manual.pdf

#####Parameters to
change#####
```

#1) Full path of file A, first column should be the list of m/zs, second column should

#the retention times

```
input_file_A<-"/Users/nicklem26/Scientific/PhD/My  
Projects/Oleuropein/Oleuropein_Mougios/Metabolomics/Serum  
/Positive/R/xMSanalyzer/Text files to be  
imported/SerumPos_Cutoff_FTMS Filt_gapped_07Nov13.txt"
```

#2) Full path of file B, first column should be the list of m/zs, second column should

#the retention times

```
input_file_B<-"/Users/nicklem26/Scientific/PhD/My  
Projects/Oleuropein/Oleuropein_Mougios/Metabolomics/Serum  
/Positive/R/xMSanalyzer/Text files to be  
imported/SerumPos_Cutoff_gapped_07Nov13.txt"
```

#3) Name of dataset A

```
name_a="A"
```

#4) Name of dataset B

```
name_b="B"
```

```
dataA<-read.table(input_file_A,sep="\t",header=TRUE)
```

```
dataB<-read.table(input_file_B,sep="\t",header=TRUE)
```

Supplementary Method 2.**sPCA**

```

data(crossover)

X <- crossover$individuality # Using one data set only

## PCA: data were centered but not scaled

result <- pca(X, ncomp = 3, center = TRUE, scale. =
FALSE);

## sPCA: we are selecting 100 variables on each of the
PCs

result <- spca(X, ncomp = 3, center = TRUE, scale. =
TRUE, keepX = rep(100,3));

```

sPLS-DA

```

data(crossover)

X <- as.matrix(crossover$week)

Y <- as.factor(crossover$treatment[, 4])

## PLS-DA function

result <- plsda(X, Y, ncomp = 3) # where ncomp is the
number of components wanted

## sPLS-DA function

result <- splsda(X, Y, ncomp = 3, keepX = c(100, 100,

```

```
100)) # where keepX is the number of variables selected
for each components
```

```
data(crossover)
```

```
X <- as.matrix(crossover$week)
```

```
Y <- as.factor(crossover$individuality[, 4])
```

```
samp <- sample(1:3, nrow(X), replace = TRUE)
```

```
test <- which(samp == 1) # Search which column in samp
has a value of 1
```

```
train <- setdiff(1:nrow(X), test) # Keeping the column
that are not in test
```

```
## For PLS-DA
```

```
plsda.train <- plsda(X[train, ], Y[train], ncomp = 3)
```

```
test.predict <- predict(plsda.train, X[test, ], method =
"max.dist")
```

```
## For sPLS-DA
```

```
splsda.train <- splsda(X[train, ], Y[train], ncomp = 3,
keepX = c(100, 100, 100))
```

```
test.predict <- predict(splsda.train, X[test, ], method =
"max.dist")
```

Supplementary Method 3.

ML-sPLS-DA

```

library(mixOmics)

EZinfo_UrinePos_WaveletsXCMS_8E4_MPLSDA_KNN_dataprocessed
<- read.csv("~/Scientific/PhD/My
Projects/Oleuropein/Oleuropein_Mougios/Metabolomics/Urine
/Positive/Lists from MZmine for
Simca/EZinfo_UrinePos_WaveletsXCMS_8E4_MPLSDA_KNN_datapro
cessed.csv", sep=";", dec=",")

View(EZinfo_UrinePos_WaveletsXCMS_8E4_MPLSDA_KNN_dataproce
ssed)

urinepos.KNN.dp<-
EZinfo_UrinePos_WaveletsXCMS_8E4_MPLSDA_KNN_dataprocessed

rm(EZinfo_UrinePos_WaveletsXCMS_8E4_MPLSDA_KNN_dataproces
sed)

View(urinepos.KNN.dp)

sample<-urinepos.KNN.dp[,1]

sample

po<-urinepos.KNN.dp[,2]

within.t.nt<-urinepos.KNN.dp[,3]

within.real<-urinepos.KNN.dp[,4]

between<-urinepos.KNN.dp[,5]

week<-urinepos.KNN.dp[,6]

data.urinepos.KNN.dp<-urinepos.KNN.dp[,-(1:6)]

View(data.urinepos.KNN.dp)

result.po.week.1level<- multilevel(data.urinepos.KNN.dp,
cond = po, sample = week, ncomp = 3, keepX = c(50,50,50),
tab.prob.gene = NULL, method = 'splsd')

plotIndiv(result.po.week.1level, col=as.factor(po),
ind.names = FALSE, pch=(1:2)

```

```

plotIndiv(result.po.week.1level, col=as.factor(po),
  ind.names = FALSE, pch=(1:2))

plotIndiv(result.po.week.1level, col=as.factor(po),
  ind.names = FALSE, pch=(1:2))

plotIndiv(result.po.week.1level, col=as.factor(po),
  ind.names = FALSE)

plotIndiv(result.po.week.1level, col=as.factor(po),
  ind.names = FALSE, pch=(1:2))

legend('topright', col = 'black', legend
  =unique(week),pch = 1:2, cex = 0.8)

legend('bottomright', col = unique(op), legend =
  levels(op), pch = 20, cex = 0.8)

legend('bottomright', col = unique(op), legend =
  levels(po), pch = 20, cex = 0.8)

legend('bottomright', col = unique(po), legend =
  levels(po), pch = 20, cex = 0.8)

result.po.between.1level<-
  multilevel(data.urinepos.KNN.dp, cond = po, sample =
  between, ncomp = 3, keepX = c(50,50,50), tab.prob.gene =
  NULL, method = 'spllda')

plotIndiv(result.po.between.1level, col=as.factor(po),
  ind.names = FALSE, pch=(1:2))

plotIndiv(result.po.between.1level, col=as.factor(po),
  ind.names = FALSE, pch=(1:9))

legend('topright', col = 'black', legend
  =unique(between),pch = 1:9, cex = 0.8)

legend('bottomright', col = unique(po), legend =
  levels(op), pch = 20, cex = 0.8)

legend('bottomright', col = unique(po), legend =
  levels(po), pch = 20, cex = 0.8)

```



```

result.po.within.real.1level<-
  multilevel(data.urinepos.KNN.dp, cond = po, sample =
    within.real, ncomp = 3, keepX = c(50,50,50),
    tab.prob.gene = NULL, method = 'spllda')

plotIndiv(result.po.within.real.1level,
  col=as.factor(po), ind.names = FALSE, pch=(1:5))

plotIndiv(result.po.within.real.1level,
  col=as.factor(po), ind.names = within.real, pch=(1:5))

plotIndiv(result.po.within.real.1level,
  col=as.factor(po), ind.names = FALSE, pch=(1:5))

legend('bottomright', col = unique(po), legend =
  levels(po), pch = 20, cex = 0.8)

legend('topright', col = 'black', legend
  =unique(within.real),pch = 1:10, cex = 0.8)

plotIndiv(result.po.within.real.1level,
  col=as.factor(po), ind.names = FALSE, pch=(1:10))

legend('topright', col = 'black', legend
  =unique(within.real),pch = 1:10, cex = 0.8)

fix(within.real)

EZinfo_UrinePos_WaveletsXCMS_8E4_MPLSDA_KNN_dataprocessed

result.po.within.t.nt.1level<-
  multilevel(data.urinepos.KNN.dp, cond = po, sample =
    within.t.nt, ncomp = 3, keepX = c(50,50,50),
    tab.prob.gene = NULL, method = 'spllda')

plotIndiv(result.po.within.t.nt.1level,
  col=as.factor(po), ind.names = FALSE, pch=(1:10))

result.week.po.1level<- multilevel(data.urinepos.KNN.dp,
  cond = week, sample = po, ncomp = 3, keepX = c(50,50,50),
  tab.prob.gene = NULL, method = 'spllda')

plotIndiv(result.po.week.1level, col=as.factor(po),

```

```

ind.names = FALSE, pch=(1:2))

plotIndiv(result.week.po.1level, col=as.factor(po),
  ind.names = FALSE, pch=(1:2))

plotIndiv(result.week.po.1level, col=as.factor(week),
  ind.names = FALSE, pch=(1:2))

legend('topright', col = 'black', legend
  =unique(week),pch = 1:2, cex = 0.8)

legend('bottomright', col = unique(week), legend =
  levels(week), pch = 20, cex = 0.8)

legend('bottomright', col = unique(week), legend =
  levels(week), pch = 20, cex = 0.8)

legend('topright', col = 'black', legend =unique(po),pch
  = 1:2, cex = 0.8)

legend('bottomright', col = unique(week), legend =
  levels(week), pch = 20, cex = 0.8)

result.between.po.1level<-
  multilevel(data.urinepos.KNN.dp, cond = between, sample =
  po, ncomp = 3, keepX = c(50,50,50), tab.prob.gene = NULL,
  method = 'splstda')

plotIndiv(result.between.po.1level,
  col=as.factor(between), ind.names = FALSE, pch=(1:2))

plotIndiv(result.between.po.1level,
  col=as.factor(between), ind.names = po, pch=(1:2))

plotIndiv(result.between.po.1level,
  col=as.factor(between), ind.names = TRUE, pch=(1:2))

plotIndiv(result.between.po.1level,
  col=as.factor(between), ind.names = FALSE, pch=(1:9))

plotIndiv(result.between.po.1level, col=as.factor(P0),
  ind.names = FALSE, pch=(1:9))

plotIndiv(result.between.po.1level, col=as.factor(po),

```

```

ind.names = FALSE, pch=(1:9))

plotIndiv(result.between.po.1level,
  col=as.factor(between), ind.names = po, pch=(1:2))

plotIndiv(result.between.po.1level,
  col=as.factor(between), ind.names = FALSE, pch=(1:2))

result.within.t.nt.po.1level<-
  multilevel(data.urinepos.KNN.dp, cond = within.t.nt,
    sample = po, ncomp = 3, keepX = c(50,50,50),
    tab.prob.gene = NULL, method = 'spllda')

plotIndiv(result.within.t.nt.po.1level,
  col=as.factor(within.t.nt), ind.names = FALSE, pch=(1:2))

po.week <- data.frame(cbind(as.character(po),
  as.character(week)))

po.week

result.po.week.within.real.2level<-
  multilevel(data.urinepos.KNN.dp, cond = po.week, sample =
    within.t.nt, ncomp = 3, keepX = c(50,50,50),
    tab.prob.gene = NULL, method = 'spllda')

plotIndiv(result.po.week.within.real.2level,
  col=as.factor(po),pch=week, ind.names = FALSE)

result.po.week.within.t.nt<-
  multilevel(data.urinepos.KNN.dp, cond = po.week, sample =
    within.t.nt, ncomp = 3, keepX = c(50,50,50),
    tab.prob.gene = NULL, method = 'spllda')

plotIndiv(result.po.week.within.t.nt.2level,
  col=as.factor(po),pch=week, ind.names = FALSE)

result.po.week.within.t.nt.2level<-
  multilevel(data.urinepos.KNN.dp, cond = po.week, sample =
    within.t.nt, ncomp = 3, keepX = c(50,50,50),
    tab.prob.gene = NULL, method = 'spllda')

plotIndiv(result.po.week.within.t.nt.2level,

```

```

col=as.factor(po),pch=week, ind.names = FALSE)

result.po.week.within.real.2level<-
multilevel(data.urinepos.KNN.dp, cond = po.week, sample =
within.real, ncomp = 3, keepX = c(50,50,50),
tab.prob.gene = NULL, method = 'spllda')

plotIndiv(result.po.week.within.real.2level,
col=as.factor(po),pch=week, ind.names = FALSE)

result.po.week.between.2level<-
multilevel(data.urinepos.KNN.dp, cond = po.week, sample =
between, ncomp = 3, keepX = c(50,50,50), tab.prob.gene =
NULL, method = 'spllda')

plotIndiv(result.po.week.between.2level,
col=as.factor(po),pch=week, ind.names = FALSE)

plotIndiv(result.po.week.between.2level,
col=as.factor(between),pch=week, ind.names = between)

plotIndiv(result.po.week.between.2level,
col=as.numeric(between),pch=week, ind.names = between)

plotIndiv(result.po.week.between.2level,
col=as.numeric(between),pch=week, ind.names = po)

plotIndiv(result.po.week.between.2level,
col=as.numeric(po),pch=week, ind.names = between)

po.within.t.nt <- data.frame(cbind(as.character(po),
as.character(within.t.nt)))

po.within.t.nt

result.po.within.t.nt.within.real.2level<-
multilevel(data.urinepos.KNN.dp, cond = po.within.t.nt,
sample = within.real, ncomp = 3, keepX = c(50,50,50),
tab.prob.gene = NULL, method = 'spllda')

plotIndiv(result.po.within.t.nt.within.real.2level,
col=as.numeric(po),pch=within.t.nt, ind.names = po)

```

```

plotIndiv(result.po.within.t.nt.within.real.2level,
  col=as.numeric(po),pch=within.t.nt, ind.names =
  within.t.nt)

plotIndiv(result.po.within.t.nt.within.real.2level,
  col=as.numeric(po),pch=within.t.nt, ind.names =
  within.real)

result.po.within.t.nt.week.2level<-
  multilevel(data.urinepos.KNN.dp, cond = po.within.t.nt,
  sample = week, ncomp = 3, keepX = c(50,50,50),
  tab.prob.gene = NULL, method = 'splsda')

plotIndiv(result.po.within.t.nt.week.2level,
  col=as.numeric(po),pch=within.t.nt, ind.names = week)

plotIndiv(result.po.within.t.nt.week.2level,
  col=as.numeric(po),pch=within.t.nt, ind.names = po)

plotIndiv(result.po.within.t.nt.week.2level,
  col=as.numeric(po),pch=within.t.nt, ind.names =
  within.t.nt)

po.between <- data.frame(cbind(as.character(po),
  as.character(between)))

result.po.between.within.t.nt.2level<-
  multilevel(data.urinepos.KNN.dp, cond = po.between,
  sample = within.t.nt, ncomp = 3, keepX = c(50,50,50),
  tab.prob.gene = NULL, method = 'splsda')

plotIndiv(result.po.between.within.t.nt.2level,
  col=as.numeric(po),pch=within.t.nt, ind.names =
  within.t.nt)

plotIndiv(result.po.between.within.t.nt.2level,
  col=as.numeric(po),pch=within.t.nt, ind.names = between)

plotIndiv(result.po.between.within.t.nt.2level,
  col=as.numeric(po),pch=po, ind.names = between)

library(mixOmics)

```

```

plotIndiv(result.po.between.within.t.nt.2level,
  col=as.numeric(po),pch=po, ind.names = between)

plotIndiv(result.po.week.1level, col=as.factor(po),
  ind.names = FALSE, pch=(1:2))

plotIndiv(result.po.week.1level, col=as.factor(po),
  ind.names = FALSE, pch=(1:2))

plotIndiv(result.po.week.1level, col=as.factor(po),
  ind.names = FALSE, pch=(1:2))

library(mixOmics)

library(mixOmics)

library(mixOmics)

fix(result.po.week.within.real.2level)

result.po.between.1level<-
  multilevel(data.urinepos.KNN.dp, cond = po, sample =
  between, ncomp = 3, keepX = c(50,50,50), tab.prob.gene =
  NULL, method = 'splsda')

plotIndiv(result.po.between.1level, col=as.factor(po),
  ind.names = FALSE)

result.po.week.within.real.2level<-
  multilevel(data.urinepos.KNN.dp, cond = po.week, sample =
  within.real, ncomp = 3, keepX = c(50,50,50),
  tab.prob.gene = NULL, method = 'splsda')

plotIndiv(result.po.week.within.real.2level,
  col=as.factor(po), ind.names = FALSE)

plotIndiv(result.po.week.within.real.2level,
  col=as.factor(po), pch=po, ind.names = FALSE)

plotIndiv(result.po.week.within.real.2level,
  col=as.factor(po), ind.names = FALSE)

result.po.within.real.1level<-
  multilevel(data.urinepos.KNN.dp, cond = po, sample =

```

```

within.real, ncomp = 3, keepX = c(50,50,50),
tab.prob.gene = NULL, method = 'spllda')

plotIndiv(result.po.within.real.1level,
col=as.factor(po), ind.names = FALSE)

result.po.week.1level<- multilevel(data.urinepos.KNN.dp,
cond = po, sample =week, ncomp = 3, keepX = c(50,50,50),
tab.prob.gene = NULL, method = 'spllda')

plotIndiv(result.po.week.1level, col=as.factor(po),
ind.names = FALSE)

library(mixOmics)

result.po.between.1level<-
multilevel(data.urinepos.KNN.dp, cond = po, sample =
between, ncomp = 3, keepX = c(50,50,50), tab.prob.gene =
NULL, method = 'spllda')

plotIndiv(result.po.between.1level, col=as.factor(po),
ind.names = FALSE)

legend('topbottom', col = unique(po), legend =
levels(po), pch = 20, cex =0.8)


result.tune.po.between <-
tune.multilevel(data.urinepos.KNN.dp,cond = po,sample =
between,ncomp=3,test.keepX=c(60, 60, 60),already.tested.X
= c(80,80),method = 'spllda',dist = 'mahalanobis.dist')

result.tune.po.between

result.tune.po.between <-
tune.multilevel(data.urinepos.KNN.dp,cond = po,sample =
between,ncomp=2,test.keepX=c(50, 50, 50),already.tested.X
= c(90),method = 'spllda',dist = 'mahalanobis.dist')

result.tune.po.between

result.tune.po.between <-

```

```

tune.multilevel(data.urinepos.KNN.dp,cond = po,sample =
between,ncomp=2,test.keepX=c(100, 100,
100),already.tested.X = c(150),method = 'spllda',dist =
'mahalanobis.dist')

result.tune.po.between

result.between.po.1level<-
multilevel(data.urinepos.KNN.dp, cond = between, sample =
po,comp = 3, keepX = c(50,50,50), tab.prob.gene = NULL,
method = 'spllda')

result.between.po.1level<-
multilevel(data.urinepos.KNN.dp, cond = between, sample =
po,comp = 3, keepX = c(50,50,50), tab.prob.gene = NULL,
method = 'spllda')

plotIndiv(result.between.po.1level, col=as.numeric(po),
ind.names =FALSE, pch=20)

plotIndiv(result.between.po.1level, col=as.numeric(po),
ind.names =between, pch=20)

plotIndiv(result.between.po.1level,
col=as.numeric(between), ind.names =between, pch=20)

plotIndiv(result.between.po.1level,
col=as.factor(between), ind.names =between, pch=20)

plotIndiv(result.between.po.1level,
col=as.numeric(between), ind.names =po, pch=20)

plotIndiv(result.between.po.1level,
col=as.numeric(between), ind.names =FALSE, pch=20)

result.between.week.1level<-
multilevel(data.urinepos.KNN.dp, cond = between, sample =
week, comp = 3, keepX = c(50,50,50), tab.prob.gene =
NULL, method = 'spllda')

plotIndiv(result.between.week.1level,
col=as.numeric(between), ind.names =FALSE, pch=20)

```



```
result.between.week.1level<-
multilevel(data.urinepos.KNN.dp, cond = between, sample =
week, comp = 3, keepX = c(100,100,100), tab.prob.gene =
NULL, method = 'splsa')
```

```
result.between.week.1level<-
multilevel(data.urinepos.KNN.dp, cond = between, sample =
week, comp = 2, keepX = c(100,100), tab.prob.gene = NULL,
method = 'splsa')
```

```
pheatmap.multilevel(result.po.individuality.1level,col_sa
mple=sample,col_stimulation=individuality,label_annotatio
n=NULL,border=FALSE,clustering_method="ward",show_colname
s = FALSE,show_rownames = TRUE,fontsize_row=2)
```

```
pheatmap.multilevel(result.po.individuality.1level,col_sa
mple=po,col_stimulation=individuality,label_annotation=NU
LL,border=FALSE,clustering_method="ward",show_colnames =
FALSE,show_rownames = TRUE,fontsize_row=2)
```

```
pheatmap.multilevel(result.po.individuality.1level,col_sa
mple=sample,col_stimulation=individuality,label_annotatio
n=NULL,border=FALSE,clustering_method="ward",show_colname
s = FALSE,show_rownames = TRUE,fontsize_row=2)
```

```
pheatmap.multilevel(result.po.individuality.1level,col_sa
mple=sample,col_stimulation=po,label_annotation=NULL,bord
er=FALSE,clustering_method="ward",show_colnames =
FALSE,show_rownames = TRUE,fontsize_row=2)
```

```
po.treatmentdays <- data.frame(cbind(as.character(po),
as.character(treatmentdays)))
```

```
result.po.treatmentdays.individuality.2level<-
multilevel(data.x, cond = po.treatmentdays, sample =
individuality, ncomp = 3, keepX = c(200,200,200),
tab.prob.gene = NULL, method = 'splsa')
```

```
col.po<-as.numeric(po)
```

```

col.po

pch.treatmentdays <- rep(20, 48)

pch.treatmentdays[treatmentdays == 't2'] <- 4

plotIndiv(result.po.treatmentdays.individuality.2level,
  col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

plotIndiv(result.po.treatmentdays.individuality.2level,
  col = col.po, pch = pch.treatmentdays, ind.names =
  sample)

plotIndiv(result.po.treatmentdays.individuality.2level,
  col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

pch.treatmentdays[treatmentdays == 't2'] <- 2

plotIndiv(result.po.treatmentdays.individuality.2level,
  col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

pch.treatmentdays[treatmentdays == 't'] <- 4

plotIndiv(result.po.treatmentdays.individuality.2level,
  col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

pch.treatmentdays[treatmentdays == 't2'] <- 4

plotIndiv(result.po.treatmentdays.individuality.2level,
  col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

legend('bottomright', col = unique(col.po), legend =
  levels(po), pch = 20, cex = 0.8)

legend('topright', col = 'black', legend =
  levels(treatmentdays), pch = unique(pch.treatmentdays),
  cex = 0.8)

legend('topright', col = 'black', legend =
  levels(treatmentdays), pch = unique(pch.treatmentdays),
  cex = 0.8)

pch.treatmentdays <- rep(10, 48)

legend('topright', col = 'black', legend =

```

```

levels(treatmentdays), pch = unique(pch.treatmentdays),
cex = 0.8)

plotIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

pch.treatmentdays <- rep(1, 18)

plotIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

pch.treatmentdays <- rep(10, 18)

plotIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

pch.treatmentdays <- rep(1, 2)

plotIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

pch.treatmentdays <- rep(1:18)

plotIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

pch.treatmentdays <- rep(1:2)

plotIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

pch.treatmentdays <- rep(10,35)

plotIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

pch.treatmentdays <- rep(10,48)

plotIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

pch.treatmentdays <- rep(20,48)

plotIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

```

```
plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE, cex =
  c(0.5,0.9))
```

```
plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE, cex =
  c(0,5,0.99))
```

```
plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE, cex =
  c(0,5,0,9))
```

```
plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE, cex =
  c(1,0.9))
```

```
plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE)
```

```
plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE, cex =
  c(1,0.8))
```

```
plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE,
  var.label=TRUE,cex = c(1,0.8))
```

```
plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE,
  var.label=TRUE)
```

```
plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE,
  keep.var=TRUE,var.label=TRUE)
```

```
plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE,
  keep.var=TRUE,var.label=TRUE,cutoff=0.5)
```

```
plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE,
  keep.var=TRUE,var.label=TRUE,cutoff=0.5, rad.in=0.9)
```

```

plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE,
  keep.var=TRUE,var.label=TRUE,cutoff=0.5, rad.in=0.8)

plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE,
  keep.var=TRUE,var.label=TRUE,cutoff=0.5,
  rad.in=0.8,cex=0.6)

plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE,
  keep.var=TRUE,var.label=TRUE,cutoff=0.5,
  rad.in=0.8,cex=0.6)

biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50)

?biplot

biplot(loadings$result.po.treatmentdays.individuality.2level.keepX50.50.50,
variates$result.po.treatmentdays.individuality.2level.keepX50.50.50)

biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50)

require(graphics)

biplot(princomp(USArrests))

biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50)

biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50)

fix(result.po.treatmentdays.individuality.2level.keepX50.
50.50)

biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50$loadings,result.po.treatmentdays.individuality.2
level.keepX50.50.50$variates)

```

```
biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50$loadings,result.po.treatmentdays.individuality.2
level.keepX50.50.50$variates)
```

```
biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50$loadings,result.po.treatmentdays.individuality.2
level.keepX50.50.50$variates)
```

```
biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50$loadings,result.po.treatmentdays.individuality.2
level.keepX50.50.50$variates,cex=0.6)
```

```
biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50,result.po.treatmentdays.individuality.2level.kee
pX50.50.50$loadings,result.po.treatmentdays.individuality
.2level.keepX50.50.50$variates)
```

```
biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50)
```

```
biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50,X,Y)
```

```
biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50$variates,result.po.treatmentdays.individuality.2
level.keepX50.50.50$loadings)
```

```
biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50,result.po.treatmentdays.individuality.2level.kee
pX50.50.50$loadings,result.po.treatmentdays.individuality
.2level.keepX50.50.50$variates)
```

```
result.po.treatmentdays.individuality.2level.keepX50.50.5
0$variates
```

```
biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50,result.po.treatmentdays.individuality.2level.kee
pX50.50.50$variates$X,result.po.treatmentdays.individuali
ty.2level.keepX50.50.50$variates$Y)
```

```
result.po.treatmentdays.individuality.2level.keepX50.50.5
0$variates$X
```

```

X1<-
  result.po.treatmentdays.individuality.2level.keepX50.50.5
  0$variates$X

X1a<-X1[,2:3]

X1a

X1a<-X1[,1:2]

X1a

result.po.treatmentdays.individuality.2level.keepX50.50.5
0$loadings

result.po.treatmentdays.individuality.2level.keepX50.50.5
0$loadings$X

y1<-
  result.po.treatmentdays.individuality.2level.keepX50.50.5
  0$loadings$X

Y1<-
  result.po.treatmentdays.individuality.2level.keepX50.50.5
  0$loadings$X

Y1

Y1a<-Y1[,2:3]

Y1a

Y1a<-Y1[,1:2]

Y1a

biplot(X1a,Y1a)

plotVar(result.po.treatmentdays.individuality.2level.keep
X50.50.50,rad.in=0.85,var.label=TRUE, cex=0.4)

plotVar(result.po.treatmentdays.individuality.2level.keep
X50.50.50,rad.in=0.85,var.label=TRUE, cex=0.3)

library(mixOmics)

```

```

plotVar(result.po.treatmentdays.individuality.2level.keep
X50.50.50,rad.in=0.85,var.label=TRUE, cex=0.4)

biplot(result.po.treatmentdays.individuality.2level.keepX
50.50.50)

result.po.treatmentdays.individuality.2level.keepX50.50.5
0$loadings

plotVar(result.po.treatmentdays.individuality.2level.keep
X50.50.50,rad.in=0.85,var.label=TRUE, cex=0.2)

plotVar(result.po.treatmentdays.individuality.2level.keep
X50.50.50,rad.in=0.85,var.label=TRUE, cex=0.1)

write.csv(result.po.treatmentdays.individuality.2level.ke
epX50.50.50$loadings$X)

write.csv(result.po.treatmentdays.individuality.2level.ke
epX50.50.50$loadings$X,"VIP")

write.csv(result.po.treatmentdays.individuality.2level.ke
epX50.50.50$loadings$X,file="VIP")

write.csv(result.po.treatmentdays.individuality.2level.ke
epX50.50.50$loadings$X,file=VIP)

write.csv(result.po.treatmentdays.individuality.2level.ke
epX50.50.50$loadings$X,file="VIP")

write.csv(result.po.treatmentdays.individuality.2level.ke
epX50.50.50$loadings$X,file="VIP.csv")

library(mixOmics)

plotVar(result.po.treatmentdays.individuality.2level.keep
X50.50.50,rad.in=0.85,var.label=TRUE, cex=0.1)

plotIndiv(result.po.individuality.1level,
col=as.factor(po), ind.names = FALSE, pch=(1:10))

plotVar(result.po.individuality.1level, comp = 1:2,
rad.in = 0.9,

```



```

plot3d(result.po.individuality.1level$loadings)

result.po.individuality.treatmentdays.2level<-
multilevel(data.x, cond = po.individuality, sample =
treatmentdays, ncomp = 3, keepX = c(50,50,50),
tab.prob.gene = NULL, method = 'spllda')

plotVar(result.po.individuality.1level, comp = 1:2,
rad.in = 0.9,

font = NULL)

repeat.simu2 = rep(c(1:6), 8)

fix(repeat.simu2)

rm(repeat.simu2)

fix(result.po.treatmentdays.1level)

pheatmap.multilevel(result.po.individuality.1level,col_sa
mple=sample,col_stimulation=po,label_annotation=NULL,bord
er=FALSE,clustering_method="ward",show_colnames =
FALSE,show_rownames = TRUE,fontsize_row=2)

plotIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

plotIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names =
sample)

plotIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

plotIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

legend('bottomright', col = unique(col.po),legend =
levels(po), pch = 20, cex = 0.8)

legend('topright', col = 'black', legend =
levels(treatmentdays),pch = unique(pch.treatmentdays),
cex = 0.8)

```

```

legend('topright', col = 'black', legend =
  levels(treatmentdays), pch = unique(pch.treatmentdays),
  cex = 0.8)

pch.treatmentdays <- rep(10, 48)

legend('topright', col = 'black', legend =
  levels(treatmentdays), pch = unique(pch.treatmentdays),
  cex = 0.8)

plotIndiv(result.po.treatmentdays.individuality.2level,
  col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

plotIndiv(result.po.treatmentdays.individuality.2level,
  col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

plotIndiv(result.po.treatmentdays.individuality.2level,
  col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

plotIndiv(result.po.individuality.1level,
  col=as.factor(po), ind.names = FALSE, pch=(1:10))

plotIndiv(result.po.individuality.1level,
  col=as.factor(po), ind.names = FALSE, pch=(1:10))

plotVar(result.po.treatmentdays.individuality.2level,
  comp = 1:2, X.label = TRUE, Y.label = TRUE, cex =
  c(0.5,0.9))

plotIndiv(result.po.treatmentdays.individuality.2level,
  col = col.po, pch = pch.treatmentdays, ind.names = FALSE)

plotIndiv(result.po.treatmentdays.individuality.2level,
  col = col.po, pch = pch.treatmentdays, ind.names =
  as.numeric(treatmentdays))

names(result.po.treatmentdays.individuality.2level)

result.po.treatmentdays.individuality.2level.keepX50.50.5
0<- multilevel$mat.c

result.po.treatmentdays.individuality.2level.keepX50.50.5
0<- $mat.c

```

```

result.po.treatmentdays.individuality.2level.keepX50.50.5
0$mat.c

fix(result.po.treatmentdays.individuality.2level.keepX50.
50.50)

write.csv(result.po.treatmentdays.individuality.2level.ke
epX50.50.50$mat.c,file="VIP.coefficient.csv")

library(mixOmics)

plotVar(result.po.treatmentdays.individuality.2level.keep
X50.50.50,rad.in=0.85,var.label=TRUE, cex=0.15)

plot3dIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names =
as.numeric(treatmentdays))

plotVar(result.po.treatmentdays.individuality.2level,
comp = 1:2, X.label = TRUE, Y.label = TRUE, cex =
c(0.5,0.9))

plotIndiv(result.po.treatmentdays.individuality.2level,
col = col.po, pch = pch.treatmentdays, ind.names =
as.numeric(treatmentdays))

plotVar(result.po.treatmentdays.individuality.2level,
comp = 1:2, X.label = TRUE, Y.label = TRUE, cex =
c(0.5,0.9))

plotVar(result.po.treatmentdays.individuality.2level,
comp = 1:2, X.label = TRUE, Y.label = TRUE, cex =
c(0.5,0.9))

plot3dVar(result.po.treatmentdays.individuality.2level.ke
epX50.50.50,cutoff=NULL,var.label=TRUE, cex=0.5,
rad.in=0)

plotVar(result.po.treatmentdays.individuality.2level.keep
X50.50.50,rad.in=0.85,var.label=TRUE, cex=0.15)

plot3dVar(result.po.treatmentdays.individuality.2level.ke
epX50.50.50,rad.in=0.85,var.label=TRUE, cex=0.15)

```

5

Metabolomics study of hydroxytyrosol's administration for the treatment of metabolic syndrome in a rat model

One part of the chapter was submitted for publication in British Journal of Pharmacology and one part will be submitted

ABSTRACT

Metabolic syndrome (MetS) is a clustering of interrelated risk factors for cardiovascular disease and diabetes. The Mediterranean diet has been proposed as an important dietary pattern to confer cardioprotection by attenuating risk factors of MetS. Hydroxytyrosol (HT) is a single phenol and it is present in large amount in olive byproducts and in lower amounts in olive oil and in olive fruits, which are basic constituents of the Mediterranean diet. In the context of identifying the features of HT that are responsible for its effects on the MetS, an experimental protocol has been setup encompassing isolated HT administration in a diet induced model of MetS in young Wistar rats. Rats were randomly divided into four groups as cornstarch (CS), cornstarch + HT (CSHT), high-carbohydrate high-fat (HF) and high-carbohydrate high-fat + HT (HFHT) (n=6/group). HT (20 mg/kg/d oral gavage, water vehicle) was administered for 8 weeks on the basal diet (CS or HF) for a further 8 weeks. In order to gain inside on the metabolic effects of HT administration on the total biochemical plasma profile (metabolome), an untargeted approach (metabolomics) has been developed and attempted on the plasma samples from the MetS model. The new biomarkers discovered from multivariate analysis have been identified by comparison of public domain metabolome libraries but also by confirmation using HRMS/MS, highlighting the differences in the biochemistry of HFHT animals versus controls as well as versus CSHT. HT supplementation for 8 weeks reduced visceral obesity and was associated with improved left ventricular structure and function, reduced blood pressure, improved glucose disposal and reduced hepatic steatosis and infiltration of inflammatory cells into the liver. Specific biomarkers were determined through the metabolomics approach supporting the pharmacological findings.

KEYWORDS

Hydroxytyrosol, metabolomics, metabolic pathway analysis, metabolic syndrome

INTRODUCTION

Metabolic syndrome (MetS) is defined as a clustering of interrelated risk factors for cardiovascular disease and diabetes. These factors include hyperglycemia, hypertension, dyslipidemia, central adiposity, non-alcoholic fatty liver disease, insulin resistance and type II diabetes, chronic low grade inflammation and abdominal obesity¹. Excess fat mass is often conjugated with MetS² and thus in recent years the prevalence of MetS has increased directly with the epidemic of obesity³. It is well known that people with MetS are at high risk of developing both cardiovascular disease (CVD)⁴ and type II diabetes⁵, a fact that led to an important public health concern. Therefore, several organizations have attempted to formulate simple criteria for its diagnosis and treatment⁶. Although there is experimental evidence, which indicate that dysfunctional adipose tissue has an unfavorable effect on metabolism and thereby seems to cause some of the obesity associated metabolic morbidities such as insulin resistance and type II diabetes^{7,8}, nutritional factors may also contribute to the development of MetS⁹.

Changes in lifestyle, characterized by reduced physical activity and increased consumption of energy-dense food, have resulted in the dramatic rise of obesity incidences and related conditions worldwide^{10,11}. According to current guidelines, a critical step in the management of MetS is encompassing a list of therapeutic lifestyle modifications¹². Although there is general agreement about reducing weight and increasing physical activity, a uniform consensus about a “health promoting” diet is lacking¹³. Diets with increased carbohydrate content, mainly as fructose¹⁴, increased fats¹⁵ or a combination of both¹⁶ have been tested to a rat model in order to mimic the human metabolic syndrome. In addition, there are epidemiological evidence

suggesting that diets characterized by low consumption of saturated and trans fatty acids, cholesterol, sodium, monosaccharides etc., but with increased intake of vegetables fruits, fish etc. are beneficial for individuals with MetS¹³. These features adhere to the principles of the traditional Mediterranean diet¹⁷ and are associated with longer survival.

During the last few decades the Mediterranean diet has been proposed as an important dietary pattern to confer cardioprotection by attenuating risk factors of MetS such as adiposity, insulin resistance and type II diabetes, atherogenic dyslipidemia, hypertension and chronic low grade inflammation¹⁸, translated to lower rates of morbidity and mortality. Furthermore, Mediterranean diet has gained in popularity worldwide due to its documented contribution for better health and quality of life of those who adhere to it¹⁷. The relationship between the Mediterranean diet adherence and the annihilation of MetS has been also investigated in a few intervention studies¹⁹, indicating its beneficial role for the prevention of the syndrome.

It is well known that high consumption of monounsaturated fatty acids (>70% dietary fat) is the hallmark of the Mediterranean diet, with olive oil as the primary fat source containing oleic as its major fatty acid²⁰. According to the literature, the risk of developing CVD is lower in cohorts consuming large amounts of olive oil²¹. Studies in cell cultures, animals and humans have confirmed the cardioprotective action of the Mediterranean diet and the olive oil constituents which is facilitated by attenuating the risk factors of MetS^{18,22,23}. Although these effects have been attributed, at least in part, to the high-oleic acid content of olive oil in the Mediterranean diet, more recent evidence emphasizes the role of minor but highly bioactive compounds of virgin olive oil, such as squalene, sterols and phenolic compounds^{24–28}. Due to this chemical

composition, virgin olive oil favors a better lipid profile, reduces blood pressure (BP) levels, endothelial dysfunction, and the inflammatory and prothrombotic environment²⁵.

The phenolic compound hydroxytyrosol (HT) or (3,4-Dihydroxyphenyl) ethanol) and its precursor oleuropein, are two highly bioactive mediators of the cardioprotective activity of olive oil²⁹. HT comes from the hydrolysis of oleuropein, which originates during the maturation of the olives, storage of the oil, and preparation of table olives. After ingestion, oleuropein is hydrolyzed in the gastrointestinal tract to HT, elenolic acid, oleuropein aglycone and glucose^{30,31}. The biological responses of HT are proposed to be mediated through its potent anti-inflammatory and free radical scavenging properties²⁹. HT lowers the expression of vascular and intracellular cell adhesion molecules by blocking the activation of the transcription factor NF- κ B, reduces LDL oxidation, and inhibits the expression of pro-inflammatory genes including those for nitric oxide synthase and cyclooxygenase-2²⁹. Furthermore, HT up-regulates components of the antioxidant pathways such as Nrf-2 and heme-oxygenase-1 (HO-1), most likely through the upstream activation of PI3K/Akt and ERK1/2 pathways²⁹. The pharmacological effects of HT in MetS have been implied from the reduced risk factors for cardiovascular diseases by the nutritional supplementation of extra virgin olive oil rich in phenolic compounds or the phenolic fraction of olive oil^{16,32,33}, rather than from studies on the isolated compound itself.

Due to the fact that HT comes from the hydrolysis of oleuropein, which originates during the maturation of the olives, storage of the oil, and preparation of table olives, byproducts from olive oil production constitute a major source of HT³⁴. There is growing interest in the health effects of HT, for its potential application in the

treatment of inflammatory conditions such as CVD³⁵. Current knowledge indicates that HT can fortify health³⁶. Many studies of this compound have demonstrated its powerful antioxidant capacity³⁷, as well as its anti-inflammatory³⁸ and antiplatelet aggregation action³⁹, which may help counteract the development of chronic diseases such as cardiovascular disease and cancer⁴⁰. For these reasons, there are new evidence and strategies regarding the use of HT as a natural drug for the prevention and treatment of diseases with high incidences in Western countries³⁶.

Recently, a health claim on “olive oil polyphenols” was established and endorsed⁴¹. The health claim was spelled as: “Olive oil polyphenols contribute to the protection of blood lipids from oxidative stress”. The claim may be used only for olive oil, which contains at least 5 mg of HT and its derivatives (e.g. oleuropein complex and tyrosol) per 20 g of olive oil. Furthermore, information is given to the consumer that “the beneficial effect is obtained with a daily intake of 20 g of olive oil”.

It has been previously reported that a high-carbohydrate, high fat-diet feeding in rat produces various symptoms of MetS including impaired glucose and insulin tolerance, elevated plasma lipid concentrations, hypertension, abdominal adiposity and remodeling of the heart, liver, kidneys and pancreas therefore mimicking the human condition^{16,42}. In this model of diet-induced MetS, olive leaf extract enriched in olive oil phenolic compounds has been found to improve glucose tolerance, normalize abdominal adiposity, reduce plasma triglyceride and total cholesterol concentrations, reduce plasma uric acid and malondialdehyde concentrations, normalize inflammation and fibrosis in the heart and the liver, and improve left ventricular and hepatic function without changing blood pressure²³.

In order to study complex biological mechanisms, such as the ones described, and study the underline biochemistry, a systems biology approach is required. Modern science uses the principles of omics technologies, especially metabolomics, in order to tackle such issues. Metabolomics is the unbiased identification and quantification of small molecules in biological fluids and it deals with the sum of metabolites, also referred as the metabolome, within a cell, an organism, or a system. Metabolomics provides an approach to explore the metabolic effects of potentially therapeutic substances on specific biological systems⁴³. It gives a snapshot of the metabolome in a particular condition and it also aids towards the identification of metabolites that are perturbed in response to the factor under study. The pipeline of metabolomics analysis primarily consists of sample preparation, instrumental analysis of the processed samples (data acquisition), data preprocessing and processing, and finally biological interpretation of the obtained results^{44,45}. Due to the fact that metabolomics can be considered as the ideal methodology for measuring small-molecule metabolite profiles in biological matrices, it potentially allows the systematic characterization of the metabolome variation associated with MetS as well as the effect of treatment with HT. Such an approach can augment the understanding of MetS from a biological point of view thereof^{46,47}.

Numerous tools and methodologies have been proposed and employed in metabolomics. Among them, mass spectrometry-based analysis is considered as the method of choice. Its variation, namely ultra performance liquid chromatography coupled with high- resolution mass spectrometry (UHPLC-HRMS), is a robust method, which provides broad coverage of metabolites and high sensitivity, thus it is widely used to build metabolic profiles in animal or human studies⁴⁸.

Metabolic pathway analysis (MPA) is a vital part in a metabolomic study. It is performed by combining advanced pathway enrichment analysis with the pathway topological analysis to help identify the most relevant metabolic pathways involved in diseases and cellular processes⁴⁹. Over the past decade, computer-assisted pathway analysis methods have gained widespread acceptance in research involving “omics” data analysis^{50,51}.

In the current study, the effects of HT on cardiovascular, hepatic and metabolic parameters in a high-carbohydrate (fructose), high-fat (tallow) diet with a corn starch-rich, low fat diet serving as the control diet, was been used in order to establish a MetS model in rats. Following, the responses of 8 weeks HT supplementation on body composition, abdominal adiposity, glucose and insulin tolerance, systolic blood pressure (SBP), vascular reactivity, plasma lipid, and structure and function of the heart and the liver was assessed. Furthermore, an untargeted UHPLC-HRMS based plasma metabolomic approach coupled with MPA was applied in order to characterize the global metabolic perturbations of MetS induced by HT, using a MetS rat model. The experimental procedure of the current study is depicted in Figure 1.

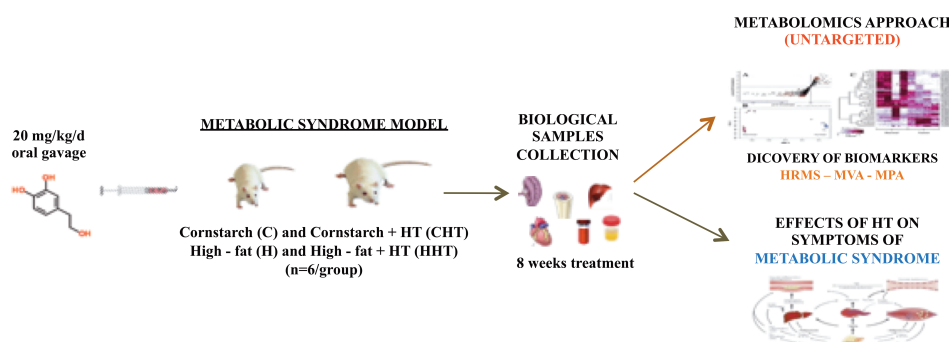


Figure 1: Schematic representation of the experimental workflow.

The aim of the study was to provide insights concerning the effect of HT treatment on a MetS induced rat model, exploring the plasma metabolome. In addition, the current project aims to unravel the underlying concerning the pathogenesis of MetS at the metabolic level and to discover specific blood biomarkers associated with MetS, which may reveal etiological pathways and help to identify individuals at risk for MetS.

MATERIALS AND METHODS

Isolation of hydroxytyrosol

HT (Figure 1) was isolated from olive mill wastewater (OMWW), as described below. Its purity was found to be more than 98% (HPLC-UV, Surveyor system, Thermo Scientific, Germany) and its structure was identified by HRMS (LTQ Orbitrap Discovery XL mass spectrometer, Thermo Scientific, Germany) and 1 & 2D NMR [Bruker Avance III spectrometer operating at 600.11 MHz (Bruker Biospin GmbH, Reinsteten, Germany)]. All the solvents used for the extraction and isolation of hydroxytyrosol i.e. cyclohexane (c-hex), ethyl acetate (EtOAc), methanol (MeOH), water (H₂O), acetone and dichloromethane (CH₂Cl₂) were supplied by Merck. The Amberlite XAD-7HP and XAD-16 resins used in the absorption experiment was supplied by Sigma-Aldrich. Si gel 60 (15–40 mm) used for the isolation of hydroxytyrosol was purchased from Merck. Preparative hydrostatic CCC separations were carried out on a Kromaton FCPC instrument equipped with a rotor of 1000 mL. The OMWW fractions were injected into the system through a 50 mL loop and fractions were collected with a Buchi B-684 fraction collector. Evaporation of the

samples in the hydroxytyrosol extraction and purification study was performed with the aid of a Buchi evaporator system consisting of a rotavapor R-210, a vacuum pump V-700 and a vacuum controller V-850 (Buchi, Switzerland).

In details, two protocols were followed for the isolation of HT. The first one was performed using the CCC approach, whereas the second one by the aid of the MPLC technique. It should be noted, that the initial extraction step was common for both isolation processes. Hence, the initial HT extract was isolated from OMWW based on previously described in-house procedure⁵². Briefly, the experimental protocol consisted of three parts. Initially, quantity of OMMW was filtered consecutively using stainless steel Waterscreen rotating screens with different slots' width (250 and 120 μm respectively) and then it was passed through a Pall bag filter with 25 μm pore sizes in order clean up the material. Subsequently, the filtered OMWW was passed through a series of adsorbent resins, XAD-16 and XAD-7HP and the absorbents containing the polyphenols were collected. All the absorbents were filtered again in order to remove suspended solids from the resin and then were evaporated until dry in order to remove the solvents.

In the first approach for the isolation of HT, 32 g of the remaining residue was submitted to countercurrent chromatography, using the FCPC apparatus. The system of the solvents used in this procedure was cyclohexane/EtOAc/ MeOH/H₂O: 4/6/4/6. The capacity of the column was 1 L, the rapidity of the rotation was 900 rpm, and the flow rate was 18 mL/ min. Following the “ascending” and the “descending” mode, the obtained fractions were collected and 6 g of hydroxytyrosol (purity 95%) were isolated by the above-mentioned process. The procedure was repeated for four times in order to isolated the appropriate quantity of the hydroxytyrosol for the completion of the experiments.

In the second approach, 400 g of the remaining residue was diluted in 1 L acetone, remained 2h for sonication and then filtered. The acetone extract was evaporated and subsequently the evaporated acetone extract was diluted in 1 L MeOH for a second extraction. After sonication for 2h, the methanol extract filtrated and then evaporated. 300 g of the final extract was subjected to reverse phase medium-performance liquid chromatography (MPLC) with Si gel 60 Merck (15–40 mm), using the H₂O-MeOH gradient as the eluent and by decreasing the polarity. The obtained water fraction was evaporated, giving an oily residue, which was identified by NMR spectroscopy as a mixture of mainly hydroxytyrosol, tyrosol, and a lactone (3:1:1). Then, the water fraction (150 g) was subjected to chromatographic fractionation in a normal phase MPLC with a CH₂Cl₂/MeOH system of increasing polarity, and 20 g of hydroxytyrosol were isolated.

Study subjects-intervention study

The experimental groups consisted of 48 male Wistar rats (9-10 weeks old) supplied by the Biological Resources unit of the University of Queensland (Australia) and individually housed in a temperature-controlled ($20 \pm 2^{\circ}\text{C}$), 12-hour light/dark cycle environment with *ad libitum* access to water and the group-specific rat diet at the Animal House of the University of Southern Queensland. The rats were randomly divided into four separate groups ($n = 7$ each) and fed with corn starch (C; 330 ± 3 g), corn starch + HT (CHT; 324 ± 3 g), high-carbohydrate, high-fat (H; 336 ± 2 g), high-carbohydrate, high-fat + HT (HHT; 336 ± 2 g).

Body weight and food and water intakes were measured daily and feed efficiency (%) was calculated as described in a previous study⁵³. The preparation and macronutrient composition of basal diets, including the dietary fatty acid profiles,

have been described in detail previously^{23,42,53}. HT (20 mg/kg) dissolved in distilled water was administered daily by oral gavage for 8 weeks starting 8 weeks after the initiation of the corn starch or high-carbohydrate, high-fat diet in CHT and HHT groups, as depicted in Figure 2. The drinking water in all high-carbohydrate, high-fat diet-fed groups was augmented with 25% fructose for the duration of the study.

All experimentation was approved by the Animal Experimentation Ethics Committees of The University of Queensland and the University of Southern Queensland under the guidelines of the National Health and Medical Research Council of Australia.

Biological samples were acquired and analyzed from 28 animals.

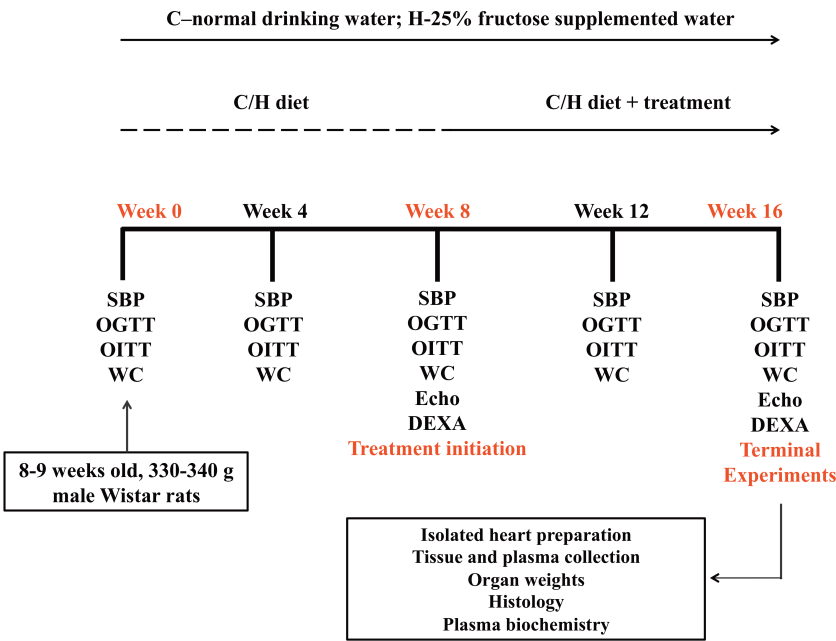


Figure 2: Schematic representation of the intervention study. SBP, systolic blood pressure; OGTT, oral glucose tolerance test; OITT, oral insulin tolerance test; WC, waist/abdominal circumference; Echo, echocardiography; DEXA, dual energy X-ray absorptiometry.

Biofluid collection

After the completion of the protocol (16 weeks), blood samples were collected into appropriate heparinized tubes, centrifuged at 5,000 g for 15 min within 30 min of the collection. Plasma was separated and transferred to microcentrifuge tubes for storage at -80 °C before analysis.

Samples

Twenty-eight rats were divided into four separate groups, i.e. C, CHT, H and HHT (n=7 each group). The corresponding plasma samples from each group were coded appropriately. For the metabolomics study, 28 plasma samples (7 from each group) have been analyzed.

Pharmacological evaluation

Pharmacological evaluation was performed in the School of Biomedical Sciences and Department of Biological and Physical Sciences at the University of Queensland (Brisbane, Australia) by Dr. Hemant Poudyal, under the supervision of Prof. Lindsay Brown. Briefly, the following experimental measurements were acquired. Cardiovascular measurements i.e. systolic blood pressure (SBP), echocardiographic examination (echo); oral glucose and insulin tolerance tests (OGTT and OITT); body composition measurements i.e. dual energy X-ray absorptiometric (DEXA) measurements; organ weights i.e. liver, heart, spleen, brain, adipose tissue, brown adipose tissue, skeletal muscle, small intestine ileum and kidney; organs histology; plasma biochemistry i.e. plasma alanine transaminase (ALT) and aspartate transaminase (AST), plasma alkaline phosphatase (ALP) and lactate dehydrogenase (LDH), bilirubin, albumin and globulin; and waist/abdominal circumference (WC).

Chemical, reagents and Instrumentation

Acetonitrile (ACN), methanol (MeOH), water and formic acid, were purchased from Fluka/Riedel-de Haën (Switzerland). UHPLC for the HUPLC-orbitrap metabolomic analysis was performed employing an Accela system (Thermo Scientific, Germany) equipped with a binary pump, an autosampler, an online vacuum degasser and a temperature-controlled column compartment, whereas high-resolution mass spectrometry was performed on a hybrid LTQ Orbitrap Discovery XL mass spectrometer (Thermo Scientific, Germany). UHPLC for the UHPLC-qQToF metabolomic analysis was performed employing an Acquity UPLC™ system (Waters Corp., Milford, MA, USA) equipped with a binary solvent manager and a temperature and light-controlled autosampler (sample manager) capable of maintaining the sample temperature from -5 to 40 °C and a column heater, whereas high-resolution mass spectrometry was performed on a Waters Micromass (Milford, USA) hybrid quadrupole-time of flight mass spectrometer (QTOF Premier) equipped with an electrospray ionization (ESI) interface. Centrifuging of the plasma samples during the sample pretreatment protocol stage was performed by a Mikro 200R centrifuge (Hettich Lab Technology, Germany) whereas their evaporation was performed with the aid of a GeneVac HT-4X EZ-2 series evaporator Lyospeed ENABLED (Genevac Ltd, UK).

Sample pretreatment

The rat plasma samples (28 rat samples, 7 from each studied group C, CHT, H and HHT) were prepared, following a sample preparation protocol⁵⁴, described below. The plasma samples were allowed initially to thaw on ice at 4 °C for 30–60 min. Following, a 200 μ L aliquot from each plasma sample were placed into a labeled 1.5

mL eppendorf tube and 600 μ L of MeOH were added followed by vortexing for 15 s. The samples were subsequently centrifuged at 15800 g for 15 min at room temperature to pellet the protein precipitate, and 185 μ L aliquots from the supernatant were transferred into two separate labeled 1.5 mL eppendorf tubes and finally evaporated in the GeneVac for 90 min at 50°C. Consequently, 100 μ L of water was added to each dry sample, vortexed for 15 s, centrifuged at 15800 g for 15 min and transferred to 200 μ L inserts placed in appropriate screw capped autosampler vials.

UHPLC-HRMS Analysis

UHPLC-orbitrap. An ACQUITY UPLC BEH C18 (2.1 \times 100 mm, 1.7 μ m) reversed phase column (Waters Corp., Milford, MSA, USA) preceded by a precolumn (Waters VanGuard 5 \times 2.1 mm, 1.7 μ m) was used for the chromatographic separation in the plasma samples. The mobile phase consisted of solvents A: aq. 0.1% (v/v) formic acid and B: methanol-0.1% (v/v) formic acid. Different gradient elutions⁵⁴ were performed for positive and negative ion mode detection, as described in Table 1, with flow rates of 0.36 and 0.40 mL/min for positive and negative ion mode detection, respectively.

In detail, a gradient method (total run time of 24 min) was used for the profiling of the plasma samples in positive ion mode as follows: 0 to 1 min: 100% A: 0% B, 1 to 16 min: from 100% A: 0% B to 0%A: 100%B, 16 to 20 min: 0% A: 100% B, 20 to 22 min: from 0% A: 100% B to 100%A: 0% B, 22 to 24 min: 100% A: 0% B. For the profiling of the plasma samples in the negative ion mode, a slightly different gradient method (total run time of 26 min) was used as follows: 0 to 2 min: 100% A: 0% B, 2 to 17 min: from 100% A: 0% B to 0%A: 100%B, 17 to 22 min: 0% A: 100% B, 22 to

24 min: from 0% A: 100% B to 100%A: 0% B, 24 to 26 min: 100% A: 0% B.

Column temperature was kept at 50 °C throughout all experiments while the autosampler tray temperature was set at 4 °C and the injection volume was 10 µL.

LC parameters

Column: ACQUITY UPLC BEH C18 (2.1 × 100 mm, 1.7 µm) reversed phase

Mobile phase: water, 0.1% (v/v) formic acid (A) and methanol, 0.1% (v/v) formic acid (B)

Flow rate: 0.36 mL min⁻¹ ESI (+) / 0.40 mL min⁻¹ ESI (-)

Column temperature: 50 °C

Tray temperature: 4 °C

Injection volume: 10 µL

ESI (+)	Gradient Program		
	Time (min)	A%	B%
	0	100	0
	1	100	0
	16	0	100
	20	0	100
	22	100	0
	24	100	0
ESI (-)	Gradient Program		
	Time (min)	A%	B%
	0	100	0
	2	100	0
	17	0	100
	22	0	100
	24	100	0
	26	100	0

MS parameters

ESI mode

30000 resolution

Centroid mode

ESI (+)	Capillary temperature (°C): 356
	Capillary voltage (V): -60
	Tube lens (V): 110
	Spray voltage (kV): 3.50
	Sheath gas flow (arb. units): 30
	Aux gas flow (arb. units): 10
ESI (-)	Capillary temperature (°C): 356
	Capillary voltage (V): 20
	Tube lens (V): -49
	Spray voltage (kV): 3.10
	Sheath gas flow (arb. units): 30
	Aux gas flow (arb. units): 10

Table 1: UHPLC-HRMS (orbitrap) parameters used in the plasma metabolomics study. Gradient elution programs and mass spectrometry parameters applied for UHPLC-HRMS (orbitrap) analysis for ESI (+) and ESI (-) modes.

A hybrid HRMS (LTQ-Orbitrap Discovery) instrument, operated separately in ESI positive and negative ion mode has been used. The conditions for the HRMS in each ionization mode were set as follow: for the ESI (-) mode, capillary temperature, 356°C; capillary voltage, 20 V; tube lens, -49 V; source voltage, 3.10 kV; sheath gas flow, 30 arb. units; aux gas flow, 10 arb. units, while for the ESI (+) mode, capillary temperature, 356°C; capillary voltage, -60 V; tube lens, 110 V; source voltage, 3.50 kV; sheath gas flow, 30 arb. units; aux gas flow, 10 arb. units (Table 1). Analysis was performed using the Fourier transform mass spectrometry (FTMS) full scan ion mode, applying a resolution of 30000, while acquisition of the mass spectra was performed in every case using the centroid mode.

UHPLC-qQToF. In the case of UHPLC-qQToF analysis, the chromatographic separation in the plasma samples was performed in the same manner as in the case of UHPLC-orbitrap, as described above. However, the HRMS analysis was accomplished with the aid of the qQToF mass spectrometer as follow: the electrospray voltage was kept at 1.5 kV and 2.5 kV for the ESI (-) and (+) mode respectively, the sample cone voltage was maintained at 35 V affording the highest sensitivity, whereas the extraction cone voltage was set at 3 V and the MCP plates were operated at 1850 V. Nitrogen was used both as the desolvation and cone gas and was set at 650 and 50 L/h, respectively. The desolvation temperature was set at 350 °C and the source temperature was 120°C. The TOF analyzer was operated in the V optics mode affording a resolution of 9500. The collision gas was argon. For the

acquisition of the full-scan mass spectra the collision energy was set at 2 V. Full-scan mass spectra were acquired in continuum and centroid mode in the mass range of 50–1000 amu, whereas the scan time was 0.4 s with an inter scan delay of 0.02 s. In order to improve mass measurement accuracy a lock mass compound (500 pg/ μ L solution of leucine enkephalin; m/z 556.2771) was continuously infused at a 2 μ L/min using the built-in syringe pump of the instrument.

Quality control samples

It is well known that ideal quality control (QC) samples should be identical biological samples, with a metabolic and sample matrix composition similar to those of the unknowns⁵⁴. An ideal QC sample regarding a small sample set would be either a pooled sample derived from the mixture of the unknowns or a commercially available biofluid. Pooled QC samples exhibit a number of advantages, as they are closely matching the composition of the unknown samples, whereas they are well suited for small studies in which all of the samples are available in a sufficient quantity (e.g. animal studies or small clinical trials).

In the current study, pooled plasma samples were used as QC samples. As described above, a QC is usually prepared by thoroughly mixing equal aliquots from all the test samples of the experimental sample set and is subsequently processed following the developed sample preparation protocol. Thus, equal aliquots from the 28 studied samples were thoroughly mixed and the mixture was subsequently prepared in the same manner as the plasma samples (see sample pretreatment section).

Five QC samples were injected at the beginning of each analytical batch, one QC every ten samples throughout the run and five QC samples at the end of the batch (Figure 3). In order to eliminate the instrumental concerns, the ion transfer tube has

been removed and cleaned every 50 injections. Furthermore, software-based correction has been applied to the data, applying data normalization algorithms. Following this approach, the obtained data sets could be easily evaluated in terms of repeatability, possible outliers can be detected, whereas deterioration of the instrumental performance could be recorded.

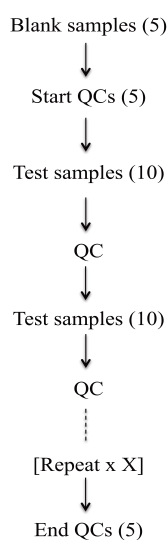


Figure 3: Schematic representation of a typical run in UHPLC-HRMS-based metabolomics study.

Orbitrap-based metabolomics

Data pre-processing. Prior to the multivariate statistical analysis data preprocessing was performed employing Xcalibur[®] (Thermo Fisher Scientific), MZmine 2.10 (<http://mzmine.sourceforge.net>), Microsoft Excel 2012 and the web-based MetaboAnalyst 2.0 suite (<http://www.metaboanalyst.ca>). Initially, Xcalibur[®] was used to convert the instrument native Xcalibur data files (*.raw) to the cdf data format (*.cdf) with the aid of the embedded to the Xcalibur[®] Xconvert program. Data were then imported to MZmine 2.10 and processed applying the baseline correction, peak detection, deconvolution, deisotoping, alignment and gap filling procedures.

Plasma Data Sets		
	ESI (+)	ESI (-)
Baseline correction		
Type	Base peak chromatogram	Base peak chromatogram
Smoothing	500	500
Asymmetry	0.001	0.001
<i>m/z</i> bin width	1	1
Mass detection		
Algorithm	Centroid	Centroid
Noise level (cps)	1.2E5	1E5
Chromatogram builder		
Minimum time spam (min)	0.05	0.05
Minimum height (cps)	1.2E5	3E4
<i>m/z</i> tolerance (ppm)	5	5
Chromatogram deconvolution		
Algorithm	Wavelets (XCMS)	Wavelets (XCMS)
S/N threshold	5	5
Wavelet scales (min)	0.1-5	0.1-6
Peak duration range (min)	0.1-1	0.05-1
Deisotoping		
Algorithm	Isotopic peaks grouper	Isotopic peaks grouper
<i>m/z</i> tolerance (ppm)	5	5
<i>t_R</i> tolerance (min)	0.05	0.05
Identification of adducts	√	√
Identification of peak complexes	√	√
Alignment		
Algorithm	RANSAC	RANSAC
<i>m/z</i> tolerance (ppm)	5	5
<i>t_R</i> tolerance (min)	0.03	0.03
<i>t_R</i> tolerance (min) after correction	0.05	0.05
RANSAC iterations	15000	15000
Minimum number of points (%)	20	20
Threshold	4	4
Model	non linear model	non linear model
Gap filling		
Algorithm	Peak finder	Peak finder
Intensity tolerance (%)	80	80
<i>m/z</i> tolerance (ppm)	5	5
<i>t_R</i> tolerance (min)	0.05	0.05

Table 2: MZmine parameters for the data pre-processing of both ESI (+) and ESI (-) plasma data sets.

At the end of the procedure, the peak list (accurate mass - t_R vs. intensity) generated was exported as .csv file. All the parameters used for the aforementioned processes in MZmine, are summarized in Table 2. The generated .csv files (accurate mass - t_R vs. intensity) have been imported to Microsoft Excel 2012 and manipulated appropriately using the CONCATANATE, ROUND and TRANSPPOSE commands.

Metabolomic data processing and statistical analysis were performed for the plasma LC-HRMS chromatograms by following the workflow of MetaboAnalyst 2.0. MetaboAnalyst is a web-based comprehensive tool suite for metabolomic data analysis, which incorporates a set of nearly 500 commands of the R statistical language. Accordingly, the generated peak list was imported to MetaboAnalyst according to the instructions⁵⁵. Thus, (a). possible missing values were imputed applying the k-nearest neighbor (KNN) algorithm and subsequently (b). data filtering has been performed applying the interquantile range (IQR), as a robust estimator of the noise, eliminating 25-40% of the near-constant throughout the experiment conditions values, (c). row-wise normalization was achieved by averaging all the samples (pseudo-reference sample) in the control group in each case whereas either unit variance (UV) or Pareto scaling was evaluated in order to approximate as close as possible the normal distribution. The generated data was used for further analysis, including exploratory multivariate techniques e.g. Principal Component Analysis (PCA) and Partial Least Square-Discriminant Analysis (PLS-DA).

Chemometric data analysis. The first objective in the data analysis process is to reduce the dimensionality of the complex data set to enable easy visualization of any metabolic clustering of the different groups of samples. Thus, PCA has been initially used to examine the intrinsic variation in the datasets, detect possible outliers using

the 2D scores plots as well as to examine the quality characteristics of the method by the visual inspection of the QC clustering in the aforementioned plot.

Hierarchical clustering of signature metabolites altered in HHT group compared to H group was performed in MetaboAnalyst as an unsupervised clustering methodology. Hierarchical clustering begins with each sample as separate cluster and then proceeds to combine them until all samples belong to one cluster. The result was presented as a heatmap, which has been constructed in MetaboAnalyst platform applying the Pearson distance measure, the Ward clustering algorithm and displaying the thirty top features ranked by the t-test and ANOVA.

PLS-DA has been sequentially used to determine any underline possible classification between the examined groups as well as to explore the important features that are responsible for that clustering. The performance of the PLS-DA models was evaluated using a series of validation tests, namely the k-fold cross validation ($k=10$), the leave one out cross validation (LOOCV), named also as Jackknife cross validation and permutation testing^{56,57}. R^2 and Q^2 were considered in order to choose the optimal model parameters and also for testing the predictability of the model concerning k-fold procedure, whereas these parameters were also used in LCOOV testing for assessing the robustness of the achieved classification. Permutation testing was performed allowing of 100 permutations and applying the prediction accuracy during training statistic test.

Following, the variable importance in projection (VIP) values were calculated for different PLS-DA approaches concerning the inspection of the data, i.e. C vs. H, C vs. CHT and H vs. HHT groups. The VIP values were used in order to indicate the most influential parameters to the classification of the aforementioned groups. The ten first-

variables in each component, all of them exhibiting VIP values >1.5 , were considered important in discriminating between the groups.

Moreover, other multivariate methods, i.e. orthogonal projection to latent structures–discriminant analysis (OPLS-DA), were also used, employing SIMCA P+ 10.5 (Umetrics, Umea, Sweden) and EZinfo 2.0 (Umetrics, Umea, Sweden). VIP and loading plots for every different OPLS-DA approach e.g. H vs. HHT groups, were constructed in order to evaluate the most influential metabolites. The quality of the model was determined by the goodness of fit and variation in the X (R^2X) and Y (R^2Y) variables and the predictability based on the fraction correctly predicted in a 1/7 cross-validation (Q^2). Seven-round cross-validation was performed to eliminate the risk of overfit with only 1 predictive component for 2 classes (H and HHT groups).

Additionally, in order to reevaluate the produced results univariate data analysis has been performed using the significance analysis of microarray (SAM), t-test, fold change (FC) monitoring and their combination in the form of the volcano plot. For this purpose, the FC threshold was set at a value of 2 throughout all analyses, while the significance threshold was set at 95% ($p\text{-value}=0.05$).

Identification and structural elucidation of potential target metabolites. The important features, which correspond to accurate mass - t_R , selected using PLS-DA approaches were putatively assigned to specific metabolites by match the selected features in terms of mass accuracy in online databases. More specifically, three publicly available databases of MS spectra of metabolites: (a) Human Metabolome Data Base (HMDB) (<http://www.hmdb.ca/>), (b) METLIN Metabolomics Database (<http://metlin.scripps.edu/index.php>), (c) Kyoto Encyclopedia of Genes and Genomes

(KEGG) (<http://www.genome.jp/kegg/>), (d) ChemSpider free chemical structure database (<http://www.chemspider.com>), (e) MassBank database (<http://www.massbank.jp>) and (f) LIPID MAPS (<http://www.lipidmaps.org>) lipidomics gateway were used and furthermore the data were compared with those from previously published literature. Furthermore, additional data were used to confirm the proposed structures, such as the isotopic pattern and the RDB information. The matching criterion for the comparison of the expected isotopic pattern with the measured one was kept to 5% of matching, whereas the RDB, which provides the degree of molecule's unsaturation, was considered. Additionally, a deconvolution step has been performed to the corresponding spectra, using the appropriate software i.e. the Mass Frontier 5.0.1 and its freely available counterpart AMDIS (<http://chemdata.nist.gov/dokuwiki/doku.php?id=chemdata:amdis>), in order to gain information about a possible fragmentation that takes place in the source. All the putatively identified metabolites were further estimated through a t-test involving the abundance of the signal in the H and HHT groups, with the aim to reveal any statistically significant differences concerning their means.

QqToF-based metabolomics

Data processing. The MassLynx V 4.1 (SCN 703) software (Waters) was used for the data acquisition and initial data manipulation, whereas the application manager MarkerLynx 2.0 has been used for the data processing, namely the peak picking, deconvolution, adduct or fragment formation recognition and isotope pattern clustering procedures. The data have been exported to EZinfo 2.0 (Umetrics, Umea, Sweden) or to SIMCA P+ 10.5 (Umetrics, Umea, Sweden) for further multivariate

analysis. There was a back-transfer to the in order to MassLynx and MarkerLynx in order to reevaluate the chromatographic peaks and the corresponding spectra.

Chemometric data analysis. The multivariate data analysis has been performed in the same manner as in the case of the orbitrap-based metabolomics (see section orbitrap-based metabolomics/chemometric data analysis). The metabolites recognized as significant ones by the multivariate approaches (PLS-DA and OPLS-DA) as well as by the univariate analysis (volcano plot), were search against the same online data bases as in the case of the orbitrap-based metabolomics, as well as against a custom made database embedded to MarkerLynx. The resulting identified metabolites were treated in the same was as in the orbitrap-based metabolomics in order to perform MPA.

Metabolite set enrichment analysis

Metabolite set enrichment analysis was performed by the aid of over representation analysis (ORA) in the enrichment analysis toolbox (<http://www.metaboanalyst.ca/MetaboAnalyst/faces/>) of the MetaboAnalyst platform, accessing on the database source including the KEGG, HMDB, METLIN and SMPD (<http://www.smpdb.ca/>). The above algorithm requires a list of compound names as input. Thus, the selected significant metabolites from the previous uni- and multivariate approaches, from both orbitrap- and qQToF-based metabolomics, were uploaded as compounds names or HMDB/KEGG ID's to the ORA, and the “pathway-associated metabolite sets filter” was applied to the data, which contains 88 metabolite sets based on normal metabolic pathways. Furthermore, in order to associate the findings with specific disorders, the results have also been analyzed

using the “disease-associated metabolite sets (blood) filter”, which contains 416 metabolite sets reported in human blood. The results were visualized both graphically and in a detailed table. The p-value indicates the probability of observing at least a particular number of metabolites from a certain metabolite set in the given compound list.

Metabolic pathway analysis

MPA was performed with MetPA (<http://metpa.metabolomics.ca/MetPA/faces/Home.jsp>) accessing on the KEGG, HMDB, SMPD and METLIN, to identify the affected metabolic pathways analysis and facilitate the visualization of the results. The significant identified features obtained from both approaches (orbitrap- and qQToF-based metabolomics) were imported as compound names or HMDB/KEGG ID's and possible associated biological pathways were investigated. Three parameters needed to be specified for pathway analysis: the specific pathway library, the algorithm for pathway enrichment analysis described above, and the algorithm for topological analysis⁵⁵. From the list of pathway libraries, *Rattus norvegicus* (rat), which contains 81 pathways, was selected for MPA. The p-value (from pathway enrichment analysis) indicates the statistical significance of association of the altered metabolites with the pathway and the pathway impact value (from pathway topological analysis) is calculated as the sum of the importance measures of the matched metabolites normalized by the sum of the importance measures of all metabolites in the pathway^{49,58}. Those pathways with $p < 0.05$ or an impact value > 0.2 were considered and filtered out as potential target pathways.

RESULTS AND DISCUSSION

Pharmacological evaluation

In this study, we have shown that, in a diet-induced model of the metabolic syndrome, treatment with 20 mg/kg/d hydroxytyrosol for 8 weeks decreased the adiposity indexes, including total body and abdominal fat mass, as it can be found in Table 3 along with a series of measurements, such as dietary intakes, body composition and “anthropometrics”, organ wet weights, and plasma biochemistry in C, CHT, H and HHT diet-fed groups.

Furthermore, systolic blood pressure was reduced in obese rats after the treatment with HT (H vs. HHT group), as it can be found in Table 4. Additionally, HT improved LV function seen as decreased diastolic stiffness, normalized LV developed pressure and ventricular dP/dt (Table 4). Changes in cardiovascular structure and function in C, CHT, H and HHT diet-fed groups can be found in Table 4.

Moreover, H diet feeding for 16 weeks increased plasma insulin concentration compared to C rats and this was effectively normalized by HT treatment (Table 3). H diet feeding decreased the tolerance to exogenous glucose (Figure 4A) and insulin (Figure 4B) and these effects were attenuated by HT treatment.

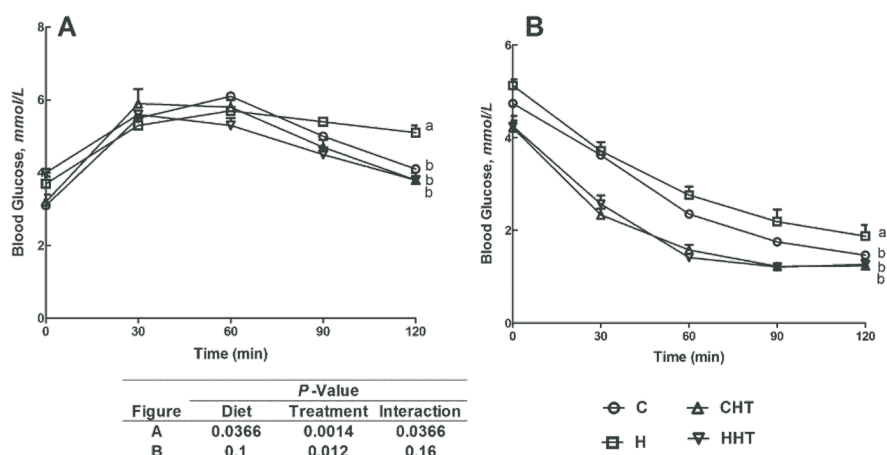


Figure 4: Glucose (2g/kg) (**A**) and insulin (0.75 IU/kg) (**B**) tolerance in cornstarch (C), cornstarch + hydroxytyrosol (CHT), high-carbohydrate, high-fat diet (H), and high-carbohydrate, high-fat diet + hydroxytyrosol (HHT) diet-fed groups. Data shown as means \pm SEM. Means of values at 120 min without a common alphabet significantly differ. n = 8-10/group.

Besides, it has been found that plasma markers of liver damage were reduced after the treatment with HT (Table 5). These results were accompanied by reduced infiltration of inflammatory cells into the LV and the liver with reduced collagen deposition and fat accumulation in the heart and the liver, respectively.

Finally, H diet feeding diminished endothelium-dependent relaxation to acetylcholine (Figure 5C), endothelium-independent relaxation to sodium nitroprusside (Figure 5B) and the α 1-adrenoceptor-mediated vascular contraction to noradrenaline (Figure 5A) in isolated thoracic aortic rings compared to C rats. HT treatment increased responses to acetylcholine (Figure 5C), sodium nitroprusside (Figure 5B) and noradrenaline in isolated thoracic aortic rings of HHT group (Table 4).

Variable	C	CHT	H	HHT	Diet <i>P</i> -Value	Treatment <i>P</i> -Value	Diet X Treatment <i>P</i> -Value
Food intake, g/d (<i>n</i> =10-12)	30.7±1.3 ^a	27.2±0.7 ^b	21.5±0.9 ^c	17.2±0.4 ^d	<0.0001	<0.0001	0.64
Water intake, ml/d (<i>n</i> =10-12)	26.0±3.1 ^a	20.6±0.7 ^{ab}	21.0±1.3 ^{ab}	17.9±0.8 ^b	0.0235	0.013	0.49
Energy intake, kJ/d (<i>n</i> =10-12)	344.1±14.9 ^b	300.9±7.3 ^c	451.8±22.9 ^a	346.6±8.7 ^b	<0.0001	<0.0001	0.0313
Feed conversion efficiency, % (<i>n</i> =10-12)	1.8±0.7 ^{bc}	0.5±0.2 ^c	6.7±0.7 ^a	2.9±0.3 ^b	<0.0001	<0.0001	0.0149
Body weight gain (8-16 weeks), % (<i>n</i> =10-12)	6.8±2.4 ^b	2.0±0.7 ^c	16.1±1.8 ^a	8.5±0.8 ^b	<0.0001	0.0002	0.35
Visceral adiposity index, % (<i>n</i> =9)	4.0±0.4 ^b	3.4±0.2 ^b	7.4±0.7 ^a	5.0±0.7 ^b	<0.0001	0.0094	0.11
Body mass index, kg/m ² (<i>n</i> =7-10)	5.1±0.1 ^b	5.1±0.1 ^b	6.2±0.2 ^a	5.2±0.2 ^b	0.0005	0.0027	0.0027
Bone mineral content, g (<i>n</i> =7-10)	12.9±0.3 ^b	12.6±0.3 ^b	15.2±0.6 ^a	15.2±0.5 ^a	<0.0001	0.74	0.74
Total body lean mass, g (<i>n</i> =7-10)	311.5±5.9 ^{ab}	336.0±6.7 ^a	326.5±9.9 ^{ab}	300.3±8.4 ^b	0.2	0.92	0.0032
Total body fat mass, g (<i>n</i> =7-10)	75.0±7.7 ^c	65.6±4.8 ^c	160.5±15.6 ^a	129.8±11.4 ^b	<0.0001	0.07	0.33
Abdominal adipose tissue [†] , mg/mm (<i>n</i> =9)							
Retroperitoneal	158.7±17.7 ^c	128.4±10.0 ^c	325.7±24.0 ^a	224.4±28.8 ^b	<0.0001	0.0042	0.11
Epididymal	102.1±9.2 ^{bc}	84.2±4.9 ^c	196.4±22.7 ^a	140.4±14.7 ^b	<0.0001	0.0158	0.2
Omental	82.2±15.4 ^b	60.1±7.4 ^b	174.4±21.8 ^a	120.2±23.0 ^b	0.0002	0.0419	0.38
Plasma total cholesterol, mmol/L (<i>n</i> =10)	1.6±0.1	1.7±0.1	1.9±0.2	1.8±0.1	0.14	1	0.45
Plasma triglyceride, mmol/L (<i>n</i> =10)	0.5±0.1	0.6±0.1	0.7±0.1	0.6±0.0	0.26	1	0.26
Plasma NEFA, mmol/L (<i>n</i> =9-10)	1.3±0.3	1.2±0.3	2.0±0.2	1.2±0.3	0.22	0.11	0.22
Plasma Insulin, μmol/L (<i>n</i> =8-10)	1.0±0.6 ^b	1.3±0.8 ^b	4.6±1.6 ^a	2.0±1.1 ^b	0.06	0.3	0.19

[†] Normalized against tibial length. Each value is a mean ± SEM. Number of repetitive experiments indicated within parenthesis.

Table 3: Dietary intakes, body composition and anthropometrics, organ wet weights, and plasma biochemistry in cornstarch (C), cornstarch + hydroxytyrosol (CHT), high-carbohydrate, high-fat diet (H), and high-carbohydrate, high-fat diet + hydroxytyrosol (HHT) diet-fed groups.

Variable	C	CHT	H	HHT	Diet <i>P</i> -Value	Treatment <i>P</i> -Value	Diet X Treatment <i>P</i> -Value
Heart rate, <i>bpm</i> (<i>n</i> =10)	239±18 ^b	350±21 ^a	235±17 ^b	262±12 ^b	0.0117	0.0003	0.0204
IVSd, <i>mm</i> (<i>n</i> =10)	1.8±0.1	2.0±0.0	1.8±0.1	1.8±0.1	0.26	0.26	0.26
LVIDd, <i>mm</i> (<i>n</i> =10)	7.1±0.3 ^b	8.0±0.2 ^a	8.1±0.2 ^a	8.1±0.2 ^a	0.0217	0.06	0.06
LVPWd, <i>mm</i> (<i>n</i> =10)	1.8±0.1	1.9±0.0	1.8±0.1	1.7±0.1	0.26	1	0.26
IVSs, <i>mm</i> (<i>n</i> =10)	3.0±0.1 ^{ab}	3.3±0.1 ^a	2.9±0.1 ^b	3.0±0.1 ^{ab}	0.05	0.05	0.32
LVIDs, <i>mm</i> (<i>n</i> =10)	3.8±0.2 ^b	4.0±0.2 ^b	4.4±0.2 ^b	4.7±0.3 ^a	0.0074	0.28	0.83
LVPWs, <i>mm</i> (<i>n</i> =10)	2.8±0.1	3.0±0.1	2.8±0.1	2.7±0.1	0.14	0.62	0.14
Fractional shortening, % (<i>n</i> =10)	55.5±2.4 ^a	49.7±1.5 ^{ab}	44.5±1.5 ^b	50.4±1.4 ^{ab}	0.0056	0.98	0.0019
Ejection fraction, % (<i>n</i> =10)	83.3±2.2	87.0±1.2	83.3±1.4	78.4±4.4	0.11	0.82	0.11
Ejection time, <i>msec</i> (<i>n</i> =10)	91.7±2.3	87.5±2.9	86.2±9.8	91.2±3.0	0.87	0.94	0.4
Diastolic volume, μ L (<i>n</i> =10)	383.0±43.0 ^b	544.0±41.8 ^a	567.4±43.4 ^a	574.5±43.0 ^a	0.0167	0.06	0.08
Systolic volume, μ L (<i>n</i> =10)	64.7±10.3 ^b	70.9±8.2 ^b	94.0±10.0 ^b	118.0±19.6 ^a	0.0051	0.25	0.49
Stroke volume, μ L (<i>n</i> =10)	318.3±37.1 ^b	473.1±37.8 ^a	473.4±38.3 ^a	456.4±47.6 ^a	0.1	0.1	0.0405

Variable	C	CHT	H	HHT	Diet <i>P</i> -Value	Treatment <i>P</i> -Value	Diet X Treatment <i>P</i> -Value
Cardiac output <i>mL</i> (<i>n</i> =10)	76.0±11.4 ^b	164.5±14.3 ^a	111.3±12.7 ^b	120.2±16.2 ^b	0.75	0.0011	0.0065
Estimated LV Mass, <i>g</i> (<i>n</i> =10))	0.9±0.1	1.1±0.0	1.0±0.1	1.0±0.0	1	0.17	0.17
LV + septum wet weight [†] , <i>mg/mm</i> (<i>n</i> =9)	16.5±0.9 ^{ab}	15.2±0.8 ^b	18.2±0.7 ^a	15.0±0.6 ^b	0.33	0.0056	0.22
Relative wall thickness, (<i>n</i> =10)	0.5±0.01 ^a	0.5±0.02 ^a	0.5±0.01 ^a	0.4±0.02 ^b	0.0032	0.0032	0.0032
Systolic wall stress, <i>mmHg</i> (<i>n</i> =10)	81.2±5.4 ^b	79.0±5.2 ^b	123.6±6.4 ^a	123.3±11.4 ^a	<0.0001	0.87	0.9
Systolic blood pressure, <i>mmHg</i> (<i>n</i> =10)	117.9±0.5 ^c	115.9±0.7 ^c	157.3±1.7 ^a	137.9±2.4 ^b	<0.0001	<0.0001	<0.0001
LV developed pressure, <i>mmHg</i> (<i>n</i> =7)	71.5±4.9 ^a	66.3±5.7 ^a	44.4±3.4 ^b	57.4±3.6 ^{ab}	0.0545	0.3948	0.0005
+dP/dt, <i>mmHg.s⁻¹</i> (<i>n</i> =7)	1164.3±98.2 ^a	1201.5±101.6 ^a	806.4±76.0 ^b	1083.1±58.4 ^a	0.1735	0.0784	0.0102
-dP/dt, <i>mmHg.s⁻¹</i> (<i>n</i> =7)	-662.9±61.9 ^b	-709.0±78.6 ^b	-426.3±51.1 ^a	-658.7±41.7 ^b	0.1331	0.0289	0.0248
Diastolic stiffness (k), (<i>n</i> =7-9)	23.4±0.6 ^b	24.5±0.6 ^b	28.8±1.1 ^a	25.6±1.1 ^b	0.0276	0.266	0.0015

[†] Normalised against tibial length. Each value is a mean ± SEM. Number of repetitive experiments indicated within parenthesis.

Table 4: Changes in cardiovascular structure and function in cornstarch (C), cornstarch + hydroxytyrosol (CHT), high-carbohydrate, high-fat diet (H), and high-carbohydrate, high-fat diet + hydroxytyrosol (HHT) diet-fed group.

Variable	C	CHT	H	HHT	Diet <i>P</i> -Value	Treatment <i>P</i> -Value	Diet X Treatment <i>P</i> -Value
Liver wet weight [†] , <i>mg/mm</i> (<i>n</i> =9)	236.0±15.7	252.0±12.7	284.2±13.8	278.7±9.0	0.0071	0.69	0.42
Plasma ALT, <i>U/L</i> (<i>n</i> =10)	33.5±3.7 ^b	39.7±5.0 ^b	56.1±6.8 ^a	39.0±3.9 ^b	0.0352	0.28	0.0256
Plasma AST, <i>U/L</i> (<i>n</i> =10)	69.6±5.4 ^b	60.5±6.5 ^b	114.7±18.2 ^a	76.3±6.0 ^b	0.0062	0.0294	0.17
Plasma LDH, <i>U/L</i> (<i>n</i> =10)	240.3±40.6	210.6±33.9	341.7±59.6	428.3±143.5	0.06	0.73	0.48
Plasma ALP, <i>U/L</i> (<i>n</i> =10)	142.1±17.0	152.0±11.9	177.5±11.4	156.9±5.8	0.11	0.66	0.22
Bilirubin, <i>μmol/L</i> (<i>n</i> =9-10)	2.1±0.1	2.1±0.2	2.5±0.2	2.1±0.2	0.27	0.27	0.27
Albumin, <i>g/L</i> (<i>n</i> =10)	28.3±0.5 ^b	30.5±0.4 ^a	28.1±0.3 ^b	28.9±0.4 ^b	0.0331	0.0007	0.09
Globulin, <i>g/L</i> (<i>n</i> =10)	27.1±1.2	27.4±0.6	28.4±0.7	27.9±0.6	0.28	0.9	0.63

[†] Normalised against tibial length. Each value is a mean ± SEM. Number of repetitive experiments indicated within parenthesis.

Table 5: Changes in hepatic structure and function in cornstarch (C), cornstarch + hydroxytyrosol (CHT), high-carbohydrate, high-fat diet (H), and high-carbohydrate, high-fat diet + hydroxytyrosol (HHT) diet-fed group.

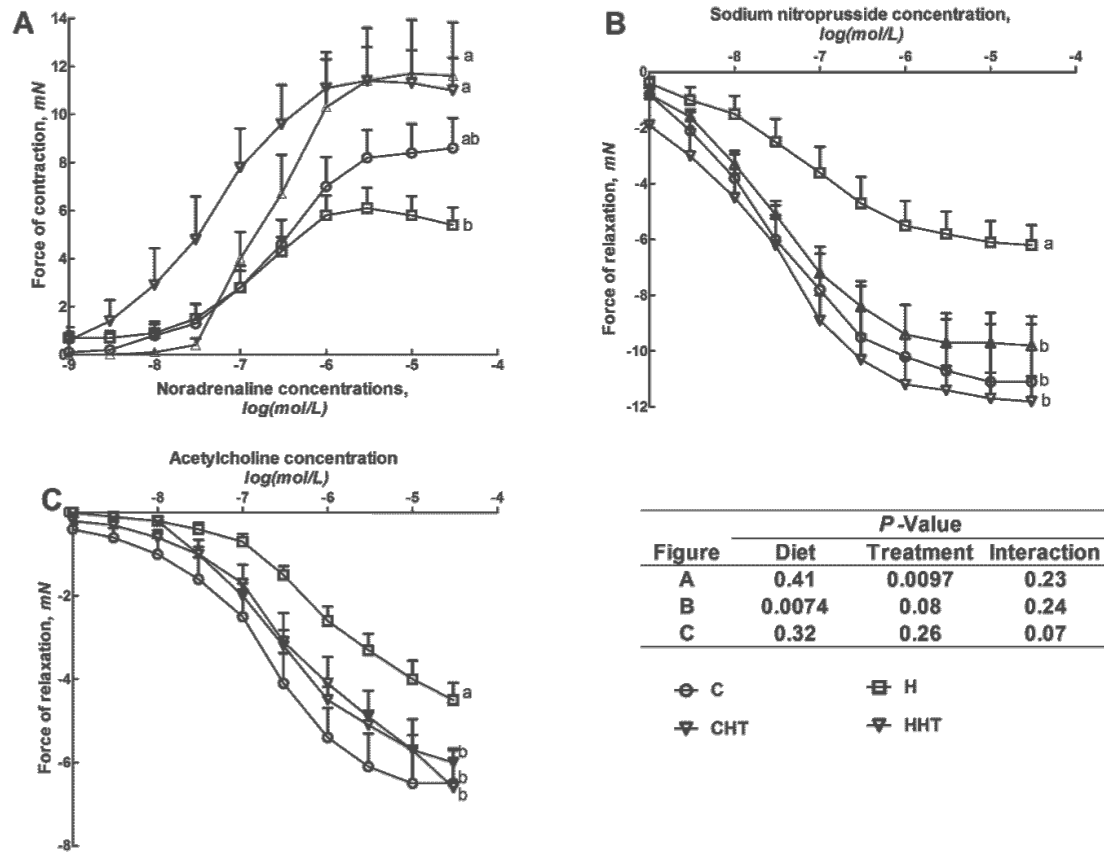


Figure 5: Cumulative concentration-response curves for noradrenaline (A), sodium nitroprusside (B) and acetylcholine (C) in thoracic aortic rings from cornstarch (C), cornstarch + hydroxytyrosol (CHT), high-carbohydrate, high-fat diet (H), and high-carbohydrate, high-fat diet + hydroxytyrosol (HHT) diet-fed groups. Data shown as means \pm SEM. Endpoint means without a common alphabet in each data set significantly differ. n = 8-10/group.

Orbitrap-based metabolomics

In order to set some more light to the underline biochemical mechanisms that invoked the aforementioned pharmacological effects, a metabolomics study has been performed. Using the optimal LC-HRMS (orbitrap) conditions described above, the representative base peak intensity (BPI) chromatograms of rat plasma samples from the HHT group collected in positive and negative ESI modes are presented in Figure 6. The metabolites could be separated well due to the small diameter particles (1.7 μm) of the UHPLC column.

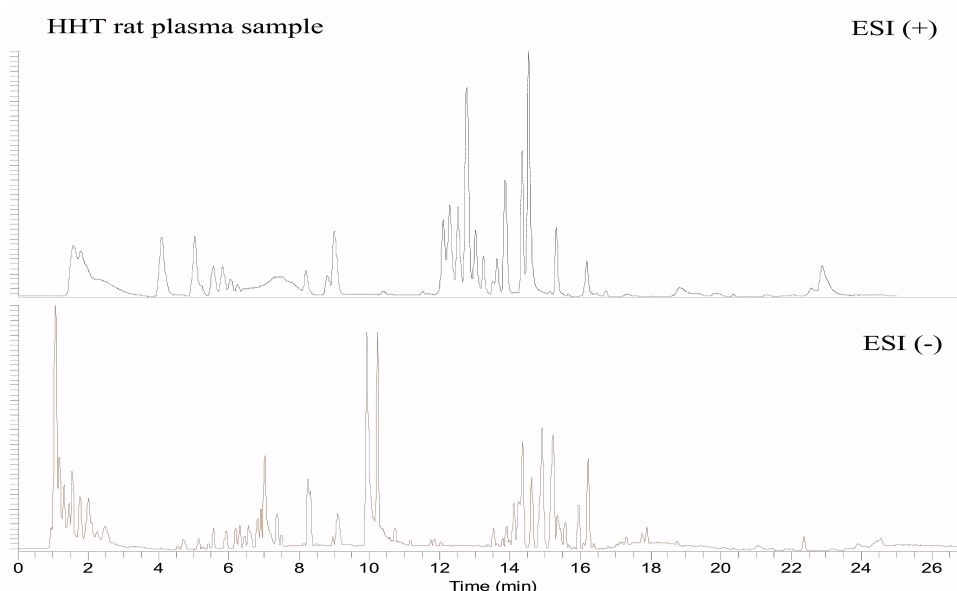


Figure 6: A BPI UHPLC–HRMS (orbitrap) chromatogram of a HHT plasma sample run in positive ESI mode (top) and negative ESI mode (below). Positive and negative ESI modes can offer complementary information regarding the metabolite profile.

Data preprocessing. A crucial parameter in the data analysis workflow is the choice of the normalization methodology. For this purpose, a software generated pseudo pooled sample from the control group (H group) has been used for the raw-wise normalization process and the distribution of the intensity regarding the 50 first features before and after the normalization is depicted in Figure 7.

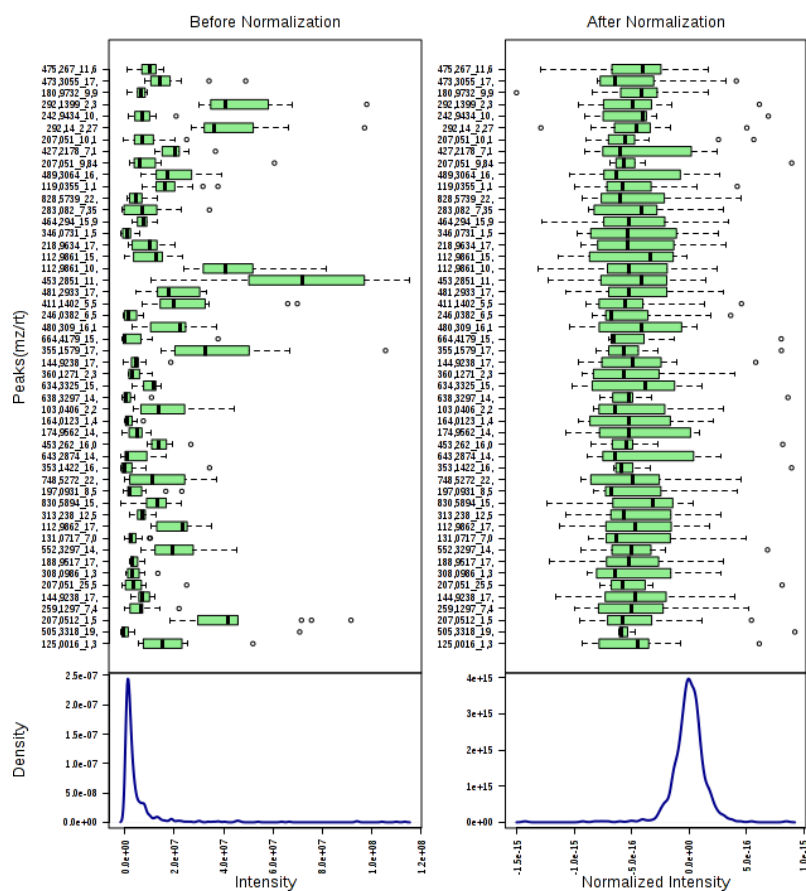


Figure 7: Data normalization view. The graph summarizes the distribution of input data values before and after normalization. The box plots on the top show the concentration distributions of individual compounds, whereas the bottom plots show the overall concentration distribution based on kernel density estimation.

Furthermore, two methodologies have been followed in terms of data scaling, namely the UV scaling (auto-scaling) and the Pareto scaling. The first one forces all metabolites to become equal important but on the hand it enhances the baseline noise, whereas the second one stays closer to the original measurements but on the other hand is sensitive to large changes⁵⁹. Evaluation of the generated results following the two aforementioned methodologies, showed that the UV scaling exhibited PCA clustering in some extent between the H and HHT groups (Figure 9) compared to the Pareto scaling. Furthermore, the PLS-DA cross validation showed better Q^2 values as

well as reduced number of misclassifications (Figure 8). R^2 values were found to be similar in both cases. Overall, the UV scaling has been used throughout all the data processing employing 5 components as the optimum value.

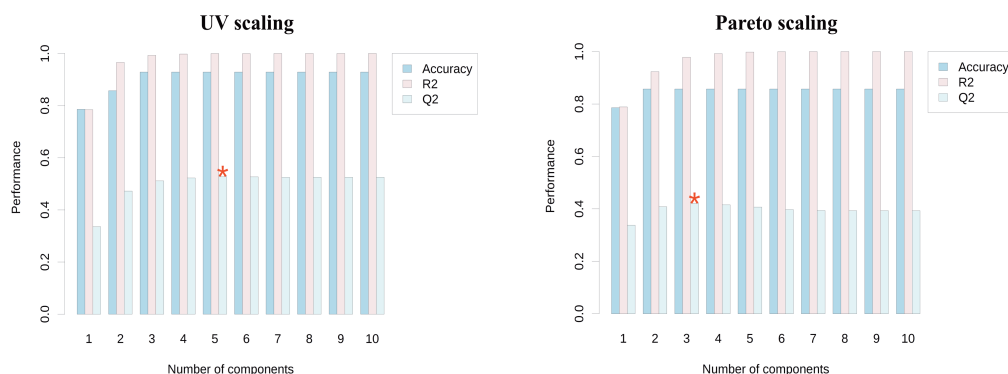


Figure 8: Multivariate analysis of the data obtained from the UHPLC(-)ESI-HRMS analysis of the H and HHT rat plasma samples. PLS-DA cross validation after UV scaling (left) and after Pareto scaling (right). Bar plots showing the three performance measures (prediction accuracy, R^2 and Q^2) using different number of components. The red ‘*’ indicates the best values of the currently selected measures (Q^2).

Chemometric data analysis. Various uni- and multi-variate analysis methodologies such as t-test, fold change analysis and PCA, PLS-DA and OPLS-DA respectively, were employed after the initial data preprocessing described above.

The PCA methodology was initially used to explore possible clustering and detect putative outliers but also to investigate the clustering of the QC samples in the corresponding analytical runs, providing a means of quality control assessment regarding the process. As the current intervention study concerns the administration of HT in obese rats with the aid to examine its impact to the MetS, the statistical analysis focused to the comparison of the H and HHT groups. For the PCA, various parameters have been used, such as different scaling e.g. uv or Pareto, and different combinations of Q^2 and R^2 in order to assess the number of factors that described as

much variance as possible without increasing the complexity of the model excessively. The uv scaling method has been finally selected as it afforded tighter groups with larger separation between them (Figure 9).

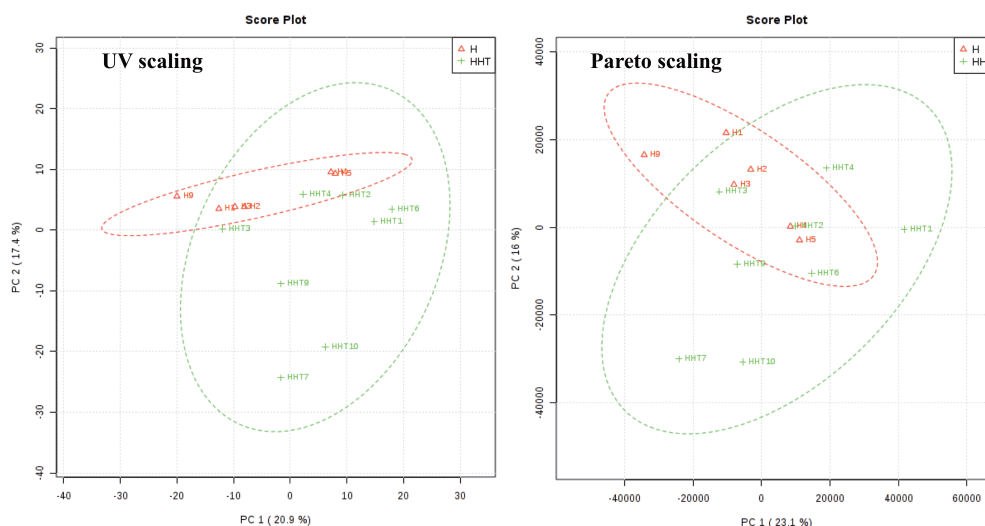


Figure 9: Multivariate analysis of the data obtained from UHPLC-(-)ESI-HRMS (orbitrap) analysis of the H and HHT rat plasma samples. PCA scores plots of the rat plasma samples obtained after UV scaling (left) and Pareto scaling (right). PCA analysis after UV scaling shows clearer grouping of the samples compared to PCA after Pareto scaling.

In order to keep the simplest PCA model, the selection of the pc's terminated to the fifth pc, which could explain more than 65% of the total variance whereas the fifth pc itself explained 7.5% of the variance. This has been verified by the scree plot, which showed the inflection after the fifth pc (Figure 10). Incorporation of more pc's led to minor improvement of the model and thus the five first pc's have been taken into account.

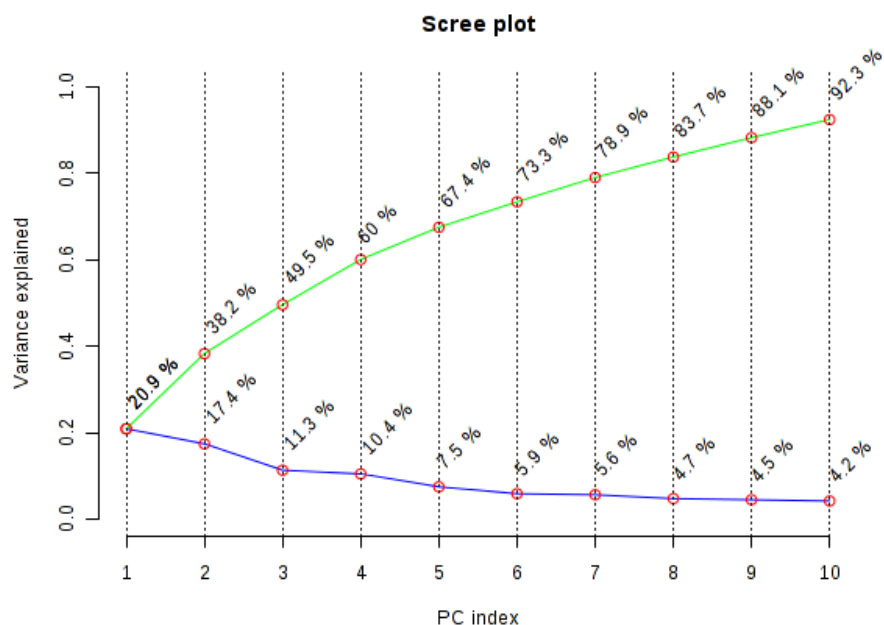


Figure 10: PCA scree plot after UV scaling obtained from UHPLC(-)ESI-HRMS analysis of the H and HHT rat plasma samples. The first five pcs have been used for further evaluation.

After evaluation of the PCA models, no outliers were detected in both datasets (positive and negative ESI) as can be found in the corresponding PCA scores plots Figure 11. Furthermore, some degree of clustering between the examined groups in the scores plot was achieved, as evident in the left part of Figure 9, but it was considered as non-satisfactory. Finally, with the aid of the PCA approach, the overall UHPLC-ESI-HRMS analyses were deemed to be reliable for downstream statistical evaluation since the QC samples were clustered tightly together in the scores plots (Figure 11).

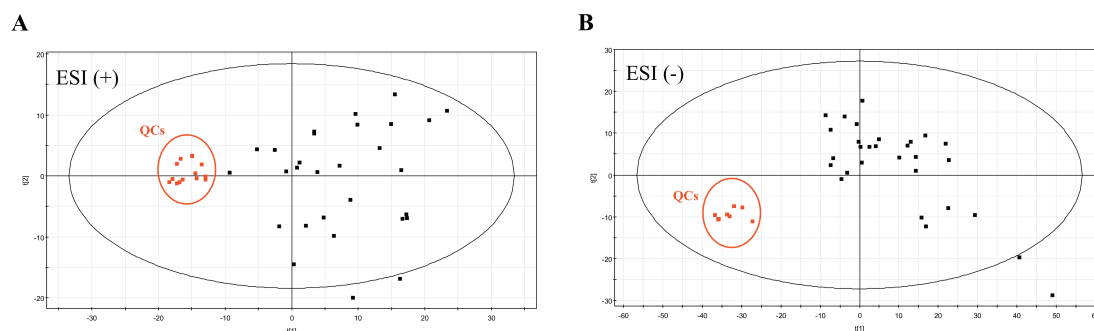


Figure 11: PCA scores plots of the rat plasma samples (C, CHT, H and HHT samples) after UV scaling obtained from UHPLC-HRMS (orbitrap) analysis in **(A)** ESI (+) and **(B)** ESI (-). No outliers were detected in both datasets. Black boxes represent the tested samples. Red boxes (QC samples) are clustered tightly together in both datasets, proving the validity of the analysis.

Unsupervised hierarchical clustering analysis of significant metabolites ($n=30$) that distinguish H vs. HHT subjects ($FC>2$, $p<0.05$, $FDR <10\%$) was visualized in a heatmap. In the heatmap representation of the most significant metabolomic data, the distinct segregation between the H and HHT groups is clearly shown, where the correlations of the features (columns) with the samples (rows) is shown in Figure 12. Shades of brown and blue represent an increase and decrease of a metabolite (see color scale).

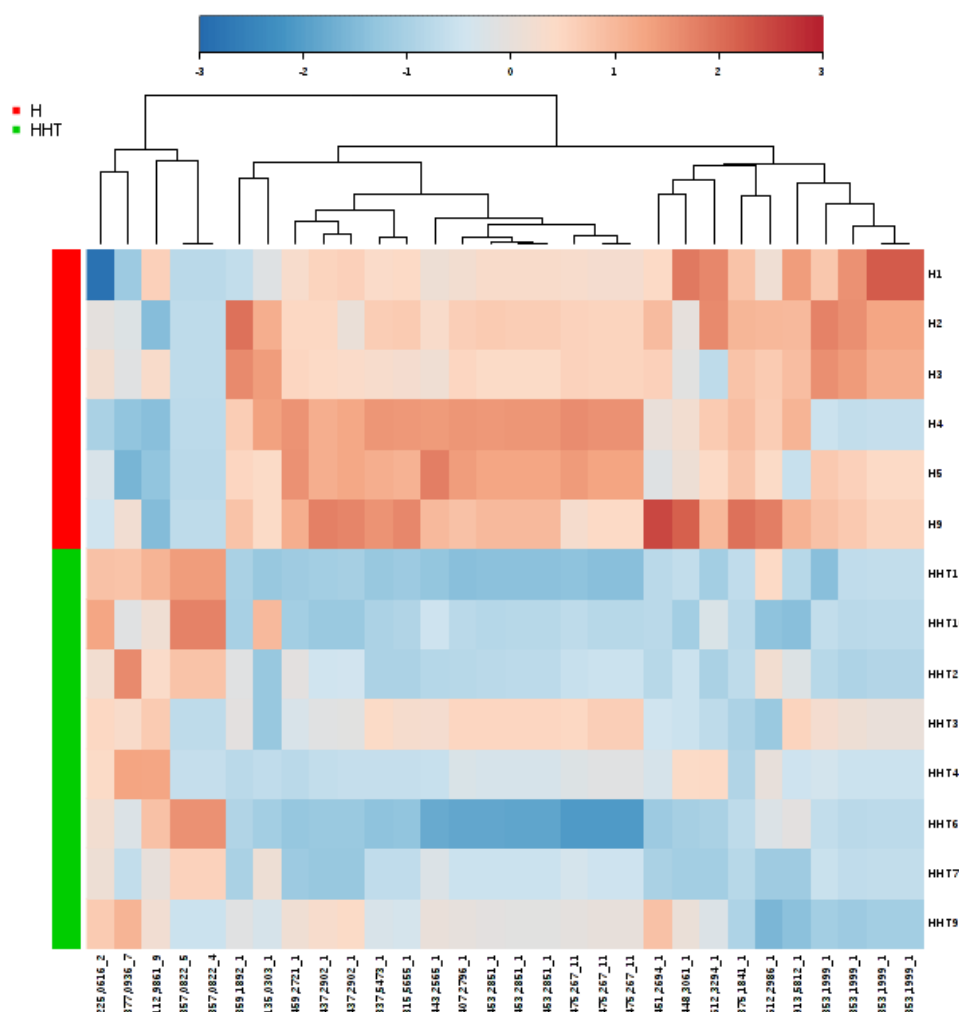


Figure 12: Heatmap visualization of the intervention effects of HT in obese rats (H vs. HHT group). Rows: samples from H and HHT groups; Columns: the first 30 significant features; Color key indicates metabolite expression value, blue: lowest, dark red: highest. The differences in the concentrations of the selected features are evident between the two groups. The dendrogram represent the unsupervised hierarchical clustering of the significant features (n=30) (FC>2, $p < 0.05$, FDR <10%).

Generally, unsupervised multivariate analysis is the method of choice for treating high dimensionality data, as they require no intervention from the analyst, thus minimizing any imposed bias. Due to the fact that PCA-based approaches frequently fail to produce any meaningful or appropriate clustering (as in the current case) and thus alternatively supervised methodologies, such as PLS-DA or OPLS-DA, are employed. In such supervised models, the user assigns the samples into predefined groups, which creates the risk of either “artificially”-biased clustering or overfitting.

In the current case, using the PLS-DA approach clear grouping between the H and HHT groups has been observed, as it can be clearly shown in the corresponding scores plot (Figure 13) obtained by ESI (-) analysis. The results show that the samples of the H group (red open triangles) are clustered separately from the samples of the HHT group (green crosses).

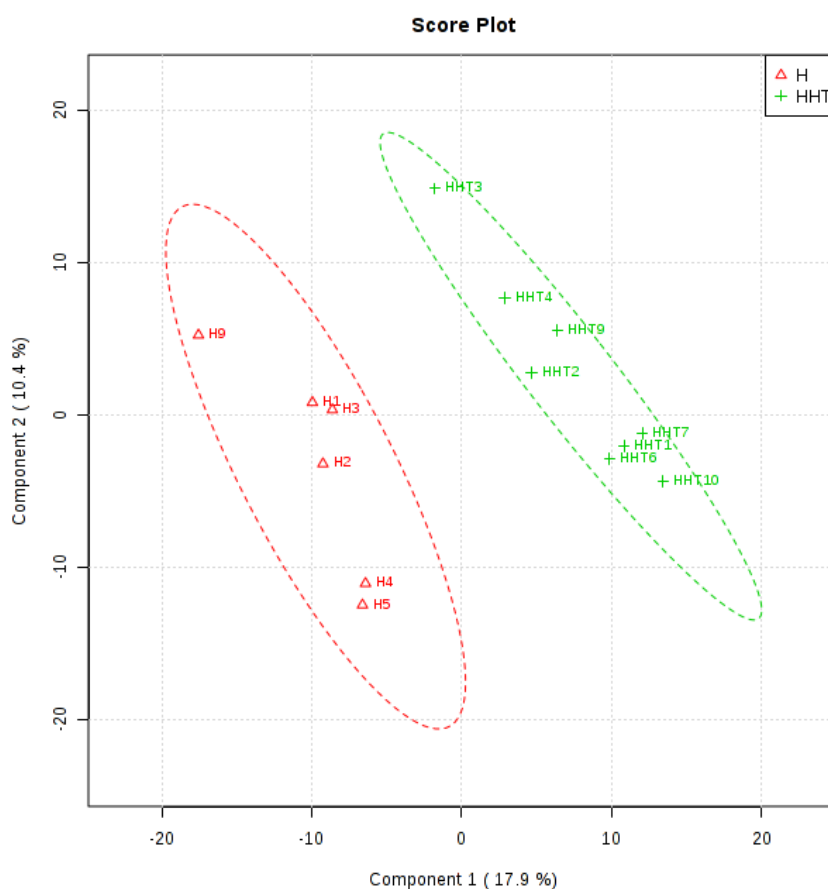


Figure 13: PLS-DA scores plot of the animals from H and HHT groups, obtained by UHPLC(-)ESI-HRMS (orbitrap). A clear separation between the samples of the H group (red open triangles) and the samples of the HHT group (green crosses) is shown.

It should be stressed out that PLS-DA “is eager to please”; hence thorough cross validation procedures should be followed in order to secure the validity of the model and verify the acceptability of the clustering. R^2 is defined as the proportion of variance in the data explained by the model and indicates the goodness of fit, whereas

Q^2 is defined as the proportion of variance in the data, which could be correctly predicted by the model. In the case of the ideal model, R^2 and Q^2 values should reach 1. In the current case, the R^2 , Q^2 and accuracy values obtained from 10-fold cross validation were 0.9996, 0.5511 and 0.9286 respectively, which indicates the predictive ability of the model, whereas the model's robustness, indicated by the respective R^2 , Q^2 and accuracy values obtained from LOOCV, was acceptable (0.9987, 0.5485 and 0.9167 for the R^2 , Q^2 and accuracy values respectively). Permutation testing used in order to assess the significance of the classification, was performed allowing for 100 permutations, employing the prediction accuracy during training statistic test. The histogram of misclassification distribution can be found in Figure 14, where the p-value was <0.01 , indicating high significance of classification. Overall, the abovementioned PLS-DA model can be considered as valid. It should be noted that similar results and trends were observed for the (+) ESI mode.

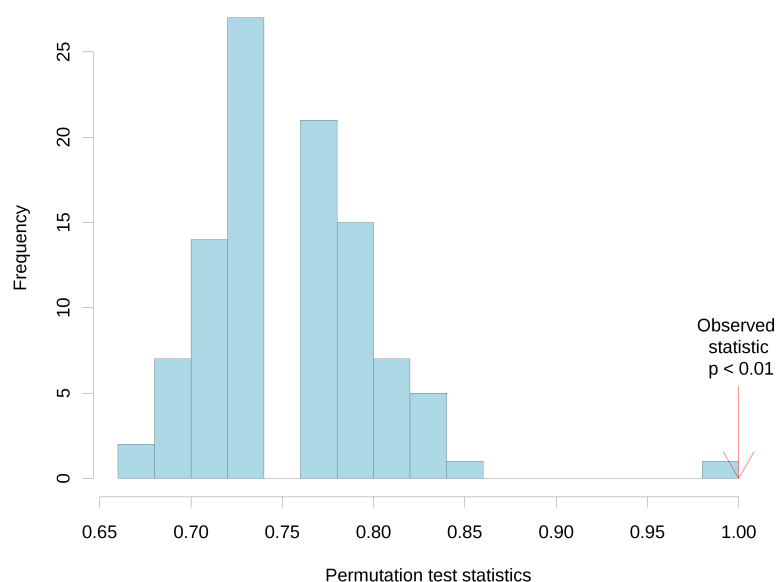


Figure 14: A histogram of misclassification distribution (100 permutations, p -value <0.01) of the PLS-DA model obtained by UHPLC-(-)ESI-HRMS analysis of the H and HHT samples. The results of the permutation testing indicate the high significance of classification.

In the next step, in order to discover the metabolites that are responsible for the classification between the H and HHT groups, the VIP values have been considered, as it represents a weighted sum of squares of the PLS loadings that takes into account the amount of explained Y-variance of each component. As an example, the ten first-VIP variables from the first component can be visualized in Figure 15, where the m/z_{t_R} values is correlated with VIP values in a scatter plot diagram. The red-green coding stands for the up and down-regulated metabolites respectively in the H and HHT groups, as it can be found in the embedded legend. Hence, all the VIP variables for each component, from the two data sets (positive and negative ESI modes), were accordingly selected and submitted to the identification process.

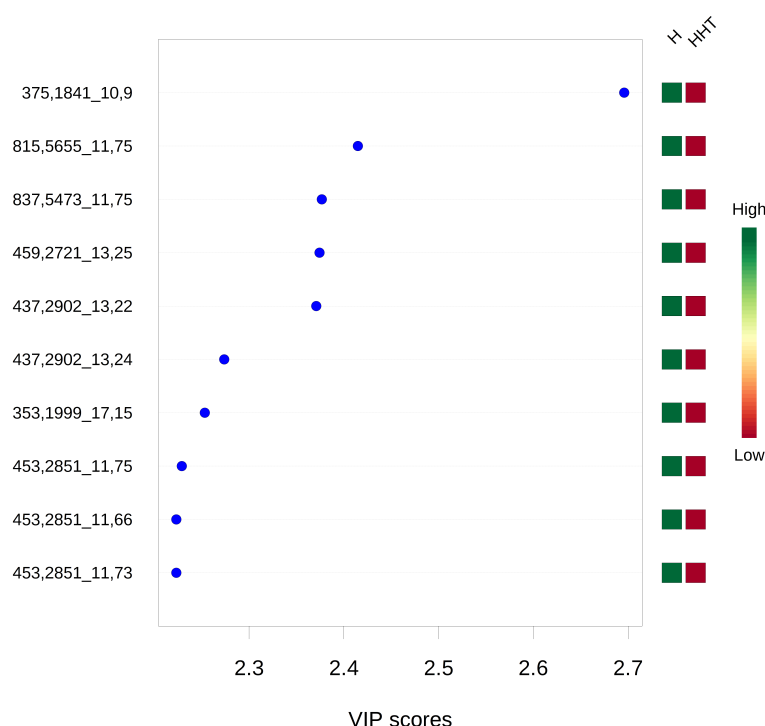


Figure 15: Top 10 variables ranked by VIP scores from the PLS-DA obtained by UHPLC(-)ESI-HRMS analysis of the H and HHT samples. Red and green indicate increased and decreased levels of the corresponding significant features respectively between the two groups.

Following, the supervised OPLS-DA models were used to enhance biomarker discovery efforts and to obtain better discrimination between the control and treated groups. Specifically, the OPLS-DA for the H and HHT groups revealed the differences in their metabolic profiles, suggesting the metabolic profiles have significantly changed as a result of HT administration in the obese rats. This can be clearly shown in the corresponding OPLS-DA scores plot obtained by (+) ESI (Figure 16A) and (-) ESI respectively (Figure 17A), where the H group (red open triangles) is clustered separately from the HHT group (green crosses).

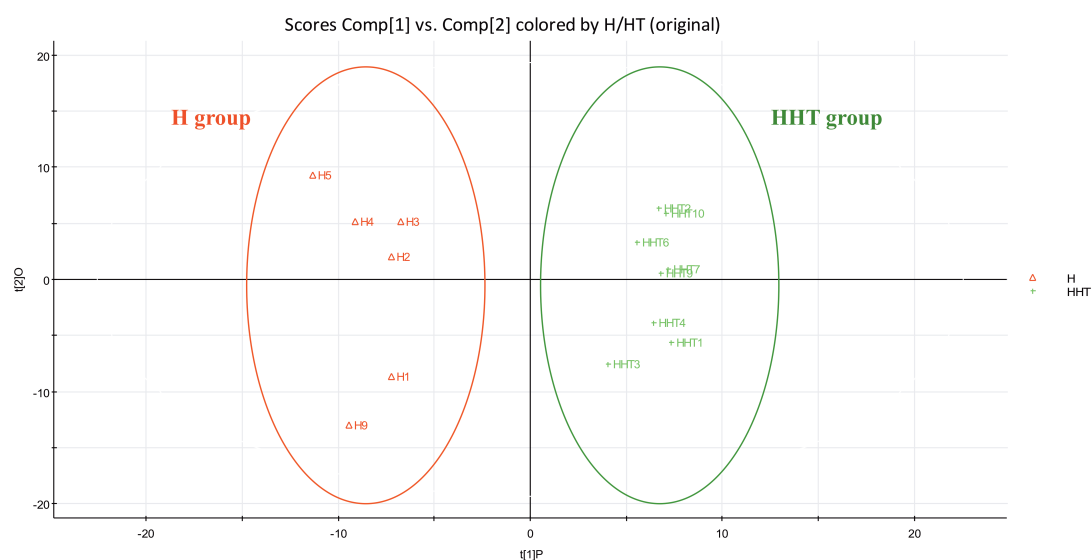
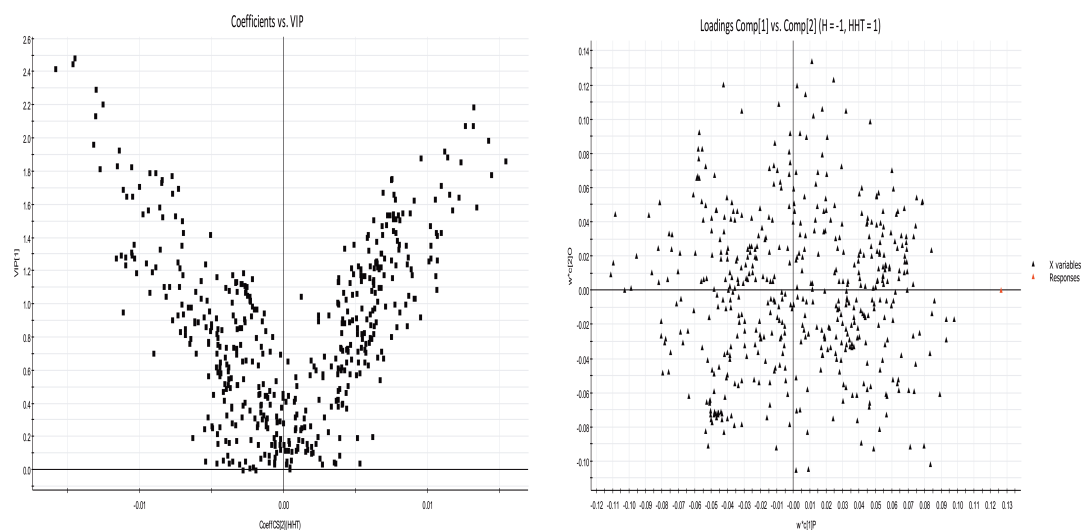
A**B**

Figure 16: OPLS-DA model results for H and HHT groups obtained by UHPLC-(+)ESI-HRMS (orbitrap). Significant metabolomic difference is shown between the H group (orange open triangles) and the HHT group (green crosses) in the OPLS-DA scores plot (A). One data point stands for one subject. (B). VIP-plot (left panel) and loadings plot (right panel) of OPLS-DA. The OPLS-DA loadings plot indicates variables positively which correlates with the corresponding subjects in the scores plots (A). In the OPLS-DA loadings plot the distance of each point from the origin represents the impact of the significantly different metabolites on the clustering results. In the VIP-plot, the further away from the origin is a black point, the higher the VIP value of the metabolites. Furthermore, the two clusters represents the two groups (H vs. HHT group).

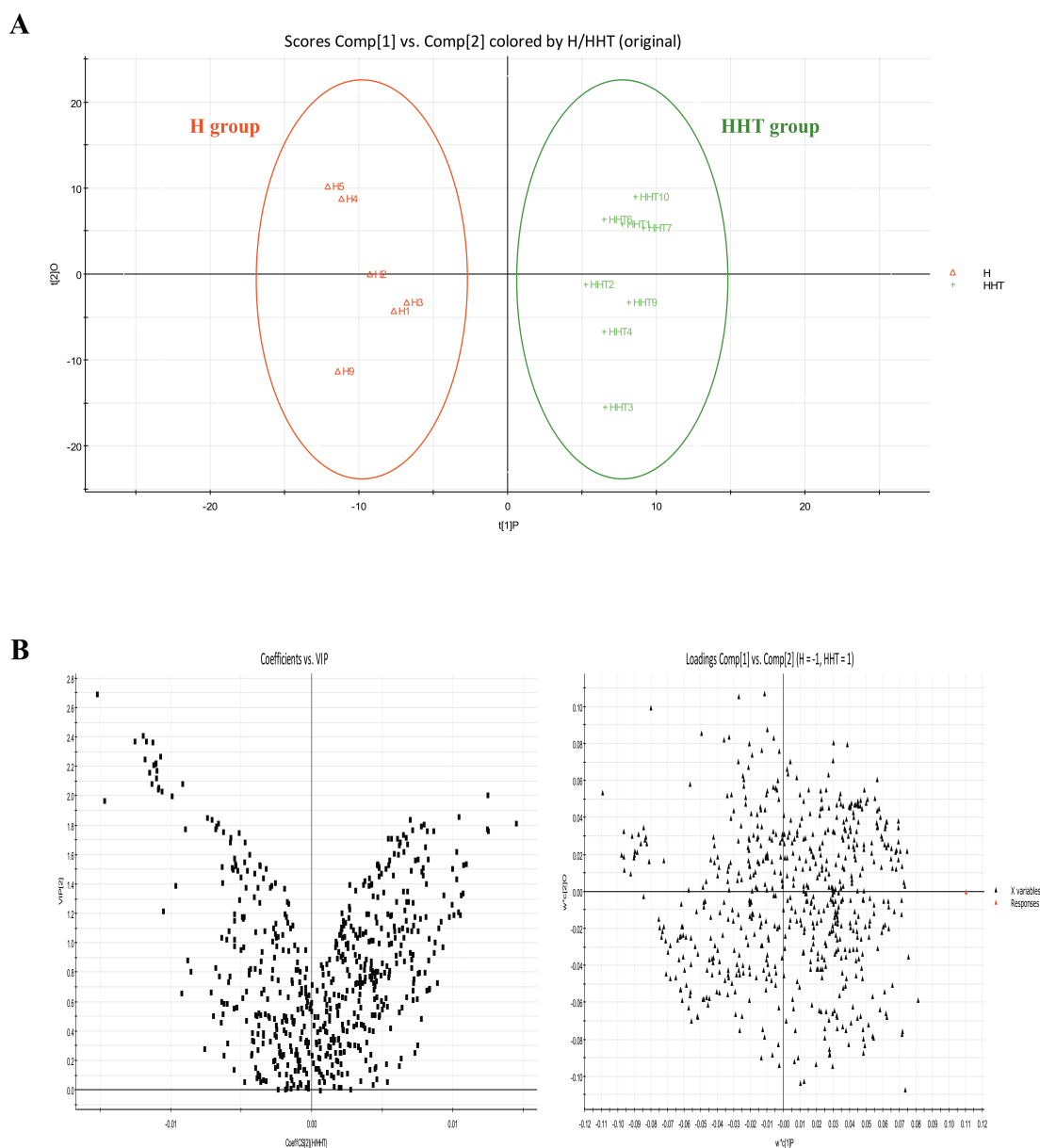


Figure 17: OPLS-DA model results for H and HHT groups obtained by UHPLC-(-)ESI-HRMS (orbitrap). Significant metabolomic difference is shown between the H group (orange open triangles) and the HHT group (green crosses) in the OPLS-DA scores plot (A). One data point stands for one subject. (B). VIP-plot (left panel) and loadings plot (right panel) of OPLS-DA. The OPLS-DA loadings plot displayed variables positively correlated with score plots (A). The loadings plot represents the impact of the significantly different metabolites on the clustering results. The ions most responsible for the variance are indicated on the loadings plot by their distance from the origin. In the VIP-plot, the farther away from the origin is a black point, the higher the VIP value of the metabolites. Furthermore, the two clusters represents the two groups (H vs. HHT group).

The features that showed significant difference ($-0.01 < p(\text{corr})$ or $p(\text{corr}) > 0.01$), in terms of abundance between the H and HHT groups, and contributed to the observed clustering were selected from the respective loadings plots and VIP plots of OPLS-DA as potential markers in (+) and (-) ESI modes respectively (Figure 16B and Figure 17B). R^2 and Q^2 values were found to be 0.970 and 0.686 respectively for the OPLS-DA model obtained from the UHPLC-(+)ESI-HRMS analysis whereas for the respective model obtained from the UHPLC-(-)ESI-HRMS analysis the R^2 and Q^2 values were found to be 0.965 and 0.635 respectively. These values showed excellent goodness of fit and significant prediction ability of the OPLS-DA models.

Furthermore, in order to confirm the produced results concerning the selection of the VIP's from both PLS-DA and OPLS-DA models, univariate data analysis has been performed using SAM, t-test, FC monitoring and their combination in the form of the volcano plot. In other words, statistical hypothesis testing has been used in order to assess the difference of the selected metabolites level between the H and HHT groups.

SAM is a popular method employed in microarray data analysis⁶⁰ and is advised for a small sample size⁵⁸. It is designed to address false discovery rate (FDR) when running multiple tests on high-dimensional data. SAM assigns a significance score to each variable based on its change relative to the standard deviation of repeated measurements. For variable with scores greater than an adjustable threshold, its relative difference is compared to the distribution estimated by random permutations of the class labels. For each threshold, a certain proportion of the variables in the permutation set will be found to be significant by chance. The number is used to calculate the FDR. The delta value, one of the most crucial parameters in SAM, was set to control the FDR and selection of important features. Figure 18A shows the

relation between delta and FDR as well as the relation between delta and the number of significant compounds. In order to minimize FDR, the delta value was set to 0.5, giving an FDR value of 0.034, which show the best performance. Thus, the selected variables, which are shown in green circles (Figure 18B), were further evaluated as candidates of potential biomarkers.

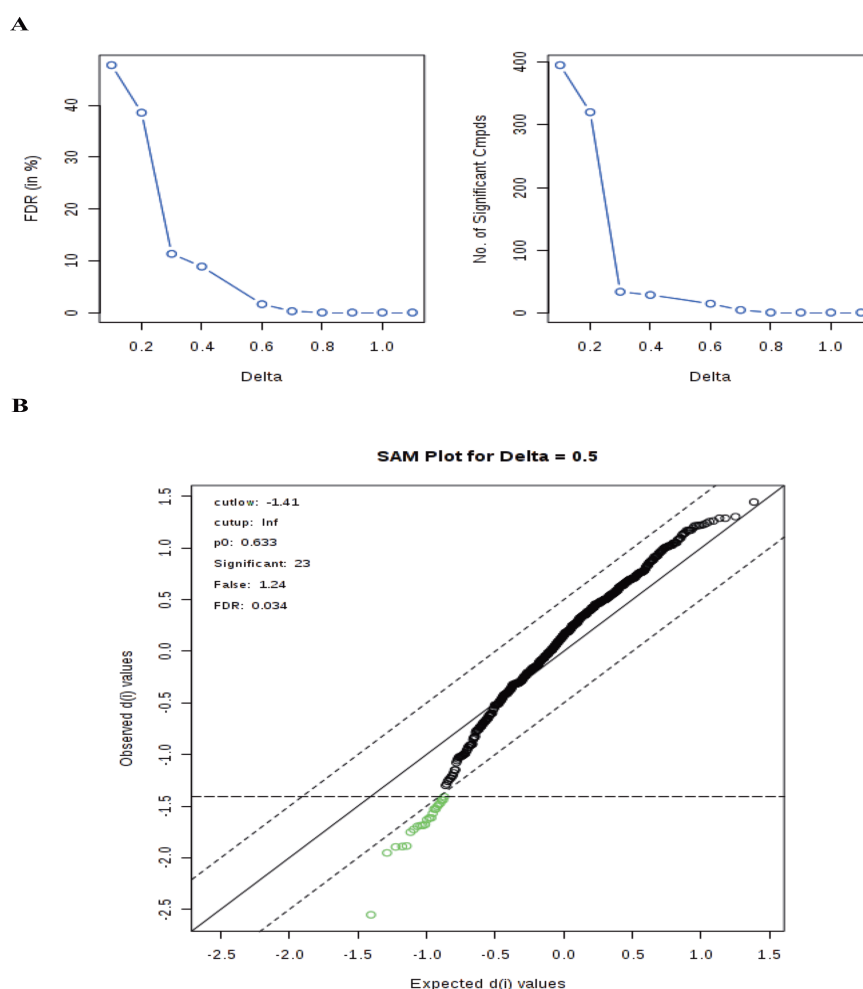


Figure 18: Identification of metabolites with significant changes in expression by SAM. **(A)** The relation between delta, FDR, and the number of significant compounds. The left plot shows delta versus the FDR and the right plot shows delta versus the number of identified compounds. **(B)** The result of identified variables by SAM with a delta value of 0.5. The SAM plot is a scatter plot of the observed relative difference versus the expected relative difference estimated by data permutation. The solid diagonal line indicates where these two measures are the same. The dotted lines are drawn at a distance of delta from the solid line. The significant variables are highlighted in green.

FC analysis is used to compare the absolute value change between two group means. Since column-wise normalization in the preprocessing step significantly changes the absolute values, FC is calculated as the ratio between the two group means using data before column-wise normalization was applied. A volcano plot is the combination of t-test and FC analysis. For the construction of the diagram, it should be noted that the FC values are log transformed (x-axis) in order for both up and down-regulated features to appear symmetrically whereas for the t-test, the y-axis is $-\log_{10}(\text{p values})$. The variables which had a FC threshold value $0.5 > \text{FC} > 2$ and a p-value < 0.05 in the volcano plot (Figure 19) were selected as significantly differentially expressed and submitted to the identification process.

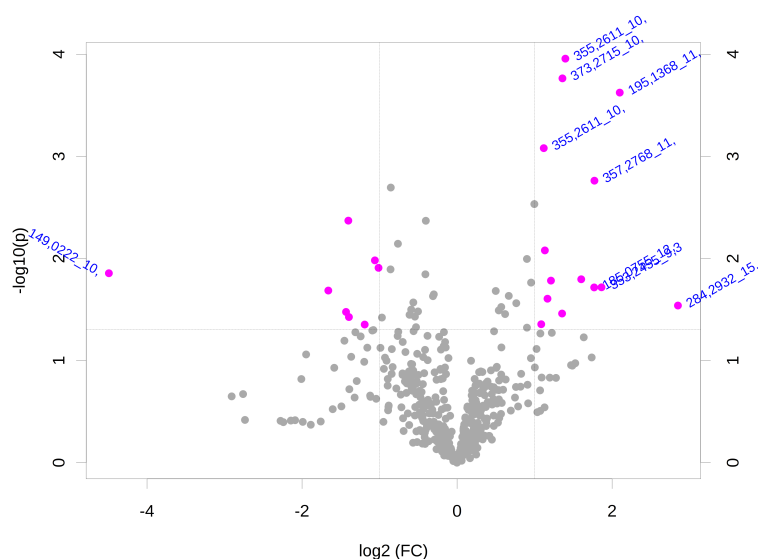


Figure 17: Volcano plot obtained by UHPLC-(+)ESI-HRMS analysis of the H and HHT samples. The variables with an FC threshold value $0.5 > \text{FC} > 2$ and a p-value < 0.05 in the volcano plot (light purple points) were selected as significantly differentially expressed and were sequentially submitted to the identification process.

Combining the aforementioned results from the univariate and multivariate analysis described obtained by both (+) and (-) ESI modes, a list of candidate

metabolites were selected for the structural identification of the biomarkers. The features selected are summarized in Table 6 as a. metabolite id., b. m/z feature, c. t^R (min), d. experimental m/z , e. theoretical m/z , f. Δm (ppm), g. corresponding RDB value, h. possible identification name, i. possible molecular formula, j. corresponding monoisotopic exact mass and k. change trend in HHT group compared with H group (up or down regulation) and l. corresponding dataset (positive or negative ESI dataset). Furthermore, an analysis for possible adducts (M-H₂O-H, M-H, M+Na-2H, M+Cl, M+K-2H, M+FA-H, M+Br, 2M-H, 2M+FA-H) has been performed for all the features in order to verify the possibility of occurrence of chromatographic peaks due to adduct formation. These features can also be identified as present or not in the raw data.

Metabolite id.	<i>m/z</i>	<i>t_R</i> (min)	Experimental <i>m/z</i>	Theoretical <i>m/z</i>	Δm (ppm)	RDB	Name	Molecular Formula	Monoisotopic Exact Mass	Trend ^a	Data set
M1	284,2932	15,15	284,2932	284,2948	-5,6	0,5	Octadecanamide	C ₁₈ H ₃₇ NO	283,2870	↓	(+)
M2	195,1368	11,24	195,1368	195,1380	-5,9	3,5	Fatty acid ester	C ₁₂ H ₁₈ O ₂	194,1301	↓	(+)
M3	357,2774	11,24	357,2774	357,2788	-3,9	6,5	Unsaturated fatty acid/C24 bile acid (sterol lipids)/ ω-3 polyunsaturated fatty acid ethyl ester	C ₂₄ H ₃₆ O ₂	356,2710	↓	(+)
M4	355,2611	10,17	355,2611	355,2632	-5,5	7,5	Unsaturated fatty acid/C24 bile acid (sterol lipids)	C ₂₄ H ₃₄ O ₂	354,2553	↓	(+)
M5	373,2728	10,11	373,2728	373,2737	-2,5	6,5	C24 bile acid (sterol lipids)	C ₂₄ H ₃₆ O ₃	372,2659	↓	(+)
M6	522,3538	13,01	522,3538	522,3554	-3,1	1,5	1-alkyl,2- acylglycerophosphocholine s (glycerophospholipids)	C ₂₆ H ₅₂ NO ₇ P	521,3476	↓	(+)

Metabolite id.	m/z	t_R (min)	Experimental m/z	Theoretical m/z	Δm (ppm)	RDB	Name	Molecular Formula	Monoisotopic Exact Mass	Trend ^a	Data set
M7	299,1989	10,65	299,1989	299,2006	-5,6	7,5	Retinoid (prenol lipids)	C ₂₀ H ₂₆ O ₂	298,1927	↓	(+)
M8	282,2776	14,77	282,2776	282,2791	-5,2	2,0	Oleamide Primary amides (fatty acyls)	C ₁₈ H ₃₅ NO	281,2713	↓	(+)
M9	568,3375	12,20	568,3375	568,3398	-4,0	6,5	Monoacylglycerophosphoc holine (glycerophospholipids)	C ₃₀ H ₅₀ NO ₇ P	567,3319	↓	(+)
M10	102,0904	1,67	102,0904	102,0913	-5,3	0,5	Primary amide (fatty acyls)	C ₅ H ₁₁ NO	101,0835	↑	(+)
M11	221,1524	10,30					Unknown			↑	(+)
M12	375,1812	10,88	375,1812	375,1802	2,6	8,5	18-Oxocortisol	C ₂₁ H ₂₆ O ₆	374,1724	↓	(-)

Metabolite id.	<i>m/z</i>	<i>t_R</i> (min)	Experimental <i>m/z</i>	Theoretical <i>m/z</i>	Δm (ppm)	RDB	Name	Molecular Formula	Monoisotopic Exact Mass	Trend ^a	Data set
M13	837,5473	11,75	837,5473	837,5488	-1,7	3,5	Diacylglycerophosphoinositol (glycerophospholipidis)	C ₄₃ H ₈₃ O ₁₃ P	838,5566	↓	(-)
M14	437,2902	13,22	437,2902	437,2898	0,6	5,5	3beta-(3-methylbutanoyloxy)-villanovane-13alpha,17-diol C20 isoprenoid (prenol lipids)	C ₂₅ H ₄₂ O ₆	438,2976	↓	(-)
M15	353,1982	17,15	353,1982	353,1959	6,6	5,5	5-hydroperoxy-7-[3,5-epidioxo-2-(2-octenyl)-cyclopentyl]-6-heptenoic acid Hydroperoxy fatty acid (fatty acyls)	C ₁₉ H ₃₀ O ₆	354,2037	↓	(-)
M16	407,2796	11,75	407,2796	407,2792	0,5	4,5	C24 bile acid (sterol lipids)	C ₂₄ H ₄₀ O ₅	408,2870	↓	(-)
M17	913,5812	19,00	913,5812	913,5801	0,4	7,5	Diacylglycerophosphoinositol (glycerophospholipidis)	C ₄₉ H ₈₇ O ₁₃ P	914,5879	↓	(-)

Metabolite id.	<i>m/z</i>	<i>t_R</i> (min)	Experimental <i>m/z</i>	Theoretical <i>m/z</i>	Δm (ppm)	RDB	Name	Molecular Formula	Monoisotopic Exact Mass	Trend ^a	Data set
M18	443,2565	11,75	443,2565	443,2557	1,5	4,5	sn-3-O-(geranylgeranyl)glycerol 1-phosphate Monoalkylglycerophosphate (glycerophospholipids)	C ₂₃ H ₄₁ O ₆ P	444,2635	↓	(-)
M19	165,0557	7,50	165,0557	165,0558	7,3	5,5	3-Methoxy-4-hydroxyphenylacetaldehyde	C ₉ H ₉ O ₃	166,0624	↑	(-)
M20	199,1683	10,23	199,1683	199,1693	-4,8	1,5	Lauric acid	C ₁₂ H ₂₄ O ₂	200,771	↓	(-)
M21	279,2305	12,52	279,2305	279,2319	-4,9	3,5	Linoleic acid	C ₁₈ H ₃₂ O ₂	280,2397	↓	(-)
M22	281,2481	13,14	281,2481	281,2475	2,1	2,5	Oleic acid	C ₁₈ H ₃₄ O ₂	282,2553	↓	(-)
M23	283,2621	14,96	283,2621	283,2632	-3,7	1,5	Stearic acid	C ₁₈ H ₃₆ O ₂	284,2710	↓	(-)

Metabolite id.	<i>m/z</i>	<i>t_R</i> (min)	Experimental <i>m/z</i>	Theoretical <i>m/z</i>	Δm (ppm)	RDB	Name	Molecular Formula	Monoisotopic Exact Mass	Trend ^a	Data set
M24	391,2850	16,10	391,2850	391,2843	1,8	5,5	3,7-Dihydroxycholelan-24-oic acid	C ₂₄ H ₄₀ O ₄	392,2921	↓	(-)
M25	409,3472	14,5 (or 17,36)	409,3472	409,3465	1,7	7,5	(3beta,5alpha)-4,4-Dimethylcholesta-8,14,24-trien-3-ol	C ₂₉ H ₄₆ O	410,3543	↓	(-)
M26	227,1995	10,76	227,1995	227,2006	-4,6	1,5	Myristic acid	C ₁₄ H ₂₈ O ₂	228,2084	↓	(-)
M27	253,2154	11,12	253,2154	253,2162	-3,2	2,5	Palmitelaidic acid	C ₁₆ H ₃₀ O ₂	254,2240	↓	(-)
M28	305,2486	17,63	305,2486	305,2475	3,6	4,5	11,14,17-Eicosatrienoic acid/8,11,14-Eicosatrienoic acid	C ₂₀ H ₃₄ O ₂	306,2553	↓	(-)
M29	303,2326	16,95	303,2326	303,2319	2,5	5,5	Arachidonic acid/Cis-8,11,14,17-Eicosatetraenoic acid	C ₂₀ H ₃₂ O ₂	304,2397	↓	(-)

Metabolite id.	m/z	t_R (min)	Experimental m/z	Theoretical m/z	Δm (ppm)	RDB	Name	Molecular Formula	Monoisotopic Exact Mass	Trend ^a	Data set
M30	91,0383	1,05	91,0383	91,0390	-7,3	0,5	Glycerol	C ₃ H ₈ O ₃	92,0468	↑	(-)
M31	165,0539	1,20	165,0539	165,0546	1,7	7,0	3-(3-Hydroxyphenyl) propanoic acid	C ₉ H ₁₀ O ₃	166,0624	↑	(-)

Table 6: Candidate metabolites identified in H and HHT groups from both orbitrap-based and qQToF-based metabolomic approaches. The features are presented as a. metabolite id., b. m/z feature, c. t_R (min), d. experimental m/z , e. theoretical m/z , f. Δm (ppm), g. corresponding RDB value, h. possible identification name, i. possible molecular formula, j. corresponding monoisotopic exact mass and k. change trend in HHT group compared with H group (up or down regulation) and l. corresponding dataset (positive or negative ESI dataset).

^a (↑): up-regulated. (↓): down-regulated.

The next step was to identify the selected features, or at least, a subset of the most dominant ones, as these are clearly important metabolites for distinguishing the H and HHT groups. The identification approach was to match the selected features in terms of mass accuracy in online databases, such as METLIN, MassBank, ChemSpider and HMDB, in order to recognize possible metabolites. Furthermore, additional data were used to confirm the proposed structures, such as the isotopic pattern and the RDB information as well as possible pseudo MS/MS data that exist in the corresponding spectra. Following this approach, 11 in (+) ESI mode and 8 in (-) ESI mode significant features were found and tabulated in Table 6. Hence, these metabolites presumably represent the metabolic pathways that are differentially affected MetS rats on HT treatment.

Identification and structural elucidation of potential target metabolites. The robust UHPLC-ESI-HRMS metabolomics platform provides the accurate mass and the corresponding t_R (min) along with the area of the selected features in the different samples analyzed. This information was used for the structural identification of biomarkers. The accurate molecular mass was determined within measurement errors (<10 ppm) of the LTQ-Orbitrap, meanwhile the potential elemental composition, degree of unsaturation and the suggested isotope pattern of compounds were obtained with the aid of the Xcalibur[®] software. The candidate molecular formulas or exact masses were searched against the METLIN, Chemspider, HMDB, KEGG, LIPID MAPS and MassBank databases to identify the possible chemical entities. Furthermore, pseudo MS/MS data, which were obtained from the Mass Frontier, were screened to determine the potential structures of the examined ions. Setting one candidate feature as an example, the identification procedure was as follows.

According to the protocol detailed above, the feature M1 from the (+) ESI data set at $t_R=15.15$ min with $m/z=284.2932$ exhibits a high VIP value as well as an FC value >2 and a $p<0.05$. Respecting the nitrogen rule, this ion might contain an odd number of nitrogen atoms because its monoisotopic exact molecular weight is 283.2870 whereas its molecular formula was speculated as $C_{18}H_{37}NO$ from the analysis of its elemental composition and its isotopic pattern. The RDB value was calculated as 0.5, indicating that it was a saturated compound. According to the previous observations the molecule was speculated as stearamide, a primary amide belonging to the fatty acyls group, after searching in the online databases. All the potential metabolites identified are shown in Table 6. According to the protocol detailed above, 19 endogenous metabolites contributing to the separation of the HHT and H groups were detected in the samples. The significantly down-regulated M1-M9 and M12-M18, and the up-regulated M10-M11 and M19 were observed in the obese rats following treatment with HT (HHT group) compared with the control group H where the changes in their relative concentrations identified in the groups have been determined (Figure 20). This difference of the metabolites in the H and HHT groups may denote their potential as targeted biomarkers in treatment of the MetS with HT. Monitoring the changes concerning these metabolites may set some light in the prediction and biochemical mechanism of the MetS.

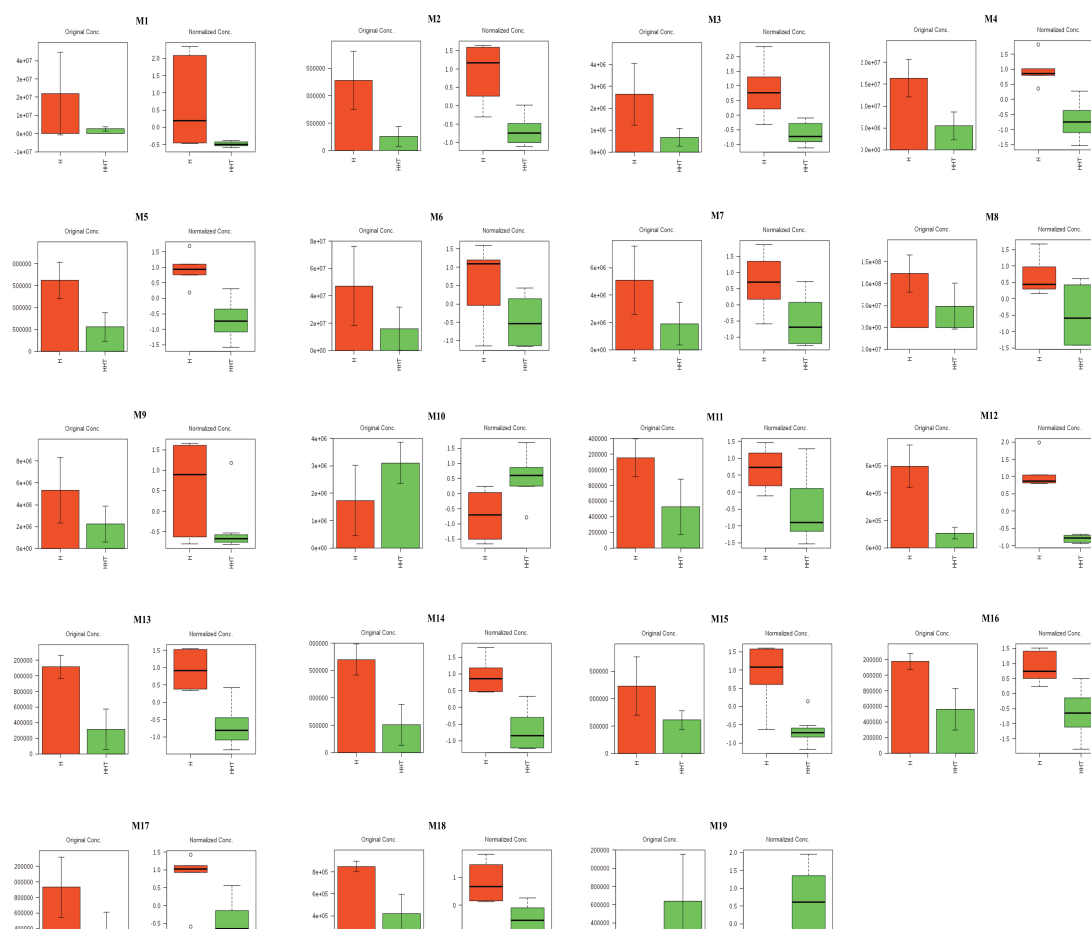


Figure 20: Histograms represent the original concentrations (area) in the left panel of each feature and box-and-whisker plots represent the normalized concentrations in the right panel of the selected features (y-axis) in the H (orange) and HHT (green) groups respectively (x-axis). Each bar graph, generated using the MetaboAnalyst platform, represents the range (upper and lower limit), median (bold line drawn across the bar) and the standard error (error bar).

qQToF-based metabolomics

The resulting multivariate models using the qQToF-based approach exhibit the same trends as the orbitrap-based approach (small degree of clustering for the PCA model but clear separation between the groups for PLS-DA and OPLS-DA). A representative OPLS-DA scores plot of H vs. HHT groups along with the corresponding S-plot can be found in Figure 21. The quality assessment of the method was expressed by the R^2 and Q^2 coefficients, which were found to be 0.987 and 0.612

respectively. The selection of VIP's as well as their identification has been performed in the same manner as in the case of the orbitrap-based metabolomic analysis. Following this approach, 12 endogenous metabolites contributing to the separation of the HHT and H groups were detected in the samples. The significantly down-regulated M20-M29, and the up-regulated M30-M31 were observed in the obese rats following treatment with HT (HHT group) compared with the control group H (Table 6).

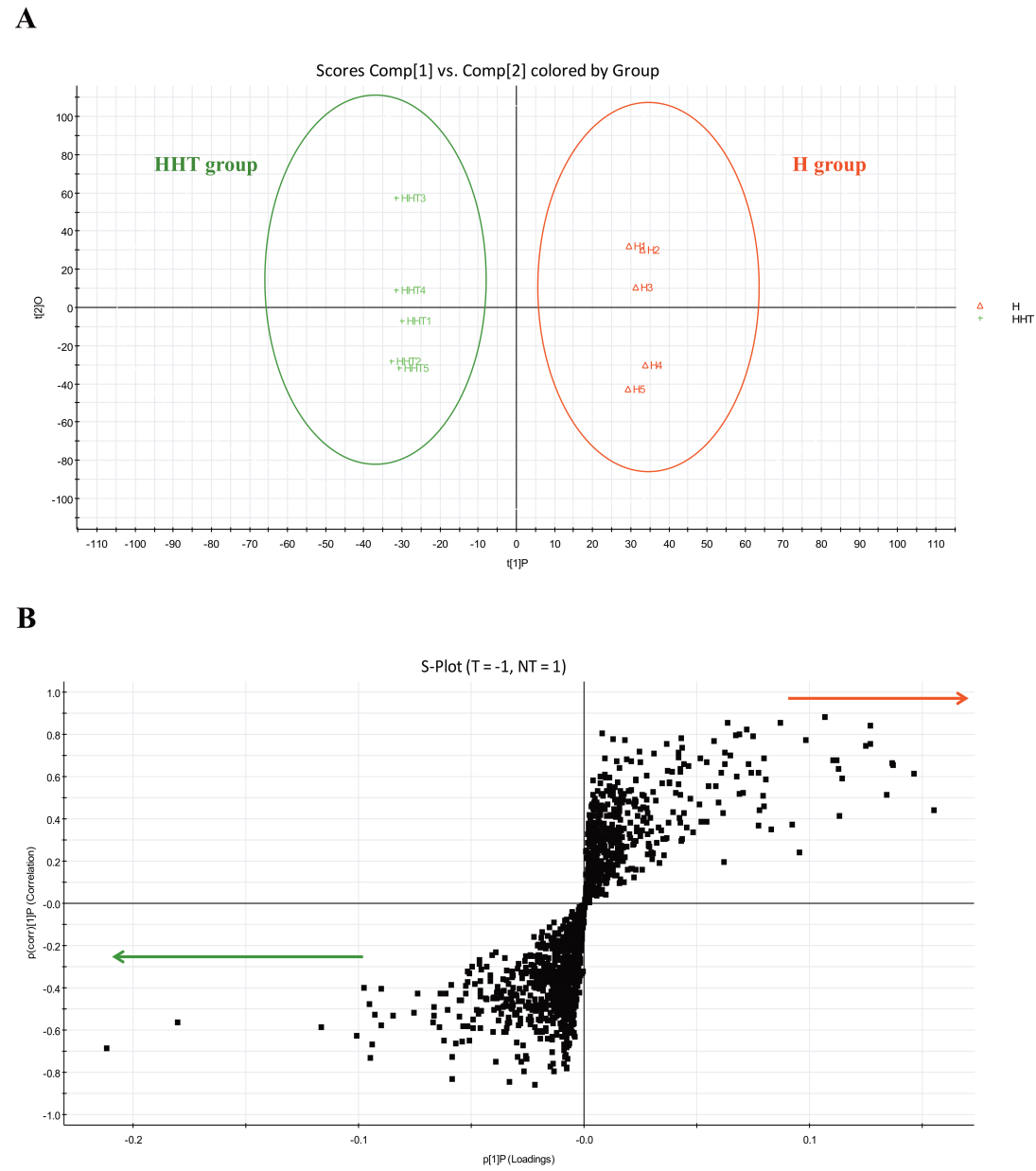


Figure 21: OPLS-DA model results for H and HHT groups obtained by UHPLC-(-)ESI-HRMS (qQToF). Significant clustering is shown between the H group (orange open triangles) and the HHT group (green crosses) in the OPLS-DA scores plot (A). S-plot of the OPLS-DA model. Points across the green and orange arrows represent the metabolites that show the highest concentration change between the two groups (B).

Metabolite set enrichment analysis

In the enrichment analysis using the “pathway-associated metabolite sets filter”, the top two metabolic pathways that appear to be associated with the MetS are the alpha linolenic acid and linoleic acid metabolism, as well as the beta-oxidation of very long chain fatty acids (Figure 22). Both are up-regulated in the H group (obese rats). It is well known, that plasma free (non esterified) fatty acids can mediate insulin resistance and are considered to arise from increased adipose tissue mass. The results show that as a high concentration of free fatty acids is observed in H group (obese rats), it is evident that there is a dysfunction positively correlated with insulin resistance, which is one of the symptoms of MetS. The decrease of the plasma levels of circulating free fatty acids in the HHT group can be correlated with the alleviation of the symptoms of the MetS. This effect could possibly disrupt the vicious cycle of insulin resistance.

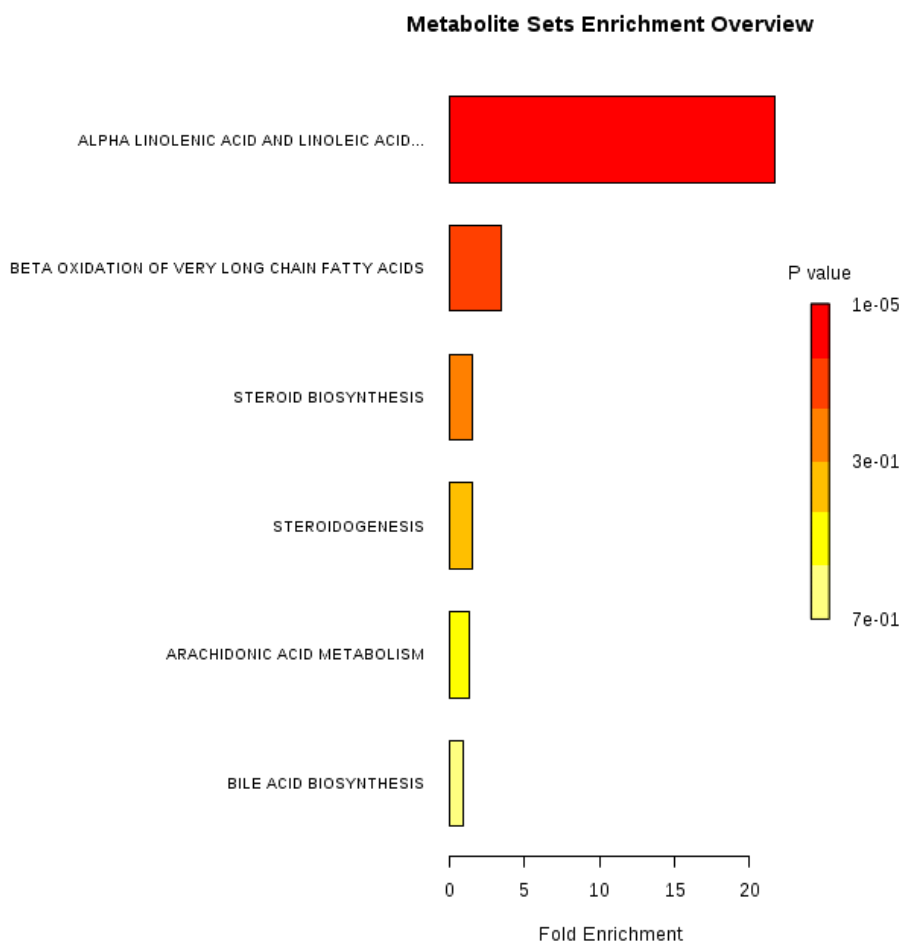


Figure 22: Over representation analysis (ORA) performed using metabolite set enrichment analysis (MSEA). ORA is based on the global test algorithm, which uses a generalized linear model to estimate a Q-statistic for each metabolite set. As significant metabolites, were considered the ones that were up-regulated in the H group. They identified from both metabolomic approaches (orbitrap- and qQToF-based) in both positive and negative ESI modes. These were inputted into MSEA and enrichment was assessed using the MSEA pathway-associated metabolite sets filter (n=88), and custom metabolite sets derived from KEGG. Pathways were considered enriched when $p < 0.05$ and $FDR < 5\%$.

Additionally, applying the “disease-associated metabolite sets (blood) filter”, possible metabolite sets reported in human blood that are associated with diseases, were investigated. It has found that hypertension and diabetes mellitus (Figure 23) associated metabolites (major characteristics of MetS) are “normalized” to the HHT group (obese rats after treatment with HT), indicating the positive effect of HT in the treatment of the disease.

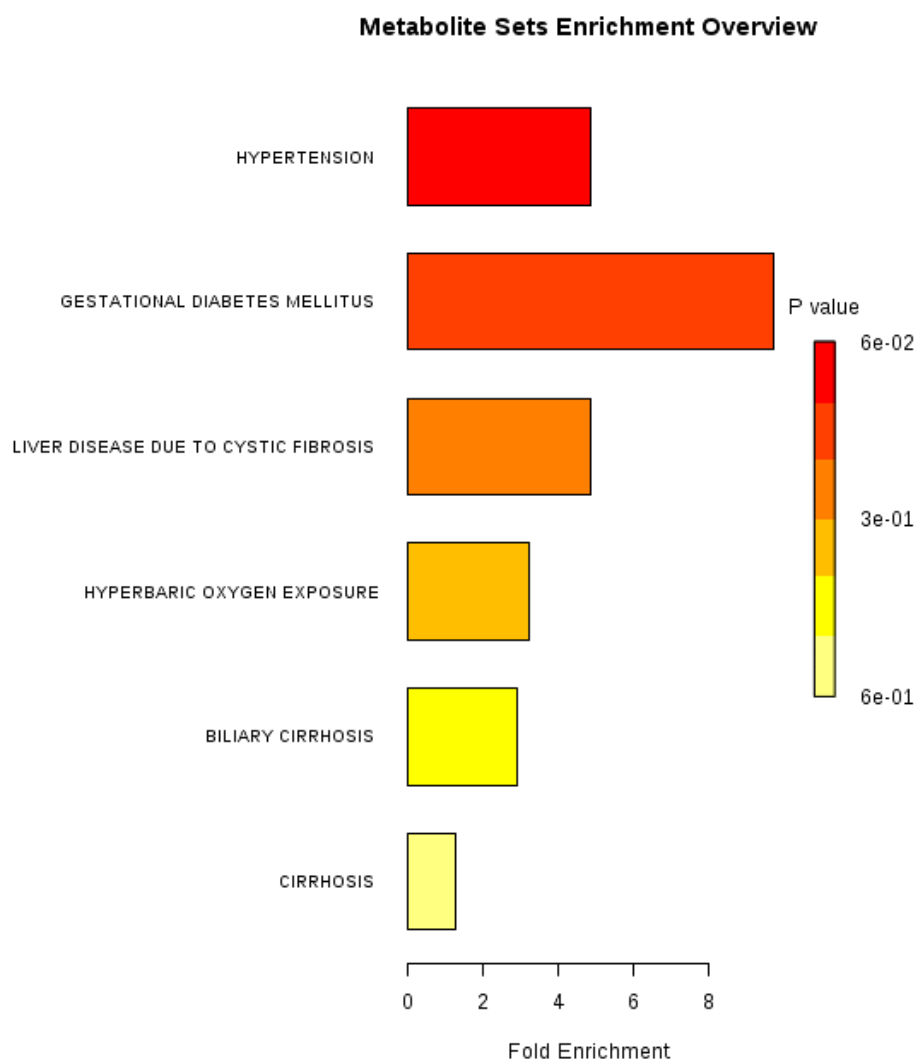


Figure 23: Over representation analysis (OPA) performed using metabolite set enrichment analysis (MSEA). Significant metabolites identified from both metabolomic approaches (orbitrap- and qQToF-based) in both positive and negative ESI modes, which were found to be up-regulated in H group (obese rats), were inputted into MSEA and enrichment was assessed using the MSEA disease-associated metabolite sets filter (n=416), and custom metabolite sets reported in human blood were found. Pathways were considered enriched when $p < 0.05$ and $FDR < 5\%$.

Metabolic pathway analysis

More detailed analysis of the pathways that influenced by the HT treatment in the rat MetS model was performed by MPA. In order to identify the possible biochemical pathways affected by the treatment in the obese rat group, the selected significant

features were analyzed using MetPA. MetPA is a free, web-based tool that combines results from the pathway enrichment analysis with the pathway topology analysis to aid towards identifying the most relevant pathways involved in the conditions under study.

MPA by MetPA respecting the double criterion $p < 0.5$ or impact value > 0.2 , revealed that the metabolic pathways involved are: biosynthesis of unsaturated fatty acids, fatty acid biosynthesis, linoleic acid metabolism, retinol metabolism, sphingolipid metabolism and arachidonic acid metabolism (Table 7), which are down-regulated in HHT groups, whereas glycerolipid metabolism is the main up-regulated metabolic pathways in the HHT group (Table 8).

	Total	Expected	Hits	Raw p	Holm p	FDR	Impact
Biosynthesis of unsaturated fatty acids	42	0,51	7	2.02E-07	1.64E-05	1.64E-05	0.00
Fatty acid biosynthesis	43	0,52	4	1.37E-03	1.09E-01	5.54E-02	0.00
Linoleic acid metabolism	5	0,06	1	5.93E-02	1.00E+00	1.00E+00	1.00
Phenylalanine metabolism	9	0,11	1	1.04E-01	1.00E+00	1.00E+00	0.00
Retinol metabolism	17	0,21	1	1.88E-01	1.00E+00	1.00E+00	0.22
Sphingolipid metabolism	21	0,25	1	2.27E-01	1.00E+00	1.00E+00	0.28
Steroid biosynthesis	35	0,42	1	3.51E-01	1.00E+00	1.00E+00	0.12
Arachidonic acid metabolism	36	0,44	1	3.59E-01	1.00E+00	1.00E+00	0.33

Primary bile acid biosynthesis	46	0,56	1	4.35E-01	1.00E+00	1.00E+00	0.00
Steroid hormone biosynthesis	70	0,85	1	5.83E-01	1.00E+00	1.00E+00	0.02

Table 7: Metabolic pathways down-regulated in HHT rats, obtained by the MPA. Total is the total number of compounds in the pathway; the Hits is the actually matched number from the user uploaded data; the Raw p is the original p value calculated from the enrichment analysis; the FDR is the false discovery rate; the Impact is the pathway impact value calculated from pathway topology analysis. The significantly altered pathways by the treatment according to the double criterion ($p < 0.05$ or $\text{impact} > 0.2$) are grayed out.

	Total	Expected	Hits	Raw p	Holm p	FDR	Impact
Glycerolipid metabolism	18	0,04	1	3.81E-02	1.00E+00	1.00E+00	0.28
Galactose metabolism	26	0,06	1	5.46E-02	1.00E+00	1.00E+00	0.00
Tyrosine metabolism	42	0,09	1	8.73E-02	1.00E+00	1.00E+00	0.00

Table 8: Metabolic pathways up-regulated in HHT rats, obtained by the MPA. Total is the total number of compounds in the pathway; the Hits is the actually matched number from the user uploaded data; the Raw p is the original p value calculated from the enrichment analysis; the FDR is the false discovery rate; the Impact is the pathway impact value calculated from pathway topology analysis. The significantly altered pathways by the treatment according to the double criterion ($p < 0.05$ or $\text{impact} > 0.2$) are grayed out.

The summary of the pathway analysis is also represented as an illustration in Figure 24A (down-regulated in HHT group) and Figure 24B (up-regulated in HHT group), in which the color and size of each circle in the figures is based on the p values and pathway impact values, respectively.

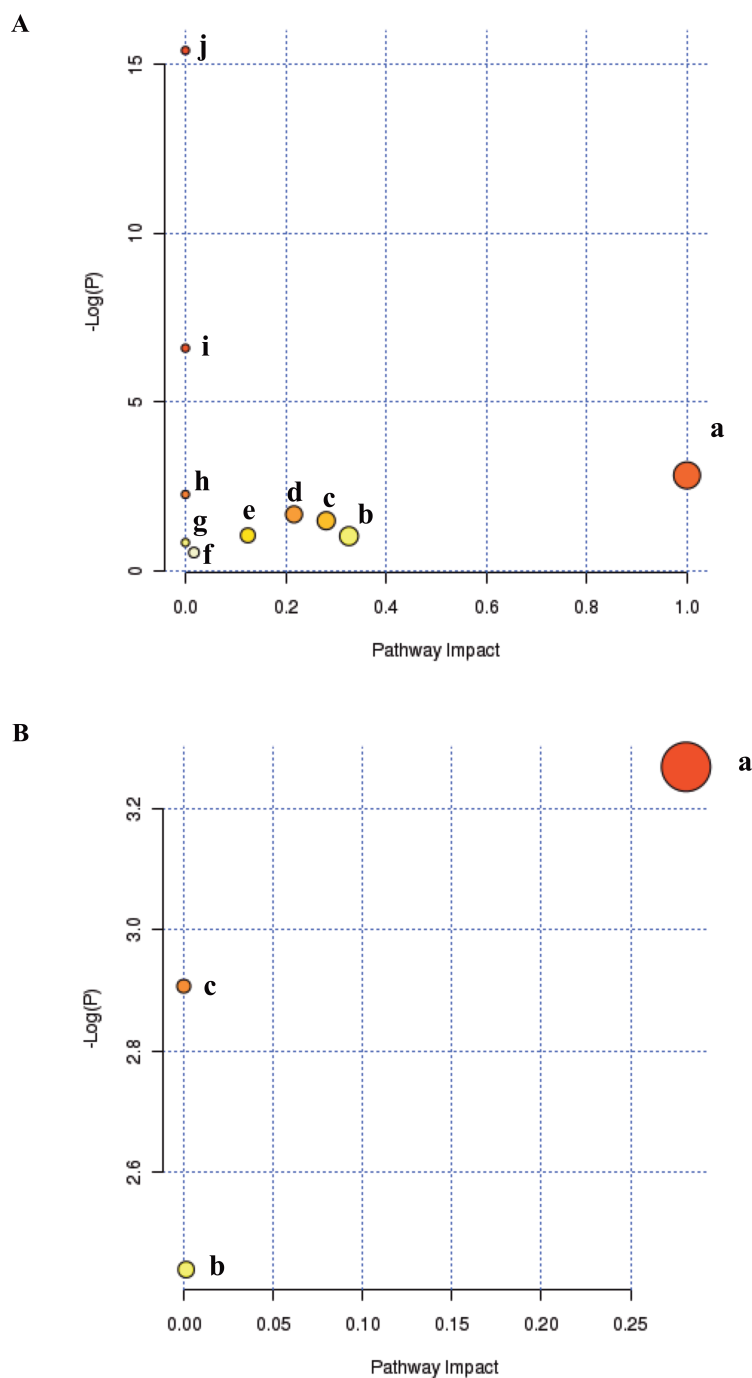


Figure 24: Summary of the pathway analysis for the HT treatment to a metabolic syndrome rat model. Pathways down-regulated (**A**) and up-regulated (**B**) in the HHT group (after treatment). All the matched pathways are represented as circles. The color intensity and the radius of the circle indicate the p value and pathway impact value respectively. (**A**): a. linoleic acid metabolism, b. arachidonic acid metabolism, c. sphingolipid metabolism, d. retinol metabolism, e. steroid biosynthesis, f. steroid hormone biosynthesis, g. primary bile acid biosynthesis, h. phenylalanine metabolism, i. fatty acid biosynthesis and j. biosynthesis of unsaturated fatty acids. (**B**): a. glycerolipid metabolism, b. tyrosine metabolism and c. galactose metabolism.

The down-regulation of the linoleic acid metabolism (Figure 25), and the concomitant down-regulation of the arachidonic acid metabolism (Figure 26), maybe associated with reduced prostaglandins biosynthesis. It is well known that prostaglandins mediate low-level chronic inflammation, which is directly associated with atherosclerosis prevalence, one of the symptoms of MetS. Treatment with HT seems to reduce this risk. Furthermore, the elevated levels of unsaturated fatty acids in obese rats (H group) compared to the obese treated rats (HHT group), denote the alleviation of the high risk of insulin resistance, which is common to diabetes type II. It should be reminded that diabetes type II is also another causative effect of MetS.

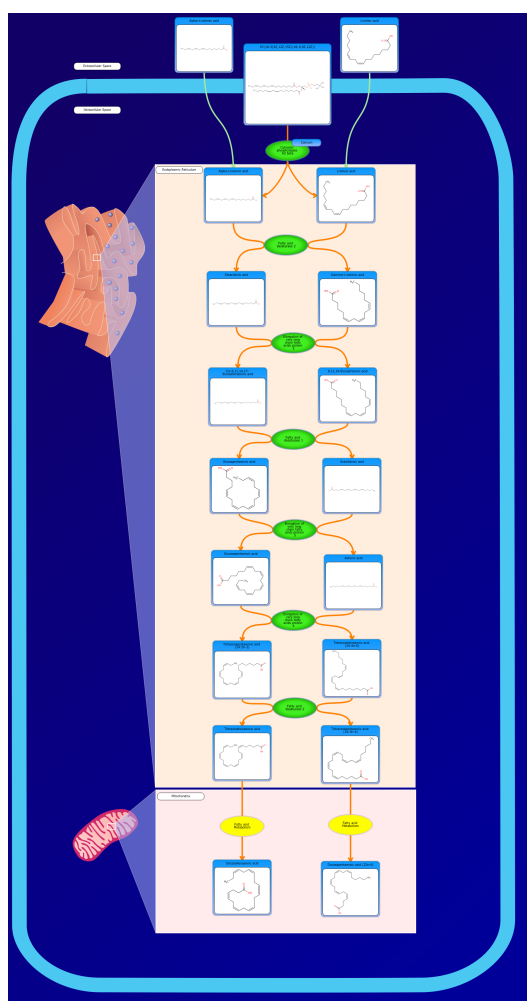


Figure 25: Linoleic acid metabolism.

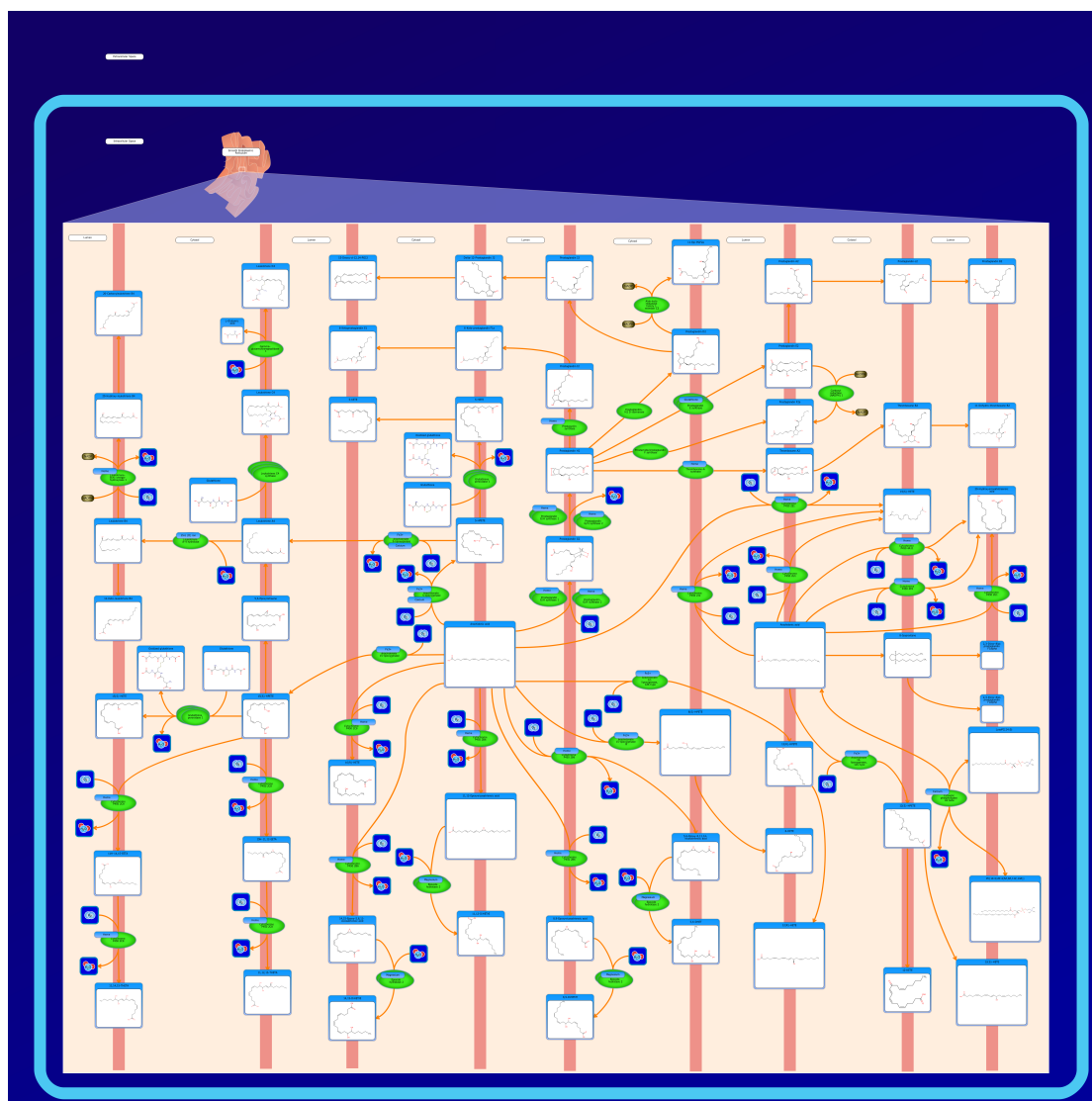


Figure 26: Arachidonic acid metabolism.

It is also of high importance that the reduced levels of plasma circulating fatty acids in the obese rats after HT treatment (HHT group) can be positively correlated with their efficient metabolism, which in turn is negatively correlated with nonalcoholic fatty liver disease. In other words, the cells could efficiently metabolize the free fatty acids after the treatment with HT and this effect leads to lower levels of plasma fatty acids. These results are in accordance with the pharmacological evaluation (pharmacological evaluation section), especially the normalization of the plasma alanine transaminase (ALT) and aspartate transaminase (AST) levels (Table

5) after treatment (HHT group), which reduces the occurrence of nonalcoholic fatty liver disease.

Retinol metabolism (Figure 27) is found to be reduced by the activity of HT in the treated group (HHT group). Concerning this observation, it should be stress out that retinoids are of significant importance in the development of diseases, such as pancreatic dysfunction, insulin resistance and lipid metabolism. These actions are mediated by the rexinoids, which seem to function as insulin sensitizers but also their effects are mediated through the dimerization with peroxisome proliferator-activated receptors γ (PPAR γ)⁶¹.

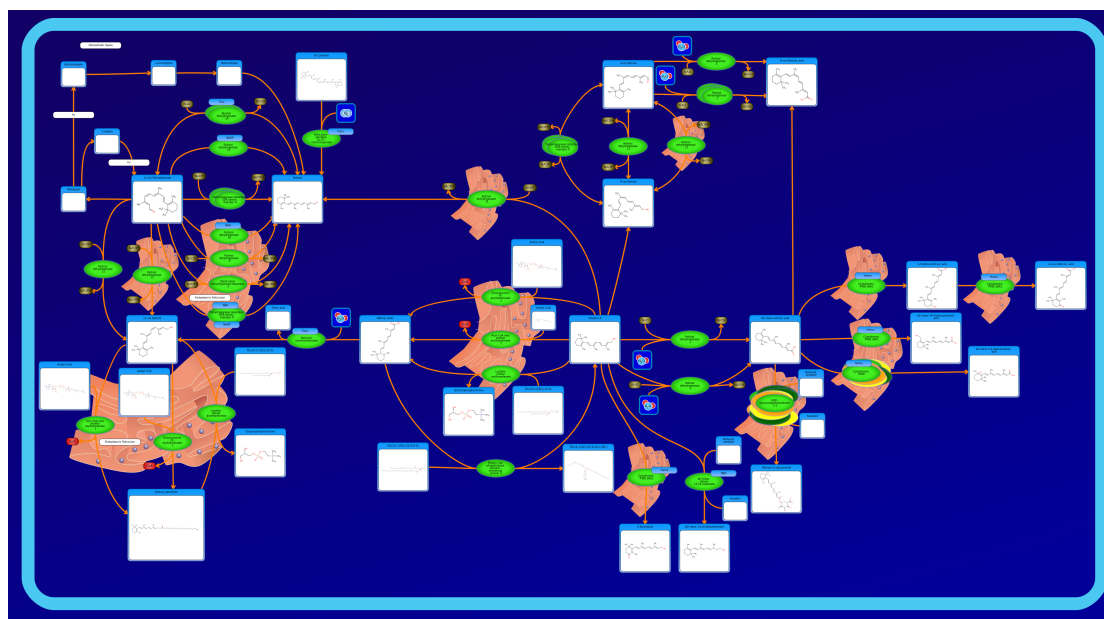


Figure 27: Retinol metabolism.

Elevated plasma sphingolipids were found as biomarkers of MetS in monkeys, whereas it seems that in rodents the accumulation of the sphingolipid ceramide in insulin-sensitive tissues, may trigger the disease^{62,63}. HT treatment seems to normalize these effects, as it has been found circulating concentrations of sphingolipids were reduced in HHT group.

Finally, the elevated plasma levels of glycerol of treated rats (HHT group) is an indication of a higher rate of lipolysis in adipose tissue, which in term indicates a higher rate of energy production through lipolysis. In other words, glycerolipids (Figure 28) that are stored in adipose cells are promoted to be utilized for the production of energy, which indicates a more efficient lipid movement after the treatment with HT.

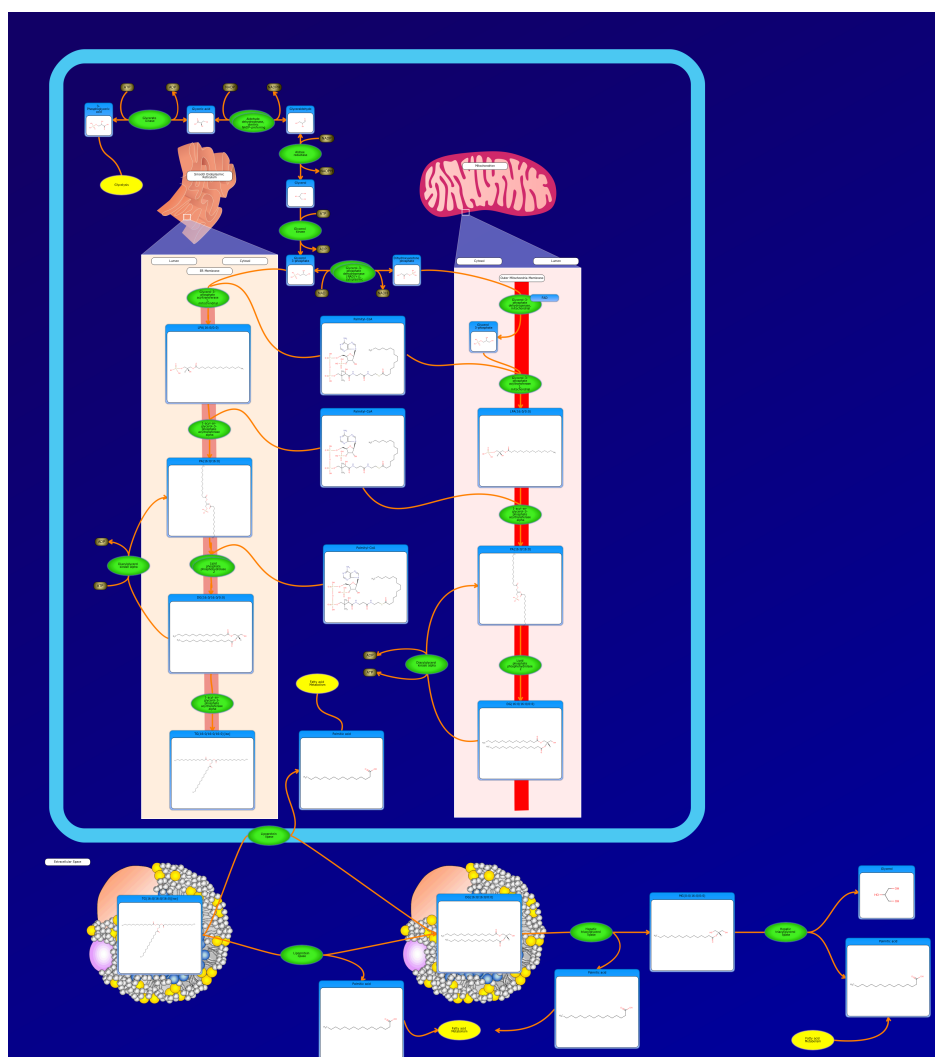


Figure 28: Glycerolipid metabolism.

CONCLUSION

Metabolomics is a powerful new technology that allows for the assessment of global metabolic profiles biosamples, such as plasma and urine, and it is frequently employed for biomarker discovery and/or assessment between two different statuses, i.e. treated/non-treated subjects. In the present study, an integrated mass spectrometry-based metabolomic approach has been utilized in order to investigate possible metabolic changes in a MetS rat model after 8-weeks treatment with HT. After data processing and careful validation of the results, significant and clear clustering between H and HHT group in both PCA and PLS-DA/OPLS-DA models has been found.

Furthermore, it is evident that MPA of the plasma metabolites from MetS is highly important for assessing inherent network properties in reconstructed biochemical reaction networks. In the present study, metabolomic analysis with the aid of MPA, regarding MetS and the response after HT treatment was inferred from changes in the intermediates during substance administration. Thus, 31 differentially expressed metabolites were identified revealing pathway-specific expression profiles.

Overall, treatment with HT has been found to alter the plasma metabolome of obese rats (HHT group) compared to those without treatment (H group), demonstrating that HT could have some therapeutic effect on the MetS. Importantly, the findings support that HT treatment normalized an array of factors concerning crucial biochemical pathways that are implicated in the development and establishment of MetS. Thus, the biosynthesis of unsaturated fatty acids and the metabolism of linoleic acid, arachidonic acid, sphingolipids and retinol were

normalized after the treatment with HT (HHT group), whereas factors involved in efficient lipid utilization are promoted, such as glycerolipid metabolism.

These results are consistent with the assumption that HT can be considered as a pharmacologically active compound for the treatment of MetS. These findings are consistent with the beneficial effects of olive oil in humans concerning the reduction of obesity, hypertension, diabetes and pro-inflammation, largely contributing to a lower risk of cardiovascular diseases^{25,64}.

REFERENCES

1. Alberti, K. G. M. M. *et al.* Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International . *Circulation* **120**, 1640–5 (2009).
2. Grundy, S. M. *et al.* Diagnosis and management of the metabolic syndrome: an American heart association/national heart, lung, and blood institute scientific statement. *Curr Opin Cardiol* **21**, 1–6 (2006).
3. Park, Y.-W. *et al.* The metabolic syndrome: prevalence and associated risk factor findings in the US population from the third national health and nutrition examination survey. *Arch. Intern. Med.* **163**, 427 (2003).
4. Gami, A. S. *et al.* Metabolic syndrome and risk of incident cardiovascular events and death: a systematic review and meta-analysis of longitudinal studies. *J. Am. Coll. Cardiol.* **49**, 403–14 (2007).
5. Ford, E. S., Li, C. & Sattar, N. Metabolic syndrome and incident diabetes: current state of the evidence. *Diabetes Care* **31**, 1898–1904 (2008).
6. Desroches, S. & Lamarche, B. The evolving definitions and increasing prevalence of the metabolic syndrome. *Appl. Physiol. Nutr. Metab.* **32**, 23–32 (2007).
7. Guilherme, A., Virbasius, J. V, Puri, V. & Czech, M. P. Adipocyte dysfunctions linking obesity to insulin resistance and type 2 diabetes. *Nat. Rev. Mol. Cell Biol.* **9**, 367–77 (2008).
8. Mayneris-Perxachs, J. *et al.* Effects of 1-year intervention with a mediterranean diet on plasma Fatty Acid composition and metabolic syndrome in a population at high cardiovascular risk. *PLoS One* **9**, e85202 (2014).
9. Abete, I., Astrup, A., Martínez, J. A., Thorsdottir, I. & Zulet, M. A. Obesity and the metabolic syndrome: role of different dietary macronutrient distribution patterns and specific nutritional components on weight loss and maintenance. *Nutr. Rev.* **68**, 214–31 (2010).
10. Cameron, A. J., Shaw, J. E. & Zimmet, P. Z. The metabolic syndrome: prevalence in worldwide populations. *Endocrinol. Metab. Clin. North Am.* **33**, 351–75, table of contents (2004).
11. Cordain, L. *et al.* Origins and evolution of the Western diet: health implications for the 21st century. *Am. J. Clin. Nutr.* **81**, 341–54 (2005).
12. Grundy, S. M. *et al.* Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. *Circulation* **112**, 2735–52 (2005).
13. Baxter, A. J., Coyne, T. & McClintock, C. Dietary patterns and metabolic syndrome--a review of epidemiologic evidence. *Asia Pac. J. Clin. Nutr.* **15**, 134–42 (2006).

14. Stanhope, K. L. & Havel, P. J. Fructose consumption: potential mechanisms for its effects to increase visceral adiposity and induce dyslipidemia and insulin resistance. *Curr. Opin. Lipidol.* **19**, 16–24 (2008).
15. Buettner, R., Schölmerich, J. & Bollheimer, L. C. High-fat diets: modeling the metabolic disorders of human obesity in rodents. *Obesity (Silver Spring)*. **15**, 798–808 (2007).
16. Poudyal, H., Campbell, F. & Brown, L. Olive leaf extract attenuates cardiac, hepatic, and metabolic changes in high carbohydrate-, high fat-fed rats. *J. Nutr.* **140**, 946–53 (2010).
17. Pérez-López, F. R., Chedraui, P., Haya, J. & Cuadros, J. L. Effects of the Mediterranean diet on longevity and age-related morbid conditions. *Maturitas* **64**, 67–79 (2009).
18. Babio, N., Bulló, M. & Salas-Salvadó, J. Mediterranean diet and metabolic syndrome: the evidence. *Public Health Nutr.* **12**, 1607–17 (2009).
19. Esposito, K. & Giugliano, D. Mediterranean diet and the metabolic syndrome: the end of the beginning. *Metab. Syndr. Relat. Disord.* **8**, 197–200 (2010).
20. Gillingham, L. G., Harris-Janz, S. & Jones, P. J. H. Dietary monounsaturated fatty acids are protective against metabolic syndrome and cardiovascular disease risk factors. *Lipids* **46**, 209–28 (2011).
21. Keys, A. *et al.* The diet and 15-year death rate in the seven countries study. *Am. J. Epidemiol.* **124**, 903–15 (1986).
22. Panagiotakos, D. B. *et al.* Can a Mediterranean diet moderate the development and clinical progression of coronary heart disease? A systematic review. *Med. Sci. Monit.* **10**, RA193–8 (2004).
23. Poudyal, H., Campbell, F. & Brown, L. Olive leaf extract attenuates cardiac, hepatic, and metabolic changes in high carbohydrate-, high fat-fed rats. *J. Nutr.* **140**, 946–53 (2010).
24. Visioli, F. & Bernardini, E. Extra virgin olive oil's polyphenols: biological activities. *Curr. Pharm. Des.* **17**, 786–804 (2011).
25. López-Miranda, J. *et al.* Olive oil and health: summary of the II international conference on olive oil and health consensus report, Jaén and Córdoba (Spain) 2008. *Nutr. Metab. Cardiovasc. Dis.* **20**, 284–94 (2010).
26. Pérez-Jiménez, F., Ruano, J., Perez-Martinez, P., Lopez-Segura, F. & Lopez-Miranda, J. The influence of olive oil on human health: not a question of fat alone. *Mol. Nutr. Food Res.* **51**, 1199–208 (2007).
27. Mastralexi, A., Nenadis, N. & Tsimidou, M. Z. Addressing Analytical Requirements To Support Health Claims on “Olive Oil Polyphenols” (EC Regulation 432/2012). *J. Agric. Food Chem.* **62**, 2459–61 (2014).
28. Ciceralo, S., Conlan, X. A., Sinclair, A. J. & Keast, R. S. J. Chemistry and health of olive oil phenolics. *Crit. Rev. Food Sci. Nutr.* **49**, 218–36 (2009).

29. Rafehi, H., Ververis, K. & Karagiannis, T. C. Mechanisms of action of phenolic compounds in olive. *J. Diet. Suppl.* **9**, 96–109 (2012).
30. Corona, G. *et al.* The fate of olive oil polyphenols in the gastrointestinal tract: implications of gastric and colonic microflora-dependent biotransformation. *Free Radic. Res.* **40**, 647–58 (2006).
31. Manna, C. *et al.* Oleuropein prevents oxidative myocardial injury induced by ischemia and reperfusion. *J. Nutr. Biochem.* **15**, 461–6 (2004).
32. Castañer, O. *et al.* Protection of LDL from oxidation by olive oil polyphenols is associated with a downregulation of CD40-ligand expression and its downstream products in vivo in humans. *Am. J. Clin. Nutr.* **95**, 1238–44 (2012).
33. Dudley, J. I. *et al.* Does white wine qualify for French paradox? Comparison of the cardioprotective effects of red and white wines and their constituents: resveratrol, tyrosol, and hydroxytyrosol. *J. Agric. Food Chem.* **56**, 9362–73 (2008).
34. Bonoli, M., Bendini, A., Cerretani, L., Lercker, G. & Toschi, T. G. Qualitative and semiquantitative analysis of phenolic compounds in extra virgin olive oils as a function of the ripening degree of olive fruits by different analytical techniques. *J. Agric. Food Chem.* **52**, 7026–32 (2004).
35. Vázquez-Velasco, M. *et al.* Effects of hydroxytyrosol-enriched sunflower oil consumption on CVD risk factors. *Br. J. Nutr.* **105**, 1448–52 (2011).
36. Granados-Principal, S., Quiles, J. L., Ramirez-Tortosa, C. L., Sanchez-Rovira, P. & Ramirez-Tortosa, M. C. Hydroxytyrosol: from laboratory investigations to future clinical trials. *Nutr. Rev.* **68**, 191–206 (2010).
37. Tutino, V., Caruso, M. G., Messa, C., Perri, E. & Notarnicola, M. Antiproliferative, antioxidant and anti-inflammatory effects of hydroxytyrosol on human hepatoma HepG2 and Hep3B cell lines. *Anticancer Res.* **32**, 5371–7 (2012).
38. Takeda, Y. *et al.* Influence of olive-derived hydroxytyrosol on the toll-like receptor 4-dependent inflammatory response of mouse peritoneal macrophages. *Biochem. Biophys. Res. Commun.* (2014). doi:10.1016/j.bbrc.2014.03.094
39. De Roos, B. *et al.* Anti-platelet effects of olive oil extract: in vitro functional and proteomic studies. *Eur. J. Nutr.* **50**, 553–62 (2011).
40. Zhao, B. *et al.* Hydroxytyrosol, a natural molecule from olive oil, suppresses the growth of human hepatocellular carcinoma cells via inactivating AKT and nuclear factor-kappa B pathways. *Cancer Lett.* **347**, 79–87 (2014).
41. European Community. Council Regulation No 432/2012 of 16 May 2012 establishing a list of permitted health claims made on foods, other than those referring to the reduction of disease risk and to children's development and health. *Off. J. Eur. Comm.* L 136; 1–40 (2012).
at <http://faolex.fao.org/cgi-bin/faolex.exe?rec_id=112829&database=faolex&search_type=link&table=result&lang=en&format_name=@ERALL>

42. Panchal, S. K. *et al.* High-carbohydrate, high-fat diet-induced metabolic syndrome and cardiovascular remodeling in rats. *J. Cardiovasc. Pharmacol.* **57**, 611–24 (2011).
43. Wang, X. *et al.* Metabolite profiling and pathway analysis of acute hepatitis rats by UPLC-ESI MS combined with pattern recognition methods. *Liver Int.* (2013). doi:10.1111/liv.12301
44. Ryan, D. & Robards, K. Metabolomics: The greatest omics of them all? *Anal. Chem.* **78**, 7954–8 (2006).
45. Stanstrup, J., Gerlich, M., Dragsted, L. O. & Neumann, S. Metabolite profiling and beyond: approaches for the rapid processing and annotation of human blood serum mass spectrometry data. *Anal. Bioanal. Chem.* **405**, 5037–48 (2013).
46. Setoyama, D., Fujimura, Y. & Miura, D. Metabolomics reveals that carnitine palmitoyltransferase-1 is a novel target for oxidative inactivation in human cells. *Genes Cells* (2013). doi:10.1111/gtc.12098
47. Kirkwood, J. S., Legette, L. L., Miranda, C. L., Jiang, Y. & Stevens, J. F. A metabolomics-driven elucidation of the anti-obesity mechanisms of xanthohumol. *J. Biol. Chem.* **288**, 19000–13 (2013).
48. Snyder, N. W., Khezam, M., Mesaros, C. A., Worth, A. & Blair, I. A. Untargeted metabolomics from biological sources using ultraperformance liquid chromatography-high resolution mass spectrometry (UPLC-HRMS). *J. Vis. Exp.* e50433 (2013). doi:10.3791/50433
49. Xia, J. & Wishart, D. S. MetPA: a web-based metabolomics tool for pathway analysis and visualization. *Bioinformatics* **26**, 2342–4 (2010).
50. Draghici, S. *et al.* A systems biology approach for pathway level analysis. *Genome Res.* **17**, 1537–45 (2007).
51. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–50 (2005).
52. Agalias, A. *et al.* A new process for the management of olive oil mill waste water and recovery of natural antioxidants. *J. Agric. Food Chem.* **55**, 2671–6 (2007).
53. Poudyal, H., Panchal, S. K., Waanders, J., Ward, L. & Brown, L. Lipid redistribution by α -linolenic acid-rich chia seed inhibits stearoyl-CoA desaturase-1 and induces cardiac and hepatic protection in diet-induced obese rats. *J. Nutr. Biochem.* **23**, 153–62 (2012).
54. Dunn, W., Broadhurst, D. & Begley, P. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. ...* **6**, 1060–83 (2011).
55. Xia, J. & Wishart, D. S. Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nat. Protoc.* **6**, 743–60 (2011).
56. Xia, J., Mandal, R., Sinelnikov, I. V., Broadhurst, D. & Wishart, D. S. MetaboAnalyst 2.0—a comprehensive server for metabolomic data analysis. *Nucleic Acids Res.* **40**, W127–33 (2012).

57. Xia, J., Psychogios, N., Young, N. & Wishart, D. S. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* **37**, W652–60 (2009).
58. Xia, J. & Wishart, D. S. Metabolomic data processing, analysis, and interpretation using MetaboAnalyst. *Curr. Protoc. Bioinformatics* **Chapter 14**, Unit 14.10 (2011).
59. Van den Berg, R. A., Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, A. K. & van der Werf, M. J. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* **7**, 142 (2006).
60. Larsson, O., Wahlestedt, C. & Timmons, J. A. Considerations when using the significance analysis of microarrays (SAM) algorithm. *BMC Bioinformatics* **6**, 129 (2005).
61. Rhee, E.-J. & Plutzky, J. Retinoid metabolism and diabetes mellitus. *Diabetes Metab. J.* **36**, 167–80 (2012).
62. Brozinick, J. T. *et al.* Plasma sphingolipids are biomarkers of metabolic syndrome in non-human primates maintained on a Western-style diet. *Int. J. Obes. (Lond)*. **37**, 1064–70 (2013).
63. Ohkawa, R. *et al.* [Sphingolipids, possible biomarkers for atherosclerotic disorders]. *Rinsho Byori*. **61**, 795–802 (2013).
64. Santos-Buelga, C. *et al.* Mediterranean diet polyphenols reduce inflammatory angiogenesis through MMP-9 and COX-2 inhibition in human vascular endothelial cells: A potentially protective mechanism in atherosclerotic vascular disease and cancer. *Arch. Biochem. Biophys.* **527**, 81–89 (2012).
65. D'Angelo, S. *et al.* Pharmacokinetics and metabolism of hydroxytyrosol, a natural antioxidant from olive oil. *Drug Metab. Dispos.* **29**, 1492–8 (2001).
66. Mateos, R., Goya, L. & Bravo, L. Metabolism of the olive oil phenols hydroxytyrosol, tyrosol, and hydroxytyrosyl acetate by human hepatoma HepG2 cells. *J. Agric. Food Chem.* **53**, 9897–905 (2005).
67. Visioli, F. *et al.* Hydroxytyrosol excretion differs between rats and humans and depends on the vehicle of administration. *J. Nutr.* **133**, 2612–5 (2003).

6

Metabolism of hydroxytyrosol after its administration to a metabolic syndrome rat model

This chapter will prepared for publication

ABSTRACT

Metabolic syndrome is a clustering of interrelated risk factors for cardiovascular disease and diabetes. Hydroxytyrosol (HT) is a phenol present in large amount in byproducts from olives. An experimental protocol has been setup encompassing isolated HT administration in a diet induced rat model of metabolic syndrome, which consisted of two groups; one fed with normal and the other with high-fat diet. As HT has not been possible to be detected in the plasma of the cohorts, it was pertinent to investigate its metabolic fate. Thus, targeted identification of circulating in plasma HT metabolites was performed using UHPLC-(-)ESI-HRMS and employing software-based identification. Twenty-five HT metabolites were detected in rat plasma, six of them exhibiting statistically significant different concentration between the two groups, with the normal diet group showing elevating levels.

INTRODUCTION

As mentioned to the previous Chapter “Metabolomics study of hydroxytyrosol’s administration for the treatment of metabolic syndrome in a rat model”, hydroxytyrosol (HT) seems to alleviate symptoms associated with metabolic syndrome (MetS). Thus, the metabolism profile of HT is of paramount importance for the clarification of its therapeutic mode of action and therefore should be investigated. Nowadays, mass spectrometry-based metabolism offers a cutting edge analytical tool for accomplishing this task¹. The evaluation of the results requires specialized software in order to be accomplished². Such pieces of software seek for predetermined mass sifts in the acquired raw data taking into advantage the high-resolution capabilities and the enhanced mass accuracy of modern mass spectrometers, such as the orbitrap.

HT is metabolized by the main enzymes of the adrenaline metabolic pathway³. Thus, HT could initially be methylated by the action of catechol-O-methyltransferase (COMT) affording homovanillic alcohol. The last one can be oxidased affording homovanillic aldehyde, which can be further oxidased by aldehyde dehydrogenase affording homovanillic acid. Alternatively, HT can be initially oxidased through 2,4-dihydroxyphenylaldehyde to 3,4-dihydroxyphenylacetic acid, which can be methylated by COMT again to homovanillic acid. Nevertheless, other HT metabolites have been also found, such as HT glucuronide, HT acetate, HT sulfate and HT acetate sulfate⁴⁻⁷. It is known from the literature, that once absorbed, HT can undergo phase II conjugation metabolism by sulfotransferase enzymes in the intestine and liver, producing HT sulphate⁶. Methylation, sulfation and glucuronidation represent the main phase II metabolic biotransformations for phenolic compounds.

Acetylation has also been described in xenobiotics by a group of enzymes called acyltransferases, which catalyzes the transfer of acetyl groups from acetyl-CoA⁸. These enzymes catalyze either N-acetylation or O-acetylation.

In the present study, effort has been spend to discern the metabolism of HT after its administration to rats, either in a healthy state or suffering from induced MetS and identify possible new HT metabolites. Furthermore, comparison of the HT metabolites levels has been attempted between the two groups, in order to clarify whether the syndrome imposes different metabolism patterns.

EXPERIMENTAL

Chemical and Reagents

HT (M1, Figure 3) was isolated from olive mill wastewater (OMWW), as described in Chapter “Metabolomics study of hydroxytyrosol’s administration for the treatment of metabolic syndrome in a rat model”. Its purity was found to be more than 98% (HPLC-UV) and its structure was identified by HRMS and 1 & 2D NMR. All solvents used in the investigation of the HT metabolism study were of LC-MS grade. Methanol (MeOH), water and formic acid, were purchased from Fluka/Riedel-de Haën (Switzerland).

Equipment

Ultra high performance liquid chromatography - (-) electrospray ionization – high resolution mass spectrometry [UHPLC- (-) ESI-HRMS] analysis of HT metabolites was carried out on a hybrid LTQ Orbitrap Discovery XL mass spectrometer (Thermo

Scientific, Germany) coupled with an Accela ultra HPLC (UHPLC) system (Thermo Scientific, Germany) equipped with a binary pump, an autosampler, an online vacuum degasser and a temperature-controlled column compartment. Centrifuging of the plasma samples was performed using a Mikro 200R centrifuge (Hettich Lab Technology, Germany). Evaporation of the plasma samples was performed with the aid of a GeneVac HT-4X EZ-2 series evaporator Lyospeed ENABLED (Genevac Ltd, UK).

Study subjects-intervention study

The experimental groups consisted of 48 male Wistar rats (9-10 weeks old) supplied by the Biological Resources unit of the University of Queensland (Australia) and individually housed in a temperature-controlled ($20 \pm 2^\circ\text{C}$), 12-hour light/dark cycle environment with *ad libitum* access to water and the group-specific rat diet at the Animal House of the University of Southern Queensland. The rats were randomly divided into four separate groups ($n = 7$ each) and fed with corn starch (C; 330 ± 3 g), corn starch + HT (CHT; 324 ± 3 g), high-carbohydrate, high-fat (H; 336 ± 2 g), high-carbohydrate, high-fat + HT (HHT; 336 ± 2 g).

Body weight and food and water intakes were measured daily and feed efficiency (%) was calculated as described in a previous study⁹. The preparation and macronutrient composition of basal diets, including the dietary fatty acid profiles, have been described in detail previously⁹⁻¹¹. HT (20 mg/kg) dissolved in distilled water was administered daily by oral gavage for 8 weeks starting 8 weeks after the initiation of the corn starch or high-carbohydrate, high-fat diet in CHT and HHT groups, as depicted in Figure 2. The drinking water in all high-carbohydrate, high-fat diet-fed groups was augmented with 25% fructose for the duration of the study.

All experimentation was approved by the Animal Experimentation Ethics Committees of The University of Queensland and the University of Southern Queensland under the guidelines of the National Health and Medical Research Council of Australia.

Biological samples were acquired and analyzed from 28 animals.

Sample pretreatment

Plasma samples (n=7) from each group (C, CHT, H and HHT), 28 plasma samples in total, were prepared for the UHPLC- (-) ESI-HRMS analysis following a universal sample preparation protocol, as described in the literature¹². Thus, the rat plasma samples were initially allowed to thaw on ice at 4 °C for 30–60 min. Then, a 200 µL aliquot of each plasma sample were placed into a labeled 2.0-mL eppendorf tube, 600 µL of ice-cold MeOH were added, followed by vortexing for 15 s. The samples were subsequently centrifuged at 15800g for 15 min at room temperature to pellet the protein precipitate, 185 µL aliquots from the supernatant were transferred into appropriately labeled 2.0-mL eppendorf tubes and evaporated in the GeneVac for 90 min at 35°C with 500g rotor speed and using a ramp in the vacuum pressure (from 500 to 0 mbar). Subsequently 100 µL water were added to each dry sample, vortex for 15 s and centrifuged at 15800g for 15 min and transferred to 200 µL inserts prior to LC- (-) ESI-HRMS analysis for the investigation of the HT metabolism.

UHPLC- (-) ESI-HRMS Analysis

An ACQUITY UPLC BEH C18 (2.1 × 100 mm, 1.7 µm) reversed phase column (Waters Corp., Milford, MA, USA) preceded by a precolumn (Waters VanGuard 5 x 2.1 mm, 1.7 µm) was used for the chromatographic separation of the plasma samples.

The mobile phase consisted of solvents A: water, 0.1% (v/v) formic acid and B: methanol, 0.1% (v/v) formic acid. A gradient elution (total run time of 26 min) was performed for negative ion mode detection with a flow rate of 0.40 mL/min. In detail, the gradient methodology using the following elution program: 0 to 2 min: 100% A: 0% B, 2 to 17 min: from 100% A: 0% B to 0%A: 100%B, 17 to 22 min: 0% A: 100% B, 22 to 24 min: from 0% A: 100% B to 100%A: 0% B, 24 to 26 min: 100% A: 0% B. The column temperature was kept at 50 °C throughout all experiments while the autosampler tray temperature was set at 4 °C and the injection volume was 10 µL.

A hybrid HRMS (LTQ-Orbitrap Discovery) instrument operating in ESI negative ion mode was used and the conditions of the HRMS were set as following: capillary temperature, 456°C; capillary voltage, 11 V; tube lens, -48 V; spray voltage, 3.10 kV; sheath gas flow, 30 arb. units; aux gas flow, 10 arb. units. Analysis was performed in using a resolution of 30000, while acquisition of the mass spectra was performed in the centroid mode. The UHPLC-HRMS parameters used in the HT plasma metabolism study are summarized in Table 1.

LC parameters

Column: ACQUITY UPLC BEH C18 (2.1 × 100 mm, 1.7 µm) reversed phase

Mobile phase: water, 0.1% (v/v) formic acid (A) and methanol, 0.1% (v/v) formic acid (B) Flow rate: 0.4 mL min⁻¹

Column temperature: 50 °C

Tray temperature: 4 °C

Injection volume: 10 µL.

Gradient Program

Time (min)	A%	B%
0	100	0
2	100	0
17	0	100
22	0	100
24	100	0
26	100	0

MS parameters

ESI negative mode

30000 resolution

Centroid mode	
	Capillary temperature (°C): 356
	Capillary voltage (V): 20
	Tube lens (V): -49
	Spray voltage (kV): 3.10
	Sheath gas flow (arb. units): 30
	Aux gas flow (arb. units): 10

Table 1: UHPLC-HRMS (orbitrap) parameters used in the HT metabolism study. Gradient elution program and mass spectrometry parameters applied for UHPLC-HRMS (orbitrap) analysis in (-) ESI mode.

Data processing

Instrument control and data collection were performed using the Xcalibur[®] software version 2.0.7 (Thermo Fisher Scientific Inc., San Jose, USA). Metabolism of HT was investigated using the MetWorks[™] 1.3 software (Thermo Fisher Scientific, Inc., San Jose, CA, USA), which was used to query the possible metabolites of the molecule, based on mammalian phase I and phase II patterns of metabolism. Ninety-six in total possible metabolic modifications of HT were explored and evaluated. The aim of this procedure was to calculate the accurate mass values of the analyte and its putative metabolites by computing the corresponding due to metabolism mass shifts, and determine the presence of extracted ion chromatograms (XICs) from the raw data. The abovementioned resulting information, used for calculating the metabolic change for each metabolite, are its elemental composition, the rings plus double bonds (RDBs) and its isotope pattern. During data processing, MetWorks[™] software automatically generated XICs for each expected metabolite. Additionally, the isotopic pattern has been also analysed and compared to its theoretical values with a margin of 1% difference. Because the plasma samples were of biological origin, subtraction of the biological matrix (control plasma samples) from the treated plasma samples was performed prior to the detection of the potential HT metabolites. Thus, the samples

from C group were subtracted from the CCT group, whereas the samples from the H group were subtracted from the corresponding ones from the HHT group (Figure 1).

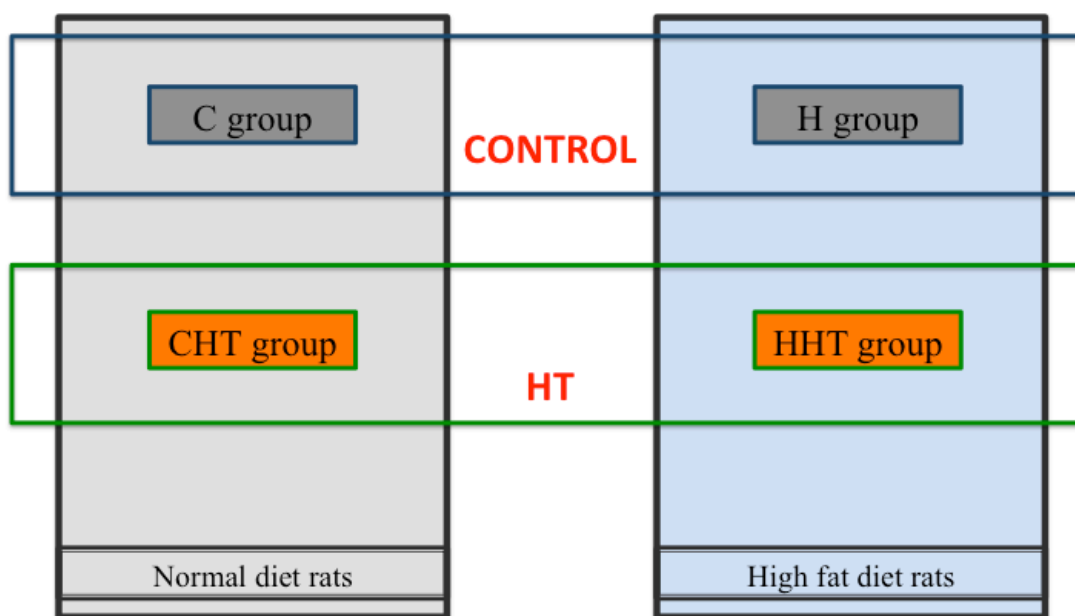


Figure 1: Subtraction workflow.

Following this approach, the chemical noise due to the biological matrix (control plasma samples) was removed from the treated plasma samples, as it is shown in Figure 2. The area of the possible HT metabolite XICs based on the raw data from the two different treated groups (CHT vs. HHT) were compared employing a t-test. The significant level has been set to 95%.

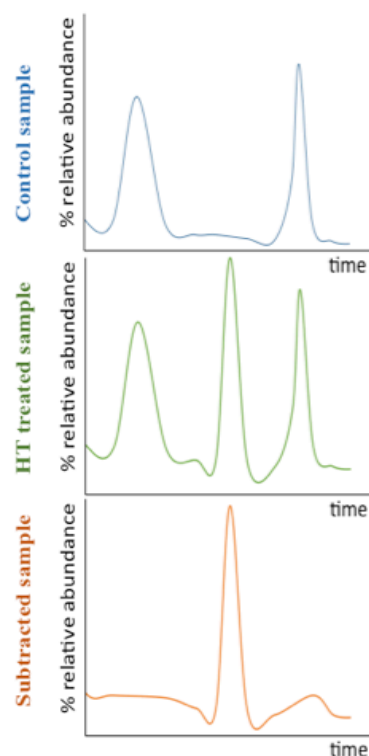


Figure 2: Subtraction scheme. A control sample in the upper part, a treated sample in the middle and the subtracted derived one in the down part of the Figure.

RESULTS AND DISCUSSION

The parent molecule (HT) and 96 of its possible metabolic modifications have been investigated. Twenty-five possible metabolites of HT have been recognized using the aforementioned methodology and are tabulated in Table 2 along with their accurate mass values, the predicted molecular formulas, the metabolic changes and metabolic generation (Phase I or II metabolism) for all the 25 metabolites. However, only six of them, namely the M6, M7, M9, M11, M20 and M24, have been found to be statistically different between the two treated groups CHT and HHT, according to the t-test results. More specifically, the significant metabolites were derived from sulfation, acetylation, combination of acetylation and sulfation of the HT, acetylation

of the homovanillic acid, transformation of homovanillic alcohol to aldehyde and glucuronidation of the 3,4-diphenylacetic acid. The *t*-test results for the six HT significant metabolites are shown in Table 3.

HT Metabolite	CHT	HHT	<i>t</i> -test <i>p</i> value
	mean \pm SD (n=7)	mean \pm SD (n=7)	
M6	21552 \pm 18251	NF	0.01
M7	1201201 \pm 124632	870219 \pm 317658	0.03
M9	959796 \pm 623547	230734 \pm 66197	0.02
M11	2543166 \pm 1251367	741307 \pm 376668	0.006
M20	92962 \pm 22802	56741 \pm 21508	0.008
M24	293331 \pm 55989	206004 \pm 71461	0.02

Table 3. *t*-test results for the six significant HT metabolites.

Id.	Metabolites	Metabolic Change shift ^a	Mass Shift (Da)	Theoretical [M-H ⁺] ⁻	Measured [M-H ⁺] ⁻	Δm (ppm)	RDB	t _R (min)	Molecular Formula [M-H ⁺] ⁻	Metabolic generation
M1	Hydroxytyrosol			153,0557	153,0559	8,1	4,5	7,18	C ₈ H ₉ O ₃	
M2	Hydroxytyrosol Double Oxidation	- 2 [H]	-2,0157	151,0401	151,0403	9	5,5	7,24	C ₈ H ₇ O ₃	Phase I
M3	Hydroxytyrosol 2-ethoxyl to acid	- [CH ₄] + [O]	-0,0364	153,0193	153,0196	8,8	5,5	6,66	C ₇ H ₅ O ₄	Phase I
M4	Hydroxytyrosol Glucuronidation	+ [C ₆ H ₈ O ₆]	176,0321	329,0876	329,0907	12,2	6,5	9,08	C ₁₄ H ₁₇ O ₉	Phase II
M5	Hydroxytyrosol Glutathione conjugation	+ [C ₁₀ H ₁₇ N ₃ O ₆ S]	307,0838	460,1395	460,1459	16,2	7,5	1,56	C ₁₈ H ₂₆ O ₉ N ₃ S	Phase II
M6	Hydroxytyrosol Sulfation	+ [SO ₃]	79,9568	233,0125	233,0123	3,8	4,5	4,72 & 4,95	C ₈ H ₉ O ₆ S	Phase I
M7	Hydroxytyrosol Acetylation	+ [COCH ₂]	42,0106	195,0663	195,064	-6,2	5,5	7,37	C ₁₀ H ₁₁ O ₄	Phase II

Id.	Metabolites	Metabolic Change shift ^a	Mass Shift (Da)	Theoretical [M-H ⁺] ⁻	Measured [M-H ⁺] ⁻	Δm (ppm)	RDB	t _R (min)	Molecular Formula [M-H ⁺] ⁻	Metabolic generation
M8	Hydroxytyrosol N- acetylcysteine conjugation	+ [C ₅ H ₇ NO ₃ S]	161,0147	314,0704	314,0754	-6,2	5,5	7,28	C ₁₃ H ₁₆ NO ₆ S	Phase II
M9	Hydroxytyrosol Acetylation + Sulfation	+ [C ₂ H ₂ SO ₄]	121,9674	275,0231	275,0228	3,1	5,5	6,24	C ₁₀ H ₁₁ O ₇ S	Combination
M10	Hydroxytyrosol Methylation (Homovanillic alcohol)	+ [CH ₂]	14,0157	167,0714					C ₉ H ₁₁ O ₃	Phase I
M11	Homovanillic alcohol First alcohol to aldehyde	- [H ₂]	-2,0157	165,0557	165,0558	7,3	5,5	7,5	C ₉ H ₉ O ₃	Phase I
M12	Homovanillic alcohol Sulfation	+ [SO ₃]	79,9568	247,0282	247,0279	3,5	4,5	6,01 & 6,15	C ₉ H ₁₁ O ₆ S	Phase I
M13	Homovanillic alcohol Methylation	+ [CH ₂]	14,0157	181,0870	181,0868	4,7	4,5	9,14	C ₁₀ H ₁₃ O ₃	Phase I

Id.	Metabolites	Metabolic Change shift ^a	Mass Shift (Da)	Theoretical [M-H ⁺] ⁻	Measured [M-H ⁺] ⁻	Δm (ppm)	RDB	t _R (min)	Molecular Formula [M-H ⁺] ⁻	Metabolic generation
M14	Homovanillic alcohol Acetylation	+ [COCH ₂]	42,0106	209,0819	209,0977	-6,5	5,5	8,24	C ₁₁ H ₁₃ O ₄	Phase II
M15	Homovanillic acid	+ [CO]	27,9949	181,0506	181,0507	6,2	5,5	5,57	C ₉ H ₉ O ₄	Phase II
M16	Homovanillic acid Aromatic hydroxylation	+ [O]	15,9949	197,0455	197,0431	-6,8	5,5	6,2	C ₉ H ₉ O ₅	Phase I
M17	Homovanillic acid Glucuronidation	+ [C ₆ H ₈ O ₆]	176,0321	357,0827	357,0822	1,5	7,5	4,59 & 5,02	C ₁₅ H ₁₇ O ₁₀	Phase II
M18	Homovanillic acid Sulfation	+ [SO ₃]	79,9568	261,0074	261,0072	3,4	5,5	5,14 & 5,93	C ₉ H ₉ O ₇ S	Phase I
M19	Homovanillic acid Methylation	+ [CH ₂]	14,0157	195,0663	195,064	-6,2	5,5	7,37	C ₁₀ H ₁₁ O ₄	Phase II
M20	Homovanillic acid Acetylation	+ [COCH ₂]	42,0106	223,0612	223,0587	-6,1	6,5	6,35	C ₁₁ H ₁₁ O ₅	Phase I

Id.	Metabolites	Metabolic Change shift ^a	Mass Shift (Da)	Theoretical [M-H ⁺] ⁻	Measured [M-H ⁺] ⁻	Δm (ppm)	RDB	t _R (min)	Molecular Formula [M-H ⁺] ⁻	Metabolic generation
M21	Homovanillic acid Glycine conjugation (carboxylic acid)	- [OH] + [C ₂ H ₄ NO ₂]	57,0215	238,0721	238,0721	2,4	6,5	1,64 & 6,97	C ₁₁ H ₁₁ NO ₅	Phase II
M22	Homovanillic acid Hydroxylation + Methylation	+ [CH ₂ O]	30,0106	211,0612	211,0611	2,3	5,5	7,42	C ₁₀ H ₁₁ O ₅	Phase II
M23	3,4-Diphenylacetic acid	- [H ₂] + [O]	13,9793	167,0350	167,0352	4,3	5,5	7,37	C ₈ H ₇ O ₄	Phase I
M24	3,4-Diphenylacetic acid Glucuronidation	+ [C ₆ H ₈ O ₆]	176,0321	343,0671	343,0617	-13,9	7,5	5,26	C ₁₄ H ₁₅ O ₁₀	Phase II
M25	3,4-Diphenylacetic acid Glycine conjugation (carboxylic acid)	- [OH] + [C ₂ H ₄ NO ₂]	57,0215	224,0564	224,0563	2	6,5	6,46	C ₁₀ H ₁₀ NO ₅	Phase II

^a - indicates loss, while + indicates addition of the respective moieties

Table 2: Possible metabolites of HT found in plasma of the two different treated groups CHT and HHT. The accurate mass values, the predicted molecular formulas, the metabolic changes and the metabolite generation for the twenty-four metabolites are presented.

It should be noted that no absolute quantitative analysis has been attempted due to the lack of the corresponding standards substances. As shown in Table 2, there were nine phase I metabolites (M2, M3, M6, M10, M11, M12, M16, M18 and M20), fourteen phase II metabolites (M4, M5, M7, M8, M13, M14, M15, M17, M19, M21, M22, M23, M24 and M25) and one combination of phase I and II (M9). Their proposed structures are given in Figure 3, which were justified by the accurate mass measurements, the RDB values and the isotopic pattern ratio.

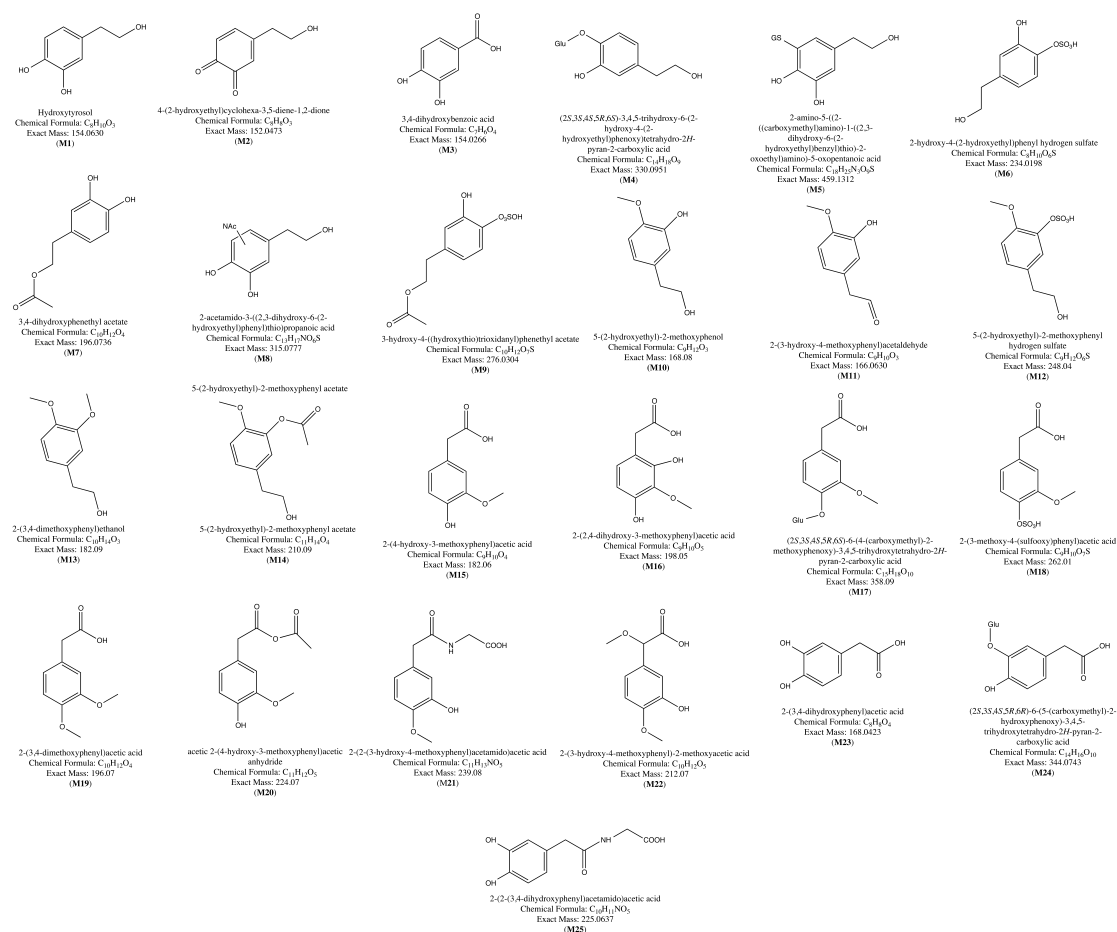


Figure 3: Proposed structures of the possible HT metabolites recognized by the developed UHPLC(-) ESI-HRMS-MetWorks methodology. Chemical structures, IUPAC names, corresponding chemical formulas and exact masses are given.

A comparative study of the significant metabolites in the two different treated groups (CHT and HHT) is depicted in Figure 4. It should be noted, that the heights of the bars correspond to the area of the corresponding metabolites found in the different groups, whereas the error bars signify the uncertainty of the measurements in terms of mean of standard error. Notably, all the significantly differentially secreted metabolites between the two groups were higher in the CHT group compared to the HHT group. A possible explanation could be either due to differences of the HT distribution between the examined groups or due to reduced metabolic capacity of the HHT group (obese rats). In the first hypothesis, HT could be distributed in higher amounts to the tissues - a pharmacodynamics effect. This could result in lower amounts of circulating HT, which presumably lead to lower concentration of its corresponding metabolites. For example, the nearly double percentage of adipose tissue in the HHT rats could indicate that HT is distributed in higher amounts to it. In the second hypothesis, the HHT rats exist in a disease state having a nearly cirrhotic liver, which could result in reduced metabolism efficiency.

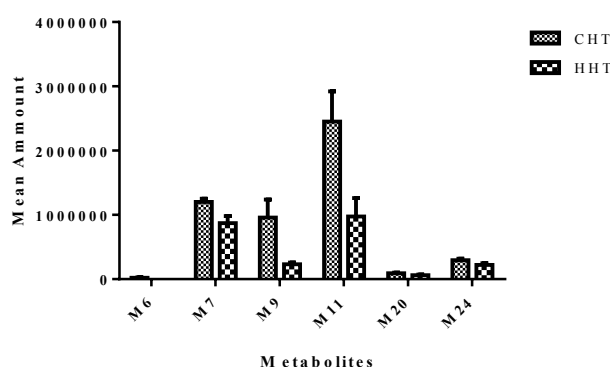


Figure 4: Mean amount of the significant metabolites in the two different treated groups CHT and HHT. The heights of the bars correspond to the area of the corresponding metabolites found in the

different groups. The error bars signify the uncertainty of the measurements in terms of the mean of standard error.

Three points should be also taken under consideration. The sulphate and the glucuronide modifications are generated probably for facilitating the excretion of HT whereas the acetylation reactions are occurring for the detoxification of the organism. Furthermore, the methoxy- analogs are produced by the action of the catechol-O-methyl transferase (COMT) enzymes whereas the aldehyde containing molecules could be produced from the action of monoamine oxidase (MAO). In addition, the acidic metabolites are potentially generated by the action of aldehyde dehydrogenase. All the three enzymes are part of catecholamine metabolism and therefore are broadly distributed in the body. Finally, it should be noted that the homovanillic moieties, such as the homovanillic aldehydes are produced in higher amounts. This is evident from their high areas found, albeit they exhibit lower ionization efficiency compared to dihydroxy counterparts.

CONCLUSION

In the current study progress was made in identifying possible HT metabolites circulating in rat plasma, after 8-weeks treatment with HT in normal (C group) as well as high fat diet (H group) subjects. This animal model can be considered as closely resembling the human MetS. The application of the HRMS-based metabolism approach for the clarification of the HT metabolic fate, revealed the occurrence of the extended metabolism of the target molecule. It has been possible to recognize 25 HT phase I and II metabolites, six of them over expressed in the CHT group (normal diet

rats). Further investigation concerning possible accumulation to the tissues is needed, in order to gain deeper insight on the HT bioavailability.

REFERENCES

1. Ma, S. & Chowdhury, S. K. Data acquisition and data mining techniques for metabolite identification using LC coupled to high-resolution MS. *Bioanalysis* **5**, 1285–97 (2013).
2. Pähler, A. & Brink, A. Software aided approaches to structure-based metabolite identification in drug discovery and development. *Drug Discov. Today. Technol.* **10**, e207–17 (2013).
3. Vissers, M. N., Zock, P. L. & Katan, M. B. Bioavailability and antioxidant effects of olive oil phenols in humans: a review. *Eur. J. Clin. Nutr.* **58**, 955–65 (2004).
4. Rubió, L. *et al.* A new hydroxytyrosol metabolite identified in human plasma: hydroxytyrosol acetate sulphate. *Food Chem.* **134**, 1132–6 (2012).
5. Suárez, M. *et al.* Bioavailability of phenols from a phenol-enriched olive oil. *Br. J. Nutr.* **106**, 1691–701 (2011).
6. De la Torre, R. Bioavailability of olive oil phenolic compounds in humans. *Inflammopharmacology* **16**, 245–7 (2008).
7. Serra, A. *et al.* Distribution of olive oil phenolic compounds in rat tissues after administration of a phenolic extract from olive cake. *Mol. Nutr. Food Res.* **56**, 486–96 (2012).
8. Levsen, K. *et al.* Structure elucidation of phase II metabolites by tandem mass spectrometry: an overview. *J. Chromatogr. A* **1067**, 55–72 (2005).
9. Poudyal, H., Panchal, S. K., Waanders, J., Ward, L. & Brown, L. Lipid redistribution by α -linolenic acid-rich chia seed inhibits stearoyl-CoA desaturase-1 and induces cardiac and hepatic protection in diet-induced obese rats. *J. Nutr. Biochem.* **23**, 153–62 (2012).
10. Panchal, S. K. *et al.* High-carbohydrate, high-fat diet-induced metabolic syndrome and cardiovascular remodeling in rats. *J. Cardiovasc. Pharmacol.* **57**, 611–24 (2011).
11. Poudyal, H., Campbell, F. & Brown, L. Olive leaf extract attenuates cardiac, hepatic, and metabolic changes in high carbohydrate-, high fat-fed rats. *J. Nutr.* **140**, 946–53 (2010).
12. Dunn, W. B. *et al.* Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* **6**, 1060–83 (2011).

7

Metabolomics-based time series analysis of hydroxytyrosol in rat plasma

This chapter is under preparation for publication

INTRODUCTION

Hydroxytyrosol (HT) or 3,4-dihydroxyphenyl-ethanol is one of the main components of virgin olive oil and olive mill waste¹ that has demonstrated the strongest radical-scavenging properties *in vitro* among all the olive oil polyphenols. In the previous study (Chapter “Metabolomics study of hydroxytyrosol’s administration for the treatment of metabolic syndrome in a rat model”), it has been demonstrated that supplementation with 20 mg/kg/daily of purified HT for 8 weeks in a rat metabolic syndrome (MetS) model, improved dramatically the symptoms of the syndrome. These results suggested that HT could be a natural antioxidant with a possible role in the prevention of MetS and thus its absorption profile should be studied.

According to the literature many studies describing an effect of olive oil polyphenols and especially HT, have been carried out with nutritional doses²⁻⁵. However, less attention has been paid to its effects when administered as a pure substance⁶⁻⁹.

Metabolomics datasets are becoming more and more complex. It is not uncommon to measure a multiple of metabolites in body fluids of several animals at different points in time with an underlying experimental design, e.g. different dose groups^{10,11}. This calls for data analysis methods specifically suited for time-resolved (or ‘longitudinal’), multigroup, multisubject (containing data of multiple animals) and multivariate data¹².

Besides the classical pharmacokinetic approach (PK) - the measurement of the target molecule in plasma, its effect on the organism - how the molecule affects the organism as a whole, represents the new scientific approach. Thus, the aim of the

current study was to explore the systems biology of the administered subject, from a metabolomics perspective besides the classical PK approach that has been also investigated.

MATERIAL AND METHODS

Isolation of hydroxytyrosol

HT used for the completion of the experiments, was isolated from olive mill wastewater, as described in the Chapter “Metabolomics study of hydroxytyrosol’s administration for the treatment of metabolic syndrome in a rat model”.

Pharmacokinetic study design

The PK was conducted in a clinical environment under special supervision, in the Biological Resources unit of the University of Queensland, Australia, under supervision of Prof. Lindsay Brown. In this study, 100 matched rats (n=50 rats/group) at the age 8-10 weeks and of 350-380 g weight, were used. The subjects were individually housed in a temperature-controlled ($20 \pm 2^{\circ}\text{C}$), 12-hour light/dark cycle environment with *ad libitum* access to water and diet at the Animal House of the University of Southern Queensland. Prior to the experiment the rats were fasted overnight.

The animals were randomly divided in two groups, the treated and the control group. Hydroxytyrosol was administered via oral gavage at the dose of 20 mg/kg in the treated group, whereas the control group was administered with the vehicle (water). The rats from both groups were sacrificed at 0.08, 0.25, 0.5, 1, 2, 4, 6, 8, 10,

and 12 h after hydroxytyrosol administration (n=5/time point). All blood samples were harvested for every time point, appropriately stored and subsequently plasma samples were prepared. Thus, after the completion of the PK protocol, blood samples were collected into appropriate heparinized tubes, centrifuged at 5,000 g for 15 min within 30 min of the collection. Plasma was separated and transferred to microcentrifuge tubes for storage at -80 °C before analysis.

Samples for the PK study

A hundred rats were used for the PK study (n=50/group, n=5/time point) and the rat plasma samples were appropriately marked.

Instrumentation

Two approaches have been employed, namely a targeted approach for the quantification of the analyte (HT) and an untargeted approach for the metabolomic PK study. Thus, for the targeted PK study an initial screening methodology has been developed in order to decipher the existence of circulating HT in the rat plasma. This methodology has been based on an LC-(-)ESI-HRMS/MS method exploiting a hybrid LTQ Orbitrap Discovery XL mass spectrometer (Thermo Scientific, Germany) coupled to an Accela UHPLC system (Thermo Scientific, Germany) and applying the selected reaction monitoring (SRM) methodology. For the metabolomic PK study, an Acquity UPLCTM system (Waters Corp., Milford, MA, USA) equipped with a binary solvent manager and a temperature and light-controlled autosampler (sample manager) capable of maintaining the sample temperature from -5 to 40 °C and a column heater, whereas high-resolution mass spectrometry was performed on a Waters Micromass (Milford, USA) hybrid quadrupole-time of flight mass

spectrometer (QTOF Premier) equipped with an electrospray ionization (ESI) interface. Centrifuging of the plasma samples was performed by a Mikro 200R centrifuge (Hettich Lab Technology, Germany). Evaporation of the plasma samples was performed with the aid of a GeneVac HT-4X EZ-2 series evaporator Lyospeed ENABLED (Genevac Ltd, UK).

Sample pretreatment for the targeted quantification of HT in plasma

For the targeted PK study, 20 representative plasma samples from all corresponding time points 0.08, 0.25, 0.5, 1, 2, 4, 6, 8, 10, and 12 h, were prepared for the UHPLC-(-)ESI-HRMS/MS analysis. Sample preparation has been accomplished using a solid phase extraction (SPE) methodology¹³. Briefly, the cleanup of the plasma samples was performed using Oasis HLB® Cartridges (30 mg, 1 cm³) obtained from Waters (Milford, MA, USA). Each cartridge was conditioned with 500 µL MeOH and equilibrated with 1000 µL of water. A 65 µL aliquot of plasma sample was carefully loaded onto the cartridge along with 10 µL of an 1 µg/mL IS solution. A washing step with 200 µL of water 0.1%(v/v) formic acid solution followed, and the SPE cartridge was dried under vacuum for 1 min. The analytes were eluted with 1000 µL of ACN:MeOH (50:50) and the eluate was evaporated to dryness by means of the GeneVac apparatus. The residue was reconstituted in 65 µL initial mobile phase [aq. 0.1% formic acid:ACN, 9:1 (v/v)] and was analyzed with the developed LC-ESI-HRMS/MS method described below.

Sample pretreatment for the metabolomics PK study

The rat plasma samples were prepared for the metabolomics PK analysis, following a universal plasma sample preparation protocol, described below. Thus, the

plasma samples were allowed initially to thaw on ice at 4 °C for 30–60 min. Following, a 200 µL aliquot from each plasma sample were placed into a labeled 1.5 mL eppendorf tube and 600 µL of MeOH were added followed by vortexing for 15 s. The samples were subsequently centrifuged at 15800 g for 15 min at room temperature to pellet the protein precipitate, and 185 µL aliquots from the supernatant were transferred into two separate labeled 1.5 mL eppendorf tubes and finally evaporated in the GeneVac for 90 min at 50°C. Consequently, 100 µL of water was added to each dry sample, vortexed for 15 s, centrifuged at 15800 g for 15 min and transferred to 200 µL inserts placed in appropriate screw capped autosampler vials.

Targeted quantification of HT in rat plasma

An RP-C18 Hypersil Gold column (50 x 2.1 mm, 1.9 µm) (Supelco Analytical, Germany) was used at a flow rate of 0.45 mL/min for the chromatographic separation of HT and IS. The mobile phase consisted of A: aq. 0.1% formic acid, (v/v) and B: ACN. An isocratic method (total run time of 1.1 min) was used in 30:70 ratio (A:B). The column temperature was kept at 40 °C throughout all experiments while the autosampler tray temperature was set at 4 °C. The injection volume was 5 µL. The ESI-HRMS parameters (capillary temperature, capillary voltage, spray voltage, sheath gas and auxiliary gas) were optimized for the SRM mode in order to achieve optimal sensitivity and selectivity. Thus the HRMS/MS was operated in the negative ion mode under the following conditions: capillary temperature, 356°C; capillary voltage, 11 V; tube lens, -49 V; source voltage, 3.10 kV; sheath gas flow, 30 arb. units; aux gas flow, 10 arb. units. Analysis was performed in the SRM mode, using a resolution of 30000 whereas the isolation width was set at 2 amu. The IS methodology has been employed for the quantification of HT, based to the calculation of the peak-area ratio of the

analyte vs. that of the IS. All compounds were analyzed as their respective deprotonated ions $[M-H]^-$.

Stock and working solutions of the analyte (HT) and the IS, were prepared daily and stored in the dark at 4 °C. A partial validation scheme has been used for the quality control of the results. The dynamic range of the HT calibration curve was 0.001 - 5 µg/mL (0.001, 0.005, 0.01, 0.05, 0.075, 0.1, 0.4, 0.5, 1, 2, 3 and 5), whereas the concentration of IS was kept at 1 µg/mL. Three concentration levels were prepared, in order to be used as QC samples namely a low (LQC 0.013 µg/mL), a medium (MQC 0.75 µg/mL) and a high (HQC 2.5 µg/mL).

Quantification was performed employing the SRM negative ion mode for HT (MW: 153.1837 g/mol) and IS (MW: 184.0115 g/mol). The SRM transitions selected were m/z 153.1 \rightarrow 123.1 (for HT) and m/z 183.0 \rightarrow 153.0 (for IS) and the fragmentation was optimized by adjustment of the normalized collision energy using the collision induced dissociation (CID) option.

Metabolomics PK study

An ACQUITY UPLC BEH C18 (2.1 × 100 mm, 1.7 µm) reversed phase column (Waters Corp., Milford, MA, USA) preceded by a precolumn (Waters VanGuard 5 × 2.1 mm, 1.7 µm) was used for the chromatographic separation in the PK plasma samples. The mobile phase consisted of solvents A: aq. 0.1% formic acid, (v/v) and B: methanol 0.1% formic acid, (v/v). A gradient elution was performed, with a flow rate of 0.40 mL/min.

In detail, a gradient method (total run time of 26 min) was used as follows: 0 to 2 min: 100% A: 0% B, 2 to 17 min: from 100% A: 0% B to 0%A: 100%B, 17 to 22 min: 0% A: 100% B, 22 to 24 min: from 0% A: 100% B to 100%A: 0% B, 24 to 26

min: 100% A: 0% B. Column temperature was kept at 50 °C throughout all experiments while the autosampler tray temperature was set at 4 °C and the injection volume was 10 µL.

A qQToF mass spectrometer has been used as follow: the electrospray voltage was kept at 2.5 kV for the (-) ESI mode, the sample cone voltage was maintained at 35 V affording the highest sensitivity, whereas the extraction cone voltage was set at 3 V and the MCP plates were operated at 1850 V. Nitrogen was used both as the desolvation and cone gas and was set at 650 and 50 L/h, respectively. The desolvation temperature was set at 350 °C and the source temperature was 120 °C. The TOF analyzer was operated in the V optics mode affording a resolution of 9500. The collision gas was argon. For the acquisition of the full-scan mass spectra the collision energy was set at 2 V. Full-scan mass spectra were acquired in continuum and centroid mode in the mass range of 50–1000 amu, whereas the scan time was 0.4 s with an inter scan delay of 0.02 s. In order to improve mass measurement accuracy a lock mass compound (500 pg/µL solution of leucine enkephalin; m/z 556.2771) was continuously infused at a 2 µL/min using the built-in syringe pump of the instrument.

Time Series analysis

In order to identify the change of metabolome *vs.* time, a time series analysis has been employed with the aid of the MetaboAnalyst 2.0 web-based software (<http://www.metaboanalyst.ca>). The data preprocessing has been conducted in the same manner as described in Chapter “Metabolomics study of hydroxytyrosol’s administration for the treatment of metabolic syndrome in a rat model”. Thus, baseline correction, peak detection, deconvolution, deisotoping, alignment and gap filling procedures were initially performed in MZmine 2.10 software

(<http://mzmine.sourceforge.net>), following by data manipulation in Microsoft Excel 2012. The generated peak list was finally imported to the MetaboAnalyst platform for the data processing (missing value estimation, normalization, scaling, transformation) and for the completion of the time series analysis.

Using the Heatmap visualization the correlation of the metabolites' concentration *vs.* time can be accomplished. Thus the Pearson correlation has been used as the distance measure whereas the average linkage has been used as the clustering algorithm. Following, the analysis of variance-simultaneous component analysis (ASCA) has been employed, in order to facilitate the monitoring of the significance of the metabolome change *vs.* time. In addition, the multivariate empirical Bayes analysis of variance (MB-MANOVA) has been employed for the comparison of time-course profiles of significant metabolites.

RESULTS AND DISCUSSION

Targeted quantification of HT in rat plasma

A novel method has been proposed for the quantification of HT in plasma, using an UHPLC-(-)ESI-HRMS/MS methodology. This methodology is characterized by an extremely short analysis time, i.e. 1.1 min overall time, along with high specificity facilitated by the resolving power from the mass spectrometer (orbitrap). It should be also noted that the method exhibited high sensitivity with an LLOQ of 1 ng/mL, a fact that is considered as crucial for the analysis of HT. Nevertheless, measurements of the rat plasma samples after administration of HT as described before, did not reveal the existence of circulating HT in concentrations above LLOQ. This fact can be

presumably attributed to the extensive metabolism of the target molecule. In order to alleviate this pitfall, a metabolomics PK approach has been attempted.

Metabolomics PK study

The main goal of this untargeted approach was to examine the possible alteration of the total plasma metabolome, after the administration of HT, since it was found that no HT was circulated in rat plasma. In order to clarify if the molecule has some kind of effect on the rat's organism or whether is only excreted with a minimal perturbation or extensively distributed, a metabolome study was considered essential. Moreover the time course of the metabolome change could potentially reveal significant information. Thus the time series approach has been employed to the dataset in order to investigate this assumption.

In a first step, a Heatmap has been constructed showing the variance of its metabolite vs. time (Figure 1).

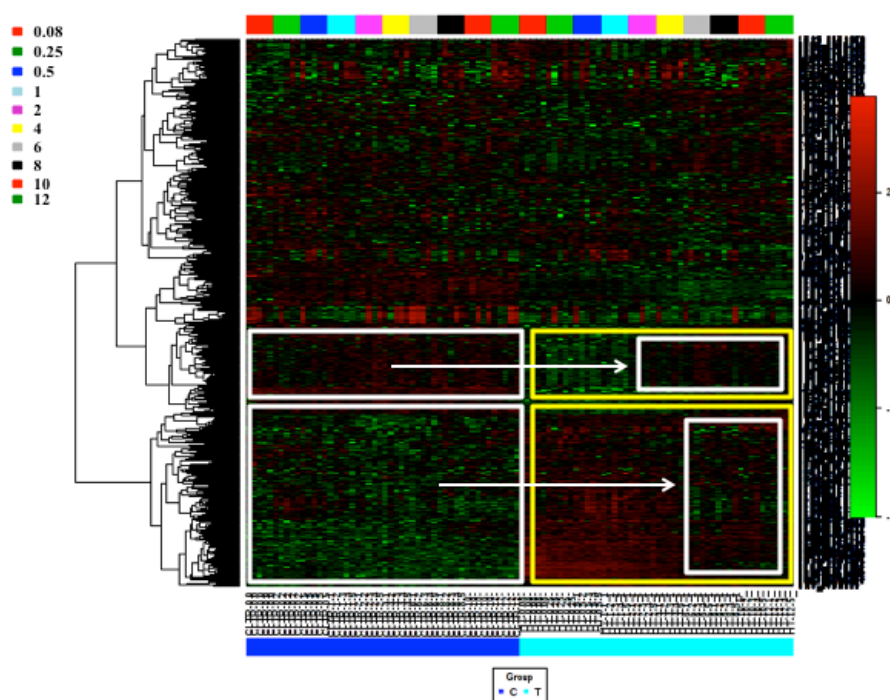


Figure 1. Heatmap representation of the metabolome time progression. The X-axis (down) represents the two groups, i.e. C control (blue) and T treated (light blue). The X-axis (upper) represents the time points, i.e. 0.08, 0.25, 0.5, 1, 2, 4, 6, 8, 10, and 12 h (see color coding at the left legend). The Y-axis (right) represents the significant metabolite ids whereas the left one is the dendrogram representation of the metabolites' clustering. The color-coding at the right legend represents the up- and down-regulated levels of the metabolites. The C group metabolome seems to be markedly different from the corresponding of the T group. The metabolome in the C group seems to be constant with time in contrast with the T group, where it changes rapidly after the administration of HT but seems to reset gradually in time resembling that of the C group. The differences along with the “metabolome reset” are highlighted in white and yellow boxes. The inserted white boxes in the yellows, signifies the metabolome reset.

It is evident that there are significant differences between the control (C) and the treated group (T). Furthermore, it is readily observable that the metabolome of the C group changes insignificantly in time (0-12 h), which does not hold for the metabolome of the T group. Hence, it is easy to observe that there is a highly significant change in the metabolome immediately after the administration of the HT in the T group. It is suppressible, that although there is no circulating HT in the rat plasma (targeted quantification), its effect on the plasma metabolome is enormous (Figure 1).

It is also interesting, that this change in the metabolome of the T group resets with the progress of time (after 12 h) to the normal level, equal to the C group. This observation signifies the high impact of HT to the metabolic profiles of the rats.

In order to identify the metabolites that have been changed most significant, the MB-MANOVA approach has been applied. Table 1 and Figure 2 shows the 6 most significant metabolites that have been changes during the time. From the results, it is clear that the significant metabolites have been rapidly up- or- down regulated in the T group, but this concentration change seems to be vanishing with the course of time (Figure 2).

Metabolites	Hotelling-T ²
17.76_483.2941	1745,2
19.96_885.5422	1566,3
17.89_361.2344	1507,8
13.97_217.6382	1168,3
12.94_588.3240	877,39
16.00_568.3573	829,24

Table 1. The six most significant metabolites identified by the MB-MANOVA approach. Hotelling-T² distribution, tests for the univariate distribution between the multivariate means of different populations.

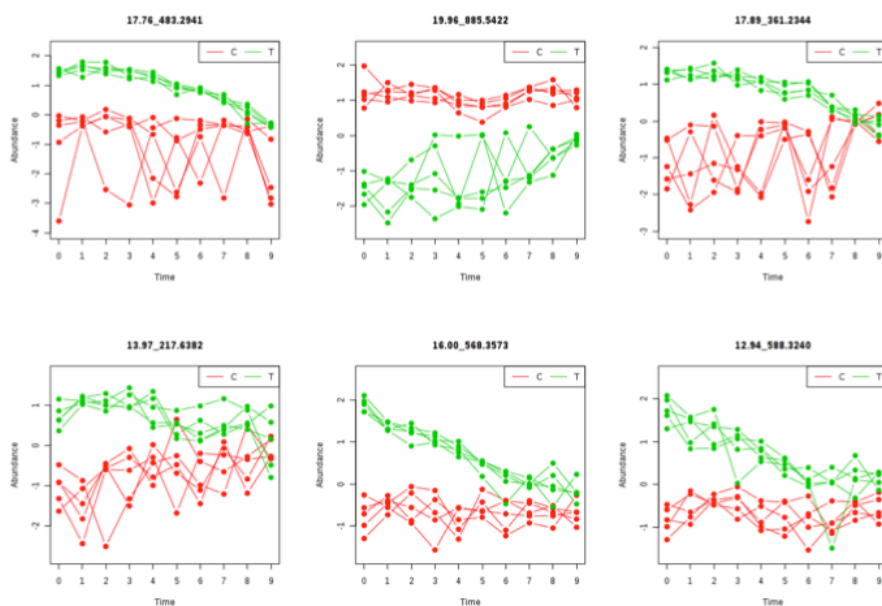


Figure 2. MB-MANOVA diagrams of the six most significant metabolites. Green line represents the T group metabolites vs. time whereas the red one represents the C group metabolites vs. time. Up- and down-regulation of the metabolites is shown.

CONCLUSION

After HT administration in rats, no appreciable amount of the molecule seems to be circulating in the plasma, denoting extensively excretion or distribution.

Nevertheless, it seems that HT interacts significantly with the plasma metabolome, as the last changes rapidly after the administration. Furthermore, this effect seems to be diminishing with time when the organism returns to its previous metabolic state.

REFERENCES

1. Owen, R. W. *et al.* Phenolic compounds and squalene in olive oils: the concentration and antioxidant potential of total phenols, simple phenols, secoiridoids, lignans and squalene. *Food Chem. Toxicol.* **38**, 647–59 (2000).
2. De La Cruz, J. P. *et al.* Antithrombotic potential of olive oil administration in rabbits with elevated cholesterol. *Thromb. Res.* **100**, 305–15 (2000).
3. Granados-Principal, S. *et al.* Hydroxytyrosol ameliorates oxidative stress and mitochondrial dysfunction in doxorubicin-induced cardiotoxicity in rats with breast cancer. *Biochem. Pharmacol.* (2014). doi:10.1016/j.bcp.2014.04.001
4. Granados-Principal, S., Quiles, J. L., Ramirez-Tortosa, C. L., Sanchez-Rovira, P. & Ramirez-Tortosa, M. C. Hydroxytyrosol: from laboratory investigations to future clinical trials. *Nutr. Rev.* **68**, 191–206 (2010).
5. Ochoa, J. J., Quiles, J. L., Ramirez-Tortosa, M. C., Mataix, J. & Huertas, J. R. Dietary oils high in oleic acid but with different unsaponifiable fraction contents have different effects in fatty acid composition and peroxidation in rabbit LDL. *Nutrition* **18**, 60–5 (2002).
6. Sarsour, E. H., Goswami, M., Kalen, A. L., Lafin, J. T. & Goswami, P. C. Hydroxytyrosol inhibits chemokine C-C motif ligand 5 mediated aged quiescent fibroblast-induced stimulation of breast cancer cell proliferation. *Age (Dordr)*. (2014). doi:10.1007/s11357-014-9645-0
7. Wiseman, S. A., Mathot, J. N., de Fouw, N. J. & Tijburg, L. B. Dietary non-tocopherol antioxidants present in extra virgin olive oil increase the resistance of low density lipoproteins to oxidation in rabbits. *Atherosclerosis* **120**, 15–23 (1996).
8. Zhao, B. *et al.* Hydroxytyrosol, a natural molecule from olive oil, suppresses the growth of human hepatocellular carcinoma cells via inactivating AKT and nuclear factor-kappa B pathways. *Cancer Lett.* **347**, 79–87 (2014).
9. Tutino, V., Caruso, M. G., Messa, C., Perri, E. & Notarnicola, M. Antiproliferative, antioxidant and anti-inflammatory effects of hydroxytyrosol on human hepatoma HepG2 and Hep3B cell lines. *Anticancer Res.* **32**, 5371–7 (2012).
10. Keun, H. C. *et al.* Geometric trajectory analysis of metabolic responses to toxicity can define treatment specific profiles. *Chem. Res. Toxicol.* **17**, 579–87 (2004).
11. Lamers, R.-J. A. N. *et al.* Identification of disease- and nutrient-related metabolic fingerprints in osteoarthritic Guinea pigs. *J. Nutr.* **133**, 1776–80 (2003).
12. Smilde, A. K. *et al.* ANOVA-simultaneous component analysis (ASCA): a new tool for analyzing designed metabolomics data. *Bioinformatics* **21**, 3043–8 (2005).
13. Bazoti, F. N., Gikas, E. & Tsarbopoulos, A. Simultaneous quantification of oleuropein and its metabolites in rat plasma by liquid chromatography electrospray ionization tandem mass spectrometry. *Biomed. Chromatogr.* **24**, 506–15 (2010).

8

Impact of a functionalized olive oil extract on the uterus and the bone in a model of postmenopausal osteoporosis

This chapter is published as: Keiler AM, Zierau O, Bernhardt R, Scharnweber D, Lemonakis N, Termetzi A, Skaltsounis L, Vollmer G, Halabalaki M

Eur J Nutr. 2013 Oct 30. DOI 10.1007/s00394-013-0609-4.

My contribution is on the isolation, preparation and characterization of the Total Polyphenolic Fraction from Extra Virgin Olive Oil

ABSTRACT

Purpose

The Mediterranean diet rich in fruits, vegetables and olive oil has been related to a lower osteoporosis incidence and accordingly to a reduced fracture risk. These observations might be mediated by the active constituents of extra virgin olive oil, and especially polyphenols. In the context of exploring the features of olive oil active constituents on postmenopausal osteoporosis, an extra virgin olive oil total polyphenolic fraction (TPF) was isolated and its effect on the bone loss attenuation was investigated.

Methods

Female Lewis rats were ovariectomized and fed a diet enriched with a total phenolic extract of extra virgin olive oil in a concentration of 800 mg/kg diet.

Results

Oleocanthal, one compound of the polyphenolic fraction, showed a higher relative estrogen receptor binding affinity to the ER α compared to the ER β . While the TPF only slightly induced the uterine wet weight (490.7 ± 53.7 vs. 432.7 ± 23 , $p = 0.058$), TPF regulated estrogen response genes in the uterus (progesterone receptor, anti- gen identified by monoclonal antibody Ki67, complement C3). Comparing the quantified bone parameters, the oral TPF substitution did not attenuate the ovariectomy-induced bone loss.

Conclusions

The administration of extra virgin olive oil polyphenols regulated uterine estrogen response marker genes in an E2-agonistic manner. The bone loss induced by estrogen ablation was not mitigated by treatment with the polyphenolic extract.

KEYWORDS

Extra Virgin Olive oil, osteoporosis, polyphenolic fraction, oleacein, oleocanthal

ABBREVIATIONS

EVOO: Extra Virgin Olive Oil

TPF: Total Polyphenolic Fraction

INTRODUCTION

The so-called Mediterranean diet is characterized by a high intake in fruits, vegetables and olive oil. Beneficial effects on diseases like cancer^{1, 2}, cardiovascular complaints^{3, 4} and metabolic syndrome^{5, 6} are ascribed to adherence to this nutrition. Besides, the incidence of osteoporosis and associated fractures is found to be lower in countries where the Mediterranean diet is predominant^{7, 8, 9}. Extra Virgin Olive Oil (EVOO), as a key constituent of the Mediterranean diet has been shown to exert several pharmacological properties. Specifically, it has been shown to possess cardioprotective effects¹⁰, anti-inflammatory properties, as well as promoting effect to endothelial function and insulin sensitivity¹¹. Several human studies showed positive effects of phenol-enriched olive oil on oxidative plasma markers, with an effective phenol concentration between 150 and 800 mg/kg^{12, 13}. EFSA recently published an opinion, concerning the health beneficial effects of EVOO consumption.

Today, it is known that the health promoting effects of EVOO could be attributed to its high content in small molecules (MW < 1000). More than 250 small molecules have been identified in olive oil so far while most of them belong to the chemical group of phenols generally referred as olive polyphenols. The most known compounds of olive phenols are the phenyl alcohols tyrosol (**1**) and hydroxytyrosol (**2**) (Figure 1). Both are also abundant as esters of secoiridoids mainly of elenolic acid and analogous giving raise to close, open and decarboxymethyl forms e.g. oleocanthal (**3**) and oleacein (**4**) (Figure 1). Glycosilated derivatives of the later compounds such as ligstroside (**5**) and oleuropein (**6**) respectively are also found but manly in olive leaves.

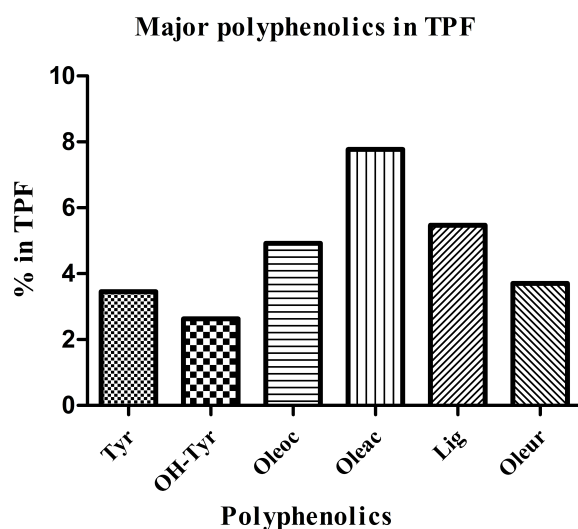
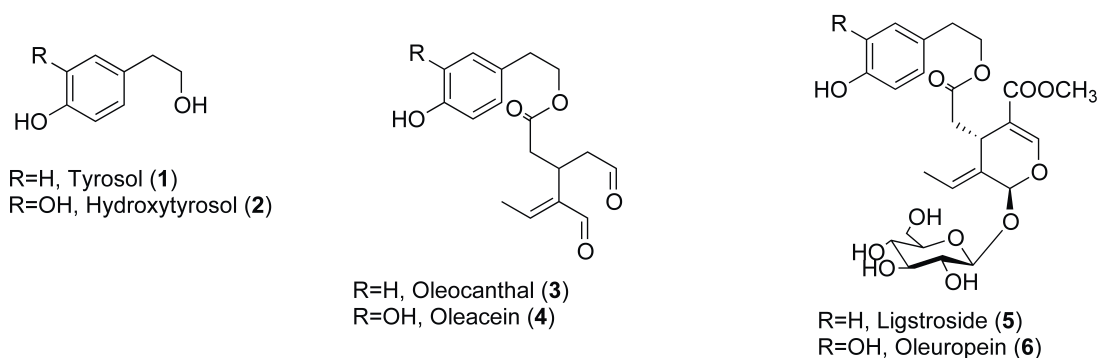


Figure 1. Chemical structure and relative quantification of tyrosol (1 Tyr), hydroxytyrosol (2 OH-Tyr), oleocanthal (3 Oleoc), oleacein (4 Oleac), ligstroside (5 Lig) and oleuropein (6 Oleur) contained in the total polyphenolic fraction.

Oleuropein, together with EVOO and olives, was proved to be effective in a rat model of senile osteoporosis with a dosage of 150 mg/kg diet^{14–17}. Moreover, hydroxytyrosol (2) and oleuropein (6) as well as olive oil were found to attenuate bone loss in two animal models for postmenopausal osteoporosis^{18,19}. Recently, Santiago-Mora et al. observed that oleuropein shifts the differentiation of mesenchymal stromal cells in favor of osteoblastogenesis²⁰. On the other hand, oleocanthal (3) responsible for the oral pungency of some oils has also been described as potent anti-inflammatory compound in vitro at concentrations of 25 and 50 μM ^{21–25}. Furthermore, it was shown to have neuroprotective effects by altering the

oligomerization of amyloid-beta-derived diffusible ligands or by inhibiting tau fibrillization²⁶⁻³⁰. Additionally, two groups described the antiproliferative effects of oleocanthal (**3**) on several tumor cell lines indicating its potential chemopreventive properties^{31, 32}. Oleacein (**4**), the hydroxylated analogue of oleocanthal, has recently drowned much attention while so far limited information is available in the literature regarding its biological role. However, could be assumed that its potency could be comparable to oleocanthal or even higher in a similar manner that hydroxytyrosol (**2**) is a stronger antioxidant agent compared to the structurally related tyrosol (**1**). Up to now, most studies highlight its potency as antioxidant³³.

Other groups of bioactive components in the EVOO are the terpenoids, with maslinic and oleanolic acids being predominant, while also phenolic acids and flavonoids have been detected³⁴. It has to be mentioned that the qualitative and quantitative content of EVOO bioactives varies and highly depends on the region of growth, the maturation time of the olive fruits at harvest and the olive oil production process¹³.

The deficiency of estrogens as a result of the cessation of ovarian function through peri- to postmenopause is associated with characteristic complaints such as hot flushes, depression, anxiety and vaginal dryness³⁵. In addition, low serum estradiol levels impact the bone homeostasis leading to osteoporosis. This systemic clinical disorder is characterized by an increased fracture risk resulting from an accelerated bone loss and deterioration of trabecular micro-architecture³⁶⁻³⁸.

In the present study, an extract of EVOO enriched in phenolic constituents, named Total Polyphenolic Fraction (TPF), has been evaluated with regard to potential bone sparing effects. Therefore, the ovariectomized rat as small animal model for postmenopausal osteoporosis was chosen³⁹⁻⁴². The removal of ovaries of skeletally

mature rats being about 12 months old induces a state of estrogen deficiency comparable to the human postmenopausal situation⁴². Afterwards post-ovariectomy, the olive extract was orally administered over a twelve-week period. Beside bone effects, potential estrogenic activities of the extract on the uterus were evaluated by determining the mRNA expression levels of marker genes.

MATERIALS AND METHODS

Chemicals and Reagents

Syringaldehyde (98% purity) used as internal standard (IS) of the analytical methodology was purchased from Sigma-Aldrich (Steinheim, Germany). Tyrosol (**1**), hydroxytyrosol (**2**), oleocanthal (**3**), oleacein (**4**), ligstroside (**5**) and oleuropein (**6**) used for the quantitation study have been previously isolated from our group⁴³⁻⁴⁵. The purity of the isolated compounds was established by high performance liquid chromatography (HPLC) coupled with a high resolution mass spectrometer (HRMS) and found to be more than 95%.

Instrumentation

The High resolution mass spectrometry (HRMS) analysis for the establishment of the purity of the standard compounds was performed on a hybrid LTQ-Orbitrap Discovery Mass Spectrometer (Thermo Scientific, Bremen, Germany). An Acquity Ultra-high Pressure Liquid Chromatography (UPLC™) system (Waters Corp., Milford, MA, USA) comprised by a Binary Solvent Manager, an autosampler (Sample Manager) capable of maintaining the sample temperature from –5 to 40 °C

and coupled to a Photo Diode Array (PDA) detector system was used for the quantification purposes.

Total Polyphenolic Fraction from Extra Virgin Olive Oil

The EVOO used in this study is from *Olea europea* var. Koroneiki originated from the Viotia area of Greece and supplied by Renieris olive oil production unit. The TPF was extracted from the EVOO following an in house previously established procedure⁴⁴. Briefly, 170 L of EVOO were subjected to 25 Kg of the adsorbent XAD – 7HP resin for polyphenols enrichment purposes. The resin activated with water and ethanol and EVOO was remained in the XAD - 7HP resin for two days with controlled and smoothed shaking and filtered. The resin was washed with 30 L cyclohexene (cHex) for the removal of the lipophilic constituents and then the polyphenol enriched extract was obtained by extraction of the resin with approx. 40L of ethanol (EtOH). Further purification, was achieved by liquid – liquid extraction using cHex & EtOH and the obtained ethanolic fraction was filtered through paper and evaporated until dry affording 150 g of TPF.

Determination of phenolic compounds by reverse phase UPLC-PDA analysis

The molecules of interest and specifically tyrosol, hydroxytyrosol, oleocanthal, oleacein, ligstroside and oleuropein, were detected and quantified at 240 nm applying the internal standard methodology. A full scan (220-500 nm) in the PDA detector was used for the identification of other EVOO constituents based on the retention time (Rt) and UV characteristics. A RP-C18, (Waters Acquity, 100 mm × 2.1 mm, 1.7 µm), preceded by a precolumn (Waters VanGuard 5 mm × 2.1 mm, 1.7 µm) was used for the analysis and the flow rate was set to 0.7 mL/min. The mobile phase consisted of

solvents A: water (H₂O) and B: acetonitrile (ACN). A gradient method (total run time of 8 min) was used for the quantification of the compounds of interest: 0 to 0.6 min: 95% A: 5% B, 0.6 to 1.3 min: 90% A: 10% B, 1.3 to 4.3 min: 75% A: 25% B, 4.3 to 6 min: 5% A: 95% B, 6 to 8 min: 95% A: 5% B. The column temperature was kept at 40 °C throughout all experiments and the injection volume was 10 µL. To carry out the quantification of the selected compounds of TPF six calibration curves were constructed one for each molecule, in a range of 0.5, 1.0, 2.0, 10.0, 25.0 and 50.0 µg/mL.

Relative Estrogen Receptor Binding Affinity

In order to assess the relative affinity of TPF constituents to both estrogen receptor subtypes, the PolarScreen™ ERα and ERβ Competitor Assay, Green (Life Technologies, Darmstadt, Germany) was used. Briefly, the respective concentration that inhibited the binding of a fluorescence-labeled estrogen (ES2, 1 nM) to recombinant human ERα or ERβ by 50% (IC₅₀) was determined for 17β-estradiol and TPF. The relative binding affinities of TPF constituents were subsequently calculated relative to that of 17β-E2.

Animals

Female pluriparous Lewis rats with an average age of 12 ± 1 months (Harlan, Borchon, Germany) were housed under controlled conditions (20 ± 1 °C, 50–80 % of relative humidity; 12:12-h light–dark cycles) in groups of five animals per cage with free access to water and a phytoestrogen-free diet (2019 Teklad Global Extruded Rodent Diet, Harlan, Borchon, Germany). Animals were subjected either to

ovariectomy (ovx, $n = 10$) or to sham surgery (sham, $n = 5$). A part of the ovariectomized animals were subsequently set on a phytoestrogen-free diet containing a polyphenolic extract in a concentration of 800 mg/kg diet (approximation to 150 mg/kg diet of oleocanthal; ovx + OL, $n = 5$). The dietary concentration of the TPF has been chosen with the aim to adjust the content of oleocanthal to the content reported in the two studies of Puel and colleagues investigating the effect of oleuropein^{14, 15}. Twelve weeks post surgery, necropsy was performed using CO₂ inhalation subsequent to a light O₂/CO₂ anesthesia including blood and organ collection. In order to verify the hormonal status, the uterine wet weights were determined and uteri were frozen in liquid nitrogen for further analysis. The left tibiae were dissected and stored in 70 % ethanol for ex vivo ICT scanning. The animal numbers and experimental setup were planned in cooperation with the Institute for Medical Informatics and Biometry of the Technische Universität Dresden and conducted to adhere closely to the 3R principle of animal welfare. All animal handling and the experimental techniques were in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki as well as licensed and carried out according to the Institutional Animal Care and Use Committee guidelines as regulated by the German federal law governing animal welfare.

RNA isolation and mRNA quantification

Total RNA was extracted from the rat uteri using peqGOLD TriFast™ according to the manufacturer's instructions (PEQLAB Biotechnologie GmbH, Erlangen, Germany). RNAs from the same group were pooled and DNA contamination was eliminated by subsequent enzymatic digestion (RQ1 DNase, Promega, Karlsruhe, Germany). Success of the DNA removal was checked by PCR.

First-strand cDNA synthesis was performed by mixing 3 µg of digested RNA with MMLV reverse transcriptase (Promega Corp., Madison, USA) and Oligo (dT)₁₂₋₁₈ primers.

Quantitative real-time PCR was applied for mRNA amplification with SybrGreen I (Sigma-Aldrich, Chemie GmbH, Steinheim, Germany) as detection dye using the iCycler iQ™ Real-Time PCR Detection System (Bio-Rad Laboratories GmbH, München, Germany).

The expression of all genes was normalized to the housekeeping gene *ribosomal protein S18* (*Rps18*) as internal control²³. Primers used for the quantitative PCR are shown in Table 1. To quantify the expression relative to the sham operated or the ovx group, the $\Delta\Delta C_T$ method was used.

Primer	Direction	Sequence	Amplicon Size [bp]
<i>Rps18</i>	Fwd	5'-CGT GAA GGA TGG GAA GTA TAG C-3'	218
	Rev	5'-TAT TAA CAG CAA AGG CCC AAA G-3'	
<i>C3</i>	Fwd	5'-ACA GCC TTC CCG GGA GCA TCA ACA-3'	275
	Rev	5'-AGC GCA CCA CAG GAG GCA CAG AGT C-3'	
<i>Cabp9k</i>	Fwd	5'-TGT CTG ACT CTG GCA GCA CTC ACT G-3'	180
	Rev	5'-CCT TCA GGA GGC TGG GGA ACT CTG-3'	
<i>Esr1</i>	Fwd	5'-GGA AGC ACA AGC GTC AGA GAG AT-3'	382
	Rev	5'-AGA CCA GAC CAA TCA TCA GGA T-3'	
<i>Esr2</i>	Fwd	5'-CTA CAG AGA GAT GGT CAA AAG TGG A-3'	215
	Rev	5'-GGG CAA GGA GAC AGA AAG TAA GT-3'	
<i>Mki67</i>	Fwd	5'-AAC CAG GAC TTT GTG CTC TGT AA-3'	208
	Rev	5'-CTC TTT TGG CTT CCA TTT CTT C-3'	
<i>Pgr</i>	Fwd	5'-CCC AGA CGA AAA GAC ACA AAA T-3'	221
	Rev	5'-CCA AAG AGA CAC CAA GAA GTG AT-3'	

Table 1. Primer pair sequences and amplicon sizes: *Rps18*-ribosomal protein S18; *C3*-complement C3; *CaBP9 k*-calcium-binding protein D-9 k; *Esr1*-estrogen receptor 1; *Esr2*-estrogen receptor 2; *Mki67*-antigen identified by monoclonal antibody Ki67; *Pgr*-progesterone receptor.

Micro-computed tomography (μ CT)

Micro-computed tomography (vivaCT 75, SCANCO Medical, Brüttisellen, Switzerland) on the excised tibiae was performed using an isotropic voxel size of 20 μ m (55 keV, 110 μ A, 220 ms, 1000 projections). Calibration of the scanner took place weekly using hydroxyapatite phantoms. Morphological analysis of a region of 2 mm thickness was performed 1 mm distal to the tibial growth plate. Contours were drawn automatically. For the analysis Image Programming Language software (IPL) from SCANCO Medical was used. The threshold of cancellous bone was set to 330 mg HA/cm³.

Bone volume ratio (BV/TV), trabecular number (Tb.N), trabecular thickness (Tb.Th), trabecular separation (Tb.Sp) and connectivity density (Conn.D) were assessed⁴⁶.

Statistical Analysis

The data are presented as arithmetic means \pm standard deviation. Statistical analysis included one-way analysis of variance (ANOVA) followed by Bonferroni post hoc test in order to determine significant differences. Statistically significant differences of the uterine gene expression profiles were tested using the Student's *t*-test. Statistical significance of the results was considered at $p \leq 0.05$.

RESULTS

Detection of phenolic compounds in the samples using the UPLC-DAD methodology

A full PDA scan methodology was established and applied for the analyses of the TPF sample for the identification and quantitation of its constituents. Identification was performed based on retention time (Rt) and UV characteristics of each compound under study. Figure 2 illustrates the UPLC-DAD chromatogram of the TPF, where the basic compounds were identified after comparison of their Rts and UV-Vis absorbance with the standard compounds.

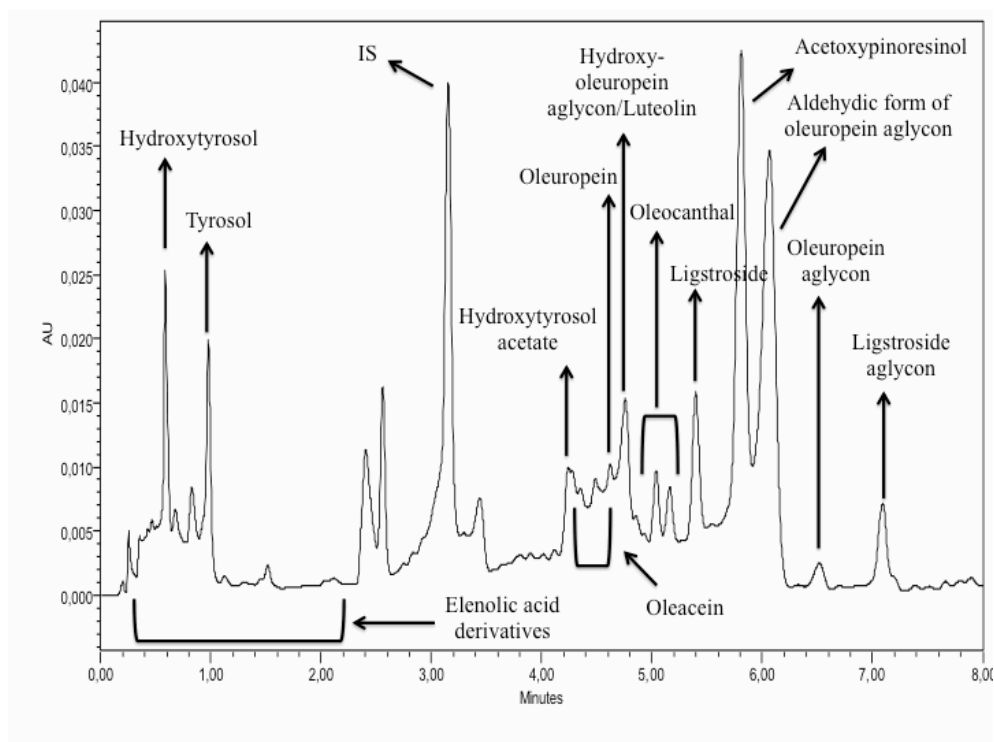


Figure 2. UPLC–PDA profiling of the total polyphenolic fraction (TPF) of EVOO. Representative compounds are annotated.

Results showed that the TPF fraction was rich in polyphenolic constituents. Among major peaks, derivatives of elenolic acid, hydroxytyrosol acetate and several oleuropein and ligstroside aglycons were detected. Tyrosol (1), hydroxytyrosol (2), oleocanthal (3), oleacein (4), ligstroside (5) and oleuropein (6) were selected for relative quantification. The quantification process was carried out applying the internal standard methodology and the respective pure standard compounds were used

for the construction of the calibration curves, consisting of six points. A second order regression analysis was employed with least squares line for the construction of the calibration curves. Quantification results are presented in Figure 1. More specifically, the TPF was found to contain 4% of tyrosol (**1**) and 3% of hydroxytyrosol (**2**). Oleuropein (**6**) content was at 3.5%, while the tyrosol analogue, ligstroside (**5**) content was found slightly higher, at 4.5%. The levels of oleocanthal (**3**), the decarboxymethyl form of ligstroside aglycon was found at similar levels (4.5%). The compound among those studied that was found to be the most abundant inside the TPF was oleacein (**4**), approaching 8%.

Relative Estrogen Receptor Binding Affinity

The estrogen receptor binding affinities were determined using a competition binding assay with purified human ER α and ER β and fluorescein-labeled E2.

The TPF extract displayed a 6.7-fold higher affinity to the estrogen receptor α with a Relative Estrogen Receptor Binding Affinity (RBA) of 0.102% compared to 0.0166% being determined as the binding affinity to ER β relative to that of 17 β -E2 (Table 2).

	ER α		ER β	
	IC ₅₀	RBA [%]	IC ₅₀	RBA [%]
17 β -estradiol	6.99 nM	100	7.85 nM	100
TPF	6.87 μ M	0.102	47.17 μ M	0.0166

Table 2. Binding affinity of oleocanthal to ER α and ER β relative to that of 17 β -estradiol.

Uterine weight

In response to the depletion of endogenous estrogens, the relative uterine wet weight (UWW) was significantly decreased in the ovx group compared to the sham animals (432.7 ± 23.7 vs. 2257.35 ± 405.9 , $p < 0.001$, Figure 3). A slightly though not significant uterotrophic effect was observed in response to the 3-month oral polyphenolic extract treatment compared to the ovx control animals (490.7 ± 53.7 vs. 432.7 ± 23 , $p = 0.058$).

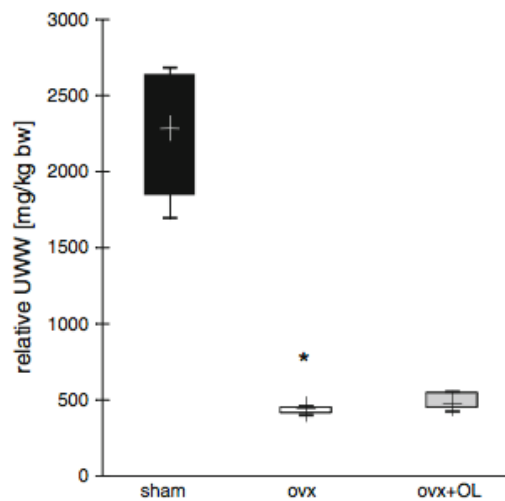


Figure 3. Relative uterine wet weights. *Box plots* show the 10 and 90 % percentile, the median (+), the maximum and minimum values. Significant differences compared to the sham or the ovx group are denoted by *asterisk* (*), *plus symbol* (+), respectively ($p < 0.05$).

Uterine gene expression analysis

The uterine gene expression pattern of both estrogen receptors (*Esr1* and *Esr2*) and the progesterone receptor (*Pgr*) are shown in Figure 4a. Compared to the ovx group, the relative mRNA levels of these three steroid receptors are found to be significantly lower in the sham animals. Furthermore, the expression of the proliferation marker gene *Mki67* among the sham group was significantly decreased compared to the ovx

group (Figure 4a). In addition, the mRNA levels of the two estrogen response marker genes *C3* and *Cabp9k* were significantly increased in the sham group compared to the ovx group (Figure 4b).

The *Esr1* and *Esr2* mRNA levels were found to be decreased in response to a 12-week oral treatment with the polyphenolic extract though statistical significance compared to the ovx group was only observed for *Esr1* (0.43 ± 0.10 , $p=0.01$; Figure 4a). The peroral treatment resulted in a slight, non-significant reduction of the *Pgr* gene expression. The *Mki67* expression rate was found to be significantly reduced in the animals treated with the polyphenolic extract (0.48 ± 0.17 , $p=0.03$). Moreover, the relative mRNA levels of *C3* and *Cabp9k* in the ovx+OL group were increased compared to the ovx control group though the elevation of the *C3* mRNA was statistically significant (2.63 ± 0.60 , $p=0.04$ and 1.93 ± 0.69 , $p=0.14$; Figure 4b).

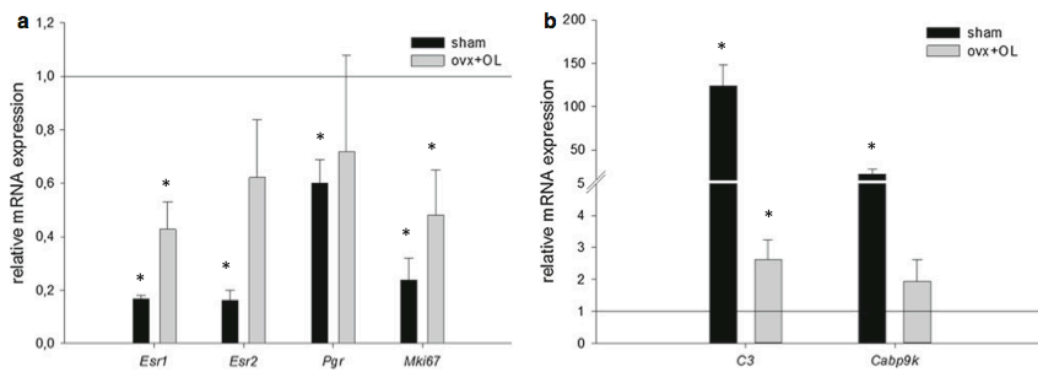


Figure 4. Uterine gene expression pattern of sham animals and ovariectomized animals that received the polyphenolic enriched diet (ovx+OL). Relative mRNA levels of *Esr1*, *Esr2*, *Pgr* and *Mki67* (a) as well as of *C3* and *Cabp9k* (b) are shown relative to the ovx group (ovx = 1). Statistical significance is compared to the ovx group is denoted by * ($p < 0.05$).

Bone morphology

The micro-architecture of the proximal tibia metaphysis has been analyzed by μ CT. The values of the measured parameters bone volume fraction (BV/TV),

connectivity density (Conn.D), trabecular number (Tb.N), trabecular thickness (Tb.Th) and trabecular separation (Tb.Sp) are shown in Table 3. In response to the 12-week depletion of the endogenous estrogens, a significant 50% reduction of BV/TV and Conn.D was observed compared to the sham group (Table 3). Furthermore, a non-significant decrease of the trabecular number and an increase of the trabecular separation were observed. Tb.Th was found to be significantly reduced among the ovx animals compared to the sham group (0.110 ± 0.007 vs. 0.122 ± 0.005 , $p = 0.015$). No changes were observed in response to the dietary treatment of the ovx animals with the polyphenolic extract (Table 3).

	sham	ovx	ovx + OL
BV/TV [%]	30.2±8.7	17.4±1.6 [<0.001*]	18.4±2.2
Conn.D [1/mm ³]	32.6±14.5	15.6±3.6 [0.034*]	18.5±4.2
Tb.N [1/mm]	3.05±0.73	2.40±0.11	2.43±0.29
Tb.Th [mm]	0.122±0.005	0.110±0.007 [0.015*]	0.111±0.005
Tb.Sp [mm]	0.308±0.105	0.386±0.010	0.392±0.057

Table 3. Micro-architectural bone parameters of the tibial metaphysis. Data are shown as mean ± SD (n=5). p-values result from one-way ANOVA representing significant differences from the sham group ([], * $p < 0.05$).

DISCUSSION

Epidemiological studies highlighted a reduced osteoporosis incidence and associated fracture risk in countries where the Mediterranean diet is prevalent⁷⁻⁹. This alimentation is characterized by high fruit, vegetable and olive oil intake. In the present study, we isolated the TPF of EVOO by removal of lipophilic and highly hydrophobic compounds. This functionalized extract (TPF) was found to be rich in tyrosol (1), hydroxytyrosol (2), oleocanthal (3), oleacein (4), ligstroside (5) and oleuropein (6). Among those constituents, oleacein, oleocanthal and ligstroside was at the highest levels in the TPF (approx. 8% of dry weight) (Figure 1). In order to address the question, whether these EVOO polyphenols possess bone sparing effects, the TPF was administered orally to ovariectomized rats over a period of 12 weeks.

As expected, the E2 deprivation in the ovx group induced an atrophy of the uterus being one-fifth of the relative uterine wet weight of the sham animals 12 weeks after removal of the ovaries (Figure 3). Though statistically not significant, the uterine wet weight in the group receiving the TPF-enriched diet was 13% higher compared to the ovx group (Figure 3). Regarding the uterine gene expression pattern for both ERs showed an E2-agonistic effect in the ovx+OL group, containing the phenolics. Both mRNA levels were found to be decreased, though statistical significance was only observed for the *Esr1* (Figure 4a). Besides, the gene expression pattern of the proliferation marker *Mki67* as well as of the estrogen response markers *C3* and *Cabp9k* were comparable in the ovx+OL group though statistical significance was only observed for the former one (Figure 4a, 4b).

A possible explanation for the observed impact on estrogen response genes might be the binding of the polyphenols to estrogen receptors because of the aromatic ring

structurally related to 17 β -Estradiol. Recently, Sirianni and colleagues showed that hydroxytyrosol but not oleuropein or tyrosol induced reporter gene expression by ER α -binding in MCF-7 cells⁴⁷. Furthermore, we observed that EVOO polyphenols bind to both estrogen receptors with a 6.7-fold higher affinity for the ER α (Table 2). This is in contrast to a number of phytoestrogens like genistein and coumestrol for which a preferential binding affinity to the ER β has been described^{48,49}.

Regarding the ovx-induced changes of the bone micro-architecture, the administration of the TPF had no impact on the quantified parameters in the proximal tibial metaphysis (Table 3). So far, only two studies have investigated the impact of olive oil or the single polyphenols on the bone in ovariectomized rats or mice. Saleh and Saleh observed an attenuation of ovx-induced bone loss in Wistar rats by olive oil¹⁸. In contrast to our study, the described phenolic content of the olive oil was twice as high as in our investigation. Furthermore, it cannot be excluded that the protective effects could be at least partially ascribable to the contained fatty acids^{50,51}, which were not part of our extract. In another study, it had been shown that orally administered hydroxytyrosol and oleuropein but not tyrosol protected from ovx-induced bone loss in mice¹⁹. The mice in this study have been exposed to 10 mg/kg bodyweight for each single compound. In contrast to this, the exposure of the rats in the present study has been calculated as 10 mg/ kg bodyweight per day for the TPF containing six different compounds taken together (Figure 2). So, the concentration of hydroxytyrosol and oleuropein might have been too low to prevent bone loss in the rat model. Further studies have described bone sparing potential of either single olive polyphenols or olives in inflammation-induced models¹⁴⁻¹⁷.

In summary, the present study showed E2-agonistic effects of a polyphenolic extract of EVOO on uterine gene expression. Furthermore, we observed estrogen

receptor binding affinity of EVOO polyphenolic extract. Regarding the bone, the polyphenolic enriched extract showed no attenuation of bone loss in a model of postmenopausal osteoporosis.

REFERENCES

1. Buckland G, Travier N, Cottet V et al (2012) Adherence to the Mediterranean diet and risk of breast cancer in the European prospective investigation into cancer and nutrition cohort study. *Int J Cancer*. doi:10.1002/ijc.27958
2. Yusof AS, Isa ZM, Shah SA (2012) Dietary patterns and risk of colorectal cancer: a systematic review of cohort studies (2000–2011). *Asian Pac J Cancer Prev* 13:4713–4717
3. Dilis V, Katsoulis M, Lagiou P et al (2012) Mediterranean diet and CHD: the Greek European prospective investigation into cancer and nutrition cohort. *Br J Nutr* 108:699–709. doi:10.1017/S0007114512001821
4. Buckland G, Travier N, Barricarte A et al (2012) Olive oil intake and CHD in the European prospective investigation into cancer and nutrition Spanish cohort. *Br J Nutr* 108:2075–2082. doi:10.1017/S000711451200298X
5. Ahluwalia N, Andreeva VA, Kesse-Guyot E, Hercberg S (2012) Dietary patterns, inflammation and the metabolic syndrome. *Diabetes Metab*. doi:10.1016/j.diabet.2012.08.007
6. Viscogliosi G, Cipriani E, Liguori ML et al (2013) Mediterranean dietary pattern adherence: associations with prediabetes, metabolic syndrome, and related micro inflammation. *Metab Syndr Relat Disord*. doi:10.1089/met.2012.0168
7. Johnell O, Gullberg B, Allander E, Kanis DJA (1992) The apparent incidence of hip fracture in Europe: a study of national register sources. *Osteoporos Int* 2:298–302. doi:10.1007/BF01623186
8. Kanis JA (1993) The incidence of hip fracture in Europe. *Osteoporos Int* 3:10–15. doi:10.1007/BF01621853
9. Benetou V, Orfanos P, Pettersson-Kymmer U et al (2012) Mediterranean diet and incidence of hip fractures in a European cohort. *Osteoporos Int* 24:1587–1598. doi:10.1007/s00198-012-2187-3
10. Perez-Martinez P, Garcia-Rios A, Delgado-Lista J et al (2011) Mediterranean diet rich in olive oil and obesity, metabolic syndrome and diabetes mellitus. *Curr Pharm Des* 17:769–777. doi:10.2174/138161211795428948
11. Lopez S, Bermudez B, Varela LM et al (2012) Olives and olive oil: diet and health impacts. *CAB Rev*. doi:10.1079/PAVSNNR20127034
12. Covas M-I (2007) Olive oil and the cardiovascular system. *Pharmacol Res* 55:175–186. doi:10.1016/j.phrs.2007.01.01013. Cicerale S, Conlan XA, Sinclair AJ, Keast RSJ (2008) Chemistry and health of olive oil phenolics. *Crit Rev Food Sci Nutr* 49:218–236. doi:10.1080/1040839070185622314. Puel C, Quintin A, Agalias A et al (2004) Olive oil and its main phenolic micronutrient (oleuropein) prevent inflammation-induced bone loss in the ovariectomised rat. *Br J Nutr* 92:119–127. doi:10.1079/BJN20041181

15. Puel C, Mathey J, Agalias A et al (2006) Dose–response study of effect of oleuropein, an olive oil polyphenol, in an ovariectomy/ inflammation experimental model of bone loss in the rat. *Clin Nutr* 25:859–868. doi:10.1016/j.clnu.2006.03.009
16. Puel C, Mardon J, Kati-coulibaly S et al (2007) Black Lucques olives prevented bone loss caused by ovariectomy and talc granulomatosis in rats. *Br J Nutr* 97:1012–1020. doi:10.1017/S0007114507659030
17. Puel C, Mardon J, Agalias A et al (2008) Major phenolic compounds in olive oil modulate bone loss in an ovariectomy/ inflammation experimental model. *J Agric Food Chem* 56:9417–9422. doi:10.1021/jf801794q
18. Saleh NK, Saleh HA (2011) Olive oil effectively mitigates ovariectomy-induced osteoporosis in rats. *BMC Complem Altern Med* 11:10. doi:10.1186/1472-6882-11-10
19. Hagiwara K, Goto T, Araki M et al (2011) Olive polyphenol hydroxytyrosol prevents bone loss. *Eur J Pharmacol* 662:78–84. doi:10.1016/j.ejphar.2011.04.023
20. Santiago-Mora R, Casado-Díaz A, Castro MDD, Quesada-Go´mez JM (2011) Oleuropein enhances osteoblastogenesis and inhibits adipogenesis: the effect on differentiation in stem cells derived from bone marrow. *Osteoporos Int* 22:675–684. doi:10.1007/s00198-010-1270-x
21. Beauchamp GK, Keast RSJ, Morel D et al (2005) Phytochemistry: ibuprofen-like activity in extra-virgin olive oil. *Nature* 437:45–46. doi:10.1038/437045a
22. Smith AB, Han Q, Breslin PAS, Beauchamp GK (2005) Synthesis and assignment of absolute configuration of (-)-oleocanthal: a potent, naturally occurring non-steroidal anti-inflammatory and anti-oxidant agent derived from extra virgin olive oils. *Org Lett* 7:5075–5078. doi:10.1021/ol052106a
23. Iacono A, Go´mez R, Sperry J et al (2010) Effect of oleocanthal and its derivatives on inflammatory response induced by lipopolysaccharide in a murine chondrocyte cell line. *Arthritis Rheum* 62:1675–1682. doi:10.1002/art.27437
24. Scotece M, Go´mez R, Conde J et al (2012) Further evidence for the anti-inflammatory activity of oleocanthal: inhibition of MIP-1a and IL-6 in J774 macrophages and in ATDC5 chondrocytes. *Life Sci* 91:1229–1235. doi:10.1016/j.lfs.2012.09.012
25. Cicerale S, Lucas L, Keast R (2012) Antimicrobial, antioxidant and anti-inflammatory phenolic activities in extra virgin olive oil. *Curr Opin Biotechnol* 23:129–135. doi:10.1016/j.copbio.2011.09.006
26. Pitt J, Roth W, Lacor P et al (2009) Alzheimer’s-associated Ab oligomers show altered structure, immunoreactivity and synaptotoxicity with low doses of oleocanthal. *Toxicol Appl Pharmacol* 240:189–197. doi:10.1016/j.taap.2009.07.018
27. Li W, Sperry JB, Crowe A et al (2009) Inhibition of tau fibrilization by oleocanthal via reaction with the amino groups of tau. *J Neurochem* 110:1339–1351. doi:10.1111/j.1471-4159.2009.06224.x
28. Monti MC, Margarucci L, Tosco A et al (2011) New insights on the interaction mechanism between tau protein and oleocanthal, an extra-virgin olive-oil bioactive component. *Food Funct* 2:423–428. doi:10.1039/C1FO10064E

29. Monti MC, Margarucci L, Riccio R, Casapullo A (2012) Modulation of Tau protein fibrillization by oleocanthal. *J Nat Prod* 75:1584–1588. doi:10.1021/np300384h
30. Abuznait AH, Qosa H, Busnena BA et al (2013) Olive-oil-derived oleocanthal enhances β -amyloid clearance as a potential neuro-protective mechanism against Alzheimer's disease. In vitro and in vivo studies. *ACS Chem Neurosci*. doi:10.1021/cn400024q
31. Khanal P, Oh W-K, Yun HJ et al (2011) p-HPEA-EDA, a phenolic compound of virgin olive oil, activates AMP-activated protein kinase to inhibit carcinogenesis. *Carcinogenesis* 32:545–553. doi:10.1093/carcin/bgr001
32. Elnagar A, Sylvester P, El Sayed K (2011) (-)-Oleocanthal as a c-Met inhibitor for the control of metastatic breast and prostate cancers. *Planta Med* 77:1013–1019. doi:10.1055/s-0030-1270724
33. Parzonko A, Naruszewicz M (2012) Oleacein protects endothelial progenitor cells against angiotensin II—induced oxidative stress. *Planta Med*. doi:10.1055/s-0032-1320446
34. Kanakis P, Termentzi A, Michel T et al (2012) From drupes to olive oil: how do bioactives vary during a single production procedure? *Planta Med*. doi:10.1055/s-0032-1321201
35. Nelson HD Menopause. *The Lancet* 371:760–770. doi: 10.1016/ S0140-6736(08)60346-3
36. (1994) Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO study group. *World Health Organ Tech Rep Ser* 843:1–129
37. (2000) Osteoporosis prevention, diagnosis, and therapy. *NIH Consens Statement* 17:1–45
38. Kanis JA, McCloskey EV, Johansson H et al (2008) A reference standard for the description of osteoporosis. *Bone* 42:467–475. doi:10.1016/j.bone.2007.11.001
39. Kalu DN (1991) The ovariectomized rat model of postmenopausal bone loss. *Bone Miner* 15:175–191
40. Turner RT, Maran A, Lotinun S et al (2001) Animal models for osteoporosis. *Rev Endocr Metab Disord* 2:117–127
41. Jee WS, Yao W (2001) Overview: animal models of osteopenia and osteoporosis. *J Musculoskelet Neuronal Interact* 1:193–207
42. Lelovas PP, Xanthos TT, Thoma SE et al (2008) The laboratory rat as an animal model for osteoporosis research. *Comp Med* 58:424–430
43. Andreadou I, Iliodromitis EK, Mikros E et al (2006) The olive constituent oleuropein exhibits anti-ischemic, antioxidative, and hypolipidemic effects in anesthetized rabbits. *J Nutr* 136:2213–2219
44. Agalias A, Magiatis P, Skaltsounis A-L et al (2007) A new process for the management of olive oil mill waste water and recovery of natural antioxidants. *J Agric Food Chem* 55:2671–2676. doi:10.1021/jf063091d
45. Karkoula E, Skantzari A, Melliou E, Magiatis P (2012) Direct measurement of oleocanthal and oleacein levels in olive oil by quantitative ^1H NMR. establishment of a new index for the characterization of extra virgin olive oils. *J Agric Food Chem* 60:11696–11703. doi:10.1021/jf3032765

46. Bouxsein ML, Boyd SK, Christiansen BA et al (2010) Guidelines for assessment of bone microstructure in rodents using micro- computed tomography. *J Bone Miner Res* 25:1468–1486. doi:10. 1002/jbmr.141
47. Sirianni R, Chimento A, De Luca A et al (2010) Oleuropein and hydroxytyrosol inhibit MCF-7 breast cancer cell proliferation interfering with ERK1/2 activation. *Mol Nutr Food Res* 54:833–840. doi:10.1002/mnfr.200900111
48. Kuiper GGJM, Carlsson B, Grandien K et al (1997) Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors α and β . *Endocrinology* 138:863–870. doi:10.1210/en.138.3.863
49. Kuiper GGJM, Lemmen JG, Carlsson B et al (1998) Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor β . *Endocrinology* 139:4252–4263. doi:10.1210/en.139. 10.4252
50. Trichopoulou A, Georgiou E, Bassiakos Y et al (1997) Energy intake and monounsaturated fat in relation to bone mineral density among women and men in Greece. *Prev Med* 26:395–400. doi:10.1006/pmed1997.0160
51. Wauquier F, Philippe C, Leotoing L et al (2013) The free fatty acid receptor G protein-coupled receptor 40 (GPR40) Protects from bone loss through inhibition of osteoclast differentiation. *J Biol Chem* 288:6542–6551. doi:10.1074/jbc.M112.429084

A

Acknowledgments

Nikolaos Lemonakis

This study was carried out at the Department of Pharmacognosy and Natural products chemistry, School of Pharmacy, National and Kapodistrian University of Athens. I am very pleased to have the opportunity to work in this high-class research unit.

I had the opportunity to work in a challenging and modern field, this of analysis, combining the last one with the natural products chemistry. I had the chance to meet the field of “olive oil”, a very popular and important field, the field of isolation and characterization of natural products and extracts, targeted quantitation using state of the art analytical instruments, such as UHPLC and HRMS (orbitrap and ToF), the design of experiments field, the multivariate analysis (metabolomics) etc. I collaborated with the elite of the scientific community in various projects, and I learned about the setup of integrated studies.

I sincerely thank the members of the examining scientific committee: Prof. A.L. Skaltsounis, Assistant Prof. E. Gikas, Researcher B' C. Tamvakopoulos, Prof. E. Mikros, Associate Prof. A. Tsarbopoulos, Associate Prof N. Thomaidis and Prof. J.L. Wolfender. I want to express my gratefulness, for agree to participate in the procedure.

I would like to thank the three-members scientific committee of my PhD Thesis, Professor Alexios-Leandros Skaltsounis, Associate Professor Evangelos Gikas and Researcher B' Constantinos Tamvakopoulos. I have been privileged to work under your scientific expertise. I am very grateful for your excellent guidance and interest to my work, but also freedom to work in my own way.

I want to express my warm thanks to Professor Leandros, the director of the laboratory and of the School of Pharmacy, for being my principle supervisor and

providing so much time and kind support throughout my postgraduate studies. I greatly appreciate that I got the change to complete this PhD period. Without his unconditional support, patient and understanding, financial support, it would have been a mission impossible. He gave me the opportunity to participate in great projects, to travel around the world for conferences, to participate in a European Marie Curie Project and travel for that purpose to Brazil.

I greatly owe much gratitude to my mentor, Vangelis. I sincerely thank him for all his scientific support, the insights and inspiring discussions about our experiments, our data, as well as for the many valuable ideas, instructions and strategies for handling complex data analysis problems. He taught me metabolomics, many related software, design of experiments, validation of analytical and bioanalytical methodologies etc. Without these essential contributions it would not have been possible for me to incorporate the natural products ideas into the appropriate biological/metabolomics context. His commitment and enthusiasm were contagious and characteristic, and inspired me to explore new insights and analysis approaches throughout the PhD period. Hence, my doctoral research would not have been possible without the vision, energy and effort of Vangelis. Last but not least, my biggest thank for his friendship. I have learned so much from you.

I want to thank Maria Halabalaki, for being part of my supervising team. I am very grateful for your support, encouragement and scientific contribution. I thank you for the enthusiastic, inspiring, indefatigable and instructive supervision, understanding and patient. There were many moments when I was lost in experiments and data and I did not know how to go on and with what, and she had the experience to drive me in the right way. Besides the close scientific

interactions, I also enjoyed many other social moments, especially our travel to Brazil. I really look forward for the next one.

I thank Professor Emmanouil Mikros for improving my overall understanding of the metabolomics in the crossover studies. In many scientific discussions we often succeeded to find rational metabolomics solutions for analyzing and handling my data, concerning this crossover study.

I thank Associate Professor Adonis Tsarbopoulos, which he kindly provided to me the ToF mass spectrometer in the bioanalytical laboratory of the Goulandris Museum. I worked there for 2 months and I really learned a lot.

I would like also to thank all the team's members that participated in the projects: Professor Vasilios Mougios, Professor Lindsay Brown, Dr. Hemant Poudyal, Associate Professor Annekathrin Martina Keilera and Professor Günter Vollmera. It is really with great pleasure that I look back on the past three years. Without you, we would not be able to perform these studies. Thank you all for the generated knowledge. I am very proud that our efforts led to joint publications.

I also would like to thank all my laboratory colleagues and friends for the pleasant atmosphere and support during my PhD period and especially: Eirini K, Katerina T, Eirini B, Eliza C, Tina K, Maria L, Job T, Panagiotis S and George P. Without your kindness help, understanding, support and friendship it would have been impossible for me to complete this PhD period. Furthermore, I would like to thank all the people from the lab, who contributed in one way or another to this Thesis, but have not been mentioned.

Eirini K, I will never forget your special involvement and interest in my work, but also for your mental support and friendship. I am also grateful for being my fellow traveller in Brazil.

Katerina T, thank you for your collaboration and for your insightful advises in the olive research group's projects.

I want to express my warm thanks to Katerina Makropoulou and Fwteini Kapsali, the secretaries of our laboratory, for their kindly support in organization and financial issues.

All my dear friends deserve my warmest gratitude. Christos, thank you for your support, the whole-hearted and ever lasting friendship. Our friendship is something special! Violeta, thank you for everything! Your support and your understanding made the things easier. Thank you!

I also express my loving thanks to my aunt and uncle, Toula and Vangelis, for all support and help. My deepest gratitude goes to my parents, Fani and Giwrgos, for teaching me what truly matters in life and providing me care and support unconditionally. I warmly thank my sister, Chrysa, and the almost new member of our family and my tremendous love, Maritina, for support and for bringing so much joy in my life and of course for their selfless love. You matter to me the most!

Ευχαριστώ!

Athens, May 2014

Níkolaos Lemonakís

P

Publications

Nikolaos Lemonakis

Lemonakis N, Magiatis P, Kostomitsopoulos N, Skaltsounis AL, Tamvakopoulos C. Oral administration of Chios mastic gum or extracts in mice: Quantification of triterpenic acids by liquid chromatography - tandem mass spectrometry. *Planta Med* 2011; 77: 1916–1923.

Bosi C, Rosa D, Grougnet R, Lemonakis N, Halabalaki M, Skaltsounis AL, Biavatti M. Pyrrolizidine alkaloids in medicinal tea of *Ageratum conyzoides*. *Revista Brasileira de Farmacognosia*. 2013 23 (3): 425–432.

Lemonakis N, Gikas E, Halabalaki M, Skaltsounis AL. Development and validation of a combined methodology for assessing the total quality control of herbal medicinal products – Application to oleuropein preparations. *PLoS One*. 2013 Oct 21;8(10):e78277.

Keiler AK, Zierau O, Bernhardt R, Scharnweber D, Lemonakis N, Termetzi E, Skaltsounis AL, Vollmer G and Halabalaki M. Impact of a functionalized olive oil extract on the uterus and the bone in a model of postmenopausal osteoporosis. *Eur J Nutr*. 2013 Oct 30

Optimization of parameters affecting signal intensity to an LTQ-Orbitrap:

Lemonakis N, Skaltsounis AL, Gikas E. A DOE approach using oleuropein and hydroxytyrosol as model compounds. *Submitted*

Lemonakis N, Gikas E, Mougios V, Halabalaki M, Skaltsounis AL. Validation of a bioanalytical methodology based on UHPLC-HRMS/MS for the quantification of oleuropein in human serum. Application to a pharmacokinetic study. *Submitted*

Poudyal H, Lemonakis N, Gikas E, Halabalaki M, Skaltsounis AL, Brown L.
Effects of hydroxytyrosol on the risk factors of metabolic syndrome in diet-induced
obese rats. *Submitted*

CV

Curriculum Vitae

Nikolaos Lemonakis



Personal information

First name / Surname **Nikolaos LEMONAKIS**
Address 28 VIRONOS St., Athens, 15231, Greece
Telephones Personal: (+30-210) 6744567 Mobile: (+30) 6972847635
E-mail nicklem26@gmail.com
Nationalities Greek
Date of birth 26.12.1983
Gender Male

Work experience

Dates Mar 2012 – May 2012
Occupation or position held Researcher at Federal University of Santa Catarina through an FP7 European Research Grand
Main activities and responsibilities Research on analysis of natural products, research on bioanalysis, seminars on Analysis of natural products
Name and address of employer Pr.Maique W. Biavatti, Federal University of Santa Catarina
Type of business or sector Laboratory in University, Faculty of Pharmacy

Dates Jan 2010 – May 2014
Occupation or position held PhD candidate student
Main activities and responsibilities Research on analysis of natural products, research on bioanalysis, seminars on Ethnopharmacologie
Name and address of employer Pr.A.Skaltsounis, Departement of Pharmacognosy, Faculty of Pharmacy, National and Kapodistrian University of Athens
Type of business or sector Laboratory in University, Faculty of Pharmacy

Dates Sep 2008 – Dec 2009
Occupation or position held Post-graduate Research master's degree (MSc)
Main activities and responsibilities Research in Pharmacology and analysis
Name and address of employe Dr. Constantin Tamvakopoulos, Biomedical Research Foundation Academy of Athens, BRFAA, Division of Pharmacology & Pharmacotechnology - Basic Research I, Athens
Type of business or sector Laboratory in Research Institute, Division of Pharmacology & Pharmacotechnology

Dates Sep 2007 – Sep 2008
Occupation or position held Post-graduate Research master's degree (MSc)
Main activities and responsibilities Research on isolation of natural products (extracts and pure compounds), chromatography
Name and address of employer Pr.A.Skaltsounis, Departement of Pharmacognosy, Faculty of Pharmacy, National and Kapodistrian University of Athens
Type of business or sector Laboratory in University, Faculty of Pharmacy

Dates Jun 2004 – Aug 2004
Occupation or position held Assistant of the director of the chemistry laboratory
Main activities and responsibilities Analysis od drugs, waste waters, drinks and food
Name and address of employer General Chemical State Laboratory of Greece, Alexandroupolis Department
Type of business or sector Directorate General - Greek Ministry of Finance

Education and training

Dates Jan 2010 – present

Title of qualification awarded PhD candidate

Principal subjects/occupational skills covered Thesis title: “Quantification of natural bioactive compounds in biological fluids”

Name and type of organisation providing education and training Departement of Pharmacognosy, Faculty of Pharmacy, National and Kapodistrian University of Athens
Pr.A.Skaltsounis, Dr. C. Tamvakopoulos, Lecturer E. Gikas

Dates September 2007 – Dec 2009

Title of qualification awarded Post-graduate Research master’s degree (MSc)

Principal subjects/occupational skills covered Master title: “Bioactive compounds from Chios mastic gum - Isolation and Pharmacokinetic studies using Liquid Chromatography - Tandem Mass Spectrometry (LC-MS/MS)”

Name and type of organisation providing education and training Departement of Pharmacognosy, Faculty of Pharmacy, National and Kapodistrian University of Athens,
Biomedical Research Foundation Academy of Athens, BRFAA, Division of Pharmacology & Pharmacotechnology - Basic Research I, Athens

Dates 2001 - 2006

Title of qualification awarded Bachelor of Chemistry

Name and type of organisation providing education and training University of Ioannina, Department of Chemistry - Ioannina, Greece. Oct 2001 – Oct 2006

Publications

Lemonakis N, Magiatis P, Kostomitsopoulos N, Skaltsounis AL, Tamvakopoulos C. Oral administration of Chios mastic gum or extracts in mice: Quantification of triterpenic acids by liquid chromatography - tandem mass spectrometry. *Planta Med* 2011; 77: 1916–1923.

Bosi C, Rosa D, Grougnet R, Lemonakis N, Halabalaki M, Skaltsounis AL, Biavatti M. Pyrrolizidine alkaloids in medicinal tea of *Ageratum conyzoides*. *Revista Brasileira de Farmacognosia*. 2013 23 (3): 425–432.

Lemonakis N, Gikas E, Halabalaki M, Skaltsounis AL. Development and validation of a combined methodology for assessing the total quality control of herbal medicinal products – Application to oleuropein preparations. *PLoS One*. 2013 Oct 21;8(10):e78277.

Keiler AK, Zierau O, Bernhardt R, Scharnweber D, Lemonakis N, Termetzi E, Skaltsounis AL, Vollmer G and Halabalaki M. Impact of a functionalized olive oil extract on the uterus and the bone in a model of postmenopausal osteoporosis. *Eur J Nutr*. 2013 Oct 30

Optimization of parameters affecting signal intensity to an LTQ-Orbitrap:
Lemonakis N, Skaltsounis AL, Gikas E. A DOE approach using oleuropein and hydroxytyrosol as model compounds. *Submitted*

Lemonakis N, Gikas E, Mougios V, Halabalaki M, Skaltsounis AL. Validation of a bioanalytical methodology based on UHPLC-HRMS/MS for the quantification of oleuropein in human serum. Application to a pharmacokinetic study. *Submitted*

Poudyal H, Lemonakis N, Gikas E, Halabalaki M, Skaltsounis AL, Brown L. Effects of hydroxytyrosol on the risk factors of metabolic syndrome in diet-induced obese rats. *Submitted*

Presentation in Federal University of Santa Catarina – FP7 Seminar (ChemBioFight) - “The “black box” of natural products analysis. What are the new analytical trends? Three different case studies”, Florianopolis, Brazil, Apr 2012

Poster in 2011 AAPS Annual Meeting and Exposition with title: “Absorption Study of Chios Mastic Gum Triterpenoids after Oral Administration in Mice Using a Novel LC-MS/MS Approach”, Washington, USA, Oct 2011

Presentation in 59th International Congress and Annual Meeting of the Society for Medicinal Plant and Natural Product Research with title: “A rapid LC-MS Method for the Simultaneous Quantification of Oleuropein and its main metabolite, Hydroxytyrosol, in clinical samples after oral administration of commercial herb medicinal products”, Antalya, Turkey – **Presentation Award in Young Researchers, Sep 2011**

Poster in 1st International Proteomics Conference on Crete –IPCC01h with title: “Applications of Mass Spectrometry in Natural Product Chemistry”, Heraklion Crete, Greece, Oct 2010

Personal skills and competences

Sailing, Music, Sports

Mother tongue(s) **Greek**

Other language(s)

Self-assessment

Understanding

Speaking

Writing

European level (*)

Listening

Reading

Spoken interaction

Spoken production

English

C1 Proficient user

B2 Independent user

C1 Proficient user

C2 Proficient user

C1 Proficient user

Spanish

C1 Independent user

B2 Independent user

C2 Independent user

C2 Independent user

C1 Independent user

(*) Common European Framework of Reference for Languages

Technical skills and competences	<p>Familiar with a variety of techniques such as</p> <p>Separation and isolation of natural products using chromatographic techniques (MPLC, FCPC, HPLC)</p> <p>Structure elucidation of natural products via spectroscopic methods (1D&2D NMR, MS)</p> <p>Research of natural products using analytical techniques (LC-MS/MS, LC-UV, LC-ELCD)</p> <p>Development of novel LC-MS/MS methods for quantification of bioactive compounds in biological fluids</p> <p>Metabolomics</p> <p>Multilevel metabolomics</p> <p>Design of experiments</p> <p>Drug Metabolism, PK/PD</p> <p>Experience in <i>in vivo</i> and <i>in vitro</i> protocols</p> <p>Handling of animal models</p>
Computer skills and competences	Microsoft Windows – Microsoft Office 2007 (Word, Excel, PowerPoint), ECDL (Word, Excel, Internet)
Driving licence	Category A

Nature shows us only the tail of the lion. But I do not doubt that the lion belongs to it even though he cannot at once reveal himself because of his enormous size

Albert Einstein