

**Development of State-of-the-art mass spectrometry-  
based analytical methodologies for the study of  
bioactive natural products fate after dietary  
administration. Bioavailability,  
metabolism/metabolomics studies**

**Baira Eirini**

**Department of Pharmaceutical Chemistry**

**School of Pharmacy**



**HELLENIC REPUBLIC  
National and Kapodistrian  
University of Athens**

**March 2016**

**This dissertation is submitted to the National and  
Kapodistrian University of Athens**

**For the degree of Doctor of Philosophy**

**Baira Eirini**

**Development of State-of-the-art mass  
spectrometry-based analytical methodologies  
for the study of bioactive natural products fate  
after dietary administration. Bioavailability,  
metabolism/metabolomics studies**

Department of Pharmaceutical Chemistry

School of Pharmacy

National and Kapodistrian University of Athens

Greece

2016

**Development of State-of-the-art mass spectrometry-  
based analytical methodologies for the study of  
bioactive natural products fate after dietary  
administration. Bioavailability,  
metabolism/metabolomics studies**

**Baira Eirini**

**Examining Committee of the PhD Thesis**

- Assistant Prof. E. Gikas<sup>1</sup>
- Prof. A.L. Skaltsounis<sup>2</sup>
- Prof. S. Mitakou<sup>2</sup>
- Prof. E. Mikros<sup>1</sup>
- Prof. G. Theodoridis<sup>3</sup>
- Principal Researcher M.I Klapa<sup>4</sup>
- Prof. G. K. Bonn<sup>5</sup>

<sup>1</sup>Department of Pharmaceutical Chemistry, Faculty of Pharmacy, National and Kapodistrian University of Athens, Greece

<sup>2</sup>Department of Pharmacognosy and Natural Products Chemistry, Faculty of Pharmacy, National and Kapodistrian University of Athens, Greece

<sup>3</sup>Department of Chemistry, Aristotle University of Thessaloniki, Greece

<sup>4</sup>Metabolic Engineering and Systems Biology Laboratory, Institute of Chemical Engineering Sciences (ICE-HT), Foundation for Research and Technology, Hellas (FORTH), Patras, Greece

<sup>5</sup>Institute of Analytical Chemistry and Radiochemistry, Leopold-Franzens University of Innsbruck, Innsbruck, Austria

## Table of Contents

<b>Abstract</b>		vi
<b>Περίληψη</b>		viii
<b>General Introduction- Aim of the study</b>		10
<b>Chapter 1</b>	Validated UHPLC-ESI(-)-HRMS methodology for the simultaneous quantitative determination of hesperidin, hesperetin, naringin and naringenin in chicken plasma	13
<b>Chapter 2</b>	Metabolism of naringin, hesperidin and their aglycones in plasma samples after dietary supplementation of naringin and hesperidin in chickens	40
<b>Chapter 3</b>	Development and validation of an UHPLC-ESI(-)-HRMS methodology for the simultaneous quantification of hesperidin, naringin and their aglycones in chicken tissue samples	59

<b>Chapter 4</b>	An alternative approach for data processing in untargeted metabolomics by UHPLC-ESI(-)-HRMS. The use of post acquisition spectral stitching	77
<b>Chapter 5</b>	UHPLC-HRMS-based plasma metabolomics study of naringin and hesperidin after dietary supplementation in chickens	113
<b>Chapter 6</b>	UHPLC-HRMS-based tissue metabolomics study of naringin and hesperidin after dietary supplementation in chickens	170
<b>General Discussion</b>		238
<b>Acknowledgment</b>		242
<b>CV</b>		245

## **Abstract**

Flavonoids are a class of polyphenolic molecules present in food, such as fruits, vegetables and plant derived juices. They represent one of the most important classes of biologically active compounds.

In the present study the fate as well as the putative biochemical effects of naringin and hesperidin was examined after their dietary supplementation in Ross 308 broiler chickens, in order to increase the antioxidant capacity of their tissues.

Two fully validated methodologies were developed for the simultaneous quantitation of hesperidin, naringin, hesperetin and naringenin in plasma and tissue samples. According to the results, in plasma samples only naringin concentration levels were above the LLOQ after 30 days of administration whereas in tissue samples no detectable amounts have been discovered.

Due to the low concentration levels of the analytes in plasma a High Resolution Mass Spectrometry based metabolism study was performed in order to examine the possibility of rapid and extensive metabolism. Two metabolite modifications were detected for naringin, for naringenin and for hesperidin, respectively. Nevertheless, no metabolite modification was detected for hesperetin.

Into the next step, the effect of the substances in the whole biochemical profile of the organism has been investigated. Due to the complexity of such analyses, novel data processing methodologies have to be developed and validated. In this context, a new strategy in the data processing pipeline of MS based metabolomics was developed with the aim to increase the number of the reliable metabolite identifications. This methodology was applied to the untargeted metabolomics study in chicken plasma samples that were administered either with hesperidin or naringin for 30 days. Four variables were detected to discriminate control group from the group after naringin administration whereas only one substance could effectively discriminate the hesperidin treated from the control group.

Additionally in order to further characterize the metabolic changes that occur in tissues as a response to naringin and hesperidin dietary supplementation in chickens, two metabolomics studies (for naringin and hesperidin) has been held. Six variables were detected as discriminating naringin the dietary supplementation group and twelve variables corresponding hesperidin supplementation group from their respective control group.

## Περίληψη

Τα φλαβονοειδή είναι μια κατηγορία πολυφαινολικών μορίων που περιέχονται σε τρόφιμα, όπως φρούτα, λαχανικά και παράγωγα φυτικών χυμών. Αντιπροσωπεύουν μία από τις πιο σημαντικές κατηγορίες βιολογικών ενώσεων με ενδιαφέρουσα δράση.

Στην παρούσα εργασία, εξετάστηκε η τύχη και οι βιοχημικές επιδράσεις της ναργινίνης και εσπεριδίνης μετά τη χορήγηση τους μέσω της τροφής σε Ross 308 κοτόπουλα, μελετώντας την αύξηση της αντιοξειδωτικής ικανότητας τους στους ιστούς.

Δύο πλήρως επικυρωμένες μεθοδολογίες αναπτύχθηκαν για την ταυτόχρονη ποσοτικοποίηση της εσπεριδίνης, ναργινίνης, εσπερετίνης και ναριγενίνης σε δείγματα πλάσματος και ιστών. Σύμφωνα με τα αποτελέσματα, στα δείγματα πλάσματος μόνο τα επίπεδα συγκέντρωσης της ναργινίνης ήταν πάνω από το κατώτατο όριο ποσοτικοποίησης (LLOQ) μετά από 30 ημέρες χορήγησης, ενώ στα δείγματα ιστών δεν ανιχνεύτηκαν οι ουσίες.

Λόγω των χαμηλών επιπέδων συγκέντρωσης των ουσιών στο πλάσμα, διεξήχθη μελέτη μεταβολισμού βασισμένη σε υψηλής ανάλυσης φασματομετρία μάζας, με σκοπό να μελετηθεί η πιθανότητα γρήγορου κι εκτεναμένου μεταβολισμού των ουσιών. Δύο μεταβολικές τροποποιήσεις εντοπίστηκαν για τη ναργινίνη, ναριγενίνη και εσπεριδίνη, αντίστοιχα. Καθώς και καμία μεταβολική τροποποίηση δεν ανιχνεύθηκε για την εσπερετίνη.

Στο επόμενο στάδιο, διερευνήθηκε η επίδραση των ουσιών στο ολικό βιοχημικό προφίλ του οργανισμού. Λόγω της πολυπλοκότητας των αναλύσεων αυτών, κρίνεται αναγκαία η ανάπτυξη και η επικύρωση καινούργιων μεθοδολογιών επεξεργασίας των δεδομένων. Κάτω από αυτό το πλαίσιο, μια νέα στρατηγική για την επεξεργασία των μεταβολομικών δεδομένων από φασματόμετρο μάζας αναπτύχθηκε, με στόχο να αυξηθεί ο αριθμός των μεταβολιτών που ανιχνεύονται με αξιοπιστία. Αυτή η μεθοδολογία εφαρμόστηκε στη μεταβολομική μελέτη των δειγμάτων πλάσματος κοτόπουλου, στα οποία χορηγήθηκαν εσπεριδίνη ή ναργινίνη για

30 ημέρες. Τέσσερις μεταβλητές εμφανίστηκαν να διαχωρίζουν την ομάδα ελέγχου από την ομάδα χορήγησης της ναργινίνης, ενώ μόνο μία μεταβλητή συνέβαλλε στο διαχωρισμό της ομάδας ελέγχου από την ομάδα χορήγησης της εσπεριδίνης.

Επιπλέον, δυο μεταβολομικές μελέτες πραγματοποιήθηκαν (μία για τη ναργινίνη και μία για την εσπεριδίνη) σε δείγματα ιστών. Μελετήθηκαν οι μεταβολομικές αλλαγές της απόκρισης που συμβαίνουν στους ιστούς, λόγω της χορήγησης της ναργινίνης και εσπεριδίνης μέσω της τροφής. Έξι μεταβλητές εντοπίστηκαν να διαχωρίζουν την ομάδα ελέγχου από την ομάδα χορήγησης της ναργινίνης και δώδεκα μεταβλητές να διαχωρίζουν την ομάδα ελέγχου από την ομάδα χορήγησης της εσπεριδίνης.

## **General Introduction-Aim of the study**

Flavonoids are an important class of naturally occurring compounds present in foods, such as fruits, vegetables and plant derived juices including tea, coffee and wine. Hesperidin and naringin are two citrus flavonoids which present a broad spectrum of biological activities. In the past decade an increasing number of studies regarding the positive effects of these natural compounds have been reported. In the present study the fate as well the putative biochemical effects of naringin and hesperidin were examined after their dietary supplementation in Ross 308 broiler chickens, in order to increase the antioxidant capacity of their tissues thereof.

A fully validated UHPLC-HRMS methodology for the simultaneous quantitation of naringin, hesperidin and their aglycones, naringenin and hesperetin, was initially developed and applied in chicken plasma samples that were administrated with either hesperidin or naringin on their nutrition. The quantitation was performed at 3 time points, after 4 and 8 hours, and after 30 days of dietary supplementation. According to the results, only naringin concentration levels were above the lower limit of quantification (LLOQ) after 30 days of administration. Due to the low concentration levels of the analytes, the fate of the administered substances has been examined in plasma samples after 30 days of administration, employing a High Resolution Mass Spectrometry based metabolism study. For naringin two metabolite modifications, namely hydroxylation-methylation and methylation were detected whereas for its aglycone naringenin, desaturation and glutathione conjugation were detected as possible metabolite modifications. For hesperidin, the metabolite modifications methylation and double bond reduction were detected whereas no metabolite modifications were detected for hesperetin.

In order to assess the impact of the administered substances on the total metabolic profile of the experimental animals a full metabolomics study was deemed as imperative. As there is an increasing need of more efficient methodologies in order to accomplish the untargeted total metabolome mapping of such difficult samples as plasma or tissues a new methodology in the data processing pipeline of MS based metabolomics was developed with

the aim to increase the number of the reliable metabolite identifications. Thus shredding of the LC-MS chromatograms into multiple  $m/z$  ranges increased the number of the identified features. This methodology was applied to the untargeted metabolomics study in the chicken plasma samples that were administered either with hesperidin or naringin for 30 days. Four variables were detected to discriminate control group from the group after naringin administration. These variables were identified as 4-hydroxyphenylacetate, cAMP, glycerophospholipid and threonine along with an unidentified variable  $m/z$  526.9852. Moreover, only one substance namely 2,4-octadecadienoic acid could effectively discriminate the hesperidin treated from the control group.

Furthermore, a fully validated methodology for the simultaneous quantitative determination of hesperidin, naringin, hesperetin and naringenin in chicken tissue samples has been developed employing UHPLC-HRMS (Orbitrap). This methodology was applied in chicken tissue samples after 30 days of dietary administration with hesperidin and naringin. Nevertheless, no detectable amounts have been discovered in the samples.

Hence, in order to characterize the metabolic changes that occur in tissues in response to naringin and hesperidin dietary supplementation in chickens two metabolomics studies (for naringin and hesperidin) has been held in tissue samples. Six variables were detected to discriminate naringin dietary supplementation group from control group and twelve variables hesperidin supplementation from control group.

Overall two new fully validated UHPLC-HRMS methods have been developed for the targeted quantification of the analytes in plasma and tissue, whereas two HRMS based full metabolism analyses were undertaken in plasma. Furthermore, four metabolomic studies were conducted with the aim of clarifying the effect of the administered substances in plasma and tissues. Finally one novel metabolomics data processing methodology has been developed, found to greatly facilitate the discovery of new metabolites.

# Chapter 1

## **Validated UHPLC-ESI(-)-HRMS methodology for the simultaneous quantitative determination of hesperidin, hesperetin, naringin and naringenin in chicken plasma**

Our first goal was to establish a fully validated methodology for the simultaneously quantitation of hesperidin, naringin, hesperetin and naringenin in chicken plasma samples, in order to assess their antioxidant capacity correlated with experiments held by the Agricultural University of Athens.

## 1. Introduction

Flavonoids are an important class of naturally occurring compounds, low molecular weight, widely distributed in the plants and plant based products as secondary metabolites. They represent one of the most important and interesting classes of biologically active compounds and occur both in the free state and as glycosides [28]. Flavonoids are associated with a broad spectrum of health promoting effects.

Hesperidin (3,5,7 – trihydroxyflavanone 7- rhamnoglucoside) is a member of the flavanone group of flavonoids, that is found abundantly in oranges and other citrus fruits and herbal products such as lemons, grapes etc. As Knekt et al. (2002) has reported, hesperidin is the most highly consumed flavonoid with an intake of 28.3 mg per day, equivalent to 30% of the total flavonoid intake [12]. It presents a broad spectrum of well established biological activities like antioxidant, anti-inflammatory, anticarcinogenic and antiallergic [7]. It is metabolized in the small intestine to leave aglycone bioflavonoid hesperetin [29]. Considering the wide range of pharmacological properties of hesperidin a large number of preclinical studies and clinical trials demonstrated therapeutical effects in various diseases, such as neurological disorders, psychiatric disorders and cardiovascular diseases [14]. Hesperidin has a limited bioavailability due to its low water solubility and its disposition by phase II enzymes. A pharmacokinetic study in rats showed that the consumption of 18.9 mg/kg of hesperidin resulted into a maximum blood concentration ( $C_{max}$ ) of 0.35  $\mu\text{M}$  at a time-to-maximum ( $T_{max}$ ) of 6 hours [16]. Furthermore, in a double-blind, randomized, crossover study, healthy volunteers consumed orange juice containing hesperidin either 0.93 mg/kg body weight or 2.92 mg/kg body weight. The maximum blood concentration after subjects consumed the high dose of hesperidin juice was 1.05  $\mu\text{M}$  at 7.4 hours. This value was 2-fold higher compared with those consuming the low dose of hesperidin juice (0.48  $\mu\text{M}$  at 7.0 hours).

Naringin is a flavonoid found in many citrus and grape fruits, beans, cherries, cocoa, oregano and tomatoes [21, 5]. It is present in grapefruit juice up to concentrations of 800 mg/L [22]. Like most flavonoids, naringin possesses a wide spectrum of beneficial effects as it exhibits anti-inflammatory [13], anti-ulcer [20], superoxide, and antioxidation activities [3]. It can be metabolized into naringenin and naringenin glucuronide [8-10]. Naringenin, the aglycone of naringin, has also been found to exhibit anti-ulcer [20], antioxidant [3] effects, and inhibit the breast cancer proliferation [25]. Naringin has been reported to inhibit P-gp [27] as well as CYP3A1/2 [31] in rats. Extensive studies to elucidate the pharmacokinetic properties of naringin have been performed on rats [20, 17, 25, 28, 7], rabbits [29, 8], dogs [6, 2], and humans [1, 19, 8]. According to Hsiu et al. the  $C_{max}$  of naringenin and its conjugates occurred around 10 min, whereas for naringin administration, it occurred around 90 min [8].

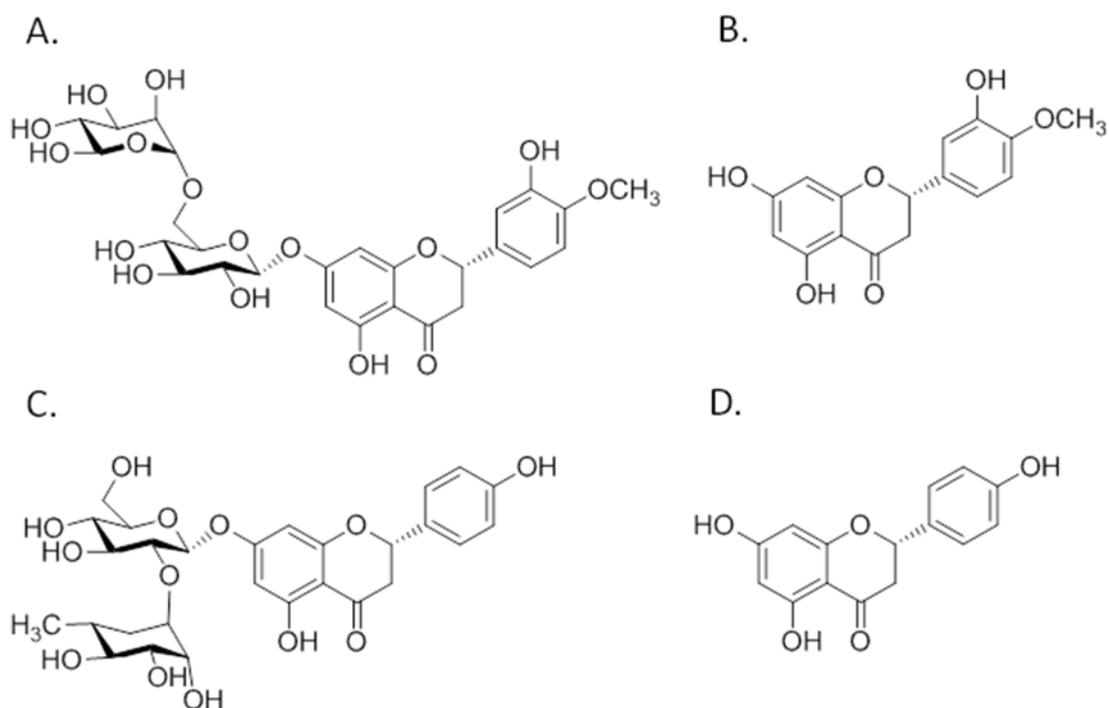
In order for flavonoids to exert their health effects *in vivo*, it is essential that they are bioavailable and absorbed from the gastrointestinal tract into the circulatory system. LC-MS/MS has been demonstrated to be the most valuable tool for pharmacokinetic studies because of its higher sensitivity and specificity compared to other analytical tools, such as HPLC—UV [18, 26, 17, 15].

Therefore, the goal of the present study was to establish and validate a sensitive, accurate and reliable UHPLC-HRMS methodology for the simultaneous quantitation of naringin, hesperidin, naringenin and hesperetin in chicken plasma. Chickens (*Gallus gallus domesticus*) have been used extensively as research models for normal human biology as well as pathological disease processes. To the best of our knowledge, no biological analytical method for the simultaneous determination of the hesperidin, naringin and their aglycones in chicken plasma has been reported so far.

## 2. Materials and methods

### 2.1 Chemicals and Reagents

All solvents were of LC-MS grade. Acetonitrile and water was purchased from Fluka/Riedel-de Haën (Switzerland). Glacial acetic acid, methanol, 4-iodophenol and acetone were purchased from Sigma-Aldrich (Steinheim, German). Naringin was purchased from Alfa Aesar GmbH & Co KG (Germany) and hesperidin from TSI Europe NV (Belgium). The molecular structures of hesperidin, hesperetin, naringin, and naringenin are displayed in figure 1.



**Figure 1** The chemical structures of hesperidin (A), hesperetin (B), naringin (C), and naringenin (D)

### 2.2 Samples treatment and study design

Fifty 1-day-old Ross 308 broiler chickens were randomly and equally divided into 5 groups. The chickens were obtained from a commercial hatchery and

were reared in pens, of a surface area of 2 m<sup>2</sup> in a controlled environment. The lighting program consisted of 23L:1D on arrival, and was decreased to 18L:6D at day 7, remained constant until day 35, and thereafter gradually increased to 23L:1D at slaughter, with access to feed, in mash form, and water *ad libitum*. The treatment groups were: N1 and N2, given diets supplemented with 0.75 g/kg (low-dose naringin treatment) and 1.5 g/kg (high-dose naringin treatment) of naringin respectively, E1 and E2, given diets supplemented with 0.75 g/kg (low-dose hesperidin treatment) and 1.5 g/kg (high-dose hesperidin treatment) of hesperidin respectively. The control group was given commercial basal diets. The administration of naringin and hesperidin started from the 11<sup>th</sup> day of age until slaughter at the age of 42 days. Blood samples were collected from each group after 4 and 8 hours and 30 days after administration into heparinized tubes and centrifuged at 2,500 rpm for 10 minutes at room temperature to recover plasma. Plasma samples were kept at -80<sup>o</sup>C until subsequent analysis.

All experimentation was carried out in strict accordance with the guidelines of “Council Directive 86/609/EEC regarding the protection of animals used for experimental and other scientific purposes”. The protocol was approved by the Bioethical Committee of the Agricultural University of Athens (Permit Number: 20/20032013).

### **2.3 Preparation of the unknown samples**

For the protein precipitation of plasma proteins, cold acetone was added to reach a final ratio of 1/8 plasma/acetone (v/v). Extraction was conducted by vortexing for 1 min and subsequently the mixture was left for 10 minutes at -20<sup>o</sup>C in order to achieve total precipitation. The extracts were then centrifuged at 12,500 rpm for 10 minutes at 4<sup>o</sup>C and the supernatant was collected and evaporated to dryness under vacuum. The residue was reconstituted with 50 µl of methanol/water (50/50, v/v) and centrifuged at 12,500 rpm for 8 minutes at 4<sup>o</sup>C. A 10 µl aliquot was injected into the UHPLC-HRMS.

## 2.4 Instrumentation

An ESI-LTQ-Orbitrap Velos (Thermo Scientific, Bremen, Germany) connected to an Accela UHPLC system was used for the analysis. The UHPLC system was consisted by an Accela quaternary pump, an Accela autosampler, a vacuum degasser and a temperature-controlled column compartment. The analysis was performed in the negative ion mode. The parameters for ion generation, transmission and detection were optimized to achieve optimal sensitivity and selectivity. Operational parameters of the mass spectrometer were set as follows: sheath gas, 35 (arbitrary units); source voltage, 3 kV; auxiliary gas, 30 (arbitrary units); S lens RF level, 60 (%) and capillary temperature, 350°C.

The Orbitrap resolution was set at 30,000 FWHM and the isolation width was set at 2 amu. Plasma samples were injected onto a reversed-phase INTERCHIM UHPLC C18 column (1.7 µm particle size, 2.1 mm x 100 mm) maintained throughout all experiments at 40°C at a flow rate of 0.3 mL/min. The auto-sampler was maintained at 4°C and the injection volume was 10 µL. The mobile phase consisted of aq. 0.1% glacial acetic acid (A), acetonitrile (B) and isopropanol/acetonitrile/acetone (58/40/2, v/v) (C). The gradient elution program was conducted as follows: 0 to 0.1 min: 95% A: 5% B: 0% C, 0.1 to 1.1 min: 80% A: 20% B: 0% C, 1.1 to 3.1 min: 70% A: 30% B: 0% C, 3.1 to 5.3 min: 50% A: 50% B: 0% C, 5.3 to 5.8 min: 0% A: 0% B: 100% C, 5.8 to 12 min: 95% A: 5% B: 0% C. The chromatographic data were acquired and processed using the Xcalibur software (Version 2.1).

Centrifuging of the plasma samples during the sample preparation protocol was performed by a Mikro 200R centrifuge (Hettich Lab Technology, Germany). Evaporation of the samples was performed by a GeneVac HT-4X EZ-2 series evaporator Lyospeed ENABLED (Genevac Ltd, UK).

## **2.5 Preparation of standard solutions, calibration curves and quality controls samples**

Stock solutions of naringin, naringenin, hesperidin, hesperetin and 4-iodophenol (IS) were prepared with methanol at 1 mg/mL. Working solutions were daily prepared by diluting and mixing appropriate volumes of stock solutions in the initial mobile phase ratio to achieve concentrations of 0.005, 0.010, 0.020, 0.050, 0.070, 0.100, 0.250, 0.400, 0.500, 0.700 and 1 µg/mL. 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 2 µg/mL concentration level.

### **2.5.1 Preparation of pre-spike calibration curves**

Ten µl of each standard working solution and 100 µl of the IS working solution were placed into 1.5 mL centrifuge tube. An aliquot of control plasma (50 µL) was added, and the tube was vortexed for 30 sec. Proteins were precipitated by addition of 400 µL of cold acetone. Extraction was conducted by vortexing for 1 min and subsequently the mixture was left for 10 minutes at -20°C in order to achieve total precipitation. The extracts were then centrifuged at 12,500 rpm for 10 minutes at 4°C and the supernatant was collected and evaporated to dryness under vacuum using a Lyospeed HT-4X GenVac. The residue was reconstituted with 50 µL of methanol/water (50/50, v/v) and centrifuged at 12,500 rpm for 8 minutes at 4°C. Totally, 3 pre-spike calibration curves were prepared ranging from 0.005 to 1 µg/mL.

### **2.5.2 Preparation of post-spike calibration curves**

Ten µl of methanol/water (50/50, v/v) were added to 50 µL of control plasma. Proteins were precipitated by addition of 400 µL of acetone. Extraction was conducted by vortexing for 1 min and subsequently the mixture was left for 10 minutes at -20°C. The extracts were then centrifuged at 12,500 rpm for 10 minutes at 4°C. Ten µL of each standard working solution and 100 µl of the IS

working solution was added to the supernatant. Subsequently, evaporation was achieved under vacuum using a Lyospeed HT-4X GenVac. The residue was reconstituted with 50  $\mu\text{L}$  of methanol/water (50/50, v/v) and centrifuged at 12,500 rpm for 8 minutes at 4<sup>0</sup>C. Totally, 3 post-spike calibration curves were prepared ranging from 0.005 to 1  $\mu\text{g}/\text{mL}$ .

### **2.5.3 Preparation of quality control (QC) samples**

Different sets of solutions were prepared individually, in order to be used as QC samples, by serial dilution at five concentration levels of low (LQC:0.015  $\mu\text{g}/\text{mL}$ ), medium (MQC:0.13  $\mu\text{g}/\text{mL}$ ), high (HQC:0.45  $\mu\text{g}/\text{mL}$ ), lower limit of quantitation (LLOQ:0.005  $\mu\text{g}/\text{mL}$ ) and upper limit of quantitation (ULOQ:1  $\mu\text{g}/\text{mL}$ ). A total of six replicates were prepared for each QC sample.

Ten  $\mu\text{L}$  of QC samples were added to 50  $\mu\text{L}$  of control plasma. Proteins were precipitated by addition of 400  $\mu\text{L}$  of cold acetone. Extraction was conducted by vortexing for 1 min and subsequently the mixture was left for 10 minutes at -20<sup>0</sup>C. The extracts were then centrifuged at 12,500 rpm for 10 minutes at 4<sup>0</sup>C. The supernatant was collected and evaporated to dryness under vacuum. The residue was reconstituted with 50  $\mu\text{L}$  of methanol/water (50/50, v/v) and centrifuged at 12,500 rpm for 8 minutes at 4<sup>0</sup>C.

### **2.5.4 Preparation of standard calibration curve in solvent**

A standard calibration curve was constructed in methanol/water (50/50, v/v) using the following concentrations: 0.005, 0.010, 0.020, 0.050, 0.070, 0.100, 0.250, 0.400, 0.500, 0.700, and 1  $\mu\text{g}/\text{mL}$ .

## **2.6. Bioanalytical method validation**

A full method validation was performed according to ICH Q2 R1 [31], the FDA [23] and the EMA CHMP guidelines for bioanalytical methodologies [30] by evaluating the specificity, linearity, recovery, matrix effect, repeatability precision, accuracy, and lower limit of quantification.

## **3. Results and discussion**

### **3.1 Comparison of FTMS-based vs IT-based Orbitrap based quantitation**

For the quantitation of the analytes in plasma samples a comparison of the quantitation ability was performed between the Fourier transform (FTMS) and the Ion Trap (IT) mass analyzer. FTMS is an ion trap mass analyzer of high mass accuracy but of low scan speed. IT on the other hand, is also an ion trap of low resolving power, but of higher scan speed and larger capacity for accommodating ions due to its much larger size, a fact that could have a significant effect on the sensitivity obtained. The Thermo Orbitrap<sup>®</sup> mass spectrometer combines an IT with an FTMS analyzer. In the literature no definitive data are presented on the quantitation aspects using an Orbitrap, an issue that should actually need to be clarified. Thus a comparison study was performed between the quantitative capability of IT and FTMS. For this comparison between the two mass analyzers, three criteria were taken under consideration: A. the sensitivity, B. the repeatability of the same samples in each of the two analyzers and C. the number of scans obtained by the same samples on IT and FTMS.

The sensitivity comparison between the two analyzers was based on the slope comparison of 4 calibration curves i.e. from naringin, naringenin, hesperidin and hesperetin. Each calibration curve was constructed from 11 standard working samples from each analyte. Thus four calibration curves were evaluated for the IT and four for the FT. For the comparison of the sensitivities obtained from each detection mode, the linear range of the

calibration curves for each compound has been taken into account. The criterion of linearity has been the  $R^2$  coefficient, readjusting the number of points to be included for its estimation. As it is well known the slope of a calibration curve denotes the sensitivity of the respective detection. Thus the ratio of the linear equation slopes for each one of the 4 analytes between the IT and the FTMS, exhibits the detection ability of each analyzer compared to its counterpart (Tables 1, 2).

Analyte	FTMS-based Equation	IT-based Equation
Naringin	$Y=722(\pm 41)X+1732(\pm 1654)$	$Y=644(\pm 7)X+(-119)(\pm 307)$
Hesperetin	$Y=1777(\pm 214)X+13548(\pm 11719)$	$Y=20629(\pm 556)X+48416(\pm 30441)$
Hesperidin	$Y=1306(\pm 39)X+12613(\pm 4276)$	$Y=12755(\pm 282)X+(-10960)(\pm 30279)$
Naringenin	$Y=3190(\pm 147)X+19522(\pm 8066)$	$Y=28180(\pm 373)X+61257(\pm 20407)$

**Table 1** Equations of FTMS-based quantitation methodology and IT-based quantitation methodology for the analytes naringin, hesperetin, hesperidin and naringenin

Analyte	FTMS-IT Slope Ratio	IT-FTMS Slope Ratio
Naringin	1.12	0.89
Hesperetin	0.08	11.60
Hesperidin	0.10	9.76
Naringenin	0.11	8.83

**Table 2** For the comparison of the sensitivity of the two methodologies (FTMS-based vs IT-based quantitation) the linear range of the calibration curves for each compound has been taken into account. the ratio of the linear equation slopes for each one of the 4 analytes between the IT and the FTMS, exhibits the detection ability of each analyzer compared to its counterpart

According to the results, the IT mass analyzer presented enhanced sensitivity compared to FTMS for the analytes hesperidin, naringin and naringenin whereas FTMS was more sensitive for naringin.

In order to examine the repeatability of the two analyzers, 6 replicates of 5 quality controls samples at concentrations levels of 0.005 (LLOQ), 0.015 (LQC), 0.13 (MQC), 0.45 (HQC), and 1 µg/mL (ULOQ) were analyzed and compared. The results were expressed as the relative standard deviation (%RSD). The excluded criteria was set as %RSD<20 for the LLOQ and <15 for the QC samples (Table 3).

<b>Analyte</b>	<b>Quality Control Samples</b>	<b>FTMS-based method (%RSD)</b>	<b>IT-based method (%RSD)</b>
Naringin	<b>ULOQ</b>	6.97	4.55
	<b>HQC</b>	9.74	8.30
	<b>MQC</b>	4.65	7.47
	<b>LQC</b>	8.73	11.05
	<b>LLOQ</b>	19.89	19.25
Hesperidin	<b>ULOQ</b>	5.84	5.45
	<b>HQC</b>	11.38	10.01
	<b>MQC</b>	4.26	7.00
	<b>LQC</b>	11.75	12.28
	<b>LLOQ</b>	17.61	18.65
Naringenin	<b>ULOQ</b>	8.68	4.09
	<b>HQC</b>	14.57	13.45
	<b>MQC</b>	3.76	8.39
	<b>LQC</b>	13.69	12.31
	<b>LLOQ</b>	15.21	11.76
Hesperetin	<b>ULOQ</b>	11.22	5.81

	<b>HQC</b>	14.89	12.10
	<b>MQC</b>	6.37	8.18
	<b>LQC</b>	11.80	12.73
	<b>LLOQ</b>	19.33	17.63

**Table 3** Comparison of the repeatability of the FTMS-based quantitation methodology and the IT-based quantitation methodology after the analysis of 6 replicates of 5 QC levels at concentrations levels of 0.005, 0.015, 0.13, 0.45, and 1 µg/mL. The results were expressed as the relative standard deviation (%RSD). The excluded criteria was set as %RSD<20 for the LLOQ and <15 for the QC samples

As table 3 showed the FTMS-based and the IT-based methodology presented similar levels of repeatability.

The number of data points (number of scans) obtained by the same samples were calculated in order to define a chromatographic peak on a low concentration level (70 ng/mL) and on a high concentration level (700 ng/mL) for the 4 analytes. As tables 4 and 5 depicted the two methodologies presented similar number of scans.

<b>Analyte</b>	<b>Number of Scans</b>	
	<b>FTMS-based method</b>	<b>IT-based method</b>
Naringin	11	10
Hesperidin	11	10
Naringenin	10	10
Hesperetin	11	10

**Table 4** Number of scans for the FTMS-based vs IT-based methodology at the concentration level of 70 ng/mL for the analytes naringin, hesperetin, hesperidin and naringenin

Analyte	Number of Scans	
	FTMS-based method	IT-based method
Naringin	10	11
Hesperidin	10	11
Naringenin	10	10
Hesperetin	11	10

**Table 5** Number of scans for the FTMS-based vs IT-based methodology at the concentration level of 700 ng/mL for the analytes naringin, hesperetin, hesperidin and naringenin

After the comparison according to the criteria of the sensitivity, repeatability and the number of scans, the IT-based methodology was preferred for the quantitation of the chicken plasma samples as it presented enhanced sensitivity though similar levels of repeatability and number of scans compared to FTMS-based methodology.

### 3.2 UPLC-ESI(-)-HRMS optimization

Mass spectrometric parameters were optimized by examining various parameters (mobile phase, flow rate, type of column chemistry etc). Consequently, the described methodology was selected because of the successful baseline separation and the symmetrical chromatographic peaks that were achieved, leading to enhanced signal intensity for the four analytes. Table 6 demonstrates the optimized collision energy that was used to afford well-described transitions and highly stable signal for the four analytes.

Analyte	Parent mass ( <i>m/z</i> )	Product mass ( <i>m/z</i> )	Isolation width (amu)	CID (V)
Naringin	579	459.11±0.5	2	30
Hesperidin	609	301.07±0.5	2	21
Naringenin	271	151.00±0.5	2	45
Hesperetin	301	286.05±0.5	2	49

**Table 6** Mass spectrometry parameters for the quantitation of naringin, hesperidin, naringenin and hesperetin in chicken plasma samples

### 3.3 Method Validation

#### 3.3.1 Specificity, carry-over and linearity

Specificity of the proposed methodology, for the quantitation of the four analytes, was evaluated by analyzing six blank plasma samples. The blank plasma samples were prepared as has been described previously. No significant interfering peaks from endogenous compounds of plasma were presented at the corresponding retention time of the analytes.

Blank samples were analyzed right after that of plasma with the highest spiked concentration, for a total of six times, in the carry-over test. The absence of carry-over effect was observed as the injection of the blank samples showed no peaks at the corresponding retention time of the analytes.

The linearity of the calibration curve was checked by regression analysis and the goodness of the regression by calculating the Pearson's determination coefficient  $R^2$ . A strong correlation between the peak area ratio and concentration of each analyte in the range of 0.005-1  $\mu\text{g/mL}$  was illustrated by the high  $R^2$  of  $> 0.990$ .

### 3.3.2 Recovery, matrix effect

Recovery represents the extraction efficiency of the analytical methodology. It was assessed by analyzing 3 calibration curves from pre-spike samples and 3 curves from post-spike samples. One curve was constructed from the calibration curves of pre-spike samples and one curve from post-spike samples for each analyte. The slope ratio of the two calibration curves was compared. As Table 7 presents the slope ratios were found to be greater than 80% for each analyte. This showed that the extraction protocol was efficient and suitable for the analysis of the four analytes in chicken plasma.

Analyte	Slope ratio
Naringin	0.88
Hesperidin	0.91
Naringenin	0.86
Hesperetin	0.92

**Table 7** Recovery of the methodology was determined after comparing the slope ratios of the calibration curve from pre-spike samples and the calibration curve from post-spike samples ranging from 0.005 to 1 µg/mL

Matrix effect arises due to effects of endogenous components of the plasma on the ionization process of the analytes of interest. The matrix effects were investigated by comparing the slopes of the calibration curve prepared in methanol/water (50/50, v/v) and the calibration curve prepared by the post-spike samples. The slope ratios were found to be greater than 90% for each analyte and therefore matrix effect would not adversely affect the accuracy and precision of the methodology (Table 8).

Analyte	Slope ratio
Naringin	0.99
Hesperidin	0.96

Naringenin	0.93
Hesperetin	0.95

**Table 8** The matrix effect was validated by comparing the slopes of the calibration curve prepared in methanol/water (50/50, v/v) and the calibration curve prepared by pre-spike samples

### 3.3.3 Repeatability, accuracy and intermediate precision

The repeatability (within-run), intermediate precision (between-run) and accuracy of the methodology for the quantitation of the 4 analytes in chicken plasma samples were assessed by analyzing 6 replicates of 5 QC levels at concentrations levels of 0.005, 0.015, 0.13, 0.45, and 1 µg/mL. The results for the intermediate precision and repeatability were expressed as the relative standard deviation (%RSD). The excluded criteria was set as %RSD<20 for the LLOQ and <15 for the QC samples (Tables 9, 10).

Analyte	LLOQ (0.005)	LQC (0.015)	MQC (0.13)	HQC (0.45)	ULOQ (1)
Hesperidin	10.87%	14.92%	5.69%	11.00%	6.09%
Naringin	11.25%	8.66%	7.02%	6.39%	5.78%
Hesperetin	18.11%	6.37%	9.11%	12.70%	6.01%
Naringenin	16.05%	7.81%	9.55%	11.55%	4.89%

**Table 9** Validation results from the intermediate precision of the proposed methodology after the analysis of 6 replicates of 5 QC levels at concentrations levels of 0.005, 0.015, 0.13, 0.45, and 1 µg/mL. The results were expressed as the relative standard deviation (%RSD). The excluded criteria was set as %RSD<20 for the LLOQ and <15 for the QC samples

	<b>LLOQ</b>	<b>LQC</b>	<b>MQC</b>	<b>HQC</b>	<b>ULOQ</b>
<b>Analyte</b>	<b>(0.005)</b>	<b>(0.015 )</b>	<b>(0.13)</b>	<b>(0.45)</b>	<b>(1)</b>
Hesperidin	9.55%	12.29%	7.00%	10.02%	5.46%
Naringin	15.85%	11.06%	7.48%	8.31%	4.55%
Hesperetin	19.17%	6.09%	8.40%	11.03%	4.10%
Naringenin	17.45%	7.55%	8.40%	11.87%	4.10%

**Table 10** Validation results from the repeatability of the proposed methodology after the analysis of 6 replicates of 5 QC levels at concentrations levels of 0.005, 0.015, 0.13, 0.45, and 1 µg/mL. The results were expressed as the relative standard deviation (%RSD). The excluded criteria was set as %RSD<20 for the LLOQ and <15 for the QC samples

Accuracy was expressed as the %standard error (%Er) between the mean concentration and the nominal concentration. The excluded criteria was set as %Er<20 for the LLOQ and <15 for the QC samples (Table 11).

	<b>LLOQ</b>	<b>LQC</b>	<b>MQC</b>	<b>HQC</b>	<b>ULOQ</b>
<b>Analyte</b>	<b>(0.005)</b>	<b>(0.015 )</b>	<b>(0.13)</b>	<b>(0.45)</b>	<b>(1)</b>
Hesperidin	11.98%	10.33%	7.33%	6.72%	4.99%
Naringin	19.02%	9.16%	7.08%	5.10%	3.11%
Hesperetin	13.17%	13.43%	9.51%	9.13%	5.87%
Naringenin	11.38%	12.05%	11.32%	9.07%	6.09%

**Table 11** Validation results from the accuracy of the proposed methodology after the analysis of 6 replicates of 5 QC levels at concentrations levels of 0.005, 0.015, 0.13, 0.45, and 1 µg/mL. The results were expressed as the %standard error (%Er) between the mean concentration and the nominal concentration. The excluded criteria was set as %Er<20 for the LLOQ and <15 for the QC samples.

Taken together, the repeatability did not exhibit values more than 19.17% for the LLOQ and 12.29% for the QC samples whereas the precision did not exhibit values more than 18.11% for the LLOQ and 14.92% for the QC samples. The %Er of the accuracy did not exhibit values more than 19.02% for the LLOQ and 13.43% for the QC samples. These results indicate that the proposed methodology for the quantitation of the 4 analytes in chicken plasma samples is reliable.

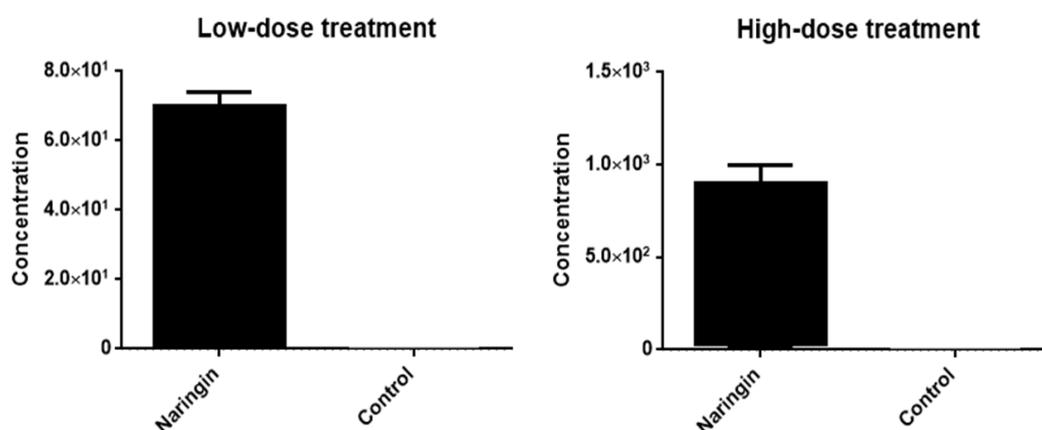
#### **3.3.4 Lower limit of quantification**

The lower limit of quantification (LLOQ) is the lowest concentration of an analyte in a sample which can be quantified reliably. The analyte signal of the LLOQ sample should be at least 5 times the signal of a blank sample and to yield a peak with a signal/noise ratio of at least 10. The LLOQ was found to be for NN, HD, NG and NN at least 0.005 µg/mL with an acceptable accuracy and precision at this concentration level according to the ICH rules.

#### **3.4 Application of the validated UPLC-ESI(-)-HRMS methodology**

The developed methodology for the simultaneous quantitation of naringin, hesperidin, naringenin and hesperetin was applied in plasma samples from chickens that were administrated with hesperidin and naringin on their nutrition. The treatment groups were: N1 and N2, supplemented with 0.75 and 1.5 g of naringin per kg of feed respectively, E1 and E2, supplemented with 0.75 and 1.5 g of hesperidin per kg of feed respectively, and control group with no feed additive. The quantitation in the chicken plasma samples was

performed on 3 time points, after 4 hours of dietary supplementation, after 8 hours and after 30 days. The quantitation was based on standard curves that were prepared for each analyte on the day of the analysis, ranging from 0.005 to 1  $\mu\text{g}/\text{mL}$ . According to the results, naringin concentration levels after dietary supplementation with 0.75 g/kg of feed (low-dose treatment) and 1.5 g/kg of feed (high-dose treatment) were above the LLOQ only after 30 days of administration (Figure 2). Nevertheless, measurements of the chicken plasma samples after hesperidin administration did not reveal the existence of circulating either hesperidin or hesperetin in concentrations above LLOQ. Similar results were observed for naringenin after naringin supplementation.



**Figure 2** Naringin concentration levels after dietary supplementation with 0.75 g/kg of feed (low-dose treatment) and 1.5 g/kg of feed (high-dose treatment) were above LLOQ after 30 days of administration

#### 4. Concluding remarks

A highly sensitive and specific methodology was developed for the simultaneous determination of hesperidin, hesperetin, naringin and naringenin in chicken plasma employing UHPLC-HRMS (Orbitrap Velos). Plasma samples were prepared by protein precipitation with cold acetone. Detection was performed by means of electrospray ionization (ESI) in the negative ion mode. The calibration curves for the 4 analytes exhibited good linearity ( $r^2 > 0.990$ ) over the concentration range of 0.005 to 1  $\mu\text{g}/\text{mL}$  with a lower limit of quantification (LLOQ) of 0.005  $\mu\text{g}/\text{mL}$ . The repeatability and

precision of the proposed methodology were within %RSD<20 and accuracy within %Error<20. No matrix effect and carry over was observed. The validated methodology was applied in plasma samples after dietary supplementation with 0.75 g/kg of feed (low-dose treatment) and 1.5 g/kg of feed (high-dose treatment) of hesperidin and naringin in Ross 308 broiler chickens.

Only naringin concentration levels were detected after 30 days of administration. However, no circulation levels of hesperidin, hesperetin and naringenin were observed above the LLOQ.

Two essential factors influence flavanones bioavailability: (i) the physical form in which they are ingested (e.g., juice, soluble extract, capsule) and (ii) the biotransformation in the gut by intestinal microflora (phase I, phase II, and microbiota metabolism). The difference in the circulating flavanone concentrations observed between these two molecules in chicken plasma remains quite surprising since these two flavanones exhibit quite similar chemical structure, lipophilicity and both presented high intestinal and biliary secretion [24]. It has been previously reported that flavonoid bioavailability tightly depends on the efficiency of their transfer through the brush border, on the capacity of the intestine to secrete conjugated metabolites and on the intensity of their splanchnic metabolism [4]. Therefore, we can assume that bacterial rhamnosidases may present differentiated activity between hesperidin and naringin. Furthermore, hesperidin, as a rutinoside, must pass onto the colon and be fermented by intestinal microflora to an alternate form that is more readily absorbed [11].

Moreover, once absorbed into enterocytes, flavanones can be conjugated and efficiently transported back into the gut lumen by active transporters. This process may be responsible for the limited access of flavanone to systemic circulation.

Hence, in accordance to the findings from the current study, it has been reported that free flavanone aglycones, such as naringenin and hesperetin, are detected in trace amounts after the consumption of flavanone-rich

products [6, 2, 1, 19]. Therefore the metabolism investigation of naringin and hesperidin is pertinent in order to decipher their fate in the organism whereas their metabolomic investigation will be highly desirable in order to shed light to their biochemical impact.

## References

1. Brett GM, Hollands W, Needs PW, Teucher B, Dainty JR, Davis BD, Brodbelt JS, Kroon PA. (2009) Absorption, metabolism and excretion of flavanones from single portions of orange fruit and juice and effects of anthropometric variables and contraceptive pill use on flavanone excretion. *Br J Nutr.* 101, 664-675.
2. Bugianesi R, Catasta G, Spigno P, D'Uva A, Maiani G. (2002) Naringenin from cooked tomato paste is bioavailable in men. *J Nutr.* 132, 3349-3352.
3. Chen Y.T., Zheng R.T., Jia ZJ., Ju Y. (1990) Flavonoids as superoxide scavengers and antioxidants. *Free Rad. Biol. Med.* 9, 19–21.
4. Crespy V, Morand C, Besson C, Cotelle N, Vezin H, Demigne C, Rémésy C. (2003) The splanchnic metabolism of flavonoids highly differed according to the nature of the compound. *Am J Physiol Gastrointest Liver Physiol.* 284, G980–988.
5. del Baño MJ, Lorente J, Castillo J, Benavente-García O, del Río JA, Ortuño A, Quirin KW, Gerard D. (2003) Phenolic diterpenes, flavones, and rosmarinic acid distribution during the development of leaves, flowers, stems, and roots of *Rosmarinus officinalis*. *Antioxidant activity.* 51, 4247-4253.
6. Felgines C, Texier O, Morand C, Manach C, Scalbert A, Régerat F, Rémésy C. (2000) Bioavailability of the flavanone naringenin and its glycosides in rats. *Am J Physiol Gastrointest Liver Physiol.* 279, G1148-1154.
7. Garg A., Garg S. Zaneveld L.J., Singla A.K. (2001) Chemistry and pharmacology of the Citrus bioflavonoid hesperidin. *Phytother Res.* 15, 655–669.

8. Hsiu SL, Huang TY, Hou YC, Chin DH, Chao PD (2002) Comparison of metabolic pharmacokinetics of naringin and naringenin in rabbits. *Life Sci.* 70, 1481–1489.
9. Ishii K., Furuta T., Kasuya Y. (1996) Determination of naringin and naringenin in human plasma by high-performance liquid chromatography. *J. Chromatogr. B.* 683, 225–229.
10. Ishii K., Furuta T., Kasuya Y. (1997) Determination of naringin and naringenin in human urine by high-performance liquid chromatography utilizing solid-phase extraction. *J. Chromatogr. B.* 704, 299–305.
11. Jin M.J., Kim U., Kim I.S., Kim Y., Kim D.H., Han S.B., Kim D.H., Kwon O.S., Yoo H.H. (2010) Effects of gut microflora on pharmacokinetics of hesperidin: a study on nonantibiotic and pseudo-germ-free rats. *J Toxicol Environ Health A.* 73, 1441-1450.
12. Knekt P., Kumpulainen J., Jarvinen R., Rissanen H., Heliövaara M., Reunanen A., Hakulinen T., Aromaa A. (2002). Flavonoid intake and risk of chronic diseases. *Am J Clin Nutr.* 76, 560-568.
13. Lambev I., Krushkov I., Zheliazkov D., Nikolov N. (1980) Antiexudative effect of naringin in experimental pulmonary edema and peritonitis. *Eksp. Med. Morfol.* 19, 207–212.
14. Li C, Schluesener H. (2015) Health-promoting Effects of the Citrus Flavanone Hesperidin. *Crit Rev Food Sci Nutr.* 12:0.
15. Li GL, Han WL, Jiang W, Zhang DM, Ye WC, Chen XJ (2013). Quantitative determination of arenobufagin in rat plasma by ultra fast liquid chromatography–tandem mass spectrometry and its application in a pharmacokinetic study. *J Chromatogr B.* 939, 86–91.
16. Li Y.M., Li X.M., Li G.M., Du W.C., Zhang J., Li W.X., Xu J., Hu M., Zhu Z. (2008) In vivo pharmacokinetics of hesperidin are affected by

- treatment with glucosidase-like BglA protein isolated from yeasts. *J Agric Food Chem.* 56, 5550-5557.
17. Liu AC, Zhao LX, Xing J, Gao J, Lou HX. (2013) LC–MS/MS method for the determination of a new puerarin derivative and its application in pharmacokinetic studies in rats. *Chin J Nat Med.* 11, 566–571.
  18. Liu QW, Wang JS, Yang L, Jia YW, Kong LY. (2013) A rapid and sensitive LC–MS/MS assay for the determination of berbamine in rat plasma with application to preclinical pharmacokinetic study. *J Chromatogr B.* 929, 70–75.
  19. Manach C, Morand C, Gil-Izquierdo A, Bouteloup-Demange C, Rémésy C. (2003) Bioavailability in humans of the flavanones hesperidin and narirutin after the ingestion of two doses of orange juice. *Eur J Clin Nutr.* 57, 235-242.
  20. Martin M.J., Marhuenda E., Perez-Guerrero C., France J.M. (1994) Antiulcer effect of naringin on gastric lesions induced by ethanol in rats. *Pharmacology.* 49 144–150.
  21. Mouly P., Gaydou E.M., Estienne J. (1993) Column liquid chromatographic determination of flavanone glycosides in Citrus. Application to grapefruit and sour orange juice adulterations. *J. Chromatogr.* 634, 129–134.
  22. Rouseff RL, Martin SF, Youtsey CO. (1987) Quantitative survey of narirutin, naringin, hesperidin, and neohesperidin in citrus. *J Agric Food Chem.* 35, 1027–1030.
  23. Shah V.P. (2000) Guidance for Industry. Bioanalytical Method Validation. *Pharmaceutical Research.* 17.
  24. Silberberg M, Morand C, Mathevon T, Besson C, Manach C, Scalbert A, Remesy C. (2006) The bioavailability of polyphenols is highly

- governed by the capacity of the intestine and of the liver to secrete conjugated metabolites. *Eur J Nutr.* 45, 88–96.
25. So F.V., Guthrie N., Chambers A.F., Moussa M., Carroll K.K. (1996) Inhibition of human breast cancer cell proliferation and delay of mammary tumorigenesis by flavonoids and citrus juices. *Nutr. Cancer.* 26, 167–181.
  26. Song M, Lee D, Lee T, Lee S. (2013) Determination of leelamine in mouse plasma by LC–MS/MS and its pharmacokinetics. *J Chromatogr B.* 931, 170–173.
  27. Tsai TH, Lee CH, Yeh PH. (2001) Effect of P-glycoprotein modulators on the pharmacokinetics of camptothecin using microdialysis. *Br. J. Pharmacol.* 134, 1245–1252.
  28. Tsao R. (2010) Chemistry and biochemistry of dietary polyphenols. *Nutrients.* 2, 1231–1246.
  29. Watanabe M, Matsumoto N, Takeba Y, Kumai T, Tanaka M, Tatsunami S, Takenoshita-Nakaya S, Harimoto Y, Kinoshita Y, Kobayashi S. (2011) Orange juice and its component, hesperidin, decrease the expression of multidrug resistance-associated protein 2 in rat small intestine and liver. *J Biomed Biotechnol.* 2011, 2011:502057.
  30. Wharf C. and Kingdom U. (2012) Guideline on bioanalytical method validation. Table of contents. 44.
  31. Zhang H, Wong CW, Coville PF, Wanwimolruk S. (2000) Effect of the grapefruit flavonoid naringin on pharmacokinetics of quinine in rats. *Drug Metabol. Drug Interact.* 17, 351–363.
  32. 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](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q2_R1/Step4/Q2_R1_Guideline.pdf). Accessed 2013 Oct 03).

## Chapter abstract

A highly sensitive and specific methodology has been developed for the simultaneous determination of hesperidin, hesperetin, naringin and naringenin in chicken plasma employing UHPLC-HRMS (Orbitrap Velos). Plasma samples were prepared by protein precipitation with cold acetone. Analysis was carried out on an INTERCHIM UHPLC C18 column using water, acetonitrile and isopropanol/acetonitrile/acetone (58/40/2, v/v) as the mobile phase. Detection was performed by means of electrospray ionization (ESI) in the negative ion. All calibration curves exhibited good linearity ( $r^2 > 0.990$ ) over the concentration range of 0.005 to 1  $\mu\text{g/mL}$  with a lower limit of quantification (LLOQ) of 0.005  $\mu\text{g/mL}$  for the 4 analytes. The repeatability and precision were within %RSD<20 and accuracy within %Error<20. No matrix effect and carry over was observed on the proposed methodology. The validated methodology was applied in plasma samples after dietary supplementation with 0.75 g/kg of feed (low-dose treatment) and 1.5 g/kg of feed (high-dose treatment) of hesperidin and naringin in Ross 308 broiler chickens.

**Keywords:** naringin-hesperidin-validation-pharmacokinetics-plasma

## Chapter 2

### **Metabolism of naringin, hesperidin and their aglycones in plasma samples after dietary supplementation of naringin and hesperidin in chickens**

Due to the low concentration levels of hesperidin, naringin and their aglycones in chicken plasma, the fate of the administered substances has been examined in plasma samples after 30 days of administration, employing a High Resolution Mass Spectrometry based metabolism study.

## 1. Introduction

Findings from the literature show that flavonoids are first biotransformed in the gut by intestinal microflora, where naringin and hesperidin (glycosides) are hydrolyzed into the free aglycones naringenin and hesperetin moieties, and both glycosides and aglycones are absorbed [8]. The intestine is known to participate in the phase II conjugation of flavanone aglycones during first-pass metabolism. The hydrophobic aglycones are able to passively diffuse through the permeable gut mucosa and conjugated within mucosal cells [3]. However, absorption into systematic circulation is limited because of the efflux of these conjugates back into the intestinal lumen by specific transporters [4, 11]. Flavanones that enter the systemic circulation may undergo metabolism in the liver by CYP450, through hydroxylation and/or O-demethylation reactions, and then subjected to glucuronidation, sulfation, and O-methylation reactions catalyzed by phase II enzymes [15]. The majority of the identified metabolites have been glucuronide and sulfate conjugates. Some sulfoglucuronide and diglucuronide conjugates have also been described but in lower concentration [14, 5]. This liver metabolism may limit the bioavailability of flavonoids in plasma *in vivo*. There is limited evidence that flavanones may also be subjected to phase I metabolism *in vivo*.

Naringin is rapidly metabolized in the liver and converted into glucuronide intermediates [6, 9, 10]. It was reported that after a single dose of naringenin and naringin administration on rats via an i.v. bolus and oral route, naringenin and naringin sulfates and glucuronides were found almost exclusively in the bloodstream. Moreover, the concentration of naringenin sulfates was much higher than that of naringenin glucuronides [18]. In a study after oral administration of naringin in rats and dogs at doses of 42 mg/kg and 12.4 mg/kg respectively, naringin underwent extensive phase I and phase II metabolism and in total twenty-two metabolites were identified in dogs and 17 metabolites were detected in rats. The observed routes of naringin metabolism were found to be hydroxylation, methylation, acetylation, hydrogenation, deglycosylation, dehydrogenation, glucuronidation, sulfation,

glucosylation, ring-fission, oxidation, glycine conjugation and dehydroxylation [12].

Hesperidin is mostly absorbed in the large intestine after deglycosylation and release of the aglycone hesperetin, which is then absorbed and further metabolized by phase II enzymes, forming several glucuronide and sulfate conjugates [14]. Additionally, it was demonstrated that the aglycone can be further degraded by the colon microbiota to render hydroxyl-phenyl propionic acid derivatives and benzoic acid derivatives that are then absorbed and detected in plasma and urine [16]. Animal studies indicate that hesperidin is absorbed as its aglycone hesperetin, after removing the disaccharide, which could be recovered in urine and bile as glucuronide [2, 7]. Manach et al. has found in human plasma hesperetin-monoglucuronides and hesperetin-sulfoglucuronides after ingestion of orange juice naturally containing hesperidin [13]. Thus, the presence of glucuronides of aglycone flavanones has been reported in human urine [1]. Moreover, Hackett et al. [7] has found only traces of hesperetin in the feces of rats after oral administration of hesperidin or hesperetin. In a randomized, placebo-controlled crossover trial, 16 participants received orange juice or a hesperidin supplement (both providing 320 mg hesperidin) or control. At baseline and 5 h post-intake plasma flavanone metabolites were assessed. Total plasma flavanone metabolite concentrations were significantly higher 5 h after the orange juice intervention than after control with hesperidin-glucuronide, contributing 47% to the total plasma flavanone concentration [17].

Nowadays, high resolution mass spectrometers seem to be the method of choice for the identification and characterization of metabolites on biofluids and tissues due to their high resolution capabilities and the enhanced mass accuracy. In the present study the metabolism of naringin, naringenin, hesperidin and hesperetin was investigated in plasma after hesperidin and naringin dietary supplementation in Ross 308 broiler chickens by high resolution MS using an Orbitrap analyzer for the structural determination of the resulting analytes.

# 1. Materials and methods

## 2.1 Chemicals and Reagents

All solvents were of LC-MS grade. Acetonitrile and water was purchased from Fluka/Riedel-de Haën (Switzerland). Glacial acetic acid, methanol and acetone were purchased from Sigma-Aldrich (Steinheim, German). Naringin was purchased from Alfa Aesar GmbH & Co KG (Germany) and hesperidin from TSI Europe NV (Belgium).

## 2.2 Animal treatment and study design

Thirty 1-day-old Ross 308 broiler chickens obtained from a commercial hatchery were reared in pens, of a surface area of 2 m<sup>2</sup> in a controlled environment. The lighting program consisted of 23L:1D on arrival, and was decreased to 18L:6D at day 7, remained constant until day 35, and thereafter gradually increased to 23L:1D at slaughter, with access to feed, in mash form, and water *ad libitum*. The chickens were randomly and equally divided into 3 groups (n=10 per group). The experimental groups consisted of 10 chickens given diets supplemented with 1.5 g/kg of feed of naringin and 10 chickens given diets supplemented with 1.5 g/kg of feed of hesperidin. The control group was given commercial basal diets. The administration was started from the 11<sup>th</sup> day of age until slaughter at the age of 42 days. Blood samples were collected on 42 days of age after 30 days of the administration into heparinized tubes and centrifuged at 2,500 rpm for 10 minutes to recover plasma. Plasma samples were kept at -80°C until the analysis.

All experimentation was carried out in strict accordance with the guidelines of "Council Directive 86/609/EEC regarding the protection of animals used for experimental and other scientific purposes". The protocol was approved by the Bioethical Committee of the Agricultural University of Athens (Permit Number: 20/20032013).

### 2.3 Sample Preparation

For the protein precipitation of plasma proteins, cold acetone was added to reach a final ratio of 1/8 plasma/acetone (v/v). Extraction was conducted by vortexing for 1 min and subsequently the mixture was left for 10 minutes at -20°C in order to achieve total precipitation. The extracts were then centrifuged at 12,500 rpm for 10 minutes at 4°C to pellet the protein precipitate and the supernatant was transferred into eppendorf tubes and evaporated to dryness under vacuum. Subsequently, the residue was reconstituted with 50 µL of methanol/water (50/50, v/v) and centrifuged at 12,500 rpm for 8 minutes at 4°C. A 2 µL aliquot was injected into the UHPLC-HRMS for the investigation of naringin, hesperidin and their aglycones metabolism.

### 2.4 Instrumentation

An UPLC-ESI(-)-HRMS analysis was performed on an LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific, Bremen, Germany) connected to an Accela UHPLC system controlled by the Xcalibur 2.1 software. The UHPLC system was equipped with an Accela quaternary pump, an Accela autosampler, a vacuum degasser and a temperature-controlled column compartment.

Operational parameters of the mass spectrometer were set as follows: sheath gas, 35 (arbitrary units); source voltage, 3 kV; auxiliary gas, 30 (arbitrary units); S lens RF level, 60 (%); and capillary temperature, 350°C in the negative ion mode. The resolution was set at 30,000 FWHM with a scan range of 100 to 900 Da in centroid mode. Blood samples were injected onto a reversed-phase INTERCHIM UHPLC C18 column (1.7 µm particle size, 2.1 mm x 100 mm) maintained throughout all experiments at 40°C. The mobile phase consisted of aq. 0.1% glacial acetic acid (A) and acetonitrile (B). A gradient methodology of 32 minutes was employed as follows: 0 to 24 min:

95% A: 5% B, 24 to 28 min: 5% A: 95% B, 28 to 32 min: 95% A: 5% B, at a flow rate of 0.36 mL/min.

Plasma samples centrifugation was performed using a Mikro 200R centrifuge (Hettich Lab Technology, Germany). Evaporation of the samples was performed by a GeneVac HT-4X EZ-2 series evaporator Lyospeed ENABLED (Genevac Ltd, UK).

## 2.5 Data Analysis

The metabolism was investigated by the MetWorks™ 1.3 software (Thermo Fisher Scientific, Inc., San Jose, CA, USA). The raw mass spectral data obtained from the LC-MS analysis were interrogated for potential metabolites of the analytes based on phase I and phase II patterns of metabolism. In total ninety-six possible metabolic modifications of naringin, hesperidin, naringenin and hesperetin were explored and evaluated. The accurate mass values of the analytes and their putative metabolites were calculated by computing the corresponding due to metabolism mass shifts, and determine the presence of extracted ion chromatograms (XICs) from the raw data.

The Mass defect filter (MDF) approach attempts to segregate characteristic ions that correspond to possible metabolites from other ions that are interfering the tested samples in a full-scan MS data set, by imposing specific criteria. All data files were processed using multiple MDFs for the elemental composition, the number of ring plus double bonds (RDBs) equivalents and the isotope pattern. The criteria for the numbers of elements present in the potential metabolites were:  $^{12}\text{C}$ : 0–40,  $^1\text{H}$ : 0–60,  $^{14}\text{N}$ : 0–5;  $^{16}\text{O}$ : 0–5, and for RDB equivalents from 0 to 20. The MDF width was set to 5 ppm around the accurate  $m/z$  of the de-protonated molecule. Moreover, the isotopic pattern was analyzed and compared to its theoretical value with tolerance at 1% difference. Once a metabolite is found by this processing methodology, its extracted ion chromatogram with the corresponding mass spectrum is

automatically generated by MetWorks™. In addition, in order to minimize the effects of the endogenous components of plasma matrix to metabolite identification, subtraction of the control plasma samples from the treated samples was performed prior to the detection of the potential metabolites. The area of the potential metabolite from the raw data was compared to the area of the control group employing a t-test. The significant level was set to 95.

For the prediction of the possible sites of metabolism and the design of the chemical structures of the possible metabolites, the online platform MetaPrint2 and the MarvinSketch drawing software were employed [20]. The structures were justified by the accurate mass measurements, the RDB values and the isotopic pattern ratio.

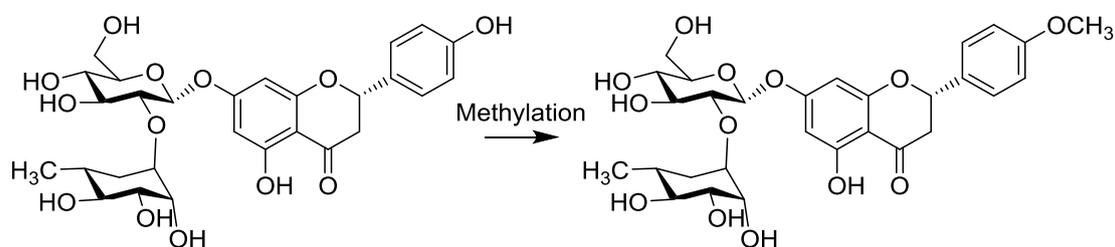
## 2. Results and discussion

### 3.1 Naringin metabolism in chicken plasma samples

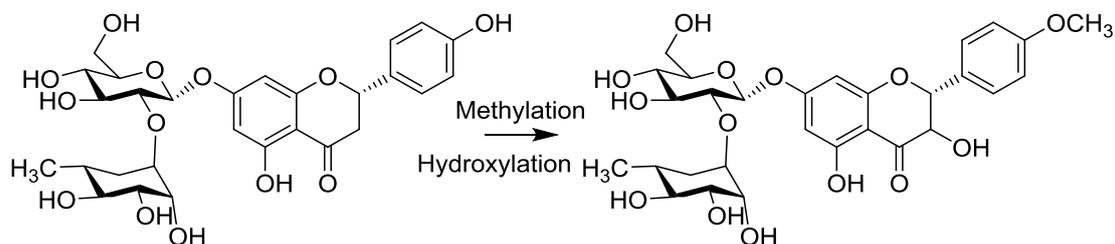
Naringin was used as a template to investigate the ninety-six possible metabolites modifications in chicken plasma samples. Two possible metabolites of naringin were identified after the data processing procedure, as described above. Table 1 summarizes the theoretical and measured mass values  $m/z$ , the number of the RDBs, the retention time ( $t_R$ ), the predicted molecular formulas, the metabolic change shift and the metabolic generation for the two possible metabolites. Their proposed chemical structures according to the accurate mass, the RDB equivalents and the isotope pattern ratio are presented in Figures 1, 2 using the online platform MetaPrint2.

Metabolites	Metabolic Change Shift	Mass Shift (Da)	Theoretical [M-H] <sup>-</sup>	Measured [M-H] <sup>-</sup>	$\Delta m$ (ppm)	RDBeq	$t_R$ (min)	Molecular Formula	Metabolic generation
Methylation	+ [CH <sub>2</sub> ]	14.0157	593.1876	593.1860	2	7	8.8	C <sub>28</sub> H <sub>34</sub> O <sub>14</sub>	Phase II
Hydroxylation-Methylation	+ [CH <sub>3</sub> O]	30.0106	609.1825	609.1805	3	7	7.1	C <sub>28</sub> H <sub>35</sub> O <sub>15</sub>	Phase II

**Table 1** Possible metabolites of naringin detected in chicken plasma samples after dietary supplementation with 1.5 g/kg of feed of naringin employing MetWorks™ 1.3 software

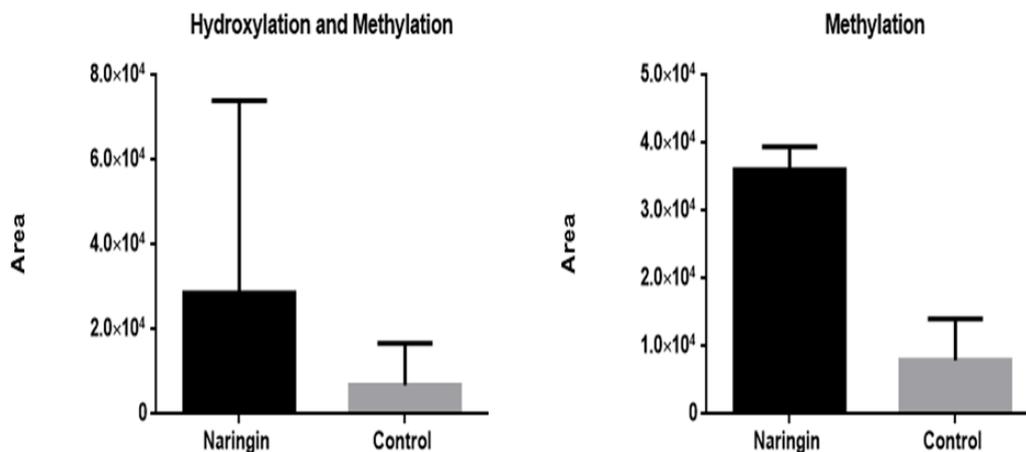


**Figure 1** The proposed chemical structure of naringin methylation according to the accurate mass, the RDB equivalents and the isotope pattern ratio using the online platform MetaPrint2



**Figure 2** The proposed chemical structure of naringin methylation and hydroxylation according to the accurate mass, the RDB equivalents and the isotope pattern ratio using the online platform MetaPrint2

A comparative study of the significant metabolites in the two groups (treated and untreated) is depicted in Figure 3.



**Figure 3** The significant metabolites hydroxylation-methylation and methylation in the control group and the group after naringin administration. The heights of the bars correspond to the area of the corresponding metabolites found in the different groups

These two metabolites modifications of naringin were also identified by Liu et al. after orally administration of naringin in rats and dogs at doses of 42 mg/kg and 12.4 mg/kg, respectively [12].

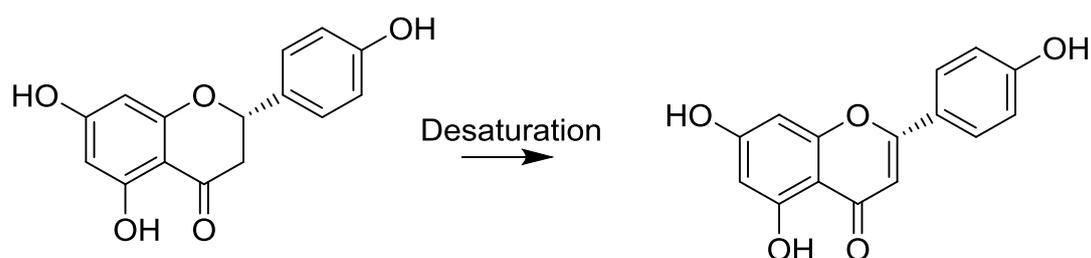
### 3.2 Naringenin metabolism in chicken plasma samples

Naringenin metabolites were investigated in chicken plasma samples. Two possible metabolites were identified after the data processing procedure. Table 2 summarizes the theoretical and measured mass values, the number of the RDBs, the retention time ( $t_R$ ), the predicted molecular formulas, the metabolic change shift and the metabolic generation for the possible metabolites.

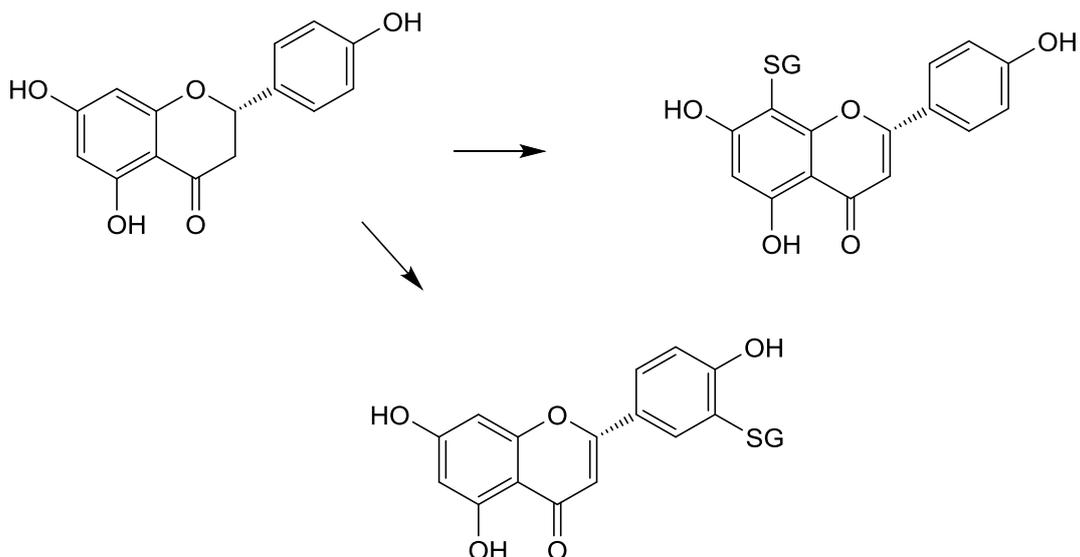
Metabolites	Metabolic Change Shift	Mass Shift (Da)	Theoretical [M-H] <sup>-</sup>	Measured [M-H] <sup>-</sup>	Δm (ppm)	RDBeq	t <sub>R</sub> (min)	Molecular Formula	Metabolic generation
Desaturation	-2H	2.0157	269.0455	269.0443	4	2	8.9	C <sub>15</sub> H <sub>10</sub> O <sub>5</sub>	Phase I
Glutathione conjugation	+GSH	307.0838	578.1450	578.1470	3	1	6.9	C <sub>25</sub> H <sub>29</sub> O <sub>11</sub> N <sub>3</sub> S	Phase II

**Table 2** Possible metabolites of naringenin detected in chicken plasma samples after dietary supplementation with 1.5 g/kg of feed of naringin employing MetWorks™ 1.3 software

Figures 4 and 5 demonstrate the proposed chemical structures employing the online platform MetaPrint2 and the MarvinSketch drawing software.

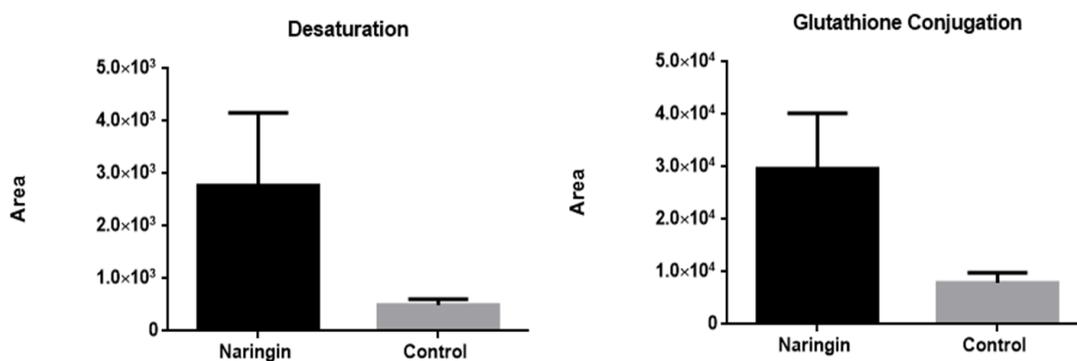


**Figure 4** The proposed chemical structure of naringenin desaturation employed MarvinSketch drawing software according to the accurate mass, the RDB equivalents and the isotope pattern ratio



**Figure 5** The proposed chemical structures of naringenin glutathione conjugation employed MarvinSketch drawing software according to the accurate mass, the RDB equivalents and the isotope pattern ratio.

Figure 6 presents the significant of the detected metabolites in the two groups (treated and untreated).



**Figure 6** The significant metabolites desaturation and glutathione conjugation in the control group and the group after naringenin administration. The heights of the bars correspond to the area of the corresponding metabolites found in the different groups

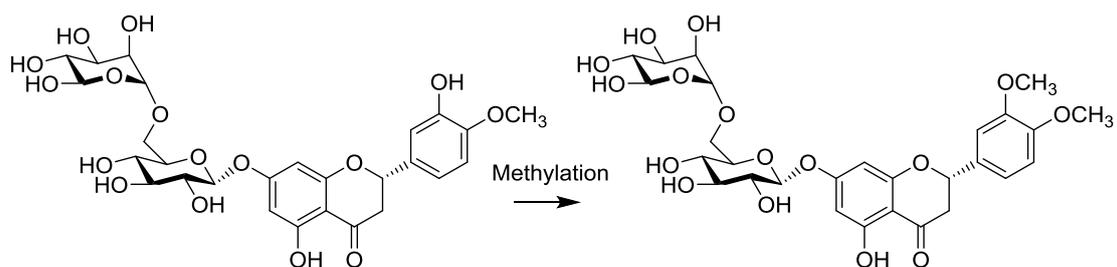
The desaturation modification of naringenin metabolism, apigenin, was also identified by Welford et al. after the incubation of anthocyanidin synthase (ANS) with the not naturally occurring substrate (+/-)-naringenin [19].

### 3.3 Hesperidin metabolism in chicken plasma samples

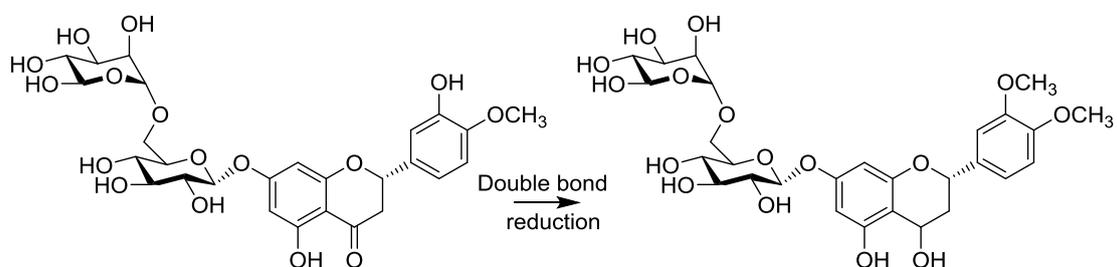
Hesperidin and its ninety-six possible metabolites modifications were investigated in chicken plasma samples after dietary supplementation with 1.5 g/kg of feed of hesperidin. Two possible metabolites of hesperidin were identified after the data processing procedure. Table 3 summarizes the theoretical and measured mass values, the number of the RDBs, the retention time ( $t_R$ ), the predicted molecular formulas, the metabolic change shift and the metabolic generation for the two possible metabolites. Their proposed chemical structures according to the accurate mass, the RDB equivalents and the isotope pattern ratio are presented in Figures 7 and 8.

Metabolites	Metabolic Change Shift	Mass Shift (Da)	Theoretical [M+H] <sup>+</sup>	Measured [M+H] <sup>+</sup>	$\Delta m$ (ppm)	RDDeq	$t_R$ (min)	Molecular Formula	Metabolic generation
Methylation	+ [CH <sub>2</sub> ]	14.0157	623.1981	623.2001	3	7	23.5	C <sub>29</sub> H <sub>36</sub> O <sub>15</sub>	Phase II
Double-bond reduction	CH=CH → CH <sub>2</sub> -CH <sub>2</sub>	2.0157	611.1981	611.2007	4	6	23.4	C <sub>28</sub> H <sub>36</sub> O <sub>15</sub>	Phase I

**Table 3** Possible metabolites of hesperidin detected in chicken plasma samples after dietary supplementation with 1.5 g/kg of feed of hesperidin employing MetWorks™ 1.3 software

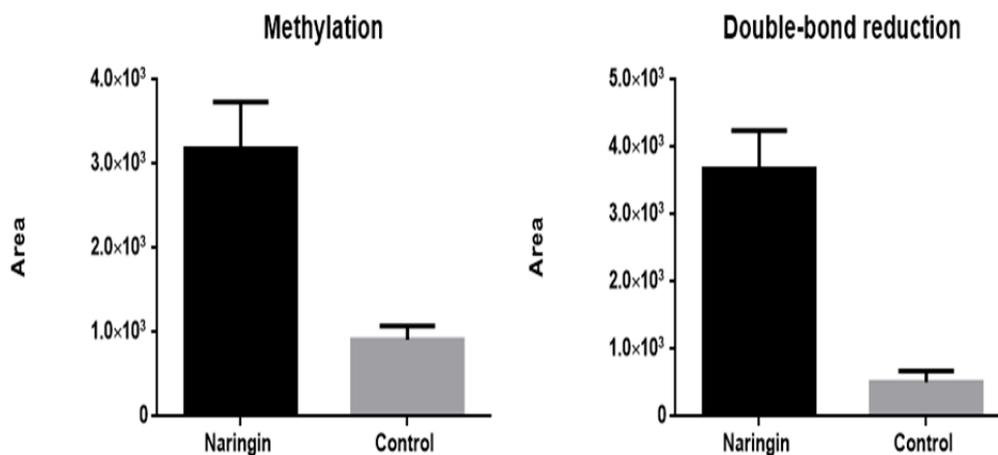


**Figure 7** The proposed chemical structure of hesperidin methylation according to the accurate mass, the RDB equivalents and the isotope pattern ratio using the online platform MetaPrint2



**Figure 8** The proposed chemical structure of hesperidin double-bond reduction according to the accurate mass, the RDB equivalents and the isotope pattern ratio using the online platform MetaPrint2

A comparative study of the significant metabolites in the two groups (treated and untreated) is depicted in Figure 9.



**Figure 9** The significant metabolites methylation and double-bond reduction in the control group and the group after hesperidin administration. The heights of the bars correspond to the area of the corresponding metabolites found in the different groups

The flavanone structure is highly reactive and have been reported to undergo O-methylation reactions [2].

### 3.4 Hesperetin metabolism in chicken plasma samples

The ninety-six possible metabolite modifications of hesperetin were investigated in chicken plasma samples after dietary supplementation with 1.5 g/kg of feed of hesperidin. Nevertheless, no metabolite modifications were detected according to the criteria of the accurate mass, the RDB equivalents and the isotope pattern ratio.

## 3. Concluding remarks

In this study the metabolism of naringin, hesperidin, naringenin and hesperetin was investigated in chicken plasma samples after 30 days of dietary supplementation with 1.5 g/kg of feed of naringin and with 1.5 g/kg of

feed of hesperidin. For the detection of the metabolite modifications the MetWorks<sup>TM</sup> software was employing. The setting criteria were the elemental composition, number of ring plus double bonds (RDBs) equivalents, isotope pattern and the peak width of the accurate  $m/z$  of the de-protonated molecule. For naringin two metabolite modifications, namely hydroxylation-methylation and methylation were detected. Furthermore, for its aglycone naringenin desaturation and glutathione conjugation were detected as possible metabolite modifications. For hesperidin, the metabolite modifications methylation and double bond reduction were detected. No metabolite modifications were detected for hesperetin.

## References

1. Ameer B, Weintraub RA, Johnson JV, Yost RA, Rouseff RL. (1996) Flavanone absorption after naringin, hesperidin, and citrus administration. *Clin Pharmac Ther.* 60, 34-40.
2. Booth AN, Jones FT, DeEds F. (1958) Metabolic fate of hesperidin, eriodictyol, homoeriodictyol, and diosmin. *J Biol Chem.* 230, 661-668.
3. Brand W, Boersma MG, Bik H, Hoek-van den Hil EF, Vervoort J, Barron D, Meinh W, Glatt H, Williamson G, van Bladeren PJ, Rietjens IM. (2010) Phase II metabolism of hesperetin by individual UDP-glucuronosyltransferases and sulfotransferases and rat and human tissue samples. *Drug Metab Dispos.* 38, 617-625.
4. Brand W, van der Wel PA, Rein MJ, Barron D, Williamson G, van Bladeren PJ, Rietjens IM. (2008) Metabolism and transport of the citrus flavonoid hesperetin in Caco-2 cell monolayers. *Drug Metab Dispos.* 36, 1794-1802.
5. Brett GM, Hollands W, Needs PW, Teucher B, Dainty JR, Davis BD, Brodbelt JS, Kroon PA. (2009) Absorption, metabolism and excretion of flavanones from single portions of orange fruit and juice and effects of anthropometric variables and contraceptive pill use on flavanone excretion. *Br J Nutr.* 101, 664-675.
6. Fuhr U, Kummert AL. (1995) The fate of naringin in humans: a key to grapefruit juice-drug interactions? *Clin Pharmacol Ther.* 58, 365–373.
7. Hackett AM, Marsh I, Barrow A, Griffiths LA. (1979) The biliary excretion of flavanones in the rat. *Xenobiotica.* 9:491-502
8. Hodek P., Trefil P., Stiborova M. Flavonoids-potent and versatile biologically active compounds interacting with cytochromes P450. *Chem Biol Interact.* 139, 1-21.

9. Ishii K, Furuta T, Kasuya Y. (1997) Determination of naringin and naringenin in human urine by high-performance liquid chromatography utilizing solid-phase extraction. *J Chromatogr B Biomed Sci.* 704,299–305.
10. Lee YS, Reidenberg MM. (1998) A method for measuring naringenin in biological fluids and its disposition from grapefruit juice by man. *Pharmacology.* 56,314–317.
11. Liu Z, Hu M. (2007). Natural polyphenol disposition via coupled metabolic pathways. *Expert Opin Drug Metab Toxicol.* 3, 389-406.
12. Liu M, Zou W, Yang C, Peng W, Su W. (2012). Metabolism and excretion studies of oral administered naringin, a putative antitussive, in rats and dogs. *Biopharm Drug Dispos.* 33, 123-134.
13. Manach C, Morand C, Gil-Izquierdo A, Bouteloup-Demange C, Remesy C. (2003) Bioavailability in humans of the flavanones hesperidin and narirutin after the ingestion of two doses of orange juice. *Eur J Clin Nutr.* 57, 235-242.
14. Mullen W., Archeveque M. A., Edwards C. A., Matsumoto H., Crozier A. (2008) Bioavailability and metabolism of orange juice flavanones in humans: impact of a full-fat yogurt. *J. Agric. Food Chem.* 56, 11157–11164.
15. Rice-Evans C. (2001). Flavonoid antioxidants. *Curr Med Chem.* 8, 797-807.
16. Roowi S., Mullen W., Edwards C. A., Crozier A. (2009) Yoghurt impacts on the excretion of phenolic acids derived from colonic breakdown of orange juice flavanones in humans. *Mol. Nutr. Food Res.* 53, S68– S75.
17. Schär MY, Curtis PJ, Hazim S, Ostertag LM, Kay CD, Potter JF, Cassidy A. (2015) Orange juice-derived flavanone and phenolic

metabolites do not acutely affect cardiovascular risk biomarkers: a randomized, placebo-controlled, crossover trial in men at moderate risk of cardiovascular disease. *Am J Clin Nutr.* 101, 931-938.

18. Wang M, Chao P, Hou Y, Hsiu S, Wen K, Tsai S. (2006) Pharmacokinetics and conjugation metabolism of naringin and naringenin in rats after single dose and multiple dose administrations. *J Food Drug Anal.* 14, 247–253.
19. Welford RW, Clifton IJ, Turnbull JJ, Wilson SC, Schofield CJ (2005). Structural and mechanistic studies on anthocyanidin synthase catalysed oxidation of flavanone substrates: the effect of C-2 stereochemistry on product selectivity and mechanism. *Org Biomol Chem.* 3, 3117-3126.
20. <http://www-metaprint2d.ch.cam.ac.uk/>

## Chapter abstract

Hesperidin and naringin are two citrus flavonoids which present a broad spectrum of biological activities. Naringin is mainly present in grapefruits and sour oranges, and hesperidin is present in lemons, oranges and other citrus fruits. The colon microflora has the ability to hydrolyze hesperidin and naringin to generate the corresponding aglycones called naringenin and hesperetin, respectively. The aim of the present study was to investigate the metabolism of naringin, hesperidin and their aglycones in plasma samples after their administration in chickens. Thirty 1-day-old Ross 308 broiler chickens were randomly divided into 3 groups. The experimental groups consisted of chickens given diets supplemented with 1.5 g/kg of feed of naringin (n=10), 1.5 g/kg of feed of hesperidin (n=10) and the control group that was given commercial basal diets (n=10). Blood samples were collected 30 days after the supplementation. The MetWorks<sup>TM</sup> 1.3 software was employed for the detection of potential metabolites of the analytes based on phase I and phase II patterns of metabolism. The setting criteria were the elemental composition, the number of ring plus double bonds (RDBs) equivalents, isotope pattern and the peak width of the accurate *m/z* of the de-protonated molecule. For naringin two metabolite modifications, hydroxylation-methylation and methylation were detected. Furthermore for its aglycone naringenin, desaturation and glutathione conjugation were detected. For hesperidin, the metabolite modifications methylation and double bond reduction were detected. No metabolite modification was observed for hesperetin. For the prediction of the possible sites of metabolism and the design of the chemical structures of the possible metabolites, the online platform MetaPrint2 and the MarvinSketch drawing software were employed.

**Keywords:** naringin-hesperidin-aglycones- plasma-metabolism

## Chapter 3

# **Development and validation of an UHPLC-ESI(-)- HRMS methodology for the simultaneous quantification of hesperidin, naringin and their aglycones in chicken tissue samples**

Due to the low concentration levels of the flavonoids in plasma, the study of their tissue distribution was deemed as compulsory. Therefore, a fully validated methodology for the simultaneous quantitation of hesperidin, naringin, hesperetin and naringenin was developed and applied in chicken tissue samples that were administrated with either hesperidin or naringin to their nutrition in order to assess their putative concentration levels.

## 1. Introduction

Citrus fruits are among the most commonly consumed fruits in Western countries, and are significant sources of flavanones belonging to the class of flavonoids. Flavonoids are a group of polyphenolic phytochemicals that act as antioxidants that sequester free radicals and reactive oxygen species [5, 2]. The two main flavanones are hesperidin in oranges (between 31 and 43.2 mg/100 g) and naringin in grapefruit (11.0–14.5 mg/100 g) [11, 7].

Hesperidin is a flavanone glycoside which is normally found in highly nutritious foods, such as oranges, tangelos, tangerines, grapefruits, and other citrus fruits. It presents a broad spectrum of biological activities like antioxidant, anti-inflammatory, anticarcinogenic and antiallergic which are well-known and promising [4]. It is metabolized in the small intestine to leave aglycone bioflavonoid hesperetin [17]. Takumi et al. has reported the tissue distribution of hesperetin after its dietary administration to rats for 4 weeks. According to his research the highest concentration level was found in the liver and the second highest in the aorta [16].

Naringin is a flavanone glycoside found in grapes and citrus fruits. It possesses the distinct bitter taste of grapefruit juice. The biological activities of naringin are related to its antioxidant nature. Moreover, naringin has been found to influence and modify several molecular signaling pathways in metabolic syndrome, obesity, and related cardiovascular complications. Thus, it also prevents the TNF- $\alpha$ -mediated inflammatory process and tissue damage in liver and vasculature [1]. Naringenin, the aglycone of naringin, has also been demonstrated to exhibit antioxidant [6] and anti-ulcer [12] effects. Previous studies report tissue distribution of naringin or naringenin after a single dose via intravenous or oral route [3, 19, 13] and after repeated dosing [8] in rat's model.

Nevertheless, to the best of our knowledge, so far there are no reports about simultaneous quantitative determination of hesperidin, naringin, hesperetin

and naringenin in chicken tissue samples after dietary supplementation with hesperidin and naringin.

After the investigation of the flavonoid levels in plasma and the exploration of their metabolism it became evident that only low levels of these substances were found in plasma. In order to explain their metabolic fate a new hypothesis was examined e.g. that they distribute to the tissues. Consequently, the aim of the current study was to develop a validated UHPLC-HRMS methodology for the determination of the 4 analytes in chicken tissue samples for pharmacokinetic studies. LC-MS/MS has been demonstrated to be the most valuable tool for pharmacokinetic studies because of its higher sensitivity and specificity compared with other analytical tools [10, 15, 18]. The bioanalytical methodology was validated in accordance with Food and Drug Administration (FDA) guidelines [14] and the EMA CHMP guidelines for bioanalytical methodologies [18] considering the specificity, linearity, recovery, matrix effect, repeatability precision, accuracy, and lower limit of quantification.

## **2. Materials and methods**

### **2.1 Chemicals and Reagents**

All solvents used were of LC-MS grade. Acetonitrile and water was purchased from Fluka/Riedel-de Haën (Switzerland). Glacial acetic acid, methanol, 4-iodophenol, chloroform and acetone were purchased from Sigma-Aldrich (Steinheim, German). Naringin was purchased from Alfa Aesar GmbH & Co KG (Germany) and Hesperidin from TSI Europe NV (Belgium).

### **2.2 Samples and study design**

Forty-nine 1-day-old Ross 308 broiler chickens were randomly divided into 3 groups. The lighting program consisted of 23L:1D on arrival, and was

decreased to 18L:6D at day 7, remained constant until day 35, and thereafter gradually increased to 23L:1D at slaughter, with access to feed, in mash form, and water *ad libitum*. The experimental groups consisted of 15 chickens given diets supplemented with 1.5 g/kg of feed of hesperidin and 17 chickens given diets supplemented with 1.5 g/kg of feed of naringin. The control group consisted of 17 chickens that were given commercial basal diets. The administration of naringin and hesperidin was started from the 11<sup>th</sup> day of age until slaughter at the age of 42 days. Chicken carcasses were then chilled at 4°C for 24h. Tissue samples were kept at -80°C until subsequent analysis.

All experimentation was carried out in strict accordance with the guidelines of “Council Directive 86/609/EEC regarding the protection of animals used for experimental and other scientific purposes”. The protocol was approved by the Bioethical Committee of the Agricultural University of Athens (Permit Number: 20/20032013).

### **2.3 Instrumentation**

An UPLC-ESI(-)-HRMS quantitation analysis was performed on an LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific, Bremen, Germany) connected to an Accela UHPLC system. The UHPLC system was equipped with an Accela quaternary pump, an Accela autosampler, a vacuum degasser and a temperature-controlled column compartment using the IT analyzer due to its enhanced sensitivity. The mass spectrometer was operated in the negative ion mode using the conditions as follows: sheath gas, 35 (arbitrary units); source voltage, 3 kV; auxiliary gas, 30 (arbitrary units); S lens RF level, 60 (%) and capillary temperature, 350°C.

The Orbitrap resolution was set at 30,000 FWHM and the isolation width was set at 2 amu. A reversed-phase INTERCHIM UPLC C18 column (1.7 µm particle size, 2.1 mm x 100 mm) at a flow rate of 0.3 mL/min was used for the chromatographic separation maintained throughout all experiments at 40°C.

The mobile phase consisted of: solvent A: aq. 0.1% glacial acetic acid, solvent B: acetonitrile and solvent C: isopropanol/acetonitrile/acetone (58/40/2, v/v). A gradient elution methodology was used as follows: 0 to 0.1 min: 95% A: 5% B: 0% C, 0.1 to 1.1 min: 80% A: 20% B: 0% C, 1.1 to 3.1 min: 70% A: 30% B: 0% C, 3.1 to 5.3 min: 50% A: 50% B: 0% C, 5.3 to 5.8 min: 0% A: 0% B: 100% C, 5.8 to 12 min: 95% A: 5% B: 0% C. The auto-sampler was kept at 4<sup>0</sup>C and the injection volume was 10 µL. The chromatographic data was acquired and processed using Xcalibur software (Version 2.1).

Centrifuging of the plasma samples during the sample preparation protocol was performed by a Mikro 200R centrifuge (Hettich Lab Technology, Germany). Evaporation of the samples was performed by a GeneVac HT-4X EZ-2 series evaporator Lyospeed ENABLED (Genevac Ltd, UK). Tissue homogenization was performed by a Kinematica Polytron PT 1200C homogenizer (Brinkmann, Westbury, NY).

#### **2.4 Preparation of standard solutions, calibration curves and quality controls samples**

Stock solutions of naringin (NN), naringenin (NG), hesperidin (HD), hesperetin (HT) and 4-iodophenol (IS) were prepared on methanol at 1 mg/mL and stored at 4<sup>0</sup>C. The stock solution of IS was daily diluted with the initial mobile phase to prepare the working solution (2 µg/mL) while stock solutions of all compounds were also appropriately diluted with the initial mobile. Thereafter, stock and diluted solutions of the compounds were appropriately mixed to afford six combined spiking solutions of all compounds with final concentrations of 0.125, 1.25, 5, 10, 17, and 25 µg/gr of tissue. Subsequently, 40 µL of each of these combined solutions were individually added to 0.020 gr of chicken tissue samples in order to construct the spike calibration curves.

#### **2.4.1 Preparation of pre-spike calibration curves**

In 0.020 gr of lyophilized chicken tissue samples, 100  $\mu\text{L}$  of the IS working solution (2  $\mu\text{g}/\text{mL}$ ) and 40  $\mu\text{L}$  of the calibration standards with concentrations from 0.125  $\mu\text{g}/\text{gr}$  to 25  $\mu\text{g}/\text{gr}$  of tissue, were added. Thus, a solution of 0.1M acetic acid-ammonium acetate (pH=5), were added to reach a final ratio of 1:10 tissue/solution (v/v) for the homogenization. The samples were then homogenized for at least 5 min. 1 mL of cold chloroform was added to the homogenizer mixture and the samples were kept at  $-20^{\circ}\text{C}$  for 10 min, in order to achieve total lipid extraction. After the removal of the lipid phase, 1 mL of cold acetone was added for the protein precipitation. Extraction was conducted by vortexing for 1 min. The extracts were centrifuged at 13,500 rpm at  $4^{\circ}\text{C}$  for 10 min. After phase separation, the aqueous layer (acetone) was collected and additional of 0.2 mL of cold acetone was added to the tissue pellet. Samples were vortexed and centrifuged as previously was reported. The supernatants were mixed and evaporated to dryness under vacuum. The residue was reconstituted with 0.2 mL of methanol/water (50/50, v/v) and centrifuged at 12,500 rpm for 8 minutes at  $4^{\circ}\text{C}$ . A 10  $\mu\text{L}$  aliquot was injected into the UHPLC-HRMS. Totally, 3 pre-spike calibration curves were constructed.

#### **2.4.2 Preparation of post-spike calibration curves**

In 0.020 gr of lyophilized chicken tissue samples, a solution of 0.1M acetic acid-ammonium acetate (pH=5), was added to reach a final ratio of 1:10 tissue/solution (v/v) for the homogenization. The samples were homogenized for at least 5 min. 1 mL of cold chloroform was added to the mixture and the samples were kept at  $-20^{\circ}\text{C}$  for 10 min. After the removal of the lipid phase, 1 mL of cold acetone was added for the protein precipitation. Extraction was conducted by vortexing for 1 min. The extracts were centrifuged at 13,500 rpm at  $4^{\circ}\text{C}$  for 10 min. After phase separation, the aqueous layer (acetone) was collected and additional of 0.2 mL of cold acetone was added to the tissue pellet. Samples were vortexed and centrifuged as previously was reported. The supernatants were mixed and 100  $\mu\text{L}$  of the IS working solution

(2 µg/mL) and 40 µL of the calibration standards were added, following the evaporation to dryness under vacuum. The residue was reconstituted with 0.2 mL of methanol/water (50/50, v/v) and centrifuged at 12,500 rpm for 8 minutes at 4<sup>0</sup>C. A 10 µL aliquot was injected into the UHPLC-HRMS. Totally, 3 post-spike calibration curves were constructed.

### **2.4.3 Preparation of quality control (QC) samples**

Different sets of solutions were prepared independently, in order to be used as QC samples, by serial dilution at four concentration levels of low (LQC:0.5 µg/g), medium (MQC:2.5 µg/g), high (HQC:25 µg/g), and lower limit of quantitation (LLOQ:0.125 µg/g). A total of six replicates were prepared for each QC sample.

In 0.020 gr of lyophilized tissue samples, 100 µl of the IS working solution (2 µg/mL) and 40 µL of the QC samples were added. Thus, a solution of 0.1M acetic acid-ammonium acetate (pH=5), was added to reach a final ratio of 1:10 tissue/solution (v/v) for the homogenization. The samples were homogenized by tissue homogenizer for at least 5 min. 1 mL of cold chloroform was added to the homogenizer mixture and the samples were kept at -20<sup>0</sup>C for 10 min. After the removal of the lipid phase, 1 mL of cold acetone was added for the protein precipitation. Extraction was conducted by vortexing for 1 min. The extracts were centrifuged at 13,500 rpm at 4<sup>0</sup>C for 10 min. After phase separation, the aqueous layer (acetone) was collected and additional of 0.2 mL of cold acetone was added to the tissue pellet. Samples were vortexed and centrifuged as previously was reported. The supernatants were mixed and evaporated to dryness under vacuum using a Lyospeed HT-4X GenVac. The residue was reconstituted with 0.2 mL of methanol/water (50/50, v/v) and centrifuged at 12,500 rpm for 8 minutes at 4<sup>0</sup>C. A 10 µL aliquot was injected into the UHPLC-HRMS.

#### **2.4.4 Preparation of standard calibration curve in solvent**

A standard calibration curve was constructed in methanol/water (50/50, v/v) using the following concentrations: 0.125, 1.25, 5, 10, 17, and 25 µg/gr of tissue.

#### **2.5 Sample preparation of unknown tissue samples**

0.020 gr of lyophilized tissue samples were used. Briefly, 100 µl of the IS working solution (2 µg/mL) and a solution of 0.1M acetic acid-ammonium acetate (pH=5), were added to reach a final ratio of 1:10 tissue/solution (v/v) for the homogenization. The samples were then homogenized for at least 5 min. 1 mL of cold chloroform was added to the homogenizer mixture and the samples were kept at -20<sup>0</sup>C for 10 min, in order to achieve total lipid extraction. After the removal of the lipid phase, 1 mL of cold acetone was added for the protein precipitation. Extraction was conducted by vortexing for 1 min. The extracts were centrifuged at 13,500 rpm at 4<sup>0</sup>C for 10 min. After phase separation, the aqueous layer (acetone) was collected and additional of 0.2 mL of cold acetone was added to the tissue pellet. Samples were vortexed and centrifuged as previously was reported. The supernatants were mixed and evaporated to dryness under vacuum. The residue was reconstituted with 0.2 mL of methanol/water (50/50, v/v) and centrifuged at 12,500 rpm for 8 minutes at 4<sup>0</sup>C. A 10 µL aliquot was injected into the UHPLC-HRMS.

#### **2.6. Bioanalytical method validation**

A full method validation was performed according to the FDA [1] and the EMA CHMP guidelines for bioanalytical methodologies [2] by evaluating the specificity, linearity, recovery, matrix effect, repeatability, precision, accuracy, and lower limit of quantification.

### **3. Results and discussion**

#### **3.1 UPLC-ESI(-)-HRMS optimization**

The mass spectrometric parameters were the same as described on Chapter 1 “Validated UHPLC-ESI(-)-HRMS methodology for the simultaneous quantitative determination of hesperidin, hesperetin, naringin and naringenin in chicken plasma”, resulting to enhance signal intensity for the four analytes.

#### **3.2 Method validation**

##### **3.2.1 Specificity, carry-over and linearity**

For the study of specificity, i.e. the ability to differentiate between target analytes and interference, six blank tissue samples were analyzed. The samples were prepared as previously has been described. It was indicated that no interference peaks could be detected in the expected retention time window for the four target analytes.

Carry-over was addressed by injecting, for a total of six times, blank samples after a high concentration sample. The absence of carry-over effect was shown as the injection of the blank samples showed no peaks at the corresponding retention time of the analytes.

Calibration curves ranging from 0.125 to 25 µg/g of tissue for the four analytes of interest were constructed and analyzed on three separate days. The linearity of the calibration curves was checked by regression analysis and the goodness of the regression by calculating the Pearson’s determination coefficient  $R^2$ . A strong correlation between the peak area ratio and concentration of each analyte in the tissue samples was illustrated by the high  $R^2$  of > 0.990.

### 3.2.2 Recovery, matrix effect

For the recovery study, 3 calibration curves from pre-spike samples and 3 curves from post-spike samples were analyzed and compared. Specifically, one curve was constructed from the calibration curves of pre-spike samples and one curve from post-spike samples for each analyte. The slopes of the two calibration curves were compared in order to assess the recovery. As shown in Table 1 the slope ratios were found to be greater than 0.94 for each analyte. This showed that the extraction protocol was suitable for the analysis of the four analytes in chicken tissue homogenates.

Analyte	Slope ratio
Naringin	0.94
Hesperidin	1.05
Naringenin	0.99
Hesperetin	1.04

**Table 1** Recovery of the methodology was determined after comparing the slope ratios of the calibration curve from pre-spike samples and the calibration curve from post-spike chicken tissues samples

Matrix effect arises due to effects of endogenous components of tissues on the ionization process of the analytes of interest. The matrix effect was investigated by comparing the slopes of the calibration curve prepared in methanol/water (50/50, v/v) and the calibration curve prepared by the post-spike samples. The slope ratios were found to be greater than 0.90 for each analyte and therefore the matrix effect could not adversely affect the accuracy and precision of the methodology (Table 2).

Analyte	Slope ratio
Naringin	0.99
Hesperidin	0.96

Naringenin	0.93
Hesperetin	0.95

**Table 2** Matrix effect was validated by comparing the slopes of the calibration curve prepared in methanol/water (50/50, v/v) and the calibration curve prepared by pre-spike chicken tissue samples

### 3.2.3 Lower limit of quantification

The lower limit of quantification (LLOQ) was defined as the lowest concentration at which both precision and accuracy were less than or equal to 20% in order an analyte in a sample to be quantified reliably. The analyte signal of the LLOQ sample should be at least 5 times the signal of a blank sample and to yield a peak with a signal/noise ratio of 10. The current assay offered an LOQ for NN, HD, NG and NN of 0.125 µg/g of tissue.

### 3.2.4 Repeatability, accuracy and intermediate precision

The repeatability (within-run), intermediate precision (between-run) and accuracy of the proposed methodology were examined by analyzing 6 replicates of 4 QC at concentrations levels 0.125, 0.5, 2.5, and 25 µg/g of tissue. The results for the intermediate precision and repeatability were expressed as the relative standard deviation (%RSD). The excluded criteria was set as %RSD<20 for the LLOQ and <15 for the QC samples (Tables 3, 4).

Analyte	LLOQ (0.125)	LQC (0.5)	MQC (2.5)	HQC (25)
Hesperidin	16.00%	11.32%	10.35%	5.81%
Naringin	19.68%	12.95%	15.06%	4.99%

<b>Hesperetin</b>	4.85%	3.48%	6.88%	4.73%
<b>Naringenin</b>	13.78%	7.29%	2.61%	4.13%

**Table 3** Validation results from the intermediate precision of the proposed methodology after the analysis of 6 replicates of 4 QC levels at nominal concentrations of 0.125, 0.5, 2.5, and 25 µg/g of tissue. The results were expressed as the relative standard deviation (%RSD). The excluded criteria was set as %RSD<20 for the LLOQ and <15 for the QC samples

<b>Analyte</b>	<b>LLOQ (0.125)</b>	<b>LQC (0.5)</b>	<b>MQC (2.5)</b>	<b>HQC (25)</b>
<b>Hesperidin</b>	11.52%	11.01%	6.18%	7.20%
<b>Naringin</b>	17.90%	9.19%	7.22%	7.11%
<b>Hesperetin</b>	16.20%	10.52%	9.71%	8.73%
<b>Naringenin</b>	17.05%	15.31%	12.64%	9.78%

**Table 4** Validation results from the repeatability of the proposed methodology after the analysis of 6 replicates of 4 QC levels at nominal concentrations of 0.125, 0.5, 2.5, and 25 µg/g of tissue. The results were expressed as the relative standard deviation (%RSD). The excluded criteria was set as %RSD<20 for the LLOQ and <15 for the QC samples

Accuracy was expressed as the %standard error (%Er) between the mean concentration and the nominal concentration. The excluded criteria was set as %Er<20 for the LLOQ and <15 for the QC samples (Table 5).

Analyte	LLOQ (0.125)	LQC (0.5)	MQC (2.5)	HQC (25)
Hesperidin	11.78%	6.30%	4.09%	4.33%
Naringin	12.60%	12.21%	1.45%	2.09%
Hesperetin	1.45%	15.76%	13.87%	14.68%
Naringenin	14.32%	14.44%	14.50%	15.77%

**Table 5** Validation results from the accuracy of the proposed methodology after the analysis of 6 replicates of 4 QC levels at nominal concentrations of 0.125, 0.5, 2.5, and 25 µg/g of tissue. The results were expressed as the %standard error (%Er) between the mean concentration and the nominal concentration. The excluded criteria was set as %Er<20 for the LLOQ and <15 for the QC samples

The results showed the repeatability did not exhibit values more than 17.9% for the LLOQ and 15.31% for the QC samples whereas the precision did not exhibit values more than 19.68% for the LLOQ and 15.06% for the QC samples. The %Er of the accuracy did not exhibit values more than 14.32% for the LLOQ and 15.77% for the QC samples. These results demonstrate that the method was accurate and precise for the quantitation of the 4 analytes in chicken tissue samples.

### 3.3 Application of the validated UPLC-ESI(-)-HRMS methodology on a pharmacokinetic study

The useful methodology was applied on tissue samples from chickens after dietary supplementation with hesperidin and naringin. The experimental groups consisted of 15 chickens given diets supplemented with 1.5 g/kg of feed of hesperidin and 17 chickens given diets supplemented with 1.5 g/kg of feed of naringin. The control group, consisted of 17 chickens, was given

commercial basal diets. Feed additives were supplemented from 11<sup>th</sup> day until slaughter at the age of 42 days. The quantitation was based on standard curves that were prepared for hesperidin, naringin, hesperetin and naringenin on the day of the analysis, ranging from 0.125 to 25 µg/g of tissue. Nevertheless, no detectable amounts of the analytes hesperidin, naringin and their aglycone have been discovered in the chicken tissue samples above the LLOQ.

#### **4. Concluding remarks**

A highly sensitive and specific methodology was developed for the simultaneous determination of hesperidin, hesperetin, naringin and naringenin in chicken tissues samples employing UHPLC-HRMS (Orbitrap Velos). Protein precipitation of tissue samples was performed using cold acetone and lipid extraction using cold chloroform. Detection was performed by means of electrospray ionization (ESI) in the negative ion mode. The calibration curves for the 4 analytes exhibited good linearity ( $r^2 > 0.990$ ) over the concentration range from 0.125 µg/g of tissue to 25 µg/g of tissue. Lower limit of quantification (LLOQ) for the 4 analytes was 0.125 µg/g of tissue. The repeatability and precision were within %RSD<20 and accuracy within %Error<20. No matrix effect and carry over was observed on the proposed methodology. The validated methodology was applied for the quantitative determination of hesperidin, naringin and their aglycones in tissue samples after dietary supplementation with 1.5 g/kg of feed of hesperidin and 1.5 g/kg of feed of naringin in Ross 308 broiler chickens. However, after the administration on chickens, no detectable amounts of the analytes above the LLOQ seem to be distributed to tissues. Therefore in order to investigate any possible effects of the analytes on tissues despite the lack of detectable levels, a metabolomics approach was performed, which is discussed on Chapter 6.

## References

1. Alam MA, Subhan N, Rahman MM, Uddin SJ, Reza HM, Sarker SD. (2014) Effect of citrus flavonoids, naringin and naringenin, on metabolic syndrome and their mechanisms of action. *Adv Nutr.* 5, 404-417.
2. Duthie G, Crozier A. (2000) Plant-derived phenolic antioxidants. *Curr Opin Lipidol.* 11, 43-47.
3. El Mohsen MA, Marks J, Kuhnle G, Rice-Evans C, Moore K, Gibson G. (2004) The differential tissue distribution of the citrus flavanone naringenin following gastric instillation. *Free Radic Res.* 38, 1329–1340.
4. Garg A., Garg S., Zaneveld L.J.D., Singla A.K. (2001) Chemistry and pharmacology of the Citrus bioflavonoid hesperidin, *Phytother. Res.* 15, 655–669.
5. Hou M, Man M, Man W, Zhu W, Hupe M, Park K, Crumrine D, Elias PM, Man MQ. (2012) Topical hesperidin improves epidermal permeability barrier function and epidermal differentiation in normal murine skin. *Exp Dermatol.* 21, 337-340.
6. Jeon S.M., Park Y.B., Choi M.S. (2004) Antihypercholesterolemic property of naringin alters plasma and tissue lipids, cholesterol-regulating enzymes, fecal sterol and tissue morphologh in rabbits. *Clinical Nutrition.* 23, 1025–1034.
7. Kyle JAM, Duthie GG. (2006) *Flavonoids in foods.* Boca Raton: CRC Press. 319–370.
8. Lin SP, Hou YC, Tsai SY, Wang MJ, Chao PD. (2014) Tissue distribution of naringenin conjugated metabolites following repeated dosing of naringin to rats. *Biomedicine (Taipei).* 4, 16.
9. Liu AC, Zhao LX, Xing J, Gao J, Lou HX.(2013) LC–MS/MS method for the determination of a new puerarin derivative and its application in pharmacokinetic studies in rats. *Chin J Nat Med.*11, 566–571.

10. Liu QW, Wang JS, Yang L, Jia YW, Kong LY. (2013) A rapid and sensitive LC–MS/MS assay for the determination of berbamine in rat plasma with application to preclinical pharmacokinetic study. *J Chromatogr B*. 929, 70–75.
11. Manach C, Scalbert A, Morand C, Remesy C, Jimenez L. (2004) Polyphenols: food sources and bioavailability. *Am J Clin Nutr*. 79, 727–747.
12. Martin M.J., Marhuenda E., Perez-Guerrero C., France J.M. (1994) Antiulcer effect of naringin on gastric lesions induced by ethanol in rats. *Pharmacology*. 49, 144–150.
13. Peng HW, Cheng FC, Huang YT, Chen CF, Tsai TH. (1998) Determination of naringenin and its glucuronide conjugate in rat plasma and brain tissue by high-performance liquid chromatography. *J Chromatogr B Biomed Sci Appl*. 714, 369–374.
14. Shah V.P. (2000) *Guidance for Industry. Bioanalytical Method Validation. Pharmaceutical Research*. 17.
15. Song M, Lee D, Lee T, Lee S. (2013) Determination of leelamine in mouse plasma by LC–MS/MS and its pharmacokinetics. *J Chromatogr B*. 931, 170–173.
16. Takumi H., Mukai R., Ishiduka S., Kometani T., Terao J. (2011) Tissue distribution of hesperetin in rats after a dietary intake. *Biosci Biotechnol Biochem*. 75, 1608-1610.
17. Watanabe M, Matsumoto N, Takeba Y, Kumai T, Tanaka M, Tatsunami S, Takenoshita-Nakaya S, Harimoto Y, Kinoshita Y, Kobayashi S. (2011) Orange juice and its component, hesperidin, decrease the expression of multidrug resistance-associated protein 2 in rat small intestine and liver. *J Biomed Biotechnol*. 2011, 502057.

18. Wharf C. and Kingdom U. (2012) Guideline on bioanalytical method validation. Table of contents. 44.
19. Zou W, Yang C, Liu M, Su W. (2012) Tissue distribution study of naringin in rats by liquid chromatography-tandem mass spectrometry. *Arzneimittelforschung*. 62, 181–186.

## Chapter abstract

A sensitive and specific methodology for the simultaneous quantitative determination of hesperidin, naringin, hesperetin and naringenin in chicken tissue samples employing UHPLC-HRMS (Orbitrap Velos) was developed. Sample pretreatment employed protein precipitation with cold acetone, preceded by a defating step with cold chloroform. Chromatography was performed on an INTERCHIM UHPLC C18 column using water, acetonitrile and isopropanol/acetonitrile/acetone (58/40/2, v/v) as mobile phase. Detection was performed by means of electrospray ionization (ESI) in the negative ion mode. Calibration curves for the 4 analytes exhibited good linearity ( $r^2 > 0.990$ ) over the concentration range from 0.125  $\mu\text{g/g}$  of tissue to 25  $\mu\text{g/g}$  of tissue. The lower limit of quantification (LLOQ) for the 4 analytes was 0.125  $\mu\text{g/g}$  of tissue. The repeatability and precision were within %RSD<20 and accuracy within %Error<20. No matrix effect and carry over was observed on the proposed methodology. The validated methodology was applied for the quantitative determination of hesperidin, naringin, hesperetin and naringenin in tissue samples after dietary supplementation with 1.5 g/kg of feed of hesperidin and 1.5 g/kg of feed of naringin in Ross 308 broiler chickens.

**Keywords:** naringin-hesperidin-validation-pharmacokinetics-tissue

## Chapter 4

### **An alternative approach for data processing in untargeted metabolomics by UHPLC-ESI(-)- HRMS. The use of post acquisition spectral stitching**

A fundamental problem concerning the metabolomics methodologies is the large number of false positives which frequently leads to non-conclusive statistical evaluation. Therefore a new strategy was developed in the course of the data processing pipeline of MS based metabolomics. This strategy highlights the number of the reliable metabolite identifications, unraveling putative metabolites that would be overssen by the classical methodologies.

## 1. Introduction

Metabolomics is one of the most recent “omics” sciences, defined as the discipline that measures ideally all the metabolite levels of a biological system, affording the chemical fingerprint of the metabolism. Metabolites are low molecular weight compounds typically <1500 Da possessing a wide range of physicochemical characteristics [4]. Metabolomic analyses can be classified either as targeted or untargeted. In the first mode, several *a priori* selected metabolites (often referred as known knowns) are analyzed, whereas in the untargeted approach, hundreds or even thousands of new molecules are detected and the differentially expressed ones are identified, aiming to reveal the metabolic pathways that are primarily affected [11,13]. Some of these metabolites might be characterized as biomarkers and can be used in diagnosis of a disease, to reveal a specific biochemical pattern etc [8, 5].

Metabolomic analysis is performed by various techniques with the aim of detecting as many features as possible. The two most commonly used techniques are nuclear magnetic resonance (NMR) and mass spectrometry coupled to liquid chromatography (LC-MS) [12, 19]. An equally important role besides the instrumental analysis is devoted to the assessment of the results, employing a large number of software based data analysis procedures. Especially for the case of MS based metabolomics, a vast array of software tools encompassing a wide range of algorithms are used for the primary processing in the biomarker discovery pipeline, such as apLCMS, Mzmine, XCMS, MAIT, XCMS online, MetAlign etc [21, 14, 15, 3, 18, 10], which enable feature detection in the extremely complex datasets produced by LC-MS analysis.

However, a fundamental problem concerning these methodologies is the fact that they produce a large number of false positives, mainly due to chemical noise, which frequently leads to non-conclusive statistical evaluation. This often occurs due to the high sensitivity of the technique, which efficiently detects the chemical background giving rise to spurious signals without

conclusive biochemical/biological meaning associated with the investigated issue. Downstream statistical analysis will in turn use this redundant information and afford models with low predictive ability, suffering also from the lack of consistency and rigid biological interpretation. As an example Tautenhahn [17] describing the centwave algorithm, showed that in order to efficiently pick peak 13 substances along with all of their possible adducts, they had to integrate a three-fold amount of peaks, leading to the generation of a vast number of artifacts, a fact that could potentially lead to erroneous statistical treatment.

The aim of the study was to tackle this problem developing a new strategy in the data processing pipeline of MS based metabolomics. This strategy aims to increase the number of chromatographically and spectroscopically reliable features (CSFR's) that are effectively detected, thus increasing the corresponding number of putative true positives. These features are obtained after the data processing in multiple narrow mass ranges instead of the processing over a wide mass range. This is of considerable importance in metabolomics, due to the fact that the discovery and identification of unknown significant features remains one of the most substantial analytical challenges.

## **2. Materials and methods**

### **2.1 Chemicals and Reagents**

All solvents were of LC-MS grade. Acetonitrile and water were purchased from Fluka/Riedel-de Haën (Switzerland). Glacial acetic acid, methanol and acetone were purchased from Sigma-Aldrich (Steinheim, Germany). Naringin was purchased from Alfa Aesar GmbH & Co KG (Germany).

## 2.2 Animal treatment

Twenty nine 1-day-old Ross 308 broiler chickens were randomly divided into 2 groups. The chickens were obtained from a commercial hatchery and were reared in pens, of a surface area of 2 m<sup>2</sup> in a controlled environment. The lighting program consisted of 23L:1D on arrival, and was decreased to 18L:6D at day 7, remained constant until day 35, and thereafter gradually increased to 23L:1D at slaughter, with access to feed, in mash form, and water *ad libitum*. The treatment groups were: 20 chickens, which were administrated with naringin (1.5 g/kg) in their nutrition and 9 control chickens that were given commercial basal diets. The administration of naringin was started from the 11<sup>th</sup> day of age until slaughter at the age of 42 days.

All experimentation was carried out in strict accordance with the guidelines of “Council Directive 86/609/EEC regarding the protection of animals used for experimental and other scientific purposes”. The protocol was approved by the Bioethical Committee of the Agricultural University of Athens (Permit Number: 20/20032013).

## 2.3 Sample treatment

Blood samples were collected on 42 days of age after 30 days of the administration into heparinized tubes and centrifuged at 2,500 rpm for 10 minutes to recover plasma. Plasma samples were kept at -80°C until the analysis.

For the protein precipitation of plasma proteins, cold acetone was added to plasma, reaching a final ratio of 1/8 plasma/acetone (v/v). Extraction was conducted by vortexing for 1 min and subsequently the mixture was left for 10 minutes at -20°C in order to achieve total precipitation. The extracts were then centrifuged at 12,500 rpm for 10 minutes at 4°C and the supernatant was collected and evaporated to dryness under vacuum using a Lyospeed HT-4X GenVac. The residue was reconstituted with 50 µL of methanol/water (50/50,

v/v) and centrifuged at 12,500 rpm for 8 minutes at 4<sup>0</sup>C. A 2 µL aliquot was injected into the UHPLC-HRMS.

## 2.4 Instrumentation

An ESI-LTQ-Orbitrap Velos (Thermo Scientific, Bremen, Germany) connected to an Accela UHPLC system controlled by the Xcalibur 2.1 software was used for the metabolomic analysis. The UHPLC system consisted of an Accela quaternary pump, an Accela autosampler, a vacuum degasser and a temperature-controlled column compartment. Operational parameters of the mass spectrometer were set as follows: sheath gas, 35 (arbitrary units); source voltage, 3 kV; auxiliary gas, 30 (arbitrary units); S lens RF level, 60 (%) and capillary temperature, 350<sup>0</sup>C in the negative ion mode. The Orbitrap resolution was set at 30,000 FWHM with a scan range of 100 to 900 (*m/z*) in the centroid mode. Plasma samples were injected onto a reversed-phase INTERCHIM UHPLC C18 column (1.7 µm particle size, 2.1 mm x 100 mm) maintained at 40<sup>0</sup>C at a flow rate of 0.36 mL/min. The mobile phase consisted of aq. 0.1% glacial acetic acid (A) and acetonitrile (B). A gradient methodology was employed as follows: 0 to 24 min: 95% A: 5% B, 24 to 28 min: 5% A: 95% B, 28 to 32 min: 95% A: 5% B.

Plasma samples centrifugation was performed using a Mikro 200R centrifuge (Hettich Lab Technology, Germany). Sample evaporation was performed by a GeneVac HT-4X EZ-2 series evaporator Lyospeed ENABLED (Genevac Ltd, UK).

## 2.5 Data Analysis

After the analysis of the plasma samples on the Orbitrap Velos, the raw files were imported to the open source software package Mzmine 2.18.1 (<http://mzmine.sourceforge.net/>) for baseline correction, peak detection,

alignment, deconvolution and retention time normalization. The noise amplitude algorithm has been employed for the peak detection, as it estimates the noise for each chromatogram separately, according to a selected segment with the highest number of data points which corresponds to the noise. The generated peak-lists were exported as comma-separated values (CSV) files to Excel 2007 (Microsoft Office 2007) and manipulated using the concatenate, round and transpose commands. The MetaboAnalyst web platform (<http://www.metaboanalyst.ca/MetaboAnalyst/>) was used for initial data processing. Normalization was performed using a reference control sample whereas missing values estimation was achieved by removing the features with at least 50% of missing values and replacing the remaining ones with the half of the minimum positive value of the original data. The Simca-P+11.5 software (Umetrics, Umeå, Sweden) was employed for multivariate analysis (MVA), i.e. principal component analysis (PCA) and projections to latent structures – discriminant analysis (PLS-DA) with the treated vs untreated data being the discriminant, in order to detect the metabolic differences. According to the structure of each dataset, different normalization methodologies have been employed i.e Pareto or UV, to the mean centered data. The unpaired t-test employing the uneven variance approach was performed between the control group and the group which was administrated with naringin for all features detected. The most important metabolites discovered by PLS-DA, were ranked by the variable importance in projection (VIP) scores. Finally in order to evaluate the possibility of random classification during PLS-DA analysis, permutation testing using 100 random permutations was applied.

### **2.5.1 Criteria for the evaluation of the detected features**

The features retained, referred hereafter as chromatographically and spectroscopically reliable features (CSFR's), should encompass three criteria i.e. a peak width of at least seven scans, mass tolerance lower than 5 ppm and signal-to-noise (S/N) ratio of at least 10. Given the fact that a 30,000

resolution scan in the Orbitrap Velos has in average scan time of 1 sec, the resulting peak width has an approximate value of 7 sec, which is the expected peak width in UHPLC chromatography. The use of 12 or 16 data points as recommended to quantitative analysis would result in unacceptable wide peaks, which would in turn exclude a substantial number of chromatographically narrower peaks from downstream statistical analysis.

### **3. Results and discussion**

A crucial issue in metabolomics is the large number of false positives, i.e. features that do not correspond to chromatographic peaks but are rather artifacts resulting from the computational procedure. This means that these features possess the characteristics attributed to an LC/MS feature in terms of e.g. S/N ratio, peak width or m/z tolerance etc. however, either the peak shape does not correspond to a chromatographic peak or it corresponds to a poorly chromatogram substance which can't be reliably integrated and quantified. Usually in the majority of cases these criteria are set in rather loose way, in an effort to include as many peaks as possible thus keeping the informational content as rich as possible

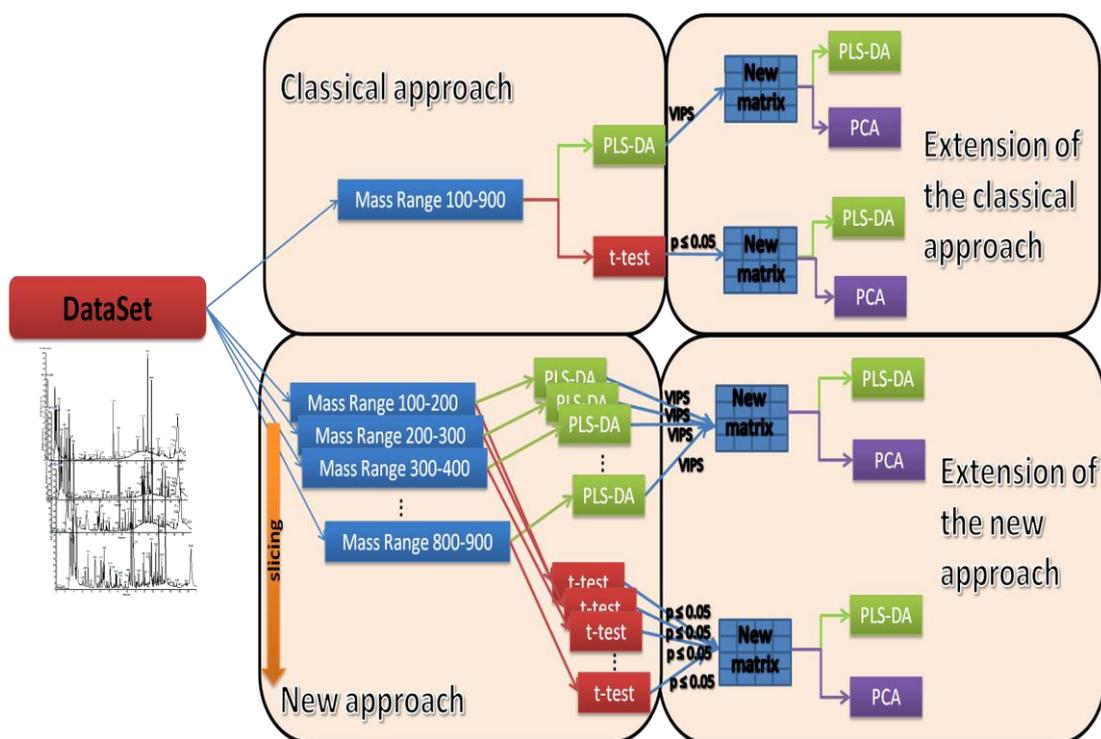
The frequency distribution of organic substances is shifted towards the low mass range with a maximum at approximately 400 Da [6]. This observation denotes a natural discrimination of masses for the low mass range. Furthermore, due to the inherent trend of mass spectrometers to induce fragmentation, high molecular weight substances survive in a lesser extent than their low molecular weight counterparts. As Tautenhahn [17] has reported, from a mixture of 14 compounds, 296 features are theoretically produced. Nevertheless, out of the 296 features only 122 were detected as fragments of the original compounds, adducts and their isotopic peaks. Moreover, it has been mentioned that in order to detect more than 100 out of the 122 features, 450 false positive features produced by solvents, background noise e.g. should additionally be detected.

Yet another problem is the selectivity of the Orbitrap mass analyzer for the most abundant ions, which as stressed out before, carry a higher probability to lie within low mass range. In other words, due to the automatic gain control (AGC) embedded to the instrument's controlling algorithm the LTQ trap fills with a finite number of  $m/z$  features and in cases of highly abundant ions those are the primarily trapped ones. In contrast, the quadrupole analyzer spends equal time to the less abundant ions and thus the sensitivity against them is increased.

All the features that are detected from an LC-MS run (true and false positives) are in turn used for MVA, resulting either in the loss of valuable information or to the production of erroneous results. Therefore the possibility of incorrect classification increases by the number of the detected features, especially in the case of supervised methodologies such as PLS-DA, where random classifications may occur, as pointed out by Westerhuis [20]. In other words as Westerhuis stated, "PLS-DA is eager to please", i.e. the larger the number of features detected, the larger the chance to discover some that lead to the correct classification despite the fact that may be artifacts or noise. Thus using Orbitrap based mass analysis the driving force for classification is the low molecular weight substances, which could usually be chemical noise due either to the presence of contaminants or non-controlled fragmentation imposed by the hardware. Therefore a significant challenge in metabolomics is to develop computational methodologies and processing pipelines that increase the number of the CSRF's, especially considering large-scale approaches.

In order to confront these problems a novel approach has been adopted. The new approach (slicing) has been compared to the existing metabolomics data processing pipeline, which will refer hereafter as the classical approach. In the latter, all the data of the acquisition mass range (i.e in the  $m/z$  100-900 for the current set of experiments) were subjected to the data processing procedure. The classical methodology has been extended by the construction of two new matrices, consisted of the VIP values resulting from the PLS-DA analysis or

the t-test values from the univariate comparison of the treated vs untreated groups. The resulting matrices have been further evaluated by PCA and PLS-DA (Scheme 1) and the detected features were validated using the criteria described in Section 2.5.1. According to our slicing approach, the dataset was shred in 100 Da slices, generating 8 datasets. These were treated using two approaches (similar to the extension of the classical approach) (Scheme 1 – New approach Panel). In the first one 8 PLS-DA models were developed – one for each slice. The variable ID's, used as their accurate m/z value-retention time ordered pairs (m/z\_tR) with the highest scoring VIP values (greater than 1) [1, 2] resulting from each model, were merged together in a reduced matrix. This new matrix has been subjected to MVA (PCA and PLS-DA). Instead of using the VIP values, the ID's that formed the reduced matrix, were determined by performing t-tests for all the detected features of each slice and selecting the statistically different ones (p-values  $\leq 0.05$ ). MVA i.e. PCA and PLS-DA has been employed in order to discriminate between the classes. The detected features were evaluated using the criteria described in Section 2.5.1.

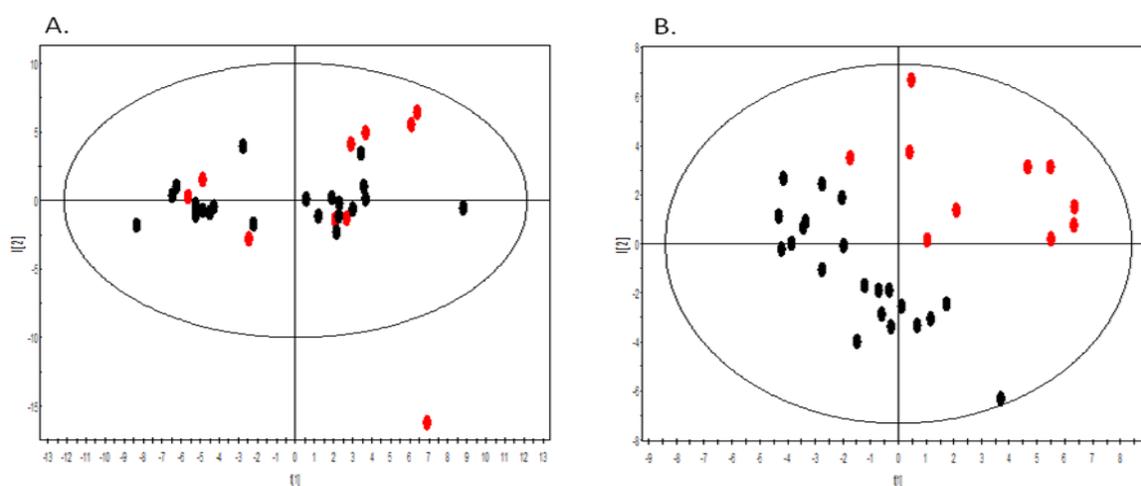


**Scheme 1** Two methodologies have been applied for data processing. In the first one (classical approach), all data in the 100-900  $m/z$  were included in the data processing procedure. The classical methodology has been extended by the construction of two new matrices, consisted either from the VIP from the PLS-DA analysis or the t-test values from the univariate comparison of the treated vs untreated group. The resulting matrices have been further evaluated by PCA and PLS-DA analysis. To the newly developed methodology, the data were shredded in 100 Da slices generating 8 datasets, which has been subjected to the downstream MS data processing. Each dataset was treated as separate and  $m/z_{tR}$  features obtained by either the VIP's or the t-test values were used as input for the construction of the general model. This new matrices has been subjected to MVA

### 3.1 Processing of raw data at the mass range of $m/z$ 100-900 (classical approach)

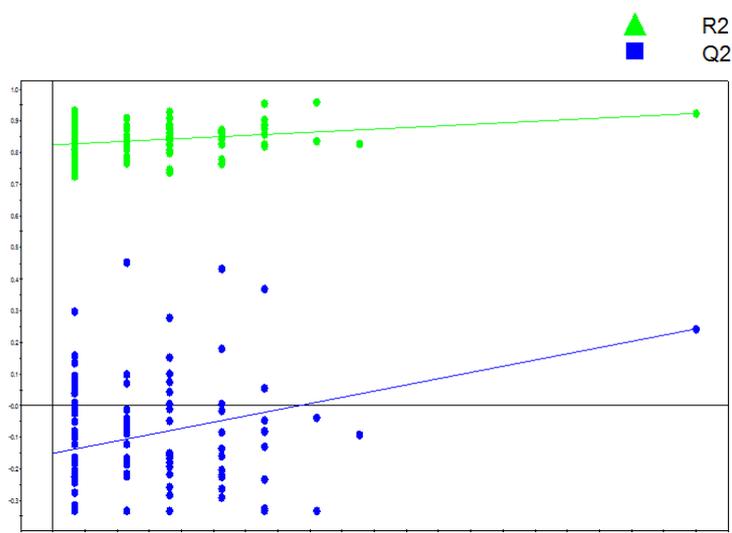
Initially the data were processed in the mass range of  $m/z$  100-900 (Supplementary Table S1). The generated peaklist, containing the detected features as their accurate  $m/z$  value-retention time ordered pairs and the

corresponding peak areas for each sample was exported as a comma-separated values (CSV) file to an Excel 2007 data sheet. After using the concatenate, round and transpose commands, the file was uploaded to the Metaboanalyst web platform for missing values estimation and normalization. Finally the resulting file was imported to Simca-P+11.5. MVA has been performed employing PCA and PLS-DA. Various parameters have been used, such as different scaling e.g. UV or Pareto producing combinations of  $Q^2$  and  $R^2$  in order to assess as much variance as possible without excessively increasing the complexity of the model. The UV scaling method has been finally selected as it afforded tighter groups with larger separation. No clustering between the two groups was observed using the PCA analysis (Figure 1A). However, employing PLS-DA, a clear separation between the control and treated group was observed (Figure 1B).



**Figure 1** Score plots from PCA (A) and PLS-DA (B) after processing at the  $m/z$  100-900 mass range of plasma samples from chickens that were administrated with naringin (black dots) and control samples (red dots) employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. A clear separation between the samples of the two groups was observed only for PLS-DA. The  $R^2$ ,  $Q^2$  values were for the PCA methodology 31.1% and 0.0581 respectively, and for the PLS-DA 29.4%, 92.4% and 0.242 respectively, employing 2 principal components for PCA and 3 for PLS-DA analysis

In order to explore the possibility that the PLS-DA model resulted by chance, permutation testing using 100 random permutations was performed. It had been demonstrated that few of the 100 permuted models presented better fit ( $R^2$ ) and predictive capability ( $Q^2$ ) than those of the original model, denoting that the obtained model was not reliable (Figure 2).

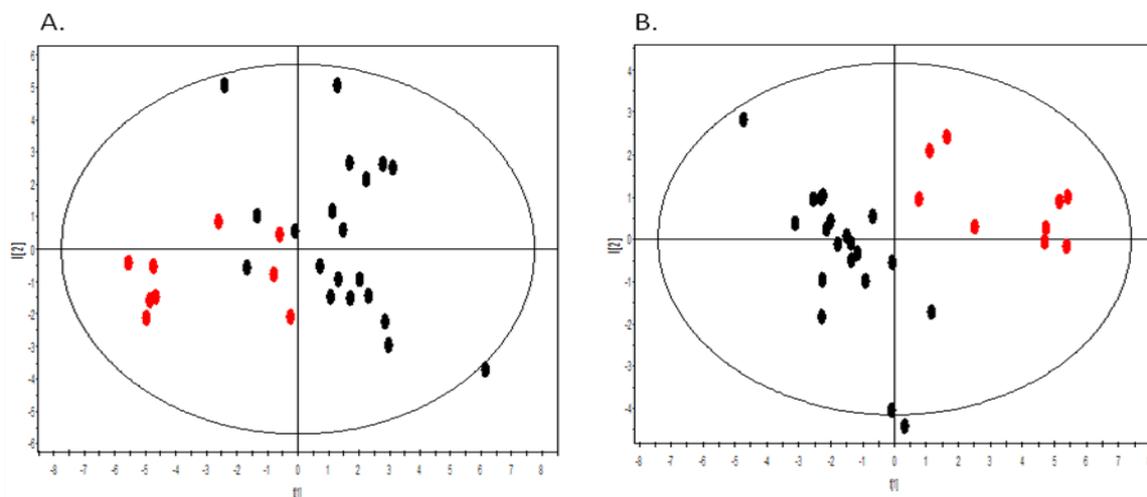


**Figure 2** Permutation testing allowing of 100 permutations of the PLS-DA model obtained after processing at the  $m/z$  100-900 mass range of plasma samples from chickens that were administrated with naringin and control samples employing UHPLC-ESI-HRMS (Orbitrap) analysis, in the negative ion mode. Few of the 100 permuted models presented better fit ( $R^2$ ) and predictive capability ( $Q^2$ ) than those of the original model

### 3.1.1 Extension of the classical approach. MVA analysis of the $m/z$ 100-900 mass range using the unpaired t-test values

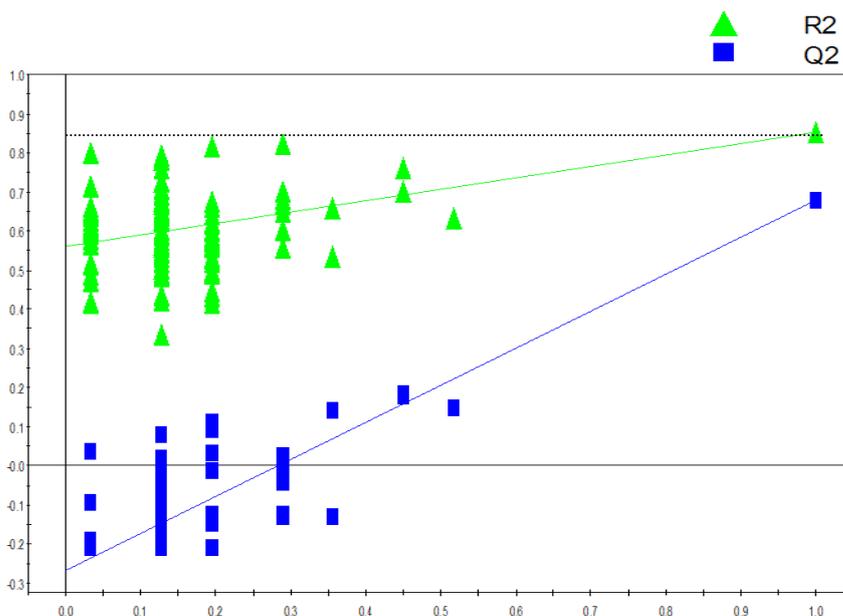
In order to exclude features resulting by the noise in the range of the  $m/z$  100-900 span and to improve the fitness ( $R^2$ ) and the predictive capability ( $Q^2$ ) of the model, an unpaired t-test was initially performed, aiming to filter out the statistically insignificant features between the two groups (p-values of  $\leq 0.05$  was considered statistically significant). Using this approach 35 features were

considered as statistically significant between the two groups which have been used for the construction of a new reduced matrix. PCA and PLS-DA analysis have in turn been applied (UV scaling) to the new dataset. As shown in figure 3, a clear separation was observed for the PLS-DA model whereas in PCA the clustering was much less obvious.



**Figure 3** Score plots from PCA (A) and PLS-DA (B) after the processing procedure of plasma samples from chickens that were administrated with naringin (black dots) and control samples (red dots) employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. The models were constructed using the 35 features that were detected as statistically significant between the two groups after using the unpaired t-test ( $p$ -values  $\leq 0.05$ ) at the  $m/z$  100-900 mass range. A clear separation between the samples of the two groups was observed only for PLS-DA. The  $R^2$ ,  $Q^2$  values were for the PCA methodology, 35.1% and 0.0457, respectively and for the PLS-DA 31.3%, 85.4% and 0.678 respectively, employing 2 principal components

Permutation testing, used to confirm the model's validity, demonstrated the goodness of fit and predictive ability ( $R^2/Q^2$ ) of the original model (Figure 4).



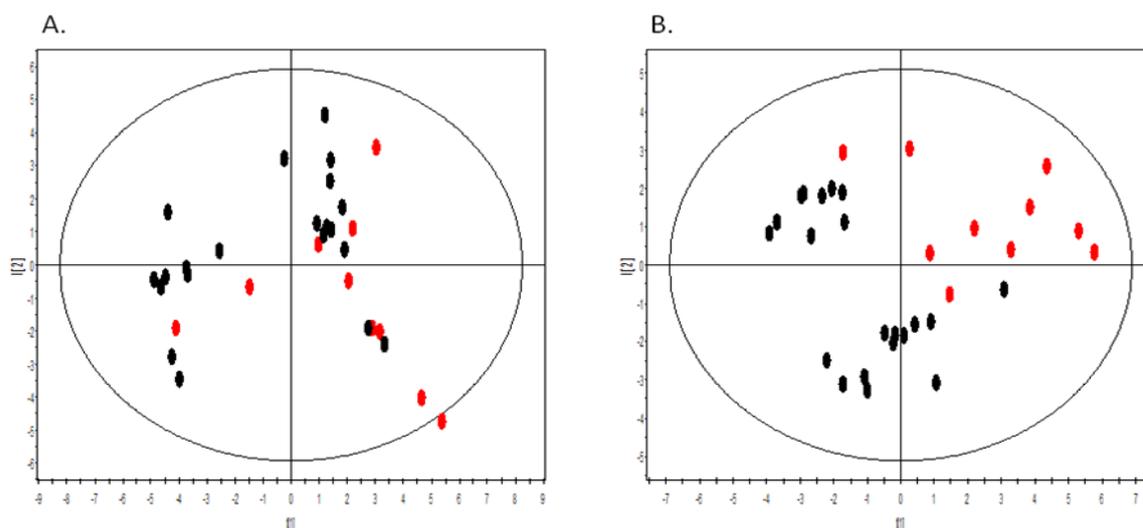
**Figure 4** Permutation testing allowing of 100 permutations of the PLS-DA model that was constructed using the 35 features that were detected as statistically significant between the two groups after using the unpaired t-test ( $p$ -values  $\leq 0.05$ ) at the  $m/z$  100-900 mass range employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. The goodness of fit ( $R^2$ ) and predictive capability ( $Q^2$ ) of the original model are indicated on the far right and remain higher than those of the 100 permuted models to the left

These 35 features were further evaluated according to the filter criteria described above (Section 2.5.1). Out of the 35 statistically significant features only 13 fulfilled those criteria which were considered as the chromatographically reliable features.

### 3.1.2 Extension of the classical approach. MVA analysis of the $m/z$ 100-900 mass range using the Variable Importance in Projection Scores (VIP)

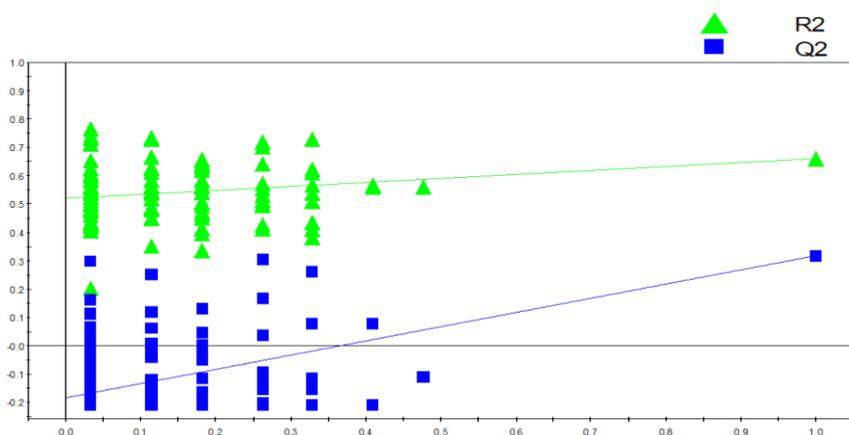
After data processing using Mzmine, the CSV file has been analyzed employing PLS-DA in order to identify those features that contributed the most

to the clustering. The 45 features that had VIP values  $>1$  were selected as deemed significant [1]. In order to examine the possible classification with this filtered dataset, PCA and PLS-DA analysis was performed, using UV scaling. Two distinct groups were observed only for the PLS-DA (Figure 5).



**Figure 5** Score plots from PCA (A) and PLS-DA (B) after the processing procedure of chicken plasma samples that were administrated with naringin (black dots) and control samples (red dots) employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. The models were constructed using 45 features that were identified as significant according to the VIP values (VIP value of  $>1$  was regarded as significant) at the mass range  $m/z$  100-900. A clear separation between samples of the two groups was observed only for PLS-DA. The  $R^2$  and  $Q^2$  values were 54.9% and 0.0929 for the PCA methodology, and 32.9%, 66.1% and 0.318 for the PLS-DA respectively, employing 4 principal components for PCA and 2 for PLS-DA

Permutation testing applied to the PLS-DA using 100 random permutations demonstrated that few of the permuted models had better fit ( $R^2$ ) and predictive ability ( $Q^2$ ) than those of the original model, implying some degree of instability for the produced model (Figure 6).



**Figure 6** Permutation testing allowing of 100 permutations of the PLS-DA model that was constructed using 45 features that were detected as significant according to the VIP values (VIP value of >1 was regarded as significant) at the mass range of  $m/z$  100-900 employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. Few of the permuted models had better fit ( $R^2$ ) and predictive ability ( $Q^2$ ) than those of the original model

The 45 features were further evaluated and only 11 fulfilled the criteria described on section 2.5.1, which were considered in turn as the chromatographically reliable features.

### 3.2 New approach for the processing of the raw data using multiple $m/z$ mass ranges (crop-filtering)

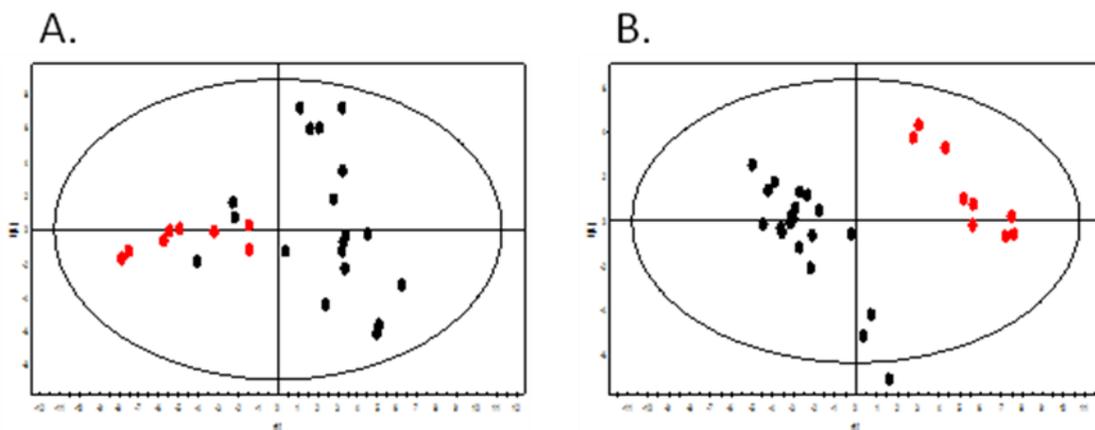
The raw mass spectral data obtained from the LC-MS analysis were imported to Mzmine and shred using the crop filter command in 100 Da slices generating 8 datasets ( $m/z$  100-200,  $m/z$  200-300 etc). For each  $m/z$  range, appropriate noise level thresholds had been calculated according to the dataset (Supplementary Table S2). PCA and PLS-DA analysis had been employed for each dataset, applying UV scaling. No significant differences between samples were observed on the PCA model (Supplementary Fig. S1).

Nevertheless, a clear separation was readily observable on the PLS-DA models using UV scaling, based on the administration status (treated vs. untreated) up to the mass range of  $m/z$  500-600 (Supplementary Fig. S2). For the three remaining mass ranges ( $m/z$  600-700,  $m/z$  700-800,  $m/z$  800-900), no separation was achieved due to the fact that there are a very limited number of features for these areas.

Permutation testing (100 random permutations) demonstrated that few of the permuted models had better fit ( $R^2$ ) and predictive ability ( $Q^2$ ) than those of the original models, implying some degree of random classification (Supplementary Fig. S3).

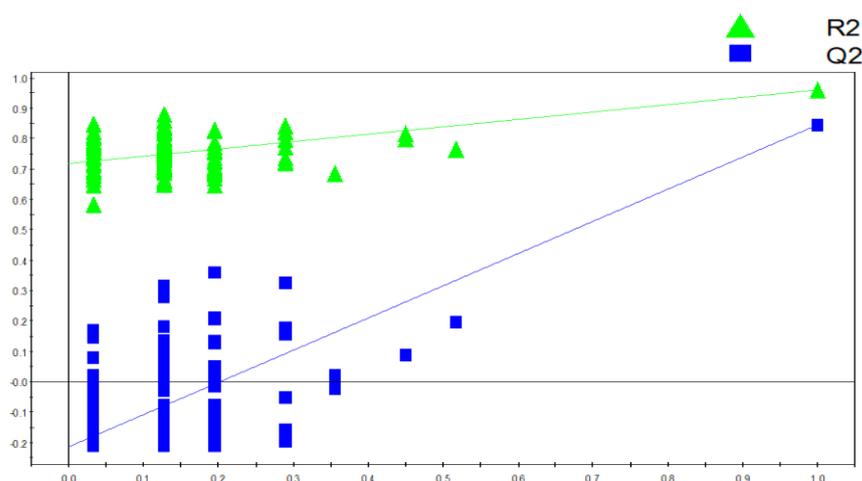
### **3.2.1 Extension of the new approach. MVA analysis of multiple $m/z$ mass ranges employing the unpaired t-test values**

The differentially expressed metabolites between the two groups were discovered using the unpaired t-test, adopting a  $p < 5\%$  significance level for all features detected in each  $m/z$  mass range slice. In summary, 91 statistically significantly differentiated features were detected and were in turn merged into a new reduced matrix, analyzed by PCA and PLS-DA (UV scaling). PLS-DA was able to accurately differentiate between the groups, whereas PCA indicated only a tendency towards the group discrimination (Figure 7).



**Figure 7** Score plots from PCA (A) and PLS-DA (B) after the processing procedure of plasma samples from chickens that were administrated with naringin (black dots) and control samples (red dots) employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. The models were constructed using the 91 features that were detected as statistically significantly differentiated between the two groups after using the unpaired t-test ( $p$ -values  $\leq 0.05$ ) on multiple  $m/z$  mass ranges. A clear separation between samples of the two groups was observed on PLS-DA. The  $R^2$ ,  $Q^2$  values were for the PCA methodology, 30.7% and 0.14 respectively, and for the PLS-DA 25.3%, 96% and 0.85 respectively, employing 2 principal components

Again permutation testing (100 random permutations) has been employed, (Figure 8) affording a valid model.

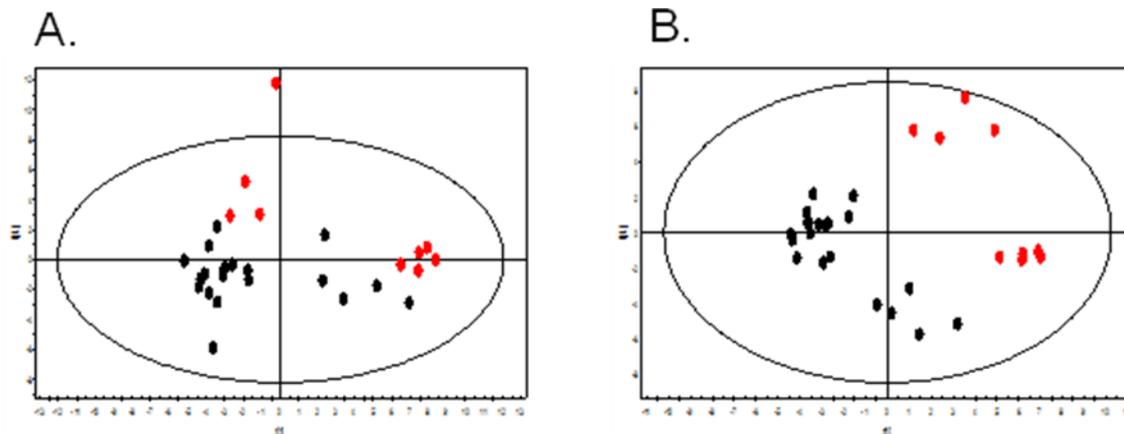


**Figure 8** Permutation testing allowing of 100 random permutations of the PLS-DA model that was constructed by the 91 features that were detected as statistically significantly differentiated between the two groups after using the unpaired t-test ( $p$ -values  $\leq 0.05$ ) on multiple  $m/z$  mass ranges. The  $R^2$  and  $Q^2$  values decreased upon 100 permutations which verified that the model was not by chance.

From the 91 statistically significant features, 47 were considered as chromatographically reliable (Section 2.5.1).

### 3.2.2 Extension of the new approach. MVA analysis of multiple $m/z$ mass ranges employing the Variable Importance in Projection Scores (VIP)

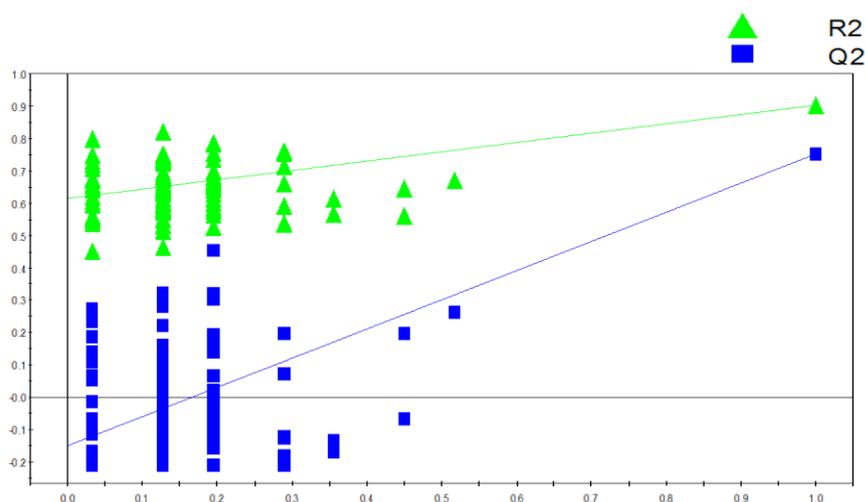
Each  $m/z$  mass range slice has been analyzed employing PLS-DA and a new dataset has been constructed from merging a total of 83 features that were detected with VIP values  $>1$  from each slice. The resulting matrix has been analyzed as described before, showing clear distinction for PLS-DA in contrast to small separation tendency for PCA (Figure 9).



**Figure 9** Score plots from PCA (A) and PLS-DA (B) after the processing procedure of plasma samples from chickens that were administrated with naringin (black dots) and control samples (red dots) in UHPLC-ESI-HRMS (Orbitrap) analysis, in the negative ion mode. The models were constructed from a reduced matrix of 83 features that were detected to contribute to the discrimination between the two groups according to VIP scoring (VIP value of  $>1$ ) on multiple  $m/z$  mass ranges. A clear separation between samples of the two groups was observed on both plots.

The  $R^2$ ,  $Q^2$  values were for the PCA methodology, 36.9% and 0.16 respectively, and for the PLS-DA 35.3%, 90.2% and 0.75 respectively, employing 2 principal components

Permutation testing (100 random permutations) of the PLS-DA model demonstrated that the classification model was accurate and reliable (Figure 10).



**Figure 10** Permutation testing allowing of 100 random permutations of the PLS-DA model that was constructed based on the 83 features that were detected to contribute to the discrimination between the two groups according to VIP scoring (VIP value of  $>1$ ) on multiple  $m/z$  mass ranges employing UHPLC-ESI-HRMS (Orbitrap) analysis, in the negative ion mode. The goodness of fit ( $R^2$ ) and predictive capability ( $Q^2$ ) of the original model are indicated on the far right and remain higher than those of the 100 permuted models to the left

Out of the 83 detected features, 48 were considered as chromatographically reliable features (Section 2.5.1).

### 3.3 Comparison of the two approaches

In almost all the metabolomics studies, except of the methodology of spectral stitching, [16] a wide mass range has being applied [11, 9, 7]. However, this

procedure seems to result to the detection of a significant number of false positive features.

A main goal of the current work was to investigate whether the number of the detected CSRF's is increased when the processing procedure is applied to multiple narrow mass ranges instead of the processing over a wide mass range since the parameters that will be used are specific for each  $m/z$  slice.

### **3.3.1 Extension of the MVA approaches**

The extension of the MVA classical methodology even though it can be considered as redundant, nevertheless it has been integrated to the proposed workflow for comparing the results to the corresponding ones from the merged dataset of the new approach. In other words a new matrix has been formed from either the VIP or the t-test discovered features in the later case. In order to analyze this matrix, it is evident some kind of MVA should be employed. Thus it has become imperative that the classical approach derived data should undergo a similar procedure in order for the outcome to be comparative.

### **3.3.2 MVA comparison**

The results exhibit that the PLS-DA methodology in every case provided better separation than PCA as expected. In the classical approach, the extension of the methodology, the t-test-based analysis afforded acceptable separation retaining good predictability in contrast to the VIP-based, which showed very low  $Q^2$ . However, the new approach provided excellent model fit and prediction ability in both cases, thus indicating its superior performance.

### 3.3.3 Significance of features

In order to assess the effectiveness of the new metabolomics processing approach in terms of discovering yet undetected features with the classical approach, it is imperative to assess the number of chromatographically reliable features.

By the classical methodology employing the t-test approach 35 features have been detected but only 13 of them fulfilled the posed criteria as the chromatographically reliable features. In order to further evaluate the validity of the model, the positive predictive value (PPV, precision) and the false discovery rate (FDR) were calculated according to the following equations:

$$\text{PPV} = \frac{\sum \text{True positive}}{\sum \text{Test outcome positive}}$$

$$\text{FDR} = \frac{\sum \text{False positive}}{\sum \text{Test outcome positive}}$$

Interestingly, employing the new methodology with the t-test approach the number of features that were discovered has been increased to 91; 47 of them being chromatographically reliable features. Thus the rate of reliable feature detection has been increased from 37.1% to 51.6% (PPV) exhibiting a substantial augmentation of the discovery rate by 14.5%. The respective FDR rate is calculated to be 62.9% for the classical and 48.4% for the new methodology.

In the case of the VIP – based approach a total of 45 features have been detected with the classical methodology, 11 of them being chromatographically, whereas 83 features were discovered by the new approach with 48 of them being the reliable attributes. The respective discovery rate has been elevated from 24.4% to 57.8% (PPV) a 33.3% increase. The relevant FDR rate is calculated to be 75.6% for the classical and 42.2% for the new methodology.

This increase proves that the new methodology could deliver more trustworthy results. Furthermore the absolute number of chromatographically reliable

features has increased by 34 for the t-test approach and 37 by the VIP. The discovery rate of the chromatographically reliable features has increased by 361.5% employing the t-test based new methodology whereas the respective % ratio for the VIP values was calculated to be 436.3 %. However, it is important to calculate if the ratio of the non-reliable features also increases by the same rate, which would prove that this approach only decreases noise level without discriminating in some way between the features. The calculated increased was 200.0% for the t-test based approach but interestingly only 102.9 % for the VIP. This observation proves that effectively the positive identification rate increases (especially for the VIP case), which essentially means that 4-fold number of putative metabolites are detected by the proposed approach. Therefore the VIP based methodology is the recommended one by the new approach.

Figures 11 and 12 summarize the numbers of the chromatographically reliable features that were detected according to the t-test and the VIPs scoring for each  $m/z$  slice compare to the wide mass range of  $m/z$  100-900.

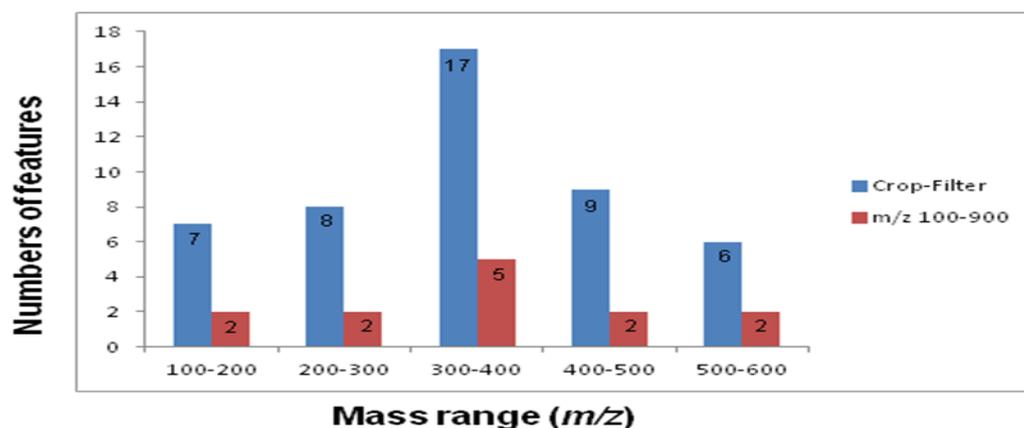
By the new methodology the noise level is calculated individually for each mass slice. The calculation was performed to a chromatographic area where no peak was detected for all the chromatograms and the S/N ratio has been determined as the standard deviation of the baseline based to the calculation of the TIC, measured for 1 min. This new approach enables improved assessment of the features, as individual S/N ratio levels specific for each slice are applied, instead of using one, which can be considered much less specific as in the case of the wide mass range. E.g. the noise levels for the  $m/z$  100-900 range has been calculated in terms of standard deviation (SD) to be 51.735 a value 24 fold larger than the one calculated for the  $m/z$  500-600 range which calculated to be 2.185.

Thus, keeping the  $m/z$  100-900 S/N ratio for the whole chromatogram would lead to substantial loss of information for the  $m/z$  500-600 range. Note that the normalization level for all features detected in the  $m/z$  500-600 range is well

below (e.g. 554.2994 intensity 13500, 526.9852 intensity192000) the noise threshold set by the wide range, thus it would render these features undetected.

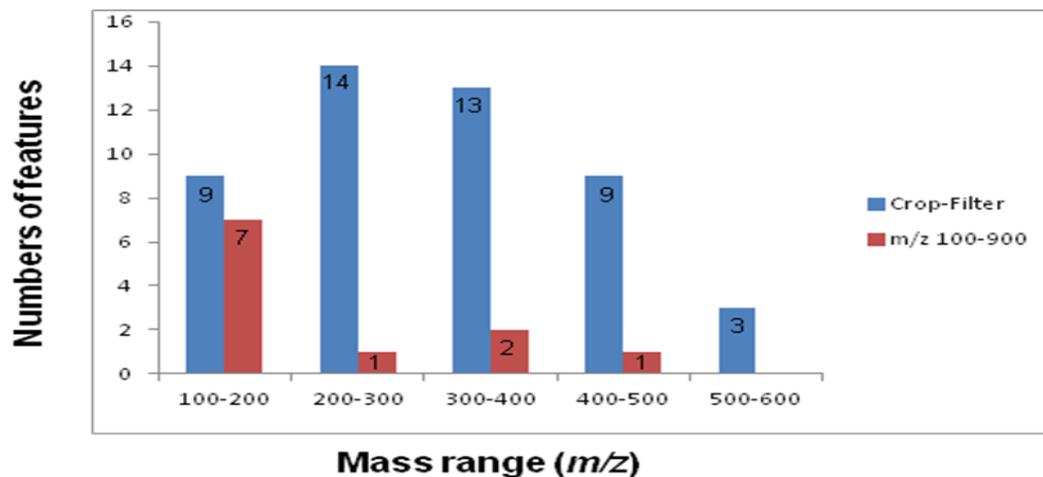
It should also be noted that in the case of the VIP values the differentiation of the results between the classical and the new approach could also be partly explained from the fact that new significant variables (according to the VIP score) emerge. This alters the multivariate space i.e. the analysis is based to different features between the classical and the new approach. In the t-test case the differentiation results only to the different S/N levels.

Furthermore, the new methodology could be considered as a weighted approach between the  $m/z$  of an LC/MS chromatogram. Thus in the wide area based methodology, the density of the  $m/z$  values is much larger at the low region and these constitute the driving force for group separation by the MVA methodologies. On the contrary in the case of the new methodology, as the slices are treated independently, the statistical weight is equally divided in all of them. It is noted again that the larger the number of the existing variables in a multivariate methodology, the larger the chance to find some, that actually produce separation even this is not consistent but rather random.



**Figure 11** Histograms of the numbers of chromatographically reliable features that were detected according to the t-test for each  $m/z$  range compared to the wide mass range of  $m/z$  100-900. For the three remaining mass ranges ( $m/z$  600-700,  $m/z$  700-

800,  $m/z$  800-900) no separation was achieved due to the fact that there was a very limited number of detected features for these areas



**Figure 12** Histograms of the numbers of chromatographically reliable features that were detected according to VIPs scoring for each  $m/z$  range compared to the wide mass range of  $m/z$  100-900. For the three remaining mass ranges ( $m/z$  600-700,  $m/z$  700-800,  $m/z$  800-900) no separation was achieved due to the fact that there was a very limited number of detected features for these areas

#### 4. Concluding remarks

In this study a new strategy was reported on the detection of chromatographically reliable features during a metabolomic approach. The shredding of the LC-MS chromatograms into multiple  $m/z$  ranges increased the number of the identified chromatographically reliable features. Moreover, using either the statistical significant features from the unpaired t-test or the variables after VIPs scoring a valid model can be produced. To our concern, this is the first time that an untargeted metabolomic analysis using multiple  $m/z$  ranges is being performed.

## References

1. Andersen CM, & Bro R. (2010) Variable selection in regression—a tutorial. *J Chemometr.* 24, 728–737.
2. Barker M., & Rayens W. (2003) Partial least squares for discrimination. *Journal of Chemometrics.* 17, 166–173.
3. Fernández-Albert F, Llorach R, Andrés-Lacueva C, Perera A. (2014) An R package to analyse LC/MS metabolomic data: MAIT (Metabolite Automatic Identification Toolkit). *Bioinformatics.* 30, 1937-1939.
4. Goodacre R. (2005) Making sense of the metabolome using evolutionary computation: seeing the wood with the trees. *J Exp Bot.* 56, 245-254.
5. Kell DB. (2007) Metabolomic biomarkers: search, discovery and validation. *Expert Rev Mol Diagn.* 7, 329-333.
6. Kind T, Fiehn O. (2007) Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics.* 8, 105.
7. Kind T, Tolstikov V, Fiehn O, Weiss RH. (2007) A comprehensive urinary metabolomic approach for identifying kidney components. *Anal Biochem.* 363, 185-195.
8. Lindon JC, Nicholson JK. (2008) Spectroscopic and statistical techniques for information recovery in metabolomics and metabonomics. *Annu Rev Anal Chem (Palo Alto Calif).* 1, 45-69.
9. Liu SY, Zhang RL, Kang H, Fan ZJ, Du Z. (2013) Human liver tissue metabolic profiling research on hepatitis B virus-related hepatocellular carcinoma. *World J Gastroenterol.* 19, 3423-3432.

10. Lommen A, Kools HJ. (2012) MetAlign 3.0: performance enhancement by efficient use of advances in computer hardware. *Metabolomics*. 8, 719-726.
11. Mueller DC, Piller M, Niessner R, Scherer M, Scherer G. (2014) Untargeted metabolomic profiling in saliva of smokers and nonsmokers by a validated GC-TOF-MS method. *J Proteome Res*, 13, 1602-1613.
12. Naz.S, Vallejo M, Garcia A, Barbas C. (2014) Method validation strategies involved in non-targeted metabolomics. *J Chromatogr A*, 1353, 99-105.
13. Patti GJ, Yanes O, Siuzdak G. (2012) Metabolomics: the apogee of the omic trilogy. *Nat Rev Mol Cell Biol*, 13, 263-269.
14. Pluskal T, Castillo S, Alejandro Villar-Briones A, Oresic M. (2010) Mzmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics*. 11, 395.
15. Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. (2006) XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem*. 78, 779-787.
16. Southam AD, Payne TG, Cooper HJ, Arvanitis TN, Viant MR. (2007) Dynamic range and mass accuracy of wide-scan direct infusion nanoelectrospray fourier transform ion cyclotron resonance mass spectrometry-based metabolomics increased by the spectral stitching method. *Anal Chem*, 79, 4595-4602.
17. Tautenhahn R, Böttcher C, Neumann S. (2008) Highly sensitive feature detection for high resolution LC/MS. *BMC Bioinformatics*. 9, 504.

18. Tautenhahn R, Patti GJ, Rinehart D, Siuzdak G. (2012). XCMS Online: a web-based platform to process untargeted metabolomic data. *Anal Chem*, 84, 5035-5039.
19. Van der Kloet FM, Hendriks M, Hankemeier T, Reijmers T. (2013) A new approach to untargeted integration of high resolution liquid chromatography-mass spectrometry data. *Anal Chim Acta*, 801, 34-42.
20. Westerhuis JA, Hoefsloot HCJ, et al (2008) Assessment of PLS-DA cross validation. *Metabolomics*. 4, 81–89.
21. Yu T, Park Y, Johnson JM, Jones DP. (2009) apLCMS—adaptive processing of high-resolution LC/MS data. *Bioinformatics*. 25, 1930-1936.

## Chapter abstract

In the case of the MS-based metabolomics, the large number of false positives remains a fundamental issue. The aim of this study was to develop a new strategy, which highlights the number of the reliable features i.e. the detected features that correspond to a consistent peak according to chromatographic and mass spectrometric criteria. Two methodologies have been compared for data processing. In the first one (classical approach), all data in the 100-900  $m/z$  mass-to charge range were included for the data processing procedure whereas for the newly developed methodology, the data were shred in 100 Da slices generating 8 datasets, which have been then subjected to the downstream MS data processing. Each dataset was treated separately and the  $m/z_{tR}$  features obtained by either VIP's or t-test values were merged and used as the input for the construction of the general model. Overall the new methodology resulted to a 4-fold increase of the peaks that could be considered chromatographically and mass spectrometrically valid. For the analysis blood samples from 20 chickens, which were administrated with naringin in their nutrition and 9 samples from control, were analyzed by UHPLC-HRMS (Orbitrap Velos).

**Keywords:** Mzmine – PCA – PLS-DA – permutation test – VIPs – t-test

## Supporting Information for Chapter 4

Supplementary Table S1

<b>Parameters</b>	<b>Mass range 100-900 (<i>m/z</i>)</b>
<b>Noise level</b>	<b>5 10<sup>5</sup></b>
<b>Minimum height</b>	<b>7 10<sup>5</sup></b>
<b>Chromatograph deconvolution</b>	<b>Noise amplitude</b>
<b>Retention time tolerance (min)</b>	<b>0.08</b>
<b>Alignment</b>	<b>Join aligner</b>

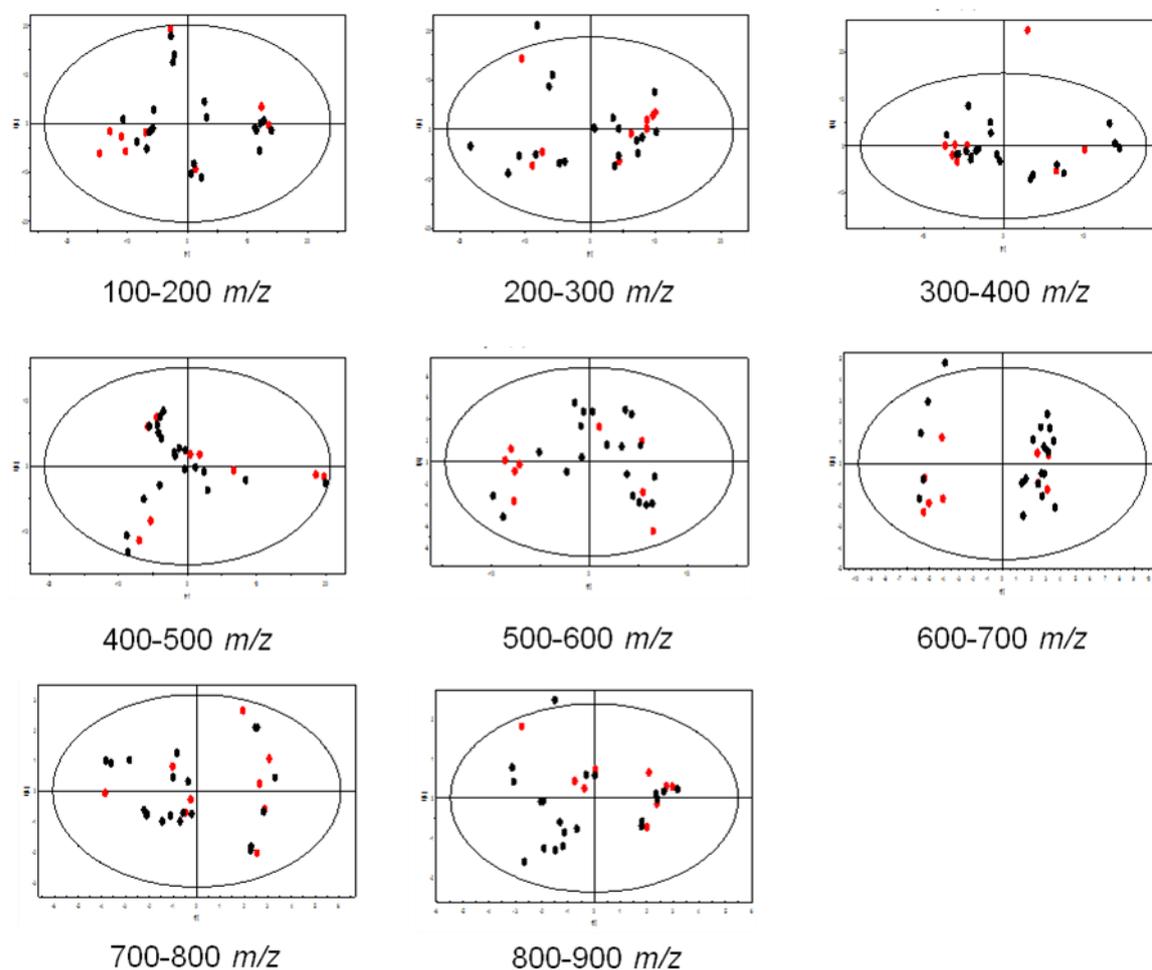
**Table S1** Parameters that were used for the processing procedure of the data in the mass range of *m/z* 100-900 by the open software Mzmine 2.18.1

**Supplementary Table S2**

Parameters	Mass range ( <i>m/z</i> )							
	100-200	200-300	300-400	400-500	500-600	600-700	700-800	800-900
Noise level	10 <sup>4</sup>	2 10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>4</sup>	2.5 10 <sup>4</sup>	2 10 <sup>4</sup>	2.5 10 <sup>4</sup>
Minimum height	2 10 <sup>4</sup>	4 10 <sup>4</sup>	2 10 <sup>4</sup>	2 10 <sup>4</sup>	2 10 <sup>4</sup>	5 10 <sup>4</sup>	4 10 <sup>4</sup>	5 10 <sup>4</sup>
Chromatograph deconvolution	Noise amplitude	Noise amplitude	Noise amplitude	Noise amplitude	Noise amplitude	Noise amplitude	Noise amplitude	Noise amplitude
Retention time tolerance (min)	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
Alignment	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner

**Table S2** Parameters that were used for the processing procedure of the data using multiple *m/z* mass ranges by the open software Mzmine 2.18.1

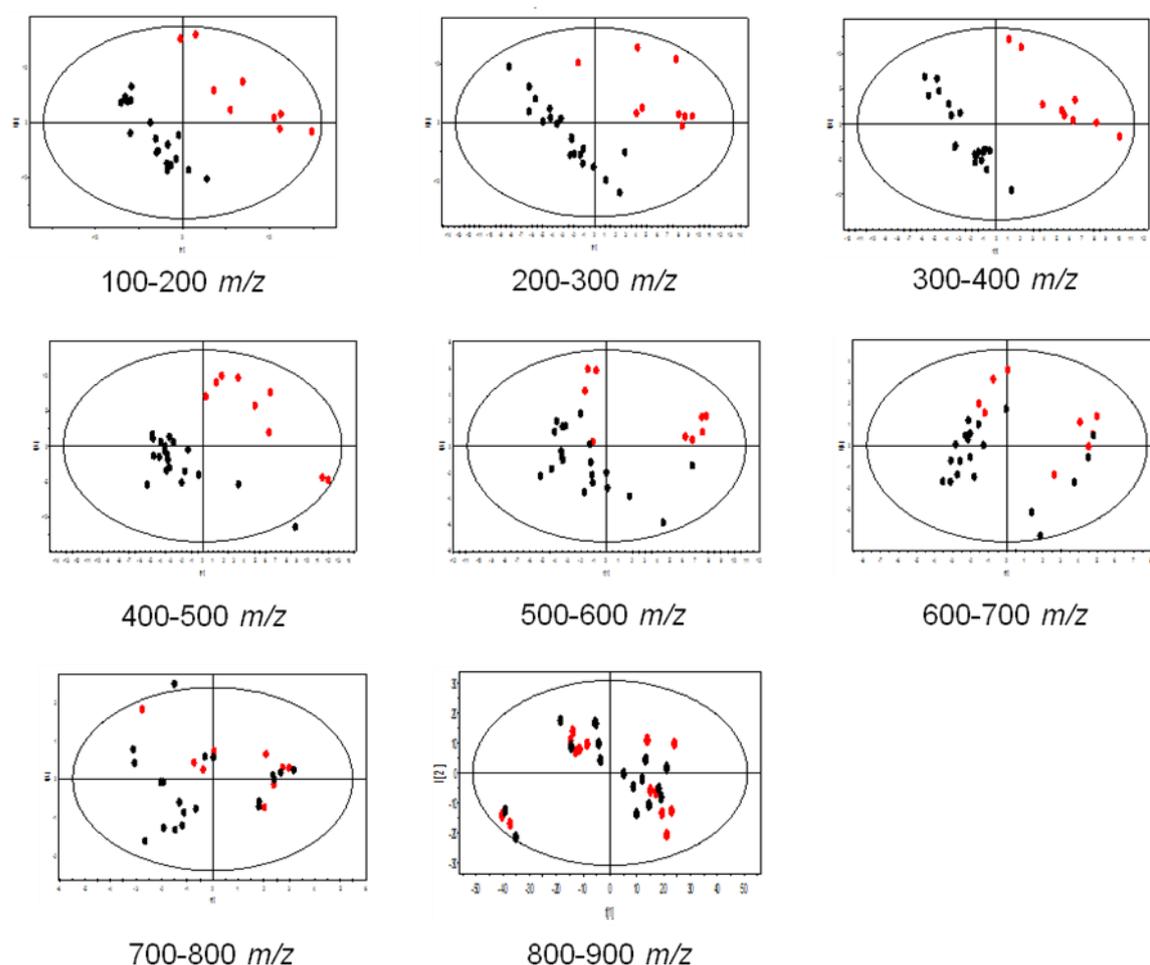
## Supplementary Fig S1



**Fig S1** Score plots from PCA after processing of the raw data using multiple mass ranges of *m/z* 100 Da from chicken plasma samples that were administrated with naringin (black dots) and control samples (red dots) employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. No significant differences between the two groups were observed. The  $R^2$ ,  $Q^2$  values were for the *m/z* 100-200 region, 22.2% and 0.0687 respectively, for the *m/z* 200-300 region, 24.7% and 0.0527 respectively, for the *m/z* 300-400 region, 20.8% and -0.00239 respectively, for the *m/z* 400-500 region, 32.7% and 0.0797 respectively, for the *m/z* 500-600 region, 36.5% and 0.205 respectively, for the *m/z* 600-700 region, 56.3% and 0.375 respectively, for the *m/z* 700-800 region, 67.4% and 0.359 respectively, and for the

*m/z* 800-900 region, 44.4% and 0.239 respectively, employing 2 principal components

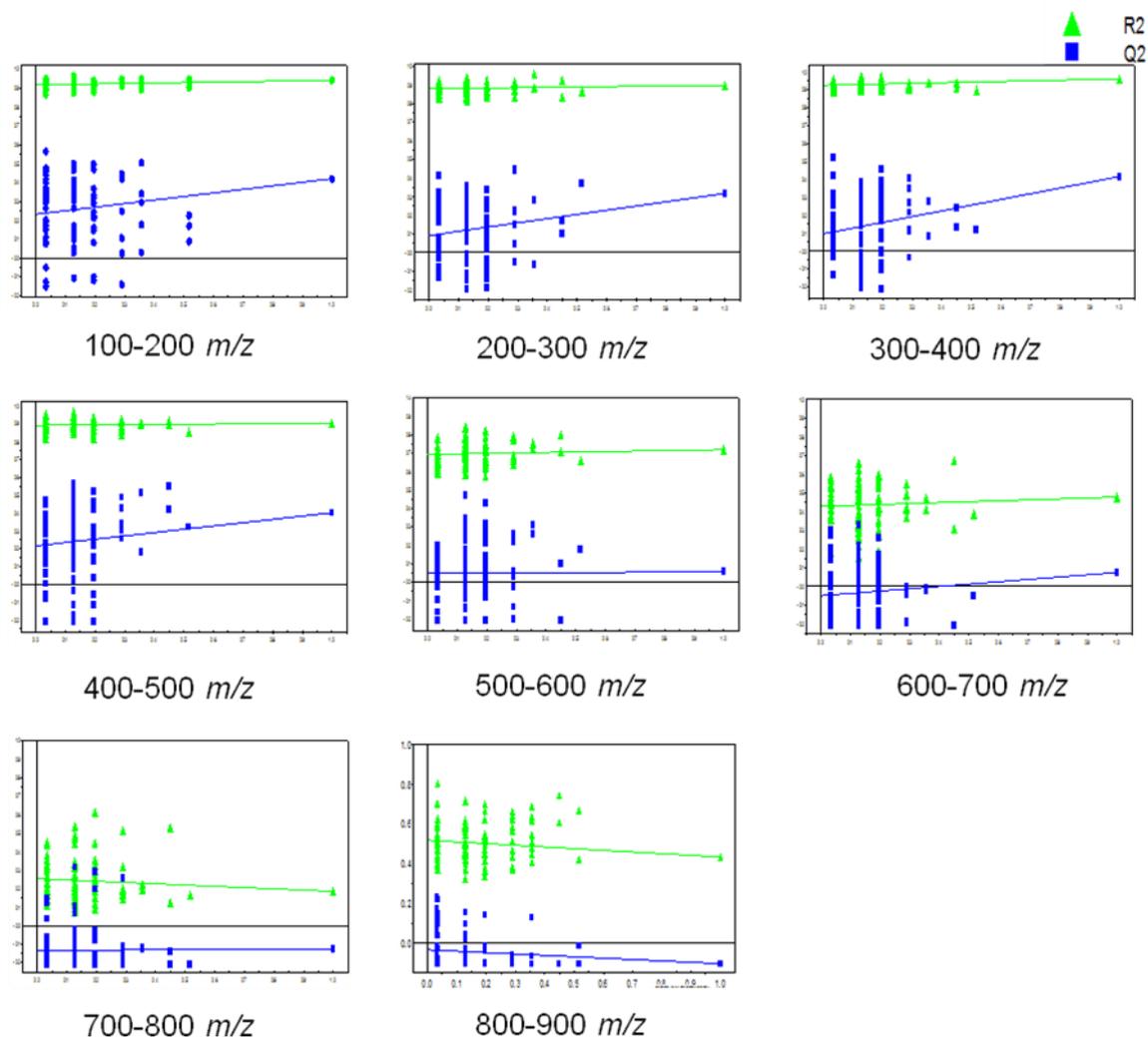
## Supplementary Fig S2



**Fig S2** Score plots from PLS-DA analysis after processing of the raw data using multiple mass ranges of *m/z* 100 Da from chicken plasma samples that were administrated with naringin (black dots) and control samples (red dots) employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. Two distinct groups were observed based on the administration scheme (treated vs. untreated) up to the mass range of *m/z* 500-600. The  $R^2$ ,  $Q^2$  values were for the *m/z* 100-200 region, 26.7%, 99.7% and 0.49 respectively, for the *m/z* 200-300 region, 40.6%, 1% and 0.802 respectively, for the *m/z* 300-400 region, 14.4%, 96.1% and 0.415 respectively, for the *m/z* 400-500 region, 20%, 90.3% and 0.4 respectively, for the *m/z* 500-600 region, 43.7%, 95.1% and 0.172 respectively, for the *m/z* 600-700 region, 54.3%,

47.8% and 0.0741 respectively, for the  $m/z$  700-800 region, 62%, 18.5% and -0.127 respectively, and for the  $m/z$  800-900 region, 50.1%, 21.5% and -0.228 respectively

### Supplementary Fig S3



**Fig S3** Permutation testing allowing of 100 random permutations of the PLS-DA models after processing of the raw data using multiple mass ranges of  $m/z$  100 Da from chicken plasma samples that were administrated with naringin and control samples employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. Few of the 100 permuted models presented better fit ( $R^2$ ) and predictive capability ( $Q^2$ ) than those of the original models

## Chapter 5

### **UHPLC-HRMS-based plasma metabolomics study of naringin and hesperidin after dietary supplementation in chickens**

The new methodology, as described in Chapter 4, was applied to an untargeted metabolomics study in chicken plasma samples that were administered either with hesperidin or naringin for 30 days, in order to assess the impact of the administered substances on the total metabolic profile of the experimental animals.

## 1. Introduction

Flavonoids are a class of polyphenolic molecules present in foods, such as fruits, vegetables and plant derived juices including tea, coffee and wine. They represent one of the most important classes of biologically active compounds and occur both in the free state and as glycosides [22]. In the past decade an increasing number of studies regarding the positive effects of these natural compounds on human health have been reported [25]. Flavonoids are categorized into six major subclasses based on their range structural complexity: flavonols, flavones; flavan-3-ols, flavanones, anthocyanins and isoflavones [19].

Hesperidin is a flavanone glycoside that is found abundantly in citrus fruits such as oranges, lemons, grapes etc. It presents a broad spectrum of biological activities like antioxidant, anti-inflammatory, anticarcinogenic and antiallergic which are well-known and promising. Hesperidin deficiency has been linked with abnormal capillary leakiness as well as pain in the extremities causing aches, weakness and night leg cramps [7]. Naringin is a well-known flavanone glycoside of grapefruits and citrus fruits [11]. It exerts a variety of biological and pharmacological effects such as anti-inflammatory, antioxidant and lipid-lowering activities [3].

Hesperidin and naringin are metabolized by human intestinal bacterial microflora to the aglycones hesperetin and naringenin, respectively [6]. Over the past decade, a large number of studies were conducted to determine the molecular targets and underlying mechanisms of hesperidin and naringin as well as their metabolites [17, 2].

Metabolomics is a high throughput analysis allowing the identification and quantification of endogenous low molecular weight compounds (e.g., <1 kDa) in a biological system [10]. Two analytical platforms are currently used for metabolomic analyses: Mass Spectrometry and NMR-based systems. These technologies are both complementary and, therefore, often used in parallel in

metabolomic research. Compared to NMR, mass spectrometry is a more sensitive technique [7].

The metabolomics approach, using mass spectrometry combined to a separation technique e.g. LC-MS, allows simultaneous detection of hundreds of low molecular weight metabolites within a biological matrix [23]. The common metabolomic analysis pipeline includes sample preparation; instrumental analysis of the samples (data acquisition); data processing and biological interpretation of the results [20]. The pattern recognition is based on multivariate statistics for elucidating metabolite signatures. In the future, metabolomic analysis might become the preferred tool for diagnostics and prediction of diseases. To date, a significant number of potential metabolite biomarkers have been discovered by profiling the human metabolome [8].

In the past few years, there is a growing interest in dietary bioactive compounds that protect against or mitigate chronic diseases without the undesirable side effects. Based on the biological activities of hesperidin and naringin, these two natural compounds are very promising candidates as alternatives to synthetic supplements. However, no data support a metabolomic strategy after dietary supplementation of hesperidin and naringin.

In the current study, an untargeted UHPLC-HRMS based plasma metabolomic approach was performed in order to characterize the metabolomic fingerprinting of hesperidin and naringin after dietary supplementation in chickens.

## **2. Materials and Methods**

### **2.1 Chemicals and Reagents**

All solvents were of LC-MS grade. Acetonitrile and water was purchased from Fluka/Riedel-de Haën (Switzerland). Glacial acetic acid, methanol and

acetone were purchased from Sigma-Aldrich (Steinheim, German). Naringin was purchased from Alfa Aesar GmbH & Co KG (Germany) and hesperidin from TSI Europe NV (Belgium).

## **2.2 Animals and samples studies**

Forty nine 1-day-old Ross 308 broiler chickens were randomly divided into 3 groups. The chickens were obtained from a commercial hatchery and The lighting program consisted of 23L:1D on arrival, and was decreased to 18L:6D at day 7, remained constant until day 35, and thereafter gradually increased to 23L:1D at slaughter, with access to feed, in mash form, and water *ad libitum*. The treatment groups were: 20 chickens, after dietary supplementation with naringin (1.5 g/kg of feed), 20 chickens after dietary supplementation with hesperidin (1.5 g/kg of feed) and 9 control chickens that were given commercial basal diets. The administration of naringin and hesperidin was started from the 11<sup>th</sup> day of age until slaughter at the age of 42 days.

Blood samples were collected on 42<sup>th</sup> day of age, after 30 days of the administration, into heparin tubes and centrifuged for 10 minutes at 2,500 rpm to obtain plasma. The plasma samples were stored at -80<sup>0</sup>C until further use.

All experimentation was carried out in strict accordance with the guidelines of “Council Directive 86/609/EEC regarding the protection of animals used for experimental and other scientific purposes”. The protocol was approved by the Bioethical Committee of the Agricultural University of Athens (Permit Number: 20/20032013).

## **2.3 Sample Preparation**

50 µL of plasma samples were used for the analysis. For the precipitation of the plasma proteins, cold acetone was added to reach a final ratio of 1/8 plasma/acetone (v/v). Extraction was conducted by vortexing for 1 min and

the mixture was left for 10 minutes at  $-20^{\circ}\text{C}$  for total precipitation achievement. The extracts were then centrifuged at 12,500 rpm for 10 minutes at  $4^{\circ}\text{C}$  and the supernatant was collected and evaporated to dryness under vacuum. The residue was reconstituted with 50  $\mu\text{L}$  of methanol/water (50/50, v/v) and centrifuged at 12,500 rpm for 8 minutes at  $4^{\circ}\text{C}$ . The supernatant of 2  $\mu\text{L}$  was injected into the UHPLC-HRMS.

## 2.4 Instrumentation

An ESI-LTQ-Orbitrap Velos (Thermo Scientific, Germany) connected to an Accela UHPLC system controlled by the Xcalibur 2.1 software was used for the metabolomic analysis. Electrospray ionization was performed in both positive and negative ion mode. The UHPLC system was equipped with a binary pump, an autosampler, a vacuum degasser and a temperature-controlled column compartment.

Operational parameters of the mass spectrometer were set for the negative mode at: sheath gas, 35 (arbitrary units); source voltage, 3 kV; auxiliary gas, 30 (arbitrary units); S lens RF level, 60 (%) and capillary temperature,  $350^{\circ}\text{C}$ . For the positive mode the operational parameters were set at: sheath gas, 35 (arbitrary units); spray voltage, 3 kV; capillary voltage, -60 (V); auxiliary gas, 30 (arbitrary units); capillary temperature,  $275^{\circ}\text{C}$ ; and source temperature  $200^{\circ}\text{C}$ . The Orbitrap resolution was set at 30,000 FWHM and the data were collected in centroid mode from 100 to 900 Da in full scan mode. For the analysis 2  $\mu\text{L}$  of each sample was injected onto a reversed-phase INTERCHIM UHPLC C18 column (1.7  $\mu\text{m}$  particle size, 2.1 mm x 100 mm, Waters Corp. Milford, MSA, USA) maintained at  $40^{\circ}\text{C}$  at a flow rate of 0.36 mL/min.

The mobile phase consisted of 0.1% (v/v) glacial acetic acid (solvent A) and acetonitrile (solvent B). A gradient methodology of 32 minutes was used for

the metabolomic analysis as follows: 0 to 24 min: 95% A: 5% B, 24 to 28 min: 5% A: 95% B, 28 to 32 min: 95% A: 5% B.

Plasma samples centrifugation was performed by a Mikro 200R centrifuge (Hettich Lab Technology, Germany). Sample evaporation was performed by a GeneVac HT-4X EZ-2 series evaporator Lyospeed ENABLED (Genevac Ltd, UK).

## 2.5 Data Analysis

The raw mass spectral data were analyzed by the open source software Mzmine 2.18.2 (<http://mzmine.sourceforge.net/>) for baseline correction, peak detection, chromatograph deconvolution and normalization. Mzmine allows the detection of the mass, retention time and intensity of peaks eluting from each chromatogram.

According to the Chapter 4 “An alternative approach for the data analysis of the data processing procedure in untargeted metabolomics by UPLC-ESI (-)-HRMS. The use of post acquisition spectral stitching”, the raw mass spectra data were shred using the crop filtering command in 100 Da slices generating 8 datasets. The peaklists (accurate mass\_retention time vs. intensity) for each dataset generated by Mzmine were exported as comma-separated values (CSV) files to Excel 2007 (Microsoft Office 2007) and manipulated using the concatenate, round and transpose commands. The Web-based software MetaboAnalyst 3.0 (<http://www.metaboanalyst.ca/MetaboAnalyst/>) was used for the estimation of the missing values, data filtering and normalization [16]. Features with at least 50% of missing values were removed whereas the missing values in features with less than 50% of missing data were replaced with half of the lowest positive value in the original data. Furthermore, data filtering was performed using the relative standard deviation and normalization was achieved by averaging all the samples in the control group, in both ion modes.

The files were then imported into Simca-P+11.5 (Umetrics, Umeå, Sweden) performing multivariate statistics and the projections to latent structures – discriminant analysis (PLS-DA) with the treated vs. untreated data being the discriminant has been used as the method of choice. 8 PLS-DA models were developed – one for each slice. According to the PLS-DA models, the variable IDs (accurate mass\_retention time) that achieved a variable importance in projection scoring (VIP scoring) of 1.0 or higher [1] were regarded as the primary drivers of the calculated discrimination and were merged together in a reduced matrix. This new matrix was again subjected to multivariate analysis (MVA) analysis i.e. principal component analysis (PCA), PLS-DA analysis and Optimized Potentials for Liquid Simulations (OPLS). Different normalization methodologies had been employed i.e. Pareto or UV (auto-scaling) scaling, to the mean centered data. The variables from the PLS-DA model were subsequently ranked according to VIP scoring in order to detect the variables that contributed to the discrimination between the two groups (treated vs. untreated). For the statistical evaluation of the detected variables, normality tests were performed using the XLSTAT 2010 software. Univariate statistical analysis was performed to verify the significance of the variables that were discovered by MVA. For t-test verification, the value of  $p \leq 0.05$  was considered to indicate a statistically significant difference between the two groups. For the variables with non-normally distributed data the Mann-Whitney test was performed as an equivalent tool of t-test, using the XLSTAT 2010. Furthermore, univariate plots (box plots) were constructed by the aid of XLSTAT 2010 software. ROC analysis was performed using the online ROC CET platform (<http://www.roccet.ca/ROCCET/>). Low-quality variables, such as those containing empty/zero/missing values >50% and those with 5% near constant variables based on the relative standard deviation (RSD) were removed. The classical univariate ROC curve analysis was used. The Matlab software (MathWorks Inc, Natick, MA, USA) was used to calculate Pearson's correlation coefficient scores for the detected variables for identifying low energy pseudo MS/MS spectra. For the prediction of the molecular formulas of the detected variables, the Rdisop package from the R statistical language,

the Mzmine 2.18.2 software, the Sirius 3.1 software and the mMass 5.5.0 were employed. Moreover, the seven golden rules software was used for the identification of possible adducts on the detected variables. Online databases such as METLIN ([www.metlin.scripps.edu](http://www.metlin.scripps.edu)), KEGG ([www.genome.jp/kegg](http://www.genome.jp/kegg)), Chemspider (<http://www.chemspider.com/>) and LIPID MAPS (<http://www.lipidmaps.org/>) were used to match the selected variables in terms of mass accuracy to specific metabolites.

### **3. Results and Discussion**

#### **3.1 Data processing of the plasma samples after dietary supplementation with naringin and hesperidin**

The raw mass spectral data that were obtained after the analysis of the plasma samples on the Orbitrap Velos were shred using the crop filtering command of the Mzmine 2.18.1 software in 8 regions of  $m/z$  100 Da slices. For each region different parameters were used (Supplementary S1-S4). The data were processed according to baseline correction, peak detection, deconvolution, deisotoping, alignment and gap filling procedures.

Each generated peaklist was exported as comma-separated values (CSV) file to Excel 2007 for concatenate, round and transpose command. Subsequently, the generated .csv files were imported to Metabaoanalyst 3.0 according to the instruction for missing values estimation, data filtering and normalization [24].

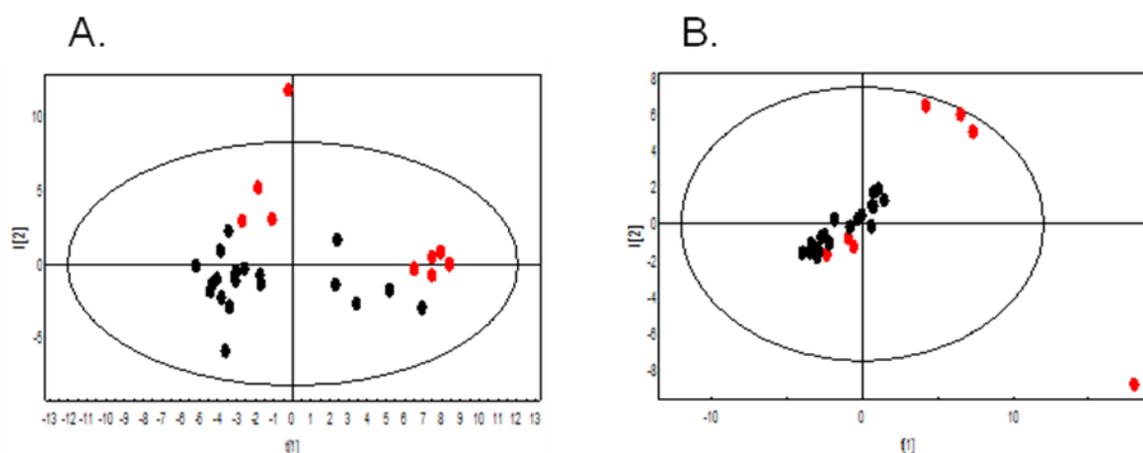
The generated .csv files after the Metaboanalyst process were imported to SIMCA-P+11.5 for PLS-DA analysis. Thus, 8 PLS-DA models were calculated. The variables from each model were subsequently ranked according to their VIP scoring. Variables that achieved a VIP score of 1.0 or higher, were regarded as the primary drivers of the calculated discrimination [1].

Totally, 75 variables were detected in the negative mode and 60 in the positive mode to discriminate naringin supplement and control group according to VIP scoring. Additionally, 68 variables were detected in the

negative mode and 60 in the positive mode to discriminate hesperidin supplement and control group. These variables were merged together in a reduced matrix and subjected to MVA.

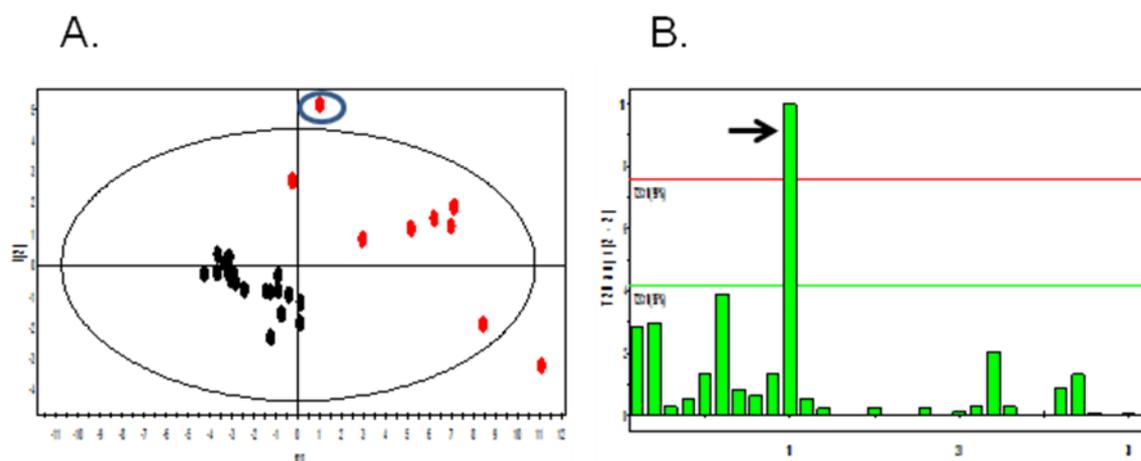
### 3.2 Multivariate statistical analysis of the chicken plasma samples after dietary supplementation with naringin

PCA was initially applied to investigate whether the group after dietary supplementation with naringin could be separated from the control group, in both ion modes. As shown in figure 1 PCA analysis showed a tendency to separate only in the negative ion mode after UV scaling.  $R^2X$  and  $Q^2$  (cum) are usually used for the evaluation of PCA model, and values of these parameters close to 1.0 indicating a good fitness and predictability of the constructed model [21]. The  $R^2$  and  $Q^2$  parameters for the classification were 0.369 and 0.157 respectively, in the negative ion mode and 0.444 and 0.018 respectively, in the positive ion mode indicating the low modeling quality of PCA models.



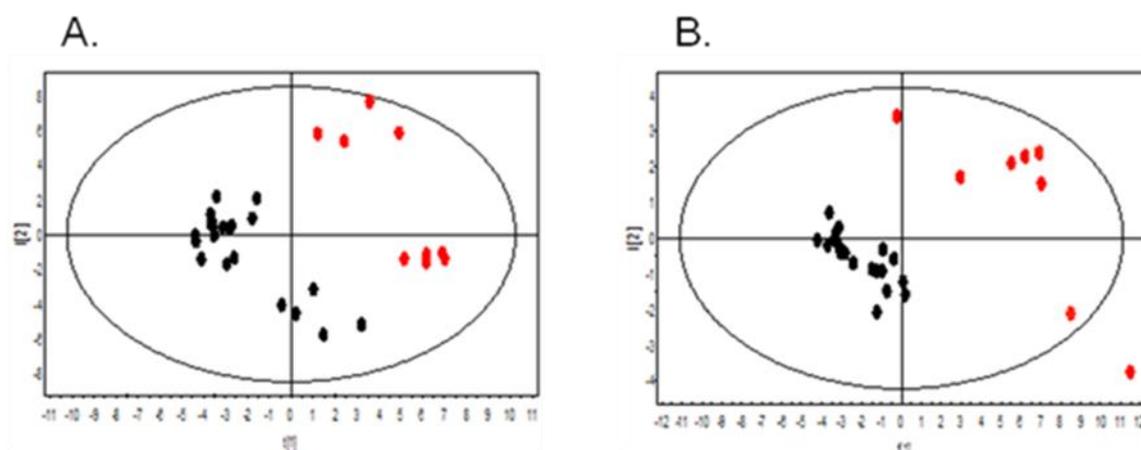
**Figure 1** Score plots from PCA based on the 75 VIPs after the analysis in the negative ion mode (A) and on the 60 VIPs in the positive ion mode (B), employing UHPLC-HRMS (Orbitrap) analysis. Naringin administration is represented by black dots and red dots depict controls. The  $R^2$ ,  $Q^2$  values were 0.369 and 0.157 respectively, in the negative ion mode and 0.444 and 0.018 respectively, in the positive mode employing 2 principal components in the negative ion mode and 5 components in the positive ion mode

In order to further characterize the differences between the two groups, PLS-DA analysis was performed to obtain better discrimination between the two groups. Nevertheless, in the positive ion mode according to the Hotelling's T2Range plot, one control sample was considered as an outlying observation (Figure 2).



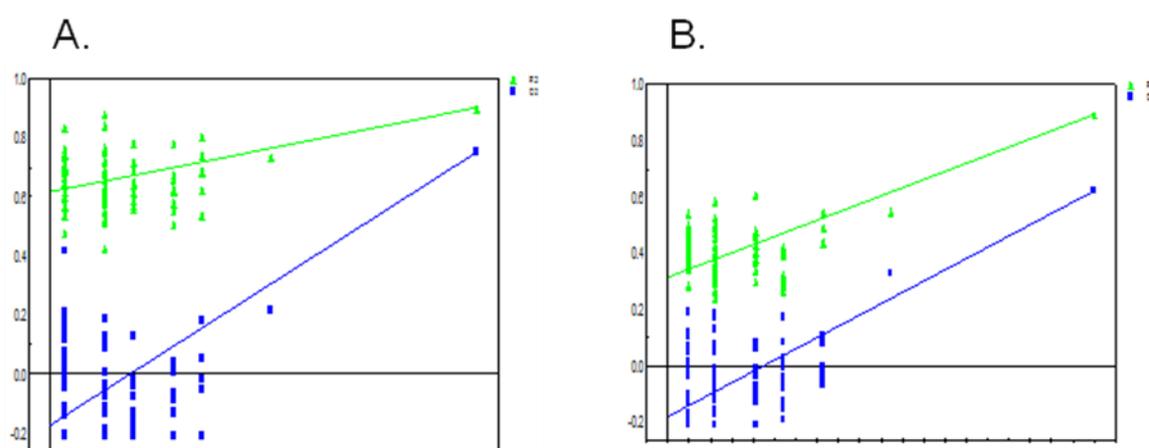
**Figure 2** Score plot from the PLS-DA (A) and Hotelling's T2 range plot (B) based on the 60 VIPs, after naringin administration employing UHPLC-HRMS (Orbitrap) analysis, in the positive ion mode. Naringin administration is represented by black dots and red dots depict controls. The  $R^2$ ,  $Q^2$  values of the PLS-DA analysis were 0.359, 0.897 and 0.624 respectively. According to the T2 range plot, which displays the distance from the origin in the model plane (score space) for each selected observation, one control sample (black arrow) was considered as outlier since its value was larger than the red limit (0.01 levels). Blue circle on PLS-DA plot indicates the selective observation

The model was re-calculated and the control sample was not included in the new workset. After UV scaling clear separation between the two classes was observed on both ion modes (Figure 3).



**Figure 3** Score plots from PLS-DA based on the 75 VIPs after the analysis in the negative mode (A) and on the 60 VIPs in the positive ion mode (B), employing UHPLC-HRMS (Orbitrap) analysis. Naringin administration is represented by black dots and red dots depict controls. The  $R^2$ ,  $Q^2$  values were 0.353, 0.902 and 0.752 respectively, in the negative ion mode and 0.538, 0.977 and 0.631 respectively, in the positive ion mode employing 2 principal components in the negative ion mode and 4 in the positive ion mode

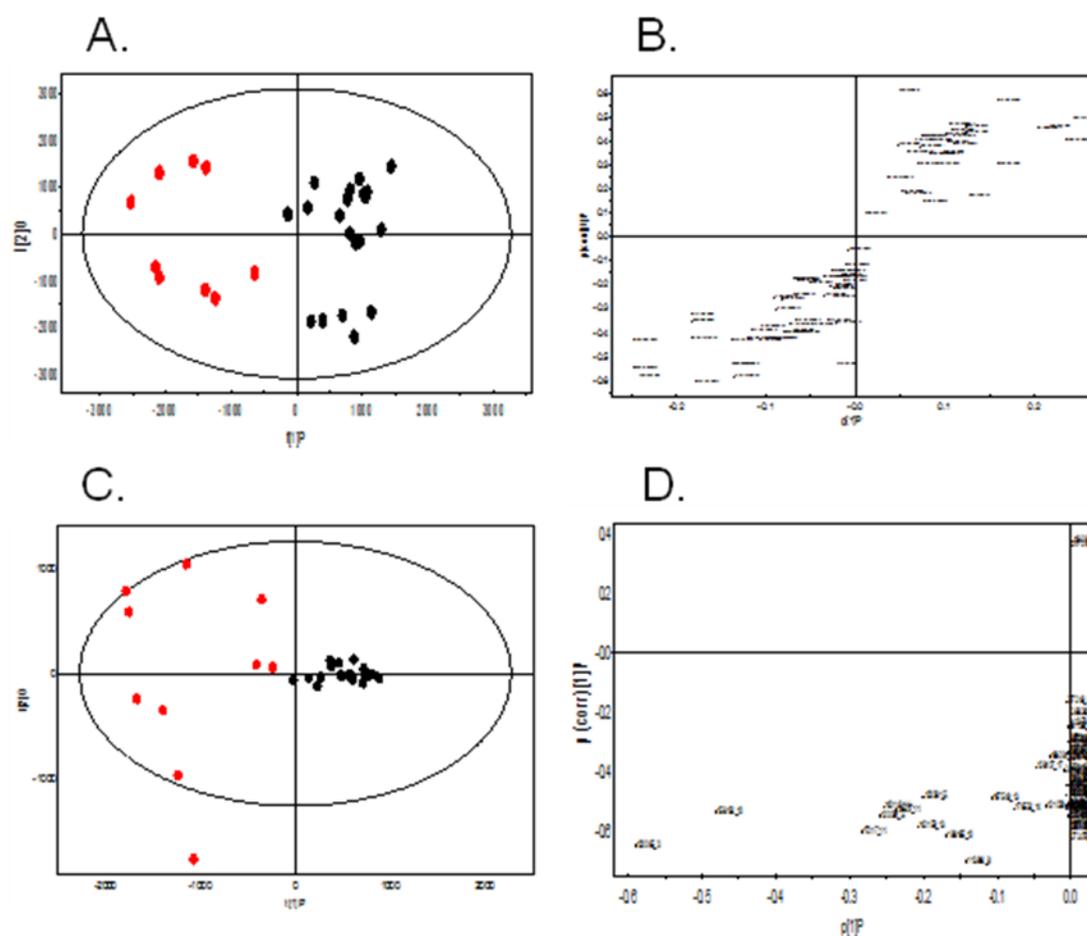
Thus, permutation testing was performed allowing of 100 random permutations in order to verify the prediction accuracy during statistic test. Permutation testing demonstrated that the goodness of fit and predictive ability ( $R^2/Q^2$ ) of the original models was higher than those of the permuted models in both ion modes (Figure 4).



**Figure 4** Statistical validation of the PLS-DA model by permutation analysis using 100 random permutations in the negative ion mode (A) and in the positive ion mode

(B) employing UHPLC-HRMS (Orbitrap) analysis. The goodness of fit ( $R^2$ ) and predictive capability ( $Q^2$ ) of the original models are indicated on the far right and remain higher than those of the 100 permuted models to the left

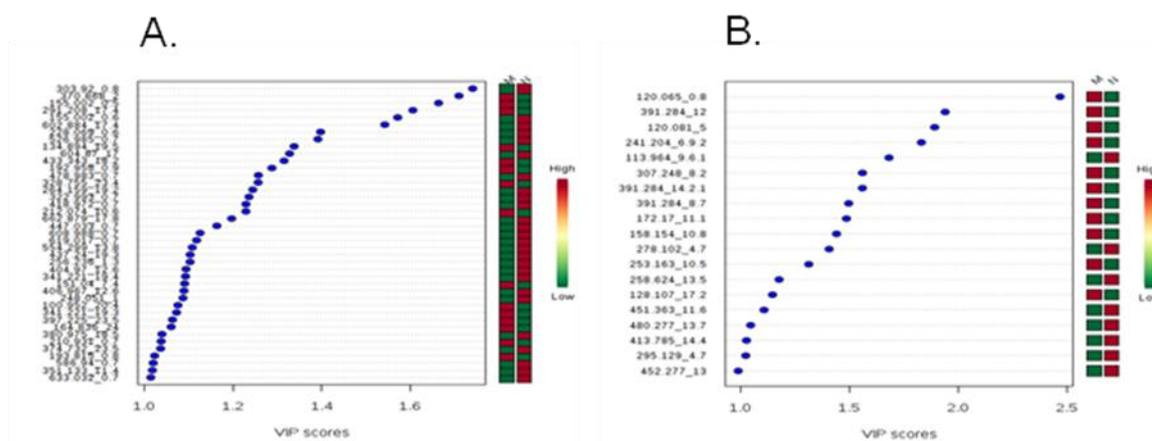
OPLS models indicated that the differences in metabolic components between the two groups after the administration with naringin, in the negative and positive ion modes were significant, since the treated group was clearly clustered separately from the untreated group. Additionally, S-plot was employed to visualize the variables that were responsible for the discrimination. Variables that are distant away from the origin and close to the vertical axis of S-plot are responsible for the clustering (Figure 5).



**Figure 5** Score plot from OPLS (A) and S-plot (B) based on the 75 VIPs employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. The  $R^2$ ,  $Q^2$  values were

0.323, 0.864 and 0.623 respectively. Score plot from OPLS (C) and S-plot (D) based on the 60 VIPs employing UHPLC-HRMS (Orbitrap) analysis, in the positive ion mode. The  $R^2$ ,  $Q^2$  values were 0.92, 0.81 and 0.681 respectively. Naringin administration is represented by black dots and red dots depict controls. Variables that are distant away from the origin and close to the vertical axis of S-plot are responsible for the clustering

The variable importance in projection (VIP) scoring was used to rank the variables in order to detect the ones that contribute to the separation between the groups.  $VIP > 1.0$  was considered sufficient for group discrimination [1]. Totally, 41 variables were detected in the negative ion mode and 19 variables in the positive ion mode (Figure 6).



**Figure 6** 41 variables in the negative ion mode (A) and 19 variables in the positive ion mode (B), ranked by VIP scoring greater than 1 from the PLS-DA models after the analysis of chicken plasma samples administrated with naringin and control plasma samples. Red color indicates the increased levels of the corresponding variables between the two groups, whereas green color indicates the decreased levels

### 3.2.1 Verification of the detected variables ranked by VIP scoring of the chicken plasma samples after dietary supplementation with naringin

The variables (accurate mass\_retention time) that were detected according to the VIP scoring, were further verified based on 3 filter criteria: A. a peak width of at least seven scans, since the Orbitrap Velos has in average scan time of 1 sec, the resulting peak width has an approximate value of 7 sec, B. mass tolerance lower than 5 ppm and C. signal-to-noise (S/N) ratio of at least 10. For the 41 variables in the negative ion mode and 19 in the positive ion mode, only 4 variables fulfilled those criteria in the negative ion mode (Table 1) and 1 variable in the positive ion mode (Table 2).

<b>Accurate mass (m/z)</b>	<b>Retention Time (min)</b>	<b>VIP value</b>
151.0400	7.4	1.124
404.9700	12.6	1.042
554.2994	13.8	1.039
526.9852	0.7	1.306

**Table 1** Table with the variables (accurate mass and retention time) that discriminated the control group from the group that was administrated with naringin, employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. The VIP values were calculated according to the PLS-DA model

<b>Accurate mass (m/z)</b>	<b>Retention Time (min)</b>	<b>VIP value</b>
120.065	0.8	1.178

**Table 2** Table with the variable (accurate mass and retention time) that discriminated the control group from the group that was administrated with naringin employing UHPLC-HRMS (Orbitrap) analysis, in the positive ion mode. The VIP value was calculated according to the PLS-DA model

### 3.2.2 Statistical evaluation of the detected variables ranked by VIP scoring of the chicken plasma samples after dietary supplementation with naringin

The detected variables, according to the VIP scoring, were evaluated according to their distribution (normally distributed data or non-normally distributed data) in the control group and the group after naringin dietary supplementation.

For the variable 151.0400\_7.4 the data, on the control and the group after naringin administration, were normally distributed according to the Lilliefors test in which the parameters of the distribution, the mean and the variance are not known and have to be estimated (Tables 3, 4).

#### Lilliefors test

D	0.120
D (standardized)	0.361
p-value	0.967
alpha	0.05

#### Test interpretation:

**H<sub>0</sub>:** The variable from which the sample was extracted follows a Normal distribution.

**H<sub>a</sub>:** The variable from which the sample was extracted does not follow a Normal distribution.

As the computed p-value is greater than the significance level  $\alpha=0.05$ , one cannot reject the null hypothesis H<sub>0</sub>.

The risk to reject the null hypothesis H<sub>0</sub> while it is true is 97.71%.

**Table 3** The Lilliefors test, is a normality test in which the parameters of the distribution, the mean and the variance are not known and have to be estimated. According to the results the data of the variable 151.0400\_7.4 follows normal distribution on the control group employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode

### Lilliefors test

D	0.173
D (standardized)	0.775
p-value	0.119
alpha	0.05

### Test interpretation:

**H<sub>0</sub>:** The variable from which the sample was extracted follows a Normal distribution.

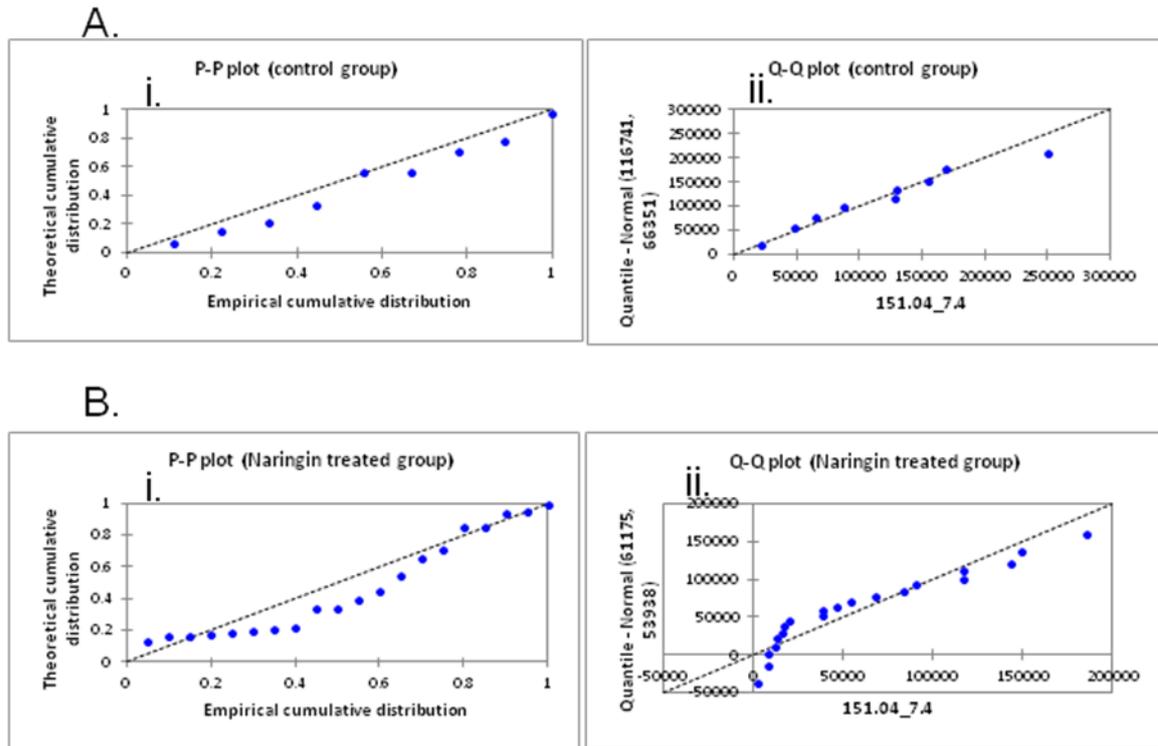
**H<sub>a</sub>:** The variable from which the sample was extracted does not follow a Normal distribution.

As the computed p-value is greater than the significance level  $\alpha=0.05$ , one cannot reject the null hypothesis H<sub>0</sub>.

The risk to reject the null hypothesis H<sub>0</sub> while it is true is 11.93%.

**Table 4 The** Lilliefors test, is a normality test where the parameters of the distribution, the mean and the variance are not known and have to be estimated. According to the results, the data of the variable 151.0400\_7.4 follows normal distribution on the group that was administrated with naringin employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode

In order to confirm by visual inspection if the variable 151.0400\_7.4 follows a normal distribution, P-P plots and Q-Q plots were calculated for the control group (Figure 7A) and the treated group (Figure 7B), in the negative ion mode.



**Figure 7** P-P (Probability-Probability) plots (Ai, Bi) and Q-Q (Quantile-Quantile) plots (Aii, Bii) for the control group and the group after naringin administration, of the variable 151.0400\_7.4 were used to compare the empirical distribution function and the quantities of the variable with that of a sample distributed according to a normal distribution of the same mean and variance. The variable follows a normal distribution since the points lie along the first bisector of the plan

For the variable 404.97\_12.6 the data follows a normal distribution only on the control group according to the Lilliefors test (Tables 5, 6).

**Lilliefors test**

D	0.251
D (standardized)	0.754
p-value	0.103
alpha	0.05

**Test interpretation:**

---

**H0:** The variable from which the sample was extracted follows a Normal distribution.

**Ha:** The variable from which the sample was extracted does not follow a Normal distribution.

As the computed p-value is greater than the significance level  $\alpha=0.05$ , one cannot reject the null hypothesis H0.

The risk to reject the null hypothesis H0 while it is true is 10.33%.

**Table 5** The Lilliefors test, is a normality test where the parameters of the distribution, the mean and the variance are not known and have to be estimated. According to the results the data of the variable 404.97\_12.6 follows a normal distribution on the control group employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode

**Lilliefors test**

---

D	0.229
D (standardized)	1.024
p-value	0.007
alpha	0.05

---

**Test interpretation:**

---

**H0:** The variable from which the sample was extracted follows a Normal distribution.

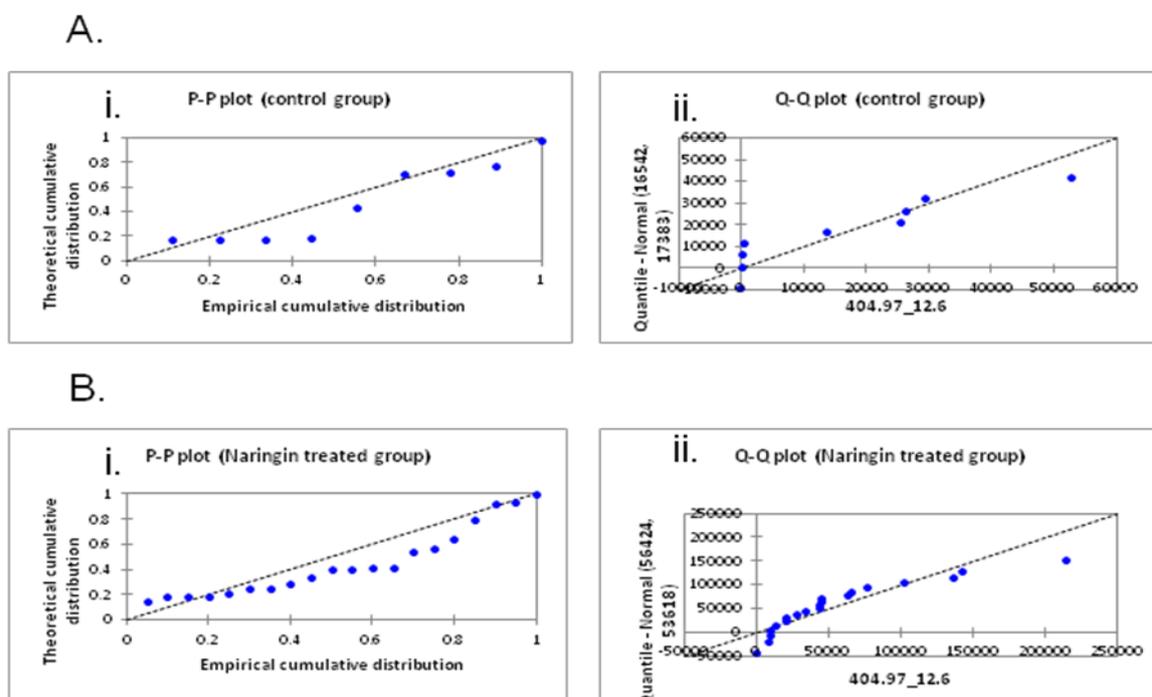
**Ha:** The variable from which the sample was extracted does not follow a Normal distribution.

As the computed p-value is lower than the significance level  $\alpha=0.05$ , one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 0.73%.

**Table 6** The Lilliefors test, is a normality test where the parameters of the distribution, the mean and the variance are not known and have to be estimated. According to the results the data of the variable 404.97\_12.6 were non-normally distributed on the group after dietary supplementation with naringin employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode

P-P plots and Q-Q plots were calculated for the variable 404.9700\_12.6 to visualize the normal distribution of the data in treated and untreated groups. The data in the treated group were non-normally distributed (Figure 8).



**Figure 8** P-P (Probability-Probability) plots (Ai, Bi) and Q-Q (Quantile-Quantile) plots (Aii, Bii) for the control group and the group after naringin administration of the variable 404.9700\_12.6 were used to compare the empirical distribution function and the quantities of the variable with that of a sample distributed according to a normal distribution of the same mean and variance. The variable follows a normal distribution only in the control group since the points lie along the first bisector of the plan

The data of the variable 554.2994\_13.8 follows a normal distribution only on the group after naringin supplementation according to the Lilliefors test (Tables 7, 8).

### Lilliefors test

---

D	0.289
D (standardized)	0.866
p-value	0.030
alpha	0.05

---

### Test interpretation:

---

**H0:** The variable from which the sample was extracted follows a Normal distribution.

**Ha:** The variable from which the sample was extracted does not follow a Normal distribution.

As the computed p-value is lower than the significance level  $\alpha=0.05$ , one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is 2.96%.

**Table 7** The Lilliefors test, is a normality test where the parameters of the distribution, the mean and the variance are not known and have to be estimated. According to the results the data of the variable 554.2994\_13.8 were non-normally distributed on the control group employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode

### Lilliefors test

---

D	0.136
D (standardized)	0.610
p-value	0.426
alpha	0.05

---

### Test interpretation:

---

**H0:** The variable from which the sample was extracted follows a Normal distribution.

**Ha:** The variable from which the sample was extracted does not follow a Normal

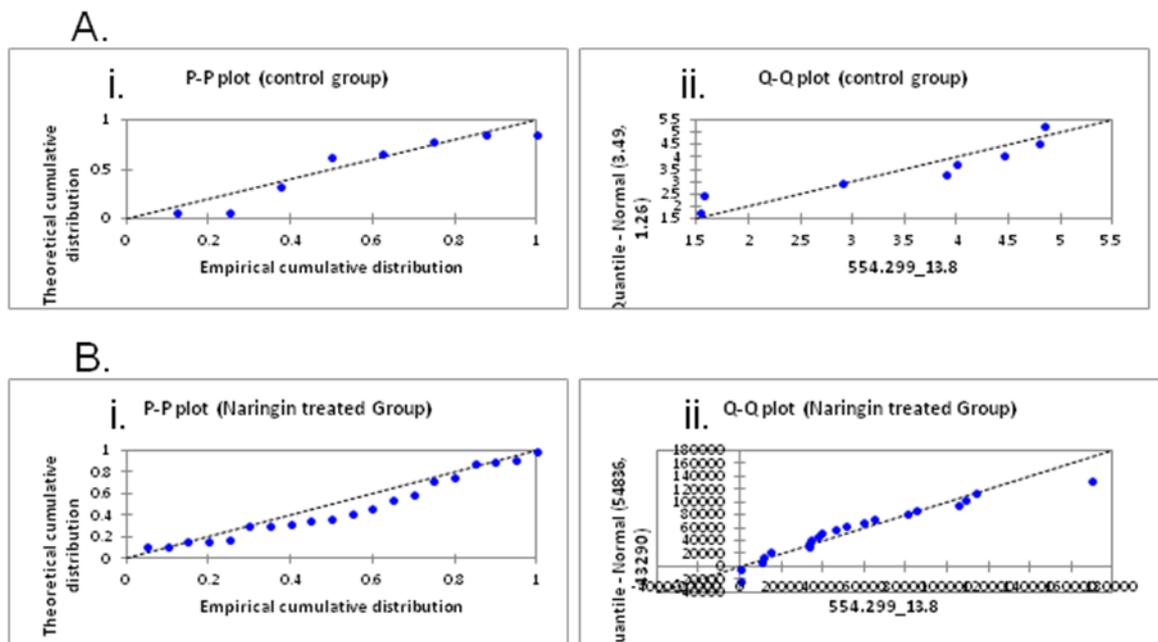
distribution.

As the computed p-value is greater than the significance level  $\alpha=0.05$ , one cannot reject the null hypothesis  $H_0$ .

The risk to reject the null hypothesis  $H_0$  while it is true is 42.62%.

**Table 8** The Lilliefors test, is a normality test where the parameters of the distribution, the mean and the variance are not known and have to be estimated. According to the results the data of the variable 554.2994\_13.8 were normally distributed on the group after naringin administration employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode

Figure 9 showed the P-P plots and Q-Q plots for the data of the variable 554.2994\_13.8. The data in the control group were non-normally distributed.



**Figure 9** P-P (Probability-Probability) plots (Ai, Bi) and Q-Q (Quantile-Quantile) plots (Aii, Bii) for the control group and the group after naringin administration of the variable 554.2994\_13.8 were used to compare the empirical distribution function and the quantities of the variable with that of a sample distributed according to a normal distribution of the same mean and variance. The variable follows a normal distribution only on the treated group since the points lie along the first bisector of the plan.

The data for the variable 526.9852\_0.7 follows a normal distribution only on the group after naringin supplementation according to the Lilliefors test (Tables 9, 10).

**Lilliefors test**

D	0.353
D (standardized)	1.060
p-value	0.002
alpha	0.05

**Test interpretation:**

**H0:** The variable from which the sample was extracted follows a Normal distribution.

**Ha:** The variable from which the sample was extracted does not follow a Normal distribution.

As the computed p-value is lower than the significance level  $\alpha=0.05$ , one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is 0.19%.

**Table 9** The Lilliefors test, is a normality test where the parameters of the distribution, the mean and the variance are not known and have to be estimated. According to the results the data of the variable 526.9852\_0.7 were non-normally distributed on the control group employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode

**Lilliefors test**

D	0.180
D (standardized)	0.804
p-value	0.089
alpha	0.05

## Test interpretation:

**H0:** The variable from which the sample was extracted follows a Normal distribution.

**Ha:** The variable from which the sample was extracted does not follow a Normal distribution.

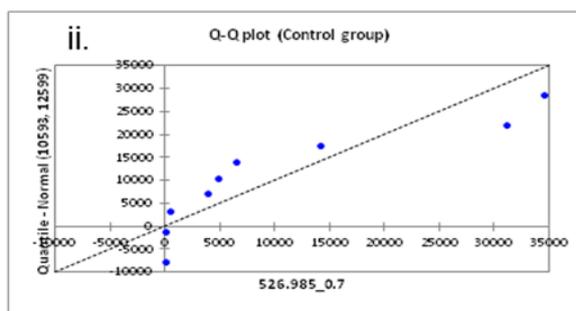
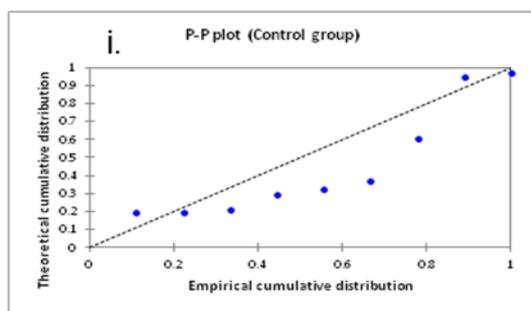
As the computed p-value is greater than the significance level  $\alpha=0.05$ , one cannot reject the null hypothesis H0.

The risk to reject the null hypothesis H0 while it is true is 8.93%.

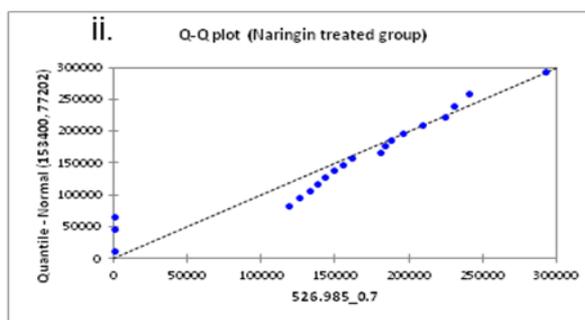
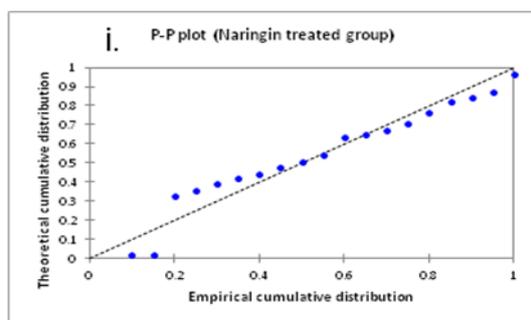
**Table 10** The Lilliefors test, is a normality test where the parameters of the distribution, the mean and the variance are not known and have to be estimated. According to the results the data of the variable 526.9852\_0.7 were normally distributed on the group after naringin administration employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode

P-P plots and Q-Q plots of the variable 526.9852\_0.7 are presented in figure 10. The data in the control group were non-normally distributed.

A.



B.



**Figure 10** P-P (Probability-Probability) plots (Ai, Bi) and Q-Q (Quantile-Quantile) plots (Aii, Bii) for the control group and the group after naringin administration of the variable 526.9852\_0.7 were used to compare the empirical distribution function and the quantities of the variable with that of a sample distributed according to a normal distribution of the same mean and variance. The variable follows a normal distribution only on the treated group since the points lie along the first bisector of the plan

The data for the variable 120.065\_0.8 were non-normally distributed on control group and on the group after naringin supplementation, in the positive ion mode according to the Lilliefors test (Tables 11, 12).

**Lilliefors test**

D	0.322
D (standardized)	1.020
p-value	0.004
alpha	0.05

**Test interpretation:**

**H0:** The variable from which the sample was extracted follows a Normal distribution.

**Ha:** The variable from which the sample was extracted does not follow a Normal distribution.

As the computed p-value is lower than the significance level  $\alpha=0.05$ , one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is 0.40%.

**Table 11** The Lilliefors test, is a normality test where the parameters of the distribution, the mean and the variance are not known and have to be estimated. According to the results the data of the variable 120.065\_0.8 were non-normally distributed on the control group employing UHPLC-HRMS (Orbitrap) analysis, in the positive ion mode

### Lilliefors test

---

D	0.514
D (standardized)	2.355
p-value	< 0.0001
alpha	0.05

---

### Test interpretation:

---

**H0:** The variable from which the sample was extracted follows a Normal distribution.

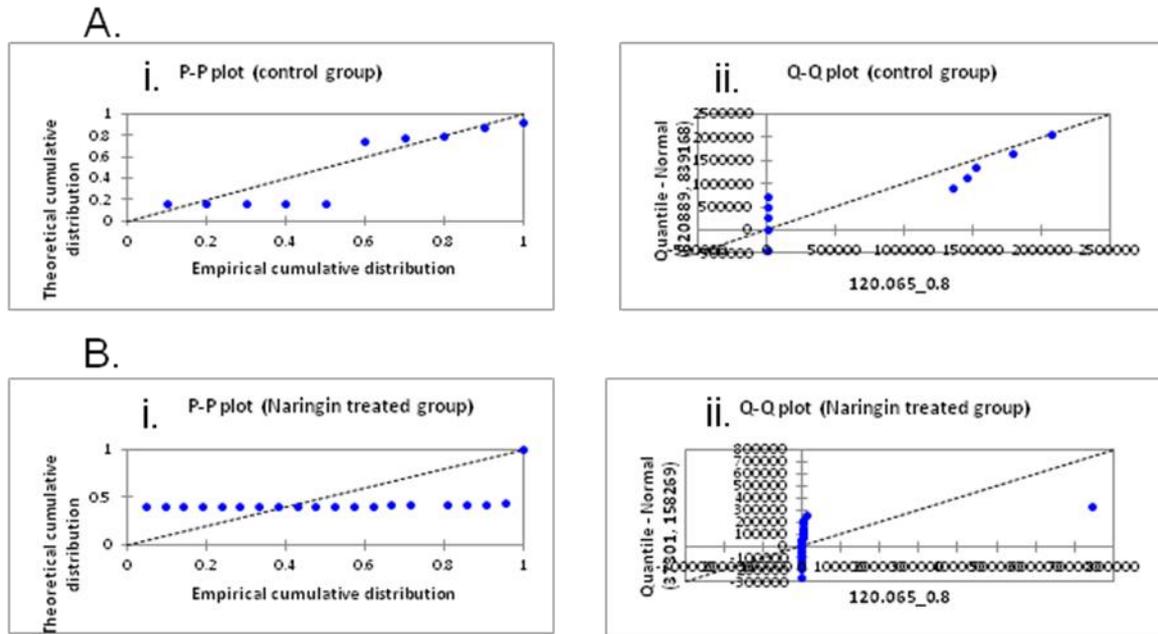
**Ha:** The variable from which the sample was extracted does not follow a Normal distribution.

As the computed p-value is greater than the significance level  $\alpha=0.05$ , one cannot reject the null hypothesis H0.

The risk to reject the null hypothesis H0 while it is true is 0.01%.

**Table 12** The Lilliefors test, is a normality test where the parameters of the distribution, the mean and the variance are not known and have to be estimated. According to the results the data of the variable 120.065\_0.8 were non-normally distributed on the treated group employing UHPLC-HRMS (Orbitrap) analysis, in the positive ion mode

P-P plots and Q-Q plots of the variable 120.065\_0.8 are presented in figure 11. The data in the control group and the group after naringin administration were non-normally distributed.



**Figure 11** P-P (Probability-Probability) plots (Ai, Bi) and Q-Q (Quantile-Quantile) plots (Aii, Bii) for the control group and the group after naringin administration of the variable 120.065\_0.8 were used to compare the empirical distribution function and the quantities of the variable with that of a sample distributed according to a normal distribution of the same mean and variance. The data on the control group and the group after naringin administration were non-normally distributed since the points did not lie along the first bisector of the plan

### 3.2.3 Univariate statistical analysis of the detected variables ranked by VIP scoring of the chicken plasma samples after dietary supplementation with naringin

The unpaired t-test employing the uneven variance approach was applied to the variables that were detected according to the VIP scoring, in order to confirm their statistical significance.  $P \leq 0.05$  was considered to indicate a statistically significant difference between the control group and the group after dietary administration with naringin. Nevertheless, the t-test was performed only for the variable 151.0400\_7.4 as the data were normally distributed on control group and the group after dietary naringin administration (Table 13).

<b>Accurate mass (m/z)</b>	<b>Retention Time (min)</b>	<b>p value</b>
151.0400	7.4	0.0281

**Table 13** Table for the variable 151.0400\_7.4 that discriminated the control group from the group that was administrated with naringin according to the VIP scoring, employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. T-test result confirmed statistically the significance of the variable

For the variables 404.9700\_12.6, 554.2994\_13.8 and 526.9852\_0.7 in the negative ion mode and the variable 120.065\_0.8 in the positive ion mode, the data were non-normally distributed. As a result the equivalent tool Mann-Whitney test was used to confirm the results from the VIP scoring. According to the results the variables on the control group were significant different from the group after dietary administration with naringin (Tables 14-17).

**Mann-Whitney test / Two-tailed test:**

---

U	43.000
Expected value	90.000
Variance (U)	450.000
p-value (Two-tailed)	0.028
alpha	0.05

**Test interpretation:**

---

**H<sub>0</sub>:** The difference of location between the samples is equal to 0.

**H<sub>a</sub>:** The difference of location between the samples is different from 0.

As the computed p-value is lower than the significance level  $\alpha=0.05$ , one should reject the null hypothesis H<sub>0</sub>, and accept the alternative hypothesis H<sub>a</sub>.

The risk to reject the null hypothesis  $H_0$  while it is true is lower than 2.84%.

**Table 14** The Mann-Whitney test was employed for the variable 404.9700\_12.6 to compare the control group and the group after naringin administration employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. According to the results, the null hypothesis of equality was rejected

**Mann-Whitney test / Two-tailed test:**

---

U	43.000
Expected value	90.000
Variance (U)	450.000
p-value (Two-tailed)	0.028
alpha	0.05

**Test interpretation:**

---

**$H_0$ :** The difference of location between the samples is equal to 0.

**$H_a$ :** The difference of location between the samples is different from 0.

As the computed p-value is lower than the significance level  $\alpha=0.05$ , one should reject the null hypothesis  $H_0$ , and accept the alternative hypothesis  $H_a$ .

The risk to reject the null hypothesis  $H_0$  while it is true is lower than 2.84%.

**Table 15** The Mann-Whitney test was employed for the variable 554.2994\_13.8 to compare the control group and the group after naringin administration. According to the results, the null hypothesis of equality was rejected

**Mann-Whitney test / Two-tailed test:**

---

U	40.000
Expected value	90.000
Variance (U)	446.121
p-value (Two-tailed)	0.019
alpha	0.05

**Test interpretation:**

---

**H0:** The difference of location between the samples is equal to 0.

**Ha:** The difference of location between the samples is different from 0.

As the computed p-value is lower than the significance level  $\alpha=0.05$ , one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 1.91%.

**Table 16** The Mann-Whitney test was employed for the variable 526.9852\_0.7 to compare the control group and the group after naringin administration employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. According to the results, the null hypothesis of equality was rejected

**Mann-Whitney test / Two-tailed test:**

---

U	167.000
Expected value	105.000
Variance (U)	559.887
p-value (Two-tailed)	0.009
alpha	0.05

**Test interpretation:**

---

**H0:** The difference of location between the samples is equal to 0.

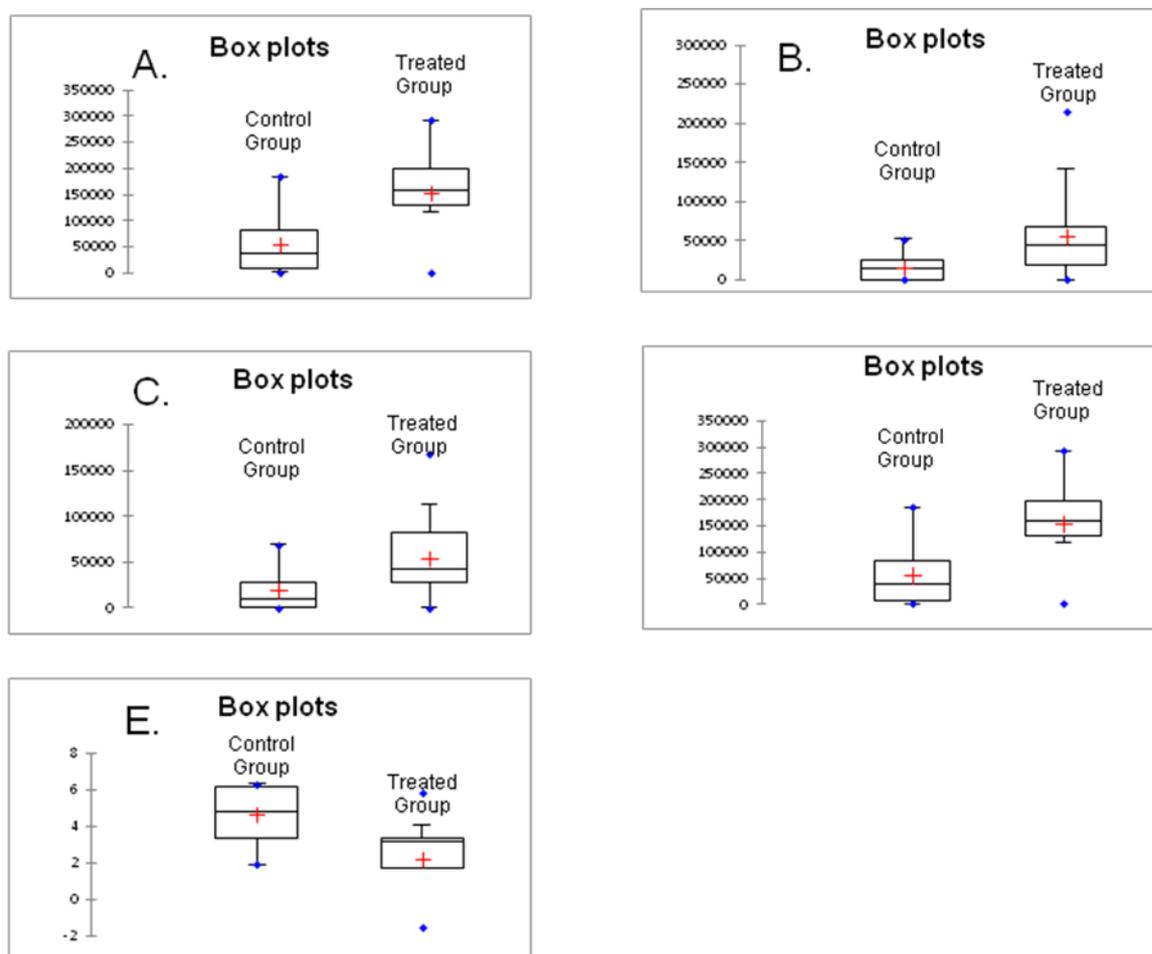
**Ha:** The difference of location between the samples is different from 0.

As the computed p-value is lower than the significance level  $\alpha=0.05$ , one should reject the null hypothesis H0, and accept the alternative hypothesis Ha.

The risk to reject the null hypothesis H0 while it is true is lower than 0.93%

**Table 17** The Mann-Whitney test employed for the variable 120.065\_0.8 to compare the control group and the group after naringin administration, employing UHPLC-HRMS (Orbitrap) analysis, in the positive ion mode. According to the results, the null hypothesis of equality was rejected

The discriminant variables 151.0400\_7.4 (A), 404.9700\_12.6 (B), 554.2994\_13.8 (C) and 526.9852\_0.7 (D) were increased in chicken plasma samples after the administration, in the negative ion mode. The variable 120.065\_0.8 (E) was decreased in chicken plasma samples after naringin administration, in the positive ion mode (Figure 12).



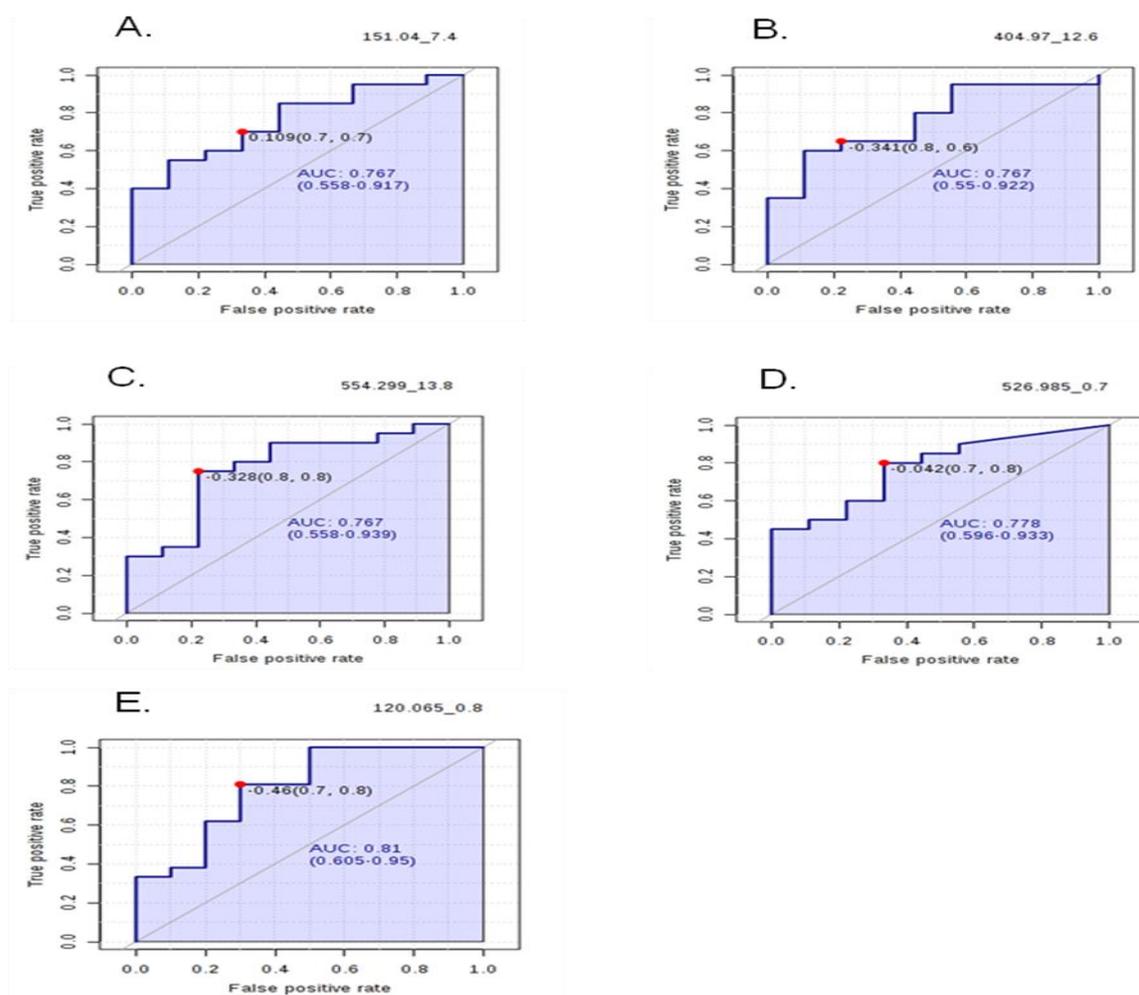
**Figure 12** Box plots of the variables 151.0400\_7.4 (A), 404.9700\_12.6 (B), 554.299\_13.8 (C), 526.985\_0.7 (D) and 120.065\_0.8 (E) derived from the VIP scoring and the univariate statistical analysis between the control group and the

group after naringin administration, employing UHPLC-HRMS (Orbitrap) analysis. The variables 151.0400\_7.4 (A) 404.9700\_12.6 (B), 554.2994\_13.8 (C) and 526.9852\_0.7 (D) were increased in chicken plasma samples after the administration, in the negative ion mode. The variable 120.065\_0.8 (E) was decreased in chicken plasma samples after naringin administration, in the positive ion mode. The red crosses correspond to the means. The central horizontal bars are the medians. The lower and upper limits of the box are the first and third quartiles, respectively. Points are minimum and maximum for each group

### **3.2.4 Evaluation of the detected variables of the chicken plasma samples after dietary supplementation with naringin**

For a metabolomic analysis, the evaluation of the variables that are discovered to discriminate groups is fundamental. Receiver operating characteristic (ROC) curve and the area under the ROC curve (AUC) provide an effective measure of accuracy. AUC value from ROC is usually between 1.0 and 0.5. An AUC value of 0.5 or less for a variable indicates discrimination by chance, whereas an AUC value of 1.0 indicates that the variable is ideal for the differentiation between groups. AUC value of 0 means incorrect classification[9].

ROC analysis was performed to evaluate the efficacy of the variables to discriminate naringin supplement and control group, in both ion modes. ROC curves for the 5 detected variables were calculated using the online software ROCCT. As showed in figure 13 the variables 151.0400\_7.4 (A), 404.9700\_12.6 (B), 554.299\_13.8 (C) and 526.985\_0.7 (D) had an acceptable AUC value of 0.767, 0.767, 0.767 and 0.778 respectively, in the negative ion mode. In the positive ion mode the variable 120.065\_0.8 (E) showed an acceptable AUC value of 0.81. These results indicate that the detected significant variables in both ion modes could be used to discriminate control group from the group after naringin dietary supplementation in chickens.



**Figure 13** ROC curves were calculated using the online software ROCCEt for the 5 variables that were detected according to the VIP scoring and the uneven variance approach to discriminate naringin group from control group employing UHPLC-HRMS (Orbitrap) analysis. The variables 151.0400\_7.4 (A), 404.9700\_12.6 (B), 554.299\_13.8 (C) and 526.985\_0.7 (D) had an acceptable AUC value of 0.767, 0.767, 0.767 and 0.778 respectively, in the negative ion mode. The variable 120.065\_0.8 (E) had an acceptable AUC value of 0.81 in the positive ion mode

### 3.2.5 Identification of the detected variables of the chicken plasma samples after dietary supplementation with naringin

For the prediction of the molecular formulas of the detected variables various software tools were employed. For the Mzmine software the element count heuristics (H/C ratio, NOPS/C ratios, Multiple element counts), RDBE restrictions (RDBE range from -1 to 40) and isotope pattern filter (isotope  $m/z$

tolerance: 0.001  $m/z$  or 5 ppm, minimum absolute intensity:  $10^3$ , minimum score: 90%) were taken under consideration. For the Sirius software  $m/z$  tolerance of 5 ppm was applied. The Rdisop package in R language using the accurate masses and the relative intensities of the first and second isotopes were used. Furthermore, the mMass software was employed using 10 ppm tolerance and included the composition rules of H/C, NOPS/C, RDBE and the isotopic pattern. The adduct calculator from the seven golden rules [14] and the mutual mass table differences for adducts [15] were used to identify adducts. The online databases Metlin, Chempider, Kegg and LIPID MAPS were used to recognize possible metabolites.

### 3.2.5.1 Identification of the detected variable 151.0400\_7.4 after dietary supplementation with naringin, in the negative ion mode

Comparing the results from the processing procedures for the variable 151.0400\_7.4 one possible molecular formula was identified (Table 18).

<b>Accurate mass</b>	<b>Molecular Formula</b>	<b><math>\Delta m</math> (ppm)</b>
151.0400	$C_8H_8O_3$	0

**Table 18** For the variable 151.0400\_7.4, one possible molecular formula was identified according to the software Mzmine, Sirius, mMass and the R package Rdisop

According to the KEGG databases 22 possible identifications were detected. Nevertheless, the 4-Hydroxyphenylacetate with mass different 0 ppm and isotopic pattern score 99% was considered as the possible identification of the detected variable. According to Kim et al. [13] naringin is metabolized to its corresponding aglycone at early time and then to 4-hydroxyphenylacetic acid.

### 3.2.5.2 Identification of the detected variable 404.9700\_12.6 after dietary supplementation with naringin, in the negative ion mode

For the variable 404.9700\_12.6 the Pearson' s correlation coefficient score was calculated using the Matlab software [15]. According to the results the variable had a high correlation score (score=1) with the variable 406.967\_12.6. The  $m/z$  value and the  $m/z$  difference between the peaks were compared with characteristic fragments, adducts, neutral losses and modifications [15]. A match was found from their  $m/z$  different ( $m/z$  1.997) to correspond to [M+Br] and [M+2K-H]. Consequently, the detected variable 404.9700\_12.6 was considered as adduct [M+2K-H] of the variable  $m/z$  328.0492. According to the Metlin database, cAMP [M=329.0525] was considered as the possible identification since the transition ion of the cAMP which was provided by Metlin applying collision energy (CE) of 10 V [M=134.0453], was identified on the chromatogram ( $\Delta$ ppm =13) as a low energy fragmentation product (pseudo MS/MS) (Table 19).

<b>Accurate mass</b>	<b>Molecular Formula</b>	<b><math>\Delta</math>m (ppm)</b>
328.0492	C <sub>10</sub> H <sub>12</sub> N <sub>5</sub> O <sub>6</sub> P	11

**Table 19** The variable 404.9700 after the Pearson' s correlation coefficient computation and comparison with characteristic adducts was identified as adduct [M+2K-H] of the variable  $m/z$  328.0492. According to Metlin database the cAMP was considered as the possible identification

cAMP is a second messenger that plays a fundamental role in many biological processes. Frequently, membrane-associated adenylyl cyclases (Acs) generate the diffusible second messenger cAMP in response to activation of multiple G-protein coupled receptors. Subsequently activation of protein kinase A (PKA) by cAMP leads to the phosphorylation of numerous downstream targets [18].

### 3.2.5.3 Identification of the detected variable 554.2994\_13.8 after dietary supplementation with naringin, in the negative ion mode

For the variable 554.2994\_13.8 the mMass, the Sirius and the Rdisop software proposed the molecular formula  $C_{30}H_{37}N_9O_2$  (Table 20).

<b>Accurate mass</b>	<b>Molecular Formula</b>	<b><math>\Delta m</math> (ppm)</b>
554.2994_13.8	$C_{30}H_{37}N_9O_2$	11

**Table 20** For the variable 554.2994\_13.8 one possible molecular formula was identified according to the software Mzmine, Sirius, mMass and the R package Rdisop

According to the chemspider network database, 11 possible identifications were detected. However, these identifications there were not assigned to any biochemical molecule.

Using the adduct calculator from the seven golden rules [14] we concluded that the variable  $m/z$  554.2994\_13.8 was adduct [M+Cl] of the  $m/z$  519.3299 (Table 21).

<b>Accurate mass</b>	<b>Molecular Formula</b>	<b><math>\Delta m</math> (ppm)</b>
519.3299	$C_{26}H_{50}NO_7P$	5

**Table 21** The variable 554.2994\_13.8 was considered as adduct [M+Cl] of the  $m/z$  519.3299 according to the adduct calculator of the seven golden rules

From the LIPID Metabolites and Pathways Strategy (LIPID MAPS) database the variable  $m/z$  519.3299 was identified as a glycerophospholipid (PC(18:2)) [M=519.3325] since its fragment 2,4-octadecadienoic acid [M=279.2330] was detected on the chromatogram ( $\Delta ppm=7$ ) (Figure 24).

Glycerophospholipids are one of the major components of cellular membranes. They are synthesized from glycerol-3-phosphate (G3P) by a de novo pathway that initially produces phosphatidic acid (PA) and diacylglycerol

(DAG) or cytidine diphosphate-DAG (CDP-DAG) [4]. Glycerophospholipids provide the membrane with a suitable environment, fluidity and ion permeability. The most abundant glycerophospholipids of mammalian tissues are phosphatidylcholine (PtdCho), phosphatidylethanolamine (PtdEtn), phosphatidylserine (PtdSer) and phosphatidylinositol (PtdIns) [12]. Once glycerophospholipids organized in bilayers they undergo reactions and activities from phospholipases that are responsible for the turnover, compositional maintenance and rearrangements in membranes. These processes results in the modulation of membrane function. Furthermore, glycerophospholipids belong to the signal transduction network as precursors for lipid mediators. The latter possess a fundamental role in internal and external communication transferring signals from the cell surface to the interior, modulating gene expression, growth, differentiation, adhesion, migration, apoptosis and ion channels [12, 5].

#### 3.2.5.4 Identification of the detected variable 526.9852\_0.7 after dietary supplementation with naringin, in the negative ion mode

For the variable 526.9852\_0.7 the Mzmine, mMass, Sirius and the Rdisop software proposed the molecular formula  $C_{21}H_{87}N_2O_{15}$  (Table 22).

<b>Accurate mass</b>	<b>Molecular Formula</b>	<b><math>\Delta m</math> (ppm)</b>
526.9852_0.7	$C_{21}H_{87}N_2O_{15}$	0

**Table 22** For the variable 526.9852\_0.7 one possible molecular formula was identified according to the software Mzmine, Sirius, mMass and the R package Rdisop

Nevertheless, no recognition was possible.

### 3.2.5.5 Identification of the detected variable 120.065\_0.8 after dietary supplementation with naringin, in the positive ion mode

For the 120.065\_0.8 variable in the positive ion mode, Sirius, mMass and Rdisop proposed the molecular formula C<sub>4</sub>H<sub>9</sub>NO<sub>3</sub> (Table 23).

<i>Accurate mass</i>	<b>Molecular Formula</b>	<b>Δm (ppm)</b>
120.065_0.8	C <sub>4</sub> H <sub>9</sub> NO <sub>3</sub>	11

**Table 23** For the variable 120.065\_0.8 one possible molecular formula was identified according to the software Sirius, mMass and the R package Rdisop

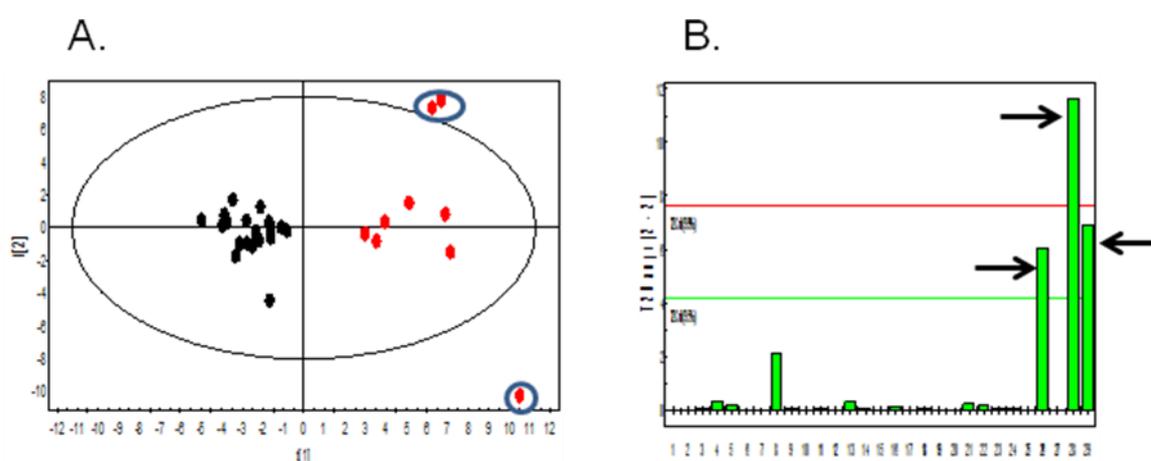
According to Metlin database, threonine [M= 119.0582] was considered as the possible identification since its fragment (CE=0) [M=102.0545] was detected on the chromatogram (Δppm=0).

Threonine is an essential amino acid which is produced by microbial synthesis using glucose or sucrose as substrates. It supports cardiovascular, liver, central nervous, and immune system function. Threonine is needed to create glycine and serine, two amino acids that are necessary for the production of collagen, elastin, and muscle tissue. It combines with the amino acids aspartic acid and methionine to help the liver with lipotropic function or the digestion of fats and fatty acids. Therefore, threonine deficiency could cause liver failure [16].

### 3.3 Multivariate statistical analysis of the plasma samples after dietary supplementation with hesperidin

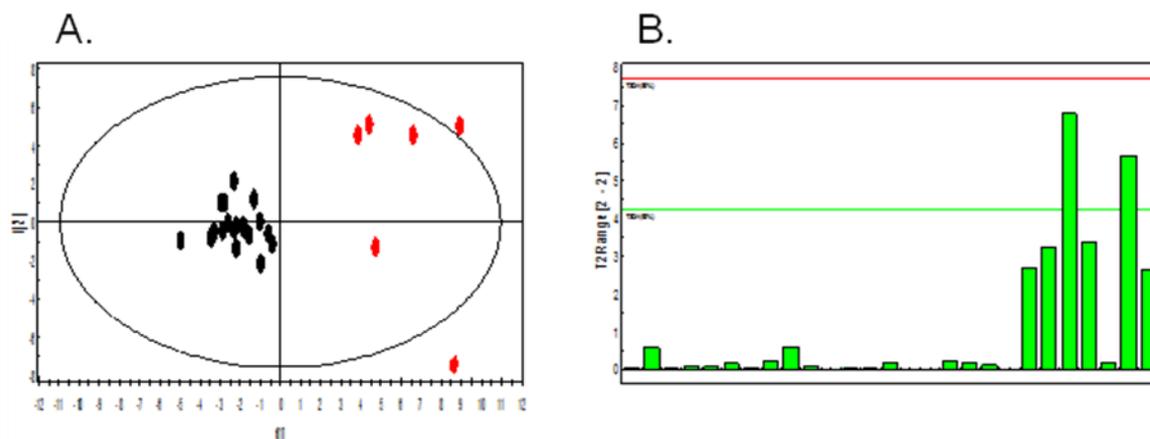
MVA was performed in the same manner as in the case of plasma samples after dietary supplementation with naringin. According to the VIP scoring (VIP>1), 68 variables in the negative mode and 60 variables in the positive mode were detected and merged together in a reduced matrix for the analysis.

In the negative mode, for the MVA analysis different parameters had been employed, such as different scaling e.g. UV or Pareto, and different combinations of  $Q^2$  and  $R^2$  to determine the number of factors that described the model without increasing the complexity. The UV scaling method was selected for PCA and PLS-DA analysis as it afforded tighter groups with larger separation. As figure 15 showed, a clear separation between the two groups is observed even in the unsupervised analysis (PCA). However, 3 outliers were observed which were confirmed by the Hotelling's T2Range plot (Figure 15B).



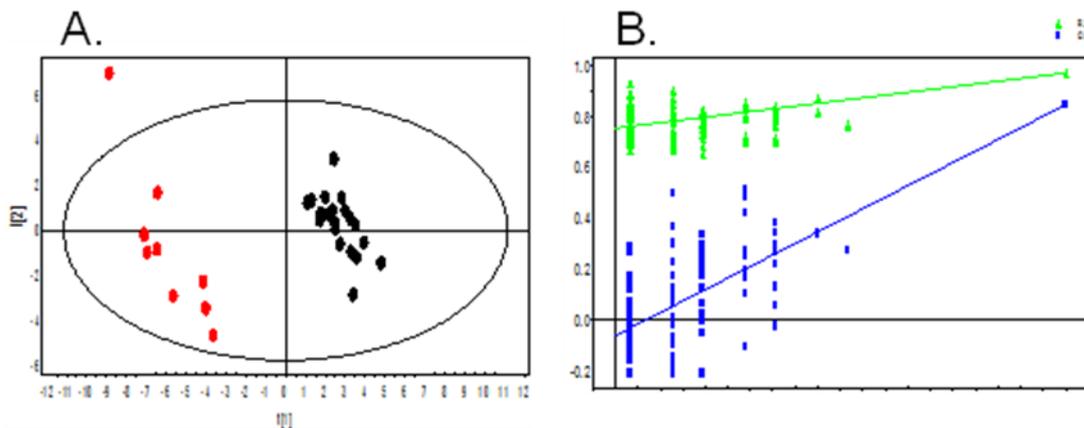
**Figure 15** Score plot from PCA (A) and Hotelling's T2 range plot (B) based on the 68 VIPs, after hesperidin administration employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. Hesperidin administration is represented by black dots and red dots depict controls. The  $R^2$ ,  $Q^2$  values for the PCA analysis were 0.305 and 0.0398 respectively, employing 2 principal components. According to the T2 range plot, which displays the distance from the origin in the model plane (score space) for each selected observation, one control sample was considered as outlier since its value was larger than the red limit (0.01 levels) and two control samples were considered as suspected since their values were larger than the green limit (0.05 levels). Blue circles on PLS-DA plot indicate the selective observations

The PCA model was re-calculated and the control samples were not included in the new workset (Figure 16A). According to the Hotelling's T2 range plot no values larger than the red limit (0.01 levels) were observed (Figure 16B).



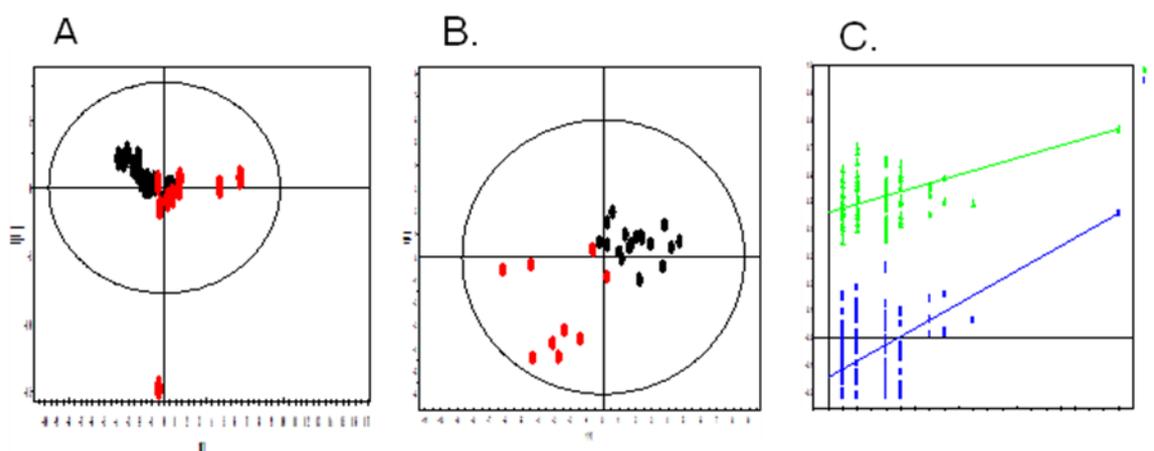
**Figure 16** Score plot from PCA (A) and Hotelling's T2 range plot (B) based on the 68 VIPs, after hesperidin administration employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. The control samples which were considered as outliers were not included in the new workset. Hesperidin administration is represented by black dots and red dots depict controls. The  $R^2$ ,  $Q^2$  values for the PCA analysis were 0.28 and 0.0125 respectively, employing 2 principal 151omponents. According to the T2 range plot, which displays the distance from the origin in the model plane (score space) for each selected observation, no values larger than the red limit (0.01 levels) were observed

The  $R^2$ ,  $Q^2$  values for the PLS-DA methodology were found to be 0.274, 0.971 and 0.845 respectively, indicating that the variance between the two groups could be explained by this model. Thus, according to the permutation testing the classification model was accurate and reliable (Figure 17).



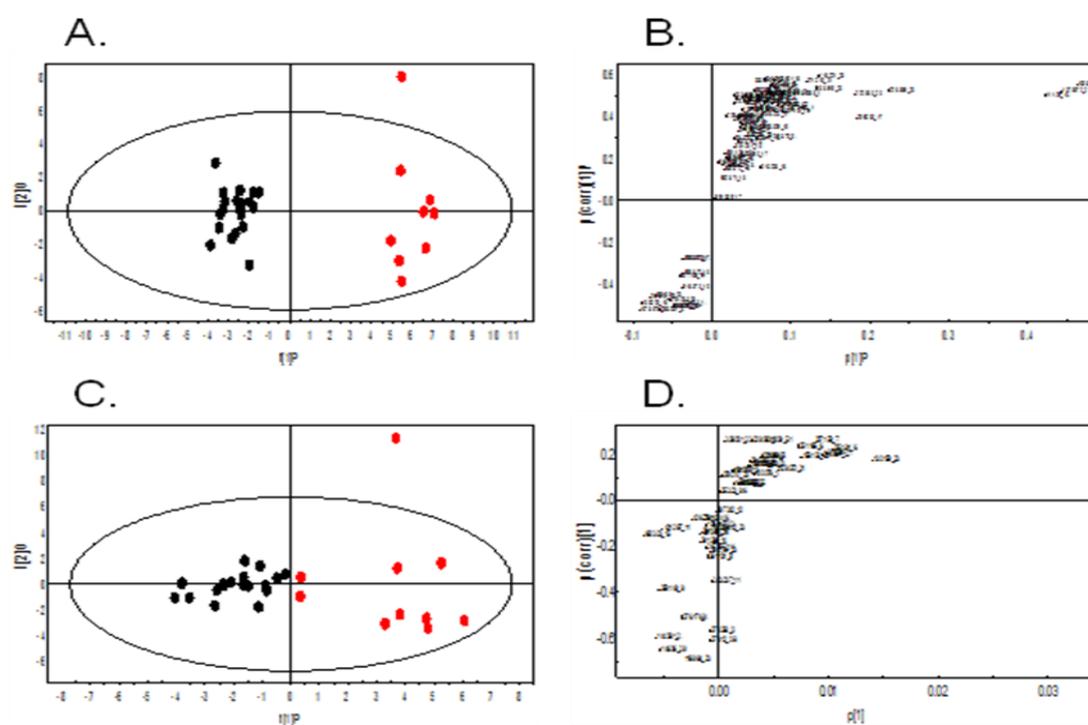
**Figure 17** Score plot from PLS-DA (A) and permutation testing (B) after the processing procedure of plasma samples from chickens that were administrated with hesperidin (black dots) and control samples (red dots) employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. A clear separation between samples of the two groups is observed. The  $R^2$ ,  $Q^2$  values for the PLS-DA methodology were 0.274, 0.971 and 0.845 respectively, employing 2 principal components. The  $R^2$  and  $Q^2$  values of the permutation testing were decreased upon 100 permutations which demonstrated that the model was not by chance

Additionally, in the positive ion mode the UV scaling method was selected. As figure 19 showed a clear separation between samples of the two groups is observed only by PLS-DA. Thus the permutation testing illustrated that the classification model was accurate and reliable (Figure 18).



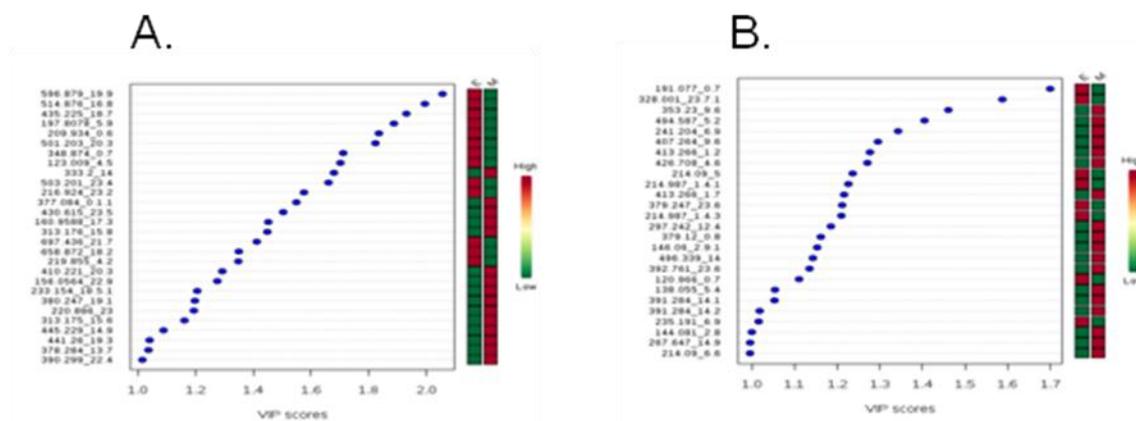
**Figure 18** Score plots from PCA (A) and PLS-DA (B) and permutation testing (C) after the processing procedure of plasma samples from chickens that were administrated with hesperidin (black dots) and control samples (red dots) employing UHPLC-HRMS (Orbitrap) analysis, in the positive ion mode. A clear separation between samples of the two groups is observed on PLS-DA methodology. The  $R^2$ ,  $Q^2$  values were for the PCA methodology 0.488 and -0.00237 respectively, and for the PLS-DA methodology 0.323, 0.769 and 0.459 respectively, employing 2 principal components. Permutation testing allowing of 100 random permutations of the PLS-DA model demonstrated that the model was not by chance

OPLS models and S-plots were employed in both ion modes to indicate the differences (Figure 19).



**Figure 19** Score plots from OPLS (A) and S-plot (B) based on the 68 VIPs employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. The  $R^2$ ,  $Q^2$  values were 0.274, 0.971 and 0.864 respectively. Score plots from OPLS-DA (C) and S plot (D) based on the 60 VIPs employing UHPLC-HRMS (Orbitrap) analysis, in the positive ion mode. The  $R^2$ ,  $Q^2$  values were 0.323, 0.769 and 0.409 respectively. Naringin administration is represented by black dots and red dots depict controls. Variables that are distant away from the origin and close to the vertical axis of S-plot are responsible for the clustering

The variables that discriminated the control group from the group after hesperidin administration were defined by VIP scoring of the PLS-DA model.  $VIP > 1.0$  was set as the threshold [1]. Generally, 28 variables in the negative ion mode and 26 variables in the positive ion mode were ranked with a VIP scoring higher than 1 (Figure 20).



**Figure 20** 28 variables in the negative ion mode (A) and 26 variables in the positive ion mode (B), ranked by VIP scoring greater than 1 from the PLS-DA model after the analysis of chicken plasma samples after hesperidin supplementation employing UHPLC-HRMS (Orbitrap) analysis. Red color indicates the increased levels of the corresponding variables between the two groups, whereas green color indicates the decreased levels

### 3.3.1 Verification of the detected variables ranked by VIP scoring of the chicken plasma samples after dietary supplementation with hesperidin

The variables that were ranked by VIP scoring higher than 1 from the PLS-DA models were further evaluated, as was mentioned above, according to the peak width, the S/N ratio, and the mass tolerance. From the 28 variables in the negative mode and 26 in the positive mode, only 1 variable in the negative mode fulfilled those criteria (Table 24).

<b>Accurate mass (m/z)</b>	<b>Retention Time (min)</b>	<b>VIP value</b>
279.2317	17.2	2.4606

**Table 24** Table with the variable (accurate mass and retention time) that discriminated the control group from the group that was administrated with hesperidin, employing UHPLC-ESI-HRMS (Orbitrap) analysis, in the negative ion mode. The VIP value was calculated according to the PLS-DA model

### 3.3.2 Statistical evaluation of the detected variable of the chicken plasma samples after dietary supplementation with hesperidin

The data of the variable 279.2317\_17.2 follows a normal distribution according to the Lilliefors test on control group and on the group after hesperidin administration (Tables 25, 26).

#### Lilliefors test

D	0.175
D (standardized)	0.525
p-value	0.591
alpha	0.05

#### Test interpretation:

**H0:** The variable from which the sample was extracted follows a Normal distribution.

**Ha:** The variable from which the sample was extracted does not follow a Normal distribution.

As the computed p-value is greater than the significance level  $\alpha=0.05$ , one cannot reject the null hypothesis H0.

The risk to reject the null hypothesis H0 while it is true is 59.08%.

**Table 25** The Lilliefors test, is a normality test where the parameters of the distribution, the mean and the variance are not known and have to be estimated. According to the results for the control group, the data of the variable 279.2317\_17.2 follows a normal distribution employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode

**Lilliefors test**

D	0.108
D (standardized)	0.481
p-value	0.788
alpha	0.05

**Test interpretation:**

**H0:** The variable from which the sample was extracted follows a Normal distribution.

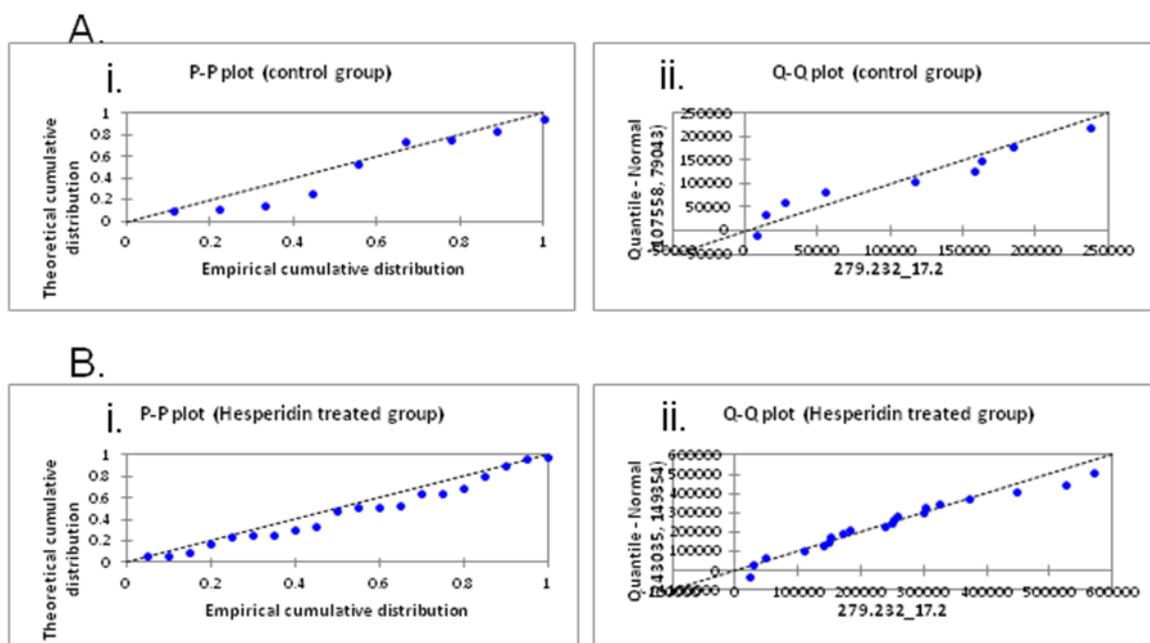
**Ha:** The variable from which the sample was extracted does not follow a Normal distribution.

As the computed p-value is greater than the significance level  $\alpha=0.05$ , one cannot reject the null hypothesis H0.

The risk to reject the null hypothesis H0 while it is true is 78.76%.

**Table 26** The Lilliefors test, is a normality test where the parameters of the distribution, the mean and the variance are not known and have to be estimated. According to the results for the group, which was administrated with hesperidin, the data of the variable 279.2317\_17.2 follows a normal distribution employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode

P-P plots and Q-Q plots of the variable 279.2317\_17.2 were calculated for the control group (Figure 21A) and the treated group (Figure 21B).



**Figure 21** P-P (Probability-Probability) plots (Ai, Bi) and Q-Q (Quantile-Quantile) plots (Aii, Bii) for the control group and the group after hesperidin administration of the variable 279.2317\_17.2 were used to compare the empirical distribution function and the quantities of the variable with that of a sample distributed according to a normal distribution of the same mean and variance. The variable follows a normal distribution since the points lie along the first bisector of the plan

### 3.3.3 Univariate statistical analysis of the detected variable of the chicken plasma samples after dietary supplementation with hesperidin

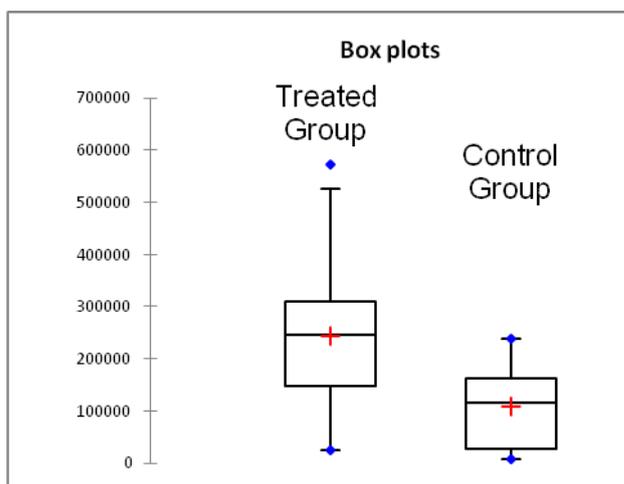
The unpaired t-test for the variable 279.2317\_17.2 was employed between the control group and the group after hesperidin dietary supplementation. The result confirmed the statistical significance of the variable (Table 27).

<i>Accurate mass (m/z)</i>	<i>Retention Time (min)</i>	<i>p value</i>
279.2317	17.2	0.0199

**Table 27** Table with the variable (accurate mass and retention time) that discriminated the control group from the group that was administrated with hesperidin

according to the VIP scoring, employing UHPLC-HRMS (Orbitrap) analysis in the negative ion mode. T-test result confirmed statistically the significance of the variable

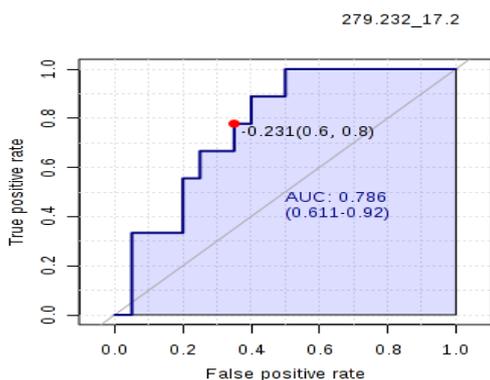
The discriminant variable 279.2317\_17.2 was increased in chicken plasma after hesperidin administration, in the negative ion mode (Figure 22).



**Figure 22** Box plot of the variable 279.2317\_17.2 derived from the VIP scoring and the uneven variance approach between the treated and the untreated groups in chicken plasma after hesperidin supplementation employing UHPLC-HRMS (Orbitrap), in the negative ion mode. The variable was increased in plasma after hesperidin administration. The red crosses correspond to the means. The central horizontal bars are the medians. The lower and upper limits of the box are the first and third quartiles, respectively. Points are minimum and maximum for each group

### 3.3.4 Evaluation of the detected variables of the chicken plasma samples after dietary supplementation with hesperidin

ROC curves were calculated for the accuracy assessment of the discovered variable using the online software ROCCET. The classical univariate ROC curve analysis was applied. The ROC curve for the variable 279.2317\_17.2 yielded an AUC value of 0.786 demonstrated highly reliability for discriminating control plasma samples from plasma samples after dietary supplementation with hesperidin (Figure 23).



**Figure 23** ROC curve was calculated using the online software ROC CET for the variable that was detected according to the VIP scoring and the uneven variance approach to discriminate hesperidin group from control group employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode. The variable 279.2317\_17.2 had an acceptable AUC value of 0.786

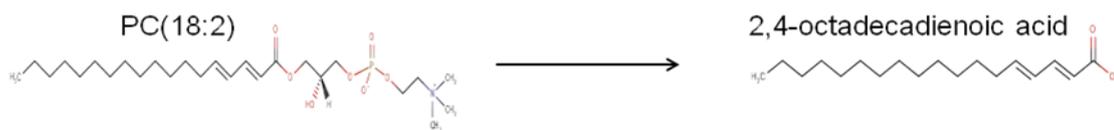
### 3.3.5 Identification of the variable 279.2317\_17.2 after dietary supplementation with hesperidin, in the negative ion mode

For the identification of the detected variable 279.2317\_17.2 the software Mzmine, Sirius, Rdisop and mMass were employed using the same parameters as was mentioned above. According to the Mzmine, Sirius and mMass software one possible molecular formula was identified (Table 28).

<i>Accurate mass</i>	<b>Molecular Formula</b>	<b><math>\Delta m</math> (ppm)</b>
279.2317_17.2	$C_{18}H_{32}O_2$	0

**Table 28** For the variable 279.2317\_17.2 one possible molecular formula was identified according to the software Mzmine, Sirius and mMass

According to the LIPID Metabolites and Pathways Strategy (LIPID MAPS) the possible identification was the 2,4-octadecadienoic acid [M= 280.2402]. The 2,4-octadecadienoic acid could be produced from the combined acid hydrolysis with enzymatic release of glycerophosphocholine PC(18:2) (Figure 24).



**Figure 24** The 2,4-octadecadienoic acid could be produced from the combined acid hydrolysis with enzymatic release of glycerophosphocholine PC(18:2)

#### 4. Concluding remarks

A UHPLC-HRMS-based plasma metabolomics study was performed after dietary supplementation of naringin and hesperidin in chickens. Shredding the LC-MS chromatograms into multiple  $m/z$  ranges and merging together in a reduced matrix the variables with VIP scoring greater than 1, four variables in the negative ion mode and one variable in the positive ion mode, were detected to discriminate naringin treated group from control group. According to multiple software and online databases the variables were identified as 4-hydroxyphenylacetate, cAMP, glycerophospholipid and an unidentified variable (526.9852\_0.7), in the negative ion mode and threonine in the positive ion mode. Furthermore, 2,4-octadecadienoic acid was identified to discriminate hesperidin treated group from control group, in the negative ion mode. Moreover, the metabolomic study was performed using the wide  $m/z$  range of 100-900 Da. According to the results only 2 metabolites were identified after naringin supplementation in the negative ion mode, which were identified with the crop-filter approach (404.97\_12.6, 151.004\_7.4). To our concern, this is the first time that an untargeted metabolomic analysis using multiple  $m/z$  ranges is being performed in plasma samples after naringin and hesperidin dietary administration.

## References

1. Andersen C.M., Bro R. (2010) Variable selection in regression — a tutorial. *J Chemometr.* 24,728–737.
2. Chanet A., Milenkovic D., Deval C., Potier M., Constans J., Mazur A., Bennetau-Pelissero C., Morand C., Bérard AM. (2012) Naringin, the major grapefruit flavonoid, specifically affects atherosclerosis development in diet-induced hypercholesterolemia in mice. *J Nutr Biochem.* 23, 469-477.
3. Chanet A., Milenkovic D., Manach C., Mazur A., Morand C. (2012) Citrus flavanones: What is their role in cardiovascular protection? *J Agric Food Chem.* 60, 8809–8822.
4. Farooqui AA, Horrocks LA, Farooqui T. (2000) Glycerophospholipids in brain: their metabolism, incorporation into membranes, functions, and involvement in neurological disorders. *Chem Phys Lipids.* 106, 1-29.
5. Farooqui AA, Ong WY, Farooqui T. (2010) Lipid mediators in the nucleus: their potential contribution to Alzheimer's disease. *Biochim Biophys Acta.* 1801, 906–916.
6. Franke A.A., Cooney R.V., Henning S.M., Custer L.J. (2005) Bioavailability and antioxidant effects of orange juice components in humans. *J Agric Food Chem.* 53, 5170–5178.
7. Garg A., Garg S., Zaneveld L.J., Singla A.K. (2001) Chemistry and pharmacology of the Citrus bioflavonoid hesperidin. *Phytother Res.* 15, 655-669.
8. Gika H.G., Theodoridis G.A., Plumb R.S., Wilson I.D. (2014) Current practice of liquid chromatography-mass spectrometry in metabolomics and metabonomics. *J Pharm Biomed Anal.* 87, 12-25.

9. Hajian – Tilaki K. (2013) Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian J Intern Med.* 2, 627-635.
10. Heuberger A.L., Robison F.M., Lyons S.M., Broeckling C.D., Prenni J.E. (2014) Evaluating plant immunity using mass spectrometry-based metabolomics workflows. *Front. Plant Sci.* 5, 1–11.
11. Jagetia G.C., Reddy T.K. (2002) The grapefruit flavanone naringin protects against the radiation – induced genomic instability in the mice bone marrow: a micronucleus study. *Mutat Res.* 519, 37-48.
12. Kennedy E. P. (1956) The biological synthesis of phospholipids. *Can. J. Biochem. Physiol.* 34, 334–348.
13. Kim DH, Jung EA, Sohng IS, Han JA, Kim TH, Han MJ. Mariana C. Sousa, Rodolpho C. Braga, Bertilha A.S. Cintra, Valéria de Oliveira, Carolina H. Andrade. (1998) Intestinal bacterial metabolism of flavonoids and its relation to some biological activities. In silico metabolism studies of dietary flavonoids by CYP1A2 and CYP2C9. *Arch Pharm Res.* 21, 17-23.
14. Kind T, Fiehn O. (2007) Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics.* 8, 105.
15. Lynn KS, Cheng ML, Chen YR, Hsu C, Chen A, Lih TM, Chang HY, Huang CJ, Shiao MS, Pan WH, Sung TY, Hsu WL. (2015) Metabolite identification for mass spectrometrybased metabolomics using multiple types of correlated ion information. *Anal Chem.* 87, 2143-2151.
16. Mumby MC, Walter G. (1993) Protein serine/threonine phosphatases: structure, regulation, and functions in cell growth. *Physiol Rev.* 73, 673-699.

17. Roohbakhsh A., Parhiz H., Soltani F., Rezaee R., Iranshahi M. (2015) Molecular mechanisms behind the biological effects of hesperidin and hesperetin for the prevention of cancer and cardiovascular diseases. *Life Sci.* 124, 64-74.
18. Roskoski R Jr. (2015) A historical overview of protein kinases and their targeted small molecule inhibitors. *Pharmacol Res.* 100, 1-23.
19. Shiota C., Abe T., Kawai N., Ohno A., Teshima – Kondo S., Mori H., Terao J., Tanaka E., Nikawa T. (2015) Flavones Inhibit LPS-Induced Atrogin-1/MAFbx Expression in Mouse C2C12 Skeletal Myotubes. *J Nutr Sci Vitaminol.* 61,188-194.
20. Stanstrup J., Gerlich M., Dragsted L.O., Neumann S. (2013) Metabolite profiling and beyond: approaches for the rapid processing and annotation of human blood serum mass spectrometry data. *Anal Bioanal Chem.* 405, 5037-5048.
21. Su Z.H., Li S.Q., Zou G.A., Yu C.Y., Sun Y.G., Zhang H.W., Gu Y., Zou Z.M. (2011) Urinary metabonomics study of anti-depressive effect of Chaihu-Shu-Gan-San on an experimental model of depression induced by chronic variable stress in rats. *J. Pharm. Biomed. Anal.* 55, 533–539.
22. Tsao R. (2010) Chemistry and biochemistry of dietary polyphenols. *Nutrients.* 2, 1231–1246.
23. Viant, M.R., Sommer U. (2013) Mass spectrometry based environmental metabolomics: a primer and review. *Metabolomics.* 9, S144–S158.
24. Xia J., Wishart D.S. (2011) Web-based inference of biological patterns, functions and pathways from metabolomics data using MetaboAnalyst. *Nat Protoc.* 6,743-760.

25. Yao L.H., Jiang Y.M., Tomas-Barberan F.A., Datta N., Singanusong R., Chen S.S. (2004) Flavonoids in food and their health benefits. *Plant Food Hum Nutr.* 59, 113–122.

## Chapter abstract

Naringin and hesperidin are flavonoids with a broad spectrum of biological activities. The aim of this study was to characterize the metabolomic fingerprint of these compounds after dietary supplementation in chickens. Forty nine 1-day-old Ross 308 broiler chickens were randomly divided into 3 groups; 20 chickens, after dietary supplementation with naringin (1.5 g/kg of feed), 20 chickens after dietary supplementation with hesperidin (1.5 g/kg of feed) and 9 control chickens that were given commercial basal diets. Blood samples were collected on 42 days of age after 30 days of the administration. The samples were analyzed by UHPLC-HRMS (Orbitrap Velos). For the processing procedure the data were shred in 100 Da slices generating 8 datasets. Each dataset was treated separately and the  $m/z_{tR}$  features obtained by VIP's values were merged and used as the input for the general models. Multivariate analysis was applied to pinpoint metabolic changes related to the administration. Four variables were detected to discriminate control group from the group after naringin administration. These variables were identified applied various software tools and online databases as 4-hydroxyphenylacetate, cAMP, glycerophospholipids and an unidentified variable 526.9852\_0.7 ( $m/z_{tR}$ ), in the negative ion mode and threonine in the positive ion mode. Furthermore, 2,4-octadecadienoic acid was identified to discriminate hesperidin treated group from control group, in the negative ion mode.

**Keywords:** Mzmine – VIPs – MVA – software – online databases

## Supporting Information for Chapter 5

### Supplementary S1

Parameters	Mass range ( <i>m/z</i> )							
	100-200	200-300	300-400	400-500	500-600	600-700	700-800	800-900
Noise level	10 <sup>4</sup>	2 10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>4</sup>	2.5 10 <sup>4</sup>	2 10 <sup>4</sup>	2.5 10 <sup>4</sup>
Minimum height	2 10 <sup>4</sup>	4 10 <sup>4</sup>	2 10 <sup>4</sup>	2 10 <sup>4</sup>	210 <sup>4</sup>	5 10 <sup>4</sup>	4 10 <sup>4</sup>	5 10 <sup>4</sup>
Chromatograph deconvolution (Algorithm)	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off
Retention time tolerance (min)	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
Alignment	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner

**S1** Parameters that were used for the processing procedure of the data from the group after naringin supplementation and control group employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode using multiple *m/z* mass ranges by the open software Mzmine 2.18.2

## Supplementary S2

Parameters	Mass range ( <i>m/z</i> )							
	100-200	200-300	300-400	400-500	500-600	600-700	700-800	800-900
Noise level	10 <sup>4</sup>	2 10 <sup>4</sup>	3 10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>4</sup>	2.5 10 <sup>4</sup>	2 10 <sup>4</sup>	2.5 10 <sup>4</sup>
Minimum height	2.5 10 <sup>4</sup>	4E <sup>4</sup>	5 10 <sup>4</sup>	2 10 <sup>4</sup>	2 10 <sup>4</sup>	5 10 <sup>4</sup>	4 10 <sup>4</sup>	5 10 <sup>4</sup>
Chromatograph deconvolution (Algorithm)	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off
Retention time tolerance (min)	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
Alignment	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner

**S2** Parameters that were used for the processing procedure of the data from the group after naringin supplementation and control group employing UHPLC-HRMS (Orbitrap) analysis, in the positive ion mode using multiple *m/z* mass ranges by the open software Mzmine 2.18.2

### Supplementary S3

Parameters	Mass range ( <i>m/z</i> )							
	100-200	200-300	300-400	400-500	500-600	600-700	700-800	800-900
Noise level	10 <sup>5</sup>	2 10 <sup>4</sup>	3 10 <sup>4</sup>	2 10 <sup>4</sup>	10 <sup>4</sup>	2.5 10 <sup>4</sup>	10 <sup>4</sup>	2.5 10 <sup>4</sup>
Minimum height	2 10 <sup>5</sup>	4 10 <sup>4</sup>	6 10 <sup>4</sup>	2.5 10 <sup>4</sup>	2 10 <sup>4</sup>	4.5 10 <sup>4</sup>	2 10 <sup>4</sup>	5 10 <sup>4</sup>
Chromatograph deconvolution (Algorithm)	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off
Retention time tolerance (min)	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
Alignment	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner

**S3** Parameters that were used for the processing procedure of the data from the group after hesperidin supplementation and control group employing UHPLC-HRMS (Orbitrap) analysis, in the negative ion mode using multiple *m/z* mass ranges by the open software Mzmine 2.18.2

### Supplementary S4

Parameters	Mass range ( <i>m/z</i> )							
	100-200	200-300	300-400	400-500	500-600	600-700	700-800	800-900
Noise level	10 <sup>5</sup>	2 10 <sup>4</sup>	3.5 10 <sup>4</sup>	10 <sup>4</sup>	2 10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>4</sup>	10 <sup>4</sup>
Minimum height	20 <sup>5</sup>	4 10 <sup>4</sup>	6 10 <sup>4</sup>	2 10 <sup>4</sup>	4 10 <sup>4</sup>	2 10 <sup>4</sup>	2 10 <sup>4</sup>	2 10 <sup>4</sup>
Chromatograph deconvolution (Algorithm)	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off	Baseline cut-off
Retention time tolerance (min)	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
Alignment	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner	Join aligner

**S4** Parameters that were used for the processing procedure of the data from the group after hesperidin supplementation and control group employing UHPLC-HRMS (Orbitrap) analysis, in the positive ion mode using multiple *m/z* mass ranges by the open software Mzmine 2.18.2

## Chapter 6

### **UHPLC-HRMS-based tissue metabolomics study of naringin and hesperidin after dietary supplementation in chickens**

As referred on Chapter 3 a metabolomic study in tissue samples was deemed as necessary in order to discover the possible biochemical alteration caused by naringin and hesperidin dietary supplementation. Therefore, an untargeted metabolomics study has been undertaken in order to complete this task.

## 1. Introduction

Metabolomics is the latest “omics” strategy that offers quantitative and qualitative analysis of endogenous low-molecular-weight compounds (e.g., <1 kDa) in a biological system [13]. Over the past decade, metabolomics has been widely used to describe and study certain pathological and physiological states with the aim of biomarkers identification in biological fluids and tissues with many powerful applications in clinical and nutritional research.

In 1998, the National Institutes of Health Biomarkers Definitions Working Group described a biomarker as “a characteristic that is objectively measured and evaluated as an indicator of normal biological courses, pathogenic progressions, or pharmacologic responses to a therapeutic intervention” [36]. In metabolomics, biomarkers can be measured in any biological sample, e.g., blood, tissue or urine [18, 10].

The use of metabolomics in the identification of biomarkers is based on the hypothesis that abnormalities like diseases, environmental alterations or changes on dietary pattern, cause changes on the biochemical pathways in response, leading to a metabolic profile characteristic of the disease. Identification of biomarkers for a disease diagnosis is one of the most important applications of metabolomics. To date, a significant number of potential metabolite biomarkers have been discovered by profiling the human metabolome [8].

Metabolomics covers a wide range of applications on diverse research areas like drug discovery [15], nutrition [7] and the study of human diseases [14]. Additionally, has been used as a promising tool to discriminate different dietary patterns [28] and also to identify dietary biomarkers [27].

Predominantly, two analytic platforms are used for data acquisition including mass spectrometry (MS) coupled to a separation technique e.g. LC-MS and nuclear magnetic resonance (NMR) spectrometry. These platforms are considered as the most effective techniques in quantitative and qualitative

metabolic analysis. Both techniques have the ultimate goal of identifying small molecules that are responsible for a particular activity or phenotype [25]. The choice of either one of these technologies depends on the sensitivity and selectivity requirements of the study.

Nowadays, a variety of software tools have been developed for the primary processing procedure after the data analysis such as apLCMS, Mzmine, XCMS, MAIT, XCMS online, MetAlign etc [42, 31, 34, 6, 38, 20]. Following data processing, various statistical tools are typically used to identify significant metabolic differences between distinct groups. Finally, the metabolites represented by the selected peaks (accurate mass<sub>tR</sub>) are verified and identified by searching on online databases.

In this chapter, we present an analytical and computational method to characterize the metabolic changes that occur in tissues in response to the administration of naringin and hesperidin in the nutrition of chickens is presented. Naringin is the major flavanone glycoside in grapefruit, citrus fruits and exerts a variety of pharmacological effects such as antioxidant activity, anti-inflammatory, anti-mutagenic and analgesic [29]. Additionally, hesperidin belongs to the flavanone class of flavonoids that is found abundantly in citrus fruits [26] and reported to possess antioxidant and anti-inflammatory activity [12, 43]. To our concern, this is the first metabolomic study in tissue samples after dietary supplementation of hesperidin and naringin in chickens.

Initially, the optimal parameters settings for the XCMS analysis were discovered using the package IPO ('Isotopologue Parameter Optimization') of R language [17]. The open-source package XCMS is widely used for metabolomics analysis. Its general workflow involves peak finding, peaks grouping and retention time alignment [34]. MetaboAnalyst, a web-based platform for metabolomic data processing, was used for the estimation of missing values [40].

After the analysis, the data were normalized according to multiple internal standards in order to remove undesired variation arising from various sources

and to identify only the differentially expressed metabolites. For the normalization a variety of methods were employed. The selection of the normalization method was based on the statistical analysis (multivariate analysis) and the internal validation measures of Connectivity, Silhouette width and Dunn index. The Silhouette index validates the clustering performance based on the pair wise difference between and within-cluster distances, with well-clustered observations having values near 1 and poorly clustered observations having values near -1 [21]. Dunn's index is the ratio of the smallest distance between observations not in the same cluster to the largest intra-cluster distance and has a value between zero and infinity, and should be maximized [5]. Connectivity relates to what extent observations are placed in the same cluster as their nearest neighbors in the data space and has a value between zero and infinity, and should be minimized [9]. The cross-contribution compensating multiple standard normalization method (ccmn) method was chosen to process the data [4]. As Redestig et al. has presented, the ccmn method can correct the cross-contribution between native metabolites and the internal standards and thereby increase the accuracy of the metabolite profiles [32].

Statistical evaluation of the results was performed by multivariate analysis (MVA) e.g. principal component analysis (PCA), projections to latent structures – discriminant analysis (PLS-DA), optimized potentials for liquid simulations (OPLS) and other statistical tools (cat score ranking) in order to discover the variables that contributed the most to the discrimination between the data sets.

For every untargeted metabolomic analysis, validation of the detected variables is required. The ROC curve and the area under the ROC curve (AUC) provide a numerical value of the relationship between the specificity and sensitivity of a biomarker. AUC values from ROC is usually between 1.0 and 0.5. An AUC value of 0.5 or less for a biomarker indicates no information and discrimination within the test, whereas an AUC value of 1.0 indicates a complete separation of two groups, and the samples can be classified with

100% sensitivity (no false negatives) and 100% specificity (no false positives) [30, 24].

After the discovery and the evaluation of the variables that were responsible for the separation between treated and untreated groups, the corresponding accurate mass\_ $t_R$  (min) was used for the structural elucidation in order to provide biological meaning to the results. For the identification various software tools were employed. Moreover, online databases e.g. METLIN ([www.metlin.scripps.edu](http://www.metlin.scripps.edu)), KEGG ([www.genome.jp/kegg](http://www.genome.jp/kegg)), Chempider (<http://www.chemspider.com/>) and LIPID MAPS (<http://www.lipidmaps.org/>) within the mass tolerance of 30 ppm were used to identify possible chemical entities. Additional data were used to confirm the proposal structures such as the computation of the Pearson's correlation coefficient (PCC) of the detected variables according to Matlab program [23], and the results from the package CAMERA of R statistical language which allows the grouping of the detected features based on the calculation of known mass differences, on isotope pattern and adduct detection [16].

## **2. Materials and method**

### **2.1 Chemicals and Reagents**

All solvents were of LC-MS grade. Acetonitrile water and formic acid was purchased from Fluka/Riedel-de Haën (Switzerland). Acetone and chloroform were purchased from Sigma-Aldrich (Steinheim, Germany). Naringin was purchased from Alfa Aesar GmbH & Co KG (Germany) and hesperidin from TSI Europe NV (Belgium).

### **2.2 Animal and samples treatment**

Forty nine 1-day-old Ross 308 broiler chickens were randomly divided into 3 groups. The chickens were obtained from a commercial hatchery and were

reared in pens, of a surface area of 2 m<sup>2</sup> in a controlled environment. The lighting program consisted of 23L:1D on arrival, and was decreased to 18L:6D at day 7, remained constant until day 35, and thereafter gradually increased to 23L:1D at slaughter, with access to feed, in mash form, and water *ad libitum*. The treatment groups were: 17 chickens, after dietary supplementation with naringin (1.5 g/kg of feed), 15 chickens after dietary supplementation with hesperidin (1.5 g/kg of feed) and 17 chickens that were given commercial basal diets (control chickens). The administration of naringin was started from the 11<sup>th</sup> day of age until slaughter at the age of 42 days.

At the age of 42 days, the chickens were sacrificed. Tissue samples were collected and dried by lyophilization. Subsequently the samples were stored at -80°C until the analysis.

All experimentation was carried out in strict accordance with the guidelines of “Council Directive 86/609/EEC regarding the protection of animals used for experimental and other scientific purposes”. The protocol was approved by the Bioethical Committee of the Agricultural University of Athens (Permit Number: 20/20032013).

### **2.3 Sample Preparation**

0.020 gr of lyophilized tissue samples were used for the metabolomic analysis. Briefly, a solution of 0.1M acetic acid-ammonium acetate (pH=5), was added to reach a final ratio of 1:10 tissue/solution (v/v) for the tissue homogenization. The samples were then homogenized for at least 5 min. 1 mL of cold chloroform was added to the homogenizer and the samples were kept at -20°C for 10 min, in order to achieve total lipid extraction. After the removal of the lipid phase, 1 mL of cold acetone was added for the protein precipitation. Extraction was conducted by vortexing for 1 min. The extracts were centrifuged at 13,500 rpm at 4°C for 10 min. After phase separation, the aqueous layer (acetone) was collected and additional of 0.2 mL of cold

acetone was added to the tissue pellet. Samples were vortexed and centrifuged as previously was reported. The supernatants were mixed and evaporated to dryness under vacuum.

The residue was reconstituted with 0.2 mL with the three internal standard mixture (1 µg/mL). For the internal standard mixture the compounds reserpine, 2-aminophenol and yohimbine in methanol/water (50/50, v/v) were used.

A pooled quality control (QC) sample in order to monitor LC and MS performance across sample runs was generated by combining equal small aliquots from all the samples of the experimental set as recommended by Sangster et al. [22]. QC sample was prepared by mixing the supernatants of all the samples before the evaporation procedure. Subsequently the mixture was evaporated to dryness. The QC sample was then used throughout the experiment as a process control.

## **2.4 Instrumentation**

The metabolomic analysis of tissue samples was performed employing an ESI-LTQ-Orbitrap Discovery XL mass spectrometry (Thermo Scientific, Germany) connected to an Accela UHPLC system (Thermo Scientific, Germany). The UHPLC system was equipped by an autosampler, a vacuum degasser, a binary pump and a temperature-controlled column. An ACQUITY UPLC BEH C18 (2.1 x 100 mm, 1.7 µm) reversed phase column (Waters Corp. Milford, MSA, USA) was used for the metabolomic analysis.

The system was run in a binary gradient solvent mode consisting of 0.01% (v/v) formic acid/water (solvent A) and acetonitrile (solvent B). Sample analysis was carried out in both positive (ESI+) and negative (ESI-) ion modes. The flow rate was 0.4 mL/min. A gradient method of 32 minutes was used for the metabolomic analysis as follows: 0 to 24 min: 95% A: 5% B, 24 to 28 min: 5% A: 95% B, 28 to 32 min: 95% A: 5% B. The column temperature

was maintained at 40°C while the autosampler tray temperature was kept at 8°C. The injection volume was 5 µL.

For the positive ion mode, the capillary temperature and voltage was set at 356°C and -60 V, respectively. The sheath gas flow was set to 30 arb. Units and the aux gas flow to 10 arb. Units. The spray voltage was set to 3.5 kV and tube lens voltage to 110 V. For the negative ion mode, the capillary temperature and voltage was set at 356°C and 20 V, respectively. The sheath gas flow was set to 30 arb. Units and the aux gas flow to 10 arb. Units. The spray voltage was set to 3.1 kV and tube lens voltage to -49 V. In both positive and negative ion modes, analysis was performed using the Fourier transform mass spectrometry (FTMS) full scan ion mode, applying a mass scan range of  $m/z$  100-1000 and a resolution of 30,000 FWHM while spectra were acquired in centroid mode.

Plasma samples centrifugation was performed by a Mikro 200R centrifuge (Hettich Lab Technology, Germany). Evaporation of the samples was performed by GeneVac HT-4X EZ-2 series evaporator Lyospeed ENABLED (Genevac Ltd, UK). Tissue homogenization was performed by a Kinematica Polytron PT 1200C homogenizer (Brinkmann, Westbury, NY).

## 2.5 Software analysis

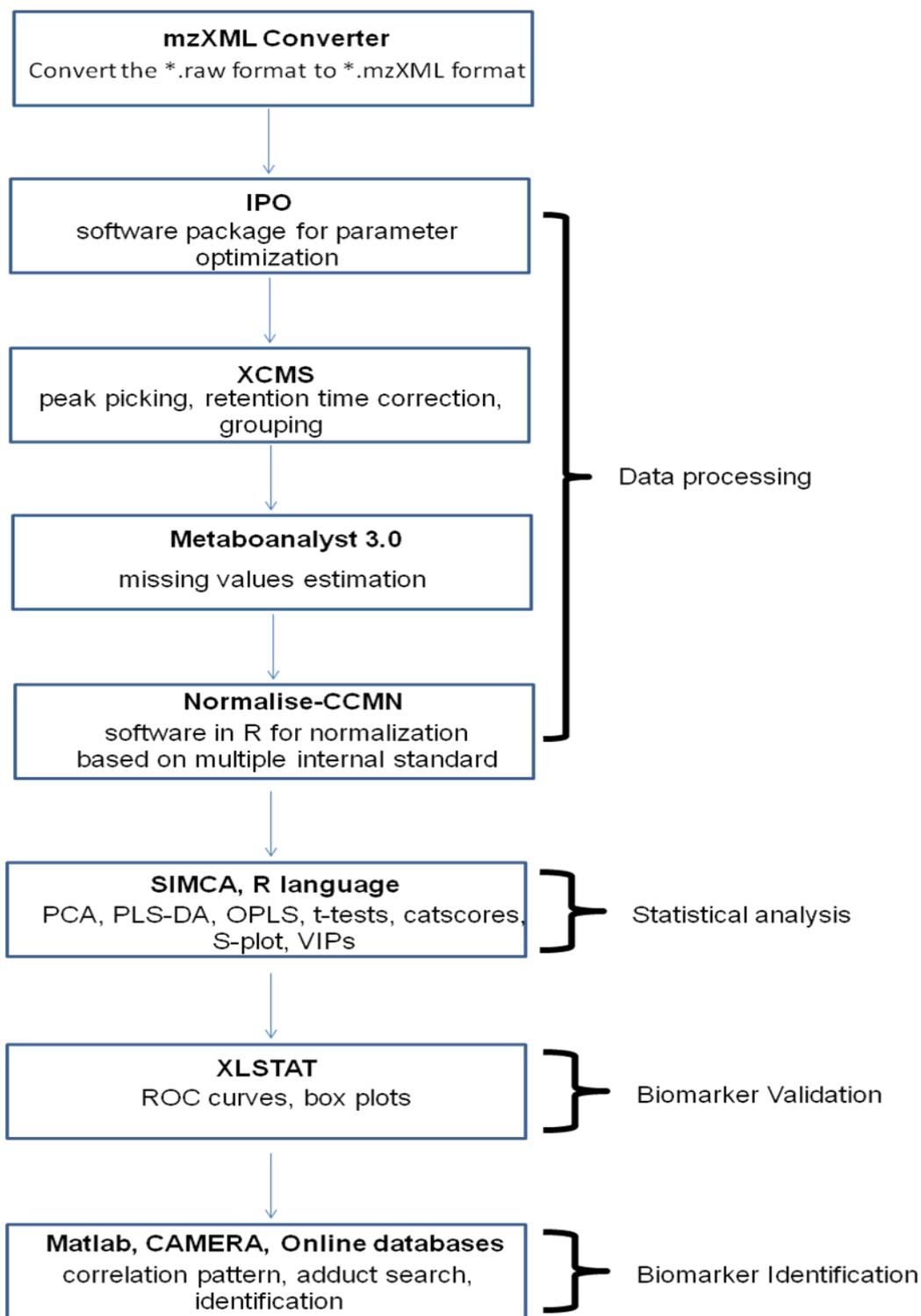
For the processing procedure, the statistical analysis and the evaluation of the data, a variety of software were employed. Xcalibur version 2.1 (Thermo Fisher Scientific), the mzXML converter (<http://proteowizard.sourceforge.net/downloads.shtml>), the web-based platform Metaboanalyst 3.0 (<http://www.metaboanalyst.ca/>), the statistical programming language R using multiple packages and the Matlab software (The MathWorks Inc, Natick, MA, USA) were used for the processing procedure.

For the statistical analysis the XLSTAT 2010 software, SIMCA-P + 11.5 software (Umetrics, Umeå, Sweden) and the R language, were employed.

For the identification of the molecular formulas of the detected variables Rdisop package from the R statistical language (WW), Sirius 3.1 software [2] and mMass 5.5.0 [37] were employed. Online databases such as METLIN ([www.metlin.scripps.edu](http://www.metlin.scripps.edu)), KEGG ([www.genome.jp/kegg](http://www.genome.jp/kegg)), Chempider (<http://www.chemspider.com/>) and Metabolites and Pathways Strategy (LIPID MAPS) (<http://www.lipidmaps.org/>) were used to match the selected variables in terms of mass accuracy to specific metabolites.

## 2.6 Experimental Design

Figure 1, shows the methodological pipeline of the untargeted metabolomic study that was developed for the processing procedure, the statistical analysis of the data and the validation and identification of the detected variables.



**Figure 1** Outline of the methodology developed for the metabolomic analysis in both ion modes

### 3. Results and Discussion

#### 3.1 Data processing

Initially, an mzXML converter was used to convert the instrument Xcalibur data files (\*.raw) to the mzXML data format (\*.mzXML) in order to be compatible for use on the next steps of the processing procedure. The software package IPO (Isotopologue Parameter Optimization) of the statistical programming language R was used to optimize the peak picking parameters, the retention time correction and the grouping parameters of XCMS (Supplementary S1). For each data set (hesperinin vs. control, naringin vs. control) different parameter settings were calculated in IPO according to the analyzed data. Briefly, for the data set hesperidin vs. control after imports in IPO the data from the 15 chickens that were administrated with hesperidin in their nutrition, the optimized parameters for the XCMS in the negative mode were:

```
xset<-xcmsSet(method="centWave", peakwidth=c(15, 40), ppm=4.6, noise=0,
snthresh=10, mzdif=0.0045, prefilter=c(3, 7500), mzCenterFun="wMean",
integrate=1, fitgauss=FALSE, verbose.columns=FALSE, nSlaves=4)
```

```
xset<-retcor(xset, method="obiwarp", plotype="none", distFunc="cor_opt",
profStep=1, center=13, response=1, gapInit=0.928, gapExtend=2.1024,
factorDiag=2, factorGap=1, localAlignment=0)
```

```
xset<-group(xset, method="density", bw=0.25, mzwid=0.01066, minfrac=1,
minsamp=1, max=50)
```

and for the positive mode were:

```
xset<-xcmsSet(method="centWave", peakwidth=c(16.364, 60), ppm=6,
noise=0, snthresh=10, mzdif=0.00406, prefilter=c(3, 7500),
mzCenterFun="wMean", integrate=1, fitgauss=FALSE,
verbose.columns=FALSE, nSlaves=4)
```

```
xset<-group(xset, method="density", bw=22, mzwid=0.015, minfrac=0.7, minsamp=1, max=50)
```

```
xset<-retcor(xset, method="obiwarp", plotype="none", distFunc="cor_opt", profStep=1, center=4, response=1, gapInIt=0.2, gapExtend=2.4, factorDiag=2, factorGap=1, localAlignment=0)
```

```
xset <- fillPeaks(xset, nSlaves=4)
```

For the data set naringin vs. control after import in IPO the data from the 17 chickens that were administrated with naringin in their nutrition, the optimized parameters for the XCMS in the negative mode were:

```
xset<-xcmsSet(method="centWave", peakwidth=c(15, 40), ppm=4.6, noise=0, snthresh=10, mzdif=0.0045, prefilter=c(3, 7500), mzCenterFun="wMean", integrate=1, fitgauss=FALSE, verbose.columns=FALSE, nSlaves=4)
```

```
xset<-retcor(xset, method="obiwarp", plotype="none", distFunc="cor_opt", profStep=1, center=13, response=1, gapInIt=0.928, gapExtend=2.1024, factorDiag=2, factorGap=1, localAlignment=0)
```

```
xset<-group(xset, method="density", bw=0.25, mzwid=0.01066, minfrac=1, minsamp=1, max=50)
```

and for the positive ion mode were:

```
xset<-xcmsSet(method="centWave", peakwidth=c(18.85, 50), ppm=5.6, noise=0, snthresh=10, mzdif=0.00395, prefilter=c(3, 7500), mzCenterFun="wMean", integrate=1, fitgauss=FALSE, verbose.columns=FALSE, nSlaves=4)
```

```
xset<-retcor(xset, method="obiwarp", plotype="none", distFunc="cor_opt", profStep=1, center=4, response=1, gapInIt=0.64, gapExtend=2.028, factorDiag=2, factorGap=1, localAlignment=0)
```

```
xset<-group(xset, method="density", bw=0.8799999999999999, mzwid=0.0265, minfrac=1, minsamp=1, max=50)
```

At the end of the procedure, \*.tsv files were generated which including the accurate masses, the retention time and the intensity of the compounds that were detected to accomplish the criteria. The \*.tsv files were imported to Microsoft Excel 2007 and manipulated using the round, concatenate and transpose commands. Metabaoanalyst 3.0 was employed according to the instruction [40] for missing values estimation by removing the features with at least 50% of missing values and replacing the remaining ones with half of the positive value in the original data. Afterwards, the data were normalized using the package Metabolomics in R environment and particularly the Normalise algorithm. The goal of normalization is to remove unwanted systematic biases, so that only biologically significant differences are present in the data.

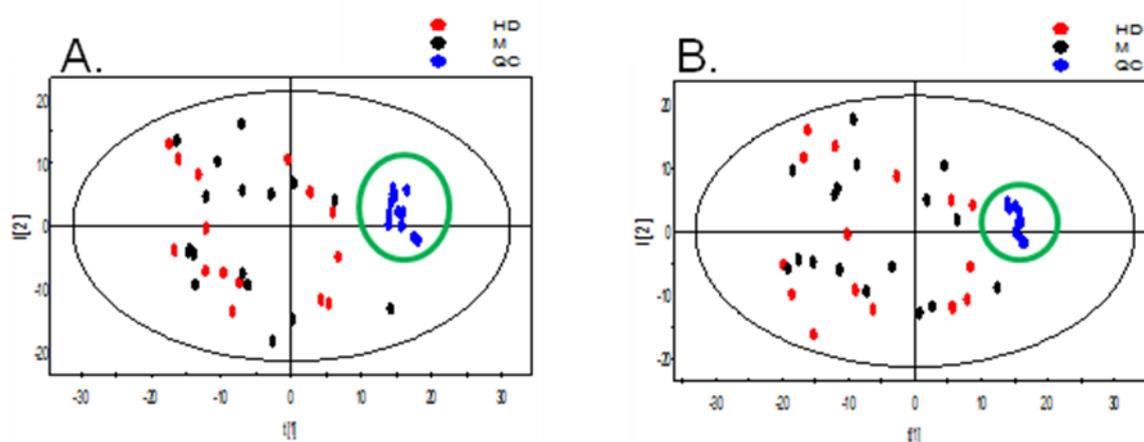
After testing different normalization methods (e.g. by the sum; median; ccmn method; nomis method), the selection of the appropriate normalization method was according to the statistical analysis (Supplementary S2, S3) and the values of the internal validation measures of Connectivity, Silhouette width, and Dunn index using the clvalid package on R environment (Supplementary S4, S5). According to the results, the ccmn method was chosen for the normalization of the data using the internal standards reserpine, 2-aminophenol and yohimbine for all the data set and ion modes (Supplementary S6).

### **3.2 Statistical analysis**

The normalized data were import to SIMCA-P + 11.5 software for multivariate analysis (MVA), i.e. PCA, PLS-DA and OPLS. According to the structure of each dataset, different normalization methodologies have been employed i.e Pareto or UV, to the mean centered data. Permutation testing for the PLS-DA models was performed in order to verify the prediction accuracy.

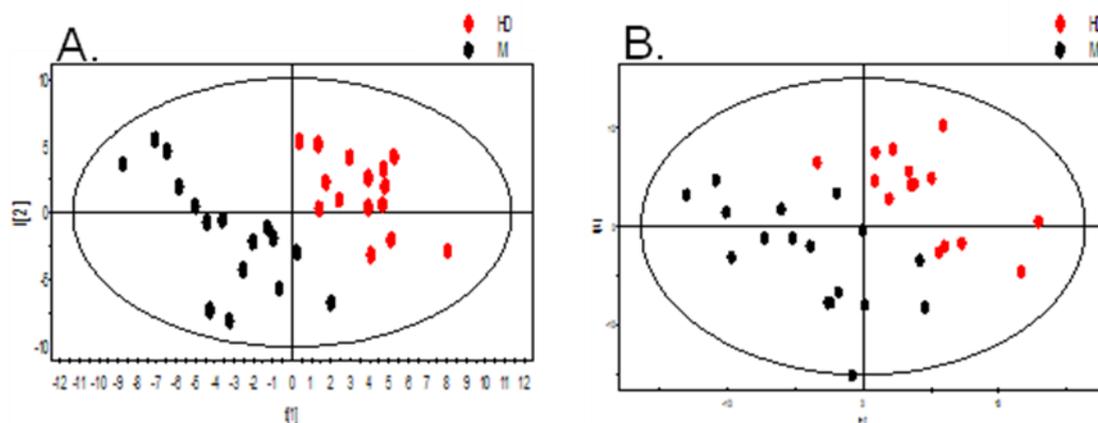
### 3.2.1 MVA of chicken tissue samples after hesperidin supplementation

For the data set hesperidin vs. control, the UV scaling was selected for both negative and positive ion modes, as it afforded tighter groups with a larger separation. The PCA methodology was initially used to examine the QC samples clustering. As illustrated in figure 2 the QC samples were clustered tightly in PCA score plots but the clustering between the two groups was non-satisfactory.



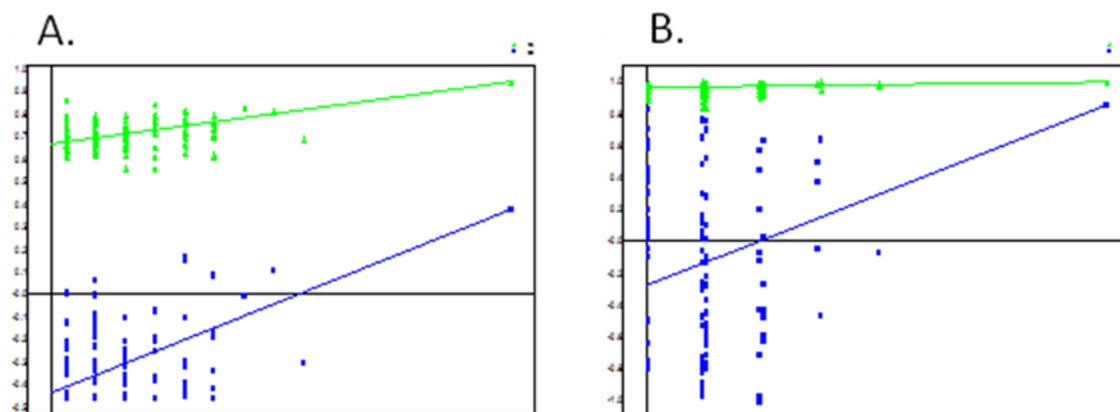
**Figure 2** Score plots from PCA in the negative ion mode (A) and in the positive ion mode (B) after the processing procedure of tissues samples from chickens that were administrated with hesperidin (HD, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. The QC samples (QC, blue dots) were clustered tightly together in both ion modes, proving the validity of the analysis

PCA-based approaches reveal separation between group sets only when the variability within group is sufficiently less than between group variability. As a consequen, PCA-based approaches frequently fail to produce any separation between groups, as shown in the current case [39]. Alternatively, PLS-DA was employed between the two groups (treated vs. untreated) and revealed a characteristic clustering, indicating a different in metabolic profile (Figure 3).



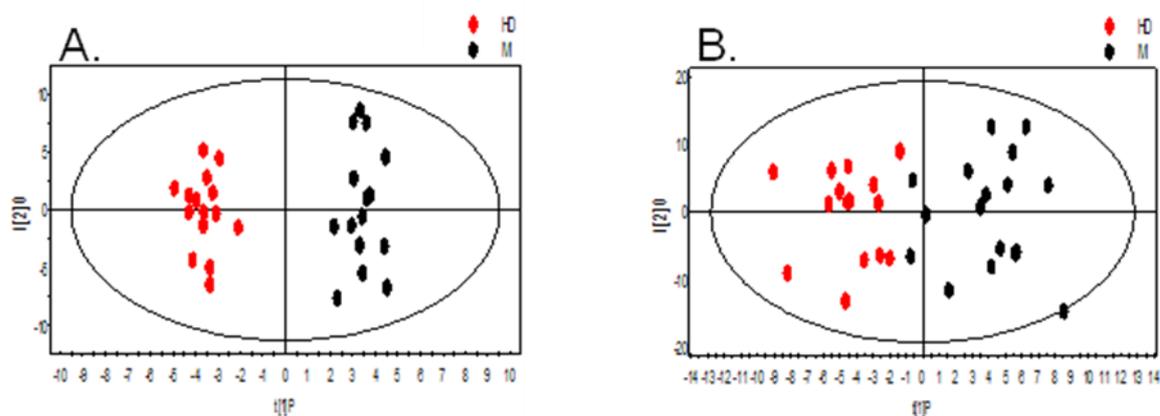
**Figure 3** Score plots from PLS-DA in the negative ion mode (A) and in the positive ion mode (A) after the processing procedure of tissues samples from chickens that were administrated with hesperidin (HD, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. A clear separation between samples of the two groups is observed. The  $R^2$ ,  $Q^2$  values were 0.574, 0.945 and 0.378 respectively, employing 4 principal components in the negative ion mode, and 0.766, 0.998 and 0.853 respectively, employing 8 principal components in the positive ion mode

Moreover, permutation testing was performed allowing of 100 random permutations in order to verify the prediction accuracy during statistic test (Figure 4). Permutation testing demonstrated that the goodness of fit and predictive ability ( $R^2/Q^2$ ) of the original model was higher than those of the permuted models in both ion modes.



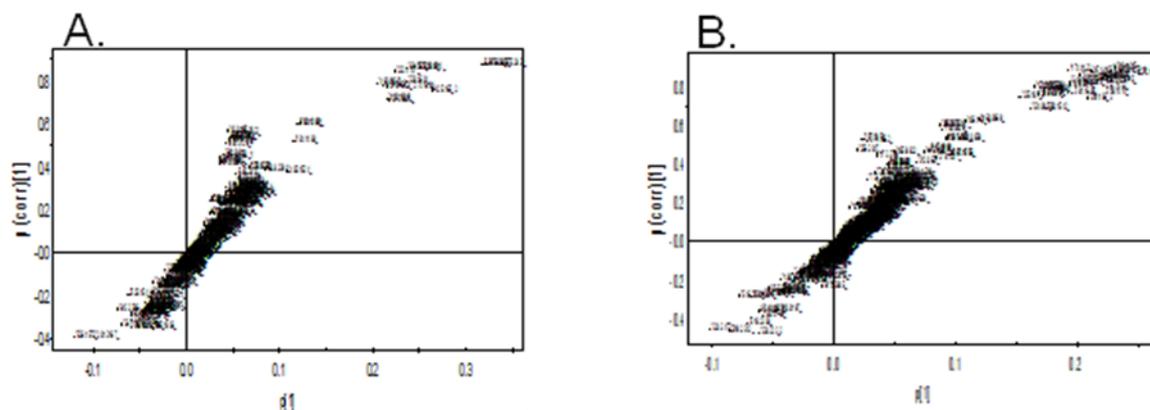
**Figure 4** Permutation testing allowing of 100 permutations of the PLS-DA model obtained by the analysis from tissues samples from chickens that were administrated with hesperidin and control samples employing UHPLC-HRMS (Orbitrap) analysis in the negative (A) and in the positive ion mode (B). The goodness of fit and predictive ability ( $R^2/Q^2$ ) of the original model was higher than those of the permuted models in both ion modes

The OPLS model was employed for the identification of the most discriminated variables, in both ion modes. The results indicated that the differences between the two groups after hesperidin administration were significant, since the treated group was clearly clustered separately from the untreated group (Figure 5).



**Figure 5** Score plots from OPLS in the negative ion mode (A) and in the positive ion mode (B) from treated with hesperidin (HD, red dots) and untreated groups (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. The  $R^2$ ,  $Q^2$  values were 0.648, 0.967 and 0.78 respectively, in the negative ion mode and 0.12, 0.762, 0.33 respectively, in the positive ion mode

Useful findings obtain from the OPLS model are the variable importance in projection (VIP) scores and S-plot. In particular, VIP values allow the identification of the variables that contributed to the separation between the two groups. Variables with a VIP score  $>1$  were taken into consideration [1]. Additionally, S-plot was employed to visualize the variables that are responsible for the discrimination (Figure 6). Variables that are distant from the origin and close to the vertical axis of S-plot are responsible for the clustering in OPLS score plot.

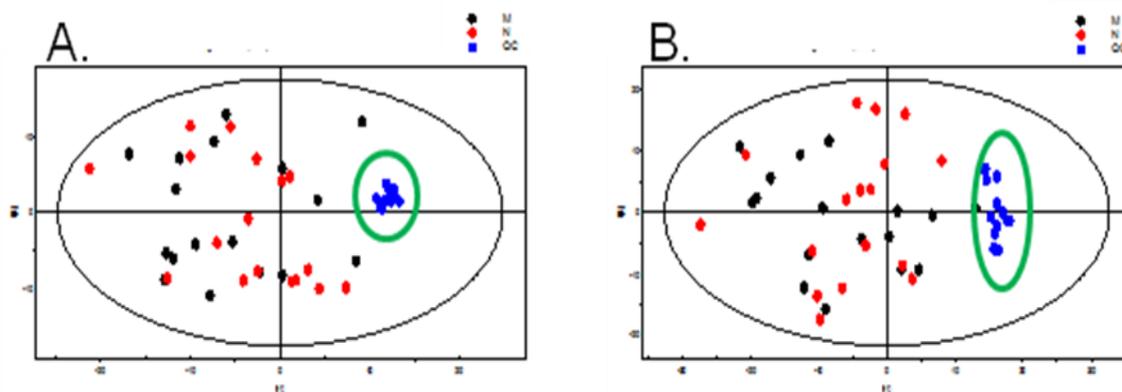


**Figure 6** S-plot of OPLS shows the relative contribution of each variable to clustering between the treated with hesperidin and untreated group after the analysis employing UHPLC-HRMS (Orbitrap) analysis, in the negative (A) and positive (B) ion modes

### 3.2.2 MVA of chicken tissue samples after naringin supplementation

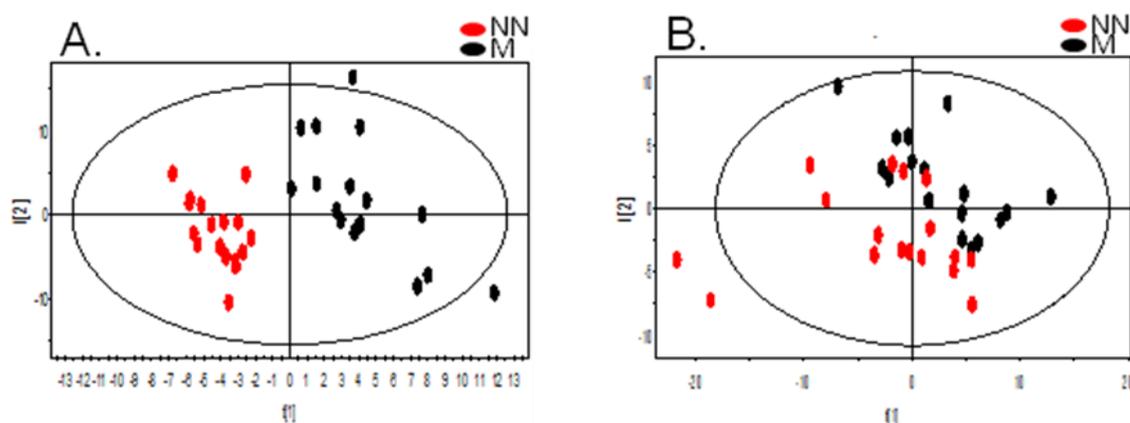
For the data set naringin vs. control the UV scaling method was selected, for both negative and positive ion mode, as it showed better clustering

separation. QC samples clustering were examined employing PCA analysis in order to investigate possible analytical variation. As shows in figure 7 the QC samples were clustered tightly in PCA scores plots in both ion modes. Nevertheless, there was no separation between the two groups.



**Figure 7** Score plots from PCA in negative (A) and positive (B) ion modes after the analysis of naringin administrated samples (NN, red dots) and control samples (M, blank dots), employing UHPLC-HRMS (Orbitrap) analysis. QC samples (QC, blue dots) were clustered tightly together in both ion modes, unlike the datasets naringin and control

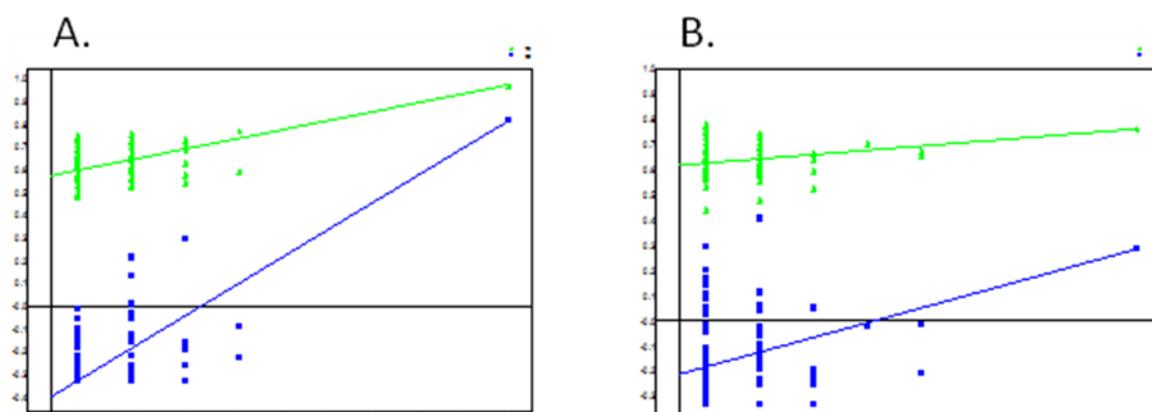
Furthermore, the PLS-DA methodology was employed to examine the separation of the two groups (treated vs. untreated). As showed on figure 8, the distribution areas for the two groups were separated, confirming the metabolomic differences.



**Figure 8** Scores plots from PLS-DA in the negative ion mode (A) and in the positive ion mode (A) after the processing procedure of tissues samples from chickens that

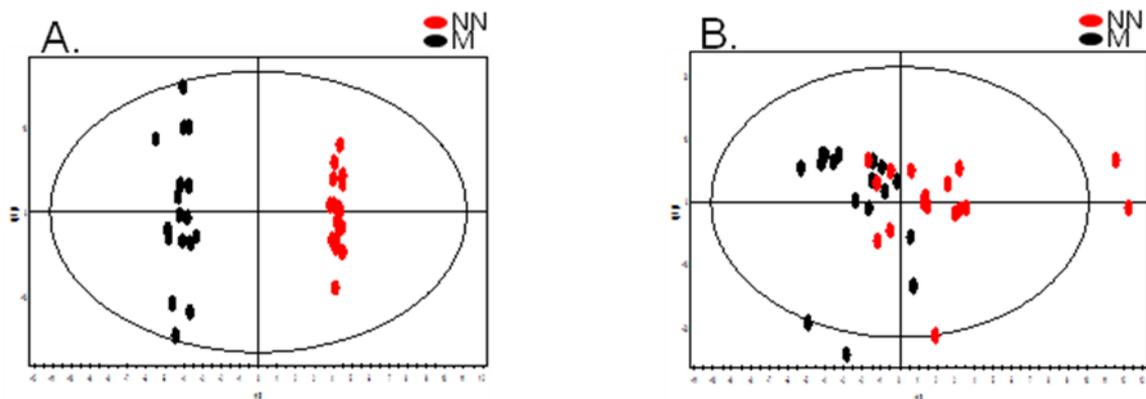
were administrated with naringin (NN, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. A clear separation between the samples of the two groups is observed. The  $R^2$ ,  $Q^2$  values were 0.404, 0.975 and 0.817 respectively, employing 3 principal components in negative ion mode. The  $R^2$ ,  $Q^2$  values were 0.514, 0.763 and 0.287 respectively, employing 3 principal components in positive ion mode.

Permutation testing was performed allowing of 100 random permutations in order to evaluate the corresponding PLS-DA model (Figure 9). Permutation testing revealed great predictability and goodness of fit in both ion modes.



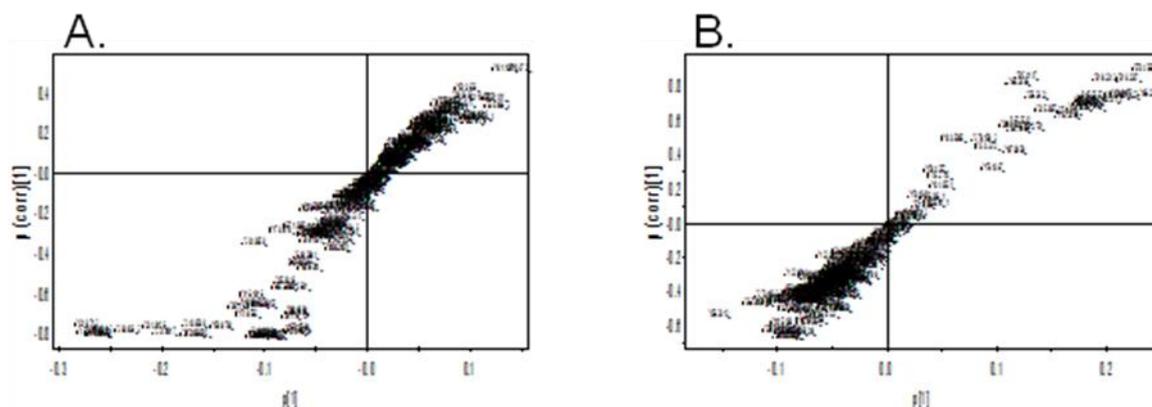
**Figure 9** Statistical validation of the PLS-DA models from tissue samples after naringin administration and control samples employing UHPLC-HRMS (Orbitrap) analysis by permutation analysis using 100 different model permutations in negative (A) and in positive ion mode (B). The goodness of fit ( $R^2$ ) and predictive capability ( $Q^2$ ) of the original model are indicated on the far right and remain higher than those of the 100 permuted models to the left

To investigate which variables significantly contribute to the observed separation, OPLS models were employed in both models. As shown in figure 10, a clear clustering between the two groups was observed in both negative and positive ion modes.



**Figure 10** Score plots from OPLS in the negative (A) and in the positive (B) ion mode, employing UHPLC-HRMS (Orbitrap) analysis. The plots showed the variation between the control group (M, black dots) and group after naringin administration (NN, red dots). The statistical parameters ( $R^2$  and  $Q^2$ ) of the models for the negative and the positive ion mode were 0.456, 0.991, 0.894 and 0.456, 0.383, 0.11 respectively

The S-plot and the variable importance in projection (VIP) scores of the OPLS models were calculated to identify potentially significant variables that contribute to the separation clustering observed by OPLS (Figure 11).



**Figure 11** S-plots of the OPLS models showed the relative contribution of each variable to the clustering between the treated with naringin and the untreated group employing UHPLC-HRMS (Orbitrap) analysis, in the negative (A) and in the positive (B) ion mode

### 3.2.3 Cat scores

In order to detect the variables that contributed the most to the clustering between the two groups, numerous statistical tools were employed for their evaluation. Correlation-adjusted t'-scores, or for short 'cat' scores are ranking criteria for biomarkers recognition. It is a corresponding shrinkage procedure which can be employed in high-dimensional settings with a comparatively small number of samples [44]. These scores are derived from a predictive perspective by exploiting a close link between gene ranking and two-class linear discriminant analysis (LDA). Cat scores measure the individual contribution of each variable to the separation of the groups, after removing the effect of all other variables. The features which are ranked with the highest score contribute the most to the classification between the training sets.

In this study the sda package and particularly the sda.ranking method from the statistical programming language R was employed for the cat scores ranking (Supplementary S7). For the analysis the normalized data after the missing values estimation were used. The variables that were ranking were responsible for the discrimination between the hesperidin vs. control group and naringin vs. control group respectively, in the negative and positive ion modes.

In order to detect only the variables that contributed the most to the classification between the treated and untreated groups, the variables that were common between cat score ranking, t-test and S-plot of OPLS were selected and evaluated (Tables 1, 2).

#### **Negative Ion Mode**

*Accurate Mass\_Retention Time*

253.1441\_18.7

128.0359\_3.7

#### **Positive Ion Mode**

*Accurate Mass\_Retention Time*

157.1413\_4.5

279.159\_24.2

503.1775\_2.9

181.1024\_2.8

**Table 1** Variables (accurate mass\_retention time) that were responsible for the clustering between the naringin group vs. control group in the negative and positive ion modes. The variables were selected according to the cat score ranking, t-test and S-plot of OPLS

**Negative Ion Mode**

*Accurate Mass\_Retention Time*

400.1836\_3.4

675.3823\_3.3

373.1648\_3.5

337.1882\_3.5

462.2027\_3.3

325.1252\_13.8

**Positive Ion Mode**

*Accurate Mass\_Retention Time*

579.2693\_16.7

371.1963\_12.9

273.1922\_3.6

193.0679\_3.6

150.091\_16.7

492.1753\_4.4

405.1688\_3.8

**Table 2** Variables (accurate mass\_retention time) that were responsible for the clustering between the hesperidin group vs. control group in the negative and positive ion modes. The variables were selected according to the cat score ranking, t-test and S-plot of OPLS

**3.3 Validation of the detected variables**

After the statistical analysis and the detection of the variables that were responsible for the separation of the treated vs. untreated groups, the package CAMERA of the R environment was employed. This package allows

the grouping of the detected features, after the XCMS analysis, into compound spectra based on retention time, calculation of known mass differences, isotope pattern and adducts detection. The compound spectra contain one or more related features, which originate from the same metabolite [16]. The required correlation across samples was set to 0.75. Table 3 summarizes the variables that contribute the most to the discrimination between naringin vs. control after the annotation for isotopes and adducts. For the hesperidin vs. control groups no isotopes and adducts were detected for the identified variables.

<i>Accurate Mass_Retention Time</i>	<i>Isotopes</i>	<i>Adduct</i>
503.1775_2.9	-	[M-2H+K],M=466.234
157.1413_4.5	[M+1]- , M=156.1380	-

**Table 3** Annotation of the variables that were responsible for the clustering between treated with naringin and control group in the negative and positive ion modes, employing package CAMERA of the R statistical language. The variables were annotated for isotopes and adducts

Furthermore, the Matlab software was used to calculate the PCC score of the detected variables. Correlation coefficients whose magnitude was higher than 0.7 indicated variables, which were considered highly correlated (Tables 4, 5).

<i>Variable</i>	<i>Correlated Variable</i>	<i>Score</i>
<i>(Accurate Mass_Retention Time)</i>	<i>(Accurate Mass)</i>	
181.1024_2.8	225.0995	0.80

**Table 4** Pearson's correlation coefficient score computation using Matlab software of the variable that was responsible for the clustering between treated with naringin and control samples in the negative ion mode

<b>Variable</b>	<b>Correlated Variable</b>	<b>Score</b>
<b>(Accurate Mass _Retention Time)</b>	<b>(Accurate Mass)</b>	
373.1648_3.5	337.1882	0.82

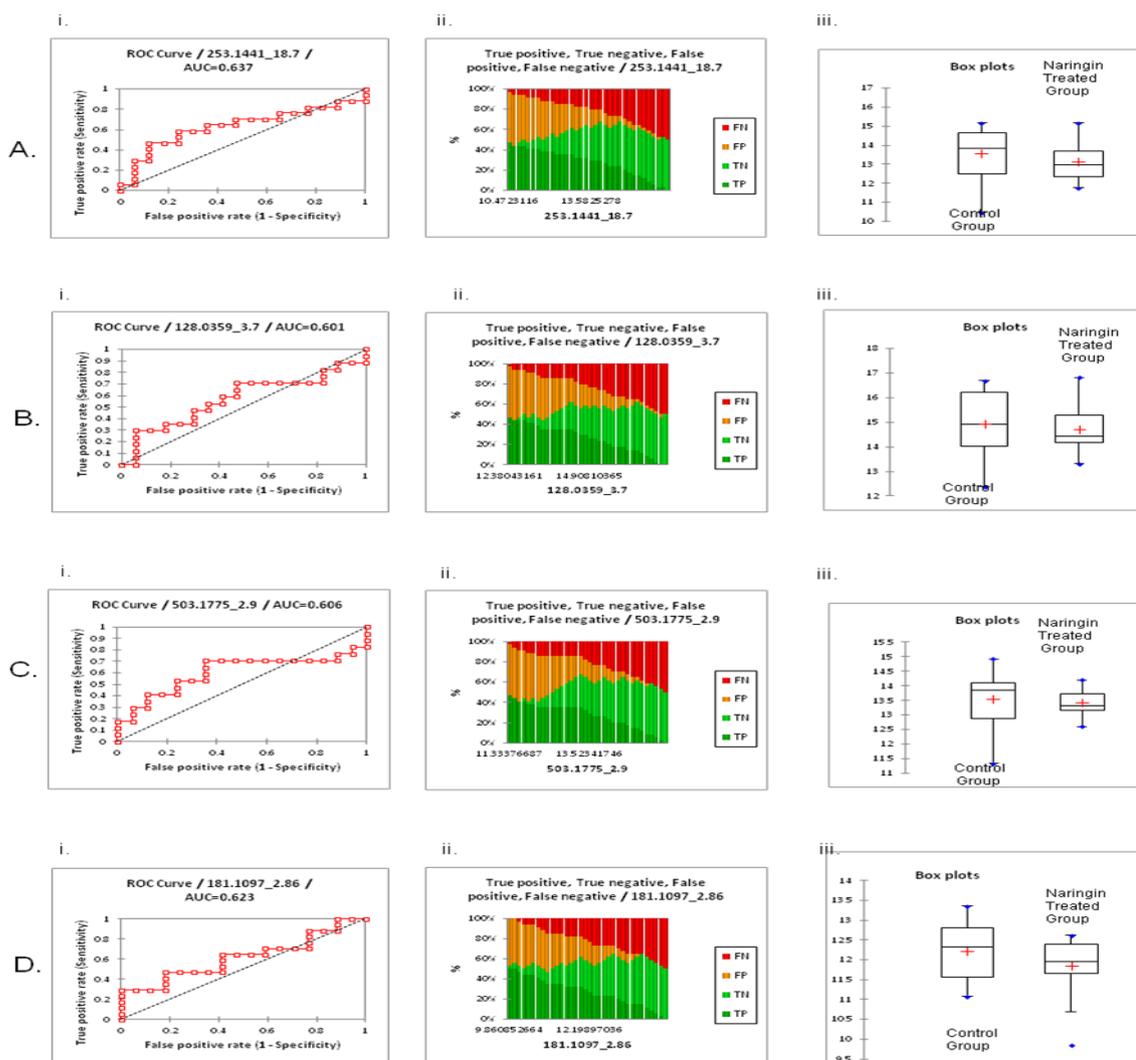
**Table 5** Pearson's correlation coefficient score computation using Matlab software of the variables that were responsible for the clustering between treated with hesperidin and control samples in the negative ion mode

### 3.3.1 Evaluation of the detected variables of the chicken tissues samples

The next step was to calculate the Receiver operating characteristic (ROC) curve in order to evaluate the sensitivity and specificity of the selected variables according to the area under the curve (AUC). Furthermore, univariate plots (box plots) and graphs of the number of TP (true positives), TN (true negative), FP (false positives) and FN (false negatives) of the detected variables were calculated using the software XLSTAT.

#### 3.3.1.1 Evaluation of the detected variables of the chicken tissues samples after dietary supplementation with naringin

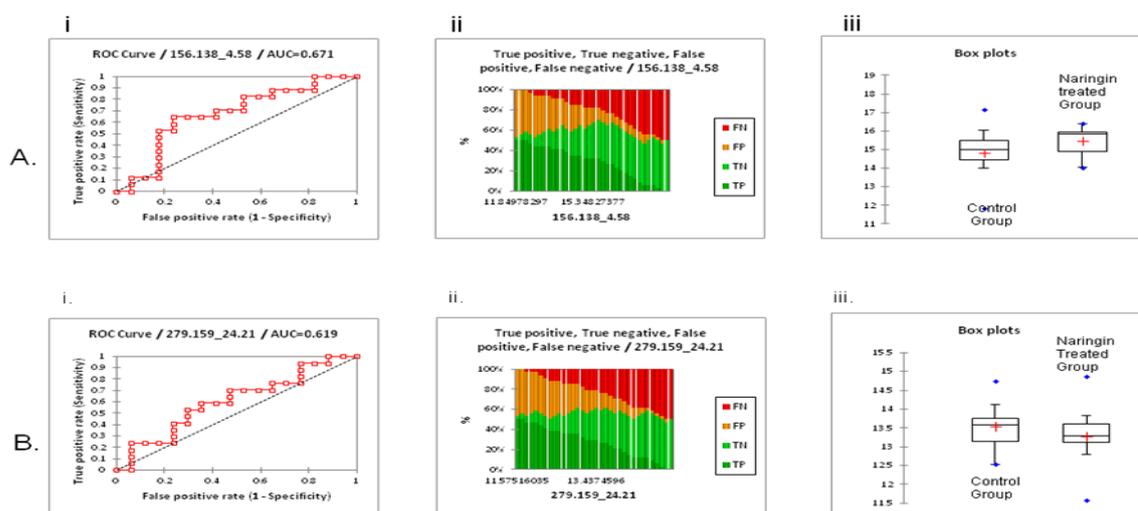
The variables 253.1441\_18.7, 128.0359\_3.7, 503.1775\_2.9 and 181.1024\_2.8 that, according to the statistical analysis contributed the most to the separation of naringin treated group and control group in the negative ion mode, had AUC value 0.6-0.7 (Figure 12Ai-Di). This indicated the low discriminating power of the variables. A box plot comparison of the range of peak heights of the variables among sample groups is shown in figure 12 (Aiii-Diii). The variables were decreased in chicken tissue samples after naringin administration. Moreover, the number of TP (true positives), TN (true negative), FP (false positives) and FN (false negatives) are shown in figure 12 (Aii-Dii).



**Figure 12** Receiver operator characteristic (ROC) curves (i), box plot (iii) and graph of the number of TP (true positives), TN (true negative), FP (false positives) and FN (false negatives) (ii) of the variables 253.1441\_18.7, 128.0359\_3.7, 503.1775\_2.9 and 181.1024\_2.8, selected by S-plot of OPLS, cat score ranking and t-test to contribute the most to the discrimination between naringin group and control group, in the negative ion mode. The AUC's were calculated according to the Bamber method. The red crosses on box plot correspond to the means. The central horizontal bars are the medians. The lower and upper limits of the box are the first and third quartiles, respectively. Points in blue are minimum and maximum for each group

For the variables that contributed the most to the separation of naringin treated group and control group in the positive ion mode, the AUC values were calculated. The variables 156.1380\_4.5 and 279.159\_24.2 showed an

AUC value 0.6-0.7 indicating the low discrimination ability of the variables. Moreover, the variable 156.1380\_4.5 was increased in chicken tissue samples after naringin administration whereas the variable 279.159\_24.2 was decreased in tissue samples (Figure 13).

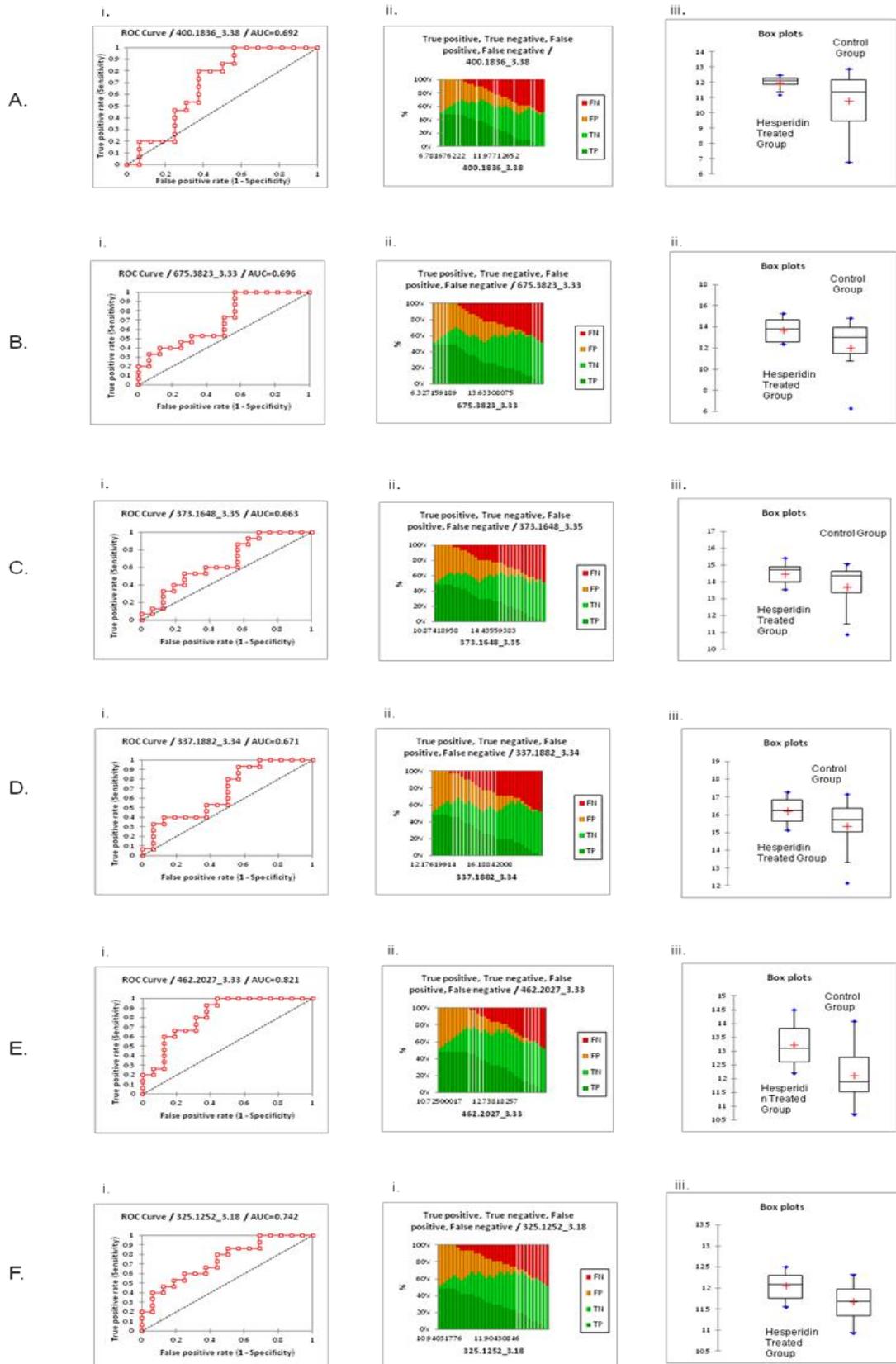


**Figure 13** Receiver operator characteristic (ROC) curves (Ai, Bi), box plots (Aiii, Biii) and graphs of the number of TP (true positives), TN (true negative), FP (false positives) and FN (false negatives) (Aii, Bii) for the variables 156.1380\_4.5 (A) and 279.159\_24.2 (B) respectively, selected by S-plot of OPLS, cat score ranking and t-test to contribute the most to the discrimination between naringin group and control group, in the positive ion mode. The AUC's were calculated according to the Bamber method. The red crosses on box plot correspond to the means. The central horizontal bars are the medians. The lower and upper limits of the box are the first and third quartiles, respectively. Points in blue are minimum and maximum for each group

### 3.3.1.2 Evaluation of the detected variables of the chicken tissues samples after dietary supplementation with hesperidin

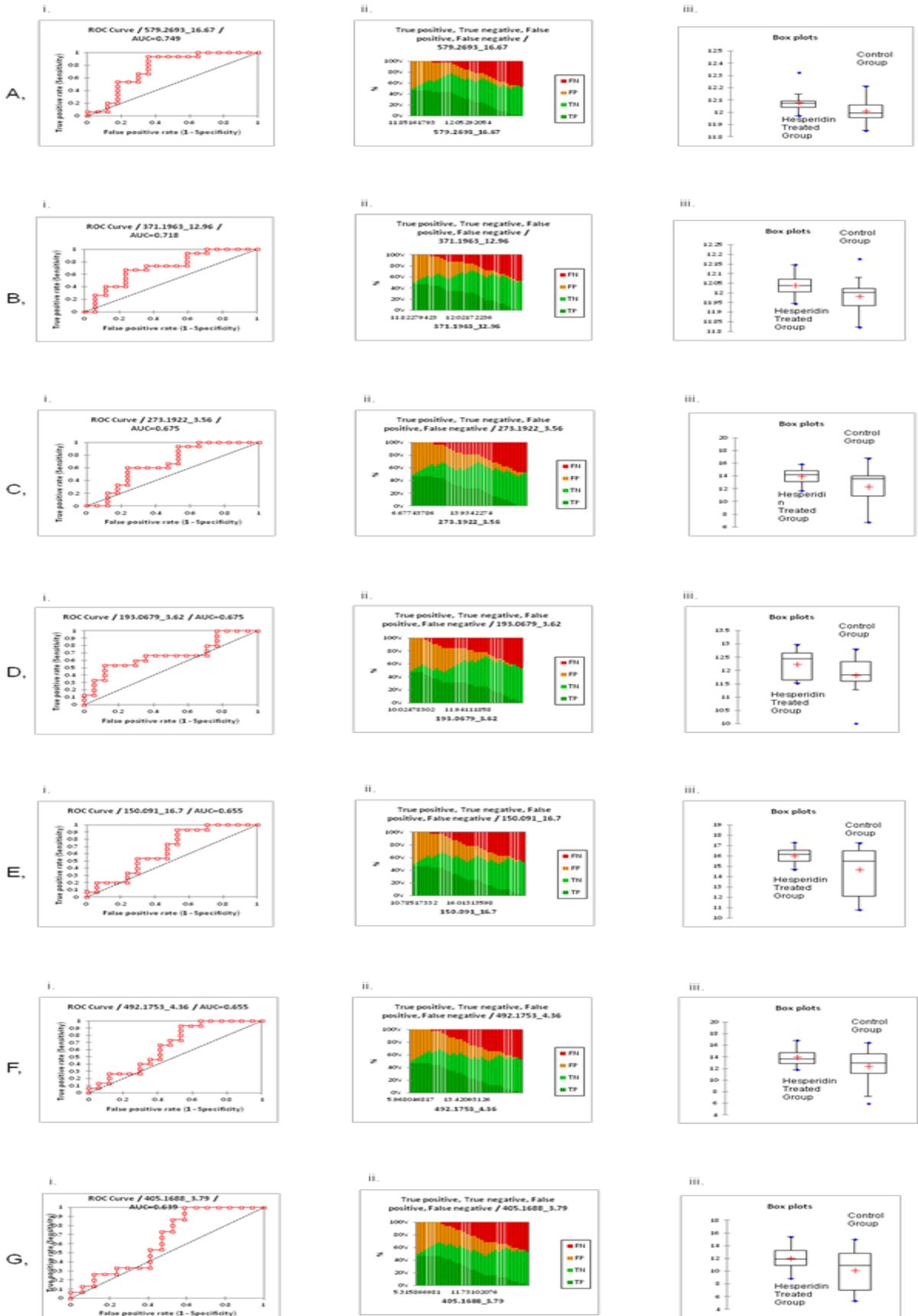
As showed in figure 14, the variables 400.1836\_3.4, 675.3823\_3.3, 373.1648\_3.5 and 337.1882\_3.5 that contributed the most to the separation of hesperidin treated group and control group in negative ion mode had an AUC value from 0.6 to 0.7 indicating low discriminating ability whereas the

variables 462.2027\_3.3 and 325.1252\_13.8 had an acceptable AUC value higher than 0.8 and 0.7 respectively. Additionally, the variable 400.1836\_3.38 was decreased in chicken tissue samples after hesperidin administration while variables 675.3823\_3.3, 373.1648\_3.5, 337.1882\_3.5, 462.2027\_3.3 and 325.1252\_13.8 were increased in tissue samples (Figure 14).



**Figure 14** Receiver operator characteristic (ROC) curves (Ai-Fi), box plots (Aiii-Fiii) and graphs of the number of TP (true positives), TN (true negative), FP (false positives) and FN (false negatives) (Aii-Fii) of the variables 400.1836\_3.4 (A), 675.3823\_3.3 (B), 373.1648\_3.5 (C), 337.1882\_3.5 (D), 462.2027\_3.3 (E) and 325.1252\_13.8 (F), selected by S-plot of OPLS, cat score ranking and t-test to contribute the most to the discrimination between hesperidin group and control group, in the negative ion mode. The AUC's were calculated according to the Bamber method. The red crosses on box plot correspond to the means. The central horizontal bars are the medians. The lower and upper limits of the box are the first and third quartiles, respectively. Points in blue are minimum and maximum for each group

The variables 579.2693\_16.7, 371.1963\_12.9, 273.1922\_3.6, 193.0679\_3.6, 150.091\_16.7, 492.1753\_4.4 and 405.1688\_3.8 that contributed the most to the discrimination between hesperidin treated group and control group in the positive ion mode. The variables 273.1922\_3.56, 193.0679\_3.62, 150.091\_16.7, 492.1753\_4.36 and 405.1688\_3.79 had an AUC value from 0.6 to 0.7. Nevertheless, the variables 579.2693\_16.67 and 371.1963\_12.96 had an AUC valuable above of 0.7. After hesperidin supplementation the variables were increased in chicken tissue samples (Figure 15).



**Figure 15** Receiver operator characteristic (ROC) curves (Ai-Gi), box plots (Aiii-Giii) and graphs of the number of TP (true positives), TN (true negative), FP (false positives) and FN (false negatives) (Aii-Gii) of the variables 579.2693\_16.7 (A), 371.1963\_12.9 (B), 273.1922\_3.6 (C), 193.0679\_3.6 (D), 150.091\_16.7 (E), 492.1753\_4.4 (F) and 405.1688\_3.8 (G), selected by S-plot of OPLS, cat score ranking and t-test to contribute the most to the discrimination between hesperidin group and control group, in the positive ion mode. The AUC's were calculated according to the Bamber method. The red crosses on box plot correspond to the means. The central horizontal bars are the medians. The lower and upper limits of the box are the first and third quartiles, respectively. Points in blue are minimum and maximum for each group

### **3.4 Identification of the detected variables of the chicken tissues samples after dietary supplementation with naringin and hesperidin**

For the prediction of the molecular formulas of the detected variables various software tools were employed. The Sirius software tool proposed the sum formula of a detected variable according to its mass and the natural distribution of its isotopes [2]. For Sirius computation analysis, an  $m/z$  tolerance of 5 ppm was applied. The Rdisop package in R language predicts the molecular formula of a variable using the accurate mass and the relative intensities of the first and second isotopes [35]. Furthermore, the mMass software tool, a cross-platform environment for the precise analysis of individual mass spectra, was employed to predict the molecular formula of the variables using 10 ppm tolerance and included the composition rules of H/C, NOPS/C, RDBE and the isotopic pattern [37]. The online databases Metlin, Kegg, Chemspider and LIPID MAPS were used to recognize the possible metabolites applied  $m/z$  tolerance of 30 ppm. Tables 6 and 7 summarize the detected variables that contribute the most to the discrimination between control group and the group after naringin-hesperidin administration, respectively. The tables depict the proposed molecular formulas and the possible identifications of the detected variables.

ID	Accurate mass_t <sub>R</sub>	VIP score	Cat score	p-value	Molecular Formula	Identification	[M]	Δm (ppm)	Mode
M1	253.1441_18.7	2.1972	7.9032	0.0493	C <sub>12</sub> H <sub>26</sub> NO <sub>2</sub>	Heptanoylcholine	[M+K-2H]- M=216.1964	3	Negative
M2	128.0359_3.7	2.3814	8.3210	0.0498	C <sub>5</sub> H <sub>9</sub> NO <sub>4</sub>	N-Methyl-D-aspartic acid	[M-H <sub>2</sub> O-H]- M=147.0532	1	Negative
M3	503.1775_2.9	1.5261	8.4560	0.0411	C <sub>20</sub> H <sub>38</sub> NO <sub>9</sub> P	Glycerophospholipids	[M-H]- M=467.2884	27	Negative
M4	181.1024_2.8	2.6902	7.8931	0.0333	-	unknown	-	-	Negative
M5	157.1413_4.5	1.5440	7.2928	0.0189	C <sub>9</sub> H <sub>17</sub> NO <sub>4</sub>	Fatty Acyl Carnitines	[M+3xHydroxylat ion] M=203.1158	0	Positive
M6	279.159_24.2	1.9722	8.4403	0.0239	C <sub>16</sub> H <sub>22</sub> O <sub>4</sub>	Prenol Lipids	[M+H]+ M=278.1581	0	Positive

**Table 6** The variables that were detected to contribute the most to the discrimination between the control group and the group after naringin administration in negative and positive ion mode. The variables were identified according to the results of the R package CAMERA, the results of PCC scoring and the proposed molecular formulas according to Sirius, Rdisop R package and Mmass software. The table depicts the VIP scoring, cat score ranking, p-value, the proposed molecular formula, the identification, the accurate mass and Δppm of the detected variables

ID	Accurate mass_t <sub>R</sub>	VIP score	Cat score	p-value	Molecular Formula	Identification	[M]	Δm (ppm)	Mode
M7	400.1836_3.4	1.6956	4.3278	0.0195	C <sub>14</sub> H <sub>30</sub> NO <sub>7</sub> P	Glycerophospholipids	[M+FA-H]- M=355.1760	22	Negative
M8	675.3823_3.3	1.6354	5.6517	0.0396	C <sub>35</sub> H <sub>61</sub> O <sub>8</sub> P	Glycerophospholipids	[M+Cl]- M=640.4104	3	Negative
M9	373.1648_3.5	1.4354	2.3529	0.0498	C <sub>16</sub> H <sub>31</sub> NO <sub>5</sub>	3-hydroxynonanoyl carnitine	[M+Na-2H]- M=317.2202	1	Negative
M10	337.1882_3.5	1.5465	4.5447	0.0478	C <sub>16</sub> H <sub>31</sub> NO <sub>5</sub>	3-hydroxynonanoyl carnitine	[M+Na-2H]- M=317.2202	1	Negative
M11	462.2027_3.3	2.1905	17.6618	0.0012	C <sub>19</sub> H <sub>40</sub> NO <sub>7</sub> P	Glycerophospholipids	[M+K-2H]- M=425.2542	1	Negative
M12	325.1252_13.8	1.7652	10.6111	0.0100	C <sub>18</sub> H <sub>24</sub> O <sub>3</sub>	Sterol lipids	[M+K-2H]- M=228.1725	11	Negative
M13	579.2693_16.7	2.8155	-	0.0309	C <sub>28</sub> H <sub>45</sub> O <sub>9</sub> P	Glycerophospholipids	[M+Na]+ M=556.2801	0	Positive
M14	371.1963_12.9	2.1262	7.6801	0.0432	C <sub>21</sub> H <sub>32</sub> O <sub>3</sub>	Sterol lipids	[M+K]+ M=332.2351	6	Positive
M15	273.1922_3.6	2.12585	4.2088	0.0402	C <sub>14</sub> H <sub>27</sub> NO <sub>4</sub>	Fatty Acyls Carnitines	[M+H]+ M=274.2013	6	Positive
M16	193.0679_3.6	2.4082	13.8606	0.0058	-	unknown	-	-	Positive

M17	150.0910_16.7	2.4802	18.4926	0.0350	C <sub>4</sub> H <sub>8</sub> N <sub>2</sub> O <sub>3</sub>	L-Asparagine	[M+NH <sub>4</sub> ] <sup>+</sup> M=132.0535	17	Positive
M18	492.1753_4.4	1.9839	4.0434	0.0498	-	unknown	-	-	Positive
M19	405.1688_3.8	2.1281	7.8323	0.0473	C <sub>20</sub> H <sub>30</sub> O <sub>6</sub>	Fatty Acyls	[M+K] <sup>+</sup> M=366.2042	1	Positive

**Table 7** The variables that were detected to contribute the most to the discrimination between the control group and the group after hesperidin administration in negative and positive ion mode. The variables were identified according to the results of the R package CAMERA, the results of PCC scoring and the proposed molecular formulas according to Sirius, Rdisop R package and Mmass software. The table depicts the VIP scoring, cat score ranking, p-value, the proposed molecular formula, the identification, the accurate mass and  $\Delta$ ppm of the detected variables

Specifically, for the variable M1 the mMass software tool proposed the molecular formula  $C_{12}H_{26}NO_2$ , according to Metlin the possible identification was the Heptanoylcholine [M=216.1964] with  $\Delta ppm=3$ . For the M2 variable according to Metlin the possible identification was the N-Methyl-D-aspartic acid [M=147.0532] with  $\Delta ppm=1$  since its fragment ions (calculated fragment ions) were identified on the chromatogram of the M2 variable (experimental fragment ions) (Table 8).

Calculated Fragment Ions ( <i>m/z</i> )	Experimental Fragment Ions ( <i>m/z</i> )	$\Delta m$ (ppm)
146.04	146.04	0
115.00	115.00	0
102.05	102.05	0
128.03	128.03	0

**Table 8** The MS/MS ions of the N-Methyl-D-aspartic acid according to Metlin database was detected on the chromatogram of the variable 128.0359\_3.7

The M3 variable according to the R package CAMERA was adduct [M-2H+K]- of the variable 466.234. According to Metlin and Lipid MAPS databases the variable was identified as a glycerophosphoserine (PS(14:1)) [M=467.2884] with  $\Delta ppm=27$ .

The M4 variable was highly correlated (score: 0.8) with the variable 225.0995. The *m/z* values and the *m/z* difference between the peaks were compared with characteristic fragments, adducts, neutral losses and modifications (YY). A match was found from their *m/z* different (*m/z* 43.9898) to correspond to [M+FA-H] and [M+H] ( $\Delta ppm=7$ ). Consequently, the detected variable 181.1024\_2.8 was considered as adduct [M+H] of the variable *m/z* 180.1024. Nevertheless, the proposed identifications from Metlin and Chempider databases were not assigned to any biochemical molecule (unknown).

The M5 variable according to the package CAMERA of R statistic language, was isotope [M+1]+ of the variable 156.1380 [M]. Employing the software Mmass and Sirius one possible molecular formula was proposed  $C_9H_{17}NO$ . According to the LIPID MAPS the possible identification was the fatty acyl

carnitine (Acetyl-D-carnitine) [M=203.1158] with  $\Delta\text{ppm}=0$  after 3x hydroxylation.

For the M6 variable the Rdisop, Sirius and Mmass proposed the molecular formula  $\text{C}_{16}\text{H}_{22}\text{O}_4$ . According to Lipid Mass the possible identification was a prenol lipid ( $\alpha$ -CEHC) [M=278.1581] with  $\Delta\text{ppm}=0$ .

The proposed molecular formula for the M7 variable employing Mmass and Sirius software tools was  $\text{C}_{14}\text{H}_{30}\text{NO}_7\text{P}$ . According to Metlin the possible identification was a glycerophospholipid (PC(6:0)) [M=355.1760] with  $\Delta\text{ppm}=22$ .

For the M8 variable according to mMass the proposed formula was  $\text{C}_{35}\text{H}_{61}\text{O}_8\text{P}$ . According to Lipid MAPS databases the possible identification was a glycerophospholipid (PA(18:4)) [M=640.4104] with  $\Delta\text{ppm}=3$ .

The variables M9 and M10 were highly correlated based on PCC score (score: 0.82). The  $m/z$  difference ( $m/z$  35.9766) between the peaks were compared with characteristic fragments, adducts, neutral losses and modifications [23]. A match was found to correspond to [M-H] and [M+Cl] ( $\Delta\text{ppm}=1$ ). Consequently, the detected variables 373.1648\_3.35 and 337.1882\_3.34 were considered as adducts of the variable  $m/z$  338.1954. The proposed identification from Metlin and Chemspider databases was 3-hydroxynonanoyl carnitine [M=317.2202] with  $\Delta\text{ppm}=1$  and proposed molecular formula  $\text{C}_{16}\text{H}_{31}\text{NO}_5$ .

The proposed molecular formula based on the Mmass, the Sirius and the Rdisop R package for the variable M11 was  $\text{C}_{19}\text{H}_{40}\text{NO}_7\text{P}$ . According to Metlin and Lipid MAPS the possible identification was a a glycerophospholipid (PC(11:0)) [M=425.2542] with  $\Delta\text{ppm}=1$ .

For the variable M12 the proposed identification according to Metlin and Lipid MAPS was a sterol lipid (2-hydroxyestradiol) [M=228.1725] with  $\Delta\text{ppm}=11$  and molecular formula  $\text{C}_{18}\text{H}_{24}\text{O}_3$ . The fragment pattern (calculated fragment ions) of 2-hydroxyestradiol was identified on the chromatogram of the M12 variable (experimental fragment ions) (Table 9).

Calculated Fragment Ions ( <i>m/z</i> )	Experimental Fragment Ions ( <i>m/z</i> )	$\Delta m(\text{ppm})$
147.04	144.04	0
162.07	162.07	0
173.04	173.04	0

**Table 9** The MS/MS ions of the 2-hydroxyestradiol according to Metlin database was detected on the chromatogram of the variable 325.1252\_13.8

For the M13 variable the Rdisop, Sirius and Mmass software tools proposed the molecular formula  $C_{28}H_{45}O_9P$ . According to Metlin and Lipid Mass database the possible identification was a glycerophospholipid (PG(22:6)) [M=556.2801] with  $\Delta \text{ppm}=0$ .

According to Metlin and Lipid MAPS the possible identification for the variable M14 was a sterol lipid (21-hydroxypregnenolone) [M=332.2351] with  $\Delta \text{ppm}=6$ . The proposed molecular formula was  $C_{21}H_{32}O_3$ . The MS/MS pattern of 21-hydroxypregnenolone was detected on the chromatogram of the variable M14 (Table 10).

Calculated Fragment Ions ( <i>m/z</i> )	Experimental Fragment Ions ( <i>m/z</i> )	$\Delta m(\text{ppm})$
297.22	297.22	0
279.20	279.20	0

**Table 10** The MS/MS ions of the 21-hydroxypregnenolone according to Metlin database was detected on the chromatogram of the variable 371.1963\_12.9

For the M15 variable the Sirius and Mmass software tools proposed the molecular formula  $C_{14}H_{27}NO_4$ . According to Lipid Mass the possible identification was a fatty acyl carnitine (Heptanoylcarnitine) [M=274.2013] with  $\Delta \text{ppm}=6$ .

For the variable M16 there was no proposed molecular formula or identification from the databases (unknown).

For the variable M17 according to Metlin databases the proposed identification was L-Asparagine [M=132.0535] with proposed molecular formula  $C_4H_8N_2O_3$  and with  $\Delta \text{ppm}=17$ . Furthermore, these findings were also

confirmed by comparing the theoretical MS/MS data of L-Asparagine with the experimental ones (Table 11).

Calculated Fragment Ions ( <i>m/z</i> )	Experimental Fragment Ions ( <i>m/z</i> )	$\Delta m(\text{ppm})$
133.06	133.06	0
116.03	116.03	0

**Table 11** The MS/MS pattern of L-asparagine according to Metlin database was detected on the chromatogram of the variable 150.091\_16.7

For the variable M18 there was no proposed molecular formula or identification from the databases (unknown).

For the M19 variable the Sirius and Mmass software tools proposed the molecular formula  $C_{20}H_{30}O_6$ . According to Metlin and Lipid Mass the possible identification was a fatty acyl (ox-LGD2) [ $M=366.2042$ ] with  $\Delta ppm=1$ .

### 3.5 Biochemical analysis of the variables after naringin administration

According to the statistical analysis six variables were detected to contribute the most to the discrimination of the group after naringin dietary supplementation from the control group.

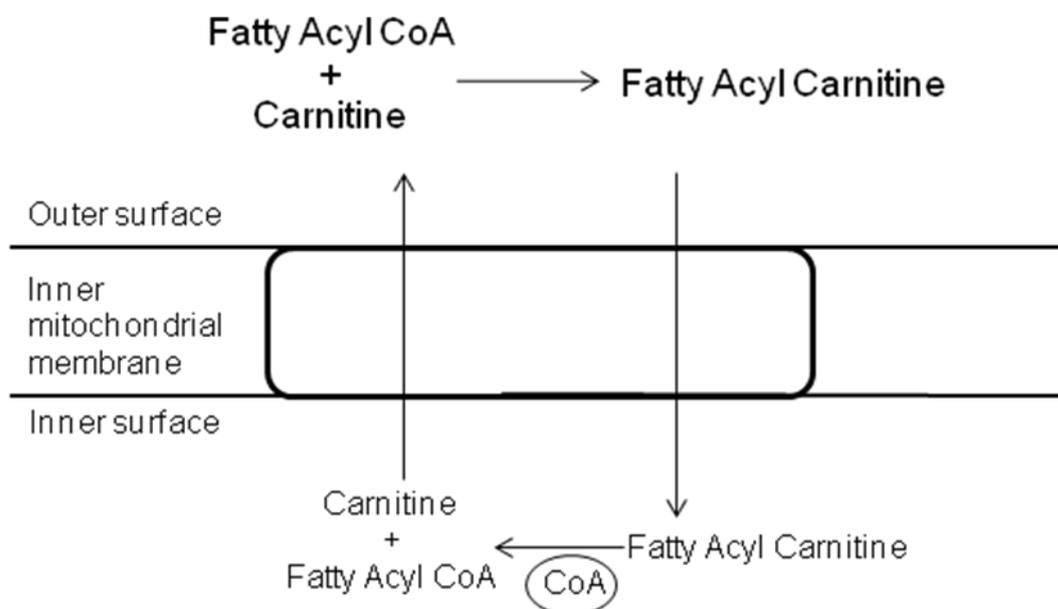
The M1 variable was identified as heptanoylcholine. Heptanoylcholine is the ester product of heptanoic acid and choline.

The variable M2 was identified as N-Methyl-D-aspartic acid. N-Methyl-D-aspartic acid (NMDA) is an amino acid derivative acting as a specific agonist for a class of glutamate receptors, the NMDA type, and therefore mimics the action of the neurotransmitter glutamate on that receptor. Synthetic NMDA has been found to elicit very strong activity for the induction of hypothalamic factors and hypophyseal hormones in mammals. Endogenous NMDA has been found in nervous tissues of gallus gallus at a concentration range of  $2.20 \pm 0.20$  nmol/g tissues [3].

The variable M3 was identified as a glycerophospholipid. Glycerophospholipids are key component of biological membrane as well as being involved in metabolism and signaling. Examples of glycerophospholipids found in

biological membranes are phosphatidylcholine (PC), phosphatidylethanolamine (PE) and phosphatidylserine (PS). In addition to serving as a primary component of cellular membranes and binding sites for intra- and intercellular proteins, some glycerophospholipids in eukaryotic cells are either precursors of, or are themselves, membrane-derived second messengers

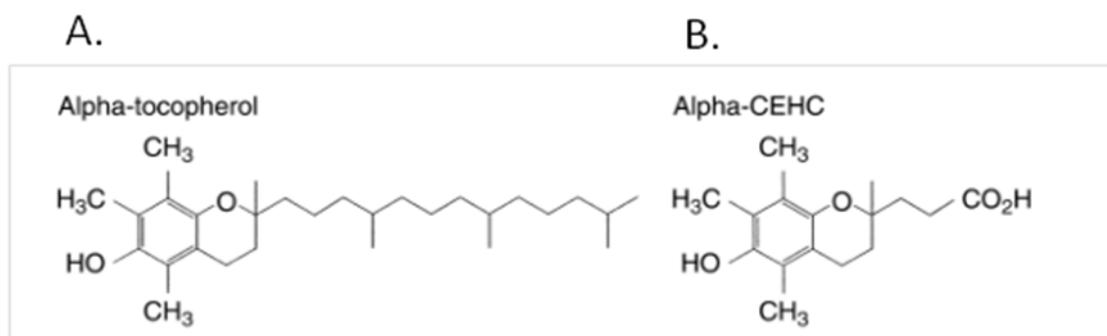
The variable M5 was identified as fatty acyl carnitine. Carnitine is an amino acid derivative involved in lipid metabolism in mammals and other eukaryotes. In eukaryotic cells, in particular, it is required for the transport of fatty acids from the intermembraneous space in the mitochondria into the inner mitochondrial membrane during the catabolism of lipids for the generation of metabolic energy [33] (Figure 16).



**Figure 16** Fatty acyl CoA is impermeable to the inner mitochondrial membrane, so it is carried in the form of fatty acyl carnitine. Carnitine is required for the transport of fatty acids from the intermembraneous space in the mitochondria into the inner mitochondrial membrane during the catabolism of lipids for the generation of metabolic energy

The variable M6 was identified as prenol lipid ( $\alpha$ -CEHC). The latter is the product of the metabolism of Alpha-tocopherol (Vitamin E) (Figure 17), [11].

The antioxidant effects of Alpha-tocopherol supplementation on chicken muscle are well established [41]. For that reason, for experimental purpose, Vitamine E was administrated on control group in order to compare its antioxidant properties in broiler meat with hesperidin supplementation.



**Figure 17** Structures of alpha-tocopherol (A) and alpha-CEHC (B). Alpha-CEHC is the product of alpha-tocopherol metabolism

### 3.6 Biochemical evaluation of the variables after hesperidin administration

Twelve variables were detected to contribute the most to the discrimination of the group after hesperidin dietary supplementation from the control group. Most of them were identified as glycerophospholipis (M7, M8, M11, M13) two as sterol lipids (M14, M12) and one as fatty acyls (M19). All of them are categories of lipids. Polar lipids are structural components of cell membranes, where they participate in the formation of the permeability barrier of cells and subcellular organelles in the form of a lipid bilayer. The major lipid type defining this bilayer in almost all membranes is glycerol-based phospholipid [19]. The importance of the membrane lipid physical (phase) state is evidenced by the fact that lipids may control the physiological state of a membrane organelle by modifying its biophysical aspects, such as the polarity and permeability. The main biological functions of lipids include energy storage, acting as structural components of cell membrane and participating as important signaling molecules.

The variables M9 and M15 were identified as 3-hydroxynonanoyl carnitine and heptanoylcarnitine respectively, which belong to the family of acyl carnitines. The latter are essential compounds for the metabolism of fatty acids (Figure 16).

The variable M17 was identified as L-Asparagine, which is a common natural amino acid. It is considered as a non-essential amino acid that means that it can be synthesized from the pathway L-asparagine biosynthesis, which is part of Amino-acid biosynthesis.

#### **4. Conclusion Remarks**

An untargeted metabolomic study was performed in tissue samples after dietary supplementation with naringin and hesperidin in chickens. The samples were analyzed using the package XCMS in R environment. After the normalization using a mixture of internal standards, MVA analysis was performed to detect the variables that contributed the most to the discrimination between treated vs. untreated groups. Six variables were detected to discriminate naringin dietary supplementation from control group and twelve variables hesperidin supplementation from control group. The variables were identified using various software tools and online databases.

## References

1. Andersen CM, Bro R. (2010) Variable selection in regression—a tutorial. *J Chemometr.* 24, 728–737.
2. Böcker S, Letzel MC, Lipták Z, Pervukhin A. (2009) SIRIUS: decomposing isotope patterns for metabolite identification. *Bioinformatics.* 25, 218-224.
3. D’Aniello G, Tolino A, D’Aniello A, Errico F, Fisher GH, Di Fiore MM. (2000) The role of D-aspartic acid and N-methyl-D-aspartic acid in the regulation of prolactin release. *Endocrinology.* 141, 3862-3870.
4. De Livera AM, Dias DA, De Souza D, Rupasinghe T, Pyke J, Tull D, Roessner U, McConville M, Speed TP. (2012) Normalizing and integrating metabolomics data. *Anal Chem.* 84, 10768-10776.
5. Dunn JC. (1974) Well separated clusters and optimal fuzzy-partitions. *Journal of Cybernetics.* 4, 95–104.
6. Fernández-Albert F, Llorach R, Andrés-Lacueva C, Perera A. (2014) An R package to analyse LC/MS metabolomic data: MAIT (Metabolite Automatic Identification Toolkit). *Bioinformatics.* 30, 1937-1939.
7. Gibbons H, O’Gorman A, Brennan L. (2015) Metabolomics as a tool in nutritional research. *Curr Opin Lipidol.* 26, 30-34.
8. Gika HG, Theodoridis GA, Plumb RS, Wilson ID. (2014) Current practice of liquid chromatography-mass spectrometry in metabolomics and metabonomics. *J Pharm Biomed Anal* 87, 12-25.
9. Handl J, Knowles J, Kell DB (2005). “Computational Cluster Validation in Post-Genomic Data Analysis.” *Bioinformatics*, 21, 3201–3212.
10. Hanhineva K, Brunius C, Andersson A, Marklund M, Juvonen R, Keski-Rahkonen P, Auriola S, Landberg R. (2015) Discovery of

urinary biomarkers of whole grain rye intake in free-living subjects using non-targeted LC-MS metabolite profiling. *Mol Nutr Food Res.* 59, 2315-2325.

11. Himmelfarb J, Kane J, McMonagle E, Zaltas E, Bobzin S, Boddupalli S, Phinney S, Miller G. (2003) Alpha and gamma tocopherol metabolism in healthy subjects and patients with end-stage renal disease. *Kidney Int.* 64, 978-991.
12. Hirata A, Murakami Y, Shoji M, Kadoma Y, Fujisawa S. (2005) Kinetics of radical-scavenging activity of hesperetin and hesperidin and their inhibitory activity on COX-2 expression. *Anticancer Res.* 25, 3367-3374.
13. Holmes E, Wilson ID, Nicholson JK. (2008) Metabolic phenotyping in health and disease. *Cell.* 134, 714-717.
14. James EL, Parkinson EK. (2015) Serum metabolomics in animal models and human disease. *18*, 478-483.
15. Kell DB, Goodacre R. (2014) Metabolomics and systems pharmacology: why and how to model the human metabolic network for drug discovery. *Drug Discov Today.* 19, 171-182.
16. Kuhl C, Tautenhahn R, Böttcher C, Larson TR, Neumann S. (2012) CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data set. *Anal Chem.* 84, 283-289.
17. Libiseller G, Dvorzak M, Kleb U, Gander E, Eisenberg T, Madeo F, Neumann S, Trausinger G, Sinner F, Pieber T, Magnes C. (2015) IPO: a tool for automated optimization of XCMS parameters. *BMC Bioinformatics.* 16, 118.
18. Liesenfeld DB, Grapov D, Fahrman JF, Salou M, Scherer D, Toth R, Habermann N, Böhm J, Schrotz-King P, Gigic B, Schneider

- M, Ulrich A, Herpel E, Schirmacher P, Fiehn O, Lampe JW, Ulrich CM. (2015) Metabolomics and transcriptomics identify pathway differences between visceral and subcutaneous adipose tissue in colorectal cancer patients: the ColoCare study. *Am J Clin Nutr.* 102, 433-443.
19. Lodish H, Berk A, Zipursky SL, et al. *Molecular Cell Biology*. 4th edition. New York: W. H. Freeman; 2000. Section 5.3, Biomembranes: Structural Organization and Basic Functions.
  20. Lommen A, Kools HJ. (2012) MetAlign 3.0: performance enhancement by efficient use of advances in computer hardware. *Metabolomics.* 8, 719-726.
  21. Lovmar L, Ahlford A, Jonsson M, Syvänen AC. (2005) Silhouette scores for assessment of SNP genotype clusters. *BMC Genomics.* 6, 35.
  22. Sangster T, Major H, Plumb R, Wilson AJ, Wilson ID. (2006) A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabolomic analysis *Analyst.* 131, 1075-1078.
  23. Lynn KS, Cheng ML, Chen YR, Hsu C, Chen A, Lih TM, Chang HY, Huang CJ, Shiao MS, Pan WH, Sung TY, Hsu WL. (2015) Metabolite identification for mass spectrometry-based metabolomics using multiple types of correlated ion information *Anal Chem.* 87, 2143-2151.
  24. Metz CE. (1978) Basic principles of ROC analysis. *Semin Nucl Med.* 8, 283-298.
  25. Nicholson JK, Connelly J, Lindon JC, Holmes E. (2002). Metabonomics: a platform for studying drug toxicity and gene function. *Nat Rev Drug Discov.* 1, 153-161.

26. Nielsen SE, Freese R, Kleemola P, Mutanen M. (2002) Flavonoids in human urine as biomarkers for intake of fruits and vegetables. *Cancer Epidemiol Biomarkers Prev.* 11, 459-466.
27. O’Gorman A, Gibbons H, Brennan L. (2013) Metabolomics in the identification of biomarkers of dietary intake. *Comput Struct Biotechnol J.* 4, e201301004.
28. O’Sullivan A., Gibney M.J., Brennan L. (2011) Dietary intake patterns are reflected in metabolomic profiles: Potential role in dietary assessment studies. *Am. J. Clin. Nutr.* 93, 314–321.
29. Pang WY, Wang XL, Mok SK, Lai WP, Chow HK, Leung PC, Yao XS, Wong MS (2010) Naringin improves bone properties in ovariectomized mice and exerts oestrogen-like activities in rat osteoblast-like (UMR-106) cells., *Br J Pharmacol.* 159, 1693-1703.
30. Park SH, Goo JM, Jo CH (2004) Receiver operating characteristic (ROC) curve: practical review for radiologists. *Korean J Radiol.* 5, 11-18.
31. Pluskal T., Castillo S., Villar-Briones A., Orešič M. (2010) Mzmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometrybased molecular profile data. *BMC Bioinformatics.* 11, 395.
32. Redestig H, Fukushima A, Stenlund H, Moritz T, Arita M, Saito K, Kusano M. (2009) Compensation for systematic cross-contribution improves normalization of mass spectrometry based metabolomics data. *Anal Chem.* 81, 7974-7980.
33. Schooneman MG, Vaz FM, Houten SM, Soeters MR. (2013) Acylcarnitines: reflecting or inflicting insulin resistance? *Diabetes.* 62, 1-8.

34. Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. (2006) XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem.* 78, 779-787.
35. Stanstrup J, Gerlich M, Dragsted LO, Neumann S. (2013) Metabolite profiling and beyond: approaches for the rapid processing and annotation of human blood serum mass spectrometry data. *Anal Bioanal Chem.* 405, 5037-5048.
36. Strimbu K, Tavel JA. (2010) What are biomarkers? *Curr Opin HIV AIDS.* 5, 463-466.
37. Strohal M, Kavan D, Novák P, Volný M, Havlíček V. (2010) mMass 3: a cross-platform software environment for precise analysis of mass spectrometric data. *Anal Chem.* 82, 4648-4651.
38. Tautenhahn R, Patti GJ, Rinehart D, Siuzdak G. (2012) XCMS Online: a web-based platform to process untargeted metabolomic data. *Anal Chem.* 84, 5035-5039.
39. Worley B, Halouska S, Powers R. (2013) Utilities for quantifying separation in PCS/PLS-DA scores plots. *Anal Biochem.* 433, 102-104.
40. Xia J, Wishart DS. (2011) Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nat Protoc.* 6, 743-760.
41. Young JF, Stagsted J, Jensen SK, Karlsson AH, Henckel P. (2003) Ascorbic acid, alpha-tocopherol, and oregano supplements reduce stress-induced deterioration of chicken meat quality. *Poult Sci.* 82, 1343-1351.
42. Yu T, Park Y, Johnson JM, Jones DP. (2009) apLCMS—adaptive processing of high-resolution LC/MS data. *Bioinformatics.* 25, 1930-1936.

43. Zhang X, Hung TM, Phuong PT, Ngoc TM, Min BS, Song KS, Seong YH, Bae K. (2006) Anti-inflammatory activity of flavonoids from *Populus davidiana*. *Arch Pharm Res.* 29, 1102-1108.
44. Zuber V, Strimmer K. (2009) Gene ranking and biomarker discovery under correlation. *Bioinformatics.* 25, 2700-2707.

## Chapter abstract

To date numerous metabolomic studies has been performed in order to detect changes on the biochemical pathways in response to diseases, environmental alterations or changes on dietary pattern. Naringin and hesperidin are two flavonoids with a broad spectrum of biological activities. The aim of this study was to present an analytical and computational methodology in order to characterize the metabolic changes that occur in tissues in response to naringin and hesperidin dietary supplementation in chickens. Forty nine 1-day-old Ross 308 broiler chickens were randomly divided into 3 groups: 17 chickens, given diets supplemented with naringin (1.5 g/kg of feed), 15 chickens given diets supplemented with hesperidin (1.5 g/kg of feed) and 17 chickens that were given commercial basal diets (control chickens). After 30 days of administration the chickens were slaughtered and tissue samples were collected and dried by lyophilization. The samples were analyzed by UHPLC-HRMS (Orbitrap Discovery XL). For the processing procedure the data were analyzed using R statistical language. Multivariate analysis was applied to detect the variables that contributed the most to the discrimination between the treated vs. the untreated groups. Six variables were detected to discriminate control group from the group after naringin administration and twelve variables the control group from the group after hesperidin supplementation. These variables were identified applied various software tools and online databases.

**Keywords:** R statistical language – MVA –software tools – online databases

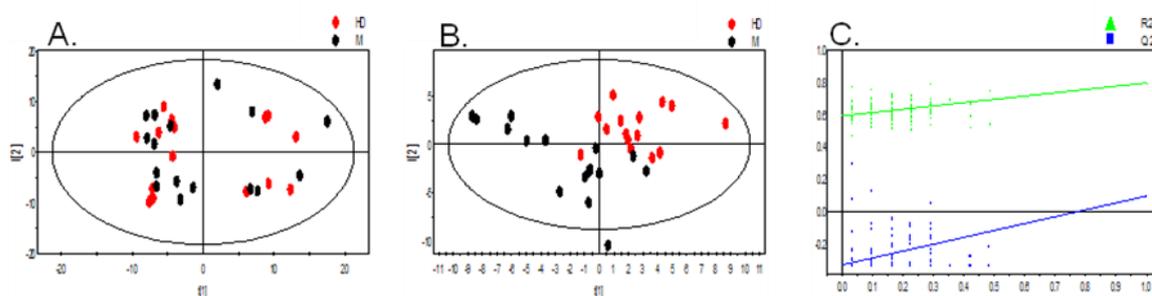
## Supporting Information for Chapter 5

### Supplementary S1

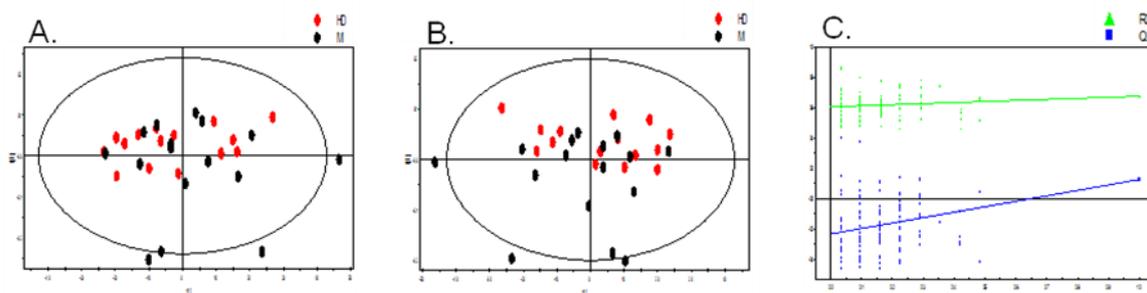
```
>setwd()
> peakpickingParameters <- getDefaultXcmsSetStartingParams('centWave')
> peakpickingParameters$min_peakwidth <- c(4,15)
> peakpickingParameters$max_peakwidth <- c(20,40)
> peakpickingParameters$ppm <- c(1,5)
> peakpickingParameters$value_of_prefilter <- 7500
>resultPeakpicking<-optimizeXcmsSet(files=file,
params=peakpickingParameters, nSlaves=4, subdir='rsmDirectory')
>optimizedXcmsSetObject <- resultPeakpicking$best_settings$xset
>resultPeakpicking$best_settings
>retcorGroupParameters <- getDefaultRetGroupStartingParams()
> retcorGroupParameters$profStep <- 1
>resultRetcorGroup<-optimizeRetGroup(xset=optimizedXcmsSetObject,
params=retcorGroupParameters,nSlaves=4, subdir="rsmDirectory")
>writeRScript(resultPeakpicking$best_settings$parameters,
resultRetcorGroup$best_settings, nSlaves=4)
```

**S1** Commands that were used to the package IPO of the statistical programming language R, in order to optimize XCMS peak picking parameters, grouping parameters and correct retention time. XCMS parameters were used for the processing of the datasets hesperidin vs. control and naringin vs. control in the negative and positive ion modes

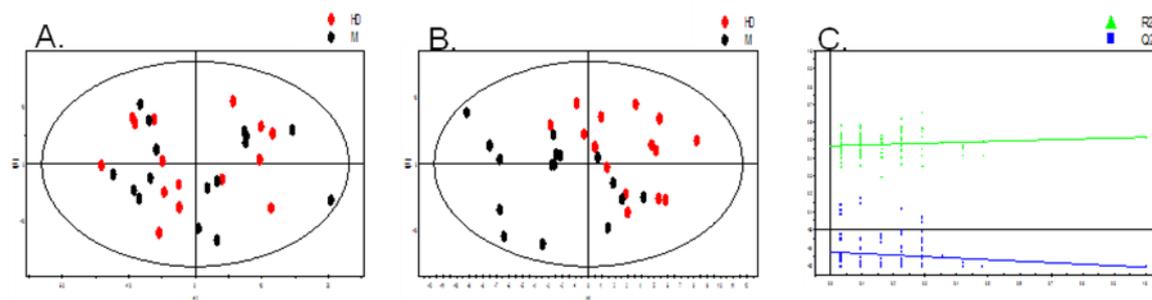
## Supplementary S2



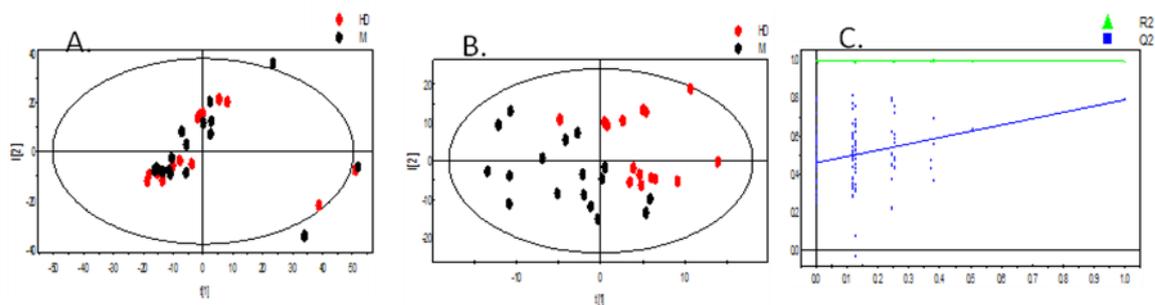
**Figure 1** Score plots from PCA (A), PLS-DA (B) and permutation testing (C) in the negative ion mode after the normalization processing procedure by median, on tissues samples from chickens that were administrated with hesperidin (HD, red dots) and control tissues samples (M, blank dots), employing UHPLC-HRMS (Orbitrap) analysis. A separation between samples is observed only in PLS-DA plot. The  $R^2$ ,  $Q^2$  values for PCA and PLS-DA were 0.529, 0.318, 0.355 and 0.8, 0.097 respectively, employing 3 principal components. The permutation test using 100 different model permutations shows that some of the permuted models presented better predictive capability ( $Q^2$ ) than the original model



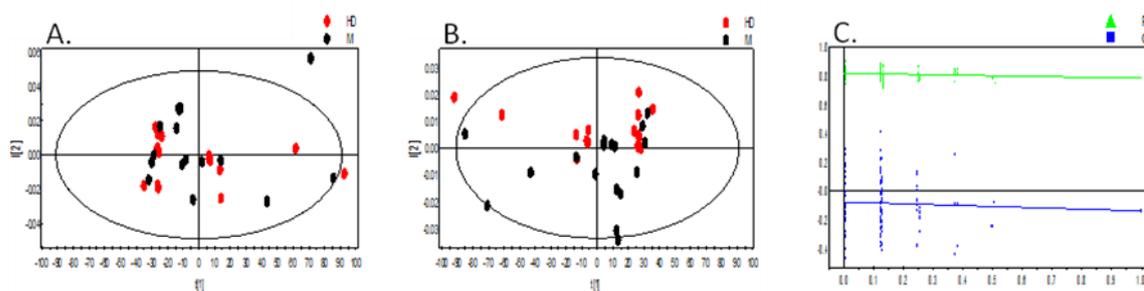
**Figure 2** Score plots from PCA (A), PLS-DA (B) and permutation testing (C) in the negative ion mode after the normalization processing procedure by sum, on tissues samples from chickens that were administrated with hesperidin (HD, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. As plots illustrated, no separation between samples is observed. The  $R^2$ ,  $Q^2$  values for PLS-DA and PCA were 1, 1, 1 and 0.676, 0.125 respectively, employing 2 principal components for PCA and 4 for PLS-DA. The permutation testing using 100 different model permutations shows that some of the permuted models presented better predictive capability ( $Q^2$ ) and fit ability ( $R^2$ ) than the original model



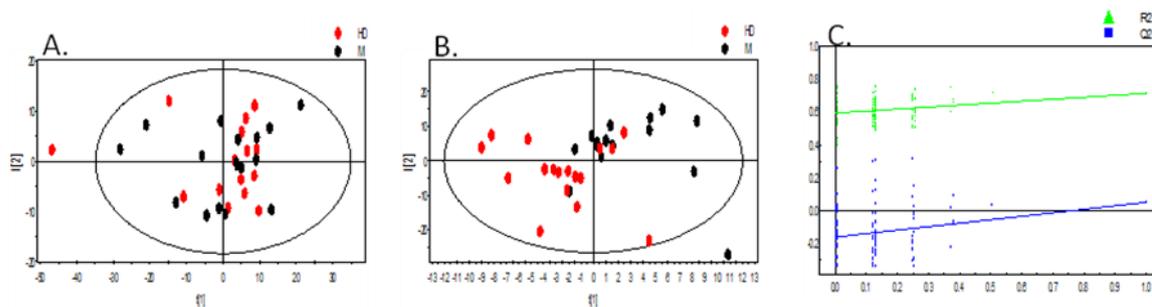
**Figure 3** Score plots from PCA (A), PLS-DA (B) and permutation testing (C) in the negative ion mode after the normalization processing procedure by nomis, on tissues samples from chickens that were administrated with hesperidin (HD, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. As plots illustrated, no separation between the two groups is observed. The  $R^2$ ,  $Q^2$  values for PLS-DA and PCA were 0.716, 0.503 and 0.336, 0.517, -0.21 respectively, employing 5 principal components for PCA and 2 for PLS-DA. The permutation testing using 100 different model permutations shows that the permuted models presented better predictive capability ( $Q^2$ ) and fit ability ( $R^2$ ) than the original model



**Figure 4** Score plots from PCA (A), PLS-DA (B) and permutation testing (C) in positive ion mode after the normalization processing procedure by median, on tissues samples from chickens that were administrated with hesperidin (HD, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. As plots illustrated, only on PLS-DA plot a separation between the two groups is observed. The  $R^2$ ,  $Q^2$  values for PCA and PLS-DA were 0.632, 0.408 and 0.725, 0.996, 0.793 respectively, employing 4 principal components for PCA and 8 for PLS-DA. The permutation testing using 100 different model permutations shows that some of the permuted models presented better predictive capability ( $Q^2$ ) than the original model



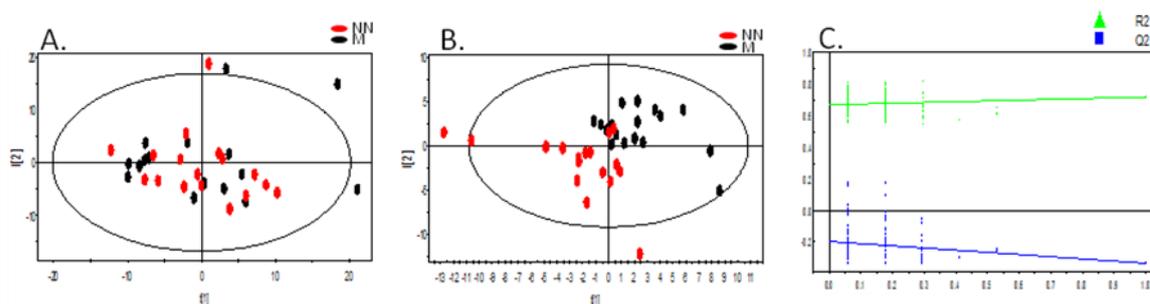
**Figure 5** Score plots from PCA (A), PLS-DA (B) and permutation testing (C) in the positive ion mode after the normalization processing procedure by sum, on tissues samples from chickens that were administrated with hesperidin (HD, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. As plots illustrated, no separation between the two groups is observed. The  $R^2$ ,  $Q^2$  values for PCA and PLS-DA were 1, 1 and 0.784, 0.996, -0.136 respectively, employing 2 principal components for PCA and 4 for PLS-DA. The permutation testing using 100 different model permutations shows that some of the permuted models presented better predictive capability ( $Q^2$ ) and fit ability ( $R^2$ ) than the original model



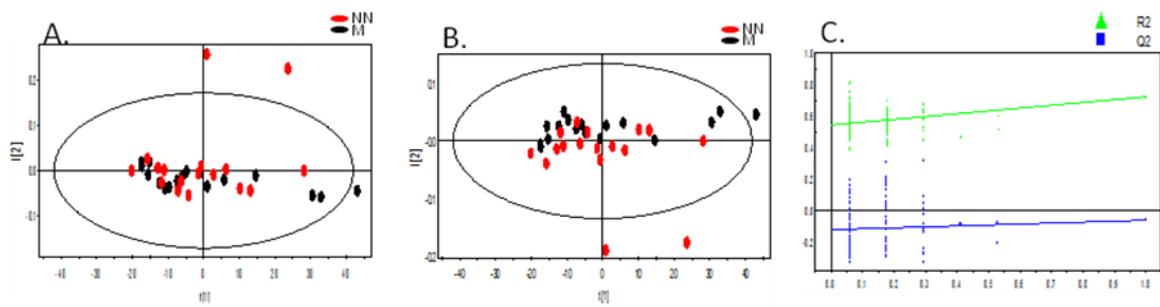
**Figure 6** Score plots from PCA (A), PLS-DA (B) and permutation testing (C) in the positive ion mode after the normalization processing procedure by nomis algorithm, on tissues samples from chickens that were administrated with hesperidin (HD, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. As plots illustrated, only on PLS-DA plot a separation between the two groups is observed. The  $R^2$ ,  $Q^2$  values for PCA and PLS-DA were 0.726, 0.545 and 0.642, 0.962, 0.521 respectively, employing 5 principal components for PCA and PLS-DA. The permutation testing using 100 different model permutations shows that some of the permuted models presented better predictive capability ( $Q^2$ ) than the original model

**S2** Score plots from PCA, PLS-DA and permutation testing after the normalization by median, sum and nomis algorithm of tissues samples from chickens that were administrated with hesperidin (HD, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis in the negative and positive ion mode

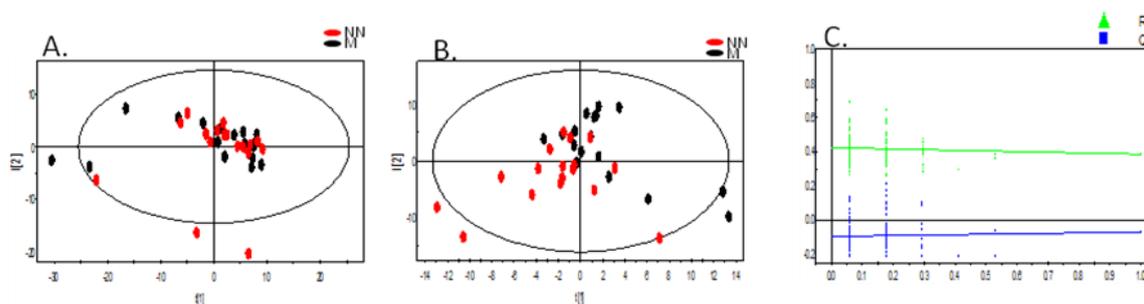
### Supplementary S3



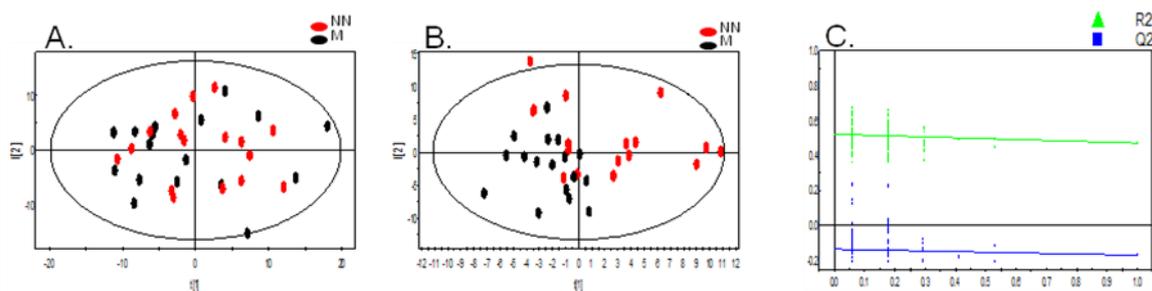
**Figure 7** Score plots from PCA (A), PLS-DA (B) and permutation testing (C) in the positive ion mode after the normalization processing procedure by median, on tissues samples from chickens that were administrated with naringin (NN, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. As plots illustrated, only on PLS-DA plot a separation between the two groups is observed. The  $R^2$ ,  $Q^2$  values for PCA and PLS-DA were 0.723, 0.4 and 0.307, 0.718, -0.331 respectively, employing 6 principal components for PCA and 3 for PLS-DA. The permutation testing using 100 different model permutations shows that some of the permuted models presented better predictive capability ( $Q^2$ ) and fit ability ( $R^2$ ) than the original model



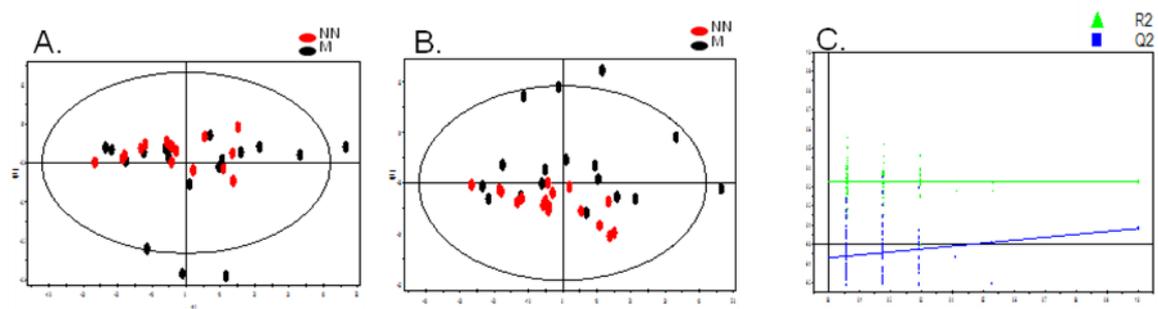
**Figure 8** Score plots from PCA (A), PLS-DA (B) and permutation testing (C) in the positive ion mode after the normalization processing procedure by sum, on tissues samples from chickens that were administrated with naringin (NN, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. As plots illustrated, no separation between the two groups is observed. The  $R^2$ ,  $Q^2$  values for PCA and PLS-DA were 1, 1 and 1, 0.721, -0.0638 respectively, employing 2 principal components for PCA and 3 for PLS-DA. The permutation testing using 100 different model permutations shows that some of the permuted models presented better predictive capability ( $Q^2$ ) and fit ability ( $R^2$ ) than the original model



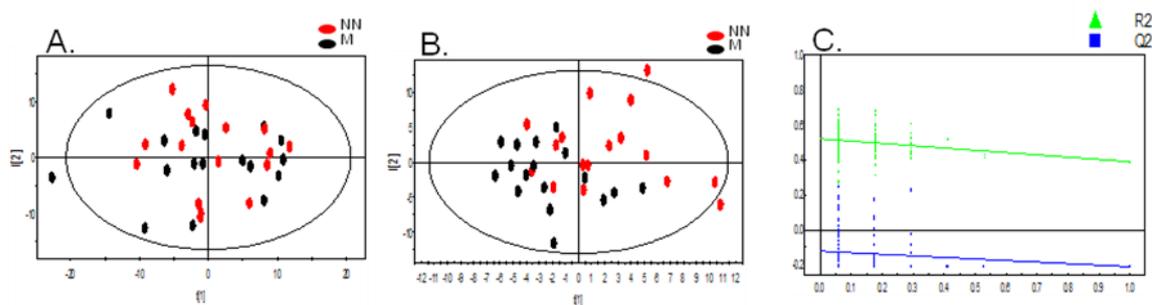
**Figure 9** Score plots from PCA (A), PLS-DA (B) and permutation testing (C) in the positive ion mode after the normalization processing procedure by nomis algorithm, on tissues samples from chickens that were administrated with naringin (NN, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. As plots illustrated, no separation between the two groups is observed. The  $R^2$ ,  $Q^2$  values for PCA and PLS-DA were 0.814, 0.52 and 0.455, 0.385, 0.0718 respectively, employing 7 principal components for PCA and 2 for PLS-DA. The permutation testing using 100 different model permutations shows that some of the permuted models presented better predictive capability ( $Q^2$ ) and fit ability ( $R^2$ ) than the original model



**Figure 10** Score plots from PCA (A), PLS-DA (B) and permutation testing (C) in negative ion mode after the normalization processing procedure by median, on tissues samples from chickens that were administrated with naringin (NN, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. As plots illustrated, only in PLS-DA plot a separation between the two groups is observed. The  $R^2$ ,  $Q^2$  values for PCA and PLS-DA were 0.483, 0.291 and 0.268, 0.471, -0.172 respectively, employing 3 principal components for PCA and 2 for PLS-DA. The permutation testing using 100 different model permutations shows that some of the permuted models presented better predictive capability ( $Q^2$ ) and fit ability ( $R^2$ ) than the original model



**Figure 11** Score plots from PCA (A), PLS-DA (B) and permutation testing (C) in the negative ion mode after the normalization processing procedure by sum, on tissues samples from chickens that were administrated with naringin (NN, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. As plots illustrated, only in PLS-DA plot a separation between the two groups is observed. The  $R^2$ ,  $Q^2$  values for PCA and PLS-DA were 1, 1 and 1, 0.329, 0.0817 respectively, employing 2 principal components for PCA and PLS-DA. The permutation testing using 100 different model permutations shows that some of the permuted models presented better predictive capability ( $Q^2$ ) and fit ability ( $R^2$ ) than the original model



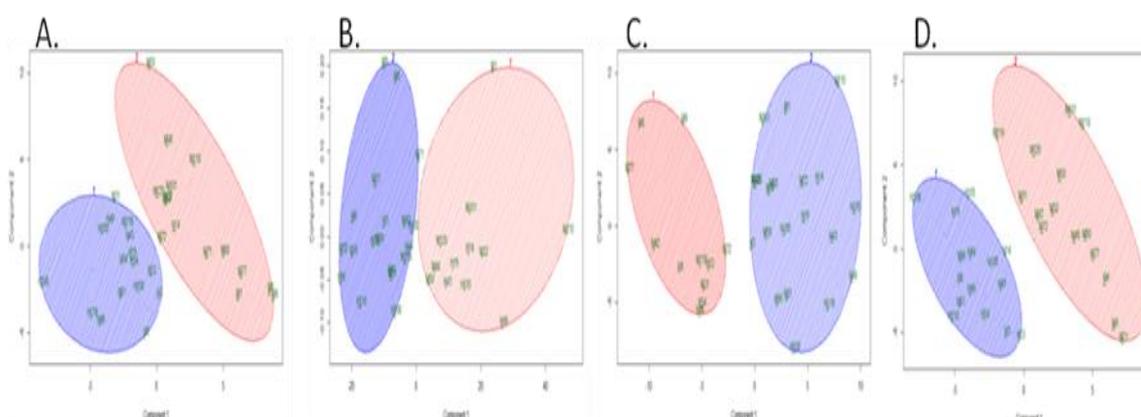
**Figure 12** Score plots from PCA (A), PLS-DA (B) and permutation testing (C) in the negative ion mode after the normalization processing procedure by nomis method, on tissues samples from chickens that were administrated with naringin (NN, red dots) and control tissues samples (M, blank dots) employing UHPLC-HRMS (Orbitrap) analysis. As plots illustrated, only in PLS-DA plot a separation between the two groups is observed. The  $R^2$ ,  $Q^2$  values for PCA and PLS-DA were 0.575, 0.319 and 0.277, 0.388, -0.21 respectively, employing 4 principal components for PCA and 2 for PLS-DA. The permutation testing using 100 different model permutations shows that some of the permuted models presented better predictive capability ( $Q^2$ ) and fit ability ( $R^2$ ) than the original model

**S3** Score plot from PCA, PLS-DA and permutation testing after the normalization by median, sum and nomis algorithm on tissues samples from chickens that were administrated with naringin (NN, red dots) and control tissues samples (M, blank dots) in employing UHPLC-HRMS (Orbitrap) analysis in the negative and positive ion modes

## Supplementary S4

Hesperidin vs. Control –Negative Ion Mode			
Normalization method	Internal validation measures		
	Connectivity	Dunn	Silhouette
Median	11.0060	0.0796	0.2565
Sum	4.0286	0.0726	0.6041
Nomis	5.6683	0.1372	0.4328
ccmn	4.3563	0.1983	0.4237

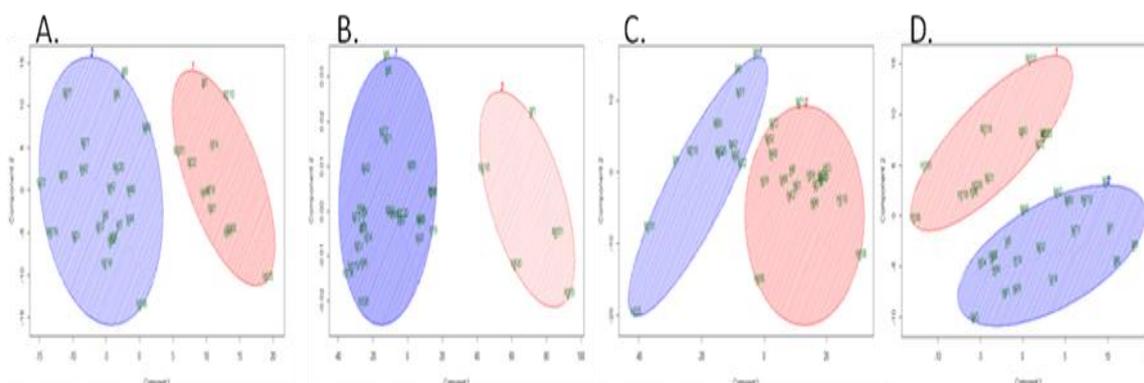
**Table 1** Values of the internal validation measures of Connectivity, Silhouette width, and Dunn index using the clustering algorithm kmeans for the treated with hesperidin vs. untreated groups in the negative ion mode



**Figure 13** Figures from the clustering of the hesperidin vs. control groups, in the negative ion mode, from the clvalid package in R environment using the clustering algorithm kmeans after the normalization by median (A), sum (B), nomis (C) and ccmn (D) methods

Hesperidin vs. Control –Positive Ion Mode			
Normalization method	Internal validation measures		
	Connectivity	Dunn	Silhouette
Median	4.9286	0.2031	0.4751
Sum	4.7115	0.5911	0.7544
Nomis	4.5829	0.1432	0.5084
ccmn	8.5881	0.1921	0.3842

**Table 2** Values of the internal validation measures of Connectivity, Silhouette width, and Dunn index using the clustering algorithm kmeans for the treated with hesperidin vs. control groups in the positive ion mode



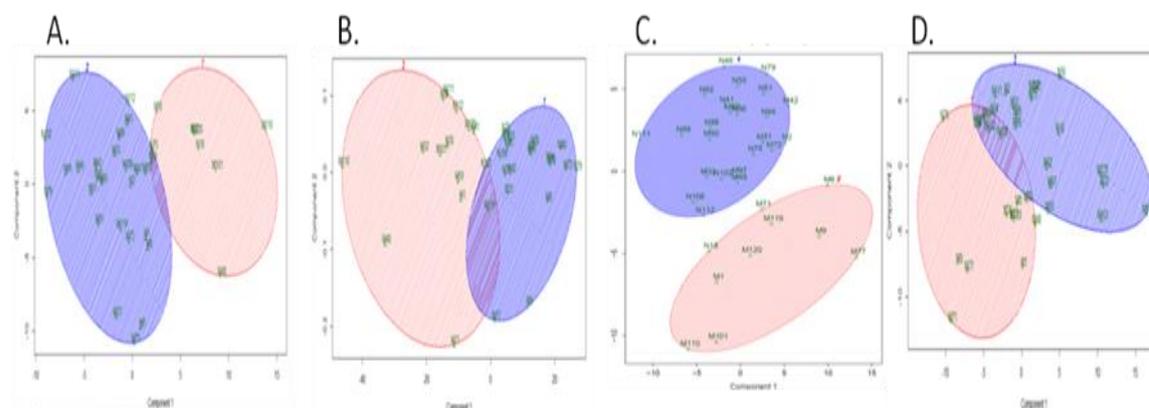
**Figure 14** Figures from the clustering of the hesperidin vs. control groups, in the positive ion mode, from the clvalid package in R environment using the clustering algorithm kmeans after the normalization by median (A), sum (B), nomis (C) and ccmn (D) methods

**S4** Comparison of the 4 normalization methods (median, sum, nomis, ccmn) according to the values of the internal validation measures of Connectivity, Silhouette width, and Dunn index using the clustering algorithm kmeans between the treated with hesperidin and untreated groups, in the positive and negative ion modes

## Supplementary S5

Naringin vs. Control – Negative Ion Mode			
Normalization method	Internal validation measures		
	Connectivity	Dunn	Silhouette
Median	8.9198	0.1380	0.4049
Sum	2.5429	0.0821	0.5934
Nomis	4.7647	0.2882	0.4358
ccmn	4.5159	0.2727	0.4561

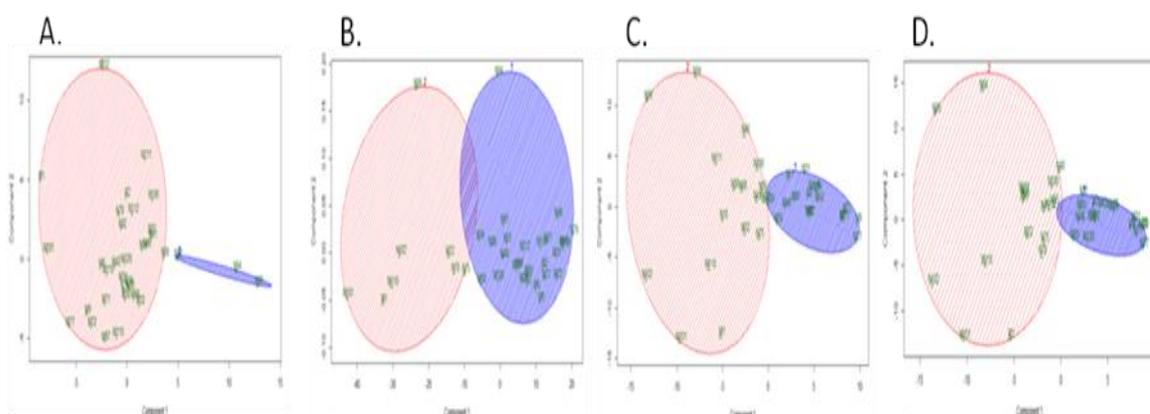
**Table 3** Values of the internal validation measures of Connectivity, Silhouette width, and Dunn index using the clustering algorithm kmeans for the treated with naringin vs. control groups in the negative ion mode



**Figure 15** Figures from the clustering of the naringin vs. control groups, in the negative ion mode, from the cvalid package in R environment using the clustering algorithm kmeans after the normalization by median (A), sum (B), nomis (C) and ccmn (D) methods

Naringin vs. Control –Positive Ion Mode			
Normalization method	Internal validation measures		
	Connectivity	Dunn	Silhouette
Median	11.5270	0.1152	0.3282
Sum	6.9714	0.1198	0.6560
Nomis	15.5024	0.0629	0.3972
ccmn	3.8579	0.5139	0.5688

**Table 4** Values of the internal validation measures of Connectivity, Silhouette width, and Dunn index using the clustering algorithm kmeans for the treated with naringin vs. control groups in the positive ion mode



**Figure 16** Figures from the clustering of naringin vs. control groups, in the positive ion mode, from the clvalid package in R environment using the clustering algorithm kmeans after the normalization by median (A), sum (B), nomis (C) and ccmn (D) methods

**S5** Comparison of the 4 normalization methods (median, sum, nomis, ccmn) according to the values of the internal validation measures of Connectivity, Silhouette width, and Dunn index using the clustering algorithm kmeans between the treated with naringin and untreated groups, in the positive and negative ion mode

## Supplementary S6

```
>assaydata2 <- read.delim("L:/MS/negative/assaydata2.txt")
> View(assaydata2)
>exprsFile2 <- "L:/MS/negative/assaydata2.txt"
>exprs2<-as.matrix(read.table(exprsFile2,header=TRUE,
sep="\t",row.names=1,as.is=TRUE))
>minimalSet2 <- ExpressionSet(assayData=exprs2)
>phenodata2 <- read.delim("L:/MS/negative/phenodata2.txt")
> View(phenodata2)
>pDataFile2 <-("L:/MS/negative/phenodata2.txt")
>pData2 <- read.table(pDataFile2,row.names=1, header=TRUE, sep="\t")
>all(rownames(pData2)==colnames(exprs2))
>metadata2<-data.frame(labelDescription=c("Type","Treatment"),
row.names=c("Type", "Treatment"))
>phenoData2<-new("AnnotatedDataFrame",data=pData2,
varMetadata=metadata2)
>annotation2 <- "uplc-orbi"
>featuredata2 <- read.delim("L:/MS/negative/featuredata2.txt")
> View(featuredata2)
> fDataFile2 <- ("L:/MS/negative/featuredata2.txt")
>fData2 <- read.table(fDataFile2,row.names=1, header=TRUE, sep="\t")
>metadata2f<-data.frame(labelDescription=("internalstandards")
,row.names=("Tag"))
>featureData2<-new("AnnotatedDataFrame",data=fData2,
varMetadata=metadata2f)
>experimentData<-new("MIAME",name="Irene",
title="Chicken_metabolomics", other=list(notes="Created from text files"))
```

```

>exampleSet2 <- ExpressionSet(assayData=exprs2,phenoData=phenoData2,
experimentData=experimentData,featureData=featureData2,annotation="uplc-
orbi")
>Y2 <- log(exprs(exampleSet2))
>inputdata2<- data.frame(pData(exampleSet2)$Type, t(Y2))
>nc <- which(with(fData(exampleSet2), Tag == "IS")==TRUE)
>norm_nomis<-Normalise(inputdata2, method = "nomis", nc=nc, saveoutput=
TRUE)
>norm_ccmn<-Normalise(inputdata2, method = "ccmn", nc = nc, ncomp = 4,
saveoutput = TRUE)
>norm_med <- Normalise(inputdata2, method = "median", saveoutput =
TRUE )
> norm_sum <- Normalise(inputdata2, method = "sum", saveoutput = TRUE)

```

**S6** Commands that were used to the package Metabolomics of the statistical programming language R, for the normalization of the datasets hesperidin vs. control and naringin vs. control in the negative and ion mode

## Supplementary S7

```
>normalized_ccmn<-read.delim("normalized_ccmn.txt")
> View(normalized_ccmn)
>test_4<-"normalized_ccmn.txt"
>Xchicken_2<-as.matrix(read.table(test_4,header=TRUE,sep="\t",
row.names=1 ,as.is=TRUE))
>Ychicken_2 <- c(rep("N",17), rep("M", 17))
>Ychicken_2 <- factor(Ychicken_2)
> ranking.DDA_2 = sda.ranking(Xchicken_2, Ychicken_2, diagonal=TRUE)
> ranking.DDA_2[1:30,]
> ranking.LDA_2 = sda.ranking(Xchicken_2, Ychicken_2, diagonal=FALSE)
> ranking.LDA_2[1:30,]
```

**S7** Commands that were used to the sda package and particularly to the sda.ranking method of the statistical programming language R, for the cat score estimation

# **General Discussion**

Hesperidin and naringin are two flavonoids with a broad spectrum of biological activities like anti-inflammatory, antioxidant, anticancer etc. The aim of the present study was to examine the effects of dietary supplementation of these substances in Ross 308 broiler chickens, in order to increase the antioxidant capacity of their tissues.

A fully validated methodology was developed for the simultaneous quantitation of hesperidin, naringin, hesperetin and naringenin in chicken plasma samples. Initially, as there is a gap in the understanding of the quantitative performance of a hybrid linear ion trap Orbitrap instrument a comparison study of its quantitation ability has been performed between the FTMS and the Ion Trap analyzer. It has been concluded that IT presented enhanced sensitivity and similar levels of repeatability and number of scans compared to the FTMS analyzer. Thus the developed methodology was applied to a pharmacokinetic study after 4 and 8 hours, and 30 days of dietary supplementation with the aforementioned flavonoids. According to the results, only naringin concentration levels were above the LLOQ after 30 days of administration. Thus a question arises about the limited access of the flavonoids to systemic circulation, which may be explained by either their rapid and extensive metabolism or by their tissue distribution.

In order to answer the first question posed, the metabolism of naringin, hesperidin and their aglycones was examined in plasma after 30 days of administration. For naringin two metabolite modifications, namely hydroxylation-methylation and methylation were detected, whereas for its aglycone naringenin, desaturation and glutathione conjugation were recognized as possible metabolite modifications. For hesperidin, the metabolite modifications, methylation and double bond reduction, were detected whereas no metabolite modifications were found for hesperetin. The identification of these metabolite modifications could explain the insufficient quantitation results, since flavonoids that do enter the systemic circulation undergo metabolism in the liver.

Furthermore, their tissue distribution was examined. A validated methodology was developed and applied after 30 days of dietary administration with hesperidin and naringin. Nevertheless, no detectable amounts have been discovered in the tissue samples. Therefore, in order to examine if the substances caused any biochemical effects to the tissues, a metabolomic study was performed to characterize the metabolic changes that occur in response to the dietary administration. Six variables were identified to discriminate naringin dietary supplementation group from control group and twelve variables hesperidin supplementation from control group.

Another problem that has been arisen during the current study (but also recognized as generally occurring issue in mass spectrometry based metabolomic analysis) is the difficulty in the detection of differentially expressed metabolites between the experimental groups caused by the larger data density in the low mass region of the mass spectra. This frequently causes an unweighted discrimination towards these data. Hence, a new methodology in the data processing procedure of MS based metabolomics was developed. The shredding of the LC-MS chromatograms into multiple  $m/z$  ranges increased the number of the identified chromatographically / spectroscopically reliable features, since the noise level is calculated individually for each mass slice. Thus the statistical weight is equally divided in all of the slices. This new methodology was applied to the chicken plasma samples after 30 days of dietary administration with naringin and hesperidin, in order to assess the impact of the administrated substances on the total metabolic profile. Four variables were identified to discriminate control group from the group after naringin administration, whereas only one variable could effectively discriminate the hesperidin treated from the control group. On the contrary, the study performed to the full  $m/z$  range, disclosed only two variables that could discriminate naringin from the control group. Consequently, this methodology could be further applied to metabolomic studies in general. Moreover, the results from the metabolomic study depicted that the metabolic changes in plasma samples in response to the

administration is more pronounced after naringin administration compared to hesperidin.

Overall, hesperidin administration presented low concentration levels in plasma samples due to the rapid metabolism and due to its tissue distribution. This could be potentially attributed to pharmacodynamic effects that it caused in tissues, since twelve variables were identified to discriminate hesperidin group from control group. On the other hand, naringin presented higher plasma levels, compared to hesperidin but low tissue distribution. This might be explained due to higher lipophilicity that hesperidin presents compared to naringin.

# **Acknowledgment**

Completion of this doctoral dissertation was possible with the support of several people. I would like to express my sincere gratitude to all of them. First of all, I would like to express my special appreciation and thanks to my supervisor Assistant Professor Dr. Evangelos Gikas, who has been a tremendous mentor for me and was always my support in the moments when there was no one to answer my queries. I would like to thank him for encouraging my research and for allowing me to grow as a research scientist. I consider it as a great opportunity to do my doctoral programme under his guidance and to learn from his research expertise.

I sincerely thank the members of the examining scientific committee, Professor Alexios-Leandros Skaltsounis, Professor Sofia Mitakou, Professor Emmanouil Mikros, Professor Georgios Theodoridis, and Principal Researcher Dr. Klapa Maria for the academic support, the brilliant comments and suggestions and the facilities provided to carry out this research work.

My colleagues have been kind enough to extend their help at various phases of this research, whenever I approached them, and mainly I would like to thank them for their friendship and I do hereby acknowledge all of them, Dr. Kouloura Eirini, Dr. Termetzi Katerina, Dr. Stathopoulos Panagiotis, Dr. Eliza Kourampie, Dr. Benaki Dimitra, Dr. Lambrinidis and Dr. Tchoumtchoua Job. Special thanks to Dagla Ioanna who became one of my best friends, for the time that we spent together in the lab and her encourage during tough times. My time at the University was enjoyable due to them who became a part of my life.

I want to express my warm thank to Milioni Alkisti-Ioanna, Fadel Anna-Maria and Efentakis Panagiotis for their support, their friendship and for all the good times that we had inside and outside the lab.

I would also like to thanks my research team from the National Hellenic Research Foundation, Institute of Biology, Medicinal Chemistry and Biotechnology, Katerina, Thaleia, Kostas for spending too many hours at the lab and helping me dealing with my MS problems. A special thanks to Dimitra

Lantzouraki for her support in a very special way, her suggestions at various points of my research programme and her friendship.

I am especially grateful to Eleni Siapi who helps me at various phases of this research. She has been very kind and patient and always willing to lend her service whenever I approached her. The joy and enthusiasm she has for my research was contagious and motivational for me, even during tough times in the PhD pursuit.

I would also like to thank Makropoulou Katerina, Lionaki Cryssa and Kapsali Foteini, the secretaries of our laboratory, for their kindly support. Especially I would like to thank Lemonaki Chryssa for all the good and bad times we spent together.

A special thanks to Dr. Lemonakis Nikolaos for his valuable guidance and consistent encouragement I received throughout this research work. Without him I would not have manage to complete this doctoral dissertation. But besides the laboratory work, he became one of the most important people in my life.

I would also like to thank all of my friends who supported me in writing, and incented me to strive towards my goal.

Lastly, I would like to thanks and dedicate this dissertation to my family. Words cannot express how grateful I am to my mother Maria, my father Dimitrios and by sister Marianthi for all of the sacrifices that they have made on my behalf, for all their love and encouragement.

# **Curriculum Vitae**

# EIRINI BAIRA, MSc

---

## Personal Details

---

Home address Akarnanias 12, Telephone +30 210 77 83 736  
11526, Athens, Greece +30 694 75 85 817

Date of birth 20.08.1982 e-mail [ebaira@pharm.uoa.gr](mailto:ebaira@pharm.uoa.gr)

---

## Education

---

- 2013 - 2016: **PhD Student**  
School of Pharmacy,  
Department of Pharmaceutical Chemistry,  
National and Kapodistrian University of Athens, Greece  
**Research Project:** "Development of State-of-the-art mass spectrometry-based analytical methodologies for the study of bioactive natural products fate after dietary administration. Bioavailability, metabolism/metabolomics studies". Supervised by Assistant Professor Gikas Evangelos
- 2009 - 2010: **MSc in Clinical Pharmacology-Therapeutics with distinction**  
Department of Medicine,  
Democritus University of Thrace,  
Alexandroupolis, Greece  
**Research Project:** "Design of bioactive compounds against novel anticancer targets. Focus on pharmacokinetics/bioavailability and drug characteristics (solubility, in vitro/in vivo stability)".  
Supervised by Professor Kolios George and Assistant Professor Tamvakopoulos Constantin, Grade: 10/10.
- 2004 - 2008: **BSc in Molecular Biology and Genetics, Grade: 6.82/10 'Very Good'**  
Democritus University of Thrace,  
Alexandroupolis, Greece  
**Research Project:**  
Genomic and Genetic Alterations of Ultraconserved RNA in Human Cancer  
Establish Stable Transfected Cell Lines.  
Supervised by Research Assistant Professor Zhang Lin, Grade: 9/10.
- 2000 - 2004: **BSc in Medical Laboratories, Grade: 7.85/10 'Very Good'**  
School of Health and Care,  
Technological Educational Institution of Athens, Greece  
**Research Project:** "Laboratory Diagnosis and Examination of Anemias with New Methods of Biotechnology".  
Supervised by Dr. Papakonstantinou Maria, Grade:10/10
-

---

**Previous Laboratory Experience**

---

February 2015-June 2015:

**Research Assistant**

«Application of state-of-the-art green technology for the development of high added value cosmeceuticals based on the Greek flora»

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

January 2013- December 2014:

**Research Assistant**

« The effects of antioxidant's dietary supplementation on animal product quality»

Department of Animal Breeding and Husbandry, Faculty of Animal Science and Aquaculture, Agricultural University of Athens, Greece

In this project UHPLC-HRMS (Orbitrap) methodologies were developed and validated for the quantification of flavonoids in chicken plasma and tissues. Moreover, metabolism/metabolomics studies were performed in plasma and tissues to reveal the differential pattern of biomolecules after dietary supplementation.

January 2011-December 2012:

**ESPA Administrative Project Manager**

ESPA Grant «Functional genomics in the pathogenesis of pulmonary fibrosis. Identification and investigation of pathogenetic mechanisms and novel therapeutic targets»

Department of Medicine, Pulmonary Clinic, Democritus University of Thrace, Greece

Responsible for administrative duties, coordination between partners and project progress reporting.

December 2010-December 2012:

**Research Assistant**

Department of Medicine, Laboratory of Immunopharmacology,

Democritus University of Thrace, Greece

Focused on establishing lab protocols for:

1.stem cell isolation from human adipose tissues

2.isolation and culture of myofibroblasts from human lung biopsies. Optimization of culture conditions and development of assays for their characterization and applications

December 2010- October 2011:

**Lab Technician**

Laboratory of Pharmacology-Toxicology, General Hospital of Alexandroupolis, Greece

Routine lab tests for determination of drug concentrations in sera and urine samples obtained from patients as well as identification of prohibited narcotics in urine samples.

- February 2010-September 2010: **MSc Dissertation**  
"Design of bioactive compounds against novel anticancer targets. Focus on pharmacokinetics/bioavailability and drug characteristics (solubility, in vitro/in vivo stability)".  
Department of Pharmacology-Pharmacotechnology,  
Biomedical Research Foundation Academy of Athens,  
Greece  
Development of LC-ESI-MS/MS methodologies for the qualitative and quantitative determination of small anti-cancer compounds in mice. Studies of in vitro and in vivo metabolism were also performed and previously unidentified metabolites were discovered.
- July 2009-November 2009: **Post-graduate Intern**  
Institute of Biological Research and Biotechnology,  
Biomedical Applications Unit,  
National Hellenic Research Foundation, Athens, Greece  
Training on molecular biology techniques.
- March 2008-September 2008: **Undergraduate Dissertation**  
"Genomic and Genetic Alterations of Ultraconserved RNA in Human Cancer. Establish Stable Transfected Cell Lines"  
Department of Obstetrics and Gynecology,  
Ovarian Cancer Research Center,  
University of Pennsylvania, Philadelphia, USA  
Study of differential mRNA expression levels in plasma isolated from ovarian cancer patients. Training on molecular and cell biology techniques involving subcloning of fluorescent tags in vectors expressing genes associated with cancer progression and their stable transfection in multiple ovarian cell lines.
- March 2004-September 2004: **Undergraduate Intern**  
General Hospital of Athens, "G. Genimatas",  
Departments of Biochemistry, Hematology, Immunology,  
Pathology and Microbiology, Athens, Greece  
Training on general routine assays: biochemical analyses on blood, urine, sperm and fecal samples, microbial cultures, antibiotic resistance charts, preparation of paraffin blocks of embedded biopsies.

---

**Publications**

**Baira E**, Siapi E, Zoumpoulakis P, Deligeorgis SG, Skaltsounis AL, Gikas E. An alternative approach for the data analysis of the data processing procedure in untargeted metabolomics by UPLC-ESI(-)-HRMS. The use of post acquisition spectral stitching. Submitted

Tzouvelekis A, Koliakos G, Ntoliou P, **Baira I**, Bouros E, Oikonomou A, Zissimopoulos A, Koliou G, Kakagia D, Paspaliaris V, Kotsianidis I, Froudarakis M, Bouros D. Stem cell therapy for Idiopathic Pulmonary Fibrosis: A protocol proposal.. J Transl Med. 2011;9:182

**Baira E**, Greshock J, Coukos G, Zhang L. Ultraconserved elements: Genomics, function and disease. RNA Biol. 2008;5:132-4.

---

**Laboratory  
Techniques****Bioanalytical  
Techniques**

Method development and validation for quantitative and qualitative characterization of small molecules and natural products in biological fluids by LC-MS/MS, sample preparation in tissues and plasma (liquid/liquid and solid phase extraction)  
Ultra High Performance Liquid Chromatography Systems (Acquity UPLC system, Waters and Accela UHPLC system)  
High Performance Liquid Chromatography Systems (HPLC Agilent 1100)  
Mass Spectrometry Systems (Sciex 4000 QTRAP, Applied Biosystems, ESI-LTQ-Orbitrap Velos, Thermo Scientific and ESI-LTQ-Orbitrap Discovery XL, Thermo Scientific)

**Pharmacokinetic  
Studies**

In vivo pharmacokinetic studies (monitoring of small molecules and natural products in plasma and tissues of animal models by LC-MS)  
In vitro/in vivo metabolite identification

**Metabolomics**

Metabolomic studies (data processing, statistical analysis, biomarker identification/validation)

**Cell Biology**

Cell Cultures (freeze/thaw cell lines and primary cells, expansion of cultures),  
Stable transfection in cell lines, wound healing assay, isolation of primary cells (myofibroblasts)

**Molecular Biology**

Isolation and purification of nucleic acids (DNA, RNA),  
polymerase chain reaction (conventional PCR, real-time PCR, reverse transcription PCR), electrophoretic analysis of macromolecules (proteins and nucleic acids), molecular cloning (genetic engineering, transformation, bacterial cultures)

---

**Computer Skills**

---

ECDL (European Computer Driving License) diploma with the following thematic modules:

Databases

Concepts of IT

Word Processing

Presentations

Information and Communication

Spreadsheets

Using the Computer Managing Files

Very good knowledge of:

MZmine

R statistic language

Simca

MetWorks

---

**Languages**

---

First Certificate in English, University of Cambridge

Certificate of Competency in English, University of Michigan

---

**Personal Interests**

---

Music, European cinema, Modern literature, Yoga, Pilates

---

**Recommendation Letters**

---

Evangelos Gikas, Assistant Professor, Laboratory of Pharmaceutical Analysis, School of Pharmacy, National and Kapodistrian University of Athens, Greece. Phone number: +30-210 7274850 [vgikas@pharm.uoa.gr](mailto:vgikas@pharm.uoa.gr).

Tamvakopoulos Constantin, Researcher B'- Assistant Professor Level, Biomedical Research Foundation Academy of Athens, Department of Pharmacology-Pharmacotechnology, Athens, Greece. Phone number: +30-210- 6597475 [ctamvakop@ bioacademy.gr](mailto:ctamvakop@bioacademy.gr)

Lin Zhang, Research Assistant Professor, Center for Research on the Early Detection & Cure of Ovarian Cancer, Center for Research on Reproduction and Women's Health, Department of Obstetrics and Gynecology, University of Pennsylvania School of Medicine. Office: 1209 BRB II/III; Lab: 1331 BRB II/III 421 Curie Blvd. Philadelphia, PA 19104 Phone number: 215-573-4780 (office), 215-746-5137 (Lab) Fax: 215-573-[linzhang@mail.med.upenn.edu](mailto:linzhang@mail.med.upenn.edu)

---